

**ADVANCES IN HIGH BANDWIDTH NANOMECHANICAL FORCE
SENSORS WITH INTEGRATED ACTUATION**

**A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF MECHANICAL
ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY**

Joseph Carl Doll
July 2012

© Copyright by Joseph Carl Doll 2012

All Rights Reserved

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Beth Pruitt) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Miriam Goodman)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Anthony Ricci)

Approved for the University Committee on Graduate Studies.

Abstract

Life is built upon mechanical forces, which play a central role in everything from cell division to embryonic development. Rather than acting as passive mechanical elements, cells and molecules sense and actively respond to mechanical loads. One example of cellular force sensing is mechanotransduction, the conversion of mechanical energy into an electrical signal, which underlies our senses of hearing, touch and balance. For example, the cochlear hair cells in your inner ear are exquisitely sensitive and fast, capable of sensing piconewton-scale forces at the microsecond-time scale. But in order to understand such fast mechanotransduction processes we must first be able to apply and measure small, fast forces.

A variety of instruments have been developed for the precise measurement of atomic-scale forces and displacements in the past 25 years. The most commonly used techniques are atomic force microscopy, magnetic tweezers, and optical tweezers. Each provides a tradeoff in force, displacement and time resolution, but none of them are capable of applying and detecting forces fast enough for the study of cochlear hair cells.

In order to address this technological gap we have developed microfabricated force probes for the application and measurement of forces at the piconewton- and microsecond-scale. In order to simultaneously achieve a high resonant frequency (20-400 kHz in air, 10-100 kHz in water), low spring constant (0.3-40 mN/m) and low minimum detectable force (1-100 pN), the probes are roughly 300 nm thick, 1 micron wide and 30-200 microns long. Force applied to the cantilever tip is transduced into a voltage by a piezoresistive silicon strain gauge that is embedded in the beam. Actuation is accomplished through a piezoelectric aluminum nitride film or a resistively heated aluminum film to enable high-speed operation without spurious resonant modes. The probes are mass produced on

silicon wafers using conventional batch fabrication techniques, and their dimensions are individually adjusted lithographically to accommodate a wide range of desired force and time resolutions. Optics are not required for sensing or actuation so the probes can be integrated with any standard upright up inverted microscope.

This thesis presents the design, fabrication and characterization of the force probes. Several enabling technologies and techniques were developed in the process. We will begin by discussing the mechanical, electrical and thermal design of the force probes with an emphasis on piezoresistor design. Numerical design optimization is utilized to satisfy the numerous design and performance constraints. Next, the seven- and nine-mask fabrication processes used to manufacture the thermally and piezoelectrically actuated probes will be presented. The sensing and actuation performance of the probes will be individually addressed before discussing their integration, particularly crosstalk compensation. Finally, preliminary data on the measurement of mammalian hair cell kinetics will be presented and possible future directions will be discussed.

The improved design, fabrication and circuit methodologies described here enable numerous performance improvements over prior work. For example, piezoresistive cantilever force resolution is improved 10-20 fold over prior cantilevers of comparable thickness. Similarly, the crosstalk between the piezoelectric actuator and piezoresistor sensor are 10-fold smaller than the best results reported to date.

Acknowledgements

Six years is a long time and many, many people helped this thesis come together. I'd like to start off by thanking my advisor, Beth Pruitt. When I came to Stanford she took a chance by funding me when it seemed like nobody else would. When I was burned out she provided encouragement and flexibility, and the rest of the time she provided absolute freedom.

Miriam Goodman and Tony Ricci provided guidance along the way. Their broad knowledge and interest in the engineering details repeatedly steered me in the right direction. Roger Howe provided research and career advice from my first quarter at Stanford until the very end when he chaired my oral exam committee. Ellen Kuhl was kind enough to serve on both my qualifying and oral exam committees and she taught me cell mechanics.

I would like to thank all of my Pruitt group labmates, in particular: Alvin Barlian, Alex Haemmerli, Vikram Mukundan, Bryan Petzold, Chelsey Simmons and Bex Taylor. My thesis only briefly mentions *C. elegans*, but the Worm Club attendees provided helpful advice over the years.

The piezoresistor gang was instrumental to my research. Nahid Harjee taught me how to work in the SNF and provided moral support. Sung-Jin Park mentored me when I first arrived at Stanford, he taught me how to design and build PCBs, and we worked extensively on cantilever optimization together. Joe Mallon and Ali Rastegar taught me piezoresistor and circuit fundamentals. I leaned heavily on the work of earlier Stanford piezoresistor researchers and hope that this thesis will be as helpful to those that follow.

I would like to thank all of my collaborators. In particular, it was a pleasure to work with Haneesh Kesari in probing the nature of sticky contacts. We spent long hours together at the AFM, which I fondly look back on. Tony Ricci and Anton Peng helped define the

force probe design and performed all of the actual hair cell experiments.

Although working in the SNF sometimes felt like an obstacle course of broken tools, I'd like to thank the SNF staff and labmates who helped, in particular: Nancy Latta, Mahnaz Mansourpour, Ed Myers, John Schott, Gary Sosa, Maurice Stevens, Mary Tang, Uli Thumser, Mario Vilanova of the SNF and Keith Best, Ping Ding and Linda Ohara from ASML. Rishi Kant developed Jobscripter, which shaved a few weeks off of my jobfile development. Ginel Hill helped double my cantilever yield with her oxide cracks and provided advice on working with soft cantilevers. Eric Perozziello, J Provine, and Usha Raghuram provided process advice. The second fabrication run was greatly aided by Sandrine Martin and Rob Hower of the University of Michigan, Tina Lamers and Keri Williams of Avago, and Valeriy Felmetsger of OEM Group.

I'm grateful to Brad Nelson of ETH Zurich for the privilege of working his lab for several months, where the cantilever optimization work got started. I'd also like to thank my undergraduate research advisor, Luke Lee, for helping me to get started in research, making me write an undergraduate thesis (which was good practice), pointing me in the direction of Beth once I got into Stanford, and for his advice that "if you don't publish it, you might as well have never done it". I haven't forgotten it.

I was lucky enough to receive financial support from the National Science Foundation's Graduate Research Fellowship Program (NSF GRFP) and a National Defense Science and Engineering Graduate Fellowship (NDSEG). Without their help I would have spent more time worrying about funding and less time doing research.

My friends and family have provided unwavering support over all of these years. I have the utmost gratitude and love for my grandmother, parents and siblings.

Finally, and most importantly, I'd like to thank Cassie. She has worked behind the scenes in all of this: proofreading papers aloud, providing cool-headed advice, and making sure that my cartoons were up to snuff. She put up with me when I was spending every waking hour in the fab for months on end. She has been the invisible co-author on every one of my papers.

Contents

Abstract	v
Acknowledgements	vii
1 Introduction	1
1.1 Cochlear hair cell mechanics and physiology	1
1.1.1 Bundle mechanics	8
1.1.2 Channel kinetics	10
1.2 Force probe design requirements	13
1.3 The limits of existing force probes	16
1.4 Microscale transduction techniques	21
1.4.1 Force sensing techniques	21
1.4.2 Actuation techniques	30
1.4.3 Probe topology and prior work	36
1.5 Thesis outline	45
2 Piezoresistor fundamentals	47
2.1 Notation fundamentals	48
2.1.1 Crystal structure	49
2.1.2 Stress, strain and tensors	50
2.2 Piezoresistivity	52
2.2.1 Crystallographic orientation	52
2.2.2 Piezoresistance factor	57
2.2.3 Nonlinearity	65

2.2.4	Theoretical models	69
2.3	Signal conditioning	73
2.3.1	Process variation and temperature compensation	76
2.3.2	Heterodyne biasing	80
2.4	Sensor resolution	81
2.4.1	Noise	81
2.4.2	Sensitivity	103
2.4.3	Calculating sensor resolution	111
2.5	Piezoresistor fabrication and modeling	113
2.5.1	Processing fundamentals	116
2.5.2	Predeposition	119
2.5.3	Ion implantation	127
2.5.4	Epitaxy	137
2.5.5	Polysilicon	142
2.5.6	Electrical modeling	143
2.5.7	Experimental validation techniques	146
2.6	Temperature effects	148
2.6.1	Piezoresistor sensitivity	149
2.6.2	Temperature coefficient of resistance	150
2.6.3	Piezoresistor self-heating	151
2.6.4	Power dissipation and measurement resolution	167
2.7	Design optimization	171
2.8	Practical notes	179
2.8.1	Wafer selection	179
2.8.2	Dopant-induced lattice strain	181
2.8.3	Piezoresistor surface damage	187
2.8.4	Lateral dopant diffusion	188
3	Force probe design	191
3.1	Sensor design	193
3.1.1	Mechanics	194

3.1.2	Piezoresistor optimization	197
3.1.3	Sensor designs	199
3.2	Actuator design	202
3.2.1	Calculating tip deflection	202
3.2.2	Actuator dimensions	203
3.2.3	Thermal actuator	205
3.2.4	Piezoelectric actuator	209
3.3	Sensor and actuator integration	212
3.3.1	Mechanics of the combined system	214
3.3.2	Higher order eigenmodes	220
3.3.3	Actuator-sensor crosstalk	220
3.4	Practical design issues	225
3.4.1	Die layout	226
3.4.2	Depletion layers	227
3.4.3	Excess electrical resistance	228
3.5	Circuit design	231
4	Fabrication	235
4.1	Processing fundamentals	235
4.2	Device design	240
4.2.1	Sensor-only probes	241
4.2.2	Thermally actuated probes	244
4.2.3	Piezoelectrically actuated probes	245
4.2.4	Miscellaneous designs	247
4.3	Wafer layout	249
4.4	Process flow	250
4.5	Process issues and iterations	257
4.5.1	DRIE and BOX release steps	257
4.5.2	BOX cracking and probe yield	263
4.5.3	Sheet resistance and contact resistivity issues	264
4.5.4	Assorted etch issues	276

5 Characterization	282
5.1 Finished devices	282
5.2 Die attach and wirebonding	283
5.3 Passivation	286
5.4 Probe mechanics	296
5.5 Piezoresistive sensor	307
5.5.1 Electrical characteristics and resistance	308
5.5.2 Dopant concentration profile	311
5.5.3 Noise	314
5.5.4 Sensitivity	316
5.5.5 Self-heating	319
5.5.6 Displacement and force resolution	323
5.6 Actuator performance	328
5.6.1 Thermal actuator	329
5.6.2 Piezoelectric actuator	332
6 Cochlear hair cell experiments	346
6.1 Experimental setup	346
6.2 Patch clamp crosstalk	348
6.3 Hair cell measurements	351
7 Conclusions	353
7.1 Design optimization is a Swiss Army knife	353
7.2 Fundamental force sensing limits	355
7.3 Future directions	357
A Glossary of mathematical symbols	363
B List of publications	369
C Circuit details	373
D Reticle and wafer layouts	382

E Fabrication process details	389
E.1 Wafer handling	389
E.2 Runsheets	390
F The first fabrication run	408
F.1 Device design	408
F.2 Reticle and wafer layouts	412
F.3 Process flow	412
F.4 Results and discussion	421
F.4.1 Sensor-only probes	421
F.4.2 Piezoelectric-only cantilevers	433
F.4.3 Piezoelectrically actuated probes	441
G Ion implantation lookup tables	450
H Code	463
H.1 TSUPREM simulations	463
H.1.1 simulation.template	465
H.1.2 simulationControl.py	467
H.1.3 postProcessTables.m	470
H.1.4 calculateProfileProperties.m	474
H.1.5 richterModel.m	476
H.1.6 resistivity.m	477
H.2 Optimization code	480
H.2.1 sampleCode.m	482
H.2.2 cantilever.m	485
H.2.3 cantileverDiffusion.m	567
H.2.4 cantileverEpitaxy.m	574
H.2.5 cantileverImplantation.m	579
Bibliography	588

List of Tables

1.1	Summary of probe performance requirements	13
2.1	Piezoresistive coefficients for silicon	55
2.2	Longitudinal and transverse piezoresistive coefficients for silicon	55
2.3	Fitting parameters for piezoresistance factor models	65
2.4	Noise coefficients for example instrumentation amplifier options	97
2.5	Summary of piezoresistor fabrication techniques	115
2.6	Summary of the dopant/process combinations modeled here	116
2.7	Best fit parameters for the Tsai predeposition model	125
2.8	TSUPREM simulation process parameters	130
2.9	The effect of dopant diffusion on epitaxial piezoresistor performance	140
2.10	Temperature dependent carrier mobility modeling parameters	144
2.11	Thermal conductivities of commonly used materials	154
2.12	Summary of the cantilevers used to investigate heat transfer	162
2.13	Model parameters for simplified resolution/power dissipation analysis	169
2.14	Design matrix summarizing the interplay between design and performance .	172
3.1	Summary of the force probe sensor designs	201
3.2	Summary of the second generation PRT actuator designs	210
3.3	Summary of the second generation PRPE actuator designs	213
4.1	Probe yield improvement with patterned cracks	264
4.2	Piezoresistor and contact thinning data	280
5.1	Measured and predicted frequency responses in air and water	308

5.2	Performance summary of several example devices	327
A.1	Glossary of mathematical symbols	363
A.1	Glossary of mathematical symbols (continued).	364
A.1	Glossary of mathematical symbols (continued).	365
A.1	Glossary of mathematical symbols (continued).	366
A.1	Glossary of mathematical symbols (continued).	367
A.1	Glossary of mathematical symbols (continued).	368
E.1	Runsheet for the PR process from the first fabrication run	394
E.2	Runsheet for the PE process from the first fabrication run	395
E.2	Runsheet for the first fabrication PE process (continued).	396
E.3	Runsheet for the PRPE process from the first fabrication run	397
E.3	Runsheet for the first fabrication PRPE process (continued).	398
E.4	Runsheet for the second fabrication run	399
E.4	Runsheet for the second fabrication run (continued).	400
E.4	Runsheet for the second fabrication run (continued).	401
E.4	Runsheet for the second fabrication run (continued).	402
E.5	Runsheet for the fabrication of POI wafers	403
E.6	Aluminum etch program	404
E.7	Aluminum nitride etch program	404
E.8	Molybdenum etch program	405
E.9	Silicon etch program	405
E.10	Optical emission wavelengths used for RIE endpoint detection	405
E.11	Endpoint detection programs	406
F.1	AlN deposition parameters from the first fabrication run	417
F.2	Film thicknesses in the first piezoelectric fabrication run	418
F.3	Summary of the first generation sensor-only designs	448
F.4	Summary of the first generation actuator-only designs	449
F.5	Summary of the first generation integrated probe designs	449
G.1	Look-up table for β_1^* (-) and β_2^* (μm) for boron ion implantation.	451

G.2	Look-up table for β_1^* (-) and β_2^* (μm) for phosphorus ion implantation.	452
G.3	Look-up table for β_1^* (-) and β_2^* (μm) for arsenic ion implantation.	453
G.4	Look-up table for R_s (Ω/\square) for boron ion implantation.	454
G.5	Look-up table for R_s (Ω/\square) for phosphorus ion implantation.	455
G.6	Look-up table for R_s (Ω/\square) for arsenic ion implantation.	456
G.7	Look-up table for N_z (μm^{-2}) for boron ion implantation.	457
G.8	Look-up table for N_z (μm^{-2}) for phosphorus ion implantation.	458
G.9	Look-up table for N_z (μm^{-2}) for arsenic ion implantation.	459
G.10	Look-up table for t_j (μm) for boron ion implantation.	460
G.11	Look-up table for t_j (μm) for phosphorus ion implantation.	461
G.12	Look-up table for t_j (μm) for arsenic ion implantation.	462

List of Figures

1.1	Structure of the ear	3
1.2	Mechanotransduction in cochlear hair cells	4
1.3	The arrangement of hair cells within the cochlea	5
1.4	Diagram of mechanotransduction in a pair of stereocilia	6
1.5	Adaptation and force generation processes	7
1.6	Measuring bundle mechanics	9
1.7	Measuring channel kinetics	11
1.8	Applications of fast force probes to cochlear hair cells	12
1.9	Hair cell experimental setup	15
1.10	Spring constants and time resolutions for conventional techniques	17
1.11	The ideal force probe is a one-dimensional cantilever beam.	19
1.12	Stiffness and rise time for cantilever beams of varying thickness	20
1.13	Optical transduction of cantilever beam deflection	22
1.14	Challenges with optical transduction	23
1.15	Summary of the all-electronic techniques for sensing cantilever deflection .	24
1.16	Differential capacitive force sensing	25
1.17	A six-axis capacitive sensor with μN force and $\text{nN}\cdot\text{m}$ torque resolution . . .	27
1.18	Piezoresistive and piezoelectric transduction	28
1.19	Summary of microcantilever actuator technologies	30
1.20	SEM micrograph of the first PRPE scanning probes	37
1.21	Cross-section of the first PRPE scanning probes	38
1.22	PZT-based PRPE probe	39
1.23	Alternative probe layouts to reduce capacitive crosstalk	41

1.24	Schematic of a piezoresistive cantilever with thermal actuation	42
1.25	Examples of 89 nm thick piezoresistive cantilevers	44
2.1	Strain gauges transduce strain into a resistance change	47
2.2	Silicon crystal structure and wafer orientations	49
2.3	General stress elements on a unit cell	50
2.4	Smith's original piezoresistance test configuration	52
2.5	Piezoresistive coefficients as a function of direction	56
2.6	A piezoresistance test chip based upon circular piezoresistors	58
2.7	Kanda's model for p- and n-type piezoresistance	59
2.8	Comparison between Kanda's theoretical model and experimental data . . .	61
2.9	Comparison between Harley's empirical fit and experimental data	63
2.10	Comparison between Richter's theoretical model and experimental data . .	64
2.11	Comparison between the modified Richter model and experimental data . .	66
2.12	Variation in piezoresistive coefficients with strain	67
2.13	Energy band splitting with strain in diamond-type lattices	70
2.14	Hole mobility enhancement in semiconductors as a function of stress . . .	72
2.15	Measurement circuits to transduce a resistance change into a voltage change	73
2.16	Example of CMOS-integrated piezoresistive cantilevers	78
2.17	Experimental noise PSD example	79
2.18	A modulation-demodulation sensor readout circuit	80
2.19	Example noise measurement system	85
2.20	Summary of the Hooge and McWhorter models for 1/f noise	87
2.21	The Hooge parameter as a function of diffusion length	92
2.22	Schematic representation of noise cascading through amplifier stages . . .	95
2.23	A typical noise spectrum of a piezoresistive sensor and measurement circuit	101
2.24	Noise spectra for varying bridge bias voltages	102
2.25	Schematic of an end loaded cantilever beam	106
2.26	Dopant concentration profiles from diffusion, ion implantation and epitaxy .	114
2.27	Example of piezoresistor junction isolation	117
2.28	Total and active concentration limits as functions of temperature	119

2.29	Intrinsic carrier concentration as a function of processing temperature	122
2.30	Simulated dopant concentration profiles for phosphorus predeposition	122
2.31	Simulated and measured predeposition profiles	125
2.32	Simulated and measured predeposition sheet resistances	126
2.33	Phosphorus diffusion through an oxide mask	126
2.34	Variations in as-implanted boron dopant profiles with energy and dose	132
2.35	Implanted phosphorus dopant profiles before and after annealing	133
2.36	As, B and P dopant profiles with inert and passivated anneal processes	135
2.37	P and B dopant profiles with and without post-implant oxidation	136
2.38	The effect of dopant diffusion on epitaxial piezoresistor performance	141
2.39	Resistivity and mobility as functions of dopant concentration	145
2.40	TCS as a function of dopant concentration	149
2.41	TCR as a function of dopant concentration	150
2.42	Qualitative example of piezoresistor self-heating	153
2.43	The thermal conductivity of silicon as a function of film thickness	155
2.44	The thermal conductivity of silicon as a function of dopant concentration . .	156
2.45	The thermal conductivity of silicon as a function of temperature	157
2.46	Example of piezoresistive cantilever self-heating	161
2.47	Simulated and measured cantilever temperature profiles	164
2.48	Force resolution as functions of piezoresistor temperature in air and water .	165
2.49	A simplified analytical model for maximum and tip cantilever temperatures	167
2.50	Theoretical variation in force resolution and power dissipation	170
2.51	Examples of convex and non-convex optimization problems	176
2.52	Optimizer convergence from three random designs	177
2.53	Numerical optimizer convergence statistics	178
2.54	Example of a curved cantilever due to dopant-induced lattice strain	181
2.55	Cantilever tip deflection as a function of length	183
2.56	Variation in sheet resistance with oxidation	186
2.57	Example of surface damage from RIE	187
2.58	Lateral dopant diffusion effects	189

3.1	Device layouts	192
3.2	Thermal actuator thickness selection	206
3.3	Example of the thermal actuator steady-state response	207
3.4	Thermal actuator FEA	208
3.5	Relative placement of the actuator and sensor	214
3.6	Force probe spring constant as a function of actuator length	218
3.7	Force probe resonant frequency as a function of actuator length	219
3.8	First five eigenmodes of an example probe	221
3.9	Example of thermal crosstalk within a thermally actuated force probe	226
3.10	Upright microscope compatibility	227
3.11	Contact resistivity as a function of dopant concentration	229
3.12	Summary of the sensor circuit	231
3.13	Voltage reference noise performance	233
3.14	Circuit noise and source impedance	234
4.1	Reticle layout overview	236
4.2	A jobfile maps images between the reticle and wafer	237
4.3	Image stitching summary	239
4.4	Snap tab design and operation	240
4.5	PR device layout	241
4.6	Scalebar and greyhound chip art	243
4.7	PRT device layout	244
4.8	PRPE device layout	246
4.9	Additional device designs	247
4.10	Example exposure grid and test features	248
4.11	Wafer layout	249
4.12	Second generation PR process flow	251
4.13	Second generation PRT process flow	254
4.14	Second generation PRPE process flow	255
4.15	Initial backside etch results using the DRIE tool in the SNF	258
4.16	Backside undercut for three different DRIE processes	259

4.17	Example of metal corroding when the carrier wafer is debonded in hot water	260
4.18	Examples of polymer buildup on released probes	261
4.19	Polymer cleanup process	261
4.20	Example of oxide flaps after etching the BOX via RIE	262
4.21	Patterned cracks protect force probes from spontaneous BOX cracks	263
4.22	TLM data from the first batch of second generation wafers	265
4.23	Contact resistivity data from lightly doped piezoresistors	266
4.24	Current-voltage data from lightly doped piezoresistors	267
4.25	Contact resistivity data before and after FGA	268
4.26	Variation in contact resistivity with metallization	269
4.27	Examples of Al spiking through the device layer after an FGA	271
4.28	One-step and two-step diffusion processes	272
4.29	Test wafer with the PSG stripping step	273
4.30	Variation in contact resistivity with doping level	274
4.31	Example dopant concentration profiles	275
4.32	Example of electrode thinning and microtrenching in the AlN etch	277
4.33	Example of AlN stringers left on a finished force probe	278
4.34	An example of stringers after the interconnect etch step	279
4.35	Example of condensation on POCl ₃ doped wafers after several hours in air	280
4.36	Examples of color changes in processing thin SOI wafers	281
5.1	Optical micrographs of finished probes	283
5.2	Scanning electron micrographs of finished probes	284
5.3	Example of a finished device before mounting	285
5.4	Top and side views of a fully assembled device	286
5.5	SEM cross section of a parylene coated sample	288
5.6	Mechanical effects of passivation coatings	289
5.7	Passivation measurement setup	290
5.8	Examples of ALD corrosion	292
5.9	Summary of parylene passivation issues	293
5.10	Example of device operation in air and liquid with parylene passivation	294

5.11	Example of parylene-induced stiction at air-water interfaces	295
5.12	Example of force probe thermomechanical noise spectrum	300
5.13	Predicted and measured resonant frequencies in air	301
5.14	Predicted and measured spring constants for unpassivated force probes . . .	302
5.15	Laser-induced heating during thermomechanical noise measurements . . .	303
5.16	Quality factor scaling in air	304
5.17	Thermomechanical noise spectra before and after parylene coating	305
5.18	Experimental data on the effects of parylene passivation	306
5.19	Measured frequency responses in air and water	307
5.20	Current-voltage data from a typical released piezoresistor	309
5.21	Electrical resistance as a function of piezoresistor length	310
5.22	Comparison between SIMS and SRA data for SOI and POI samples . . .	311
5.23	Example noise spectra	313
5.24	Distribution of Hooge factor measurements	315
5.25	The Hooge factor is independent of bridge bias and piezoresistor size . . .	316
5.26	Optical micrograph of the displacement sensitivity calibration setup . .	317
5.27	Example of sensitivity calibration data	318
5.28	Predicted and measured displacement sensitivities	319
5.29	Piezoresistor self-heating during operation	321
5.30	Thermal impact of parylene passivation	322
5.31	Force resolution variation with measurement bandwidth	324
5.32	Force resolution comparison with prior work	325
5.33	Force noise as a function of bridge bias	326
5.34	Thermal actuator frequency and step responses	330
5.35	Simultaneous thermal actuation and sensing	331
5.36	Piezoelectric actuator performance summary	333
5.37	Typical film crystallinity measurement data	334
5.38	Cross-sectional SEM of an actuator with a 150 nm thick AlN film . . .	335
5.39	Frequency response of a piezoelectrically actuated probe in air	336
5.40	Actuator matching between main and compensation probes	337
5.41	Piezoelectric actuator-sensor crosstalk summary	337

5.42	Piezoelectric actuator-sensor crosstalk comparison	338
5.43	Capacitive crosstalk compensation circuit	339
5.44	Crosstalk comparison with a reference PCB	341
5.45	Actuator-sensor crosstalk increases with Wheatstone bridge bias	342
5.46	Reduced crosstalk with differential actuator driving	343
5.47	Time-domain mechanical crosstalk data	344
6.1	Experimental setup summary	347
6.2	Micrograph of a force probe with hair cells	348
6.3	Process for attaching beads to the force probes	349
6.4	Patchclamp crosstalk schematic	350
6.5	Example of capacitive coupling with the patch clamp	351
6.6	Example mechanoreceptor current from a rat inner hair cell	352
7.1	Force resolution for two different design methodologies	356
C.1	Amplifier and crosstalk compensation board schematic	377
C.2	Amplifier and crosstalk compensation board layout	378
C.3	Four-layer cantilever board layout	379
C.4	Example of an assembled amplifier circuit	380
C.5	Example of an assembled cantilever cable	381
C.6	Example of an assembled power cable	381
D.1	Reticle fast-pr-1	383
D.2	Reticle fast-pr-2	384
D.3	Reticle fast-probe-1	385
D.4	Reticle fast-probe-2	386
D.5	Reticle fast-probe-3	387
D.6	Reticle fast-probe-4	388
F.1	Device layouts from the first fabrication run	409
F.2	Wafer layout	413
F.3	Sensor-only process flow	414

F.4	Piezoelectric-only process flow	416
F.5	Process flow for piezoelectrically actuated probes	420
F.6	Interactions between Ti and forming gas	421
F.7	A finished 10 μm thick sensor-only probe	422
F.8	Finished 340 nm thick sensor-only probes	423
F.9	Backside release alignment issues	424
F.10	BOX buckling and cracking	425
F.11	Comparison between measured and predicted resonant frequencies	426
F.12	Piezoresistor electrical characteristics	427
F.13	Example of thermal runaway in piezoresistive cantilevers	428
F.14	Example displacement sensitivity data	429
F.15	Piezoresistor noise characteristics	430
F.16	Piezoresistor noise spectra in air, DI water and PBS	432
F.17	Excess piezoresistor noise from displacement noise	433
F.18	SEMs of piezoelectric-only probes	434
F.19	AlN film texture characterization	435
F.20	Variation in rocking curves with deposition parameters	436
F.21	Example AlN atomic composition data	437
F.22	Variation in piezoelectric coefficients with rocking curve FWHM	438
F.23	Issues with wet etching AlN	439
F.24	Example AlN current-voltage data	440
F.25	SEMs of piezoelectrically actuated probes	441
F.26	Piezoresistor damage due to AlN deposition	442
F.27	Non-Ohmic contacts in damaged piezoresistors	443
F.28	Surface stress and potato chipping	444
F.29	Mechanical nonlinearity caused by transverse bending	445
F.30	Actuator-sensor capacitive crosstalk	446

Chapter 1

Introduction

This thesis will present the development of a force probe capable of sensing and delivering high-speed atomic-scale forces and displacements. In this chapter we will first introduce the biological applications that motivated the development of the probe. Based upon these applications we will develop a set of performance requirements for a high-speed probe and compare them with the capabilities of conventional scientific force measurement instruments. The mismatch between requirements and capabilities will motivate the need for a new type of force probe. We will present the high level design of the probe, and discuss the alternative sensing and actuation technologies that could have been incorporated into the probe. Once we select sensing and actuation techniques we will review prior work before outlining the remainder of the thesis.

1.1 Cochlear hair cell mechanics and physiology

Mechanical forces are central to biological processes. Examples range from macroscale skeletal development to the movement of molecular motors and proteolysis at the microscale [1, 2]. Cells and molecules are not passive mechanical components, however; they sense, adapt, and respond to forces. One way in which cells sense force is by coupling a mechanical load to the opening of an ion channel. The conversion of mechanical energy into an electrochemical signal (mechanotransduction) underlies our senses of hearing, proprioception and touch amongst many other important processes. If our goal is to

understand biology then, it is necessary to measure the forces that cells generate and respond to.

Sensory hair cells are an important and widely studied example system for mechanotransduction. The hair cells located in the cochlea and vestibular system are responsible for the ability of vertebrates to sense sound and head orientation, respectively. Hair cells are exquisitely sensitive and fast, transducing nanometer- and piconewton-scale deflections and forces into ionic currents at the microsecond scale. The speed of mechanotransduction varies between organisms and determines the upper frequency limit of hearing. Examples of upper frequency limits include turtles (1 kHz), bullfrogs (3 kHz), most birds (7-12 kHz), humans (23 kHz), dogs (45 kHz), mice (91 kHz) and bats (110 kHz) [3].

While the measurement of small, fast forces is a general challenge in mechanobiology and molecular force spectroscopy, we will focus on hair cells because the fast force probes developed in this thesis were specifically designed for the study of hair cells. Beyond biology there are numerous applications for high bandwidth scanning probes with pN-scale force resolution (e.g. magnetometry [4]) that the probes developed in this thesis could be used for with little to no modification.

In this section we will briefly review hair cell mechanics and physiology to motivate the experiments that the fast force probes were designed for; more in-depth discussion and background information can be found in Ref. [5]. We will focus on two specific hair cell experiments that would benefit from faster measurements: bundle mechanics and channel kinetics.

The structure of the human ear is shown in Figure 1.1. Incoming sound waves vibrate the tympanum (labeled 't' in Figure 1.1), which is mechanically coupled to the middle ear. The middle ear consists of three bones, the malleus (m), incus (i) and stapes (s), which couple the tympanum vibration to the inner ear. The middle ear also serves as an acoustic impedance transformer between the low impedance air in the outer ear and the high impedance fluid in the cochlea. Movement of the stapes initiates a traveling wave in the cochlea that induces motion in the basilar membrane, which is where the cochlear hair cells are located. The stapes presses on the oval window and deformation of the round window (rw) enables the fluid flow within the cochlea. The location of maximum vibration varies with frequency, with the highest frequencies nearest to the oval window. In other

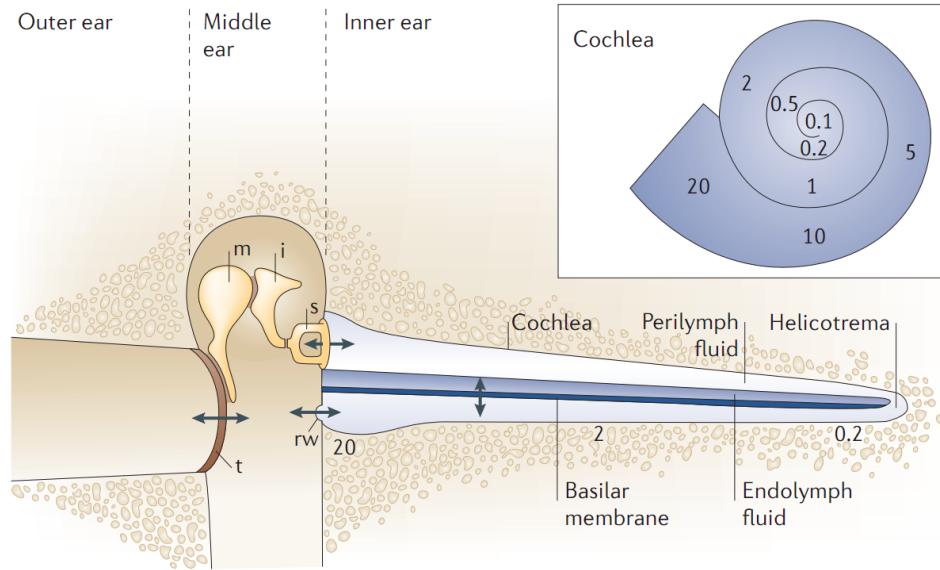


Figure 1.1: The structure of the ear. Sound waves vibrate the tympanum and are relayed to the cochlea through the middle ear, which acts as an acoustic impedance transformer between the air in the outer ear and fluid in the inner ear. Movement of the stapes initiates a propagating pressure wave in the cochlea. The cochlear hair cells sit on the basilar membrane. (inset) The location of maximum vibration varies with frequency (in kHz), with the highest frequencies nearest to the middle ear. Reprinted from Ref. [6] with permission from Nature Publishing Group.

words, the cochlea functions as a spatially-multiplexed spectrum analyzer.

The hair bundle of a turtle cochlear hair cell is shown in Figure 1.2a. The hair bundle consists of many stereocilia arranged into rows. Deflection of the bundle towards the tallest row opens mechanosensitive ion channels located at the tip of each stereocilium. The mechanotransducer currents evoked by a step deflection in a turtle hair cell are shown in Figure 1.2b. The response appears to be band-pass filtered; the initial current rise is limited by the bandwidth of the hair cell while the current drops over time due to adaptation processes within the bundle.

A cross-section through the cochlea is shown in Figure 1.3, illustrating the location of hair cells within the cochlea. Basilar membrane vibration is transduced into hair bundle deflection by the tectorial membrane, which the hair bundles press against or have their tips embedded in. The relative motion between the hair cell body and tectorial membrane

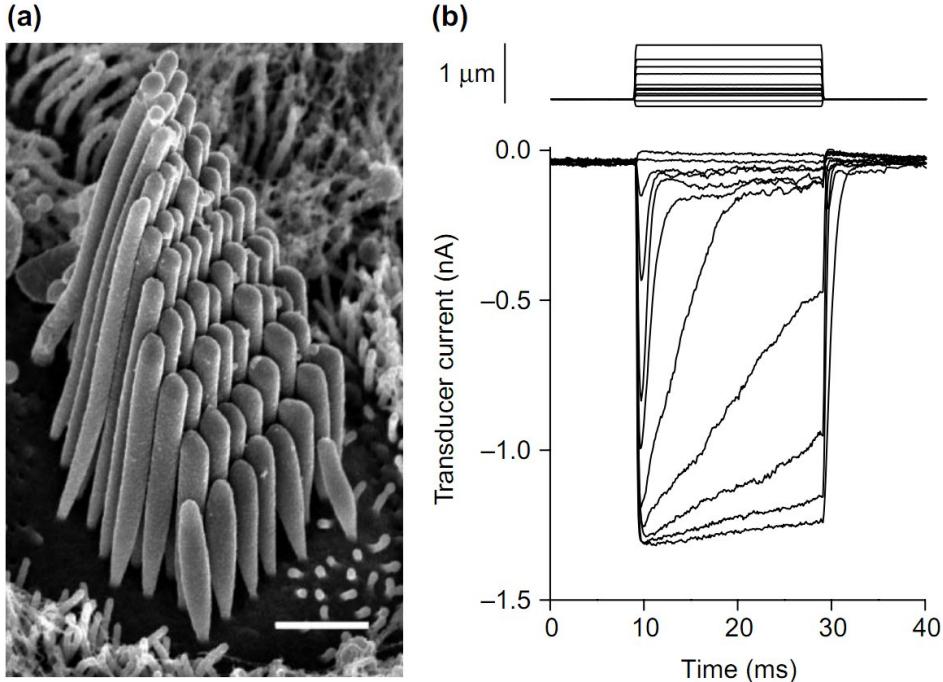


Figure 1.2: Mechanotransduction in cochlear hair cells. (a) Scanning electron micrograph showing the hair bundle of a turtle hair cell. The height of the stereocilia increases between rows. The scale bar is 1 μm . (b) Mechanotransducer currents in response to a step displacement for a voltage-clamped turtle hair cell. Deflections towards the tallest hair bundle row (positive) evoke an inward current. The current rapidly adapts to a constant force, although the adaptation rate depends the amplitude of the bundle deflection. Reprinted from Ref. [7] with permission from Elsevier. ©2001 Elsevier.

deflects the bundles laterally, opening the mechanotransduction channels located in the bundle stereocilia. The open channels allow potassium and calcium ions to enter the cell. Cell depolarization opens additional voltage-gated channels, eventually leading to neurotransmitter-release.

In the mammalian cochlea, hair cells are arranged into three rows of outer hair cells (OHCs) and one row of inner hair cells (IHCs). Other vertebrates (such as bullfrogs and turtles) only have a single type of cochlear hair cell but can have many hair cell rows. Sound perception is driven by IHCs while OHCs contribute frequency tuning and signal amplification to the system. The OHCs can couple electrical signals back into mechanical motion through prestin, a voltage-dependent motor protein in the cell body, and

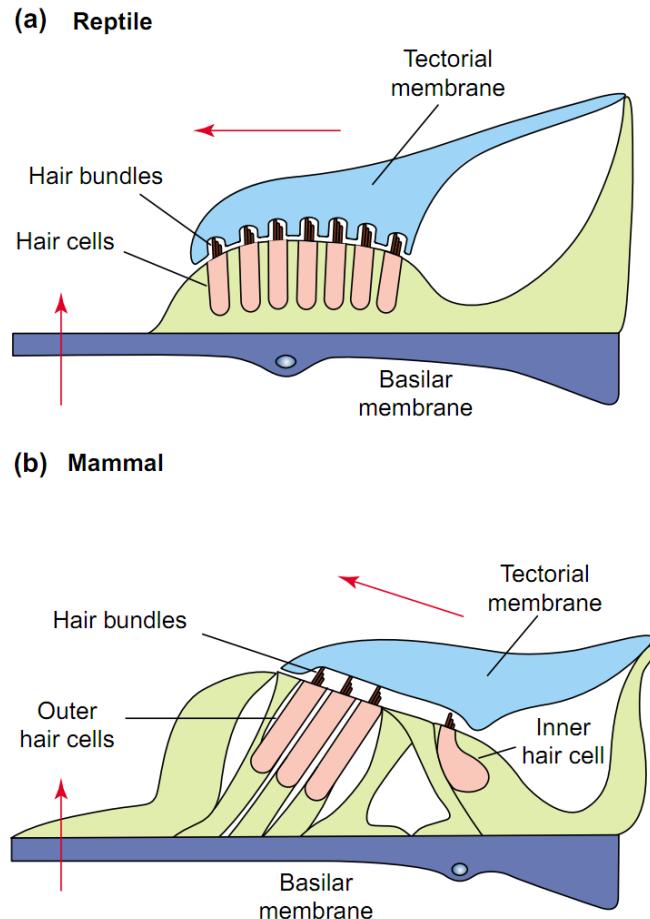


Figure 1.3: The arrangement of hair cells within the cochlea. (a) Reptiles only possess one type of hair cell while (b) mammals have both inner and outer hair cells. Both types of hair cells have distinctive stereocilia bundles that protrude from their top surfaces. Sound-induced vibration the basilar membrane (indicated by the red vertical arrows) deflects the hair cell bundles with respect to the tectorial membrane (indicated by the red horizontal arrows). The bundle deflection opens mechanosensitive ion channels in the bundle tips, converting the mechanical sound pressure into an electrochemical signal. Reprinted from Ref. [7] with permission from Elsevier. ©2001 Elsevier.

through motor proteins in the stereocilia that are coupled to channel gating. The relative contributions of prestin and channel gating is controversial, and is one of the applications where higher-speed force measurements would be helpful [6, 7].

The structure of the hair cell bundle is shown in more detail in Figure 1.4. Motion of the bundle is transduced into the opening of ion channels due to the height difference between

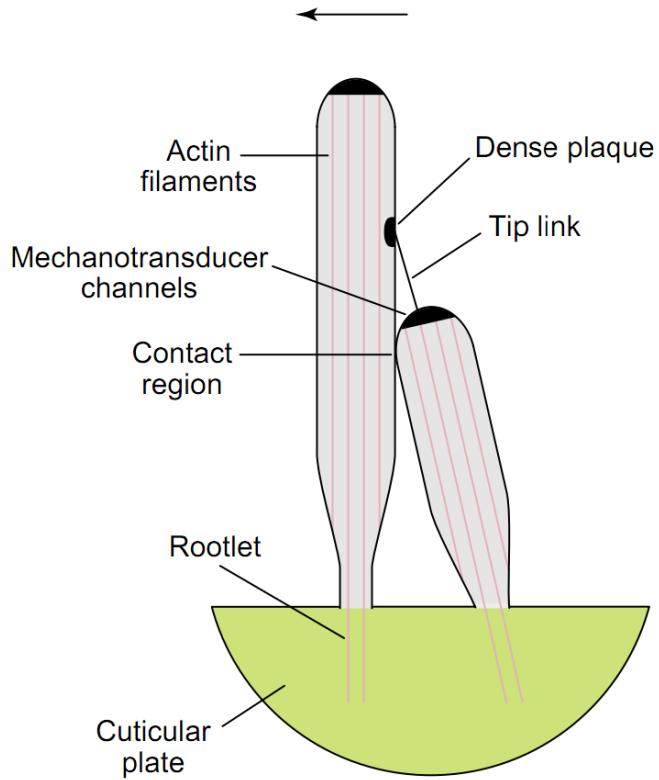


Figure 1.4: Diagram of mechanotransduction in a pair of stereocilia. Deflection towards the taller stereocilium opens the mechanotransducer channels, located at the tip of each stereocilium, due to their height difference. The stereocilia rotate about their bases in response to a load because they are packed with a dense network of actin filaments, making them relatively stiff. The upper end of the tip link is attached to myosin motor proteins, which move along the actin filaments to allow the adaptation to continuous loads. Reprinted from Ref. [7] with permission from Elsevier. ©2001 Elsevier.

stereocilia rows and the tip links that connect them. The mechanotransduction channels are located in the tip of each stereocilium. There can be up to six channels per tip link with more than 200 channels per bundle. The conductance of a single channel is fairly high ($\approx 100 \text{ pS}$), yielding peak mechanoreceptor currents on the order of 1 nA assuming 200 channels and a resting potential of -80 mV [8]. The height and diameter of the stereocilia as well as the number of rows varies between species and tonotopically within the cochlea, with hair cells nearest to the incoming pressure wave sensing the highest frequencies.

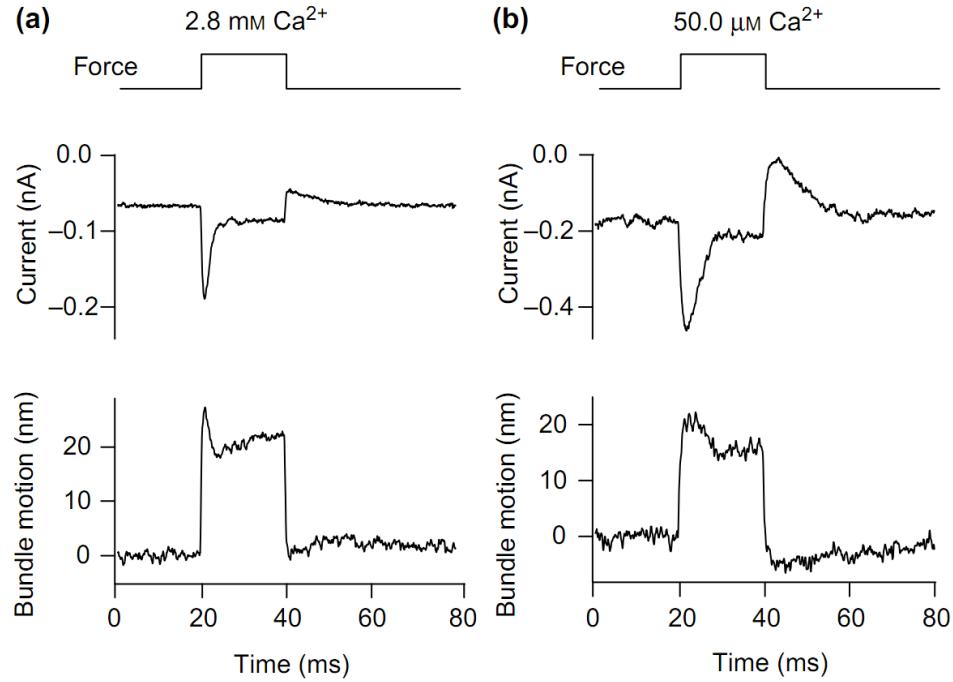


Figure 1.5: Adaptation and force generation processes. (a) A force step of approximately 28 pN is delivered to a turtle hair cell bundle via a flexible fiber. Mechanoreceptor currents adapt to the continuous load. Bundle deflection initially overshoots the steady-state position due to variation in the bundle stiffness over time. (b) Reducing the concentration of calcium ions in the surrounding fluid from 2.8 mM to 50 μ M slows current adaptation and reduces both the speed and magnitude of the mechanical adaptation. Reprinted from Ref. [7] with permission from Elsevier. ©2001 Elsevier.

Hair cell mechanoreceptor currents adapt to static loads through fast and slow adaptation processes [8, 9]. Fast adaptation is due to ion channel inactivation, while slow adaptation is due to the movement of molecular motors. In the case of fast adaptation, incoming calcium ions bind to the transduction channel and stabilize the closed channel conformation, shifting the open probability distribution of the channel. Slow adaptation is caused by the continuous climbing of motor proteins within the stereocilium that maintain tension in the tip link.

1.1.1 Bundle mechanics

The first measurement that would benefit from faster stimulus delivery is measuring bundle mechanics. Hair cell mechanics also adapt to loads. In Figure 1.5 a flexible glass fiber applies an approximately 28 pN step force to a turtle hair cell bundle. The load induces approximately 20 nm deflection of the bundle, measured by projecting the image of the flexible fiber onto a pair of photodiodes and assuming that the fiber remains in contact with the bundle. The deflection of the bundle is initially quite high and decays to its steady-state position as the channels close, suggesting that changes in channel conformation are coupled to changes in bundle mechanics. Reducing the concentration of calcium ions in the surrounding fluid from 2.8 mM to 50 μ M slows current adaptation and reduces both the speed and magnitude of the active component of the bundle mechanical response.

Figure 1.6 illustrates the conventional approach to measuring bundle mechanics and variation in bundle mechanics with time in more detail. As seen in Figure 1.6a, a silicone bead is attached to the tip of a flexible glass fiber. A small bead (\approx 3 micron diameter) is typically used to interface with the V- or W- shaped bundle of outer hair cells. The shape and size of the bundles vary tonotopically throughout the cochlea due to the variation in sound frequencies that the bundles are optimized to detect. The glass fiber is mounted on a macroscale piezoelectric actuator, enabling micron-scale movement of the fiber. The probability that one of the several hundred ion channels in the bundle is open varies with the deflection of the bundle (Figure 1.6b). The response typically saturates for a deflection on the order of 500 nm, saturating at slightly smaller deflections when the stimulus is applied more rapidly. As noted earlier, fiber and bundle motion is measured by projecting the image of the fiber or bead onto a position sensitive photodiode.

A flexible fiber with stiffness comparable to the bundle (\approx 1 pN/nm) is used in order to allow the applied force and deflection of the bundle to be simultaneously measured. Force is calculated by calibrating the spring constant of the fiber and calculating the fiber deflection from the difference between the motion of the piezoelectric stack and the tip of the flexible fiber. However, the small spring constant and large mass of the bead and fiber leads to the first-resonant mode of the system having a fairly low frequency. The fastest flexible fiber system reported to date had a -3 dB bandwidth on the order of 2 kHz in water,

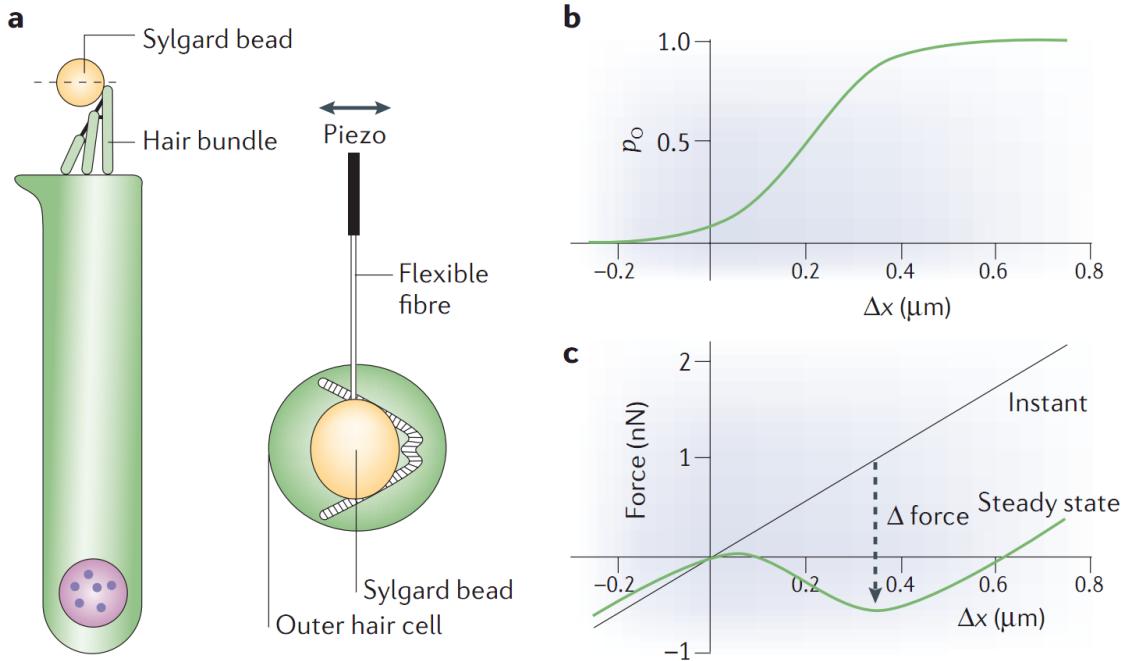


Figure 1.6: Measuring bundle mechanics. (a) A flexible glass fiber is mounted on a piezoelectric actuator with an ≈ 3 micron diameter silicone bead on its tip. The piezoelectric drive signal is filtered at 1.5-2 kHz in order to avoid exciting the first-resonant mode of the fiber. (b) Channel opening probability varies with bundle deflection and saturates for a roughly 500 nm deflection. (c) Bundle mechanics vary with time. When the bundle is deflected faster than the channel gating latency, the passive mechanical response of the bundle dominates. When measured more slowly, both the passive and active components of the bundle mechanical response are measured. Reprinted from Ref. [6] with permission from Nature Publishing Group.

yielding a 10-90% rise time of 150 μ s [10].

The bundle mechanics vary with both time and bundle deflection (Figure 1.6c). When measured on a timescale shorter than channel gating, the bundle appears to be a passive mechanical element. But when measured more slowly, the bundle stiffness varies with deflection due to active processes. In Figure 1.6c a region of negative stiffness is shown, which would lead to positive-feedback and mechanical amplification. Whether or not the effective stiffness of the bundle becomes negative, the bundle stiffness varies over time due to mechanical adaptation which could contribute to the cochlear amplifier. The other amplification mechanism (prestin in the cell body) is based upon changes in cell voltage,

so is filtered by the capacitance of the cell membrane at approximately 1 kHz and thus may become less important as the frequency of the incoming sound wave increases [10].

Bundle mechanics are typically modeled as a combination of passive and active processes according to

$$F = xk_p - Ap_o(x, t) + F_0 \quad (1.1)$$

$$k = k_p - A \frac{\partial p_o(x, t)}{\partial x} \quad (1.2)$$

where k_p is the passive bundle stiffness, A is a scaling constant and p_o is the open probability of the channel, which varies with both bundle deflection and time due to adaptation processes. Channel gating generates a force in the same direction as the applied force, leading to variation in bundle stiffness with deflection and time. Bundle stiffness reaches a minimum when the channel open probability varies most rapidly with deflection (Figure 1.6b and c).

The channels in turtle and frog OHCs are slow enough to enable the instantaneous measurement of bundle stiffness with a flexible fiber, but mammalian OHCs, which operate to much higher frequencies, are faster than conventional measurement techniques. The maximum frequency at which bundle force generation can contribute to signal amplification within the mammalian cochlea is unclear due to the limited speed of the flexible fiber approach.

1.1.2 Channel kinetics

The second measurement that would benefit from faster stimulus delivery is measuring ion channel kinetics (Figure 1.7). As noted earlier it takes a finite amount of time for the ion channels in the bundle to open (activation), after which they immediately begin to close due to the binding of calcium ions to the channel and the movement of molecular motors (adaptation). If stimuli are delivered to the bundle much faster than the activation and adaptation rates of the hair cell then the channel kinetics can be modeled. While the kinetics of bullfrog and turtle hair cells have been measured [12, 13], mammalian hair cells

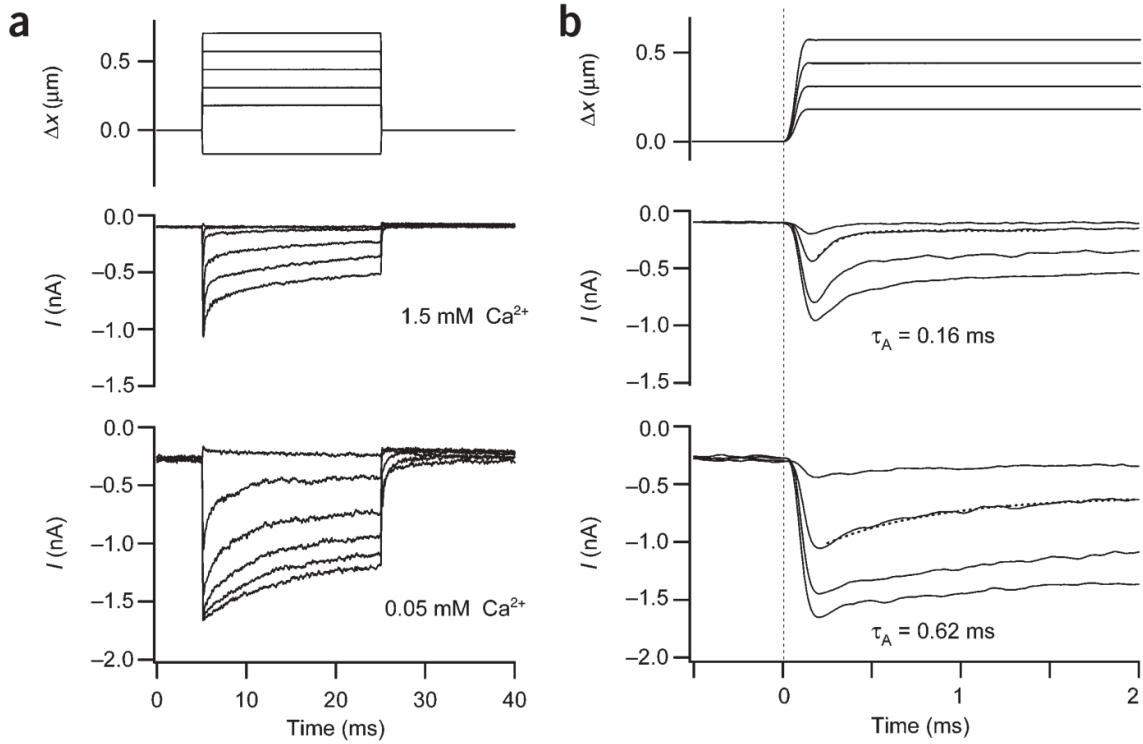


Figure 1.7: Measuring channel kinetics. Bundles are deflected using a stiff glass probe mounted on a piezoelectric stack. The increased stiffness of the probe allows the actuation signal to be filtered at 5–10 kHz, yielding probe rise times of 50–100 μs . Probe deflection is shown in the first row and mechanoreceptor current is shown in the second and third rows for high (1.5 mM) and low (50 μM) concentrations of calcium ions. The first column (a) shows the entire response of the cell while the second column (b) focuses on the stimulus onset. The speed of channel activation is limited by the speed of stimulus delivery ($50 \pm 10 \mu\text{s}$). Reprinted from Ref. [11] with permission from Nature Publishing Group.

are significantly faster and their activation kinetics have not been successfully measured to date.

A stiff glass probe is typically used to stimulate the hair bundle when focusing on channel kinetics. If the probe is much stiffer than the bundle then the probe does not deflect during stimulus delivery and the bundle tracks the motion of the piezoelectric actuator. While bundle mechanics can not be measured accurately in this configuration, it enables the actuator signal to be filtered at a higher frequency. For example, the actuator signal was filtered at 5–10 kHz in Ref. [11], yielding a 10–90% rise time of 50–100 μs .

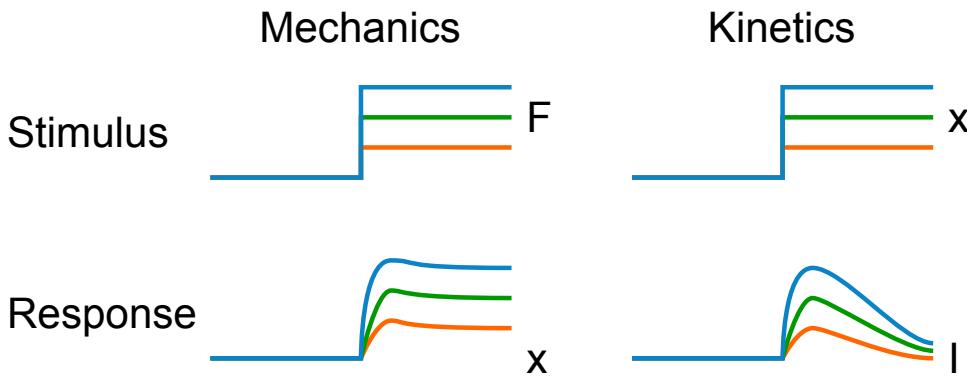


Figure 1.8: Applications of fast force probes to cochlear hair cells. There are two immediate cochlear hair cell measurements that would benefit from faster stimulus delivery: mechanics and kinetics. Both measurements have been performed on frog and turtle hair cells, but mammalian hair cells are faster than conventional methods that employ macroscale glass probes. Increasing the rate of stimulus delivery would enable new cochlear hair cell science.

However, mammalian hair cell activation is faster than 50 μs and only hair cell adaptation kinetics could be extracted from those measurements. The ideal force probe for kinetics measurements would have a much faster rise time than the hair cells being studied, ideally on the order of 1-10 μs . In order to measure such rapidly changing mechanoreceptor currents the speed of the patch clamp would need to be increased as well.

In summary, there are two immediate cochlear hair cell measurements that would benefit from faster stimulus delivery: mechanics and kinetics (Figure 1.8). Active mechanical processes in mammalian OHCs contribute to cochlear amplification but it is unclear what frequency active bundle mechanics can operate up to. Existing methods have used flexible glass fibers and have been limited to rise times of approximately 150 μs . Channel activation has been measured at higher speeds (50 μs) but stimulus delivery remains too slow to extract the channel kinetics. In both cases a bundle deflection of less than 1 micron is required to saturate the hair cell response. Other measurements include measuring prestin-generated OHC forces and deflections, but bundle mechanics and channel kinetics have been the driving force behind the development of a faster force probe.

Table 1.1: Summary of probe performance requirements. The main difference between the mechanics and kinetics experiments is the maximum probe stiffness. In the case of mechanics, the stiffness of the probe should roughly match the stiffness of the hair cell bundle in order to balance force and displacement resolution for the calculation of bundle stiffness. The mechanics probe should have as high of a resonant frequency as possible while satisfying the stiffness constraint. There is no stiffness constraint in the case of measuring channel kinetics - the probe should simply apply the stimulus as rapidly as possible. The rise times to date are from Refs. [10] and [11].

	Mechanics	Kinetics
Hair cell type(s)	OHCs	OHCs and IHCs
Probe stiffness (pN/nm)	1-5	10-100
Minimum probe deflection (nm)	300	300
Best rise time to date (μ s)	150	50
Target rise time (μ s)	20-50	1-10
Displacement resolution (nm)	1-10	1-10
Force resolution (pN)	1-100	1-100

1.2 Force probe design requirements

In this section we will discuss the performance and design requirements for an improved high speed force probe. We will focus on the hair cell measurements, but want to emphasize that numerous other applications in biology and surface science could benefit from fast, soft force probes. Table 1.1 summarizes the performance of existing hair cell probes and the ideal performance of a new force probe. In particular, the spring constant of a probe designed for studying bundle mechanics should closely match the stiffness of the bundle (1-5 pN/nm) in order to balance the force and displacement resolution of the system. Mechanics experiments have been performed up to approximately 2 kHz to date ($\approx 150 \mu$ s rise time) [10]. Decreasing the rise time to 20-40 μ s, corresponding to a resonant frequency of roughly 8-16 kHz, should provide new information about the upper speed limits of active bundle processes.

We will relate the 10-90% rise time of a system to its resonant frequency several times in this chapter and repeatedly throughout the rest of the thesis. The rise time of a second-order system can be calculated from its resonant frequency (f_0) and quality factor (Q) as

$$t_r = \frac{1.76\zeta^3 - 0.417\zeta^2 + 1.039\zeta + 1}{2\pi f_0} \quad (1.3)$$

where $\zeta = 1/2Q$. The rise time calculated from (1.3) is accurate to 0.5% for ζ values ranging from 0 to 0.9 [14]. The rise time of a critically damped system ($Q = 1/\sqrt{2}$) is roughly equal to $1/3f_0$. For the force probes developed in this thesis, the quality factor varies from approximately 2 to 10 in air and 0.5 to 1 in water. In most situations we will conservatively approximate the rise time of a system as $t_r \approx 1/3f_0$. In the limit of $Q \rightarrow \infty$ the rise time approaches $1/6f_0$ so our conservative assumption is off by a factor of two at most.

Kinetics measurements are not limited by probe stiffness because the only goal is to stimulate the bundle faster than the ion channels can open. Increasing the probe stiffness allows for faster stimulus delivery. However, in order to increase the speed of stimulus delivery the length of the probe must be decreased. A shorter probe leads to less deflection at the tip, and our minimum target deflection of 300 nm prevents the use of extremely short force probes. A rise time of 1-10 μ s translates to a resonant frequency in water of roughly 30-300 kHz. Due to viscous damping, these correspond to resonant frequencies in air of 50-500 kHz depending on the geometry of the probe.

The force and displacement resolution of the probe should be on the order of 1-100 pN and 1-10 nm for both measurements. Achieving that level of force resolution becomes increasingly difficult as the measurement bandwidth increases due to the simultaneous increase in the probe spring constant and integrated noise. Both measurements require stimulus delivery at relatively high frequencies without any spurious resonant modes below the resonant frequency of the probe, making it necessary to integrate the probe with the actuator rather than relying on an external, macroscale actuator.

The typical experimental setup for studying cochlear hair cell physiology is shown in Figure 1.9. A section of the cochlea is excised from an animal, or hopefully soon, produced from embryonic or induced pluripotent stem cells as in Ref. [15]. The tissue is mounted in a chamber on an upright microscope in order to enable the continuous perfusion of a salt solution. The salt solution is intended to mimic the perilymph solution within the intact cochlea in order to keep the tissue alive for several hours. Experiments can only be

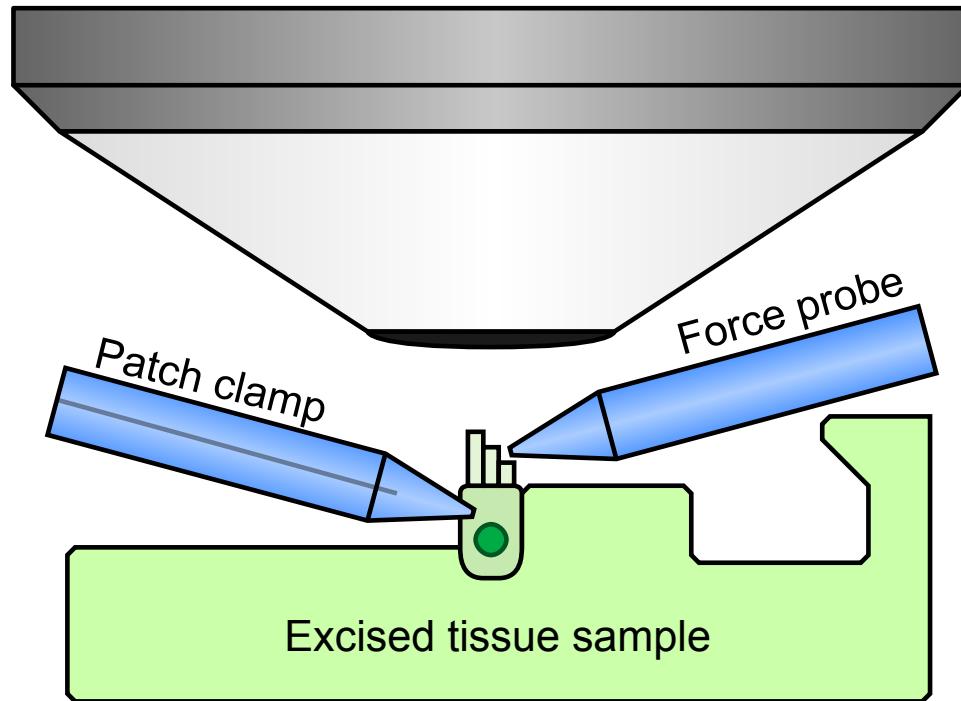


Figure 1.9: Hair cell experimental setup. A section of the cochlea is excised and mounted on an upright microscope in a perfusion chamber. The chamber is continuously perfused with a salt solution (not shown). The mechanoreceptor currents are recorded with a patch clamp electrode in a whole-cell voltage clamp configuration. The bundle is stimulated with a force probe, depicted as a stiff glass fiber in this case. The force probe must approach from a 10-15° angle in order to avoid crashing into the tissue and the microscope objective. A 60x or 100x objective is typically used with a working distance of 1-2 mm.

performed while the tissue is alive and in a physiologically relevant state [5]. After patching onto the hair cell in a whole-cell voltage clamp configuration, stimuli can be delivered to the bundle. The salt solution is grounded using a reference electrode (e.g. Ag/AgCl) in order to control the voltage of the cell with respect to the solution.

Experiments need to be performed on an upright rather than inverted microscope due to the presence of the tissue, making it significantly more challenging to fit the patch clamp pipette and force probe into the experimental setup. The magnification and working distance of the objective are usually 40-100x and 1-3mm, respectively. Depending on whether mechanics or kinetics are being studied the force probe may be a flexible glass fiber or stiff glass probe. If OHCs are being studied, the probe may need to have a 2-5

micron diameter bead attached to its tip to fit into the V- or W-shaped bundle.

The experimental setup adds several constraints to the design of our fast force probe. First, the probe needs operate in a grounded salt solution without corroding. Second, the probe needs to electrically insulated from the solution and should inject less than 500 pA of peak current through capacitive coupling in order to avoid interfering with the mechanoreceptor current measurements (Figure 1.2). Third, the tip of the probe should be flat for IHC measurements and round for OHC measurements. Fourth, the silicon die that the force probe is attached to needs to be relatively long and narrow (at least 5-10 mm long and less than 1 mm wide) in order to fit between the microscope objective and tissue sample. Fifth, the probe needs to allow for both actuation and sensing at frequencies of up to 10-20 kHz for mechanics measurements and 100-200 kHz for kinetics

1.3 The limits of existing force probes

We can examine conventional force probes based upon these requirements. Devices sensitive enough to measure the force required to break a single hydrogen bond, to stall a single molecular motor, and to flip the spin of a single electron have been developed [16–18]. The main issue limiting their application to cochlear hair cells is not force or displacement resolution, but the tradeoff between spring constant and time resolution.

Figure 1.10 plots the probe spring constant against time resolution for optical tweezers (OT), magnetic tweezers (MT), atomic force microscope cantilevers (AFM) and the macroscale probes used for most hair cell measurements to date (M). The range of performance goals discussed in the last section is plotted for comparison. Optical tweezers, magnetic tweezers and atomic force microscope cantilevers have been applied to measure the mechanical properties of cells in hundreds of papers and we quote performance ranges from two recent review papers for succinctness [19, 20]. The definition of time resolution varies between papers, but we will generally equate time resolution to the 10-90% rise time of the measurement system, calculated from (1.3).

Time resolution and spring constant are directly related. Each of the measurement techniques in Figure 1.10 can be approximated as an undamped simple harmonic oscillator (SHO) in order to calculate resonant frequency as

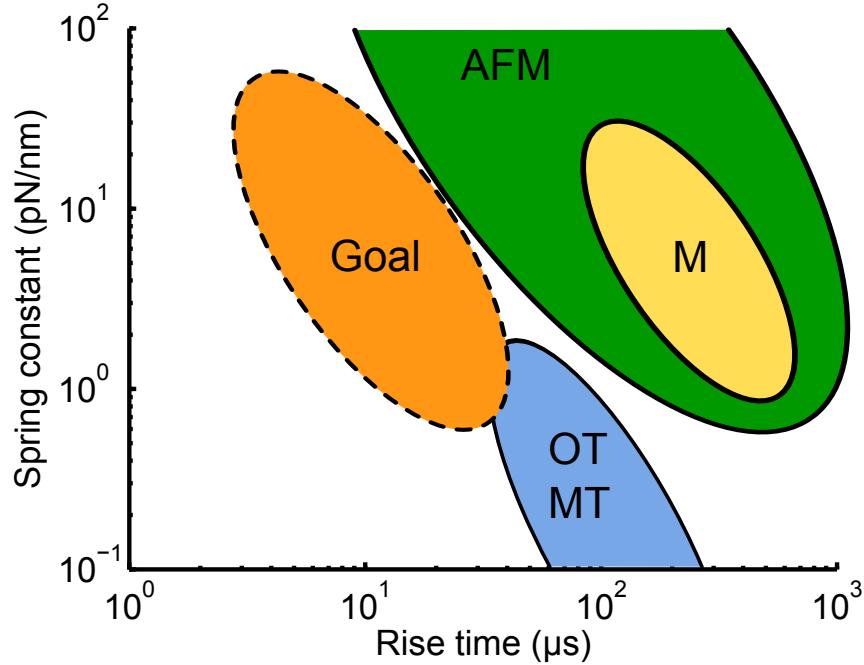


Figure 1.10: Existing force measurement techniques do not satisfy the spring constant and time resolution combination needed for the study of cochlear hair cells. Optical tweezers (OT), magnetic tweezers (MT), atomic force microscope cantilevers (AFM) and macroscale probes (M) are plotted against the cochlear hair cell application requirements for comparison.

$$f_0 = \frac{1}{2\pi} \sqrt{\frac{k_c}{m_{\text{eff}}}} \quad (1.4)$$

where k_c is the spring constant and m_{eff} is the effective mass of the system. Substituting (1.4) into (1.3) we can calculate the rise time as

$$t_r \propto \sqrt{\frac{m_{\text{eff}}}{k_c}}. \quad (1.5)$$

In order to decrease the rise time either the mass needs to decrease or the spring constant needs to increase.

Optical and magnetic tweezers are limited primarily by the maximum trap stiffness. Trap stiffness scales with laser light intensity, numerical aperture, bead dielectric constant and bead size in the case of optical tweezers and the magnetic field gradient, bead

permittivity and bead size in the case of magnetic tweezers [21]. While bead mass and spring constant are coupled, the maximum stiffness is limited to less than 1 pN/nm in both cases [20].

Atomic force microscope cantilevers are limited by their minimum size. More precisely, micromachined cantilevers that detect forces optically have a minimum possible size of 5-10 μm due to the spot size of the laser beam bounced off of the cantilever [22]. While their dimensions can be reduced further through the addition of metallized light reflecting paddles [17], micromachined cantilevers with the spring constant and resonant frequency combination needed to study cochlear hair cells have not been produced to date.

We can analyze the cantilever beam geometry in further detail. The spring constant and effective mass of an end-loaded cantilever beam can be calculated from

$$k_c = \frac{E_c w_c t_c^3}{4l_c^3} \quad (1.6)$$

$$m_{\text{eff}} = 0.243 \rho_c l_c w_c t_c. \quad (1.7)$$

where E_c and ρ_c are the elastic modulus and density of the cantilever beam and l_c , w_c and t_c are its length, width and thickness, respectively.

The product of spring constant and rise time, which we would like to minimize, scales as

$$k_c t_r \propto \frac{\sqrt{E_c \rho_c} w_c t_c^2}{l_c} \quad (1.8)$$

revealing that the $k_c t_r$ product scales as $[L]^2$ where $[L]$ is the overall length scale of the device. This model of cantilever mechanics, while fairly simple, captures the key challenge in improving force probe performance. Second order details, such as fluid damping and mesoscale effects on cantilever material properties, will be discussed in the following chapters.

From this analysis we can determine what the ideal force probe would look like (Figure 1.11). From (1.5) we know that it should have the smallest possible size and should allow the spring constant to be adjusted over a wide range. The smallest possible probe would

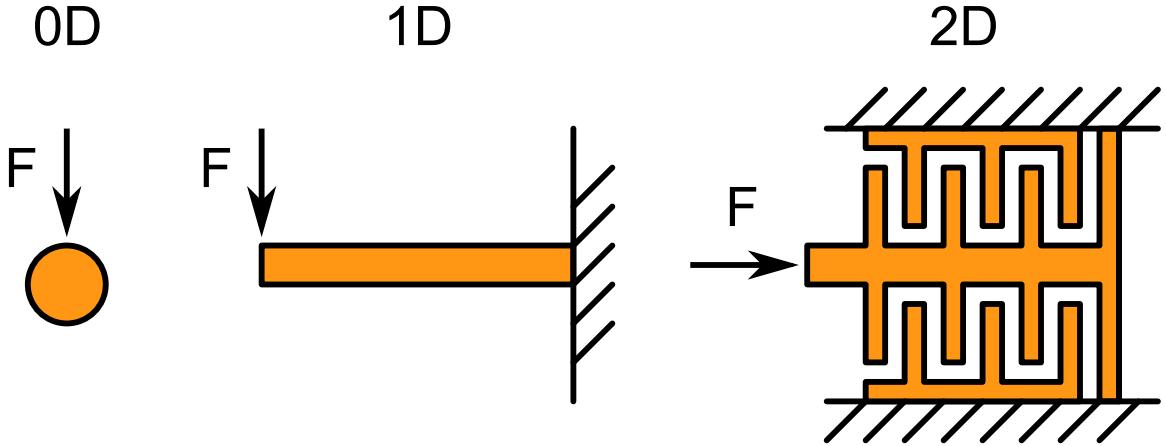


Figure 1.11: The ideal force probe is a one-dimensional beam. While a 1D force probe is more massive than a 0D probe (i.e. a bead), the length of the 1D probe can be adjusted in order to yield a large range of spring constants. Shifting the probe from 1D to 2D (e.g. to add tines for capacitive sensing and actuation) adds additional mass to the system, reducing system performance.

be infinitesimally small in all three dimensions, i.e. a bead. But as we have discussed, while the mass of a bead can be made quite small, its maximum possible spring constant is generally limited because it can not be adjusted geometrically.

Adding an additional dimension to the probe allows the spring constant and mass of the probe to be adjusted over an arbitrary range. Increasing the dimensionality of the probe any further, such as the addition of tines for capacitive sensing and actuation, only increases the mass of the probe further and necessarily increases the k_{ct_f} product. We would argue that the cantilever beam is the smallest possible geometry that allows the geometric adjustment of probe spring constant; any other dimensions or any other shape adds mass to the system which reduces the performance envelope of the force probe. We will focus our attention on the cantilever beam geometry for the remainder of the thesis.

We can quickly estimate how small the cantilever would need to be in order to satisfy the cochlear hair cell spring constant and rise time requirements. Figure 1.12 plots the spring constant as a function of the rise time for silicon cantilever beams of varying thickness. The stiffness and rise time are calculated from (1.3) and (1.6) for a single crystal silicon beam oriented in the $\langle 100 \rangle$ direction. The beam width in the calculations is equal to twice the beam thickness (in order to suppress lateral oscillations) and the length is varied

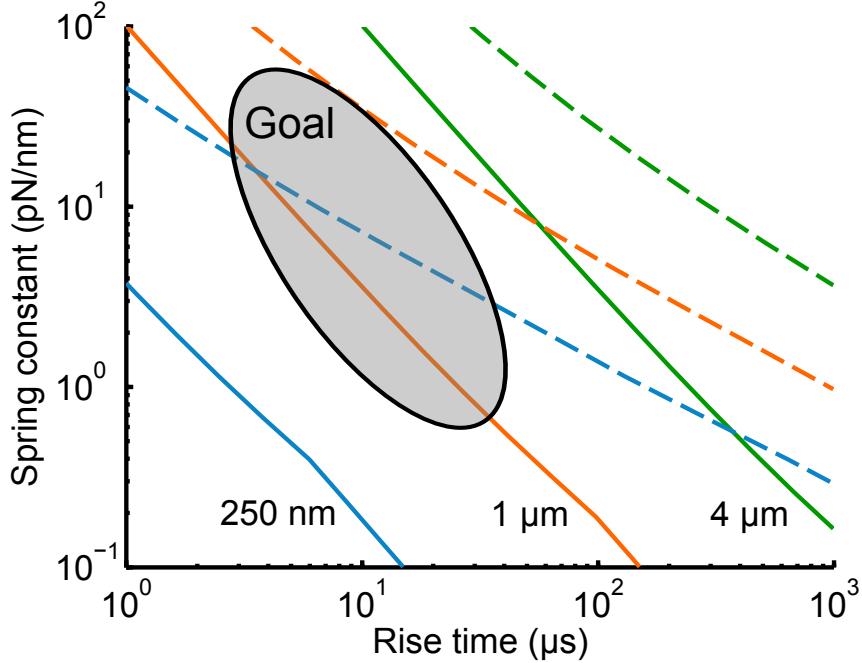


Figure 1.12: Stiffness and rise time for cantilever beams of varying thickness in air (solid lines) and water (dashed lines). While a force probe for studying channel kinetics ($t_r \approx 1 \mu\text{s}$ and any spring constant) could be fabricated from any beam thickness, a bundle mechanics probe ($t_r < 50 \mu\text{s}$, $k_c \approx 1 \text{ pN/nm}$) would require a roughly 250 nm thick beam. Resonant frequencies and quality factors in both air and water are calculated from Ref. [23].

in order to sweep the possible combinations of rise time and spring constant. The resonant frequencies are calculated assuming operation in either air or water using the lookup tables in Ref. [23]. We will discuss the dynamics of cantilever motion in water in more detail in the next chapter, but for now, we simply want to illustrate the effect of water operation on rise time.

While a force probe for studying channel kinetics ($t_r \approx 1 \mu\text{s}$ and any spring constant) could be fabricated from any beam thickness, a bundle mechanics probe ($t_r < 50 \mu\text{s}$, $k_c \approx 1 \text{ pN/nm}$) requires a relatively thin beam in order to satisfy both design constraints. The results indicate that if the cantilever were operating in air the cantilever beam could be up to 1 μm thick, but water operation requires a thinner, shorter cantilever due to the increased damping and mass loading.

In summary, the ideal force probe would be as small as feasibly possible and

manufactured from a material with the smallest possible product of elastic modulus and density. A cantilever beam geometry allows the spring constant adjustment over a wide range without any excess mass. If the cantilever beam is fabricated from silicon, it would need to be roughly 250 nm thick and 500 nm wide in order to meet the spring constant and rise time requirements of the bundle mechanics application. In the next chapter we will expand upon the force probe design in much more detail.

1.4 Microscale transduction techniques

In this section we will briefly review the most common microscale sensing and actuation techniques. We will discuss the performance, scaling, and limitations of each technique in the context of the cochlear hair cell experiments. Performance will mostly be discussed qualitatively because we have not discussed models for noise, sensitivity and resolution yet. We hope to provide at least an imperfect understanding of the relative strengths and weaknesses of each technique in order to motivate the design used for the fast force probe. We refer the interested reader to Refs. [24–27] for more comprehensive reviews.

1.4.1 Force sensing techniques

There are numerous techniques for transducing mechanical loads into an electrical signal. The four most common sensing techniques are optical, piezoresistive, piezoelectric and capacitive [28]. All four techniques can be integrated into micromachined cantilever with varying degrees of complexity.

The optical detection of micromachined cantilever motion was pioneered by Binnig, Quate and Gerber in 1986 [29]. Its predecessor, the scanning tunneling microscope transduced probe-sample spacing from the electron tunneling current between a sharp tip and conductive sample [30]. Shifting to optical transduction allowed the analysis of insulating samples and lead to the widespread use of the atomic force microscope.

Optical detection is still the dominant measurement technique for sensing the deflection of micromachined cantilevers. Deflections can be measured optically using either interferometry or the optical lever technique. For simplicity we will focus on measuring

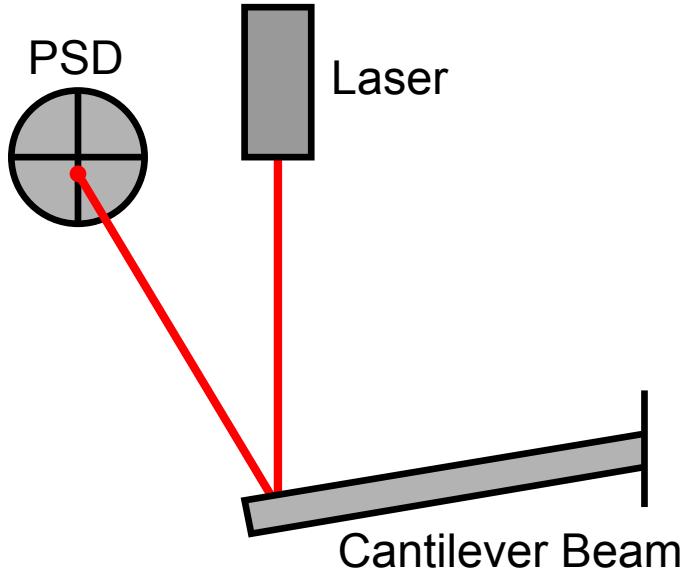


Figure 1.13: Optical transduction of cantilever beam deflection. A laser beam is reflected off of the back of the cantilever and onto a position sensitive photodetector (PSPD). The laser spot is centered onto the PSPD using a mirror (not shown). The PSPD is typically a two- or four-segment reverse biased photodiode, and movement of the laser spot leads to changes in current between the PSPD segments. The current is tranduced to a voltage using a transimpedance amplifier.

the deflection of a cantilever beam via the optical lever principle (Figure 1.13). Both approaches are discussed extensively by Solgaard [31]. By increasing the distance between the cantilever beam and the photodetector the force and displacement resolution of the system can be improved far beyond the limits of simply projecting the image of the cantilever directly onto the photodetector.

Optical detection has two major advantages over the other techniques. First, it yields the best possible displacement and force resolution. Without going into the details yet, the measurement resolution of any sensor is fundamentally limited in the best possible case by Brownian motion, otherwise known as thermomechanical noise [32]. Other noise sources, whether from electronic or mechanical sources, simply add to the thermomechanical noise floor. While the three other sensing techniques (piezoresistive, piezoelectric and capacitive) have certain benefits over optical transduction, they have greater displacement and force noise in most circumstances. The second major advantage of optical detection is that the

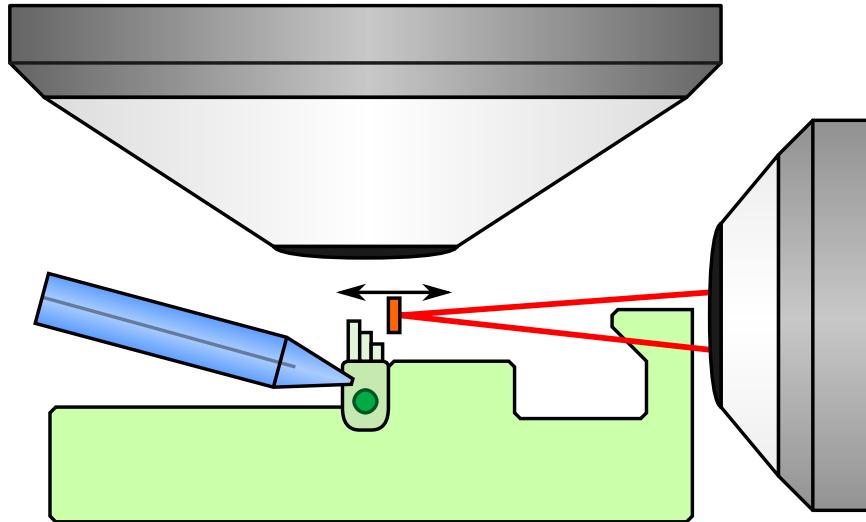


Figure 1.14: The hair bundle must be deflected parallel to the top surface of the cell. Measuring the deflection of the cantilever (shown in cross-section) via laser beam bounce would require the addition of a second, side-view microscope objective. This would pose practical challenges in terms of space and obtaining a clear view around the surrounding tissue. Measuring the cantilever deflection by projecting its image directly onto the photodetector is possible but still challenging and prone to spurious signals.

force probe fabrication would be relatively simple. All of the electronics and complexity in using optical readout is located far away from the cantilever.

The challenge in applying optical detection to cochlear hair cell experiments is that the hair bundles must be deflected parallel to the top surface of the cell (Figure 1.14). This necessitates either adding a second microscope objective to the system for a side-view [33] or projecting the image of the cantilever beam as seen from the top directly onto the photodetector. The first option is challenging in practice and the second option would lead to degraded force and displacement resolution compared to the beam bounce method and other transduction techniques [34].

A second challenge, although not specific to the cochlear hair cell problem, is that the cantilever beam must reflect sufficient laser light. Silicon and silicon nitride, two commonly used materials, are not particularly reflective and a substantial fraction of the incident light will simply pass through the cantilever if it is sufficiently thin ($< 1 \mu\text{m}$ thick). The cantilever can be coated with a thin layer of metal, then the bilayer beam structure

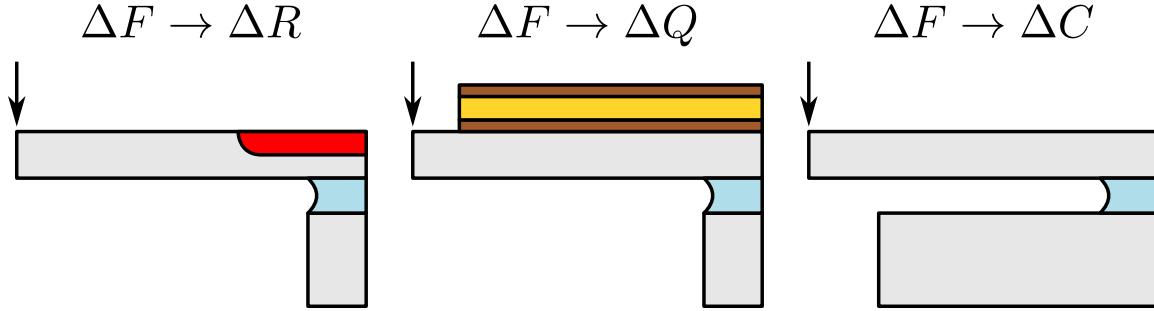


Figure 1.15: All-electronic techniques for transducing cantilever deflection. Piezoresistive, piezoelectric and capacitive sensing techniques transduce a change in the force applied to the tip of the cantilever into a change in the resistance of the piezoresistor, the polarization of the piezoelectric film, or a change in the capacitance between the cantilever and a stationary surface, respectively.

will then deflect in response to changes in temperature due to the coefficient of thermal expansion mismatch. The spot size of most laser beam bounce system is in the order of 5-10 μm while deep sub-micron cantilevers can be patterned using optical or electron beam lithography. A small paddle at the tip of the cantilever is typically used to reflect sufficient light when cantilevers are operated in air or vacuum [4, 17], but the paddle would lead to increased damping for liquid measurements.

It would clearly be desirable to detect the deflection of the cantilever electronically rather than optically. Piezoresistive [35, 36], piezoelectric [37, 38] and capacitive [39, 40] sensing techniques have all been developed to enable the detection of cantilever deflection. The three techniques and typical implementation geometries are summarized in Figure 1.15. Piezoresistive, piezoelectric and capacitive sensing techniques transduce a change in the force applied to the tip of the cantilever into a change in the resistance of the piezoresistor, the polarization of the piezoelectric film, or a change in the capacitance between the cantilever and a stationary surface, respectively.

Capacitive sensors are similar to optical sensors in that their output is proportional to deflection. The capacitance between two parallel plate electrodes is proportional to their overlap area and inversely proportional to their separation distance. A deflection can be coupled to a capacitance change using either interdigitated electrodes or parallel plates (as in Figure 1.15). Capacitance is inversely proportional to electrical impedance, so

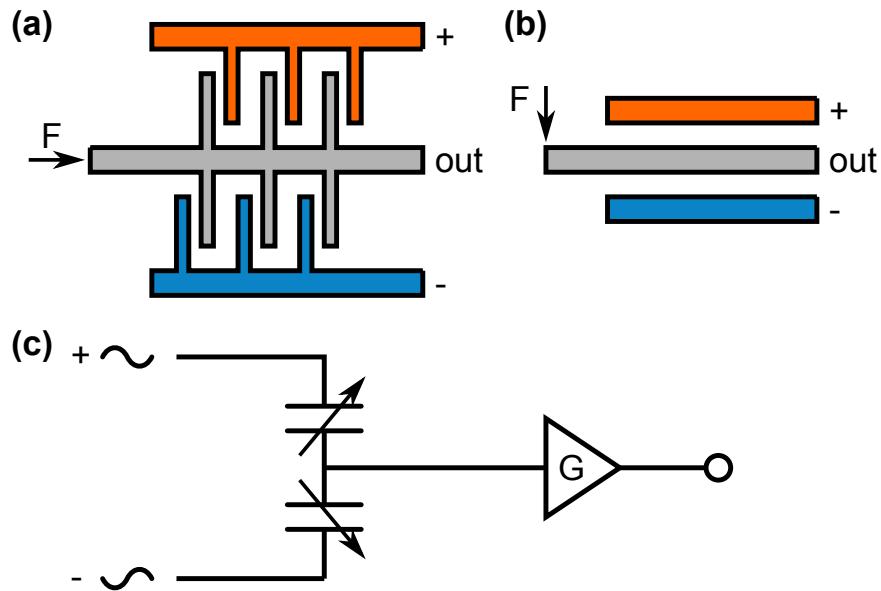


Figure 1.16: Capacitive force sensors typically utilize a differential measurement scheme. (a) Deflection of the middle electrode is coupled to an increase between the middle and top electrode and a decrease in the capacitance between the middle and bottom electrode. The top and bottom electrodes are stationary while the middle electrode is mounted on a set of flexures and the deflection to the right is determined by $x = F/k$. (b) Similarly, a force to a cantilever beam can be transduced from the capacitance change between it and a pair of stationary electrodes located above and below it. Alternatively, a single electrode could be used for two cantilevers: a probe cantilever and a reference cantilever. (c) An AC-bridge is typically used to transduce a differential capacitance change into an output signal. The amplifier output is demodulated and measured to calculate deflections and loads.

deflections can be inferred from impedance changes.

Two examples of the parallel plate arrangement are shown in Figure 1.16. In Figure 1.16a, the middle electrode deflects to the right in response to a load. The deflection increases the capacitance between the middle and top electrode while decreasing the capacitance between the middle and bottom electrode. Biasing the top and bottom electrodes 180° out of phase an AC-bridge is formed and the voltage at the middle electrode is proportional the deflection. Similarly in Figure 1.16b, a cantilever beam is used as the middle electrode and stationary electrodes above and below it are used to form a differential pair of capacitors.

However, we plan to utilize a cantilever beam geometry for the force probe, making the

parallel plate approach (Figure 1.16b) the only feasible option. Further, in order to simplify the fabrication of the device only a single stationary electrode is typically used [39]. A differential measurement is desirable so that parasitic capacitance variations (e.g. due to cable movements) can be nulled out. The baseline capacitance of capacitive cantilevers is extremely small (e.g. 100 fF for a 1 μm initial gap in Ref. [39]) because it scales with cantilever area, and the changes in capacitance with deflection are smaller still. Differential measurements can be performed with a single electrode per cantilever by combining the force probe cantilever with a stationary reference cantilever.

Capacitive sensors have several advantages over piezoresistive and piezoelectric sensing. First, their fabrication is relatively simple because they do not require carefully controlled doping steps or the deposition of exotic films. Second, capacitive transduction can be used for both actuation and sensing by applying an AC sensing bias far above the resonant frequency of the cantilever. Third, capacitive sensors have relatively high electrical impedance and do not typically dissipate much heat. Fourth, the amplifier is the only significant electrical noise source in a capacitive sensor, making it possible to achieve thermomechanical noise limited performance if the capacitor is large enough. Finally, capacitive transduction scales extremely well to measuring forces along multiple axes, as demonstrated by the six-axis force and torque sensor shown in Figure 1.17.

However, there are three major challenges with capacitive sensing that make it unworkable for the cochlear hair cell application.

First, the displacement and force sensitivity of capacitive cantilevers is quite low due to their small capacitances and large impact of parasitics, yielding force resolution on the order of nN-scale [39, 40]. For example, the six-axis force and torque sensor in Ref. [41] had a noise floor of 1.4 μN in a 15 Hz measurement bandwidth; both several orders of magnitude from the performance required for this application. Capacitive sensing is ubiquitous in microelectromechanical systems (MEMS) but does not scale to nanoelectromechanical systems (NEMS) quite as well as piezoresistive and piezoelectric sensing due to fundamental physical scaling laws.

Second, sensing and actuating over a 0.5-1 μm range would require an initial gap of at least 3 μm in order to avoid pull-in instability [42]. With such a large initial gap, the initial capacitance would be particularly small and large voltages (e.g. > 20 V) would be required

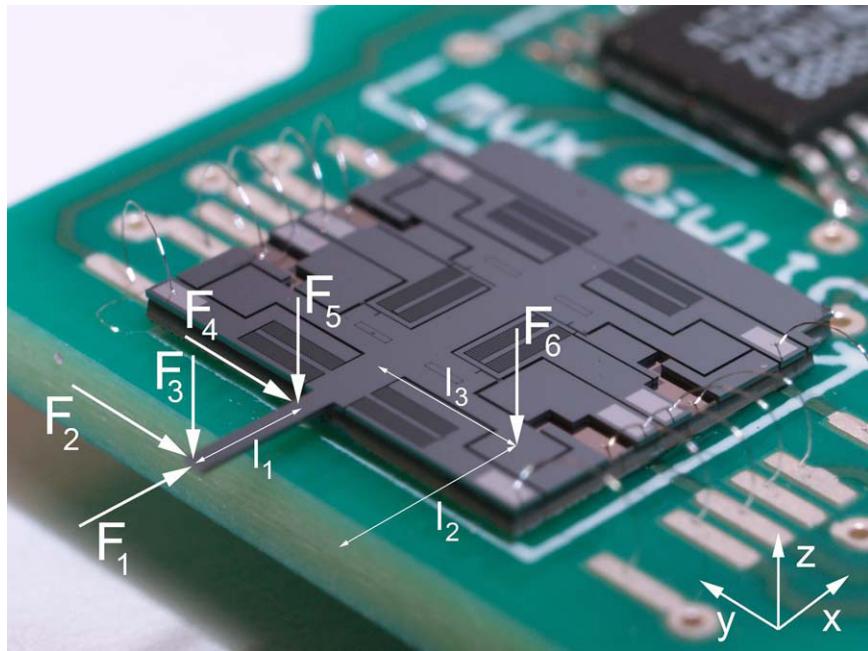


Figure 1.17: A six-axis capacitive sensor enabled by three sets of interdigitated electrodes and four parallel plate capacitors. The parallel plate capacitors are formed by a set of buried aluminum pads underneath the top silicon wafer, which was bonded to the handle wafer using low temperature (250°C) bonding. The sensor has force and torque resolutions of $1.4\ \mu\text{N}$ and $3.6\ \text{nN}\cdot\text{m}$ in a $15\ \text{Hz}$ measurement bandwidth. Reprinted from Beyeler et al. [41]. ©2009 IEEE.

for actuation and sensing. The gap could be reduced by extending the cantilever far beyond the end of the stationary electrode (taking advantage of mechanical amplification) at the expense of a smaller capacitance.

Third, the presence of a stationary electrode underneath the electrode would cause two problems for liquid operation: stiction and squeezed film damping. It is possible that they could be alleviated by coating the area with a superhydrophobic film to prevent liquid from entering the gap.

Due to the issues with optical and capacitive sensing, piezoresistive and piezoelectric sensing are the two remaining options (Figure 1.18). Piezoresistive sensors transduce a change in the force applied to the cantilever into a change in the resistance of a doped silicon resistor according to

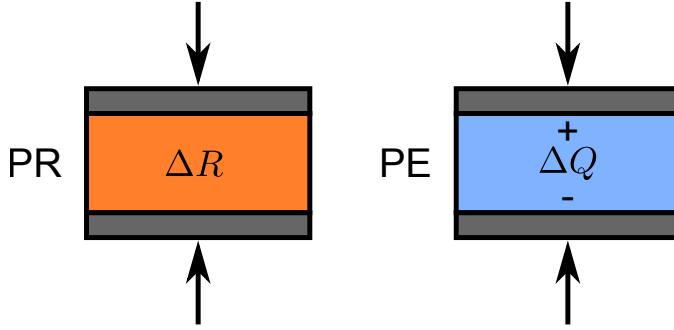


Figure 1.18: Piezoresistive and piezoelectric transduction. Piezoresistive sensing transduces a mechanical load into a change in resistance while piezoelectric sensing transduces the load into charge polarization. Both detect stress whereas optical and capacitive sensing respond to deflections.

$$\frac{\Delta R}{R} = \sigma\pi \quad (1.9)$$

where π is the piezoresistive coefficient and σ is the mechanical stress. The change in resistance is transduced into a voltage using a Wheatstone bridge.

Piezoelectric sensors transduce a change in force into a polarization charge on a piezoelectric film. The polarization charge can be calculated from

$$\Delta Q = d_{31}\sigma A$$

where d_{31} is the transverse piezoelectric coefficient of the film and A is the film area. The charge can be measured by sandwiching the piezoelectric film between metal electrodes and connecting them to a measurement circuit. Either the charge polarization or the voltage drop across the film can be measured, with the latter approach working particularly well for the low permittivity piezoelectric materials commonly used in MEMS (e.g. AlN and ZnO). The voltage drop across the film can be calculated from

$$\Delta V = \frac{\Delta Q}{C} = \frac{\sigma d_{31}A}{\epsilon_{pe}A/t_{pe}} = \frac{\sigma d_{31}t_{pe}}{\epsilon_{pe}}. \quad (1.10)$$

While the charge scales with the area of the film and decreases as the size of the film decreases, the voltage is invariant to film area and actually increases as the film is made

thinner. Voltage-sensing is preferable for most nanoelectromechanical systems (NEMS) [43]. But whereas charge-sensing is insensitive to parasitic capacitances, the performance of voltage-sensed piezoelectric films degrades with parasitic capacitances.

Both piezoresistive and piezoelectric sensors respond to stresses whereas optical and capacitive sensors respond to deflections. Both types of sensors can be integrated into the cantilever beam without any stationary reference electrode (Figure 1.15) and both can be integrated with $< 1 \mu\text{m}$ thick cantilever beams.

Each technique has practical advantages and disadvantages. Both operate at fairly low voltages ($< 1 \text{ V}$) and can be coated with a thin dielectric layer to electrically isolate the device from the patch clamp electrode. Piezoresistors can be fabricated using common cleanroom equipment whereas the deposition of thin piezoelectric films uses fairly specialized equipment and processes. But piezoresistors continuously dissipate electrical power, raising the possibility of damaging the sample being studied, whereas piezoelectric sensors essentially do not dissipate any power.

The deciding factor between piezoresistive and piezoelectric sensing is the resolution of the sensor. We have written a detailed comparison of piezoresistive and piezoelectric sensing for broadband cantilever-based force sensors [43] and found that piezoresistive sensing yields significantly better force resolution for cantilevers less than $1 \mu\text{m}$ thick. We will not discuss the piezoelectric readout circuit or comparison in detail in this work for brevity, but refer the interested reader to Ref. [43].

The major distinction between piezoresistive and piezoelectric sensing is the impact of parasitics on their measurement resolution. Parasitic resistances decrease the sensitivity of piezoresistive sensors (1.9) while parasitic capacitances do the same for piezoelectric sensors (1.10). If parasitics are large relative to the piezoresistor resistance or piezoelectric capacitance then it becomes impossible to achieve the theoretical force and displacement resolution that the sensor is capable of. We have focused on the mechanics of the force probe so far (i.e. spring constant and rise time) but achieving the force and displacement resolutions required for the cochlear hair cell application is nontrivial in its own right. Of the two types of parasitics, resistive parasitics are obscenely simpler to minimize [42, 44] while parasitic capacitances are anything but, making piezoresistive sensing the clear choice. Piezoelectric sensing is often used in cantilevers that are operated on resonance

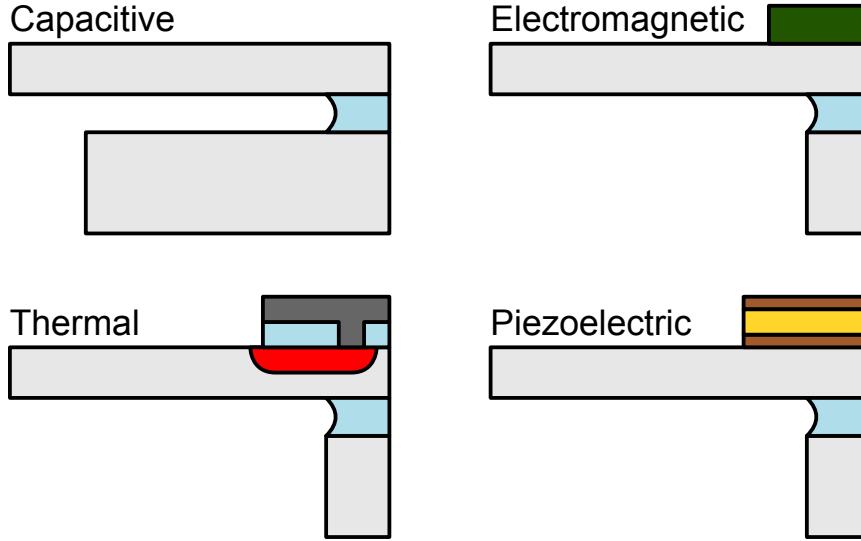


Figure 1.19: Summary of microcantilever actuator technologies. Microcantilevers are commonly actuated capacitively, electromagnetically, thermally and piezoelectrically. Capacitive actuation utilizes a stationary electrode to apply an electrostatic force to the cantilever, while electromagnetic actuation uses a magnetic material on the cantilever and external magnetic field gradient (or vice versa) to actuate the cantilever. Thermal actuation is accomplished by using a bilayer cantilever with a coefficient of thermal expansion mismatch (e.g. aluminum and silicon) and a resistive heater. Piezoelectric actuation uses an electric field to induce strain in a piezoelectric film.

because the sensor resolution requirements are not as stringent [45, 46]. In the next chapter we will discuss the fundamentals of piezoresistive sensing in much more detail and consider the design optimization of piezoresistive cantilevers.

1.4.2 Actuation techniques

The force probe we are designing requires both a sensor and actuator. Integrating the actuator directly into the probe rather than using an external, macroscale actuator ensures that the probe has a flat frequency response over the entire measurement bandwidth. For example, the main limitation in measuring hair cell kinetics to date has not been the frequency response of the glass probe but that of the actuator and mounting hardware. In this section we will investigate the actuator technologies that could be integrated with the piezoresistive sensor.

There are four common microscale actuation techniques: capacitive, electromagnetic, thermal and piezoelectric (Figure 1.19). The four actuators are located at the base of the cantilever on the assumption that the piezoresistive sensor will be located beyond the end of the actuator, closer to the tip. We will discuss this design choice in more detail in the next chapter.

We discussed capacitive sensing in the last section. Capacitive actuation uses the same arrangement of a cantilever beam and stationary electrode. Rather than applying an electrical potential between the cantilever and electrode in order to infer their spacing, the potential is applied in order to electrostatically pull the cantilever towards the electrode. While parasitic capacitances are no longer a major issue, the other issues with capacitive transduction that we discussed in the last section (i.e. pull-in, large voltages, squeezed film damping, stiction and passivation) still hold. Capacitive sensing and actuation is viable for certain biological measurements [47, 48] but is not a particularly attractive option for this application.

Electromagnetic actuation takes advantage of magnetic-field mediated forces to deflect the cantilever. There are several possible electromagnetic actuator implementations. A magnetic film could be deposited onto the cantilever, allowing the cantilever actuation using a macroscale electromagnetic. Alternatively, the on-chip portion of the actuator could be an electromagnet in the form of a resistive loop.

The first option (permanent magnet on the cantilever) is not convenient because high permeability magnetic materials are not CMOS compatible [49]. Iron, nickel, lead and other heavy metals reduce carrier lifetime in silicon and diffuse quickly at high temperature, limiting the tools available in most cleanrooms for processing wafers contaminated with heavy metals.

The second option (electromagnet on the cantilever) is a more practical option. Device fabrication would be straightforward and actuation frequencies could easily extend into the MHz range due to the relatively small inductance of microscale electromagnets. Ref. [50] is an excellent recent example of a cantilever that utilizes electromagnetic actuation. The large external DC magnetic field would not interfere with the patch clamp or piezoresistor signals so long as it was fairly stable. The main challenge in electromagnetic actuation is generating sufficient force. Force can be generated from either magnetic moments or the

Lorentz force.

The force on a magnetic moment in a magnetic field is

$$F_M = \nabla(m \cdot B)$$

where m is the magnetic moment and B is the external magnetic field. The magnetic moment of a current carrying loop is $m = IA$ where I is the current and A is the area of the loop. The force on the cantilever in the out-of-plane direction can then be calculated from

$$F_M = IA \frac{\partial B}{\partial z}.$$

Assuming an external magnetic field gradient of 100 T/m [51], a current of 1 mA (limited by heating and electromigration), and a 5 μm x 40 μm loop, electromagnetic actuation could generate a force of about 20 pN. However, the actuator is applying the force near the base of the cantilever. The lever arm effect from this arrangement reduces the effective force by the ratio of the total probe length to the actuator length. For example, if the actuator were 40 μm long and the entire length of the probe was 140 μm , the effective force would be reduced 7-fold, yielding an effective force of about 3 pN.

Alternatively, the cantilever could be actuated using the Lorentz force. This approach has been used for several microcantilever-based systems and integrated with piezoresistive force sensing [50, 52, 53]. The force experienced by a charge carrier moving in a magnetic field is

$$F_L = q\vec{v} \times \vec{B}$$

where v is the carrier velocity. An out-of-plane force on a cantilever beam can be generated from a current loop by orienting that external magnetic field along the longitudinal axis of the beam so that the carriers are moving perpendicular to the magnetic field at the end of the current-carrying loop. We can rewrite the Lorentz force experienced by the cantilever beam as

$$F_L = Iw_a B_x$$

where I is the total current, w_a is the extent of the current loop in the y -direction and B_x is the magnitude of the magnetic field in the x -direction. Assuming 1 mA of current, a 5 μm wide loop and external field of 0.2 T [52] the Lorentz force could generate 1 nN of force. Assuming once more that the loop extended 40 μm along the length of a 140 μm long probe the effective tip force would be about 285 pN. While smaller than the ideal range of 0.3 to 15 nN (Table 1.1), electromagnetic actuation could probably generate large enough deflections with sufficient design optimization, particularly if the electromagnet extended the entire length of the force probe.

However, we opted against electromagnetic actuation for the force probe. While capacitive, thermal and piezoelectric actuators are self-contained, electromagnetic actuation requires the generation of a large external magnetic field. Self-contained actuators have a substantial advantage in that they can be calibrated and characterized once before being used on a standard microscope in any research laboratory. Most importantly, integrating either a solenoid, permanent magnet or electromagnetic micromanipulator with the hair cell measurement setup would have been nontrivial [51, 54].

Piezoelectric and thermal actuation both operate on the principle of a bilayer beam (Figure 1.19). Expansion and contraction of the actuator induces deflections at the tip of the cantilever beam. Whereas electromagnetic actuators generate tip deflection from a force, piezoelectric and thermal actuators generate tip deflection from strain. Piezoelectric actuator strain is induced by applying an electric field across a piezoelectric film, while thermal actuator strain is induced by resistively heating a bilayer film composed of two layers with a large mismatch in their coefficients of thermal expansion.

Due to their similarities, we will discuss the two techniques in parallel. We will briefly compare the deflection range of piezoelectric and thermal actuators with our design requirements before discussing the actuation bandwidth and fabrication complexity of each technique.

We can estimate the tip deflection for piezoelectric and thermal actuation by making a few simplifying assumptions. We will assume that the actuator extends l_a along the length of the beam, which has a total length of l_c . We will assume that the thickness of the actuator is t_a and that uniform strain of ϵ_a is generated throughout the thickness of the actuator. Finally, we will assume that the actuator and silicon portions of the beam have

similar bending rigidities so that the neutral axis is located at the interface between the two.

From these assumptions, we can approximate the curvature along the length of the actuator as

$$C = \frac{M}{EI} = \frac{w_a E_a \epsilon (t_a/2)^2}{E_a w_a (2t_a)^3 / 12} = \frac{3\epsilon}{8t_a}.$$

The curvature can be integrated along the length of the actuator to calculate the tip deflection as

$$v_{\text{tip}} = \frac{1}{2} Cl_a^2 + Cl_a(l_c - l_a). \quad (1.11)$$

We will present more detailed multilayer beam bending models in Chapter 3 but will use (1.11) for now in order to quickly estimate the capability of each actuation technique.

Piezoelectric actuator strain can be calculated from

$$\epsilon = \frac{d_{31}V}{t_a}$$

where d_{31} is the transverse piezoelectric coefficient of the material and V is the potential drop across the film. Thermal actuator strain can similarly be calculated from

$$\epsilon = \Delta\alpha\Delta T$$

where $\Delta\alpha$ is the mismatch between the cantilever and actuator coefficients of thermal expansion while ΔT is the average temperature change of the actuator with respect to its initial temperature.

We can calculate the approximate tip deflections for piezoelectric and thermal actuators using our simplified bilayer beam model (1.11). We will once again assume that the probe consists of a 40 μm long actuator at the base of a 100 μm long force sensor. Based upon our earlier estimate of the required beam thickness (Figure 1.12) we will assume that the actuator is 250 nm thick.

The piezoelectric actuator would induce a 900 nm tip deflection assuming $d_{31} = 2 \text{ pm/V}$ and $V = 10 \text{ V}$. The d_{31} coefficient we assume is representative of thin aluminum nitride (AlN) or zinc oxide (ZnO) films. In the case of the thermal actuator, we will assume an

aluminum actuator ($\alpha = 23 \text{ ppm/K}$), silicon beam ($\alpha = 3 \text{ ppm/K}$) and temperature change of 5 K to yield an identical tip deflection of 900 nm. The equivalent force at the tip depends on the spring constant of the force probe; assuming spring constants ranging from 1 to 50 pN/nm the equivalent force would range from 0.9 to 45 nN, several orders of magnitude larger than electromagnetic actuation. The voltage required for each technique can be reduced by either decreasing the actuator thickness (piezoelectric) or decreasing the heater resistance (thermal) and can be scaled over a wide range.

Both thermal and piezoelectric actuation are clearly capable of providing sufficiently large deflections for the force probe. A second requirement for an actuator is providing sufficiently fast actuation. Both piezoelectric and thermal actuators are first-order systems, and their time constants can be calculated from

$$\tau = 2\pi RC.$$

where R and C are the electrical resistance and capacitance (piezoelectric actuator) or thermal resistance and capacitance (thermal actuator).

The time constant of a piezoelectric actuator is straightforward to calculate. Assuming R and C of 50Ω and 1 nF , the piezoelectric actuator time constant would be only $0.3 \mu\text{s}$, substantially faster than even the fastest hair cell experiments we hope to do. Piezoelectric actuator speed would only be limited by the overall dynamics of the force probe.

The time constant of a thermal actuator is slightly more complicated to calculate. We can conservatively estimate τ by assuming that heat only flows along the length of the cantilever and not directly into the surrounding air or liquid (i.e. that the cantilever is operating in vacuum). In that case, R and C can be calculated from

$$R = \frac{l_a}{2w_a t_a k}$$

$$C = \rho c l_a w_a t_a$$

where k , ρ and c are the effective thermal conductivity, density and specific thermal capacity of the actuator. Assuming the material properties of aluminum (ρ , c and k of

2700 kg/m^3 , $900 \text{ J/kg}\cdot\text{K}$ and $200 \text{ W/m}\cdot\text{K}$), the thermal time constant of the actuator would be about $50 \mu\text{s}$. While slower than ideal, the assumptions used in calculating it are fairly conservative, suggesting that it might be fast enough for at least the bundle mechanics experiments if not the kinetics experiments.

Both thermal [55, 56] and piezoelectric [45, 46, 57] actuators have been successfully integrated with piezoresistive sensors. However, piezoelectric actuators are slightly more complicated to fabricate due to issues with obtaining high d_{31} coefficients from thin film and the specialized deposition and etch processes required. A substantial advantage of thermal actuation is that it does not require any additional processing steps beyond those required for piezoresistive sensor fabrication. We will discuss their fabrication in Chapter 4, but for now simply want to note that both actuation techniques have previously been integrated with piezoresistive sensing.

1.4.3 Probe topology and prior work

Based upon this brief analysis of microscale sensing and actuation techniques, we will fabricate force probes that combine piezoresistive force detection with piezoelectric and thermal actuation. Both piezoresistive/piezoelectric (PRPE) and piezoresistive/thermal (PRT) scanning probes have been fabricated by other research groups. In this section we will briefly summarize prior work on scanning probes with integrated sensing and actuation and distinguish how our probe design will differ from them.

Minne, Manalis and Quate fabricated the first PRPE probes in 1995 [58]. Their cantilevers were $3.5 \mu\text{m}$ thick, $75 \mu\text{m}$ wide and on the order of $420 \mu\text{m}$ long with spring constants ranging from 0.6 to 7.1 N/m and resonant frequencies in air ranging from 20 to 70 kHz (Figure 1.20). They used ZnO deposited on a Ti/Au electrode for actuation and an ion implanted boron piezoresistor for force readout. The displacement resolution of the piezoresistor was 2.2 nm (equivalent to a force resolution of 1.3 to 15.6 nN) in a 10 Hz - 1 kHz bandwidth. The DC deflection of the cantilever ranged from 15 - 58 nm/V .

Their fabrication process is illustrated in Figure 1.21. The process consisted of four phases: tip formation and sharpening, piezoresistor implantation and activation, piezoelectric deposition and patterning, and cantilever release. There are two notable

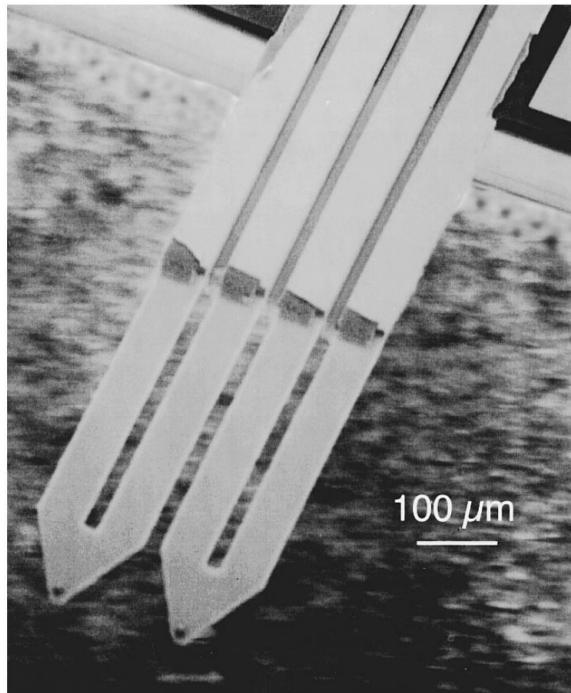


Figure 1.20: SEM micrograph of the first scanning probes to integrate piezoelectric actuation with piezoresistive force readout. The probes were developed by Minne, Manalis and Quate in 1995 for constant-force topography scanning using an array of scanning probes. The probes were roughly $3.5\text{ }\mu\text{m}$ thick, $75\text{ }\mu\text{m}$ wide and $420\text{ }\mu\text{m}$ long with spring constants ranging from 0.6 to 7.1 N/m . Reprinted from Ref. [58] with permission from the American Institute of Physics.

features of this first PRPE fabrication process. First, the piezoresistor extended the entire length of the cantilever. We will see in the next chapter that a shorter piezoresistive loop is generally desirable. Second, the piezoresistor extended underneath the piezoelectric actuator. This configuration was modified by subsequent researchers to limit capacitive and mechanical coupling between the actuator and sensor.

The Quate group improved their PRPE probes several times in the following years. In 1996, Minne, Manalis, Atalar and Quate reported two improvements [59]. First, they increased the doping of the piezoresistor underneath the piezoelectric actuator. This modification reduced the excess piezoresistor resistance and reduced mechanical actuator-sensor crosstalk, although it did not reduce capacitive crosstalk. Second, they added an additional trace of doped silicon in order to allow tip biasing for scanning probe

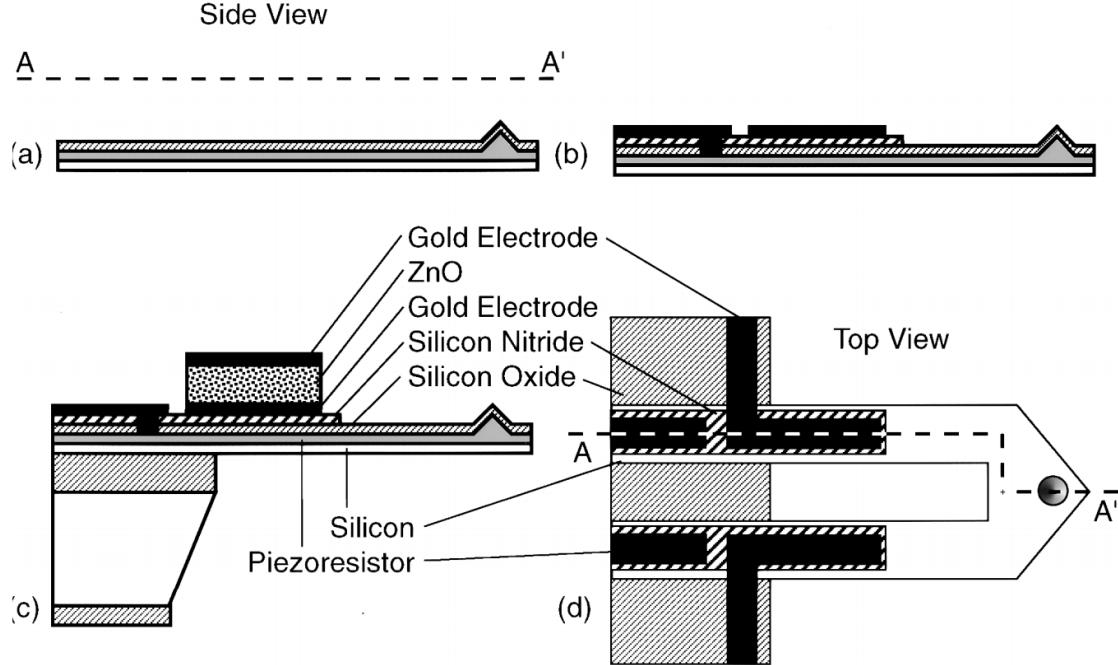


Figure 1.21: Cross-section of the first PRPE scanning probes. The fabrication process consisted of four phases: tip, piezoresistor, piezoelectric and cantilever release. The piezoresistor extended along the entire length of the cantilever, with the air gap between the piezoresistor legs forming the piezoresistor loop. The piezoelectric actuator consisted of a ZnO film deposited on Cr/Au electrodes. Notably, the piezoresistor runs underneath the piezoelectric actuator, in contrast with later piezoresistor-actuator (PRA) probes. Reprinted from Ref. [58] with permission from the American Institute of Physics.

lithography. Later that year, Manalis, Minne and Quate reported an improved piezoresistor readout scheme [57]. They measured the piezoresistor signal using a lock-in amplifier to eliminate the impact of capacitively coupling between the actuator and sensor. By biasing the piezoresistor at a much higher frequency than the actuator bandwidth (130 kHz vs. 8 kHz) they were able to reduce the impact of capacitive crosstalk.

In 1998, Minne et al. presented an updated PRPE probe [60]. They made two important design improvements. First, they improved the grounding of the bottom PE electrode to reduce capacitive crosstalk. Second, they reduced the length of the piezoresistor so that it did not extend the entire length of the cantilever, improving the deflection resolution of the system. They fabricated a 50 x 1 array of PRPE probes in order to enable parallel,

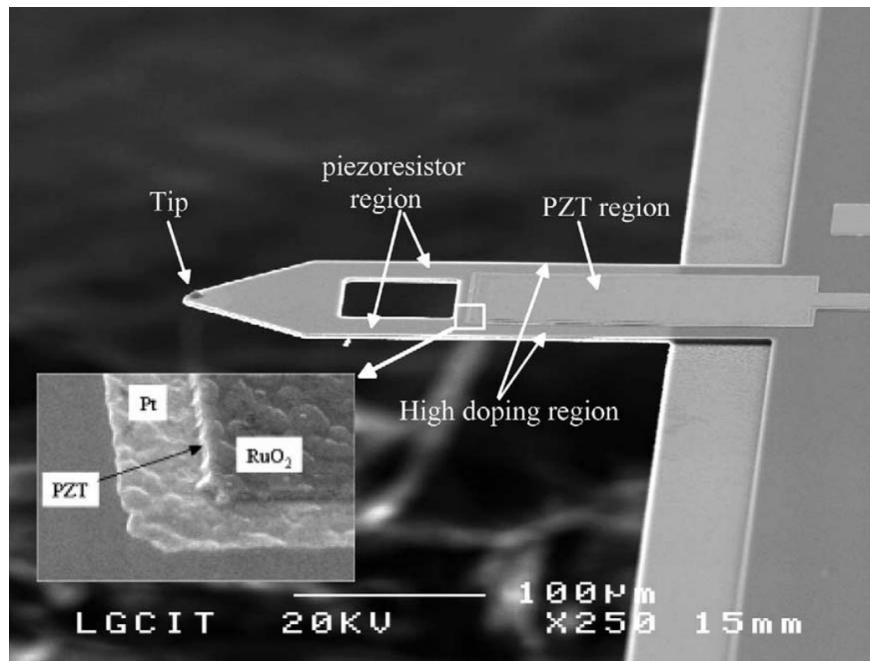


Figure 1.22: A PZT-based PRPE probe developed by Kim et al. [65]. The probes incorporated several important improvements from earlier PRPE designs. First, the piezoresistor did not run underneath the actuator, reducing capacitive crosstalk. Second, they used PZT rather than ZnO for the actuator enabling a reduction in probe size and actuator voltages. Reprinted from Ref. [65] with permission from Elsevier.

high-speed atomic force microscopy and lithography [61]. In parallel, Manalis and Sulcuk developed an interferometric optical readout scheme for parallel, high-speed atomic force microscopy, although they did not integrate individual probe actuation into these devices [62,63].

In 2000, Sulcuk et al. applied on-chip piezoelectric actuation to operation in liquid [64]. They used a thin coating of polydimethylsiloxane (PDMS) to passivate the ZnO actuator from the liquid. These probes did not include a piezoresistor and used the conventional laser beam bounce approach to detect cantilever deflection. Sulcuk et al. observed a reduction in resonant frequency (56 to 35 kHz) and quality factor (110 to 3) for cantilever operation in liquid rather than air. While the PDMS coating would have been too thick for application to sub-micron thick cantilevers, their use of a soft polymer was notable.

In 2003, Kim et al. presented a PRPE probe utilizing a PZT actuator [65]. They made two major improvements over earlier PRPE scanning probe designs. The first improvement was switching from ZnO to PZT for the piezoelectric actuator. The much higher d_{31} coefficient of PZT allowed for lower actuation voltages and the fabrication of shorter, higher frequency probes while still enabling tip deflections of up to 10 μm .

The second improvement that they made was changing the piezoresistor layout. They compared three different configurations. In the first configuration the piezoresistor extended underneath the PZT actuator. In the second, they reduced the width of actuator so that it did not overlap with the piezoresistor. In the third configuration, they only fabricated the piezoresistor beyond the end of the PZT actuator and used metal interconnects to make contact with it. The third configuration represented a 5-fold improvement over the work by Minne et al. [60]. The three piezoresistor configurations are shown in Figure 1.23. Note that the third configuration yielded the best performance due to a reduction in the actuator-sensor capacitance and the lower resistance of the metal interconnects, because the sensor resistance is what transduces the capacitively injected current into a voltage. We will discuss actuator-sensor crosstalk mechanisms in more detail in the next chapter.

Shin et al. developed a similar PRPE probe in 2010 utilizing a PZT actuator and placing the silicon piezoresistor beyond the end of the actuator [66]. They also included a reference cantilever in the Wheatstone bridge in order to reduce common-mode signals such as capacitive coupling and temperature changes, although they did not report the level of capacitive crosstalk.

In parallel with the development of PRPE probes from 1995 to 2010, several groups developed PRT probes. Early in 2003, Ivanov et al. reported a piezoresistive probe utilizing the thermal expansion of an aluminum film for actuation [67]. Later that year the Rangelow group reported additional details of their PRT probes in Pedrak et al. [68]. The silicon resistor used to heat the aluminum actuator was biased with the sum of two signals. The resonant mode of the cantilever was excited by the first signal (AC) while a lower frequency signal (DC) was used to adjust the deflection of the probe to maintain constant cantilever deflection based upon the piezoresistor signal. The beauty of the aluminum-based PRT process was that it did not require any substantial modification from a piezoresistor-only fabrication process. In fact, their 2003 process was based upon a piezoresistive cantilever

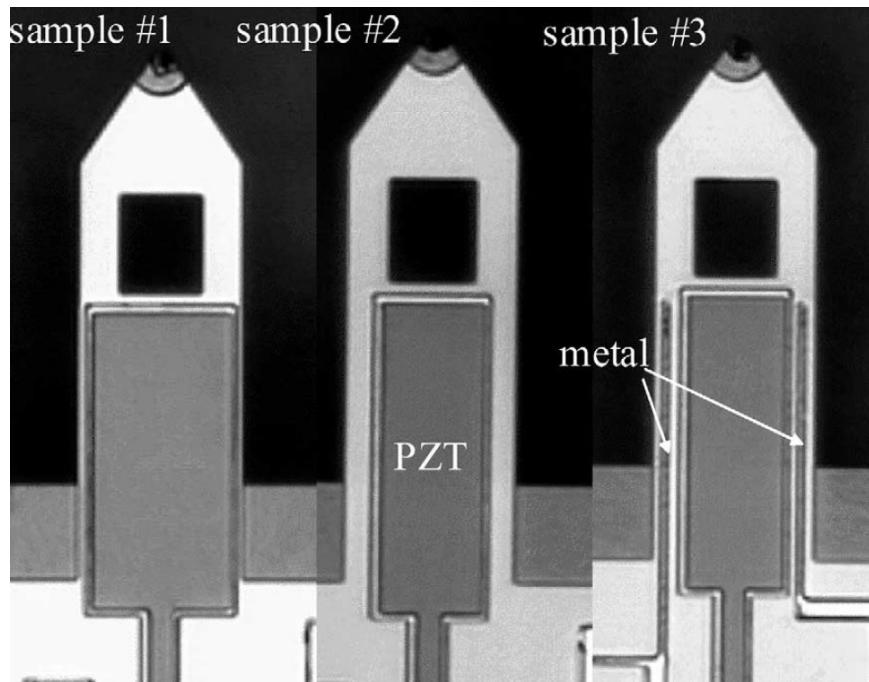


Figure 1.23: Alternative probe layouts to reduce capacitive crosstalk. In sample #1 the silicon piezoresistor runs underneath the piezoelectric actuator. In sample #2, the width of the actuator has been reduced in order to reduce the overlap capacitance between the actuator and sensor. In sample #3, which yielded a nearly 10-fold reduction in crosstalk from the other designs, metal interconnects are used to contact the piezoresistor, which is located beyond the end of the actuator. Reprinted from Ref. [65] with permission from Elsevier.

process developed in 1996 by some of the same authors [69].

Fantner et al. reported an updated PRT probe in 2009 and applied it to topographical scans of yeast cells in liquid [55]. A schematic of their probe is shown in Figure 1.24. Their cantilevers were 3-5 μm thick, 110 μm wide and 320 μm long with spring constants ranging from 2 to 11 N/m. A typical probe had resonant frequencies of 80 kHz and 13 kHz in air and water, respectively. The probes were coated with a thin layer of formvar dissolved in ethylene dichloride in order to passivate the aluminum from the surrounding liquid.

As in the 2003 PRT devices, the first resonant mode of the cantilever was excited by the thermal actuator (ω) while the piezoresistor signal was measured at twice the

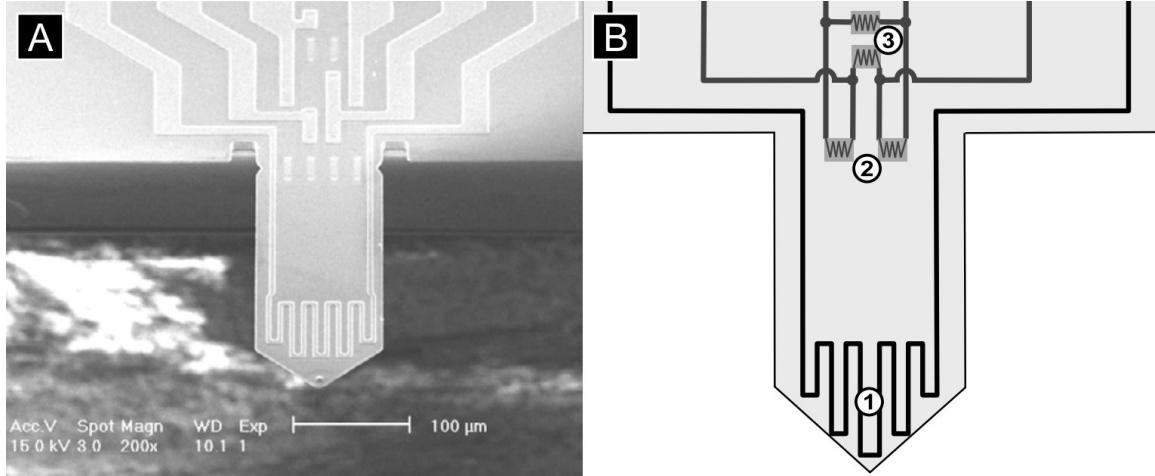


Figure 1.24: Schematic of a piezoresistive cantilever with thermal actuation. A Wheatstone bridge consisting of four piezoresistors is fabricated near the base of the cantilever with two resistors actively sensing stress. The thermal actuator is located near the tip of the probe. The cantilever is driven on resonance and a constant force is maintained using the thermal actuator based upon the piezoresistor signal amplitude. Reprinted from Ref. [55] with permission from IOP Publishing.

thermal actuator frequency due to the square-law nature of thermal actuation (2ω) [67, 68]. Assuming an actuation signal of the form

$$V(t) = V_{AC} \sin(\omega t) + V_{DC}$$

the actuator power dissipation can be calculated from

$$\begin{aligned} W &= \frac{V^2}{R} \\ &= \frac{1}{R} \left[V_{AC}^2 \left(\frac{1 - \cos(2\omega t)}{2} \right) + 2V_{AC}V_{DC} \sin(\omega t) + V_{DC}^2 \right] \\ &= \frac{1}{R} \left[-\frac{V_{AC}^2}{2} \cos(2\omega t) + 2V_{AC}V_{DC} \sin(\omega t) + \left(\frac{V_{AC}^2}{2} + V_{DC}^2 \right) \right] \end{aligned}$$

where R is the heater resistance and W is the power dissipation. By driving V at ω and measuring the cantilever deflection at 2ω the capacitive crosstalk signal can be isolated from the piezoresistor signal, although thermal and mechanical crosstalk will remain. If

$V_{DC} \neq 0$ then there will be cantilever deflection at ω as well. The spectral separation between the voltage and actuation signals is ideal for tapping mode probe operation but can not be applied to measuring the step or impulse response of a sample.

The work by Fantner et al. demonstrated the two major advantages of thermal actuation over piezoelectric actuation. First, the fabrication process is fairly straightforward. Second, if the PRT probe is operated in tapping mode then the capacitive crosstalk signal can be isolated. This is not an option for piezoelectric actuation, because piezoelectric actuation is linear while thermal actuation is quadratic.

The following year, Hafizovic et al. demonstrated thermally actuated piezoresistive cantilevers with nN force resolution that were monolithically integrated with CMOS in order to enable the independent and parallel operation of multiple probes [70]. The actuation bandwidth of the thermal actuators was limited to about 300 Hz due to the relatively large size of the probes [71].

Sarov et al. developed the thermally actuated probe topology further in 2011 in order to reduce actuator-sensor crosstalk. They used metal rather than silicon for the resistive heater. Reducing the heater resistance allowed for the same actuation power with a smaller voltage. By reducing the heater resistance 30-fold they were able to reduce the required actuator voltage 5-fold [56].

In this section we have discussed the PRPE and PRT probes that have been fabricated to date. All of them have been much larger ($> 2.5 \mu\text{m}$ thick) and stiffer ($> 0.6 \text{ N/m}$) than required for the study of cochlear hair cell mechanics. The reason for their larger size is simply that they were designed for conventional AFM topography scanning rather than the study of soft biological samples. If we expand our review of prior work to include probes without integrated actuation, there are numerous examples of smaller and softer probes.

For example, Ando et al. reported an AFM system in 2001 that was capable of imaging cells and biomolecules at 12.5 frames per second (fps) in liquid [72]. They utilized optical beam bounce readout with 140 nm thick, 2 μm wide and $\approx 10 \mu\text{m}$ long silicon nitride cantilevers. The cantilevers were coated with a thin metal film in order to increase their reflectivity. Cantilever stiffnesses ranged from 150 to 280 pN/m with resonant frequencies ranging from 1.3 to 1.8 MHz in air and 450 to 650 kHz in water. If their probes had been slightly longer, their mechanical characteristics would have been ideal for measurements

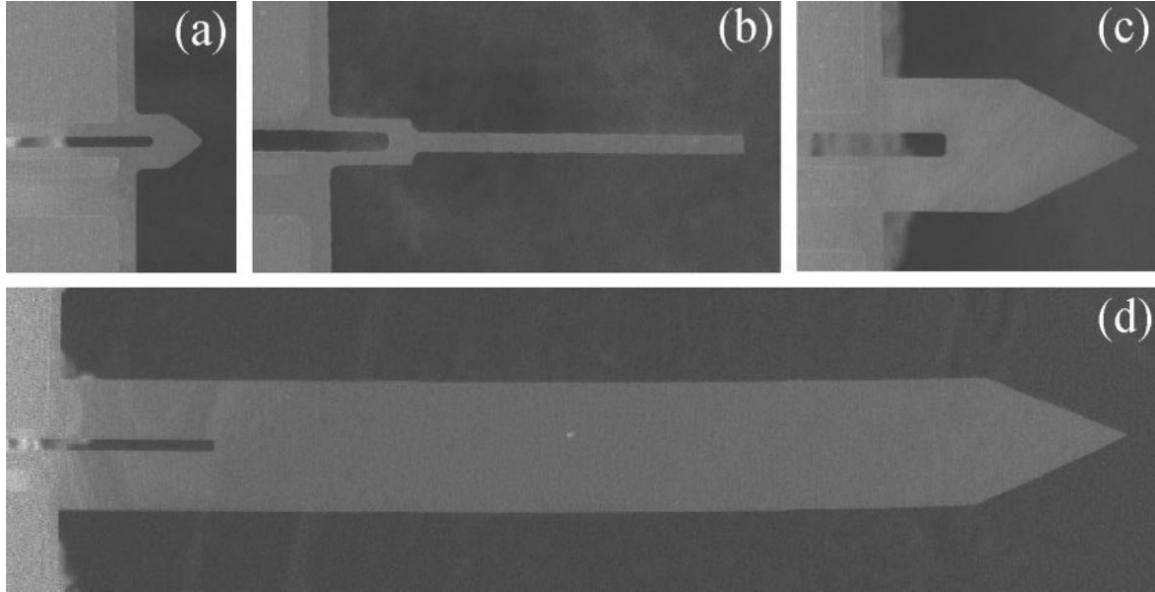


Figure 1.25: Examples of 89 nm thick piezoresistive cantilevers. The cantilevers are all 89 nm thick and range from 2 to 44 μm in width and 10 to 350 μm in length. While several of the cantilevers have the mechanical properties required for the study of cochlear hair cell mechanics, they do not include integrated actuation for the high-speed operation. Reprinted from Ref. [36] with permission from the American Institute of Physics.

of bundle mechanics besides the challenge in applying beam bounce readout to the hair cell problem.

The first piezoresistive cantilevers, reported by Tortonese et al. in 1993 [35], were relatively large. Their cantilevers were 4.5 μm thick, 10 to 50 μm wide and 75 to 400 μm long with spring constants ranging from 5 to 100 N/m and resonant frequencies in air ranging from 40 to 800 kHz. But small, soft piezoresistive cantilevers have also been developed. Chui et al. reported 1 μm thick cantilevers in 1996 [73] while Ried et al. reported 340 nm thick cantilevers in 1997 [74]. In 1999, Harley and Kenny reported 89 nm thick epitaxial piezoresistive cantilevers [36]. Several of the probes that they fabricated are shown in Figure 1.25, with lengths ranging from 10 to 350 μm and widths ranging from 2 to 44 μm , respectively. Harley and Kenny published their seminal work on piezoresistor noise and design optimization the following year, which we will draw heavily upon in the next chapter [75].

To summarize this section, while numerous probes with integrated actuation have been

fabricated to date, none have combined the stiffness, bandwidth and sensor resolution performance required for studying high-speed molecular and cellular processes, particularly cochlear hair cell mechanics and kinetics. There is a clear need for a faster, higher performance force probe; the remainder of the thesis will focus on the design, fabrication and application of such a probe.

1.5 Thesis outline

The rest of the thesis is organized as follows:

- **Chapter 2: Piezoresistor fundamentals.** Piezoresistive sensor design requires a good understanding of the mechanical, electrical and thermal processes involved. In this chapter we will discuss the fundamentals of piezoresistive sensors, including signal conditioning, fabrication methods and numerical design optimization.
- **Chapter 3: Probe design.** We will apply the numerical models developed in the last chapter to design optimized force probe sensors. Next we will discuss the design of thermal and piezoelectric unimorph actuators and generate the combined sensor-actuator force probe designs. We will close by discussing actuator drive and sensor readout circuits.
- **Chapter 4: Fabrication.** The mask layout, fabrication processes and preliminary characterization of the force probes will be presented.
- **Chapter 5: Characterization.** The performance of the finished force probes will be characterized using a variety of techniques. We will discuss the mechanical, sensing, and actuation performance of the probes and compare them with our design goals.
- **Chapter 6: Hair cell experiments.** We will apply our force probes to the study of cochlear hair cell mechanics and kinetics. Capacitive crosstalk between the on-chip actuator and patch clamp electrode will be discussed and solutions will be presented. Preliminary measurements on a mammalian inner hair cell will be shown.

- **Chapter 7: Summary.** We will summarize the most important results from the rest of the thesis and suggest future research directions.

A handful of appendices are included at the end of the thesis. Appendix A summarizes the mathematic symbols used throughout this work. Appendix B lists the publication that I contribute to while at Stanford. Appendix C includes a detailed discussion of the circuit designs, including schematics and printed circuit board layouts. Appendix D reproduces the masks used to fabricate the force probes while Appendix E includes fabrication runsheets and process details. Appendix F describes the first iteration of the force probe fabrication process. Appendix G provides lookup tables that can be used to design ion implanted piezoresistive sensors. Finally, Appendix H reproduces the code used to design and model the piezoresistive cantilevers.

Chapter 2

Piezoresistor fundamentals

In this chapter we will discuss the modeling and design concepts that form the basis of piezoresistive sensor design. First we will consider resistive strain sensing in general and the piezoresistive effect in particular. Next, we will discuss signal conditioning, process modeling, noise sources and other fundamental modeling issues. We will close by discussing temperature effects and numerical design optimization. Parts of this chapter were adapted from an earlier review of piezoresistance [24].

Consider a homogenous resistor in the shape of a square prism (Figure 2.1). The electrical resistance (R) measured between its ends is

$$R = \frac{\rho l}{w^2} \quad (2.1)$$

where l is the length, w is the width of each side, and ρ is the electrical resistivity.

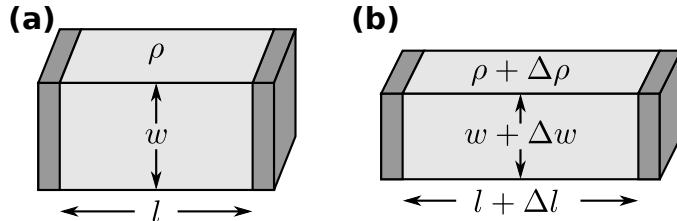


Figure 2.1: (a) The electrical resistance of a homogenous square prism depends on its dimensions (l and w) and resistivity (ρ). (b) When subjected to mechanical loading, all three can potentially change, leading to a change in the measured electrical resistance.

A mechanical load applied to one face of the resistor while holding the other end stationary will mechanically strain the material. For a change in length equal to Δl , the longitudinal strain in the material is $\varepsilon_l = \Delta l / l$ to first-order approximation.

The width will also change due to contraction of material in the transverse direction in proportion to the Poisson's ratio of the material (ν). For a sufficiently small strain the change in width is $\Delta w = -\nu \varepsilon_l w$. ν ranges from 0.20 to 0.35 for most polycrystalline metals. For single crystal silicon, which is anisotropic, the Poisson's ratio ranges from 0.06 to 0.36 depending on the crystal orientation [76, 77]. The lower and upper limits for ν of an isotropic material are -1.0 and 0.5 [78].

We can take the natural log and partial derivatives of both sides of (2.1) to obtain

$$\frac{\partial R}{R} = \frac{\partial \rho}{\rho} + \frac{\partial l}{l} - 2 \frac{\partial w}{w} \quad (2.2)$$

and can write the relative resistance change in terms of the strain [79] as

$$\frac{\Delta R}{R} = \frac{\Delta \rho}{\rho} + (1 + 2\nu) \varepsilon_l \quad (2.3)$$

The gauge factor (GF) of a strain gauge is defined as

$$GF = \frac{\Delta R/R}{\varepsilon_l}. \quad (2.4)$$

Combining (2.3) and (2.4) we see that geometric effects alone provide a gauge factor of approximately 1.4 to 2.0 depending on the Poisson's ratio. Metals experience a small resistivity change with strain, which typically increases the gauge factor by an additional 0.3 or so. However, for silicon and other semiconductors, $\Delta \rho / \rho$ can be 50 to 100 times larger than the geometric term due to the piezoresistive effect.

2.1 Notation fundamentals

In this chapter we will explore the piezoresistive effect in detail. After introducing crystal structure and notation fundamentals we will describe how the magnitude and origins of piezoresistivity, and how the magnitude of the effect varies with crystal orientation, dopant

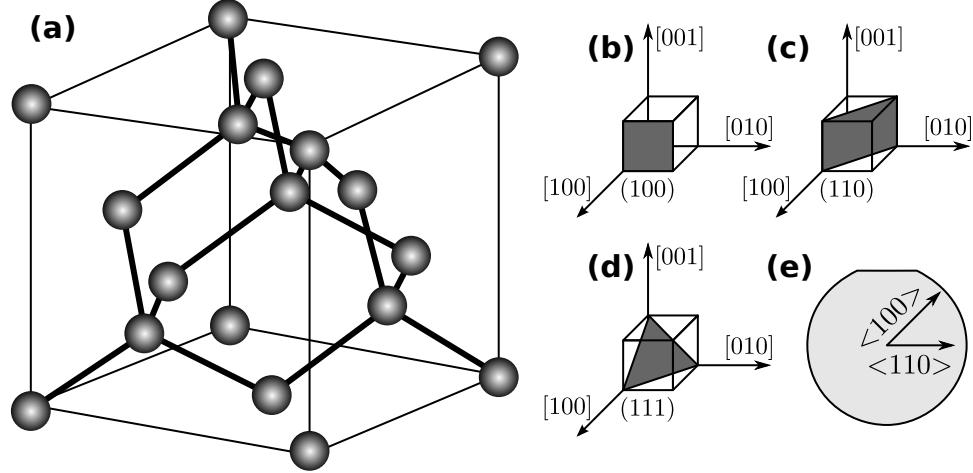


Figure 2.2: (a) Covalently bonded diamond cubic structure of silicon. Silicon has four covalent bonds and coordinates itself tetrahedrally. (b-d) Commonly employed crystal planes of silicon include the (100), (110) and (111) planes. (e) Wafers smaller than 200 mm in diameter have a major flat that indicates the $\langle 110 \rangle$ direction. Piezoresistors are typically oriented in the $\langle 100 \rangle$ or $\langle 110 \rangle$ directions for n- or p-type doping, respectively. Following Ref. [24]. ©2009 IEEE.

concentration and temperature.

2.1.1 Crystal structure

Crystals are periodic arrangements of atoms arranged in one of 14 lattice types and complete reviews are available elsewhere [80,81]. The Miller indices specify crystal planes by n-tuples. A direction index $[hkl]$ denotes a vector normal to a plane described by (hkl) . Angle-bracketed indices, like $\langle hkl \rangle$, represent all directions equivalent to $[hkl]$ by symmetry.

In a hexagonal crystal, as found in most silicon carbide polytypes, the Bravais-Miller index scheme is commonly adopted where four indices are used to represent the intercept-reciprocals corresponding to the four principal crystal axes (a_1 , a_2 , a_3 , and c). The axes a_1 , a_2 , and a_3 are on the same plane and 120° apart from one another while c is perpendicular to the a-plane defined by the (a_1, a_2, a_3) triplet.

Crystalline silicon forms a covalently bonded diamond-cubic structure with lattice constant $a=5.43$ Å (Figure 2.2a). The diamond-cubic structure is equivalent to two

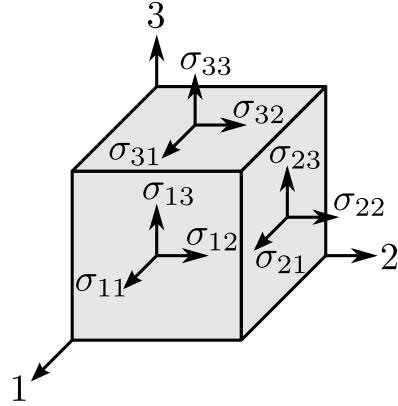


Figure 2.3: Nine components, σ_{ij} , of stress on an infinitesimal unit element. For clarity, stresses on negative faces are not depicted. Following Ref. [24]. ©2009 IEEE.

interpenetrating face-centered-cubic (FCC) lattices with basis atoms offset by $\frac{1}{4}a$ in the three orthogonal directions [80]. Silicon's diamond-cubic lattice is relatively sparse (34% packing density) compared to a regular face-centered-cubic (FCC) lattice (74% packing density).

Commonly used wafer surface orientations in micromachining include (100), (111), and (110) (Figure 2.2b). The {111} planes, oriented 54.74° from {100} planes, are the most densely packed. As alluded to earlier, the magnitude of the piezoresistive effect varies with the crystal orientation of the resistor. In order to maximize sensitivity piezoresistors are typically oriented in the <100> direction for n-type doping and <110> direction for p-type doping. Variation in piezoresistive coefficients with orientation will be discussed in more detail shortly.

2.1.2 Stress, strain and tensors

To define the state of stress for a unit element (Figure 2.3), nine components, σ_{ij} , must be specified, as in:

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \quad (2.5)$$

The first index i denotes the direction of the vector normal to the surface that the stress is applied to, while j indicates the direction of the force or stress. If $i = j$, the stress is normal to the specified surface, while $i \neq j$ indicates a shear stress on face i (Figure 2.3). From static equilibrium requirements that forces and moments sum to zero, a stress tensor is always symmetric, that is $\sigma_{ij} = \sigma_{ji}$, and thus the stress tensor contains only six independent components. Strain, ε_{ij} , is also directional. For an isotropic, homogeneous material, stress is related to strain by Hooke's Law, $\sigma = \varepsilon E$ [82].

Although effective values of Young's modulus and Poisson's ratio for a single direction are often employed for simple loading situations, a tensor is required to fully describe the stiffness of an anisotropic material such as silicon [42, 83, 84]. The stress and strain are related by the elastic stiffness matrix, C , where $\sigma_{ij} = C_{ijkl}\varepsilon_{kl}$, or equivalently by the inverse compliance matrix, S , where $\varepsilon_{ij} = S_{ijkl}\sigma_{kl}$:

$$\begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46} \\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56} \\ c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66} \end{pmatrix} \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{pmatrix} \quad (2.6)$$

and

$$\begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} \\ s_{12} & s_{22} & s_{23} & s_{24} & s_{25} & s_{26} \\ s_{13} & s_{23} & s_{33} & s_{34} & s_{35} & s_{36} \\ s_{14} & s_{24} & s_{34} & s_{44} & s_{45} & s_{46} \\ s_{15} & s_{25} & s_{35} & s_{45} & s_{55} & s_{56} \\ s_{16} & s_{26} & s_{36} & s_{46} & s_{56} & s_{66} \end{pmatrix} \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{pmatrix} \quad (2.7)$$

Collapsed notation reduces each pair of subscripts to one number: $11 \rightarrow 1$, $22 \rightarrow 2$, $33 \rightarrow 3$, $23 \rightarrow 4$, $13 \rightarrow 5$, $12 \rightarrow 6$, e.g. σ_{11} becomes σ_1 , ε_{12} becomes ε_6 , c_{1111} becomes c_{11} and s_{2323} becomes s_{44} .

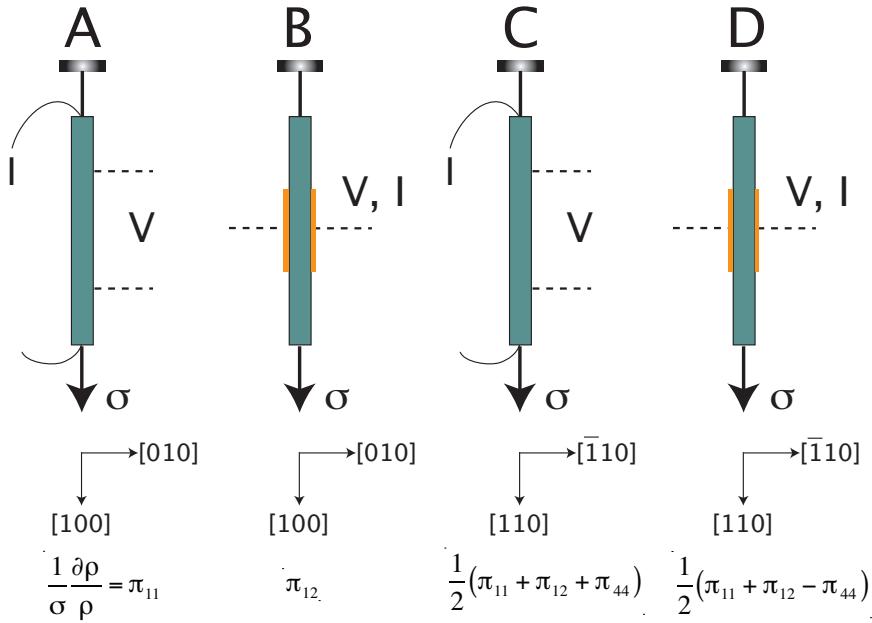


Figure 2.4: Notation for Smith’s test configurations. Configurations A and C measured longitudinal piezoresistance, while configurations B and D provided transverse coefficients. Voltage drops between the electrodes (dotted lines) were measured while uniaxial tensile stress, σ , was applied to the test sample by hanging a weight. The experiments were done in constant-current mode in a light-tight enclosure with controlled temperature ($25 \pm 1^\circ\text{C}$). After Smith [85]. ©1954 American Physical Society.

2.2 Piezoresistivity

Single crystal germanium and silicon were the first materials widely used as piezoresistors. Smith reported the first measurements of large piezoresistive coefficients in these semiconductor crystals in 1954 noting that work by Bardeen and Shockley, and later Herring, could explain the phenomena [85]. Smith applied Bridgman’s tensor notation [86] in defining the piezoresistive coefficients and geometry of his test configurations (Figure 2.4).

2.2.1 Crystallographic orientation

The piezoresistive coefficients (π) require four subscripts because they relate two second-rank tensors of stress and resistivity. In general current is forced across the piezoresistor in one direction, the voltage drop is measured along another, and the stress is applied in a

separate arbitrary direction.

The first subscript refers to the direction of electrical potential measurement, the second to the direction of current flow, in the resistor, and the third and fourth to the stress (recall that stress has two directional components). For conciseness, the subscripts of each tensor are also collapsed [87], e.g. $\pi_{1111} \rightarrow \pi_{11}$, $\pi_{1122} \rightarrow \pi_{12}$, $\pi_{2323} \rightarrow \pi_{44}$. Kanda later generalized these relations for a fixed voltage and current orientation (ω) as a function of stress (λ) [88]:

$$\frac{\Delta\rho_\omega}{\rho} = \sum_{\lambda=1}^6 \pi_{\omega\lambda} \sigma_\lambda. \quad (2.8)$$

Smith determined these coefficients for relatively lightly doped silicon and germanium samples with resistivities ranging from 1.5 to 22.7 $\Omega\text{-cm}$ [85]. He measured the piezoresistive coefficients for (100) samples along the $\langle 100 \rangle$ and $\langle 110 \rangle$ crystal directions. Longitudinal and transverse coefficients for the fundamental crystal axes were determined directly. Shear piezoresistive coefficients were inferred. By these measurements and considering the crystal symmetry, Smith fully characterized the piezoresistive tensor of 7.8 $\Omega\text{-cm}$ p-type silicon as

$$\pi_{\omega\lambda} = \begin{pmatrix} 6.6 & -1.1 & -1.1 & 0 & 0 & 0 \\ -1.1 & 6.6 & -1.1 & 0 & 0 & 0 \\ -1.1 & -1.1 & 6.6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 138.1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 138.1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 138.1 \end{pmatrix} \times 10^{-11}$$

where π has units of Pa^{-1} and takes the general form of

$$\pi_{\omega\lambda} = \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{22} & \pi_{12} & 0 & 0 & 0 \\ \pi_{12} & \pi_{12} & \pi_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \pi_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \pi_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \pi_{44} \end{pmatrix}$$

As illustrated in Figure 2.4, there are four particularly important piezoresistive measurement configurations. The direction of current flow and applied stress can be parallel (longitudinal) or perpendicular (transverse) to each other, and they can be oriented in the $\langle 100 \rangle$ or $\langle 110 \rangle$ directions of the crystal lattice.

The net relative resistivity change is usually written as a sum of the longitudinal and transverse effects as

$$\frac{\partial \rho}{\rho} = \pi_l \sigma_l + \pi_t \sigma_t \quad (2.9)$$

where σ_l is the stress parallel to the current flow (longitudinal direction), σ_t is the stress perpendicular to the current flow (transverse direction), and π_l and π_t are the longitudinal and transverse piezoresistive coefficients.

The relationship between π_l , π_t and the fundamental piezoresistive coefficients depends on the crystallographic orientation of the resistor. In the $\langle 100 \rangle$ direction, the relationships are

$$\pi_l = \pi_{11} \quad (2.10)$$

$$\pi_t = \pi_{12} \quad (2.11)$$

while in the $\langle 110 \rangle$ direction they are

Table 2.1: Piezoresistive coefficients for n- and p-type single crystal silicon in the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions, respectively, near room temperature and at low dopant concentrations ($< 10^{17} \text{ cm}^{-3}$).

Doping Type	$\pi_{11} (10^{-11} \text{ Pa}^{-1})$	$\pi_{12} (10^{-11} \text{ Pa}^{-1})$	$\pi_{44} (10^{-11} \text{ Pa}^{-1})$
n-type	-102.2	53.4	-13.6
p-type	6.6	-1.1	138.1

Table 2.2: Longitudinal and transverse piezoresistive coefficients for single crystal silicon in the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions.

Doping Type	Orientation	$\pi_l (10^{-11} \text{ Pa}^{-1})$	$\pi_t (10^{-11} \text{ Pa}^{-1})$
n-type	$\langle 100 \rangle$	-102	53.4
	$\langle 110 \rangle$	-31.6	-17.6
p-type	$\langle 100 \rangle$	6.6	-1.1
	$\langle 110 \rangle$	71.8	-66.3

$$\pi_l = 1/2(\pi_{11} + \pi_{12} + \pi_{44}) \quad (2.12)$$

$$\pi_t = 1/2(\pi_{11} + \pi_{12} - \pi_{44}). \quad (2.13)$$

The piezoresistive coefficients for lightly doped ($< 10^{17} \text{ cm}^{-3}$), room temperature single crystal silicon piezoresistors are summarized in Table 2.1. Note the large π_{11} and π_{44} coefficients for n- and p-type silicon, respectively. The longitudinal and transverse piezoresistive coefficients in the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions are summarized in Table 2.2.

Note that the $\langle 110 \rangle$ orientation with transverse stress can be used to reduce the resistivity of both n- and p-type resistors. In CMOS devices, where mobility enhancement is desired, the $\langle 110 \rangle$ orientation is often used to obtain mobility enhancement for both n- and p-type channels [89]. An alternative approach, which maximizes the mobility enhancement for both transistor types, is to orient n- and p-type devices in the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions with longitudinal tensile and compressive stresses, respectively [90].

The coefficients illustrate why n- and p-type resistors should be oriented in the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions, respectively, in order to maximize the magnitude of the

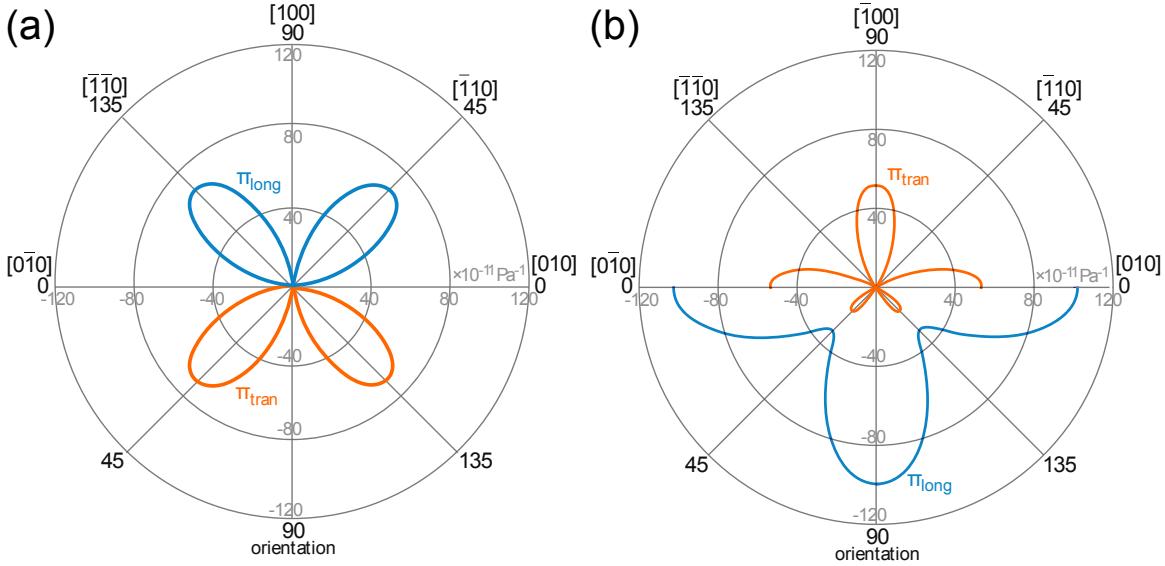


Figure 2.5: Room temperature piezoresistive coefficients in the (100) plane of (a) p-type silicon and (b) n-type silicon. After Kanda [88], ©1982 IEEE.

piezoresistive effect. Notably, $\pi_l \approx -\pi_t$ for p-type piezoresistors, while $\pi_l \approx -2\pi_t$ for n-type piezoresistors. In certain applications, the large transverse effect of p-type piezoresistance can be utilized to improve sensitivity (e.g. full Wheatstone bridge devices) while in others the transverse effect leads to reduced sensitivity (e.g. surface stress sensors).

It is important to emphasize that the polarity of the piezoresistive effects differs for n-type and p-type silicon. A longitudinal tensile stress decreases the resistivity of n-type resistors and increases the resistivity of p-type resistors, while a transverse stress leads to the opposite effects.

Although the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions are optimal, the longitudinal and transverse piezoresistive coefficients can be calculated for an arbitrary direction from [87]

$$\pi_l = \pi_{11} - 2(\pi_{11} - \pi_{12} - \pi_{44})(l_1^2 m_1^2 + l_1^2 n_1^2 + m_1^2 n_1^2) \quad (2.14)$$

$$\pi_t = \pi_{12} + 2(\pi_{11} - \pi_{12} - \pi_{44})(l_1^2 l_2^2 + m_1^2 m_2^2 + n_1^2 n_2^2) \quad (2.15)$$

where l , m , and n are the direction cosines of the direction associated with π_l or π_t , with respect to the crystallographic axes. Kanda [88] presented a graphical representation of

the longitudinal and transverse piezoresistive coefficients in arbitrary directions for (100) silicon (Figure 2.5) as well as the less commonly used (110) and (211) wafer orientations.

There are two common arrangements for measuring piezoresistive coefficients. Smith, Tufte and Stelzer, and others used a tensile test bar (Figure 2.4). This method works particular well for uniformly doped samples. More recent experiments have used four-point bending to apply a known stress to the piezoresistors [89, 91–93]. Four-point bending applies uniform stress to any piezoresistors located between the two inner supports, and can be used to test arbitrary samples without the need for sample gripping. An additional benefit of four-point bending is that the π_{11} , π_{12} and π_{44} coefficients can be readily determined from the same sample by orienting piezoresistors along the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions.

An alternative method for determining the piezoresistive coefficients in arbitrary directions was developed by Richter et al. [94, 95]. Their method used a circular piezoresistor with multiple contacts along the resistor annulus to arbitrarily set the current flow orientation. The circular piezoresistor and typical four-point bending specimen are shown in Figure 2.6.

2.2.2 Piezoresistance factor

Variation in the resistivity of silicon with respect to strain and temperature depends on dopant concentration. Initial experiments by Smith used bars of silicon cut from wafers that were doped while growing the single-crystal ingot [85]. Smith's samples were relatively lightly doped, ranging from 5×10^{14} to 10^{16} cm^{-3} in dopant concentration.

Later, Pfann and Thurston [96] proposed diffusion techniques to integrate doped piezoresistors on the sensor surface. Tufte and Stelzer [97] fabricated diffused piezoresistors and investigated their properties, observing a large reduction in strain sensitivity at high concentration. In contrast with Smith, their dopant concentrations varied from 3×10^{18} to $2 \times 10^{21} \text{ cm}^{-3}$. Tufte and Stelzer reported empirical data on piezoresistive coefficients as a function of both dopant concentration and temperature, which they varied from -200°C to 100°C. Kurtz and Gravel replotted their data and noted that the piezoresistive coefficients decreased approximately with the logarithm of the surface concentration [98].

Kanda presented a theoretical model for variation in piezoresistive coefficients with

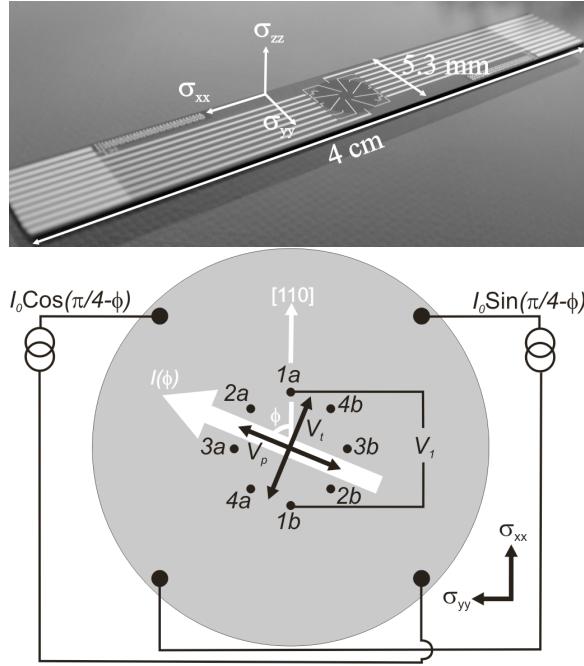


Figure 2.6: (a) Stress sensor chip with a p-type circular piezoresistor in the middle of the chip. (b) Schematic diagram of the circular piezoresistor with a radius of 1700 μm . From Richter et al. [95]. ©2007 IEEE.

dopant concentration and temperature (Figure 2.7) [88]. He also introduced the piezoresistance factor notation for calculating piezoresistive coefficients for arbitrary dopant concentration and temperature,

$$\pi_l = P(n, T)\pi_l^0 \quad (2.16)$$

$$\pi_t = P(n, T)\pi_t^0 \quad (2.17)$$

where π_l^0 and π_t^0 are the longitudinal and transverse piezoresistive coefficients for lightly doped silicon at 300 K for the chosen dopant type and crystallographic orientation (Table 2.2). The piezoresistance factor, P , accounts for variation in the piezoresistive coefficients with concentration and temperature, and in Kanda's model is calculated from

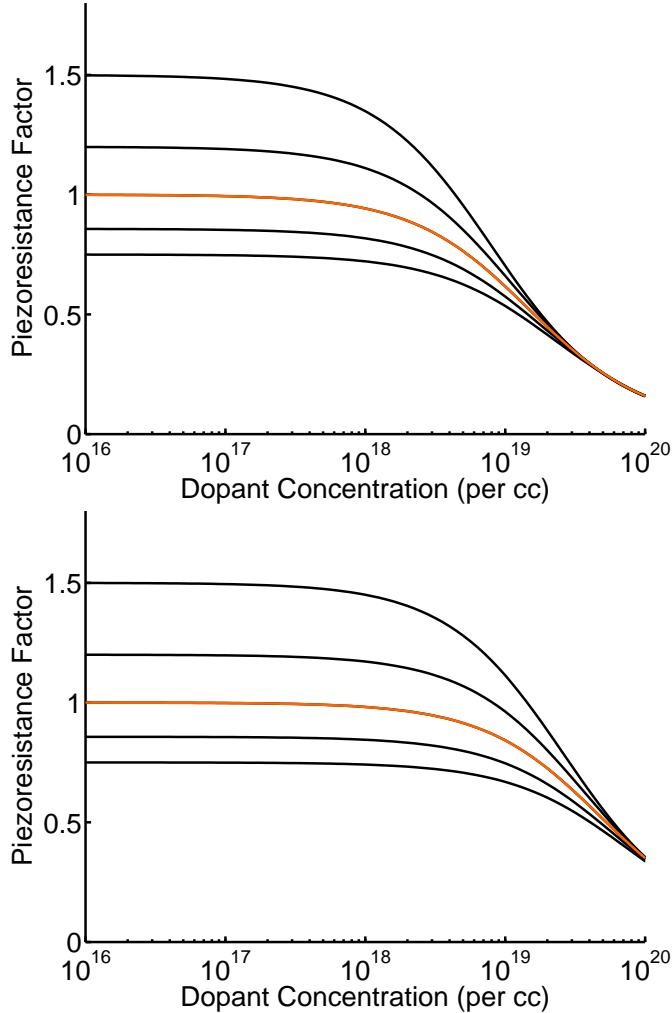


Figure 2.7: Theoretical predictions for the piezoresistance factor for p-type and n-type silicon from Kanda's model, calculated from (2.18). The piezoresistance factor is plotted as a function of dopant concentration for several operating temperatures, and it decreases with increasing dopant concentration and temperature. The piezoresistance factor decreases more quickly with dopant concentration for p-type than n-type silicon in Kanda's model.

$$P = \frac{300}{T} \frac{F'_{s+(1/2)}}{F_{s+(1/2)}} \quad (2.18)$$

where E_F is the Fermi level measured from the conduction or valence band and s is the scattering exponent, while F_s and F'_s are the Fermi integral and its derivative with respect

to E_F . For $s = -\frac{1}{2}$, which corresponds to phonon rather than ionized impurity dominated carrier scattering, the Fermi integral is approximately

$$F_{s+(1/2)} = \ln(1 + \exp(E_F/k_b T)) \quad (2.19)$$

$$F'_{s+(1/2)} = \frac{1}{1 + \exp(-E_F/k_b T)} \quad (2.20)$$

In order to properly account for the density of states available for conduction, the Fermi level is calculated from

$$n = v\sqrt{2} \left(\frac{m_d^* k_b T}{\pi h} \right)^{3/2} F_{1/2}(E_F/k_b T) \quad (2.21)$$

where n is the carrier concentration, h is Planck's constant, v is the number of carrier valleys, and m_d^* is the density of states effective mass [99]. There are 6 valleys for both n-type and p-type silicon, while m_d^* is 1.08 or 0.49 for electrons (n-type) or holes (p-type), respectively [99, 100]. In this problem, n is known and we are solving for E_F . In general, the Fermi integral is

$$F_{s+(1/2)} = (k_b T)^{s+(3/2)} \int_0^\infty \frac{E^{s+(1/2)}}{1 + \exp(E - E_F/k_b T)} dE. \quad (2.22)$$

Solving (2.21) for E_F when $s = 0$ would normally require iterative numerical integration, but accurate analytical approximations for the inverse Fermi integral have been developed. We use the method from Nilsson, which is accurate to within 0.5% for $-10 < E_F/k_b T < 20$ [101, 102]. E_F is calculated from

$$\frac{E_F}{k_b T} = \frac{\ln(u)}{1 - u^2} + v - \frac{v}{1 + (0.24 + 1.08v)^2} \quad (2.23)$$

where

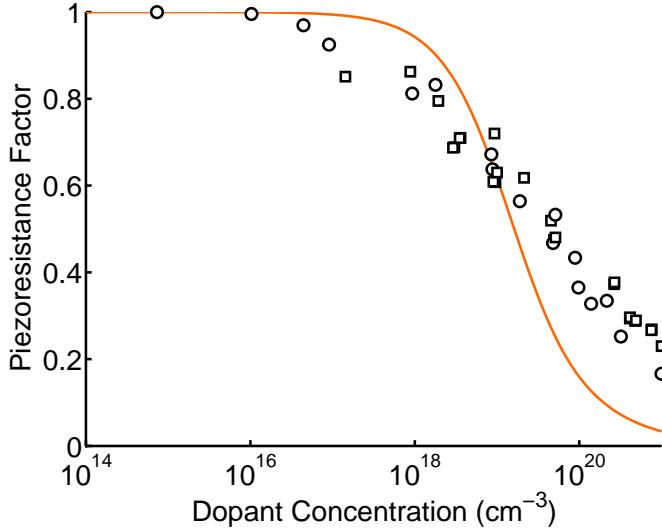


Figure 2.8: Comparison between Kanda’s theoretical piezoresistance factor and experimental data at 300 K. The theory, plotted for p-type silicon, overpredicts P at intermediate dopant concentrations (10^{17} - 10^{18} cm^{-3}) and underpredicts P for high concentrations (10^{19} cm^{-3}). Experimental data is drawn from Refs. [97, 98, 103–106], and p-type and n-type data points are plotted as squares and circles, respectively.

$$u = \frac{\pi^2 N_d}{v\sqrt{2}} \left(\frac{m_d^* k_b T}{h} \right)^{-3/2} \quad (2.24)$$

$$v = \left(\frac{3\sqrt{\pi}u}{4} \right)^{2/3} \quad (2.25)$$

Kanda’s piezoresistance factor, calculated from (2.18), is plotted in Figure 2.7. The model predicts a reduction in the piezoresistance factor with increasing temperature and dopant concentration. Additionally, it predicts a more rapid reduction in P with dopant concentration for p-type silicon than for n-type silicon. The model also predicts that the temperature coefficient of sensitivity (TCS), $\partial P/\partial T$, decreases with increasing dopant concentration, which will be discussed in more detail later.

Comparing Kanda’s model with experimental data (Figure 2.8) we see two things. First, the model significantly underpredicts P at high concentration. Second, there is no substantial difference in P between p- and n-type silicon. Kanda’s calculated values

of P agreed well with the experimental values obtained by Mason [104] for doping concentrations less than $1 \times 10^{17} \text{ cm}^{-3}$ from -50 to 150°C, but they are 21% lower than experimental results at $3 \times 10^{19} \text{ cm}^{-3}$. His calculations only consider carrier scattering with the crystal lattice, and the deviation from experimental data was attributed to additional scattering by ionized dopants at high concentrations.

The experimental data in Figure 2.8 is drawn from Tufte and Stelzer [97, 103], Mason [104], Kerr [105], Kurtz and Gravel [98], and Richter [106]. There is no obvious difference in P between p-type and n-type silicon, although Kanda's theory predicts a higher P for n-type silicon at high dopant concentrations. Both Tufte and Kurtz used diffusion doping to form their piezoresistors, so the coefficients are plotted against the surface dopant concentration. This can lead to an error of up to 15% due to contributions from the more lightly doped silicon deeper in the sample [105]. In contrast, Richter used an SOI structure with uniform doping throughout the device layer to eliminate this effect.

In 2000, Harley developed an empirical fit for the piezoresistance factor based upon the p-type data from Tufte and Stelzer, Mason, and Kerr [75]. In Harley's fit, P is calculated from

$$P = \log_{10} \left(\left[\frac{b}{n} \right]^a \right) \quad (2.26)$$

where $a = 0.2014$, $b = 1.53 \times 10^{22} \text{ cm}^{-3}$, and n is the dopant concentration.

Harley's fit is plotted against the experimental data in Figure 2.9. The fit is accurate for dopant concentrations above $1 \times 10^{18} \text{ cm}^{-3}$ but overpredicts P for lower concentrations. Additionally, the fit is only for $P(N)$ and not $P(N, T)$, so can't be used for devices operating at temperatures above or below 300 K and does not provide information about the temperature coefficient of sensitivity.

Richter et al. presented an updated model in 2008 for the piezoresistance factor of p-type silicon as a function of dopant concentration and temperature [106]. The model was developed from first principles and included both phonon and ionized impurity scattering, in contrast with Kanda's model, which only included phonon scattering. The complete model was reduced to an approximate analytical formula at the end of the 2008 paper,

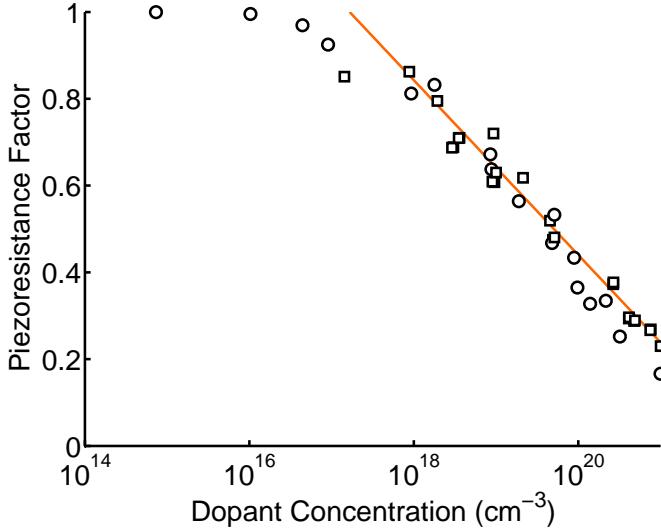


Figure 2.9: Comparison between Harley’s empirical fit and experimental data at 300 K. The fit is accurate for dopant concentrations greater than $1 \times 10^{18} \text{ cm}^{-3}$ but overpredicts the piezoresistance factor for lower concentrations. The fit does not include any temperature dependency.

$$P = \frac{T_n^{-\theta}}{1 + \left(\frac{N}{N_b}\right)^\alpha T_n^{-\beta} + \left(\frac{N}{N_c}\right)^\gamma T_n^{-\eta}} \quad (2.27)$$

where $T_n = T/300$ is the normalized temperature. The fitting coefficients θ , N_b , N_c , α , β , γ and η are presented in Table 2.3.

Richter’s model is compared with the experimental data at 300 K in Figure 2.10a. The fit is excellent at high concentrations and $P \rightarrow 1$ at low concentrations. Although the model was developed specifically for p-type silicon it fits n-type accurately as well.

However, many piezoresistive sensors do not operate at room temperature. In contrast with Harley’s fit, Richter’s model is temperature dependent. Experimental data for temperatures other than 300 K is available from Richter et al. (300, 325 and 350 K) and Tufte and Stelzer (77, 125, 160, 200, 250, 300 and 350 K) [103, 106]. The Richter model is plotted against the experimental piezoresistance factor data in Figure 2.10b for the temperatures investigated by Tufte and Stelzer.

The piezoresistance factor predicted by Richter’s model is accurate over the entire

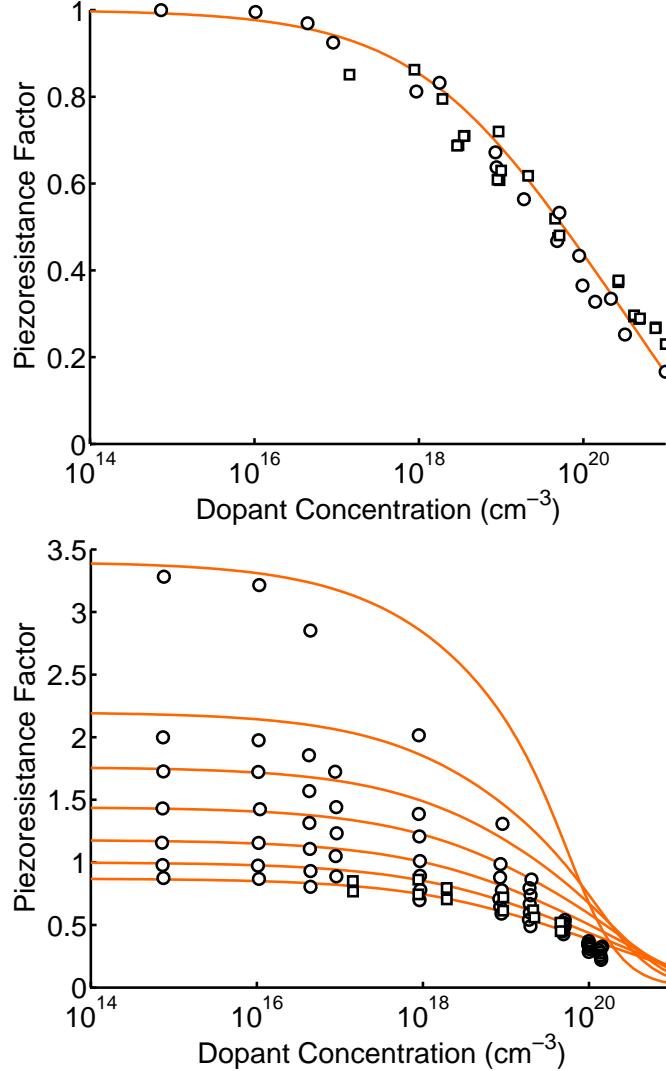


Figure 2.10: (a) Comparison between Richter's theoretical model and experimental data at 300 K. The fit is excellent over the entire range of dopant concentrations. (b) Richter's model compared with experimental data from 77K to 350K. Data from Tufte and Stelzer is for diffusion doped n-type silicon and is plotted as circles [103], while data from Richter et al. is for uniformly doped p-type silicon and is plotted as squares [106]. The model overpredicts the piezoresistance factor for dopant concentrations greater than 10^{16} cm^{-3} and temperatures below 200 K.

temperature range for dopant concentrations below 10^{16} cm^{-3} . As in Kanda's theory, the piezoresistance factor scales as roughly $1/T$ for lightly doped silicon. However, the

Table 2.3: Fitting parameters for the Richter model and the modified Richter model to calculate the piezoresistance factor as a function of dopant concentration as temperature.

Parameter	Richter [106]	Modified Richter
θ (-)	0.9	0.95
N_b (cm^{-3})	6×10^{19}	4.9×10^{19}
N_c (cm^{-3})	7×10^{20}	2.6×10^{20}
α (-)	0.43	0.39
β (-)	1.6	1.35
γ (-)	0.1	0.94
η (-)	3	4.55

predicted value is too large for higher dopant concentrations when the temperature is below 200 K.

In order to provide a more accurate model for low temperature piezoresistor operation, we fit Richter's model (2.27) to the experimental data in Refs. [103, 106] by minimizing the squared sum of the residual error between the theory and data. The resulting modified Richter model is plotted in Figure 2.11 at 300 K and the same temperatures plotted in Figure 2.10. The fit is substantially improved at low temperature without any reduction in accuracy at higher temperatures.

We will use the standard Richter model to design the force probes in the next chapter, but include the note about the modified Richter model in case it proves useful to other designers. For applications below 77 K, it would be worth visiting the data from Morin et al. who measured the piezoresistance factor of p- and n-type silicon of very lightly doped silicon ($< 10^{14} \text{ cm}^{-3}$) at temperatures as low as 5 K [107]. However, their data is not included in the fit because the vast majority of piezoresistance sensors utilized doping levels four to six orders of magnitude higher.

2.2.3 Nonlinearity

Piezoresistive sensors are linear for small mechanical loads. However, several sources of nonlinearity become significant as the load increases: piezoresistive nonlinearity, circuit nonlinearity, and structural nonlinearity. These nonlinear effects have been characterized

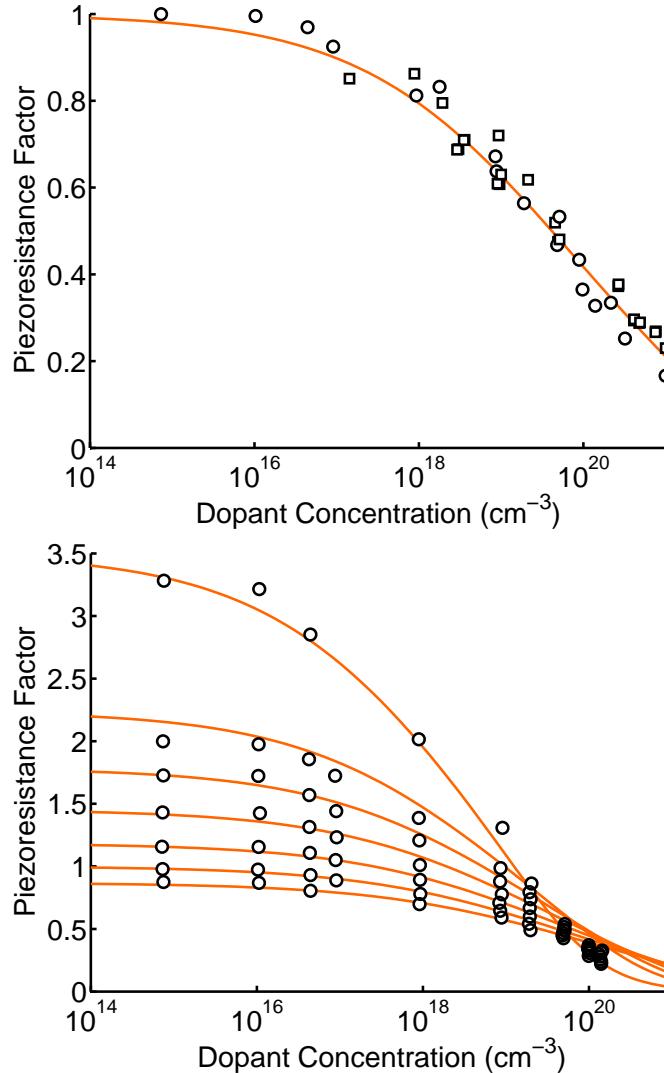


Figure 2.11: Comparison between the modified Richter model and experimental data at (a) 300 K and (b) 77-350 K. The model predicts the piezoresistance factor relatively accurately over the entire temperature and concentration range, while continuing to accurately predict the performance at 300 K.

and can be accounted for when designing precision sensors. Circuit and structural nonlinearity depend greatly on application specifics, but nonlinearity in the piezoresistance coefficients is a fundamental design issue.

The change in resistance with longitudinal stress can be expanded from a linear model to a third order polynomial of the form

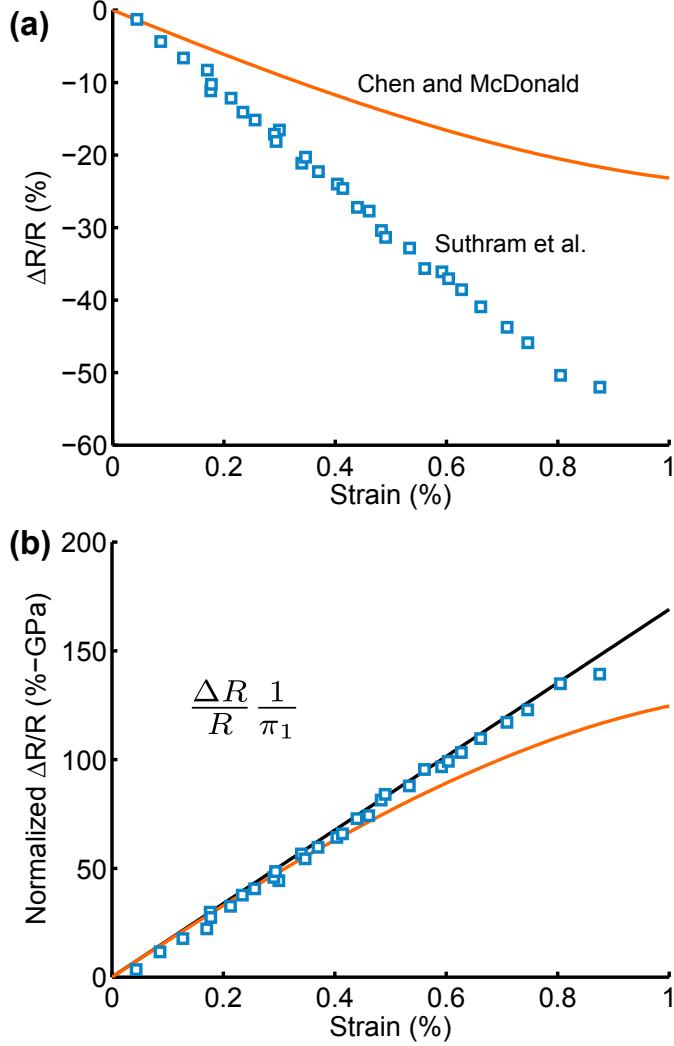


Figure 2.12: Piezoresistive nonlinearity. (a) Relative resistance changes for n-type, $\langle 110 \rangle$ oriented silicon as a function of strain. The samples were doped to 1×10^{18} and $2 \times 10^{20} \text{ cm}^{-3}$ by Suthram and Chen, respectively [89, 108]. (b) By normalizing the resistance change to the first-order piezoresistivity coefficient (π_1), we can account for the substantial difference in dopant concentration and illustrate the deviation from linearity, which increases with dopant concentration.

$$\frac{\Delta R}{R} = \pi_1 \sigma_l + \pi_2 \sigma_l^2 + \pi_3 \sigma_l^3 \quad (2.28)$$

where π_1 , π_2 and π_3 are the first-, second- and third-order coefficients.

Matsuda et al. [109, 110] calculated and measured the piezoresistive coefficients and third-order effects for both p-type and n-type silicon for the three major crystallographic orientations with strain up to 0.1%. Higher strain levels were difficult to measure at the time due to surface defects in the silicon. Measuring higher strain values is necessary to accurately quantify the second- and third-order coefficients.

Addressing this problem, Chen and MacDonald [108] co-fabricated a microactuator and a 150 μm long, 150 nm diameter single-crystal silicon fiber, allowing larger strains to be measured. With the increased range of strain, the second and third order fit for piezoresistive coefficients were quantified more accurately. The fiber was $\langle 110 \rangle$ oriented and doped to $2 \times 10^{20} \text{ cm}^{-3}$. Their best-fit coefficients were $\pi_1 = 1.86 \pm 0.01 \times 10^{-10} \text{ Pa}^{-1}$, $\pi_2 = 0.12 \pm 0.01 \times 10^{-19} \text{ Pa}^{-2}$ and $\pi_3 = 0.100 \pm 0.003 \times 10^{-28} \text{ Pa}^{-3}$.

Similarly, Suthram et al. measured the mobility enhancement in $\langle 110 \rangle$ oriented nMOS channels doped to $1 \times 10^{18} \text{ cm}^{-3}$ up to 1.5 GPa of tensile stress [89]. They used a four-point bending system to apply large strains to their samples, and corrected for metal interconnect resistance changes in calculating the silicon mobility enhancement.

The relative resistance changes ($\Delta R/R$) measured by Chen and Suthram are plotted in Figure 2.12a. Suthram et. al measured substantially larger resistance changes due to the lower doping of their samples. As noted earlier, $\langle 110 \rangle$ oriented n-type silicon does not maximize the piezoresistive coefficients (Table 2.2).

The first-order coefficient (π_1) varies substantially with dopant concentration and temperature, so we can normalize the expansion to obtain

$$\frac{\Delta R}{R} \frac{1}{\pi_1} = \sigma_l + \pi'_2 \sigma_l^2 + \pi'_3 \sigma_l^3 \quad (2.29)$$

where $\pi'_2 = \pi_2/\pi_1$ and $\pi'_3 = \pi_3/\pi_1$. The relative second- and third-order coefficients can be used to approximate the magnitude of nonlinearity regardless of variations in dopant concentration and temperature. Using Chen and McDonald's best-fit parameters, $\pi'_2 = 6.45 \times 10^{-11} \text{ Pa}^{-1}$ and $\pi'_3 = 5.38 \times 10^{-20} \text{ Pa}^{-2}$.

The normalized resistance changes ($\Delta R/R/\pi_1$) measured by Chen and Suthram are presented in Figure 2.12b. No substantial deviation from linearity was observed by Suthram et al. for strains up to 0.8% (1.35 GPa). Substantially greater deviation was observed by

Chen and McDonald, and a 24% decrease in the longitudinal piezoresistive coefficient was observed at 1% strain. One possible explanation for the results is increased piezoresistive nonlinearity with increasing dopant concentration.

2.2.4 Theoretical models

The discovery of such large piezoresistive effects demanded a theory of the underlying physics. This section discusses the prevailing theories at the time of Smith's measurements as well as more recent advances.

The theories of semiconductor piezoresistance are grounded in one-dimensional descriptions of electron and hole transport in crystalline structures under strain, with some extensions to three dimensions, crystal defects, electric potentials and temperature effects. The various models require some understanding of bandgap energy models, wave mechanics, and quantum effects; the interested reader is referred to [80, 111–113].

At the time of Smith's piezoresistance measurements, existing theories were based on shifts in bandgap energies. The band structure of diamond (Figure 2.13) was first calculated by Kimball in 1935 [114], and that of silicon by Mullaney in 1944 [115]. In 1950, Bardeen and Shockley presented a model for mobility changes in semiconductors subjected to deformation potentials and compared both predicted and measured conductivity changes in the bandgap with dilation [116]. This work served as the basis for later analyses, such as that of Herring [117, 118] and Long [119].

The mobilities and effective masses of electrons and holes are significantly different from one another and fluctuate under strain. Piezoresistors with n- and p-type doping exhibit opposite trends in resistivity change with strain and different direction-dependent resistivity change magnitudes under stress (Tables 2.2 and 2.1). The magnitudes and signs of the piezoresistive coefficients depend on a number of factors including dopant concentration, temperature, crystallographic direction, and the relative directions of the voltage, current and stress to one another and the crystallographic axes.

The relationship between carrier characteristics and strain has been investigated both experimentally [85, 87, 97] and analytically [88, 96, 116, 118, 120, 121]. Focusing on n-type silicon, these early studies utilized either effective mass or energy band calculations

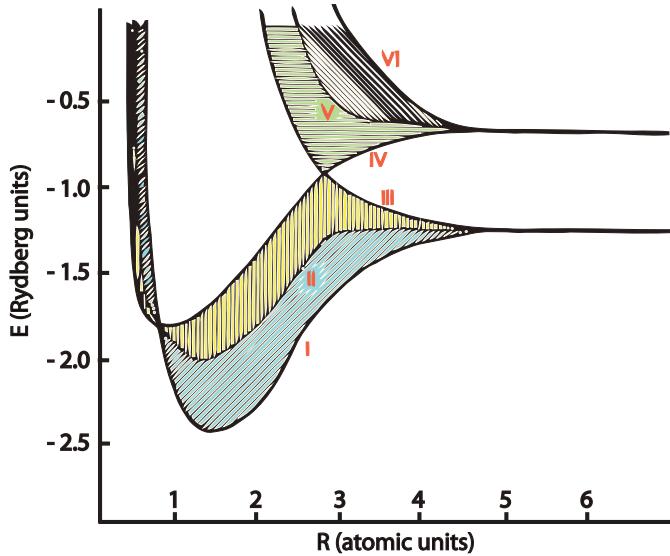


Figure 2.13: Energy band splitting with strain in diamond-type lattices. The energy bands are plotted for a diamond lattice as a function of strain. Besides the four shaded bands there are four bands of zero energy width, two following curve IV and two following curve VI. After Kimball [114]. ©1935 American Institute of Physics.

with wave propagation in one direction at a time. The change in mobility (and thus, conductivity) with lattice strain is attributed to band warping or bending and the non-uniform density of states.

The implications for the related large mobility and resistance changes were not realized prior to Smith's discovery [122, 123]. Following Bardeen and Shockley's models for mobility changes with deformation potentials, more refined models of transport and energy band structure based on new experimental work became available. In 1955, Herring proposed his Many-Valley model, which adequately explained piezoresistance for n-type silicon and germanium [96, 116, 118, 120, 121, 124–127].

Herring's Many-Valley model for n-type silicon proposes three symmetrical valleys along the $\langle 100 \rangle$ direction [118]. His model projects the band energy minima in three orthogonal directions as locations of constant minimum energy. The minimum energy of each valley lies along the centerline of the constant energy ellipsoid. Electrons have a higher mobility along the direction perpendicular to the long axes of the ellipsoids. Since electrons occupy lower energy states first, they are found in these regions bounded by

ellipsoids of constant low-energy. These ellipsoids, bounded by higher-energy regions, are referred to as valleys. With strain however, the symmetry is broken and the ellipsoids are asymmetrically dilated or constricted. This results in an anisotropic change in conductivity proportional to strain.

Most models represent the direction dependence of bandgap and electron energies by either a wave vector (k) or momentum (p) and the effective masses of the carriers. The energy surfaces for electron mobility are accordingly represented in k - or p -space. The wave propagation is confined to quantum states by the periodicity of the lattice, and edges in the band diagrams correspond to the edges of the Brillouin zone (smallest primitive cell, or unit cell, of the reciprocal lattice) oriented in a direction of interest [80].

In the unstrained silicon crystal, the lowest conduction band energies (valleys) or highest mobility orientations are aligned with the $\langle 100 \rangle$ directions. The conduction electrons are thus imagined to be lying in six equal groups aligned with the three $\langle 100 \rangle$ directions. For any valley, the mobility is the lowest when parallel to the valley direction, and the highest when perpendicular to the valley. For example, an electron in the z valley has higher mobility in the x and y directions than in the z direction. The effective mobility is the average from the three valleys [127]. Net electron conductivity is the sum of the conductivity components along the three valley orientations and is independent of direction in the unstrained crystal.

Uniaxial elongation increases the band energy of the valley parallel to the strain and transfers electrons to perpendicular valleys, where they have higher mobility in the strain direction. Tension shifts the electron distribution amongst the valleys, increasing the conductivity in the direction of the tension (longitudinal effect) and reducing the conductivity in the directions perpendicular to the tension (transverse effect). Compression has the opposite effect. The many-valley model accurately models n-type silicon, and recent extensions improve the accuracy at high transverse electric fields and low temperature [128].

The piezoresistance theory for n-type semiconductors continued to be refined from 1954 onward, but until recently piezoresistance in p-type silicon was not fully understood [132, 133]. However, recent computational advances have enabled an improved understanding of p-type piezoresistance [113, 134–136]. This is important because most

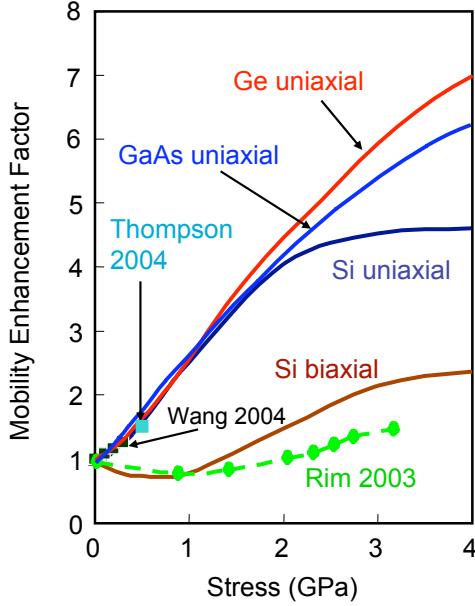


Figure 2.14: Hole mobility enhancement in semiconductors as a function of stress. Sun et al. compared their experimental results with those from several other research groups [129–131] and noted that a hole mobility enhancements up to 2 GPa in Si and up to 4 GPa in Ge and GaAs. Reprinted with permission from Sun [113], ©2007 American Institute of Physics.

research and commercial piezoresistive devices are p-type and models had been largely based on empirical results. Theoretical studies based on the strain Hamiltonian [137–139], deformation potentials in strained silicon, and cyclotron resonance results have revealed several factors that affect hole mobilities in semiconductors, such as band warping, splitting, and effective mass changes [140–144].

Historically, piezoresistive technology drew from mainstream IC research and continues to do so. Now, with the strong interest in strain engineering to increase carrier mobility in ICs, the situation has reversed and mainstream semiconductor technology is drawing on findings of piezoresistive research. Strain engineered materials, such as silicon-germanium ($\text{Si}_{1-x}\text{Ge}_x$), can increase the channel mobility in MOSFET devices [89, 113, 145, 146]. Suthram et al. [89] applied large uniaxial stress to nMOS transistors and showed that piezoresistive coefficients were constant while the electron mobility enhancement increased linearly for stresses up to 1.5 GPa. Figure 2.14 shows plotted hole

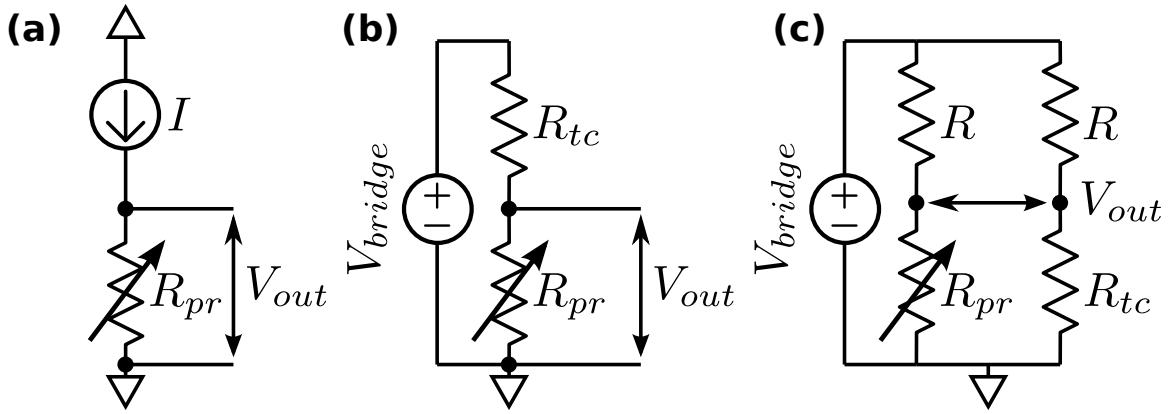


Figure 2.15: A resistance change can be transduced into a voltage change using a (a) four-wire resistance measurement, (b) voltage divider or (c) Wheatstone bridge. The Wheatstone bridge allows circuit output nulling and temperature compensation.

mobility enhancement factor for several semiconductors as a function of stress.

2.3 Signal conditioning

The electrical resistance of a piezoresistor can be monitored using a variety of techniques. The most direct method is to force a known current through the piezoresistor (R_{pr}) and measure the voltage drop. This is typically accomplished using a four-wire measurement (Figure 2.15a), where one pair of the wires is connected to a current source and the other pair is connected to a voltmeter. As long as the input impedance of the voltmeter is much larger than the piezoresistor resistance, negligible current will flow in the sensing wires and their resistance will not affect the measurement.

The resistance of the piezoresistor depends upon the temperature and stress according to

$$R_{pr} = R_0(1 + \alpha\Delta T + \pi_l\Delta\sigma_l + \pi_t\Delta\sigma_t) \quad (2.30)$$

where R_0 is the resistance for a reference temperature and stress, α is the temperature coefficient of resistance (TCR), π_l and π_t are the longitudinal and transverse piezoresistive coefficients of the piezoresistor, and ΔT , $\Delta\sigma_l$ and $\Delta\sigma_t$ are the changes in temperature,

longitudinal stress and transverse stress from their reference values.

The initial output voltage of the circuit is $V_{\text{out}} = IR_0$, while the output change can be calculated from $\Delta V_{\text{out}} = IR_{\text{pr}} = IR_0(1 + \alpha\Delta T + \pi_l\Delta\sigma_l + \pi_t\Delta\sigma_t)$. Assuming that the piezoresistor is uniaxially loaded along its longitudinal axis ($\sigma_t = 0$) then the relative change in output voltage is

$$\frac{\Delta V_{\text{out}}}{V_{\text{out}}} = 1 + \alpha\Delta T + \pi_l\Delta\sigma_l. \quad (2.31)$$

The relative change in output voltage due to temperature and stress changes can be calculated by taking the partial derivatives of (2.31), yielding

$$\frac{\partial}{\partial T} \left(\frac{\Delta V_{\text{out}}}{V_{\text{out}}} \right) = \alpha \quad (2.32)$$

$$\frac{\partial}{\partial \sigma_l} \left(\frac{\Delta V_{\text{out}}}{V_{\text{out}}} \right) = \pi_l. \quad (2.33)$$

Depending on the magnitude of temperature fluctuations, the temperature induced change in output may be substantially larger than that due to mechanical loading.

Temperature compensation can be improved by modifying the circuit layout. The most common temperature compensation technique is to use two identical piezoresistors in a configuration where the common-mode resistance change is rejected. For example, the piezoresistor that experiences mechanical loading (R_{pr}) can be placed in a voltage divider with an identical, temperature compensation resistor that does not experience any loading (R_{tc}) as shown in Figure 2.15b.

The output from the voltage divider is $V_{\text{out}} = V_{\text{bridge}}R_{\text{pr}}/(R_{\text{pr}} + R_{\text{tc}})$, and assuming that R_{pr} and R_{tc} are perfectly matched the initial output will be $V_{\text{bridge}}/2$. The relative change in output voltage, assuming uniaxial longitudinal loading once more, is

$$\frac{\Delta V_{\text{out}}}{V_{\text{out}}} = \frac{1 + \alpha\Delta T + \pi_l\Delta\sigma_l}{1 + \alpha\Delta T + \frac{1}{2}\pi_l\Delta\sigma_l}. \quad (2.34)$$

Taking partial derivatives of (2.34) with respect to T and σ_l , we obtain

$$\frac{\partial}{\partial T} \left(\frac{\Delta V_{\text{out}}}{V_{\text{out}}} \right) = \frac{\alpha(1 + \alpha\Delta T + \frac{1}{2}\pi_l\Delta\sigma_l) - \alpha(1 + \alpha\Delta T + \pi_l\Delta\sigma_l)}{(1 + \alpha\Delta T + \frac{1}{2}\pi_l\Delta\sigma_l)^2}$$

$$\frac{\partial}{\partial \sigma_l} \left(\frac{\Delta V_{\text{out}}}{V_{\text{out}}} \right) = \frac{\pi_l(1 + \alpha\Delta T + \frac{1}{2}\pi_l\Delta\sigma_l) - \frac{1}{2}\pi_l(1 + \alpha\Delta T + \pi_l\Delta\sigma_l)}{(1 + \alpha\Delta T + \frac{1}{2}\pi_l\Delta\sigma_l)^2}.$$

For small changes in temperature and stress ($\alpha\Delta T \ll 1$ and $\pi_l\Delta\sigma_l \ll 1$) we can simplify the results to obtain

$$\frac{\partial}{\partial T} \left(\frac{\Delta V_{\text{out}}}{V_{\text{out}}} \right) = 0 \quad (2.35)$$

$$\frac{\partial}{\partial \sigma_l} \left(\frac{\Delta V_{\text{out}}}{V_{\text{out}}} \right) = \frac{1}{2}\pi_l. \quad (2.36)$$

$$\frac{\partial}{\partial \sigma_l} \left(\frac{\Delta V_{\text{out}}}{V_{\text{out}}} \right) = \frac{1}{2}\pi_l. \quad (2.37)$$

The voltage divider circuit eliminates first-order output voltage temperature changes, although the sensitivity is a factor of two smaller. The fact that $\pi_l \approx -\pi_t$ for p-type silicon piezoresistors can be utilized by placing both piezoresistors on the mechanically loaded element, one oriented in the longitudinal direction and the other in the transverse direction, to restore the original stress sensitivity of the sensor.

However, in both the 4-wire and voltage divider circuits the initial output offset is much larger than the stress-induced voltage changes. If data acquisition circuits had infinite resolution this would not pose a problem, however the dynamic range in a practical measurement circuit is limited by the analog-to-digital converter (ADC) resolution, and the ratio between the measurement noise floor and full scale output.

The most common measurement circuit, which achieves temperature compensation and allows nulling of the initial output offset is the Wheatstone bridge (Figure 2.15c). If only one resistor in the bridge is mechanically loaded (quarter-active configuration) then the stress sensitivity of the sensor is reduced by a factor of four from the 4-wire configuration (2.33). As in the voltage divider case, the sensitivity can be increased by using a half- or full-bridge configuration with either two or four piezoresistors placed on the mechanically

loaded element. In the half-bridge configuration the temperature and stress sensitivities are identical to (2.35) and (2.37). We emphasize that voltage divider and Wheatstone bridge circuits are inherently nonlinear, with the magnitude depending on how many resistors in the circuit change resistance with strain, and both circuit and piezoresistive nonlinearity may need to be compensated for in precision applications [147].

The Wheatstone bridge can be biased with either a voltage or current source, and feedback resistors can be added to the bridge for improved temperature compensation [148]. Additionally, a pair of differential amplifiers can be used instead of a Wheatstone bridge [149] at the cost of additional complexity and increased noise depending on the particular design of the piezoresistors and differential amplifier.

Pfann and Thurston [96] recognized the benefits of using transverse and shear piezoresistance effects in conjunction with longitudinal piezoresistance for devices. Many of their geometries employed a full Wheatstone bridge with two longitudinal and two transverse piezoresistors to increase sensitivity and compensate for resistance changes due to temperature. Notably, they proposed integrating the piezoresistors with the force collecting structure and discussed the advantages and disadvantages of a number of geometries for various types of measurements. They anticipated most of the geometries widely employed today.

2.3.1 Process variation and temperature compensation

In practice, fabrication process variations give rise to mismatch in R_0 , α and π_l between nominally identical silicon resistors so perfect temperature compensation is never achieved. Mismatch can be minimized by placing all of the resistors as close as possible to each other to minimize variations, and output from the signal conditioning circuit can be calibrated and compensated to achieve high accuracy. Signal conditioning electronics typically perform amplification, filtering, additional temperature compensation and nonlinearity correction, and can be monolithically integrated with the piezoresistive sensors or fabricated separately.

Modern electronics can ultimately correct all repeatable errors. If a piezoresistive sensor is heated and then cooled to the initial temperature, then the output should be

the same for the same input. However, small differences are usually observed between temperature cycles. This thermal non-repeatability is one of the fundamental limits of sensor accuracy, not correctable with signal conditioning circuits.

Prior to 1980 most of the temperature compensation circuits for piezoresistive sensors employed trim resistors with or without low noise bipolar junction transistor (BJT) based amplifiers. Laser-trimmed resistors were used to adjust the offset, span, nonlinearity and other errors of piezoresistive sensors.

CMOS circuitry became the dominant source of signal conditioning after 1990. The need for even smaller, more accurate, and cheaper sensors was an impetus for the transition to CMOS. The bipolar technology, an analog technology, does not offer the functionality of a digital technology (CMOS) measured in terms of cost per power per functionality. The push toward CMOS technology evolved with the availability of non-volatile memory (NVM). The laser-trimmed resistors were then replaced with digital-to-analog converters (DAC) and memory. By use of double correlated sampling, offset and low frequency noise of the CMOS circuit are sampled and stored on a capacitor and in the next cycle they are subtracted from the original signal. Hence rendering the CMOS amplifier almost ideal in the low frequency region relative to the sampling frequency.

In CMOS, the need for digital output is easily addressed by integrating the ADC with the sensor. A majority of integrated designs incorporate sigma-delta converters (Figure 2.16) as the primary ADC architecture due to its inherent robustness [150, 151]. Through the combination of CMOS integration and NVM, the need for laser trimming as a means of sensor compensation was eliminated and the power of digital technology was used to compensate and calibrate the piezoresistive sensors. This technology enabled unprecedented sensor accuracy at very low cost [152].

There are two main architectures for piezoresistor temperature compensation: fully digital compensation and digitally controlled analog compensation [153]. The digital signal path architecture uses an ADC to digitize the Wheatstone bridge signal and a temperature sensor signal, then uses a predefined model to perform signal compensation. If an analog output is needed then the compensated digital data is fed to a DAC. This architecture is the most flexible but has some inherent problems that limit its use in control loops. One of the main drawbacks is the delay time from the input to output. The ADC, the microprocessor,

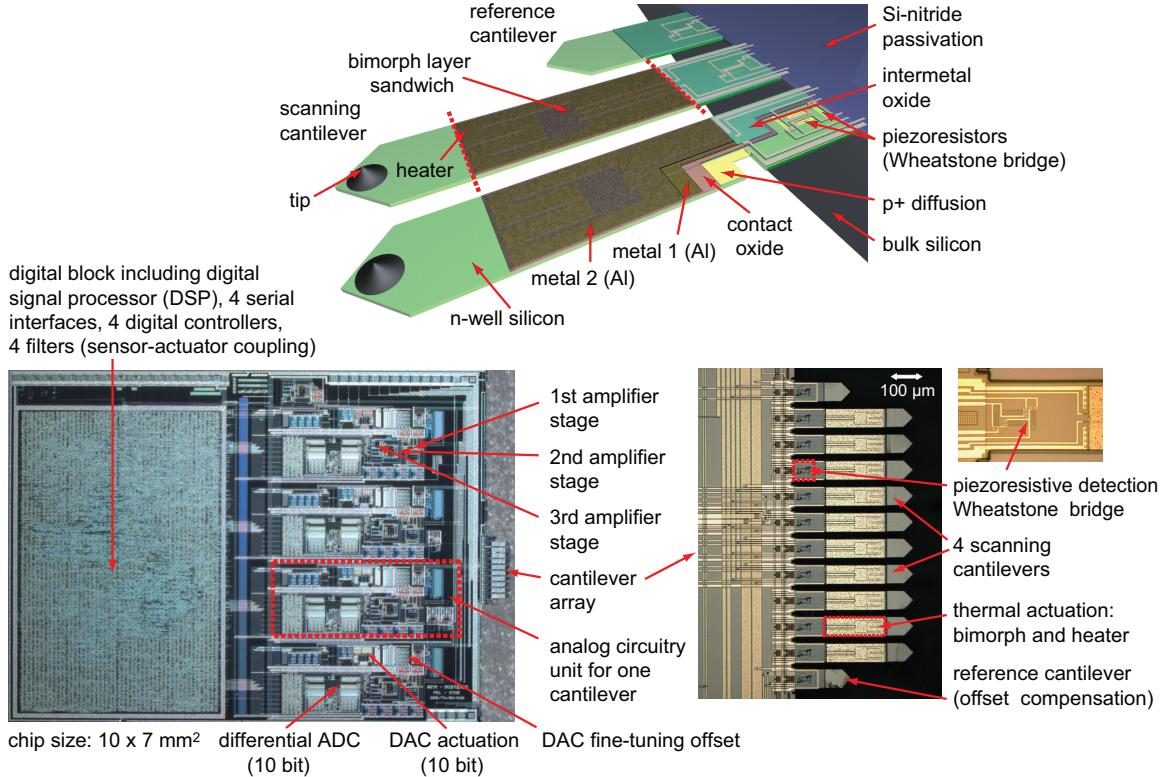


Figure 2.16: (a) CMOS integrated piezoresistive cantilever array (two scanning cantilevers and one reference cantilever) (b) Micrograph of the overall sensor CMOS signal conditioning circuit (c) Array of 12 cantilevers (the inner ten can be used for scanning while the outer two serve as a reference). The dimensions of the scanning cantilevers are $500 \mu\text{m} \times 85 \mu\text{m}$. From Hafizovic et al. [70], reprinted with permission from PNAS.

and the DAC all need processing time, this dead time may not be tolerated in feedback control. In contrast, the digitally controlled analog path architecture takes advantage of the fact that temperature is a slow signal. Hence, delay in processing of the temperature signal is not of concern. The digitized temperature signal is mathematically processed and controls an analog path by changing the gain and the offset of wide-band amplifiers, which inherently have small delays.

The question of integration of the sensor with electronics mainly depends on the application. Generic signal conditioning circuitry consists of an excitation circuit, a bridge circuit, an amplifier, and a filter [42]. These components all contribute to the noise figure of the system (Figure 2.17). Ishihara et al. developed the first CMOS integrated

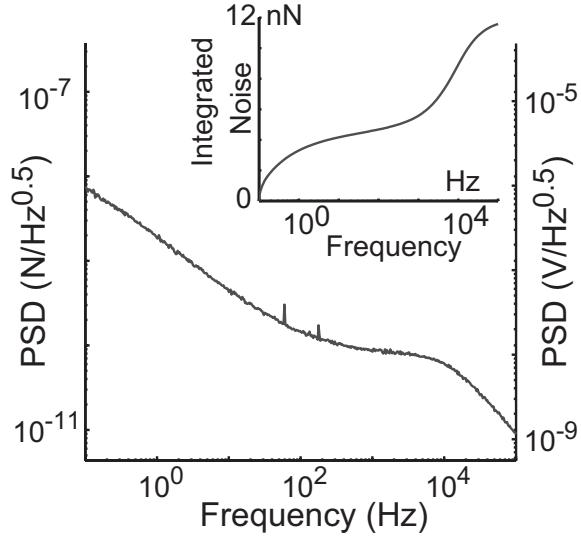


Figure 2.17: The power spectral density (PSD) and integrated force noise of a measurement system using an AD622 instrumentation amplifier and piezoresistor bridge. All components in a signal conditioning circuit contribute to the noise and resolution of the system. Courtesy of Sung-Jin Park [154], reprinted with permission from PNAS.

silicon diaphragm pressure sensors in 1987 [155]. Since then, CMOS circuitry has been integrated with piezoresistive MEMS devices, such as AFM [70, 156–162] and force or stress sensors [163–171]. Baltes et al. reviewed advances in the CMOS-based MEMS until 2002, including microsensors and packaging, and discussed some key challenges and applications for the future [172, 173].

The most common integration approach in commercial MEMS sensors is currently hybrid integration with MEMS and CMOS fabrication handled at separate foundries and package-level integration. This approach avoids the fabrication process compromises inherent to monolithic integration, and avoids wasting the relatively expensive area on the CMOS die with large MEMS structures. Current examples of products that utilize hybrid integration include the ADXL345 (in contrast with earlier monolithically integrated accelerometers from Analog Devices) and the Bosch BMP180 piezoresistive barometer.

When there is sufficient room on the sensor, such as in pressure sensors and accelerometers which benefit from a large force collection membrane or proof mass, a full-bridge configuration is typically used with four active piezoresistors. This approach provides the highest possible sensitivity and first-order temperature compensation. However, other

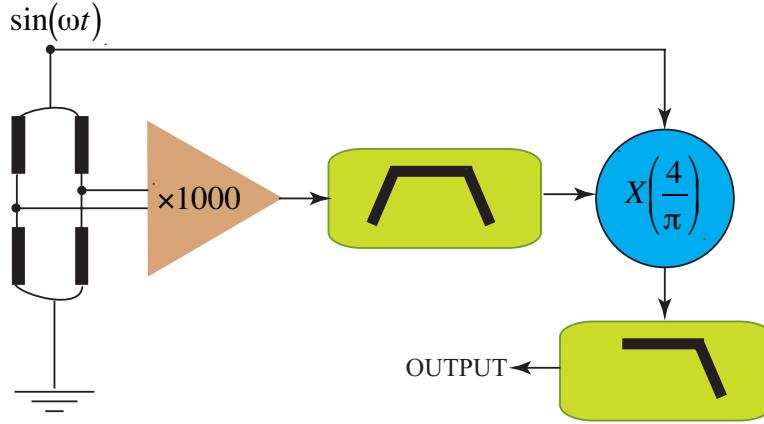


Figure 2.18: Modulation-demodulation circuit for low frequency sensor readout. Modulation-demodulation allows the sensor signal to be amplified at an arbitrary frequency, eliminating the effect of amplifier 1/f noise on the signal. Following Ref. [24]. ©2009 IEEE.

piezoresistive sensors, such as cantilevers, should typically be made as small as possible, making it challenging to optimize performance while placing all four piezoresistors on the cantilever. If only one piezoresistor can fit on the cantilever, an off-cantilever temperature compensation is typically used in a quarter-active Wheatstone bridge. More recently, Chui et al. took advantage of the insensitivity of p-type piezoresistors to strain in the $<100>$ direction by orienting one cantilever leg in the $<110>$ direction and the other in the $<100>$ direction, yielding an order of magnitude improvement in thermal disturbance rejection [174].

2.3.2 Heterodyne biasing

The Wheatstone bridge is typically biased with a DC bias. However, in certain situations the 1/f noise of the signal conditioning electronics can be greater than that of the piezoresistor sensor, which requires an alternative approach. One solution is to use a modulation-demodulation circuit, in which the bridge is sinusoidally biased at a frequency above the 1/f noise corner frequency of the electronics [175]. The amplitude of the Wheatstone bridge output remains proportional to the resistance change of the sensor, and the modulated bridge output is amplified, bandpass-filtered to reduce the effect of noise

folding, and demodulated before data acquisition. Modulation-demodulation allows the sensor signal to be amplified at an arbitrary frequency, eliminating the effect of amplifier 1/f noise on the signal or other spectrally limited noise sources (Figure 2.18).

2.4 Sensor resolution

The design of the signal conditioning circuit plays a large role in determining the resolution of a piezoresistive sensor. Sensor resolution is defined as the smallest signal that can be reliably detected. The minimum detectable signal without averaging multiple trials is commonly equated to the RMS noise of the measurand (i.e. signal-to-noise ratio of unity) according to

$$\text{Resolution} = \frac{V_{\text{noise}}}{S} \quad (2.38)$$

where V_{noise} is the RMS voltage noise and S is the voltage-referred sensitivity with respect to the measurand (e.g. force, displacement or pressure) of the sensor. In this section we will present models for calculating the noise and sensitivity of arbitrary piezoresistive sensors.

2.4.1 Noise

Noise is any unwanted disturbance that interferes with the signal being measured. A variety of noise sources lead to random fluctuations in the output of a piezoresistive sensor. The dominant noise sources in piezoresistive sensors are electrical, although there are exceptions. For example, temperature fluctuations typically limit resolution for measurements that are hours to days in length, while Brownian motion of the atoms in the piezoresistive sensor (thermomechanical noise) provides a fundamental lower limit for optimized nanomechanical sensors.

We will divide piezoresistor noise into intrinsic and extrinsic sources. Intrinsic sources, which depend solely on the design of the piezoresistive sensor and signal conditioning circuitry, are the focus of our design methods because they are readily quantified. Extrinsic sources are associated with the specific measurement environment, such as inductive or

capacitive line pickup [42], carrier injection due to high intensity light illumination, and variation in the ambient temperature or humidity [176]. Extrinsic noise sources can vary enormously with the measurement situation so are not included in the piezoresistor performance models.

Noise processes are stochastic, but can be modeled using statistical methods in both the time and frequency domains. The goal of our noise analysis will be to determine amplitude and spectral distribution of the noise sources for a general piezoresistive sensor and signal conditioning circuit.

We will discuss a generic noise source in order to introduce these concepts before moving onto piezoresistive sensor and signal conditioning noise sources.

A brief introduction to noise

Consider a random voltage noise source, v_N , that is passed through a brickwall bandpass filter with center frequency f_0 and bandwidth Δf . Due to prior knowledge, we expect the noise to have a mean value of zero and a standard deviation of σ . If we assume that the noise is normally distributed, then its probability density function (PDF) is

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right). \quad (2.39)$$

where $p(x)$ is the probability of observing $v_N = x$. We assume that the noise source is stationary so that the PDF doesn't change with time. In order to verify the PDF, we measure the noise amplitude a large number of times and compute the mean and mean-square values as

$$\overline{v_N} = \frac{1}{N} \sum_{t=1}^N v_N(t) = 0 \quad (2.40)$$

$$\overline{v_N^2} = \frac{1}{N} \sum_{t=1}^N v_N^2(t) = \sigma^2 \quad (2.41)$$

We can compare these values with the PDF model by computing the mean and mean-square values as

$$\overline{v_N} = \int_{-\infty}^{\infty} xp(x)dx = 0 \quad (2.42)$$

$$\overline{v_N^2} = \int_{-\infty}^{\infty} x^2 p(x)dx = \sigma^2 \quad (2.43)$$

to find that the PDF model does in fact accurately describe the noise source.

The RMS noise amplitude, equal to the square root of the mean-square value, can then be calculated as σ . In general, any noise source with a Gaussian probability distribution will have an RMS amplitude equal to the probability distribution's standard deviation. The peak-to-peak amplitude of the same signal is usually estimated to be 6σ , because the chance of observing a voltage outside of that range is small (0.3%).

Now suppose that we discovered v_N is actually the sum of two independent noise sources, v_{N1} and v_{N2} each with their own standard deviations, σ_1 and σ_2 . The mean and mean-square values of the total noise can be recalculated in terms of the noise components from

$$\overline{v_N} = \overline{v_{N1}} + \overline{v_{N2}} \quad (2.44)$$

$$\overline{v_N^2} = \overline{v_{N1}^2} + \overline{v_{N2}^2} + 2\overline{v_{N1}v_{N2}}. \quad (2.45)$$

The third term in (2.45) depends on the correlation between the two smaller noise sources. If they are uncorrelated, then the average value of their product is equal to zero and the term drops out. Most noise sources are uncorrelated, for example the thermal noise of two separate resistors. Correlated noise sources can arise when there is coupling between noise sources (e.g. capacitive, inductive or thermal coupling) or when the noise is due to an external source that affects a circuit in multiple locations simultaneously. The dominant noise sources in piezoresistive sensors are uncorrelated.

Noise is often expressed in units of power (which is proportional to V^2), because the total noise power increases linearly with the addition of uncorrelated noise sources. Assuming that v_{N1} and v_{N2} are uncorrelated, the RMS noise is

$$v_N^{\text{RMS}} = \sqrt{\overline{v_{N1}^2} + \overline{v_{N2}^2}}. \quad (2.46)$$

An important concept in adding uncorrelated noise sources is that the RMS noise mainly depends on the largest noise source. For example, if $\overline{v_{N1}} = \overline{v_{N2}} = 10\text{nV}$ then both contribute equally to the total noise and $v_N^{\text{RMS}} = 14.1\text{nV}$. But if $\overline{v_{N2}^2}$ decreases by a factor of 5, the total noise only decreases slightly (from 14.1 to 10.2 nV) because it is limited by the largest noise source. The fact that overall noise is mainly determined by the largest noise source will have broad implications for optimizing the design of piezoresistive sensors, particularly the tradeoff between Johnson and 1/f noise.

We have kept the frequency (f_0) and bandwidth (Δf) of the bandpass filter constant so far. But both will change depending on the frequency response of the amplifiers and filters in the signal conditioning circuit. In general, noise is generated by broadband sources over a wide frequency range, and by applying a narrow-band filter we have selected a small frequency range to observe. The mean-square amplitude of our noise source can be normalized to the measurement bandwidth to yield its power spectral density:

$$S_N(f) = \frac{\overline{v_N^2}}{\Delta f} = \frac{\sigma^2}{\Delta f} \quad (2.47)$$

The power spectral density (PSD) has units of power per unit frequency, such as V^2/Hz or A^2/Hz . The RMS amplitude of any noise source can be calculated from

$$v_N^{\text{RMS}} = \sqrt{\int_{f_{\min}}^{f_{\max}} S_N(f) df} \quad (2.48)$$

for an arbitrary frequency range if $S(f)$ of the noise source is known, where f_{\min} and f_{\max} are the lower and upper measurement frequency limits. Multiple uncorrelated noise sources can be treated by summing their power spectral densities. By integrating the power spectral density over frequency, we are effectively treating the noise at each frequency as an independent noise source and adding their mean-square values.

Assuming that the power spectral density of our example noise source is constant and independent of frequency (2.47), then its integrated RMS noise from f_{\min} to f_{\max} will be

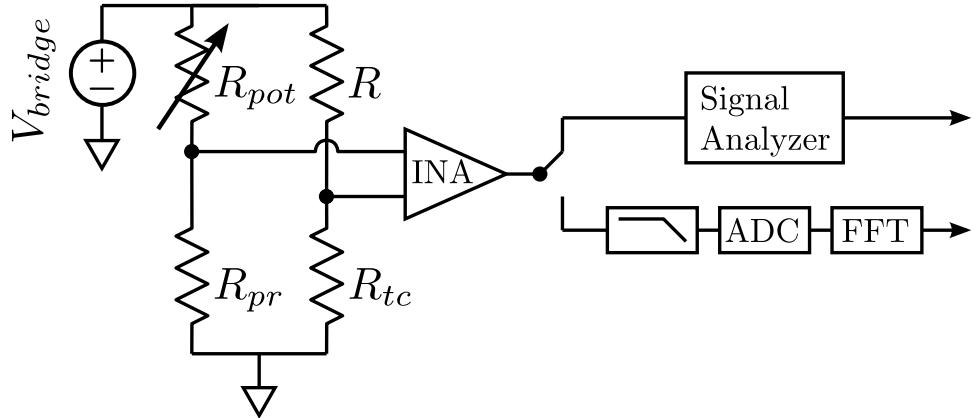


Figure 2.19: Example noise measurement system. The Wheatstone bridge is balanced using a single potentiometer, the output is amplified using a high input impedance amplifier, and the spectral content of the noise signal is measured using a signal analyzer or taking the Fourier transform of the time-domain data.

$$v_N^{\text{RMS}} = \frac{\sigma}{\sqrt{\Delta f}} \sqrt{f_{\max} - f_{\min}}. \quad (2.49)$$

The noise amplitude increases as the square root of the measurement bandwidth. In effect, the noise at two separate frequencies is uncorrelated and the integrated noise is their vector sum.

The noise source we have discussed in this section is an example of white noise, which has a flat power spectral density, typical of Johnson noise or shot noise. Other noise sources, such as 1/f noise, have power spectral densities that vary with frequency. The first step in designing any sensor is determining the measurement bandwidth and choosing appropriate filters, because the results from piezoresistive sensor design will vary depending on the measurement bandwidth. The measurement bandwidth should always be limited to the frequency range over which the signal of interest is distributed.

We will often discuss noise sources in terms of their voltage or current spectral densities rather than their power spectral densities. Voltage spectral density, V_N , is equivalent to the square root of the power spectral density, and has units of $\text{V}/\sqrt{\text{Hz}}$. The noise performance of integrated circuits is typically quoted in terms of voltage and current spectral densities by manufacturers.

Noise spectra can be measured using either a signal/spectrum analyzer or by taking the Fourier transform of time-domain data (Figure 2.19). The measurement circuit in Figure 2.19 includes two silicon resistors in a Wheatstone bridge for temperature compensation (R_{pr} and R_{tc}) and one potentiometer (R_{pot}). The potentiometer is used to null the Wheatstone bridge output before it is amplified and measured.

We will now discuss the particular noise sources that affect piezoresistive sensors.

Johnson noise

Thermal noise, also referred to as Johnson or Johnson-Nyquist noise, is universal to electrical resistors [177]. It was first observed in 1928 by Johnson [178] and theoretically explained later that year by Nyquist [179]. Johnson noise is caused by the thermal agitation of charge carriers within a conductor. Similarly, Brownian motion and thermomechanical noise are caused by the thermal agitation of atoms. Johnson noise is fundamental, exists in all resistors and modern electronic devices, and cannot be eliminated.

The power spectral density of Johnson noise is independent of frequency and can be calculated from

$$S_J = 4k_b T R \quad (2.50)$$

where k_b (J/K) is Boltzmann's constant while T (K) and R (Ω) are the temperature and electrical resistance of the resistor. We will generally model Johnson noise as a voltage noise source in series with an ideal resistor, although it can equivalently be modeled as a current noise source in parallel with the resistor with a current noise power spectral density of $4k_b T / R$.

The RMS voltage noise in a given circuit can be calculated by integrating the Johnson noise power spectral density over the measurement bandwidth to obtain

$$V_J = \sqrt{4k_b T R (f_{\text{max}} - f_{\text{min}})} \quad (2.51)$$

Thus, in order to minimize Johnson noise, the resistance, temperature and measurement bandwidth should be minimized. The temperature and measurement bandwidth are typically determined by the sensor application, so reducing the piezoresistor resistance is

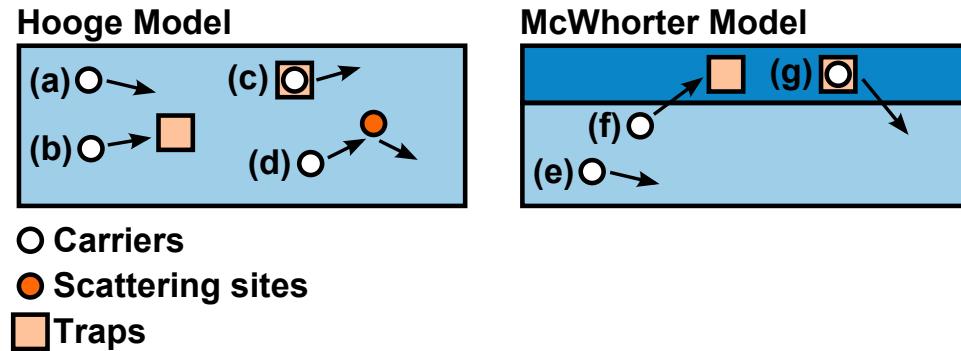


Figure 2.20: Models for semiconductor 1/f noise. The Hooge model assumes that scattering sites and traps exist on the bulk of the material. Carriers may (a) drift unimpeded in the electric field, (b) enter and (c) exit traps, or (d) scatter. Post-ion implantation annealing reduces the density of scattering sites and traps. In the McWhorter model, traps exist at material interfaces, such as between Si and SiO_2 . (e) Carriers generally drift unimpeded within the bulk of the semiconductor, but may (f) enter and (g) exit trapping sites at the material interface.

the main design option for minimizing Johnson noise.

Although (2.50) holds for a single resistor, a Wheatstone bridge contains four resistors. Assuming that the bridge is balanced and all four resistors have resistance R , then the impedance looking out from each amplifier input is $R/2$ and the Johnson noise PSD due to the resistors will be $2k_bTR$ at each amplifier input. But the noise power of uncorrelated sources adds linearly, so the total Johnson noise of the Wheatstone bridge is $4k_bTR$. In short, the Johnson noise of a balanced Wheatstone bridge is equal to the Johnson noise of one of the resistors in the bridge.

1/f noise

The power spectral density of 1/f noise, as its name implies, is inversely proportional to frequency. First discovered due to excess low frequency noise in vacuum tubes, 1/f noise is found in a wide variety of systems, from noise in field effect transistors to the resonant frequency of quartz crystal oscillators to annual variations in rainfall and temperature [180]. 1/f noise in semiconductor electronics is not completely understood and remains an active research topic [181–186]. Despite many decades of research, the underlying sources of 1/f noise are still debated [187].

The most widely used models for 1/f noise today are attributed to McWhorter and Hooge. These views are currently the leading explanations for the origin of 1/f noise in semiconductors. The McWhorter model attributes the 1/f noise to surface factors [188, 189], while the Hooge model implicates bulk defects [190, 191] (Figure 2.20).

One point that all experiments and models agree upon is that 1/f noise in resistors is caused by conductivity fluctuations [186, 191]. Hooge showed that the 1/f low-frequency noise modulated the Johnson noise even with no current flowing through the resistor [192]. This experiment demonstrated that 1/f noise is not current-generated, and that rather, current flow is only needed to transduce conductivity fluctuations into voltage noise.

Attempts to observe the lower limit of 1/f, below which the spectrum theoretically flattens, have not been successful [191]. Measurements down to 3 μ Hz (or approximately 4 days per cycle) show a noise spectrum that is still 1/f [193].

Hooge's empirical 1/f noise model predicts that the voltage power spectral density of a resistor with uniform current density is

$$S_H = V_{\text{bias}}^2 \frac{\alpha}{N_{\text{eff}} f} \quad (2.52)$$

where f , N_{eff} , and V_{bias} are the frequency, the effective number of carriers in the resistor, and the bias voltage across the resistor, respectively. A non-dimensional fitting parameter, α , is ascribed to the crystal lattice quality and typically ranges from 10^{-3} to 10^{-7} depending on the fabrication process details [75, 175, 194]. Note that the total number of carriers in the resistor is not equal to the effective number of carriers when the current density is not constant throughout the resistor. We will discuss calculating the effective number of carriers shortly.

The RMS 1/f noise in a measurement bandwidth from f_{\min} to f_{\max} , rewritten in terms of V_{bridge} rather than V_{bias} , is equal to

$$\begin{aligned}
 V_H &= \sqrt{\int_{f_{\min}}^{f_{\max}} S_H df} \\
 &= V_{\text{bridge}} \sqrt{\frac{\alpha}{4N} \ln \frac{f_{\max}}{f_{\min}}}.
 \end{aligned} \tag{2.53}$$

(2.53) and (2.52) apply to a single resistor. Although a Wheatstone bridge could contain just a single silicon piezoresistor, either two or four silicon resistors are typically used for temperature compensation (Section 2.6). The 1/f noise from discrete resistors is uncorrelated, so the integrated noise increases by a factor of $\sqrt{2}$ for every additional (nominally identical) silicon resistor in the bridge.

We also want to emphasize that the integrated 1/f noise power spectral density is constant for every decade of measurement bandwidth, whereas the Johnson noise power spectral density scales linearly with bandwidth. The result is that 1/f noise tends to be limiting factor for low frequency measurements while Johnson noise limits wide bandwidth measurements. The differences in frequency scaling behavior between the various noise sources has important implications for design optimization. For example, increasing the length of a piezoresistor reduces the 1/f noise (more carriers) but increases the Johnson noise (higher resistance) and an optimized cantilever design will generally balance the integrated noise of each noise source.

Is 1/f noise in silicon piezoresistors due to surface or bulk defects? Harley and Kenny showed that resistors with different surface to volume ratios have the same 1/f noise characteristics, and that 1/f noise scales with the resistor volume, consistent with Hooge's empirical equation [75]. The Hooge model has been applied to accurately model 1/f noise in silicon piezoresistors for diffused [195, 196], ion implanted [197, 198] and epitaxial piezoresistors [36, 199]. In summary, 1/f noise in silicon piezoresistors is predominantly due to bulk defects located near the wafer surface and we will use the Hooge model for all modeling and optimization work.

Hooge defines 1/f noise as only those spectra described by $1/f^n$ where the frequency component n ranges from 0.9 to 1.1. Noise with a different power spectral density and other frequency exponents, sometimes referred to as 1/f-like noise, is often confused with

1/f noise and is not predicted by the Hooge equation. According to Hooge, noise with a higher exponent (e.g. 1.5 or 2), indicates noise mechanisms other than mobility fluctuations that should not be considered 1/f noise.

Abnormal 1/f noise characterization can give insights into piezoresistor reliability and failure analyses. For example, Neri [200] found that the 1/f exponent is closer to 2 in metal traces that exhibit electromigration. Similarly, Vandamme [201] showed that excess 1/f noise in semiconductors can be attributed to small constrictions and current crowding. Constrictions can also lead a nonlinear response and the appearance of third harmonics in the signal output.

Park showed that excess 1/f noise in cantilevers varies with the Joule heating power and thermal conductance from cantilever to the ambient temperature [202]. For example, large power dissipation in cantilevers lead to excess noise, while the same power dissipation in identical, unreleased test structures did not exhibit excess noise. When the temperature of the silicon resistor is significantly above the ambient temperature, air flow currents modulate the resistor temperature and lead to fluctuations in the sensor output. Elevated temperature leads to an increase in sensor noise and a reduction in sensitivity, so is important to consider in the design optimization process and will be discussed in Section 2.6.

Polysilicon resistors have higher 1/f noise than their single crystal counterparts [203]. At grain boundaries, small constrictions are present, thus reducing the effective number of carriers due to current crowding. The reduction in the number of carriers depends on the polysilicon processing history due to its dependence on grain size.

As noted earlier in (2.52), the 1/f noise depends on the effective number of carriers, not the total number. When the current density varies over the resistor volume, the effective number of carriers will always be less than the total number. Examples include polysilicon and metal-silicon contacts, where current crowding is well known, but also any resistor where the dopant concentration varies throughout the thickness (e.g. ion implanted and diffused resistors). The general expression for the Hooge noise power spectral density is [204]

$$S_H = \frac{\alpha}{I^2 f} \int \frac{J^4 \rho^2}{n} dV \quad (2.54)$$

where I is the total current and J , ρ and n are the spatially varying current density, resistivity and carrier concentration. Current density contributes to the noise power by the fourth power, and local constrictions, sharp changes in current direction, or abrupt changes in dopant concentration can substantially increase the 1/f noise. The integral for arbitrary geometries can be calculated using finite element analysis.

Any piezoresistor that varies in dopant concentration throughout its depth has some degree of current crowding. This occurs because the carrier mobility varies with dopant concentration; as dopant concentration increases the mobility decreases and the current carried per dopant atom decreases. The total number of carriers per unit area (N_z^{total}) can be calculated directly from the carrier profile as

$$N_z^{\text{total}} = \int_0^{t_j} n(z) dz. \quad (2.55)$$

The total number of carriers is essentially equal to the effective number of carriers for epitaxial piezoresistors, but the total number of carriers is often erroneously applied to other fabrication methods [22]. The effective number of carriers per unit area (N_z) can be calculated from (2.54) following [205] as

$$N_z = \frac{\left(\int_0^{t_j} n \mu d z \right)^2}{\int_0^{t_j} n \mu^2 d z}. \quad (2.56)$$

For a typical diffusion dopant profile (e.g. POCL₃ at 800°C for 60 minutes) the effective number of carriers per unit area is 80% of the total number. Using the total rather than the effective number of carriers would lead to a 20% overestimate of α during noise analysis. Current crowding should also be considered when laying out devices with doped silicon traces, which will contribute to the overall 1/f noise. Current crowding is not an issue for epitaxial profiles. For the ion implanted profiles we will investigate (Section 2.5.3) the effective number of carriers ranges from 54 to 99% of the total number and is 90% on average.

A defining characteristic of Hooge-type 1/f noise is that the voltage spectral density increases linearly with the applied bias and the power spectral density scales with the bias squared. If the noise amplitude scales in another manner, then it is likely that other

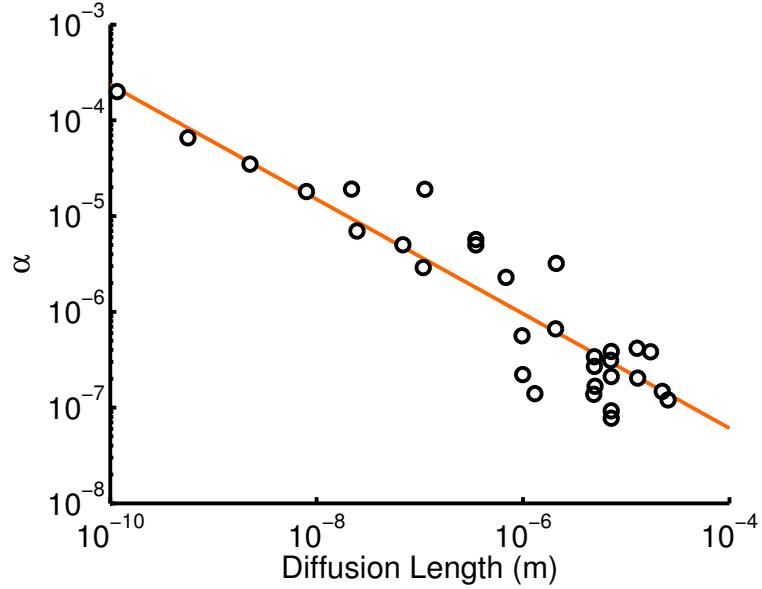


Figure 2.21: The Hooge noise parameters, α , decreases for ion implanted piezoresistors as the diffusion length of dopant atoms during the post-implantation anneal, \sqrt{Dt} , increases. Data was compiled from Refs. [35, 73, 75, 175, 206].

mechanisms, such as current crowding or temperature fluctuations, are present.

1/f noise can be reduced by either increasing N or decreasing α . Vandamme showed that α depends on crystal lattice perfection and that α can be reduced for ion implanted resistors by increasing the time and temperature of the post-implantation anneal [206]. He later showed that the α for a MOSFET channel varies with the gate voltage due to variation in α through the channel depth [207]. The 1/f noise was lower when the channel was operated in depletion mode than in accumulation mode because the current flow was further away from the surface.

Thus, the post-ion implantation anneal performs three functions: dopant activation, lattice repair and shifting conduction away from the surface. Several groups have investigated the dependence of α on the diffusion length of the dopant atoms during the anneal, \sqrt{Dt} , and shown that longer diffusion length anneals decrease the 1/f noise of silicon piezoresistive cantilevers [75, 175]. The Hooge noise parameter (α) is plotted in Figure 2.21 as a function of the dopant diffusion length during the post-ion implantation anneal based upon data from Refs. [35, 73, 75, 175, 206].

It is important to note that a surface passivation oxide layer is not necessary to achieve low 1/f noise. Harley measured the noise spectra of resistors with and without passivation oxide layers and observed no difference in 1/f noise [36]. The data suggests that the post-implantation anneal decreases 1/f noise by eliminating defects in the bulk of the resistor and driving dopants further into the bulk, not by passivating the resistor surface.

The best-fit line plotted in Figure 2.21 is

$$\alpha = 2.469 \times 10^{-10} (\sqrt{Dt})^{-0.598} \quad (2.57)$$

which is slightly different from the fit presented in 2000 by Harley [22] because it includes subsequent data from Mallon and Yu [175, 194].

We note several points of caution with regards to the empirical fit.

First, there is substantial uncertainty in calculating α due to the difficulty in calculating the effective number of carriers in the piezoresistor. Errors can be introduced either through an inaccurate estimate of the dopant concentration profile, error in calculating the resistor volume, or not accounting for current crowding in the resistor. For example, using simulated dopant profiles rather than a direct measurement of the electrically active dopant concentration profile (via spreading resistance analysis) could account for some error. However, the calculated α values are from multiple research groups spanning several decades, so random errors should presumably average out. There are several possible sources of systematic error in the data as well. In particular, using N_z^{total} rather than N_z would lead to overestimates of 10-50% in calculating α depending on the degree of current crowding.

Second, there is probably a lower bound on α . Harley initially predicted a lower bound of 10^{-6} based upon the data available at the time [22]. Jevtic theoretically investigated variation in α with dopant concentration and predicted minimum values of 1×10^{-5} and 3×10^{-5} for n- and p-type single crystal silicon, respectively [208]. However, silicon piezoresistors with significantly lower α values have been demonstrated for both dopant types. Applying (2.57) to diffusion lengths greater than 10^{-5} m without preliminary tests is not advised.

Third, the data fit only applies to ion implanted piezoresistors and all of the data points

are for boron doping. This leads to two major assumptions. First, we will assume that the relationship between α and the dopant diffusion length holds for phosphorus and arsenic as well. Due to the lower diffusivities of phosphorus and especially arsenic, this assumption leads to very different optimization results than if we had assumed that α depends on the diffusion length of the silicon atoms. Second, we will assume that $\alpha = 10^{-5}$ for diffused and epitaxial piezoresistors without any subsequent annealing.

Harley calculated a value of 2×10^{-5} for his epitaxial piezoresistors [22]. We measured α values ranging from 2×10^{-7} to 5×10^{-5} for diffused phosphorus piezoresistors with a mean value of 1.05×10^{-5} [196]. Neither technique has the inherent lattice damage of ion implantation, however lattice defects are probably still generated. Predeposition can introduce electrically inactive dopants, and epitaxy can introduce grain boundaries and other lattice defects depending on the growth conditions. In particular, we have experimentally observed interstitial phosphorus near the surface of POCl_3 doped piezoresistors which may disrupt the lattice.

Vandamme's results in Ref. [207] suggest that lattice quality is higher in the bulk of the silicon than near its surface, and that shifting electrical conduction away from the surface may be an important role of the post-implantation anneal. If the sole purpose of the anneal was to reduce implantation-induced crystal lattice damage, then piezoresistors formed from different dopant atoms (e.g. boron vs. arsenic) would exhibit the same value of α for an identical annealing process. If this were true then arsenic piezoresistors would yield much higher performance than boron or phosphorus doped piezoresistors due to its dramatically lower diffusivity.

However, personal correspondence with Vandamme suggest that the dopant diffusion length, not the time and temperature of the anneal, is critical and that the 1/f noise parameter decreases as conduction shifts further into the bulk. Without clear experimental data on the subject, we will assume that the dopant diffusion length defines the 1/f noise parameter and that arsenic-doped piezoresistors would need to be annealed much longer in order to achieve the same 1/f noise parameter as boron- and phosphorus-doped piezoresistors.

In summary, 1/f noise is fundamentally different from Johnson noise. Johnson noise is present regardless of the current density flowing through the resistor and depends only on its electrical resistance. In contrast, 1/f noise is a conductivity noise, and the voltage

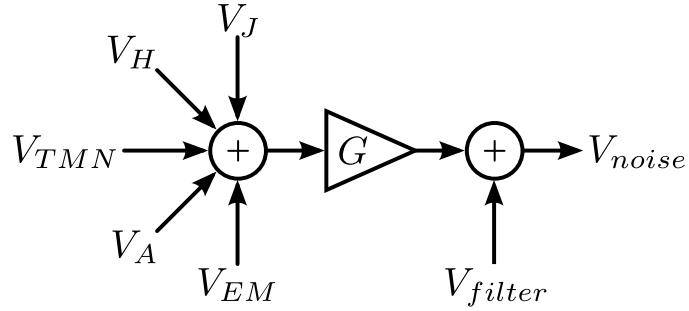


Figure 2.22: Example noise cascade illustrating the dominant role of noise at the input of the first-stage amplifier.

noise measured across the resistor scales linearly with the current flowing through the resistor. Additionally, 1/f noise depends on the dimensions and processing history of the resistor. Any residual damage from ion implantation or ion bombardment during reactive ion etching will increase the 1/f noise, but will only affect the Johnson noise to the extent that any lattice damage increases the overall electrical resistance.

Amplifier noise

The Wheatstone bridge output is amplified before being filtered and acquired (Figure 2.19). Throughout our discussion of piezoresistor design we will assume that the signal is amplified by an instrumentation amplifier, but the noise analysis can be extended to any signal conditioning configuration.

Instrumentation amplifiers typically consist of three operational amplifiers (op-amps). The first two op-amps present a high input impedance and amplify the signal at each amplifier input. This eliminates loading effects on the circuit and improves the rejection of common mode signals. The final op-amp differentially amplifies the buffered signals. A single resistor, integrated on-chip or provided by the designer, determines the amplifier gain. The high input impedance, high common mode rejection ratio, and precision with which the gain can be set make instrumentation amplifiers ideal for piezoresistor signal conditioning.

The instrumentation amplifier, which provides the first-stage of gain, is usually the dominant signal conditioning noise source. The noise contribution from second-stage

elements such as filters and variable gain amplifiers should be small compared to the first-stage if the instrumentation amplifier gain, G , is large.

This effect is illustrated in Figure 2.22. There are five noise sources that contribute to the amplifier input signal: Johnson noise (V_J), Hooge noise (V_H), thermomechanical noise (V_{TMN}), amplifier noise (V_A) and electromagnetic coupling noise (V_{EM}). The noise sources sum at the amplifier input and are amplified by G , before the amplifier output is filtered and acquired. The filter contributes additional noise V_{filter} , and could be from either active or passive filtering elements.

The overall noise for the configuration in Figure 2.22 is $V_{\text{noise}} = G(V_J + V_H + V_{TMN} + V_A + V_{EM}) + V_{\text{filter}}$. If the first term is much larger than the second term (e.g. $G \gg 1$) then signal conditioning operations after the initial signal amplification don't substantially increase the noise. If the filter was erroneously placed before the first-stage amplifier or a particularly noisy filter were used, this assumption would no longer hold.

Amplifiers contribute both voltage and current noise to the system. Both forms of noise have Johnson and 1/f spectral components. The voltage noise is independent of the amplifier source impedance; it remains the same whether the amplifier inputs are left open or shorted together. In contrast, the current noise is transduced into a voltage noise by the Wheatstone bridge. As in the case of Johnson noise, the impedance seen looking out from each amplifier input is $R/2$ for a balanced Wheatstone bridge. The voltage spectral density contributed by each amplifier input is then $A_{IJ}R/2$ where A_{IJ} is the magnitude of the Johnson component of the current noise in units of $\text{A}/\sqrt{\text{Hz}}$. The current noise at each of the amplifier inputs is uncorrelated, so the voltage spectral density is increased by a factor of $\sqrt{2}$ and the power spectral density is increased by a factor of 2. The same analysis holds true for the amplifier's 1/f-type current noise.

The overall voltage power spectral density contributed by the instrumentation amplifier can be calculated from

$$S_A = A_{VJ}^2 + 2A_{IJ}^2 \left(\frac{R}{2} \right)^2 + \frac{1}{f} \left(A_{VF}^2 + 2A_{IF}^2 \left[\frac{R}{2} \right]^2 \right). \quad (2.58)$$

where A_{VJ} , A_{IJ} , A_{VF} and A_{IF} quantify the Johnson and 1/f components of the voltage and current noise from the amplifier. We used the subscript H for piezoresistor 1/f (Hooge)

Table 2.4: Noise coefficients for example instrumentation amplifier options. The coefficients can be calculated from the noise spectra in most amplifier data sheets. The -3 dB bandwidth and voltage noise coefficients are for $G = 1000$. The 1/f noise coefficients, A_{VF} and A_{IF} , are equal to the voltage and current noise spectral density at 1 Hz, respectively. Amplifier characteristics vary with the input stage transistor type, bandwidth and cost.

Amplifier	Input Stage	Bandwidth (kHz)	A_{VJ} (nV/ $\sqrt{\text{Hz}}$)	A_{VF} (nV)	A_{IJ} (fA/ $\sqrt{\text{Hz}}$)	A_{IF} (fA)
INA103	BJT	80	1.2	6	2000	2500
INA111	FET	50	10	68	0.8	8
INA116	FET	7	28	300	0.1	1
AD622	BJT	12	9	23	100	600
AD623	BJT	2	30	130	100	590
AD8220	FET	14	14	50	1	10
AD8221	BJT	15	8	11	40	540

noise, but we use F to describe the amplifier 1/f noise because it is not necessarily due to bulk mobility fluctuations. The 1/f noise coefficients, A_{VF} and A_{IF} , are equal to the voltage and current noise spectral density at 1 Hz, respectively. The integrated amplifier voltage noise is

$$V_A = \left[\left(A_{VJ}^2 + 2A_{IJ}^2 \left[\frac{R}{2} \right]^2 \right) (f_{\max} - f_{\min}) + \left(A_{VF}^2 + 2A_{IF}^2 \left[\frac{R}{2} \right]^2 \right) \ln \left(\frac{f_{\max}}{f_{\min}} \right) \right]^{1/2}. \quad (2.59)$$

The amplifier noise coefficients can be calculated from the amplifier data sheet provided by the manufacturer. Example coefficients for several instrumentation amplifiers manufactured by Texas Instruments and Analog Devices are tabulated in Table 2.4. Note that BJT-based amplifiers have relatively low voltage noise and high current noise coefficients, while the converse is generally true for FET-based amplifiers.

An accurate description of amplifier noise is essential for sensor design and optimized designs will vary substantially depending on the amplifier noise characteristics. For example, the high current noise of the INA103 (Table 2.4) would lead to a lower optimal piezoresistor resistance than if the sensor were designed for the low current noise INA111

(2.58). The optimal amplifier for a particular application will depend upon bandwidth, cost and performance requirements as well as the sensor design constraints.

Thermomechanical noise

All atoms possess thermal energy at temperatures greater than absolute zero. Brownian motion of the atoms in the sensor is an additional noise source (thermomechanical noise) and places a lower limit on the resolution of all mechanical sensors [32]. Thermomechanical noise is a direct analogy of Johnson noise. The motion of a simple harmonic oscillator (SHO) is described by

$$m_{\text{eff}} \frac{\partial^2 x}{\partial t^2} + b \frac{\partial x}{\partial t} + k_c x = F. \quad (2.60)$$

Similarly, the charge in an RLC oscillator circuit can be written as

$$L \frac{\partial^2 Q}{\partial t^2} + R \frac{\partial Q}{\partial t} + \frac{1}{C} Q = V. \quad (2.61)$$

where Q is the charge. In the case of the electrical circuit, the equivalent voltage noise spectral density due to thermal equilibrium is $\sqrt{4k_b T R}$ [178]. By direct analogy, the equivalent force noise spectral density in a mechanical system is [32]

$$F_{\text{TMN}} = \sqrt{4k_b T b} \quad (2.62)$$

which has units of $\text{N}/\sqrt{\text{Hz}}$. The greater the damping force on a mechanical structure, the greater the thermomechanical force noise.

It is common to rewrite (2.62) so that it depends on more easily measurable experimental parameters. The damping coefficient of the SHO can be written in terms of the spring constant, resonant frequency and quality factor of the oscillator as $b = k_c / \omega_0 Q$, allowing the thermomechanical force noise spectral density to be written as

$$F_{\text{TMN}} = \sqrt{\frac{4k_c k_b T}{\omega_0 Q}} \quad (2.63)$$

where k_c , ω_0 and Q are the spring constant, resonant frequency and quality factor of

the oscillator, respectively. If the spring constant varies with respect to position on the mechanical structure, as in the case of a cantilever beam, then the equivalent force noise also varies with position.

Note that thermomechanical force noise is independent of frequency up to extremely high frequencies as in the case of Johnson noise. Also, the interactions between the SHO and any surrounding fluid are completely accounted for by (2.63) except in the case of significant nonlinear damping.

Thermomechanical noise is transduced into a voltage noise by the force sensitivity of the sensor according to

$$S_{\text{TMN}} = S_{\text{FV}} F_{\text{TMN}} \quad (2.64)$$

where S_{FV} is the force sensitivity of the sensor (with units of V/N). As the force sensitivity increases, the magnitude of the thermomechanical noise at the circuit output increases. We will discuss sensitivity in the next section in detail, but briefly, force sensitivity increases as the sensor size is reduced and the bridge bias voltage is increased, making thermomechanical noise particularly important for nanomechanical devices.

As with the other noise sources, we can calculate the RMS thermomechanical noise voltage from

$$V_{\text{TMN}} = S_{\text{FV}} \sqrt{\frac{4k_{\text{c}}k_{\text{b}}T}{\omega_0 Q}} \sqrt{f_{\max} - f_{\min}} \quad (2.65)$$

Although the thermomechanical force noise is flat with respect to frequency, the induced displacement noise is not. The thermomechanical displacement noise is equal to

$$x_{\text{TMN}} = F_{\text{TMN}} G \quad (2.66)$$

where G is the SHO transfer function. The transfer function can be calculated by taking the Laplace transform of (2.60), resulting in

$$\begin{aligned} G(s) &= \frac{X(s)}{F(x)} \\ &= \frac{1}{m_{\text{eff}}s^2 + bs + k_c}. \end{aligned}$$

Substituting $s = i\omega = i2\pi f$ we obtain

$$G(f) = \frac{1/k}{\left(1 - \frac{f^2}{f_0^2}\right) + i\frac{f}{f_0 Q}}. \quad (2.67)$$

For example, on resonance the displacement noise of the SHO will be equal to $F_{\text{TMN}}Q/k$ and 90° out of phase with the driving force due to mechanical amplification.

Another way of looking at thermomechanical noise can be obtained from the Equipartition Theorem [80], which states that if a set of eigenmodes is in thermal equilibrium with a thermal reservoir, the average energy of each mode will be $\frac{1}{2}k_b T$ where k_b is Boltzmann's constant and T is the absolute temperature. Each eigenmode must be capable of storing energy, e.g. kinetic energy ($\frac{1}{2}m_{\text{eff}}v^2$), elastic energy ($\frac{1}{2}k_c x^2$) or electrostatic energy ($\frac{1}{2}CV^2$). Accordingly, the mean-square displacement of our SHO can be calculated from $\frac{1}{2}k_c \langle x^2 \rangle = \frac{1}{2}k_b T$ as

$$\langle x^2 \rangle = \frac{k_b T}{k_c} \quad (2.68)$$

which can be equivalently calculated by integrating the thermomechanical displacement noise spectral density over each eigenmode.

Noise summary

We have discussed the most important noise sources for piezoresistive sensors: Johnson, Hooge, amplifier and thermomechanical noise. Before moving on, we will illustrate typical noise spectra and discuss practical noise measurement issues.

Figure 2.23 presents the noise spectrum of a simulated piezoresistive cantilever sensor. The simulated cantilever is 1 μm thick, 10 μm wide, and 100 μm long. Two piezoresistors are included in the Wheatstone bridge, and the bridge bias voltage is 2 V. The piezoresistor

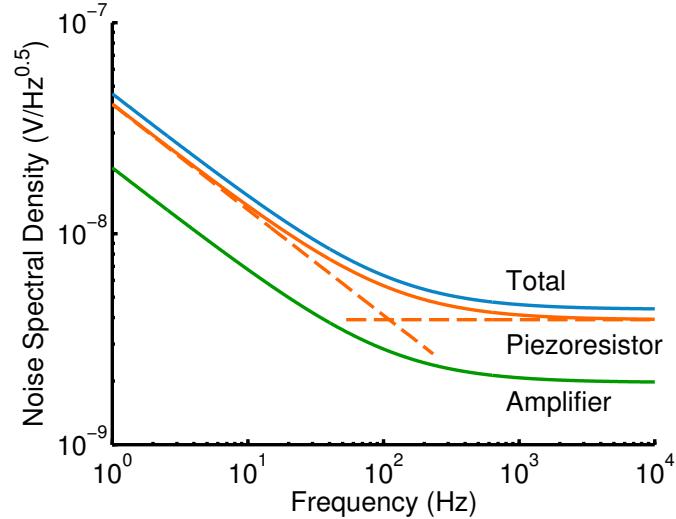


Figure 2.23: Typical noise spectrum. Piezoresistor noise is substantially larger than amplifier noise. Johnson noise and Hooge noise for the piezoresistor are shown as dashed lines. Noise is dominated by Hooge noise at frequencies below the 1/f corner frequency. Thermomechanical noise is 1-2 orders of magnitude smaller than the other noise sources depending on the damping experienced by the sensor. The relative magnitude of the noise sources will vary depending on the precise design of the sensor.

extends 30% of the cantilever length, and is an epitaxially grown phosphorus resistor with a 333 nm junction depth and $n = 10^{20} \text{ cm}^{-3}$. An INA103 is used to amplify the Wheatstone bridge output due to the low impedance of the piezoresistor (580Ω). The noise is input referred; in other words, the noise at the output of the amplifier would be measured and divided by the gain of the amplifier.

The piezoresistor noise consists of Johnson and Hooge noise. At frequencies below the 1/f corner frequency (≈ 120 Hz for this particular example), piezoresistor noise is dominated by Hooge noise. Dashed lines in Figure 2.23 illustrate the relative contribution of each noise source.

The relative magnitudes of the piezoresistor, amplifier and thermomechanical noise is typical for most applications. The thermomechanical noise would be $20 \text{ pV/Hz}^{0.5}$ in air ($Q \approx 160$) and $215 \text{ pV/Hz}^{0.5}$ in water ($Q \approx 3.9$). Thermomechanical noise can become comparable to amplifier noise when the motion of the sensor is highly damped and it has a high force sensitivity.

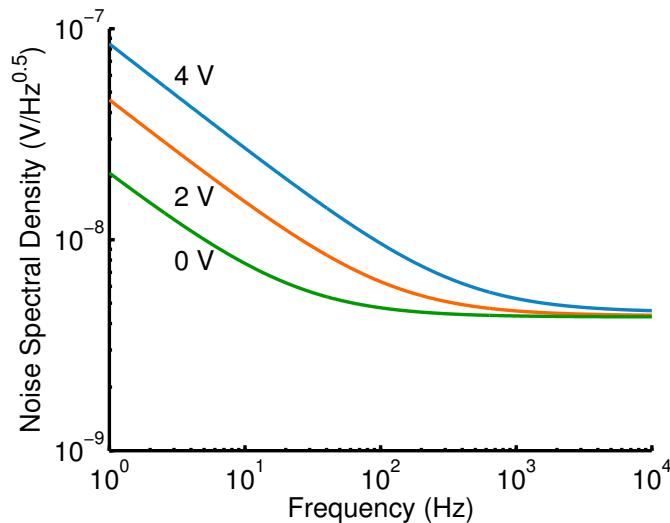


Figure 2.24: Noise spectra for varying bridge bias voltages. The Hooge noise spectral density increases linearly with the bridge bias, while the Johnson and amplifier noise are independent of the bridge bias (excluding piezoresistor self-heating). When the bridge bias goes to zero, the 1/f noise is limited by the amplifier rather than the piezoresistor. When the amplifier 1/f noise floor is larger than the piezoresistor noise, a modulation-demodulation technique can be utilized (Section 2.3.2).

The noise spectral density varies with the bias applied to the Wheatstone bridge, as illustrated in Figure 2.24. The simulation conditions are identical to those used in Figure 2.23. The Hooge noise spectral density increases linearly with the bridge bias, while the Johnson and amplifier noise are approximately independent of it if Joule heating is neglected. As the bridge bias approaches zero, the Hooge noise goes to zero but the 1/f noise of the amplifier remains constant. The only way to eliminate the amplifier 1/f noise is to use a heterodyne measurement technique (Section 2.3.2) in order to shift the piezoresistor noise signal to a frequency above the amplifier's 1/f corner frequency before amplification.

There are a number of experimental issues that can confound noise measurements and, in some cases, add additional noise to the system:

Potentiometer noise: Certain types of potentiometers can add excess noise to the measurement system. Wirewound potentiometers have negligible 1/f noise, but ceramic-based potentiometers (e.g. Cermet) can contribute excess 1/f noise. Care should be taken in potentiometer selection.

Noise floor verification: The noise floor of the system should be verified by grounding the Wheatstone bridge bias, shorting the amplifier inputs, and/or inserting a test resistor in the place of the piezoresistor in the bridge. The noise should also be measured over a range of bias voltages to ensure that the noise isn't limited by a system noise floor and that other devices (e.g. potentiometers) are not limiting the system noise.

Noise cascading: This is a straightforward but important point. When multiple signal conditioning stages are used, the highest gain stages should be closest to the initial signal. The noise that a signal conditioning stage effectively adds to the system is reduced by the cumulative system gain to that stage.

Gain accuracy: The amplifier gain is set by 1 or 2 feedback resistors depending on the circuit topology. The gain accuracy will vary from 0.1-1% depending on the accuracy of the resistors and will vary with temperature. The gain of the system should be directly measured by applying a test signal to the amplifier inputs and measuring its output.

Line noise: Noise in the bridge bias signal will couple into the circuit output due to imperfect matching between resistors in the bridge and a finite amplifier common mode rejection ratio (CMRR). In particular, noise peaks at 50 or 60 Hz and higher harmonics are often present. Noise from the power supply can be reduced by using a voltage reference (e.g. the ADR44x from Analog Devices), which improves the system line noise rejection. A voltage reference can be combined with a potentiometer controlled voltage divider and voltage buffer to provide an adjustable, low noise bridge bias voltage.

2.4.2 Sensitivity

Piezoresistive sensors transduce a mechanical signal into an electrical one according to

$$\frac{\Delta R}{R} = \pi_l \sigma_l + \pi_t \sigma_t \quad (2.69)$$

as discussed in Section 2.2.1. A Wheatstone bridge transduces a resistance change into a voltage change according to

$$\Delta V = N_{\text{pr}} \frac{V_{\text{bridge}}}{4} \frac{\Delta R}{R} \quad (2.70)$$

where N_{pr} is the number of piezoresistors in the Wheatstone bridge that are mechanically loaded and V_{bridge} is the potential drop across the bridge. This expression assumes that the bridge is balanced (i.e. all four quadrants of the bridge have the same nominal resistance) and that the relative resistance change ($\Delta R/R$) of each of the piezoresistors is equal in magnitude. If the bridge is not balanced, then the voltage change will be attenuated due to the nonlinear response of the bridge.

Sensitivity is typically reported as either a fractional change in resistance or voltage change with respect to the mechanical signal. For example, the sensitivity of a pressure sensor may be reported with units of ppm/kPa or mV/kPa for the former and latter cases, respectively. Sensitivity should always be reported for the signal at the input to the amplifier inputs (input referred) because signal conditioning gain can be arbitrarily adjusted.

Sensitivity is calculated by integrating the relative resistance change over the piezoresistor volume. However, the stress and piezoresistive coefficients will almost always vary spatially. In this section we will discuss two important general sensitivity concepts, the sensitivity (β^*) and resistance factors (γ). The resistance factor will allow us to account for excess resistances in the device. The sensitivity factor will allow us to calculate sensitivity by multiplying the maximum stress by a correction factor, simplifying the calculation process substantially.

We have discussed variations in the piezoresistance factor with dopant concentration and temperature (Section 2.2.2) and variation in dopant concentration with depth through the device (Section 2.5). In this section we will discuss how to calculate sensitivity while accounting for spatially varying dopant concentration and stress profiles. The stress profile is always either constant (e.g. strain gauges) or varies linearly (e.g. cantilevers and everything else), and we will discuss the two cases in depth.

Excess resistance and γ

Piezoresistive sensors function by measuring a relative resistance change, and sensor sensitivity is always of the form

$$S \propto \frac{\Delta R}{R}. \quad (2.71)$$

Until now, we have assumed that the resistance is composed entirely of the piezoresistor resistance, i.e. $R = R_{\text{pr}}$. However, additional resistances will always be present in a real sensor and measurement system (e.g. contact resistance, conducting traces, wirebonds), and the total resistance should actually be written as $R = R_{\text{pr}} + R_{\text{excess}}$. Clearly, the sensitivity in (2.71) will be decreased by any excess resistance that does not vary in response to a mechanical load.

We can define a resistance factor as

$$\gamma = \frac{R_{\text{pr}}}{R} \quad (2.72)$$

where γ is the ratio of the piezoresistor resistance to the total resistance, so that the relative change in resistance can be written as

$$\frac{\Delta R}{R} = \frac{\Delta R_{\text{pr}}}{R_{\text{pr}}} \gamma. \quad (2.73)$$

Combining (2.71) and (2.73), we can write the sensitivity in terms of the piezoresistor resistance and resistance factor as

$$S \propto \frac{\Delta R_{\text{pr}}}{R_{\text{pr}}} \gamma. \quad (2.74)$$

We will include the resistance factor, γ , in all sensitivity calculations to account for this resistance effect. During design optimizations we will typically define γ by assuming a fixed value of R_{excess} such as 200Ω .

The total resistance, $R = R_{\text{pr}}/\gamma$ is used to calculate Johnson and amplifier noise. Excess resistance decreases the sensitivity and increases the noise of the sensor, so should clearly be avoided. Conductive traces should be made short and from low resistivity materials (e.g. aluminum rather than doped silicon) and oxide vias should be as large as possible given the area constraints of the device.

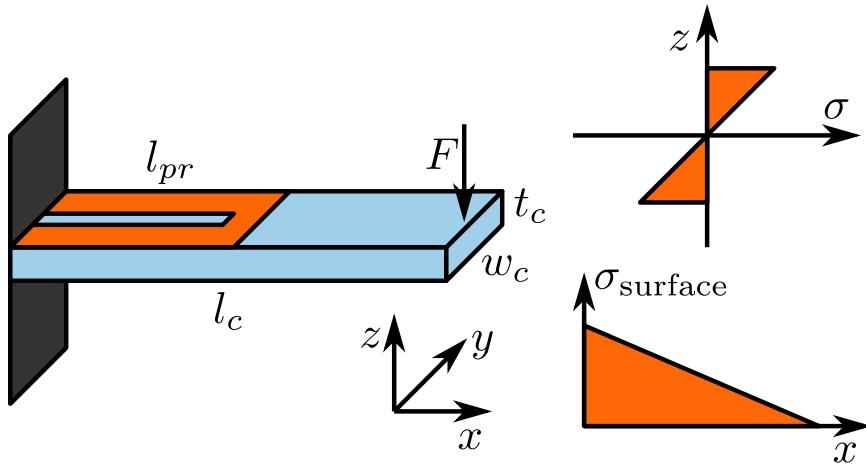


Figure 2.25: An end loaded cantilever beam. The beam has overall length, width and thickness of l_c , w_c , and t_c , respectively. The piezoresistor extends length l_{pr} from the clamped base of the beam and is divided by an infinitesimally wide gap to form a resistive loop. The material in the gap should generally be removed rather than diode isolated in order to limit shunted leakage currents. The orientation of the reference axis is shown.

Sensitivity and β^*

We will use a simple, end loaded cantilever beam to demonstrate how to calculate sensor sensitivity before generalizing the approach. The example beam is shown in Figure 2.25. The overall dimensions of the beam are noted (l_c , w_c , t_c and l_{pr}), and the beam is subjected to a point load at its tip of magnitude F .

Assuming a linear elastic beam that obeys Bernoulli beam theory, the longitudinal stress in the beam varies linearly along the x - and z - axes according to

$$\sigma_L = \frac{12z(l_c - x)}{w_c t_c^3} F. \quad (2.75)$$

We will assume that the transverse stress in the beam is negligible because $l_c \gg w_c$. The average longitudinal stress, $\bar{\sigma}_L$, will be used frequently in this section and is calculated as

$$\bar{\sigma}_L = \frac{1}{l_{pr}} \int_0^{l_{pr}} \sigma_L dx. \quad (2.76)$$

For this particular case, the average longitudinal stress is

$$\overline{\sigma_L} = \frac{12z(l_c - l_{pr}/2)}{w_c t_c^3} F. \quad (2.77)$$

The surface of the cantilever experiences the maximum stress level, $\overline{\sigma_L^{\max}}$, for $z = t_c/2$. The average maximum stress level is

$$\overline{\sigma_L^{\max}} = \frac{6(l_c - l_{pr}/2)}{w_c t_c^2} F. \quad (2.78)$$

We will first calculate the sensitivity assuming that the piezoresistor is of infinitesimal thickness and situated at the top surface of the beam, before performing the more realistic and complex calculation, following the approach taken by Harley [22].

Consider a thin, short slice of the piezoresistor with length dx and thickness dz . The resistance of this slice is

$$dR_{\text{slice}} = 2 \frac{\rho}{w_c} \frac{dx}{dz} \quad (2.79)$$

where ρ is the resistivity of the material and the factor of two comes from the fact that each leg of the piezoresistor is one half of the cantilever width. We can integrate the differential slice resistance along the length of the piezoresistor to calculate the total resistance of the slice with thickness dz ,

$$\begin{aligned} R_{\text{slice}} &= 4 \int_0^{l_{pr}} \frac{\rho}{w_c} \frac{dx}{dz} \\ &= \frac{4\rho l_{pr}}{w_c dz} \end{aligned} \quad (2.80)$$

where the additional factor of two comes from the addition of the two piezoresistor legs in series. The resistance change in the region defined by dR_{slice} is

$$\Delta(dR_{\text{slice}}) = 2 \frac{\Delta\rho}{w_c} \frac{dx}{dz}$$

which can be integrated along the length of the piezoresistor using (2.75) and the fact that $\Delta\rho = \rho\pi_L\sigma_L$ to obtain

$$\begin{aligned}\Delta R_{\text{slice}} &= 4 \int_0^{l_{\text{pr}}} \frac{\rho \pi_L \sigma_L}{w_c} \frac{1}{dz} dx \\ &= \frac{4\rho \pi_L \bar{\sigma}_L l_{\text{pr}}}{w_c dz}\end{aligned}\quad (2.81)$$

where π_L is the effective piezoresistive coefficient of the slice and the additional factor of two again comes from integrating along both piezoresistor legs. We can calculate the relative resistance change in the slice by combining (2.77), (2.80) and (2.81) to obtain

$$\begin{aligned}\frac{\Delta R_{\text{slice}}}{R_{\text{slice}}} &= \pi_L \bar{\sigma}_L \\ &= \frac{12\pi_L z(l_c - l_{\text{pr}}/2)}{w_c t_c^3} F\end{aligned}$$

If we assume that piezoresistor consists of a single slice located at the surface ($z = t_c/2$), then the force sensitivity of the cantilever beam is

$$\frac{\Delta R_{\text{pr}}}{R_{\text{pr}}} = \frac{6\pi_L (l_c - l_{\text{pr}}/2)}{w_c t_c^2} F. \quad (2.82)$$

However, (2.82) does not account for the finite thickness of the piezoresistor or piezoresistive coefficient variations through its thickness.

We can generalize the sensitivity derivation by integrating (2.80) and (2.81) over the beam thickness rather than assuming that a single slice exists at the surface. The slices through the piezoresistor are connected in parallel, so their conductivities add linearly. The conductance of each slice, G_{slice} , is equal to $1/R_{\text{slice}}$, and relative changes in resistance and conductance are related by

$$\frac{\Delta R_{\text{slice}}}{R_{\text{slice}}} = -\frac{\Delta G_{\text{slice}}}{G_{\text{slice}}}. \quad (2.83)$$

The denominator of (2.83) is calculated from

$$\begin{aligned}
G_{\text{slice}} &= \int_{-t_c/2}^{t_c/2} \frac{1}{R_{\text{slice}}} dz \\
&= \int_{-t_c/2}^{t_c/2} \frac{w_c}{4\rho l_{\text{pr}}} dz \\
&= \frac{w_c}{4l_{\text{pr}}} \int_{-t_c/2}^{t_c/2} q\mu n dz.
\end{aligned} \tag{2.84}$$

where we have substituted $\rho = 1/q\mu n$ where q is the elementary charge, μ is the carrier mobility and n is the carrier density. As in the simplified example, the factor of four follows from the width ($w_c/2$) and overall length ($2l_{\text{pr}}$) of the piezoresistor.

To calculate the numerator of (2.83), we can rearrange (2.83) to find

$$\Delta G_{\text{slice}} = -\frac{\Delta R_{\text{slice}}}{R_{\text{slice}}^2}.$$

We previously calculated both ΔR_{slice} and R_{slice} in (2.80) and (2.81). Combining them, we obtain

$$\Delta G_{\text{slice}} = -\frac{w_c}{4l_{\text{pr}}} \int_{-t_c/2}^{t_c/2} \pi_L \bar{\sigma}_L q\mu n dz. \tag{2.85}$$

Finally, we can calculate the relative resistance change of the piezoresistor using (2.84) and (2.85) as

$$\frac{\Delta R_{\text{pr}}}{R_{\text{pr}}} = \pi_L^{\max} \frac{\int_{-t_c/2}^{t_c/2} P \bar{\sigma}_L \mu n dz}{\int_{-t_c/2}^{t_c/2} \mu n dz} \tag{2.86}$$

where we have used the fact that $\pi_L = \pi_L^{\max} P$. The relative resistance change of any piezoresistive sensor can be calculated using (2.86) assuming that the piezoresistor electrical properties only vary in the z -direction. It is important to emphasize that the overall resistance change is calculated as

$$\frac{\Delta R}{R} = \gamma \frac{\Delta R_{\text{pr}}}{R_{\text{pr}}}$$

due to excess resistances in the measurement circuit. We can apply (2.86) to the specific case of an end loaded piezoresistive cantilever by substituting (2.77) for $\bar{\sigma}_L$, which yields

$$\frac{\Delta R_{pr}}{R_{pr}} = \frac{12F(l_c - l_{pr}/2)\pi_L^{\max}}{w_c t_c^3} \frac{\int_{-t_c/2}^{t_c/2} P \mu n z dz}{\int_{-t_c/2}^{t_c/2} \mu n d z}. \quad (2.87)$$

We can simplify (2.87) significantly using (2.78) to obtain

$$\frac{\Delta R_{pr}}{R_{pr}} = \bar{\sigma}_L^{\max} \pi_L^{\max} \left(\frac{2}{t_c} \frac{\int_{-t_c/2}^{t_c/2} P \mu n z dz}{\int_{-t_c/2}^{t_c/2} \mu n d z} \right) \quad (2.88)$$

where $\bar{\sigma}_L^{\max}$ is the average stress at the surface of the piezoresistor from (2.78). We can make one last simplification by defining a sensitivity factor, β^* (beta star), following Park et. al [198]. The sensitivity factor is calculated from

$$\beta^* = \frac{2}{t_c} \frac{\int_{-t_c/2}^{t_c/2} P \mu n z dz}{\int_{-t_c/2}^{t_c/2} \mu n d z}. \quad (2.89)$$

Combining (2.88) and (2.89), we obtain our final result for the relative resistance change of a piezoresistive cantilever,

$$\frac{\Delta R_{pr}}{R_{pr}} = \bar{\sigma}_L^{\max} \pi_L^{\max} \beta^*. \quad (2.90)$$

The sensitivity factor varies from 0 to 1 and defines the ratio of the actual relative resistance change of the piezoresistor to the maximum that is theoretically possible if the piezoresistor were infinitely thin, located at the surface and lightly doped.

Why use β^* rather than β as the symbol to define this ratio? Tortonese introduced β in 1993 to account for the finite piezoresistor thickness [209]. However, he included the piezoresistance factor in both the numerator and denominator, so that β represented the conductivity and piezoresistivity weighted stress. In contrast, β^* only includes the piezoresistance factor in the numerator so defines the conductivity weighted product of the stress and piezoresistance factor. The difference is slight, but β^* correctly computes the sensitive of arbitrarily doped piezoresistors.

In summary, (2.90) can be used to accurately calculate the sensitivity of any piezoresistive sensor where the stress varies linearly across its thickness. The integral in (2.89) extends from the bottom to the top of the device. For devices that are junction isolated, the integral should only extend to the junction, however the difference is usually minor because the piezoresistor is much more highly doped than the background.

Devices with a constant stress, such as strain gauges, must be handled slightly differently. Following Tufte and Stelzer [210], we can define a conductivity weighted, average piezoresistance factor as

$$\bar{P} = \frac{\int_0^{t_c} P \mu n dz}{\int_0^{t_c} \mu n dz} \quad (2.91)$$

so that the sensitivity of strain gauges can be calculated as

$$\frac{\Delta R_{pr}}{R_{pr}} = \sigma \bar{P} \pi_L^{\max}. \quad (2.92)$$

2.4.3 Calculating sensor resolution

We can calculate the resolution for any sensor by combining (2.38) with expressions for the sensor noise and sensitivity. We will use a piezoresistive cantilever as a brief example to illustrate how to calculate resolution.

The noise sources for the cantilever include Johnson, Hooge, amplifier and thermomechanical noise sources. Combining (2.51), (2.53), (2.59) and (2.65), the RMS voltage noise of the sensor is

$$V_{noise} = \sqrt{V_J^2 + V_H^2 + V_A^2 + V_{TMN}^2} \quad (2.93)$$

$$= \left[(f_{\max} - f_{\min}) \left(4k_b T R + A_{VJ}^2 + 2A_{IJ}^2 \left(\frac{R}{2} \right)^2 + S_{FV}^2 \frac{4k_c k_b T}{\omega_0 Q} \right) + \ln \left(\frac{f_{\max}}{f_{\min}} \right) \left(V_{bridge}^2 \frac{\alpha}{2N_{eff}} + A_{VF}^2 + 2A_{IF}^2 \left(\frac{R}{2} \right)^2 \right) \right]^{1/2}. \quad (2.94)$$

The noise expressions assume a balanced Wheatstone bridge (Section 2.4.1) and

two silicon piezoresistors in the bridge, one for measurement and one for temperature compensation (Section 2.4.1).

Assuming that only one of the piezoresistors is mechanically loaded, we can calculate the force and displacement sensitivities ($\Delta V/F$ and $\Delta V/x$) of the cantilever from (2.70) as

$$\Delta V = \frac{V_{\text{bridge}}}{4} \frac{\Delta R}{R}.$$

The overall relative resistance change (2.90) is

$$\frac{\Delta R}{R} = \overline{\sigma_L^{\max}} \pi_L^{\max} \gamma \beta^*.$$

where $\overline{\sigma_L^{\max}}$ (2.78) is

$$\overline{\sigma_L^{\max}} = \frac{6(l_c - l_{\text{pr}}/2)}{w_c t_c^2} F.$$

The average maximum longitudinal stress can be rewritten in terms of displacement by combining $F = k_c x$ and the spring constant of a cantilever beam ($k_c = E_c w_c t_c^3 / 4l_c^3$) to yield

$$\overline{\sigma_L^{\max}} = \frac{3E_c t_c (2 - l_{\text{pr}}/l_c)}{4l_c^2} x.$$

The maximum longitudinal piezoresistive coefficient (π_L^{\max}) can be found in Table 2.2. The resistance factor, γ , is calculated from (2.72) and depends on the excess resistance in the device from interconnects and contacts. The sensitivity factor, β^* , is calculated from (2.89) and depends on the dopant concentration profile. Both the resistance and sensitivity factors vary from 0 to 1.

We can combine all of these equations to calculate the force and displacement sensitivities for an end-loaded cantilever beam as

$$S_{\text{FV}} = \frac{V_{\text{bridge}}}{4} \frac{6(l_c - l_{\text{pr}}/2)}{w_c t_c^2} \pi_L^{\max} \gamma \beta^* \quad (2.95)$$

and

$$S_{\text{XV}} = \frac{V_{\text{bridge}}}{4} \frac{3E_c t_c (2 - l_{\text{pr}}/l_c)}{4l_c^2} \pi_L^{\max} \gamma \beta^*. \quad (2.96)$$

Finally, the minimum detectable displacement (MDD) and force (MDF) are

$$\text{MDD} = \frac{V_{\text{noise}}}{S_{\text{XV}}} \quad (2.97)$$

and

$$\text{MDF} = \frac{V_{\text{noise}}}{S_{\text{FV}}}. \quad (2.98)$$

The difference in calculating the resolution of other piezoresistive sensors (e.g. pressure, surface stress, and acceleration sensors) lies mainly in the sensitivity equations. Design optimization results, however, will vary widely between sensor types. Considering the geometry scaling in (2.95) and (2.96), we can already see that a cantilever optimized for force sensing will be substantially longer, thinner, and narrower than one optimized for displacement sensing.

In the next section, we will consider the most common piezoresistor fabrication techniques and models for dopant concentration profiles in order to accurately calculate piezoresistor noise and β^* .

2.5 Piezoresistor fabrication and modeling

A necessary condition for piezoresistive sensor design is the accurate calculation of the concentration of dopant atoms as a function of depth from the surface. The dopant concentration profile, $N(z)$, determines the charge carrier and electrical resistivity profiles, $n(z)$ and $\rho(z)$. As we saw in the last section, these quantities determine the noise, sensitivity and resolution of a sensor.

We will review the three main fabrication techniques for doping silicon: ion implantation, epitaxy, and diffusion. We will also discuss polysilicon deposition and the tradeoffs between single-crystal and polycrystalline silicon. For each fabrication method, we will discuss the models we will use in the rest of the thesis to calculate dopant concentration profiles. Finally, we will discuss models for carrier mobility and resistivity, before discussing experimental techniques to verify concentration and resistivity profiles with respect to depth.

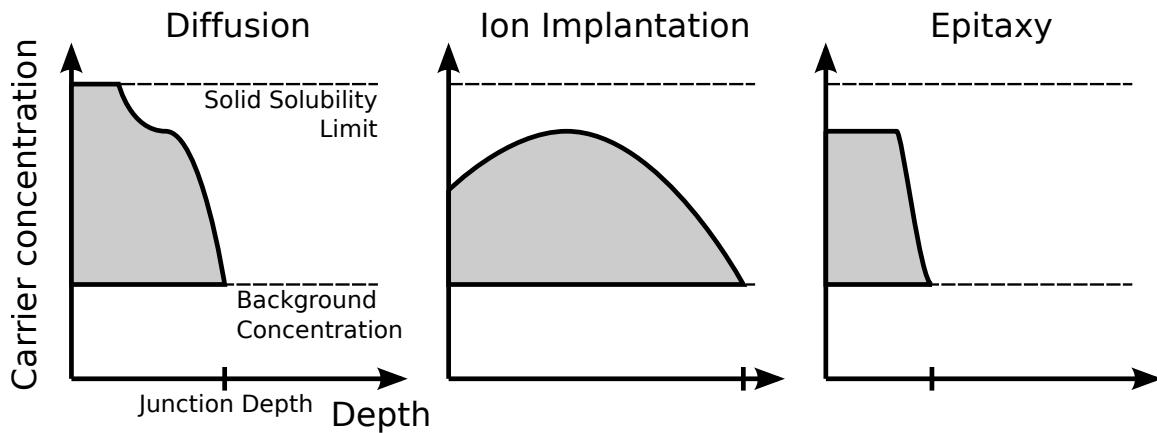


Figure 2.26: Comparison of dopant profiles produced by diffusion, ion implantation, and epitaxial doping. Diffusion is characterized by a high surface concentration and shallow junction depth before an optional drive-in step. Ion implantation can produce lower concentrations with more flexibility and accuracy, but has difficulty doping shallow layers due to diffusion. Epitaxial growth enables precise dopant profile control at the expense of increased system complexity and cost.

Predeposition, ion implantation and epitaxy result in very different dopant concentration profiles (2.26), leading to relative strengths and weaknesses for each. A complete review of doping techniques is available elsewhere [211]. The three main techniques are summarized in Table 2.5.

Ion implantation is the most common technique for fabricating piezoresistors. Advantages of ion implantation are the precise control of the dopant dose, concentration and depth. Ion implantation also exhibits the best cross-wafer uniformity of the three techniques. Downsides of ion implantation are the lattice damage intrinsic to the process and the accompanying need to anneal the wafer in order to activate the dopants and reduce lattice damage. Shallow piezoresistors are difficult to fabricate using ion implantation due to the post-implantation anneal step. Ion implanted piezoresistors have achieved the lowest Hooge factors, although data for the other two techniques is more limited.

Predeposition is a batch process while the other two techniques are generally serial processes, and it can be used to fabricate very shallow piezoresistors. It is also the simplest, most widely available technique of the three due to its longer history. However, predeposition generally yields less cross-wafer uniformity than ion implantation and the

Table 2.5: Summary of piezoresistor fabrication techniques. A hard mask capable of withstanding high temperature processing and with a low dopant diffusivity, such as SiO_2 or Si_3N_4 , is required for diffusion and epitaxy while photoresist is sufficient for masking ion implantation.

	Predeposition	Ion implantation	Epitaxy
Processing temperature ($^{\circ}\text{C}$)	750-1100	750-1100	600-1100
Processing pressure	atmospheric	high vacuum	low vacuum
Wafer throughput	high	high	low
Lattice damage	none	substantial	none
Surface concentration (cm^{-3})	$10^{20} - 10^{21}$	$10^{16} - 10^{20}$	$10^{16} - 10^{20}$
Minimum junction depth (nm)	< 100	> 500	< 100
Typical R_s nonuniformity (%)	3 – 5	< 0.5	< 1
Masking materials	hard mask	photoresist or hard mask	hard mask
Profile modeling	more complex	more complex	less complex

dopant concentration is fixed at the solid solubility concentration unless a drive-in step is included in the fabrication process. Lattice damage is minimal during predeposition, although excess dopants tend to aggregate at the surface.

Epitaxy is generally a serial process, but provides the most precise control over piezoresistor depth and concentration of the three techniques. Any damage introduced into the lattice during growth is minimal with proper growth conditions, so a post-growth anneal is not generally performed. Epitaxy has been used to fabricate the thinnest piezoresistors to date. However, processing complexity and equipment costs limit the availability of epitaxial reactors.

The most important facet of piezoresistor design and optimization is calculating the dopant concentration profile. Without an accurate predictive model, designs will not match experimental results and optimized designs will be inaccurate. We will model diffused and epitaxial piezoresistors analytically while ion implanted piezoresistors will be modeled using tabulated numerical results from TSUPREM4 (Synopsys, Mountain View, CA).

We will present models for the combinations of dopant species and fabrication types listed in Table 2.6. Predeposition dopant profiles will only be calculated for phosphorus doping. Boron and arsenic will not be investigated due to practical fabrication issues and the lack of accurate, analytical models for their dopant profiles (Section 2.5.2). Epitaxial

Table 2.6: Summary of the dopant species and fabrication technique combinations that we have developed models for. The force probes described in this thesis use diffused phosphorus piezoresistors.

	Diffusion	Ion implantation	Epitaxy
Arsenic	no	yes	yes
Boron	no	yes	yes
Phosphorus	yes	yes	yes

profiles do not vary significantly between the three dopant species, because we will focus on relatively low temperature epitaxy where diffusion lengths are short. Ion implantation, as noted earlier, will be treated numerically and so the differences in solid solubility, diffusivity, and oxide segregation between the dopant species will be handled by the software.

2.5.1 Processing fundamentals

In this section we will briefly discuss some of the fundamental concepts in piezoresistor fabrication before moving on to discuss dopant profile modeling in detail.

Dopant ionization

Throughout the thesis we will assume charge neutrality and complete dopant ionization so that $N(z) \approx n(z)$. The charge neutrality assumption only affects very shallow or narrow geometries. The Debye length, the characteristic distance over which charge screening occurs, is on the order of 10 nm in silicon so does not affect the vast majority of devices.

Incomplete dopant ionization, also known as freezeout, occurs when there is insufficient thermal energy to ionize the dopants [212]. Degenerately doped piezoresistors ($> 10^{19} \text{ cm}^{-3}$) have operated at temperatures below 4.3 K [213]. Many piezoresistive cantilevers used for cryogenic measurements have not been optimized for low temperature operation, so have dissipated more power than necessary, limiting their utility [214]. With the addition of accurate carrier freezeout modeling, the design techniques in this thesis could be applied to cryogenic temperatures with relatively little modification.

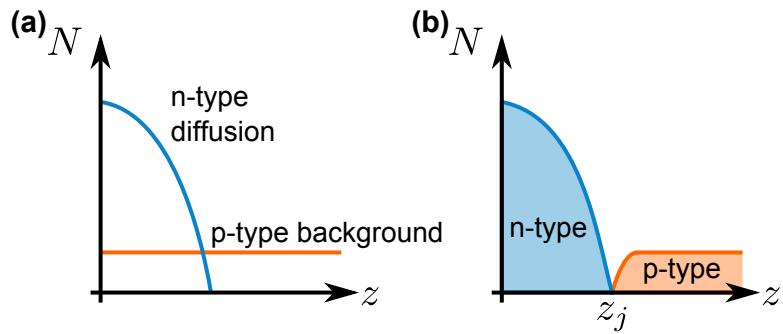


Figure 2.27: Piezoresistor formation. (a) Dopants are introduced to the wafer surface to define an n-type piezoresistor. The piezoresistor is electrically isolated from the bulk of the wafer by a pn junction. (b) The net carrier concentration drops sharply at the junction, which is depth z_j from the surface.

Finally, we will assume that the dopant concentration only varies vertically through the device. There are several potential design issues related to lateral dopant diffusion, which will be discussed in Section 2.8.4.

Junction isolation

Piezoresistor should generally be isolated from the bulk of the wafer by a pn junction. During piezoresistor formation, dopants are introduced at the wafer surface to define the piezoresistor. Electrical current is confined to the piezoresistor by forming a pn junction with the initial wafer material. For example, if the wafer is initially p-type, an n-type dopant would be introduced to the wafer (Figure 2.27a). The p-type and n-type dopants cancel one another, so the depth at which the introduced impurities are equal in concentration to the initial background concentration defines the junction depth. The net dopant concentration is shown in Figure 2.27b. By ensuring that the pn junction is always reverse biased, the depletion region of the diode can be used to electrically isolate the piezoresistor up to moderately high frequencies.

Concentration limits

Piezoresistor design optimization generally favors high dopant concentrations, as we will see. Two concentration limits are important to consider: the solid solubility limit and the

maximum electrically active concentration. These concentration limits determine the maximum possible dopant concentrations possible with diffused and epitaxial piezoresistors. Concentration limits for ion implanted dopant profiles will be handled by TSUPREM.

Dopants only contribute charge carriers when positioned in substitutional crystal lattice sites; interstitial dopant atoms do not [211]. The total dopant concentration (substitutional and interstitial sites) is always greater than or equal to the electrically active dopant concentration (substitutional sites only), and the proportion of activated dopants depends on the processing history of a wafer.

The total dopant concentration is bounded by the solid solubility limit. Above the solubility limit, dopant atoms aggregate and begin to come out of solution. The solid solubility limit varies with the dopant atom and maximum wafer processing temperature and is on the order of 10^{20} - 10^{21} cm $^{-3}$ for boron, phosphorus and arsenic at 1000°C. Solid solubility data of multiple atomic species was first tabulated by Trumbore in 1960 [215], and updated by Borisenko and Yudin in 1987 [216]. The maximum electrically active concentration also varies with temperature and can approach the solid solubility limit. Excess, electrically inactive dopants have been observed for phosphorus in particular [217, 218].

Figure 2.28 plots the solid solubility and electrically active concentration limits for boron and phosphorus as functions of temperature. Solmi et al. measured the total and electrically active concentrations for phosphorus [218]. Their best-fit models for the total and active phosphorus concentrations were

$$N_{\text{total}}^{\text{phos}} = 2.45 \times 10^{23} \text{cm}^{-3} \exp(-0.62 \text{eV}/k_b T) \quad (2.99)$$

$$N_{\text{active}}^{\text{phos}} = 9.2 \times 10^{21} \text{cm}^{-3} \exp(-0.33 \text{eV}/k_b T). \quad (2.100)$$

Solmi et al. observed the presence of electrically inactive, interstitial phosphorus for processing temperatures above 750°C. Vick et al. electrically measured the maximum boron concentration achievable by diffusion as a function of temperature [219]. We fit the Arrhenius equation to their data to obtain a best-fit model for the maximum electrically active boron concentration of

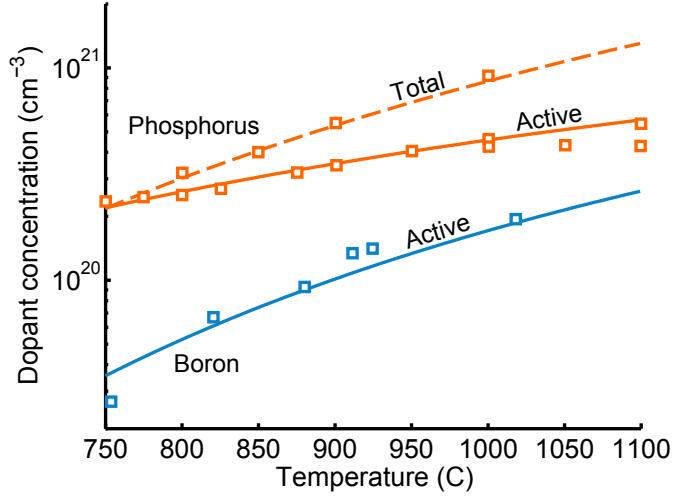


Figure 2.28: Total and electrically active boron and phosphorus concentrations vary with processing temperature. Solmi et al. measured the total and electrically active concentrations of phosphorus [218], while Vick et al. measured the electrically active boron concentration [219]. Excess, electrically inactive phosphorus was observed for anneal temperatures above 750°C.

$$N_{\text{active}}^{\text{boron}} = 1.9 \times 10^{22} \text{ cm}^{-3} \exp(-0.41 \text{ eV}/k_b T) \quad (2.101)$$

When calculating diffused and epitaxial profiles, we will take these concentration limits into account.

2.5.2 Predeposition

Diffusion is the migration of dopant atoms from a region of high concentration to a region of low concentration. The fabrication of piezoresistors using diffusion involves a pre-deposition and an optional drive-in step. During the pre-deposition step, a high concentration of dopant atoms are presented to the wafer, where they diffuse into the silicon.

Pre-deposition is typically performed at atmospheric pressure in the 750-1100°C temperature range. The dopant source can be solid-, liquid- or gas-phase [211, 220]. When a gas-phase source is used, such as diborane (B_2H_6), phosphine (PH_3), or arsine (AsH_3), it is carried in an inert gas such as N_2 or Ar. Most modern predeposition systems use

liquid- rather than gas-phase sources because they are dramatically safer and provide more consistent dopant profiles [221]. A common approach is to bubble the carrier gas through the liquid-source, most often BBr_3 or POCl_3 ; liquid-phase arsenic sources are not widely available. Solid-phase systems use solid wafer dopant sources, typically P_2O_5 or BN, which are interleaved with the device wafers. When both the wafers are heated, dopant atoms sublime and diffuse from the source wafers to the device wafers. A final option is the use of doped spin-on-glass layers rather than solid sources. [222–224].

Regardless of the dopant source, dopant atoms diffuse into the silicon lattice during pre-deposition via a two step process. First, a highly doped oxide is formed at the surface, followed by the silicon diffusion reaction [225]. For example, in the case of phosphorus doping the reaction proceeds according to



while in the case of boron doping it proceeds according to



During predeposition, the boundary condition is a constant surface concentration and the doping profile is often approximated by a complementary error function. The surface concentration is equal to the solid dopant solid solubility limit which varies with temperature, assuming that a high enough concentration of dopants is present in the furnace.

After pre-deposition, the oxide dopant source is removed in hydrofluoric acid (HF) and the dopants can be “driven-in” via high temperature annealing in either an inert or oxidizing environment. The drive-in step is performed in order to reduce the surface concentration and increase the junction depth, although it may be beneficial to skip the drive-in step

depending on the piezoresistive sensor design. The force probes developed in this thesis do not use a drive-in step in order to keep the dopants as close as possible to the surface.

One notable issue with boron pre-deposition is the formation of a $\approx 200 \text{ \AA}$ thick B:Si layer on the surface that can not be etched in hydrofluoric acid (HF). Typically after pre-deposition and before the drive-in step, if it is being performed, the doped glass on the surface should be removed in HF in order to remove the dopant source. In this case, a short oxidation step can be performed to oxidize the B:Si layer (e.g. 750°C for 20 minutes), allowing it to be removed in HF. The complete removal of the oxide and B:Si layer can be easily verified because the oxide and B:Si surfaces are hydrophilic while the bare silicon surface is hydrophobic.

Predeposition profiles are often modeled using a complementary error function. An important assumption in this approach is that the dopant diffusivity does not vary with concentration. Unfortunately, this assumption does not hold for boron, arsenic, or phosphorus when the dopant concentration exceeds the intrinsic carrier concentration, n_i , which is plotted in Figure 2.29. Above the intrinsic concentration level, the diffusivities of boron and arsenic increase approximately linearly with concentration [226]. In general, diffusivity varies with dopant concentration due to the formation of lattice defects, dopant clustering, and lattice strain [219, 227, 228]. Increased diffusivity at high dopant concentrations leads to more abrupt concentration profiles, and using a complementary error function to model piezoresistive devices will lead to substantial errors in calculating device performance. Typical predeposition concentrations are an order of magnitude larger than n_i (Figure 2.28).

We will only investigate phosphorus predeposition in this thesis because a simple, accurate dopant profile model is available and the practical issues with boron (B:Si formation) and arsenic (only gas-phase sources). Phosphorus predeposition is widely used and has been researched extensively [228–231]. Typical concentration profiles for phosphorus predeposition are shown in Figure 2.30. Phosphorus predeposition in particular yields concentration profiles with a characteristic plateau and kink, and more sophisticated modeling approaches are required to accurately model diffused piezoresistors.

We calculate dopant concentration profiles for phosphorus predeposition using Tsai's 1969 model [228]. Although more advanced models have since been developed [232, 233],

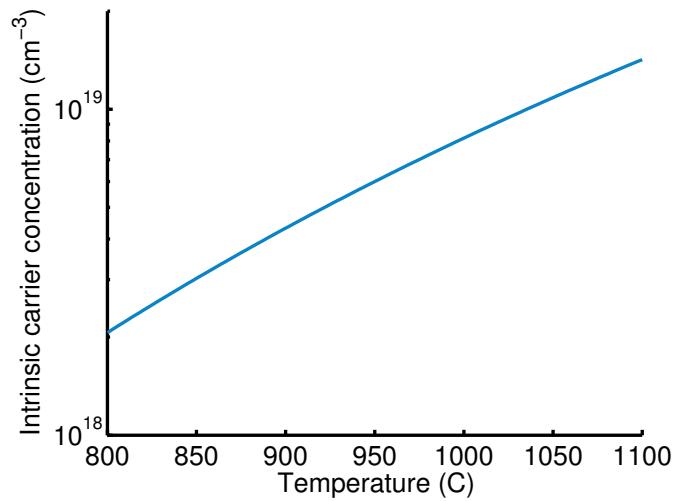


Figure 2.29: Intrinsic carrier concentration (n_i) as a function of processing temperature. Dopant diffusivity is constant below n_i but increases above it, leading to more abrupt concentration profiles than would otherwise be predicted. Predeposition can only be accurately modeled by a complementary error function for concentrations below n_i .

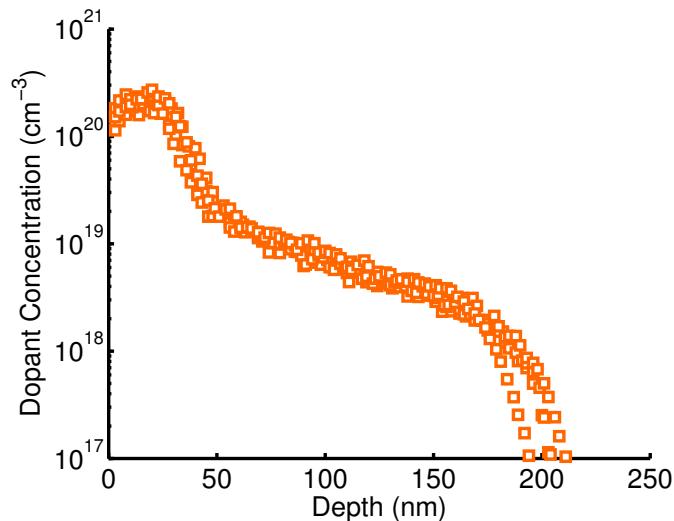


Figure 2.30: Electrically active dopant concentration profiles for phosphorus predeposition at 800°C for 35 minutes. The electrically active concentration was measured via spreading resistance analysis (SRA). The profile consists of a constant concentration surface region, a kink in the transition region, and a fast-diffusing tail due to variation in diffusivity with concentration. Four SRA profiles from separate wafers are plotted for comparison.

we have used the Tsai model to accurately model dopant concentration and sheet resistance for processing temperatures from 775 to 950°C. The model is simple to implement, computationally inexpensive, and accurate. In our experience, TSUPREM does not accurately reproduce the dopant concentration profiles measured experimentally.

In Tsai's model, the dopant profile consists of three regions: a surface layer, a transition region, and a normal diffusion region. The dopant concentration is constant in the surface region. The dopant profile consists of slow and fast diffusion processes, and the concentration beyond the surface region is determined by their sum. The slow diffusion process, with diffusivity D_A , determines the dopant concentration in the transition region. The fast diffusion process, with diffusivity D_B , determines the dopant concentration in the normal region. More specifically,

$$C(x, t) = C_S \quad \text{for } x \leq x_0 \quad (2.106)$$

$$C(x, t) = C_A(x, t) + C_B(x, t) \quad \text{for } x > x_0. \quad (2.107)$$

The depth of the surface layer is $x_0 = \alpha t$ where t is the time in seconds and α is the growth rate of the surface phase. The surface concentration, C_S , is the total dopant concentration (2.99). The concentration profile is calculated using the total phosphorus concentration before being capped at the electrically active concentration limit (2.100).

The slow and fast diffusing components are calculated from

$$C_A(x, t) = \frac{1 - \kappa}{2} C_S \exp \left[-\frac{\alpha}{2D_A} (x - \alpha t) \right] F_1(x, t) \quad (2.108)$$

$$C_B(x, t) = \frac{\kappa}{2} C_S \exp \left[-\frac{\alpha}{2D_B} (x - \alpha t) \right] F_2(x, t) \quad (2.109)$$

where $\kappa = C_B(x_0)/C_S$ and

$$F_1(x,t) = \operatorname{erfc}\left(\frac{x+\alpha t}{2\sqrt{D_A t}}\right) + \operatorname{erfc}\left(\frac{x-3\alpha t}{2\sqrt{D_A t}}\right) \quad (2.110)$$

$$F_2(x,t) = \operatorname{erfc}\left(\frac{x+\alpha t}{2\sqrt{D_B t}}\right) + \operatorname{erfc}\left(\frac{x-3\alpha t}{2\sqrt{D_B t}}\right). \quad (2.111)$$

There are four model parameters: D_A , D_B , α and $C_B(x_0)$. However, they vary with temperature and are modeled via Arrhenius kinetics as

$$D_A = D_{A0} \exp\left(-\frac{E_{DA}}{k_b T}\right) \quad (2.112)$$

$$D_B = D_{B0} \exp\left(-\frac{E_{DB}}{k_b T}\right) \quad (2.113)$$

$$\alpha = \alpha_0 \exp\left(-\frac{E_\alpha}{k_b T}\right) \quad (2.114)$$

$$C_B(x_0) = C_{B0} \exp\left(-\frac{E_{CB}}{k_b T}\right) \quad (2.115)$$

and there are a total of eight fitting parameters. In order to verify the model and provide fitting parameter estimates, Tsai doped samples from 820-1100°C via POCl_3 predeposition. He measured the electrically active dopant concentration profiles by iteratively anodically oxidizing the samples and measuring their sheet resistances. We have measured phosphorus predeposition dopant concentration profiles using spreading resistance analysis (SRA) and sheet resistance data. Our best fit parameters for POCl_3 -based predeposition from 775-950°C and 10-60 minutes are presented in Table 2.7.

Figure 2.31 compares the computed and measured electrically active concentration profiles for predeposition at 800°C for 35 minutes. The agreement between the model and experiments is excellent, and the three profile regions (plateau, kink, and tail) are captured by the model. Figure 2.32 compares the modeled and measured sheet resistance values from 775-950°C. The model error is less than 10% for all of the measured data points.

We found that α decreases more than expected for processing temperatures below 825°C. Tsai also noted lower than expected values of α at low temperatures, and did not use

Table 2.7: Best fit parameters for the Tsai model of phosphorus predeposition obtained by fitting to experimentally measured spreading resistance profiles. The α activation energy varies with temperature to accurately model low temperature predeposition processes ($< 820^{\circ}\text{C}$) where phosphorus diffusion through a surface oxide layer can become the rate limiting step.

	Prefactor	Activation energy
D_A	$200 \text{ cm}^2/\text{sec}$	3.77 eV
D_B	$2.3 \times 10^{-5} \text{ cm}^2/\text{sec}$	1.95 eV
α	0.12 cm/sec	$1.65 \text{ eV} + 0.05 \text{ eV} \exp[-0.015 \frac{1}{\text{K}}(T - 1073\text{K})]$
C_B	$3 \times 10^{23} \text{ cm}^{-3}$	0.88 eV

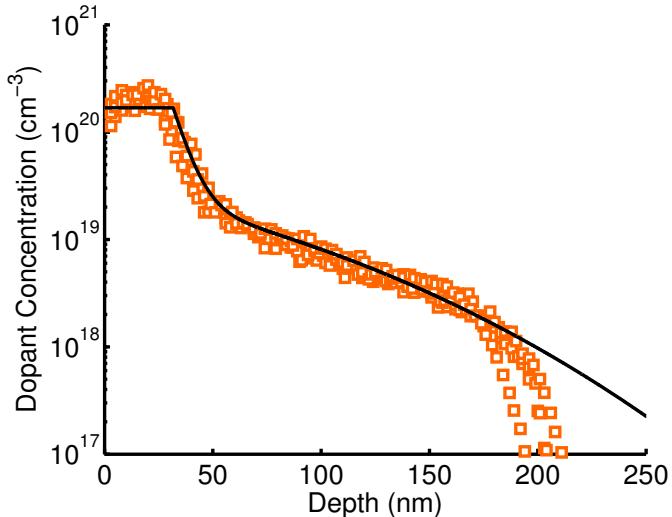


Figure 2.31: Comparison between experimental and modeled dopant profiles for phosphorus predeposition 800°C for 35 minutes. The plateau, kink, and tail regions of the profile are captured. The model overestimates the dopant concentration beyond 200 nm ($< 10^{18} \text{ cm}^{-3}$), however there is no impact on piezoresistor modeling accuracy because the vast majority of the current is carried near the surface.

the 820°C data in calculating the model parameter estimates. The most likely mechanism is that the diffusivity of phosphorus in SiO_2 is relatively slow and varies dramatically with temperature. Most phosphorus predeposition processes include an initial oxidation step (e.g. 2-5 minutes) to prevent surface pitting and roughness. Thus, diffusion through surface SiO_2 can become the rate limiting diffusion process at temperatures below 825°C .

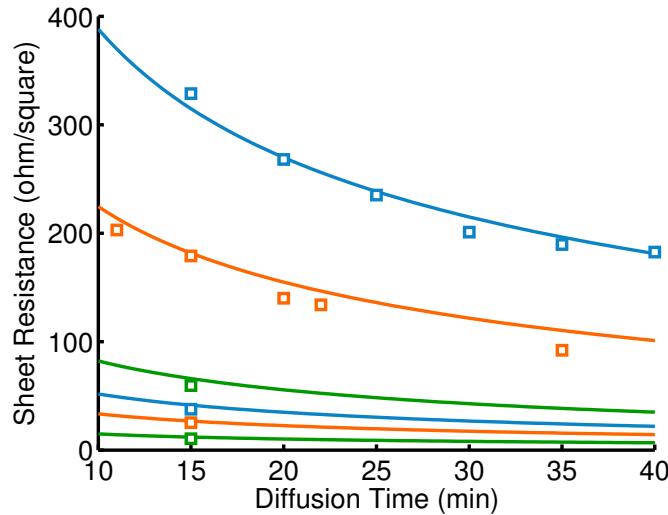


Figure 2.32: Comparison between experimental and modeled sheet resistance data. Phosphorus predeposition was performed at temperatures from 775-950°C and the sheet resistance was measured via 4-point probe. The model error is less than 10% for all of the measured data points. The curves in the plot correspond to 775, 800, 850, 875, 900 and 950°C, from top to bottom.

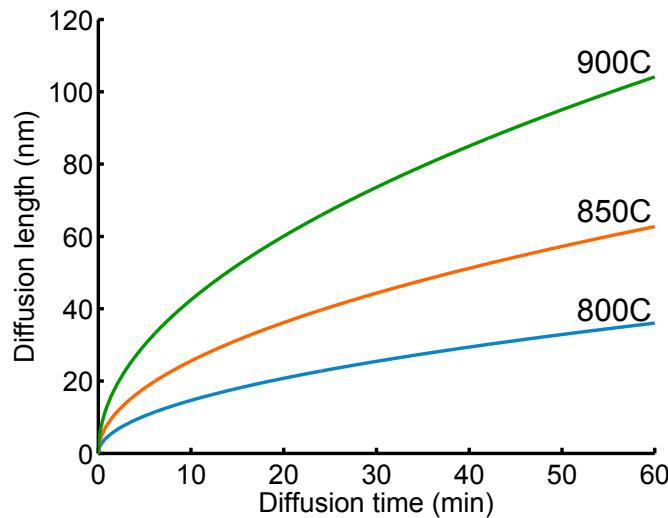


Figure 2.33: Calculated phosphorus diffusion lengths in SiO_2 as a function of predeposition time and temperature. Calculated results agree well with our experimental observations of the thickness of phosphosilicate glass (PSG) removed in a brief hydrofluoric acid (HF) dip.

The relatively slow diffusion of phosphorus through SiO_2 makes it an excellent masking material. Sah et al. measured the diffusivity of phosphorus in SiO_2 in 1959 [234]. Their data was later reanalyzed by Ghezzo and Brown and tabulated with over common silicon dopants [235]. The diffusivity of other dopants in SiO_2 and other materials can be found in an unpublished book chapter by Jones [236].

Phosphorus diffusion in SiO_2 obeys Arhenius kinetics with $E_a = 2.3 \text{ eV}$ and $D_0 = 5.73 \times 10^{-5} \text{ cm}^2/\text{s}$ [235]. The diffusion length of phosphorus in SiO_2 is plotted in Figure 2.33. Our own experimental measurements indicate a 60 nm phosphorus diffusion length during a 900°C , 20 minute diffusion. We measured the diffusion length by measuring the thickness of the thermal oxide mask before the diffusion, and then afterwards removing the deposited PSG and doped oxide layer in 50:1 hydrofluoric acid before measuring the remaining oxide thickness. Dilute HF etches undoped oxide exceedingly slowly (50 A/min) compared to phosphorus doped oxide ($> 3000 \text{ A/min}$). The minimum required oxide mask thickness (or maximum diffusion time) can be calculated by equating the diffusion length and oxide thickness, $2\sqrt{Dt} = t_{\text{oxide}}$, although we would recommend including a safety factor of at least 2 if possible to avoid shorting out the entire wafer.

2.5.3 Ion implantation

Two important downsides of pre-deposition are its high surface dopant concentration, fixed at the dopant solid solubility limit, and its relatively high dose nonuniformity due to the strong variation in diffusivity with temperature. Ion implantation was researched extensively in the 1950s and 1960s as an alternative to pre-deposition that could provide better dose uniformity and lower dopant concentrations [237–239]. The depth distribution of implanted ions and their electrical properties were studied extensively using the Hall effect and analytical theories [240–243]. The lattice damage induced by ion implantation was recognized early on [244, 245]. There are several excellent reviews on the history of ion implantation [246, 247] and modern ion implantation theory and simulation tools are discussed extensively by Plummer et al. [211]. Ion implantation gained wide use in the 1980s and remains the preferred method in commercial production today.

In ion implantation, dopant atoms are ionized and accelerated at high energy (keV to

MeV) into the substrate. The substrate is grounded, which allows the implanted dose to be measured and accurately controlled by integrating the ion current [211]. Additionally, grounding the substrate allows the ionized atoms to be neutralized immediately. The ions also leave a cascade of damage in the crystal structure [244]. Any layer thick or dense enough to block the implanted ions, such as photoresist, silicon oxide, silicon nitride, or metal, can be used to mask the implantation. Photoresist is a common choice. Typical silicon piezoresistor doses range from 1×10^{14} to $5 \times 10^{16} \text{ cm}^{-2}$, with energy ranges from 30 to 150 keV [42]. The typical sheet resistance nonuniformity is $< 0.5\%$ in modern ion implanters [248], while we typically observe 3-5% nonuniformity for POCl_3 doped wafers in an academic research facility.

Dopant distribution is approximated by a symmetric Gaussian distribution [244]. Most implants are performed with a 7° tilt of (100) silicon wafers to avoid ion channeling, a phenomenon where ions deeply traverse gaps in the lattice without scattering. Larger implant angles (7° - 45°) are sometimes used to form piezoresistors on etched sidewalls to enable in-plane strain measurements [197, 249–252].

One major disadvantage of ion implantation is the damage to the crystal. Lattice scattering leads to 1/f noise, while lattice order is mostly restored by high-temperature dopant activation and annealing [182, 206, 244]. However, shallow junctions are difficult to obtain with high crystal quality, so ion implantation is problematic for thin piezoresistive sensors where conduction needs to be tightly confined to the surface. Some researchers have used low energy [73], higher mass ions [253], or intentional lattice damage [254] to extend ion implantation to shallower piezoresistors. Parameters that affect the junction depth include the acceleration energy, the ion mass, and the stopping power of the material [241].

We will present tabulated ion implantation results for quickly and efficiently calculating the electrical properties of ion implanted piezoresistors. The damage induced by ion implantation can lead to complex dopant activation and diffusion behavior, making a numerical approach necessary [255]. Other issues that make a numerical approach necessary are oxidation enhanced diffusion, dopant segregation at the oxide interface, and diffusivities that vary with dopant concentration and lattice defect densities.

Numerical design optimization requires a large number of computational steps. Finding

the design optimum from a random initial design might require 5,000 iterations of the optimization algorithm. At each step, the algorithm determines how to improve the design by computing the gradient of the target function in the basis defined by the design parameters. If two of the parameters affect the dopant profile (e.g. predeposition temperature and time), then the dopant profile will need to be computed up to 10,000 times. If calculating the dopant profile takes 2 milliseconds, then optimizing the sensor design will take less than 30 seconds. Simple analytical models (diffusion and epitaxy) can be optimized to run extremely quickly. However, more complex models that use iterative numerical methods can take tens of seconds to complete, and nobody wants to wait 27 hours for a computation to finish.

Thus, we take the approach that we developed with Sung-Jin Park in 2010 of computing a lookup table for dopant profile dependent parameters: β^* , N_z , \bar{P} and R_s [198, 202]. The only problem is the sensitivity factor, which depends on the sensor thickness. Calculating a β^* lookup table for every possible thickness isn't practical, but β^* can be reformulated to be thickness independent. We do so by shifting the coordinate system in (2.89) from the neutral axis to the sensor surface by introducing $z' = t_c/2 - z$, yielding

$$\begin{aligned}\beta^* &= \frac{2}{t_c} \frac{\int_{-t_c/2}^{t_c/2} P \mu n(t_c/2 - z') dz}{\int_{-t_c/2}^{t_c/2} \mu n dz} \\ &= \beta_1^* - \frac{2}{t_c} \beta_2^*. \end{aligned} \quad (2.116)$$

The two new parameters, β_1^* and β_2^* , are defined by

$$\beta_1^* = \frac{\int_0^{t_j} P \mu n dz'}{\int_0^{t_j} \mu n dz'} \quad (2.117)$$

$$\beta_2^* = \frac{\int_0^{t_j} P \mu n z' dz}{\int_0^{t_j} \mu n dz}. \quad (2.118)$$

where t_j is the junction depth. The first parameter, β_1^* , does not have any position

Table 2.8: Process parameters used in the TSUPREM ion implantation simulations. A range of conditions are simulated and a lookup table is computed in order to make numerical design optimization viable.

Process	Parameters
1) Grow 250 Å screening SiO ₂	Wet O ₂ (850°C, 17 min)
2) Ion implantation	B, As, or P, 2×10^{14} to 2×10^{16} cm ⁻² and 20 to 80 keV
3) Strip screening SiO ₂	
4) Wet oxidation (optional)	900, 1000, or 1100 °C for 66, 15 or 5 minutes
5) Inert N ₂ anneal	900, 1000 or 1100 °C for 15-120 minutes

dependence, so is just a conductivity weighted average of the piezoresistance factor. If the dopant concentration is low then β_1^* will be high. The second parameter, β_2^* , is a stress and conductivity weighted average of the piezoresistance factor. If the dopant atoms are near the surface, then $\beta_2^* \rightarrow 0$ and $\beta^* \rightarrow \beta_1^*$. However, if the dopants are distributed uniformly over the sensor thickness, then $\beta_2^* \rightarrow \beta_1^* t_c / 2$ and $\beta^* \rightarrow 0$.

Separating β^* into β_1^* and β_2^* allows us to precompute β^* for arbitrary sensor thickness. By running a large number of TSUPREM simulations offline, we can calculate a lookup table for all of the parameters that vary with the dopant concentration profile (β^* , N_z , \bar{P} and R_s) and quickly look them up during numerical optimization.

If we compare (2.91) and (2.117), we note that $\bar{P} = \beta_1^*$. Thus, the data we tabulate for ion implanted piezoresistors can be used to design constant strain devices ($\bar{P} = \beta_1^*$) in addition to linear strain devices ($\beta^* = \beta_1^* - 2\beta_2^*/t_c$).

We simulated ion implanted dopant profiles using TSUPREM4 (Synopsys, Mountain View, CA). The basic fabrication process used in the simulations is noted in Table 2.8. After the wafers are ion implanted they are annealed to activate the dopants and reduce lattice damage from the implantation step. Implantation was analytically modeled using a dual Pearson distribution. The implant moment tables used to calculate the profiles included full energy, dose and tilt dependence. We chose the maximum implantation energy (80 keV) based upon the energy range of the lookup table data. The models used for dopant activation and diffusion included dopant-defect clustering, oxidation-enhanced diffusion, concentration and damage effects, transient defect generation, pairing and recombination.

We assumed a wafer tilt of 7° to avoid ion channeling and a background resistivity of 10 Ω-cm. The simulation input file is reproduced in its entirety in Appendix H for additional implementation details.

We computed ion implantation lookup tables for arsenic, boron and phosphorus over a wide range of process conditions. The data was calculated for implantation energies of 20, 50 and 80 keV, doses of 2×10^{14} , 2×10^{15} and 2×10^{16} cm⁻², post-implantation anneal temperatures of 900, 1000 and 1100 °C, and anneal times ranging from 15 to 120 minutes in 15 minute increments. A 250 Å thick oxide layer is grown before the implantation step in order to protect the silicon from contamination.

We assumed two different conditions for the anneal. In the first we assumed that the wafers were oxidized in steam to obtain a 1500-2000 Å thick passivation oxide before performing the inert anneal. In the second condition we assumed that the wafers were not oxidized and the anneal was performed only in an inert atmosphere. In the first case, the passivation oxide was grown at 900, 1000, or 1100 °C for 66, 15 or 5 minutes, respectively. If a passivation oxide will be regrown during the anneal step then the 250 Å thick protection oxide should be stripped first. However, if the wafer will not be reoxidized then the protection oxide should be left in place during the anneal and stripped afterwards in order to prevent dopant outgassing.

A total of 1296 different process flows were simulated to build the lookup tables. In order to efficiently compute the data we used two scripts and a template for the TSUPREM input file. The first script inserted the appropriate dopant, energy, dose, temperature and time into the template, generating a set of TSUPREM input files. After running the simulations, the script saved the concentration profiles, where the second script aggregated the data into a relatively small (3 MB) Matlab data file. The two scripts and template are presented in Appendix H. The lookup tables could be readily extended to an alternative process flow or larger design parameter space. The simulations took about 12 hours to run, but allow the electrical properties of ion implanted piezoresistors to be calculated in milliseconds.

There are five process parameters that determine the dopant profile for ion implanted piezoresistors: the dopant, the implanted dose and energy (N_{implant} and E_{implant}), and the time and temperature of the inert anneal (t_{anneal} and T_{anneal}). The following lookup tables

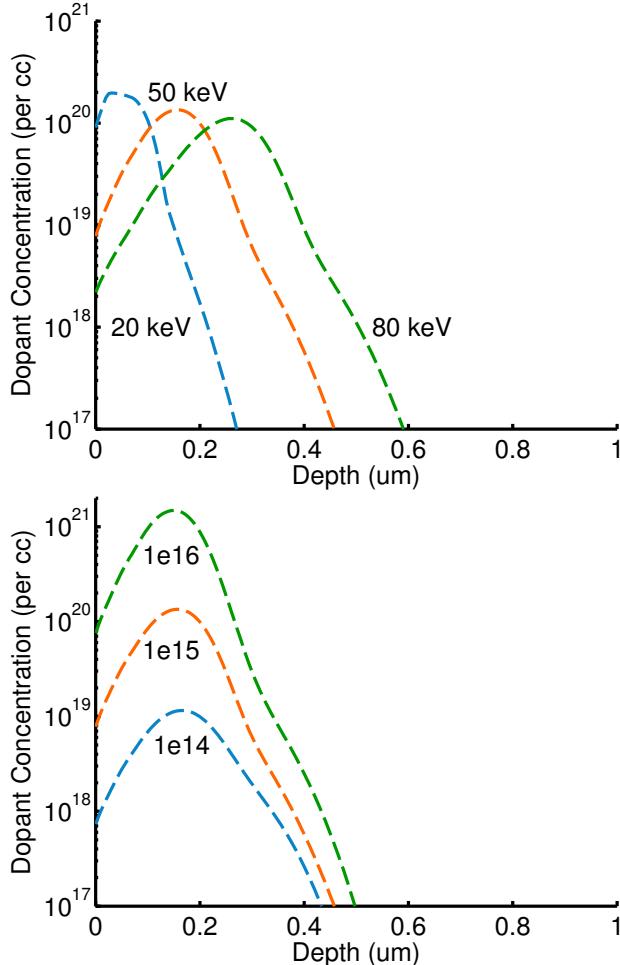


Figure 2.34: Changes in ion implantation energy (top) and dose (bottom) affect the as-implanted dopant concentration profile. The simulation results are for boron implanted at energies of 20, 50 and 80 keV for a dose of $2 \times 10^{15} \text{ cm}^{-2}$ (top) and for an energy of 50 keV and doses of 2×10^{14} , 2×10^{15} and $2 \times 10^{16} \text{ cm}^{-2}$ (bottom). The concentration and mean projected range scale roughly linearly with dose and energy, respectively. The initial dopant concentration profile is modified substantially by the post-implantation anneal.

provide β_1^* , β_2^* , R_s and N_z over a wide process parameter range. Note that N_z is not simply equal to the dose due to the possibility of incomplete dopant activation and because the dopant concentration is not constant so the current density varies. Similarly, the sheet resistance can not be calculated directly from the process parameters and needs to be simulated.

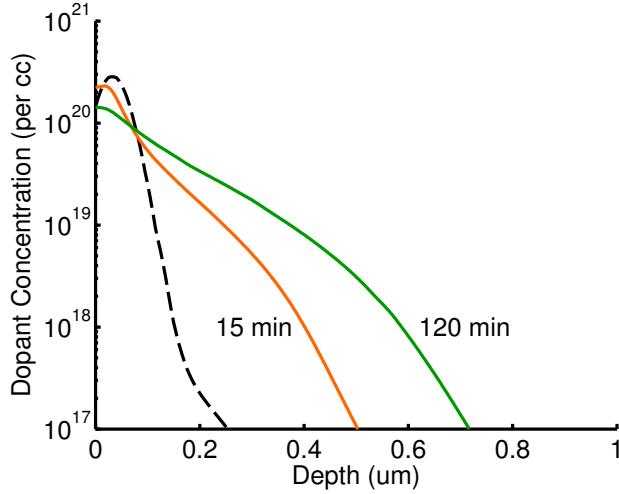


Figure 2.35: Comparison between the initial phosphorus concentration profile (dashed line) and concentration profiles after 15 and 120 minutes of annealing at 900 °C . The profiles are calculated for a 50 keV implant with a dose of $2 \times 10^{15} \text{ cm}^{-2}$. The dopants are not electrically active initially, and the anneal integrates them into the crystal lattice and reduces implantation-induced lattice damage. Increasing the temperature or duration of the anneal impacts the noise and sensitivity of the sensor in a variety of ways.

Variations in the initial, as-implanted dopant concentration profile with changes in energy and dose are shown in Figure 2.34. The simulations were run for boron implanted at energies of 20, 50 and 80 keV at a dose of $2 \times 10^{15} \text{ cm}^{-2}$ and for an energy of 50 keV with doses of 2×10^{14} , 2×10^{15} and $2 \times 10^{16} \text{ cm}^{-2}$. The concentration and mean projected range scale roughly linearly with the dose and ion energy, respectively, over the range that we will simulate for piezoresistor fabrication [241].

Figure 2.35 demonstrates the evolution of the dopant concentration profile over time. For the simulation, phosphorus ions were implanted at 50 keV and a dose of $2 \times 10^{15} \text{ cm}^{-2}$ before being annealed at 900 °C for 15 or 120 minutes. The passivation oxide was not regrown before the anneal, but the protection oxide from the implantation step was left in place as discussed earlier. The dopant atoms are initially located within the first 200 nm from the surface (dashed line) but they are not electrically active. The anneal is required to activate them and anneal out lattice damage induced by the ion implantation. At high implantation doses (e.g. $2 \times 10^{16} \text{ cm}^{-2}$) the damage can be great enough to make the crystal amorphous. In that case, the anneal restores the crystallinity of the lattice through epitaxial

regrowth by using undamaged material as a growth template.

The parameters chosen for the anneal affect sensor performance in several ways. Increasing either the time or temperature of the anneal increases the diffusion length, which reduces 1/f noise by reducing the Hooge factor (Section 2.4.1). Increasing the temperature leads to an increase in the solid solubility limit (Section 2.5.1), which can increase the dopant concentration near the surface if the dose is high enough that the dopant concentration is limited by solubility. A higher temperature, assuming that the time is reduced to maintain a constant diffusion length, accordingly leads to a smaller piezoresistance factor but shifts conduction closer to the surface, decreasing both β_1^* and β_2^* .

Increasing the duration of the anneal, in contrast, reduces the concentration near the surface and shifts the conduction further from the surface, increasing both β_1^* and β_2^* . Sensors with constant strain through their thickness (uniaxial strain gauges) do not depend on β_2^* for their operation, so a longer anneal and lower dopant concentration is good from a sensitivity standpoint but bad from the standpoint of temperature stability. Sensors with a linear strain gradient through their thickness (all other sensors) will experience either increased or decreased sensitivity depending on the thickness of the sensor, which determines the relative weighting of β_1^* and β_2^* (2.116).

Figure 2.36 compares dopant concentration profiles for arsenic, boron and phosphorus in processes with and without the growth of a passivation oxide. The profiles are calculated for 50 keV, $2 \times 10^{15} \text{ cm}^{-2}$ implants that are annealed in an inert atmosphere at 900 °C for 60 minutes. In the passivated process, a 1500-2000 Å thick passivation oxide is grown before the inert anneal. The depth of the piezoresistor increases going from arsenic to phosphorus to boron as the atomic mass of the dopant decreases (75, 31 and 10.8 amu, respectively) and the dopant diffusivity increases. The peak concentration of boron is lower than either arsenic or phosphorus due to its lower solid solubility limit and higher diffusivity. The oxidation step modifies the concentration profiles through oxidation-enhanced diffusion, dopant segregation at the Si-SiO₂ interface and the increased time at high temperature.

Figure 2.37 compares dopant profiles for phosphorus and boron for anneals with and without the growth of a passivation oxide. Both profiles extend slightly deeper into the silicon due to oxidation-enhanced diffusion and the increased time at high temperature.

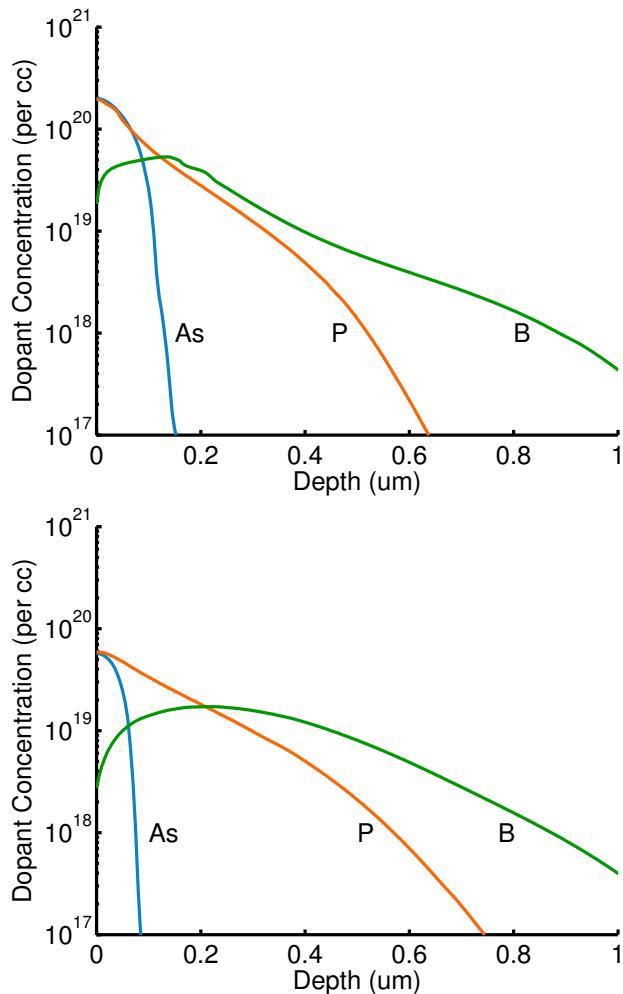


Figure 2.36: Comparison between As, B and P concentration profiles for inert (top) and passivated (bottom) processes. The profiles are calculated for 50 keV, $2 \times 10^{15} \text{ cm}^{-2}$ implants that are annealed in an inert atmosphere at 900 °C for 60 minutes. The depth of the piezoresistor increases going from arsenic to phosphorus to boron as the atomic mass of the dopant decreases and the dopant diffusivity increases. The oxidation step modifies the concentration profiles through oxidation-enhanced diffusion, dopant segregation and the increased time at high temperature.

But the difference is particular significant for boron, where the surface concentration has decreased approximately 20-fold.

Oxidation particularly causes the dopants to redistribute at the surface. Silicon at the surface is consumed during the oxidation, and the dopant atoms once contained in those

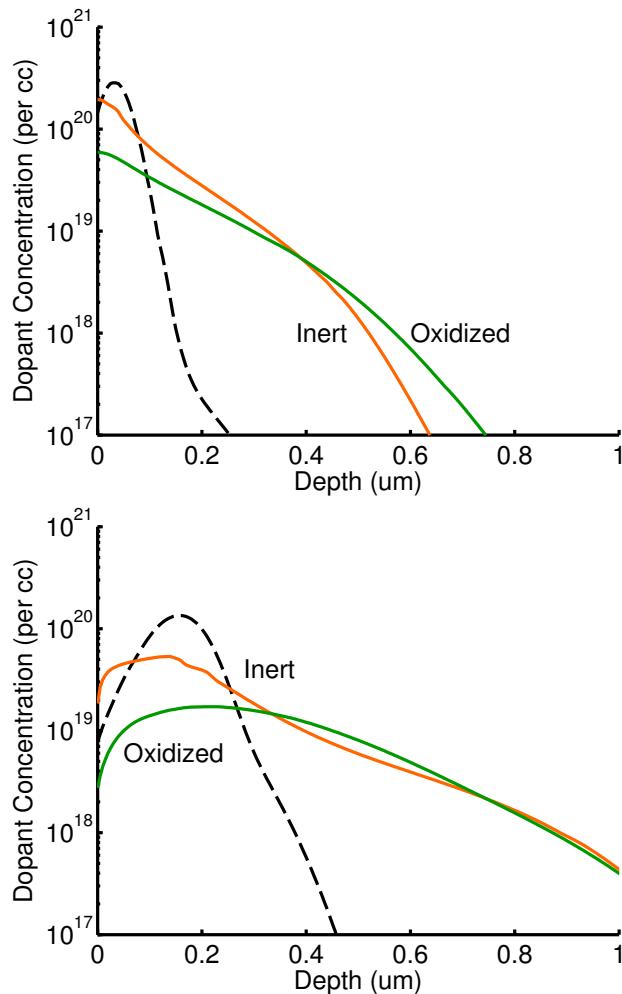


Figure 2.37: Phosphorus (top) and boron (bottom) dopant concentration profiles with and without the growth of a passivation oxide immediately before the anneal. The profiles are calculated assuming a 50 keV, $2 \times 10^{15} \text{ cm}^{-2}$ implant and an inert anneal at 900 °C for 60 minutes. Oxidation modifies the surface concentration due to dopant segregation and the diffusion length increases due to oxidation-enhanced diffusion and the additional time at high temperature.

layers either preferentially shift into the oxide or follow the silicon interface, causing pileup at the surface. A segregation coefficient less than one indicates that the dopants will segregate into the oxide, while a coefficient greater than one indicates that the dopants will pileup at the silicon interface. The segregation coefficient of boron is about 0.3 while the coefficients for arsenic and phosphorus are more than 1000 [236, 256]. Including a

reoxidation step reduces the surface concentration of boron 20-fold while only reducing the surface concentrations of arsenic and phosphorus by a factor of two. While the presence of the 250 Å thick protection oxide (in the case where the oxide is not regrown) does modify the profile slightly, the impact is much less than the reoxidation case.

The Tables in Appendix G can be used to calculate β_1^* , β_2^* , R_s , N_z and t_j for arsenic, boron and phosphorus implantation over a wide design parameter range. For design optimization, we allow the process parameters to vary continuously by linearly interpolating between the simulated parameter combinations. The lookup table used for optimization also includes the dopant concentration up to 5 μm from the wafer surface at 10 nm intervals, allowing the raw simulation data to be recovered.

2.5.4 Epitaxy

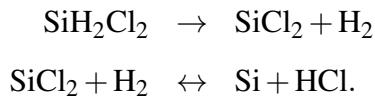
The primary performance limitation of ion implantation for piezoresistor fabrication is the diffusion that necessarily accompanies dopant activation and annealing, which is addressed through epitaxial piezoresistor growth. Conventional epitaxial growth is performed at high temperatures (800-1300°C) and reduced pressure (10-200 Torr). The four most commonly used silicon gas sources are silicon tetrachloride ($SiCl_4$), trichlorosilane ($SiHCl_3$), dichlorosilane (SiH_2Cl_2), and silane (SiH_4) [211]. A clean surface is necessary to obtain a high quality epitaxial layer. Contaminants and native oxide will prevent single-crystal growth. An in situ HCl clean can remove wafer contaminants and native oxide. Halide source gases are used to grow silicon with the advantage that chlorine is one of the net byproducts. The chlorine removes metal contaminants from the deposited silicon film, resulting in better quality single-crystal silicon.

Selective epitaxy is the deposition of single crystal silicon only on silicon surfaces rather than all exposed surfaces. Joyce and Baldrey first demonstrated the selective deposition of silicon epitaxial layers in 1962 using silicon tetrachloride at 1200°C and an oxide mask [257]. The presence of hydrochloric acid (HCl), whether from silicon source gas as in silicon tetrachloride or in a separate gas flow, acts to suppress the nucleation of silicon clusters on surfaces less favorable for growth than bare silicon. Silicon dioxide is generally used as a masking layer, although silicon nitride has also been demonstrated.

Selective epitaxy has been used both for undoped and various doping levels up the solid solubility limits of the dopant atoms. The selective growth of epitaxial layers on existing single crystal silicon is possible using a standard LPCVD epitaxial reactor. The HCl continuously etches the silicon, and due to the higher concentration of silicon atoms required on a dielectric surface to nucleate a grain, it is possible to continuously remove silicon from atop a dielectric layer while depositing on the exposed single crystal silicon. Since its initial development, selective epitaxy has been optimized to work at lower temperatures, allowing for less dopant diffusion [211, 258–262].

Silicon tetrachloride is widely used because it is stable, has a low vapor pressure and leaves little excess coating on the reaction chamber. However, it requires a high temperature for the reaction to proceed (1100 - 1300°C). Silane is widely used for the deposition of polysilicon, but is not favored for epitaxy or selective epitaxy because of its lower deposition rate at any given temperature compared to chlorinated silicon sources [260]. Of the remaining gases, dichlorosilane (DCS) is optimal for selective silicon epitaxy due to its low deposition temperature, high rate of deposition, low nucleation rate on silicon dioxide, and good selectivity at reduced pressure [263]. Epitaxial silicon films may be doped during the deposition by introducing appropriate dopant source gases such as AsH₃, PH₃, or B₂H₆ into the chamber along with the silicon source gases.

The reactions for the deposition of silicon from dichlorosilane are



The second reaction is reversible, and the growth rate is reduced by an increase in the HCl concentration. Selectivity improves as the temperature, pressure, or growth rate are decreased, or as the HCl concentration is increased [259]. Selective epitaxy at atmospheric pressure typically requires the addition of HCl, with a typical combination of gases being DCS, hydrogen, hydrogen chloride. At pressures of less than 100 torr, the addition of HCl is not required because the critical size for grain growth of the silicon on the dielectric is larger [264].

Epitaxial silicon does not require annealing to activate the dopants or reduce lattice disorder. Without an additional annealing step, dopants diffuse less and the dopant concentration profile can approximate a sharp step. Harley [36] and Liang [265] demonstrated the use of epitaxially grown doped silicon to form piezoresistors in ultra-thin cantilevers less than 100 nm thick.

The dopant concentration profile of an epitaxial piezoresistor, including dopant diffusion, can be calculated from

$$N = \frac{N_{\text{epi}}}{2} \operatorname{erfc} \left(\frac{z - t_{\text{pr}}}{2\sqrt{Dt}} \right) + N_{\text{B}} \quad (2.119)$$

where N_{epi} is the dopant concentration of the epitaxial layer, t_{pr} is the piezoresistor thickness, D is the dopant diffusivity, t is the processing time and N_{B} is the background concentration.

Regarding notation, we will use t_{pr} throughout the thesis to denote the nominal thickness of epitaxial piezoresistors and t_c to denote the total sensor thickness in order to avoid confusion with t , which will be consistently used to denote time. The piezoresistor thickness is not well defined for diffused or ion implanted piezoresistors, so we will only use t_{pr} in discussing epitaxial piezoresistors.

When the diffusion length is relatively short compared to the piezoresistor thickness, the concentration profile can be approximated as a concentration step where

$$N(z) = N_{\text{epi}} \text{ for } z \leq t_{\text{pr}} \quad (2.120)$$

$$= N_{\text{B}} \text{ for } z > t_{\text{pr}}. \quad (2.121)$$

However, if there is significant dopant diffusion from either a low epitaxial deposition rate or subsequent high temperature processing, the simplified step model is not accurate and (2.119) should be used instead or the profile should be modeled using TSUPREM.

Figure 2.38a compares epitaxial dopant profiles over a range of deposition rates for a 100 nm thick piezoresistor grown at 900°C and doped to $5 \times 10^{19} \text{ cm}^{-3}$. Curves are plotted for growth times of 10, 100 and 1000 minutes. Concentration profiles were calculated from (2.119). The growth rate, diffusion length, sensitivity factor and sheet resistance for each

Table 2.9: The sensitivity factor and sheet resistance for three different boron epitaxy growth rates. The sensitivity factor remains roughly constant until the diffusion length during epitaxial deposition is comparable to the piezoresistor thickness.

	10 minutes	100 minutes	1000 minutes
Growth rate (nm/min)	10	1	0.1
Diffusion length, $2\sqrt{Dt}$ (nm)	13	42	132
Sensitivity factor, β^* (-)	0.33	0.31	0.17
Sheet resistance (ohm/sq)	185	175	148

profile is presented in Table 2.9. The growth rates used are all relatively slow; growth rates on the order of 300 nm/min have been reported for selective epitaxy at 900°C [264].

There is almost no change in β^* until the diffusion length is on the order of the piezoresistor thickness (Figure 2.38b). Similarly, there is only a small change in the sheet resistance (due to the higher mobility at lower concentrations). Thus, a step profile can be used to model epitaxial piezoresistor with negligible error (< 5%). More detailed, numerical models (e.g. TSUPREM) can be used once a design has been optimized using the approximate model. Diffusion during epitaxial growth has an insignificant impact on piezoresistor sensitivity unless the growth rate is exceptionally low or additional thermal processes are performed.

The electrical and piezoresistive properties of an epitaxial piezoresistor with negligible diffusion can be calculated as

$$\beta^* = P(1 - t_{pr}/t_c) \quad (2.122)$$

$$\bar{P} = P \quad (2.123)$$

$$R_s = \rho / t_{pr} \quad (2.124)$$

$$N_z = N_{epit} t_{pr} \quad (2.125)$$

where P and ρ are calculated as functions of piezoresistor dopant concentration from (2.27) and (2.129), respectively. Epitaxial piezoresistors are far simpler to model than their diffused or ion implanted counterparts, which makes them useful for quickly designing a

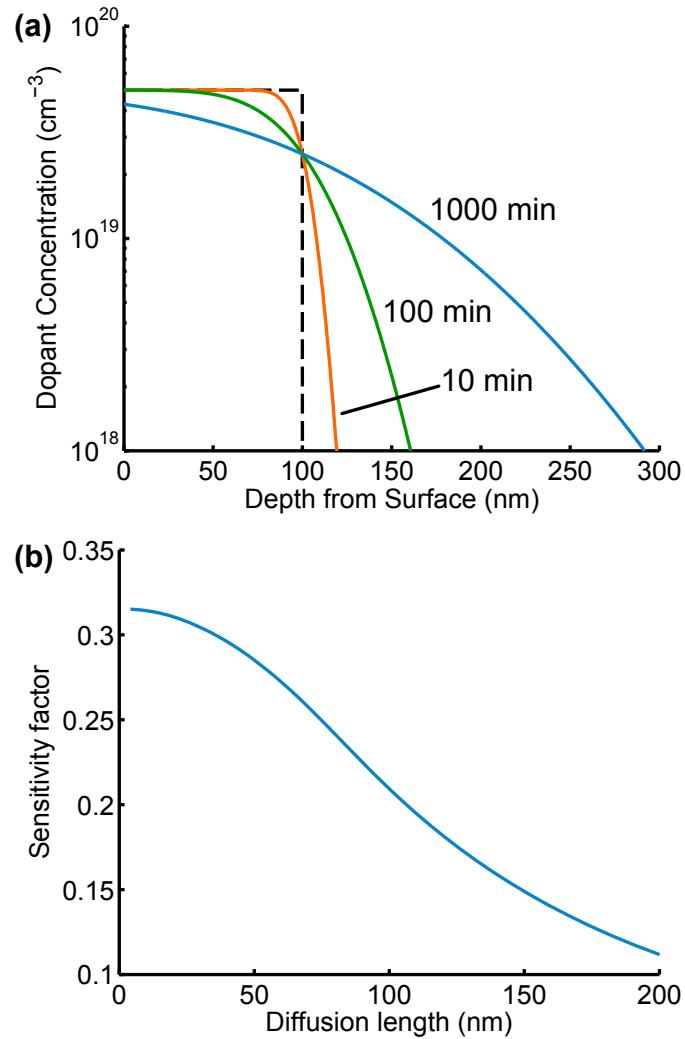


Figure 2.38: Dopant diffusion during epitaxial growth. (a) Concentration profiles for a phosphorus doped 100 nm thick epitaxial resistors grown at 900°C for 10, 100 and 1000 minutes. The dashed line denotes the ideal step concentration profile. (b) The sensitivity factor, β^* , decreases as the diffusion length increases from its maximum possible value of β_0^* if no diffusion occurred.

device before shifting to more complicated numerical models.

2.5.5 Polysilicon

The preceding discussion of diffusion, ion implantation and epitaxy implicitly focused on single crystal silicon. However, piezoresistors are also often fabricated from polycrystalline silicon films, usually referred to as polysilicon or poly. Polysilicon is deposited using the same LPCVD processes as epitaxial silicon, but at lower temperatures or higher rates so that there is insufficient time or thermal energy to eliminate grain boundaries.

Polysilicon can be doped using diffusion, ion implantation, or in-situ doping by introducing gas-phase dopants with the precursor polysilicon gases during the CVD process. However, the introduction of dopant gases affects the polysilicon deposition rate, thickness uniformity, grain size and film stress [211]. Depending upon the polysilicon deposition temperature, it can be amorphous, microcrystalline or polycrystalline. After polysilicon deposition the wafers are often annealed at a high temperature (800-1100°C) in order to relieve stress and alter the grain structure. While the properties of single crystal silicon piezoresistors are well known and reproducible, polysilicon piezoresistor properties vary significantly with the deposition, anneal and doping conditions, because they determine the orientation distribution of the grains.

Piezoresistive effects in polysilicon were studied extensively in the 1970s and 1980s. Onuma and Sekiya [266] first reported the piezoresistive properties of polysilicon in 1972. Two years later, Gurtler and Zwernemann filed a patent for polysilicon piezoresistive pressure sensors [267]. Seto further reported on the electrical and piezoresistive properties of polysilicon in 1975 and 1976 [268,269]. In 1981, Lu and Mandurah separately reported on models for electrical conduction in polysilicon [270–273], and Mikoshiba reported on the stress sensitivity of CMOS devices with polysilicon gates [274]. Germer [275] and Erskine [276] reported on piezoresistive devices utilizing polysilicon piezoresistors in 1983. Finally, in 1984 [277] and 1989 [278], French and Evans presented theoretical models for piezoresistance in polysilicon as a function of doping, grain size, and orientation and proposed an optimum set of processing parameters for a given grain size.

We primarily focus on the design of single crystal silicon piezoresistors in this thesis. Mainly because they are more widely used in both research and commercial production, but also because all of the models and methods can be applied to polysilicon piezoresistors

with slight modifications to the noise and sensitivity formulae.

2.5.6 Electrical modeling

In the last section we calculated the dopant concentration profile, $N(z)$. In this section, we will use the dopant concentration profile to calculate the carrier concentration and resistivity profiles, $n(z)$ and $\rho(z)$. Minority carriers do not affect the operation of piezoresistive sensors, and we will ignore them for the most part.

Both intrinsic and extrinsic carriers can potentially contribute to the carrier concentration. The intrinsic carrier concentration is exceedingly small near room temperature ($\approx 10^{10} \text{ cm}^{-3}$) but contributes to the concentration of carriers at elevated temperatures (e.g. $n_i = 10^{17} \text{ cm}^{-3}$ at $T = 900 \text{ K}$).

The intrinsic carrier concentration is calculated from the empirical fit of Trupke et al. as [279]

$$n_i = 2.9135 \times 10^{15} T^{1.6} \exp\left(\frac{-E_g}{2k_b T}\right) \quad (2.126)$$

where E_g is the bandgap energy, which also varies with temperature and is calculated following Thurmond as [280]

$$E_g = E_g^0 - \frac{\alpha T^2}{T + \beta} \quad (2.127)$$

where $E_g^0 = 1.170 \text{ eV}$, $\alpha = 4.730 \times 10^{-4} \text{ eV/K}$, and $\beta = 636 \text{ K}$. We neglect bandgap narrowing for simplicity.

Assuming complete dopant ionization, the majority carrier concentration is [112]

$$n = \left| \frac{N_D - N_A}{2} \right| + \sqrt{\left(\frac{N_D - N_A}{2} \right)^2 + n_i^2}. \quad (2.128)$$

This expression is only valid for non-degenerately doped silicon, because it is based upon the mass action law ($np = n_i^2$)

We calculate the resistivity from the dopant concentration as

Table 2.10: Temperature dependent carrier mobility modeling parameters. Reproduced from Ref. [281], ©2002 IEEE.

Parameters	Phosphorus	Arsenic	Boron
μ_{\max} (cm ² /V-sec)	1441	1441	470.5
c (-)	0.07	0.07	0.0
γ (-)	2.45	2.45	2.16
μ_{0d} (cm ² /V-sec)	$62.2T_n^{-0.7}$	$55T_n^{-0.6}$	$90.0T_n^{-1.3}$
μ_{0a} (cm ² /V-sec)	$132.0T_n^{-1.3}$	$132.0T_n^{-1.3}$	$44.0T_n^{-0.7}$
μ_{1d} (cm ² /V-sec)	$48.6T_n^{-0.7}$	$42.4T_n^{-0.5}$	$28.2T_n^{-2.0}$
μ_{1a} (cm ² /V-sec)	$73.5T_n^{-1.25}$	$73.5T_n^{-1.25}$	$28.2T_n^{-0.8}$
C_{r1} (cm ⁻³)	$8.50 \times 10^{16}T_n^{3.65}$	$8.90 \times 10^{16}T_n^{3.65}$	$1.30 \times 10^{18}T_n^{2.2}$
C_{r2} (cm ⁻³)	$1.22 \times 10^{17}T_n^{2.65}$	$1.22 \times 10^{17}T_n^{2.65}$	$2.45 \times 10^{17}T_n^{3.1}$
C_{s1} (cm ⁻³)	4.00×10^{20}	2.9×10^{20}	$1.1 \times 10^{18}T_n^{6.2}$
C_{s2} (cm ⁻³)	7.00×10^{20}	7.00×10^{20}	6.10×10^{20}
α_1 (-)	0.68	0.68	0.77
α_2 (-)	0.72	0.72	0.719

$$\rho = \frac{1}{q\mu n} \quad (2.129)$$

where q is the single electron charge and μ is the carrier mobility. Both the carrier density and mobility vary through the depth of the piezoresistor. We calculate the mobility as a function of the donor and acceptor concentration and temperature following Reggiani et al. [281] as

$$\mu = \mu_0 + \frac{\mu_L - \mu_0}{1 + \left(\frac{N_D}{C_{r1}}\right)^{\alpha_1} + \left(\frac{N_A}{C_{r2}}\right)^{\alpha_2}} - \frac{\mu_1}{1 + \left(\frac{N_D}{C_{s1}} + \frac{N_A}{C_{s2}}\right)^{-2}} 0 \quad (2.130)$$

and

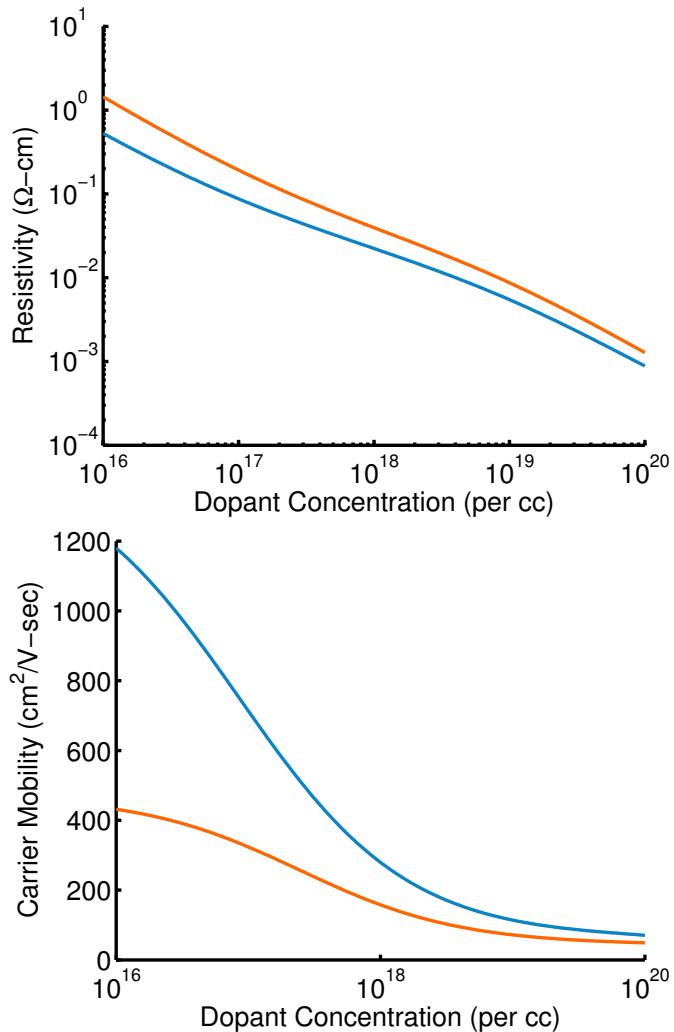


Figure 2.39: Resistivity and mobility as functions of dopant concentration for boron (orange) and phosphorus (blue) doped single crystal silicon.

$$\mu_L = \mu_{\max} T_n^{-\gamma + cT_n} \quad (2.131)$$

$$\mu_0 = \frac{\mu_{0d}N_D + \mu_{0a}N_A}{N_D + N_A} \quad (2.132)$$

$$\mu_1 = \frac{\mu_{1d}N_D + \mu_{1a}N_A}{N_D + N_A} \quad (2.133)$$

where μ_L is the lattice mobility due to phonon scattering while μ_0 and μ_1 account for

impurity scattering. The model parameters are presented in Table 2.10. The temperature dependencies in Table 2.10 and (2.131) are provided in terms of a normalized temperature, $T_n = T/300\text{K}$.

The resistivity and carrier mobility for phosphorus and boron doped silicon at 300 K are plotted in Figure 2.39 as functions of dopant concentration. The mobility decreases with dopant concentration due to ionized impurity scattering.

2.5.7 Experimental validation techniques

There are four common experimental techniques for measuring the dopant concentration profile: secondary ion mass spectrometry, spreading resistance analysis, capacitance voltage profiling, and anodic oxidation. [282–285].

Secondary ion mass spectrometry (SIMS) is used to quantify the total concentration of a dopants, including both substitutional and interstitial positions. By directing a beam of ions at the sample atoms from the sample are sputtered from the surface and separated based upon their charge and mass. The initial data obtained from SIMS are the count rates of several ions as a function of time. In order to quantify the dopant profile, the time must be converted to depth, and the count rate to concentration. The sputtering rate can be measured mechanically using a stylus profilometer or atomic force microscope, or optically using interferometry.

Quantitative concentrations are obtained from the relative count rate of the element of interest compared to that of the major element in the sample, typically silicon [286]. The relative sensitivity factor, used to calculate concentration from the intensity ratio, is calibrated using a reference sample. Ion implantation is typically used to generate the reference sample, because the integrated count rate can be compared to the implanted dose in order to calculate the relative sensitivity factor. SIMS has excellent depth resolution, on the order of several nanometers depending on the energy and angle of the primary ion beam, and resolution down to 10^{15} cm^{-3} . We refer the reader to Ref. [287] for a longer review article and Ref. [288] for a comprehensive book on SIMS.

Spreading resistance analysis (SRA) is used to quantify the electrically active dopant concentration profile. SRA reconstructs the dopant profile based upon the resistance

measured over the depth of the sample using a four-point probe arrangement. In theory, one could simply cleave the sample to be probe and measure the resistance over its cross-section. However, the typical depth of an electronic device ranges from nanometers to microns and is far too small to directly probe. In SRA, the sample is first beveled at a shallow angle, θ . This serves two functions; the resistance at a particular depth can now be measured by moving the sample laterally rather than vertically, and size of the features to be probed is magnified by a factor of $1/\sin(\theta)$. This enables a four-point-probe instrument to measure a resistance profile by stepping laterally across the sample. The resistivity profile is calculated from the resistance profile using numerical methods [289, 290], and the resistivity profile is converted to a concentration profile based upon reference standards. SRA directly probes the resistivity profile of the sample, which is of prime importance for piezoresistive sensors, but accurate measurements within the first 20-50 nm of the sample surface can prove challenging due to edge effects.

Capacitance voltage (C-V) profiling infers the dopant concentration profile by measuring the capacitance as a function of voltage between the sample being probed and an electrode. As carriers near the surface are depleted, the measured capacitance decreases, and C-V analysis is widely used in the semiconductor industry to probe gate oxide defects.

The measurement electrode can stationary or mobile. When the sharp tip of a scanning probe is used as the electrode, scanning capacitance microscopy (SCM) allows for sub-micron lateral resolution in addition to nanometer scale depth resolution [291]. One challenge with C-V and SCM measurements is that the field required to deplete highly doped silicon can lead to electrical breakdown. Although both techniques have been used to probe sub-100nm junctions doped to 10^{20} cm^{-3} , depleting deep, highly doped structures can prove difficult [292]. Additionally, trapped charge or interface states can reduce the measurement accuracy. Scanning Kelvin probe microscopy (SKPM) is a technique related to SCM that infers the surface concentration from the work function difference between a scanned tip and the sample [293].

The final approach, anodic oxidation, encompasses a variety of techniques that use room temperature anodic oxidation for dopant profiling [282]. In the original technique, developed in the 1960s, the silicon wafer is placed in an anodizing solution and electrically biased to induce oxidation. The thickness of the oxide (and the amount of silicon removed)

is controlled by applying a constant bias (e.g. 100 V) and terminating the oxidation step once the current reached a predetermined final value. Afterwards, the oxide is removed in HF and the sheet resistance is measured using a conventional four point probe system. By iterating between the oxidation and measurement steps, the resistivity profile can be assembled. The thickness of silicon removed is controlled by the bias potential and can be quite small, e.g. 194 Å [228]. Although the process is relatively slow, it has been automated [294]. A more recent approach to anodic oxidation is to measure the electrical potential while controlling the current bias, allowing the dopant profile to be measured continuously [295]. However, this latter technique only works for p-type doping.

2.6 Temperature effects

Temperature fluctuations alter the carrier density and mobility in silicon piezoresistors, leading to variation in electrical resistivity and piezoresistive coefficients [107, 281]. Consequently, doped silicon can be used for accurate temperature sensing as in resistance temperature detectors (RTDs). A typical commercial piezoresistive pressure sensor has a temperature induced resistance change equal to ten times the full-scale stressed resistance change over a temperature range of 55°C.

Kurtz [296] presented data and discussed the trends of the piezoresistive coefficient (π), temperature coefficient of piezoresistive sensitivity (TCS), resistivity (ρ), temperature coefficient of resistance (TCR) and strain nonlinearity as functions of dopant concentration. Tufte and Stelzer [97] subsequently presented detailed measurements of these parameters for diffused layers over a wide range of dopant concentrations (10^{18} - 10^{21} atoms cm $^{-3}$) and temperatures (-90°C to 100°C). They also showed that the piezoresistive coefficient was relatively insensitive to the diffusion depth for a diffused layer. Kerr and Milnes [105], showed that the surface dopant concentration could be used as an adequate proxy for the average effective concentration in modeling the piezoresistivity of piezoresistors formed through diffusion/predeposition.

Kurtz was the first to clearly highlight the advantages of using higher doping levels for piezoresistors. The temperature coefficient of sensitivity decreases with increasing surface concentration. This trend is desirable, except that increasing surface concentration

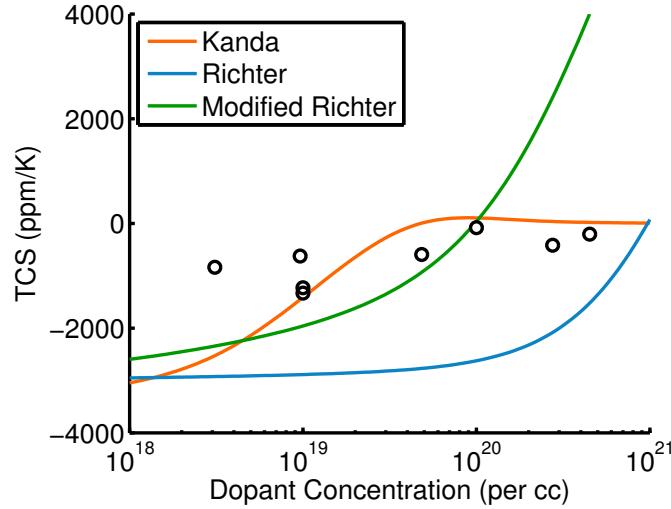


Figure 2.40: Temperature coefficient of sensitivity (TCS) calculated at 300 K using the Kanda, Richter, and modified Richter models for the piezoresistance factor. Experimental data, tabulated from Refs. [98, 297, 298], indicates that the TCS magnitude decreases from more than 1000 ppm/K to less than 30 ppm/K as the dopant concentration is increased from 10^{18} to 10^{20} cm^{-3} . Of the three piezoresistance factor models, Kanda’s model predicts the TCS variation with dopant concentration most accurately at 300 K.

sacrifices the sensitivity of the piezoresistors. However, the temperature coefficient of sensitivity rolls off faster than sensitivity with concentration. Other parameters that decrease with dopant concentration are strain and temperature sensitivity nonlinearities and the temperature coefficient of resistance. Some piezoresistive pressure sensor manufacturers, such as Kulite Semiconductor Products, Merit Sensors, and GE NovaSensor manufacture high dose piezoresistors, taking advantage of this reduced temperature sensitivity. Ultimately, some temperature sensitivity is inevitable in silicon strain sensors and must be compensated using a Wheatstone bridge with two or four active elements and additional conditioning circuitry.

2.6.1 Piezoresistor sensitivity

The models we have presented for calculating the piezoresistance factor and electrical resistivity are temperature dependent, and can be used to calculate the temperature-dependent properties of piezoresistors. The temperature coefficient of sensitivity (TCS) can

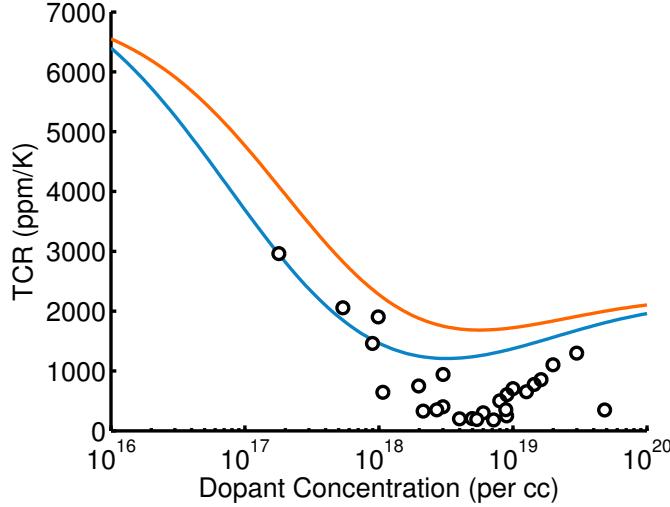


Figure 2.41: Temperature coefficient of resistance (TCR) at 300 K as a function of dopant concentration. Experimental data from p-type silicon resistors are plotted for comparison [299, 300]. The TCR reaches a minimum near $5 \times 10^{18} \text{ cm}^{-3}$ before increasing again slightly. Experimentally measured TCR data is substantially lower than predicted by theory, and research is ongoing [299].

be calculated from the temperature dependent piezoresistance factor models as

$$\text{TCS}(N, T) = \frac{1}{P(N, T)} \frac{\partial P(N, T)}{\partial T}. \quad (2.134)$$

The TCS at 300 K is calculated from the Kanda, Richter, and modified Richter models for the piezoresistance factor (Figure 2.40). Experimental data from Refs. [98, 297, 298] is plotted for comparison, and they agree relatively well with Kanda's model. The magnitude of the TCS decreases from 1000 ppm/K below 10^{18} cm^{-3} to less than 30 ppm/K above 10^{20} cm^{-3} , making temperature compensation substantially easier or even unnecessary for highly doped piezoresistors. The Kanda model is the most accurate option for modeling the TCS of piezoresistive sensors as a function of dopant concentration.

2.6.2 Temperature coefficient of resistance

The temperature coefficient of resistance (TCR) can be calculated from (2.129) as

$$\alpha = \frac{1}{R} \frac{\partial R}{\partial T} \quad (2.135)$$

$$\approx \frac{1}{\rho} \frac{\partial \rho}{\partial T} \quad (2.136)$$

where α has units of $1/K$ and is commonly expressed as ppm/K . The TCR decreases sharply as the dopant concentration is increased until it reaches a minimum near $3 \times 10^{18} \text{ cm}^{-3}$ and begins increasing again (Figure 2.41).

An additional complication with TCR modeling is that temperature change is transduced into a resistance change both electrically and mechanically. Thermal expansion mismatches within the piezoresistive device and the surface it is mounted to transduce temperature changes into resistance changes. Silicon has a particularly small coefficient of thermal expansion (CTE) of about 3 ppm/K compared with most metals (e.g. 23 ppm/K for aluminum) at room temperature. The use of a low modulus die attach material, such as room temperature vulcanization silicone, can reduce the transmission of shear strain between the sensor and its surroundings.

2.6.3 Piezoresistor self-heating

Self-heating is an inherent byproduct of piezoresistor operation. There are three notable design issues with Joule heating in piezoresistors. First, performance degrades as the piezoresistor temperature increases due to a smaller piezoresistance factor and larger Johnson, amplifier and thermomechanical noise. Second, the maximum sensor temperature is constrained in many applications (e.g. force probes for studying hair cells). Third, in applications where the temperature is not constrained, the maximum power dissipation might be constrained. Examples include operation at cryogenic temperatures and battery powered devices.

In this section we will discuss heat transfer models for piezoresistive sensors. Using these models, we will be able to calculate the temperature distribution of a sensor during operation and eventually optimize designs with respect to temperature and power dissipation constraints. First, we will provide a brief example of heat transfer in a

piezoresistive sensor. Next, we will describe the thermal conductivity of common sensor materials, including variation in the thermal conductivity of silicon with film thickness and dopant concentration. Finally, we will develop models to analytically and numerically calculate the temperature distribution within a sensor and compare the results with experimental results.

A qualitative example

A balanced Wheatstone bridge with bridge bias of V_{bridge} dissipates electrical power (W) equal to $V_{\text{bridge}}^2/4R$ in each resistor. The heat is conducted away from the piezoresistor via conduction through the sensor structure and convection into the surrounding air or liquid. Electromagnetic radiation is typically small, on the order of a few percent even at temperatures as high as 600°C [301, 302].

Joule heating induces a temperature rise in the sensor structure. The magnitude of the self-heating effect depends on the sensor geometry, the thermal conductivity of the materials used, and the thermal properties of any surrounding fluid. Figure 2.42a illustrates a cross-section through a pressure sensor and the approximate steady-state temperature distribution. Power is dissipated in the two piezoresistors located at the edges of the diaphragm. The heat is conducted away from the piezoresistors via conduction to the rim of the sensor and via convection directly to the fluid on either side of the diaphragm (Figure 2.42b).

The temperature reaches a peak near the center of each piezoresistor and decreases toward the center of the diaphragm and the edge of the device. The temperature decay towards the center is due to convection; if the device were operated in vacuum then the temperature would remain approximately constant across the diaphragm. The sharp reduction in the slope of the temperature at the edge of the diaphragm is due to the sudden increase in cross-sectional area available for heat conduction.

The most direct way to decrease self-heating is to simply reduce the power dissipation. However, we'll see in the next section (Section 2.6.4) that decreasing the power dissipation leads to worse sensor resolution. Alternatively, the sensor design can be modified for increased heat transfer efficiency. From a heat transfer standpoint we want to use thick films and high thermal conductivity materials in the device, but from a mechanics standpoint

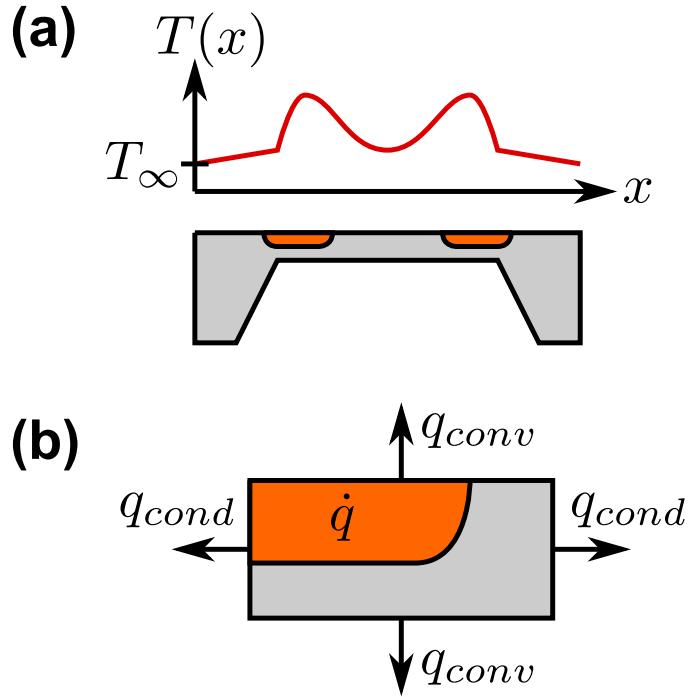


Figure 2.42: Self-heating during piezoresistor operation. (a) Piezoresistors embedded in a diaphragm dissipate power, which is transferred away to the edge of the device and into its surroundings. Temperature variation with position is qualitatively plotted. The temperature rise is substantially higher in the diaphragm than the rest of the device due to its reduced thickness. (b) The heat generated in the piezoresistor (\dot{q}) is conducted away via conduction through the device and convection directly into the surrounding fluid (q_{cond} and q_{conv} , respectively).

we want to use thin films without any excess layers for thermal management. The tradeoff between power dissipation, the resulting temperature increase, and piezoresistor performance is relatively complex and we will develop the building blocks in this section.

Thermal conductivity of typical materials

Thermal conductivity (k) varies widely between the materials commonly used in piezoresistive sensors (Table 2.11). Note that we use k and k_c to denote thermal conductivity and the spring constant of a cantilever beam, respectively. Metals, with the exception of Ti, conduct heat more than two orders of magnitude better than dielectrics. Silicon is a fairly good conductor of heat, although polysilicon is much less so. Polysilicon is a much less

Table 2.11: Thermal conductivities of materials commonly used in piezoresistive silicon sensors. Data is accurate for lightly doped bulk samples near room temperature. Thermal conductivity is reduced by phonon scattering at film boundaries, grain boundaries, and with dopant atoms.

Material	k (W/m-K)
Al	200
Cu	400
Ti	22
Si	142-148
Si (poly)	10-50
SiO ₂	1.4
Si ₃ N ₄	30
Parylene	0.08

efficient conductor of heat due to phonon scattering at grain boundaries, and illustrates an important caveat for the data in Table 2.11; thermal conductivity varies significantly with film morphology and processing history.

In general, thermal conductivity varies significantly with film morphology and thickness due to phonon scattering at film and grain boundaries. It also varies substantially in doped silicon due to phonon scattering with ionized dopants. Finally, it varies with temperature for all materials, and generally reaches a maximum at fairly low temperatures (30-100 K) due to insufficient thermal energy and increasing phonon-phonon scattering at low and high temperatures, respectively. Despite these limitations, we will use the data in Table 2.11 for all of the materials except for silicon, which we will model in more detail. The thermal conductivity of Al is taken from Ref. [303] while the other film properties are taken from the reference literature [304]. We refer the reader to the literature for more detailed models of particular materials and to Kittel for a more general introduction to solid state physics [80].

The thermal conductivity of silicon has been studied extensively for its important to the IC industry. While the properties of polycrystalline materials vary enormously with processing history, the reproducibility of single crystal silicon makes it amenable to study. The thermal conductivity of single crystal silicon depends on three factors: film thickness, temperature, and dopant concentration.

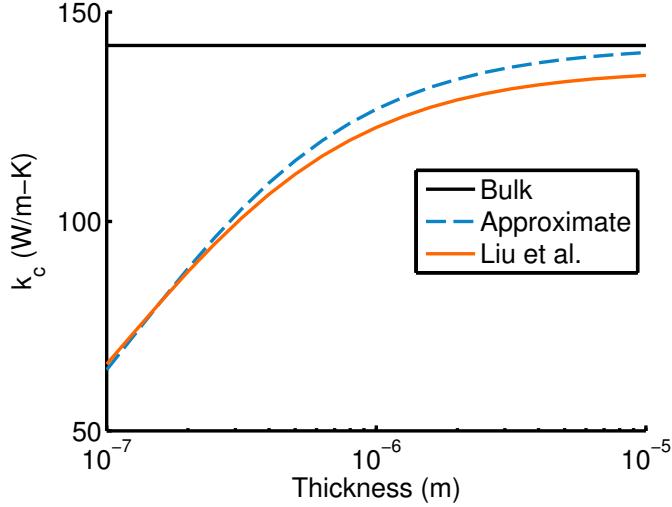


Figure 2.43: Silicon thermal conductivity as a function of film thickness. Data was calculated for lightly doped (10^{15} cm^{-3}) single crystal silicon using our approximate fit to experimental data (2.137) and the more detailed model of Liu et al. (Ref. [306]). The approximate model is accurate for devices with a lightly doped background layer operating near room temperature.

We will use two models for the thermal conductivity of silicon. The first is derived from the data that Asheghi et al. measured for lightly doped SOI device layers in 1997 [305]. They measured thermal conductivity as a function of film thickness, holding the temperature and dopant concentration constant. We performed a simple first-order fit to their room temperature (300 K) data, yielding

$$k_{\text{Si}} = k_{\text{bulk}} \frac{t_c}{t_c + t_0} \quad (2.137)$$

where $t_0 = 120 \text{ nm}$. For most devices the film thickness is the smallest dimension and is the dominant source of phonon-boundary scattering, but the smallest dimension of the structure should always be used in (2.137). Although this simplified model does not account for variation in thermal conductivity with temperature and dopant concentration, thickness is the most important source of thermal conductivity variation for most piezoresistive devices operating near room temperature.

The second model we will use was developed by Liu et al. in 2006, and includes temperature dependent phonon-impurity and phonon-boundary scattering [306]. Impurity

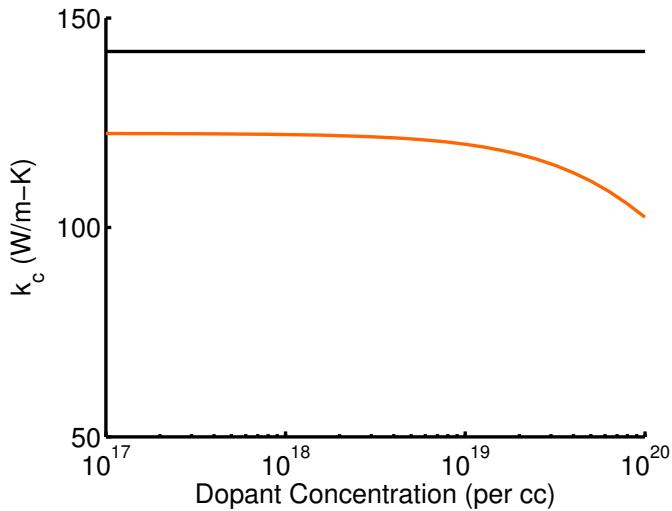


Figure 2.44: Silicon thermal conductivity as a function of dopant concentration. Data was calculated for a $1\ \mu\text{m}$ thick single crystal silicon film at 300 K using the model in Ref. [306]. Thermal conductivity does not drop noticeably until the dopant concentration reaches $10^{19}\ \text{cm}^{-3}$. In most devices the piezoresistor is relatively thin and the background dopant concentration is low, so heat transfer is relatively insensitive to the piezoresistor dopant concentration.

scattering is calculated from the dopant concentration profile, calculated using the models in Section 2.5. The model in Ref. [306] is fairly long and would introduce an excessive amount of new terminology, so we will simply present the model results here and refer the reader to the original paper and our Matlab implementation (Appendix H).

Figure 2.43 compares the bulk, approximate (2.137), and exact (Ref. [306]) thermal conductivity models for a lightly doped ($10^{15}\ \text{cm}^{-3}$) layer of single crystal silicon. Thermal conductivity decreases noticeably for thin films: 100 nm and $1\ \mu\text{m}$ thick films have 53% and 14% lower thermal conductivities than bulk samples. The approximate and exact models agree to within 5% over the entire thickness range. (2.137) is fairly accurate when devices are operated near room temperature, the piezoresistor is relatively thin, and the background dopant concentration is low.

Variation in thermal conductivity with dopant concentration is presented in Figure 2.44. Data was calculated for a $1\ \mu\text{m}$ thick film at 300 K using the detailed thermal conductivity model. Thermal conductivity does not drop noticeably until the dopant

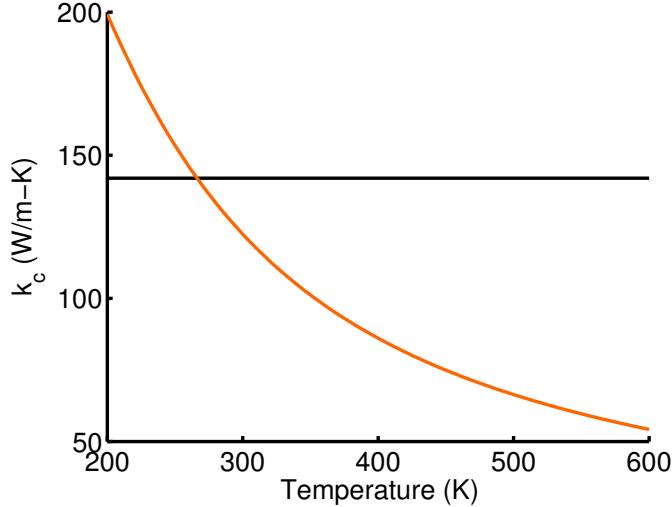


Figure 2.45: Silicon thermal conductivity as a function of temperature. Data was calculated for a 1 μm thick, lightly doped (10^{15} cm^{-3}) single crystal silicon film using the model in Ref. [306].

concentration reaches 10^{19} cm^{-3} . Considering that the piezoresistor is relatively thin in most piezoresistive devices (< 30% of the device thickness) and that the background dopant concentration far lower than 10^{19} cm^{-3} , heat transfer is relatively insensitive to the piezoresistor dopant concentration. Accordingly, (2.137) is accurate for nearly all sensors operating near room temperature.

Thermal conductivity is plotted as a function of temperature in Figure 2.45 for a 1 μm thick, lightly doped (10^{15} cm^{-3}) film using the detailed thermal conductivity model. Temperature is a strong influence on thermal conductivity. If the temperature jumps from 300 to 400 K, either by self-heating or a high ambient temperature, the thermal conductivity of the silicon drops by 30%. The decrease in thermal conductivity with temperature acts as a source of positive-feedback, exacerbating self-heating effects, particularly if the piezoresistor is biased with a current source due to silicon's positive TCR (Section 2.6.2).

To summarize, the thermal conductivity of silicon decreases when film thickness is reduced or the temperature is increased; dopant concentration plays a relatively minor role. Devices that operate from 275 - 325 K can use (2.137) to calculate the thermal conductivity with < 10% error. For larger temperature deviations the more precise model of Ref. [306] should be employed.

Modeling heat transfer

Temperature gradients develop in order to balance heat generation, \dot{q} , with conductive and convective heat transfer. Conductive heat flux, $q_{\text{cond}} = -k\nabla T$, is proportional to the material's thermal conductivity and the temperature gradient through the material. Convective heat flux, $q_{\text{conv}} = h_{\text{eff}}(T - T_{\infty})$, is proportional to the effective convection coefficient (h_{eff}) and the temperature difference between the device and surrounding fluid (T_{∞}).

The temperature profile, $T(x)$, can be calculated from the steady-state heat equation, $\dot{q} = q_{\text{conv}} + \nabla q_{\text{cond}}$ [307]. For a one-dimensional problem, the result reduces to

$$\dot{q}'(x) = h_{\text{eff}}(x)P(x)[T(x) - T_{\infty}] - \frac{\partial}{\partial x} \left[A(x)k(x)\frac{\partial T(x)}{\partial x} \right]$$

where \dot{q}' is the heat generation per unit length and P is the area per unit length available for convection (i.e. the device perimeter). If the cross-sectional area, perimeter, convective coefficient and thermal conductivity do not change appreciably with position then the result simplifies to

$$\dot{q}'(x) = h_{\text{eff}}P[T(x) - T_{\infty}] - Ak\frac{\partial^2 T(x)}{\partial x^2}. \quad (2.138)$$

We will solve for the temperature profile in one-dimension using the finite-difference method and in higher dimensions using finite element analysis.

The power dissipation per unit length is $\dot{q}'(x) = 4I^2R_s(x)/w_c$, where I is the current flow, $R_s(x)$ is the position-dependent sheet resistance. The sheet resistance as a function of temperature can be experimentally measured or calculated as in Sections 2.5.6 and 2.6.2. When the temperature increase is small the sheet resistance remains approximately constant along the piezoresistor length and the power dissipation per unit length simplifies to W/l_{pr} .

Convective heat transfer

Although we refer to heat transfer through the ambient fluid (e.g. air or water) as convection, buoyancy driven convection is negligible due to the small Rayleigh number

of most microscale devices. The Rayleigh number for free convection at a heated wall can be calculated from

$$\text{Ra} = \frac{g\beta}{\nu\alpha}(T_s - T_\infty)x^3$$

where g is the acceleration due to gravity, β , ν and α are the thermal expansion coefficient, kinematic viscosity and thermal diffusivity of the fluid, T_s and T_∞ are the temperatures of the heated surface and ambient, and x is the characteristic length of the system. The Rayleigh number is small for most microsystems due to the strong dependence on length; we calculated $\text{Ra} = 0.1$ for the piezoresistive cantilevers that we will discuss shortly. The critical Rayleigh number for free convection, above which buoyancy driven convection is significant, depends on the fluid but is on the order of 1000 typically.

Although free convection does not take place, conduction through the fluid can be modeled using an effective convection coefficient. The convection coefficient can be calculated using one of two methods. The most direct method is to calculate it from the measured thermal conductance as

$$h_{\text{eff}} = \frac{G'_f}{P} \quad (2.139)$$

where G'_f is the thermal conductance from the device to the fluid per unit length and P is the surface area per unit length exposed to the fluid [308]. A second, less direct approach to calculating h_{eff} was presented by Hu et al. in 2008 [309],

$$h_{\text{eff}} = \frac{k_f S}{A} \quad (2.140)$$

where k_f is the thermal conductivity of the surrounding fluid, S is the heat conduction shape factor [307], and A is the total area exposed to the fluid. A major challenge with the second approach is calculating an appropriate shape factor. However, it does provide two important results: h_{eff} increases when the sensor is brought near a solid surface due to an increase in S [310], and, h_{eff} scales roughly linearly with the thermal conductivity of the surrounding fluid.

The Biot number

One-dimensional heat transfer models are sufficient when the temperature gradient within a device cross-section is small. In other words, if thermal conduction within the device is faster than the conduction at its interfaces then the system can be modeled one-dimensionally. The Biot number can be calculated to determine if a one-dimensional model can be used

$$\text{Bi} = \frac{h_{\text{eff}}L_c}{k} \quad (2.141)$$

where L_c is the characteristic length of the system, commonly defined as its volume to surface area ratio. Like the Rayleigh number, the Biot number is exceedingly small for most MEMS sensors, allowing them to be modeled as one-dimensional systems. For example, $\text{Bi} < 10^{-5}$ for typical piezoresistive cantilevers.

Example experimental and modeling results

The heat transfer and thermal conductivity models we have discussed in this section were used to investigate heat transfer in piezoresistive cantilevers [302]. We were interested in developing an accurate heat transfer model in order to enable the design of temperature- rather than power-constrained piezoresistive sensors. Until that point in time, most piezoresistor design optimization work either ignored Joule heating or set a power dissipation budget that was decoupled from the device design.

But most applications are limited by temperature rise rather than power dissipation. The power dissipation is related to the temperature rise by the thermal resistance between the piezoresistor and ambient environment. Assuming that the cantilever cross-section scales as $[L]^2$ and the piezoresistor length as $[L]$, then the conductive thermal resistance from the piezoresistor to the silicon die scales as $[L]^{-1}$. Similarly, the thermal resistance due to heat transfer through the ambient fluid scales with the exposed surface area as $[L]^{-2}$. Clearly, power dissipation needs to scale with cantilever dimensions in order to maintain a constant temperature rise during operation.

Figure 2.46 shows the cantilever geometry, heat transfer model and a fabricated device. The overall length, width and thickness of the cantilever are l_c , w_c and t_c , respectively. The

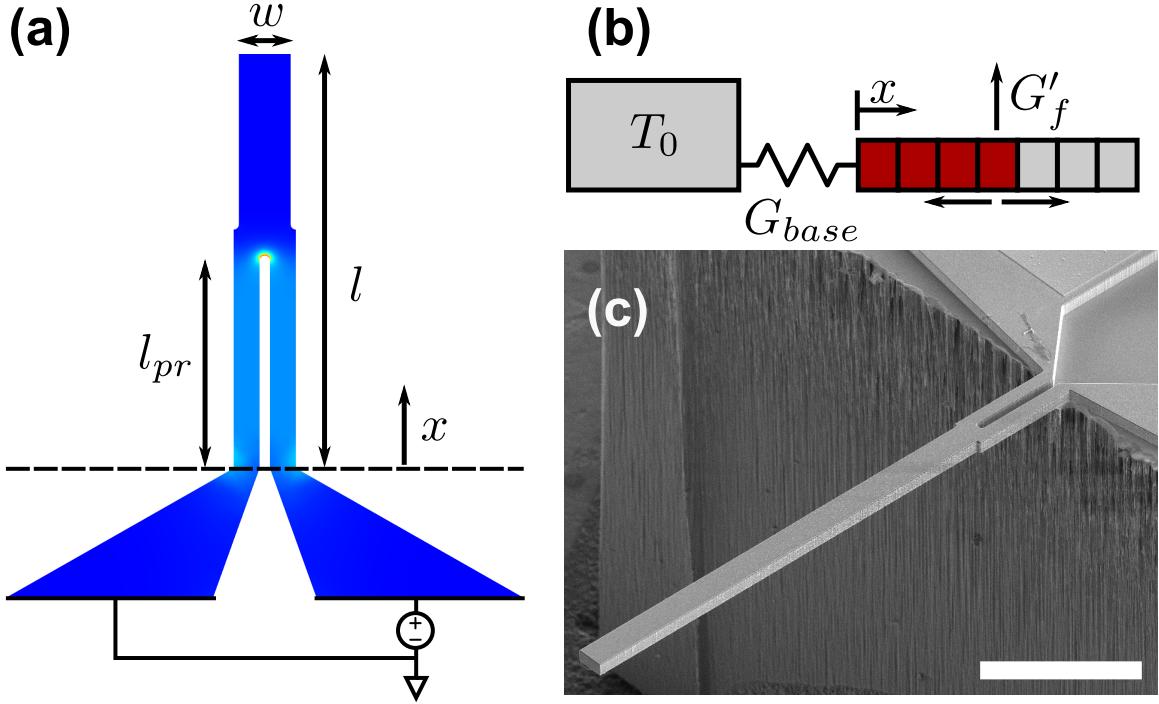


Figure 2.46: (a) The cantilever geometry and simulated power dissipation density, (b) one-dimensional finite differences model used to calculate the temperature profile, and (c) scanning electron micrograph of a typical cantilever used here. Reprinted from [302], ©2011 IEEE.

piezoresistor extends length l_{pr} from the base, encompasses the full cantilever width, and has a narrow slit to electrically isolate the loop. During operation, the resistor is electrically biased and Joule heating power W is distributed along its length.

The cantilevers were fabricated from silicon-on-insulator (SOI) wafers [196]. Briefly, the wafers were phosphorus doped at 800°C via POCl_3 predeposition. Subsequently the piezoresistor and cantilever were patterned and etched, aluminum was sputtered and patterned for electrical contacts, and the cantilever was released using deep reactive ion etching. The dopant concentration profile was measured using spreading resistance analysis and the electrically active surface concentration was $2 \times 10^{20} \text{ cm}^{-3}$.

We focused on two cantilever designs whose dimensions are summarized in Table 2.12. The piezoresistor constitutes 90% and 59% of the total resistance for designs A and B respectively. The piezoresistor self-heating depends primarily on the power dissipated in

Table 2.12: Summary of the cantilever designs used to investigate heat transfer in piezoresistive cantilevers. Calculated and experimentally determined parameters are presented in the bottom half of the table. Experimental parameter uncertainty is presented as $\mu \pm \sigma$.

	Design A	Design B
l_c (μm)	890	124
w_c (μm)	10	10
t_c (μm)	2	0.34
l_{pr} (μm)	212	32
k of Si at 300K (W/m-K)	128.2	98.9
G'_f (mW/m-K)	32.8 ± 3.3	43.4 ± 2.7
h_{eff} (W/m ² -K)	1366 ± 139	2098 ± 131
G_{base} ($\mu\text{W/K}$)	41.9 ± 16.3	17.7 ± 8.4
Λ (μm)	287.1 ± 20.7	104.6 ± 25.0

the released cantilever, so all of the power dissipation levels we will quote refer to the power dissipated in the cantilever itself. The excess resistance was measured using the transfer length method and is primarily due to the silicon interconnect resistance.

We calculated the temperature profile, $T(x)$, from (2.138) assuming an isothermal boundary condition at the silicon die (T_0) and adiabatic conditions at the cantilever tip ($\partial T / \partial x = 0$). The thermal conductance between the cantilever and die, G_{base} , is finite. We calculated $k(x)$ using Ref. [306]. Due to the high temperatures we investigated, we used Ref. [306] to calculate k and calculated $\dot{q}'(x)$ from $4I^2R_s(T)/w_c$ as discussed earlier.

There are two unknown parameters in the model (h_{eff} and G_{base}), so we measured the temperature profile of the cantilever and fit the model to the data. To fit the data, we allowed h_{eff} (or equivalently, G'_f) and G_{base} to vary while using the calculated values of k and \dot{q}' to minimize the sum of the squared relative temperature residuals. There are three common temperature calibration techniques for piezoresistive silicon sensors: Raman thermometry, TCR thermometry and IR thermometry. All three temperature measurement techniques require calibration, whether it is the the Raman shift with respect to temperature, the resistor TCR, or the emissivity of the device.

We used Raman thermometry to measure the temperature at discrete locations along the length of the cantilever as in Refs. [311, 312]. The cantilever was biased to induce

self-heating and the temperature was measured from the wavelength shift of the Stokes peak [313]. The laser wavelength, power, and spot diameter were 488nm, 45 μW and 1 μm , respectively.

We also used a second method to estimate the average piezoresistor temperature: TCR thermometry. For this method, we calibrated the piezoresistor TCR was calibrated using a temperature controlled oven (Thermotron S-1.2) and small bias power ($< 1\mu\text{W}$), allowing the average piezoresistor temperature to be calculated from subsequent current-voltage measurements. Due to the high doping level of our diffusion-formed piezoresistors ($\approx 2 \times 10^{20} \text{ cm}^{-3}$), the TCR was only $1372 \pm 171 \text{ ppm/K}$ ($\mu \pm \sigma$, $N = 4$ devices) and was linear over the range of temperatures investigated. While the uncertainty of TCR-based measurements is usually higher than Raman-based measurements, the former is relatively quick and easy to perform once the calibration is complete and does not require any specialize optics.

A third method, infrared microscopy, is often used to measure the temperature optically. However, silicon is relatively transparent to infrared light and not enough light reflected off of our thin devices (340 and 2000 nm thick) to enable accurate measurements. Raman thermometry uses a shorter wavelength of light (e.g. 488 nm) that interacts more readily with silicon.

Figure 2.47 plots the experimental and predicted temperature profiles for designs A and B operating in air. The temperature reaches a maximum approximately $\frac{2}{3}l_{\text{pr}}$ from the cantilever base and then decreases exponentially beyond it. The exponential decay can be described with a thermal healing length [314]

$$\Lambda = \sqrt{\frac{kA}{h_{\text{eff}}P}}.$$

While h_{eff} is insensitive to power dissipation, G_{base} and Λ vary considerably because they depend on k , which changes with temperature. The measured effective convection coefficients (Table 2.12) are comparable to recent simulation results (1000-2000 $\text{W/m}^2\text{-K}$) [301].

The cantilever tip temperature increases linearly with power dissipation by 2.5 and 36.7 K/mW for designs A and B, respectively. The 15-fold difference in $\partial T_{\text{tip}}/\partial W$ is primarily

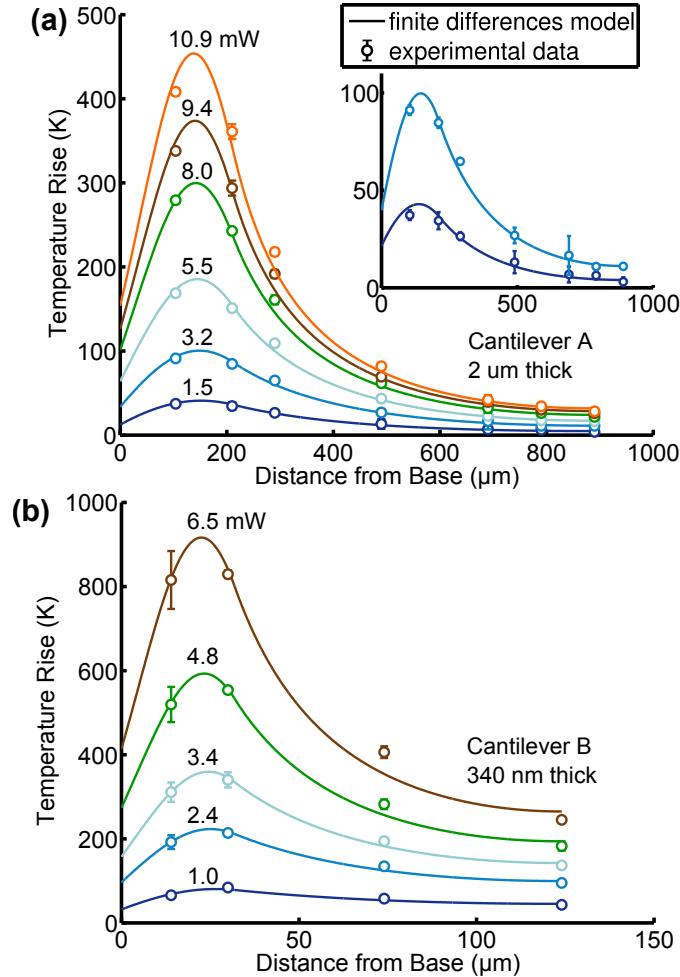


Figure 2.47: (a-b) Comparison between Raman thermometry results and the finite differences model over a range of power dissipations in air. Error bars represent the standard deviation and are smaller than some of the symbols. The average model residual for both cantilever designs is 2%. (inset) Detail of the 1.5 and 3.2 mW data for design A. Reprinted from Ref. [302], ©2011 IEEE.

due to the difference in distance between the piezoresistor and cantilever tip between the designs (2.4Λ vs. 0.9Λ), because the thermal conductance from the piezoresistor to the cantilever base varies less than 20% between the designs. In cantilever designs that are constrained by tip temperature, a shorter piezoresistor will tend to be favored during design optimization so that the distance from the piezoresistor to the tip is as large as possible.

The thermal conductivity of water is approximate 25 times greater than that in air.

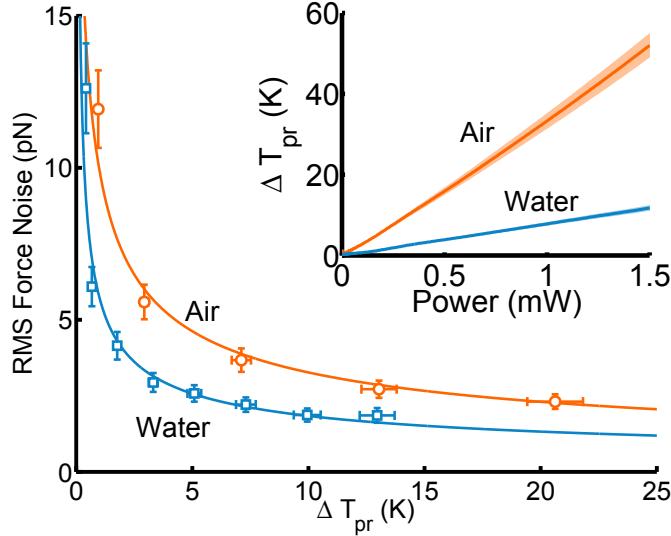


Figure 2.48: Experimental force resolution in a 10 Hz to 10 kHz bandwidth for a piezoresistive cantilever (design B in Table 2.12) in air and water. The force noise in air is 76% greater than in water for the same average piezoresistor temperature increase. (inset) Average piezoresistor temperature as a function of total power dissipation calculated from current-voltage curves in air and water. The piezoresistor TCR was calibrated in a temperature controlled oven in order to infer the piezoresistor temperature during experiments. Error bars represent the standard deviation. Reprinted from Ref. [302], ©IEEE 2011.

Based upon (2.140) we would expect a corresponding increase in h_{eff} in water. We measured the piezoresistor temperature during operation in air and water using TCR thermometry, and the results are presented in Figure 2.48. The h_{eff} we measured in water was approximately 59 kW/m²-K, or, about 28 times larger than the value in air. We will generally assume h_{eff} of 2 and 50 kW/m²-K in air and water, respectively, unless otherwise noted.

Increased heat transfer into the fluid enables greater piezoresistor power dissipation for the same temperature rise in the sensor. As will discuss in the next section, greater power dissipation corresponds to better sensor performance. For example, the force noise that we measured in air was 76% greater than in water for the same average piezoresistor temperature increase. The 28-fold increase in h_{eff} corresponds to a 5-fold decrease in the thermal healing length, so the temperature beyond the piezoresistor decays far more quickly

in water than in air. If the sensor resolution is compared between air and water for the same tip temperature then the difference is even greater; for 500 μW of power dissipation (2.6 pN in air and 2.9 pN in water), the predicted tip temperature change was 20 K in air and only 50 mK in water, a 400-fold reduction. The thermal advantage of water is slightly offset by water's increased density and viscosity which reduce the cantilever resonant frequency. Both the thermal and physical properties of the fluid will be accounted for during cantilever design optimization, and operation in water still yields better force resolution for a given tip temperature increase when fluid damping is accounted for.

Analytical temperature approximations

The results presented in Figure 2.47 suggest that a fairly simple approach could be used to approximate the maximum and tip temperatures in a piezoresistive cantilever. If we can approximate the maximum temperature in the device and calculate an appropriate thermal healing length then the temperature at any other location can be quickly calculated.

For a piezoresistive cantilever, the thermal conductance from the piezoresistor to the cantilever base, neglecting conduction to the surrounding fluid, is

$$G_{\text{pr}} \approx \frac{2w_c t_c k}{l_{\text{pr}}}.$$

If we assume that the majority of the heat is conducted to the base of the cantilever and not directly into the fluid then

$$T_{\text{max}} \approx W \frac{G_{\text{base}} + G_{\text{pr}}}{G_{\text{base}} G_{\text{pr}}}.$$

In an ideally designed and fabricated device, $G_{\text{base}} \rightarrow \infty$ and $T_{\text{max}} \rightarrow W/G_{\text{pr}}$. In reality G_{base} will always be finite, although it can be maximized by ensuring a good thermal contact between the thin sensor region and thicker inactive region of the device.

The temperature at the tip of the device can then be calculated from

$$T_{\text{tip}} \approx T_{\text{max}} e^{-(l_c - 2l_{\text{pr}}/3)/\Lambda}. \quad (2.142)$$

where l_c is the overall length of the cantilever and $2l_{\text{pr}}/3$ is the approximate distance

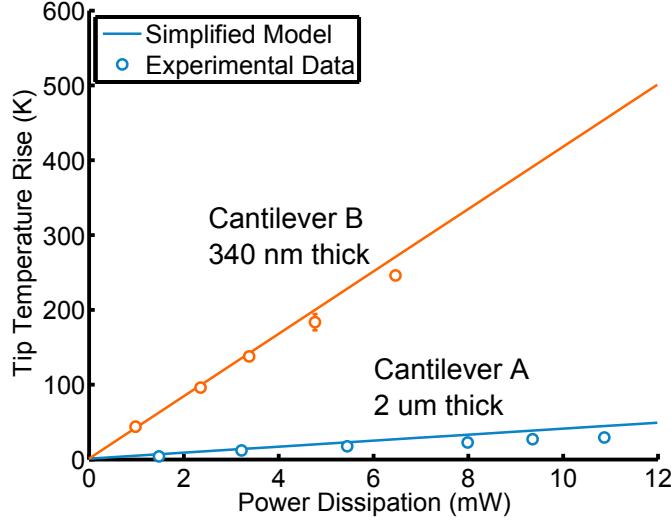


Figure 2.49: Comparison between the measured and calculated tip temperature as a function of power dissipation. The tip temperature was calculated from (2.142) using the experimentally measured parameters in Table 2.12. The residual error between the simplified model and experimental data is < 10%.

from the base of the device where the maximum temperature occurs.

Figure 2.49 compares the approximate tip temperature calculated from (2.142) and the experimental data. The agreement between the simplified model and experimental data is excellent (residual error < 10%) and we will occasionally use this simplified model during design optimization in order to quickly determine an approximate design.

In summary, we have developed and validated thermal models for calculating the temperature increase of piezoresistive sensors due to Joule heating. Self-heating is an important consideration for the design of force probes, particularly those intended for biological force measurements. Thermal modeling is incorporated into the numerical cantilever optimization code that will be discussed shortly.

2.6.4 Power dissipation and measurement resolution

Piezoresistor power dissipation affects device performance in a number of ways. Consider the force resolution of a piezoresistive cantilever (2.98). Ignoring thermomechanical and amplifier noise for the sake of simplicity, and because they do not limit measurement

resolution in most applications, the MDF is

$$\begin{aligned} \text{MDF} &= \frac{\sqrt{V_J^2 + V_H^2}}{S_{\text{FV}}} \\ &= \frac{\sqrt{4k_b T_{\text{pr}} R (f_{\max} - f_{\min}) + \alpha V_{\text{bridge}}^2 \ln(f_{\max}/f_{\min})/2N_{\text{eff}}}}{3V_{\text{bridge}}(l_c - l_{\text{pr}}/2)\pi_L^{\max}\gamma\beta^*/2w_c t_c^2} \end{aligned} \quad (2.143)$$

where we emphasize that T_{pr} is the average piezoresistor temperature. Rewriting the resolution in terms of power dissipation rather than bridge bias by using the fact that $W = V_{\text{bridge}}^2/4R$ yields

$$\text{MDF} = \frac{\sqrt{k_b T_{\text{pr}} (f_{\max} - f_{\min})/W + \alpha \ln(f_{\max}/f_{\min})/2N_{\text{eff}}}}{3(l_c - l_{\text{pr}}/2)\pi_L^{\max}\gamma\beta^*/2w_c t_c^2}. \quad (2.144)$$

The only term in (2.144) that depends on the power dissipation is Johnson noise. Increasing the power dissipation without changing any other design parameters, by increasing the bridge bias for example, will reduce the MDF as $1/\sqrt{W}$. Several earlier design optimization studies have made this observation [75, 202, 315], but is more power dissipation always good?

One important simplification in (2.144) is that self-induced Joule heating is ignored; piezoresistor temperature is independent of power dissipation. But we know that a higher piezoresistor temperature corresponds to increased Johnson noise and decreased sensitivity. We can simplify (2.144) by defining constants:

$$\begin{aligned} c_1 &= k_b(f_{\max} - f_{\min}) \\ c_2 &= \alpha \ln(f_{\max}/f_{\min})/2N_{\text{eff}} \\ c_3 &= 3(l_c - l_{\text{pr}}/2)\pi_L^{\max}\gamma\beta^*/2w_c t_c^2 \end{aligned}$$

to obtain

$$\text{MDF} = \frac{\sqrt{c_1 T_{\text{pr}}/W + c_2}}{c_3}. \quad (2.145)$$

Table 2.13: Typical model parameters for our power dissipation analysis.

Parameter	Units	Description	Typical Value
c_1	-	Hooge noise	10^{-12}
c_2	J/K	Johnson noise	5.5×10^{-18}
c_3	V/N-V	Sensitivity	10^5
c_4	W/K	Thermal conductance	2×10^{-6}
c_5	1/K	TCS	-1.2×10^{-3}
T_0	K	Ambient temperature	300

The temperature coefficient of resistance for the resistor isn't explicitly included in (2.145) here because the problem is formulated in terms of W ; the problem can be formulated in terms of V_{bridge} in order to explicitly include the TCR.

But force resolution depends on piezoresistor temperature. As the piezoresistor temperature increases, the Johnson noise increases while the piezoresistance factor decreases. We can add thermal coupling to the model with these two refinements:

$$T_{\text{pr}} = T_0 + W/c_4$$

$$S_{\text{FV}} = c_3 V_{\text{bridge}} (1 + c_5 T_{\text{pr}})$$

where c_4 is the thermal conductance from the piezoresistor to the ambient, which is fixed at temperature T_0 . For our particular example of a piezoresistive cantilever, considering only thermal conduction along the length of the cantilever, the thermal conduction coefficient is approximately

$$c_4 = 2w_c t_c k / l_{\text{pr}}$$

The final coefficient, c_5 , is the temperature coefficient of sensitivity (TCS) and accounts for the reduction in piezoresistance factor with increasing temperature. We assume a TCS of -1300 ppm/K (Figure 2.40), representative of a piezoresistor doped to about 10^{19} cm^{-3} [298].

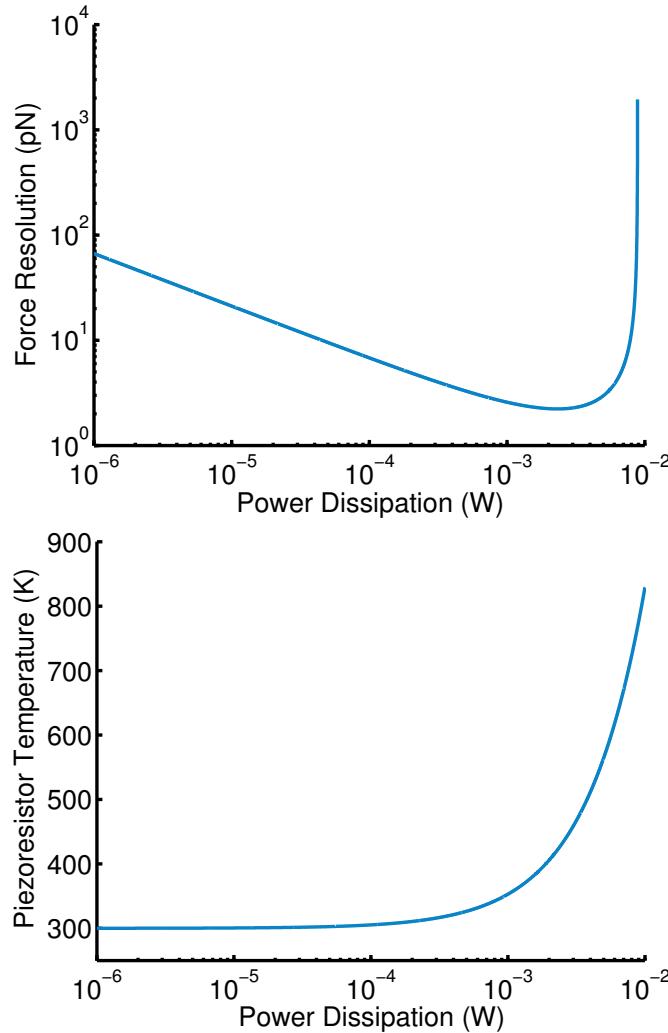


Figure 2.50: Simplified numerical model results for the (a) minimum detectable force and (b) average piezoresistor temperature as a function of power dissipation.

Substituting these new parameters into (2.145) yields

$$\text{MDF} = \frac{\sqrt{c_2 + c_1/c_4 + c_1 T_0/W}}{c_3(1 + c_5 T_0 + c_5 W/c_4)}. \quad (2.146)$$

Numerical evaluation of (2.146) using the parameters in Table 2.13 reveals that resolution does not vary monotonically with power dissipation. Figure 2.50a shows the calculated force resolution from this simple numerical model. For low powers, when T_{pr} is

close to T_0 , the force resolution improves as $1/\sqrt{W}$ and a minimum is reached at 2.58 mW, corresponding to $T_{\text{pr}} = 436\text{K}$ and $F_{\text{min}} = 2\text{ pN}$.

But above a critical power dissipation the temperature begins to take off (Figure 2.50b). The transition begins when $W \approx c_4 T_0 / 2$ and the transition rate with respect to power depends on the magnitudes of $c_4 (\partial T_{\text{pr}} / \partial W)$ and $c_5 (\partial S_{\text{FV}} / \partial T_{\text{pr}})$.

This relatively simple model captures the behavior of more complete piezoresistive sensor performance models that include thermomechanical noise, amplifier noise, and more accurate temperature modeling. Ideally, the power dissipation would be chosen to reach the minimum of (2.146). But in applications where the maximum temperature needs to be limited to avoid damaging the sample, design optimization strives to get as close as possible to the curve minimum without exceeding the temperature constraint.

In summary, there are two regimes for $\partial \text{MDF} / \partial W$; when the cantilever is operated near ambient temperature performance improves with power as $1/\sqrt{W}$, but if the piezoresistor temperature gets too far above the ambient then performance rapidly degrades. The power dissipation at which the performance transitions depends on the ambient temperature. If the ambient temperature is low (e.g. 300 mK in a helium cryostat) then the optimal power dissipation will be far smaller than if the same sensor were operated at room temperature.

This analysis bridges the gap between previous analysis that indicated higher piezoresistor power dissipation was always favorable [75, 202, 315] and the deleterious performance at high power dissipation levels that researchers have observed [316]. Regardless of the cantilever operating regime, increasing the thermal conductance from the piezoresistive sensor (e.g. operating in a higher thermal conductivity fluid) leads to better performance unless power dissipation is tightly constrained.

2.7 Design optimization

Tradeoffs are fundamental to piezoresistor design because changing a single design parameter affects sensor performance in several ways. Consider one of the most important design parameters, the peak dopant concentration of the piezoresistor. Increasing the dopant concentration reduces the noise of the piezoresistor (Section 2.4.1) but also decreases its strain sensitivity (Sections 2.2.2 and 2.4.2). Increasing the dopant concentration also

Table 2.14: Design matrix showing the effect of design parameter changes on sensor performance. As the indicated design parameter increases (while the other parameters are held constant), the observed parameters increase (\uparrow), decrease (\downarrow), or do not change substantially (-). The observed parameters are the displacement sensitivity (S_{XV}), force sensitivity (S_{FV}), 1/f noise (V_H), Johnson noise (V_J), power dissipation (W), maximum temperature increase (ΔT_{\max}), resonant frequency (f_0) and stiffness (k). Design parameters common to all fabrication methods are described first, followed by predeposition-, ion implantation-, and epitaxy-specific parameters. l , w , and t are the overall sensor length, width and thickness while l_{pr} , w_{pr} and t_{pr} are the piezoresistor length, width and thickness. V_{bridge} is the bridge bias voltage. n and t_{pr} are the dopant concentration and piezoresistor thickness of the epitaxial piezoresistor. T_{diff} and t_{diff} are the predeposition temperature and time. n_{implant} and E_{implant} are the ion implantation dose and energy, while \sqrt{Dt} is the diffusion length of the post-implantation anneal. Adapted and extended from [24].

	S_{XV} (V/m)	S_{FV} (V/N)	V_H (V/ $\sqrt{\text{Hz}}$)	V_J (V/ $\sqrt{\text{Hz}}$)	W (W)	ΔT_{\max} (K)	f_0 (Hz)	k (N/m)
l_c (m)	\downarrow	\uparrow	-	-	-	-	\downarrow	\downarrow
w_c (m)	-	\downarrow	-	-	-	\downarrow	-	\uparrow
t_c (m)	\uparrow	\downarrow	-	-	-	\downarrow	\uparrow	\uparrow
l_{pr} (m)	\uparrow	\uparrow	\downarrow	\uparrow	\downarrow	-	-	-
w_{pr} (m)	-	-	\downarrow	\downarrow	\uparrow	\uparrow	-	-
V_{bridge} (V)	\uparrow	\uparrow	\uparrow	-	\uparrow	\uparrow	-	-
N_{epi} (cm $^{-3}$)	\downarrow	\downarrow	\downarrow	\downarrow	\uparrow	\uparrow	-	-
t_{pr} (m)	\downarrow	\downarrow	\downarrow	\downarrow	\uparrow	\uparrow	-	-
T_{diff} ($^{\circ}\text{C}$)	\downarrow	\downarrow	\downarrow	\downarrow	\uparrow	\uparrow	-	-
t_{diff} (min)	\downarrow	\downarrow	\downarrow	\downarrow	\uparrow	\uparrow	-	-
N_{implant} (cm $^{-2}$)	\downarrow	\downarrow	\downarrow	\downarrow	\uparrow	\uparrow	-	-
E_{implant} (keV)	\downarrow	\downarrow	-	-	-	-	-	-
\sqrt{Dt} (m)	\downarrow	\downarrow	\downarrow	\downarrow	\uparrow	\uparrow	-	-

has secondary effects that don't directly impact resolution but are important for most applications. Increased dopant concentrations yield reduced temperature coefficients of sensitivity and resistance (Sections 2.6.1 and 2.6.2) but increased power dissipation and self-heating (Section 2.6.3). Table 2.14 summarizes the effect of individual parameter changes on several of the most important measures of sensor performance.

The tight coupling between the design parameters in Table 2.14 makes piezoresistor design non-trivial. Design is further complicated because piezoresistors are simultaneously

mechanical, thermal and electronic devices, so multiple interconnected models are required to accurately predict device performance. A final challenge is that designs typically need to be constrained in some way, either in the maximum spring constant for a contact mode scanning probe or the maximum temperature rise of a catheter pressure sensor. In this section we'll discuss the methods that can be used to optimize piezoresistive sensors; in the following chapter we will apply them to force probe design.

Design optimization refers to selecting the best available design from a set of alternatives given a complex set of fabrication and operating constraints. There are three important concepts in design optimization: optimality, parameters, and constraints.

Optimality can be defined in a variety of ways depending on the application. A force sensor for scientific research might be optimized to minimize the force noise irrespective of cost or other constraints, while a commercial pressure sensor might aim to minimize the die area while meeting minimum performance requirements [317]. The first step in design optimization then is to define a goal that can be minimized or maximized. The goal should be a smooth and continuous function that can be computed from the design parameters that define the sensor.

Parameters define the size and shape of the piezoresistive sensor. Examples include the overall sensor dimensions, piezoresistor dimensions, ion implantation dose and bridge bias voltage. The size of the design parameter set should be minimized while completely and uniquely defining the device design; in other words, the design parameter set should span the device design space. As the size of the design parameter set increases, the complexity and computational cost of design optimization increases. Accordingly the size of the design parameter set should be minimized; 5 to 20 parameters is typical. As in the case of the optimization goal function, the design parameters should be smooth and continuous to simplify the optimization process. Both analytical and numerical techniques rely on computing the gradient of the goal function with respect to the design parameters, and while discontinuous goals or parameters can be used if absolutely necessary, they generally add unnecessary complexity and can be rewritten using constraints.

Constraints define the range of design parameter values that are physically realizable. Constraints come in two varieties: parameter constraints and computed constraints. Parameter constraints apply directly to the design parameters; minimum and maximum

dimensions for the piezoresistor or Wheatstone bridge bias voltage. Parameter constraints simply limit the range of realizable design parameters. However, combinations of the design parameters are often important. For example, the sensor stiffness depends on all of the sensor dimensions, so cannot be written as simple bounds on each of the parameters, which is unfortunate because simple bounds (or box constraints) are far simpler to handle in numerical optimization [318]. Computed constraints can be nonlinear and can be either inequality or equality constraints, and can be handled by both analytical and numerical design optimization techniques. While some design parameters are flexible and can be modified during the design process (e.g. dopant concentration), others are constrained to predetermined values (e.g. measurement bandwidth, contact via size); parameters that are constrained to a single value can either be excluded from the entire optimization process or constrained with the lower and upper bounds set to the desired value.

Device designs can be optimized either analytically or numerically. Both approaches strive to find the combination of design parameters that achieve optimal performance while respecting design constraints. Both methods operate by finding a combination of design parameters where the gradient of the optimization goal, $\nabla f(\mathbf{x})$, approaches zero and the Hessian, $\nabla^2 f(\mathbf{x})$, is either positive or negative depending on whether a minimum or maximum value of the performance function is desirable. In this syntax, $f(\mathbf{x})$ may be the MDF while the vector \mathbf{x} defines the device design parameters. In analytical methods we symbolically manipulate the device design equations to find parameter combinations where $\nabla f(\mathbf{x}) \rightarrow 0$, while in numerical methods we utilize a computer to help us in the task.

The analytical optimization of piezoresistive devices has been reported numerous times with varying degrees of model complexity [75, 194, 319, 320]. As an example of analytical optimization, consider the force resolution of a piezoresistive cantilever. Beginning from (2.143), which neglects piezoresistor self-heating, amplifier noise, and thermomechanical noise, the MDF can be written as

$$\text{MDF} = \frac{\sqrt{4k_b T_{\text{pr}} R(f_{\max} - f_{\min}) + \alpha V_{\text{bridge}}^2 \ln(f_{\max}/f_{\min})/2N_{\text{eff}}}}{3V_{\text{bridge}}(l_c - l_{\text{pr}}/2)\pi_L^{\max}\gamma\beta^*/2w_c t_c^2}.$$

In the case of an epitaxial device, we can make the following simplifications

$$\begin{aligned}
R &= 4l_{\text{pr}}/q\mu n w_c t_{\text{pr}} \\
N &= nl_{\text{pr}} w_c t_{\text{pr}} \\
\beta^* &= P(1 - t_{\text{pr}}/t_c) \\
P &= \log_{10}(1.53 \times 10^{22}/n)^{0.2}
\end{aligned}$$

where β^* for epitaxial piezoresistors was derived earlier in (2.122) and we are using Harley's empirical fit for the piezoresistance factor presented earlier in (2.26). Further, we'll assume the $\gamma \approx 1$ for simplicity, although in reality it varies with R .

After making these substitutions, the final expression for the MDF is

$$\text{MDF} = \frac{\sqrt{16k_b T_{\text{pr}} l_{\text{pr}} (f_{\max} - f_{\min}) / q\mu n w_c t_{\text{pr}} + \alpha V_{\text{bridge}}^2 \ln(f_{\max}/f_{\min}) / 2nl_{\text{pr}} w_c t_{\text{pr}}}}{3V_{\text{bridge}}(l_c - l_{\text{pr}}/2)\pi_L^{\max} \log_{10}(1.53 \times 10^{22}/n)^{0.2}(1 - t_{\text{pr}}/t_c)/2w_c t_c^2}. \quad (2.147)$$

In order to analytically optimize the device design the next step would be to calculate the gradient of the MDF with respect to the design parameters. This is by no means a simple task, although the process can be simplified by recognizing certain parameters (e.g. w_{pr} , V_{bridge}) that should be as large or as small as possible. We refer the reader to Harley et al [75] for a complete design optimization of epitaxial piezoresistive cantilevers for constant power dissipation operation.

The fundamental problem with analytical optimization is one of complexity. Even (2.147), while unwieldy, is a simplified device model that neglects Joule heating, amplifier and thermomechanical noise, and all forms of design constraints. Probably most importantly, the approach does not extend to diffused or ion implanted piezoresistors, where simple expressions for R , β^* and N are not easily obtained. Other aspects of device design, such as dynamics when operating in liquid, can not be included in this approach either. Fortunately, numerical optimization methods do not suffer from these limitations and can be implemented in a very user friendly fashion.

Numerical optimization evaluates the goal function and its gradients numerically in order to automatically find the design optimum. Consider $f(\mathbf{x})$ and an initial set of design parameters, $\mathbf{x} = (x_1, x_2, \dots, x_n)$. Starting from this initial design, we can calculate the

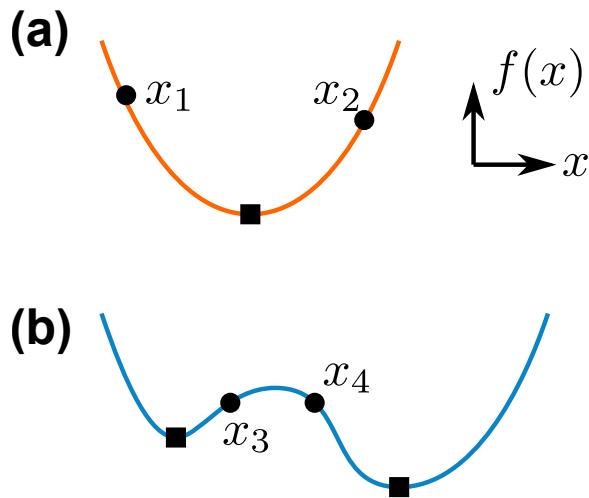


Figure 2.51: Examples of (a) convex and (b) non-convex optimization problems. In (a), a single global optimum exists and the same result will be obtained whether x_1 or x_2 is used as the optimization starting point. However, piezoresistive optimization is not usually a convex problem, and many local optima exist. In (b), the optimizer will return different results depending on whether it starts from x_3 or x_4 . The solution to this problem is seeding the optimizer with several randomly generated starting points.

gradient, $\nabla f(\mathbf{x}) = (\partial f / \partial x_1, \partial f / \partial x_2, \dots, \partial f / \partial x_n)$, with respect to the design parameters. We can then decrease $f(\mathbf{x})$ by moving along the directions where $\partial f / \partial x_i < 0$. By taking a step in the direction where the gradient is most negative, recalculating the gradient and iterating, we will eventually find a point where $\nabla f(\mathbf{x}) \rightarrow 0$. The gradient can be zero at minima, maxima, and saddle points, so the Hessian, $\nabla^2 f(\mathbf{x})$, is used to determine the nature of the inflection point. The search direction (p), step length (α), and handling of constraints varies between optimization methods.

Numerical optimizers can either calculate the gradient at each iteration or use a provided value. Allowing the optimizer to calculate the gradient numerically is slower but allows more complex models to be used and is necessary for piezoresistors fabricated by either predeposition or ion implantation. The numerical optimizer estimates the gradient and Hessian by calculating the change in $f(\mathbf{x})$ for small changes in \mathbf{x} . The optimizer will need to calculate $f(\mathbf{x})$ up to $2n$ times for each step (potentially less if any constraints are active), where n is the number of design parameters. Thus our emphasis on calculation speed in the selection of device models (e.g. ion implantation modeling in Section 2.5.3).

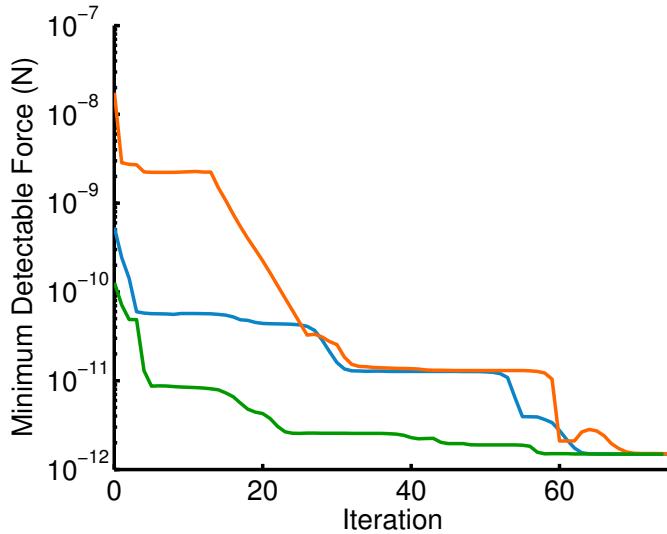


Figure 2.52: Numerical optimizer operation. Several sensor designs are randomly generated and optimized until each converges. The non-convex nature of the problem is handled by optimizing randomly generated designs until several of them converge to the same solution. Reprinted from Ref. [315], ©2009 AIP.

Although optimization is conceptually straightforward, there are numerous practical difficulties ranging from large scale and non-convex problems to stochastic and multi-objective problems. For example, optimization problems are convex if $\nabla^2 f(\mathbf{x}) > 0$ for all \mathbf{x} [321]. If the problem is non-convex then multiple minima will exist (Figure 2.51). Piezoresistor design is non-convex [315], so piezoresistive sensor optimization techniques will occasionally find local minima that are not the global minimum.

Although the problem is not strictly convex, the global optimum can reliably be found by starting from many random initial guesses until multiple iterations converge to within a small tolerance. Typical optimizer operation is illustrated in Figure 2.52. Three piezoresistive cantilever designs are randomly seeded to start the process. The initial designs are randomly generated assuming a uniform probability density function for each parameter between its upper and lower bounds. At each iteration, the optimizer evaluates the gradient and Hessian matrices and chooses a step direction that would improve the design performance (e.g. minimizing the MDF). The MDF generally decreases every iteration, unless the optimizer runs into a constraint, in which case it may need to briefly

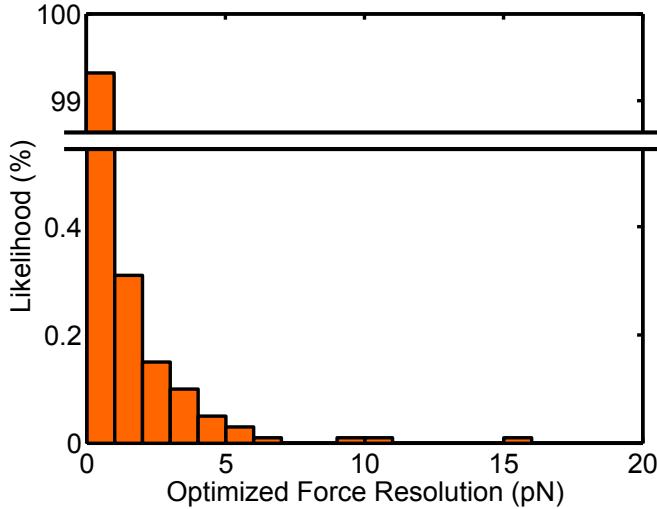


Figure 2.53: Numerical optimizer convergence statistics. The global optimum is found 99.3% of the time in the case of an epitaxial piezoresistor. Running the optimizer with randomly generated designs until three of them converge to the same result ensures that the global optimum is missed less than once every million runs.

increase the MDF. All three initial designs reach the same result and the performance improves up to four orders of magnitude compared with the initial, randomly generated designs.

The statistics of the numerical optimizer are presented in Figure 2.53. The optimizer finds the global optimum 99.3% of the time in the case of an epitaxial piezoresistive cantilever. The other 0.7% of the time it finds a local optimum which is slightly inferior to the global optimum. By running the optimizer with randomly generated designs until several of them converge to the same result the global optimum can reliably be found. For example, if the optimizer misses the global optimum 0.7% of the time, waiting until three optimization runs converge to the same result ensures that the global optimum is missed less than once every million runs.

We won't go into numerical design optimization much further than this, instead relying on widely available optimizer implementations, but we refer the reader to Nocedal and Wright [322] for a comprehensive treatment of numerical optimization. We use the interior-point method with the gradient approximated by finite-differences and Quasi-Newton Hessian updates as implemented by `fmincon` in the Matlab Optimization Toolbox

(Mathworks, Cambridge, MA). Open source alternatives are available as well, including one in SciPy, an alternative to Matlab. Optimization typically takes less than a minute for the analysis models presented in this work.

Numerical optimization has been used by several other piezoresistor research groups. Researchers at Bosch have presented several numerical optimization techniques for piezoresistive accelerometers [317, 323]. The Bosch researchers used gradient-based optimization algorithms to find local optima while they used genetic algorithms to aid in finding the global optimum. We will use gradient-based methods, but rather than using genetic algorithms, we will generate randomized seed designs in order to find the global optimum.

Their 2010 paper [317] addresses some of the practical issues with numerical optimization that we did not have room to discuss in our 2009 paper [315]. Most notably, they discuss parameter scaling in detail. In brief, the wide variation in design parameter scale (e.g. $t_{pr} = 10^{-7}$ m while $n = 10^{20}$ cm $^{-3}$) makes it difficult for the optimizer to choose an appropriate step size for gradient evaluation. To address this issue, we scale the design parameters to O(1) at every point in the process where the optimizer needs to choose a step direction and step length. Our optimizer implementation at the time of thesis publication is reproduced in full in Appendix H. The optimizer is fully integrated into the cantilever modeling code base, so that the electrical, mechanical and thermal models discussed in this chapter are properly accounted for.

2.8 Practical notes

Before moving on to the particular case of designing a force probe in the next chapter, there are a few general issues with piezoresistors that are worth describing.

2.8.1 Wafer selection

Piezoresistive sensors are typically fabricated by bulk micromachining so require the use of double-side polished wafers to allow both front- and backside-patterning. Silicon-on-insulator (SOI) wafers are often the preferred wafer option because the buried SiO₂ layer

can be used for a reliable etch stop and electrical insulation. The sensor thickness can be controlled fairly tightly in SOI wafers as well, although they are more expensive than double-side polished wafers. Although most piezoresistive sensors are bulk micromachined, surface micromachined sensors are becoming more popular. Single crystal silicon membranes can be surface micromachined using the porous silicon process [324, 325], enabling the pre-CMOS fabrication of piezoresistive pressure sensors.

SOI wafers come in two varieties: bonded and etched-back SOIs (BESOIs) and ion implantation-based SOIs. Although there are several books on SOI fabrication techniques [326, 327], we are providing a brief summary here with emphasis on issues for piezoresistive sensors.

BESOI wafers are fabricated by bonding a wafer to an oxidized handle wafer and thinning the bonded wafer to the desired device layer thickness through grinding and chemical mechanical planarization (CMP). Variation in the BESOI device layer thickness is typically 0.5 μm or more. Accordingly, they are typically used for device layers more than 1 μm thick due to the strong thickness dependence of sensor mechanics and sensitivity.

Thinner device layers, used for high performance sensors and fully depleted ICs, require an alternative fabrication method. Rather than polishing back the bonded wafer, a thin layer of oxygen (e.g. the SIMOX process) or hydrogen (e.g. the Smart Cut process) is ion implanted into the oxidized device wafer before it is bonded to the handle wafer. Annealing the bonded wafer stack forms either a buried oxide (SIMOX) or a weakened layer that allows the device layer to be mechanically cleaved before a final CMP cleanup step (Smart Cut). In either case the thickness of the device layer is set by the projected range of the implanted ions so only relatively thin device layers can be formed (e.g. < 500 nm thick). Thickness control is excellent with 3σ thickness variation on the order of 30 nm or less.

Particular care should be taken in the selection of SOI vendor when fabrication piezoresistive sensors, because contamination during the fabrication process can lead to buried impurity layers that interfere with piezoresistor operation. Nafari et al. investigated wafers from three SOI manufacturers and found boron contamination with wafers from one of them [328]. At least one researcher experienced boron contamination from the same vendor, which shorted out the piezoresistors on his wafers and required another fabrication run to be performed [329]. Nafari et al. recommend ensuring that the SOI

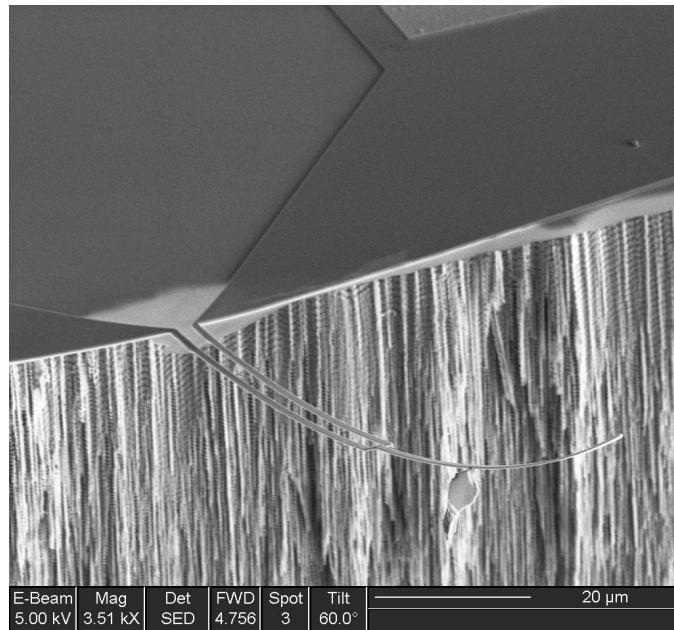


Figure 2.54: Example of dopant-induced lattice strain leading to device curvature. The high dopant concentration near the device surface, which maximizes piezoresistor sensitivity (β^*) leads to a strain gradient through the device thickness.

vendor oxidizes both the device and handle wafer, using the highest possible background dopant concentration and maximizing the device thickness to avoid contamination issues near the BOX. Wafer quality can be verified by either SIMS or SRA before committing a batch of wafers.

2.8.2 Dopant-induced lattice strain

Piezoresistive sensors tend to be thin, making them more susceptible to residual strain induced bending. Bending can be caused by a residual strain or strain gradients within silicon, dielectric or metal layers on the sensor. Strain in sputtered films can be minimized by careful process development, and strain gradients in the device layer itself can be minimized through a long, high temperature anneal [330]. We typically use a 20 hour, 1100°C anneal in an inert atmosphere to reduce strain and strain gradients in SOI wafers and have observed a noticeable curvature reduction in released cantilevers.

Dopant atoms can add residual strain to the device as well. The addition of smaller

or larger dopant atoms to the crystal lattice will induce tensile or compressive strain, respectively. Boron (0.88Å) and phosphorus (1.10Å) are both smaller than silicon (1.17Å) while arsenic (1.20Å) is slightly larger. Numerous researchers have measured the strain induced by the addition of dopants to the silicon lattice and have observed a linear reduction in average lattice constant with increasing dopant concentration for boron and phosphorus. The lattice strain can be calculated from

$$\epsilon = \beta_1 n \quad (2.148)$$

where n in this case is the total, not just electrically active, dopant concentration and β_1 is the lattice contraction coefficient. Celotti et al. measured contraction coefficients of $4.5 \times 10^{-24} \text{ cm}^3$ and $1.8 \times 10^{-24} \text{ cm}^{-3}$ for boron and phosphorus, respectively, via X-ray diffraction (XRD) rocking curves [331]. Lee et al. measured the contraction coefficient for phosphorus doping via micromachined mechanical test structures and XRD and inferred a contraction coefficient of $4.5 \times 10^{-24} \text{ cm}^{-3}$ [332]. McQuhae and Brown measured contraction coefficients of 5.2×10^{-24} and $1 \times 10^{-24} \text{ cm}^{-3}$ for boron and phosphorus, respectively, via measurements of wafer curvature [333].

We fabricated thin piezoresistive cantilevers using POCl_3 predeposition and also observed substantial lattice strain (Figure 2.54). The dopant concentration gradient within the device (i.e. $\beta^* > 0$) is essential to the operation of piezoresistive devices but directly induces an accompanying strain gradient. By measuring the tip deflection of cantilevers with varying lengths we inferred a lattice contraction coefficient of $4 \times 10^{-24} \text{ cm}^{-3}$ (Figure 2.55). Our measurement technique was based upon the concentration of electrically active dopants rather than their total concentration, but the results are inline with the other data. We will use the data from Celotti et al. in calculating the curvature of boron and phosphorus doped piezoresistive sensors.

Lattice strain is important for the design of our thin force probes, so we will briefly go over the theory here. Hubbard and Wylde investigated residual strain effects in micromachined cantilever beams and presented analytical theory for calculating bending for arbitrary strain profiles [334]. We assume that no additional films are present on the sensor where the piezoresistor is located and that the dopant concentration profile, $n(z)$, is

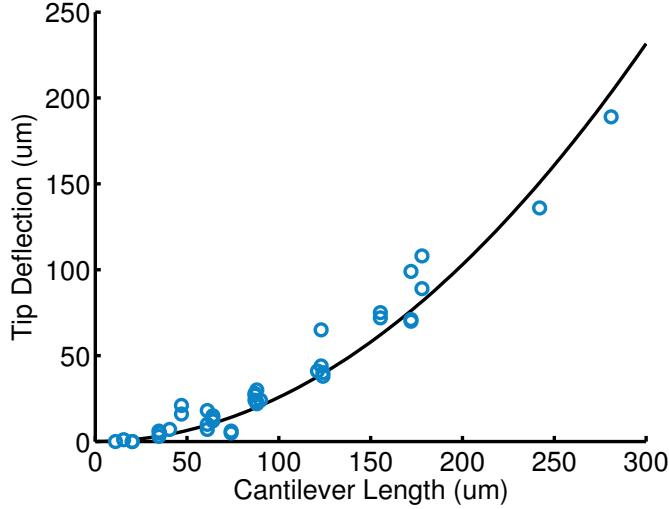


Figure 2.55: Example experimental cantilever bending results. The entire length of the cantilever was doped via POCl_3 inducing uniform bending along its length.

known. Using the concentration profile we can calculate the stress profile,

$$\sigma(z) = E\varepsilon(z) = E\beta_1 n(z),$$

and calculate the corresponding centroid of the stress distribution as

$$\bar{z} = \frac{\int_0^{t_c} z \sigma(z) dz}{\int_0^{t_c} \sigma(z) dz}$$

where the origin of the coordinate system is at the bottom of the sensor. We can treat the beam as a bilayer structure with the interface between the two layers located at \bar{z} . The layer thicknesses are

$$t_1 = t_c - \bar{z}$$

$$t_2 = \bar{z}$$

where layer 2 is at the bottom of the cantilever beam, nearest to the origin. The average axial stress of each layer is

$$\sigma_1 = \frac{1}{t_1} \int_{t_2}^{t_c} \sigma(z) dz$$

$$\sigma_2 = \frac{1}{t_2} \int_0^{t_2} \sigma(z) dz$$

and the centroid of the stress for each layer is

$$\bar{z}_1 = \frac{\int_{t_2}^{t_c} z \sigma(z) dz}{\int_{t_2}^{t_c} \sigma(z) dz}$$

$$\bar{z}_2 = \frac{\int_0^{t_2} z \sigma(z) dz}{\int_0^{t_c} \sigma(z) dz}.$$

We can calculate the equivalent moment induced by each layer about the \bar{z} from

$$M_1 = \sigma_1 t_1 w (\bar{z}_1 - \bar{z})$$

$$M_2 = \sigma_2 t_2 w (\bar{z} - \bar{z}_2)$$

where w is the width of the cantilever beam. The curvature of the beam can be calculated, assuming Euler-Bernoulli beam theory and small angular deflections, from

$$C = \frac{M_1 + M_2}{EI}$$

where we note that term cancellation makes the curvature independent of the beam width and elastic modulus. The curvature can be translated to angle and deflection from

$$\theta(x) = \int_0^x C(x) dx \quad (2.149)$$

$$v(x) = \int_0^x \theta(x) dx. \quad (2.150)$$

where x is the distance from the base of the cantilever. Note that $C(x) = 0$ beyond the

end of the piezoresistor or any other doping so that $\theta(x > l_{\text{pr}}) = \theta(l_{\text{pr}})$. The deflection at the tip of the cantilever beam, assuming that it is doped along the entire length, is $Cl_c^2/2$. The tip deflection is inversely proportional to the bending rigidity of the beam, so dopant-stress becomes less of an issue as the sensor thickness increases.

Dopant-induced bending can be minimized by minimizing the dopant concentration, maximizing the bending rigidity of the sensor, minimizing the area that is doped, or adding a stress compensation film. Modifying the dopant concentration or bending rigidity are usually sub-optimal solutions, and if the area that is doped is already limited to just the piezoresistors then the only option if bending remains an issue is to add a compensation film. Dopant stress can be compensated by adding a film with compressive (for boron or phosphorus doping) or tensile (for arsenic doping) stress directly above the piezoresistor. The ideal compensation material would have a low elastic modulus, high residual stress, and could be deposited with minimal additional thermal budget so that only a thin layer would be necessary and it would negligibly affect the sensor mechanics or electrical properties.

The thickness of a compensation film can be determined by calculating the sum of the moments at any point on the sensor from

$$\sum M = \sum_i \int (z_i - z_n) \sigma_i dA \quad (2.151)$$

where i is the film number and z_n is the neutral axis of the device cross-section about which the moments are calculated. If thickness of the compensation film (t_{comp}) is much less than the thickness of the silicon (t_c), then we can approximate the neutral axis as $z_n = t_c/2$. Assuming that the compensation film has been chosen to ensure a flat sensor we can assume that the bending stress, $E(z - z_n)/R$ is negligible, simplifying the analysis considerably.

We assume that the stress profile in the silicon is an arbitrary function of depth, $\sigma(z) = E\varepsilon(z)$, while the stress in the compensation film is a constant value, σ_{comp} . Expanding (2.151),

$$0 = \int_0^{t_c} (z - z_n) \sigma(z) dz + \int_{t_c}^{t_c + t_{\text{comp}}} (z - z_n) \sigma_{\text{comp}} dz.$$

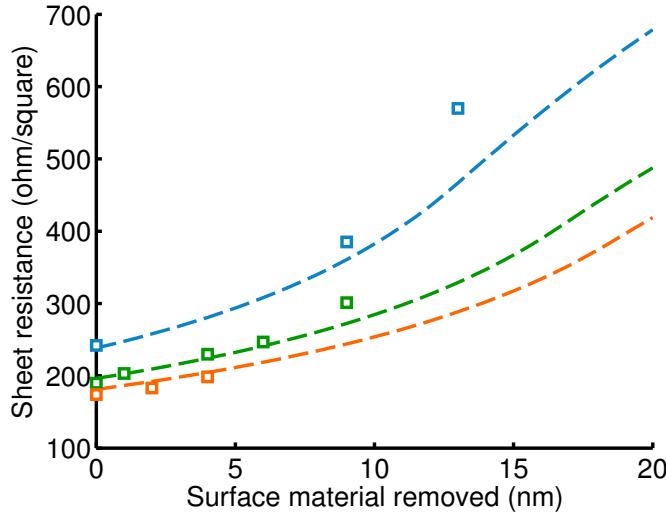


Figure 2.56: Variation in piezoresistor sheet resistance with surface material removal. SOI samples were doped with phosphorus predeposition at 775°C for 25, 35 and 40 minutes and initial sheet resistance measurements were performed. The samples were subsequently coated with LPCVD SiO₂ at 300-400° for 20-60 minutes. Surface material removal was calculated by measuring the thickness of the SOI device layer before and after the deposition step via ellipsometry.

We can continue to expand the result in order to obtain

$$t_{\text{comp}}^2 \left(\frac{\sigma_{\text{comp}}}{2} \right) + t_{\text{comp}} \left(\frac{t_c \sigma_{\text{comp}}}{2} \right) + \int_0^{t_c} (z - \frac{t_c}{2}) \sigma(z) dz = 0, \quad (2.152)$$

which can be solved using the quadratic equation and taking the positive solution for t_{comp} . For the force probes developed in this thesis, a compensation film was considered but not ultimately used. Instead, doping was confined to the relatively short piezoresistor at the base of the beam so that the static tip deflection from dopant stress was finite but relatively small. The first generation of force probes that were designed and fabricated, discussed in Appendix F, doped the cantilever beam along its entire length, resulting in much greater curvature and nonlinear mechanics.

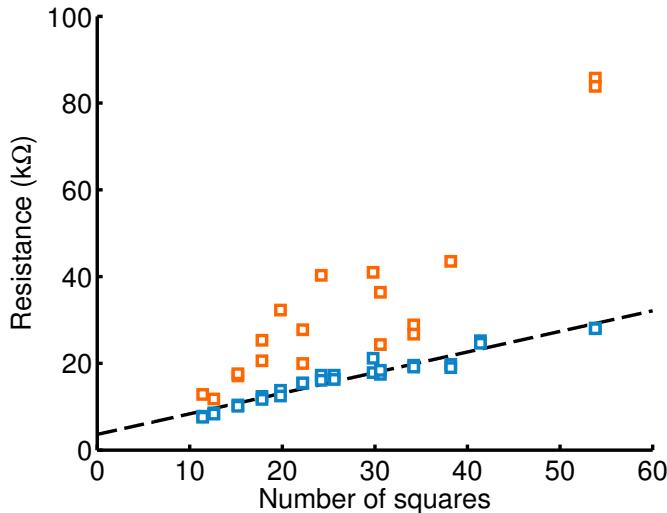


Figure 2.57: Variation in sheet resistance with reactive ion etching damage. Silicon resistors released from the BOX by a backside RIE step (orange) exhibited much higher resistances and variability than their on-chip counterparts (blue) that were not released. The increased resistance and mismatch lead to reduced performance. Switching to an HF vapor release step alleviated this problem.

2.8.3 Piezoresistor surface damage

In addition to being prone to strain effects, thin piezoresistive sensors are prone to surface damage during wafer processing. The sheet resistance of shallow piezoresistors can vary substantially during processing due to surface oxidation or damage. For example, LPCVD deposition of SiO_2 at 300-400°C usually oxidizes 5-10 nm of surface silicon. The slight reduction in thickness can affect both the mechanics and electrical properties of thin sensors. The sheet resistance and number of carriers are particularly prone to processing variation because the most highly doped silicon is generally located at the top surface of the wafer.

Variation in sheet resistance with surface oxidation can be experimentally measured and modeled (Figure 2.56). Wafers were doped with POCl_3 predeposition at 775°C for 25, 35 and 40 minutes and initial sheet resistance measurements were performed. The samples were subsequently coated with LPCVD SiO_2 at 300-400°C for 20-60 minutes. The changes in silicon thickness were measured via variable angle spectroscopic ellipsometry. Thin SOI device layers (300nm in this case) transmit enough light to enable optical measurements

of thickness, enabling the measurement of uniform changes in device layer thickness. The circles correspond to experimental data and the solid lines correspond to predictions obtained by calculating the theoretical dopant concentration and removing material from the surface. Although only several nanometers of material was removed during each step, the impact can be substantial when the piezoresistor is thin. Arlett reported a 25% sheet resistance increase from the deposition and patterning of a passivation PECVD oxide layer, corresponding to 12 nm of removed material from her 60 nm thick epitaxial piezoresistors [335].

A second source of electrical variation in thin piezoresistors can come from surface damage during reactive ion etching. We fabricated piezoresistive cantilevers from SOI wafers and released them from the BOX using a CHF_3/O_2 plasma. Despite the low nominal silicon etch rate of the process (28 Å/min), the resistors located on the cantilevers increased substantially compared to their temperature compensation counterparts located further back on the silicon die (Figure 2.57).

Changes in piezoresistor electrical properties during wafer processing can be mitigated in two ways. First, the impact of surface damage is inversely proportional to the piezoresistor thickness. Thicker sensors are less sensitive to process variation than their thinner counterparts in general, with the tradeoff being slightly worse performance. The second approach to limiting surface damage is to eliminate the damaging steps from the process. For example, sensors can be released from the SOI BOX using HF vapor rather than RIE in order to avoid problems with etch selectivity and the deposition of passivation polymer. Surface oxidation is more challenging to control, but atomic layer deposition (ALD) is a promising option for the deposition of thin nitride- or oxide-based surface passivation films at low temperature.

2.8.4 Lateral dopant diffusion

Throughout this chapter we have neglected the lateral diffusion of dopants during fabrication. Lateral diffusion results in the actual piezoresistor dimensions not matching the dimensions on the lithographic mask as illustrated in Figure 2.58. Lateral diffusion becomes a problem as the piezoresistor dimensions approach the dopant diffusion length

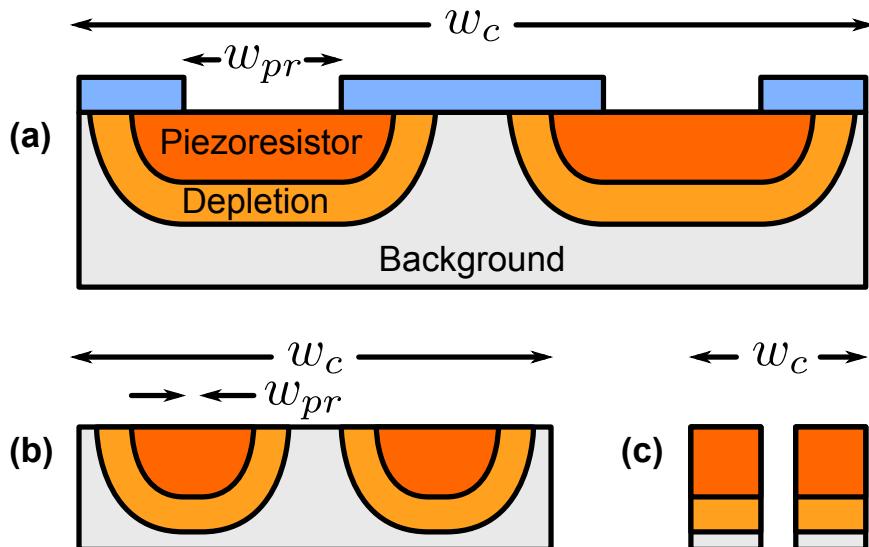


Figure 2.58: Lateral dopant diffusion effects. In this example, a piezoresistive loop is formed via phosphorus diffusion through an oxide mask. (a) If the piezoresistor dimensions are much larger than the dopant diffusion length then there is negligible error in using the lithographic mask dimensions to model the device. (b) As the piezoresistor dimensions decrease, the effective piezoresistor dimensions can become significantly larger than its nominal dimensions. Additionally, the device width is limited by the size of the depletion regions. (c) The extra material around the piezoresistor can be etched away after doping in order to control its dimensions and enable further reductions in device size.

(Figure 2.58b). A larger than intended piezoresistor results in a lower sheet resistance, more carriers in the piezoresistor and a greater risk of the depletion regions coming into contact than expected.

Consider a phosphorus predeposition process that yields a phosphorus saturated layer that is 50 nm thick, a junction depth of 300 nm and depletion regions that are 1 μm wide. For simplicity we will assume that the lateral and vertical dopant diffusion lengths are equal [236]. In reality, the lateral diffusion length can be significantly greater than expected due to oxidation enhanced diffusion and surface defects. For example, a four-fold greater lateral diffusivity than expected has been reported [336].

If each leg of the piezoresistor were 30 μm across then the error from lateral diffusion would be negligible ($2 \times 50 \text{ nm} / 30 \mu\text{m} = 0.3\%$). If the device dimensions were reduced to a nominal piezoresistor width of 1 μm then there would be significant error between

the nominal and effective piezoresistor width ($2 \times 50 \text{ nm} / 1 \mu\text{m} = 10\%$). Furthermore, the minimum spacing between the two legs of the piezoresistor would be roughly $5 \mu\text{m}$ in order to allow a $2 \mu\text{m}$ gap between the two depletion regions. This is a particularly important issue for designing sub-micron piezoresistive sensors.

Even if the depletion regions that surround each leg of the piezoresistor do not come into contact (punchthrough), electrical current leaks across the reverse biased pn junctions. If the piezoresistor is designed with a high impedance (e.g. $> 100 \text{ k}\Omega$) then the junction leakage can appreciably affect the piezoresistor electrical properties. The current that leaks from the piezoresistor to the substrate ground (or between the piezoresistor legs if the substrate is left floating) will reduce the sensor sensitivity, while carrier generation and recombination can increase the piezoresistor noise [337]. For example, Harjee observed a four-fold reduction in piezoresistor 1/f noise when the gap between two piezoresistor legs was etched away [329].

These issues can be eliminated by etching away the material that surrounds the piezoresistor (Figure 2.58c). The etch step both defines the lateral dimensions of the piezoresistor and eliminates the current leakage and noise issues with the gap between the piezoresistor legs. For applications that require a continuous structure (e.g. pressure sensors) the gap could be refilled with a dielectric afterwards. Eliminating the excess material also increases the force resolution of cantilever beam structures (due to the spring constant reduction) and we will use this approach in fabricating the force probes presented in this work.

Chapter 3

Force probe design

The performance of our microscale force sensor and actuator depends upon the optimization of numerous coupled parameters; the macroscale mechanical design of the probe, the sensitivity and noise of the piezoresistor and measurement circuit, the mechanics of the probe itself, and the speed of the actuator and driving electronics. In this chapter we will touch on all of these topics, aiming to obtain an optimized mechanical, electrical and thermal design. The sensor design relies heavily on the numerical optimization and piezoresistor models described in the previous chapter.

As discussed in Section 1.4.3, three different probe varieties were designed and fabricated. The probe designs integrate piezoresistive sensing with either no actuation, thermal actuation or piezoelectric actuation. The layout and dimensions of the thermally and piezoelectrically actuated probes are presented in Figure 3.1. The sensor-only probe variety is not shown but is similar to the thermally actuated probe except in that it lacks a resistive silicon heater.

Both the thermally and piezoelectrically actuated probes utilize a unimorph cantilever actuator, where differential expansion between layers leads to transverse tip motion. In the case of the thermally actuated probe the differential expansion arises from the combination of a change in temperature and the large mismatch between the coefficients of thermal expansion of Si and Al. In the case of the piezoelectrically actuated probe it arises from electric field-induced strain in a piezoelectric film.

All of the probes consist of a narrow, soft sensor at the end of a wide, stiff actuator.

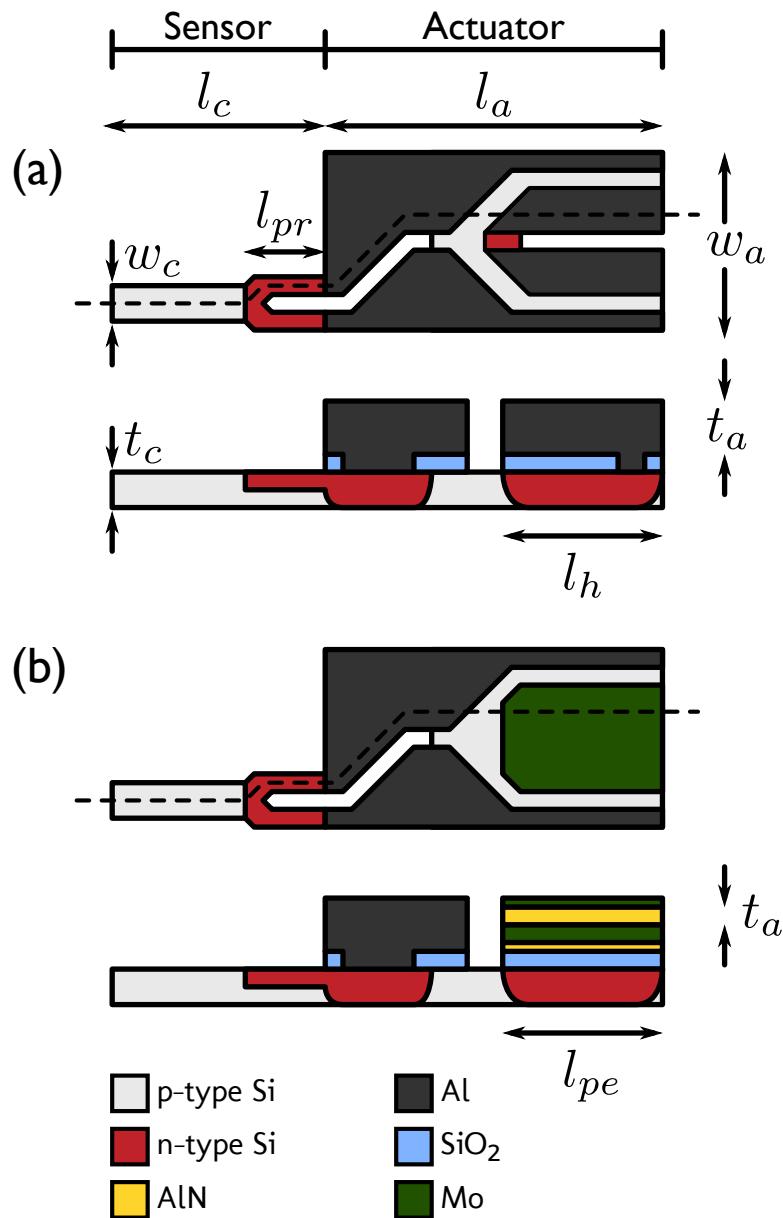


Figure 3.1: Device layouts. Force probe designs that integrate piezoresistive sensing with (a) thermal actuation or (b) piezoelectric actuation are presented. In both cases, the probe consists of a narrow, soft sensor attached to the end of a wide, stiff actuator. Placement of the actuator at the base rather than the tip leads to numerous performance benefits and decouples the mechanical and thermal design of the actuator and sensor.

The sensor-only devices do not include an actuator but retain the stiffened base in order to define the force probe length from the front of the wafer rather than during the imprecise backside release etch. We will show that placing the actuator at the base rather than the tip maximizes the tip deflection range, maximizes the overall resonant frequency of the probe, minimizes the thermal actuator time constant and decouples the mechanical design of the sensor and actuator portions.

Electrical isolation between the sensor and actuator are critical due to their close proximity (e.g. a gap of 2 μm). On the thermally actuated probe the piezoresistive sensor and silicon heater are electrically isolated at DC by reverse biased pn junctions, while the piezoelectric actuator is separated from the underlying silicon by a 100-200 nm thick layer of SiO_2 . Capacitive coupling between the actuator and sensor becomes increasingly important as the actuation frequency increases. Mechanical crosstalk is also an issue in the force probe design because the Al interconnects run alongside the actuator and bend with it. A final source of actuator-sensor crosstalk, unique to the thermally actuated probes, is thermal crosstalk between the silicon heater and piezoresistive sensor. Probe design and circuit-based compensation strategies for capacitive, mechanical and thermal crosstalk will be addressed in Section 3.3.3.

The probe dimensions are illustrated in Figure 3.1. The length, width and thickness of the sensor are l_c , w_c and t_c while the piezoresistor has length l_{pr} . The actuator has overall length and width of l_a and w_a , which includes both the active portion of the actuator and the electrical contacts between the Al interconnects and piezoresistor. The heater on the thermally actuated design and the piezoelectric stack on the piezoelectrically actuated design extend l_h and l_{pe} from the base of the force probe, respectively. We will discuss the actuator design in more detail in Section 3.2.

3.1 Sensor design

The sensor is largely designed through numerical optimization using the models that we presented in the last chapter. The designs produced by the numerical optimizer are presented in Section 3.1.3. In this section we will discuss the cantilever-specific design issues that are included in the numerical model but were not discussed in the last chapter.

3.1.1 Mechanics

As discussed in Section 1.3, force probes that are simultaneous soft and fast must be made as small as possible. While a zero-dimensional probe (i.e. a bead) minimizes the mass, its spring constant can not be easily adjusted which limits the maximum measurement bandwidth. A one-dimensional probe (i.e. a cantilever beam) remedies this limitation by allowing bandwidth adjustment via the cantilever beam length. Any additional dimensions (e.g. tines for capacitive readout) simply contribute additional mass. Accordingly, we chose to use a cantilever beam geometry with the smallest convenient width and thickness.

The cantilever spring constant to resonant frequency in vacuum ratio scales as

$$\frac{k_c}{f_0} \propto \frac{\sqrt{E_c \rho_c} w_c t_c^2}{l_c} \quad (3.1)$$

where l_c , w_c and t_c are the length, width and thickness of the cantilever while E_c and ρ_c are its elastic modulus and density. We opted for a cantilever beam thickness of 300 nm because the reduction in thermal conductivity [306] and uncertainty in the effective elastic modulus [338, 339] of thinner probes lead to diminishing performance returns for piezoresistive cantilevers [302].

This analysis only applies to a one-dimensional cantilever beam with a simple diving board geometry and constant cross-section. However, as noted in Section 1.3, the probe needs to be made as small as possible. If the entire width of the force probe is made equal to the minimum feasible size for a given fabrication technology (i.e. 500 nm wide piezoresistor legs and 1 μm overall width as in the current work) then there is negligible value in considering alternative cross-section variations (e.g. tapered) because, by definition, they would be larger than the minimum possible feature size at some point along their length.

The spring constant and measurement bandwidth of the force sensor depends on the frequency response of the sensor-portion of the force probe. The actuator-portion is intentionally designed to have a much higher spring constant and resonant frequency than the sensor. We will address the frequency response of the combined sensor-actuator system in Section 3.3.1, but for now will focus on modeling the sensor-portion of the force probe independently of the actuator.

In vacuum it is trivial to calculate resonant frequency of a fixed-free cantilever beam. However, the force probes are intended to be used in air or water. Operation in air results in a smaller quality factor and resonant frequency compared to operation in vacuum, and water operation decreases both parameters further still.

The frequency range over which the force probes can be used depends on the frequency response of the sensor. We will operate the probes over a frequency range where the frequency response is relatively flat. When the frequency response is critically damped ($Q \approx 1/2$), as is the case in water, the probes can operate up to their damped resonant frequency (f_d). In air, the probes generally can only be operated up to about $f_0/2$ depending on how flat the frequency response needs to be and if ringing is acceptable.

The resonant frequency and quality factor of the sensor in fluid are calculated using the model presented by Brumley, Willcox and Sader [23]. Qualitatively, one would expect the frequency and the quality factor of the first eigenmode for a cantilever to decrease as the density and viscosity of the surrounding fluid increases. Fortunately, there are quantitative solutions to the fluid dynamics problem of a cantilever oscillating in fluid that we can take advantage of. The behavior of cantilevers in fluid has received numerous theoretical treatments, particularly by John Sader in a series of papers dating back to 1995 [340]. We will briefly summarize the theory of cantilevers in fluid here.

The frequency response of a cantilever in fluid can be treated using continuum mechanics as long as the mean free path of the surrounding fluid (e.g. ≈ 50 nm in air, < 1 nm in water) is significantly smaller than the size of the cantilever. This assumption holds for our cantilevers (minimum dimension of 300 nm) although deviations can be seen in smaller cantilevers [341]. The fluid can be treated as incompressible as long as the wavelength of the cantilever vibration greatly exceeds the fluid flow length scale and the fluid flow length scale greatly exceeds the mean free path.

In general, the elastic deformation of a cantilever is governed by

$$EI \frac{\partial^4 v(x,t)}{\partial x^4} + \mu \frac{\partial^2 v(x,t)}{\partial t^2} = F(x,t) \quad (3.2)$$

where $v(x,t)$ is the deflection in the z -direction, μ is the mass per unit length and $F(x,t)$ is the external applied force per unit length in the z -direction. For a beam in vacuum,

$\mu = \rho_c w_c t_c$, where ρ_c , w_c and t_c are the density, width and thickness of the cantilever as usual. The Fourier transform of Eq. 3.2 is

$$EI \frac{\partial^4 \tilde{v}(\omega)}{\partial x^4} - \mu \omega^2 \frac{\partial^2 \tilde{v}(\omega)}{\partial t^2} = \tilde{F}(\omega) \quad (3.3)$$

The applied force can be broken down into the hydrodynamic load and the driving force. Components of the driving force include thermal vibrations and a load applied at the tip. The hydrodynamic force can be found by solving the Navier-Stokes equation [342], and is equal to

$$\tilde{F}(\omega) = \frac{\pi}{4} \rho_f \omega^2 w_c^2 \Gamma(\omega, n) \tilde{v}(\omega) \quad (3.4)$$

where ρ_f is the fluid density and Γ is the dimensionless hydrodynamic function. By substituting Eq. 3.4 into 3.3, the frequency and quality factor of the first resonant mode can be found using

$$\frac{f_d}{f_0} = \left(1 + \frac{3\pi\rho_f w_c}{2\rho_c t_c} \Gamma_r(\text{Re}, \kappa) \right)^{-0.5} \quad (3.5)$$

$$Q = \frac{2\rho_c t_c / 3\pi\rho_f w_c + \Gamma_i(\text{Re}, \kappa)}{\Gamma_r(\text{Re}, \kappa)} \quad (3.6)$$

where Γ_r and Γ_i are the real and imaginary parts of the hydrodynamic function. Re and κ are the Reynold's number and normalized aspect ratio, which are discussed next. It's interesting to note that a similar expression for the resonant frequency in fluid of an infinitely long cylinder in fluid assuming inviscid flow was derived by Chu in 1963 [343] as

$$\frac{f_d}{f_0} = \left(1 + \frac{\pi\rho_f w_c}{4\rho_c t_c} \right)^{-0.5}. \quad (3.7)$$

So essentially, if we know the dimensions and hydrodynamic function then we can calculate the damped resonant frequency and quality factor of a cantilever. But calculating the hydrodynamic function is nontrivial. Fortunately, a lookup table for the hydrodynamic function in terms of the Reynold's number and normalized aspect ratio was compiled by

Van Eysden and Sader in 2007 [344] and revised by Brumley, Willcox and Sader in 2010 [23].

Based upon our choice of the cantilever width w_c as the dominant fluid flow length scale, the Reynolds number (Re) and normalized aspect ratio (κ) are

$$Re = \frac{\rho_f \omega w_c^2}{4\eta} \quad (3.8)$$

and

$$\kappa = C_1 w_c / l_c \quad (3.9)$$

where η is the fluid viscosity and C_1 is the first solution to $1 + \cos(x) \cosh(x) = 0$, or about 1.875. The limit $\kappa \rightarrow 0$ corresponds to the case of an infinitely long cantilever beam. In the design optimization code, the Reynold's number and normalized aspect ratio are calculated from the cantilever dimensions and used to look up the appropriate hydrodynamic function using linear interpolation using the table presented in Ref. [23]. The lookup table and calculation method are included in the optimization code (Appendix H).

It is important to emphasize that the hydrodynamic function is a function of frequency, while the resonant frequency depends on the hydrodynamic function, resulting in a transcendental function that needs to be solved iteratively. Interestingly, the resonant frequency and quality factor of the first eigenmode of a microcantilever can be used to infer the density and viscosity of the surrounding fluid for rheology applications [345].

3.1.2 Piezoresistor optimization

A thin cantilever beam requires a shallow piezoresistor. Sub-micron thick piezoresistive cantilevers have been fabricated using epitaxy [36, 346], ion implantation [254, 347] and diffusion [195]. For our application, diffusion yields higher performance than ion implantation and is far simpler than epitaxy while matching its performance. We chose to use $POCl_3$ rather than BBr_3 predeposition to avoid the formation of a boron-rich surface layer [348], to orient the piezoresistors in the $\langle 100 \rangle$ direction to minimize E_c , and for

the performance advantage of n-type piezoresistors for force sensing applications [315]. The piezoresistors are electrically isolated by reverse biased pn junctions because the piezoresistors are positively biased and the substrate is grounded.

As discussed in the last chapter and numerous papers [43, 198, 202, 302, 315], the cantilever design parameters should be chosen to minimize the RMS force noise, equal to the minimum detectable force (MDF). The force noise is calculated by integrating the force noise spectral density over the intended measurement bandwidth of the probes. The MDF is calculated from

$$\text{MDF} = \frac{\sqrt{V_J^2 + V_H^2 + V_A^2 + V_{TMN}^2}}{S_{FV}} \quad (3.10)$$

where V_J , V_H , V_A and V_{TMN} are the integrated Johnson, 1/f, amplifier and thermomechanical voltage noise, respectively, while S_{FV} is the voltage sensitivity of the piezoresistor and signal conditioning circuit with respect to force. For the sake of discussion we can neglect the amplifier and thermomechanical noise (although the numerical model includes all of the noise sources), in which case the force resolution can be written as

$$\text{MDF} \approx \frac{\sqrt{4k_b T_{pr} R_{pr} (f_{\max} - f_{\min}) + \alpha V_{\text{bridge}}^2 \ln(f_{\max}/f_{\min})/2N_{\text{eff}}}}{3V_{\text{bridge}}(l_c - l_{pr}/2)\pi_L^{\max}\gamma\beta^*/2w_c t_c^2} \quad (3.11)$$

where T_{pr} is the average piezoresistor temperature, R_{pr} is the piezoresistor resistance, f_{\min} and f_{\max} define the measurement bandwidth, V_{bridge} is the Wheatstone bridge bias voltage, α is the Hooge factor, N_{eff} is the effective number of carriers and π_L^{\max} is the peak piezoresistive coefficient [202]. The resistance factor (γ) is the ratio of the piezoresistor resistance to the total resistance including interconnects and contacts, while the sensitivity factor (β^*) is calculated from the dopant concentration profile and accounts for both the finite depth of the piezoresistor and the decrease in piezoresistive coefficients with increasing dopant concentration [198].

Piezoresistive cantilever design, even before we account for self-heating and fluid damping, is clearly too complex for analytical design optimization. Instead we design the force probe sensor using numerical optimization [315]. Although the piezoresistor design problem is not convex, seeding the numerical optimizer with several random initial

guesses leads to a globally optimal design. The optimization code handles all of the doping profile, fluid damping and heat transfer details that we have discussed throughout the thesis. Second-order effects, such as the slight reduction in sensitivity from (3.11) due to the transverse current flow at the end of the piezoresistive loop, are also included in the numerical model and become increasingly important as piezoresistor size decreases.

The code includes an experimentally validated thermal model (Section 2.6.3) in order to enable temperature-constrained cantilever design [302]. The probes are designed to limit the maximum and tip temperature increases to approximately 5 K and 1 K during operation in water. Temperature-constrained piezoresistor design leads to a shorter piezoresistor than the more conventional power-constrained design approach. An additional benefit of a shorter piezoresistor is a reduction in the static device curvature due to the high dopant concentration near the piezoresistor surface (Section 2.8.2). One final benefit of modeling the cantilevers numerically is that the static deflection range due to intrinsic stress in the deposited films can be predicted using the Monte Carlo method.

3.1.3 Sensor designs

Using the numerical optimization code, we designed force sensors ranging from 30 to 200 μm long in order to span wide spring constant (0.3-40 pN/nm), resonant frequency (20-400 kHz in air and 1-100 kHz in water) and force resolution (1-100 pN) ranges. We opted to fabricate the force probes using optical lithography, leading to a minimum feature size of 500 nm. The overall sensor width is constant along the entire probe length and split between the two piezoresistor legs. Both 1 and 2 μm wide sensors (500 and 1000 nm wide piezoresistors) were designed and fabricated in order to mitigate the risk of fabricating devices at the resolution limit of the lithography system.

Table 3.1 presents the sensor designs that were fabricated. The designs and their performance were calculated using the modeling and optimization code (Appendix H). The sensor performance is calculated for the doping process used during fabrication, namely piezoresistor doping in POCl_3 at 825°C for 30 minutes while the contacts are doped at 900°C for 20 minutes. The resulting piezoresistor sheet resistance (R_s) is 110 Ω/\square with a contact resistance (R_{contact}) of 170 Ω . The designs assume that $\alpha = 10^{-5}$, a sensitivity

factor (β^*) of 0.233 and a carrier density (N_z) of $4.13 \times 10^6 \mu\text{m}^{-2}$. These design parameters match the experimentally measured parameters from the fabricated devices.

Performance is calculated for a Wheatstone bridge bias (V_{bridge}) of 1 V; in practice the devices are operated anywhere between 0.5 and 2 V depending on the temperature requirements of the application. The spring constants (k_c) of the force probe designs are calculated for unpassivated devices; passivation with ≈ 230 nm of parylene N, required for experiments in water, roughly doubles their spring constants. The resonant frequency in air (f_0) and water (f_d) both account for fluid damping, calculating from Ref. [23]. The minimum detectable displacement (MDD) and force (MDF) are calculated assuming a measurement bandwidth from 1 Hz to the resonant frequency in water. The maximum and tip temperature increases (ΔT_{max} and ΔT_{tip}) are calculated assuming operation in water.

The temperature coefficient of tip deflection (TCTD) for the designs in Table 3.1 are calculated assuming that there is a 20 μm long step at the base of the force probe that is stiffened with a 1 μm thick Al film (i.e. the sensor-only probe designs). The TCTD of the actuated probes will be presented in Tables 3.2 and 3.3. The high surface phosphorus concentration of the piezoresistor and the tensile stress in the Al film (180 MPa) lead to static upward tip deflection (z_0).

Table 3.1: Summary of the force probe sensor designs. The sensor dimensions are specified by the sensor length (l_c), overall width (w_c) and piezoresistor length (l_{pr}). The air gap between the piezoresistor legs is 2 μm and an effective cantilever thickness of 290 nm in all cases (slightly thinner than the nominal 300 nm due to piezoresistor thinning during the fabrication process). The temperature increases and sensor resolution are calculated assuming water operation. The self-actuated force probes combine one of the sensor designs from this table with the corresponding actuator design from Tables 3.2 and 3.3.

Design #	l_c (μm)	w_c (μm)	l_{pr} (μm)	k_c (pN/nm)	f_0 (kHz)	f_d (kHz)	R_{pr} (k Ω)	ΔT_{\max} (K)	z_0 (μm)	TCTD (mm/K)	S_{xy} (V/m)	V_{noise} (μV)	MDD (mm)	MDF (pN)	
1	142	1	12.6	0.3	16.0	1.8	6.3	4.2	0	8.9	59	166	1.44	8.7	2.4
2	96	1	9.5	0.9	36.4	5.3	4.9	4.7	0	6.1	41	342	1.71	5.0	4.5
3	75	1	7.7	1.9	60.5	10.8	4.1	4.9	0.1	4.9	33	532	1.94	3.7	6.9
4	61	1	6.6	3.5	92.3	19.4	3.6	5.1	0.3	4.2	28	768	2.21	2.9	10
5	46	1	5.2	8.1	164	41.8	3.0	5.3	0.7	3.3	22	1250	2.67	2.1	17
6	35	1	4.1	18	284	85.3	2.5	5.3	1.4	2.6	18	1960	3.25	1.7	31
7	29	1	3.6	32	414	138	2.3	5.2	1.9	2.3	15	2690	3.80	1.4	46
8	192	2	17.4	0.2	8.7	0.8	4.5	3.5	0	12	79	87.4	0.82	9.4	2.1
9	128	2	13.6	0.8	20.3	2.7	3.6	4.0	0	8.0	54	185	0.99	5.4	4.0
10	98	2	11.1	1.7	35.3	5.8	3.1	4.2	0	6.3	42	297	1.14	3.8	6.5
11	80	2	9.4	3.1	53.4	10.1	2.7	4.4	0.1	5.2	35	422	1.30	3.1	9.5
12	45	2	5.9	17	171	45.0	1.9	4.7	0.8	3.2	22	1110	1.98	1.8	31
13	34	2	4.6	40	300	89.8	1.6	4.6	1.5	2.6	17	1710	2.49	1.5	58
14	28	2	4	72	442	143	1.5	4.5	2	2.2	15	2320	2.95	1.3	91

3.2 Actuator design

Before discussing the design of the thermal and piezoelectric on-chip actuators, it is worth considering the limitations of macroscale actuation (e.g. mounting a PCB on a piezoelectric stack). On-chip actuation has two clear advantages over external actuation: speed and the suppression of spurious modes. Both are related to the fact that on-chip actuation moves a much smaller mass than macroscale actuation. If a smaller mass is actuated, then the resonant frequency for a given displacement range can be increased (i.e. a reduction in the mass allows a higher resonant frequency for a given spring constant).

A second issue with macroscale actuators, related to the movement of a much larger mass, is that all of the resonant modes in the entire mounting assembly will be excited by the actuator. As the size of the mass being actuated decreases, so does the inertial load on the mounting assembly that holds the device in place. For example, high-speed AFM scanners based upon macroscale piezoelectric stack-based actuators have been designed with mirrored stacks and a dummy proof mass in order to balance inertial loads [72]. As the size of the mass being actuated decreases, the reaction force on the rest of the mounting assembly decreases. In the case of the on-chip actuators used in the force probes described in this thesis, the spurious modes from the rest of the mounting assembly are 3-4 orders of magnitude smaller than the force probe deflections.

The remainder of this section will focus on the design of thermal and piezoelectric on-chip actuators. The design requirements for the actuators are summarized in Table 1.1, with the two most important requirements being a minimum single-sided step response of at least 300 nm and a rise time of < 10 μ s.

3.2.1 Calculating tip deflection

The force probe tip deflection depends on the curvature generated by the actuator and the relative lengths of the actuator and sensor portions of the probe. The actuator curvature in general can be calculated from [349]

$$C = -\frac{\sum_i z_i A_i \sigma_i \sum_i E_i A_i + \sum_i A_i \sigma_i \sum_i z_i E_i A_i}{\sum_i E_i A_i \sum_i E_i (I_i + A_i z_i^2) - (\sum_i z_i E_i A_i)^2} \quad (3.12)$$

where z_i , E_i , A_i , I_i and σ_i are the distance from the bottom of the force probe to the midpoint of each film, the elastic modulus, the cross-sectional area, the moment of inertia and the average stress of layer i . If there is a large stress gradient within a layer (not applicable for the probes designed here) then it could be decomposed into multiple effective layers in order to improve the accuracy of the calculation. The actuator curvature is integrated twice in order to obtain the force probe deflection according to

$$\theta(x) = \int_0^x C(x)dx$$

and

$$v(x) = \int_0^x \theta(x)dx.$$

where θ is the angle, v is the deflection of the probe, and we take advantage of the fact that $\tan(\theta) \approx \theta$ for small angles. Although the curvature drops to zero beyond the end of the actuator, the sensor mechanically amplifies the actuator motion, part of the reason for placing the actuator at the base of the probe. An important parameter for actuator design is the placement of the neutral axis, calculated from

$$z_n = \frac{\sum_i z_i E_i A_i}{\sum_i E_i A_i}. \quad (3.13)$$

The tip deflection for a given actuator stress is maximized if the actuator is close to the neutral axis but does not cross it. Actuator design is similar to piezoresistor design in that the actuator should be on one side of the neutral axis (so that the stress generated by the actuator does not partially cancel out), but in contrast with piezoresistors, the actuator should be close to the neutral axis rather than far away from it.

3.2.2 Actuator dimensions

One of the main considerations in actuator design is determining the optimal length, width and thickness of the actuator. We will show in Section 3.3.1 that it is beneficial to have a relatively wide actuator in order to maximize the resonant frequency of the entire probe. We will consider several design nuances that are specific to the actuator

transduction mechanism in Sections 3.2.3 and 3.2.4. In this section we will provide a high-level overview of the actuator thickness and length selection.

Assuming that the actuator has a constant curvature ($C = 1/R$) along its entire length, the tip deflection at the tip of the cantilever is

$$v_{\text{tip}} = v_a + l_c \frac{\partial v_a}{\partial x} \Big|_{x=l_a} \quad (3.14)$$

$$\approx \frac{1}{2} C l_a^2 + C l_a l_c \quad (3.15)$$

where we have used the fact that $C \approx \partial \theta / \partial x \approx \partial^2 v / \partial x^2$ and that $\partial v / \partial x \approx \theta$. We also know that $f_0 \propto t_a / l_a^2$ in vacuum and that $C \propto 1/t_a$ for a unimorph actuator for a constant strain. We can combine all of these equations to find

$$v_{\text{tip}} \propto \frac{1}{f_0} + \frac{l_c}{\sqrt{t_a f_0}} \quad (3.16)$$

which implies that for a given resonant frequency, and assuming a fixed sensor length (l_c), the actuator should be as thin as possible with the length adjusted to achieve the desired resonant frequency.

But this analysis neglects two points. First, for piezoelectric actuation the curvature is actually proportional to $1/t_a^2$ for a constant bias voltage because the electric field scales as $1/t_a$ and the actuator strain is proportional to the electric field. Accordingly, a thin actuator is even more favorable for piezoelectric actuation if we neglect the dependence of the d_{31} coefficient on actuator thickness.

Second, there are several competing design considerations that lead to a slightly thicker actuator design. As the actuator becomes thinner, it also must become shorter. However, the electrical contacts between the piezoresistor and the metal interconnects are formed beyond the end of the actuator, which adds length to the actuator (on the order of 20 μm for a contact resistance of $< 200 \Omega$). Additionally, the active portion of the actuator (excluding the contact region) needs to be at least 20 μm long in order to make it less sensitive to the backside release etch alignment and edge roughness (Appendix F). Finally, specific to the thermal actuator, as the actuator becomes shorter the resistive heater comes closer to the

piezoresistive sensor, which increases thermal crosstalk between the actuator and sensor. In practice, these requirements led to actuators at least 40 μm in length with the actuator thickness chosen in order to balance the tradeoffs between actuator and overall force probe performance.

3.2.3 Thermal actuator

The thermal actuator consists of a silicon heater, SiO_2 etch stop and a layer of Al (Figure 3.1). The heater increases the temperature of the actuator, deflecting the probe tip due to the mismatch in thermal expansion coefficients between Si and Al. Thermal actuator design is a tradeoff between tip deflection, operating temperature and heater time constant. For example, a longer actuator enables larger tip deflections for a given temperature but increases the heater time constant.

The thermally actuated probe is modeled as a three layer system (3.12) with Si ($i = 1$), SiO_2 ($i = 2$) and Al ($i = 3$) layers. The stress change in each layer of the thermal actuator is calculated from

$$\Delta\sigma_i(x) = \Delta T(x)\alpha_i E_i \quad (3.17)$$

where $\Delta T(x)$ is the the temperature change of the actuator a distance x from its base and α_i is the coefficient of thermal expansion (CTE) of layer i . Note that each layer may have some initial stress but that the stress change during actuation is the parameter of interest. The CTEs of the Si, SiO_2 and Al layers are assumed to be 2.6, 0.5 and 23.1 ppm/K [304].

The tip deflection is maximized for a given temperature increase by choosing the SiO_2 and Al thicknesses to place the neutral axis at roughly the Si- SiO_2 interface. The SiO_2 thickness (100 nm) is chosen to minimize the thermal resistance between the heater and Al while providing an adequate etch stop for the reactive ion etch (RIE) based Al patterning step during the fabrication process. The thermal resistance between the heater and the Al is roughly 10 times smaller than the other thermal resistances in the system.

The optimal Al actuator thickness, based purely upon actuation efficiency, would be about 250 nm (Figure 3.2). However, as we discussed in the last section (Section 3.2.2), a thicker actuator is required for a variety of other design considerations. We balanced

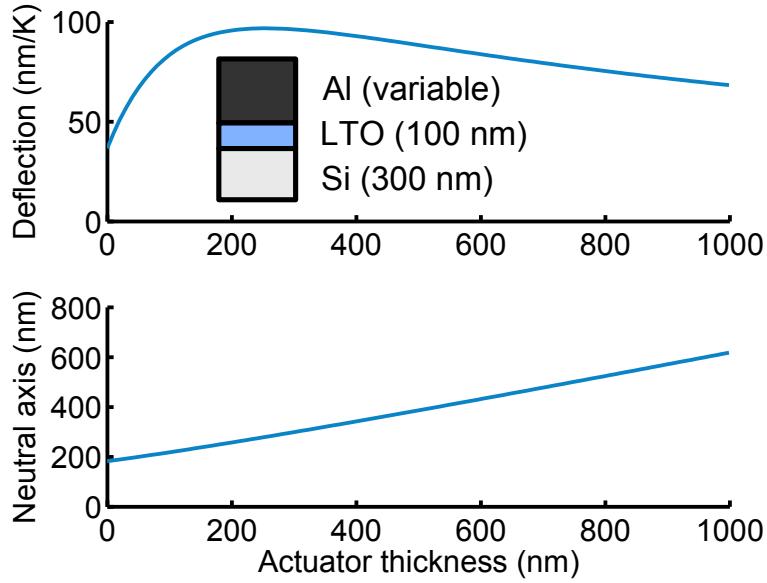


Figure 3.2: Thermal actuator thickness selection. The tip deflection (top) and the distance between the bottom of the device and the neutral axis (bottom) are plotted as a function of the thickness of the Al actuator. A 250 nm thick Al actuator would yield the maximum tip deflection and optimal neutral axis placement. However, a thin actuator would need to be impractically short for the various reasons discussed in the text. The final devices were designed around a 1 μm thick Al actuator.

these requirements by choosing an actuator length of 50-60 μm and a 1 μm thick Al film, yielding a tip deflection of 100-300 nm/K depending on the sensor length. The data in Figure 3.2 are calculated for a 61 μm long sensor at the end of a 60 μm long actuator.

Similarly, the overall width of the actuator section (20 μm) and the width of the heater (12 μm) are chosen to provide room for 2 μm wide interconnects to the piezoresistor with 2 μm wide gaps.

The steady-state temperature over the thermal actuator is calculated from the finite difference-based model presented in Section 2.6.3 and Ref. [302]. An example of the predicted steady-state response of a force probe is presented in Figure 3.3. The actuator voltage determines the temperature near the base of the force probe while a second peak in the temperature profile is generated by the piezoresistive sensor.

The dynamic response of the actuator is calculated using a combination of lumped parameter and FEA models. The lumped parameter model treats the actuator as first order

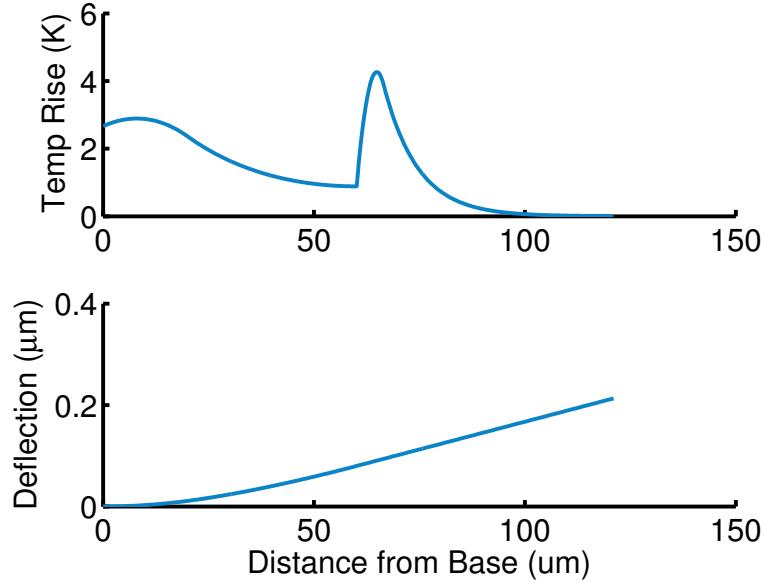


Figure 3.3: Example of the thermal actuator steady-state response. As in many of the other actuator examples, we use a 61 μm long sensor at the end of a 60 μm long actuator. The temperature profile (top) and the resulting change in deflection during actuation (bottom) are presented for a probe operated in water with Wheatstone bridge and actuator bias voltages of 1 V. The tip deflection scales linearly with the power dissipation. Initial deflection from intrinsic stress is included in the cantilever model but removed in this example.

system with an effective thermal resistance and capacitance. The overall thermal resistance (R_a) from the actuator to the ambient temperature is calculated from solid-state conduction and convection resistances operating in parallel (R_{cond} and R_{conv}) while the heat capacity (C_a) is calculated from the thermal actuator volume. Radiated power is negligible because the devices operate near room temperature [302]. The actuator working equations are

$$R_{\text{cond}} \approx \frac{l_h}{2w_a \sum_i t_i k_i} + R_{\text{base}} \quad (3.18)$$

$$R_{\text{conv}} \approx \frac{1}{2h_{\text{eff}} l_a w_a} \quad (3.19)$$

$$C_a \approx l_a w_a \sum_i t_i \rho_i c_i \quad (3.20)$$

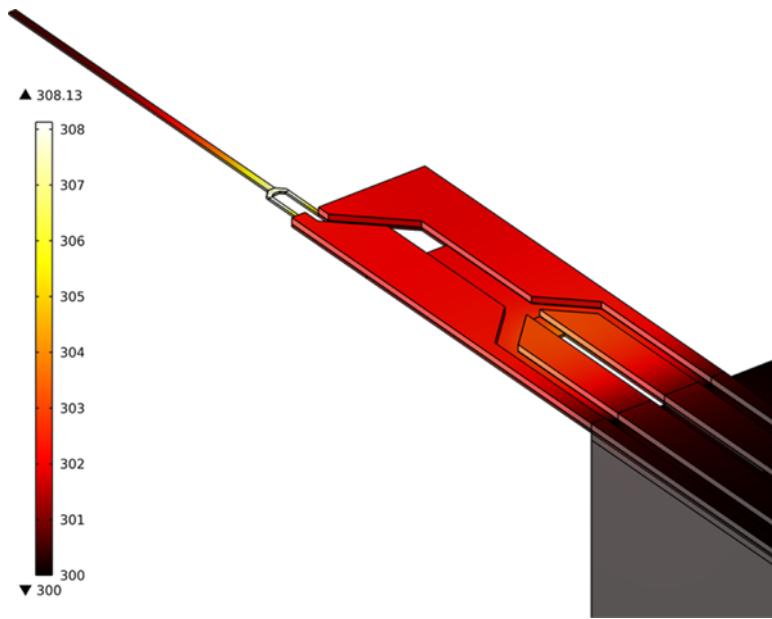


Figure 3.4: Steady state temperature distribution of a thermally actuated probe operating in air. The actuator is biased to 1 V and the sensor is operated at a bridge bias of 2 V. The temeprature at the base of the probe is slightly higher than the ambient temperature of 300 K, illustrating the finite thermal resistance at the base of the probe.

where w_a and l_a are the overall dimensions of the actuator, l_h is the length of the silicon heater, t_i , k_i , ρ_i and c_i are the thickness, thermal conductivity, density and specific heat capacity of each layer in the actuator and h_{eff} is the effective convection coefficient. The finite thermal resistance between the base of the actuator and the ambient environment (R_{base}) is calculated as 1.4 and 1.1 K/mW in air and water, respectively, from a three-dimensional conduction-convection finite element analysis (FEA) model (Figure 3.4). We assume $h_{\text{eff}} = 2,000$ and $50,000 \text{ W/m}^2\text{-K}$ in air and water (Section 2.6.3) while the thermal conductivities of the silicon and aluminum layers are taken as 100 and 200 W/m-K, respectively [303, 306]. Other material properties are taken from the literature [304].

The lumped parameter model can be used to predict the actuator time constant ($\tau = R_a C_a$), -3 db frequency ($f_{-3\text{dB}} = 1/2\pi\tau$) and 10-90% rise time ($t_r \approx 1/3f_{-3\text{dB}}$). The predicted time constant, -3 dB frequency and rise time in air are 13 μs , 13 kHz and 26 μs while in water they are 8.5 μs , 19 kHz and 17 μs , respectively. In comparison, the three-dimensional FEA model predicts rise times in air and water of 71 and 24 μs .

Table 3.2 summarizes the thermal actuators presented in this thesis. The actuators are paired with the corresponding sensor designs from Table 3.1. The sensor design, operating conditions and performance are affected negligibly by the addition of the actuator at the base of the probe. For each design only the overall length of the actuator varies (l_a). The overall actuator width remains fixed at 20 μm for all of the designs. The silicon heater is a resistor loop with 20 μm long, 5 μm wide legs and a 2 μm air gap between them. The heater dimensions determine the thermal time constant of the system. The remaining 8 μm of the actuator width are occupied by the 2 μm wide piezoresistor interconnects and 2 μm gaps.

Due to the fixed heater dimensions, the electrical resistance of the heater does not vary between the designs and is defined by the diffusion process. The electrical resistance of the heater varies from approximately 200 to 1000 Ω depending on the sheet resistance of the fabrication process. The thickness of the SiO_2 and Al films on the thermal heater also remain constant at 100 and 1000 nm, respectively.

The tip deflection of the probe is calculated with respect to the average actuator temperature (A_{xk}) and with respect to the heater power (A_{xw}), which varies depending on the thermal properties of the surrounding fluid. The actuator response is calculated assuming that the corresponding sensor design is on the end of the actuator. The variation in tip position with ambient temperature matches A_{xk} to within a few percent; an efficient thermal actuator is inherently sensitive to temperature fluctuations. The static tip deflection (z_0) due to the high phosphorus surface doping of the piezoresistor and the tensile stress in the Al is larger than in the sensor-only case due to the addition of the actuator.

3.2.4 Piezoelectric actuator

The piezoelectric actuator consists of a silicon beam, SiO_2 etch stop, AlN seed layer, bottom Mo electrode, AlN actuator and top Mo electrode (Figure 3.1). The AlN seed layer improves the grain alignment and resulting piezoelectric properties of the AlN actuator [350–353]. The Mo electrodes are used to apply an electric field to the AlN actuator, which transduces the electric field into lateral strain which deflects the tip of the force probe.

We opted to use AlN rather than ZnO [57], PZT [65] or PMN-PT [354] for maximum

Table 3.2: Summary of the second generation PRT actuator designs. Each actuator design is paired with the corresponding sensor design (Table 3.1. For each design, the actuator length (l_a), the force probe tip deflection with respect to temperature (A_{xk}) and power dissipation (A_{xw}), and the initial tip deflection (z_0) are presented.

Design #	l_a (μm)	A_{xk} (nm/K)	A_{xw} (air) (nm/mW)	A_{xw} (water) (nm/mW)	z_0 (μm)
1	60	220	2200	310	33
2	60	160	1600	220	24
3	60	130	1300	180	20
4	60	120	1100	160	18
5	60	98	950	140	15
6	55	74	730	110	11
7	50	57	580	95	9
8	60	280	3100	520	42
9	60	200	2200	380	31
10	60	160	1800	310	25
11	60	140	1500	270	21
12	60	95	1050	190	14
13	55	72	790	150	11
14	50	55	630	120	9

material compatibility with the processing equipment in the Stanford Nanofabrication Facility (SNF). We developed a process for AlN on Ti from 2008-2010 in collaboration with Tango Systems [355], but opted to use Mo electrodes for the devices that are the main subject of this thesis.

We chose to use Mo rather than Ti for three reasons. First, we developed the Ti-based process due restrictions on our ability to process Mo through the SNF. By the time of the second fabrication run in 2010-2011 the SNF contamination policies had loosened enough for us to very carefully process Mo through the SNF. Second, Mo is a superior electrode material for AlN-based actuators and sensors. It yields better AlN grain alignment, has superior electrical and thermal conductivity, and most importantly, can be integrated into a chlorine/fluorine-based dry etch process [356, 357]. Third, we found a vendor (OEM Group) with a well developed process for sub-100 nm thick AlN on Mo with good stress control and film thickness uniformity as demonstrated in Ref. [358]. Additionally, OEM Group had experience depositing high quality piezoelectric actuators on LPCVD SiO₂ from their collaborations on AlN-based resonators with the Pisano group at UC Berkeley and the

Piazza group at the University of Pennsylvania [359].

The piezoelectric actuator probe is modeled as a six layer system (3.12) with Si ($i = 1$), SiO_2 ($i = 2$), the AlN seed layer ($i = 3$), the Mo bottom electrode ($i = 4$), the AlN actuator ($i = 5$) and the Mo top electrode ($i = 6$) layers. The change in stress from piezoelectric actuation in each layer is calculated from

$$\Delta\sigma_i(x) = \frac{d_{31,i}E_i\Delta V_i}{t_i} \quad (3.21)$$

where $d_{31,i}$, ΔV_i and t_i are the transverse piezoelectric coefficient, voltage drop and thickness of layer i . None of the properties vary along the length of the actuator, in contrast with the thermal actuator. The piezoelectric coefficient is zero for each of the layers except for the AlN seed layer ($i = 3$) and AlN actuator ($i = 5$). The voltage drop across the AlN seed layer is negligible, while the voltage drop across the AlN actuator layer is V_a . During experiments, if the piezoelectric actuator is being used to apply a step displacement to a sample then the actuator is typically prebiased while the force probe is brought into contact with the sample (e.g. $V_a = -8$ V) and swung to the opposite polarity (e.g. $V_a = 8$ V) in order to maximize the range of the actuator within the voltage breakdown limits of the AlN actuator.

As the thickness of the piezoelectric actuator (t_a) decreases, the actuator stress for a given voltage increases. A reduction in V_a is desirable because it reduces the requirements on the actuator drive circuit and reduces capacitive crosstalk between the actuator and the piezoresistive sensor and (if being used to study hair cells) between the actuator and the patch clamp electrode. However, there are two competing effects. First, as the thickness of the AlN actuator decreases, so does the uniformity of the AlN grain alignment and the effective d_{31} coefficient [360]. Second, the location of the neutral axis shifts further from the actuator, resulting in less tip deflection for a given strain.

For the piezoelectric actuator design we initially assumed an effective d_{31} coefficient of -1.9 pm/V based upon several reports of approximately 100 nm thick AlN actuators deposited on Pt or Mo [361, 362]. However, the performance of the fabricated devices implied a d_{31} coefficient closer to -0.9 pm/V. There are a variety of possible reasons for the smaller than expected piezoelectric coefficient based upon the deposition conditions (e.g.

the films were deposited with a compressive stress on the order of 200 MPa in order to cancel other stresses in the film stack), the substrate roughness (e.g. the silicon underneath the LPCVD SiO₂ is highly doped in order to minimize electrical crosstalk, which may have affected the roughness of the SiO₂ layer underneath the actuator) or the fabrication process (e.g. undercut of the actuator stack during processing).

Table 3.3 summarizes the piezoelectric actuators presented in this thesis. The actuators are paired with the corresponding sensor designs from Table 3.1. The sensor design, operating conditions and performance are affected negligibly by the addition of the actuator at the base of the probe.

For each design only the overall length of the actuator varies (l_a). The overall actuator width and piezoelectric film stack are fixed at 20 μm and 12 μm respectively. The gap between the end of the actuator and the piezoelectric film stack is 22 μm for all of the designs in order to provide room for the piezoresistor contacts. The electrical time constants for all of the piezoelectric actuators are less than 0.1 μs , much faster than the force probe mechanical response. Actuator performance is calculated for an LTO/AlN/Mo/AlN/Mo stack of 80/25/75/75/50 nm and assuming a d_{31} coefficient of -0.9 pm/V.

The breakdown voltage of the 75 nm thick actuator is approximately 10 V, allowing a safe actuator swing of about 16 V (i.e. -8 V to +8 V). The actuator response (A_{xv}), temperature coefficient of tip deflection (TCTD) and initial tip offset from residual stress (z_0) are calculated for each design. The actuator response is calculated assuming that the corresponding sensor design is on the end of the actuator. The TCTD is significantly smaller than for the thermally actuated probes because of the small mismatch between the coefficients of thermal expansion of Si (2.6 ppm/K) and Mo (4.8 ppm/K) compared with Al (23.1 ppm/K). The length of the actuator decreases along with the sensor in order to ensure that the actuator and sensor are mechanically decoupled.

3.3 Sensor and actuator integration

We have discussed the design of the sensor and actuator independently of each other until now. In combining the sensor and actuator, the most important issue is whether the actuator or sensor should be at the base of the force probe. Until now we have assumed that

Table 3.3: Summary of the second generation PRPE actuator designs. Each actuator design is paired with the corresponding sensor design (Table 3.1. For each design, the actuator length (l_a), the force probe deflection with respect to voltage (A_{xv}), the temperature coefficient of tip deflection (TCTD) and the initial tip deflection (z_0) are presented.

Design #	l_a (μm)	A_{xv} (nm/V)	TCTD (nm/K)	z_0 (μm)
1	60	138	43	4.4
2	60	103	32	3.1
3	60	87	27	2.7
4	60	77	23	2.5
5	50	46	15	1.6
6	45	31	11	1.1
7	38	19	8	0.7
8	60	176	56	5.7
9	60	127	40	4.2
10	60	105	32	3.2
11	60	91	28	2.7
12	50	45	15	1.5
13	45	31	11	1.1
14	38	18	8	0.7

the actuator would be placed at the base of the probe, but now will justify that decision. Fortunately, the decision is fairly clear because there are numerous benefits to placing the actuator rather than the sensor at the base of the probe and there are no clear downsides (Figure 3.5).

There are several benefits to placing the actuator at the base. First, the actuator is inherently going to be stiffer and more massive than the sensor due to the addition of the actuator films. Accordingly, placing the actuator at the base of the probe maximizes the resonant frequency of the first force probe eigenmode for given actuator and sensor designs. Second, both piezoelectric and thermal actuators are displacement-based actuators (in contrast with electromagnetic actuators, which are force-based). The actuator generates a curvature that is integrated along the length of the force probe in order to generate a tip deflection. The maximum tip deflection range determines the maximum force that can be applied to a sample, equal to the product of the deflection and spring constant of the sensor. If the actuator is at the base of the probe then the additional length of the sensor mechanically amplifies the actuator motion, resulting in larger possible tip deflections and

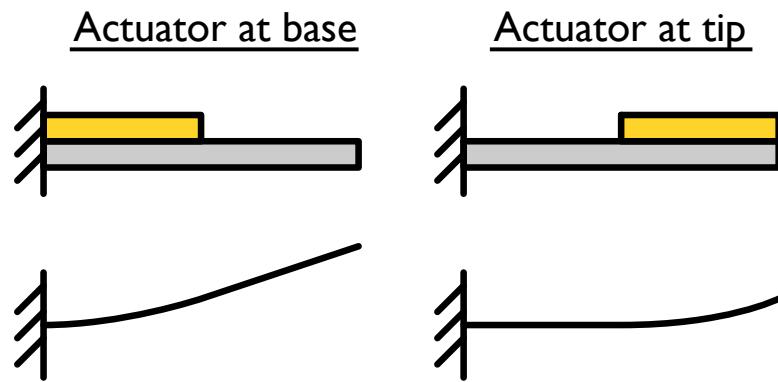


Figure 3.5: Relative placement of the actuator and sensor. Placing the actuator at the base of the force probe (left) rather than at the tip (right) maximizes the tip deflection because the actuator curvature is mechanically amplified by the sensor. Additionally, the actuator is significantly stiffer and more massive than the sensor, so placing the actuator at the base results in greater measurement bandwidth and ensures that both the effective spring constant and bandwidth are set solely by the sensor-portion of the probe.

forces. Third, the sensor needs to be made as thin and narrow as possible. If the actuator extends beyond the end of the sensor then the actuator films will need to run alongside the sensor, increasing its spring constant and reducing the stress in the piezoresistor for a given tip force.

The fact that the actuator and sensor are mechanically decoupled allows them to be designed independently of one another for the most part. The addition of the actuator to the base of the sensor only increases the maximum and tip temperatures of the sensor slightly (< 0.5 K) because the actuator presents a fairly low thermal resistance compared to the sensor. The sensor affects the actuator performance slightly due to the Al interconnects that run alongside the active component of the actuator (particularly for the piezoelectric actuator), but the reduction in actuation amplitude is slight and integrated into the actuator design.

3.3.1 Mechanics of the combined system

Although intuitively we expect the sensor and actuator to be mechanically decoupled, we can examine the overall mechanical system in more detail in order to validate this assumption. In order to model the combined system we could turn exclusively to FEA,

but there are two good reasons to develop a reasonably accurate analytical model.

First, analytical models can be computed 3-4 orders of magnitude more quickly than FEA models (e.g. 0.1 ms vs. 60 seconds). This enormous difference in speed allows quicker design iterations. Second, developing an analytical model makes the design knobs available at your disposal abundantly clear. The impact of each parameter (e.g. actuator width) on the system performance comes naturally from analytical development while relying exclusively on FEA lends itself toward guess-and-check design. FEA is valuable at the end of the design process, but after working on cantilever design for a few years I view FEA as a final design validation step, not as a design tool.

The mechanics (i.e. spring constant and resonant frequency) of the sensor or actuator can be modeled fairly easily independently of one another via Bernoulli beam analysis. But the system becomes significantly more complicated once the width, material properties and thickness of the beam start to vary along its length. Consider a beam with two sections: a thick section near the base (the actuator), and a thin section near the tip (the sensor). One might be tempted to model the beam as two springs in series to find the stiffness and then to model the resonant frequency by coming up with an effective mass for the system. Unfortunately, this approach does not work.

Spring constant

In modeling the stiffness, there are two intuitive reasons why the springs-in-series model breaks down. First, the force is applied at the tip of the sensor, so the moment seen by the actuator is larger than if the force was applied at the tip of the actuator. When you model the system as springs in series, you are modeling the system using forces and not moments, so there is no amplification effect. Second, the force causes the actuator to bend, leading to a non-zero angle at the tip of the actuator. This means that the sensor is initially at a slight angle, so the deflection at the tip of the sensor increases by the initial slope integrated along the length of the sensor.

The spring constant of the combined system can be modeled accurately by applying a virtual force to the tip of the sensor. The textbook equation for the tip deflection of a fixed-free prismatic, homogenous, linear elastic cantilever beam is

$$v(x) = -\frac{Fx^2(3(l_a + l_c) - x)}{6EI}. \quad (3.22)$$

where x is the distance from the base of the probe while l_a and l_c are the lengths of the actuator and sensor.

We can relax the prismatic and homogenous assumptions by treating the cantilever in sections. We'll assume that the cantilever has two sections. The actuator section is at the base and is composed of an arbitrary stack of multiple films. The width of the actuator section can be different from the sensor section. We can calculate the effective EI for the two sections from

$$EI_a = \sum_i E_i [I_i + A_i(z_n - z_i)^2] \quad (3.23)$$

$$EI_c = \frac{E_c w_c t_c^3}{12} \quad (3.24)$$

where EI_a and EI_c are the bending rigidities for the actuator and sensor sections, respectively, z_n is the distance from the bottom of the device to the neutral axis, and z_i is the distance from the bottom of the device to the midpoint of each film in the system.

By applying a test force, we can calculate the overall stiffness of the beam from

$$k_c = \frac{F}{v_{tip}} \quad (3.25)$$

where

$$v_{tip} = v_a + l_c \left. \frac{\partial v_a}{\partial x} \right|_{x=l_a} + v_c \quad (3.26)$$

where v_a is the relative deflection of the actuator, $\partial v_a / \partial x$ is the slope at the end of the actuator and v_c is the relative deflection of the sensor.

If we assume a point load force F applied at the tip of the sensor, then the relative deflection and slope along the length of the actuator are

$$v_a(x) = \frac{-Fx^2[3(l_a + l_c) - x]}{6EI_a} \quad (3.27)$$

$$\frac{\partial v_a(x)}{\partial x} = \frac{-Fx[6(l_a + l_c) - 3x]}{6EI_a} \quad (3.28)$$

Similarly, the relative deflection of the sensor can be calculated from

$$v_c(x) = \frac{-F(x - l_a)^2(3l_c - (x - l_a))}{6EI_c} \quad (3.29)$$

where we have maintained the coordinate system of x denoting the distance from the base of the actuator. The deflection at the tip ($x = l_c + l_a$) can be calculated by combining (3.27), (3.28) and (3.29) and inserted into (3.25) to calculate the spring constant of the combined actuator-sensor system as

$$k_c = \left[\frac{l_a^2(2l_a + 3l_c) + 3l_a l_c(l_a + 2l_c)}{6EI_a} + \frac{l_c^3}{3EI_c} \right]^{-1}. \quad (3.30)$$

Figure 3.6 compares the spring constant of the entire force probe with just the sensor portion of the probe for a 61 μm long, 1 μm wide sensor (design 4 in Table 3.1) with a 60 μm long, 20 μm wide piezoelectric actuator (design 4 in Table 3.3) at its base. The spring constant of the force probe decreases by about 4% due the addition of the actuator, which is negligible compared with the device fabrication tolerances. The small deviation (< 2%) between the FEA and analytical model is due to additional device design features that are not captured by the simplified model. In particular, the legs of the piezoresistor are tapered in order to avoid current crowding and the end of the piezoresistive loop is twice the piezoresistor leg width in order to minimize the sensitivity degradation from the loop for short piezoresistors. Figure 3.6 confirms that the spring constant of the force probe primarily depends on the sensor design, simplifying the overall probe design.

Resonant frequency

The resonant frequency of the force probe can be approximated by using the Rayleigh-Ritz method [42]. The Rayleigh-Ritz method can be used to calculate the resonant frequency

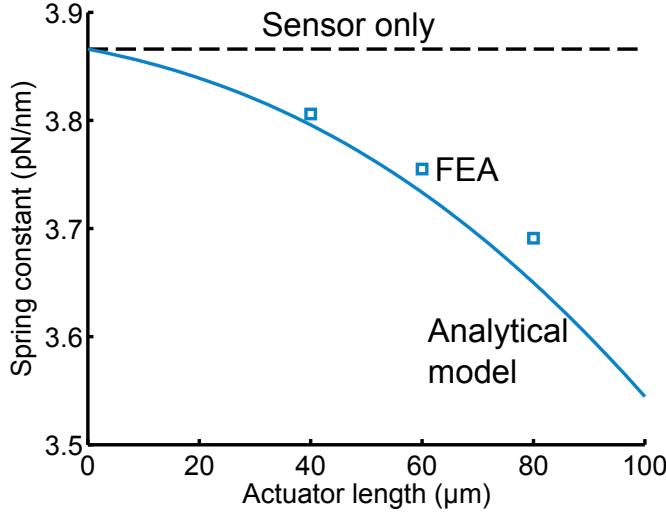


Figure 3.6: Force probe spring constant as a function of actuator length. The spring constant of the 61 μm long, 1 μm wide sensor is plotted for comparison. The overall force probe spring constant decreases by about 4% from the sensor-only spring constant when a 60 μm long, 20 μm wide piezoelectric actuator is added at its base. Finite element analysis (FEA) results agree with the analytical model.

of an arbitrary mechanical structure from the fact that the kinetic energy and elastic strain energy are equal in magnitude (and 90 degrees out of phase) on resonance. Three steps are required to calculate the resonant frequency:

1. Calculate the approximate mode shape of the beam
2. Calculate the kinetic and strain energy amplitudes
3. Find the frequency where the energies are equal

If we assume that the cantilever deflection as a function of time is $v(x,t) = v(x)e^{i\omega t}$, then its maximum velocity is $v(x)\omega$ and we can calculate the total kinetic energy of the beam from

$$U_k = \int_0^{l_a} \frac{1}{2} [\omega v(x)]^2 w_a (t_c \rho_c + t_a \rho_a) dx + \int_{l_a}^{l_c} \frac{1}{2} [\omega v(x)]^2 w_c t_c \rho_c dx. \quad (3.31)$$

where ρ_a is the average density of the actuator stack. The mode shape is approximated from (3.27) and (3.29). The strain energy is calculated from

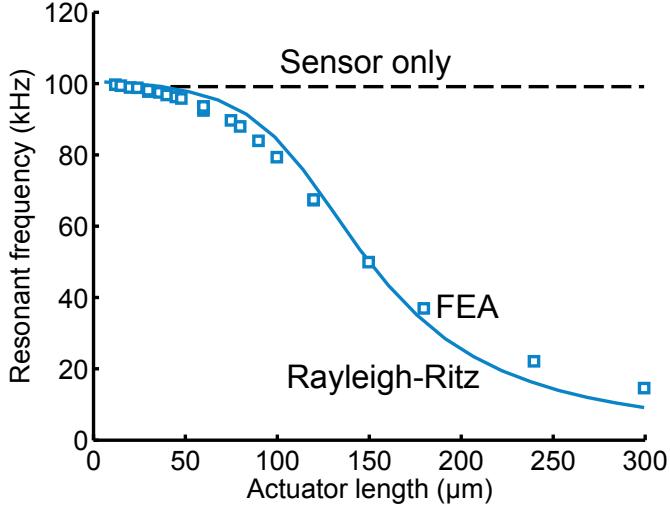


Figure 3.7: Force probe resonant frequency as a function of actuator length. The resonant frequency for a 61 μm long, 1 μm wide sensor attached to the end of a 60 μm long, 1 μm thick thermal actuator is calculated via the Rayleigh-Ritz method (solid line) and from a three-dimensional FEA model (squares). The resonant frequency drop in vacuum from the sensor-only resonant frequency (dashed line) is about 7%.

$$U_s = \int_0^{l_a} \frac{M(x)^2}{2EI_a} dx + \int_{l_a}^{l_c} \frac{M(x)^2}{2EI_c} dx. \quad (3.32)$$

The moment along the length of the beam is unaffected by the presence of the actuator, and is $M(x) = F(l_c + l_a - x)$. By equating the kinetic and strain energies, $U_k(\omega) = U_s$, we can solve for ω to find the resonant frequency of the structure. The resonant frequency calculation is performed numerically and is included in the cantilever optimization and modeling code (Appendix H).

A comparison between the Rayleigh-Ritz method and a three-dimensional FEA model is presented in Figure 3.7. The resonant frequency for a 61 μm long, 1 μm wide sensor (design 4 in Table 3.1) attached to the end of a 60 μm long, 1 μm thick thermal actuator (design 4 in Table 3.2) is modeled using the two methods.

The agreement between the Rayleigh-Ritz and FEA results is excellent. The resonant frequency drop in vacuum from the sensor-only resonant frequency (dashed line) is about 7%. In water the reduction in resonant frequency is negligible because fluid damping of the sensor portion rather than the actuator portion dominates the overall frequency response.

An important result that comes from the Rayleigh-Ritz analysis is that as the width of the actuator increases, its impact on the resonant frequency decreases. This result does not hold for a homogenous, prismatic beam and was initially unexpected. As the width of the actuator increases it deflects less and its kinetic energy at a given frequency decreases, increasing the overall resonant frequency of the system.

To conclude, this section confirms that the force probe actuator can be designed to have a negligible impact on the frequency of its first eigenmode.

3.3.2 Higher order eigenmodes

The Rayleigh-Ritz analysis only considers the first eigenmode, however. If higher order modes are excited during force probe actuation then the force measured by the piezoresistor might not correspond to the force actually delivered to the sample (e.g. due to lateral or torsional oscillations). The most effective way to investigate higher order eigenmodes is through FEA.

The first five eigenmodes of a piezoelectrically actuated force probe are shown in Figure 3.8. The force probe design is identical to the one that we considered for the spring constant and Rayleigh-Ritz analyses (i.e. a 61 μm long sensor and 60 μm long actuator). The modes were calculated using the solid mechanics module in Comsol 4.2 and using the material properties assumed throughout the thesis. The first mode (98 kHz) corresponds to the first flexural mode of the sensor. The second mode (270 kHz) is the first flexural mode of the actuator. The third mode (422 kHz) is the transverse flexural mode of the sensor. The fourth mode (637 kHz) is the second flexural mode of the sensor. Finally, the fifth mode (1150 kHz) is the first torsional mode of the actuator and sensor. The results show that undesirable higher order modes (e.g. the third eigenmode) can be suppressed by filtering the stimulus delivered to the actuator.

3.3.3 Actuator-sensor crosstalk

The major downside in bringing the force probe actuator on-chip is increased crosstalk between the actuator and sensor. Simultaneous actuation and sensing requires that any crosstalk signal from the actuator be below the noise floor of the sensor. Crosstalk,

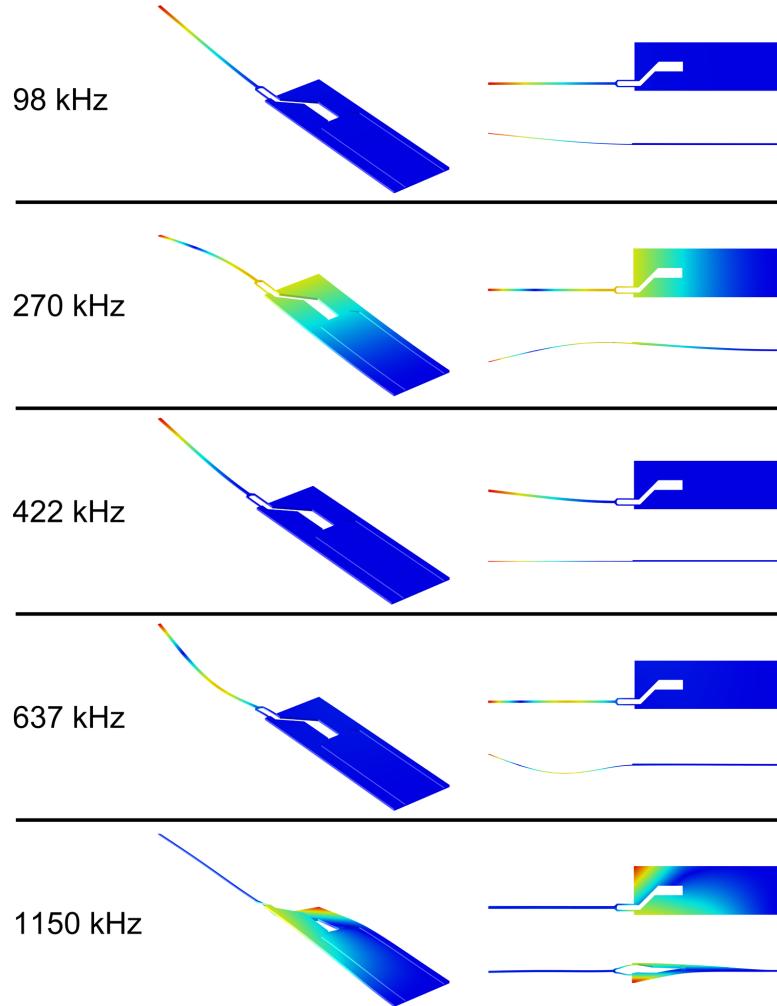


Figure 3.8: First five eigenmodes of a piezoresistive force probe with on-chip piezoelectric actuation (PRPE) calculated via finite element analysis. Shading indicates the absolute deflection of the device from its initial position. The simulation assumes a 300 nm thick, 1 μm wide and 61 μm long sensor attached to the tip of a 20 μm wide and 60 μm long actuator. The first flexural mode of the sensor (98 kHz) is well separated from higher order modes, allowing higher order mode suppression by filtering the actuator stimulus signal.

particularly capacitive crosstalk, has historically been one of the major challenges in the development of scanning probes with on-chip actuation.

For example, Manalis et al. fabricated piezoresistive cantilevers with on-chip ZnO-based piezoelectric actuation [57]. However, the capacitive crosstalk between the actuator and sensor was on the order of -70 dB at 10 kHz. Driving the ZnO actuator at 10 kHz with

a 1 V sine wave would have introduced a spurious signal in the readout circuit roughly 300-times the noise floor of the system). Manalis et al. got around this problem by biasing the piezoresistor at 130 kHz and monitoring the piezoresistor signal using a lock-in amplifier with a 6 kHz measurement bandwidth.

Lock-in based readout has two limitations. First, it can only eliminate capacitive crosstalk and not mechanical or thermal crosstalk, because the latter two are due to fluctuations in the piezoresistor resistance and are not spectrally isolated from the signal of interest. Second, modulating the bridge bias voltage and demodulating the amplifier output is nontrivial for a measurement bandwidth of 100 kHz, although feasible in theory. For example, the AD630 modulation/demodulation IC (2 MHz bandwidth, 100 dB of dynamic reserve) could be combined with the INA103 (800 kHz bandwidth) in order to modulate the Wheatstone bridge bias at several hundred kHz and low-pass filter both the actuator drive signal and the AD630 output at 100 kHz. However, this thesis demonstrates that capacitive crosstalk can be driven below the mechanical crosstalk floor up to at least 100 kHz, so the benefit of AC modulation is limited.

Alternatively, Kim et al. reported a PZT-actuated piezoresistive cantilever in 2003 that modified the device layout to reduce capacitive and mechanical crosstalk [65]. While Manalis et al. used doped silicon interconnects that ran underneath the actuator, Kim et al. used metal interconnects that ran alongside it (Figure 1.23). They also reported a detailed electrical model for capacitive crosstalk and achieved capacitive crosstalk of less than -80 dB at 10 kHz.

As a rule of thumb, capacitive crosstalk of less than -120 dB enables simultaneous actuation at 1 V and sensing with the capacitive crosstalk signal below the piezoresistor noise floor (on the order of 1 μ V). The cut off for mechanical and thermal crosstalk is less clear because it depends on the magnitude of the force being measured, which depends on the stiffness of the sample. Our aim then is to improve on the capacitive crosstalk reported in prior work by 40 dB.

The general approach taken here was to measure the piezoresistor signal differentially, using an identical force probe and instrumentation amplifier with excellent common mode rejection to cancel out the crosstalk signal. A temperature compensation resistor is an essential part of any piezoresistive readout circuit, and the concept is simply extended to

compensate for spurious signals generated by the on-chip actuator. There are three major sources of crosstalk: mechanical, electrical and thermal.

Mechanical crosstalk

Recall that the output of a Wheatstone bridge can be calculated from

$$\Delta V_{\text{out}} = \frac{1}{4} \frac{\Delta R}{R} V_{\text{bridge}}. \quad (3.33)$$

The main downside of placing the sensor beyond the end of the actuator is that the sensor interconnects must pass adjacent to the actuator. The curvature generated by the actuator induces a bending stress in the force probe which can be transduced into a resistance change by the interconnects.

Beyond the use of a compensation force probe and utilizing common mode rejection, there are three specific design strategies that can be used to minimize mechanical crosstalk. First, the interconnects should be fabricated from a material with a low gauge factor (e.g. Al) rather than a high one (e.g. Si). Second, the resistance of the interconnects should be as small as possible compared to the piezoresistor resistance so that the magnitude of the interconnect change in resistance is relatively small. Third, the interconnects should be placed as near as possible to the neutral axis of the force probe in order to minimize the stress that they see. Ideally the interconnects would span both sides of the neutral axis so that the average stress that they experience is negligible.

The force probes (Section 3.1) use relatively thick Al interconnects (400 and 1000 nm for the piezoelectrically and thermally actuated probes, respectively). The interconnects are on one side of the neutral axis due to limitations in the fabrication process.

Electrical crosstalk

In contrast with mechanical crosstalk, which is due to a resistance change in the interconnects and depends on the magnitude of the Wheatstone bridge bias (Figure 5.45), electrical crosstalk directly generates a spurious voltage signal in the measurement circuit output.

There are two types of electrical crosstalk: direct and capacitive. Direct crosstalk would result from direct electrical coupling between the actuator and sensor. For example, if the pn junction isolation between the thermal actuator and piezoresistive sensor broke down then the actuator would directly inject electrical current into the sensor circuit. Fortunately, this did not prove to be an issue with the force probes because the electrical isolation between the actuator and sensor is much larger than the piezoresistor impedance. For example, the back-to-back pn junctions on the thermally actuated probes yields an impedance of at least $50\text{ M}\Omega$ and the oxide underneath the piezoelectric actuator stack yields an impedance of greater than $1\text{ G}\Omega$ compared with a piezoresistor impedance on the order of $2\text{-}5\text{ k}\Omega$.

Capacitive crosstalk is the far larger problem. Variation in the actuator voltage injects a capacitive current into the sensor circuit, and the piezoresistor impedance transduces the electrical current into a voltage. Capacitive crosstalk occurs through both the substrate underneath the actuator and through the narrow gap separating the actuator and interconnects (Figure 3.1).

Two approaches were taken to minimize capacitive crosstalk through the substrate. First, the silicon underneath the actuators is heavily doped and grounded in order to provide a low impedance path to ground. Second, the silicon at the edge of this isolation diffusion well is etched away in order to decrease the capacitance between the n-type well and the p-type substrate. Mismatch in the lateral capacitance between the actuator and interconnects between the main and compensation probes (recall that only the mismatch matters) is inevitable, and we will discuss circuit-level compensation techniques in Section 5.6.

Thermal crosstalk

Doped silicon resistors are excellent temperature sensors. Their resistance change with temperature can be approximately calculated as

$$R = R_0(1 + \alpha\Delta T) \quad (3.34)$$

where α is the temperature coefficient of resistance (TCR). The highly doped silicon utilized for the piezoresistors exhibited a TCR between 1100 and 1500 ppm/K based

upon calibration experiments. Note that the TCR for doped silicon is positive (resistivity increases with temperature) because the carrier concentration remains fairly constant but carrier-phonon scattering increases with temperature, while it is negative for intrinsic silicon because thermal carrier activation is the dominant process. For comparison, pure aluminum and other metals of practical interest have TCRs on the order of 4000 ppm/K.

As in the case of mechanical crosstalk, temperature fluctuations are transduced into a resistance change. Consider a Wheatstone bridge composed of four resistors. If the average temperature of one of the resistors changes, the output voltage varies as

$$\Delta V_{\text{out}} = \frac{1}{4} \frac{\Delta R}{R} V_{\text{bridge}} \quad (3.35)$$

$$= \frac{1}{4} \alpha \Delta T V_{\text{bridge}}. \quad (3.36)$$

For a bridge bias of 1 V and piezoresistor TCR on the order of 1500 ppm/K, a temperature change of just 3 mK is sufficient to generate an input referred voltage change equal to the piezoresistor noise floor (approximately 1 μ V).

This calculation illustrates two issues: the need for a temperature compensation resistor for any piezoresistive device (silicon or metal), and the complications inherent in integrating thermal actuation with piezoresistive sensing. A theoretical calculation of thermal crosstalk is presented in Figure 3.9, illustrating that the average temperature change of the piezoresistor during thermal actuation is large enough (roughly 250 mK) that a common mode rejection approach is absolutely essential. In practice, relying on common mode rejection is not sufficient due to mismatch between the main and compensation force probes on a device, but resistive compensation can be readily added to the actuation circuit (Section 5.6).

3.4 Practical design issues

A number of practical issues had to be solved during the force probe design. In this section, I'd like to highlight a few of the issues that were considered.

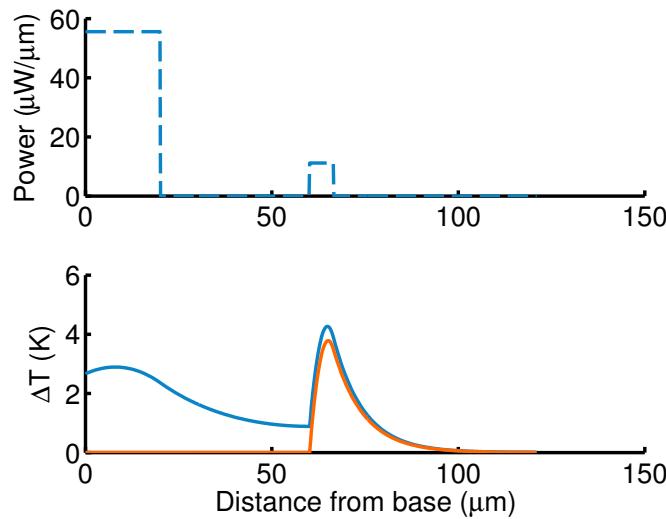


Figure 3.9: Example of thermal crosstalk within a thermally actuated force probe. As in many of the other actuator examples, we use a 61 μm long sensor at the end of a 60 μm long actuator. The power dissipation (top) and resulting temperature distribution (bottom) with the actuator turned on (blue) and off (orange) are presented for a probe operated in water with Wheatstone bridge and actuator bias voltages of 1 V. The average piezoresistor temperature change during actuation is on the order of 250 mK.

3.4.1 Die layout

One of the most important design requirements for the force probe was that it needed to be compatible with a 1 mm working distance upright microscope for the mammalian hair cell experiments. The conventional measurement setup with a pipette-based force probe was shown in Figure 1.9. This chapter has focused on the microscale design of the force probe, but microscope compatibility depends on the macroscale design of the silicon die that the force probe is attached to.

The size and shape of the silicon die was specifically designed to fit underneath a 1 mm working distance upright microscope objective (Figure 3.10). The extension is 5 mm long, 400 μm wide and 400-500 μm thick and is about six orders of magnitude stiffer than the force probes (10^3 N/m vs. 10^{-3} N/m). The extension introduces a spurious resonant mode into the force probe response near 20 kHz (Figure 5.40), but the vibration is small compared with the overall movement of the probe and is not detectable in the force probe step response. Both the silicon extension and the printed circuit board that the device is

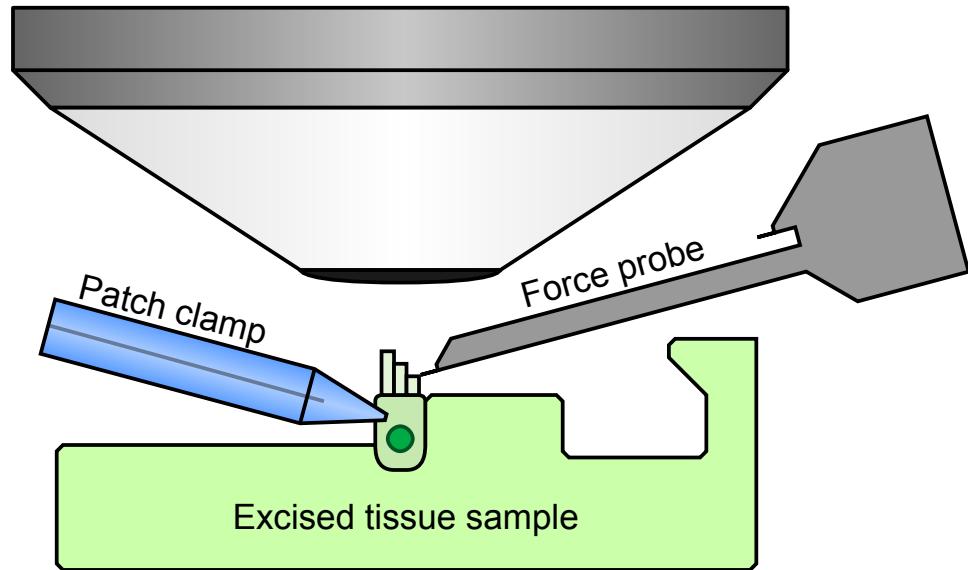


Figure 3.10: The devices are designed to fit underneath a 1 mm working distance upright objective. The force probe sits at the end of a 5 mm long, 400-500 μm thick, 400 μm wide silicon extension and is shifted as far to the left as possible so that it can be used to apply and measure either in-plane or out-of-plane forces to a sample.

mounted on (presented in Appendix C) need to be narrow enough to avoid running into the microscope objective.

In addition to being mounted on the end of a long extension, the force probe is shifted as far to the left as possible so that the probe can apply in-plane forces to a sample without the edge of the die crashing into the sample first. The first generation probes (Appendix F) did not include this feature and were not usable for hair cell experiments.

3.4.2 Depletion layers

The force probes combine a heavily doped (10^{20} cm^{-3}) n-type piezoresistor with a lightly doped (10^{15} cm^{-3}) p-type background. Accordingly, the depletion region around the piezoresistors extends roughly 2.5 μm into the lightly doped substrate (if the substrate is grounded and the piezoresistors are biased to approximately 1 V). If two n-type regions are separated by a distance comparable to the depletion width then a significant amount of electrical current will be shunted between them, also known as punchthrough. This poses

a problem for both the piezoresistor and the thermal actuator heater, where the electrical current should be forced around the resistive loops rather than shunted around them. The reader is referred to Ref. [211] for the pn junction working equations.

Ideally the background dopant concentration would be higher in order to reduce the depletion width and reduce the leakage current across the pn junction [211], but the background concentration was set by the availability of 100 mm SOI wafers with roughly 300 nm thick device layers from Soitec. The background concentration could be increased through a combination of boron ion implantation or diffusion and a long anneal at the possible expense of additional lattice damage and 1/f noise.

Instead, the force probes were simply designed with the wide depletion region in mind. Note that in Figure 3.1 there is an air gap between the two piezoresistor contacts, and the silicon between the piezoresistor legs is etched away. The air gap extends 3 μm beyond the end of the piezoresistor contacts so that the depletion regions are well separated. One benefit of the wide depletion region is a reduction in the junction capacitance, which reduces the capacitive crosstalk from the junctions.

An additional potential shunt resistance is leakage current across the reverse biased pn junction. The leakage current is proportional to the exposed area of the pn junction, which is fortunately quite small because the entire SOI device layer is n-type doped underneath the contacts and piezoresistors. In other words, for thin SOIs where the device layer is doped throughout the thickness (or where the thickness is much smaller than the depletion length) the leakage current is set by the perimeter of the doped area rather than the area underneath the doped regions. The typical reverse biased impedance between a piezoresistor and the substrate was at least $10 \text{ M}\Omega$, at least three orders of magnitude larger than the piezoresistor impedance.

3.4.3 Excess electrical resistance

In addition to maximizing the impedance of shunts in the piezoresistor circuit, the impedance of the interconnects and contacts should be minimized. Excess resistance reduces the resistance factor (Section 2.4.2) which reduces the piezoresistor sensitivity and adds additional Johnson noise. In addition, current crowding in the contacts can increase

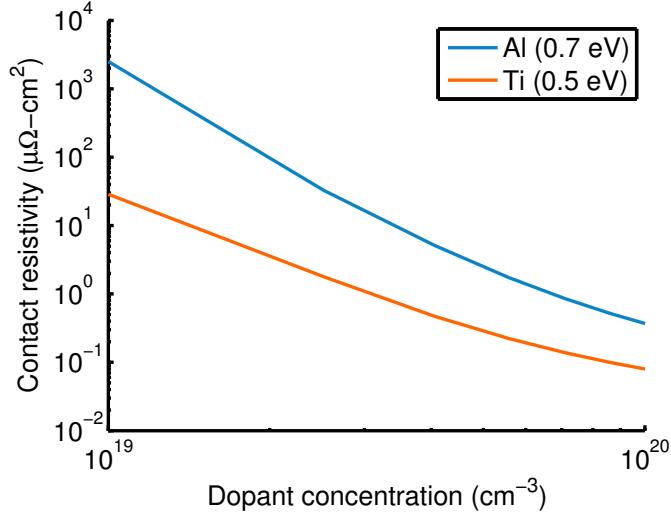


Figure 3.11: Contact resistivity as a function of n-type dopant concentration for Ti and Al metallization. The contact resistivity is calculated following Ref. [363]. In theory, both Ti and Al metallization can achieve a contact resistivity of less than $1 \mu\Omega \cdot \text{cm}^2$ for a dopant concentration of 10^{20} cm^{-3} . In practice, we found that the effective contact resistivity was on the order of $10\text{-}20 \mu\Omega \cdot \text{cm}^2$ for an optimized doping and Ti-based metallization process.

the 1/f noise of the piezoresistor [207].

The interconnects are fairly straightforward to design. They should be as thick, wide and as short as possible and be composed of a high electrical conductivity material. Aluminum serves this purpose quite well, with a resistivity on the order of $3 \mu\Omega \cdot \text{cm}$ in contrast with a minimum resistivity on the order of $1100 \mu\Omega \cdot \text{cm}$ for single crystal silicon doped to 10^{20} cm^{-3} with phosphorus. A $1 \mu\text{m}$ thick, $30 \mu\text{m}$ wide and 5 mm long Al interconnect ($30 \text{ m}\Omega/\square$) adds a series resistance of approximately 5Ω , negligible compared to the $2\text{-}5 \text{ k}\Omega$ piezoresistor impedance.

The contact resistance is more problematic, and I highly recommend Ref. [364] as an introduction to contact modeling and characterization techniques. In theory, the contact resistivity of Al to n-type Si doped to 10^{20} cm^{-3} should be less than $1 \mu\Omega \cdot \text{cm}^2$ (Figure 3.11), where the contact resistivity is calculated as a function of the contact barrier height and dopant concentration from the model presented in Ref. [363].

Ti presents a lower barrier height than Al to n-type Si [365] and doubles as a contact diffusion barrier [366, 367]. In theory, both Ti and Al metallization can achieve a contact

resistivity of less than $1 \mu\Omega\text{-cm}^2$ for a dopant concentration of 10^{20} cm^{-3} . In practice, we found that the effective contact resistivity was on the order of $10\text{-}20 \mu\Omega\text{-cm}^2$ for an optimized doping and Ti-based metallization process. Note that the contact resistivity decreases three orders of magnitude if the dopant concentration at the piezoresistor surface is increased from 10^{19} to 10^{20} cm^{-3} . If the piezoresistor is lightly doped then it is important to add a second doping step to the fabrication process in order to minimize the contact resistivity and sheet resistance of the silicon underneath the contacts.

The piezoresistor contacts on the devices fabricated in Chapter 4 ended up being approximately $3 \mu\text{m}$ wide and $16 \mu\text{m}$ long. In a simple zero-dimensional model, the contact resistance could be calculated from

$$R_{\text{contact}} = \frac{\rho_{\text{contact}}}{w_{\text{contact}} l_{\text{contact}}} \quad (3.37)$$

where w_{contact} and l_{contact} are the width and length of the contact while ρ_{contact} is the contact resistivity. Assuming a contact resistivity of $10 \mu\Omega\text{-cm}^2$, the contact resistance would be 22Ω .

However, the contacts are relatively long and narrow and the silicon underneath the contact has a much higher sheet resistance than the Al, so a one-dimensional transmission line model for the contact resistance would be more appropriate. Following Ref. [368], the contact resistance can be calculated as

$$R_{\text{contact}} = \frac{\sqrt{R_s \rho_{\text{contact}}}}{w_{\text{contact}}} \coth \left(l_{\text{contact}} \sqrt{\frac{R_s}{\rho_{\text{contact}}}} \right)$$

where R_s is the sheet resistance of the underlying silicon, assumed to be much less than the metal. A typical piezoresistor sheet resistance described in Chapter 4 is $110 \Omega/\square$, leading to a contact resistance of 110Ω . If the contact doping is increased to yield a sheet resistance of $20 \Omega/\square$ then the contact resistance drops to 50Ω , even neglecting the contact resistivity reduction that would likely accompany the increased doping.

To summarize this section, in designing a force probe you should use metal rather than silicon interconnects, maximize the area of the contacts, select an appropriate contact resistance model and dope the contacts as highly as possible.

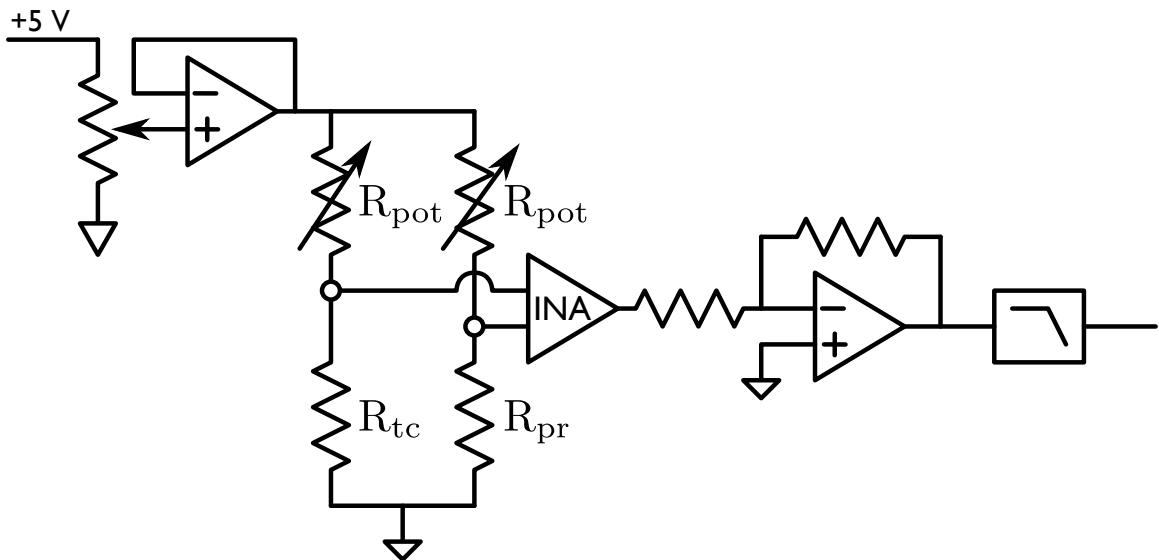


Figure 3.12: Summary of the sensor signal conditioning circuit. The Wheatstone bridge bias is generated by a 5 V voltage reference, adjusted downward using a potentiometer and buffered. The Wheatstone bridge consists of two potentiometers and the main (R_{pr}) and compensation (R_{tc}) force probes. The Wheatstone bridge output is amplified using an instrumentation amplifier, amplified further with an inverting amplifier and low pass filtered before it is sampled.

3.5 Circuit design

The force probe sensor and actuator performance depends strongly on the design of their circuit interface. Just one example of the tight coupling between the device and circuit performance is the effect of amplifier noise on the piezoresistor force resolution (Section 2.4.1). In this section we will discuss the design of the circuit used to interface with the force probes. The complete circuit schematic and printed circuit board (PCB) layout as well as important implementation details are presented in Appendix C.

The circuit has three main jobs. First, the circuit must provide an adjustable, low noise voltage source for the Wheatstone bridge. Second, the circuit needs to amplify the output of the Wheatstone bridge for downstream data acquisition. Finally, the circuit must include a high bandwidth amplifier to drive the on-chip actuators and provide a means to compensate any capacitive crosstalk between the actuator and sensor.

The bias generation and signal conditioning portions of the circuit are summarized

in Figure 3.12. The Wheatstone bridge bias is generated by a 5 V voltage reference (e.g. ADR445), adjusted downward using a potentiometer and buffered (AD8671). The Wheatstone bridge consists of two potentiometers and the main (R_{pr}) and compensation (R_{tc}) force probes. The Wheatstone bridge output is amplified using an instrumentation amplifier, amplified further with an inverting amplifier and low pass filtered before it is sampled.

A voltage reference is used to generate the bridge bias to reduce the power supply noise in the system. The voltage reference rejects ripples in the power supply below 1 kHz exceptionally well, which greatly reduces the system noise compared with the power supply used to bias the circuit (HP E3631A). A comparison between the power supply and voltage reference voltage noise spectra, as well as the corresponding piezoresistor noise spectra, are presented in Figure 3.13. The voltage reference decreases the power supply noise peaks by an order of magnitude below 1 kHz, resulting in significantly better system noise performance. Ripple in the Wheatstone bridge bias couples into the sensor output due to imperfect matching between the two sides of the bridge, which limits the common mode rejection ratio (CMRR) of the circuit.

The Wheatstone bridge output is amplified 100-fold using a high bandwidth instrumentation amplifier (INA103). The amplifier gain is set below its maximum (1000-fold) in order to maintain a -3 dB bandwidth of 800 kHz. Additional gain is provided by an inverting amplifier (THS4031). The op-amp and feedback resistors are chosen so that their noise is roughly 50-fold smaller than the system noise at this intermediate point in the circuit. A switch on the PCB allows adjustment of the feedback resistor in order to select variable gain amplifier (VGA) gains of 1-, 10- or 100-fold.

The selection of the instrumentation amplifier plays a large role in the noise performance of the system. Figure 3.14 compares the integrated input referred voltage noise of the measurement system from 1 Hz to 10 kHz for two typical instrumentation amplifiers as a function of the source impedance presented by the piezoresistor. The INA103 has low voltage noise and high current noise, so performs best with a relatively low source impedance. In contrast, the AD8221 has high voltage noise and low current noise, so yields lower noise for piezoresistors above 4 k Ω . The bandwidth of the INA103 and AD8221 are on the order of 800 and 100 kHz for G = 100. The amplifier noise and bandwidth

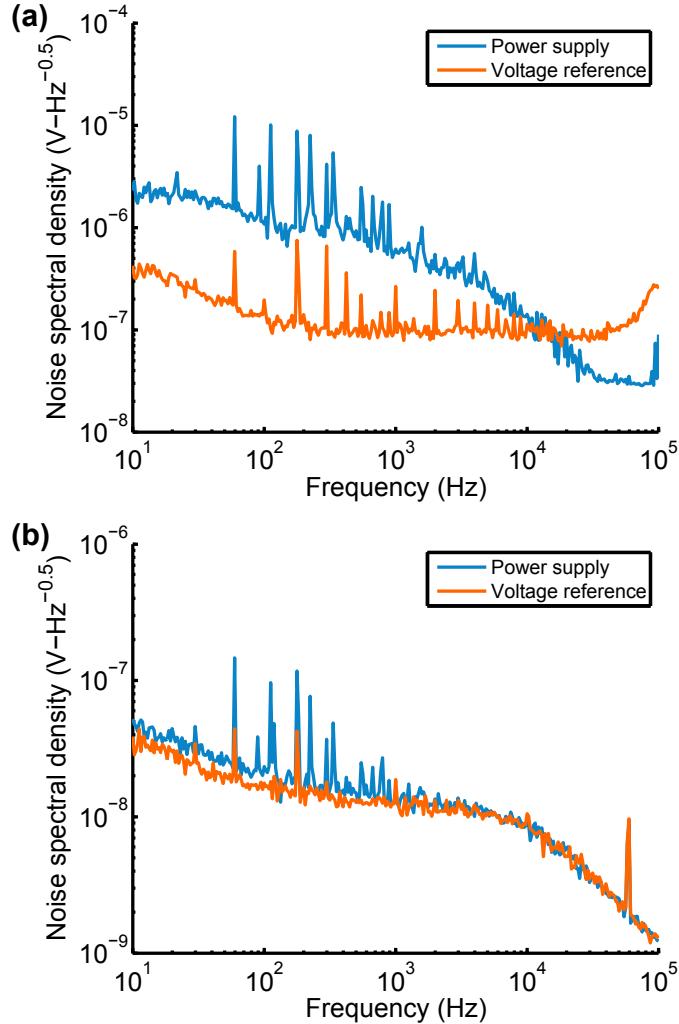


Figure 3.13: Comparison between the noise performance of power supply and voltage reference based Wheatstone bridge bias sources. (a) The voltage reference decreases the power supply noise peaks by an order of magnitude below 1 kHz, resulting in (b) significantly better system noise performance.

performance are included in the numerical cantilever optimization model.

The force probes are mounted on a custom PCB that is connected to the amplifier circuit PCB using a shielded RJ45 cable. The details of the device cable and power supply cables are presented in Appendix C. The cable shield is directly connected to the PCB ground in order to reduce capacitive crosstalk, ground the force probe PCB ground planes and reduce electromagnetic interference. Directly connecting the cable shield to the ground plane on

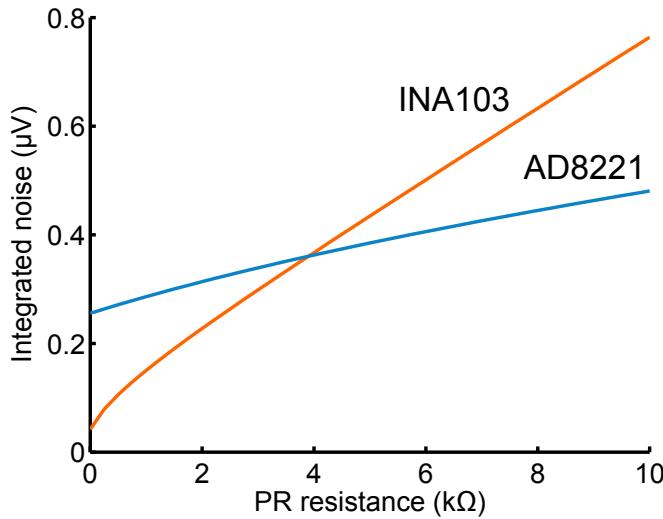


Figure 3.14: Measurement circuit noise as a function of source impedance for two typical instrumentation amplifier options. The low voltage noise and high current noise of the INA103 compared with the AD8221 leads to inferior performance above a piezoresistor resistance of 4 kΩ. The amplifier performance should be considered during piezoresistor design for high performance devices.

the circuit avoids the formation of a ground loop, which would occur if a separate cable was used to connect the two.

The design of actuator drive and crosstalk compensation portions of the circuit will be discussed in Sections 5.6.1 and 5.6.2. Essentially, the actuator drive is generated by a computer controlled digital-to-analog converter (DAC), buffered with an op amp and used to drive the actuator. The buffer is required in order to compensate for the limited slew rate and output current of the DAC. Several potentiometers and variable capacitors are added to the circuit in order to compensate for mismatch between the main and compensation force probes. The circuit can be operated in either a single ended or differential drive configuration through the addition of a fully-differential op-amp (THS4131). As for the rest of the circuit, the implementation details are presented in Appendix C.

Chapter 4

Fabrication

In this chapter we will translate the optimized force probe designs from the last chapter into working devices. We will discuss the mask layouts, wafer layouts and device layouts before moving on to the process flows. Detailed process runsheets and recipes are presented in Appendix E. Device performance will be discussed in the next chapter. Many of the fabrication and design features were developed from the lessons learned in fabricating an earlier generation of devices, described in Appendix F. We will begin by discussing some fabrication process fundamentals.

4.1 Processing fundamentals

The starting wafers for the fabrication process are 100 mm diameter, (100) oriented double-side polished p-type SOI wafers. The SOIs were sourced from Soitec and had device layer, buried oxide (BOX) and handle wafer thicknesses of 340 ± 44 nm ($\mu \pm 3\sigma$), 400 or 1000 ± 7.5 nm, and 400 to 500 μm , respectively. The background resistivity of the device layer was 10-20 $\Omega\text{-cm}$.

I had the good fortune of processing my wafers shortly after an ASML PAS 5500/60 stepper with backside alignment capability was installed in the SNF. The ASML is a 5:1 reduction stepper that combines an i-line source (365 nm) and 0.54 numerical aperture (NA) lens to achieve resolution of approximately 500 nm in practice. The frontside-to-frontside and frontside-to-backside overlays of the system are typically on the order of 15

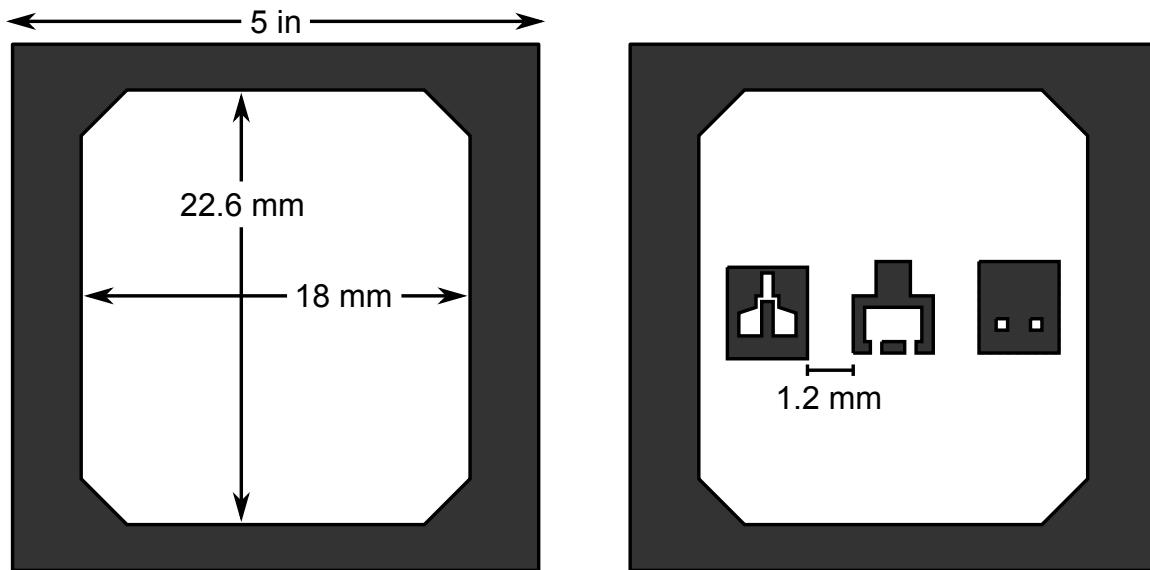


Figure 4.1: (left) The reticles used in the ASML PAS 5500 stepper are 5 inch quartz plates with a thin chrome coating. A 18 mm x 22.6 mm (at the wafer scale) region in the center of the reticle is usable for device layout. (right) Reticle layout consists of placing images on the reticle. A small gap is required between images in order to allow the stepper blades to select the image to expose. Spacing of at least 1.2 mm is recommended.

and 25 nm, respectively.

The data on the reticle consists of a series of images that, when properly combined on the wafer, can produce a single device. A jobfile specifies where each image should be located within a cell on the wafer (Figure 4.2). Each cell consists of multiple devices, and the cell is repeated across the wafer, yielding multiple copies of each device design. As drawn in Figure 4.2 there are six devices in each cell, with each device consisting of three images (frontside, backside, bondpads). Devices that are too close to the edge of the wafer (shaded red) are manually removed from the jobfile so that they are not exposed. This is a particular issue for processes that include a through-wafer etch step, where etching too close to the edge can lead to weaken the wafer. A 5 mm edge exclusion zone was used for all of the jobfiles.

The ASML uses a 5 inch quartz mask with a usable area of 18 mm x 22.6 mm at the wafer scale (90 mm x 113 mm at the reticle scale) as shown in Figure 4.1. During device layout, the reticle is drawn in a computer program (L-Edit) to define the regions on the

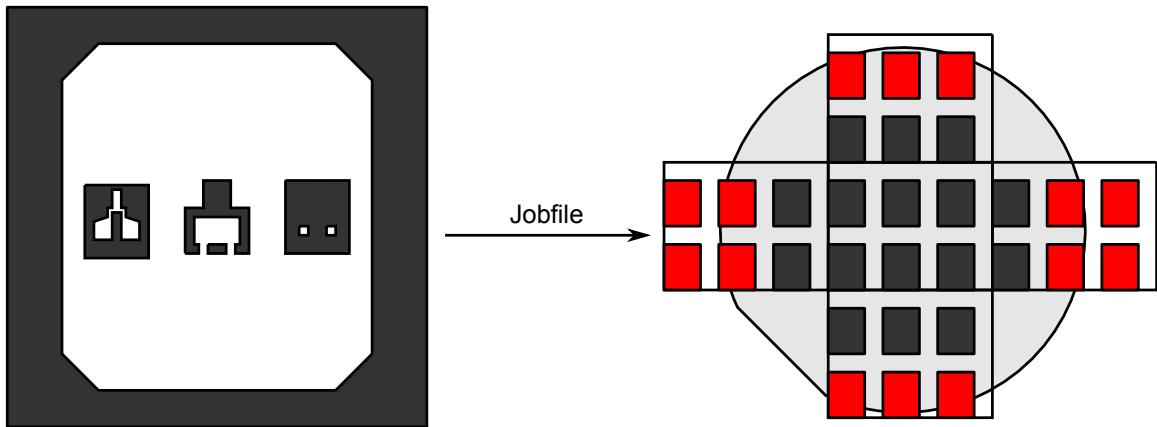


Figure 4.2: A jobfile maps images between the reticle (left) and the wafer (right). The three images on the reticle represent a frontside, backside and bondpad layers that could be used in a simple three-layer piezoresistor process. The jobfile maps each image to a layer within a cell. A cell (consisting of six devices as drawn) is repeated across the wafer. Devices that cross the edge exclusion zone of the wafer (drawn red) are manually removed from the jobfile.

quartz mask where chrome should be removed (i.e. the mask is drawn data clear). A small gap is required between the images on the reticle in order to allow the stepper to select the image to expose via a set of blades. If the images are too close to one another and/or the layer is overexposed then neighboring images can bleed into one another. The minimum image spacings recommended by ASML and the SNF are 1.24 and 1.4 mm at the wafer scale, respectively. I used an image spacing of 1.2 mm without any noticeable problems.

The number of devices that can be fabricated on each wafer depends on two main issues. First, the devices need to be spaced sufficiently far apart so that the wafer does not crack after the backside etch step. The die size (2.7 x 8.15 mm) and spacing between devices (2 mm) yielded 105 devices per wafer. Second, the backside etch area needs to be minimized in order to reduce loading during the deep reactive ion etching (DRIE) step; etching less than 10% of the wafer area is a good rule of thumb [369]. The exposed area on all of the wafers that were processed was less than 3%. The reticles and wafer layouts are shown in Appendix D. Four reticles (fast-probe-1, -2, -3 and -4) were used to fabricate the devices discussed in this chapter, while two reticles (fast-pr-1 and -2) were used for the fabrication

run discussed in Appendix F.

The use of a stepper rather than a contact aligner for lithography provides several practical benefits. The first benefit is the improvement in resolution, allowing the force probe dimensions to be minimized. The probe designs assume a 0.5 μm minimum feature size and an overlay tolerance of 0.25 μm for contact vias. The second benefit is that the reticle and wafer layouts are decoupled, increasing the flexibility of the process and allowing devices to be stitched together from multiple images as shown in Figure 4.3.

Although the force probe die are large, only a small region changes between designs. The probes were defined using a 60 x 340 μm image (Figure 4.3b). In order to stitch images together, a region was left unexposed on the large, shared image and filled in later with the smaller, unique images. The unique images were designed to include several microns of overlap when possible and were designed so that any seams would not impede device operation. The downside of image stitching is increased exposure time; rather than a device consisting of a single exposure, with stitching it could consist of up to five separate exposures. Stitching increases the flexibility of the lithography process and allows utility images (e.g. 100 x 100 μm squares) to be included on the reticle. The mechanics of the ASML backside alignment system (3D align) and image stitching are discussed in more detail by Nahid Harjee in Ref. [329].

After the wafers have been completely processed the devices need to be released from the wafer. Conventional approaches to device release include wafer sawing and, more recently, laser dicing. Unfortunately, neither technique is suitable for our extremely soft and fragile force probes. We opted to use the snap tab approach that has been used for several generations of piezoresistive cantilevers developed at Stanford.

An example of a released device and the snap tab release process is illustrated in Figure 4.4. Force applied to the silicon die induces a tensile stress in the top surface of the snap tabs. The force is usually applied with a set of tweezers or other small but not overly sharp object. Plastic tweezers (e.g. Delrin) work particularly well.

The peak tensile stress in the tabs can be roughly calculated from

$$\frac{\sigma_{\text{tab}}}{F} = \frac{3l_d}{w_{\text{tab}}t_{\text{wafer}}^2}$$

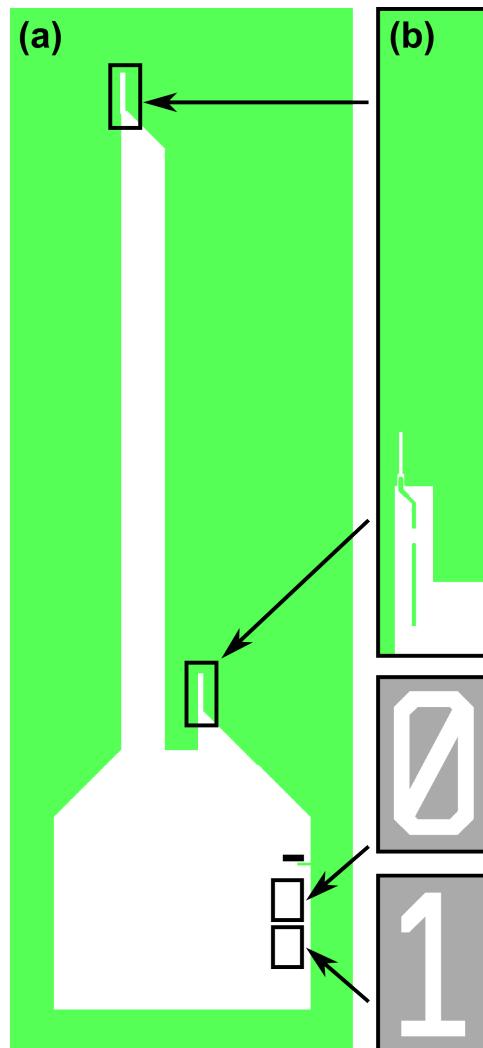


Figure 4.3: Image stitching enables the conservation of reticle space. Only a small portion of the device varies between designs, so the majority of the device can be patterned with a large, shared image (a) while the design of the force probes (b) can be adjusted by exposing them separately. The probe image overlaps the base image by several microns so that seams are not present. Multiple exposures can also be extended to die labeling; in this case the device is labeled 01 using separate number images.

where l_d is the distance from the base of the snap tabs to where the force is applied, w_{tab} is the width of each tab and t_{wafer} is the thickness of the wafer. Rather than trying to quantify the force required to break the snap tabs, we want to qualitatively note that the force required should scale as $1/l_d$, w_{tab} and t_{wafer}^2 , and should not depend heavily on the

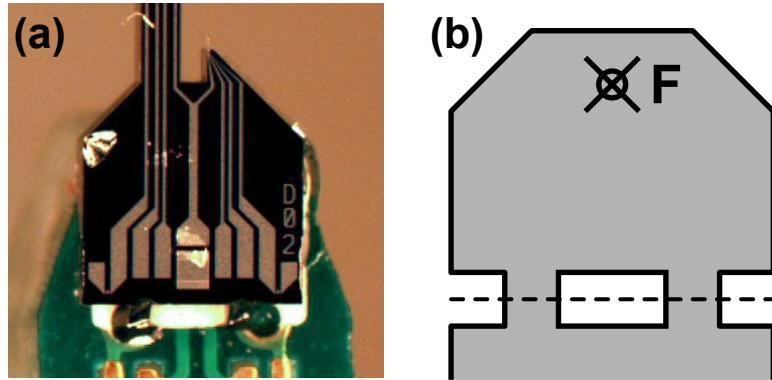


Figure 4.4: The force probes are released from the wafer using snap tabs defined during the backside etch step. (a) A released device with broken snap tabs is shown. The metal lines were slightly damaged by pressing on them with the tweezers. (b) The snap tabs break when a force is applied to the silicon die. If the tabs are too stiff then the tabs can be lightly scribed in order to initiate a crack for easier device release (dashed line).

length of the tabs.

The force probes were designed with 200 μm wide snap tabs and the force was applied about 1.5 mm away. These snap tabs were far too stiff, and excessive force led to wafers cracking and reduced device yield. In order to solve this problem, the snap tabs were lightly scribed (Figure 4.4b) in order to form cracks at the surface. The cracks allowed the devices to be removed from the wafers with much lower forces and increased device yield dramatically. Although scribing generated particles, the force probes were far enough away that the particles did not cause any issues. The device layer was not etched on top of the snap tabs; etching it may have helped in making the tabs easier to snap.

4.2 Device design

The force probes were fabricated from September 2010 through May 2011. Within the run there were several process flow iterations in order to solve problems encountered during fabrication, which will be discussed in detail. Three types of force probes were fabricated: sensor-only (PR), thermally actuated (PRT) and piezoelectrically actuated (PRPE). In this section we will discuss their layout and fabrication, drawing upon the lessons from the first fabrication run (Appendix F).

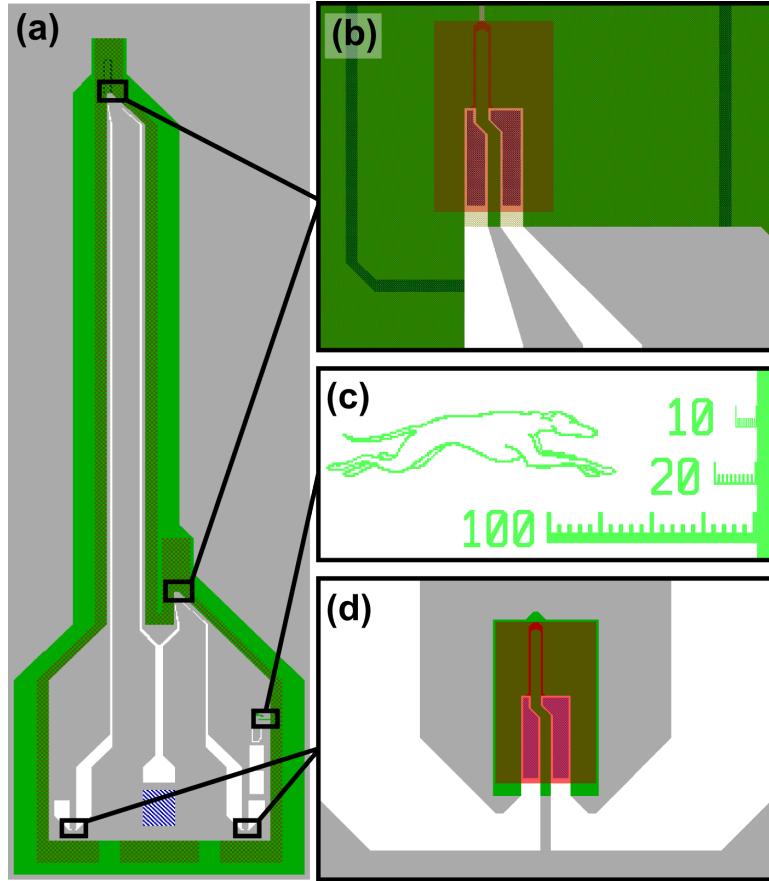


Figure 4.5: PR device layout. (a) The die is long and skinny in order to fit underneath a 1 mm working distance objective. (b) Probe layout, showing the relatively large diffusion well and Al-stiffened base. (c) Scalebars and chip art located on every device. (d) Detail of the on-chip temperature compensation resistors.

4.2.1 Sensor-only probes

An overview of the PR device layout is shown in Figure 4.5. The overall die layout is shown in Figure 4.5a. The backside etch trench that traces the silicon die is 100 μm wide. If the release trench is too narrow then the etch rate suffers and, in extreme cases, the trench can pinch itself off. If the trench is too wide then the sidewalls will be less vertical. The trench size was selected by running several test wafers through the backside DRIE release process. Multiple trench size variations (100, 200 and 300 μm) were included on the reticle for flexibility.

The Al interconnect layer is shown in gray as data clear. The interconnects connect the $100 \times 100 \mu\text{m}$ bondpads with the piezoresistors. The bondpads allow the piezoresistor resistor (PR), temperature compensation resistor (TC) and two unreleased silicon resistors that were not used in practice. The bottom bondpad in the middle of the die provides contact to the substrate, whereas the rest of the bondpads sit on a thin layer of SiO_2 for electrical isolation.

A closeup of the main and compensation force probes is shown in Figure 4.5b. The width of the piezoresistor legs is either 0.5 or 1 μm in all of the device designs. Rather than attempting to diffuse dopants into such a small gap, a larger region is doped (shaded red) and the unwanted silicon is etched away in order to form the resistive loop. By limiting the region of doped silicon to just the piezoresistor loop, rather than the entire length of the probe, curvature is kept to a minimum.

The depletion region extends about 2.5 μm into the p-type material due to the high background resistivity of the wafers (Section 3.4.2). Material is etched away 3 μm beyond the end of the diffused region at the base of the probe in order to ensure that the two depletion regions do not connect. A crack is intentionally etched into the BOX (Figure 4.5b). The crack surrounds the force probes and increases the device yield during the fabrication process.

The metal interconnects extend onto the base of the force probe, forming a 20 μm long stiffened portion. This design was chosen in order to minimize variation in the resonant frequency and stiffness of the probe with variation in back-to-front misalignment from the backside DRIE. But this approach has two downsides. The first is that the CTE mismatch between the metal and silicon leads to variation in deflection with temperature (Table 3.1). This was not a problem in practice due to the short timespan of the hair cell experiments. The stiffener could alternatively have been made from silicon at the expense of additional process complexity. The second downside of this approach is that the contacts are fairly small at 3 μm wide by 16 μm long with a 500 nm gap between the edge of the vias and the edge of the probe to allow for misalignment. As we will see, contact resistance proved to be a substantial problem.

A closeup of the scalebars and chip art included on every device is shown in Figure 4.5c. The greyhound is composed of $1 \times 1 \mu\text{m}$ squares and gives the force probes an important

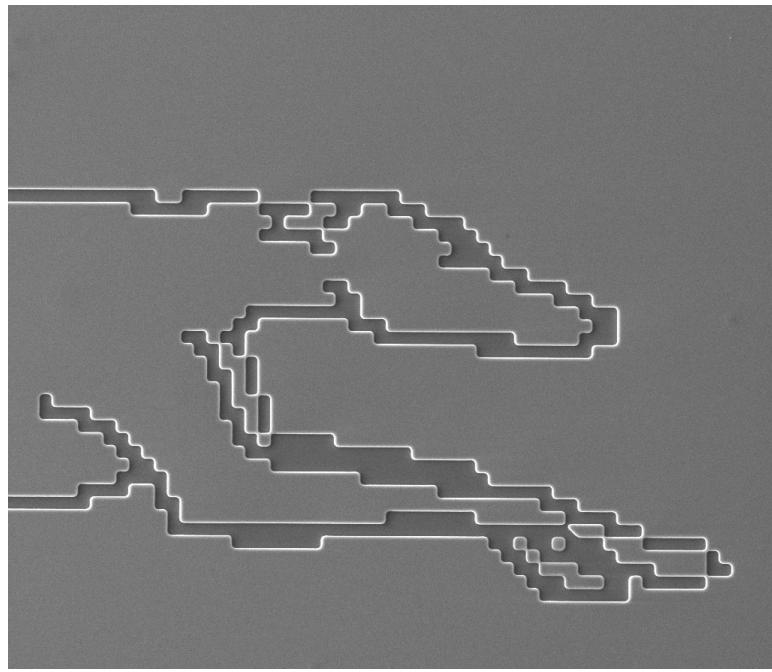


Figure 4.6: Scanning electron micrograph of a greyhound etched into a 300 nm thick SOI device layer. The greyhound is composed of $1 \times 1 \mu\text{m}$ squares and gives the force probes an important speed boost.

speed boost. A fabricated greyhound is shown in Figure 4.6. The scalebars were helpful in assessing the backside DRIE undercut because the device layer is thin enough to transmit light.

The on-chip temperature compensation resistors are shown in Figure 4.5d. The oxide via and diffusion layers from the released probes are recycled for the on-chip resistors, with just a new frontside etch image required to etch around the resistor. The contact regions on all of the piezoresistors, both released and unreleased, are identical. As noted earlier, the unreleased resistors were not used in practice. They were intended to simplify the measurement circuit, allowing a single potentiometer to balance the bridge, but we found that resistor mismatch usually necessitated the use of two potentiometers.

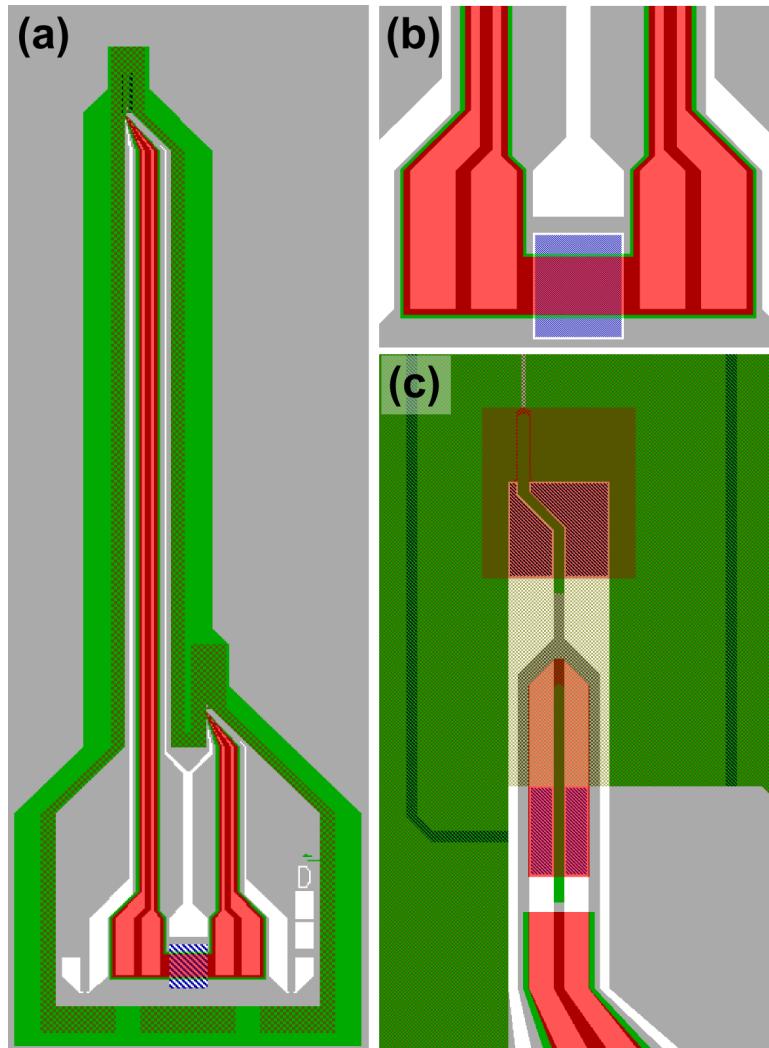


Figure 4.7: PRT device layout. (a) Overall layout of the probe. (b) Detailed view of the base of the die, showing the n-type diffusion well that runs underneath the actuator lines and the narrow trench etched around the well. The substrate contact provides an AC ground for the diffusion well. (c) Detailed view of the probe. The heater portion consists of Al interconnects that make contact with a silicon resistor that runs in a loop underneath the Al lines in the center of the probe.

4.2.2 Thermally actuated probes

The PRT device layout is shown in Figure 4.7. The overall shape and size of the silicon die is identical to the PR layout. Four bondpads and metal lines are added to the device to

provide electrical access to the resistive heaters used for actuation.

One important feature of the design is the diffusion well and etch trench that electrically isolates the actuator lines from the sensor lines (Figure 4.7b). The region underneath the actuator interconnects is doped during the formation of the piezoresistor in order to provide a low impedance path to an AC ground. The isolation well extends 20 μm on either side of the actuator lines. Although the depletion region between the well and p-type background is fairly wide (2.5 μm as noted earlier), a 15 μm trench is etched into the device layer around the isolation well in order to further reduce the capacitive coupling between the actuator and sensor lines. The isolation well is connected to the substrate contact to provide an AC ground.

The probe design is shown in Figure 4.7c. The actuator interconnects make contact with the silicon heater via 4 x 17 μm contact windows. But the actuator interconnects continue beyond the vias because the expansion of the metal is the actuation source for the PRT probe. The gap between each of the four metal lines at the base of the probe is 2 μm and the minimum width of the metal is 2 μm . The isolation well and isolation trench follow the actuator interconnects to within 7 μm of the silicon heater. The depletion regions occupy 5 μm of the gap with a 2 μm gap to allow for misalignment. The gap between the end of the silicon heater and the piezoresistor is 16 μm . The small gap between the isolation well and the heater was a source of concern during processing and we will return to it later.

4.2.3 Piezoelectrically actuated probes

The PRPE device layout is shown in Figure 4.8. The layout and design features of the PRPE probe are nearly identical to the PRPE probe. The top and bottom PE layers are shown as data dark rather than data clear for improved clarity.

The actuator is a sandwich structure, so a single broad electrode runs along the length for the long extension rather than two adjacent electrodes. As in the case of the PRT design, an isolation well and trench separate the actuator from the sensor (Figure 4.8b). The bondpads for the bottom PE electrodes are on the inside of the die while the bondpads for the top electrodes are on the outside.

A detailed view of the PRPE probe is presented in Figure 4.8c. As in the case of the

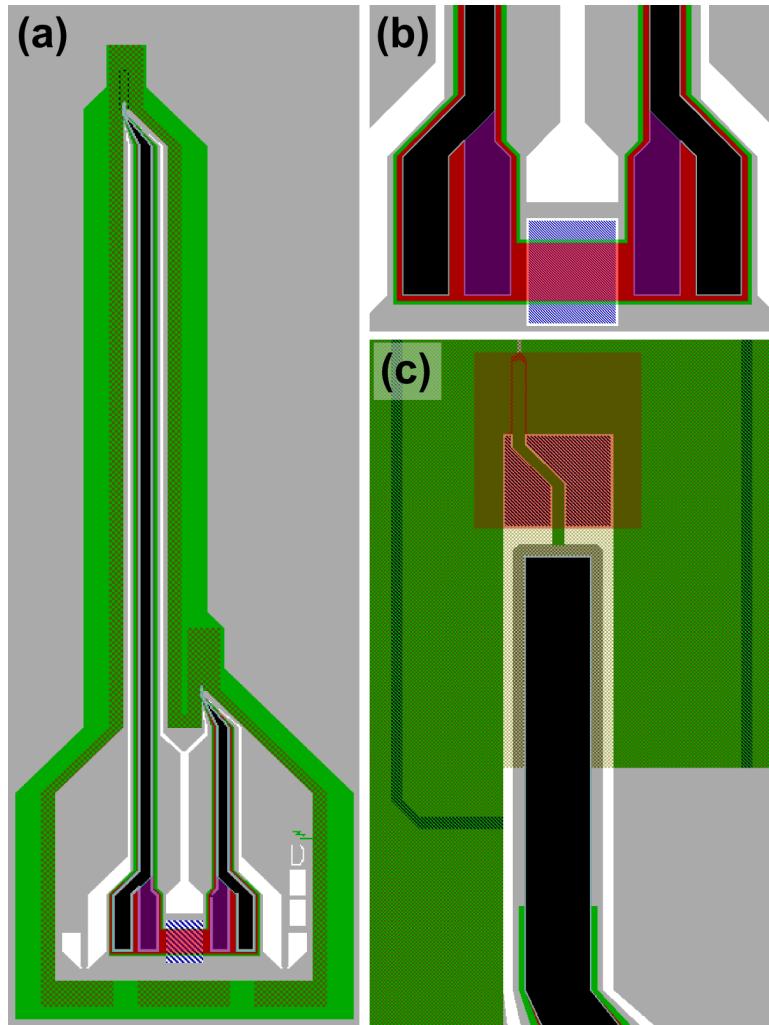


Figure 4.8: PRPE device layout. (a) Overall layout of the probe. (b) Detailed view of the base of the die, showing PE electrode bondpads, isolation well and isolation trench. (c) Detailed view of the probe. The gap between the actuator and Al is 2 μm .

PRT design, the PR interconnects run along the outside of the probe and are 2 μm wide with a 2 μm gap between them and the PE stack that runs along the center of the probe. The actuator stack is 12 μm occupying slight more than half of the width of the 20 μm wide base of the probe. The isolation trench runs to within 25 μm of the piezoresistor diffusion, while the isolation well continues to within 6 μm of the piezoresistor diffusion. As in the case of the PRT probe, this small gap became an issue during probe fabrication.

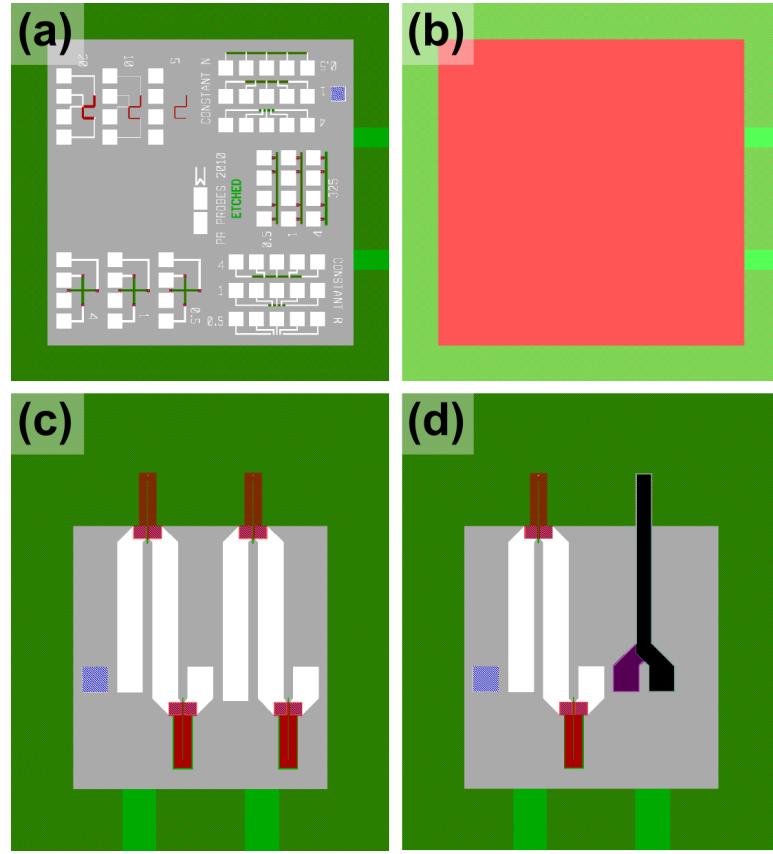


Figure 4.9: Additional device designs. Every wafer included (a) electrical test structures, (b) test coupons for SRA/SIMS and ellipsometry, and sensors intended for surface stress measurements based upon either (c) piezoresistive cantilevers or (d) a combination of piezoresistive and piezoelectric cantilevers.

4.2.4 Miscellaneous designs

Other designs that were included on every wafer are shown in Figure 4.9. The electrical test structure die is shown in Figure 4.9a and includes four types of test structures. The bottom left of the die has Van der Pauw Greek cross structures of three different sizes for measuring the sheet resistance of the piezoresistor. The top left of the die has Kelvin bridges of three different contact sizes for measuring the contact resistivity. On the right side of the die are three linear resistors of varying widths used to measure the effective resistor width including systematic biases from the lithography and etch steps. The top and bottom corners of the right side contain four resistors arranged for measuring the noise

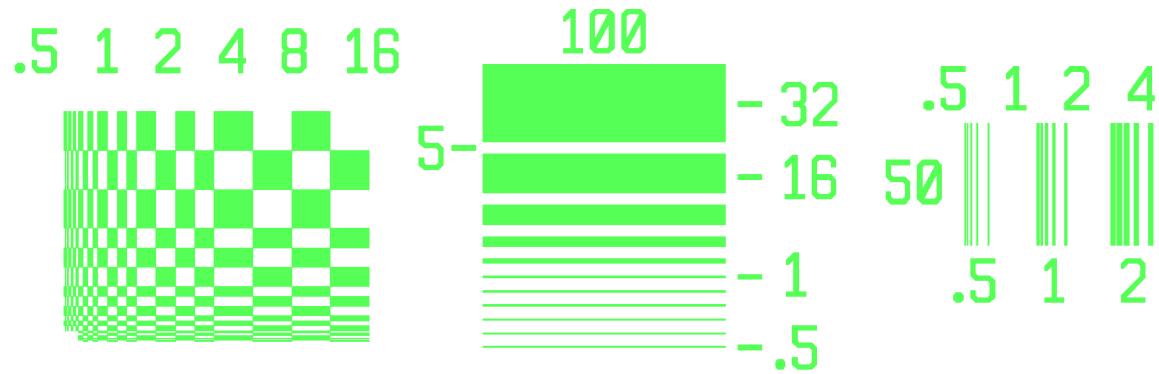


Figure 4.10: The exposure grid and test features used to calibrate the exposure dose for each layer. The ideal exposure dose varies substantially depending on the films underneath the photoresist, even for the same photoresist thickness, making it essential to characterize each lithography step. The exposure grid was particularly important due to the $0.5\text{ }\mu\text{m}$ alignment tolerance used in the device layout. The feature sizes and spacings are indicated (in μm) by the text labels.

characteristics of the diffused resistors. The Kelvin bridge and Van der Pauw structures were essential in characterizing the processed wafers and should always be included. Interested readers are referred to Refs. [370] and [285] for test structure overviews.

A blank test coupon (Figure 4.9b) was also included on the wafers and served two purposes. During processing, the coupon was used for ellipsometry measurements to monitor the oxidation and etch steps in the process and to measure the final thickness of the device layer. After processing the coupons were snapped out of the wafer and used for SRA and SIMS measurements of the dopant concentration profile. Both undoped and doped coupons were fabricated on each wafer.

Surface stress sensors for chemical sensing were fabricated during the second and third fabrication runs for a side project with Ali Rastegar. The surface stress sensor design will not be discussed in this thesis, but the optimization code can be used to design them. The device fabricated on the PR and PRT wafers is seen in Figure 4.9c and includes two released cantilevers, two unreleased TC resistors and a substrate contact. The device fabricated on the PRPE wafers is seen in Figure 4.9d and includes both a PR and PE cantilever in order to combine both surface stress- and mass-based sensing modalities.

The exposure calibration image used to calibrate the exposure dose for each layer of

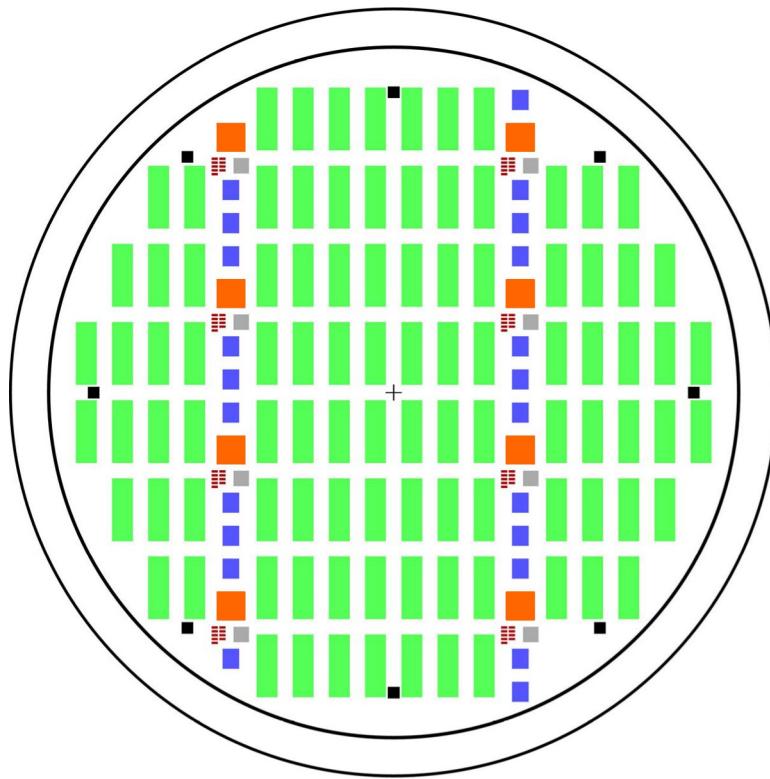


Figure 4.11: Wafer layout. The force probes are colored green, chemical sensors are blue, exposure grids are red, test die are orange, open areas for ellipsometry are gray and alignment marks are black. A 5 mm edge exclusion zone was included in the device layout. Four of the test die are used for electrical test structures and the other four are used for SRA/SIMS measurements.

the process is shown in Figure 4.10. An exposure grid, isolated lines and line arrays are included in the image on the reticle. The ideal dose depends on a number of variables, particularly the optical properties of the films underneath the photoresist, and certain layers in the process were intentionally under- or overexposed in the course of the process. The exposure dose values included in the runsheet (Table E.4) represent the optimized settings.

4.3 Wafer layout

All three device varieties used the same wafer layout, shown in Figure 4.11. The 14 device designs are placed in a 7 x 2 grid within each cell with a 2 mm gap between images.

A column of additional devices is placed at the right side of each cell containing three chemical sensor devices, an exposure grid for each layer, a 2 x 2 mm open area and either a test structure or test coupon (four of each variety were fabricated on each wafer). Each wafer yielded 105 devices. The wafers used PM alignment marks, with four marks for frontside alignment and four additional marks for backside alignment. All metal layers are removed from the 2 x 2 mm test areas so that the thickness of the Si/SiO₂ stack formed during the process can be measured. The area is large enough for access using the variable angle spectroscop ellipsometer in the lab.

The job files were defined using Jobscripter, a program developed by Rishi Kant. Jobscripter translates a spreadsheet into a job file using the ASCII import/export functionality of the ASML workstation software. The total number of images in the job (including all layers) was greater than limit allowed by the ASML software, so each layer was output as an individual job file. Generating the PR, PRT and PRPE job files manually in the ASML software would have taken 2-3 weeks even as an experienced and relatively fast user. In contrast, generating the job files in Jobscripter took about 2 days of work and a little bit of trial and error. I can not recommend the use of Jobscripter highly enough for anyone developing a moderately complex processor on an ASML stepper. Sample spreadsheets for Jobscripter are available upon request.

4.4 Process flow

In this section we will describe each of the process flows in detail. We will describe the challenges encountered during processing and the process flow iterations after describing the each of the finalized process flows. Three varieties of probes were fabricated (PR, PRT and PRPE) and types of substrates were used (SSP, POI and SOI). The entire 84-step runsheet that can be used to fabricate all three probe varieties is presented in Table E.4.

Single side polished (SSP) wafers were used to test the frontside wafer processing steps and to fabricate batches of electronic test structures in order to optimize the diffusion and metallization processes. Poly-on-insulator (POI) wafers batch manufactured in the SNF in order to test the entire process flow and to provide functional test probes. The POIs were fabricated in 25 wafer batches using the seven step process described in Table E.5 and they

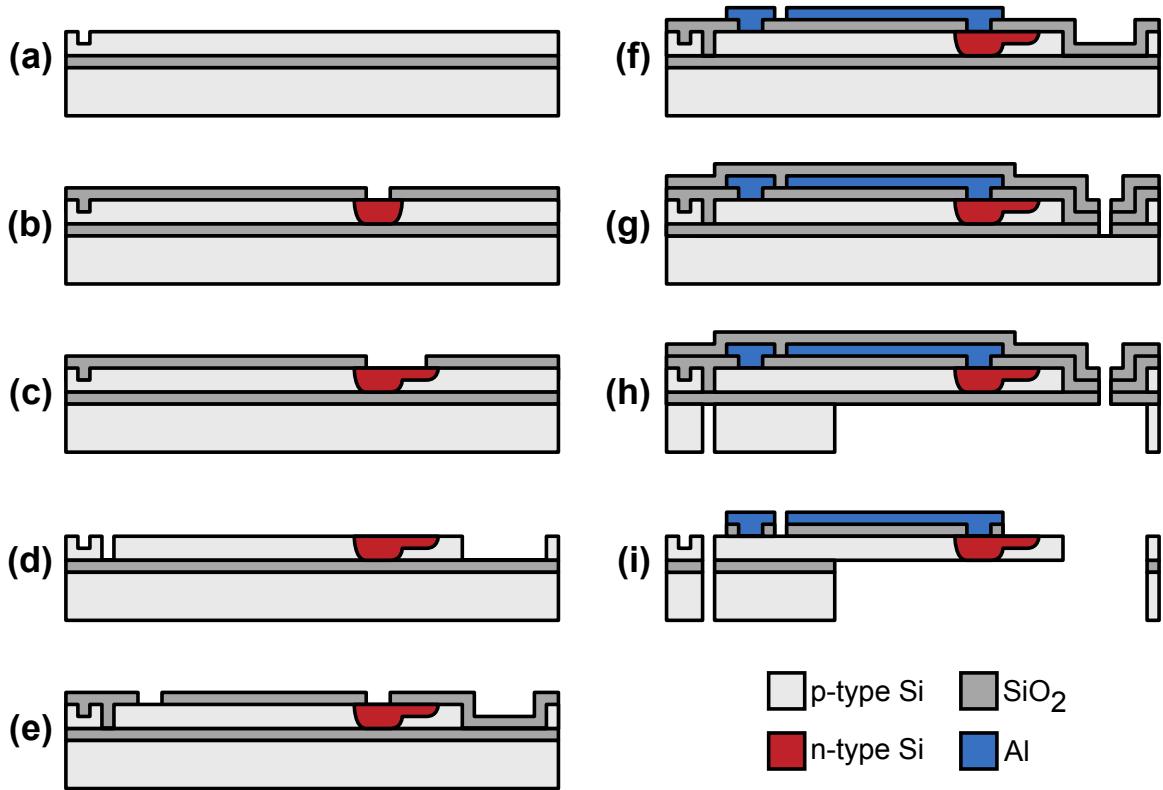


Figure 4.12: Second generation PR process flow. (a) Alignment marks are patterned and etched. (b) The wafer is oxidized, an additional layer of LTO is deposited, and the wafer is annealed. Windows in the oxide are opened for the contact diffusion. (c) The oxide windows are opened further for the piezoresistor diffusion. (d) The oxide is stripped and the frontside of the wafer is etched to define the force probe. (e) An LTO isolation layer is deposited and vias are opened before (f) the Ti/Al metal layer is deposited and etched back. (g) An additional layer of LTO is deposited to seal the metal and cracks are patterned into the LTO/BOX stack. (h) The backside of the wafer is patterned and etched via DRIE before (i) the BOX and LTO layers are removed via vapor HF.

were instrumental in developing the fabrication process and experimental methods. The SOI wafers combined a 340 nm thick device layer with either a 400 or 1000 nm thick BOX. The former variety had fewer problems with BOX cracking and oxide undercut during the release process.

The PR process is illustrated in Figure 4.12. The devices are fabricated using a seven mask process. Process details, such as wafer handling issues and etch recipes, are presented in Appendix E. At the start of the process the wafers are labeled with a diamond scribe and

cleaned in piranha to remove the particles. The thicknesses of the device layer and BOX are measured at 9-points via VASE and the overall thickness of the wafer is measured using a depth gauge.

Alignment marks are patterned in a 0.7 μm thick layer of SPR955 and etched 1200 \AA into the device layer. The wafer is rotated $45 \pm 0.001^\circ$ during the alignment mark exposure so that the probes will be oriented in the $<100>$ direction ± 0.5 degrees as set by the device layer orientation tolerance. After the marks are etched the photoresist is stripped in piranha.

The oxide diffusion mask is formed next. The wafers are diffusion cleaned and a 900 \AA thick thermal oxide layer is grown by oxidizing them in steam at 900°C. The thermal oxidation decreases the thickness of the device layer from approximately 340 to 300 nm. An additional 3200 \AA thick layer of LTO is deposited on the wafers at 400°C in order to increase the thickness of the oxide mask without any further reduction in the device layer thickness. After the LTO deposition the wafers are annealed in N₂ at 1100°C for 20 hours in order to densify the LTO and reduce strain gradients in the device layer.

Two separate diffusions are performed. In the first, the piezoresistor contacts are POCl₃ doped at 900°C for 20 minutes (Figure 4.12b). The thin layer of PSG formed on the surface is stripped in a dilute HF solution, leaving the underlying oxide mask intact. In the second diffusion the window in the oxide diffusion mask is opened further to dope the entire piezoresistor region at 825°C for 30 minutes (Figure 4.12c).

All of the remaining oxide is stripped and the device layer is patterned and etched via RIE (Figure 4.12d). Afterwards the wafer is cleaned and coated with a 2000 \AA thick layer of LTO and contact vias are wet etched (Figure 4.12e). Wet rather than dry etches are used for all steps that terminate on the piezoresistor to avoid any damage to the surface.

The wafers are coated in metal next (Figure 4.12f). In the sputterer loadlock the wafers are heated to 200°C for 3 minutes using quartz lamps to drive off any moisture before depositing 50 nm thick Ti and 1 μm thick Al layers. As discussed in Section 3.4.3, the layer of Ti reduces the contact resistance and provides a contact diffusion barrier. The metal is lithographically patterned and etched back via RIE, stopping on the LTO layer (Figure 4.12f). Dry rather than wet etching is required due to the thick Al (1 μm) and narrow metal lines (2 μm). Afterwards the wafers are rinsed in water to eliminate residual HCl on the wafer before stripping the photoresist mask. The metal RIE etch is one of the

trickiest steps in the process; it needs to be long enough to eliminate any sidewall stringers. Next a 4000 Å thick LTO layer is deposited on the wafer and a trench is patterned and etched via RIE into the oxide around the force probe (Figure 4.12g).

Afterwards the backside of the wafer is patterned and etched via DRIE, stopping on the BOX (Figure 4.12h). The wafers are shipped to the Lurie Nanofabrication Facility (LNF) at the University of Michigan for the backside etch. A staff member at the LNF bonds the device wafer to a carrier wafer using CrystalBond 555 (CB) using a custom bonding tool that eliminates air pockets in the CB. After the wafers are etched and shipped back to the SNF, the photoresist is stripped and the wafer is debonded from the carrier by soaking in 70°C water for 30 minutes. The LTO coating deposited earlier prevents metal corrosion during this step. Afterwards the wafer is cleaned in new DI water and dried.

The wafers are cleaned in a downstream O₂ plasma ashser to remove residual polymer. However a thin layer of fluorocarbon polymer remains on the wafer because it is not volatile in pure O₂ plasma. Without further cleaning a thin web of polymer remains on the wafer, rendering most of the devices unusable. A second polymer cleanup step is performed in a CF₄/O₂ plasma in order to remove the remaining fluorocarbon polymer as in Ref. [371].

Finally, the BOX and LTO layers are etched in vapor HF in order to release the force probes (Figure 4.12i). The wafers are etched from the frontside, but cracks in the BOX allow the vapor HF to attack the BOX and LTO layers from both sides at once. The top LTO coating is approximately the same thickness as the BOX and minimizes undercut of the LTO layer underneath the metal interconnects. No FGA is necessary and can actually degrade the contact resistivity depending on the metallization and diffusion processes.

The PRT process is illustrated in Figure 4.13. The fabrication process is identical to the PR process and similarly uses seven masks. The same LTO and Al thicknesses are used in both processes. The only changes to the PRT wafers are in the device layout with the addition of the isolation well, isolation trench, and silicon heater for the thermal actuator at the base of the force probe.

The PRPE process is illustrated in Figure 4.14. The beginning and end of the process are identical to the PR and PRT process flow, with the addition of several steps in the middle for the deposition and patterning of the piezoelectric stack (Figure 4.14d-g). The additional steps add two lithography steps to the process, bringing the total number of

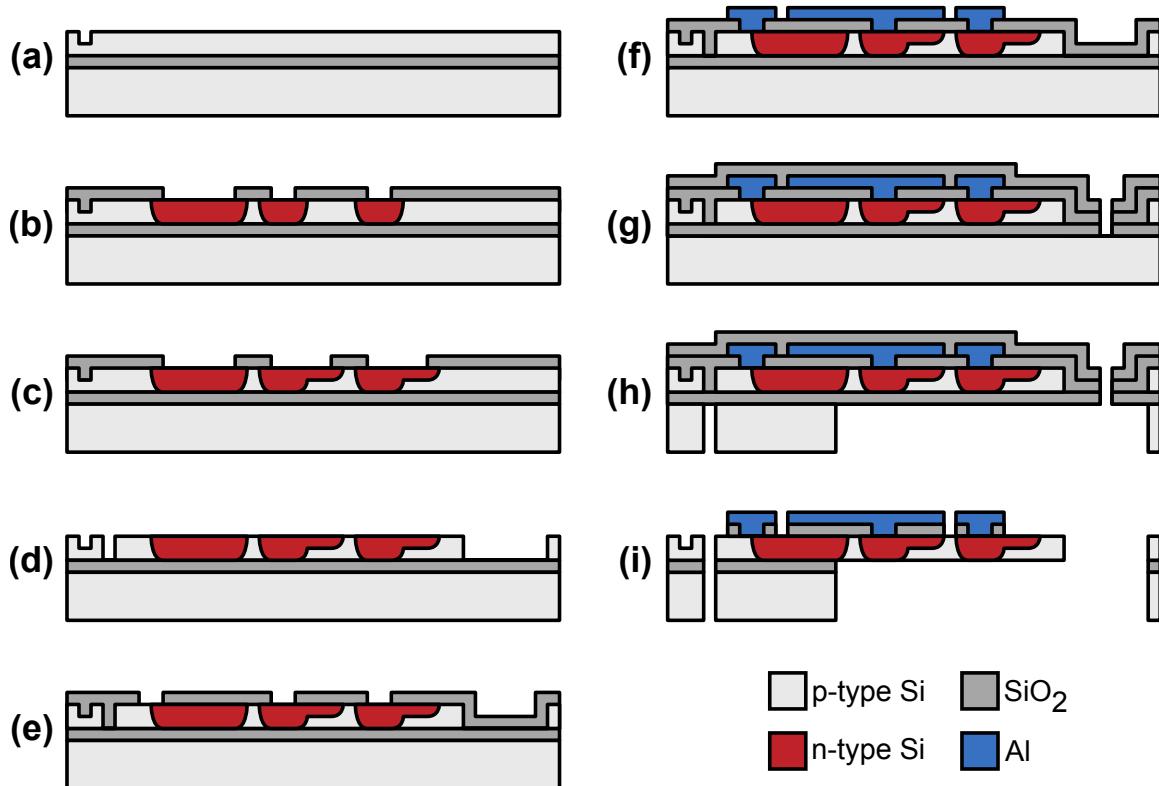


Figure 4.13: Second generation PRT process flow. (a) Alignment marks are patterned and etched. (b) The wafer is oxidized, an additional layer of LTO is deposited, and the wafer is annealed. Windows in the oxide are opened for the contact and isolation well diffusions. (c) The oxide windows are opened further for the piezoresistor and heater diffusions. (d) The oxide is stripped and the frontside of the wafer is etched to define the force probe. (e) An LTO isolation layer is deposited and vias are opened before (f) the Ti/Al metal layer is deposited and etched back. (g) An additional layer of LTO is deposited to seal the metal and cracks are patterned into the LTO/BOX stack. (h) The backside of the wafer is patterned and etched via DRIE before (h) the BOX and LTO layers are removed via vapor HF.

lithography steps up to nine.

After the frontside of the wafer is patterned (Figure 4.14c) a 100 or 250 nm thick layer of LTO is deposited and the wafers are shipped to OEM Group. Approximately 200 Å of LTO is sputtered from the surface of the wafer and the PE stack is deposited, consisting of a 250 Å thick AlN seed, 750 Å thick bottom Mo electrode, 750 Å thick AlN actuator and 500 Å thick top Mo electrode. The LTO layer serves to protect the piezoresistor from the presputter step and as an etch stop while the PE stack is patterned. The net stress of the

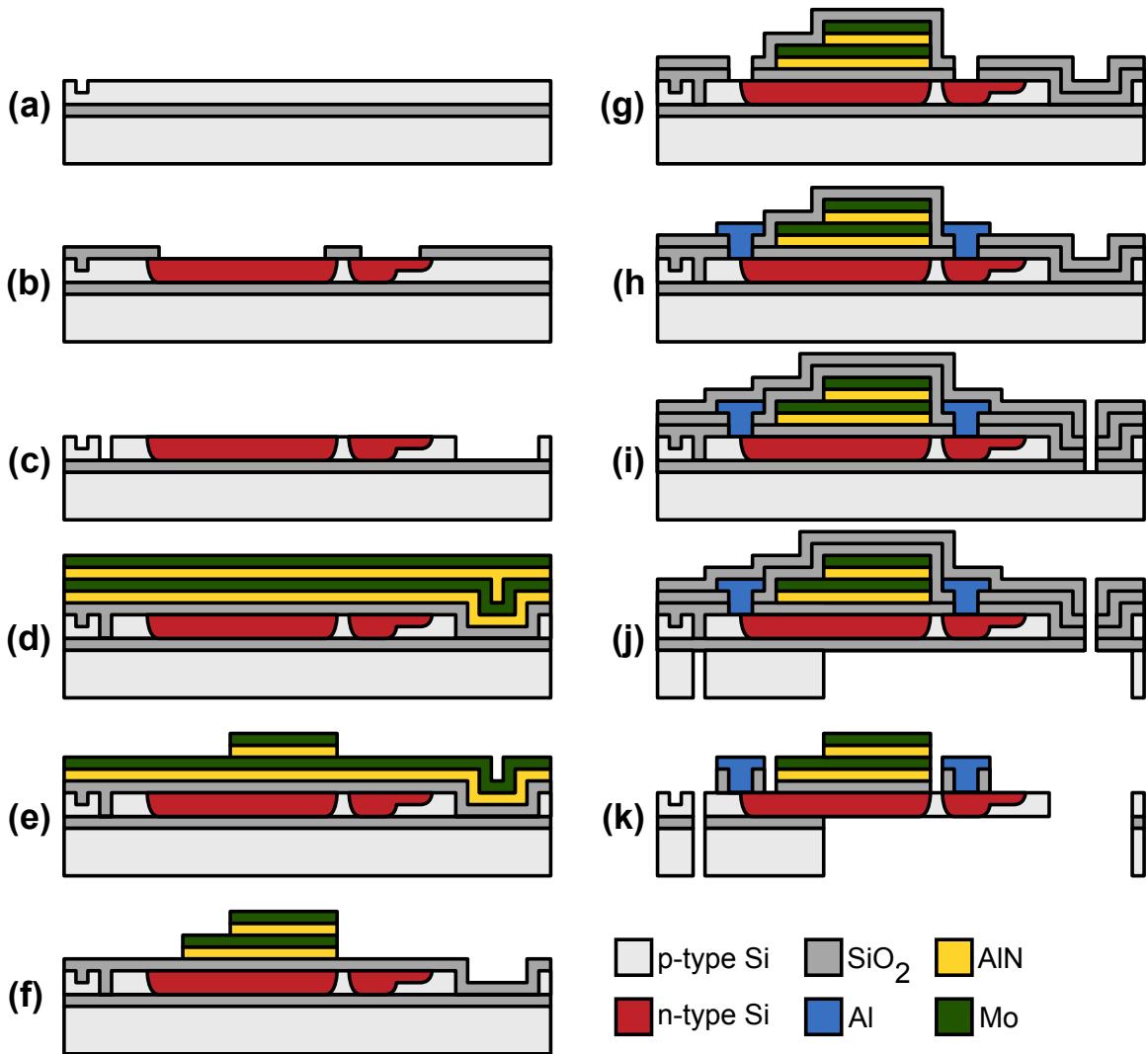


Figure 4.14: Second generation PRPE process flow. (a) Alignment marks are patterned and etched. (b) The oxide mask is formed and the contacts and piezoresistors are doped. (c) The oxide is stripped, the frontside of the wafer is etched before (d) a layer of LTO and the PE stack are deposited. (e) The top Mo and AlN layers are patterned and (f) the process is repeated for the bottom Mo and AlN layers. (g) A layer of LTO is deposited and vias are opened. (h) The Ti/Al metal stack is deposited and patterned. (i) An additional layer of LTO is deposited and cracks are patterned into the oxide stack. (j) The backside of the wafer is etched via DRIE before (k) the BOX and LTO layers are etched via vapor HF.

deposited films and an XRD rocking curve are measured by the vendor before the wafers are shipped back. The stress was typically engineered to -200 ± 100 MPa in order to

balance the compressive stress of the LTO and typical rocking curve FWHM values were 3.3-3.5 degrees.

After the PE stack is deposited (Figure 4.14d) a 1 μm thick layer of SPR955 is patterned and the top Mo electrode is etched in an SF₆ plasma (Table E.8). The endpoint of the etch is detected by monitoring the Mo emission peak, although the timing is not critical because it terminates on the 7500 \AA thick AlN layer, which etches at <10 $\text{\AA}/\text{min}$. The photoresist mask is left intact and the AlN actuator is etched in a Cl₂ plasma (Table E.7). The etch consists of a 30 second long BCl₃ breakthrough step in order to remove the surface oxide [372, 373]. While Ar can be added to the main etch step in order to increase the etch rate [374], we opted to use only Cl₂ in order to maximize the selectivity with respect to the underlying 750 \AA thick Mo layer. The AlN recipe etches AlN and Mo at about 540 and 240 $\text{\AA}/\text{min}$, respectively, making the use of OES endpoint detection important for this step (Table E.10). Afterwards the wafer is immediately rinsed in DI water in order to remove residual chlorine and HCl from the wafer and the photoresist is stripped in PRX-127 (Figure 4.14e). All wet processing steps in which Mo is exposed are performed in a quartz beaker rather than the standard wetbenches in order to prevent cross contamination. The Mo and AlN etch recipes were developed by consulting Refs. [375,376] and [372,373,377], respectively, in addition to helpful discussions with Tina Lamers and Keri Williams of Avago.

The exact same lithography and etch process is repeated for the bottom Mo electrode and AlN seed layer (Figure 4.14f). The AlN etch stops on the underlying LTO layer and is fairly insensitive to overetch; 800 or 2300 \AA of LTO remains and the recipe etches LTO at only 78 $\text{\AA}/\text{min}$ (Table E.12).

After the PE stack is completely processed a 1000 \AA thick layer of LTO is deposited and oxide vias are wet etched (Figure 4.14g). As in the rest of the oxide etch steps the wafer is pretreated with O₂ plasma in order to render the photoresist hydrophilic and the vias are etched using 20:1 BOE. The etch is timed by visually monitoring the backside of the wafer and overetching by 20%. The wafers are agitated during the etch and lifted out of the bath every few seconds to check if the backsides of the wafers are still hydrophilic or have become hydrophobic. After the wafers are cleaned in PRS-1000, the native oxide is removed with a brief 50:1 HF dip before rushing the wafers to the sputterer loadlock for

metal deposition.

As in the case of the PR and PRT processes, the wafers are heated in the load lock to drive off moisture and the Ti/Al stack is deposited. Whereas 1000 nm of Al was used for the PR and PRT devices, only 400 nm of Al is deposited on the PRPE devices. In the PR and PRT devices the Al serves as a stiffener and actuator, respectively, and benefits from being thicker. But in the case of the PRPE devices it only serves as an electrical interconnect and any excess thickness actually inhibits the PE actuator. Afterwards the Al/Ti stack is lithographically patterned and etched back via RIE as in the rest of the processes (Figure 4.14h). The remainder of the steps (Figure 4.14i-k) are unchanged from the other fabrication processes.

4.5 Process issues and iterations

The fabrication process took about nine months and consisted of three process iterations with several short loops interleaved. The process described so far in this chapter is the final, optimized process. Now we will discuss the issues encountered along the way, including: Al patterning, BOX cracking and yield, piezoresistor thinning and damage, high contact resistivity, backside DRIE and polymer cleanup.

4.5.1 DRIE and BOX release steps

The first fabrication run (Appendix F) was performed in 2008 and used the STS DRIE tool in the SNF. Tool limitations (etch rate uniformity, footing, grass and device undercut) were overlooked at the time due to other more fundamental issues with the devices. By the time of the second fabrication run (2010-2011) the performance of the tool had degraded even further. Typical results are presented in Figure 4.15. While the second generation devices were designed with frontside stiffeners in order to reduce their sensitivity to backside undercut, the 30-50 μm undercut resulted in some yield issues and would have degraded the thermal and mechanical properties of the probe too much. We consulted the DRIE literature in an attempt to improve the sidewall angle of the process [378–380], but were not able to obtain vertical sidewalls without the formation of grass. Other issues with the

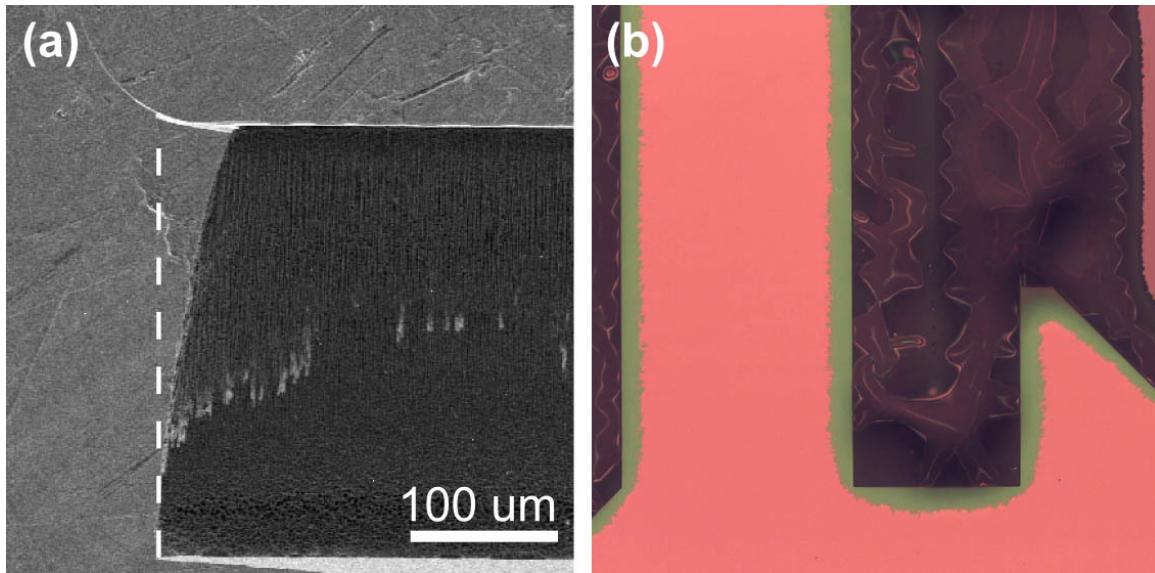


Figure 4.15: Initial backside etch results using the DRIE tool in the SNF. (a) Side view scanning electron micrograph of a test POI wafer. The backside undercut is approximately 50 μm . The device layer is extremely curved because the wafer was not annealed after the polysilicon deposition. (b) Optical micrograph from the frontside of a typical etched wafer.

DRIE tool in the SNF included lack of repeatability and frequent downtime.

After investigating outside DRIE vendor options, we opted to send the wafers to the Lurie Nanofabrication Facility (LNF) at the University of Michigan. We shipped device and carrier wafers to the LNF where they bonded them via Crystalbond 555 and etched them on their STS Pegasus. While we attempted to bond several wafers ourselves, air pockets in the Crystalbond reduced its thermal conductivity and resulted in photoresist burning; the LNF has a special bonding tool that applies heat and vacuum during the bonding process. The Pegasus etch recipe used yielded an etch rate of 12.1 $\mu\text{m}/\text{min}$, PR:Si selectivity of 200:1 and SiO_2 :Si selectivity of over 800:1.

Typical results from the SNF and LNF processes are presented in Figure 4.16. The device undercut depends on two factors: the tool and the etch trench size. Larger trenches etch faster also have increased undercut due to a wider acceptance angle of the incoming ions. Initial release tests were performed with a 300 μm wide trench. Reducing the size of the trench to 100 μm decreased the undercut in the SNF from about 50 to 30 μm (Figure 4.16a). Combining the 300 μm trench with LNF release process resulted in about 10 μm of

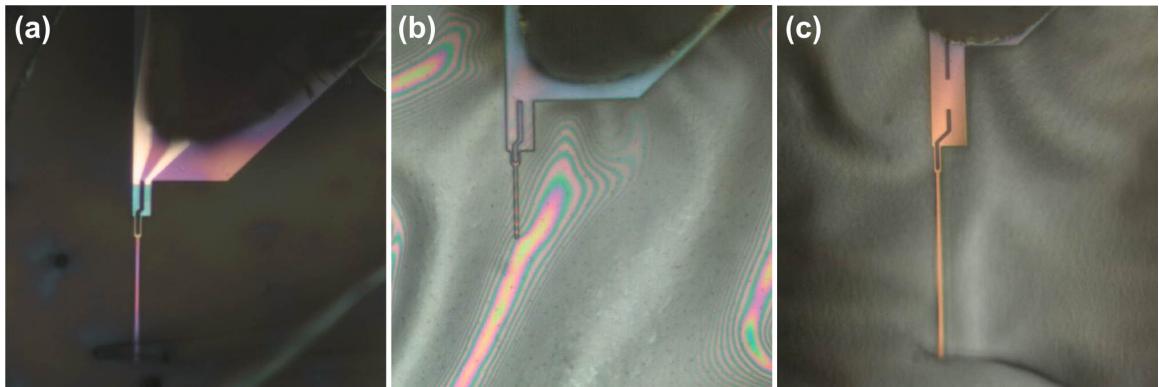


Figure 4.16: Backside undercut for three different DRIE processes. The undercut depended on the size of the etch trench and the tool that the wafers were etched on. (a) A 100 μm etch trench and the SNF DRIE tool resulted in 35 μm of undercut. (b) Combining a 300 μm etch trench and the LNF DRIE tool resulted in 10 μm of undercut and even more near the probe. (c) Switching to a 100 μm etch trench and shifting the backside mask 10 μm to the side resulted in almost no undercut.

undercut. But the undercut in the vicinity of the probe was closer to 20 μm due to undercut from the side (Figure 4.16b). Switching to a 100 μm trench and shifting the lithographic mask 10 μm to account for the sidewall undercut resulted in almost no undercut (Figure 4.16c).

Switching DRIE tools required a few changes to the process and introduced two new problems that needed to be solved. One benefit of the SNF's DRIE process is that it only requires a frontside photoresist protection layer rather than a carrier wafer. The LNF etched several test wafers without carriers, and while they survived the etch, they broke in transit from Michigan to California. The remainder of the wafers were etched with a carrier.

Upon receiving the wafers back from the LNF, the first step was to debond the device wafer from the carrier. At this point in the process it was essential to not heat the wafers above 55°C and to avoid applying vacuum to the wafer. For example, running the wafers through an O₂ plasma ash with the quartz lamps turned on resulted in the Crystalbond melting and puncturing the BOX membranes holding it in place. In light of these limitations, the first step upon receiving the wafers was to strip off the backside photoresist in PRX-127 (40°C). After rinsing the wafers in room temperature DI water the wafers were immediately transferred while still wet to a quartz beaker containing hot water

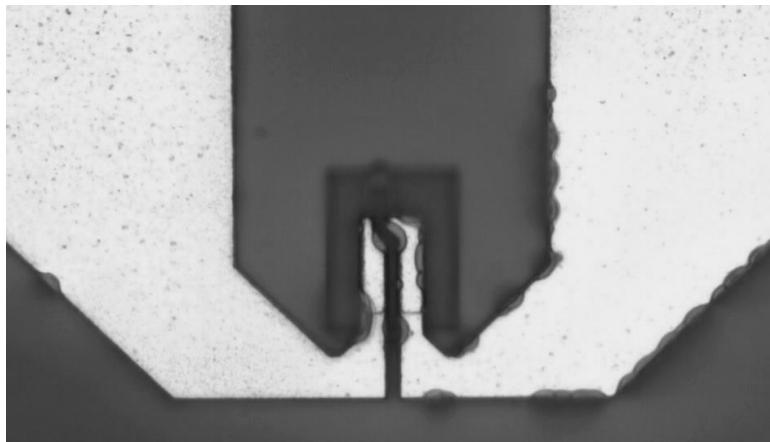


Figure 4.17: Example of metal corroding when the carrier wafer is debonded in hot water. Coating the wafer with a layer of LTO before the backside release process prevents corrosion and reduces undercut during the vapor HF release step.

($\approx 70^{\circ}\text{C}$). If the photoresist was still present on the wafer at this point in the process then it would bubble and contaminate the water bath.

Exposed metal on the wafer tends to corrode when the carrier wafer is debonded in hot water. In Figure 4.17 the Ti/Al metallization was not coated with an additional layer of LTO before undergoing the release process. Depositing an additional layer of LTO on the metal ensures that no corrosion occurs and reduces the LTO undercut during the vapor HF release process.

After sitting in the water for about 30 minutes the device wafer was lifted very carefully off of the handle wafer and delicately cleaned in DI water. For all of the wet steps after the backside of the wafer is etched the wafers should be processed in a cassette and carefully lifted up and down without any excess agitation; holding the wafers in tweezers results in too much motion normal to the cantilever surface and breaks probes. Finally the wafers are placed in a 90°C convection oven for about 15 minutes to dry them.

A major issue with the release process was polymer deposition. Polymer is deposited on the wafer during the DRIE process and the RIE process initially used to release the BOX. Polymer buildup from these two sources is illustrated in Figure 4.18.

The DRIE polymer cleanup proceeds in two steps. After debonding the device wafer from the carrier, the frontside of the wafer is covered in Crystalbond residue while the

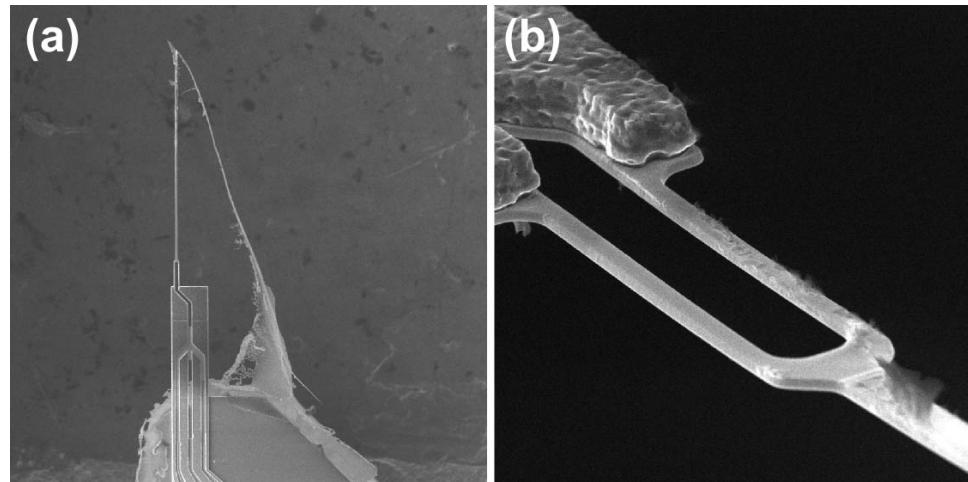


Figure 4.18: Examples of polymer buildup on released devices during the (a) DRIE and (b) RIE etch steps. Both polymer types were cleaned or eliminated from the devices in the optimized process.

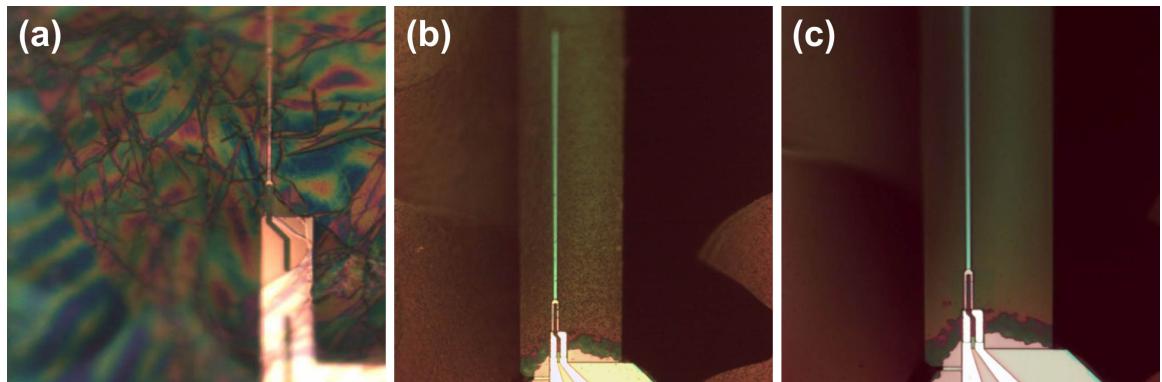


Figure 4.19: Polymer cleanup process. (a) The bulk of the polymer is removed in a heated O₂ plasma ashing step. (b) However, a thin fluorocarbon film remains on the wafer. (c) While the fluorocarbons are not volatile in a pure O₂ plasma, adding a small concentration of CF₄ to the plasma removes them from the wafer.

backside is coated in fluorocarbons (Figure 4.19a). Running the wafer through a heated O₂ plasma ashing step removes the bulk of the polymer but leaves a thin film of non-volatile fluorocarbon polymer (Figure 4.19b). As discussed in Ref. [371], fluorocarbons are not volatile in pure O₂ but can be removed by adding a small concentration of fluorine ions to the plasma. Running the wafer through a CF₄/O₂ cleanup step removes the remaining polymer (Figure 4.19c). We experimented with both SF₆- and CF₄-based cleanup steps

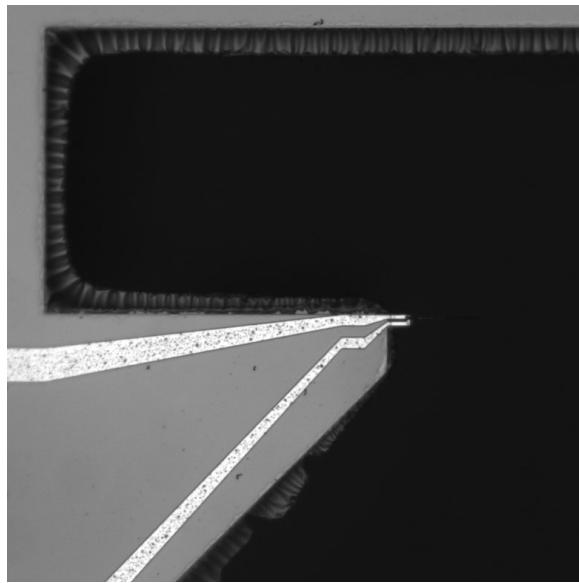


Figure 4.20: Example of oxide flaps after etching the BOX via RIE. The lower etch rate of the BOX near the edges of the DRIE trench led to the formation of oxide flaps which broke off during experiments. The snapped flaps allowed liquid to enter underneath the passivation coating and delaminate it from the device. Releasing the BOX in vapor HF eliminated this problem.

and obtained better results with the latter. The detailed cleanup recipes are in the runsheet (Table E.4). Small flecks of polymer were occasionally left on devices even in the final process that we used, and it could be optimized further.

Etching the BOX via RIE also deposited polymer on the devices (Figure 4.18b). A second, much larger issue with the RIE-based release process was piezoresistor damage. The resistances of the force probe piezoresistors increased substantially over their unreleased counterparts on the silicon die (Figure 2.57). The random nature of the surface damage also resulted in poor mismatch between piezoresistors on the same die. The oxide RIE step only etched silicon at about 28 \AA/min , but it was enough to damage the piezoresistors while simultaneously depositing polymer everywhere.

Until early 2011, the only two options for releasing the BOX were pad etch and RIE. While pad etch did not damage the Al lines, it did break almost all of the force probes on the wafer due to their low spring constants, although a critical point drying (CPD) step could have been added to the process. Fortunately, Gary Yama of Bosch installed a vapor

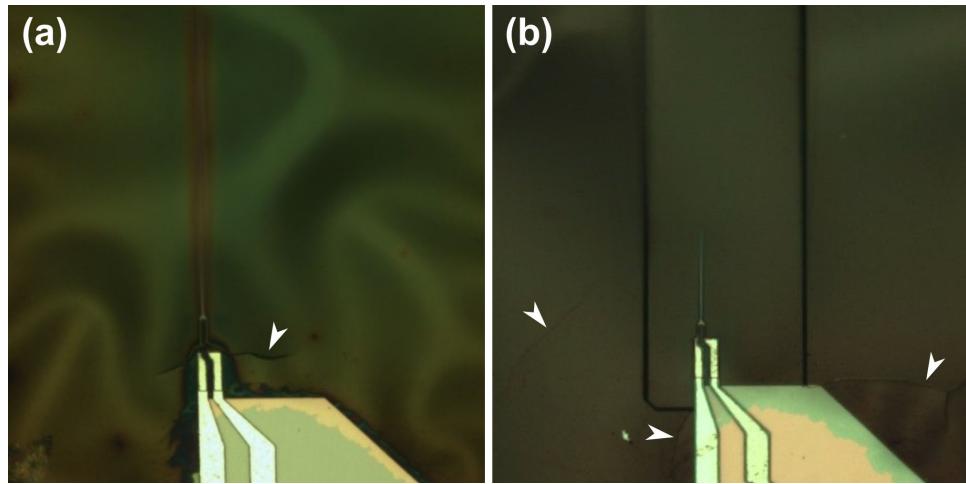


Figure 4.21: Patterned cracks protect force probes from spontaneous BOX cracks. (a) A crack in the BOX passes through the force probe, snapping it off at the end of the Al stiffener. The BOX cracked at this location between 50% (400 nm BOX) and 80% (1 μ m BOX) of the time on the initial wafers. (b) Etching a moat into the BOX stops the BOX from cracking at the end of the stiffener, nearly eliminating the yield problem.

HF system into the SNF that allowed metals. Etching the BOX in vapor HF solved both the polymer and RIE damage problems.

An additional benefit of etching the BOX in HF vapor is that the process is completely isotropic. The RIE process etched the BOX in the center of the release trench much faster than the edges and so BOX flaps remained at the edge of the die after all processing was finished (Figure 4.20). The BOX flaps tended to crash into the sample during experiments, providing a route for liquid to enter underneath the passivation coating and delaminate it from the device. Devices that were released using vapor HF did not have this problem.

4.5.2 BOX cracking and probe yield

The force probe yields from the first few wafers in the second fabrication run were exceptionally low (16-50%). The force probes were typically snapped in half at the end of the Al stiffener (Figure 4.21a). Fortunately, Ginel Hill experienced similar problems with her ultrasoft magnetic resonance force microscopy (MRFM) probes and developed a solution [381]. Etching an isolation trench into the BOX prevented cracks from crossing

Table 4.1: Probe yield improvement with patterned cracks. The force probe yield was mapped across 10 test wafers via optical microscopy. The average yield of the initial batch of wafers was 35%. Adding a patterned crack to the BOX roughly doubled the yield to 69%. By the end of the fabrication run the typical wafer yield had increased to roughly 90%. Table adapted from Ref. [381].

Wafer	Unpatterned BOX	Patterned crack
1	21%	73%
2	47%	63%
3	38%	90%
4	50%	51%
5	16%	-
6	40%	-
Mean	35%	69%

the force probe (Figure 4.21b).

The crack is 2 μm wide and the image on the reticle extends completely around the longest force probe design on the reticle. The crack was etched using a via etch RIE program with relatively poor selectivity to silicon. The etch terminates on the handle wafer, so selectivity is not an issue, and the low polymer deposition rate due to the high O₂ flow rate provides straighter sidewalls than the more selective oxide etch program. The crack was overetched by 100% based upon the nominal etch rate in order to ensure that the oxide was completely removed.

The force probe yield was carefully measured on 10 test wafers in order to assess the effect of the crack moat. The results are presented in Table 4.1. The crack moat roughly doubled the probe yield, although by the end of the fabrication run most of the wafers had approximately 90% yield due to subtle improvements in wafer handling.

4.5.3 Sheet resistance and contact resistivity issues

The first batch of wafers that was processed received a single 775°C, 35 minute POCl₃ diffusion to the contacts and piezoresistor. During metallization they were heated in the loadlock to drive off moisture and then coated with a 1 μm thick layer of Al. At the end of processing they received a 30 minute long FGA at 425°C.

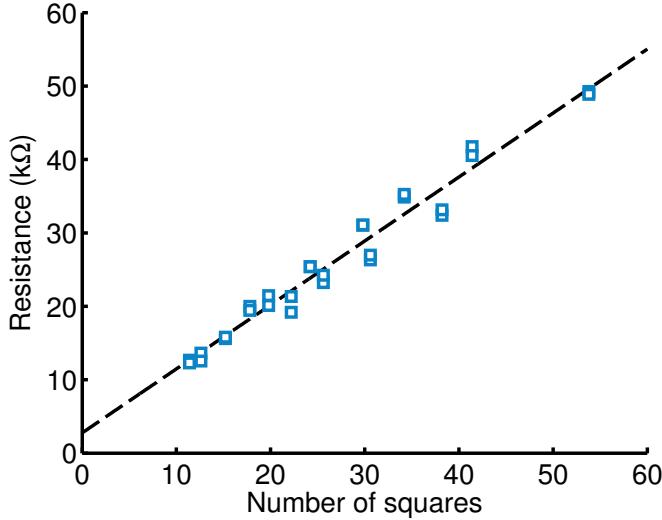


Figure 4.22: TLM data from the first batch of second generation wafers. The piezoresistor resistance at 1V bridge bias is plotted as a function of the number of squares in the resistor. The measured sheet resistance is $790 \Omega/\square$ compared with $190 \Omega/\square$ measured on test wafers, while the excess resistance from the contacts is $3.4 \text{ k}\Omega$ in contrast with our initial prediction of 64Ω .

Figure 4.22 presents TLM data from one of the SOI wafers processed in the first batch. The resistances were measured on a probe station after the metal interconnects were patterned but before backside and BOX were etched. The wafer exhibits higher than expected sheet resistance and contact resistances. Unpatterned monitor wafers that were processed alongside the device wafers indicated a piezoresistor sheet resistance of $190 \Omega/\square$ based upon 4PP measurements. In contrast, the sheet resistance of the patterned piezoresistors was $790 \Omega/\square$. Assuming a surface dopant concentration of 10^{20} cm^{-3} we would have expected a contact resistivity of $0.5 \mu\Omega\text{-cm}^2$ and for the contacts to contribute 64Ω of excess resistance to the system. In fact, they contributed $3.4 \text{ k}\Omega$ of excess resistance. Clearly, something went wrong.

The increase in sheet resistance suggested that the piezoresistor was uniformly damaged during the process performed after the diffusion. After running test wafers to investigate the possible culprits, the sheet resistance increase was traced to the LTO deposition after frontside etching. During that step the bare piezoresistor is exposed to the $300\text{-}400^\circ\text{C}$ oxidizing atmosphere of the furnace for several minutes before the silane begins flowing.

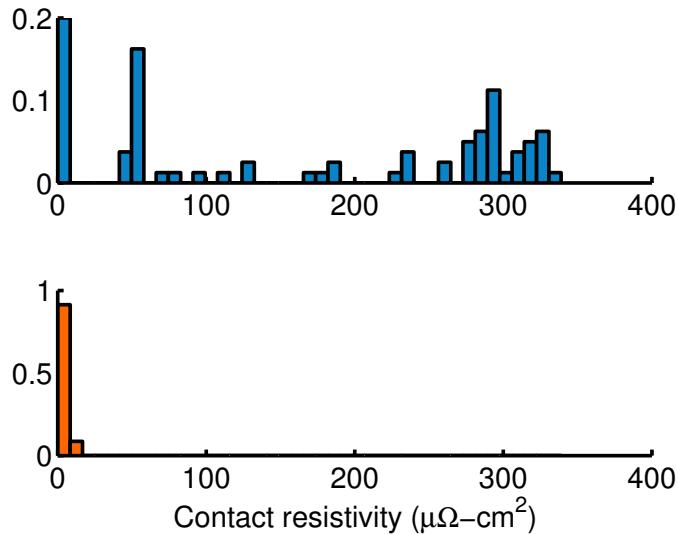


Figure 4.23: Contact resistivity probability distributions for SOI (top) and POI (bottom) wafers doped at 775°C for 35 minutes. Contact resistivities were measured from Kelvin bridge test structures.

We measured the device layer thickness of wafers via ellipsometry before and after LTO deposition in order to measure the thickness of surface silicon oxidized at the beginning of the LTO deposition. The theoretical and experimental sheet resistance data as a function of the thickness of material removed from the surface were plotted in Figure 2.56 for samples doped at 775°C for 25, 35 and 40 minutes. Depending on the deposition recipe, between 2 and 10 nm of silicon were oxidized at the beginning of the LTO deposition. Although only several nanometers of material was removed during each step, the impact can be substantial when the piezoresistor is thin.

Unfortunately it is not straightforward to avoid this surface oxidation process. In the final process we increased the piezoresistor doping (825°C for 30 minutes) and allowed for a slight increase in sheet resistance. The devices were originally designed for a sheet resistance of $180 \Omega/\square$, but the final devices had a sheet resistance of $110 \Omega/\square$. In order to maintain a constant temperature increase during operation the bridge bias voltage had to be decreased, shifting the devices slightly off of their optimal operating point. Increasing the number of dopants in the piezoresistor and reducing the bridge bias resulted in lower than intended 1/f noise and a very slight degradation (5-10%) of sensor resolution.

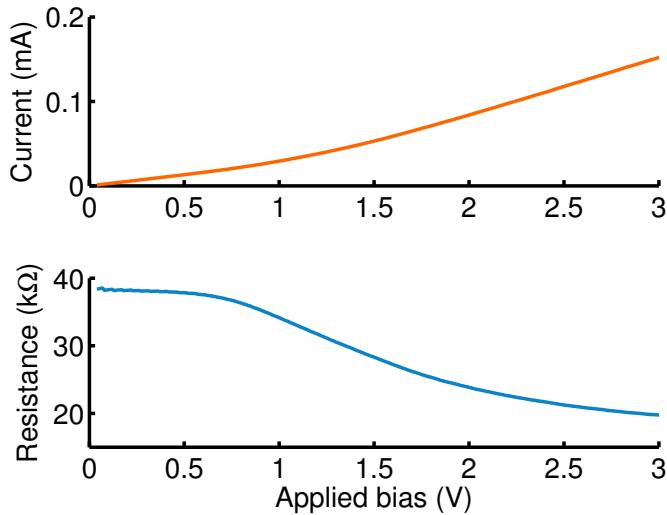


Figure 4.24: Current-voltage data from a released piezoresistor doped at 775°C for 35 minutes. The contact is non-ohmic and the contact dominates the resistance rather than the piezoresistor resistance, which should scale with V^2 due to Joule heating.

The second issue with the first batch of wafers was the higher than expected contact resistance. The Kelvin bridge test structures were used to measure the contact resistivity of the metal-Si contacts. The contact resistivity probability distributions for SOI and POI wafers from the first batch of wafers are plotted in Figure 4.23. The SOI contact resistivities ranged from 20 to 350 $\mu\Omega\text{-cm}^2$ while most of the POI contact resistivities were less than 5 $\mu\Omega\text{-cm}^2$.

The higher than expected contact resistivity could be attributed to either a lower than expected surface dopant concentration or residual oxide/contaminants at the metal-Si interface. The current-voltage characteristics of a typical piezoresistor from the first batch of wafers is shown in Figure 4.24. The piezoresistor in the data has been released from the substrate, and while the resistance should scale with V^2 due to Joule heating, in fact it decreases with increasing bias. The contact is not ohmic and the resistance decreases from nearly 40 k Ω at biases below 1 V to 20 k Ω at a 3 V bias (equivalent to a 6 V bridge bias). A small fraction of the devices improved after a large bias was applied, indicating that there was oxide in the contact that could be electrically broken down. But most devices did not benefit from applying a brief high voltage, suggesting that the problem was caused by a low surface dopant concentration.

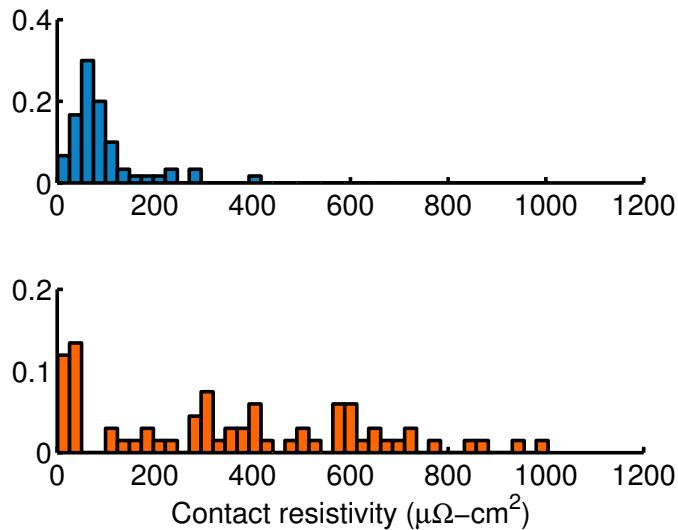


Figure 4.25: Contact resistivity probability distributions before (top) and after (bottom) a 30 minute long FGA at 425°C. The wafers were doped at 775°C for 35 minutes, heated to 200 °C for 3 minutes in the sputterer loadlock and metallized with 1 μm of pure Al.

The large difference in contact resistivity between SOI and POI wafers was traced back to the difference in diffusion rates. The depth of the constant concentration surface layer was extremely thin for the SOI substrates; on the order of 10-15 nm. Popovic experimentally and theoretically investigated the contact resistivity of shallow diffusions by POCl_3 doping samples at 850°C for several minutes or less [382]. He found that the contact resistivity increased sharply when the thicknesses of the diffused layer and depletion region at the Al-Si interface were comparable in thickness. But dopants diffuse much faster in polysilicon; anywhere from 5 to 50 times faster depending on the grain size [236]. The deeper diffused layer in the POI samples rendered them less sensitive to surface oxidation during processing and decreased the contact resistivity.

Some of the wafers from the first batch had been doped but not metallized by the time the high contact resistivity was identified. Without the option of increasing the depth or concentration of dopants underneath the contact, the only option for those wafers was to investigate metallization options. The first batch of wafers received a 30 minute long FGA at 425°C based upon Ref. [383]. But in light of the shallow diffusion and dissolution of Si into pure Al during high temperature processing, we investigated the effect of a FGA

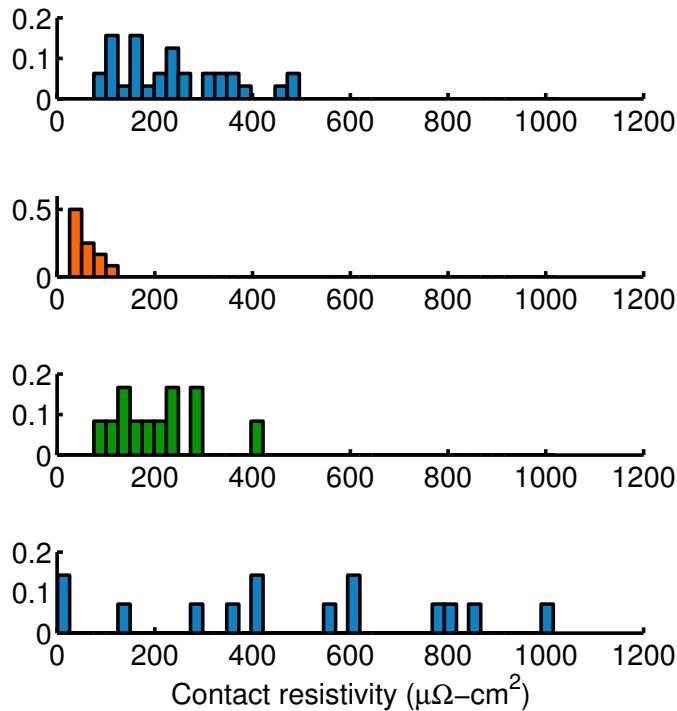


Figure 4.26: Contact resistivity probability distributions for four different metallizations. (a) The wafer is simply coated with 50 nm and 1 μm thick Ti and Al films. (b) The wafer is heated in the sputterer loadlock to 200°C for 3 minutes to drive off moisture before the Ti and Al films are deposited. (c) The wafer is heated in the loadlock but only Al is deposited. (d) After the wafer is heated, the wafer is presputtered in Ar to remove approximately 70 Å of material before the Ti and Al layers are deposited.

on wafers doped at 775°C for 35 minutes. The results are plotted in Figure 4.25. Before the FGA the contact resistivities are normally distributed around 80 $\mu\Omega\text{-cm}^2$. After the FGA the probability distribution is much broader, with contact resistivities as high as 1000 $\mu\Omega\text{-cm}^2$. The 425°C FGA clearly reduced the repeatability and quality of the contacts.

After eliminating the FGA from the process, four different metallization processes were tested on wafers doped at 775°C for 35 minutes. In the first process the wafer was simply coated with 50nm and 1 μm thick layers of Ti and Al. In the second process the wafer was heated in the sputterer loadlock to 200 °C for 3 minutes to drive off moisture before the Ti and Al films are deposited. In the third process the wafer is still heated in the loadlock but only the Al layer is deposited. In the fourth process the wafer is heated and then

presputtered in the loadlock to remove approximately 70 Å of surface material before the Ti and Al films are deposited. None of the wafers received an FGA afterwards. The wafer was maintained at 40 °C during deposition for all of the processes and the base pressure of the system reached 10^{-7} torr before the films were deposited at 3 mTorr.

The contact resistivity results are presented in Figure 4.26. The second process (heat + Ti/Al) yields the lowest and most reproducible constant resistivity results. The first and third processes (Ti/Al and heat + Al) yield similar results while the fourth process (heat + presputter + Ti/Al) yields the worst results.

The Ar sputtering is theoretically intended to remove any native oxide from the surface. However it also removes precious dopants from the surface of the contact region and damages the crystal lattice. Both presputtering and preheating were reported to increase contact resistivity in Ref. [384]. Vossen found that presputtering deposited metal oxide and other contaminants from the sputtering chamber on the wafer, while he suspected that preheating caused moisture underneath the wafer to leak out and actually contribute to surface oxidation. While his wafers were situated on a flat surface, allowing moisture on the back to act as a slow leak in the system, ours were held in a carousel which may explain the beneficial effect of preheating in our case.

The Ti coating serves two roles in the process. First, the barrier height from Ti to n-type Si is only 0.5 eV as opposed to 0.7 eV for Al [365]. Second, the Ti serves as a diffusion barrier between the Al and Si which react readily [366,367]. Ti was developed as a contact diffusion barrier for ICs in the 1980s before being replaced with more advanced silicide-based metallization schemes. Without a diffusion barrier, Al reacts readily with Si at high temperature and forms spikes that can protrude several microns into the silicon. In the case of a thin released SOI the spikes can penetrate all the way through device layer as illustrated in Figure 4.27. While we eliminated the FGA from the process flow due to its detrimental effect on the contact resistivity, the wafer is heated to 100-200°C during several O₂ plasma ashing steps performed during the release process and including a contact barrier is worthwhile.

By applying the optimized metallization process (heat + Ti/Al, no FGA) to the first batch of wafers we were able to reduce the contact resistivity substantially. However, the first batch illustrated the challenges in using a shallow, relatively lightly doped

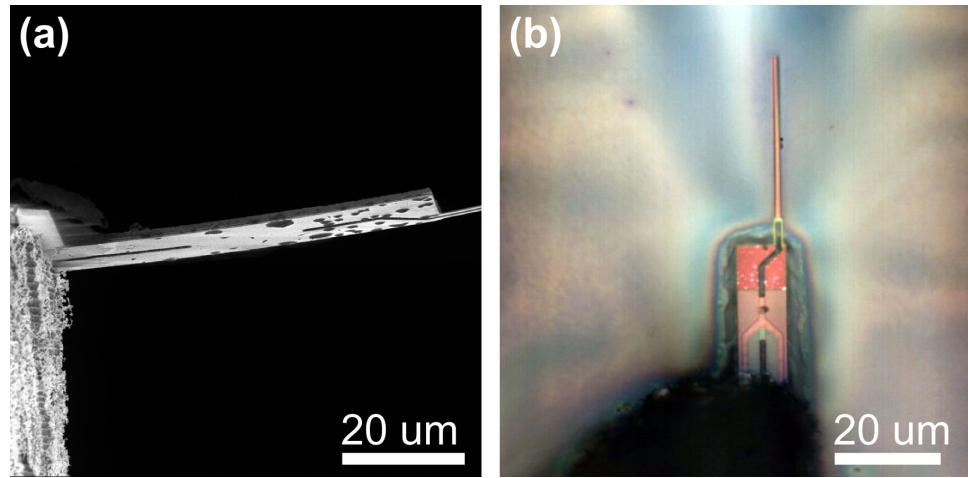


Figure 4.27: Examples of Al spiking through the device layer due to Al metallization and a 30 minute long FGA at 425°C(a) Scanning electron and (b) optical micrographs that show Al spikes protruding through the device layer

piezoresistor. For the remaining wafers (of 5 of the 15 SOIs that were originally purchased from Soitec) we developed the two-step doping process described in the main process flow and runsheet.

In the first version of the second generation process flow only a single diffusion was performed to form the piezoresistor and contact regions. The one-step diffusion process is shown in Figure 4.28a-e. After the thermal oxide mask is grown and annealed (Figure 4.28a), a photoresist mask is lithographically patterned (Figure 4.28b), the photoresist is treated with O₂ plasma and the oxide is wet etched to open the diffusion windows (Figure 4.28c). Both the contact and piezoresistor regions are doped at the same time. During the diffusion a thin layer of PSG is deposited on the wafer surface and serves as a constant concentration source of phosphorus. The phosphorus diffuses into both the SiO₂ and the Si, turning some of the SiO₂ into PSG as shown in Figure 4.28d. Finally, the PSG and remaining SiO₂ mask are stripped (Figure 4.28e).

In order to mitigate the problems that we have discussed in this section a two-step diffusion process with separate contact and piezoresistor doping steps was developed. By decoupling the two steps the contact would be heavily doped without degrading the piezoresistor sensitivity.

In the two-step process the thermal oxide is grown and additional LTO is deposited

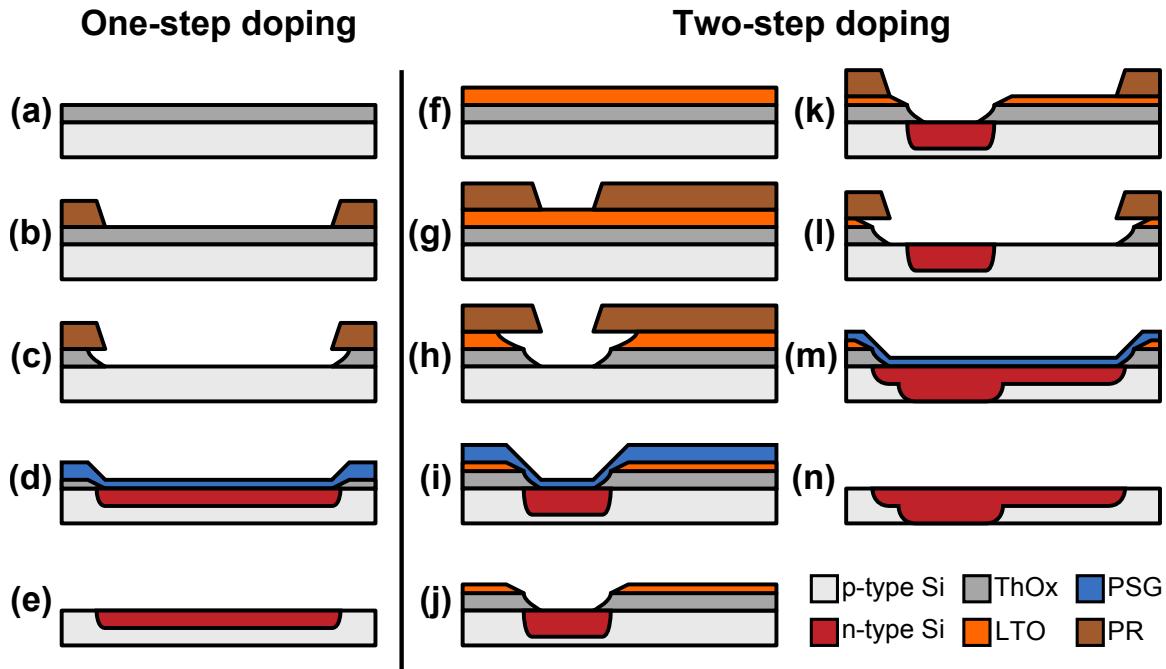


Figure 4.28: The first second generation devices were fabricated with a one-step diffusion process in which (a) a thermal oxide mask is grown, (b) the photoresist mask is patterned, (c) the oxide mask is opened for the entire piezoresistor region, (d) the wafer is doped and (e) the PSG and oxide mask are stripped. The one-step process resulted in a thin, lightly doped piezoresistor and contact which caused problems later in the process. In the two-step process (f) a thicker oxide mask is used and (g) only the contact regions are patterned and (h) opened. (i) The contacts are heavily doped and (j) the fast etching PSG is stripped to reduce mask undercut. Next the piezoresistor mask is used to (k) pattern and (l) open the entire piezoresistor window. (m) The piezoresistor is doped and (n) all of the remaining oxide is stripped, resulting in a heavily doped contact and shallow piezoresistor.

before the wafer is annealed (Figure 4.28f). The additional LTO is necessary in order to ensure that the oxide is thick enough to mask the diffusion without thinning the SOI device layer any further (Section 2.5.2). The photoresist is spun and lithographically patterned by using the contact via mask originally intended for the formation of the metal-Si vias (Figure 4.28g). The densified LTO and thermal oxide are wet etched in order to open the contact windows (Figure 4.28h). The photoresist is stripped, the wafers are diffusion cleaned and the contacts are doped (Figure 4.28i).

After the first diffusion the PSG is stripped by dipping the wafers in a dilute HF solution

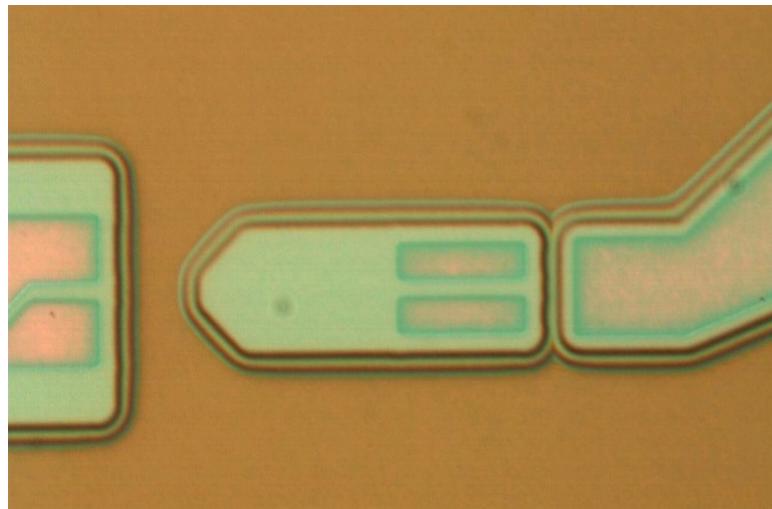


Figure 4.29: A PRT test wafer shown after the oxide mask has been opened for the piezoresistor diffusion. The PSG layer was not stripped after the contact diffusion, resulting in a 5 μm photoresist mask undercut and 2 μm larger piezoresistor regions than intended.

(Figure 4.28j). Removing the surface PSG layer between the diffusions is absolutely crucial due to the high etch rate of PSG ($> 3000 \text{ \AA/min}$) in HF. If the PSG is not removed then the photoresist mask is undercut in the next step and the oxide windows become several microns larger than intended. The process window is extremely narrow at this point due to roughly 6 μm gap between adjacent diffusions in the PRT and PRPE probe designs. A 30-40 second dip in 50:1 HF removes the PSG while only removing several nm of the oxide mask, and can be timed based upon large diffusion windows becoming hydrophobic. A PRT test wafer where the PSG layer was not stripped is shown in Figure 4.29 after opening the piezoresistor window in the oxide mask. The PSG undercut resulted in the piezoresistor and heaters being 2 μm larger than intended.

After the PSG is stripped a layer of photoresist is spun and lithographically patterned with the piezoresistor mask (Figure 4.28k). The remaining oxide mask is wet etched to define the piezoresistor region (Figure 4.28l). A second diffusion forms the piezoresistors and drives the contact dopants further into the wafer (Figure 4.28m). Finally, all of the remaining oxide is stripped from the wafer (Figure 4.28n).

Test wafers were processed with the optimized metallization process (heat + Ti/Al, no FGA) and the three different diffusion processes. In the first process a single 45 minute long

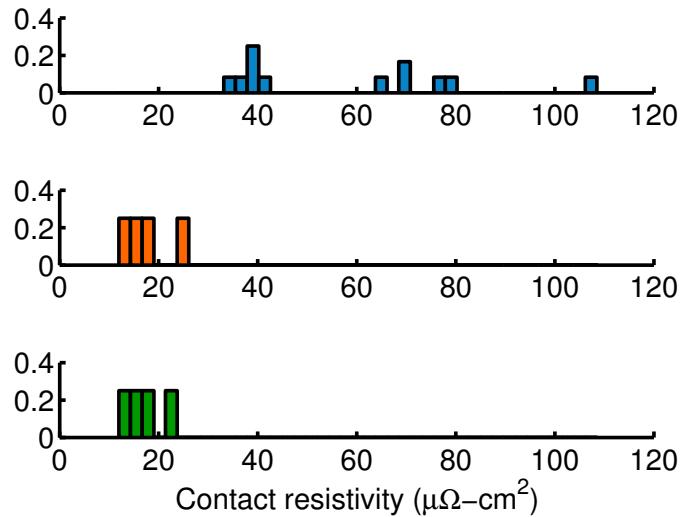


Figure 4.30: Contact resistivity probability distributions for the optimized metallization process and three different diffusion processes. The piezoresistors were doped at (top) 775°C for 45 minutes, (middle) 825°C for 30 minutes or (bottom) 920°C for 20 minutes. The contact resistivity is comparable for the latter two processes, although the lower contact sheet resistance of the third process yields a lower overall contact resistance from the 1D contact transmission line model.

775°C diffusion is performed while the second process uses a single 30 minute long 825°C diffusion. The third process used a 20 minute long contact diffusion at 900°C followed by a 30 minute long diffusion at 825°C to form the piezoresistor. The 825°C diffusion yields a surface concentration of approximately $1 \times 10^{20} \text{ cm}^{-3}$, sheet resistance of $110 \Omega/\square$ and constant concentration surface layer of approximately 50 nm. The 900°C diffusion yields a surface concentration of $4 \times 10^{20} \text{ cm}^{-3}$, sheet resistance of $20 \Omega/\square$ and surface layer of about 150 nm.

The contact resistivity probability distributions for the three processes are presented in Figure 4.30. Contact resistivities of about $20 \mu\Omega\text{-cm}^2$ are obtained for the second and third processes while the first process has a wider data spread. Based upon the results one would be tempted to simply use the one-step diffusion process at 825°C. However, the net resistance contribution depends on both the contact resistivity and sheet resistance of the underlying silicon (Section 3.4.3). Assuming a contact resistivity of $20 \mu\Omega\text{-cm}^2$ and contact dimensions of $3 \times 16 \mu\text{m}$ the 825°C process yields a total resistance of 142Ω per

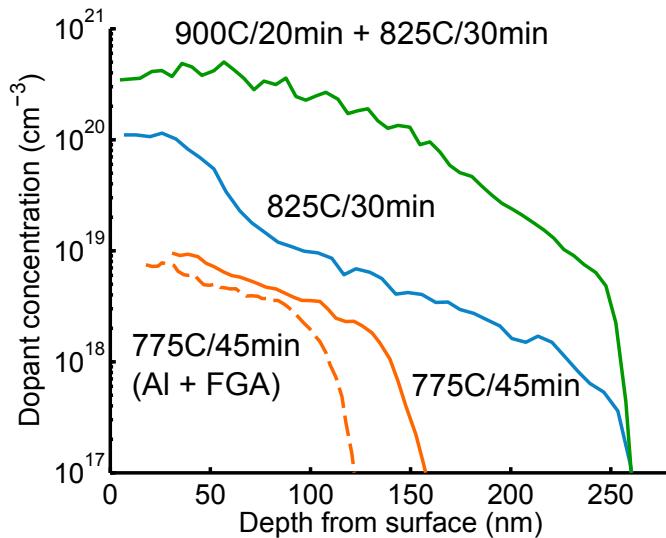


Figure 4.31: Electrically active dopant concentration profiles for three different diffusion processes. Two curves are plotted for the 775°C process. The dashed line corresponds to a sample that was metallized with pure Al and received a 425°C FGA. The three solid lines all correspond to samples that were not metallized.

contact while the 900°C + 825°C process yields a resistance of just 72 Ω. Additionally the isolation well can be doped at the same time as the contacts, and the 5-fold reduction in the isolation well sheet resistance leads to a reduction in actuator-sensor crosstalk. The average contact resistivity of the POI test wafers processed with the 900°C contact doping step was $0.2 \mu\Omega\text{-cm}^2$, about 100 times lower than their SOI counterparts. The cause of the discrepancy is unclear because the electrically active and total dopant concentrations are similar for both processes based upon SIMS and SRA data.

The SRA dopant concentration profiles for the three difference processes are summarized in Figure 4.31. The electrically active surface concentrations of the 900°C, 825°C and 775°C diffusion processes are approximately 4×10^{20} , 1×10^{20} and 1×10^{19} , respectively. Two curves are plotted for the 775°C diffusion. The solid line represents the piezoresistor while the dashed line represents the contact region. The latter sample was coated with Al and received a 30 minute FGA at 425°C before the metal is wet etched from the surface. Approximately 30 nm of Si is consumed due to its dissolution into the Al during the FGA. It is surprising that the contact resistivity of the 825°C and 900°C samples were nearly

identical despite the higher surface concentration of the latter process, suggesting that some other factor was limiting the contact resistivity. The low surface concentration of the 775°C doped samples confirms that the high contact resistivity of the initial diffusion process was caused by a low surface concentration.

The devices that will be discussed in the following chapter were fabricated with the optimized process, consisting of a two-step 900°C + 825°C diffusion, Ti/Al metallization with loadlock heat and no FGA step. We will delay discussing their electrical performance until the next chapter.

4.5.4 Assorted etch issues

Patterning the piezoelectric actuator via RIE worked exceedingly well compared with wet patterning (Appendix F). However, there were several noticeable issues worth commenting on.

First, the Cl-based AlN etch stops on the bottom Mo electrode but only has an AlN:Mo selectivity of 2.3:1 (Table E.12). The process was monitored via OES, however some overetch was both inevitable and necessary in order to reduce the presence of AlN stringers. A scanning electron micrograph of the etched stack is seen in Figure 4.32. The bottom Mo electrode is thinned by 12 nm in most places and up to 25 nm near the edge of the photoresist due to microtrenching. The thinned bottom electrode slightly increases its impedance but has no impact on actuator performance and negligible impact on actuator-sensor crosstalk.

Despite the slight overetch and electrode thinning the AlN etch process left stringers on the wafer. Stringers are a risk for any device that combines conformal deposition steps, substantial topography and anisotropic etch processes. The PE actuator stack is deposited on 300 nm of topography (the patterned SOI device layer). If the etch were perfectly anisotropic and no overetch was performed then the sidewall stringers would be 300 nm in height. An example of AlN stringers is seen in Figure 4.33. On some devices the stringers simply fell off when the LTO layer separating the stringers from the Si was removed, but on many devices the stringers and residual oxide or polymer debris remained, leading to a slight reduction in resonant frequency and increase in spring constant.

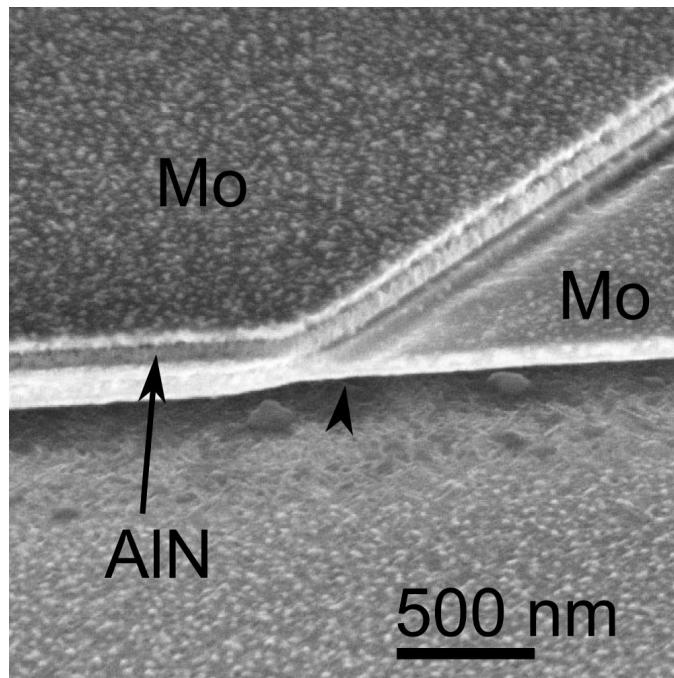


Figure 4.32: The AlN actuator etch stops on the bottom Mo electrode, resulting in electrode thinning and microtrenching. The bottom Mo electrode is thinned by about 12 nm far away from the photoresist mask and up to 25 nm near the mask due to microtrenching (indicated by the arrow head). The sample surface is rough due to the incomplete removal of the encapsulation oxide layers at the end of the process.

As noted in the main description of the fabrication process (Section 4.4), stringers were also an issue for the Al interconnect layer. The Al layer is 400 and 1000 nm thick for the PRPE and PR/PRT devices, respectively, and like the AlN layer, is deposited after the SOI device layer has been patterned. As in the case of the AlN etch, the overetch amount was balanced between the competing demands of removing the stringers and not breaking through the underlying LTO etch stop. Overetching the Al layer by 10-15% was sufficient to remove the stringers for most but not all devices (Figure 4.34). On some devices the stringers even shunted current around the piezoresistor, but applying a brief pulse of several volts was sufficient to vaporize them, and Al stringers were not an issue on the wafers used for performance characterization.

One strange and notable observation during the process was that wafers should be stripped immediately after POCl_3 doping. If they are stored in air for several hours after

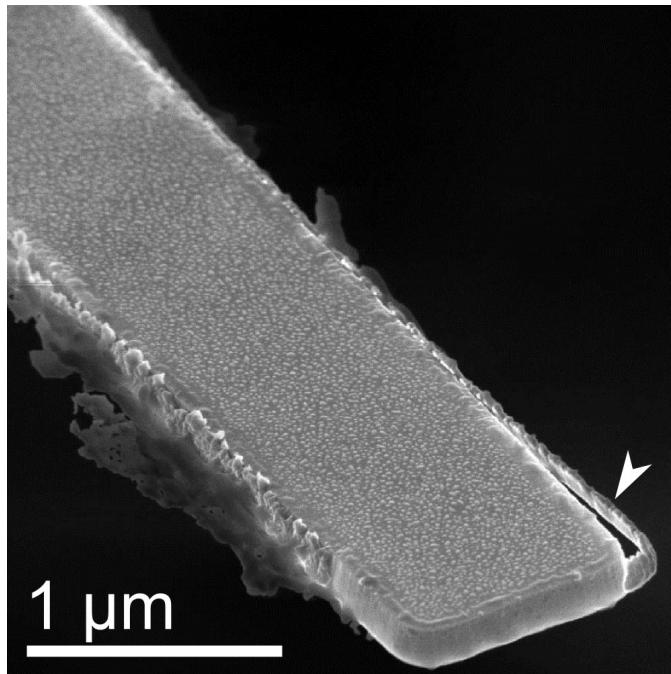


Figure 4.33: Example of AlN stringers left on a finished force probe. The stringers are separated from the Si by a thin layer of LTO, which when released, separates the stringer from the device (see arrow head). On devices where the encapsulation LTO layers were underetched (as evident by the spottled Si) the AlN stringers would not be released.

the doping step but before the PSG and oxide mask are stripped then droplets will form on the wafer (Figure 4.35). The droplets are probably either H_3PO_4 or HCl from residual $POCl_3$ or Cl on the wafers surface reacting with moisture in the air. Promptly stripping the PSG after the diffusion prevented droplet formation.

Thin force probes and SOIs can be a challenge to work with between their mechanical compliance and the delicate films that they are composed of. One major benefit, however, is that they change color during processing. An example of a PRT wafer after vias have been etched in the LTO and before metallization is shown in Figure 4.36. The undoped device layer (green), exposed BOX (blue), piezoresistor (orange) and contacts (pink) are all visible, making it easy to optically monitor the fabrication process.

But why does the piezoresistor appear a different color from the rest of the undoped device layer? The piezoresistor is slightly thinner. The $POCl_3$ diffusions are performed in an oxidizing atmosphere and exposed silicon is thinned due to oxidation. Oxygen is

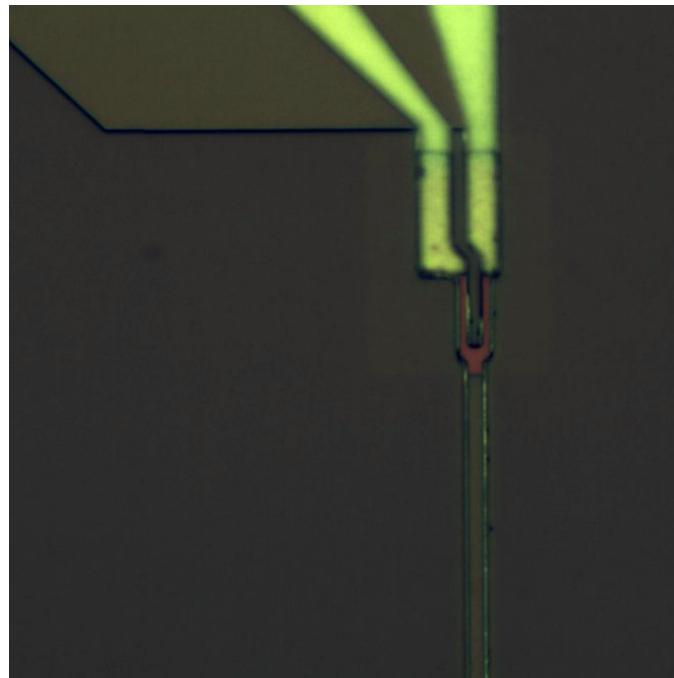


Figure 4.34: An example of stringers after the interconnect etch step. Without any overetch the metal on the device layer sidewalls remains intact, but too long of an overetch would break through the LTO layer protecting the piezoresistor. Increasing the LTO layer thickness increases the process window but degrades PRT and PRPE device performance.

necessary to form the PSG surface layer that provides the solid-state source of phosphorus atoms. While the thinning of the piezoresistor would be insignificant for thicker devices, it noticeably affects the mechanics of the force probes.

Diffusion thinning was measured via variable angle spectroscopic ellipsometry for SOI and POI substrates for the finalized piezoresistor (825°C for 30 minutes) and contact (900°C for 20 minutes followed by 825°C for 30 minutes) diffusion processes. The optical measurements were confirmed mechanically by profilometer measurements in order to ensure that thickness changes rather than the high doping levels were responsible for the changes. The data is summarized in Table 4.2. The SOI substrates were surprisingly thinned more than the POI wafers during the process, which normally oxidize more rapidly. Diffusion thinning should be considered during the design of thin piezoresistive force probes.

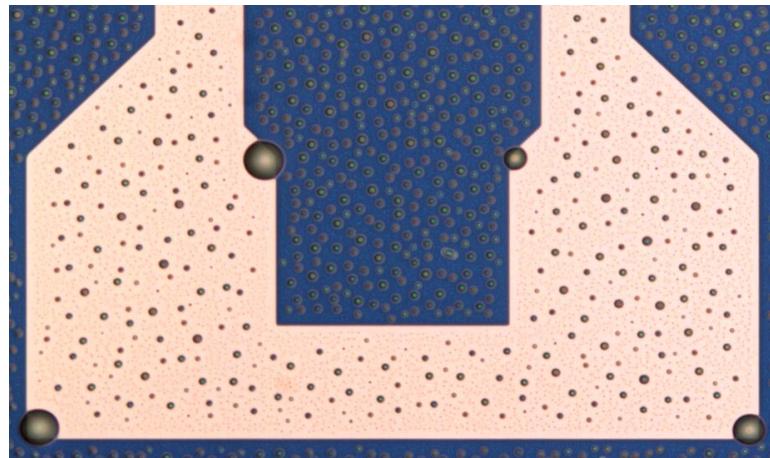


Figure 4.35: Condensation on POCl_3 doped wafers after several hours in air. The droplets are probably either H_3PO_4 or HCl from residual POCl_3 or Cl on the wafers surface reacting with moisture in the air. Promptly stripping the PSG after the diffusion prevented droplet formation.

Table 4.2: Piezoresistor and contact thinning data. The POCl_3 diffusions are performed in an oxidizing atmosphere and exposed silicon is thinned due to oxidation. Diffusion thinning was measured for SOI and POI samples for both the piezoresistor (825°C for 30 minutes) and contact (900°C for 20 minutes followed by 825°C for 30 minutes) processes. The thickness change was measured by variable angle spectroscopic ellipsometry and profilometry. Five samples were measured for each condition and the data is reported as $\mu \pm \sigma$.

Substrate	Diffusion	Thickness reduction (nm)
SOI	piezoresistor	30.8 ± 2.0
	contact	65.2 ± 2.4
POI	piezoresistor	21.7 ± 0.6
	contact	54.2 ± 1.2

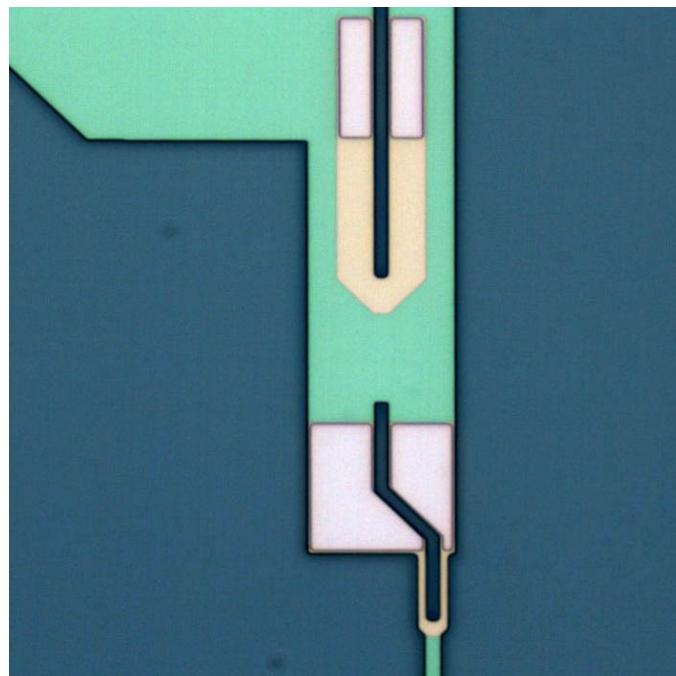


Figure 4.36: Examples of color changes in processing thin SOI wafers. The undoped device layer (green), exposed BOX (blue), piezoresistor (orange) and contacts (pink) are all visible. The PRT wafer is seen immediately before metallization and after vias have been etched in the LTO. A single-step doping process was used for this wafer; the contacts appear pink rather than orange due to the opened vias.

Chapter 5

Characterization

In this chapter we will discuss everything that happens once the wafers leave the cleanroom and before they are mounted on a microscope for experiments. Along the way we will compare the experimental results with the predictions from our cantilever performance model. After discussing the overall mechanics of the probe we will focus on the sensor and actuators individually before returning to the overall performance of the integrated actuator-sensor system.

5.1 Finished devices

When the wafers leave the cleanroom, the force probes are still held in place by snap tabs. The devices are snapped out of the wafers and stored in gel packs until they are examined under an optical microscope in order to screen low quality devices. The most common issues that prevent a device from being used are excess polymer buildup or a snapped off main or compensation force probe. Optical micrographs of finished PR, PRPE and PRT devices are shown in Figure 5.1. Scanning electron micrographs are presented in Figure 5.2.

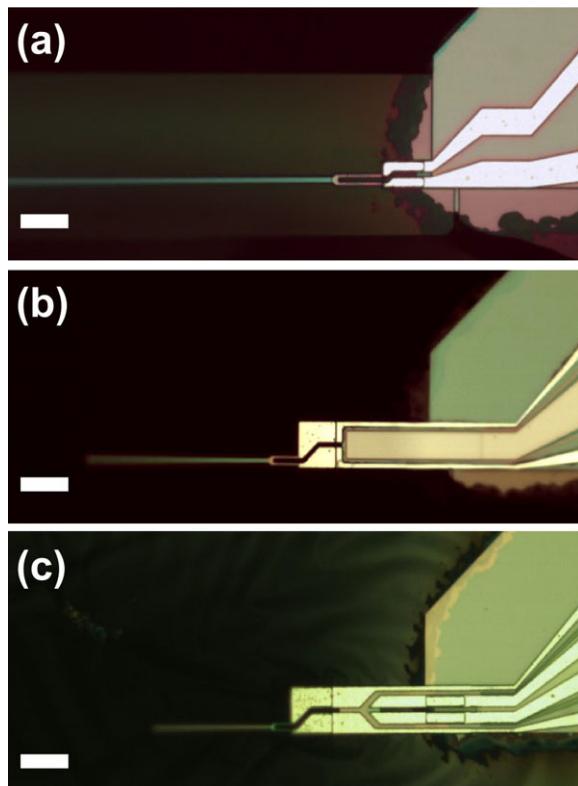


Figure 5.1: Optical micrographs of (a) sensor-only, (b) piezoelectrically actuated and (c) thermally actuated devices. The sensor-only and thermally actuated probes have not been released from the BOX yet while the piezoelectrically actuated probe is completely finished. All of the scale bars represent 20 μm .

5.2 Die attach and wirebonding

After the devices are examined, they are removed from the gel packs and attached to a custom printed circuit board (PCB) using a two-part 5-minute epoxy (Devcon, Danvers, MA). A finished, released device is seen in Figure 5.3. The bondpads on the device are connected to the PCB using an ultrasonic wedge bonder (Westbond 7400A in the Moler Lab or Westbond 7400E in the Kenny Lab at Stanford) and Al wire. The PCB is coated with immersion gold in order to allow for easy wirebonding.

The PCB, like the silicon die itself, is custom designed to fit under an upright microscope objective. The layout of the device PCB is shown in Figure C.3 in Appendix C. Two version of the PCB were built. The first generation board did not include any

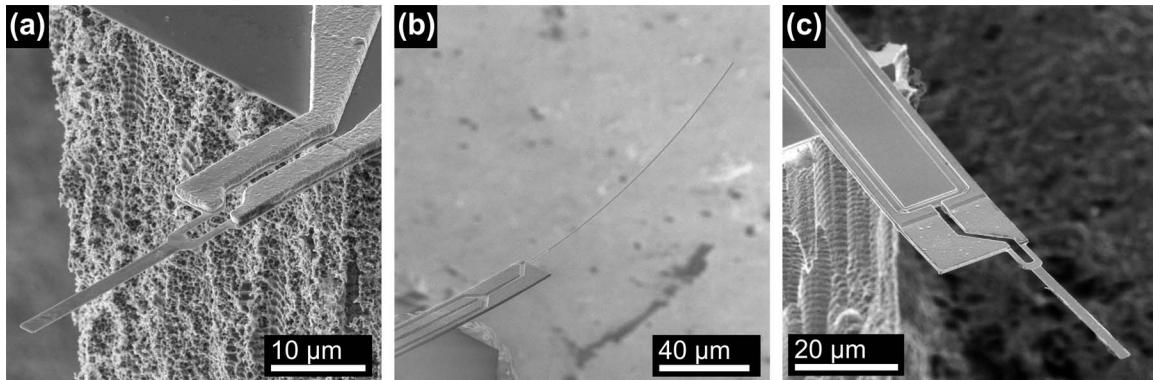


Figure 5.2: Scanning electron micrographs of completed (a) sensor-only, (b) thermally actuated and (c) piezoelectrically actuated force probes.

shielding, which led to more capacitive crosstalk between the actuator and patch clamp electrode during hair cell experiments than was acceptable. The second generation board was fabricated using a four-layer process and routed the signal layers between two ground planes, solving the capacitive crosstalk problem.

While the Al bondpads were trivial to wirebond, the Mo bondpads were much less so. Due to the thin Mo (500 Å) and AlN (750 Å) layers the top and bottom PE electrodes were shorted together about 60% of the time on the first devices that were wirebonded. If either PE actuator is shorted out then the capacitive and mechanical crosstalk compensation is not effective. If only 40% of the bonds are successful then just 16% of the devices will actually work as intended.

Our first attempt at solving the problem was using thermosonic or thermocompression wedge bonding. In ultrasonic wirebonding the Al wire and bondpad are fused together using mechanical energy derived from scrubbing the wedge across the bondpad. Alternatively the bonding energy can be provided by a combination of thermal and mechanical energy (thermosonic) or entirely thermal energy (thermocompression). Both bonding types require the use of Au wire because Al oxidizes too readily. However, the Au wire would not stick to the Mo, even if the bondpad was scratched beforehand.

We returned to ultrasonic bonding and optimized the process to improve the yield. By reducing the bonding power (from 350 to 200 mW), increasing the time (from 25 to 30 ms) and reducing the bonding force (from 50 to 20 gram-force) we were able to

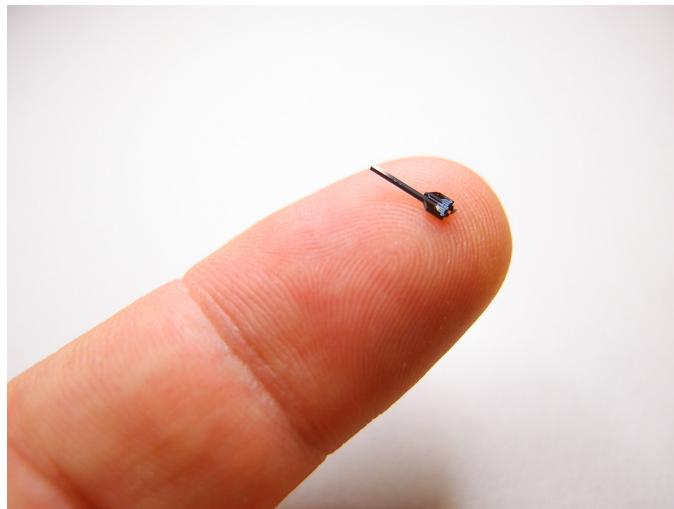


Figure 5.3: Example of a released device before it is mounted on a printed circuit board.

increase the wirebond yield up to 60-80%. While far from ideal, the wirebonding yield permitted the assembly of sufficient piezoelectrically actuated probes for characterization and experiments. It was particularly important to calibrate the force using a gram gauge. After wirebonding the nominal resistance of each connection was measured with a multimeter. The resistance of a good PE connection is on the order of $100\text{ M}\Omega$ while a shorted connection is on the order of $20\text{ }\Omega$. The resistances of the piezoresistors and heaters were likewise recorded in order to select device for further characterization later.

Afterwards the wirebonds are completely covered with epoxy. This step is crucial because the parylene used to passivate the devices for water operation is not thick enough to completely seal the wirebond-bondpad connection. A thin layer of epoxy is usually deposited on the long extension as well. By coating the extension the capacitance to the surrounding liquid is minimized and the exposed area where the passivation could fail is reduced. For all of the epoxy steps a toothpick is used and the epoxy is only worked with for several minutes before it becomes too viscous. While tedious, the yield from this part of the process was about 95% after some practice.

After the device has been epoxied, wirebonded and epoxied the entire device is passivated. A finished, parylene coated device is shown in Figure 5.4. The side of the device appears an iridescent purple color due to optical interference from the parylene

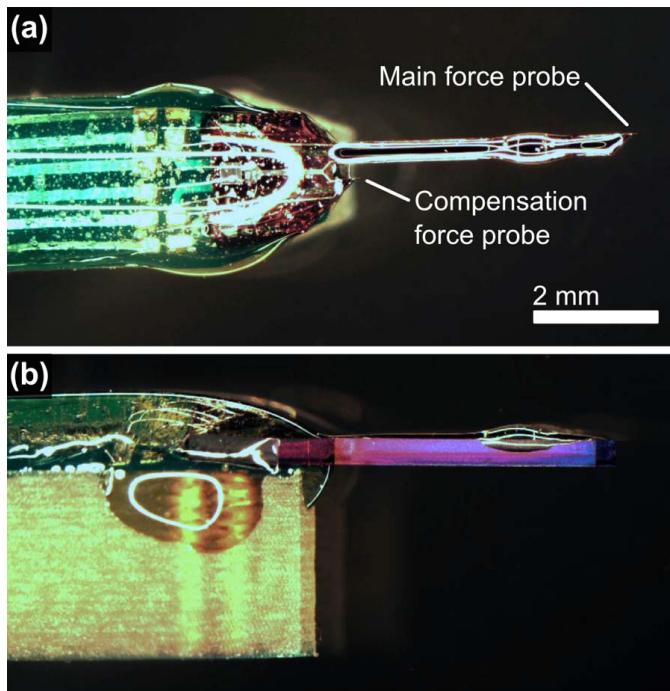


Figure 5.4: (a) Top and (b) side views of a fully assembled device. First, the silicon die is epoxied and wirebonded to the PCB. Next, the wirebonds and long silicon extension are coated with epoxy. Finally, the entire assembly is coated in a thin layer of parylene. The wirebonds are visible through the epoxy. The side of the probe is iridescent due to optical interference from the parylene coating.

coating.

5.3 Passivation

For experiments in liquid the exposed electrodes need to be coated with a thin passivation film. The passivation prevents corrosion and provides DC electrical isolation. Corrosion is a particularly large issue for Al, which corrodes within seconds if electrically biased and unpassivated. Electrical isolation is important for both sensor and actuator operation, and if the devices are being used for electrophysiology experiments, for patch clamp operation. Ref. [385] is an excellent overview of corrosion processes in microsystems and microelectronics.

The passivation film needs to be as conformal, thin and soft as possible. Dipping

the probe into epoxy or silicone would electrically isolate the device but destroy the mechanical performance of the device. We evaluated two options for passivation: atomic layer deposition (ALD) and parylene. 10-50 nm thick ALD oxides have been used to passivate CMOS biosensors [386] and steel samples [387] while parylene has been used to passivate electrodes [388] and other microfabricated devices [154]. Parylene has also been widely investigated as a low-k dielectric for multilevel Cu metallization [389, 390]. While LPCVD silicon nitride [391] and oxide [392] have also been used for passivation, they are deposited at high temperature and are less conformal than parylene or ALD films.

Nanoscale oxide and nitride films can be deposited via ALD; common examples include Al_2O_3 and HfO_2 [393, 394]. ALD enables the deposition of almost perfectly conformal films with digital film thickness control. We performed ALD deposition experiments in a thermal ALD system (Cambridge Nanotech Savannah S200). The chamber was operated at 200°C with Al and Hf precursors of trimethylaluminum and tetrakis(dimethylamido)hafnium, respectively, and water as the oxidizer. ALD materials are relatively stiff and dense. Alumina has an elastic modulus and density of 170 GPa and 3170 kg/m³ while the modulus and density of hafnia are 77 GPa and 6100 kg/m³ [395]. We focused on hafnia ALD coatings due to its relatively low elastic modulus and excellent electrical properties.

Parylene is a soft, clear polymer that is deposited via LPCVD at room temperature (SCS PDS 2010). There are several types of parylene (e.g. parylene N, parylene C) that are all based upon a p-xylylene monomer. We exclusively used parylene N because it has the highest degree of conformality and lowest deposition rate and yield of the Parylene family, making it easy to deposit in thin films with controllable thickness. In order to coat a sample, parylene dimer powder is loaded into a temperature-controlled vaporizer. The heater vaporizes the dimer (150-160 °C) which is pyrolyzed in a furnace (650 °C) and deposited on the sample at room temperature. The vaporizer temperature is controlled to maintain constant deposition pressure. The deposition rate (nm/min), deposition yield (nm/mg) and film crystallinity all depend on the deposition pressure. We typically operated the deposition at 40 mTorr above the system base pressure, yielding a deposition rate and yield of about 10 nm/min and 0.7 nm/mg, respectively. We did not characterize the crystallinity of the deposited films.

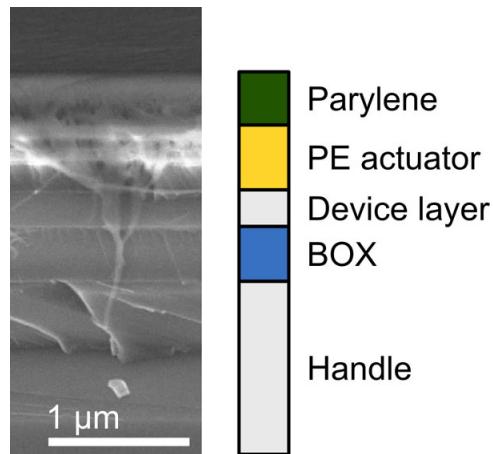


Figure 5.5: SEM cross section of a parylene coated sample. The device is coated with approximately 330 nm of parylene. The cross section is taken through the piezoelectric actuator of a PRPE device. The underlying silicon is 250 rather than 300 nm thick because it is highly doped to form an isolation well. The parylene coating is smooth, uniform in thickness and continuous.

Parylene is far softer (nominally 2.4 GPa) and less dense (1110 kg/m^3) than the ALD films. Variation in mechanical properties, deposition yield and dielectric properties with deposition pressure is discussed in Ref. [390]. Parylene film thicknesses were measured using three different methods: optical reflectometry (assuming $n = 1.651$), stylus profilometry on patterned samples, and cross-sectional scanning electron microscopy. All three techniques agreed to within $\pm 10\%$ and optical thickness measurements were used for most measurements. A SEM cross section through a coated piezoelectrically actuated probe is presented in Figure 5.5.

Coating a force probe with either hafnia or parylene changes its mechanical properties (Figure 5.6). The spring constant increases monotonically with film thickness. However the resonant frequency can either decrease or increase due to the competing effects of mass loading and stiffening, which scale linearly and cubically with passivation thickness. The rate of stiffness change increases as the passivation layer gets further from the neutral axis of the probe. In order to minimize the spring constant increase we targeted hafnia and parylene film thicknesses of 10 to 20 and 200 to 300 nm, respectively.

The relative change in spring constant and undamped resonant frequency with passivation can be calculated from

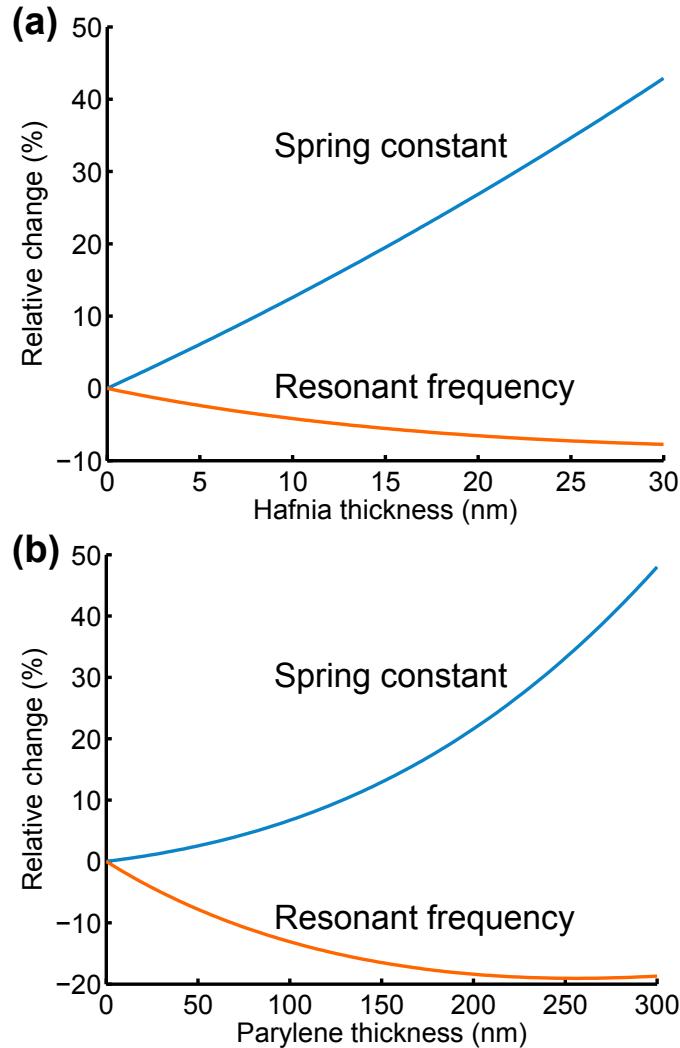


Figure 5.6: Changes in spring constant and resonant frequency with (a) hafnia and (b) parylene N coating. The spring constant increases monotonically for both passivation layers while the resonant frequency first decreases before increasing. In order to minimize the spring constant increase we targeted hafnia and parylene film thicknesses of 10 to 20 and 200 to 300 nm, respectively.

$$\frac{k_c^{\text{coated}}}{k_c^{\text{uncoated}}} = 1 + \frac{2E_p t_p [w_c t_p^2 + 3w_c(t_p + t_c)^2 + (t_c + 2t_p)^3]}{E_c w_c t_c^3} \quad (5.1)$$

and

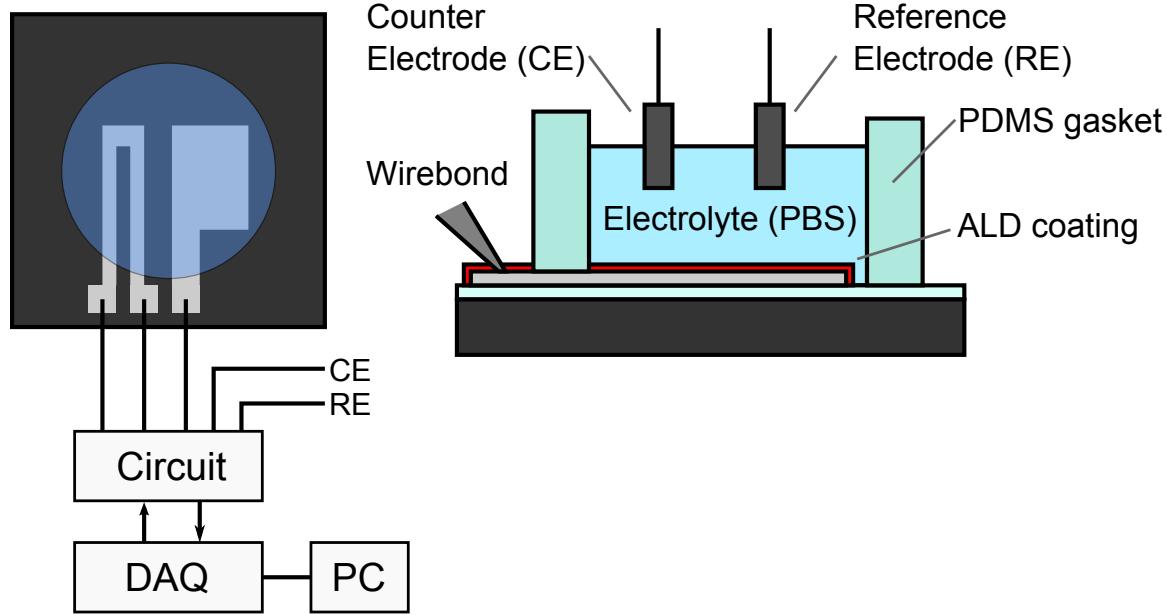


Figure 5.7: Passivation measurement setup. A custom built potentiostat controls the counter, reference and working electrodes allowing the application of a DC bias across the passivation coating while simultaneously measuring the electrical properties of the passivation layer and an on-chip resistor.

$$\frac{f_0^{\text{coated}}}{f_0^{\text{uncoated}}} = \left[\frac{k_c^{\text{coated}}}{k_c^{\text{uncoated}}} \frac{\rho_c w_c t_c}{\rho_c w_c t_c + 2\rho_p t_p (w_c + t_c + 2t_p)} \right]^{1/2} \quad (5.2)$$

where t_p , E_p and ρ_p are the thickness, elastic modulus and density of the conformally deposited passivation coating. The equations assume conformal coating and account for the parylene coating on the sides of the force probe.

The ALD passivation coatings and corrosion resistance tests were performed in collaboration with Alex Haemmerli using a three-electrode measurement setup (Figure 5.7). The potentiostat used to control the counter (Pt wire), reference (Ag/AgCl) and working (a coated test structure or device) electrodes was custom built. The resistance and capacitance of the ALD passivation layer were independently measured by controlling the potentiostat with a large DC control signal (e.g. 1 V) and relatively small (e.g. 20 mV) AC test signal.

We found that ALD coatings were capable of preventing corrosion in wet etched Al

test structures up to approximately 200 mV/nm. But when finished force probes were passivated with ALD they broke down almost immediately for fields as small as 10 mV/nm. For example, the device shown in Figure 5.8 corroded under a 250 mV bias in a grounded PBS solution. Corrosion was accompanied by a sudden change in piezoresistor resistance. The most notable difference between the test structures and the force probes was the use of wet and dry Al etching, respectively. It is plausible that sidewall polymer from the dry etch process led to incomplete ALD sidewall coverage.

Parylene initially also had some performance issues. The first issue we encountered was the quality of the deposited film (Figure 5.9a and b). The coatings were initially rough, porous and did not protect the devices sufficiently from surrounding liquid. In fact, the low quality of the films was the reason for the investigation into ALD passivation coatings. There were two sources of the deposition problems: particles in the chamber, particularly buildup in the pyrolysis furnace, and premature venting of the chamber at the end of the deposition run. At the start of every run I would recommend cleaning the chamber with cleanwipes and a handheld vacuum, pumping and venting the system, and cleaning it once more before loading the samples and dimer.

Once the parylene films were smooth and clear they still had a tendency to delaminate from the devices (Figure 5.9c). Poor adhesion between the parylene and device resulted in a single defect in the coating leading to catastrophic failure. Coating the device with silane A-174 (Sigma Aldrich) before parylene deposition solved the adhesion problem. Parylene passivation with A-174 pretreatment has been employed to passivate microelectronic devices from liquid and humidity since at least 1970 [396]. Hanley and Martin found that A-174, as a Cl-free adhesion promoter, is preferable to other Cl containing promoters which can corrode exposed Al [397]. Hassler et al. investigated the impact of deposition pressure, post processing temperatures and adhesion promoters on the parylene passivation of interdigitated electrode devices [398]. They found an enormous improvement in parylene adhesion, up to 600-fold for parylene on Si_3N_4 , and device stability over time with the application of A-174.

A-174 can be applied in the liquid or vapor phase and we tested both. Liquid deposition requires mixing the A-174 into isopropyl alcohol (IPA) which damaged the epoxy on the devices and broke the softest force probe designs. We opted for in-situ vapor phase

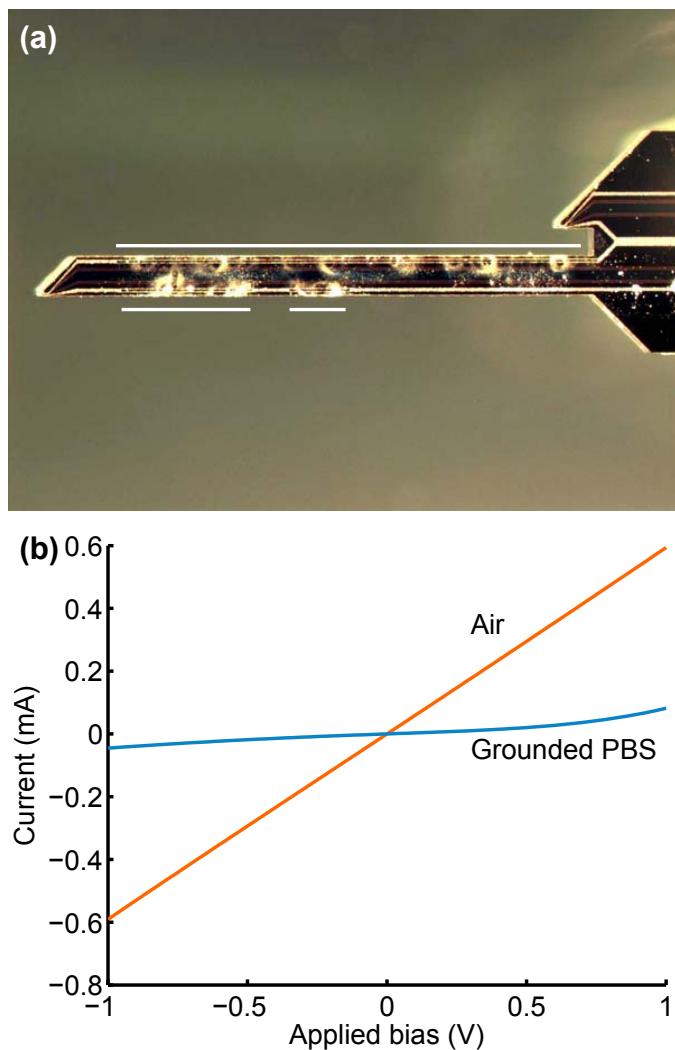


Figure 5.8: Examples of ALD corrosion. (a) A PRPE device coated with 20 nm of HfO_2 corroded when biased to 250 mV in a solution of grounded PBS. The Al interconnects have corroded away in the regions indicated by the white lines. (b) Current-voltage curves in air and grounded PBS for a device coated with 20 nm of HfO_2 .

deposition, accomplished by adding about 0.3 mL of A-174 to the parylene chamber with the devices, pumping it and delaying the deposition until all of the A-174 had evaporated. The chamber pressure is typically 15 mTorr above the base pressure while the A-174 is evaporating, which takes about 30 minutes for 0.3 mL of A-174.

The electrical behavior of a force probe operating in air and grounded phosphate

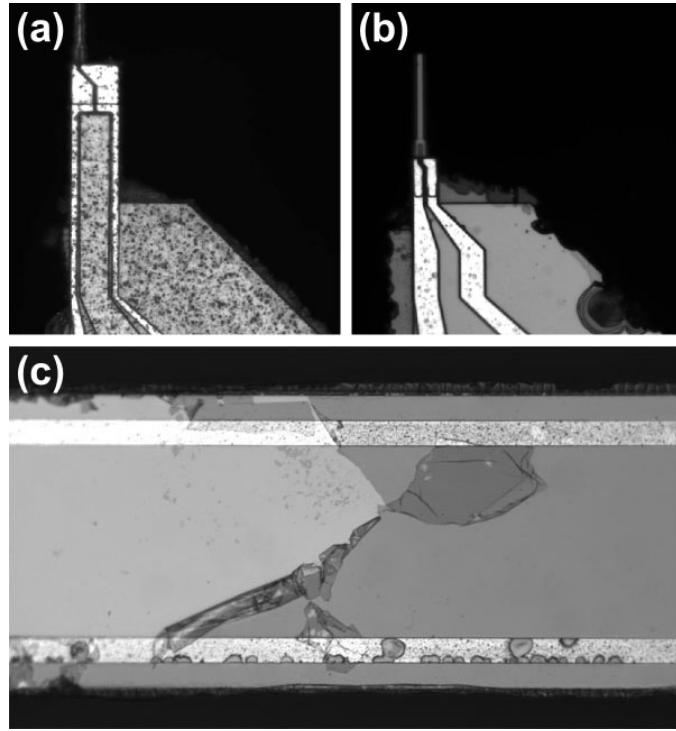


Figure 5.9: Summary of parylene quality control issues. (a) Rough, porous parylene films are deposited if the parylene chamber is not cleaned thoroughly before deposition or if the system is vented too quickly at the end of the process. (b) High quality films are nearly transparent, conformal and pinhole free. (c) Poor adhesion between the parylene and device can result in film delamination in liquid, which can be solved by coating the device with silane A-174 vapor in-situ before the parylene deposition.

buffered saline (PBS) is presented in Figure 5.10. The resistance of the force probe piezoresistor is measured as a function of applied bias from 20 mV to 2 V (equivalent to 4 V bridge bias). The resistance scales quadratically in both cases due to self-heating, illustrating the heat transfer enhancement from operating in water. The small initial resistance mismatch is due to the temperature difference between the air and water.

Many researchers have investigated the impact of high temperature processing on parylene. Depositing the parylene at 100-200°C or annealing it in nitrogen at similar temperatures afterwards decreases its moisture permeability about 10-fold [399]. Wu et al. found that parylene N is stable up to 175°C in air [400], well above the 120°C glass transition temperature of the FR-4 that our PCBs are manufactured from. However, the

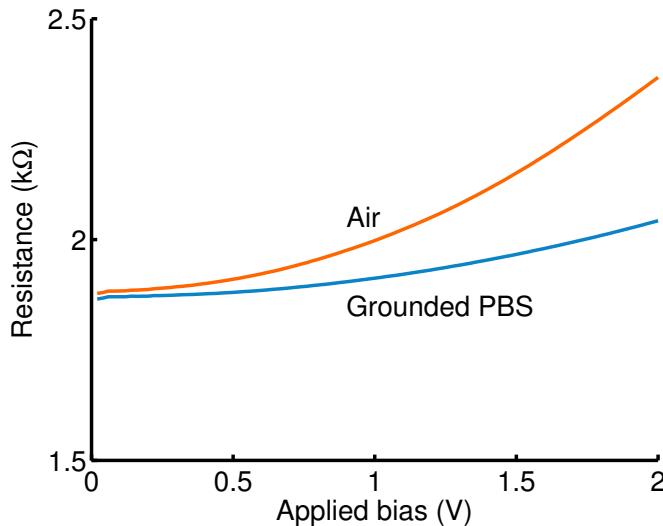


Figure 5.10: Example of device operation in air and liquid with a 300 nm thick parylene passivation layer. The resistance of the force probe piezoresistor is measured as a function of applied bias. The resistance scales quadratically in both cases due to self-heating, illustrating the heat transfer enhancement from operating in water. The small initial resistance mismatch is due to the temperature difference between the air and water.

CTE of parylene N is approximately 69 ppm/K, far higher than the rest of the materials that the force probe is composed off [401]. Huang and Tai investigated the impact of A-174 pretreatment and annealing on parylene-Si adhesion [402]. They found a 15-fold increase in the adhesive force when silicon was treated with A-174 before parylene deposition. However, annealing the sample afterwards reduced the strength of the parylene-Si adhesion 3-fold. Hassler et al. reached a same conclusion in investigating the adhesion between parylene on Si_3N_4 , Pt and Au; A-174 is good, annealing is bad [398]. Considering that annealing induces substantial stress in the parylene and degrades its adhesion to the sample we opted against exploring annealing.

The combination of A-174 adhesion promotion and 200-300 nm thick coatings of parylene N resulted in stable device passivation. The parylene is still relatively soft, however, and can scrape off if the device is rubbed against something hard during experiments. We will address the impact of parylene on probe mechanics at several points in the remainder of this chapter.

Parylene coating has one final and important impact on device performance. Soft,

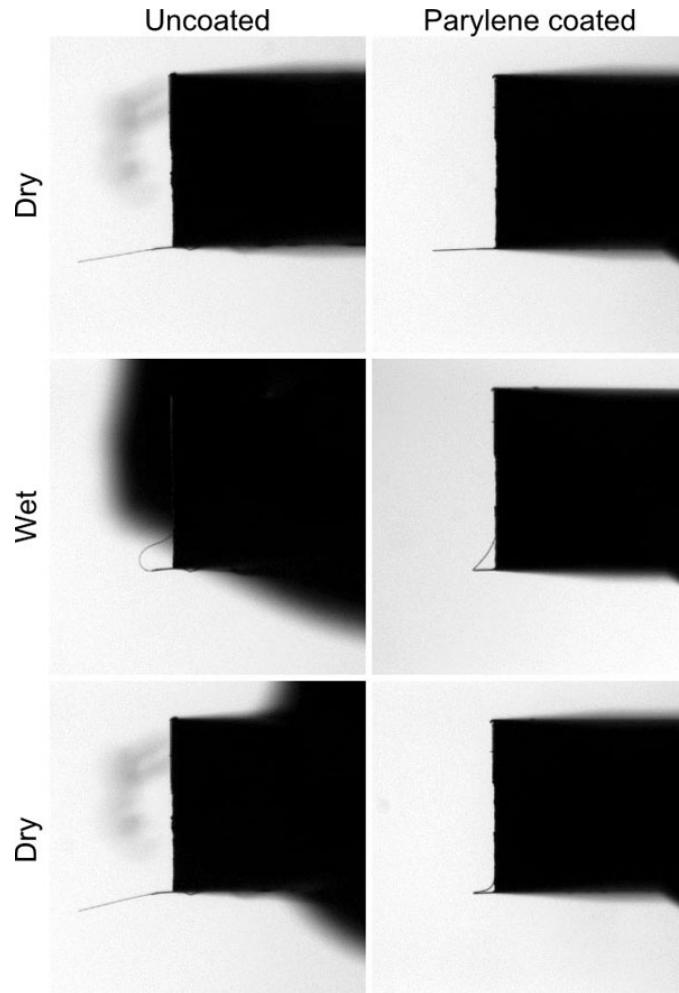


Figure 5.11: Examples of (left) uncoated and (right) parylene coated devices interacting with air-water interfaces. Uncoated devices often stick to the sidewall when removed from water but spring back to their initial shape. Parylene coated devices adhere more strongly to the sidewall and are more likely to remain permanently stuck. When coated devices come unstuck they are likely to remain permanently bent due to plastic deformation of the parylene.

parylene coated devices are far more prone to break at the air-water interface than their uncoated counterparts. In the left column of Figure 5.11 a bare silicon cantilever is brought into water and removed. The device retains its initial shape in air and while submerged in water. But while approaching the air-water interface from the water side, surface tension forces drive the probe to the sidewall of the die. Once removed from water and allowed to

sit until dry the probe springs back and regains its initial shape. Parylene coated devices (right column of Figure 5.11) are more likely to remain stuck to the sidewall. If they do come unstuck, they are often permanently bent, presumably due to plastic deformation of the parylene coating (note the sharp bend at the actuator-sensor interface for the parylene coated device in Figure 5.11).

The stiction problem only affects the softest probe designs (designs 1, 2, 8 and 9 in Table 3.1) and can be mitigated by replacing water with a lower surface tension solvent before removing the probe from the liquid. For example, the surface tensions of water-air and methanol-air interfaces at room temperature are roughly 73 and 23 mN/m, respectively. Heating the solvent decreases the surface tension even further. Methanol or ethanol are ideal for their low surface tension and relative chemical inertness (e.g. isopropyl alcohol and acetone attack the epoxy used to attach the device to the printed circuit board). We found that replacing water with methanol at the end of each experiment and carefully heating it with a hot air gun to speed evaporation allowed probe designs 2 and 9 to be routinely used for experiments.

5.4 Probe mechanics

We will discuss the mechanics of the force probes before moving onto their electrical and actuator performance in the following sections.

The frequency response of a force probe can be measured by either actively oscillating the probe or relying on thermomechanical noise. The benefit of using a forced oscillation is that the deflection of the probe can be large (>10 nm) and the probe can be driven with a chirp (a swept-frequency cosine signal) in order to simultaneously measure the amplitude and phase response of the probe. The problem with applying this approach to the force probes is that macroscale actuators are far from ideal and introduce numerous spurious modes into the measurement. Instead, if the external actuator is turned off, the force probe will oscillate on its own due to the Brownian motion of its constituent atoms, as described by Gabrielson in 1993 [32]. The deflections are far smaller, on the order of $10 \text{ pm}/\sqrt{\text{Hz}}$ for the devices presented here, but the signal is exceptionally free of any other modes. The magnitude of the deflection on resonance will depend on the spring constant and quality

factor of the cantilever and works best for micron-scale devices.

The measurement bandwidth in the case of thermomechanical noise excitation is limited only by the technique used to measure the deflection of the force probe. We measured the force probe deflection via laser doppler vibrometry (LDV) (Polytec OFV-2500). The -3 dB bandwidth of the LDV system was approximately 800 kHz for a velocity sensitivity of 100 mm/s-V. The laser spot is focused through a 100x Mitutoyo air objective with a resulting laser spot diameter of approximately 5 μm . The noise spectral density are averaged 100-200 times and recorded with a GPIB-controlled spectrum analyzer (HP 3562A). For measurements in water, a pair of adapters were purchased from Thorlabs (SM1A28 and SM1A12) in order to mount a standard M25 thread water immersion objective on the M26 threads of the LDV system.

The spring constants of microfabricated cantilevers can be measured using a variety of techniques, as reviewed in Ref. [403]. We will briefly discuss the most common techniques with an emphasis on calibrating small cantilevers. The first and probably most common approach is to simply calculate the spring constant of the cantilever using its nominal dimensions from

$$k_c = \frac{E_c w_c t_c^3}{4l_c^3}.$$

While the elastic modulus of the cantilever is well known if it is made from single crystal silicon and its crystallographic orientation is well known, the dimensions of the cantilever are usually not well known. The thickness is the most important source of spring constant variability for large cantilevers, and all three dimensions are important for our micron-scale force probes. For example, 25% variation in the cantilever thickness leads to 95% variation in its spring constant. Simply calculating the spring constant is not an option for our probes.

The second common method is to measure the resonant frequency of the cantilever and assume that any mismatch in resonant frequency between the predicted and measured resonant frequency is due to variation in the cantilever thickness. This method is commonly referred to as the Cleveland method following Ref. [404]. This approach works well for large cantilevers where the length and width tolerances are less than several percent. In the

case of very small probes (e.g. our 0.5 to 1 μm wide and 3 to 12 μm long piezoresistor legs) neither of these conditions holds. Additionally, any debris on the cantilever or imperfections from the fabrication process (e.g. backside DRIE undercut) will lead to a spring constant underestimate.

A third approach, known as the Sader method following Ref. [405] combines the resonant frequency, quality factor and nominal dimensions of the cantilever in order to provide a better spring constant estimate. However, it suffers from the same problems as the Cleveland method because it relies on accurate knowledge of the cantilever dimensions.

A fourth approach is to deflection the cantilever being studied by a second cantilever with a well known spring constant [406–408]. We have attempted to use this method for calibrating the force probes and ran into two problems. First, the softest commercial available AFM cantilevers have spring constants of roughly 30 mN/m, about 90x stiffer than our softest probes. This large spring constant mismatch leads to difficulty in accurately estimating the spring constant of our force probes. A second problem with this approach is that if the contact point between the two cantilevers slips then the results will not be reproducible. Variability in the inferred spring constants leads to the need to perform a large number of measurements, making this approach both inaccurate and tedious.

Fortunately there is an approach to calibrating the spring constant that is fast, non-contact, repeatable and does not depend on knowledge of any of the dimensions of the force probe. Thermomechanical noise, in addition to being the best approach for measuring the resonant frequency and quality factor of the force probes, is also the best option for calibrating their spring constants.

To briefly summarize thermomechanical noise theory, the force probe is excited by a white force noise. The noise spectral density of the force noise ($\text{N}/\sqrt{\text{Hz}}$) is

$$S_F = \sqrt{4k_b T b} \quad (5.3)$$

in direct analogy with Johnson noise. The force probe is modeled as a simple harmonic oscillator with effective mass, damping and spring constant terms of m_{eff} , b and k_c , respectively. It is convenient to rewrite (5.3) in terms of the resonant frequency (ω_0) and quality factor (Q) using

$$Q = \frac{\sqrt{m_{\text{eff}} k_c}}{b} \quad (5.4)$$

and

$$\omega_0 = \sqrt{\frac{k_c}{m_{\text{eff}}}} \quad (5.5)$$

in order to yield

$$S_F = \sqrt{\frac{4k_b k_c T}{Q\omega_0}}. \quad (5.6)$$

The deflection noise spectral density is calculated from the product of the force noise and transfer function of the force probe as

$$S_X = \sqrt{\frac{4k_b k_c T}{2\pi Q f_0}} \left[\frac{1}{k_c} \frac{1}{(1 - f^2/f_0^2)^2 + (f/Q f_0)^2} \right]. \quad (5.7)$$

The spring constant is calculated by fitting a model to the measured deflection noise spectral density. The model contains three uncorrelated noise terms: the Johnson noise of the LDV system, the 1/f noise of the LDV system and the thermomechanically driven SHO response (5.7). The model is fit to the experimental data in order to minimize the sum of the squared error using `fmincon` in Matlab, an approach that will be used repeatedly throughout this chapter.

The deflection noise spectral density of a typical force probe is presented in Figure 5.12. The solid line represents the entire model fit while the dashed line represents only the SHO component of the model. The deflection resolution of the LDV system scales as 1/f (due to the fact that it measures velocity and is limited by white noise) and is about 1 pm/ $\sqrt{\text{Hz}}$ at 50 kHz for the typical measurement settings.

The measured probe resonant frequencies are presented in Figure 5.13 as a function of probe length. The measured resonant frequencies are 7 and 19% lower on average than the predicted frequencies for the 1 and 2 μm wide force probes, respectively. Air damping is accounted for in calculating the predicted resonant frequencies using Ref. [23]. The slight reduction in resonant frequency is probably due to added mass from residual oxide and

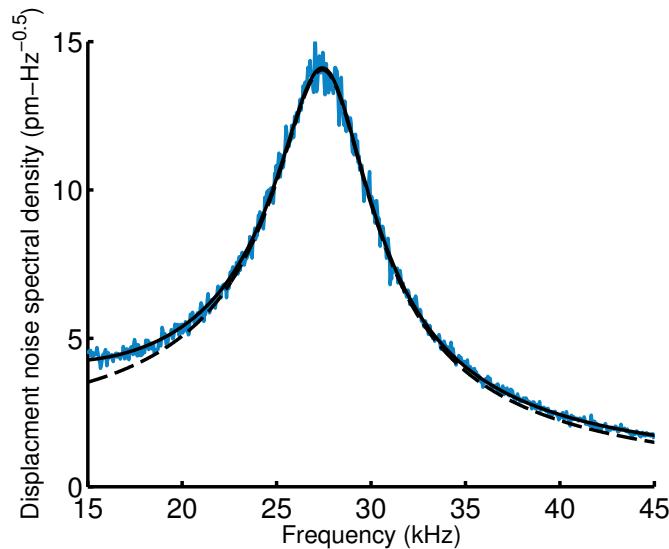


Figure 5.12: Example of force probe thermomechanical noise spectrum. The displacement noise spectral density of the probe is measured via laser doppler vibrometry (LDV) and the noise floor is approximately $1 \text{ pm}/\sqrt{\text{Hz}}$ at 50 kHz. The resonant frequency, quality factor and spring constant are extracted by numerically fitting a model to the data (solid line) consisting of a thermally agitated simple harmonic oscillator (dashed line) and noise terms.

stringers on the devices (Figure 4.34).

The measured and predicted probe spring constants are presented in Figure 5.14 as a function of probe length. The spring constants of the 2 μm wide force probes match the predictions well, deviating only +9% on average. However, the 1 μm wide probes are noticeably stiffer than we predicted (+52%). As in the case of their resonant frequencies, the spring constant deviations can be explained by the presence of residual oxide and stringers on the devices. The results highlight the importance of calibrating small force probes with thermomechanical noise rather than methods that depend on assumptions about the length, width or thickness of the devices.

We make one simplifying assumption in calculating the predicted resonant frequencies and spring constants. Although we assume that the cantilever beam is uniform in thickness along its entire length, in reality the piezoresistor is about 270 nm thick while the undoped portion of the beam is 300 nm thick (Table 4.2). In order to simplify the mechanics calculations we assume that the beam has a uniform thickness of 290 nm. This assumption

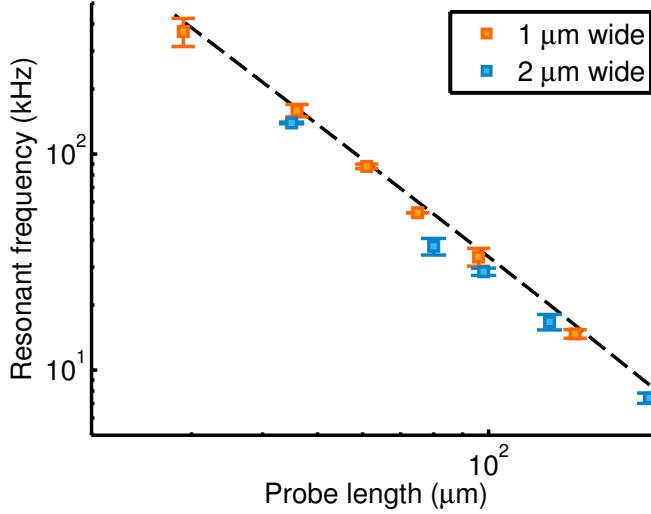


Figure 5.13: Predicted and measured resonant frequencies for 1 μm wide (orange) and 2 μm wide (blue) force probes in air. Air damping is accounted for in calculating the predicted resonant frequencies. Each data point is the aggregate of measurements from multiple devices. The resonant frequencies are 7 and 19% lower on average than predicted for the narrow and wide probes, respectively, probably due to added mass from residual material and stringers on the probes.

is based upon a three-dimensional finite element analysis (FEA) model that we used to calculate the impact of piezoresistor thinning on resonant frequency. The average deviations between the single-thickness analytical model and multiple-thickness FEA model in calculating the undamped resonant frequency and spring constant are 1.4% and 0.3%, respectively.

We considered two possible sources of error in measuring spring constants via thermomechanical noise using laser doppler vibrometry: laser heating and laser position.

We can estimate the temperature increase from some simple optical analysis. The LDV system used for thermomechanical noise calibration (Polytec OFV-534) uses a 1 mW laser operating at 633 nm. But the force probes are significantly narrower than the laser spot size (roughly 5 μm in diameter) so only a fraction of the emitted laser power actually impinges on the device. The fraction of the incident laser power that actually hits the device ranges from 12 to 25% for the 1 and 2 μm wide probes assuming that the laser spot is centered on their tips. Neglecting reflections at the air-silicon interfaces (which would

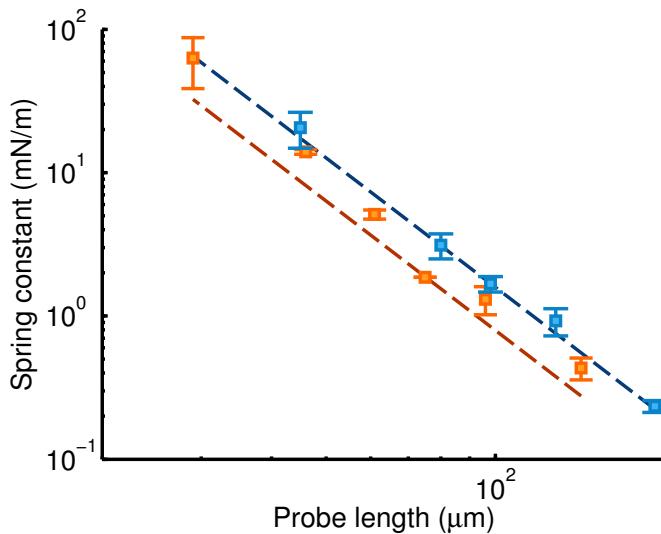


Figure 5.14: Predicted and measured spring constants for 1 μm wide (orange) and 2 μm wide (blue) unpassivated force probes. Each data point is the aggregate of measurements from multiple devices. The average spring constants of the 1 and 2 μm wide probes are 152% and 109% of their predicted values.

reduce the absorbed power), only about 8% of the incident light is absorbed by a 300 nm thick cantilever at 633 nm ($\alpha_{\text{Si}} = 2900/\text{cm}$).

We calculated the temperature increase during LDV measurements in air for the two longest force probes we developed (Figure 5.15). The maximum and average temperature increases are less than 16 and 5 K in both cases. The increased average temperature of the probe would lead to increased thermomechanical noise, resulting in a smaller spring constant estimate. However, the temperature changes from the LDV measurement are small with respect to the ambient temperature (293 K) and the inferred spring constants are only about 0.8% larger if we account for laser heating, small enough to neglect in our thermomechanical noise calculations.

The second possible issue with LDV-based spring constant measurements is the position of the laser beam on the force probe. As noted a moment ago, the diameter of the laser beam (5 μm) is fairly large compared with the probes. This leads to the deflections we measure being averaged over the beam area and the possibility that the beam is set substantially back from the tip. However, beam measurement location was not an issue for several reasons. First, the deflection of the first resonant mode of a

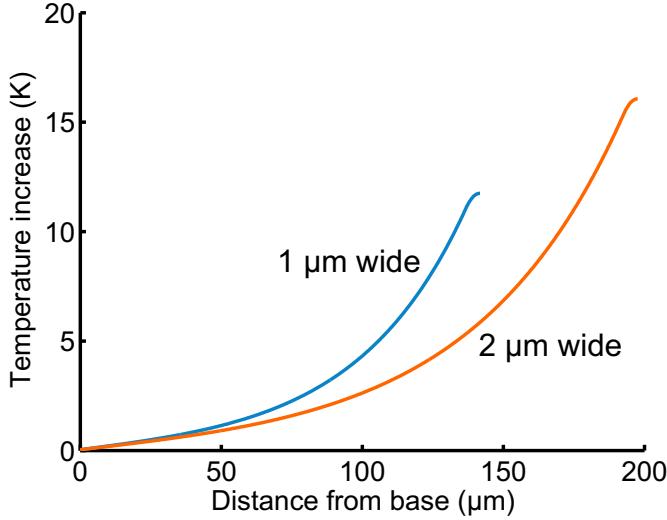


Figure 5.15: Laser-induced heating during thermomechanical noise measurements. The temperature profiles are calculated via finite differences for the two longest cantilevers that we fabricated (300 nm x 1 μm x 142 μm and 300 nm x 2 μm x 198 μm) assuming operation in air. Thermomechanical noise is measured via laser doppler vibrometry using a 1 mW, 633 nm laser which is focused to a 5 μm diameter spot using a 100x microscope objective. Only a fraction of the emitted laser power is absorbed due to low absorptivity of silicon at 633 nm and the large beam spot size; about 10 and 20 μW for the 1 and 2 μm wide devices, respectively. The maximum and average temperature increase for the 2 μm wide device are 16 and 4 K, resulting in error of less than 1% in calculating the cantilever spring constant from its thermomechanical noise.

cantilever beam scales as x^4 where x is the distance from the base of the beam, so the averaging effect shifts the effective measurement location from the center of the laser spot out towards the tip of the beam. Second, we carefully centered the laser beam on the tips of the probes for all measurements. Finally, all of the probes, particularly the devices intended for bundle mechanics measurements where accurate spring constant calibrations are particularly important, are fairly long with respect to the spot size.

We measured the quality factors of our probes in air at atmospheric pressure as part of the TMN spring constant calibration process. Experimentally measured quality factors in air are plotted as a function of measured probe resonant frequency in Figure 5.16. The 1 μm and 2 μm wide force probe designs are plotted separately and agree extremely well with the model predictions calculated from Brumley et al. [23]. Variation in quality factor

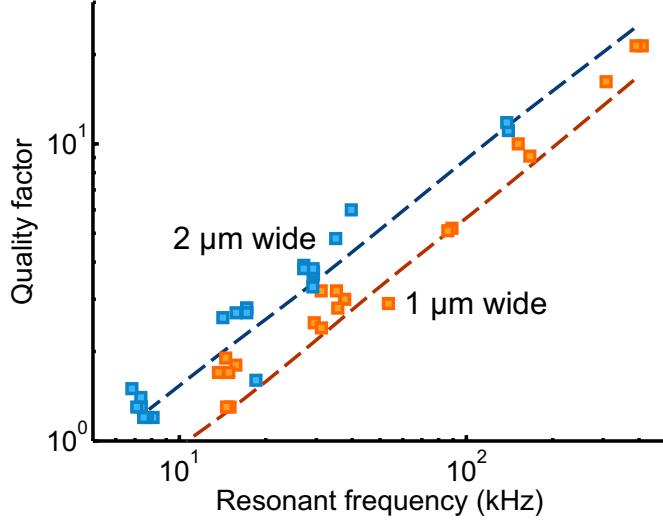


Figure 5.16: The quality factor of the force probes in air varies with their resonant frequency. Data for 1 and 2 μm wide force probes are plotted separately. The model predictions (dashed lines) are calculated from Brumley et al. [23] and show excellent agreement. The quality factor is greater for wider probes and increases roughly quadratically with resonant frequency.

with probe size follows directly from physical scaling laws because the quality factor can be written as

$$Q = \frac{\sqrt{m_{\text{eff}}k_c}}{b}.$$

Variation in m_{eff} and b with probe size and ambient pressure is complex due to the effect of the surrounding fluid and is discussed in more detail in Refs. [409] and [23]. For our probes Q scaled as $w_c^{1/2}l_c^{-3/2}$ in air.

We characterized the impact of parylene coating on probe mechanics using thermomechanical noise. Example thermomechanical noise spectra are presented in Figure 5.17 before and after the same probe is coated with 220 nm of parylene N. After coating, the resonant frequency drops from 29.2 to 27.7 kHz, the quality factor increases from 3.3 to 5.6 and spring constant increases from 1.40 to 2.57 mN/m.

We measured the change in resonant frequency and spring constant with parylene passivation for a range of parylene thicknesses (Figure 5.18). The best-fit lines are calculated assuming a parylene N elastic modulus of 5 GPa rather than the nominal

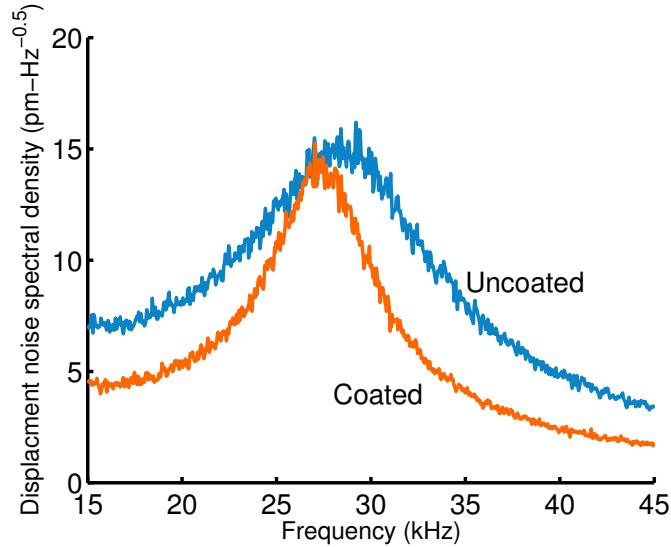


Figure 5.17: Example thermomechanical noise spectra before and after parylene coating. The same device (design 10) was measured before and after a 220 nm thick parylene N coating. After coating, the resonant frequency drops from 29.2 to 27.7 kHz, the quality factor increases from 3.3 to 5.6 and spring constant increases from 1.40 to 2.57 mN/m.

modulus of 2.4 GPa from (5.1) and (5.2). The increased effective modulus, regardless of its source, leads to a much greater increase in spring constant for a given coating thickness. For example, comparing Figures 5.6 and 5.18 one sees that a 300 nm passivation coating would increase the spring constant by 50 and 170% for 2.4 and 5 GPa films, respectively.

The increase in spring constant with parylene passivation is far more important than the slight decrease in resonant frequency. Why is the parylene less compliant than expected? Zhang et al. studied the microstructural properties of parylene N deposited over a range of process pressures [390]. They observed the presence of a 50 nm thick surface layer with 3-7x larger hardness than the bulk of the parylene film. The hardened surface layer is only apparent in our data because the parylene is comparable in thickness to the silicon beam.

We measured the frequency response of several force probes in air and water in order to validate the model from Brumley et al. [23] that we use to account for fluid damping effects. The frequency response of an example device is presented in Figure 5.19. We measured the frequency response from the thermomechanical noise spectrum of the device via LDV. The resonant frequency and quality factor of the device were calculated by numerically fitting

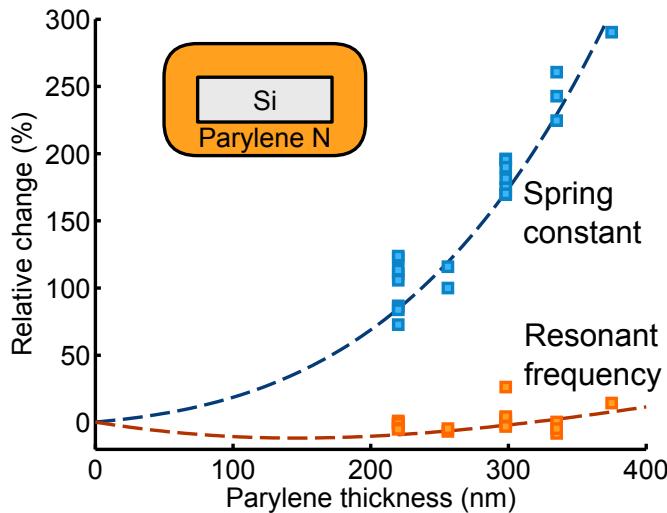


Figure 5.18: Experimental variation in spring constant and resonant frequency with varying parylene coating thicknesses. Thermomechanical noise spectra were recorded before and after parylene coating and the relative changes were calculated for matched pairs. The parylene thickness was measured via spectroscopic reflectometry with profilometer and SEM measurements for validation. The relative change in spring constant with coating thickness indicates an effective parylene N elastic modulus of 5 GPa rather than 2.4 GPa. Resonant frequency changes very little with coating thickness and the spring constant increase has far greater implications for device performance.

5.7) to the data in order to minimize the relative residual error. In air the resonant frequency and quality factor of the device are 290 kHz and 12.6 compared with model predictions of 281 kHz and 12.8. In water the resonant frequency and quality factor decrease to 85.2 kHz and 0.7 compared with model predictions of 84.3 kHz and 0.6.

The measured and predicted resonant frequencies and quality factors in air and water from two additional force probes are presented in Table 5.1. The second example device in the table corresponds to the device presented in Figure 5.19. All of the measured devices were unpassivated in order to verify the accuracy of the model. In general the model predictions agree with the measured results to within about 5% using the nominal densities and viscosities of air and water at room temperature (293 K).

In summary, we measured the spring constants and resonant frequencies of the uncoated force probes via their thermomechanical noise. Their resonant frequencies were slightly lower than predicted (81-94% of the nominal values) while their spring constants were

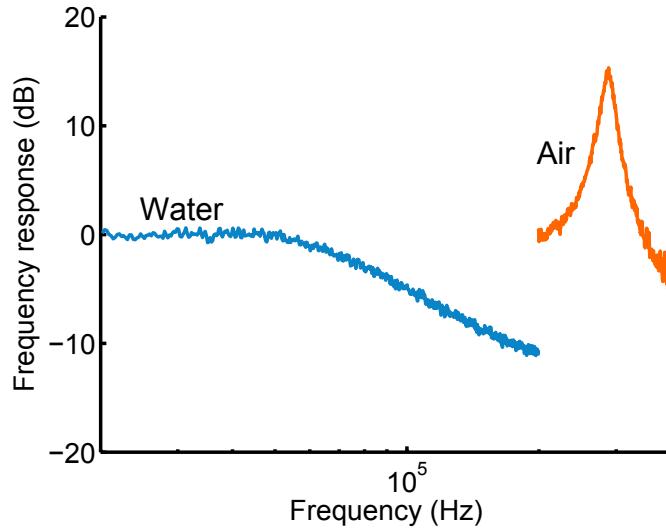


Figure 5.19: Frequency responses in air and water of a 35 μm long, 1 μm wide and 290 nm thick force probe. In air the resonant frequency and quality factor of the device are 290 kHz and 12.6 compared with model predictions of 281 kHz and 12.8. In water the resonant frequency and quality factor decrease to 85.2 kHz and 0.7 compared with model predictions of 84.3 kHz and 0.6.

slightly high (109-152% of the nominal values). We used roughly 250 nm thick parylene N films to passivate the force probes for water operation. The parylene passivation roughly doubles the spring constants of the probes while decreasing their resonant frequencies by 5-10%. The parylene coating does not affect the displacement sensitivity of the probes, but the increased spring constant decreases their force sensitivity linearly. Finally, we measured the frequency response of several devices in air and water and confirmed that the model that we use to predict force probe frequency response in water is accurate to within about 5%.

5.5 Piezoresistive sensor

In this section we will discuss the performance of the piezoresistive sensor. We will start by discussing its electrical characteristics before moving onto noise, displacement sensitivity, self-heating and resolution.

Table 5.1: Frequency response of several probes in air and water. The measured and predicted (in parentheses) resonant frequency and quality factor are reported. The predicted frequency response is calculated from Brumley et al. [23] using the nominal densities and viscosities of air and water at room temperature. The force probe frequency responses were measured from their thermomechanical noise spectra via LDV. For each spectrum, the resonant frequency and quality factor were determined by numerically fitting a SHO model to the measured data to minimize the relative residual error.

Device number	1	2	3
Design number	4	6	12
l_c (μm)	61	35	45
w_c (μm)	1	1	2
t_c (nm)	290	290	290
Air			
f_0 (kHz)	80.5 (91.1)	290 (281)	162 (170)
Q (-)	5.0 (5.3)	12.6 (12.8)	12.5 (13.4)
Water			
f_d (kHz)	15.0 (19.1)	85.2 (84.3)	44.8 (44.6)
Q (-)	0.4 (0.4)	0.7 (0.6)	0.8 (0.7)

5.5.1 Electrical characteristics and resistance

One of the most important characteristics of piezoresistors is the nature of the metal-silicon contacts. Figure 5.20 presents the electrical current and resistance as a function of applied bias for a typical device in air. The data reveals that the contact is Ohmic and that the piezoresistor resistance scales as V^2 as expected due to Joule heating. The overall electrical performance of the piezoresistor is ideal in marked contrast with the performance of the initial devices that we fabricated (Figure 4.24).

All of the devices that will be discussed in this chapter were fabricated with the optimized process flow developed in Chapter 4. Briefly, the contacts and piezoresistors are doped at 900°C for 20 minutes and 825°C for 30 minutes, respectively. During metallization the wafers are heated in the load lock to drive off moisture before Ti and Al films are successively sputter deposited. The force probes are released from the BOX using vapor HF and no FGA is performed at the end of the process.

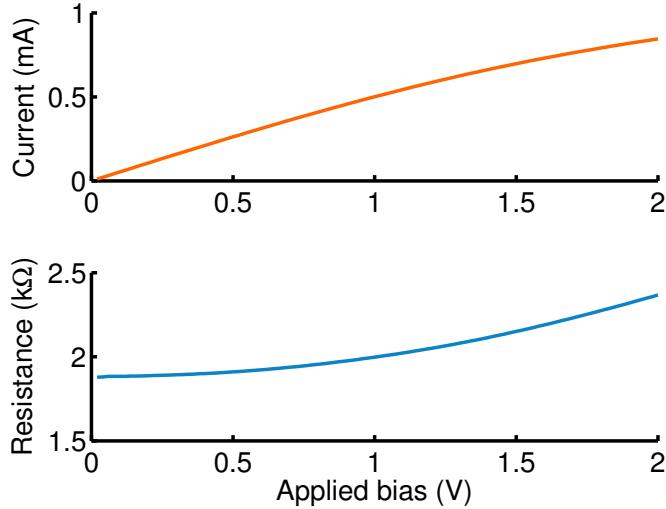


Figure 5.20: Current-voltage data from a typical released piezoresistor. The contact is ohmic and piezoresistor resistance scales with V^2 due to Joule heating as expected.

Figure 5.21 plots the electrical resistances of finished devices as a function of the number of squares in the piezoresistor. All of the piezoresistors shown here have 500 nm wide legs and their lengths are varied. The transfer length method (TLM) data indicates a sheet resistance of $110 \Omega/\square$ and excess resistance of 340Ω (170Ω per contact). Accounting for the approximately 10Ω added by the Al interconnects and cables the contact resistance is only slightly greater than our earlier prediction of 144Ω based upon the nominal sheet resistances and contact resistivity (Section 4.5.3).

Each force probe device has two released resistors and two unreleased on-chip resistors (that are not used). The released and unreleased resistors are plotted as blue and orange squares, respectively, in Figure 5.21. The excellent matching between the released and unreleased resistors verifies that the piezoresistors are not damaged by the optimized vapor HF release process.

The resistance factor is also plotted in Figure 5.21 as a function of piezoresistor length. The excess resistance from the contacts and interconnects decreases the resistance factor, γ , of the devices. The net excess resistance is equivalent to about 3.1 resistor squares, resulting in resistance factors (γ) ranging from 0.83 to 0.96 depending on the piezoresistor dimensions. Recall that excess resistance simultaneously increases the Johnson noise and

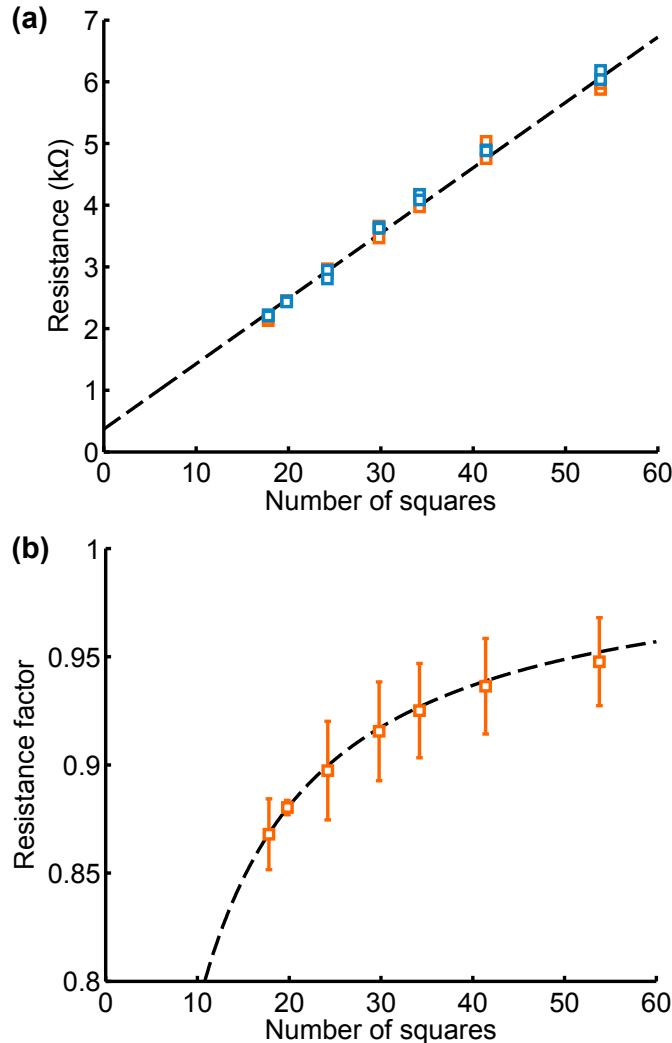


Figure 5.21: (a) Electrical resistance and (b) resistance factor as a function of piezoresistor size. The typical sheet resistance for finished wafers was $110 \Omega/\square$ and each metal-silicon contact contributed about 170Ω of excess resistance. The resistances of the released (blue) and unreleased (orange) silicon piezoresistors match well. (b) Excess interconnect and contact resistance leads to resistance factor variation between 0.83 and 0.96.

decreases the sensitivity of the piezoresistor, and increases the minimum detectable force resolution as roughly $\gamma^{-3/2}$, resulting in resolution degradation of 6 to 30% for γ ranging from 0.96 to 0.83.

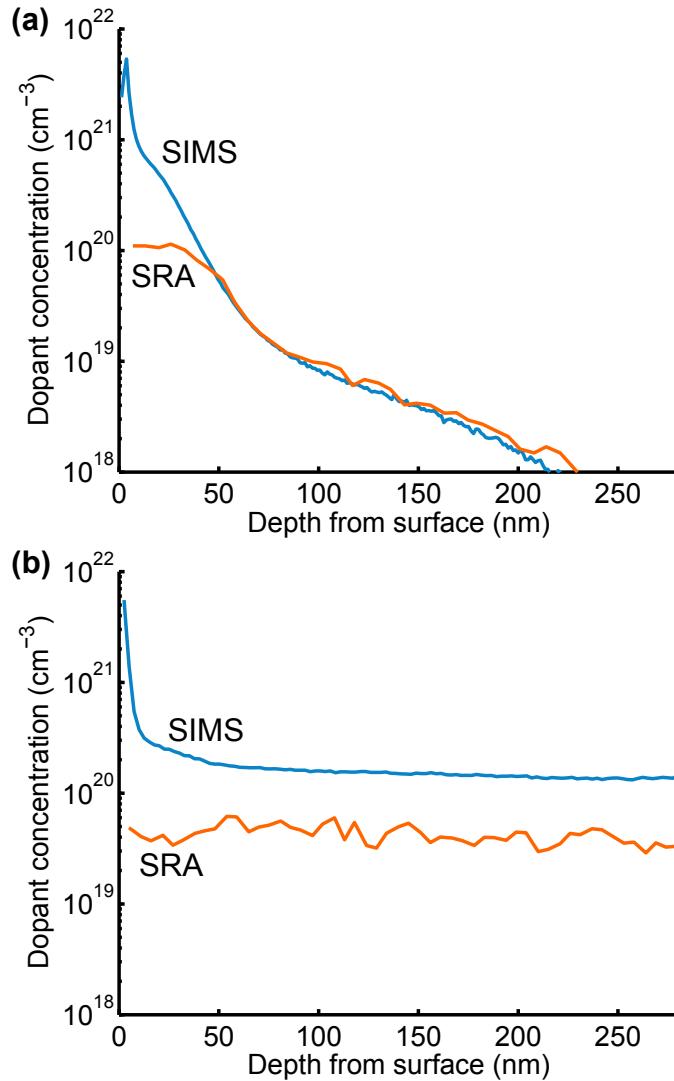


Figure 5.22: Comparison between SIMS and SRA data for (a) SOI and (b) POI samples. Both samples were POCl_3 doped at 825°C for 30 minutes. While most of the dopants are confined to within the first 50 nm of the SOI sample, the POI sample is uniformly doped throughout its thickness due to the rapid diffusion of dopants along grain boundaries.

5.5.2 Dopant concentration profile

The concentration profile of dopants through the thickness of the device layer plays a major role in determining piezoresistor noise and sensitivity. The electrically active and total dopant concentration profiles are presented in Figure 5.22 for both SOI and POI

wafers. The electrically active dopant concentration profile was measured via SRA by Solecon Laboratories while I measured the total dopant concentration profile at Stanford using a Cameca NanoSIMS 50L. The relative sensitivity factor for the SIMS data was calibrated from an ion implanted sample with known dopant concentration profile (courtesy of Pradeep Nataraj).

The SOI data indicate that there is excess interstitial phosphorus within 50 nm of the sample surface. The excess phosphorus, a well known side effect of phosphorus predeposition [217, 218, 227], introduces lattice dislocations and other defects that likely increase the overall Hooge factor of the silicon resistors. Three SOI samples were characterized using SRA in order to calculate a sensitivity factor (β^*) of $0.233 \pm 5.2\%$ and a carrier density of $4.13 \times 10^6 \pm 22.9\%$ per square micron. Note that the carrier density calculation accounts for the variation in dopant concentration and mobility (Section 2.4.1).

The POI data indicate that the dopants diffused all the way through the sample due to the high diffusivity along grain boundaries. While the SRA data indicate an electrically active dopant concentration of $5 \times 10^{19} \text{ cm}^{-3}$, the actual electrically active concentration lies somewhere between the SRA and SIMS data because of the nature of SRA measurements, which measures a resistivity profile and converts it to a concentration profile based upon a single crystal silicon reference sample. Neglecting the dopants that remain at the grain boundaries, the POI data suggest a 2.2-fold mobility reduction due to grain boundary scattering.

Three POI samples were characterized using SRA in order to calculate a sensitivity factor of $-0.04 \pm 105.1\%$ (effectively zero) and a carrier density of $2.74 \times 10^7 \pm 15\%$ per square micron. There is substantial uncertainty in the carrier density calculation. Most importantly, the POI samples are uniformly doped throughout their thickness, yielding extremely low sensitivity piezoresistors. While the POI wafers and force probes were extremely helpful in developing the fabrication process and experimental setup, the rest of the thesis will focus on the SOI devices.

The sensitivity factor (β^*) and carrier density (N_z) values determined via SRA will be used in the noise and sensitivity sections in order to interpret device performance.

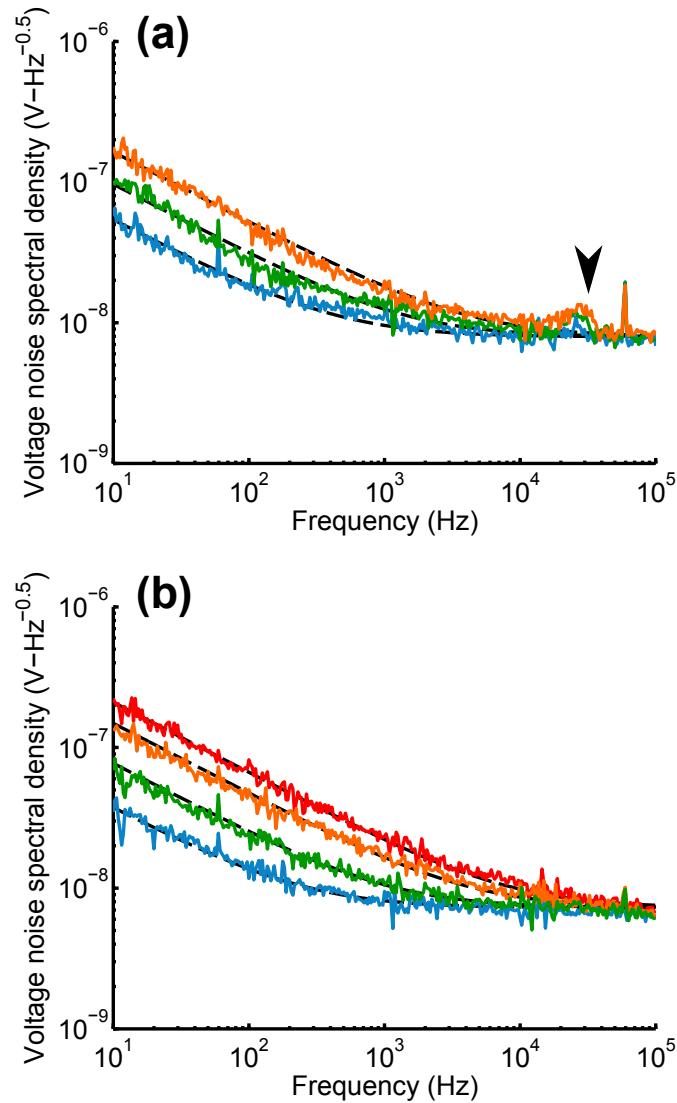


Figure 5.23: Example piezoresistor noise spectra in air. (a) Noise spectra for an SOI force probe biased to 1, 1.5 and 2 V. The probe exhibits ideal 1/f noise behavior and is capable of self-detecting its thermomechanical resonance in air (indicated). The model fits (black dashed lines) are used to calculate α . (b) Noise spectra for a POI-based PRPE device biased to 0.25, 0.5, 1 and 1.5 V. The probe exhibits ideal 1/f noise behavior but is incapable of self-detection due to the uniform doping through the polysilicon piezoresistor.

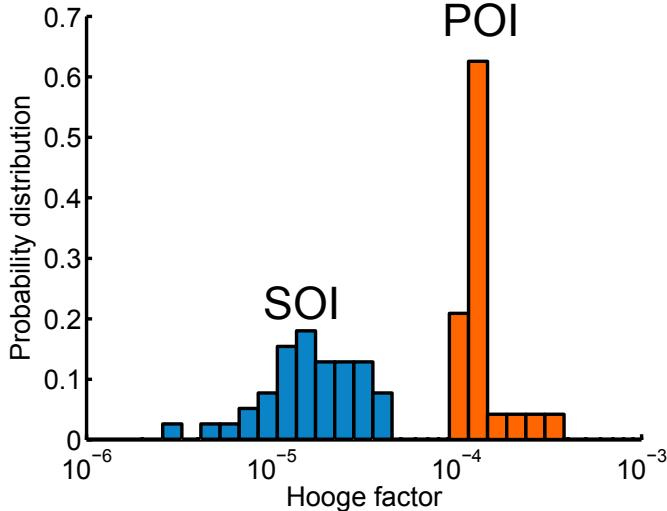
5.5.3 Noise

Example noise spectra from SOI- and POI-based force probes are presented in Figure 5.23. The Wheatstone bridge is biased to 1, 1.5 and 2 V for the SOI measurements and the power dissipation in the $2.8\text{ k}\Omega$ piezoresistor ranges from 90 to $360\text{ }\mu\text{W}$. When operated in air at these bias levels the average piezoresistor temperature increase ranges from 7 to 30 K. The device is capable of self-detecting its thermomechanical resonance ($f_0 = 30\text{ kHz}$ and $Q = 3.6$) in air.

As described in Section 3.5, the measurement circuit consists of a low-noise instrumentation amplifier (INA103, 100x gain, 800 kHz -3dB bandwidth) followed by an inverting op-amp (THS4061) variable gain amplifier (VGA) with gain settings of 1, 10 or 100x. Noise measurements were performed with an overall gain of 1000x. All measurements are referred to the instrumentation amplifier inputs. The frequency response of the signal conditioning path was measured directly by driving the instrumentation amplifier inputs with a 1-2 mV chirp from 1 kHz to 1 MHz.

The Hooge factor for the SOI force probe in Figure 5.23 is $1.1 \pm 0.4 \times 10^{-5}$ and is calculated by fitting a noise model to the experimental data to minimize the squared sum of the residual error. The noise model includes Johnson, Hooge and amplifier noise and minimizes the squared sum of the residual error (via `fmincon`) by adjusting α . The nominal resistance and number of carriers in the piezoresistor are provided to the fitting function. The number of carriers in the piezoresistor is calculated by integrating the carrier density (Section 5.5.2) over the piezoresistor loop and excludes the Al interconnects and highly doped contacts. Note that the noise peak at 60 Hz is barely visible because the Wheatstone bridge is biased with an IC voltage reference and the measurement is performed in a shielded box.

The POI device is operated in a Wheatstone bridge biased to 0.25, 0.5, 1 and 1.5 V. The power dissipation (9 to $230\text{ }\mu\text{W}$) and average piezoresistor temperature increase (1 to 35 K) ranges in the $2.8\text{ k}\Omega$ piezoresistor are comparable to the SOI device. The probe exhibits ideal $1/f$ noise behavior but is incapable of self-detection due to the uniform doping through the polysilicon piezoresistor ($\beta^* \approx 0$). Assuming complete dopant activation the Hooge factor is on the order 10^{-4} , although there is substantial uncertainty in this value.



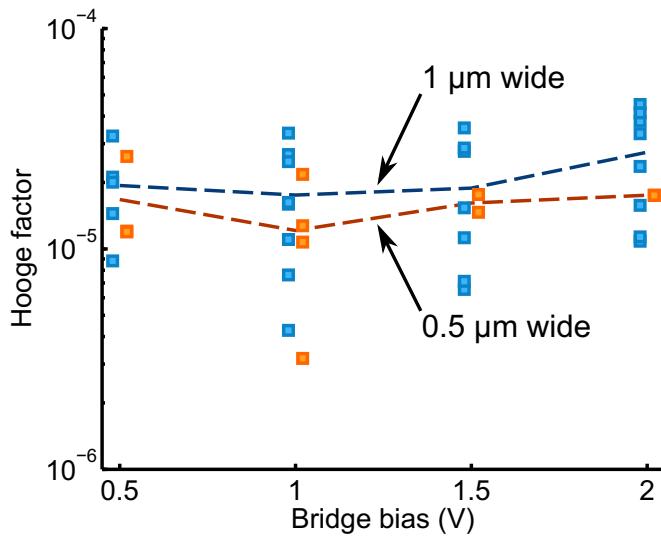


Figure 5.25: The Hooge factor is independent of Wheatstone bridge bias and piezoresistor width. The average Hooge factor for each bridge bias (dashed lines) is plotted for 0.5 and 1 μm wide piezoresistors. The results suggest that the Hooge factor is limited by diffusion-rather than etch-related defects.

5.5.4 Sensitivity

Force probe displacement sensitivity was measured by mounting the devices on the stage of an AFM (Witec Alpha300) and deflecting the tip of the force probe using a stiff (40 N/m) tipless AFM cantilever. The force probe is positioned between the objective lens and stiff AFM cantilever so that the contact point can be readily visualized to ensure that the contact point is at the tip of the force probe (Figure 5.26).

The AFM stage is piezoelectrically actuated and capacitively sensed with a bandwidth of approximately 300 Hz and typical RMS noise of 1 nm. The stage actuator, capacitive sensor and force probe piezoresistor output are controlled and recorded via Matlab using a 16-bit, 400 kHz DAQ (NI USB BNC-6212). The AFM stage is driven at 10 Hz and is controlled by one of the analog DAQ outputs. The contact between the force probe and AFM cantilever is prone to drifting over long measurements, and although we initially used a ramped deflection signal we found that sinusoidal deflection resulted in much less variability in order to restrict the measurement bandwidth. The AFM cantilever is about 4 orders of magnitude stiffer than the force probe, so we assume that the stage deflection

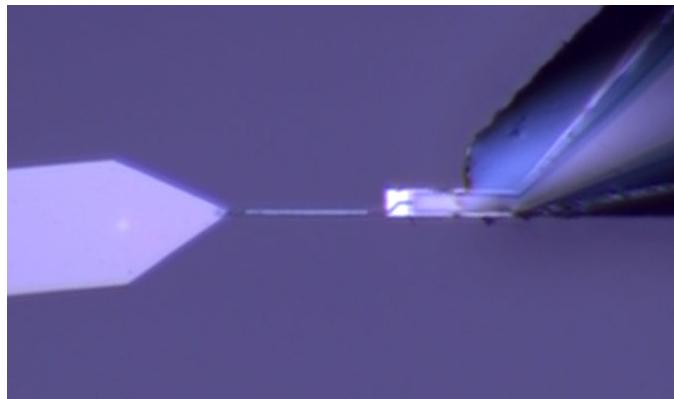


Figure 5.26: Optical micrograph of the displacement sensitivity calibration setup. The force probe (right) is mounted to a high-precision three-axis positioning stage. A stiff AFM cantilever (left) is brought into contact with the tip of the force probe. The positioning stage is actuated with a 10 Hz AC signal and both the stage position and output from the piezoresistor readout circuit are recorded.

corresponds directly to the force probe tip deflection.

We also attempted to simultaneous calibrate the displacement sensitivity and spring constant of the force probe by using a soft AFM cantilever. However, the softest commercially available tipless AFM cantilevers (30 mN/m) are still an order of magnitude stiffer than most of the force probes we fabricated, resulting in substantial spring constant uncertainty. Additionally, the contact point between the soft cantilever and force probe tended to slip during measurements. We resorted to calibrating the displacement sensitivity using a stiff AFM cantilever and then measuring the force probe spring constant via thermomechanical noise.

Typical calibration data are presented in Figure 5.27. The force probe is driven with a 1 μm peak-peak deflection at 10 Hz for 5 seconds. Both the stage deflection and sensor output signals are recorded and their spectral densities are calculated using Hanning windows. The displacement sensitivity is calculated from the ratio of the stage position and piezoresistive sensor outputs at the drive frequency. For this particular device, the displacement sensitivity was calculated as $227.1 \pm 5.6 \text{ V/m-V}$ ($N = 7$). The typical standard deviation for the displacement sensitivity protocol we developed was about 2-3%. Finally, note that all of the displacement sensitivities that we report are normalized to the Wheatstone bridge bias voltage and are referred to the instrumentation amplifier inputs (i.e. independent of

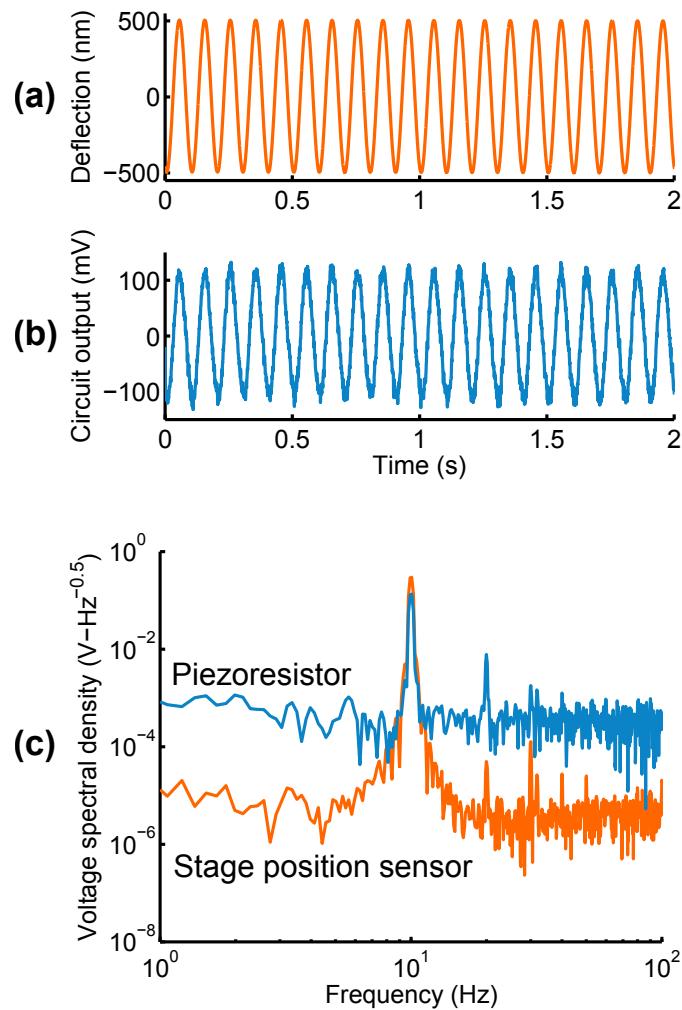


Figure 5.27: Example of sensitivity calibration. (a) The force probe is driven with a 500 nm, 10 Hz deflection using an extremely stiff (about 40 N/m) AFM cantilever. The deflection of the stage is measured using its built-in capacitive sensor. (b) The piezoresistor output in response to the known deflection is measured. (c) The deflection sensitivity is calculated from the slotage and piezoresistor spectra. For this example the resulting deflection sensitivity (S_{XV}) is 224.5 V/m-V.

amplifier gain).

The measured and predicted displacement sensitivities for the probe designs that we tested are presented in Figure 5.28 as a function of probe length. The predicted displacement sensitivities of the 1 and 2 μm wide probes diverge slightly as probe length decreases due to two mechanisms. First, current flow is transverse to the applied stress at

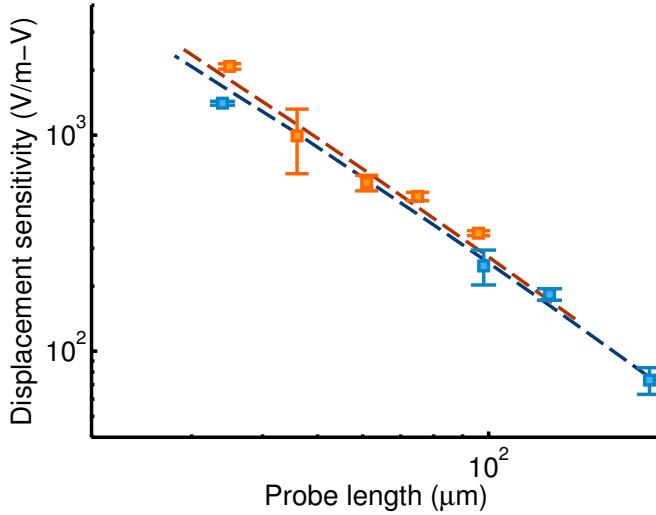


Figure 5.28: Predicted and measured displacement sensitivities as functions of force probe length. The predicted performance (dashed lines) is different for 1 (orange) and 2 (blue) μm wide probes due to differences in γ . The predicted sensitivity decreases slightly as piezoresistor length decreases due to the combined effects of γ and the transverse stress at the end of the piezoresistive loop. The agreement between the predicted and measured performance is excellent.

the end of the piezoresistor loop. The transverse portion slightly decreases the sensitivity of the piezoresistor because $\pi_t \approx \pi_l/2$ for n-type silicon. The ratio of the transverse resistance to the overall piezoresistor resistance is smaller for the 1 μm wide probes than the 2 μm wide probes, leading to a smaller reduction in sensitivity. The second mechanism for sensitivity reduction is the resistance factor (γ); as the force probe and piezoresistive loop become shorter, the ratio of the contact and interconnect resistance to the overall piezoresistor resistance increases. The narrower probes have slightly higher resistances than their wider counterparts, so their resistance factors are larger.

5.5.5 Self-heating

Piezoresistor temperature during operation is inferred from the piezoresistor temperature coefficient of resistance (TCR). The TCR is calibrated using a temperature controlled oven (Thermotron S-1.2) and small bias power ($< 1 \mu\text{W}$). The temperature is ramped from 20 to 60°C in 5°C increments and stabilized to within 0.1°C at each measurement point. We

calibrated four devices using this protocol and calculated a TCR of 1590 ± 150 ppm/K ($\mu \pm \sigma$). In contrast, the TCR of the first generation of devices that we fabricated was 1167 ± 53 ppm/K [302].

In order to determine the temperature of the force probe during operation we measure an I-V curve and calculate the average piezoresistor temperature from

$$T_{\text{PR}} = T_0 + \frac{1}{\alpha} \frac{R - R_0}{R_0}. \quad (5.8)$$

where the initial resistance is calculated from extrapolating the resistance to zero bias (the minimum bias actually applied is on the order of 50 mV). The uncertainty of this calibration approach can be calculated from

$$\frac{\sigma_{T_{\text{PR}}}}{\mu_{T_{\text{PR}}}} = \sqrt{\left(\frac{\sigma_R}{\mu_R}\right)^2 + \left(\frac{\sigma_{R_0}}{\mu_{R_0}}\right)^2 + \left(\frac{\sigma_\alpha}{\mu_\alpha}\right)^2}. \quad (5.9)$$

where σ_X/μ_X is the relative uncertainty in each relevant variable. The relative uncertainties of R and R_0 are $\pm 0.1\%$ (about 5Ω for a $5 \text{ k}\Omega$ resistor) and the uncertainty in our α calibration is $\pm 10\%$. The overall uncertainty in calculating the piezoresistor temperature from its resistance is dominated by the TCR uncertainty and limited to $\pm 10\%$.

Figure 5.29 presents the calculated temperatures in air and water for a parylene passivated probe (design 12, 45 μm long and 2 μm wide with a 5.9 μm long piezoresistor). The piezoresistor resistance as a function of the applied bias voltage is calculated from the I-V response of the piezoresistor (Figure 5.10). The relative resistance change is converted to a change in the average piezoresistor temperature from the TCR calibration data. The thermal model used in the optimization code matches the experimental results to within approximately 20% (Figure 5.29a).

The predicted temperature profile along the length of the force probe is plotted in Figure 5.29b for operation in air and water at a bridge bias of 1 V (the typical operating conditions during experiments). The maximum and tip temperatures during operation in water increase by 5 and 1 K, respectively. Our model assumes h_{eff} values in air ($2 \text{ kW/m}^2\text{-K}$) and water ($50 \text{ kW/m}^2\text{-K}$) that we measured using the first generation of devices that were fabricated (Section F) [302].

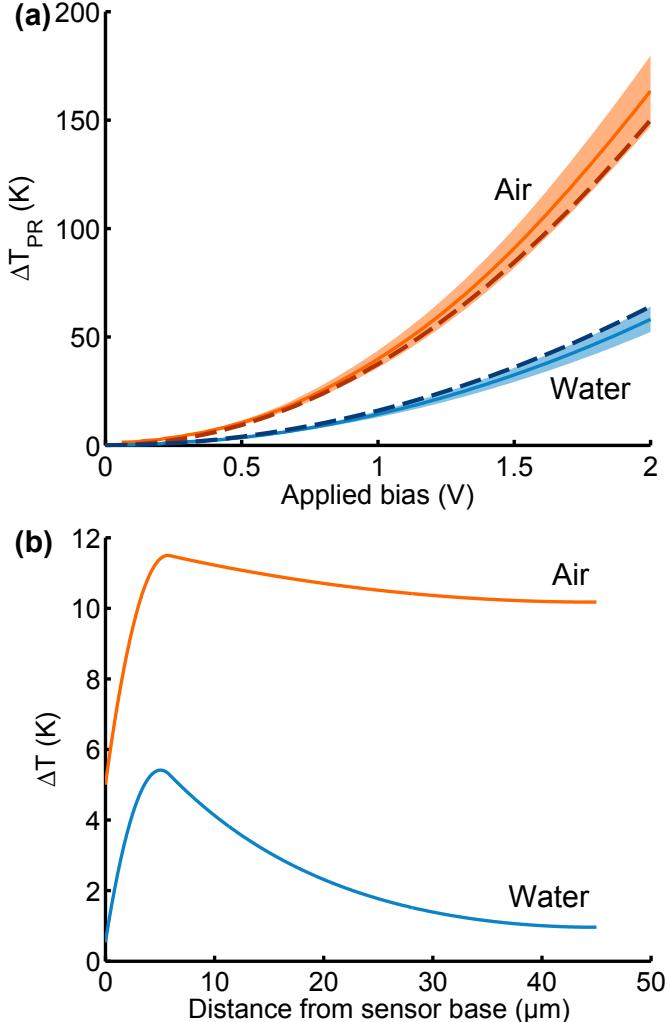


Figure 5.29: Piezoresistor self-heating during operation. (a) The change in resistance with applied bias is converted to a temperature increase (solid lines) from TCR calibration data. The average piezoresistor temperature predicted by the finite difference-based thermal model (dashed lines) diverges by about 20% from the experimental data. (b) The predicted temperature profiles in air and water for a bridge bias voltage of 1 V are plotted.

The thermal resistance at the base of the device (R_{base}) is determined from a three-dimensional conduction-convection FEA model as 20 and 5 K/mW in air and water, respectively, for sensors attached to the end of a piezoelectric actuator. The FEA model includes the force probe and a $100 \times 100 \times 100 \mu\text{m}$ volume of the silicon die at the base of the probe. Heat dissipates away from the force probe and die convectively and

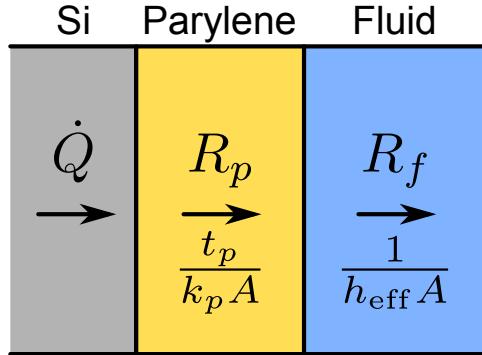


Figure 5.30: Thermal impact of parylene passivation. The parylene coating adds a series resistance between the silicon and the surrounding fluid, which can be modeled as a simple one-dimensional thermal system. A 250 nm thick parylene N coating decreases convective heat transfer by 1% and 10% in air and water, respectively.

R_{base} is calculated from the temperature and heat flux at the base of the force probe. The simulations and our earlier Raman thermometry results indicate that the widely utilized assumption that the temperature of the probe is constrained to the ambient temperature is not a particularly good one.

Until now we have focused on the thermal modeling of unpassivated force probes. However, parylene N has a fairly low thermal conductivity (0.125 W/m-K) and effectively insulates the force probe from the surrounding fluid. However, parylene has only a small effect on the self-heating of the force probes. Consider the one-dimensional thermal conduction shown in Figure 5.30. The heat generated by the silicon piezoresistor must traverse the parylene coating before escaping to the fluid. The thermal resistance of the parylene (R_p) and fluid interface (R_f) are

$$R_p = \frac{t_p}{k_p A} \quad (5.10)$$

$$R_f = \frac{1}{h_{\text{eff}} A} \quad (5.11)$$

where t_p is the parylene thickness, k_p is the parylene thermal conductivity, h_{eff} is the effective convection coefficient at the solid-fluid interface, and A is the surface area at each interface. The addition of the series resistance from the parylene reduces the temperature at

the parylene-fluid interface, resulting in a decrease in the effective value of h_{eff} according to

$$h'_{\text{eff}} = h_{\text{eff}} \left(1 - \frac{R_p}{R_f} \right) \quad (5.12)$$

$$= h_{\text{eff}} \left(1 - \frac{t_p h_{\text{eff}}}{k_p} \right). \quad (5.13)$$

The parylene coating reduces h_{eff} by 1% and 10% in air and water, respectively, assuming a 250 nm thick parylene coating. This simple analytical model does not account for details such as the increase in surface area with parylene coating, but a two-dimensional conduction-convection FEA model verifies that the simple analytical model is fairly accurate. For example, the FEA model predicts a 14% heat transfer reduction in water while the analytical model only predicts a 10% drop.

In summary, we monitor the force probe temperature during experiments from the change in temperature of the silicon piezoresistor. The thermal model of the force probe that we have developed enables the temperature along the entire length of the probe to be calculated from the average piezoresistor temperature. Although parylene passivation thermally insulates the force probe from the surrounding fluid, the effect on temperature is small.

5.5.6 Displacement and force resolution

In this section we will combine the mechanical, electrical and thermal performance of the devices in order to discuss piezoresistor measurement resolution. The measured force resolution of several devices at a bridge bias of 1 V is presented in Figure 5.31 as a function of their measurement bandwidth in air and water. Force noise increases roughly linearly with bandwidth due to the combined effects of increasing noise and decreasing force probe length.

The force resolution is calculated from

$$\text{MDF} = \frac{V_{\text{noise}}}{S_{\text{FV}}} \quad (5.14)$$

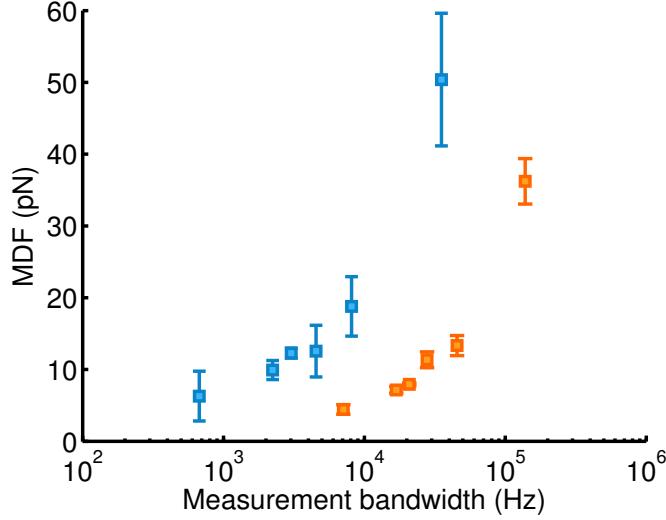


Figure 5.31: Measured force resolution for varying design resonant frequencies. The minimum detectable force increases roughly linearly with measurement bandwidth (i.e. probe resonant frequency) due to the combined effects of increasing noise and decreasing sensitivity. Force resolution and measurement bandwidth are presented for operation in air (orange) and water (blue) and account for the increased spring constant due to polyimide coating.

where V_{noise} is the integrated noise from 1 Hz to the resonant frequency of the probe and S_{FV} is the force sensitivity, calculated from

$$S_{\text{FV}} = \frac{S_{\text{XV}}}{k_c}. \quad (5.15)$$

The uncertainty in the force resolution is calculated by propagating the uncertainty in the spring constant and displacement sensitivity according to

$$\sigma_{\text{MDF}} = \mu_{\text{MDF}} \sqrt{\left(\frac{\sigma_{S_{\text{XV}}}}{\mu_{S_{\text{XV}}}} \right)^2 + \left(\frac{\sigma_{k_c}}{\mu_{k_c}} \right)^2}. \quad (5.16)$$

The uncertainty in the noise measurement is insignificant compared with the spring constant and displacement sensitivity uncertainty so is not included in the analysis for simplicity.

Figure 5.32 compares the force resolution of the presented devices with prior single crystal silicon piezoresistive cantilevers operating in air [35, 36, 154, 171, 195, 196, 202].

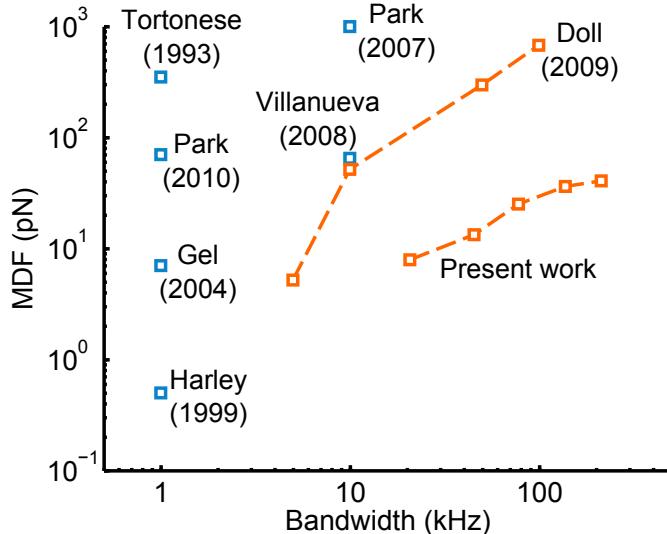


Figure 5.32: Comparison between the force resolution of the present devices with prior piezoresistive cantilevers. The force resolution of single crystal silicon piezoresistive cantilevers in air, usually from 1 Hz to some upper frequency limit, is tabulated. The cantilever dimensions, spring constant, resonant frequency, power dissipation and operating temperature vary significantly between devices.

The cantilever dimensions, spring constant, resonant frequency, power dissipation and operating temperature vary significantly between devices. For example, the cantilevers reported by Harley were 89 nm thick and operated with a maximum temperature increase of approximately 60 K in air while the cantilevers reported by Gel and Villanueva were both 340 nm thick. While piezoresistive cantilever operation in water has been reported previously, their integrated force noise was not stated [335, 410]. The performance of our devices has improved more than an order of magnitude between 2009 and 2012 due to both design optimization and fabrication improvements [196].

The force resolution of a piezoresistive sensor does not scale linearly with the Wheatstone bridge bias voltage (Figure 5.33). Recall that both the 1/f noise voltage noise spectral density (2.52) and force sensitivity (2.95) of the piezoresistor scale linearly with bridge bias while the Johnson voltage noise spectral density (2.50) is independent of bridge bias. Also recall that the force noise spectral density is equal to the voltage noise divided by the force sensitivity. Accordingly, as the bridge bias is increased the 1/f noise-limited force noise spectral density remains unchanged while the Johnson-limited

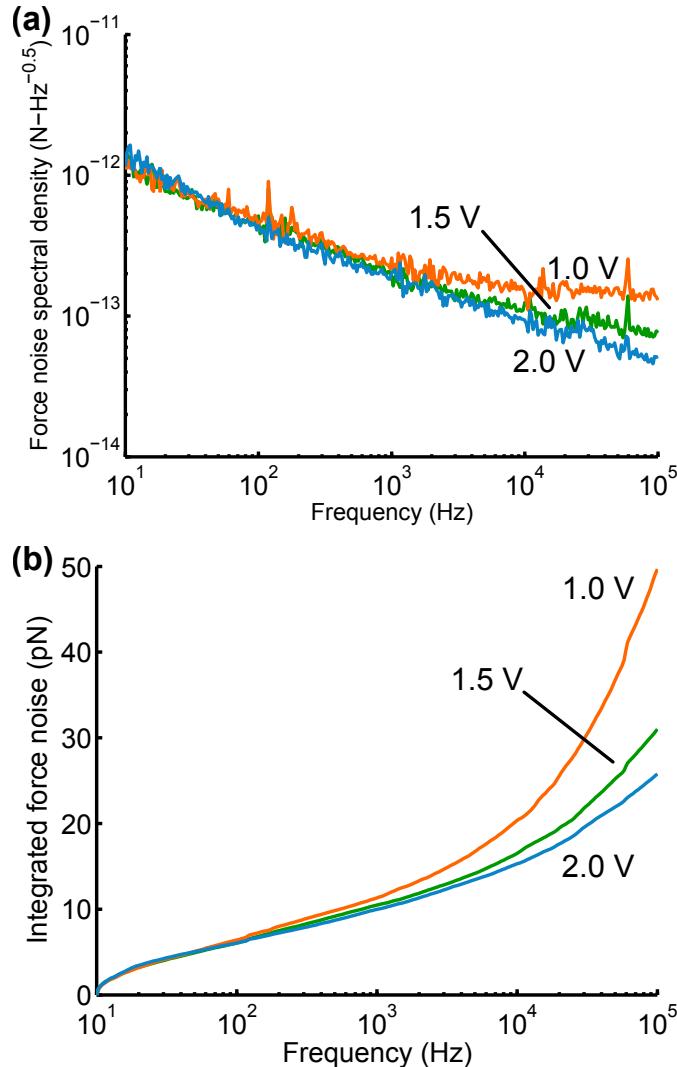


Figure 5.33: Example force noise spectra and integrated force noise. (a) The force noise above the 1/f corner frequency improves with increasing bridge bias voltage but low frequency force noise is unaffected. (b) The relative force noise contributions of 1/f and Johnson noise depend on the measurement bandwidth.

force noise spectral density decreases linearly. This leads to sub-linear improvement in integrated force/displacement noise with increasing bridge bias.

For example, in Figure 5.33b the integrated force noise below 1 kHz is relatively independent of the bridge bias. Resolution only improves with bridge bias if the devices are operated beyond the 1/f noise corner frequency (e.g. 100 kHz), which increases

Table 5.2: Summary of mechanical and sensor performance for selected devices operated at a Wheatstone bridge bias of 1 V. All of the probes are 300 nm thick and have a total width of 1 μm . Noise and resolution are calculated from 1 Hz to the resonant frequency of the sensor. Operation in water reduces the resonant frequency, quality factor and temperature increase from Joule heating while the required parylene passivation layer increases the spring constant.

Device number	1	2	3	4	5
l_c (μm)	96	75	61	46	35
l_{pr} (μm)	9.5	7.7	6.6	5.2	4.1
R ($\text{k}\Omega$)	5.16	3.97	3.94	3.15	2.65
$\alpha \times 10^{-5}$ (-)	4.3 ± 1.3	3.0 ± 0.3	3.7 ± 1.5	2.7 ± 1.6	2.2 ± 1.7
S_{XV} (V/m-V)	352 ± 9.4	509 ± 29	639 ± 18	1160 ± 13	2080 ± 60
Air operation (unpassivated)					
f_0 (kHz)	20.8	45.5	78.2	138	214
Q (-)	3.7	4.3	6.3	8.9	11.7
k_c (pN/nm)	1.12 ± 0.01	2.40 ± 0.21	4.72 ± 0.26	11.2 ± 0.97	20.8 ± 0.59
$\Delta T_{\text{max}}/\Delta T_{\text{tip}}$ (K)	7.2/1.5	7.6/2.5	7.9/3.6	8.4/5.2	8.9/6.6
V_{noise} (μV)	2.49	2.83	3.40	3.76	4.06
MDD (nm)	7.1 ± 0.2	5.6 ± 0.3	5.3 ± 0.1	3.2 ± 0.0	2.0 ± 0.1
MDF (pN)	7.9 ± 0.2	13.3 ± 1.4	25.1 ± 1.5	36.2 ± 3.2	40.6 ± 1.7
TMN floor (pN)	0.9	1.2	1.4	1.8	2.2
Water operation (220-250 nm of parylene N)					
f_d (kHz)	3.0	8.1	16.4	35.2	64.2
Q (-)	0.5	0.6	0.6	0.7	0.8
k_c (pN/nm)	2.32 ± 0.04	5.09 ± 0.37	9.91 ± 0.36	23.1 ± 0.76	43.4 ± 0.88
$\Delta T_{\text{max}}/\Delta T_{\text{tip}}$ (K)	3.0/0.0	3.4/0.0	3.6/0.0	3.9/0.0	4.1/0.2
V_{noise} (μV)	1.83	1.90	2.29	2.49	2.74
MDD (nm)	5.2 ± 0.1	3.7 ± 0.2	3.6 ± 0.1	2.1 ± 0.1	1.3 ± 0.1
MDF (pN)	12.3 ± 0.8	18.8 ± 3.5	35.6 ± 2.9	50.4 ± 3.5	57.4 ± 3.6
TMN floor (pN)	3.5	4.7	6.6	9.3	12

linearly with the bridge bias. While the sensor resolution only improves sub-linearly with increasing bridge bias, the piezoresistor power dissipation and resulting temperature increase scales quadratically with bridge bias (Section 2.6.3). In summary, increasing the Wheatstone bridge bias leads to diminishing returns in sensor resolution and a sharp

increase in device temperature, so must be adjusted with care.

These results illustrate why both piezoresistor design optimization and accurate device fabrication are essential to building high performance piezoresistive sensors. If the device fabrication process does not match the assumptions used during the design (e.g. if α or R_s deviate by an order of magnitude from their expected values) then the device will no longer be operating at its optimal point and performance may degrade substantially.

The performance of several force probes in air and water is summarized in Table 5.2. All performance parameters are directly measured or calculated from validated experimental data. The dominant sources of resolution uncertainty are the spring constant and displacement sensitivity, and the error from each is propagated through to the sensor resolution calculation. Probe mechanics are determined by the sensor; the actuator at the base of the probe is relatively decoupled due to its much higher spring constant and resonant frequency. Operation in water requires the deposition of a 200-300 nm thick parylene N coating. The coating roughly doubles the spring constants of the probes while the increase in damping reduces their resonant frequencies and quality factors 3-5x. The minimum detectable displacements and forces are calculated in a measurement bandwidth from 1 Hz to the resonant frequency of the probe. The spring constant increase and bandwidth reduction in water partly cancel each other, resulting in a relatively small change in resolution. In both air and water the force resolution of the sensor is about 5x the thermomechanical force noise floor.

5.6 Actuator performance

We developed both thermal and piezoelectric on-chip actuators. We will discuss the thermal actuator performance before moving onto the piezoelectric actuator. Their capabilities are orthogonal to one another; the thermal actuator enables micron-scale deflections (e.g. 200-300 nm/K) with rise times on the order of 100 μ s, limited by the first-order response of the actuator. In contrast, piezoelectric actuators are only capable of sub-micron deflections (e.g. 10-100 nm/V) but the rise time is limited only by the mechanical frequency response of the probe and can be sub-10 μ s in water.

5.6.1 Thermal actuator

The frequency response of a thermally actuated force probe with a 61 μm long and 1 μm wide sensor combined with a 60 μm long actuator is presented in Figure 5.34a. For the measurements, the heater is DC biased (2 V) and then driven with a swept cosine (0.1 V) so that there is a linear component to the actuation [55]. The transfer function is calculated from the power dissipation at the AC drive frequency. The measured DC actuator response of 560 nm/mW closely matches the 550 nm/mW (110 nm/K) predicted from the finite difference-based thermal model. However, the measured -3 dB bandwidth is 3.1 kHz compared with the lumped parameter model prediction of 12.7 kHz. For comparison, prior micromachined scanning probes with integrated thermal actuation have demonstrated -3 dB actuation bandwidths in air on the order of 300 Hz [71], 1 kHz [50] and 2 kHz [411].

The time-domain step responses of the same 61 μm long force probe design in air and water are presented in Figure 5.34b. The absolute deflection of the force probe tip was measured on an upright microscope by projecting the image of the tip onto a differential photodiode with a measurement bandwidth of 230 kHz. The heater power dissipation required for a given tip deflection is significantly greater in water (280 nm/mW) than in air (560 nm/mW) for a comparable tip deflection. Actuation in air is characterized by a single time constant ($57 \pm 1 \mu\text{s}$). In contrast, the response in water is characterized by fast ($10 \pm 1.4 \mu\text{s}$) and slow ($108 \pm 20 \mu\text{s}$) components due to heating of the surrounding liquid. The 10-90% rise times in air and water are 92 ± 6.7 and $80 \pm 4.9 \mu\text{s}$. Actuation is about four times slower than the lumped parameter model predictions (Section 3.2.3), indicating that variation in temperature throughout the actuator can not be neglected.

Simultaneous actuation and sensing is demonstrated using the same force probe design (Figure 5.35). The actuator is biased with a 1 V step (1.6 mW) filtered at 10 kHz with an 8-pole Bessel filter. The Wheatstone bridge is biased at 1 V, the signal is amplifier 1000-fold, filtered at 10 kHz and sampled at 100 ksps. The input referred crosstalk is on the order of 700 $\mu\text{Vpp}/\text{mW}$ if a board-level reference resistor is used to complete the Wheatstone bridge. The crosstalk decreases to 400 $\mu\text{Vpp}/\text{mW}$ if the resistor is swapped for an on-chip reference cantilever, but crosstalk is limited by electrical and thermal mismatch between the main and compensation actuators. Crosstalk decreases dramatically to 15 $\mu\text{Vpp}/\text{mW}$ if a

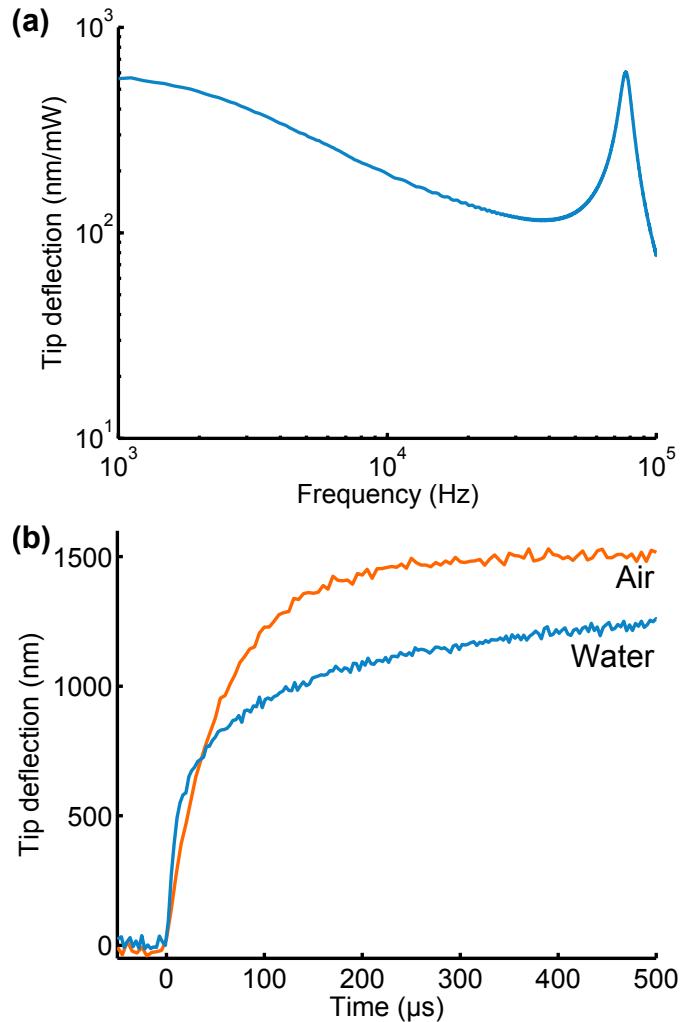


Figure 5.34: Thermal actuator frequency and step responses. (a) The frequency response in air has a significantly smaller bandwidth (3.1 kHz) compared with the lumped parameter model prediction (12.7 kHz) due to temperature variation within the actuator. (b) The step response in air is characterized by a single time constant. In contrast, the step response in water has both fast and slow components due to heating of the surrounding liquid. The resulting 10-90% rise times in air and water are 92 and 80 μ s, respectively.

resistor (50Ω) and potentiometer are added in series with the thermal actuators (nominally 600Ω) on the main and compensation probes, respectively. The residual thermal crosstalk signal corresponds to a temperature mismatch of 40 mK/mW and has a time constant (50 ms) that is several orders of magnitude slower than the actuator time constant.

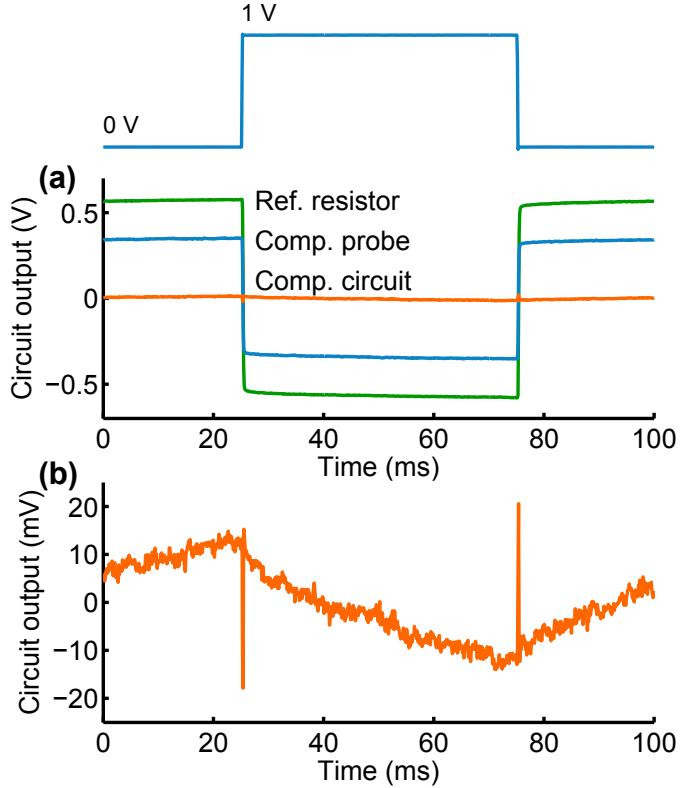


Figure 5.35: Simultaneous actuation and sensing. (a) Time-domain crosstalk between the thermal actuator and piezoresistive sensor for a 1 V (1.6 mW) step filtered at 10 kHz for three different crosstalk compensation strategies that use an off-chip reference resistor, on-chip compensation force probe, or the compensation probe combined with a resistive compensation circuit. (b) Detailed view of the fully compensated crosstalk performance. The residual input referred crosstalk is 15 μ V/mW (40 mK/mW). Capacitive coupling between the actuator and sensor is on the order of -90 dB at 10 kHz.

Once the thermal crosstalk is compensated, capacitive crosstalk becomes noticeable. The input referred capacitive crosstalk for an actuator step filtered at 10 kHz is -90 dB, smaller than the -70 dB [57] and -80 dB [65] results reported previously. Capacitive crosstalk is minimized through the combination of the compensation probe, metal rather than silicon interconnects to the piezoresistor, and the grounded, highly doped diffusion well that runs underneath the actuators for most of their length.

Recall that the force probes developed in this thesis were designed for the study of cochlear hair cells (Table 1.1). The fastest rise times to date for macroscale flexible and

rigid glass fibers are 150 μ s and 50 μ s, respectively. While the thermal actuator is a slight improvement on the flexible fiber measurement system, it is not nearly fast enough for the study of mammalian transduction channel kinetics, which would require a rise time of less than approximately 10 μ s.

5.6.2 Piezoelectric actuator

Piezoelectric actuation provides the possibility of much faster actuation. Both thermal and piezoelectric actuators are first-order systems, with the -3 dB frequency set by the electrical or thermal resistance and capacitance of the actuator. The -3 dB frequency of a typical thermal actuator was on the order of 3 kHz. In contrast, the resistance and capacitance of a typical piezoelectric actuator are on the order of 50 Ω and 800 pF, yielding a -3 dB frequency on the order of 3 MHz.

The amplifier used to drive the actuator is a possible rate limiter in the system. However, most op-amps are capable of driving an 800 pF actuator up to 100 kHz with ease; the peak current is on the order of $(100 \text{ kHz})(800 \text{ pF})(1 \text{ V}) = 80 \mu\text{A}$. For example, we often used the THS4031 or THS4131 (Texas Instruments) to drive the piezoelectric actuator (for single-ended and differential drive configurations, respectively). Both have an open loop bandwidth on the order of 100 MHz, a slew rate of 50-100 V/ μ s and peak output current of 30-90 mA.

The frequency response of a force probe in air is presented in Figure 5.36a. The DC response of the actuator is 8 nm/V while the resonant frequency and quality factor in air are 308 kHz and 27, respectively. The probes could be operated in a closed loop up to 181 kHz with a 45° phase margin.

Actuation and measurement bandwidth are limited only by the mechanical frequency response of the force probe sensor. There is a tradeoff between sensor length and tip deflection (Figure 5.36b). A longer sensor results in greater mechanical amplification of the actuator motion but a lower resonant frequency. Variation in tip deflection with sensor length implies an effective d_{31} coefficient of 1.3 pm/V in comparison with 2.4 pm/V recently reported for 100 nm thick AlN films [361].

The step response of an actuator operating in electrically grounded saline is presented

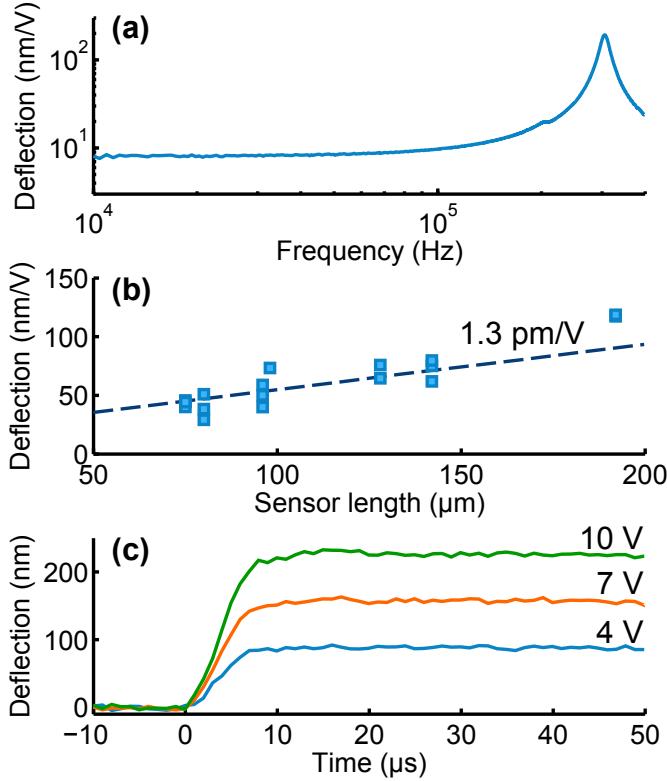


Figure 5.36: Piezoelectric actuator performance summary. (a) Example frequency response in air of a force probe with a resonant frequency of 303 kHz. (b) Tip deflection increases with sensor length due to mechanical amplification. The effective d_{31} coefficient of the 75 nm thick AlN actuator is approximately 1.3 pm/V . (c) Step response in water of a parylene passivated force probe. The 10-90% rise time is $5 \mu\text{s}$ and the tip deflection is 22 nm/V .

in Figure 5.36c. The probe deflection is measured by projecting the probe image onto a differential photodiode with a 230 kHz bandwidth readout circuit while the actuator signal is filtered at 80 kHz with an eight-pole Bessel filter. The probe response is critically damped with a 10-90% rise time of $5 \mu\text{s}$. The resistivity and breakdown voltage of the actuator were measured as approximately $10^{13} \Omega\text{-cm}$ and 10 V, respectively. The actuator is initially held at a negative potential (e.g. -5 V) to maximize the usable single-sided deflection range.

The typical film crystallinity of finished PRPE wafers is shown in Figure 5.37. The 300°C LTO deposition step after the AlN/Mo deposition process did not degrade film crystallinity based upon measurements; the rocking curve FWHM measured by the vendor immediately after film deposition was not measurably different from the final FWHM

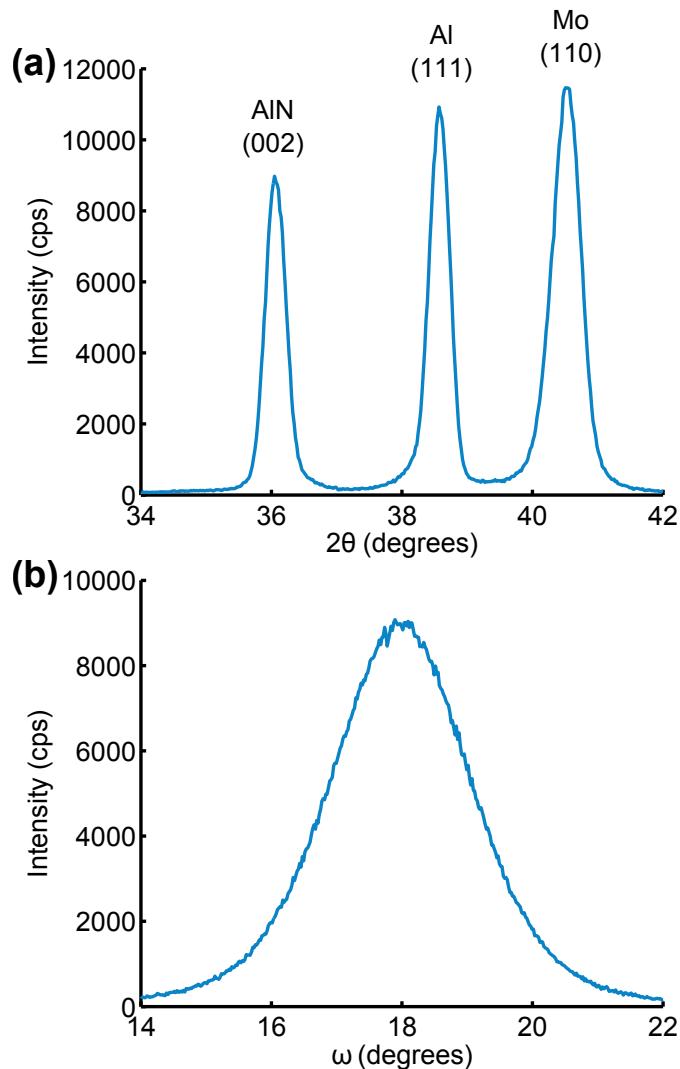


Figure 5.37: Typical film crystallinity measurement data. (a) $2\theta - \theta$ measurements show the presence of well oriented AlN, Al and Mo films. (b) Rocking curve measurements indicate excellent AlN grain orientation. The rocking curve FWHM values for 150 and 75 nm thick AlN films are 2.5 and 3.3 degrees, respectively.

measured on finished wafers. The XRD measurements were performed on an unpatterned location near the center of the wafer, which is why the AlN, Mo and Al peaks are all visible. The measured (nominal) $2\theta - \theta$ peaks for AlN (002), Al (111) and Mo (110) peaks are 36.05 (36.04), 38.57 (38.44) and 40.53 (40.60) degrees, respectively. The only other expected peaks in this angle range are Ti (100), (002) and (101) peaks but they are not

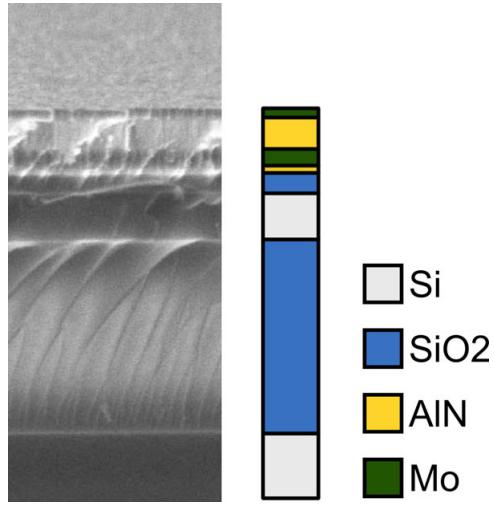


Figure 5.38: Cross-sectional SEM of an actuator with a 150 nm thick AlN film. The AlN and Mo grains are clearly oriented and each layer of the stack (Si/SiO₂/Si/SiO₂/AlN/Mo/AlN/Mo) is clearly visible. The thickness of the deposited films were verified via cross-sectional SEM and matched the nominal film thicknesses to within ± 5 nm.

detectable. The XRD rocking curve FWHM determines how uniformly the AlN grains are oriented. The 75 and 150 nm thick AlN actuators that we fabricated had FWHM values of 3.3 and 2.5 degrees, respectively. When deposited on polysilicon rather than single crystal silicon the FWHM of the 75 nm thick film increased to 3.8 degrees due to the increased substrate roughness.

A cross-sectional SEM of an actuator with a 150 nm thick AlN film is presented in Figure 5.38. The AlN and Mo grains are clearly oriented and each layer of the stack (Si/SiO₂/Si/SiO₂/AlN/Mo/AlN/Mo) is clearly visible. The thickness of the deposited films were verified via cross-sectional SEM and matched the nominal film thicknesses to within ± 5 nm.

The frequency response of a piezoelectrically actuated probe in air is presented in Figure 5.39 and illustrates the separation between the probe and actuator modes. The frequency response was measured via LDV at the tip of the probe (top curve) and the end of the actuator (bottom curve). The DC response of this probe is 50 nm/V and three resonant modes are seen in the curves. The first resonant mode (21 kHz) corresponds to the bending mode of the long silicon extension. The second mode (54 kHz) corresponds to

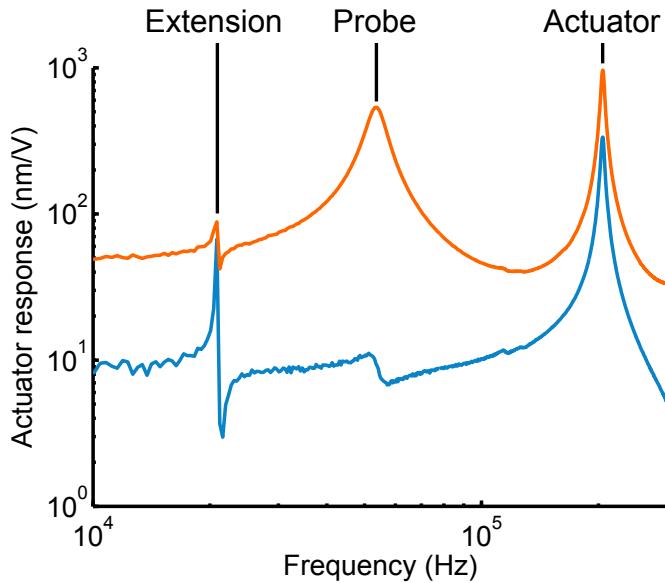


Figure 5.39: Frequency response of a piezoelectrically actuated probe in air. The frequency response is measured via LDV at the tip of the probe (top curve) and the end of the actuator (bottom curve). The results confirm the separation between the first resonant modes of the sensor and actuator.

the sensor portion of the force probe, and it nearly vanishes when the measurement laser is placed on the actuator. The third mode (205 kHz) corresponds to the actuator portion of the probe. The actuator deflection is mechanically amplified 3-fold between the actuator and the tip of the probe. The results confirm the separation between the first resonant modes of the sensor and actuator and that the higher order modes (Figure 3.8) do not come into play.

The frequency response of the main and compensation probes on the same silicon die are presented in Figure 5.40. Matching between the main and compensation actuator is critical to minimize both mechanical and capacitive crosstalk from the actuator to the piezoresistive sensor. The main probe (orange line) has a resonant mode at 20 kHz from the long silicon extension that is missing from the compensation probe response (blue line). The DC responses of the main and compensation actuators are 117 and 118 nm/V, respectively, while the actuator responses on resonance are 212 and 190 nm/V, respectively. The first resonant mode of both probes is at 6.3 kHz. The DC actuator mismatch is less than 1% and the mismatch on resonance is on the order of 10% due to quality factor mismatch.

The primary downside of moving the actuator on-chip is increased crosstalk with

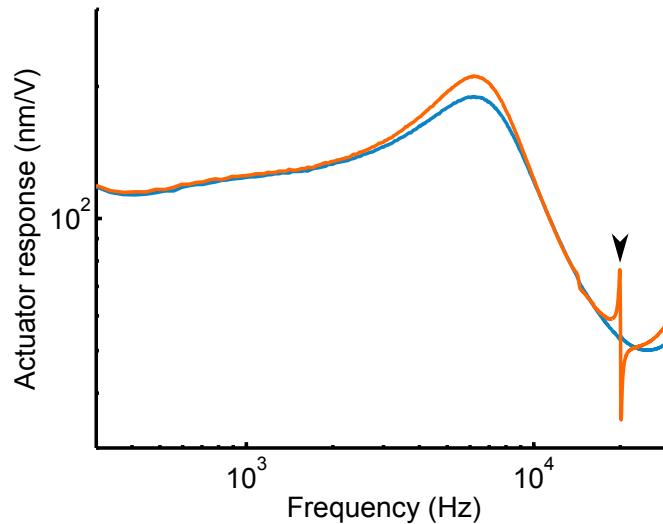


Figure 5.40: Example of actuator matching between the main and compensation force probes on a single device. The main probe (orange line) has a resonant mode at 20 kHz from the long silicon extension that is missing from the compensation probe response (blue line). The DC responses of the main and compensation actuators are 117 and 118 nm/V, respectively, while the actuator responses on resonance are 212 and 190 nm/V, respectively. The first resonant mode of both probes is at 6.3 kHz.

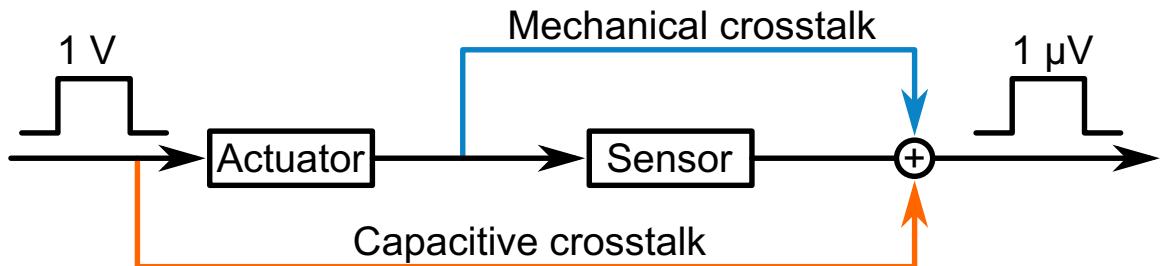


Figure 5.41: Piezoelectric actuator-sensor crosstalk summary. The actuator is mechanically and capacitively coupled to the sensor. The noise floor is the sensor is 120 dB smaller than the typical actuator signal. We address crosstalk by taking advantage of common-mode rejection, electrically isolating the actuator and sensor portions of the device, and trimming any remaining capacitive crosstalk on the readout circuit.

the piezoresistive sensor. Crosstalk comes from both mechanical and capacitive sources (Figure 5.41a). Mechanical crosstalk is from bending of the metal interconnects that run adjacent to the piezoelectric actuator, and is limited by the interconnect gauge factor.

Capacitive crosstalk stems from coupling through the substrate and directly between

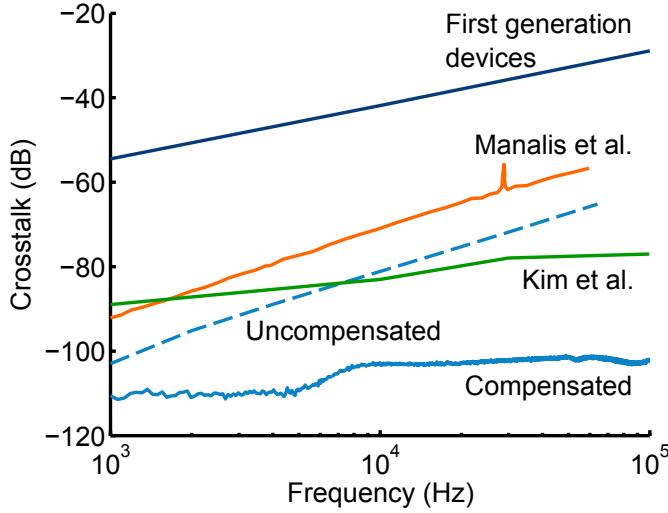


Figure 5.42: Piezoelectric actuator-sensor crosstalk comparison. The uncompensated and compensated performance of the second generation force probes is compared with the first generation probes described in Appendix F and the results reported by Manalis [57] and Kim [65]. The compensated crosstalk performance is about 20 dB better than all previous work and is limited to -115 dB by mechanical crosstalk.

the adjacent electrodes. We combined four approaches to minimize crosstalk. First, the optimal piezoresistor impedance was fairly low, ranging from 2-5 k Ω , minimizing the voltage transduced by the capacitive current. Second, we fabricated identical main and compensation force probes so that crosstalk is primarily limited by probe mismatch. Third, we highly doped and grounded the silicon underneath the piezoelectric actuator. Finally, we trimmed any capacitance mismatch between the main and compensation probes with a variable capacitor-based scheme on the circuit board used for sensor readout (Appendix C).

The force probe crosstalk performance is compared with prior work in Figure 5.42. The uncompensated and compensated performance of the second generation force probes is compared with the first generation probes described in Appendix F and the results reported by Manalis [57] and Kim [65]. Crosstalk is limited to -115 dB by mechanical crosstalk and is below -100 dB to 100 kHz. The mechanical crosstalk level corresponds to 1.8 ppm of resistance mismatch between the main and compensation probes per volt applied to the actuator. The capacitive crosstalk at all frequencies is approximately 20 dB less

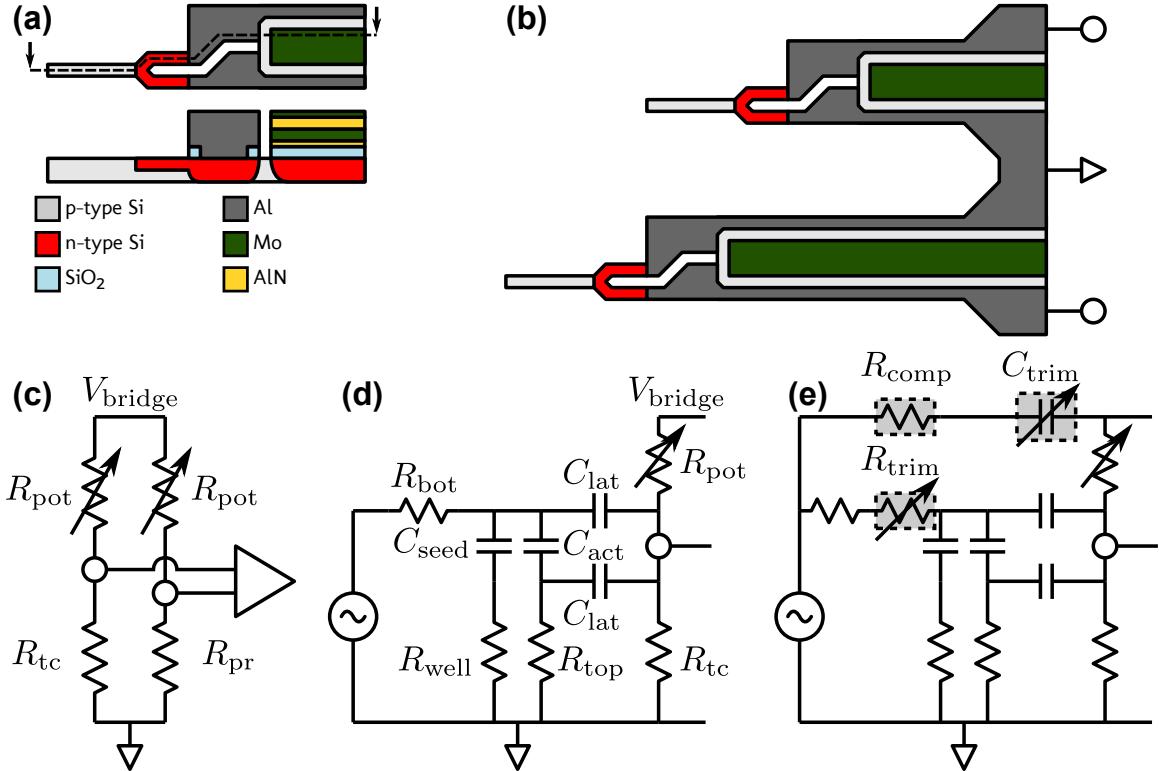


Figure 5.43: Capacitive crosstalk compensation circuit. (a) Force probe schematic layout. (b) The main and compensation probes on the silicon die are connect to the instrumentation amplifier inputs and share a common ground connection. (c) The entire piezoresistive sensing circuit with two off-chip potentiometers and two on-chip piezoresistors. (d) One half of the actuation and sensing circuit. Capacitive crosstalk is in the uncompensated circuit is limited by mismatch in the lateral capacitances between the actuator and piezoresistor electrodes. (e) The addition of two resistors and one capacitor to the actuator circuit allows the capacitance mismatch to be trimmed out.

than previous devices with integrated piezoresistive sensors and piezoelectric actuators. The crosstalk is comparable to the piezoresistor noise floor depending on the actuation bandwidth and amplitude.

The crosstalk compensation circuit is summarized in Figure 5.43. A detailed view of a single force probe and how the main and compensation probes are connected together to form the sensing circuit are presented in Figure 5.43a-c. The crosstalk and compensation circuit topologies are presented in Figure 5.43d-e.

The compensation probe side of the circuit (V_{amp}^-) has a variable capacitor (C_{trim}), fixed

series resistance (R_{comp}) and series potentiometer (R_{trim}). The main probe side of the circuit (V_{amp}^+) does not have the series potentiometer and C_{trim} is a fixed capacitor. The lateral capacitance mismatch between the two sides of the circuit is typically on the order of 2 pF, and a 3-10 pF variable capacitor is typically matched with a 6.5 pF fixed capacitor. The series resistance ensures that the time constant of the compensation capacitor and the lateral capacitances are similar in magnitude. Best results are typically obtained with $R_{\text{comp}} \approx 1 \text{ k}\Omega$. The common mode rejection ratio (CMRR) of the instrumentation amplifier was not a limiting factor in crosstalk performance; SPICE simulations indicate that the CMRR would support crosstalk of as little as -160 dB.

The series potentiometer is necessary because the main and compensation actuators are not identical in shape and size. The long extension on the main probe leads to a larger resistance and capacitance for the main actuator (roughly 45Ω and 900 pF, which includes the actuator and seed/LTO capacitance) than the compensation actuator (roughly 13Ω and 225 pF). The series potentiometer (R_{trim}) adds resistance to the bottom electrode on the compensation side of the circuit so that the main and compensation actuators have similar time constants.

Note that the crosstalk circuit assumes that there is negligible capacitive coupling between the diffusion well and the piezoresistor interconnects. This is a good approximation due to the low impedance from the diffusion well to ground and the combination of the wide depletion region between the n-type well and the isolation trench that is etched around it along most of the length of the device. If the diffusion well is left floating or not heavily doped then the crosstalk performance degrades 40-60 dB.

In order to debug the actuator-sensor crosstalk, a small reference board was built with just two surface mount resistors ($2 \text{ k}\Omega$) and capacitors (2.2 nF). The force probe and reference board crosstalk performance is compared in Figure 5.44. While the crosstalk on the reference board decreases monotonically, the force probe crosstalk flattens out below 30 kHz. The first and second eigenmodes of the sensor are visible in the crosstalk spectrum. The results clearly indicate that mechanical crosstalk is the limiting factor at moderate frequencies.

Due to the fact that mechanical crosstalk plays an important role in the system performance, crosstalk rejection depends on the Wheatstone bridge bias. Crosstalk spectra

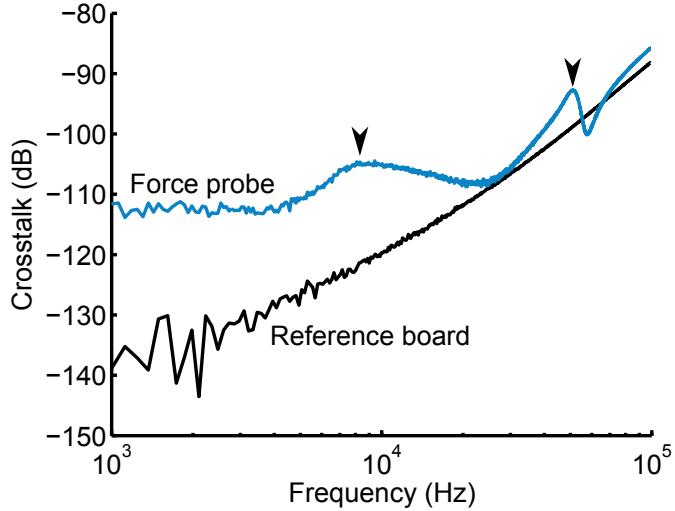


Figure 5.44: Crosstalk spectra from a PRPE force probe and a reference printed circuit board. The reference board includes a pair of surface mount resistors and capacitors in order to emulate the electrical characteristics of the force probes. The reference board crosstalk increases monotonically with frequency and is limited purely by capacitive mismatch. The force probe deviates from the reference board due to the presence of mechanical crosstalk, and the first and second eigenmodes of the force probe are visible in the spectrum.

recorded from the same device presented in Figure 5.44 were measured for bridge bias voltages of 50 mV, 250 mV and 1 V (Figure 5.45). Additionally, crosstalk performance depends on resistance matching between the two sides of the Wheatstone bridge. Even if both sides are balanced (i.e. the amplifier inputs are at $V_{\text{bridge}}/2$) there could be large capacitive crosstalk if the resistive impedance seen looking out from each amplifier input do not match. Depending on how well the main and compensation piezoresistors are matched, the crosstalk at 100 kHz can be limited by mechanical crosstalk (Figure 5.42) or capacitive crosstalk (Figure 5.45).

Mechanical crosstalk can be minimized by reducing the resistance of the interconnects that run alongside the actuator, reducing their gauge factor, and reducing their stress. These goals are orthogonal to one another and should be optimized if one wanted to improve the mechanical crosstalk performance. For example, increasing the metal interconnect thickness reduces the resistance, keeps the gauge factor constant, but increases the stress as the metal moves further from the neutral axis. Metal is superior to silicon in the first two

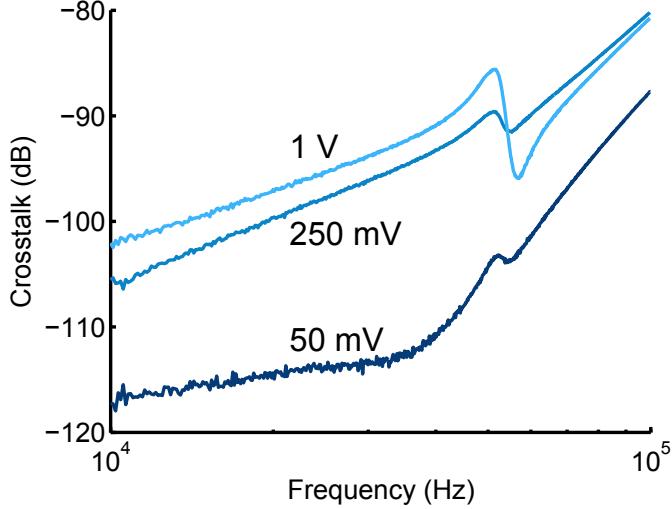


Figure 5.45: Actuator-sensor crosstalk increases with Wheatstone bridge bias. Crosstalk spectra were recorded for bridge biases of 50 mV, 250 mV and 1 V. As the bridge bias increases the mechanical crosstalk between the actuator and sensor is amplified. Capacitive crosstalk also increases due to degraded matching between the main and compensation force probes as their temperatures increase. The resonant peak and dip correspond to the second eigenmodes of the main and compensation force probes, respectively.

respects (much lower resistivity and gauge factor) although it is typically slightly further from the neutral axis than a silicon resistor could be [58].

The crosstalk circuit (Figure 5.43) illustrates that driving the actuator differential will result in less crosstalk. Mismatch in the lateral capacitances between the main and actuator probes is the main contributor of capacitive crosstalk. If the top and bottom electrodes are driven differentially then the net capacitive mismatch on each side of the circuit as well as the overall mismatch decrease. Figure 5.46 illustrates the capacitive crosstalk of a force probe that was driven with single-ended and differential signals. A differential drive circuit yields a crosstalk improvement of about 15 dB at 200 kHz and could probably be improved further. The dip in the single-ended spectrum and enhancement in the differential spectrum both correspond to the first resonant mode of the force probe; the amplitude changes due to the difference in phase between the capacitive and mechanical crosstalk signals on resonance. A differential drive circuit can be implemented fairly simply with a differential op-amp in an inverting amplifier configuration. The combined sensing, actuation and

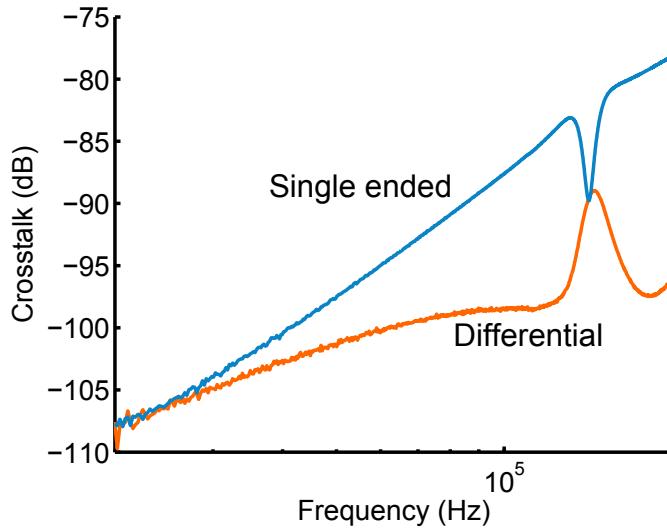


Figure 5.46: Driving the piezoelectric actuator differential leads to significantly smaller crosstalk. In the single ended drive configuration the top electrode and diffusion well are grounded while the bottom electrode is biased. In the differential configuration the diffusion well is still grounded but the top and bottom electrodes are negatively and positively biased, respectively, using a fully-differential amplifier. Partial cancellation between the positive and negative drive signals leads to a small common-mode capacitive crosstalk signal and improved crosstalk (20 dB at 200 kHz). However, the single-ended bottom electrode drive configuration is required for simultaneous force- and electrophysiology measurements as we will see in the next chapter.

compensation printed circuit includes a toggle switch to select between single-ended and differential drive (Appendix C). Both options were used for characterization, but only single-ended drive of the bottom electrode was used for hair cell experiments in order to reduce crosstalk with the patch clamp electrode.

The step response of a piezoelectrically actuated probe in the time domain is presented in Figure 5.47. In this set of curves the actuator is driven with 1, 2 or 4 V single-ended steps. The tip of the force probe is unloaded and both the actuator and sensor signals are filtered at 10 kHz with an 8-pole analog Bessel filter. Crosstalk is primarily mechanical rather than capacitive up to 10 kHz, and the steps are repeated 400 times before the results are averaged and plotted. The DC crosstalk (e.g. 2.5 mV for a 4 V actuator step) corresponds to input referred mechanical crosstalk of -125 dB and is limited by mismatch between the main and compensation force probes. The spring constant, resonant frequency and actuator

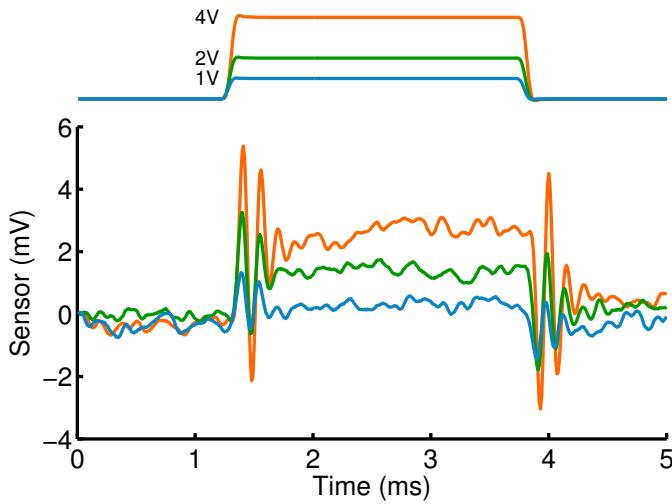


Figure 5.47: Time-domain measurements illustrate mechanical crosstalk between the actuator and sensor. The tip of the sensor was unloaded, so that only inertial forces are acting on the sensor. Both the actuator and sensor signals are filtered at 10 kHz, and probe resonant frequency is 8 kHz. The DC crosstalk corresponds to input referred mechanical crosstalk of -125 dB.

deflection response for this device are 0.65 mN/m (after parylene coating), 7.3 kHz (in air) and 75 nm/V. A sensor output change of 1 mV corresponds to a force of about 0.2 pN. The ringing in the sensor signal is due to inertial loading of the sensor and can be used to calibrate the frequency response of the device, particularly when Q is low.

A reasonable question at this point is whether the crosstalk problem is solved. Figure 5.47 illustrates that capacitive crosstalk is not detectable for force probes operated at relatively low frequencies. As the operation frequency increases, capacitive crosstalk will eventually perturb the piezoresistor signal. But capacitive crosstalk should be repeatable and, once characterized, could be subtracted from the sensor output for open loop experiments. Figure 5.47 also illustrates the magnitude of the mechanical crosstalk artifact in the signal. For a tip deflection of about 300 nm the mechanical crosstalk signal corresponds to a force artifact of 0.5 pN. If both the force probe and sample have a spring constant of 1 pN/nm then the 300 nm deflection will be split equally between them, yielding a force of 150 pN. For most applications, a 0.5 pN offset on top of a 150 pN signal would be negligible. The results suggest that neither mechanical or capacitive crosstalk will be a

limiting factor for measurement bandwidths below 10 kHz and possibly up to 100 kHz that the force probes would be used for.

For applications that require more precision, mechanical crosstalk compensation could easily be built into the measurement circuit. For example, the actuator voltage could be buffered, inverted, manually trimmed with a potentiometer to cancel the residual mechanical crosstalk, and summed at the input to the variable gain amplifier along with the sensor output. The capacitive crosstalk results demonstrate that relatively simple analog circuits can be used to compensate any repeatable artifacts in the sensor output.

Chapter 6

Cochlear hair cell experiments

In this chapter we will apply the force probes to the study of the hair cells in the mammalian cochlea. We will briefly summarize the experimental setup before we move onto the integration of the force probes with a patch clamp measurement system for simultaneous hair cell stimulation and electrical current readout. The experimental work in this chapter was largely performed by Anton Peng of the Ricci Lab at Stanford.

6.1 Experimental setup

The force probes were specifically designed for compatibility with the standard experimental setup used to study cochlear hair cells. Hair cell protocols are described in detail elsewhere (e.g. Ref. [412]) and we will only provide a high level overview of the setup.

The excised tissue sample that contains the hair cells is mounted underneath the microscope objective and the patch clamp pipette and force probe are brought in from the side at a 10-15° angle (Figure 6.1). The force probe is positioned using a three-axis micromanipulator which allows the probe to be placed in contact with the bundle before the microscale actuator is utilized. The probes are mounted on a custom PCB (Figure 5.4), which is attached to an aluminum rod using a pair of screws (Figure C.3). The rod provides an interface between the PCB and micromanipulator. The mechanical response of the rod does not affect device performance because it is far stiffer and more massive than the rest of the device and only on-chip actuation is used. A force probe is shown near a row of

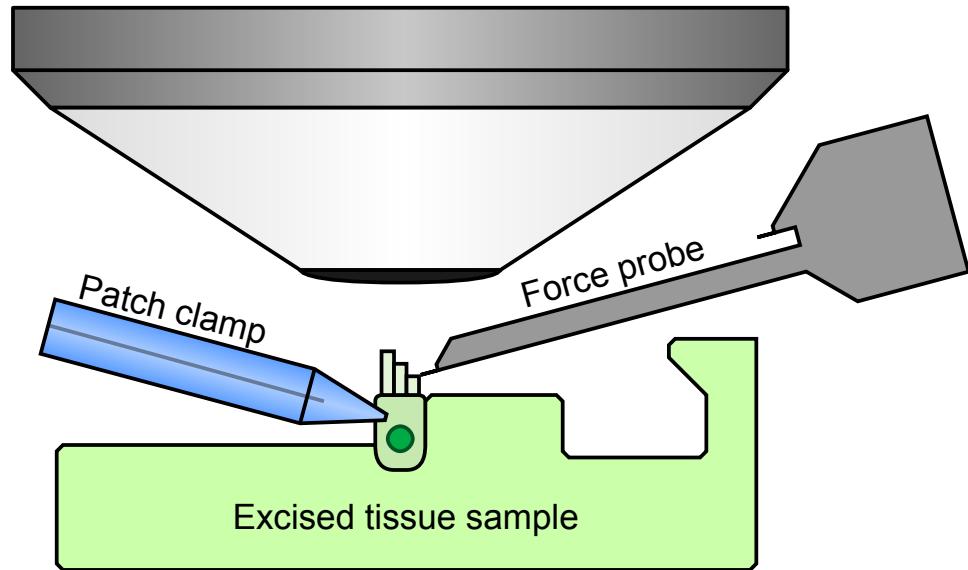


Figure 6.1: Experimental setup for studying cochlear hair cells. The excised tissue sample is mounted on an upright microscope. The patch clamp pipette and force probe are mounted on three-axis micromanipulators and are brought underneath the microscope objective from the side at a 10-15° angle.

IHCs in Figure 6.2.

After the force probe is brought into the hair cell bath and all experiments have been performed, the water is replaced with methanol to prevent stiction (Figure 5.11). While relatively stiff devices ($k_c > 10 \text{ mN/m}$) can be removed from water without incident, the softer probes intended for studying bundle mechanics would break at the air-water interface due to the high surface tension of water. The bath should be replaced with methanol 5-10 times in order to remove all water and the force probe should be remain stationary while the methanol evaporates. Heating the methanol to 40-50°C speeds evaporation and reduces the surface tension even further.

If a probe will be used to study OHCs, where the hair cell bundle has a V-shape, then a 2-10 μm diameter borosilicate glass bead (Duke Scientific) is attached to the tip of the probe after it has been parylene coated. The beads are attached to the probes on the same microscope where the experiments are performed using the three-axis micromanipulator that the force probes are mounted on. The beads are adhered to the device using a UV curable adhesive (Loctite 352) although we have also used two-part epoxies with success

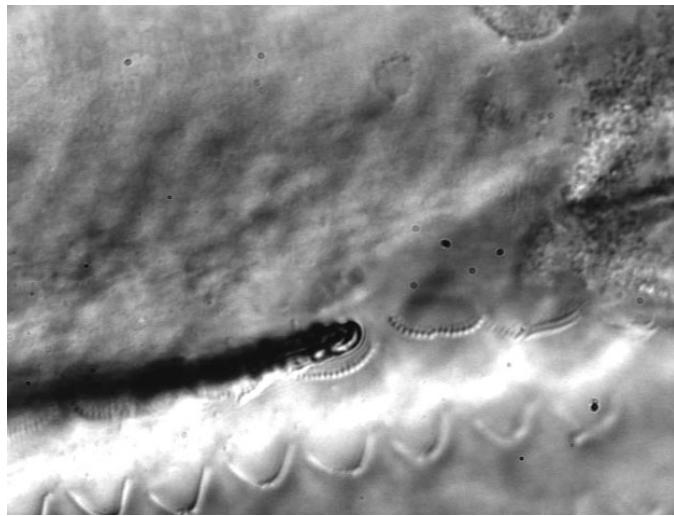


Figure 6.2: Micrograph of a force probe and IHCs. The probe is mounted on a three-axis micromanipulator with a step size of approximately 50 nm and brought into contact with the bundle before the stimulus is delivered using the on-chip actuator.

(Devcon). The UV curing variety simply provides more time for the bead attachment process. The attachment process is depicted in Figure 6.3. The beads and epoxy are spread on a glass slide. The tip of the probe is dipped in adhesive near the edge of an adhesive drop so that the probe is not pulled in too far via surface tension. Next the probe is positioned over a bead and brought into contact with it. When the probe is retracted from the glass slide the bead comes with it. Finally, the adhesive is cured for 15-20 minutes using a UV lamp (Spectroline EN-180) and allowed to sit for at least 24 hours to allow the adhesive to fully crosslink. If the probes are used too soon then the adhesive may delaminate from the underlying parylene.

6.2 Patch clamp crosstalk

In the last chapter we discussed capacitive and mechanical crosstalk between the on-chip piezoelectric actuator and piezoresistive sensor. Similarly, the on-chip actuator is capacitively coupled with the hair cell and patch clamp pipette. A top-down schematic view of the force probe pressing on a hair bundle is presented in Figure 6.4. The capacitance between the force probe and the liquid is governed by the thickness and dielectric constant

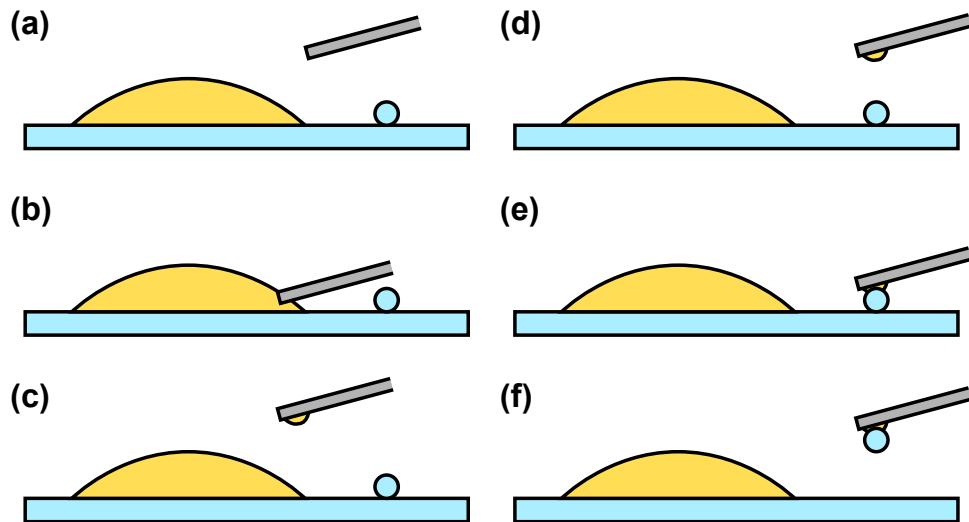


Figure 6.3: Process for attaching beads to the force probes. (a) The probe is mounted on a three-axis micromanipulator and approaches a glass slide with a drop of UV curable adhesive and glass beads on it. (b) The tip of the probe is dipped in adhesive near the edge of a drop. (c) The probe is retracted and (d) positioned over a bead. (e) The probe is brought into contact with a bead and (f) retracted from the slide, carrying the bead away with it. Afterwards the adhesive is cured for 15-20 minutes with a UV lamp.

of the parylene passivation layer. The electrically conductive bath that the experiments are performed in is grounded, but the Ag/AgCl reference electrode is relatively far away from the probe and cell. The capacitance between the force probe and liquid is dominated by the top electrode and diffusion well. A secondary source of capacitive coupling is the PCB that the force probe is mounted on. Current from the actuator enters the patch clamp signal through the cell membrane capacitance and patch pipette capacitance. The cell capacitance is typically about twice the pipette capacitance (5 pF vs. 10 pF).

The capacitive artifact in the patch clamp signal is proportional to the rate at which the actuator voltage changes and the capacitances involved. Accordingly, capacitive crosstalk is primarily an issue for kinetics measurements, both because the actuator signal is filtered at a higher corner frequency and because we are interested in the very beginning of the patch clamp signal when the channels are just beginning to open. The cell and patch clamp pipette capacitances are relatively fixed so we focused our effort on minimizing the probe capacitances.

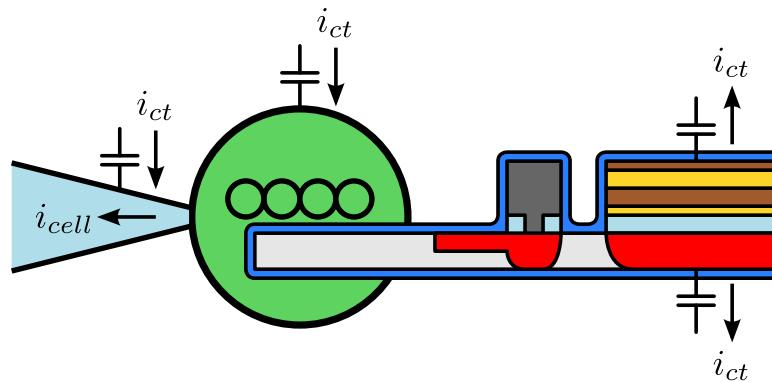


Figure 6.4: The on-chip actuator is capacitively coupled with the hair cell and patch clamp pipette. The capacitance between the force probe and surrounding liquid is dominated by the top electrode, diffusion well and PCB that the probe is mounted on. The capacitance between the patch clamp signal path and the liquid is set by the cell membrane and pipette capacitance. In order to minimize actuator to patch clamp crosstalk the top electrode and diffusion well on the force probe are grounded while the bottom electrode is biased and a four-layer PCB is used to shield the signal lines from the liquid.

We combined several approaches to minimize the size of the capacitive artifact. The most important step for minimizing crosstalk is to ground the top actuator electrode and diffusion well on the force probe and only drive the bottom actuator electrode. This effectively shields the actuator voltage between two ground planes, with only the edges of the bottom electrode exposed to the liquid. Next, we mounted the devices on a four-layer PCB with the sensor and actuator signals routed in the inner two layers and the outer planes serving as shields.

Figure 6.5 presents capacitive artifacts in the patch clamp signal for two different actuator stimulus rates. The patch clamp pipette and force probe were brought to within several microns of each other to record the capacitive artifact. The capacitive artifact does not change noticeably if the patch clamp is sitting in solution or patched on to a cell. The 8 V actuator voltage step is filtered with an 8-pole Bessel filter at either 10 kHz or 80 kHz. No capacitive artifact is present in the patch clamp signal at 10 kHz, while at 80 kHz the peak capacitive current is about 500 pA (65 pA/V). If the top piezoelectric actuator electrode is biased rather than grounded the size of the artifact increases 10 to 30-fold (e.g. 5 to 15 nA at 80kHz). Similarly, if a two-layer board without shield layers is used rather than the shielded four-layer design then the artifact increases by about 30 pA/V (e.g. 240 pA at

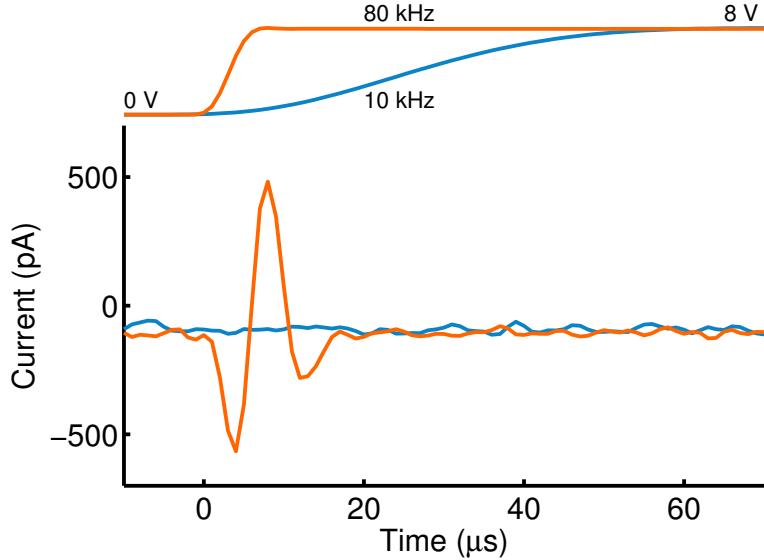


Figure 6.5: Example of capacitive coupling between the on-chip actuator and patch clamp. The force probe is positioned near but not in contact with a voltage clamped rat IHC. The on-chip actuator is driven with an 8 V voltage step filtered at 10 kHz or 80 kHz with an 8-pole Bessel filter. In order to minimize capacitive coupling, the force probe is mounted on a shielded 4-layer PCB and the bottom actuator electrode is driven while the diffusion well and top electrode are grounded. No capacitive artifact is observed at 10 kHz. At 80 kHz the artifact peak is about 500 pA and it is gone within 17 μ s.

80kHz).

6.3 Hair cell measurements

Finally, we demonstrated that the probes could be used to study the mechanics and kinetics of mammalian cochlear hair cells. In this section we will briefly present proof of concept data that were obtained using the fast force probes presented in this thesis. The mechanoreceptor current from a rat IHC is presented in Figure 6.6. The actuator stimulus is filtered at 80 kHz with an 8-pole Bessel filter and the probe movement is measured optically by projecting its shadow onto a differential photodiode. The patch clamp and photodiode signals are sampled at 1 MHz with parallel 16-bit ADCs. The force probe rise time, roughly 6 μ s, is significantly faster than the mechanoreceptor current.

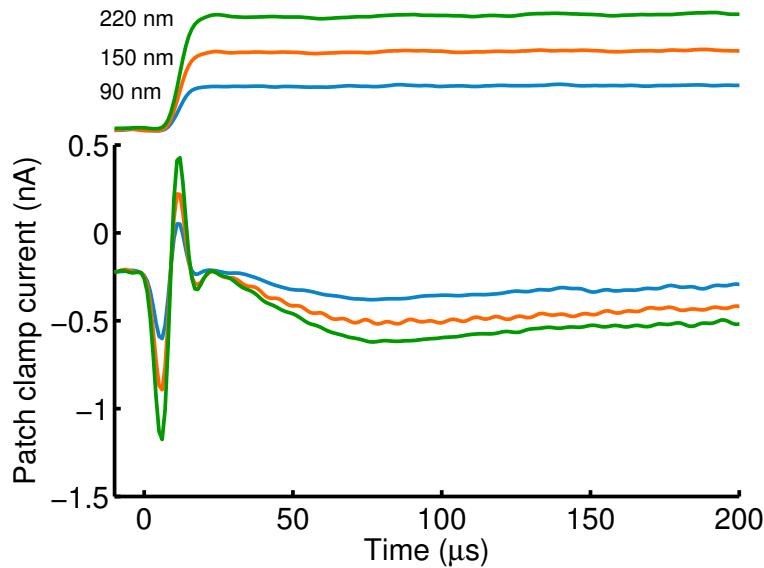


Figure 6.6: Example mechanoreceptor current from a rat IHC. The bundle is stimulated with a $45\text{ }\mu\text{m}$ long, $2\text{ }\mu\text{m}$ wide piezoelectrically actuated force probe. The actuator voltage step is filtered at 100 kHz with an 8-pole Bessel filter. The force probe movement (top traces) are measured with a differential photodiode while the mechanoreceptor current is measured with a patch clamp.

As of the thesis submission date, hair cell experiments are ongoing. The results in this chapter demonstrate that the integration challenges (e.g. simultaneous stimulus delivery and patch clamp measurements) are manageable through careful system design. Surprisingly, the speed of the patch clamp is currently the slowest component of the system and is the limiting factor for the kinetics experiments. However, the hair bundle mechanics experiment is the application for which the force probes are uniquely suited (i.e. where the force probe needs to be simultaneously soft and fast, rather than just fast) and where the probes should hopefully enable high bandwidth measurements of the active mechanical process in outer hair cells.

Chapter 7

Conclusions

The aim of this thesis was to develop a better, smaller, faster force probe for the study of cochlear hair cells. In order to achieve that goal we had to develop new design and fabrication processes along the way. In this chapter I want to look back on the project and talk about some of the ways that the devices could be improved.

7.1 Design optimization is a Swiss Army knife

Design optimization played a central role in the project, which was born in the spring of 2008. Shortly after passing my qualifying exam, I moved to Zurich for a few months and worked in Brad Nelson's microrobotics lab at ETH. When I talked with my co-workers at ETH about trying to measure pN-scale forces in a 100 kHz bandwidth they were fairly skeptical. Slightly worried, I started designing cantilevers by hand based upon Jonah Harley's 2000 design paper [75]. It looked feasible, but designing devices by hand quickly becomes unwieldy. So I wrote a set of Matlab scripts in order to make it easier to calculate device performance from a set of design parameters.

Before leaving for Zurich, Sung-Jin Park presented slides at a group meeting on his ion implantation design optimization work. He used an algorithmic approach to optimization rather than a more general gradient-based one, because at the time we all thought that numerical optimization would not have routinely found the global optimum. So I decided to try implementing numerical design optimization to see if there was a way to make it

work. Its strengths weren't obvious at the time, and it was mainly a way to double-check Harley's design guidelines.

I started work on a Python implementation of a numerical optimizer (using SciPy and NumPy) in order to use object oriented programming. This initial attempt failed, probably due to not rescaling the design parameters to $O(1)$ between iterations, so I went back to the drawing board. Coincidentally, Mathworks had just added objects to Matlab R2008a, making it feasible to implement numerical optimization in Matlab. After a few weeks of trial and error I eventually got it working and was able to confirm that the results matched Harley's 2000 paper when the design constraints were identical. But suddenly we could easily modify the design constraints and look at a much wider range of device designs.

After improving the optimizer some more and implementing randomized optimization, Sung-Jin, Beth and I decided to write it up as a paper. With the initial design optimization scaffold complete we were able to pile on the features. Over time I added amplifier and thermomechanical noise, thermal modeling, and added support for ion implantation, polysilicon, metal strain gauges and piezoelectric films. Design optimization was the glue that held the project together; it lead us to compare piezoresistive and piezoelectric sensing [43], which subsequently made us go back to piezoresistive sensing and critically examine Joule heating and thermal modeling [302].

I wanted to relate the history of the project in order to encourage the use of numerical design optimization. The initial purpose of the design optimization code was fairly modest (verify existing design guidelines) and it wasn't immediately obvious how important good design would be to the success of the project. Along those lines, numerical design optimization is an underutilized component in the MEMS designer's toolbox, probably because it seems like a large investment of time and resources without much benefit. I would argue that numerical optimization is both easier to implement than you might think (lumped parameter modeling can get you very far and you can use Appendix H as an example to get started) and that it will eventually prove more useful than you could imagine at the outset.

7.2 Fundamental force sensing limits

In retrospect, my application was actually quite similar to Jonah Harley's initial goal of obtaining a 1 pN force resolution in a 10 Hz to 100 kHz bandwidth with $k_c \approx 10$ mN/m for DNA sequencing [22]. At the time he identified several challenges with piezoresistive cantilevers: that 1/f noise increased more than expected as he reduced the cantilever dimensions (leading to the identification of Hooge noise as the dominant 1/f noise source) and that operation in liquid was troublesome due to the unpassivated metal leads corroding (which we have since solved at the expense of increased cantilever spring constant). In the end, he concludes that piezoresistive cantilevers are unlikely to meet the requirements and that the smart money would be on optical detection techniques.

Given another decade of piezoresistive cantilever design development it is worth revisiting the DNA unzipping application. As an aside, high-throughput DNA sequencing is in a very different state than it was a decade ago (e.g. nanopores) and I don't see force-based sequencing playing a role the future, but it provides a good opportunity to consider the fundamental limits of high bandwidth force sensing. Figure 7.1 compares the force resolution and several key design parameters for piezoresistive cantilevers as a function of thickness for two design cases.

The first case assumes operation in vacuum and constrains the power dissipation to 2.5 mW, representative of the state of the art circa-2000. The resonant frequency of the vacuum cantilever is constrained to 200 kHz in order to ensure a flat frequency response up to 100 kHz.

The second case assumes operation in water and constrains the maximum and tip temperature increases during operation to 10 and 1 °C, respectively. We assume a minimum resonant frequency in water of 100 kHz; the increased damping leads to a reduced quality factor and flatter frequency response, although the overall length of the cantilever must still decrease. The second design case represents the state of the art circa-2012.

In both cases we assume an epitaxial phosphorus piezoresistor and that the cantilever width is three times its thickness. Both cantilever designs account for the reduction in thermal conductivity with decreasing silicon thickness and the increase in Johnson and thermomechanical noise with cantilever temperature. We assume that elastic modulus is

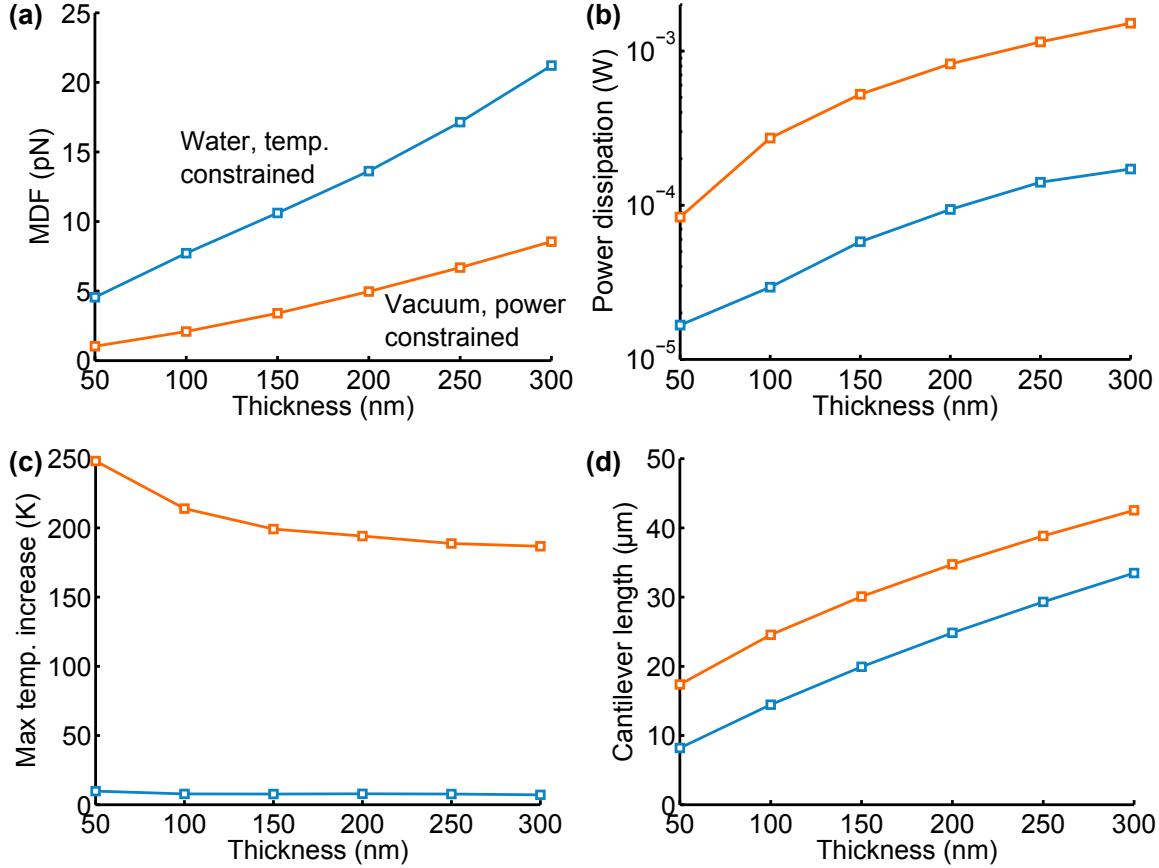


Figure 7.1: Comparison between power and temperature constrained design methods. (a) Force resolution from 10 Hz to 100 kHz as a function of cantilever thickness for two different design methodologies. The first set of designs assume operation in vacuum with a maximum power dissipation of 2.5 mW. The second design assumes operation in water with the maximum and tip temperature increases constrained to 10 and 1°C, respectively, taking advantage of improvements in cantilever design and modeling in the past decade. (b) The ideal power dissipation scales with cantilever thickness and is less than 2.5 mW in both cases due to piezoresistor self-heating, which (c) heats the cantilever in vacuum more than 200 K. (d) The cantilever in water is about 10 μm shorter than the vacuum design in all cases.

independent of cantilever thickness for simplicity.

The results illustrate that as cantilever designs have become more realistic (i.e. one cannot sequence DNA in vacuum with a cantilever tip temperature of 250°C) the performance of piezoresistive cantilevers has degraded modestly. The design improvements degrade

cantilever force resolution roughly 4-fold at any given thickness, although the cantilever tip temperature decreases 20 to 25-fold. As the cantilever thickness decreases from 300 to 50 nm, the force resolution only improves by about 4-fold in the temperature constrained design case due to a reduction in the thermal conductivity of the silicon cantilever. The degraded force resolution of the power constrained design stems from both the reduction in length necessary for water operation and the reduction in power dissipation to maintain a reasonable temperature during operation.

How can the force resolution improve in the future? Even current devices, without any further improvements, operate fairly close to the thermomechanical noise floor. For example, the thermomechanical force noise from 10 Hz to 100 kHz for a 1 mN/m cantilever operating in water ($Q \approx 1/2$) is about 1 pN. Thus, even optical readout is barely feasible for high throughput DNA unzipping. Broadband force sensing in water eventually runs into the thermomechanical force noise wall, and the only ways through it are to reduce the measurement bandwidth or reduce the damping. The latter option can be achieved by continuing to scale down cantilever dimensions or, perhaps, operating the devices on the air-water interface [413]. Sub-micron piezoresistive cantilevers can readily operate at or near the thermomechanical noise floor (depending on the precise design and operating constraints), so there is fairly little benefit in using optical readout or alternative transduction principles/materials. Broadband force sensing eventually runs into the immovable barrier of thermodynamics and the only way forward is a reduction in bandwidth through lock-in techniques. Of course, I would like to be proven wrong.

7.3 Future directions

If I were going to make another generation of fast force probes, here are some of the changes that I would make. First, if I was going to use the same mask set again...

- The DRIE polymer cleanup step could be improved. The O₂/CF₄ process that we used worked well on some wafers (Figure 4.19) but did not remove the polymer from others. It would be worthwhile to more thoroughly investigate the through silicon via (TSV) literature for aqueous cleaning solutions capable of removing the DRIE

passivation polymers. Another option would be minimizing the polymer buildup in the first place. The STS Pegasus tool used for the etches at the University of Michigan deposited a far thicker and more tenacious polymer than the older STS tool used at the SNF. It may be worth accepting more DRIE undercut in exchange for less polymer buildup.

- The use of parylene passivation films should be critically evaluated. Parylene ended up increasing force probe stiffness more than anticipated (compare Figures 5.6 and 5.18) and led to stiction problems with the softest force probe designs (Figure 5.11). In particular, I would recommend revisiting the use of ALD-deposited passivation films. We experimented with ALD passivation films (Figure 5.8) but the films appeared to fail at the Al interconnect sidewalls. Including a sidewall polymer cleanup step after the Al RIE step and increasing the ALD film thickness (from 10-20 nm to 40-50 nm) may solve the problem. In hindsight I think that ALD passivation would result in less probe stiffening, avoid the stiction problem, and could be deposited at the wafer- rather than the device-scale because it can be readily wirebonded through.
- Speaking of wirebonding, Al should be deposited on the Mo bondpads for the piezoelectrically actuated devices. The top and bottom Mo electrodes shorted out far too easily during wirebonding; even with the force, power and time of the wirebonding process optimized, actuator shorting resulted in a substantial drop in device yield. A thick Al bondpad should solve this problem and could be solved using images already included on the reticles that I used.
- All of the wafers received a roughly 400 nm thick LTO coating towards the end of the fabrication process, just before etching the cracks in the exposed oxide. This LTO layer was originally intended to protect the exposed metal from corrosion (Figure 4.17). However, it had an unintended consequence when we switched from releasing the BOX via RIE to vapor HF. Due to the oxide cracks, vapor HF attacks the oxide from both the front and back of the wafer. If the frontside LTO is much thinner than the BOX or does not exist then the Al interconnects on the frontside of the wafer will be undercut significantly. We only used a 400 nm thick LTO buffer layer but

processed wafers with both 400 and 1000 nm thick BOX layers. The latter wafer variety suffered from significant LTO undercut on the frontside, which lifted the Al interconnects off the wafer entirely in some cases. I would recommend depositing a slightly thicker layer of LTO on the wafer than the thickness of the BOX (e.g. 600 nm and 1200 nm) so that the BOX is completely removed by the time that etch starts undercutting the frontside of the wafer.

If I was going to change the device layout without modifying the process flow...

- Increase the size of the piezoresistor contact vias. Any reduction in contact resistance improves the device performance enormously. For thin piezoresistive devices I think that it will always be necessary to use metal interconnects with a heavily doped contact and lightly doped piezoresistor (Figure 4.28), but I would take a fresh look at the size of the contacts to see if they could be increased in size. The width of the contact is particularly important, and the one-dimensional transmission line model should be used for modeling the contacts.
- Use Al interconnects to contact the piezoelectric actuators. The devices I designed run the piezoelectric actuator all the way from the base of the device out to the force probe. This results in poor matching between the main and compensation actuators and much larger capacitances than necessary (e.g. 300 and 900 pF for the compensation and main actuators, respectively). Instead, I would fabricate identical piezoelectric actuators with the metal electrodes only extending 10-20 μm back from the force probe. Then I would coat the actuators with LTO and make electrical contact with them at the same time that the piezoresistor contacts are formed (Mo does not etch in HF). This approach would result in better matching for reduced crosstalk and a roughly 1000-fold increase in actuator bandwidth (not that GHz actuation is necessary for the hair cell application). It would also solve the Mo bondpad problem mentioned a moment ago.
- Finally, I would also take a fresh look at using Lorentz force-based electromagnetic actuators (Section 1.4.2), mainly if you did not want to work with piezoelectric films. In order to fabricate electromagnetic actuators with enough force generation,

the current loop would probably need to extend fairly far along the force probe, which is much easier with thicker devices. Crosstalk from the magnetic field would necessitate placing the main and compensation probes directly next to each other, and differential piezoresistor readout would become all the more important. Electromagnetic actuation combines the fabrication simplicity of thermal actuators with the speed of piezoelectric actuators. I think that there are an equivalent set of challenges in working with piezoelectric and electromagnetic actuators, so it would be worth experimenting with the latter actuator type. See Ref. [50] for a recent report on electromagnetically actuated cantilevers.

I would encourage future researchers who develop high performance force probes to focus on minimizing the steps required after fabrication and before the probes are used for an experiment. The probes presented in this work require epoxying, wirebonding, epoxying, sensitivity calibration, parylene coating and spring constant calibration at a bare minimum. One could improve this process by developing a wafer-scale passivation process (e.g. ALD), an alternative to wirebonding, and integrating spring constant and displacement sensitivity calibration onto the measurement setup itself. At that point, the preparation time per probe would drop from several hours to tens of minutes.

The piezoresistor design methodology and optimization code could be improved in a variety of ways. First, only the vertical component of the dopant concentration profile is currently included in the model. This simplification does not present a problem for most cantilevers (i.e. this work or Ref. [329]) where it is beneficial to etch away all of the silicon surrounding the piezoresistor (Section 2.8.4). However, accounting for lateral dopant diffusion when the diffusion length is comparable to the nominal piezoresistor width would improve the design of devices where the extra silicon can not be removed (e.g. pressure sensors). Second, the design code could be extended to low temperature operation by adding a more detailed analysis of dopant ionization. Third, the Hooge parameter (α) varies up to an order of magnitude between devices from the same wafer (Figure 5.25). Piezoresistor design could benefit from an improved understanding of the variation sources.

The force probes described in this thesis are relatively complex. If I were a biologist studying hair cells who wanted to dramatically simplify the force probe design, then this is what I would think about:

1. The fundamental issue with flexible glass fibers or any other macroscale force probe is that they're simply too large to simultaneously have good force resolution and time resolution. So the most important functions of a MEMS force probe are that it is small, defined lithographically and mass produced.
2. External actuation with a macroscale piezoelectric stack and optimized experimental setup can operate up to 30-40 kHz. While you can go much faster with an on-chip actuator, that degree of actuator speed is not necessary for all applications.
3. One of the sticking points in this project was complexity. The fabrication process took a long time to perfect and has a large number of steps. Devices with on-chip sensing and actuation can be used on any microscope and have a number of benefits, but the electrical connections, passivation and calibration steps add time and complexity. A probe that requires five hours of preparation and calibration is less useful for actual experiments than a probe that takes just five minutes and uses mostly external components.
4. With that in mind, if all electrical connections are eliminated (i.e. no on-chip actuation or sensing) and probe deflection is measured optically, then the force probes could be snapped out of the wafer and mounted on the macroscale actuator within minutes. Depending on the resolution of the optical sensor, the spring constant could be measured in-situ via thermomechanical noise.
5. In this design modality, microfabrication would be used solely to make mechanical probes. The fabrication process would require only two lithography steps (frontside etch and backside etch) and could use polysilicon-on-insulator wafers to cut costs. This force probe approach would work for hair cell mechanics (where force probe spring constant vs. rise time is the limiting factor) and would be an improvement over stiff glass probes for measuring hair cell kinetics.

Interestingly, at this point the force probes would be small AFM cantilevers except that the probe deflection might be measured by projecting the image of the probe directly onto a segmented photodiode rather than bouncing a laser beam off of it. In designing

a high performance force probe, the fundamental question to consider is whether the complexity should go in the probe (i.e. piezoresistive sensing, piezoelectric actuation, wirebonding and passivation) or in the experimental setup (i.e. a high-speed photodiode detector and optimized macroscale actuation). The same essential tradeoff is made in commercial atomic force microscopes, where the complexity is in the instrument and the silicon cantilevers are inexpensive and mass produced. Integrating actuation and sensing onto the microcantilever yields large performance improvements, but is not necessary for all applications.

There is no such thing as a free lunch and the complexity must go somewhere. If you choose to place the complexity in the force probe, then I hope that this thesis will prove useful.

Appendix A

Glossary of mathematical symbols

This Appendix summarizes the mathematical symbols that are used throughout the thesis. Several symbols have multiple meanings; for example, α is used to represent the temperature coefficient of resistance, the coefficient of thermal expansion, and the Hooge factor. We have tried to separate these overloaded symbols in the text and make their meaning clear whenever they are used. Symbols that are only used once or twice (e.g. the Tsai model for phosphorus predeposition or the variation in bandgap energy with temperature) are not included in the table for succinctness.

Table A.1: Glossary of mathematical symbols.

Symbol	Units	Description
α	-	Hooge factor
α	ppm/K	Coefficient of thermal expansion
α	ppm/K	Temperature coefficient of resistance
β^*	-	Sensitivity factor
β_1^*	-	Sensitivity factor (piezoresistance factor effect)
β_2^*	-	Sensitivity factor (depth effect)
ϵ	F/m	Electrical permittivity
ϵ	-	Mechanical strain
η	Pa-s	Fluid viscosity
γ	-	Resistance factor

Table A.1: Glossary of mathematical symbols (continued).

Symbol	Units	Description
Λ	m	Thermal healing length
μ	cm ² /V-sec	Carrier mobility
ω	rad/s	Angular frequency
π	1/Pa	Piezoresistive coefficient
ρ	$\Omega\text{-m}$	Electrical resistivity
ρ	kg/m ³	Density
ρ_a	kg/m ³	Actuator density
ρ_c	kg/m ³	Cantilever density
ρ_f	kg/m ³	Fluid density
ρ_p	kg/m ³	Passivation layer density
σ	Pa	Mechanical stress
σ	S/m	Electrical conductivity
τ	s	Time constant
A	m ²	Cross-sectional or film area
A_{VJ}	V/ $\sqrt{\text{Hz}}$	Amplifier Johnson voltage noise coefficient
A_{VF}	V	Amplifier 1/f voltage noise coefficient
A_{IJ}	A/ $\sqrt{\text{Hz}}$	Amplifier Johnson current noise coefficient
A_{IF}	A	Amplifier 1/f current noise coefficient
A_{XK}	m/K	Thermal actuator temperature responsivity
A_{XW}	m/W	Thermal actuator power responsivity
A_{XV}	m/V	Piezoelectric actuator responsivity
B	T	Magnetic field
b	kg/s	Drag coefficient
c	J/kg-K	Specific heat capacity
C	1/m	Cantilever beam curvature
C	Pa	Elastic stiffness tensor
C	F	Electrical capacitance
C	J/K	Thermal capacitance

Table A.1: Glossary of mathematical symbols (continued).

Symbol	Units	Description
C_a	J/K	Actuator thermal capacitance
D	m ² /s	Diffusivity
d_{31}	pm/V	Transverse piezoelectric coefficient
\sqrt{Dt}	m	Dopant diffusion length
E_a	Pa	Actuator elastic modulus
E_c	Pa	Cantilever elastic modulus
E_F	eV	Fermi energy level
E_g	eV	Bandgap energy
E_p	Pa	Passivation layer elastic modulus
F_{TMN}	N/ $\sqrt{\text{Hz}}$	Thermomechanical force noise
$f_{-3\text{dB}}$	Hz	-3 dB frequency
f_0	Hz	Undamped first resonant mode frequency
f_d	Hz	Damped first resonant mode frequency
f_{\min}	Hz	Minimum measurement frequency
f_{\max}	Hz	Maximum measurement frequency
G	-	Amplifier gain
G	m/N	Simple harmonic oscillator frequency response
G	W/K	Thermal conductance
G_{pr}	W/K	Piezoresistor thermal conductance
G_{base}	W/K	Cantilever base thermal conductance
G'_f	W/m-K	Structure-fluid thermal conductance per unit length
h_{eff}	W/m ² -K	Effective convection coefficient
h	J-sec	Planck's constant
I	m ⁴	Second moment of area
I	A	Electrical current
J	A/m ²	Electrical current density
k	W/m-K	Thermal conductivity
k	N/m	Spring constant

Table A.1: Glossary of mathematical symbols (continued).

Symbol	Units	Description
k_b	J/K	Boltzmann's constant
k_c	N/m	Force probe spring constant
k_f	W/m-K	Thermal conductivity (fluid)
k_s	W/m-K	Thermal conductivity (solid)
L	H	Inductance
l_a	m	Actuator length
l_c	m	Cantilever length
l_h	m	Thermal actuator heater length
l_{pr}	m	Piezoresistor length
l_{pe}	m	Piezoelectric actuator length
m	kg	Mass
m	A-m ²	Magnetic moment
m_{eff}	kg	Effective mass, scaled to account for eigenmode shape
MDD	m	Minimum detectable displacement
MDF	N	Minimum detectable force
N	cm ⁻³	Dopant concentration
N_a	cm ⁻³	Acceptor dopant concentration
N_B	cm ⁻³	Background dopant concentration
N_d	cm ⁻³	Donor dopant concentration
N_{epi}	cm ⁻³	Epitaxial layer dopant concentration
n	cm ⁻³	Carrier density
n_i	cm ⁻³	Intrinsic carrier density
N_{eff}	-	Effective number of carriers
N_{pr}	-	Number of piezoresistors in a Wheatstone bridge
N_z	μm ⁻²	Effective carrier density per unit area
P	m	Perimeter
P	-	Piezoresistance factor
\bar{P}	-	Average piezoresistance factor

Table A.1: Glossary of mathematical symbols (continued).

Symbol	Units	Description
q	C	Single electron charge
Q	C	Electrical polarization charge
Q	-	Quality factor
R	Ω	Total electrical resistance
R	K/W	Thermal resistance
R_a	K/W	Actuator thermal resistance
R_{base}	K/W	Cantilever base thermal resistance
R_{contact}	Ω	Contact resistance
R_{excess}	Ω	Excess electrical resistance
R_{pr}	Ω	Piezoresistor electrical resistance
R_s	Ω/\square	Sheet resistance
S	1/Pa	Elastic compliance tensor
S_{XV}	V/m	Displacement sensitivity
S_{FV}	V/N	Force sensitivity
S_A	V^2/Hz	Amplifier noise power spectral density
S_H	V^2/Hz	Hooge (1/f) noise power spectral density
S_J	V^2/Hz	Johnson noise power spectral density
S_{TMN}	V^2/Hz	Thermomechanical noise power spectral density
T	K	Temperature
T_∞	K	Ambient temperature
TCTD	m/K	Temperature coefficient of tip deflection
T_{\max}	K	Maximum cantilever temperature
T_{pr}	K	Average piezoresistor temperature
T_{tip}	K	Cantilever tip temperature
T_n	-	Temperature normalized to 300 K
t	s	Time
t_a	m	Actuator thickness
t_c	m	Cantilever thickness

Table A.1: Glossary of mathematical symbols (continued).

Symbol	Units	Description
t_j	m	Junction depth
t_p	m	Passivation layer thickness
t_{pr}	m	Epitaxial piezoresistor thickness
t_r	s	Rise time (10-90%)
v	m	Transverse beam deflection
V_b	V	Resistor bias voltage
V_{bridge}	V	Bridge bias voltage
v_{tip}	m	Cantilever beam tip deflection
V_A	V/ $\sqrt{\text{Hz}}$	Amplifier noise spectral density
V_H	V/ $\sqrt{\text{Hz}}$	Hooge (1/f) noise spectral density
V_J	V/ $\sqrt{\text{Hz}}$	Johnson noise spectral density
V_{noise}	V/ $\sqrt{\text{Hz}}$	Overall noise spectral density
V_{TMN}	V/ $\sqrt{\text{Hz}}$	Thermomechanical noise spectral density
w_a	m	Actuator width
w_c	m	Cantilever width
w_{pr}	m	Piezoresistor width
W	W	Electrical power dissipation
z_n	m	Cantilever beam neutral axis
z_0	m	Static tip deflection

Appendix B

List of publications

This Appendix lists the publications that I contributed to while at Stanford. Several of them were not discussed in this work, in particular: SU-8 force posts for biological force measurements, a detailed comparison between piezoresistive and piezoelectric sensing, and the contact mechanics of rough surfaces during adhesive elastic contact. The common theme between all of them is the measurement of microscale mechanical phenomena.

Journal publications

1. J.C. Doll and B.L. Pruitt, “High bandwidth piezoresistive force probes with integrated thermal actuation”, *Journal of Micromechanics and Microengineering*, July 2012.
2. S.L. Geffeney, J.G. Cueva, D.A. Glauser, J.C. Doll, T.H.-C. Lee, M. Montoya, S. Karania, A. Garakani, B.L. Pruitt, M.B. Goodman, “DEG/ENaC but not TRP Channels are the major mechanoelectrical transduction channels in a *C. elegans* nociceptor”, *Neuron*, September 2011.
3. G.C. Hill, J.I. Padovani, J.C. Doll, B.W. Chui, D. Rugar, H.J. Mamin, N. Harjee, B.L. Pruitt, “Patterned cracks improve yield in the release of compliant microdevices from silicon-on-insulator wafers”, *Journal of Micromechanics and Microengineering*, June 2011.

4. J.C. Doll, E.A. Corbin, W.P. King, and B.L. Pruitt, “Self-heating in piezoresistive cantilevers”, *Applied Physics Letters*, May 2011.
5. H. Kesari, J.C. Doll, B.L. Pruitt, W. Cai, A.J. Lew, “Role of surface roughness in hysteresis during adhesive elastic contact”, *Philosophical Magazine Letters*, September 2010.
6. J.C. Doll and B.L. Pruitt, “Design of Piezoresistive vs. Piezoelectric Contact Mode Scanning Probes”, *Journal of Micromechanics and Microengineering*, August 2010.
7. J.C. Doll, B.C. Petzold, B. Ninan, R. Mullapudi, B.L. Pruitt, “Aluminum Nitride on Titanium for CMOS Compatible Piezoelectric Transducers”, *Journal of Micromechanics and Microengineering*, February 2010.
8. S.-J. Park, J.C. Doll, A.J. Rastegar, and B.L. Pruitt, “Piezoresistive cantilever performance, part II: optimization,” *Journal of Microelectromechanical Systems*, February 2010.
9. S.-J. Park, J.C. Doll, and B.L. Pruitt, “Piezoresistive cantilever performance, part I: analytical model for sensitivity”, *Journal of Microelectromechanical Systems*, February 2010.
10. J.C. Doll, S.-J. Park, and B.L. Pruitt, “Design optimization of piezoresistive cantilevers for force sensing in air and water”, *Journal of Applied Physics*, September 2009.
11. J.C. Doll, N. Harjee, N. Klejwa, R. Kwon, S.M. Coulthard, B.C. Petzold, M.B. Goodman, and B.L. Pruitt, “SU-8 Force Sensing Pillar Arrays for Biological Measurements”, *Lab on a Chip*, May 2009.
12. S.R. Lockery, K.J. Lawton, J.C. Doll, S. Faumont, S.M. Coulthard, T.R. Thiele, N. Chronis, K.E. McCormick, M.B. Goodman, and B.L. Pruitt, “Artificial dirt: Microfluidic substrates for nematode neurobiology and behavior”, *Journal of Neurophysiology*, June 2008.

Conference publications

1. J.C. Doll and B.L. Pruitt, “MEMS force probes for cell mechanobiology at the microsecond scale”, Hilton Head Workshop, June 2012.
2. J.C. Doll, S.-J. Park, N. Harjee, B.C. Petzold, and B.L. Pruitt, “Piezoresistive Cantilevers for Biological Force Measurements”, Proceedings of Nanomechanical Cantilever Workshop, June 2012.
3. J.C. Doll, A. Peng, A. Ricci, and B.L. Pruitt, “New devices for investigating hair cell mechanical properties”, Mechanics of Hearing workshop, July 2011.
4. G.C. Hill, J.I. Padovani, B.W. Chui, H.J. Mamin, D. Rugar, N. Harjee, J.C. Doll, and B.L. Pruitt, “Patterned Cracks in the Buried Oxide Layer Improve Yield in Device Release from SOI Wafers”, Proceedings of IEEE MEMS, January 2011.
5. S.-J. Park, J.C. Doll, N. Harjee, B.L. Pruitt, “Optimization with Process Limits and Application Requirements for Force Sensors”, Proceedings of IEEE Sensors, November 2010.
6. J.C. Doll and B.L. Pruitt, “Design of Piezoresistive vs Piezoelectric Contact Mode Scanning Probes”, Hilton Head Workshop, June 2010.
7. J.C. Doll, B.C. Petzold, M.B. Goodman, B.L. Pruitt, “Piezoresistive cantilevers optimized for kilohertz force sensing in aqueous solutions”, Nanomechanical Cantilever Workshop, May 2010.
8. N. Harjee, A.G.F. Garcia, M. Konig, J.C. Doll, D. Goldhaber-Gordon, and B.L. Pruitt, “Coaxial tip piezoresistive scanning probes for high-resolution electrical imaging”, Proceedings of IEEE MEMS, January 2010.
9. J.C. Doll, S.-J. Park, N. Harjee, A.J. Rastegar, J.R. Mallon, B.C. Petzold, G.C. Hill, A.A. Barlian, B.L. Pruitt, “Piezoresistive Cantilever Optimization and Applications”, Materials Research Society Proceedings, December 2009.

10. J.C. Doll, B.C. Petzold, B. Ninan, R. Mullapudi, B.L. Pruitt, “A High d_{33} CMOS Compatible Process for Aluminum Nitride on Titanium”, Proceedings of Transducers, June 2009.
11. J.C. Doll, B.C. Petzold, P. Ghale, M.B. Goodman, B.L. Pruitt, “High Frequency Force Sensing with Piezoresistive Cantilevers”, Proceedings of Transducers, June 2009.
12. J.C. Doll, S. Muntwyler, F. Beyeler, S. Geffeney, M.B. Goodman, B.J. Nelson, and B.L. Pruitt, “Measuring Thresholds for Touch Sensation in *C. elegans*”, Proceedings of MEMS in Medicine and Biology, April 2009.
13. J.C. Doll, N. Harjee, N. Klejwa, R. Kwon, S.M. Coulthard, and B.L. Pruitt, “Biological measurements of *C. elegans* touch sensitivity with microfabricated force sensors”, Proceedings of MicroTAS, October 2007.

Book chapters

1. J.C. Doll, S.-J. Park, A.J. Rastegar, N. Harjee, J.R. Mallon Jr., G. Hill, A.A. Barlian and B.L. Pruitt, “Force sensing optimization and applications”, book chapter in Advanced Materials and Technologies for Micro/Nano-Devices, Sensors and Actuators, Springer, 2010.

Appendix C

Circuit details

This appendix presents several important circuit implementation details. The schematic (Figure C.1) and layout (Figure C.2) of the final amplifier board as well as the layout of the final four-layer cantilever board (Figure C.3) that I developed are presented.

First, a quick general note on the circuits in general. The sensing portion of the circuit could be implemented on a breadboard rather than a PCB without much of an impact on performance (extra capacitive and inductive parasitics probably wouldn't change the operation below 100 kHz). A PCB simply makes the system more robust and reliable, and in my experience, is well worth the time and cost. However, the crosstalk compensation portions of the circuit could not be reliably implemented on a breadboard. Variation in capacitances of a few hundred fF leads to a large change in capacitive crosstalk.

The boards were designed in Eagle and crosstalk simulations were performed in LTSpice IV. The boards were fabricated at Sierra Proto Express (Sunnyvale, CA) and Imagineering Inc. (Elk Grove Villtage, IL). Sierra enabled smaller vias (10 mils) and trace size/spacing (5 mils) which was essential for the layout of the four-layer cantilever board so that the ground planes wrapped around the bondpads (Figure C.3). Imagineering was less expensive for large boards that combined amplifier and cantilever boards using route/retain. Both vendors manufactured extremely high quality boards and I would recommend either. Power supply decoupling capacitors were chosen based upon the specification sheet of each component on the board (mainly 0.1 μF ceramic capacitors with 10 μF tantalum capacitors for the power supply cable and instrumentation amplifier).

The amplifier circuit (Figures C.1 and C.2) consists of sensor readout, actuator driver and crosstalk compensation portions. The Wheatstone bridge bias is generated by a +5V voltage reference (ADR425), trimmed down to the desired voltage with a potentiometer and buffered (AD8671). The Wheatstone bridge consists of potentiometers on each of the top two legs with the main and compensation probes on the bottom two legs. The main probe is arbitrarily connected to the positive amplifier input.

A pair of 1 k Ω resistors are included on the board as a voltage divider. The divider output can be connected to the negative amplifier input instead of the potentiometer/compensation resistors using a switch (AYZ0102). This is useful in case the compensation probe breaks (unlikely) and it makes balancing the bridge easier (i.e. switch to using the reference resistors and balance the positive side of the bridge, then switch back to using the compensation probe and finish balancing the bridge instead of using a multimeter to monitor the amplifier input voltages).

The main and compensation probes are put on opposite sides of the Wheatstone bridge so that any common-mode crosstalk from the actuator is rejected by the instrumentation amplifier. The downside with this approach is that two potentiometers are needed (a poor balancing job could be done with a single potentiometer, though, if absolutely needed). Alternatively, one could imagine putting both the main and compensation probes on one side of the bridge with a fixed resistor and single potentiometer on the other side of the bridge. This is the approach that I utilized for the first-generation probes and circuit board, and it performs extremely poorly.

The potentiometers at the top of the Wheatstone bridge actually consist of two wirewound potentiometers connected in series on each side of the bridge (Figure C.4). One potentiometer on each side is used for coarse adjustment (typically 20 k Ω) while the second potentiometer provides fine adjustment (typically 1 k Ω). You should only use wirewound potentiometers; cermet potentiometers can yield excess 1/f noise that confounds your noise experiments. Both through-hole and off-board wirewound potentiometers can be procured. I ended up using off-board potentiometers because they are easier to adjust (via fingers rather than a screwdriver).

The amplifier and cantilever circuits are connected using 10-position low profile headers (B10B-ZR, JST). The matching cable housings are attached to standard 8-wire

shielded RJ45 cables (Figure C.5). The exposed wire is coated with hot glue in order to make it sturdier. The spare connections on each end of the header are connect to the amplifier circuit ground and to the cantilever circuit ground planes. This ensures that both the RJ45 shield and cantilever circuit are grounded with the formation of any ground loops and works extremely well in practice. The power cable (Figure C.6) is assembled using standard Molex headers and cable housings. The substrate connection can either be grounded (using a jumper) or arbitrarily set (using an SMA connection).

The signals from each side of the Wheatstone bridge are connected directly to the instrumentation amplifier (INA103) inputs. The actuator drive signals are capacitively coupled to the instrumentation amplifier inputs using a fixed capacitor and variable capacitor. I tested several different surface mount variable capacitors and found that Murata (e.g. TZB4Z030AA10R00 or 490-1984-1-ND from Digikey) varcaps were the easiest to work with. Optional series resistors and capacitors to ground are included on the board before the capacitive coupling in order to better emulate the on-device crosstalk mechanisms. In practice these additional components made a negligible difference on crosstalk performance.

It is extremely important to adjust the variable capacitor using an insulated screwdriver (e.g. Murata KMDR010) so that you do not affect the capacitances while adjusting the variable capacitor; otherwise you will have an extremely difficult time canceling the < 1 pF of capacitive mismatch between the two instrumentation amplifier inputs.

The output from the instrumentation amplifier is amplified using an inverting op-amp with the feedback resistor selected using a three-position switch (AYZ0103). A second gain stage ensures that there is sufficient bandwidth for any conceivable measurement (e.g. 800 kHz for the INA103 with G=100) and can be convenient for measuring either very small or large forces if the data acquisition system does not incorporate variable gain. The output signal from the variable gain amplifier (VGA) is connected to an SMA connector. If a long cable is used (i.e. a large capacitive load), then a small snubber resistance can be added in series with the cable to eliminate oscillations. This is a common problem and is due to the fact that the output impedance of op-amps increase with frequency so that, combined with the capacitive load of the cable, the amplifier sees an RLC resonator. Adding the snubber resistor damps out the oscillation. A high-speed, low-noise op-amp with sufficient

bandwidth (e.g. THS4031 or ADA4898-1) is used for the VGA.

The final portion of the circuit is the actuator driver. The input signal is buffered with an op-amp (ADA4898-1 or AD8671) to increase the current drive capability. Many data acquisition boards only have an output drive capability of 1-2 mA, which can limit the actuator slew rate and range. The buffer input is connected to ground with a large resistor (e.g. 100 k Ω) so that the input bias current (e.g. 1-2 pA) does not saturate the buffer amplifier when a cable is not connected. The buffer drives an inverting differential op-amp (THS4131) which converts the single-ended signal into a differential one. A switch (JS202) selects between the single-ended and differential drive options. In both cases, a 20 Ω snubber resistance is added in series with the actuator drive signal in order to prevent oscillations. Additional resistance can be added to the actuator drive outputs (e.g. a 50 Ω fixed resistance for the main probe actuator and a potentiometer for the compensation probe actuator) for crosstalk compensation. This excess resistance can either be used to match the time constants of the two actuators (piezoelectric actuation) or trim any mismatch between the silicon heater on each probe (thermal actuation) and is fairly important for the overall crosstalk performance of the system.

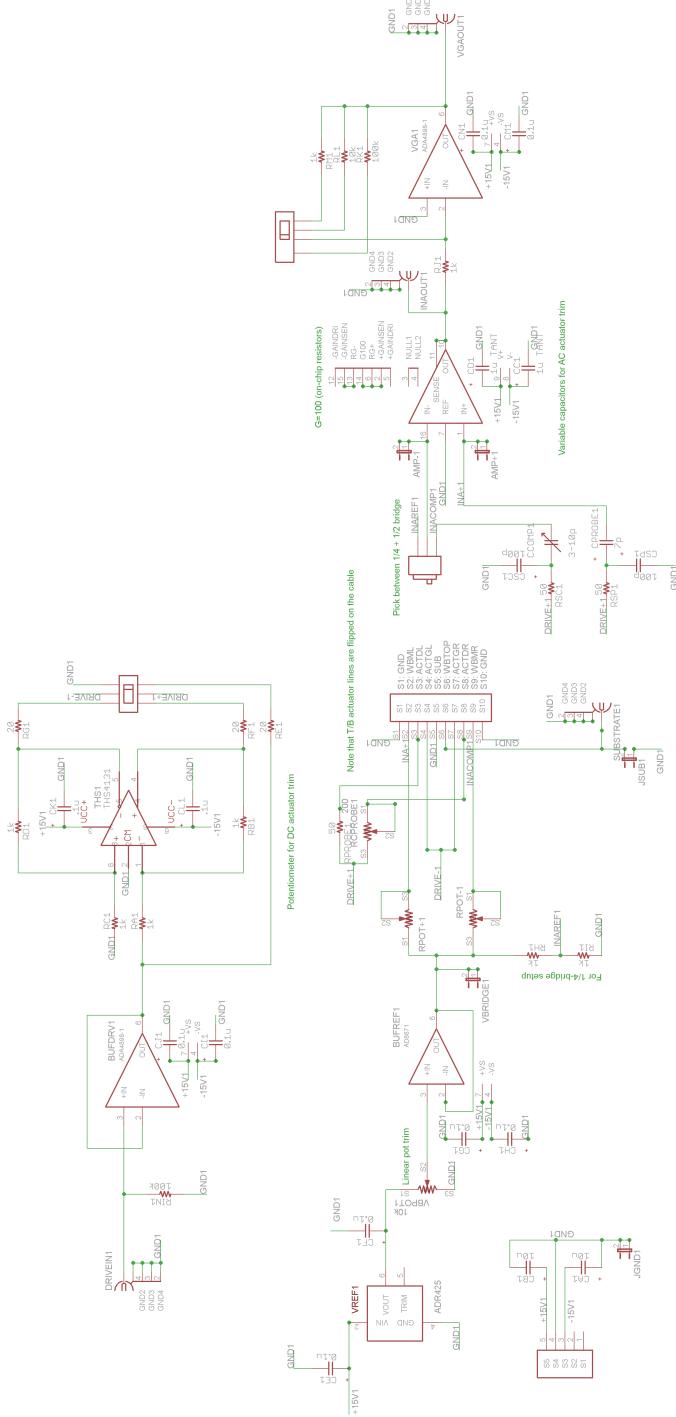


Figure C.1: Amplifier and crosstalk compensation board schematic. Circuit details are discussed in the text of the appendix.

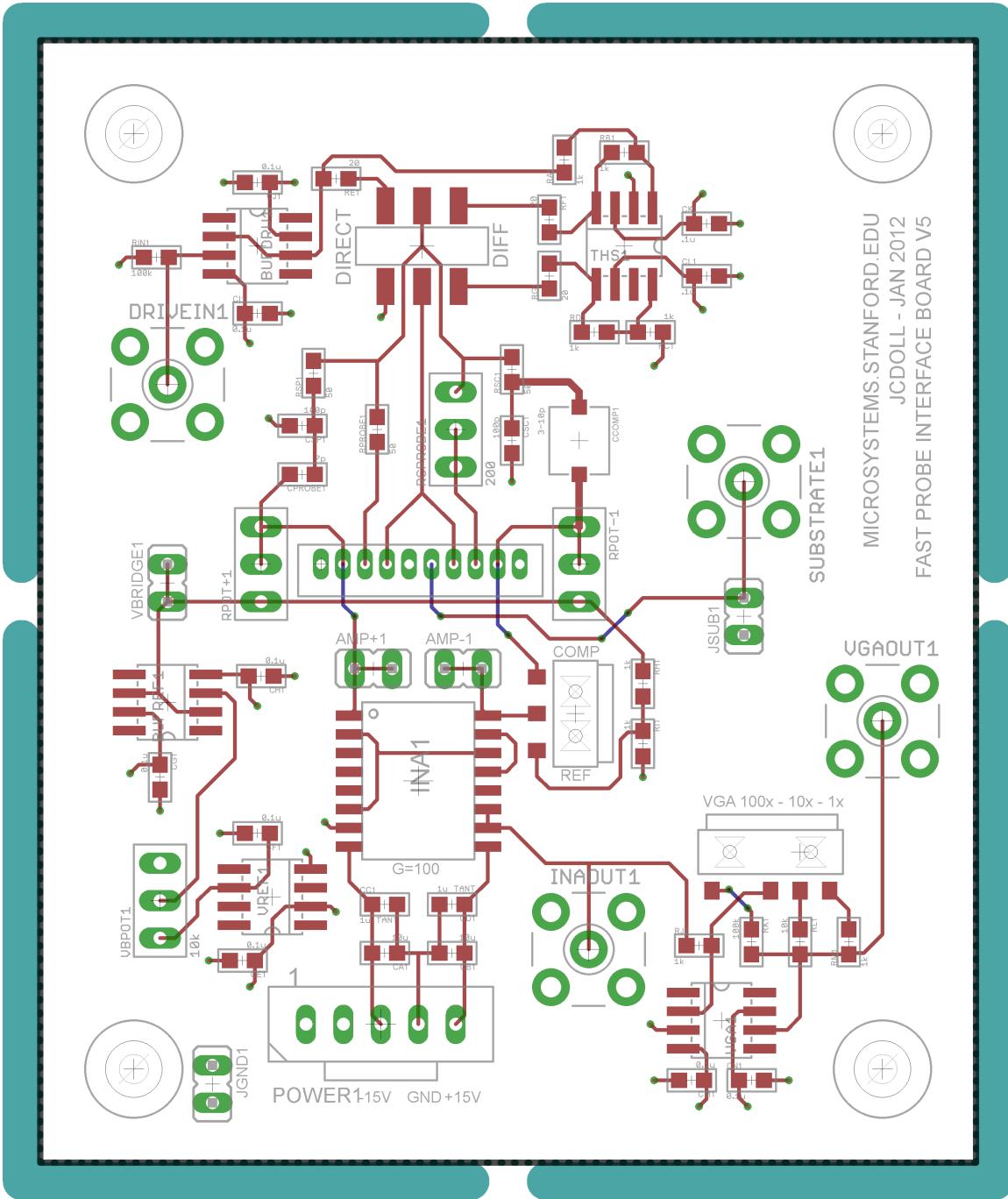


Figure C.2: Amplifier and crosstalk compensation board layout. The ground and power planes are omitted for clarity. Both two- and four-layer boards were designed and assembled. The board is routed and retained on the panel received back from the PCB vendor.

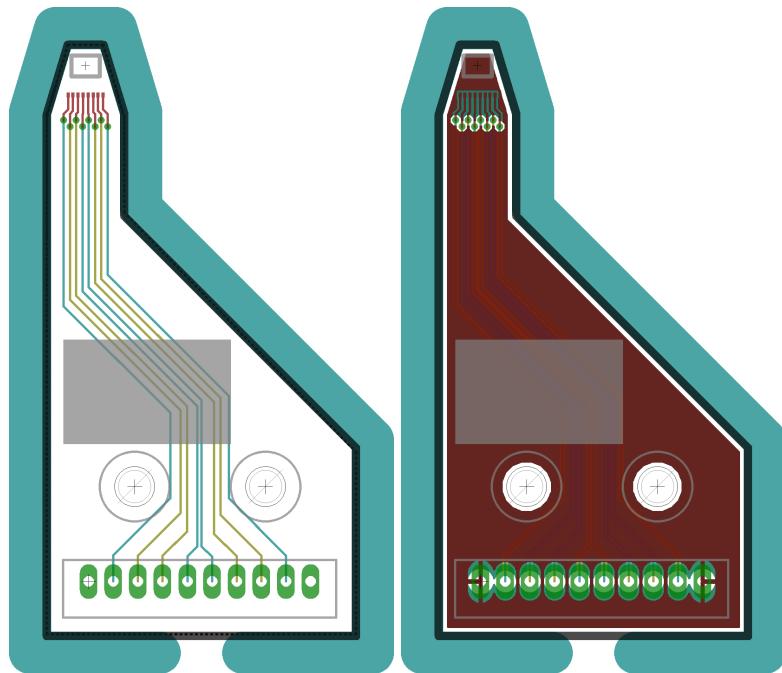


Figure C.3: Four-layer cantilever board layout. The board is shown without (left) and with (right) the ground planes shown. The signal traces are run on the interior two layers of the board. Ground planes cover the top and bottom layers of the board and run along the outside of the signal layers. The sensor (teal) and actuator (yellow) signals are run on separate planes and run symmetrically so that any capacitive crosstalk is balanced at the amplifier inputs. A silkscreen block on the frontside of the board is used to label each device and a pair of 2-56 through holes are used for board mounting. It is important that the tip of the board is as narrow as possible so that it does not interfere with operation on an upright microscope. The board layout is defined using a 100 mil routing bit with a 10 mil gap between the copper and the edge of the board.

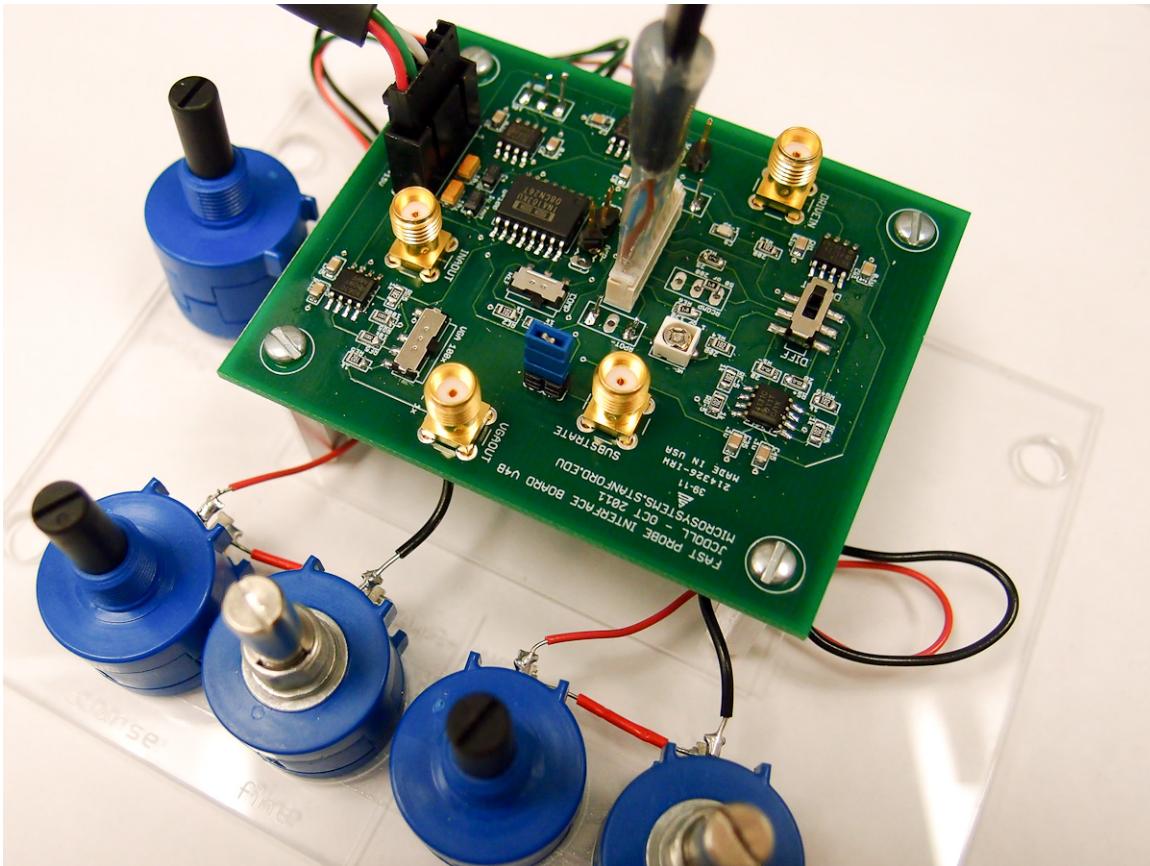


Figure C.4: Example of an assembled amplifier circuit. A custom base plate is laser cut from acrylic and text labels are engraved into it for user friendliness. The Wheatstone bridge bias and balance are adjusted using the five potentiometers shown. The amplifier board is connected to the ± 15 V power supply and cantilever board using a pair of cables (shown). The circuit inputs and outputs are accessed using SMA-BNC cables (not shown).

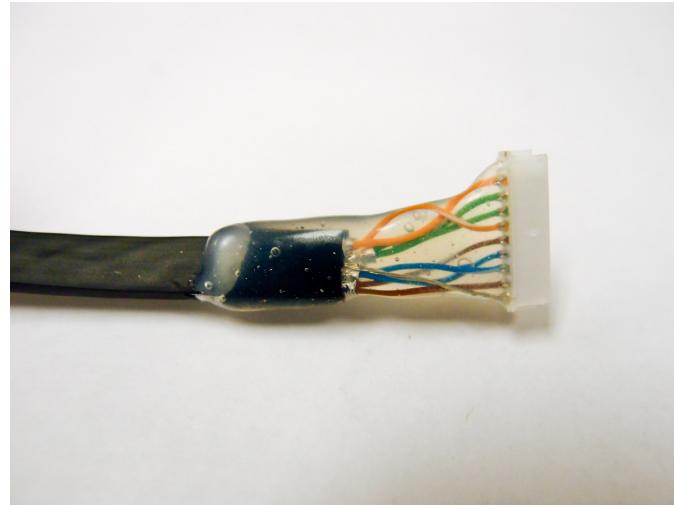


Figure C.5: Assembled cantilever cable. A standard shielded 8-wire RJ45 cable connects the amplifier and cantilever boards. The cable is attached to a 10-position housing and the exposed wires are coated with hot glue. The cable shield (bottom wire in the photo) is connected to one of the edge positions of the cable housing on both sides of the cable. This grounds the cable shield as well as the ground planes on the cantilever board that the force probe is mounted on without ground loop formation.

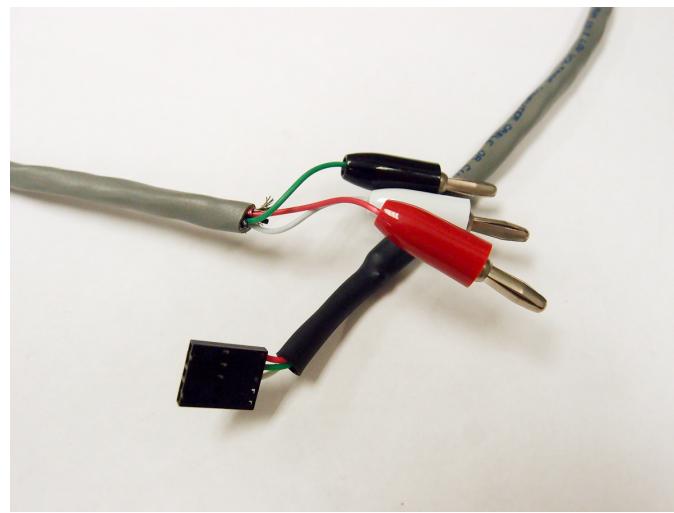


Figure C.6: Example of an assembled power cable. The amplifier circuit board is connected to a power supply using a hand-assembled power cable with Molex headers and cable housings.

Appendix D

Reticle and wafer layouts

The layouts of the reticles and wafers for each of the fabrication runs are illustrated in this section. The fast-pr-* reticles were used for the first fabrication run (Appendix F) while the fast-probe-* reticles were used for the final run presented in the body of the thesis. The fast-probe-* reticles include the boxes that denote the size of each image including the 1.2 mm gap between images. All of the reticles are drawn data clear; the shaded areas correspond to regions where chrome should be removed from the reticle.

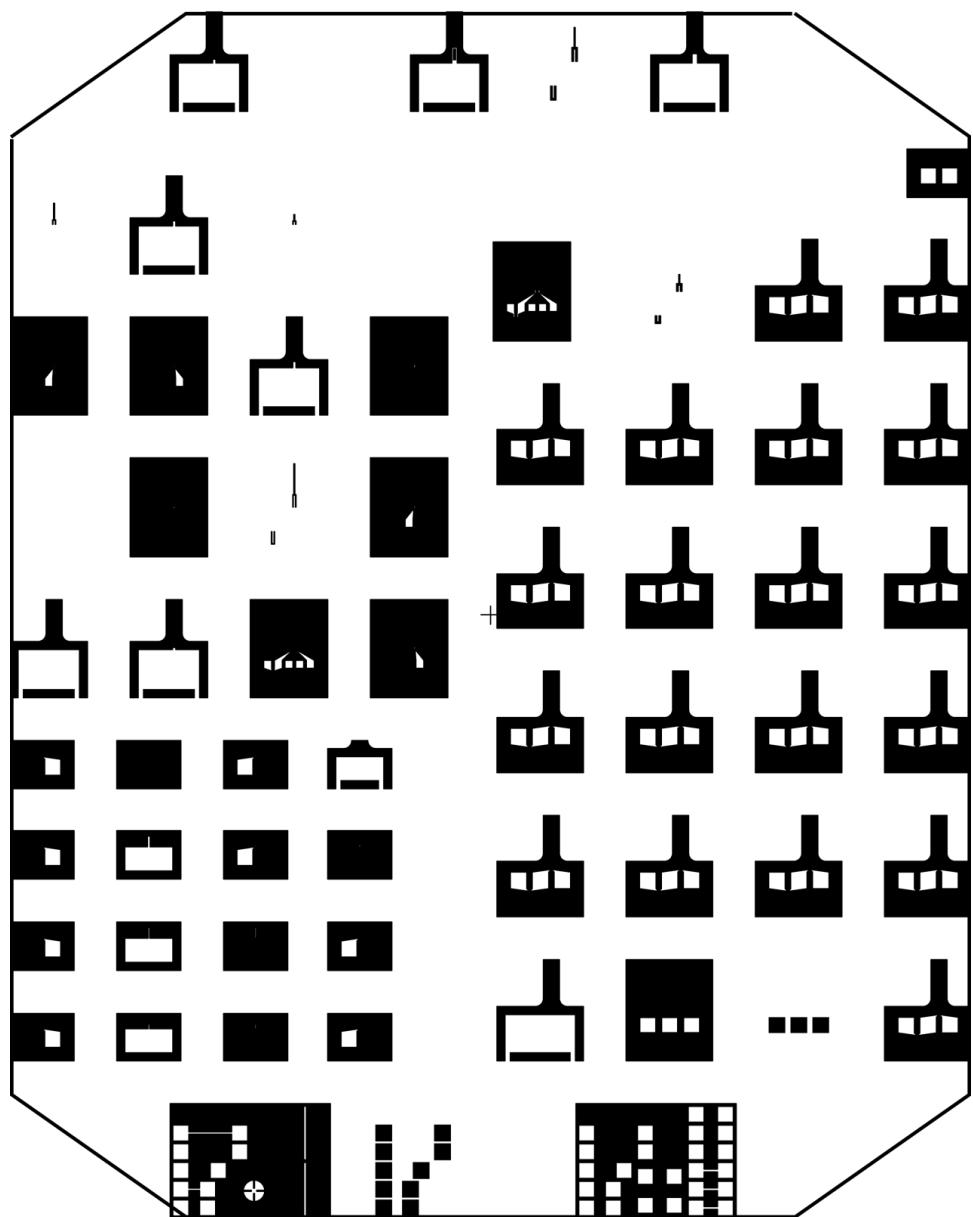


Figure D.1: Reticle fast-pr-1. The first reticle used in the 2008 fabrication run. The images used for the test structure die are near the bottom of the reticle while the 19 nearly identical images near the right side of the reticle were used to define the various piezoresistor and cantilever designs.

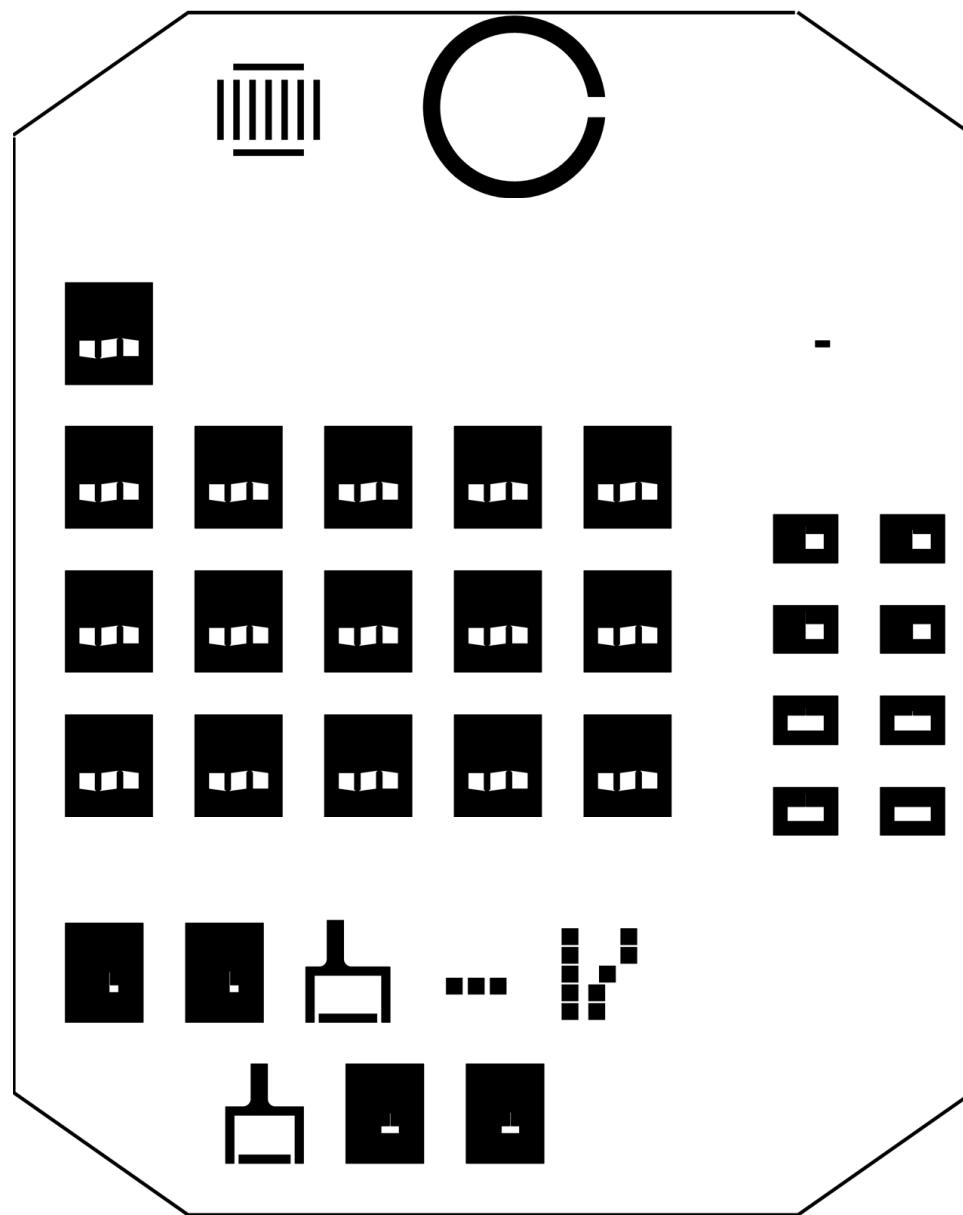


Figure D.2: Reticle fast-pr-2. The second reticle used in the 2008 fabrication run. The semicircle and trenches near the top of the reticle were used for a side project related to positioning *C. elegans* works for cross-sectional cryogenic TEM imaging.

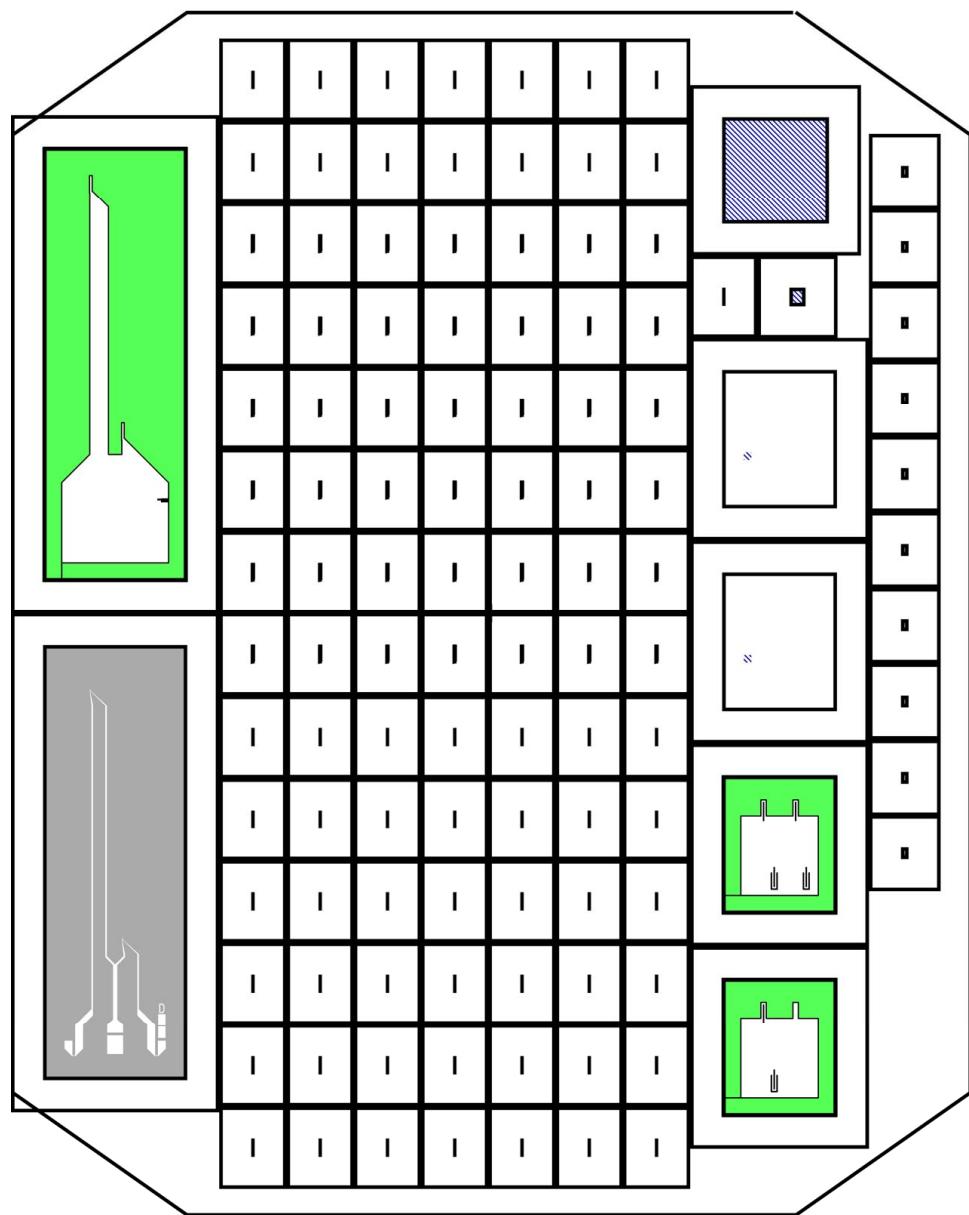


Figure D.3: Reticle fast-probe-1. The first reticle used in the 2010 and 2011 fabrication runs. The black bounding boxes denote the boundary of each image and the minimum spacing required between images and are not included on the reticle.

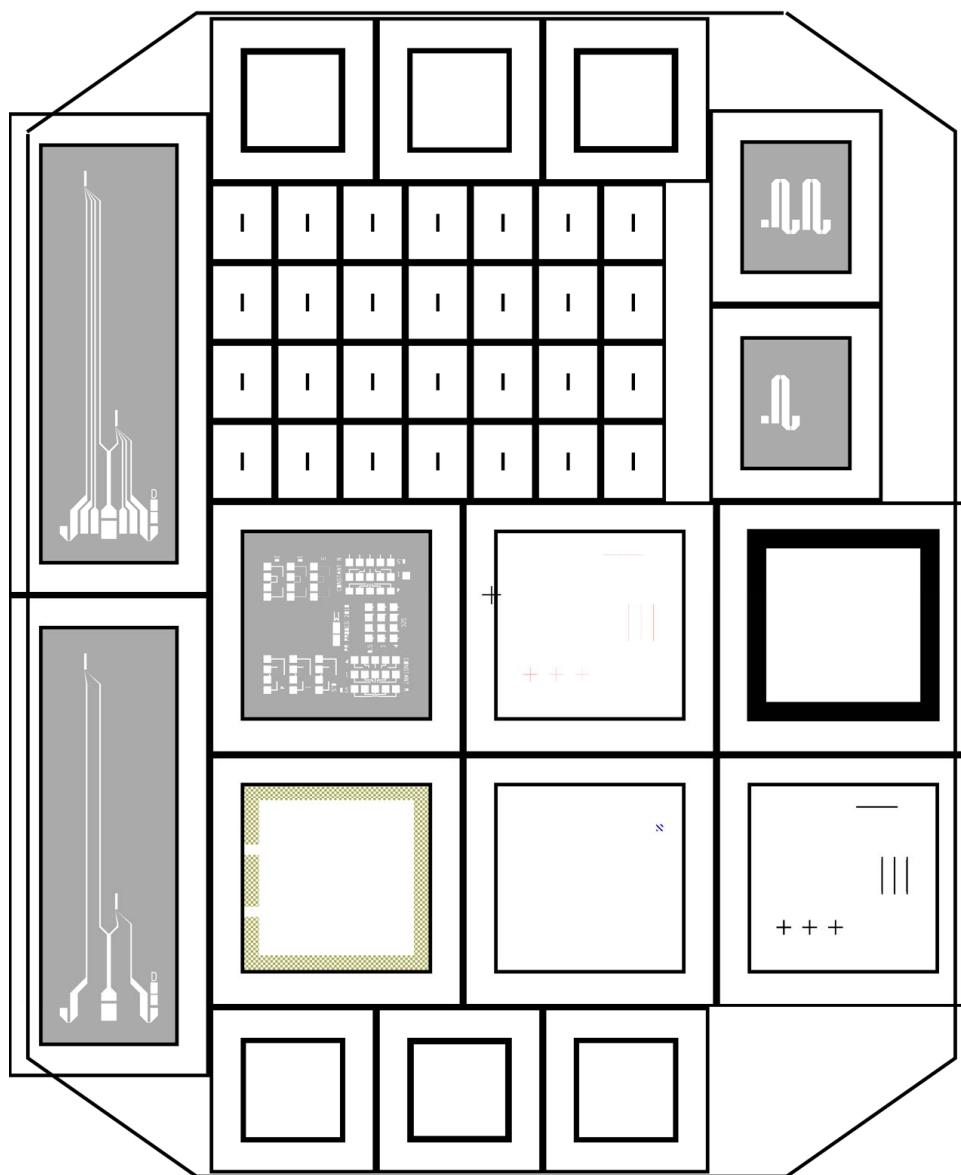


Figure D.4: Reticle fast-probe-2. The second reticle used in the 2010 and 2011 fabrication runs.

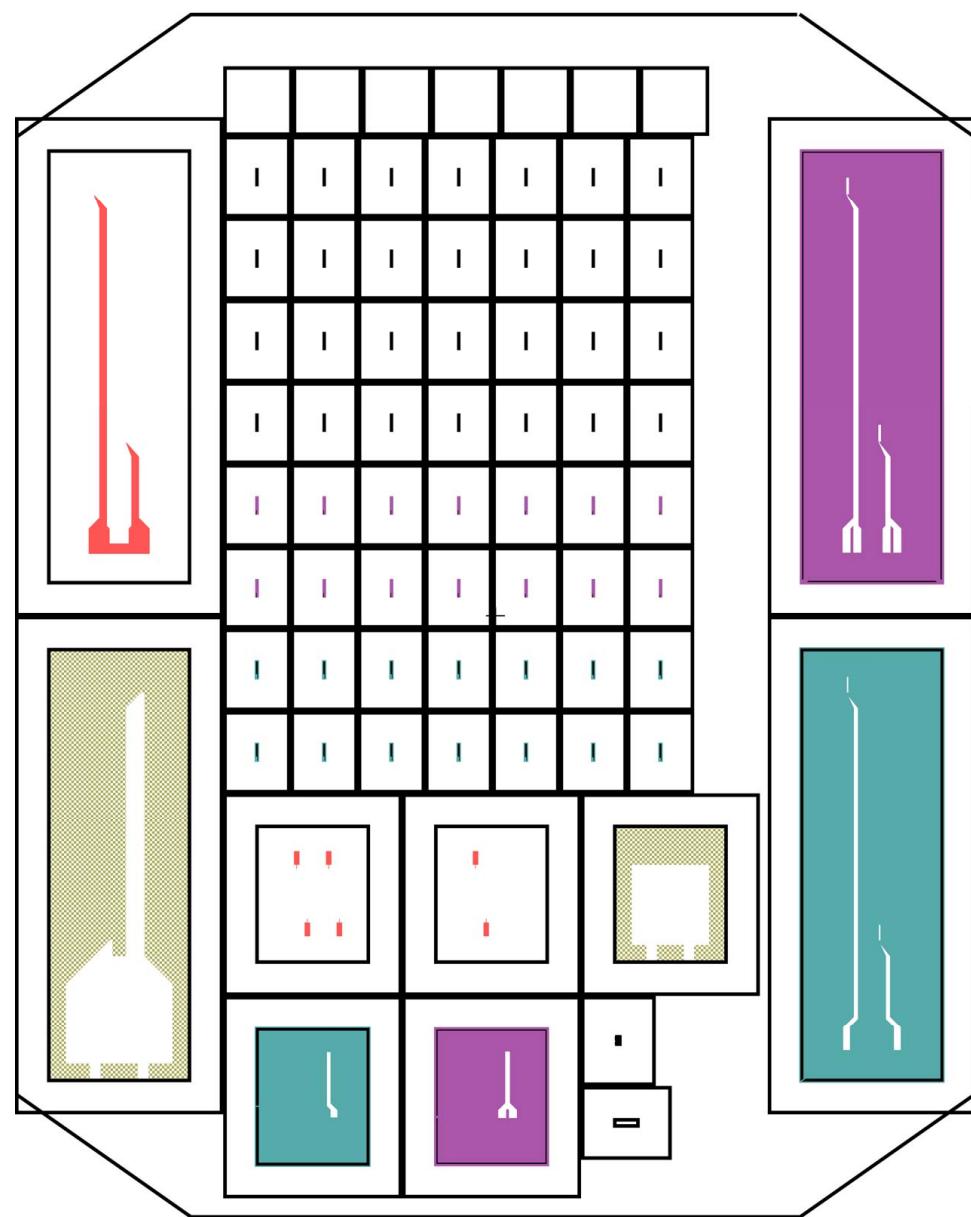


Figure D.5: Reticle fast-probe-3. The third reticle used in the 2010 and 2011 fabrication runs.

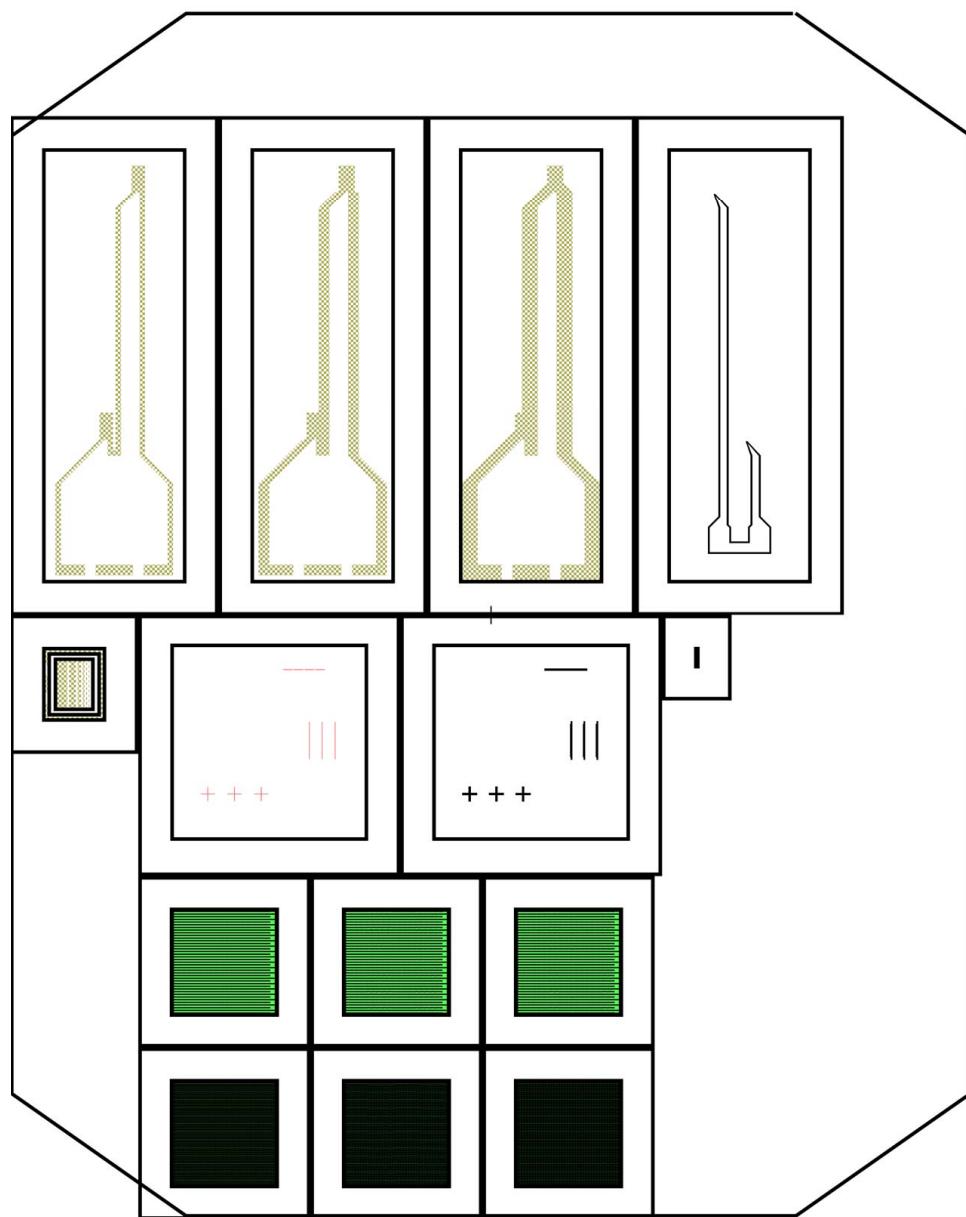


Figure D.6: Reticle fast-probe-4. The fourth reticle used in the 2010 and 2011 fabrication runs.

Appendix E

Fabrication process details

This section includes process details that should enable someone to reproduce the force probes described in this thesis. All process details are explained in

E.1 Wafer handling

The wafers are extremely fragile once the backside etch is complete. Precautions need to be taken during wafer handling. There are a few very specific handling issues that I want to mention.

First, gasonics (a downstream oxygen plasma ash) can be used on released wafers. However, the vacuum chuck may break the devices or, more likely, fail to recognize the wafer due to the lack of a vacuum seal. This problem can be resolved by loading a dummy wafer, pausing the tool once the wafer is on the main robot arm, replacing it with the device wafer, and restarting the tool. Once the oxygen plasma step is complete, the reverse process should be used.

Second, force probes can handle wet processing steps if 1) they are still attached to the BOX, 2) they are handled extremely carefully and 3) they are dried in a convection oven rather than a spin rinse dryer (SRD). Regarding the first point, the BOX prevents the probes from sticking to themselves or to the sidewalls of the devices, even if cracks have been etched into the BOX. Regarding the second point, the wafers specifically can not handle normal dump rinsers. Rather, they need to be manually rinsed in a still DI water

bath and set aside while the bath is refilled. The manual rinse step usually consists of lowering the wafers into the water (with the probes facing down), letting it sit, and very slowly raising it back up. Do not move the wafers quickly at any time. Regarding the third point, the wafers need to be air dried after any wet steps. While one could allow them to dry for 30-60 minutes at room temperature, I usually put the wafers in a 95°C convection oven for faster drying. Before putting the wafers in the oven I would recommend letting them sit out (while oriented vertically, as they should be in the entire cleaning process) so that the bulk of the water can drip off of them.

Third, the wafers can usually handle being placed facedown on flat surfaces after the probes are released as long as the surface isn't rough. For example, when etching the BOX from the backside of the wafer in AMT etcher, I have tried placing the wafers directly in the tool (there is no frontside photoresist at this point; it tends to burn) and have also tried putting the wafer in the tool on a backing wafer with kapton tape spacers between the two (to prevent direct contact). There wasn't any noticeable difference in yield between the two techniques. I would not take a released wafer and slide it across a table, but if it needs to be set facedown on a surface, the devices will probably survive.

Wafer handling after the backside etch is fairly nuanced, and very small changes in handling technique can result in dramatic changes in probe yield.

E.2 Runsheets

The runsheets for the various fabrication processes discussed in the thesis are presented in this section. The runsheets reflect the original processes used to fabricate the probes and not necessarily best practices. For example, the first fabrication run used RIE to remove the BOX and release the probes. At the time this was the only option; wet processing (e.g. pad etch) broke nearly all of the cantilevers. However, since then, a metal compatible vapor HF tool was installed in the SNF. The vapor HF tool yields far better results (e.g. no shadowing or polymer deposition) and is preferable to RIE release in all cases. Another example: the first fabrication used a forming gas anneal (FGA), but later experiments indicated a lower contact resistivity could be achieved without the FGA step and the second and third fabrication runs used a Ti/Al metal stack without an FGA. In other words, if you were to

process devices based upon these runsheets, please do not blindly follow them.

All of the process runsheets are written assuming an initial device layer thickness of roughly 340 nm and relatively soft probes (< 100 pN/nm). Thicker device layers will require DRIE rather than RIE for the frontside etch. Soft probes tend to break during wet processing after they are released from the wafer. Accordingly, all steps after the BOX removal are dry in each of the processes. If the probes are stiff enough to handle wet processing steps then steps can be rearranged and modified accordingly (e.g. pad etch rather than RIE BOX removal, assuming that HF vapor is not available). If you plan to use a wet BOX release, first render the wafer hydrophilic using an oxygen plasma RIE step or else the etchant may not enter the backside etch trenches.

Regarding the runsheets, please note that the the recipes and settings are only approximate. If you are trying to reproduce the processes it is essential that you examine the wafers before and after each step and perform additional measurements (e.g. SEM, profilometer, sheet resistance) whenever possible. Etch processes should be monitored using endpoint detection (if available) or by visually monitoring changes in wafer color, plasma color, or the hydrophobicity of the wafer. For the sake of brevity these steps are not explicitly included in the runsheets. To be perfectly clear:

- Examine the wafers before and after every step under the microscope. New cleanroom users do not do this enough and it is absolutely essential. If a layer looks strange, examine it using another tool (e.g. SEM, nanospec, ellipsometer) and rework the layer if necessary.
- Check the the thickness of every deposited film. Dielectric films can be measured via ellipsometry. The thickness of conductive films can be inferred from their sheet resistance or can be measured via profilometry.
- Check the sheet resistance of every conductive film. This data will be essential during device modeling. If the deposited film will not be large enough for 4PP sheet resistance measurements (e.g. diffusion through oxide windows), be sure to include blank process monitor wafers.
- Check the frontside-to-backside overlay using infrared microscopy if possible.

- Check critical dimensions via electron microscopy (e.g. probe width). There will always be some systematic bias between the mask layout and resulting dimensions due to etch/exposure settings.
- Check tool etch rates immediately before processing your wafers. This is particularly important for etches without endpoint detection or another clear endpoint indicator (e.g. becoming hydrophobic). This is particularly important for etching alignment marks because the amtetcher etch rate can be unstable.
- Similarly, check deposition rates whenever possible by running dummy wafers immediately before your device wafers. This is especially important for LTO deposition (tylanbpsg) because the deposition rate is unstable.

Several of the recipes will be abbreviated in the runsheets. Here are details on some of the tools and recipes that will be used:

- **amtetcher:** a batch reactive ion etcher used for etching alignment marks and silicon dioxide. We use the oxide etch and via etch recipes. Both use CHF₃ and O₂ at roughly 40 mTorr and a -530 V DC bias, but with different flow rates. The oxide etch program (6 sccm O₂ and 85 sccm CHF₃) etches oxide with high selectivity to silicon. The low oxygen flow rate leads to polymer buildup on non-SiO₂ surfaces and the high selectivity. The oxide etch program etches SiO₂ and Si at roughly 300 Å/min and 28 Å/min, respectively. The polymer buildup is a problem for thin mechanical structures. The second recipe (via etch) has lower selectivity (30 sccm O₂ and 50 sccm CHF₃) but does not deposit polymer at a significant rate. Alignment marks are etched with the via etch program. Processes that use amtetcher for BOX removal generally use the via etch program for the first half of the etch (to minimize polymer buildup) and the oxide etch program for the second half (to avoid damage to the piezoresistor).
- **gasonics:** a downstream oxygen plasma aura ash. In contrast with RIE tools, the oxygen ions in contact with the wafer have relatively low energy and this process is closer to the chemical etching side of the spectrum than the physical sputtering

side. We use gasonics to aid in the removal of photoresist. A second application is gasonics is to aid in the etching of small features. Photoresist is hydrophobic, making it difficult for water based etchants (e.g. BOE) to enter small features (e.g. $< 10 \mu\text{m}$). A short oxygen plasma pretreatment renders the photoresist hydrophilic without removing much material ($< 100 \text{ \AA}$). Gasonics is used before etching windows in oxide layers for diffusion doping and metal contact vias.

- **mitutoyo:** a depth gauge used to measure the overall thickness of a wafer
- **nanospec:** a single point spectroscopic reflectometer used to measure the thickness of single layer dielectric films. Does not work well with thin device layer SOIs due to multilayer effects.
- **p2:** a stylus profilometer.
- **prometrix:** a four point probe sheet resistance measurement tool.
- **woollam:** a variable angle spectroscopic ellipsometer (VASE) used to measure the thickness of multilayer film stacks. Works well with thin device layer SOIs but requires relatively large measurement windows (2-3 mm).

APPENDIX E. FABRICATION PROCESS DETAILS

Table E.1: Runsheet for the PR process from the first fabrication run.

Step	Purpose	Description	Tool	Recipe, parameters and notes
1	Prep	Label	wbnonmetal	
2		Remove dust	scribe	
3		Measure thickness	mitutoyo	
4		Initial film thickness	woollam	
5	Doping	Diffusion clean	wbdfusion	Perform HF dip last
6		Diffusion	tylan6	POCl_3 at 800°C for 35 min. Target 110 Ω/\square .
7		Remove PSG	wbnonmetal	30 sec 50:1 HF to remove the roughly 200 Å PSG layer
8	Marks	Pattern marks	litho	0.7 μm SPR955, 120 mJ/cm ² . Note 45° rotation
9		Etch marks	amtetcher	1200 Å target. Approx. 5 min via etch
10		Strip resist	wbnonmetal	20 min piranha
11	Metal	Diffusion clean	wbdfusion	Perform HF dip last. Rush wafers to the loadlock to minimize oxidation
12		Metal deposition	grphon	Target 3000 Å, pure Al. Do not presputter or heat in the loadlock
13		Bondpad litho	litho	0.7 μm SPR955, 120 mJ/cm ²
14		Bondpad etchback	wbmetal	Target 3000 Å, approx 40 sec in Transene Al etch A. Insensitive to overetch
15		Strip resist	wbmetal	20 min PRX-127
16	Frontside	Frontside litho	litho	0.7 μm SPR955, 120 mJ/cm ²
17		Frontside etch	p5000	Target 3400 Å, approx 90 sec. Recipe in Table E.9. Use endpoint detection
18		Strip resist	wbmetal	20 min PRX-127
19	Backside	Backside litho	litho	7 μm SPR220, 350 mJ/cm ² . Note -45° rotation for backside exposure
20		Frontside protection	litho	3 μm SPR220. No EBR, backside wash, or hotplate bake
21		Resist hardbake	oven	Bake at 95°C for 2-3 hours
22		Backside etch	stetch1	Time depends on etch rate and wafer thickness. Rotate wafer halfway
23	Release	Strip resist	gasonics	Short cleanup (e.g. 0C) to remove hardened resist. See Section E.1
24		Strip resist	wbmetal	20 min PRX-127. This is the final wet polymer cleanup
25		Clean wafers	wbmetal	DI water rinse. See Section E.1
26		BOX etch	amtetcher	400 nm target, approx. 10 min via etch and 10 min oxide etch
27		FGA	tylanfga	450°C for 30 min. Not generally recommended, use with caution

Table E.2: Runsheet for the PE process from the first fabrication run. In contrast with the rest of the processes, the PE process used relatively thick device layer wafers (5 μm) and the frontside etch is performed via DRIE rather than RIE.

Step	Purpose	Description	Tool	Recipe, parameters and notes
1	Prep	Label	scribe	
2		Remove dust	wbnonmetal	20 min piranha
3	Marks	Measure thickness	miutoyo	
4		Pattern marks	litho	0.7 μm SPR955, 120 mJ/cm ² . Note 45° rotation
5		Etch marks	amtetcher	1200 Å target. Approx. 5 min via etch
6		Strip resist	wbnonmetal	20 min piranha
7	PE films	Diffusion clean	wbdiffusion	
8		PE stack deposition	hionix	In-situ ICP clean. 100/100/250-1000/100 nm thick AlN/Ti/AlN/Ti layers
9		Wafer clean	wbmetal	20 min PRX-127
10	PE top	PE top litho	litho	1.6 μm SPR3612, 60 mJ/cm ²
11		Ti etch	wbmetal	100 nm target. 1-2 min Ti etch (20:1:1 H ₂ O:H ₂ O ₂ :HF)
12		Resist strip	wbmetal	20 min PRX-127. Resist would otherwise dissolve in the TMAH.
13		AlN etch	wbgeneral	250-1000 nm target. 10-40 min 25% TMAH. Do not agitate wafer
14	PE bot	PE bot litho	litho	1.6 μm SPR3612, 60 mJ/cm ²
15		Ti etch	wbmetal	100 nm target. 1-2 min Ti etch. If etch fails, go back and redo the top etches.
16		Resist strip	wbmetal	20 min PRX-127
17		AlN etch	wbgeneral	100 nm target. 5 min 25% TMAH. Do not agitate wafer
18	Frontside	Frontside litho	litho	1 μm SPR955, 150 mJ/cm ²
19		Frontside etch	stetch1	5 μm target. Check for oxide before etching.
20		Strip resist	wbmetal	20 min PRX-127
21		Check thickness	p2	Measure device layer thickness via profilometer. 5 μm target.
22	Backside	Backside litho	litho	7 μm SPR220, 350 mJ/cm ² . Note -45° rotation for backside exposure
23		Frontside protection	litho	3 μm SPR220. No EBR, backside wash, or hotplate bake
24		Resist hardbake	oven	Bake at 95°C for 2-3 hours
25		Backside etch	stetch1	Time depends on etch rate and wafer thickness. Rotate wafer halfway
26	Release	Strip resist	gasonics	Short cleanup (e.g. 0C) to remove hardened resist. See Section E.1

Table E.2: Runsheet for the first fabrication PE process (continued).

Step	Purpose	Description	Tool	Recipe, parameters and notes
27	Strip resist	wbmetal		20 min PRX-127. This is the final wet polymer cleanup
28	Clean wafers	wbmetal		DI water rinse. See Section E.1
29	BOX etch	amtetcher		400 nm target, approx. 10 min via etch and 10 min oxide etch

Table E.3: Runsheet for the PRPE process from the first fabrication run.

Step	Purpose	Description	Tool	Recipe, parameters and notes
1	Prep	Label	scribe	
2		Remove dust	wbnonmetal	20 min piranha
3		Measure thickness	mitutoyo	
4		Initial film thickness	woollam	
5	Marks	Pattern marks	litho	0.7 μm SPR955, 120 mJ/cm ² . Note 45° rotation
6		Etch marks	amtetcher	1200 Å target. Approx. 5 min via etch
7		Strip resist	wbnonmetal	20 min piranha
8	Doping	Diffusion clean	wbddiffusion	Perform HF dip last
9		Oxidation	thermco	Target 650 Å, wet 900°C oxidation for 20 min
10		PR litho	litho	0.7 μm SPR955, 160 mJ/cm ²
11		O ₂ plasma	gasonics	Recipe 0A (standard pretreatment for etching small oxide features)
12		Oxide etch	wbnonmetal	6:1 BOE, approx. 3 min. Monitor wafer backside and overetch 20%
13		Resist strip	wbnonmetal	20 min piranha
14		Diffusion clean	wbddiffusion	Perform HF dip last
15		Diffusion	tylan6	POCl ₃ at 800°C for 35 min. Target 110 Ω/\square .
16		Oxide strip	wbnonmetal	6:1 BOE, approx. 3 min
17	PE films	PE stack deposition	hionix	In-situ ICP clean. 100/100/250/100 nm thick AlN/Ti/AlN/Ti layers
18		Wafer clean	wbmetal	20 min PRX-127
19	PE top	PE top litho	litho	1.6 μm SPR3612, 60 mJ/cm ²
20		Ti etch	wbmetal	100 nm target. 1-2 min Ti etch (20:1:1 H ₂ O:H ₂ O ₂ :HF)
21		Resist strip	wbmetal	20 min PRX-127. Resist would otherwise dissolve in the TMAH.
22		AlN etch	wbgeneral	250 nm target. 10 min 25% TMAH. Do not agitate wafer
23	PE bot	PE bot litho	litho	1.6 μm SPR3612, 60 mJ/cm ²
24		Ti etch	wbmetal	100 nm target. 1-2 min Ti etch. If etch fails, go back and redo the top etches.
25		Resist strip	wbmetal	20 min PRX-127
26	Vias	AlN etch	wbgeneral	100 nm target. 5 min 25% TMAH. Do not agitate wafer
27		Wafer clean	wbmetal	10 min PRS-1000

APPENDIX E. FABRICATION PROCESS DETAILS

Table E.3: Runsheet for the first fabrication PRPE process (continued).

Step	Purpose	Description	Tool	Recipe, parameters and notes
28		LTO deposition	tylanbpsg	
29	Via litho	O ₂ plasma	litho	Target 1000 Å, 400°C for 8 min 0.7 µm SPR955, 160 mJ/cm ²
30		Oxide etch	gasonics	Recipe 0A (standard pretreatment for etching small oxide features)
31		Resist strip	wbmetal	6:1 BOE, approx 5 min. Monitor wafer backside and overetch 20%
32		Wafer clean	wbmetal	20 min PRX-127
33	Metal	HF dip	wbmetal	10 min PRS-1000
34		Metal deposition	grphon	30 sec 50:1 HF dip to remove native oxide from the vias
35		Bondpad litho	litho	Target 3000 Å, pure Al. Do not presputter or heat in the loadlock
36		Bondpad etchback	wbmetal	0.7 µm SPR955, 120 mJ/cm ²
37		LTO etchback	ametcher	Target 3000 Å, approx 40 sec in Transene Al etch A. Sensitive to overetch
38		Resist strip	wbmetal	Target 1000 Å, approx 4 min oxide etch. Extremely sensitive to overetch
39		Frontside litho	litho	20 min PRX-127
40	Frontside	Frontside etch	p5000	0.7 µm SPR955, 120 mJ/cm ²
41		Strip resist	wbmetal	Target 3400 Å, approx 90 sec. Recipe in Table E.9. Use endpoint detection
42		Backside litho	litho	20 min PRX-127
43	Backside	Frontside protection	litho	7 µm SPR220, 350 mJ/cm ² . Note -45° rotation for backside exposure
44		Resist hardbake	oven	3 µm SPR220. No EBR, backside wash, or hotplate bake
45		Backside etch	ststch1	Bake at 95°C for 2-3 hours
46		Strip resist	gasonics	Time depends on etch rate and wafer thickness. Rotate wafer halfway
47	Release	Strip resist	wbmetal	Short cleanup (e.g. 0C) to remove hardened resist. See Section E.1
48		Clean wafers	wbmetal	20 min PRX-127. This is the final wet polymer cleanup
49		BOX etch	ametcher	DI water rinse. See Section E.1
50		FGA	tylanfga	400 nm target, approx. 10 min via etch and 10 min oxide etch
51				450°C for 30 min. Not generally recommended, use with caution

Table E.4: Runsheet for the second fabrication run. If PR or PRT probes are being fabricated then steps 38-50 (specific to the PRPE process) can be ignored.

Step	Purpose	Description	Tool	Recipe, parameters and notes
1	Prep	Label	scribe	
2		Remove dust	wbnonmetal	20 min piranha
3		Measure thickness	mitutoyo	
4		Background resistivity	prometrix	
5	Marks	Initial film thickness	woollam	9 point measurement
6		Pattern marks	litho	0.7 μm SPR955, 120 mJ/cm^2 . Note 45° rotation
7		Etch marks	ametcher	1200 Å target. Approx. 5 min via etch
8		Strip resist	wbnonmetal	20 min piranha
9	Diff. mask	Diffusion clean	wbdiffusion	Perform the HCl step last
10		Oxidize	thermcoll/2	Target 900 Å, wet 900°C oxidation for 36 min
11		LTO deposition	tylanbpsg	Run immediately after last step. Target 320 nm, 400°C for 22 min
12		Anneal	thermcoll/2	Run immediately after last step. Inert anneal (1100°C, 20 hours).
13		Check stack	woollam	Check the oxide and device layer thicknesses.
14	Contacts	Contact litho	litho	0.7 μm SPR955, 130 mJ/cm^2
15		O_2 plasma	gasonics	Recipe 0A (standard pretreatment for etching small oxide features)
16		Oxide etch	wbnonmetal	6:1 BOE, approx. 5 min. Monitor wafer backside and overetch 20%
17		Resist strip	wbdiffusion	20 min piranha
18		Diffusion clean	tylan6	Perform HF dip last
19		Diffusion	wbnonmetal	POCl_3 at 900°C for 20 min. Target 20 Ω/\square .
20		PSG strip	woollam	50:1 HF dip, 40 sec. Critical to limit undercut in the next steps
21		Check stack	litho	Check the remaining thermal oxide and densified LTO thickness
22	PR doping	PR litho	gasonics	0.7 μm SPR955, 130 mJ/cm^2
23		O_2 plasma	wbnonmetal	Recipe 0A (standard pretreatment for etching small oxide features)
24		Oxide etch	wbnonmetal	6:1 BOE, approx. 5 min. Monitor wafer backside and overetch 20%
25		Resist strip	wbdiffusion	20 min piranha
26		Diffusion clean		Perform HF dip last

APPENDIX E. FABRICATION PROCESS DETAILS

Table E.4: Runsheet for the second fabrication run (continued).

Step	Purpose	Description	Tool	Recipe, parameters and notes
27		Diffusion	tylan6	POCl ₃ at 825°C for 30 min. Target 70 Ω/□.
28		PSG strip	wbnonmetal	50:1 HF dip, 40 sec
29		Check stack	woollam	Check the remaining thermal oxide and densified LTO thickness
30		Oxide strip	wbnonmetal	6:1 BOE, remove all remaining oxide
31		Check monitors	prometrix	Check monitor sheet resistance. Continue processing them afterwards
32	Frontside	Frontside litho	litho	0.7 μm SPR955, 120 mJ/cm ²
33		Frontside etch	p5000	Target 3000 Å, approx 75 sec. Recipe in Table E.9. Use endpoint detection
33		Strip resist	wbmetal	20 min PRX-127
34		Check probe thickness	p2	Measure near but not directly on the probes
35	LTO	Diffusion clean	wbdiffusion	Perform the HCl step last
36		LTO deposition	tylanbpg	Target 200 nm, 400°C for 13 min
37		Check stack	woollam	Measure the Si thinning (2-3 nm) and LTO thickness on test coupons
38	PE films	Stack deposition	oem group	Vendor steps: presputter/AIN/Mo/AlN/Mo 200/250/750/750/500 Å
39		Clean wafers	wbgeneral	Steps with exposed Mo are processed at wbgeneral. 20 min PRX-127
40	PE top	PE top litho	litho	1 μm SPR955, 150 mJ/cm ²
41		Mo etch	p5000	50 nm target, approx 50 sec. Recipe in Table E.8. Use endpoint detection
42		AlN etch	p5000	75 nm target, approx 120 sec. Recipe in Table E.7. Use endpoint detection
43		Strip resist	wbgeneral	Rinse in water immediately, dry, then 20 min PRX-127
44	PE bot	PE bot litho	litho	1 μm SPR955, 150 mJ/cm ²
45		Mo etch	p5000	75 nm target, approx 75 sec. Recipe in Table E.8. Use endpoint detection
46		AlN etch	p5000	20 nm target, approx 45 sec. Recipe in Table E.7. Use endpoint detection
47		Strip resist	wbgeneral	Rinse in water immediately, dry, then 20 min PRX-127
48		Check stack	woollam	Measure the LTO thickness after PE processing, typically 70-80 nm removed
49	LTO	Metal clean	wbgeneral	10 min PRS-1000
50		LTO deposition	tylanbpg	Target 100 nm, 300°C for 16 min. Mo is enclosed in LTO afterwards.
51	Vias	Via litho	litho	0.7 μm SPR955, 130 mJ/cm ² (aim for slight underexposure)

Table E.4: Runsheet for the second fabrication run (continued).

Step	Purpose	Description	Tool	Recipe, parameters and notes
52		O ₂ plasma	gasonics	Recipe 0A (standard pretreatment for etching small oxide features)
53		Oxide etch	wbmetal	20:1 BOE, approx 3 min. Monitor wafer backside and overetch by 20%
54		Resist strip	wbmetal	20 min PRX-127
55	Metal	Wafer clean	wbmetal	10 min PRS-1000
56		HF dip	wbmetal	30 sec 50:1 HF dip to remove native oxide from the vias
57		Metal deposition	grphon	50 nm Ti + 400 nm (PRPE) or 1 μm (PR/PRT) Al. No presputter or heat
58		Check monitors	prometrix	
59		Metal litho	litho	1 μm SPR955, 110 mJ/cm ² (aim for slight underexposure)
60		Metal etch	p5000	450 or 1050 nm target, 40 or 80 sec (use endpoint). Recipe in Table E.6
61		Resist strip	wbmetal	Rinse wafers in water immediately, dry, then 20 min PRX-127
62		Check Al thickness	p2	
63		Probe station check	probestation	Check resistances (TLM) and test die in case Al or Ti is remaining
64		Check stack	woollam	Measure the LTO thickness after Al etching, typically 60-70 nm removed
65	LTO	Metal clean	wbgeneral	10 min PRS-1000
66		LTO deposition	tylanbpsg	Target 400 nm, 300°C for 50 min. Match to BOX thickness
67		Check stack	woollam	Measure thickness including BOX for crack etch (can use nanospec)
68	Cracks	Crack litho	litho	1 μm SPR955, 200 mJ/cm ² (overexposure is fine)
69		Oxide etch	amtetcher	Use oxide etch program, approx 50 min. Overetch by 50-100%
70		Resist strip	wbmetal	The resist is usually quite burned. Use gasonics 044 before 20 min PRX-127
71		Frontside coat	svgcoat	Coat frontside in resist but do not expose
72		Strip backside LTO	wbmetal	Strip the backside LTO, approx 100 sec in 6:1 BOE
73		Resist strip	wbmetal	20 min PRX-127
74	Backside	Backside litho	litho	7 μm SPR220, 350 mJ/cm ² . No EBR but run backside wash for ASML
75		Resist hardbake	oven	Bake at 95°C for 2-3 hours
76		Backside etch	umichigan	Ship wafers to U. Michigan with oxidized SSP handle wafers (they bond)
77	Debond	Strip resist	wbmetal	Remove resist in PRX-127. No vacuum or heat (> 50°C) until released
78		Debond	wbgeneral	Soak wafer stack in 70°C DI water for 30 min to release

Table E.4: Runsheet for the second fabrication run (continued).

Step	Purpose	Description	Tool	Recipe, parameters and notes
79		Rinse and dry	wbmetal	Rinse 5x in clean water and dry in 90 °C oven
80	Cleanup	Remove bulk	gasonics	Run program 0444 to remove the non-fluorocarbon polymer buildup
81		CF ₄ cleanup	drytek2	100 sccm CF ₄ at 100 mTorr, 500 W, 12/6 min backside/frontside
82		SF ₆ cleanup	drytek2	100/40 sccm O ₂ /SF ₆ at 50 mTorr, 500 W, 4/2 min backside/frontside
83		Final cleanup	gasonics	Run program 044 to cleanup any final polymer bits
84	Release	HF vapor	hfvapor	45/30 °C wafer/bath, 5-10 min (60 nm/min). Etch from frontside only

Table E.5: Runsheet for the fabrication of poly-on-insulator (POI) wafers. Most processes should include POI dummy wafers for testing each process before running the SOI wafers. Note that dopants diffused much more rapidly in polysilicon than single crystal silicon and if functional piezoresistors are desired then the POI wafers will need to be doped separately using shorter predeposition times.

Step	Description	Tool	Recipe, parameters and notes
1	Diffusion clean	wbdiffusion	Perform the HCl step last
2	Oxidize	thermco1/2	Target 400 nm, wet 1000°C oxidation for 60 min
3	Amorphous poly dep	thermcopoly1	Run immediately after last step. Target 340 nm, 580°C for 100 min
4	Grain growth	thermco 1/2	Run immediately after last step. Inert anneal (1075°C, 2 hours)
5	Backside poly	drytek2	Remove backside poly. 3-4 min isotropic Si etch (SF ₆ + 20% O ₂)
6	Backside oxide	wbnonmetal	Remove backside oxide. Approx 10 min 6:1 BOE
7	Check stack	woollam	Check the remaining thermal oxide and densified LTO thickness

Table E.6: Program for etching aluminum on the Applied Materials p5000 MERIE etcher. Settings are specified for the breakthrough etch (BT), main etch (ME) and overetch (OE) steps. The main etch was automatically terminated using the endpoint program described in Table E.11. A thin Ti underlayer (< 50 nm) was included in some of the wafers from the second and third fabrication processes, and the Ti was etched during the OE step.

	BT	ME	OE
Time (s)	10	-	8
Pressure (mTorr)	30	200	200
Power (W)	300	450	500
Magnetic field (G)	0	0	0
BCl ₃ flow (sccm)	40	40	40
Cl ₂ flow (sccm)	10	30	15
N ₂ flow (sccm)	0	40	20

Table E.7: Program for etching aluminum nitride (AlN) on the Applied Materials p5000 MERIE etcher. Settings are specified for the breakthrough etch (BT), main etch (ME) and overetch (OE) steps. The main etch was manually terminated by monitoring the Cl emission peak over time.

	BT	ME	OE
Time (s)	30	-	10
Pressure (mTorr)	8	12	12
Power (W)	100	100	100
Magnetic field (G)	0	0	0
BCl ₃ flow (sccm)	15	0	0
Cl ₂ flow (sccm)	0	15	15

Table E.8: Program for etching molybdenum (Mo) on the Applied Materials p5000 MERIE etcher. Settings are specified for the main etch (ME) and overetch (OE) steps; no separate breakthrough step was required. The main etch was manually terminated by monitoring the Mo emission peak over time.

	ME	OE
Time (s)	-	10
Pressure (mTorr)	50	50
Power (W)	100	100
Magnetic field (G)	0	0
SF ₆ flow (sccm)	100	100

Table E.9: Program for etching silicon (Si) on the Applied Materials p5000 MERIE etcher. Settings are specified for the breakthrough etch (BT), main etch (ME) and overetch (OE) steps. The main etch was automatically terminated using the endpoint program described in Table E.11. Note that the HBr:Cl₂ ratio during the ME is critical for the control of the sidewall angle profile. The default program on the tool (HBr/Cl₂ = 20/20 sccm) resulted in a slight positive sidewall angle ($\approx 75^\circ$). Reducing the ratio of sidewall passivation (HBr) to etching (Cl₂) resulted in more vertical sidewalls, as described in Ref. [414].

	BT	ME	OE
Time (s)	10	-	30
Pressure (mTorr)	30	100	100
Power (W)	250	200	90
Magnetic field (G)	0	75	75
CF ₄ flow (sccm)	35	0	0
HBr flow (sccm)	0	17	30
Cl ₂ flow (sccm)	0	23	15

Table E.10: Optical emission wavelengths for the etchant gases and volatile products used in fabricating the fast force probes. Compiled from Refs. [375, 415, 416].

Wavelength (nm)	Species	Role
319.0	Mo	product
396.4	Al	product
470.5	Br and Cl	etchant

Table E.11: Endpoint detection programs for the Al and Si etch processes in the Applied Materials p5000 MERIE etcher. Other etch processes were manually terminated by monitoring the optical emission spectroscopy (OES) signal.

Algorithm	Al	Si
Wavelength (nm)	396.4	470.5
AGC level (%)	50	50
AGC time (s)	10	10
Initial dead time (s)	15	20
Window height (%)	-1	0.5
Window time (s)	1	2
No. of windows in (-)	3	3
No. of windows out (-)	4	3
Derivative time (s)	0	0
Overetch (%)	0	0
Display start level (%)	80	50
Display gain (-)	5	10
Magnetic field period (s)	2	2.1
No. of periods to average (-)	0	1

Table E.12: Etch rates are quoted in Å/min. Thermally grown SiO_2 , undensified LPCVD SiO_2 and photoresist are abbreviated as ThOx, LTO and PR, respectively. Directly measured rates are indicated in bold while other etch rates are compiled from Refs. [417–419]. Etch rates of less than 10 Å/min are abbreviated as - while etch rates greater than 10,000 Å/min are abbreviated as +. Rates that were not measured and are not cited anywhere in the literature are abbreviated as *. Note that pad etch roughens and removes approximately 500 Å of material from Al before terminating [420] and tends to lift off photoresist.

Description	Composition	Al	AlN	Mo	Si	Ti	ThOx	LTO	PR
Wet Ti etch	20:1:1 $\text{H}_2\text{O}:\text{H}_2\text{O}_2:\text{HF}$ at 25°C	150	-	-	-	8800	120	>120	-
Wet AlN etch	25% TMAH at 25°C	-	240	-	-	-	-	-	+
Wet Al etch	Transene Al etch Type A at 40°C	4000	-	-	-	-	-	-	-
Piranha	9:1 $\text{H}_2\text{SO}_4:\text{H}_2\text{O}_2$ at 120°C	1800	-	180	-	2400	-	-	+
20:1 BOE	20:1 40% $\text{NH}_4\text{F}:49\%$ HF at 25°C	1400	-	-	-	+ 210	740	-	-
Pad etch	Ashland pad etch 4 at 25°C	500	-	-	-	20	285	2200	+
HF vapor	38% HF at 30°C and wafer at 45°C	-	-	-	-	660	800	-	-
Dry Al etch	Table E.6	7500	*	*	4500	+	<540	540	3750
Dry AlN etch	Table E.7	750	540	236	400	1000	<78	78	<3750
Dry Mo etch	Table E.8	-	430	350	-	<580	580	*	-
Dry Si etch	Table E.9	*	*	2400	*	<100	<100	800	-

Appendix F

The first fabrication run

The devices described in the main body of the thesis were the second generation of force probes that I fabricated. The first devices were fabricated in 2008 from September through November as part of Stanford’s E342 course with help from Bryan Petzold. They included sensor-only (PR), actuator-only (PE) and integrated (PRPE) force probes. Piezoelectric actuation was accomplished using AlN deposited on a Ti electrode for this run.

In this appendix we will discuss their design, fabrication, performance, and the issues that came up requiring a redesign and the additional fabrication run. The first fabrication run is described in an appendix because it sidetracks the main story of the thesis, but it was an important step in refining the cantilever design process (particularly the thermal and piezoelectric modeling work [302, 355]) and established that the fabrication process used in the second generation force probes was feasible.

F.1 Device design

The layouts of the first generation PR, PE and PRPE probes are shown in Figure F.1. For clarity several of the layers (aluminum, top piezoelectric, bottom piezoelectric) are drawn as data dark rather than data clear. We will briefly discuss each of the designs without delving into too many of their shortcomings just yet.

The PR designs (Figure F.1a) use three layers: frontside, metal and backside. Rather than doping just the piezoresistors through an oxide mask, in this design the entire wafer

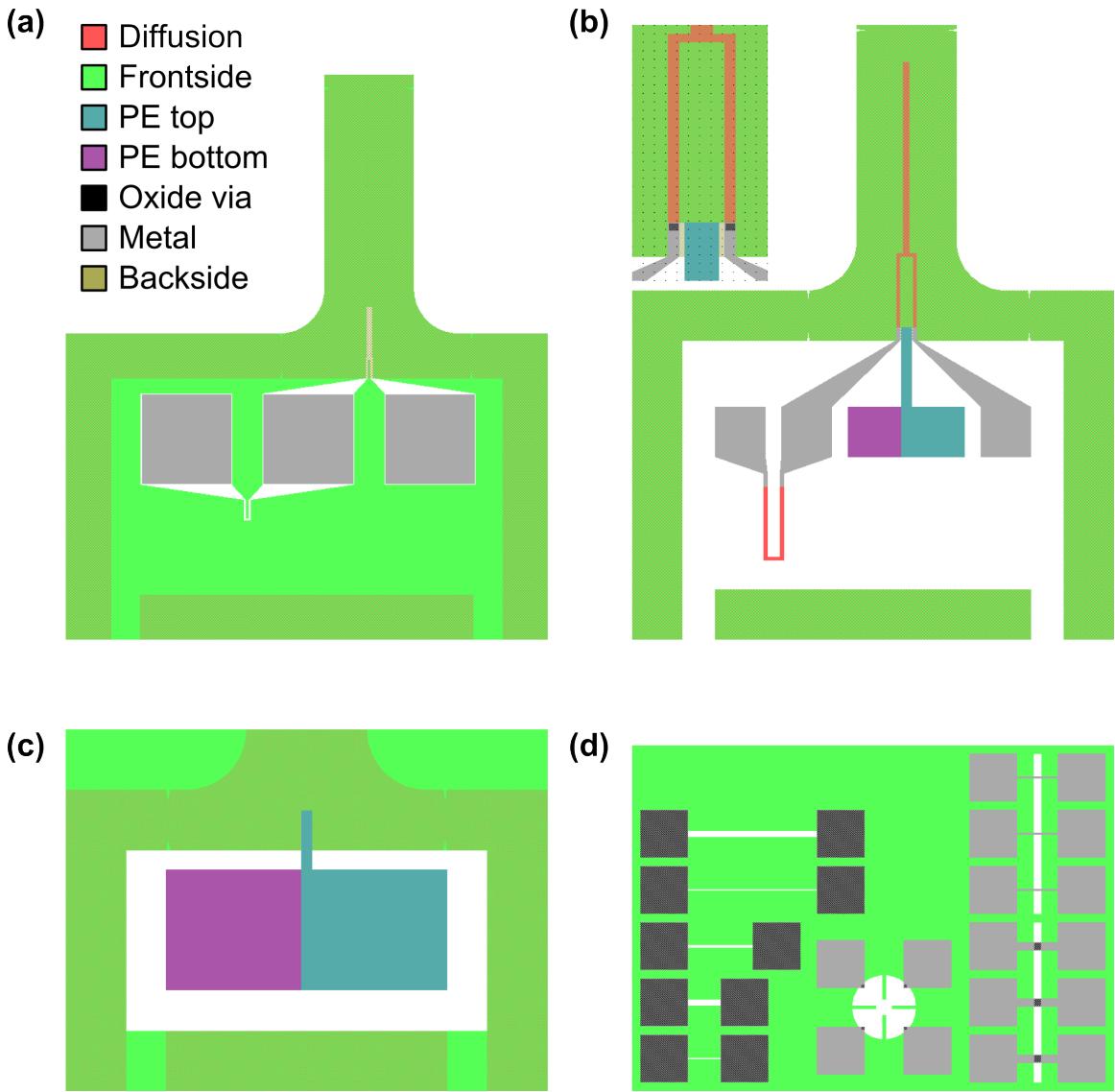


Figure F.1: Device layouts from the first fabrication run. Four different types of devices were fabricated: (a) PR probes, (b) PRPE probes, (c) PE probes and (d) test dies. The PR, PRPE and PE probes used three, seven and four layers, respectively. A detailed view of the PRPE probe is inset in (b). For clarity several of the layers (metal, PE top, PE bottom) are shown here as data dark rather than data clear.

is doped and then all of the undesired silicon is etched away. This approach simplifies the process at the expense of design flexibility. The device layout is fairly straightforward and consists of three bondpads (each 300 x 300 μm) and two silicon resistors. There is

no substrate contact because all of the silicon is doped. During operation the right and left bondpads are biased and grounded, respectively, while the middle bondpad serves as an input to the instrumentation amplifier.

The on-chip resistor is for temperature compensation and has dimensions identical to the piezoresistor portion of the force probe. The layout allows for a 5 μm gap between the metal contact and edge of the doped silicon, although it is not crucial due to the lack of a substrate contact. The design uses silicon interconnects to connect the piezoresistor to the bondpads. Each interconnect adds about 4.5 squares for a total of 9 squares of parasitic resistance. Assuming a piezoresistor sheet resistance of $140 \Omega/\square$ this corresponds to excess resistance of 1260Ω . Finally, the release point of the force probe is defined by the backside etch rather than from the front side. As noted by other piezoresistive cantilever designers, this configuration can lead to substantial variation in the effective length and corresponding stiffness and resonant frequency of the probe [22, 329]. Crack initiators are visible in all three of the probe designs but were found to be ineffective; we will discuss BOX cracking and mitigation approaches while discussing the second fabrication run.

The PRPE designs (Figure F.1b) increase the number of layers to seven: diffusion, frontside, PE top, PE bottom, oxide via, metal and backside. The sensor design is similar to the PR probes; two silicon resistors and three bondpads. However, rather than blanket doping the wafer and etching back the undesired silicon the PRPE probe design uses a selective doping process. This is necessary because the piezoelectric actuator needs to be mechanically coupled to the rest of the probe. An alternative approach would have been to use a dielectric film for this connection, but at the time we thought that it was necessary to deposit the piezoelectric stack directly onto bare silicon in order to yield a high d_{31} coefficient. We will see in the second and third fabrication processes that this is not strictly necessary.

The piezoelectric layers are patterned after the entire PE stack has been deposited. Due to the high impedance of the piezoelectric we were concerned that patterning the PE layer before depositing the top electrode risked shorting out the actuator. The top and bottom PE layers are defined such that the bottom PE electrode layer protects the top electrode and PE actuator while the bottom electrode is patterned. The downside of this design and processing approach is excess capacitance between the two electrodes due to their

overlap. The excess capacitance does not have any practical impact on the work in this thesis but would degrade sensor or higher bandwidth actuator operation. The PRPE images were drawn when we were still planning on patterning the PE layer before depositing the top electrode, and the top PE electrodes were initially going to be coated with a layer of aluminum for easier wirebonding. The top aluminum layer is not shown in Figure F.1b for clarity and because the aluminum above the PE bondpads was removed during wafer processing through the use of utility images.

Metal rather than silicon interconnects are used in the PRPE design. Contacts to the silicon piezoresistor are made through either 8 x 8 or 28 x 28 μm vias which correspond to piezoresistor legs that are 10 or 30 μm wide with a 1 μm alignment tolerance gap on either side of the contact. Assuming a sheet resistance of $140 \Omega/\square$ and contact resistivity of $20 \mu\Omega\cdot\text{cm}^2$ the 8 and 28 μm contacts correspond to contact resistances of 68 and 19 Ω using one-dimensional transmission line contact theory. In hindsight metal interconnects are clearly superior to silicon interconnects.

One important aspect of the PRPE probe design is that the piezoresistor contacts were made on either side of the piezoelectric actuator, making the piezoresistor 10 to 30 μm wide. As discussed in the last section, the transverse section of the piezoresistor should be minimized because $\pi_T \approx -\pi_L/2$, and we will see during testing that this had a notable impact on the sensitivity of short, wide PRPE designs. The gap between the PR interconnects and PE actuator is 5 μm and the PR interconnects are either 10 or 30 μm wide.

The PE designs (Figure F.1c) reduce the number of layers to four: frontside, PE top, PE bottom and backside. Only two bondpads (each 300 x 300 μm) are present on the device. As discussed while describing the PRPE design, the top and bottom PE layers are defined such that the actuator can be patterned after the entire actuator stack is deposited.

Finally, the test die (Figure F.1d) uses four layers as well: frontside, oxide via, metal and backside. Resistors of varying length are included on the left hand side of the die for inferring the sheet resistance from the transfer length method (TLM). A clover-shaped Van der Pauw structure is included in the center of the die for a separate determination of the sheet resistance. On the right hand side are two Kelvin bridge contact structures for measuring the contact resistance.

The nominal dimensions, spring constants and undamped resonant frequencies of the PR, PE and PRPE probes are presented in Tables F.3, F.5 and F.4.

F.2 Reticle and wafer layouts

The first fabrication run used two reticles, which are reproduced in Appendix D. An example PE wafer layout is shown in Figure F.2. The devices are placed in a checkerboard pattern within each cell to ensure that there is sufficient space between them to prevent wafer cracking. The spacing between each row and column is 3 mm. Approximately 140 devices were fabricated on each wafer. The wafer layout was manually defined using the standard ASML layout software. Four alignment marks were used for frontside alignment (on the left- and right-hand sides of the wafer) and four marks were used for backside alignment. The first wafers in this run were processed with both PM and XPA alignment marks and no substantial difference in alignment overlay or mark robustness was observed. Most wafers were processed with PM marks.

F.3 Process flow

In this section we will describe the PR and PE process flows before integrating the two for the PRPE process flow. The limitations and issues with the processes will be discussed in the following section (Section F.4).

The PR process is illustrated in Figure F.3. The process was previously described in Refs. [196] and [302]. The devices are fabricated using a three mask process. Process details, such as wafer handling issues and etch recipes, are presented in Appendix E. The details PR process runsheet is presented in Table E.1.

At the start of the process the wafers are labeled with a diamond scribe and cleaned. The thicknesses of the device layer and BOX are measured via variable angle spectroscopic ellipsometry (VASE) and the overall thickness of the wafer is measured using a depth gauge. Next, the wafers are diffusion cleaned and doped via POCl_3 diffusion (Figure F.3a). The diffusion temperature and times varied, but 800°C and 35 minutes was used for most of the wafers, yielding a sheet resistance of $110\text{-}120 \Omega/\square$. A thin layer of phosphosilicate

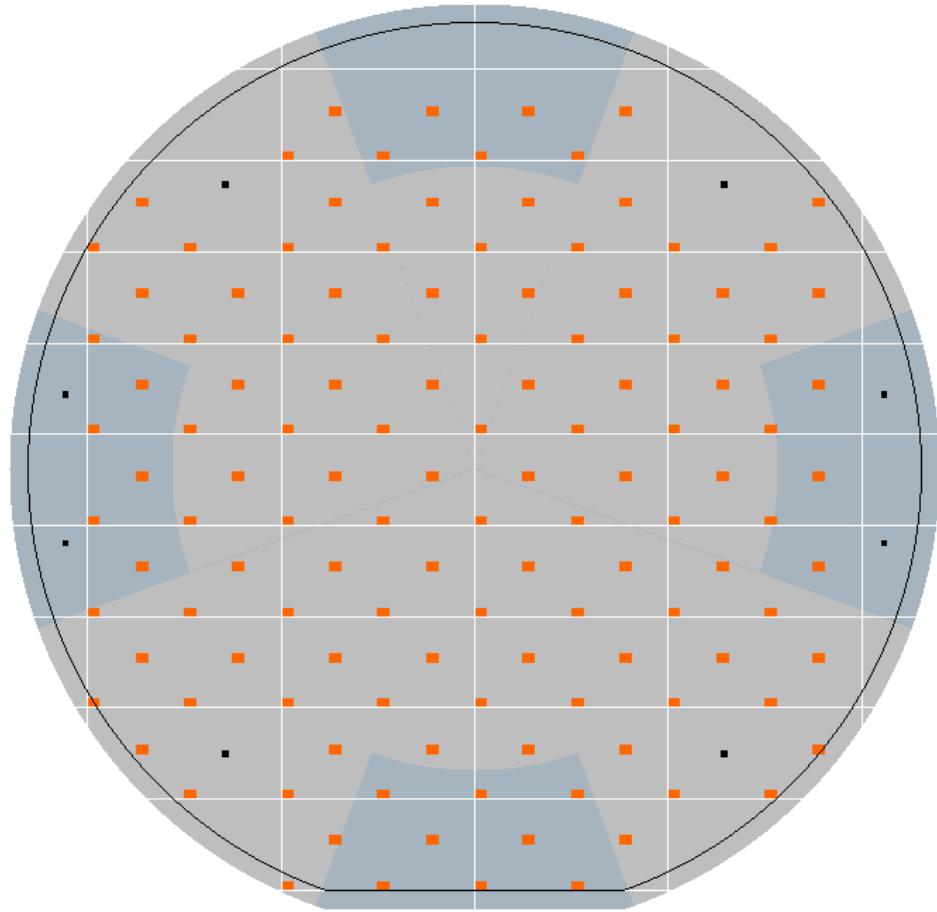


Figure F.2: An example wafer layout used in the first fabrication run. This particular layout was used for the fabrication of PE test devices. The probes are colored orange while the alignment marks are colored black. The four marks located in the gray zones on the right and left are used for frontside alignment while the other four marks are used for backside alignment.

glass (PSG) is deposited during the diffusion process, and it is stripped using a 30 second dip in 50:1 HF. Next, alignment marks are patterned into 0.7 μm thick Shipley SPR955 and etched 120 nm into the wafer using a CHF_3/O_2 plasma (Figure F.3b). The wafers are rotated 45° during the alignment mark patterning so that the piezoresistors will be oriented in the $<100>$ direction.

The ASML stepper detects the alignment marks based upon their diffraction pattern. The marks are illuminated with a 633 nm HeNe laser. Due to the relatively thin device

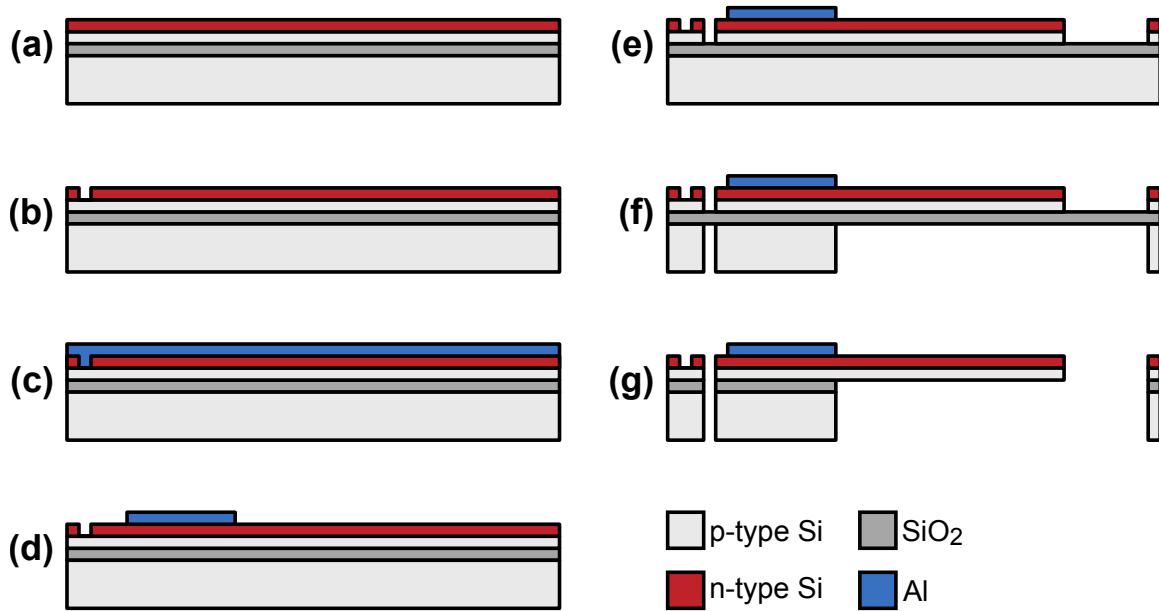


Figure F.3: Piezoresistive sensor-only process flow from the first fabrication run. (a) Wafers are blanket doped using POCl_3 diffusion at 800°C for 35 minutes. (b) Alignment marks are patterned and etched via RIE. (c) Aluminum is sputtered and (d) patterned by wet etching to form the bondpads. (e) The silicon device layer is patterned and etched via RIE. Finally, the backside of the wafer is (f) patterned and etched via DRIE and (g) the BOX is etched via RIE to release the cantilevers.

layers of our wafers, there was initially some concern that the stepper would have trouble detecting the marks. The stepper reports a wafer quality metric (WQ%) for each mark, with $\text{WQ\%} > 5\%$ required for alignment. Simulations performed by ASML staff suggested that mark depths of 80 or 180 nm would have been ideal for a 340 nm thick device layer. However, we did not encounter any issues using the standard alignment mark depth. The stepper reported WQ% values of 80-90% for standard wafers and values of 30-50% for our SOIs, more than sufficient for wafer alignment. Regardless, once the wafers are coated with metal the alignment mark quality becomes a non-issue.

A 300 nm thick layer of pure aluminum is sputtered on to the frontside of the wafer next (Figure F.3c). The wafers are cleaned before the metal deposition and a final HF dip removes the native oxide before the wafers are rushed to the sputterer load lock and pumped down. The wafers should be exposed to the air for no more than 5-10 minutes in order to ensure the formation of a good metal/Si contact. The wafers are heated to 200°C

for 3 minutes in the loadlock via a set of quartz lamps. We will discuss the impact of heat and presputtering on the resistivity of the metal/Si contact in the next section. The Al was patterned and then etched back (Figure F.3d) using a premixed chemical etchant (Transene Al etch Type A). The Al etchant has high selectivity towards Si, and is the reason that we opted to use pure Al rather than 99%/1% Al/Si for metallization. If we had opted for the latter option, the residual Si would have been left on the wafers and performing a freckle etch would have damaged the piezoresistors.

After the metal is deposited and patterned, the SOI device layer is patterned and etched (Figure F.3e). If the device layer is thin ($< 1 \mu\text{m}$) then it is etched via RIE in a Cl_2/HBr plasma (Table E.9). Several wafers with thicker device layers (2-10 μm) were processed and required the application of DRIE for its higher etch rate and selectivity with respect to photoresist. The frontside etch removes most of the silicon from the device, leaving only the silicon interconnects, piezoresistive loop and cantilever.

The backside of the wafer is patterned next (Figure F.3f) using the 3D align feature of the ASML stepper. A 7 μm thick layer of photoresist is spun onto the backside of the wafer and lithographically patterned. During the exposure the unprotected frontside of the wafer is in contact with the spincoater, hotplate and the chuck in the stepper, but the probes are not visibly damaged. Scanning probes with unprotected sharp tips have also been processed in the same fashion [329]. After the backside photoresist is patterned, the front-to-back alignment is checked using an infrared (IR) microscope before the frontside of the wafer is coated with a 3 μm thick photoresist layer. No edge bead removal (EBR) or backside rinse is used in coating the wafer frontside. The wafers are then placed in a 90°C convection oven for 2-3 hours to hard bake the photoresist layers on the front and backside. Once the photoresist has been hardened the wafers are etched from the backside, stopping on the BOX.

The probes are released by etching the BOX from the backside of the wafer (Figure F.3g). After being released from the BOX the softer probes (less than approx 20 pN/nm) break if submerged into liquid. The surface tension of the liquid can break the probes by either bending them back on themselves or adhering them to the sidewall underneath the probes. Accordingly, all of the processes after the BOX is removed must be dry. After the DRIE step, the photoresist is stripped from both sides of the wafers via O-2 plasma

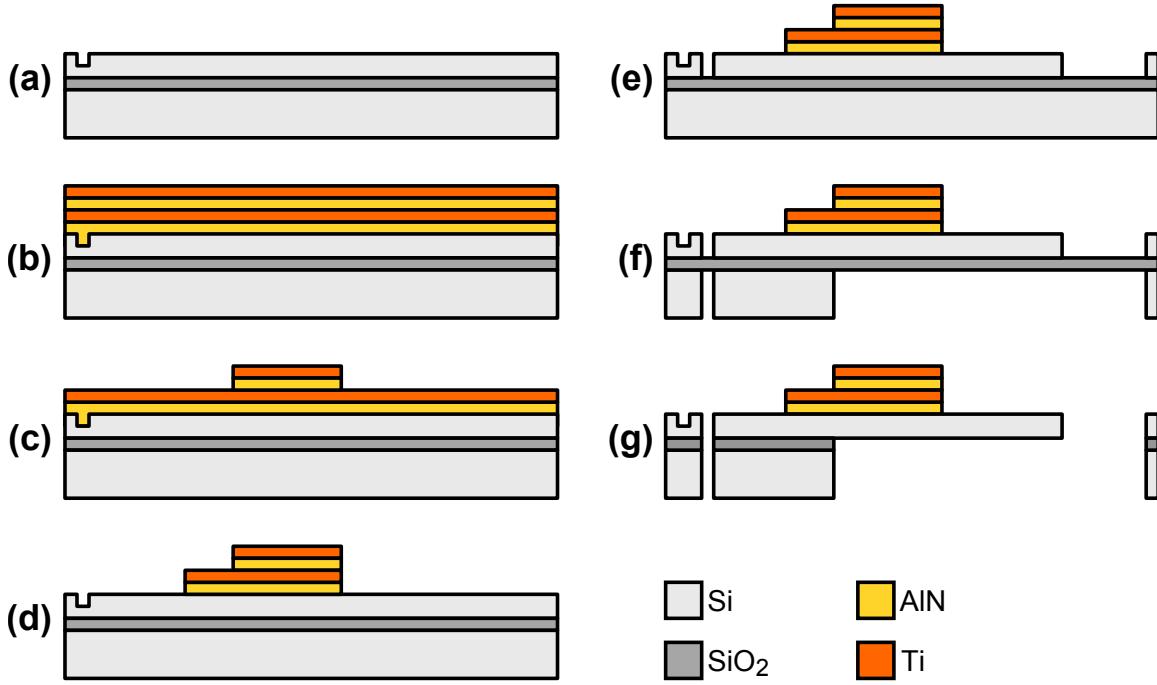


Figure F.4: Piezoelectric actuator-only process flow from the first fabrication run. (a) Alignment marks are patterned and etched via RIE and (b) the piezoelectric stack (AlN/Ti/AlN/Ti) is deposited on the wafer. The (c) top and (d) bottom Ti/AlN layers are patterned by wet etching and (e) the device layer is patterned and etched. Finally, the backside of the wafer is (f) patterned and etched via DRIE and (g) the BOX is etched via RIE to release the cantilevers.

and a metal compatible resist stripper (PRX-127). Vapor HF, the ideal option for removing the BOX, was not available at the time that these wafers were processed, so the BOX was etched via RIE in a CHF₃/O₂ plasma. Once the BOX is removed the wafers are annealed in a 4%/96% H₂/N₂ forming gas.

The PE fabrication process is illustrated in Figure F.4. Most of the steps are identical to the PR process. The two notable changes are the removal of the doping step and the addition of the PE stack deposition and patterning steps. We opted to use Ti rather than Mo or Pt for the metal electrode in the first fabrication run for tool compatibility reasons. We worked with a local vendor, Tango Systems, in order to develop the Ti/AlN process and first presented the process in Refs. [421] and [355].

Until our work, the piezoelectric properties of AlN on Ti were not particularly good.

Table F.1: AlN deposition parameters from the first fabrication run.

Parameters	Values
Base pressure (torr)	10^{-8}
Deposition pressure (torr)	5×10^{-3}
Target power (W)	5000
Substrate Bias Power (W)	200-400
Substrate bias voltage (V)	40-57
Ar:N ₂ flow rate (sccm)	10:40
Substrate temperature (°C)	200
Target-substrate distance (mm)	45
Deposition rate (nm/min)	21.2

The piezoelectric properties of polycrystalline AlN are derived from columnar grains with (002) crystal orientation. Metals that have been shown to yield reproducibly good AlN texture (e.g. Mo, Pt) are distinguished from other electrode materials (e.g. Ti, Al) partly by their small degree of lattice mismatch with AlN [422]. Metal electrode crystal planes which present hexagonal orientation for the growth of wurtzite AlN include Al (111), Pt (002), Ti (002) and Mo (110). Although the lattice mismatch of AlN with Ti and Pt is comparable, Pt had yielded better performance to date [423]. Exposure of the bottom electrode to oxygen has been shown to affect AlN grain structure and polarity [424], and an amorphous layer of AlN precedes columnar growth when deposited on TiO₂ [425]. These results suggested that the piezoelectric response of AlN on Ti could be improved by investigating the surface condition of the deposition substrate.

We chose to use Ti rather than Al for the metal electrodes for its greater performance and process compatibility. Ti (002) has less lattice mismatch (5%) with AlN than Al (111) (23%), reducing strain near the film interface [422]. Etch selectivity between AlN, the electrode metal and the Si substrate is also required. Room temperature tetramethylammonium hydroxide (TMAH) etches both Al and AlN [419], but neither Ti nor Si. Thus, it is possible to etch the AlN with TMAH and the Ti with a mixture of HF and H₂O₂ without affecting the underlying Si substrate, whereas Al and AlN can not be selectively etched with respect to each other.

The AlN and Ti films were deposited in a pulsed DC reactive sputter deposition

Table F.2: Film thicknesses in the first PE fabrication run.

Layer	Thickness (nm)	Measurements (N)
Top Ti electrode	166 ± 30	29
AlN actuator	208 ± 11, 388 ± 37, 1683 ± 54	6, 5, 10
Bottom Ti electrode	88 ± 19	26
AlN interlayer	136 ± 23	22

system (Tango Systems, San Jose, CA). Power, pressure and substrate temperature for AlN deposition were held constant at 5 kW, 5 mTorr and 200°C, respectively. The temperature and pressure were chosen to minimize N incorporation into the Ti and intrinsic stress in the AlN [423]. Target-substrate distance was fixed at 45 mm. The chamber was evacuated to a base pressure of 10^{-8} torr before the sputtering. The AlN deposition rate was 21.2 nm/min with Ar and N₂ flow rates maintained at 10 sccm and 40 sccm, respectively. The AlN deposition parameters are summarized in Table F.1. All Ti films were sputtered at 3 kW with 40 sccm Ar to yield a deposition rate of 40 nm/min.

For surface condition studies, the Si substrate was cleaned by either an inductively coupled plasma (ICP) at 800 W bias for 150 seconds or sputtered at 800 W for 100 seconds. Both options utilized an Ar flow rate of 40 sccm and remove approximately 100 Å of material from the surface. The ICP cleaning was performed in a separate chamber from the sputtering chamber. Vacuum was maintained while transferring wafers from the ICP chamber to the PVD chamber. The Ti and AlN depositions were isolated through sputtering shields to control cross-contamination. We utilized an AlN interlayer below the bottom metal electrode as demonstrated by Kamohara et al [351]. The AlN interlayer increases the alignment of the bottom metal electrode, improving the AlN alignment. The thicknesses of films in the PE stack are summarized in Table F.2.

After the PE film stack is deposited the top Ti and AlN layers are patterned (Figure F.4c). First, the top Ti electrode is lithographically patterned and etched in a room temperature solution of 20:1:1 H₂O:H₂O₂:HF. The H₂O₂ oxidizes the Ti and the TiO₂ is etched in the HF. Although etching in straight 50:1 HF may appear to etch the Ti, we often had to repeat the etch in order to completely remove the Ti. After the Ti is etched, the photoresist is stripped and the wafer is placed in room temperature 25% TMAH. Both the

Ti and AlN etches are visually monitored and overetched by roughly 20%. It was important not to agitate the wafers during the AlN etch; the entire stack lifted off on several of the wafers when agitated, particularly for the thicker AlN films.

After the top Ti and AlN are patterned and etched the process is repeated for the bottom Ti actuator and AlN interlayer (Figure F.4d). The bottom PE mask layer is designed so that it overlaps the top PE mask layer, protecting the top Ti layer while the bottom Ti is etched.

After the PE stack is patterned the device layer (Figure F.4e) and backside of the wafer (Figure F.4f) are patterned and the BOX is removed (Figure F.4g) using the standard release process. The PE device wafers ranged from 2 to 10 μm in thickness and so the only notable change from the PR process was the use of DRIE rather than RIE to define the cantilevers.

The PRPE fabrication process is illustrated in Figure F.5. We have not presented the PRPE process in any publications previously. The process introduces several new steps in order to successfully integrate the PR and PE processes.

In the PRPE process, the alignment marks are patterned via RIE (Figure F.5a) before the wafers are oxidized. The oxide is approximately 650 \AA thick, thinning the entire device layer to about 310 nm. The oxide is lithographically patterned and windows are opened via wet etching. Wet etching rather than RIE is used in order to avoid damage to the crystal lattice (Figure F.5b). Next, the piezoresistors are formed through POCl_3 predeposition at 800°C for 35 minutes. In the process the phosphorus diffuses about 25 nm through the oxide mask (Figure 2.33). After the predeposition all of the oxide is stripped from the wafer and the PE film stack is deposited (Figure F.5c). Recall that approximately 100 \AA of Si is removed through the in-situ clean performed as part of the PE stack deposition. The PE stack is patterned using the same techniques discussed for the PE process flow (Figure F.5d and e).

After the PE actuator stack has been completely processed the entire wafer is coated in a 100 nm thick layer of LTO (Figure F.5 f). Vias are opened in the LTO through wet etching in order to allow a sputtered Al film to make contact with the piezoresistor (Figure F.5g). The Al is patterned using wet etching. Before the photoresist used to pattern the Al is removed, the exposed LTO is etched away via RIE (Figure F.5h). The LTO was removed via RIE rather than pad etch due to concern that the HF in the pad etch might damage the Ti. However, the oxide RIE program etches Si at about 20 $\text{\AA}/\text{min}$ and terminates on the

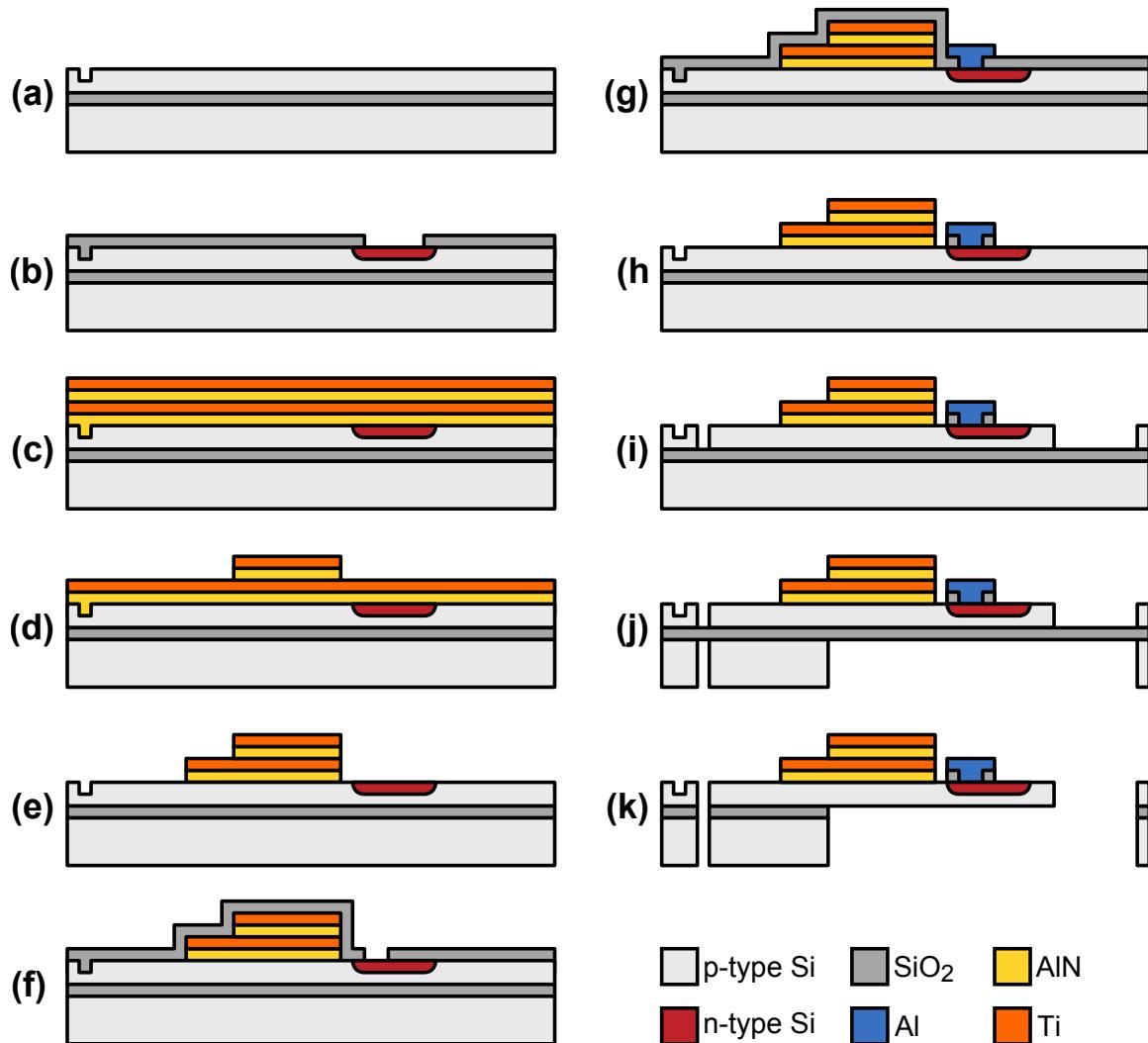


Figure F.5: Process flow for piezoresistively sensed and piezoelectrically actuated force probes from the first fabrication run. (a) Alignment marks are patterned and etched via RIE. (b) The wafer is oxidized, Windows in the oxide are opened via wet etching and the wafer is doped in POCl_3 . (c) The PE film stack is deposited and the (d) top and (e) bottom PE layers are patterned and etched. (f) A layer of LTO is deposited on the wafer and vias are wet etched. (g) Al is sputtered and patterned via wet etching in order to define the piezoresistor contacts, interconnects and bondpads. (h) Before removing the photoresist from the last step, the LTO is removed via RIE, stopping on the underlying silicon. (i) The frontside of the wafer is patterned via RIE and (j) the backside is patterned via DRIE. (k) The BOX is etched via RIE to release the cantilevers.

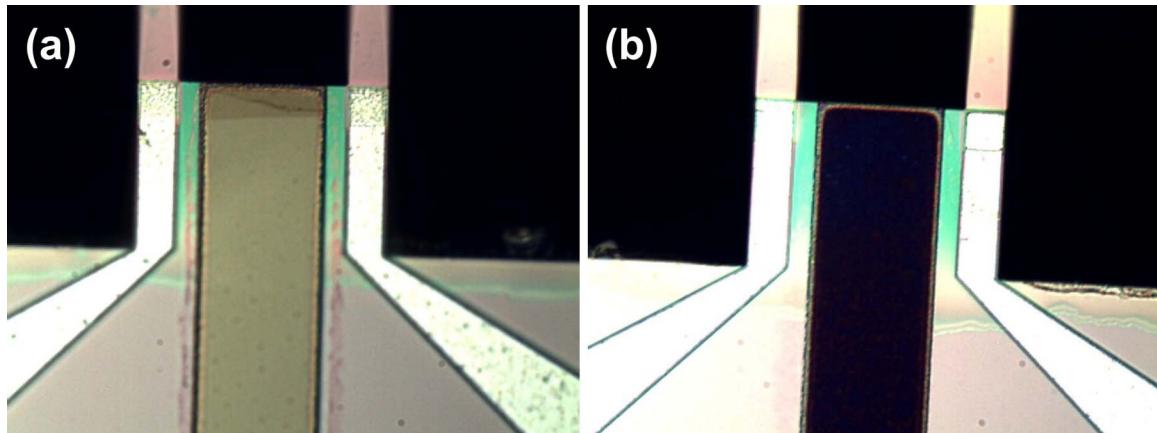


Figure F.6: Interactions between the top Ti metal electrode and forming gas anneal. A device is shown (a) before and (b) after a 30 minute 450 °C forming gas anneal. We did not investigate the change in film composition (e.g. using XPS, AES or SIMS) so can not tell exactly what happened (e.g. TiH or TiN formation) but wanted to highlight the impact of a forming gas anneal on device appearance.

piezoresistor. We will discuss the implications of piezoresistor damage from the PE in-situ clean and oxide RIE steps in the next section.

The remainder of the PRPE process follows the standard sequence of a frontside etch (Figure F.5i), backside etch (Figure F.5j) and BOX removal (Figure F.5k). A forming gas anneal was applied to the PRPE wafers as the last step. We do not recommend that researchers anneal metal contacts to thin n-type piezoresistors, which we will discuss in more detail later. But another reason to not anneal contacts is the formation of TiN or TiH on the surface (Figure F.6). Note the change in color of the Ti top electrode after the forming gas anneal.

F.4 Results and discussion

F.4.1 Sensor-only probes

A 10 μm thick PR device is seen in Figure F.7. The overall layout of the device, consisting of three Al bondpads and the temperature compensation (TC) and piezoresistive (PR) resistors is seen in Figure F.7a. A closer view of the cantilever is shown in Figure F.7b.

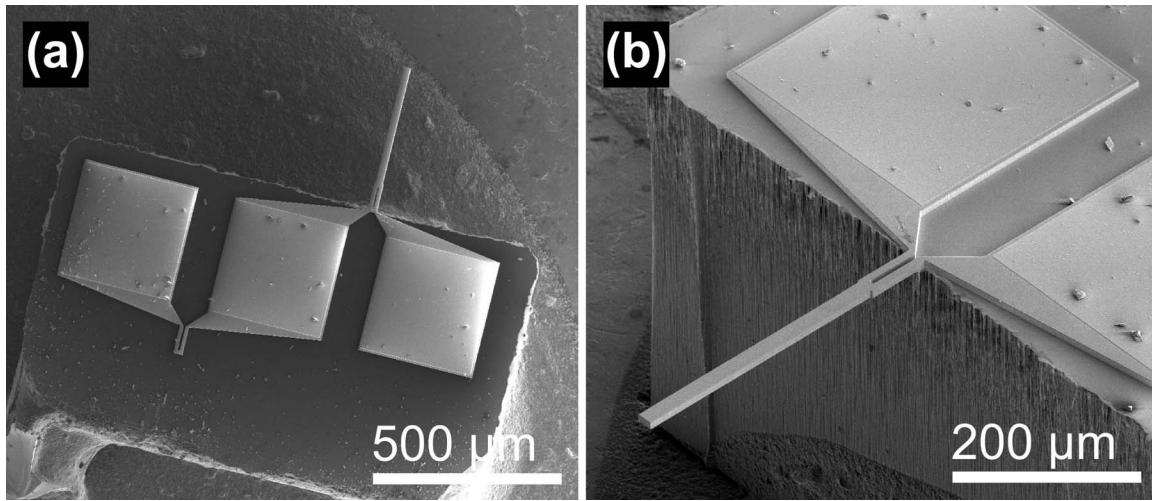


Figure F.7: A finished 10 μm thick sensor-only device. (a) Top down view of the die layout. The three bondpads are used to interface the released piezoresistor (PR) and unreleased temperature compensation resistor (TC) with the Wheatstone bridge based measurement circuit. (b) A closer view of a released cantilever. The cantilever is extremely flat and, at least on this device, the backside etch aligns well with the base of the cantilever.

The cantilever is extremely flat and the backside DRIE etch aligns fairly well with the base of the cantilever. The high density of particles in these SEMs is not typical; this was one of the first devices I removed from a wafer using snap tabs.

The predominant reason that the devices from the first fabrication run were not useful for hair cell experiments is illustrated in Figure F.7a. The die is small and rectangular, with the force probe located near the center of the die. Contrast the die layout with the schematic of the hair cell experimental setup shown in Figure 1.9. Quite simply, the die layout used in the first fabrication run was not compatible with use on an upright microscope. The roughly 0.5 mm distance from the probe to the edge of the die meant that probes could not be used at an angle. I developed the die layout from earlier piezoresistive cantilevers that were much longer (e.g. 1-2 mm vs. 30-100 μm) and were used without being flipped sideways. This fatal flaw in the probe design was solved in the following fabrication runs, but the characterization data obtained from the first generation devices was enormously important in developing and refining our design optimization techniques.

Figure F.8 presents several of the 340 nm thick cantilevers. A probe in relatively good condition is shown in Figure F.8a. The cantilever is fairly flat and although there is a small

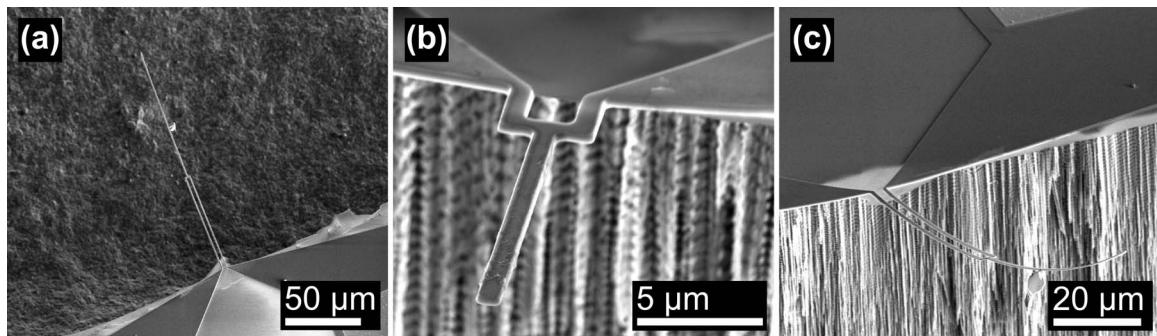


Figure F.8: Finished 340 nm thick sensor-only probes. (a) Top down view of a relatively long probe intended for low frequency (< 5 kHz) force measurements. The cantilever is in good condition besides the small particle attached to it beyond the end of the piezoresistive loop. (b) A 20 μm long probe highlighting the issues with polymer deposition and transverse resistance. (c) An 80 μm long probe highlighting the problem with strain gradients during the first fabrication run.

fleck of polymer on it halfway between the end of the piezoresistive loop and probe tip, the probe is fairly clean. Two common fabrication problems came up with the first batch of PR probes; polymer deposition and strain gradients. Figure F.8b shows a much smaller probe that is quite flat but has polymer visibly deposited on it. Figure F.8c shows a curved probe. Strain gradients necessarily develop from the formation of a highly doped piezoresistor near the surface of the device. But the probe is also noticeably curved beyond the end of the piezoresistor, due to either a thin compressively stressed layer on the backside of the probe (e.g. polymer or oxide) or an existing strain gradient in the device layer. I favor the second latter explanation because annealing the wafers at the beginning of the second and third fabrication runs (e.g. 1100°C for 20 hours) noticeably reduced the probe curvature. The curvature in thin first generation devices was not a major factor in limiting their usefulness.

One major issue in the first fabrication run for the PR, PE and PRPE devices was the alignment between the backside DRIE and the cantilever on the frontside of the wafer (Figure F.9). Depending on the location of the device on the wafer and the duration of the backside etch the probe could be under released (Figure F.9b) or over released (Figure F.9c). The etch rate of the DRIE tool used (stsetch in the SNF circa 2008) varied substantially across the wafer. A device in the center of the wafer might be perfect while the devices near the perimeter of the wafer were overetched. This variation in etch time coupled into

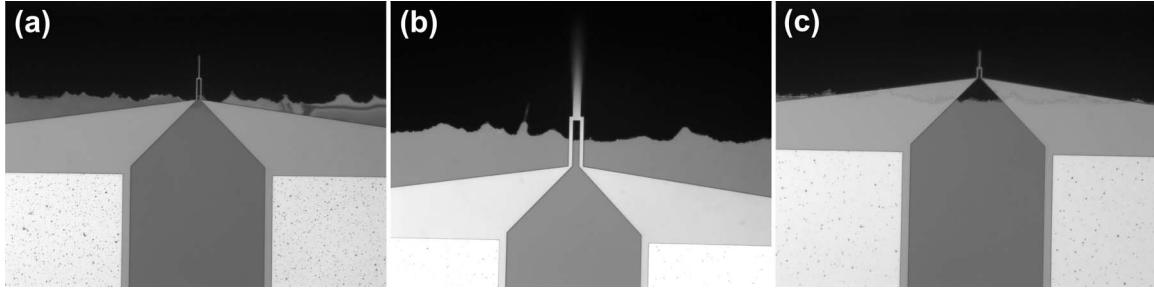


Figure F.9: Backside release alignment issues. Although the alignment was (a) good on some devices, other devices were either (b) under released or (c) over released. The former problem was more serious due to the large reduction in piezoresistor sensitivity. The latter problem led to a decrease in the probe stiffness and resonant frequency but did not substantially affect sensor resolution.

variation in the cantilever release point through footing at the Si-BOX interface.

A second, related issue with the DRIE release process was variation in the sidewall angle. Although the sidewall angle in a well tuned DRIE process is 90° , the etcher used for this process usually had a slightly negative sidewall angle (e.g. 95°), leading to a slight device undercut. We compensated for this effect by shifting the position of the backside photoresist mask. However, the line of sight nature of the RIE process used to remove the BOX limited our ability to perfectly compensate for the sidewall angle. Additionally, the same etch rate variations across the wafer (primarily due to temperature variation) coupled into sidewall angle variations.

The release point variation was mainly an issue for short cantilevers intended for high frequency measurements. Short probes are far more susceptible to under release (due to their shorter piezoresistor) as well as over release (due to their shorter length). Fortunately, a sufficient number of probes survived the release process with good back-to-front alignment that other problems limited the usefulness of the devices from the first fabrication run. Harley first noted issues with back-to-front alignment in his thesis [426] and suggested options for defining the cantilever release point from the frontside of the wafer.

A separate issue with the backside release process is shown in Figure F.10. The high compressive stress in the BOX (≈ -300 MPa) caused it to buckle and occasionally crack once released from the handle wafer by the DRIE etch. The frontside protection layer

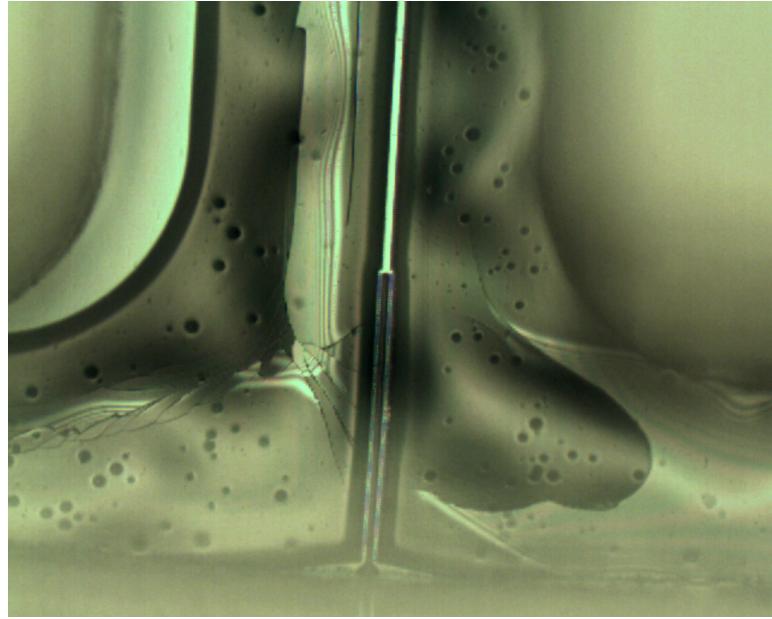


Figure F.10: The BOX noticeably buckled and cracked during the first fabrication run due to its high compressive stress. In this view from the backside of the wafer, while the BOX and frontside photoresist protection are still intact, cracks in the BOX are visible. However, the impact on probe yield was quite small, probably on the order of 2-3%.

of photoresist reduced the magnitude of the buckling and cracking. Other researchers at Stanford who did not use a frontside resist layer noted major issues with BOX cracking and reduced device yield, which actually motivated the development of the BOX moat approach used in the second and third fabrication runs [381].

Besides the device release and die layout issues the first generation PR devices performed extremely well. Their predicted and measured resonant frequencies are plotted against one another in Figure F.11. While the measured resonant frequency deviates by more than 100% in some cases, the overall agreement is excellent indicating that the intended and actual mechanical properties of the devices match each other. We used the Cleveland method for calculating probe stiffness during the first fabrication run [403]. The Cleveland method assumes that the length and width of the probe are known and that any deviation from the predicted resonant frequency is due to thickness variation. However, variation in probe length is coupled directly into variation in resonant frequency ($f_0 \propto l_c^{-2}$) and the back-to-front alignment issues that we have discussed contribute to some

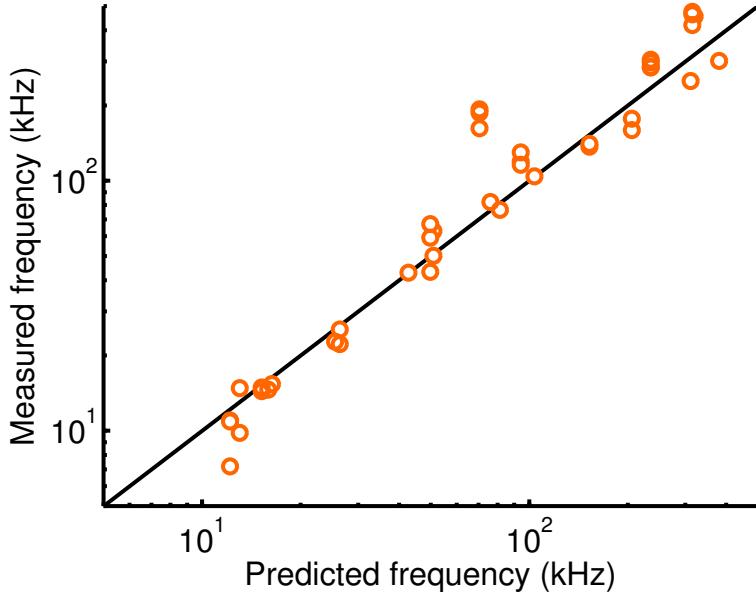


Figure F.11: Comparison between measured and predicted probe resonant frequencies. Probe spring constant was calculated directly from the measured resonant frequency data for the first run devices, so the overall agreement between predicted and actual probe mechanics was excellent. The data indicates that the intended and actual mechanical properties of the probes can match each other quite well, even for resonant frequencies of up to 500 kHz in air.

of the resonant frequency variation as well, rendering the Cleveland method not entirely accurate. The stiffness of the first generation PR devices is not critical, and we only note the issue of spring constant calibration because we will use a dimension invariant method (thermomechanical noise) for the following fabrication runs in order to avoid the issue of length variations completely.

The electrical resistance characteristics of the diffused piezoresistors are summarized in Figure F.12. The fixed resistance contribution of the interconnects was obtained via the transfer length method (TLM) in which resistances are plotted against the number of squares in the resistor (Figure F.12a). We determined an average sheet resistance of $115 \Omega/\square$ and constant interconnect resistance of $1.1 \text{ k}\Omega$. The interconnect resistance contribution (9.9 squares) agreed well with FEA simulation results. Variation in sheet resistance across an individual wafer was on the order of 3% based upon multipoint 4PP measurements. Combined with lithography and etch variations, the mismatch between the

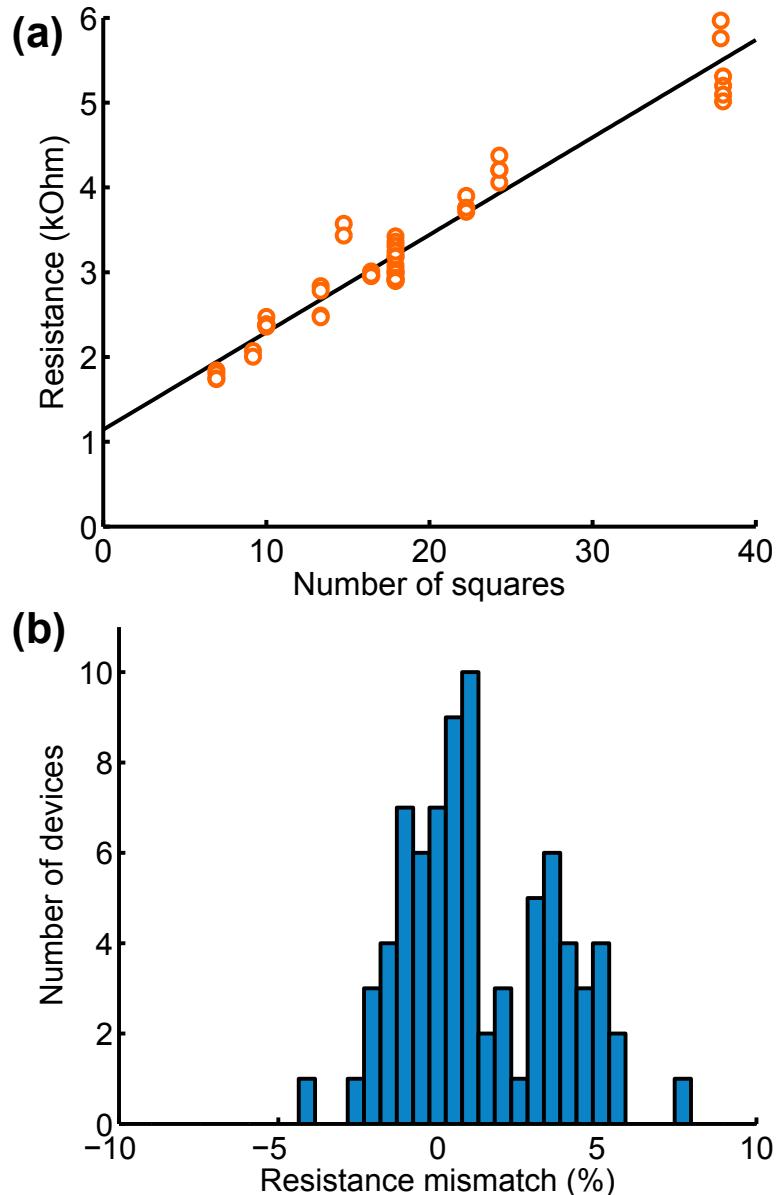


Figure F.12: Electrical resistance characteristics. (a) The transfer length method (TLM) was used to determine an average sheet resistance of $115 \Omega/\square$ and a constant resistance contribution of 9.9 squares ($1.1 \text{ k}\Omega$) from the silicon interconnects. The TLM-derived sheet resistance estimate agrees well with inline 4PP measurements. (b) Mismatch between the PR and TC resistors on a single silicon die was almost always less than 5% in magnitude.

TC and PR resistances on a single device were almost all less than 5% (Figure F.12b).

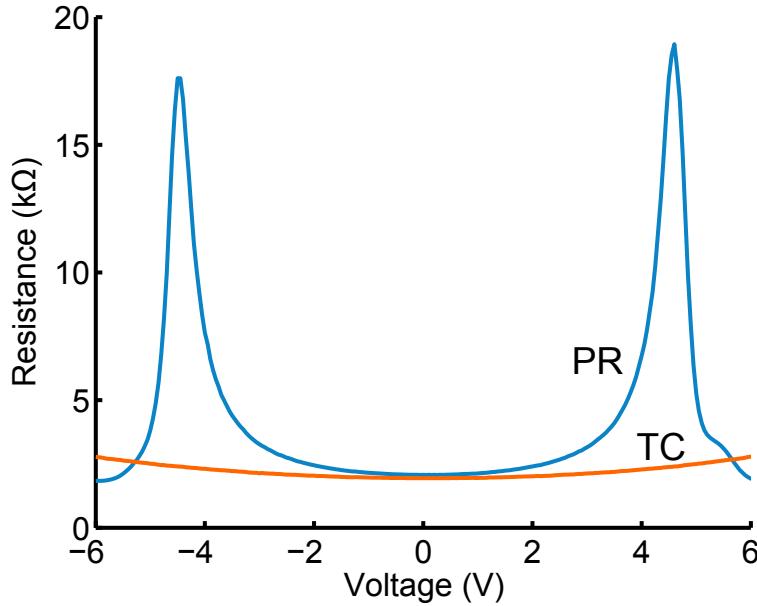


Figure F.13: Example of self-heating and thermal runaway in piezoresistive cantilevers. The resistance of both the PR and TC resistors increases quadratically for voltage biases of less than 2V. The much higher thermal resistance of the released PR resistor leads to its larger increase in temperature with power dissipation. Above 4 V the PR resistor temperature reaches 500-700°C and the concentration of thermally generated carriers becomes comparable to the dopant concentration, leading to a decrease in resistance with increasing power dissipation from that point on.

The resistance mismatch determines the resistor and potentiometer values to use in the measurement circuit.

The electrical resistance of the PR and TC resistors is plotted in Figure F.13 as a function of applied bias. For biases of 4 V or less the resistance of the PR resistor increases substantially over the TC resistor. The quadratic variation in resistance with voltage is the result of a linear TCR and quadratic variation in power dissipation with applied bias. Above 4 V the PR resistance reaches a peak and then decreases sharply. Once the piezoresistor reaches a temperature of 500-700°C the concentration of thermally generated carriers becomes comparable to the extrinsic dopant concentration leading to a decrease in resistance. Thermal runaway becomes a risk at such high operating temperatures due to positive feedback between the change in resistance with voltage ($\partial R / \partial V < 0$) and power dissipation with resistance ($\partial W / \partial R < 0$).

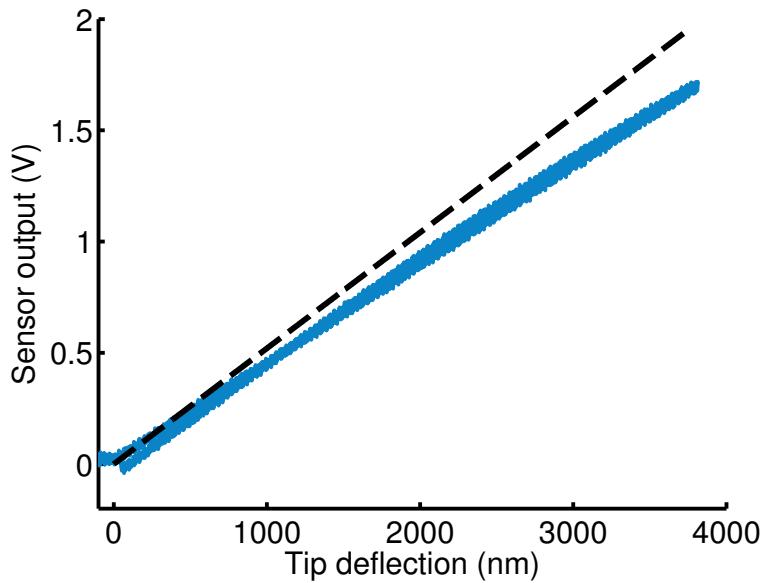


Figure F.14: Example displacement sensitivity from a 340 nm thick PR device. The measured (blue) and predicted (black) displacement sensitivities for a bridge bias of 2 V are shown. The predicted sensitivity matches the experimental data, particularly for small deflections. The reduction in sensitivity for large deflections is caused by geometric nonlinearity.

This data was the first indication that piezoresistor self-heating was substantial enough to be a problem for device operation. It also indicated that the TC resistor should be released from the silicon die in order to both electrically and thermally match the PR resistor. Most importantly, the data indicated that using a power dissipation rule of thumb was grossly inadequate, and that Joule heating should be considered in the design optimization process.

The displacement sensitivity of one of the PR devices is shown in Figure F.14. The sensitivity was measured by mounting the cantilever on high precision piezoelectric stage and deflecting the force probe with a stiff AFM cantilever. The predicted displacement sensitivity of the device agrees extremely well with the experimental results, at least for relatively long and narrow piezoresistors. We observed lower than expected sensitivities in shorter piezoresistors (as in Figure F.8b) due to the contribution from the transverse portion of the piezoresistor. The impact of transverse piezoresistance on sensitivity was incorporated into the design of the next generation probes.

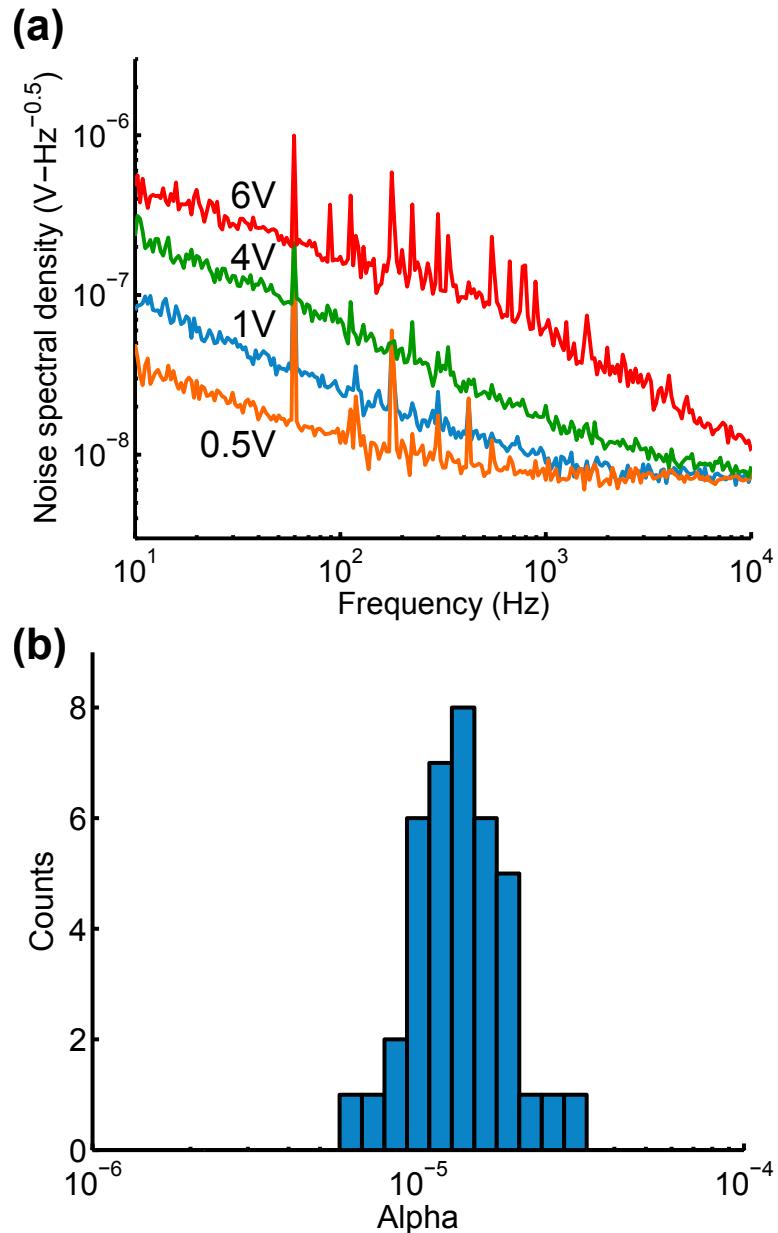


Figure F.15: Piezoresistor noise characteristics. (a) Example noise spectra from a single device at bridge biases ranging from 0.5 to 6 V. The piezoresistor shows ideal $1/f$ noise except at high bias values. The excess noise is caused by the large temperature difference between the piezoresistor and surrounding air. (b) We observed a range of Hooge factors with an average value of 10^{-5} .

The noise characteristics of the devices are shown in Figure F.15. Four example noise spectra from a single device are shown in Figure F.15a. The $1/f$ noise of the device behaves as expected (with $\alpha = 10^{-5}$) except at high power dissipation levels where a bump of excess noise is visible from 100 Hz to 10 kHz. Evidence suggests that the excess noise is from fluctuations in the piezoresistor temperature which become larger as the average temperature of the resistor increases. The excess low frequency noise at high power dissipations vanishes if the piezoresistors are not released from the wafer, as in Ref. [202]. We observed Hooge factors ranging from 5×10^{-6} to 4×10^{-5} in the first generation devices and did not notice any systematic variation in α with the piezoresistor geometry. Harley similarly observed Hooge factors of 10^{-5} on average for 30 nm thick epitaxial piezoresistors and did not observe any variation in α with the surface area to volume ratio or presence of a surface passivation oxide [36].

We emphasize that the predicted sensitivity and Hooge factor calculations were based upon SRA data. We also note that the first generation PR devices were originally designed for an epitaxial piezoresistor process. However, a leak in the epitaxial reactor led use to switch to POCl_3 predeposition midway through the process. Thus, the noise, sensitivity and SRA data obtained from the first generation PR devices was hugely important in improving our probe design methodology. The predeposition model and fitting parameters used later in design optimization (Section 2.5.2) were not available during this first run and we essentially fabricated the piezoresistors blind, simply trying to match the sheet resistance of the original epitaxial design.

We investigated several other aspects of piezoresistor noise using the first generation PR devices. First, we tried coating the devices in 200-300 nm of parylene N, immersing them in liquid, and observing their electrical noise characteristics. Figure F.16 plots noise spectra for a device in air, deionized water (DI) and phosphate buffered saline (PBS) for a bridge bias of 2 V. There is no substantial change in noise when the device is immersed in liquid besides the increased 60 Hz noise pickup due to the ungrounded liquid serving as an antenna. Grounding the liquid or placing it within a Faraday cage would mitigate the antenna effect. These initial noise data are important and verify that there is no variation in device noise between air and liquid that should be incorporated into the design process.

We also measured the noise of the same device when freely suspended and when

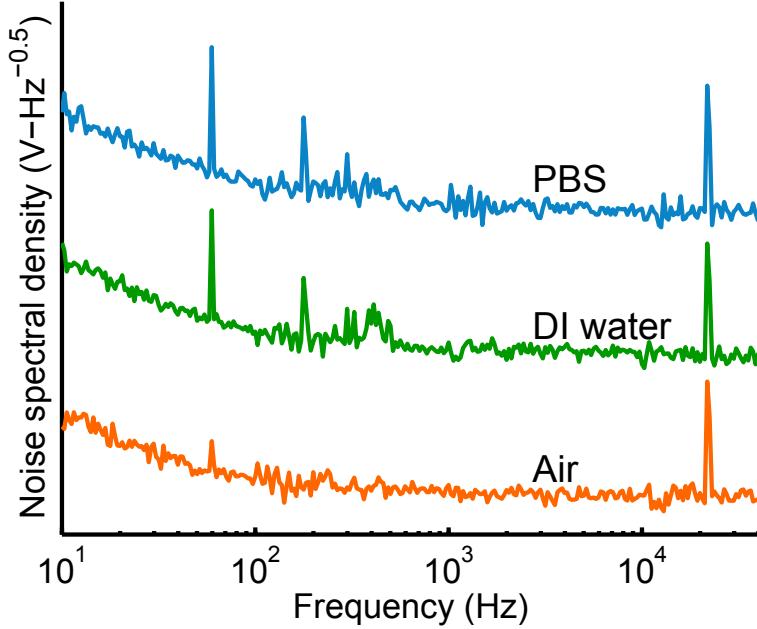


Figure F.16: Piezoresistor noise spectra in air, deionized water (DI) and phosphate buffered saline (PBS). The spectra were recorded for a single device with a bridge bias of 2 V and a 200-300 nm thick passivation coating of parylene N. No substantial difference in noise spectra between air and liquid is observed besides additional electromagnetic noise pickup at 60 Hz and higher harmonics.

brought into contact with a hard surface (Figure F.17). Relative movement between the device and the surface leads to a nearly 10-fold increase in the integrated noise between 10 and 600 Hz. The results highlight the importance of acoustic and mechanical isolation for achieving ideal sensor resolution when measuring atomic scale forces and displacements.

To summarize the PR device results: the die layout of the first generation PR devices prevented them from being used in hair cell experiments but we obtained a great deal of useful characterization data from them. We verified that the intended and actual mechanical properties of the devices closely matched. We determined that mismatch between the PR and TC resistors was small and that POCl_3 predeposition was a viable piezoresistor fabrication process. We learned that the piezoresistors could easily heat up by several hundred degrees without careful operation, and used the first generation devices in order to investigate piezoresistor self-heating [302]. The temperature modeling work from that effort was integrated into the design of the next generation probes. We measured noise

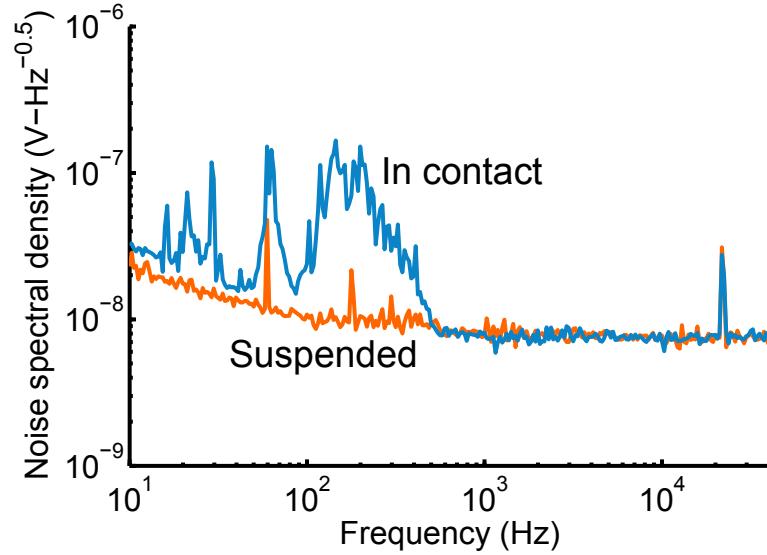


Figure F.17: Excess piezoresistor noise from displacement noise. The piezoresistor noise spectrum increases when the tip of the probe is brought into contact with a hard surface due to the relative movement between the probe and the surface. The results illustrate that acoustic and mechanical isolation of the measurement setup from the environment is essential in order to enable operation with ideal force and displacement resolution.

spectra for the devices in air and liquid, determining the average Hooge factor to expect for diffused piezoresistors. Finally, we verified that our model for piezoresistor sensitivity closely matched experimental results.

F.4.2 Piezoelectric-only cantilevers

Moving on to the PE process, SEMs of a finished device are shown in Figure F.18. The first PE devices were designed and fabricated with the sole purpose of determining the material properties of the AlN-based actuator and so the actuator stack covers most of the cantilever surface in order to maximize the induced deflections. The AlN etch, while wet rather than dry, is still fairly anisotropic due to a reduction in etch rate at the grain boundaries and the strong columnar orientation of the grains as seen in Figure F.18c. The stress in the AlN and Ti films is tensile (although we did not quantify it) and the devices bend slightly upwards. Most of the figures used to describe the PE device performance were originally presented in Ref. [355] and are reprinted here.

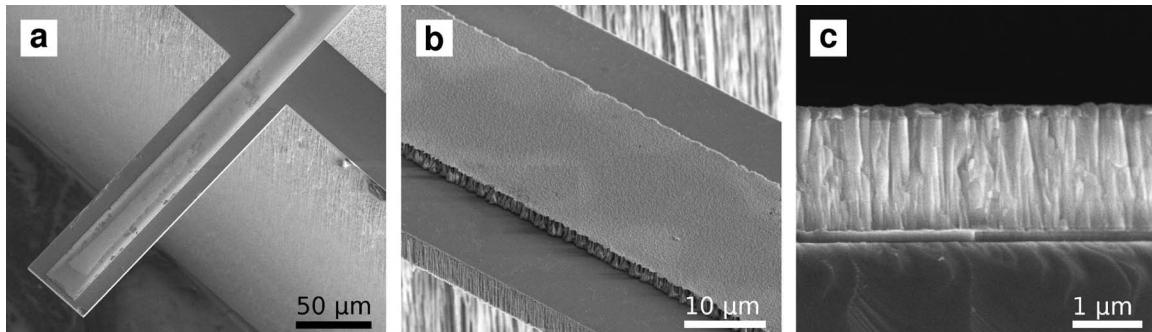


Figure F.18: SEMs of finished piezoelectric-only probes. (a) A released 5 μm thick, 50 μm wide and 200 μm long cantilever with a 30 μm wide and 196 μm long actuator. (b) Closeup of the piezoelectric stacking showing minimal undercut of the AlN by the TMAH etch. (c) Cross-sectional view of a piezoelectric stack with a 1700 nm thick AlN actuator illustrating the columnar orientation of the (002) oriented AlN grains. Reprinted from Ref. [355]. Copyright 2010 IOP.

The major issue in using AlN as a piezoelectric actuator is obtaining well aligned (002) oriented crystal grains. There is a massive body of literature on variations in grain orientation with process conditions [351,353,360,425,427,428]. Before fabricating the PE devices shown in Figure F.18 we spent some time developing the AlN deposition process in order to maximize grain orientation as measured by X-ray diffraction (XRD).

We first investigated the effect of the Si substrate surface condition on the AlN microstructure. Previous work indicated that Ti (100) rather than (002) is obtained when an oxidized Si substrate is used [429]. We found that removing the native SiO_2 layer in-situ via ICP improved the orientation of the Ti and AlN films in comparison with presputtered or untreated Si substrates (Figure F.19). All of the AlN films that we deposited exhibited pure (002) alignment without any other observable orientations.

As a brief aside, while we initially believed that removal of oxide layers was critical to obtain well aligned AlN, other researchers have obtained good AlN alignment even on layers of SiO_2 . We refer the interested reader to Felmetser et al. for details on a process for sub-100 nm thick AlN films on SiO_2 [358]. The AlN films in the second and third runs were deposited by Valeriy Felmetser of OEM Group.

Next we investigated the effect of AlN film thickness and ion energy on the degree of grain alignment. The rocking curve full width at half maximum (FWHM) depended

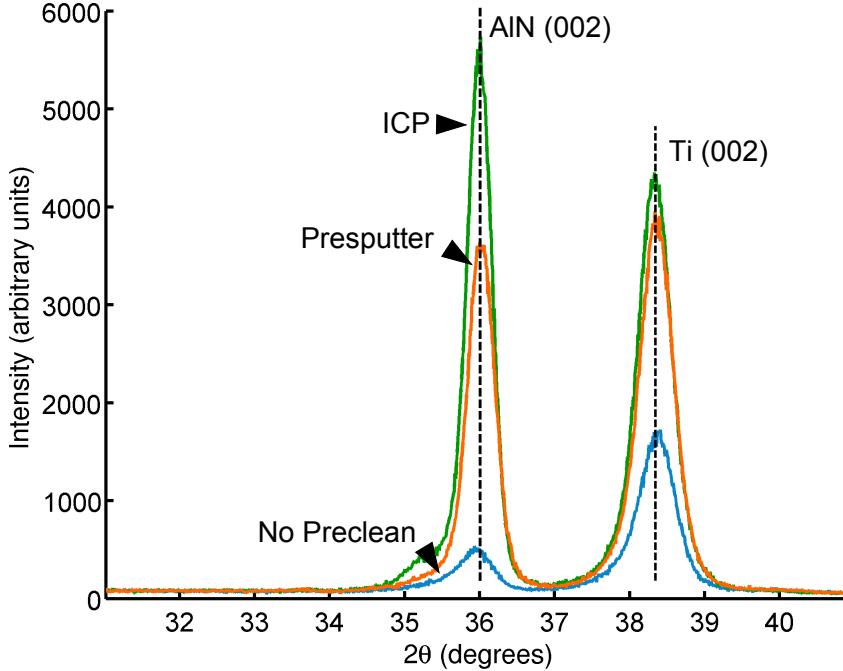


Figure F.19: AlN film texture characterization and variation with processing conditions. 100 nm of Ti and 500 nm of AlN were deposited on a Si (100) substrate in order to investigate the effect of removing the native SiO_2 layer in-situ by presputtering or an inductively coupled plasma (ICP). Precleaning the substrate in-situ increased the X-ray diffraction (XRD) intensity of the AlN and Ti (002) peaks. All cleans and depositions were performed in a single vacuum run to prevent the formation of SiO_2 or TiO_2 surface layers. Reprinted from Ref. [355]. Copyright 2010 IOP.

inversely upon the film thickness as expected (Figure F.20). Briefly, the $\Theta - 2\Theta$ scans in Figure F.19 show the relative quantity of certain crystallographic orientations while rocking curves measure the degree of alignment of a single crystallographic orientation. Increasing the RF-induced DC bias from 45 to 57 V modestly improved the FWHM while decreasing the bias energy from 45 V to 40 V almost doubled the FWHM. These values are slightly greater than previous reports of 2.6° for AlN on Ti with an AlN interlayer [352]. Although our rocking curve FWHM values are larger than reported values for Pt and Mo ($< 2^\circ$), Tonisch et al [430] found that a narrow rocking curve is far more important for resonator performance, which depends on the k_{31} coefficient, than generation actuation, which depends on the d_{31} coefficient.

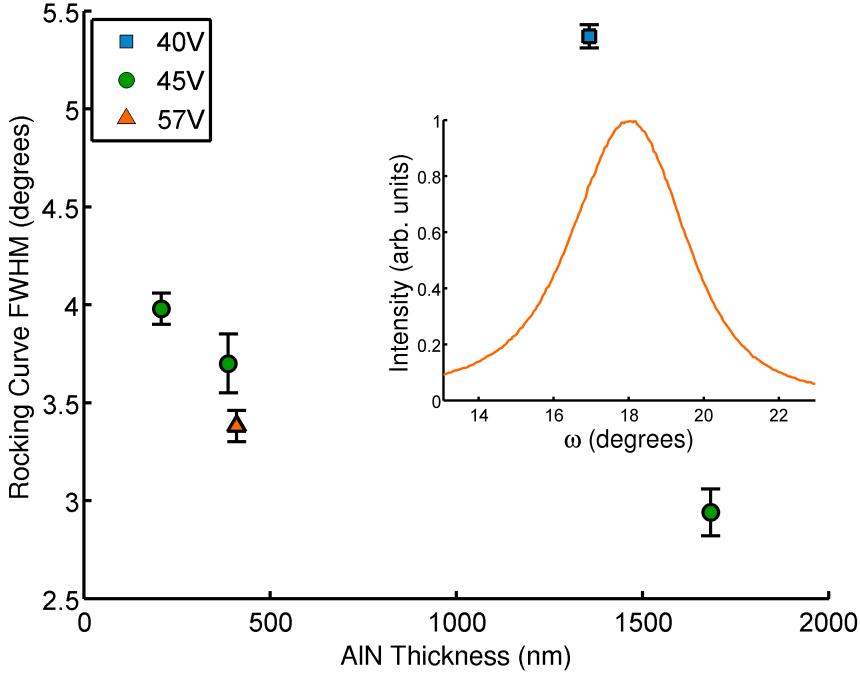


Figure F.20: The XRD rocking curve FWHM of the AlN was found to vary with film thickness and ion energy. All samples were cleaned by ICP before depositing the AlN/Ti/AlN/Ti film stack ($n = 5$ for each thickness). The orientation uniformity of the AlN (002) grains was improved by increasing both the film thickness and ion energy. We investigated the effect of RF-induced bias for several thicknesses and found that increased bias increased grain alignment. A sample rocking curve is inset (57 V bias). Reprinted from Ref. [355]. Copyright 2010 IOP.

However, film texture is an indirect measurement of piezoelectric properties because the formation of both parallel and anti-parallel grains can lead to a negligible piezoelectric response despite good apparent grain alignment [431]. For example, Ruffner et al. found that allowing the bottom metal electrode to oxidize before depositing the AlN film led to a change in the polarity and a reduction in the magnitude of the piezoelectric response [424]. Akiyama et al. observed a similar trend in the piezoelectric response by directly controlling the concentration of oxygen in the plasma during AlN deposition .

In light of the potential importance of oxygen concentration, we characterized the atomic composition of the AlN films using X-ray photoemission spectroscopy (XPS). We removed the top Ti electrode by Ar sputtering and measured the composition of the AlN

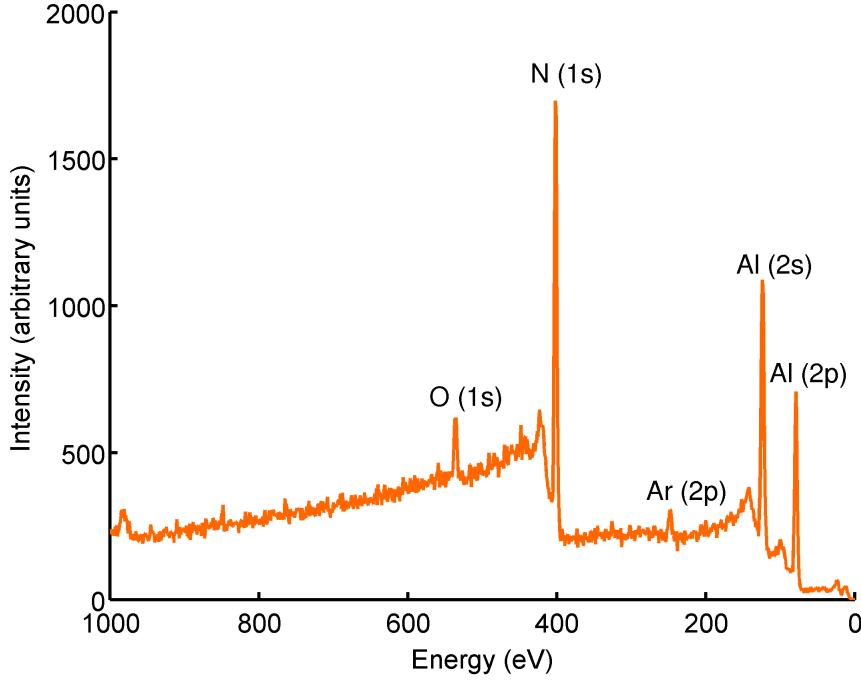


Figure F.21: Atomic composition of the AlN actuator. X-ray photoelectron spectroscopy (XPS) survey scans recorded after removing the top Ti electrode by sputtering. Six samples were measured. The atomic composition of the AlN layer was calculated to be O = 6.4% \pm 1.6%, N = 37.8% \pm 1.8%, Al = 54.0% \pm 2.2% ($\mu \pm \sigma$). Reprinted from Ref. [355]. Copyright 2010 IOP.

layer (Figure F.21). We calculated the composition of the AlN actuator layer as O = 6.4% \pm 1.6%, N = 37.8% \pm 1.8%, Al = 54.0% \pm 2.2% ($\mu \pm \sigma$) based upon six samples. The measured oxygen concentration agrees well with the concentration for peak piezoelectric response observed by Akiyama et al. in Ref. [432].

One interpretation of these results is that oxygen affects the AlN piezoelectric response by two mechanisms: substrate oxidation and bulk oxidation. At low concentration when the substrate is not uniformly oxidized, the piezoelectric response decreases and can vanish due to the mixture of + and - grain polarities. At high concentration, the piezoelectric response is reduced by oxygen-induced grain defects. XRD rocking curve measurements will be affected by the second effect but not the first. In general, grain polarity is a complex function of many parameters including surface condition and a polarity change

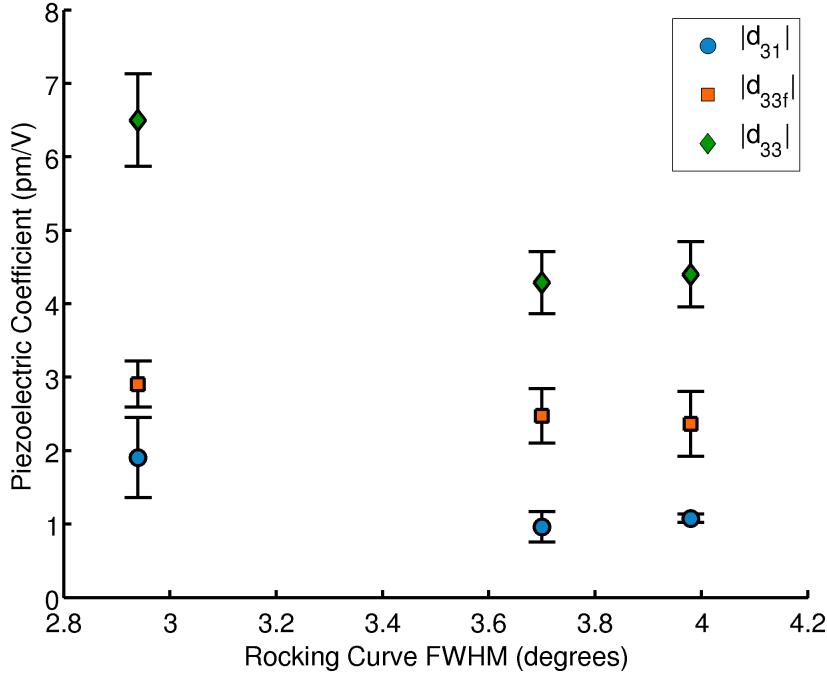


Figure F.22: Variation in the piezoelectric coefficients with rocking curve FWHM. The d_{33f} (21 measurements, 3 wafers) and d_{31} coefficients (264 measurements, 8 cantilevers) were used to compute the d_{33} coefficients. The AlN films were deposited with an RF-induced bias of 45 V. The d_{33} coefficient is calculated from d_{33f} , d_{31} and the AlN material properties. Both d_{33f} and d_{33} are positive while d_{31} is negative. Reprinted from Ref. [355]. Copyright 2010 IOP.

can be induced by varying the deposition pressure [428] or power [427]. Accordingly, it is necessary to directly measure both the sign and magnitude of the piezoelectric response rather than try to infer them from the film structure. The film morphology of AlN on Ti has been reported [429, 433], but the piezoelectric response has been less frequently measured [352, 423].

We measured the piezoelectric properties of the films we deposited by two separate methods: d_{33f} at the wafer scale and d_{31} using microfabricated cantilevers. The thin film longitudinal piezoelectric coefficient (d_{33f}) is smaller than the actual longitudinal piezoelectric coefficient of the material (d_{33}) due to the clamping effect of the substrate that the film is deposited on. The two are related by the mechanical properties of the AlN film

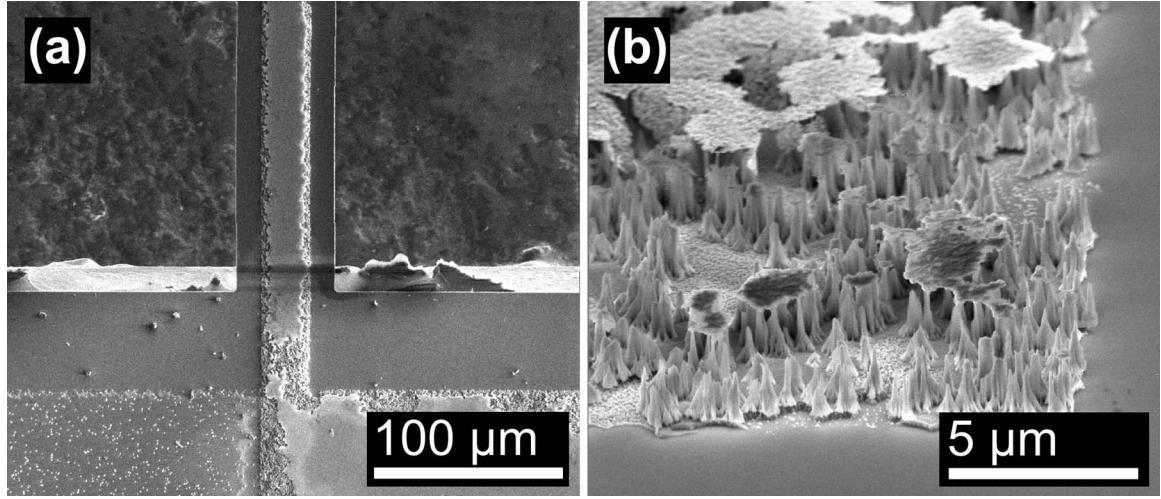


Figure F.23: Issues with wet etching AlN. (a) Mask undercut and micromasking in the wet Ti and AlN etch steps is frequency. On this device the top Ti electrode has been undercut by 10-20 μm disconnecting the metal electrode on the cantilever itself from the bondpad. The dots visible on the bottom electrode bondpad (left) are micromasked AlN grains. (b) Detailed view of the the undercut top electrode. Moving to a dry etch process in the second and third fabrication runs eliminated these etch control issues.

and the d_{31} coefficient. More details are available in Ref. [355]. We measured both the d_{33f} and d_{31} coefficients via laser doppler vibrometry (LDV) and used the combination of the two and assumed mechanical properties to calculate d_{33} . All three coefficients are presented in Figure F.22 as a function of the rocking curve FWHM. The piezoelectric coefficients that we measured were greater than the best previously reported for AlN on Ti ($d_{33f} = 2.3 \text{ pm/V}$ in Ref. [423]) and comparable to the best data to date for AlN on Pt ($d_{33} = 6.8 \text{ pm/V}$ in Ref. [360]).

We did encounter several issues with the AlN on Ti process though. First, both the Ti and AlN layers were wet etched and we encountered issues with mask undercut and micromasking. In Figure F.23a we see that the top Ti electrode on the cantilever has been electrically isolated from the top Ti bondpad due to 10-20 μm of mask undercut. Mask undercut can come from poor adhesion of the photoresist to the stack materials or, more often, from poor coverage over the actuator sidewalls after the bottom PE lithography step. The poor critical dimension (CD) control of the wet etch steps set the minimum feature size of this process to 30-40 μm , far larger than would have been ideal. A closer view of the

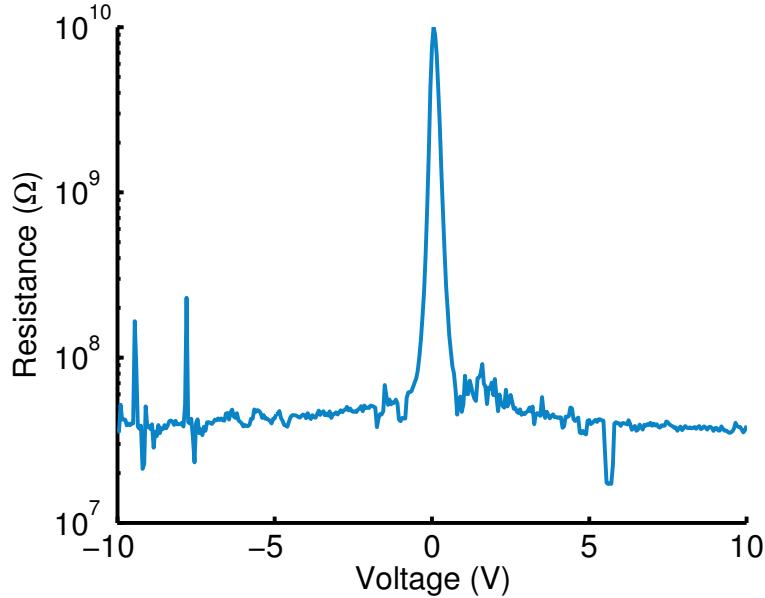


Figure F.24: Electrical impedance of an AlN actuator. The high impedance ($> 10^7 \Omega$) indicates that the wirebond did not puncture the top Ti bondpad. One devices where the device was shorted out, no actuation was possible and the electrical impedance was on the order of 20Ω .

etch nonuniformity and mask undercut is seen in Figure F.23b. Clearly the wet etch process leaves something to be desired, and it was one of the main motivations for switching to a Mo-based process for the second and third fabrication runs.

A second issue with the PE devices was wirebonding to the thin Ti electrodes. The resistance of the AlN actuator is plotted as a function of voltage in Figure F.24. The high impedance of the actuator ($> 10^7 \Omega$) indicates good electrical contact between the wirebond and bondpad. On many devices, the wirebond actually broke through the top Ti electrode making contact with the bottom electrode. In that case the actuator is shorted out ($20\text{-}50 \Omega$) and no electrical potential develops across the actuator. The tendency to break through the top electrode varied between the wirebonders we used, with one breaking through more than 90% of the time while another breaking through less than 20% of the time. We mitigated this issue by adjusting the wirebonding settings, but it would come up again in the second and third fabrication runs.

To summarize the PE device results: we first developed a deposition process to obtain

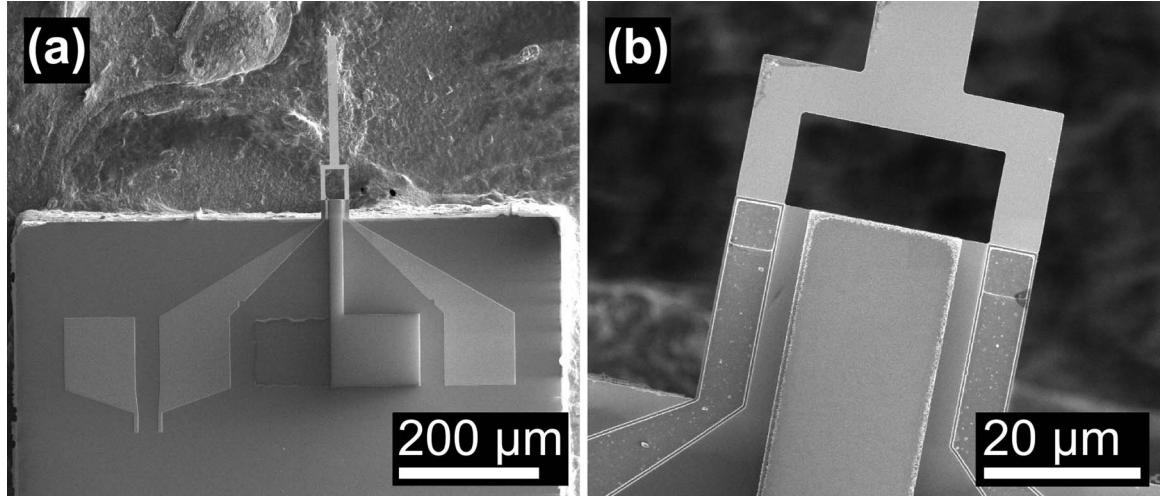


Figure F.25: SEMs of finished piezoresistively sensed and piezoelectrically actuated probes. (a) A top down view of the overall die layout and bondpads. The TC resistor is not visible due to the low contrast between the n-type piezoresistor and p-type background. (b) A closeup view of the base of the probe, with the piezoelectric actuator located in the center of the cantilever and two aluminum lines running on either side. Note the long transverse portion of the piezoresistor, which degrades performance and led to curvature problems.

good piezoelectric performance from AlN on Ti. Next, we fabricated PE cantilevers and verified that functional cantilevers were possible. Finally, we characterized the composition, structure and piezoelectric properties of the AlN actuators using XRD, XPS and LDV. AlN on Ti is a viable process flow for applications that require piezoelectric actuation but where conventional metal electrodes (e.g. Mo or Pt) are not an option.

F.4.3 Piezoelectrically actuated probes

Moving on to the PRPE process, finished devices are shown in Figure F.25. The die layout is roughly the same as the PR and PE devices, with the probe located near the center of the die. There are three bondpads for the PR sensor and two for the PE actuator. A closeup of the cantilever is seen in Figure F.25b, illustrating the aluminum interconnects and vias used to make electrical contact with the piezoresistor. Note that the transverse portion of the piezoresistor is roughly as long as the longitudinal portions. Although the devices in Figure F.25 are intact and in good condition, there were several flaws with the PRPE device

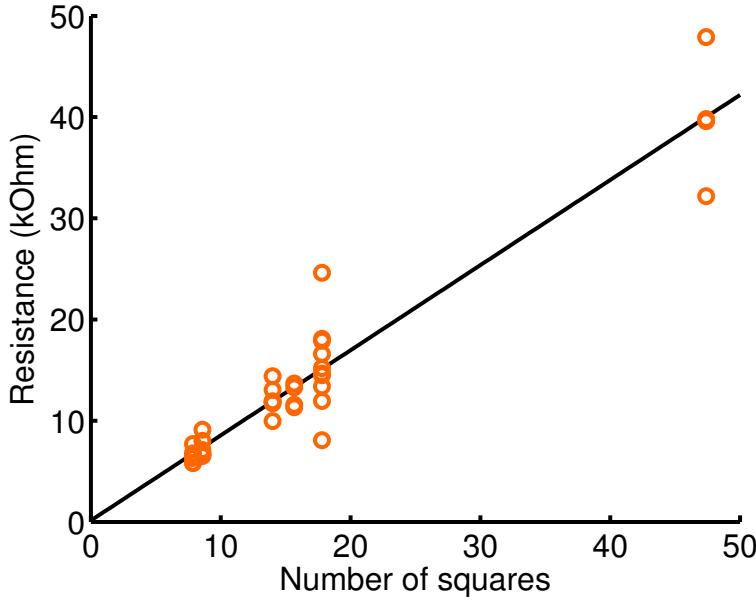


Figure F.26: Damage to the piezoresistor surface increased the electrical sheet resistance from 115 to 840 Ω/\square . The piezoresistors were damaged at two points in the process: during the in-situ AlN preclean and the RIE etchback of the LTO.

design.

The first problem was the placement of the cantilever in the center of the die. As noted at the beginning of this section, this die layout is impossible to integrate with an upright microscope for cochlear hair cell experiments (Figure 1.9).

The second major problem was piezoresistor damage during processing. Two steps of the process removed material from the surface of the piezoresistor. First, material was removed during the in-situ ICP cleaning immediately before the AlN stack deposition. Second, when the exposed LTO was etched back after the patterning of the Al interconnects and bondpads the RIE oxide etch stopped directly on the piezoresistor. The combination of the two increased the sheet resistance of the piezoresistor from 115 to 840 Ω/\square (Figure F.26). The sheet resistance increase corresponds to the removal of roughly 800 Å of material from the piezoresistor surface. On the bright side, the excess resistance in the system decreases from 9.9 squares to 0.2 squares through the use of metal interconnects.

The surface damage had several detrimental effects. First, it reduced the thickness of the force probe by approximately 80 nm, reducing the resonant frequency, stiffness,

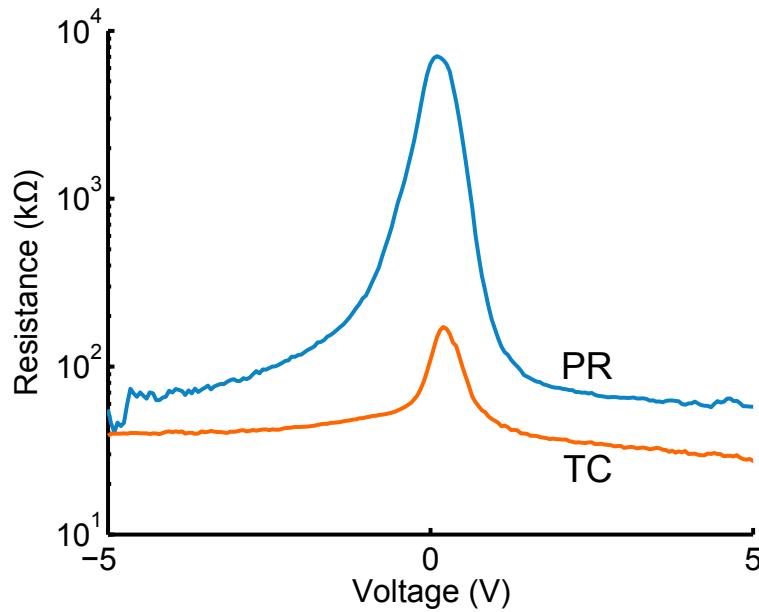


Figure F.27: Piezoresistor damage reduces the surface concentration yielding non-ohmic contacts. The resistance increases nearly 10-fold for voltage biases of less than 1 V. The resistance-voltage curve is symmetric because one of the Al-Si contacts is always forward biased while the other is reverse biased regardless of the voltage polarity.

and displacement sensitivity of the sensor. Second, it increased the Johnson noise of the piezoresistor roughly 3-fold due to the increased resistance. Third, it increased the 1/f noise of the piezoresistor nearly 10-fold due to the reduced number of carriers and increased bridge bias required to operate at the same power dissipation and temperature increase. Finally, the damage reduced the dopant concentration at the surface of the piezoresistor which yielded Schottky rather than ohmic contacts (Figure F.27). The resistance increased nearly 10-fold at low voltage biases due to the non-ideal contacts.

The third major problem with the PRPE design was mechanical hysteresis and nonlinearity. As we saw earlier there is a substantial strain gradient in the piezoresistor which curves the device (Figure F.8c). When combined with the long transverse section of the piezoresistor in the PRPE design, the surface stress caused the probe to adopt the shape of a potato chip. Figure F.28 illustrates the transverse curvature seen in most of the PRPE devices. As seen in Figure F.28b, the curvature begins at the contact vias, indicating that it is caused by the tensile surface stress of the piezoresistor.

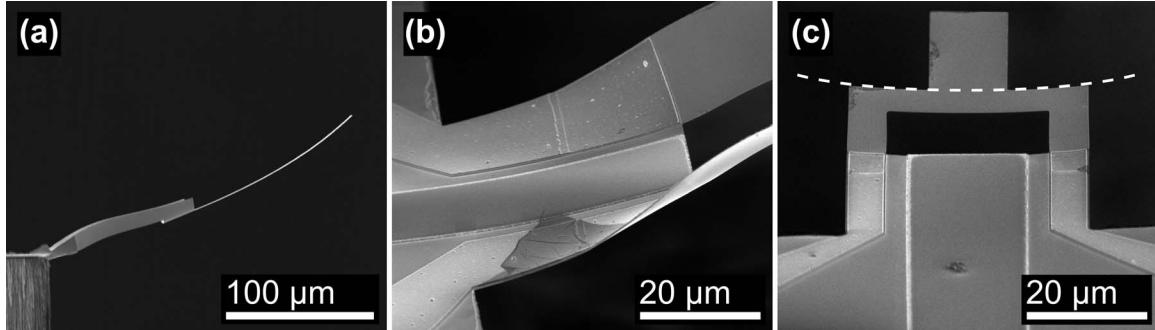


Figure F.28: Surface stress and potato chipping in the first generation PRPE devices. (a) Longitudinal and transverse bending is seen in this side view of a released probe. (b) The tensile surface stress in the piezoresistor causes the probe to adopt a potato chip-like configuration. (c) A view along the length of a released probe. The dashed line highlights the transverse curvature.

In the case of the PR devices, curvature did not affect device performance for the most part. We were not so fortunate in the case of the PRPE devices. The devices adopted a potato chip conformation due to the combination of transverse and longitudinal bending, which resulted in hysteresis and nonlinearity in the mechanical response of the probe. The effect is seen in one of the PRPE devices in Figure F.29. The piezoresistor sensor output is plotted as a function of the tip deflection of the probe for three different bridge bias voltages. The sensor output is proportional to the bending in the piezoresistor portion of the device. For deflections of less than 5 μm most of the bending is confined to the region of the probe beyond the piezoresistor. For a 5 μm tip deflection the piezoresistor suddenly flips, yielding a very high displacement sensitivity. Beyond that point most of the deflection is confined to the region of the probe beyond the piezoresistor once again.

The same effect can be seen at the macroscale by taking a playing card, curling it parallel to its long axis and then applying a force near one end. The card will be very stiff up to a critical force, at which point it will become unstable and suddenly flip.

The final issue with the PRPE devices was their poor actuator-sensor crosstalk performance. Crosstalk was not even a design consideration for the first generation PRPE devices and it shows. No attempt was made to introduce identical capacitive and mechanical crosstalk signals into both the PR and TC resistors. The frequency and time responses of the actuator-sensor crosstalk are shown in Figure F.30. The capacitive

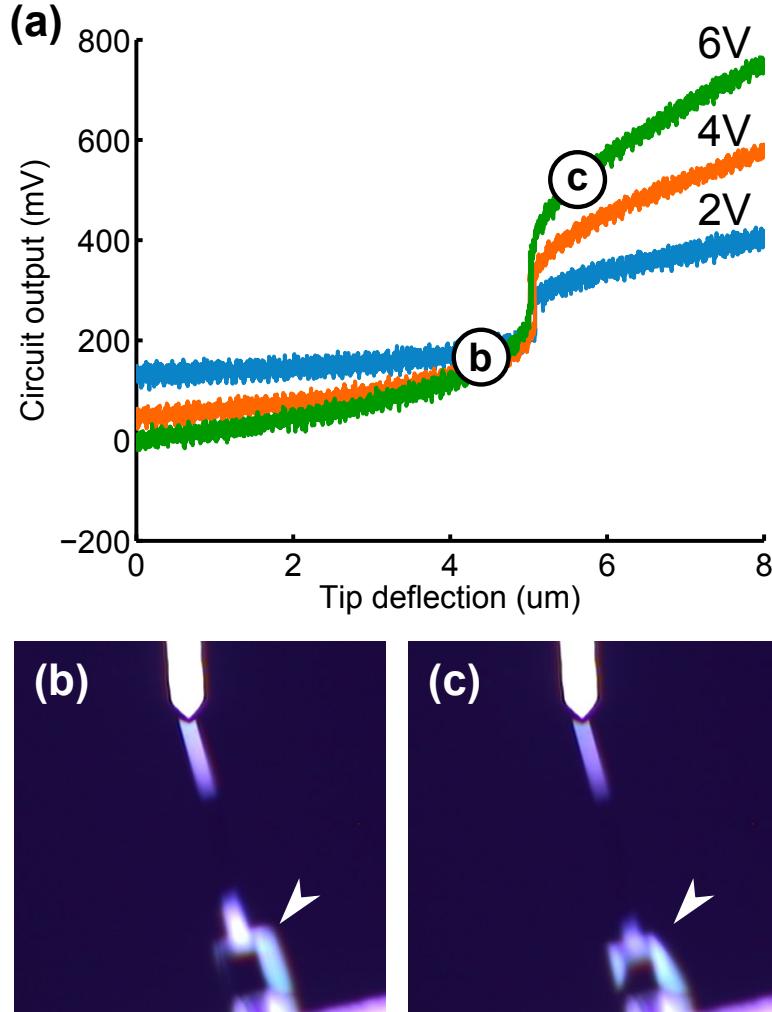


Figure F.29: Mechanical nonlinearity caused by transverse bending. (a) The piezoresistor sensor output is plotted as a function of the probe tip deflection for three different bridge bias voltages. The displacement sensitivity of the piezoresistor is low except near $5\text{ }\mu\text{m}$ of tip deflection. Micrographs of the device (b) before and (c) after the device flips are shown. For deflections less than and greater than $5\text{ }\mu\text{m}$ the cantilever beyond the end of the piezoresistor is bending rather than the piezoresistor itself.

crosstalk increases nearly linearly with frequency as expected and is 20-40 dB worse than the PRPE devices published to date in Refs. [57] and [65]. The time domain crosstalk response illustrates the enormous capacitive crosstalk and smaller mechanical crosstalk as illustrated by the 250 mV DC offset in sensor response depending on the actuator position. Ideally the PRPE device would have crosstalk of -120 dB up to 100 kHz. In that case, a

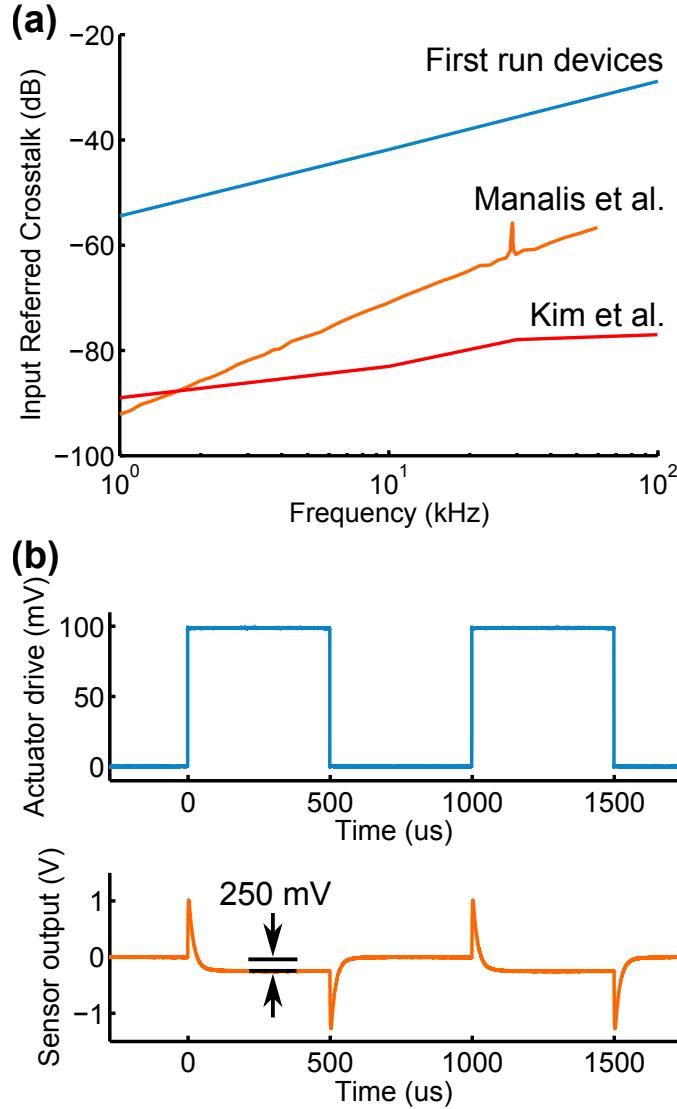


Figure F.30: Frequency and time actuator-sensor crosstalk responses in the first generation PRPE devices. (a) The capacitive crosstalk is 20-40 dB worse than prior work by Manalis [57] and Kim [65]. (b) The time domain response of the sensor in response to a 100 mV actuator signal illustrates the large capacitive (AC) and mechanical (DC) crosstalk signals.

1V actuator signal would introduce a 1 μ V crosstalk signal at the instrumentation amplifier input, comparable to the noise floor of the device.

In summary, the first generation PRPE devices had several crippling flaws. The first flaw was the die layout which rendered the probes incompatible with the cochlear hair

cell experimental setup. The second flaw was the piezoresistor surface damage. The third flaw was the potato chip effect caused by the combination of a tensile surface stress and wide piezoresistor. The fourth and final flaw was the large actuator-sensor crosstalk. An additional flaw seen in all of the devices was the poor alignment between the base of the probe and the backside release etch. The probe layouts and processes in the second fabrication run were developed in order to address these shortcomings.

But the first fabrication run yielded a great deal of useful information. The PR noise, sensitivity and self-heating data was used to refine the design optimization models. For example, the impact of the transverse portion of the piezoresistive loop was added to the models. The processes also illustrated that it was possible in principle to fabricate force probes with the force resolution and actuator performance needed for the cochlear hair cell application.

For completeness, the probe designs that were fabricated during the first run are summarized in Tables F.3, F.4 and F.5.

Table F.3: Summary of the first generation PR probe designs. For each design the overall cantilever length (l_c), overall cantilever width (w_c), piezoresistor length (l_{pr}) and air gap width (w_g) are specified. The piezoresistor width (w_{pr}) is equal to $w_c/2$ in all cases. The spring constant (k_c) and undamped resonant frequency (f_0) are calculated for a 340 nm thick cantilever.

Design #	l_c (μm)	w_c (μm)	l_{pr} (μm)	w_g (μm)	k_c (N/m)	f_0 (kHz)
1	123	6	34	5	4.1×10^{-3}	27.1
2	61	6	15	5	3.4×10^{-2}	110
3	870	6	212	5	1.2×10^{-5}	0.55
4	87	6	23	5	1.2×10^{-2}	54.2
5	890	10	212	5	1.8×10^{-5}	0.52
6	890	10	104	5	1.8×10^{-5}	0.52
7	178	10	43	5	2.3×10^{-3}	12.9
8	124	10	32	5	6.7×10^{-3}	26.7
9	88	10	24	5	1.9×10^{-2}	52.9
10	859	20	205	5	4.0×10^{-5}	0.56
11	859	20	290	5	4.0×10^{-5}	0.56
12	242	20	60	5	1.8×10^{-3}	7.00
13	172	20	45	5	5.0×10^{-3}	13.9
14	385	20	64	5	4.5×10^{-4}	2.77
15	281	10	64	5	5.8×10^{-4}	5.19
16	90	6	53	5	1.1×10^{-2}	50.6
17	121	20	34	5	1.4×10^{-2}	28.0
18	74	2	16	1	6.3×10^{-3}	74.9
19	47	2	25	1	2.5×10^{-2}	186
20	286	2	115	2	1.1×10^{-4}	5.00
21	64	2	25	2	9.8×10^{-3}	100
22	41	2	15	2	3.9×10^{-2}	250
23	29	2	10	2	1.1×10^{-1}	501
24	20	2	7	2	3.1×10^{-1}	1000
25	640	2	41	2	9.7×10^{-6}	1.00
26	64	2	9	2	9.7×10^{-3}	100
27	20	2	3	2	3.1×10^{-1}	1000
28	155	2	64	2	6.8×10^{-4}	17.0
29	35	2	16	2	6.1×10^{-2}	340
30	22	2	9	2	2.4×10^{-1}	847
31	16	2	6	2	6.9×10^{-1}	1710
32	11	2	4	2	1.9	3390
33	358	2	74	2	5.6×10^{-5}	3.19
34	35	2	6	2	6.0×10^{-2}	335
35	11	2	1.8	2	1.9	3390

Table F.4: Summary of the first generation piezoelectric actuator-only probe designs. For each design the overall cantilever length (l_c), overall cantilever width (w_c), piezoelectric actuator length (l_{pe}) and piezoelectric actuator width (w_{pe}) are specified. These probes were designed in order to characterize the first generation piezoelectric actuator process (AlN on Ti).

Design #	l_c (μm)	w_c (μm)	l_{pe} (μm)	w_{pe} (μm)
1	100	30	96	26
2	200	30	30	30
3	200	50	26	26
4	200	30	10	10

Table F.5: Summary of the first generation piezoresistively sensed and piezoelectrically actuated probe designs. For each design the length of the piezoresistive sensor (l_c), overall sensor width (w_c), piezoresistor length (l_{pr}), piezoelectric actuator length (l_{pe}) and piezoelectric actuator width (w_{pe}) are specified. The air gap width between the piezoresistor legs (w_g) is 40 μm for all of the designs. The piezoresistor width (w_{pr}) is equal to $w_c/2$ in all cases. The spring constant (k_c) and undamped resonant frequency (f_0) are calculated for a 340 nm thick cantilever.

Design #	l_c (μm)	w_c (μm)	l_{pr} (μm)	l_{pe} (μm)	w_{pe} (μm)	k_c (N/m)	f_0 (kHz)
1	800	20	212	40	40	5.0×10^{-5}	0.64
2	385	20	64	40	40	4.5×10^{-4}	2.78
3	172	20	45	40	40	5.0×10^{-3}	13.94
4	61	20	18	40	40	1.1×10^{-1}	111
5	600	60	20	60	60	3.6×10^{-4}	1.15
6	281	60	83	60	60	3.5×10^{-3}	5.22

Appendix G

Ion implantation lookup tables

The following lookup tables can be used to calculate β_1^* , β_2^* , N_z , R_s and t_j for boron, phosphorus and arsenic ion implantation processes.

In all of the tables, the first number corresponds to a process with only an inert anneal while the second number (in parentheses) corresponds to a process with a 1500 Å wet passivation oxide grown immediately before the N₂ anneal. The wet passivation oxide was grown at the indicated process temperature for 66, 15, or 5 minutes at 900, 1000, or 1100°C, respectively. Junction depths are calculated assuming a background resistivity of 10 Ω-cm.

Table G.1: Look-up table for β_1^* (-) and β_2^* (μm) for boron ion implantation.

		Time (min)													
		Time (min)													
		Time (min)													
Dose (cm^{-2})	Energy (keV)	β_1^*	15	30	45	60	75	β_1^*	β_2^*	β_1^*	β_2^*	β_1^*	β_2^*		
900	.75(.82)	.12(.18)	.75(.82)	.12(.18)	.75(.82)	.12(.18)	.75(.82)	.12(.18)	.75(.82)	.12(.18)	.75(.82)	.12(.19)	.75(.82)	.12(.19)	
20	1000	.73(.83)	.12(.23)	.73(.84)	.13(.24)	.74(.84)	.14(.24)	.74(.84)	.14(.25)	.74(.84)	.15(.25)	.74(.84)	.16(.26)	.75(.84)	.17(.27)
1100	.76(.85)	.22(.37)	.77(.86)	.26(.40)	.78(.86)	.29(.42)	.78(.86)	.33(.45)	.79(.86)	.36(.47)	.79(.86)	.38(.49)	.79(.87)	.41(.51)	.80(.87)
2×10^{14}	50	.77(.80)	.22(.26)	.77(.80)	.22(.26)	.77(.80)	.22(.26)	.77(.80)	.22(.26)	.77(.80)	.22(.26)	.77(.80)	.22(.26)	.77(.80)	.22(.26)
1100	.77(.83)	.28(.42)	.78(.84)	.31(.44)	.78(.84)	.34(.47)	.79(.84)	.37(.49)	.79(.84)	.40(.51)	.79(.85)	.42(.52)	.80(.85)	.44(.54)	.80(.85)
900	.77(.80)	.28(.31)	.77(.80)	.28(.31)	.77(.80)	.28(.31)	.77(.80)	.28(.31)	.77(.80)	.28(.31)	.77(.80)	.28(.31)	.77(.80)	.28(.31)	
80	1000	.77(.81)	.28(.34)	.77(.81)	.28(.34)	.77(.81)	.29(.35)	.77(.81)	.29(.35)	.77(.81)	.29(.36)	.78(.81)	.30(.36)	.78(.81)	.30(.37)
1100	.78(.83)	.33(.45)	.79(.83)	.36(.48)	.79(.83)	.39(.50)	.79(.83)	.41(.52)	.80(.84)	.44(.54)	.80(.84)	.46(.55)	.80(.84)	.48(.57)	.80(.84)
900	.47(.67)	.05(.16)	.47(.67)	.06(.16)	.48(.68)	.06(.16)	.49(.68)	.07(.16)	.50(.68)	.08(.16)	.51(.68)	.08(.16)	.52(.68)	.09(.16)	.52(.68)
20	1000	.53(.66)	.12(.23)	.54(.66)	.14(.24)	.55(.67)	.16(.24)	.55(.67)	.17(.25)	.56(.67)	.18(.26)	.56(.67)	.19(.26)	.57(.67)	.20(.27)
1100	.58(.69)	.25(.38)	.59(.69)	.29(.40)	.60(.70)	.33(.43)	.61(.70)	.36(.45)	.62(.71)	.40(.48)	.62(.71)	.42(.50)	.63(.71)	.45(.52)	.63(.71)
900	.55(.65)	.17(.25)	.55(.65)	.17(.25)	.56(.65)	.17(.25)	.56(.65)	.17(.25)	.56(.65)	.18(.26)	.56(.65)	.18(.26)	.57(.65)	.18(.26)	.57(.65)
20	1000	.57(.65)	.19(.29)	.57(.66)	.20(.29)	.57(.66)	.21(.30)	.57(.66)	.21(.30)	.58(.66)	.22(.31)	.58(.66)	.23(.31)	.58(.67)	.24(.32)
1100	.59(.68)	.27(.40)	.60(.68)	.31(.43)	.61(.69)	.35(.46)	.62(.69)	.38(.48)	.62(.69)	.41(.50)	.63(.70)	.44(.52)	.63(.70)	.46(.54)	.64(.70)
2×10^{15}	50	1000	.60(.65)	.27(.34)	.60(.65)	.27(.34)	.60(.65)	.27(.34)	.60(.65)	.27(.34)	.60(.66)	.27(.34)	.60(.66)	.27(.34)	.60(.66)
80	1000	.59(.66)	.26(.35)	.59(.66)	.27(.36)	.59(.66)	.27(.36)	.60(.66)	.28(.37)	.60(.66)	.29(.37)	.60(.66)	.30(.38)	.60(.66)	.30(.38)
1100	.60(.67)	.32(.45)	.61(.68)	.35(.47)	.62(.68)	.38(.50)	.62(.69)	.41(.52)	.63(.69)	.44(.54)	.63(.69)	.46(.56)	.64(.69)	.49(.57)	.64(.70)
900	.48(.51)	.09(.12)	.48(.51)	.10(.12)	.48(.50)	.10(.12)	.48(.50)	.10(.13)	.48(.50)	.11(.13)	.48(.51)	.11(.13)	.48(.51)	.11(.14)	.48(.51)
20	1000	.42(.45)	.14(.20)	.41(.46)	.17(.21)	.41(.46)	.19(.23)	.41(.47)	.21(.25)	.41(.47)	.22(.26)	.41(.47)	.24(.27)	.41(.48)	.25(.28)
1100	.38(.47)	.30(.39)	.40(.48)	.37(.44)	.41(.49)	.43(.47)	.42(.49)	.48(.51)	.43(.50)	.52(.54)	.44(.50)	.56(.57)	.45(.51)	.59(.60)	.45(.51)
900	.48(.50)	.17(.18)	.48(.50)	.17(.18)	.48(.50)	.17(.18)	.48(.50)	.17(.18)	.48(.50)	.17(.18)	.48(.50)	.17(.18)	.48(.50)	.18(.18)	.48(.50)
20	1000	.40(.42)	.17(.20)	.40(.42)	.18(.22)	.40(.42)	.20(.23)	.40(.42)	.22(.25)	.40(.42)	.24(.26)	.40(.42)	.25(.28)	.40(.42)	.27(.29)
1100	.38(.45)	.32(.41)	.40(.46)	.39(.45)	.41(.47)	.45(.49)	.42(.47)	.49(.53)	.43(.48)	.53(.56)	.44(.48)	.57(.59)	.44(.49)	.61(.62)	.45(.49)
2×10^{16}	50	1000	.49(.50)	.26(.27)	.48(.50)	.26(.28)	.48(.51)	.26(.28)	.48(.51)	.26(.28)	.48(.50)	.26(.28)	.48(.50)	.26(.28)	.48(.50)
80	1000	.40(.41)	.21(.23)	.40(.41)	.22(.24)	.40(.41)	.24(.25)	.40(.41)	.25(.27)	.40(.41)	.26(.28)	.40(.41)	.29(.30)	.40(.41)	.30(.32)
1100	.39(.45)	.34(.42)	.40(.46)	.41(.47)	.42(.46)	.46(.51)	.42(.47)	.51(.54)	.43(.47)	.55(.57)	.44(.48)	.58(.60)	.44(.48)	.62(.63)	.45(.49)

APPENDIX G. ION IMPLANTATION LOOKUP TABLES

Table G.2: Look-up table for β_1^* (-) and β_2^* (μm) for phosphorus ion implantation.

		Time (min)															
Dose (cm^{-2})	Energy (keV)	Temp. ($^\circ\text{C}$)	15	30	45	60	75	90	105	120	β_1^*	β_2^*	β_1^*				
900	.62(.80)	.03(.09)	.63(.80)	.03(.09)	.63(.80)	.04(.09)	.64(.80)	.04(.09)	.64(.80)	.04(.09)	.65(.80)	.04(.09)	.65(.80)	.04(.10)			
20	1000	.68(.86)	.07(.16)	.70(.87)	.08(.17)	.71(.87)	.10(.18)	.72(.87)	.11(.18)	.73(.87)	.12(.19)	.73(.87)	.13(.19)	.73(.87)	.13(.20)		
1100	.76(.91)	.20(.32)	.78(.91)	.25(.35)	.78(.91)	.29(.37)	.79(.92)	.32(.40)	.80(.92)	.36(.42)	.80(.92)	.39(.44)	.81(.92)	.41(.46)	.81(.92)	.44(.48)	
900	.65(.73)	.06(.10)	.65(.73)	.06(.10)	.65(.73)	.06(.10)	.65(.73)	.06(.10)	.65(.73)	.06(.10)	.65(.73)	.07(.10)	.65(.73)	.07(.10)	.66(.73)	.07(.10)	
2 × 10 ¹⁴	50	1000	.68(.78)	.08(.17)	.68(.78)	.09(.18)	.69(.79)	.10(.18)	.70(.79)	.11(.19)	.70(.79)	.12(.19)	.71(.79)	.13(.20)	.72(.79)	.14(.21)	
1100	.74(.85)	.20(.35)	.76(.85)	.25(.38)	.77(.85)	.29(.40)	.77(.86)	.33(.43)	.78(.86)	.36(.45)	.78(.86)	.39(.47)	.79(.86)	.41(.50)	.79(.86)	.44(.52)	
900	.69(.73)	.10(.12)	.69(.73)	.10(.12)	.69(.73)	.10(.12)	.69(.73)	.10(.12)	.69(.73)	.10(.12)	.69(.73)	.10(.12)	.69(.73)	.10(.12)	.69(.73)	.10(.12)	
80	1000	.70(.76)	.11(.19)	.71(.77)	.12(.19)	.71(.77)	.13(.20)	.72(.77)	.14(.20)	.72(.77)	.15(.21)	.72(.77)	.15(.21)	.73(.77)	.15(.21)	.73(.77)	.16(.22)
1100	.75(.82)	.22(.36)	.76(.83)	.26(.39)	.77(.83)	.30(.41)	.78(.83)	.33(.44)	.78(.84)	.37(.46)	.79(.84)	.39(.48)	.79(.84)	.42(.51)	.79(.84)	.45(.53)	
900	.49(.64)	.09(.09)	.49(.65)	.09(.09)	.50(.65)	.09(.09)	.50(.65)	.10(.09)	.50(.65)	.10(.09)	.51(.65)	.10(.09)	.51(.65)	.10(.09)	.51(.65)	.10(.09)	
20	1000	.51(.74)	.10(.16)	.52(.74)	.11(.17)	.53(.74)	.12(.17)	.53(.74)	.12(.18)	.54(.74)	.13(.18)	.54(.75)	.14(.19)	.55(.75)	.15(.19)	.55(.75)	.15(.20)
1100	.58(.82)	.20(.34)	.59(.82)	.24(.37)	.61(.83)	.28(.40)	.62(.83)	.31(.43)	.62(.83)	.34(.45)	.63(.84)	.37(.48)	.64(.84)	.40(.50)	.64(.84)	.42(.52)	
900	.41(.57)	.05(.12)	.43(.57)	.06(.12)	.44(.57)	.07(.12)	.45(.57)	.08(.12)	.46(.57)	.08(.13)	.46(.57)	.09(.13)	.47(.57)	.09(.13)	.47(.57)	.09(.13)	
2 × 10 ¹⁵	50	1000	.47(.59)	.09(.14)	.49(.59)	.10(.15)	.50(.60)	.11(.15)	.50(.60)	.12(.16)	.51(.60)	.13(.16)	.51(.61)	.14(.17)	.52(.61)	.15(.18)	
1100	.55(.71)	.20(.34)	.57(.72)	.24(.37)	.58(.72)	.28(.39)	.59(.73)	.31(.42)	.60(.73)	.34(.44)	.61(.73)	.37(.47)	.61(.74)	.39(.49)	.62(.74)	.42(.51)	
900	.42(.55)	.05(.13)	.42(.55)	.05(.13)	.42(.55)	.06(.13)	.42(.55)	.06(.13)	.43(.55)	.06(.13)	.43(.55)	.06(.13)	.44(.56)	.06(.13)	.44(.56)	.07(.13)	
80	1000	.47(.56)	.09(.14)	.49(.56)	.10(.15)	.50(.57)	.11(.15)	.50(.57)	.12(.16)	.51(.57)	.13(.16)	.52(.58)	.13(.17)	.52(.58)	.14(.17)	.53(.58)	.15(.18)
1100	.55(.67)	.20(.33)	.57(.68)	.24(.36)	.58(.68)	.28(.38)	.59(.69)	.31(.41)	.60(.69)	.34(.43)	.61(.70)	.37(.46)	.61(.70)	.39(.48)	.62(.70)	.42(.50)	
900	.49(.53)	.53(.28)	.49(.53)	.54(.28)	.49(.53)	.55(.28)	.49(.53)	.55(.28)	.49(.53)	.56(.28)	.48(.53)	.56(.29)	.48(.54)	.56(.29)	.48(.54)	.56(.29)	
20	1000	.39(.49)	.35(.15)	.39(.49)	.35(.15)	.40(.49)	.36(.16)	.40(.50)	.36(.16)	.40(.50)	.36(.17)	.40(.50)	.37(.18)	.40(.51)	.37(.18)	.41(.51)	.37(.19)
1100	.41(.59)	.36(.30)	.41(.60)	.38(.33)	.42(.61)	.41(.36)	.43(.62)	.43(.39)	.43(.62)	.45(.41)	.44(.63)	.47(.43)	.44(.63)	.49(.46)	.45(.63)	.51(.48)	
900	.38(.48)	.15(.59)	.39(.48)	.17(.59)	.39(.49)	.18(.59)	.39(.49)	.19(.60)	.40(.49)	.20(.59)	.40(.49)	.21(.59)	.40(.49)	.22(.59)	.40(.49)	.23(.60)	
2 × 10 ¹⁶	50	1000	.33(.39)	.22(.25)	.33(.39)	.22(.26)	.33(.39)	.23(.39)	.34(.40)	.24(.27)	.34(.40)	.24(.28)	.34(.40)	.25(.28)	.35(.40)	.26(.28)	
1100	.35(.46)	.26(.27)	.36(.47)	.30(.30)	.38(.48)	.34(.33)	.38(.49)	.37(.36)	.39(.49)	.40(.38)	.40(.50)	.42(.41)	.40(.50)	.45(.43)	.41(.51)	.47(.45)	
900	.30(.47)	.08(.59)	.31(.47)	.09(.60)	.32(.47)	.11(.60)	.33(.47)	.12(.60)	.34(.47)	.13(.61)	.34(.47)	.14(.61)	.35(.47)	.15(.61)	.35(.47)	.16(.61)	
80	1000	.31(.37)	.17(.25)	.31(.37)	.18(.26)	.32(.38)	.19(.26)	.32(.38)	.19(.27)	.33(.38)	.20(.27)	.33(.38)	.21(.28)	.33(.39)	.22(.28)	.34(.39)	.23(.29)
1100	.35(.43)	.25(.27)	.36(.45)	.29(.30)	.38(.46)	.33(.33)	.39(.46)	.36(.36)	.39(.47)	.39(.39)	.40(.48)	.41(.41)	.41(.48)	.44(.43)	.41(.49)	.46(.45)	

Table G.3: Look-up table for β_1^* (-) and β_2^* (μm) for arsenic ion implantation.

		Time (min)												
		Time (min)												
		Time (min)												
Dose (cm^{-2})	Energy (keV)	β_1^*	15	30	45	60	75	90	105	β_2^*	β_1^*	β_2^*	β_1^*	
2×10^{14}	900	.65(.84)	.01(.01)	.66(.84)	.01(.01)	.66(.84)	.01(.01)	.67(.85)	.01(.01)	.67(.85)	.01(.01)	.67(.85)	.02(.01)	
	20	1000	.71(.90)	.03(.03)	.73(.90)	.03(.03)	.74(.91)	.03(.03)	.75(.91)	.04(.04)	.75(.91)	.04(.04)	.76(.92)	.05(.04)
	1100	.79(.93)	.08(.08)	.80(.94)	.09(.09)	.81(.94)	.11(.10)	.82(.94)	.12(.11)	.83(.95)	.13(.12)	.83(.95)	.14(.13)	.83(.95)
	900	.53(.69)	.02(.02)	.54(.69)	.02(.02)	.54(.69)	.02(.02)	.55(.69)	.02(.02)	.55(.69)	.02(.02)	.55(.70)	.02(.02)	.56(.70)
	1100	.66(.81)	.08(.09)	.68(.82)	.10(.10)	.69(.83)	.11(.12)	.70(.83)	.13(.13)	.71(.83)	.14(.14)	.71(.84)	.15(.15)	.72(.84)
	80	1000	.61(.68)	.04(.04)	.62(.68)	.05(.04)	.62(.69)	.05(.05)	.63(.70)	.05(.05)	.64(.70)	.06(.05)	.64(.71)	.06(.06)
2×10^{15}	1100	.67(.76)	.08(.09)	.68(.77)	.10(.11)	.70(.78)	.12(.12)	.70(.78)	.13(.13)	.71(.79)	.14(.14)	.72(.79)	.15(.15)	.72(.80)
	900	.45(.70)	.01(.01)	.46(.71)	.01(.01)	.46(.71)	.01(.01)	.47(.71)	.02(.01)	.48(.72)	.02(.01)	.48(.72)	.02(.02)	.49(.72)
	20	1000	.54(.80)	.03(.03)	.55(.81)	.04(.03)	.56(.82)	.04(.03)	.57(.82)	.05(.04)	.57(.82)	.05(.04)	.58(.83)	.05(.04)
	1100	.62(.87)	.08(.08)	.64(.88)	.10(.10)	.65(.89)	.12(.11)	.66(.89)	.13(.12)	.67(.89)	.14(.13)	.67(.89)	.15(.14)	.68(.90)
	900	.36(.49)	.02(.01)	.36(.50)	.02(.01)	.36(.50)	.02(.02)	.36(.51)	.02(.02)	.36(.52)	.02(.02)	.37(.52)	.02(.02)	.37(.52)
	1100	.48(.67)	.10(.09)	.49(.68)	.12(.11)	.51(.69)	.14(.12)	.52(.70)	.15(.13)	.52(.71)	.16(.15)	.53(.71)	.17(.16)	.54(.72)
80	900	.37(.44)	.02(.02)	.37(.44)	.02(.02)	.37(.45)	.02(.02)	.37(.45)	.02(.02)	.37(.46)	.02(.02)	.37(.46)	.02(.02)	.37(.47)
	1000	.40(.50)	.04(.04)	.41(.51)	.05(.04)	.42(.52)	.05(.05)	.43(.53)	.06(.05)	.44(.54)	.06(.06)	.44(.54)	.07(.06)	.45(.55)
	1100	.48(.61)	.10(.10)	.50(.63)	.12(.11)	.51(.64)	.14(.13)	.52(.64)	.15(.14)	.53(.65)	.16(.15)	.53(.66)	.17(.16)	.54(.66)
	900	.33(.50)	.02(.01)	.33(.51)	.02(.01)	.34(.52)	.02(.01)	.34(.53)	.02(.01)	.34(.53)	.02(.02)	.34(.54)	.02(.02)	.34(.54)
	20	1000	.35(.62)	.04(.03)	.37(.63)	.05(.03)	.37(.64)	.05(.04)	.38(.65)	.06(.04)	.39(.65)	.06(.04)	.39(.66)	.07(.05)
	1100	.43(.73)	.10(.09)	.45(.74)	.13(.10)	.46(.75)	.14(.12)	.47(.76)	.16(.13)	.48(.76)	.17(.14)	.49(.77)	.19(.15)	.49(.77)
2×10^{16}	900	.27(.31)	.02(.02)	.28(.32)	.02(.02)	.28(.32)	.03(.02)	.28(.33)	.03(.02)	.29(.33)	.03(.03)	.29(.33)	.03(.03)	.29(.33)
	1000	.28(.35)	.04(.04)	.29(.35)	.05(.05)	.29(.36)	.06(.06)	.29(.37)	.06(.06)	.29(.37)	.07(.07)	.30(.38)	.07(.07)	.30(.39)
	1100	.31(.47)	.12(.12)	.32(.48)	.15(.14)	.33(.49)	.17(.15)	.34(.50)	.19(.17)	.34(.50)	.21(.18)	.35(.51)	.23(.19)	.35(.52)
	900	.28(.29)	.02(.03)	.28(.30)	.02(.03)	.28(.31)	.02(.03)	.28(.31)	.03(.03)	.28(.31)	.03(.03)	.28(.31)	.03(.03)	.28(.31)
	1000	.28(.32)	.04(.05)	.29(.33)	.05(.05)	.29(.33)	.06(.06)	.29(.34)	.06(.06)	.30(.34)	.07(.07)	.30(.35)	.08(.08)	.30(.35)
	1100	.31(.42)	.12(.13)	.32(.43)	.15(.15)	.33(.44)	.17(.16)	.34(.45)	.19(.18)	.35(.45)	.21(.19)	.35(.46)	.23(.20)	.36(.47)

Table G.4: Look-up table for R_s (Ω/\square) for boron ion implantation.

Dose (cm ⁻²)	Energy (keV)	Temp. (°C)	Time (min)						
			15	30	45	60	75	90	105
20	900	5.94e2(8.49e2)	5.85e2(8.50e2)	5.78e2(8.51e2)	5.72e2(8.52e2)	5.66e2(8.53e2)	5.61e2(8.53e2)	5.57e2(8.53e2)	5.53e2(8.53e2)
	1000	4.67e2(7.72e2)	4.47e2(7.67e2)	4.32e2(7.63e2)	4.21e2(7.59e2)	4.10e2(7.55e2)	4.01e2(7.52e2)	3.93e2(7.48e2)	3.86e2(7.44e2)
	1100	3.39e2(6.02e2)	3.15e2(5.87e2)	2.98e2(5.73e2)	2.86e2(5.62e2)	2.77e2(5.51e2)	2.68e2(5.42e2)	2.62e2(5.34e2)	2.56e2(5.26e2)
	900	3.68e2(4.69e2)	3.67e2(4.69e2)	3.67e2(4.69e2)	3.67e2(4.70e2)	3.66e2(4.70e2)	3.66e2(4.70e2)	3.65e2(4.70e2)	3.65e2(4.71e2)
2×10^{14}	50	1000	3.48e2(4.66e2)	3.44e2(4.65e2)	3.40e2(4.64e2)	3.36e2(4.63e2)	3.33e2(4.61e2)	3.30e2(4.60e2)	3.27e2(4.59e2)
	1100	3.03e2(4.17e2)	2.89e2(4.08e2)	2.79e2(4.00e2)	2.70e2(3.93e2)	2.63e2(3.87e2)	2.57e2(3.82e2)	2.52e2(3.77e2)	2.47e2(3.72e2)
	900	3.10e2(3.75e2)	3.10e2(3.76e2)	3.10e2(3.76e2)	3.11e2(3.76e2)	3.11e2(3.76e2)	3.11e2(3.76e2)	3.11e2(3.76e2)	3.11e2(3.77e2)
	80	1000	3.05e2(3.78e2)	3.04e2(3.78e2)	3.02e2(3.78e2)	3.00e2(3.78e2)	2.98e2(3.78e2)	2.97e2(3.78e2)	2.96e2(3.74e2)
20	1100	2.81e2(3.53e2)	2.71e2(3.47e2)	2.64e2(3.41e2)	2.58e2(3.36e2)	2.52e2(3.31e2)	2.47e2(3.27e2)	2.43e2(3.23e2)	2.39e2(3.19e2)
	900	9.80e1(2.20e2)	9.21e1(2.20e2)	8.93e1(2.20e2)	8.79e1(2.21e2)	8.83e1(2.21e2)	8.81e1(2.21e2)	8.74e1(2.22e2)	8.64e1(2.22e2)
	1000	6.83e1(1.30e2)	6.53e1(1.30e2)	6.32e1(1.29e2)	6.17e1(1.28e2)	6.05e1(1.27e2)	5.96e1(1.27e2)	5.87e1(1.26e2)	5.80e1(1.25e2)
	1100	5.50e1(1.00e2)	5.27e1(9.80e1)	5.11e1(9.60e1)	5.00e1(9.43e1)	4.90e1(9.27e1)	4.82e1(9.13e1)	4.75e1(9.00e1)	4.68e1(8.88e1)
2×10^{15}	900	6.61e1(1.04e2)	6.62e1(1.05e2)	6.64e1(1.05e2)	6.66e1(1.05e2)	6.68e1(1.05e2)	6.69e1(1.05e2)	6.70e1(1.05e2)	6.70e1(1.06e2)
	1000	6.22e1(9.71e1)	6.09e1(9.70e1)	5.99e1(9.68e1)	5.91e1(9.66e1)	5.85e1(9.63e1)	5.79e1(9.61e1)	5.74e1(9.58e1)	5.69e1(9.55e1)
	1100	5.45e1(8.40e1)	5.26e1(8.23e1)	5.13e1(8.09e1)	5.02e1(7.96e1)	4.93e1(7.85e1)	4.85e1(7.74e1)	4.78e1(7.65e1)	4.72e1(7.56e1)
	900	6.02e1(8.10e1)	6.01e1(8.11e1)	6.00e1(8.12e1)	5.99e1(8.13e1)	5.98e1(8.14e1)	5.97e1(8.15e1)	5.97e1(8.16e1)	5.96e1(8.17e1)
80	1000	5.66e1(7.96e1)	5.60e1(7.97e1)	5.54e1(7.97e1)	5.49e1(7.96e1)	5.45e1(7.95e1)	5.42e1(7.95e1)	5.39e1(7.93e1)	5.36e1(7.92e1)
	1100	5.19e1(7.37e1)	5.07e1(7.26e1)	4.97e1(7.17e1)	4.88e1(7.07e1)	4.81e1(6.99e1)	4.74e1(6.91e1)	4.68e1(6.84e1)	4.63e1(6.77e1)
	900	6.09e1(5.90e1)	5.74e1(5.70e1)	5.47e1(5.52e1)	5.25e1(5.35e1)	5.06e1(5.20e1)	4.89e1(5.10e1)	4.74e1(5.03e1)	4.60e1(4.94e1)
	20	1000	2.14e1(2.12e1)	1.81e1(2.04e1)	1.60e1(1.98e1)	1.45e1(1.93e1)	1.34e1(1.90e1)	1.25e1(1.87e1)	1.17e1(1.85e1)
2×10^{16}	1100	7.40e0(1.24e1)	7.00e0(1.22e1)	6.88e0(1.20e1)	6.79e0(1.18e1)	6.72e0(1.17e1)	6.66e0(1.16e1)	6.61e0(1.15e1)	6.57e0(1.14e1)
	900	3.59e1(3.84e1)	3.49e1(3.79e1)	3.41e1(3.75e1)	3.34e1(3.71e1)	3.28e1(3.67e1)	3.22e1(3.62e1)	3.16e1(3.58e1)	3.11e1(3.55e1)
	80	1000	1.62e1(1.53e1)	1.44e1(1.40e1)	1.31e1(1.30e1)	1.21e1(1.22e1)	1.14e1(1.15e1)	1.08e1(1.09e1)	1.03e1(1.05e1)
	1100	6.87e0(1.03e1)	6.70e0(1.01e1)	6.59e0(9.91e0)	6.51e0(9.78e0)	6.44e0(9.68e0)	6.39e0(9.59e0)	6.34e0(9.51e0)	6.30e0(9.45e0)
900	2.43e1(2.64e1)	2.38e1(2.68e1)	2.34e1(2.69e1)	2.30e1(2.68e1)	2.27e1(2.67e1)	2.25e1(2.66e1)	2.23e1(2.65e1)	2.21e1(2.63e1)	
	1000	1.29e1(1.31e1)	1.17e1(1.20e1)	1.09e1(1.12e1)	1.03e1(1.06e1)	9.77e0(9.73e0)	9.32e0(9.73e0)	8.94e0(9.35e0)	8.64e0(9.12e0)
80	1100	6.73e0(9.44e0)	6.59e0(9.30e0)	6.49e0(9.17e0)	6.41e0(9.07e0)	6.35e0(8.98e0)	6.30e0(8.90e0)	6.26e0(8.83e0)	6.22e0(8.77e0)

Table G.5: Look-up table for $R_s (\Omega/\square)$ for phosphorus ion implantation.

		Time (min)							
Dose (cm ⁻²)	Energy (keV) Temp. (°C)	15	30	45	60	75	90	105	120
2×10^{14}	900	3.99e2(7.10e2)	3.88e2(7.05e2)	3.75e2(7.01e2)	3.70e2(6.98e2)	3.66e2(6.92e2)	3.63e2(6.90e2)	3.59e2(6.88e2)	
	20	3.20e2(8.53e2)	3.04e2(8.41e2)	2.92e2(8.30e2)	2.83e2(8.20e2)	2.75e2(8.11e2)	2.68e2(8.02e2)	2.63e2(7.94e2)	2.58e2(7.86e2)
	1000	2.24e2(9.15e2)	2.07e2(8.90e2)	1.95e2(8.68e2)	1.87e2(8.51e2)	1.80e2(8.36e2)	1.74e2(8.23e2)	1.70e2(8.12e2)	1.66e2(8.02e2)
	900	2.74e2(3.29e2)	2.67e2(3.28e2)	2.63e2(3.26e2)	2.61e2(3.25e2)	2.60e2(3.24e2)	2.58e2(3.24e2)	2.57e2(3.23e2)	2.56e2(3.22e2)
	50	2.40e2(3.12e2)	2.32e2(3.09e2)	2.26e2(3.05e2)	2.21e2(3.02e2)	2.16e2(2.99e2)	2.12e2(2.96e2)	2.08e2(2.94e2)	2.05e2(2.91e2)
	1100	1.81e2(3.16e2)	1.69e2(3.05e2)	1.60e2(2.96e2)	1.53e2(2.89e2)	1.48e2(2.83e2)	1.43e2(2.77e2)	1.40e2(2.72e2)	1.37e2(2.67e2)
2×10^{15}	900	2.44e2(2.76e2)	2.39e2(2.75e2)	2.37e2(2.74e2)	2.36e2(2.74e2)	2.36e2(2.73e2)	2.35e2(2.72e2)	2.35e2(2.72e2)	2.34e2(2.71e2)
	80	1000	2.26e2(2.49e2)	2.22e2(2.46e2)	2.18e2(2.44e2)	2.14e2(2.41e2)	2.10e2(2.39e2)	2.07e2(2.37e2)	2.04e2(2.35e2)
	1100	1.82e2(2.37e2)	1.70e2(2.29e2)	1.61e2(2.23e2)	1.55e2(2.17e2)	1.50e2(2.12e2)	1.46e2(2.08e2)	1.42e2(2.05e2)	1.39e2(2.01e2)
	900	5.14e1(1.74e2)	4.95e1(1.73e2)	4.80e1(1.72e2)	4.81e1(1.72e2)	4.78e1(1.71e2)	4.75e1(1.71e2)	4.73e1(1.71e2)	4.71e1(1.70e2)
	20	1000	4.74e1(2.16e2)	4.67e1(2.14e2)	4.62e1(2.11e2)	4.56e1(2.09e2)	4.52e1(2.07e2)	4.48e1(2.05e2)	4.44e1(2.03e2)
	1100	4.20e1(2.33e2)	4.03e1(2.25e2)	3.91e1(2.19e2)	3.81e1(2.13e2)	3.73e1(2.08e2)	3.67e1(2.04e2)	3.61e1(2.00e2)	3.56e1(1.96e2)
2×10^{15}	900	4.59e1(6.61e1)	4.28e1(6.60e1)	4.11e1(6.58e1)	4.00e1(6.57e1)	3.93e1(6.56e1)	3.88e1(6.55e1)	3.84e1(6.54e1)	3.81e1(6.53e1)
	50	1000	3.79e1(6.63e1)	3.72e1(6.57e1)	3.67e1(6.51e1)	3.63e1(6.46e1)	3.59e1(6.41e1)	3.56e1(6.36e1)	3.53e1(6.32e1)
	1100	3.34e1(8.17e1)	3.22e1(7.95e1)	3.13e1(7.77e1)	3.06e1(7.61e1)	3.00e1(7.46e1)	2.95e1(7.34e1)	2.91e1(7.22e1)	2.87e1(7.11e1)
	900	4.60e1(5.34e1)	4.40e1(5.32e1)	4.31e1(5.31e1)	4.25e1(5.29e1)	4.21e1(5.28e1)	4.17e1(5.28e1)	4.14e1(5.27e1)	4.11e1(5.26e1)
	80	1000	3.91e1(5.05e1)	3.84e1(5.01e1)	3.79e1(4.97e1)	3.74e1(4.93e1)	3.70e1(4.90e1)	3.67e1(4.87e1)	3.64e1(4.84e1)
	1100	3.44e1(5.94e1)	3.31e1(5.79e1)	3.22e1(5.67e1)	3.14e1(5.56e1)	3.09e1(5.47e1)	3.03e1(5.38e1)	2.99e1(5.31e1)	2.95e1(5.24e1)
2×10^{16}	900	9.97e0(2.20e1)	9.66e0(2.19e1)	9.39e0(2.19e1)	9.14e0(2.18e1)	8.91e0(2.18e1)	8.70e0(2.18e1)	8.51e0(2.18e1)	8.32e0(2.17e1)
	20	1000	5.95e0(2.58e1)	5.83e0(2.56e1)	5.75e0(2.55e1)	5.70e0(2.53e1)	5.65e0(2.52e1)	5.62e0(2.51e1)	5.59e0(2.50e1)
	1100	5.54e0(3.22e1)	5.46e0(3.15e1)	5.40e0(3.10e1)	5.35e0(3.05e1)	5.31e0(3.01e1)	5.28e0(2.97e1)	5.25e0(2.94e1)	5.22e0(2.90e1)
	900	1.62e1(7.76e0)	1.49e1(7.70e0)	1.40e1(7.64e0)	1.32e1(7.60e0)	1.26e1(7.57e0)	1.20e1(7.54e0)	1.15e1(7.52e0)	1.10e1(7.50e0)
	50	1000	5.61e0(6.93e0)	5.27e0(6.87e0)	5.09e0(6.82e0)	4.97e0(6.78e0)	4.89e0(6.75e0)	4.82e0(6.71e0)	4.76e0(6.68e0)
	1100	4.62e0(1.10e1)	4.48e0(1.08e1)	4.40e0(1.06e1)	4.34e0(1.05e1)	4.28e0(1.04e1)	4.24e0(1.03e1)	4.21e0(1.02e1)	4.17e0(1.01e1)
2×10^{16}	900	1.60e1(6.58e0)	1.45e1(6.55e0)	1.36e1(6.50e0)	1.29e1(6.46e0)	1.24e1(6.40e0)	1.19e1(6.35e0)	1.15e1(6.31e0)	1.12e1(6.28e0)
	80	1000	6.55e0(6.14e0)	5.81e0(6.06e0)	5.50e0(6.00e0)	5.31e0(5.96e0)	5.19e0(5.92e0)	5.10e0(5.89e0)	5.03e0(5.86e0)
	1100	4.78e0(8.95e0)	4.64e0(8.80e0)	4.55e0(8.68e0)	4.48e0(8.57e0)	4.42e0(8.49e0)	4.38e0(8.41e0)	4.34e0(8.35e0)	4.31e0(8.28e0)

Table G.6: Look-up table for R_s (Ω/\square) for arsenic ion implantation.

		Time (min)							
Dose (cm ⁻²)	Energy (keV)	15	30	45	60	75	90	105	120
2×10^{14}	900	1.58e3(1.04e4)	1.53e3(9.95e3)	1.50e3(9.58e3)	1.47e3(9.28e3)	1.45e3(9.03e3)	1.43e3(8.81e3)	1.42e3(8.62e3)	1.40e3(8.46e3)
	20	1.20e3(9.6e3)	1.13e3(9.06e3)	1.08e3(8.68e3)	1.04e3(8.41e3)	1.01e3(8.19e3)	9.84e2(8.02e3)	9.61e2(7.86e3)	9.41e2(7.72e3)
	1000	7.83e2(7.13e3)	7.17e2(6.87e3)	6.74e2(6.69e3)	6.42e2(6.56e3)	6.18e2(6.47e3)	5.98e2(6.39e3)	5.81e2(6.33e3)	5.67e2(6.28e3)
	900	3.68e2(1.43e3)	3.55e2(1.39e3)	3.48e2(1.36e3)	3.44e2(1.33e3)	3.41e2(1.31e3)	3.38e2(1.29e3)	3.35e2(1.27e3)	3.33e2(1.26e3)
	50	3.05e2(1.02e3)	2.95e2(9.81e2)	2.88e2(9.50e2)	2.82e2(9.24e2)	2.77e2(9.03e2)	2.73e2(8.84e2)	2.69e2(8.68e2)	2.66e2(8.53e2)
	1100	2.40e2(8.29e2)	2.26e2(7.75e2)	2.16e2(7.35e2)	2.08e2(7.05e2)	2.02e2(6.81e2)	1.97e2(6.61e2)	1.92e2(6.44e2)	1.89e2(6.29e2)
2×10^{15}	900	3.31e2(8.71e2)	3.21e2(8.46e2)	3.17e2(8.29e2)	3.15e2(8.18e2)	3.14e2(8.10e2)	3.13e2(8.04e2)	3.12e2(7.98e2)	3.11e2(7.93e2)
	80	2.96e2(5.85e2)	2.89e2(5.68e2)	2.84e2(5.52e2)	2.79e2(5.43e2)	2.75e2(5.34e2)	2.72e2(5.25e2)	2.69e2(5.17e2)	2.66e2(5.10e2)
	1100	2.42e2(4.90e2)	2.29e2(4.62e2)	2.19e2(4.40e2)	2.11e2(4.24e2)	2.05e2(4.11e2)	2.00e2(3.99e2)	1.96e2(3.89e2)	1.92e2(3.81e2)
	900	2.69e2(2.28e3)	2.54e2(2.21e3)	2.44e2(2.16e3)	2.37e2(2.12e3)	2.32e2(2.09e3)	2.29e2(2.06e3)	2.26e2(2.04e3)	2.23e2(2.02e3)
	20	2.00e2(2.57e3)	1.93e2(2.42e3)	1.88e2(2.30e3)	1.84e2(2.22e3)	1.82e2(2.15e3)	1.79e2(2.09e3)	1.77e2(2.04e3)	1.75e2(2.00e3)
	1100	1.60e2(1.91e3)	1.52e2(1.78e3)	1.46e2(1.69e3)	1.41e2(1.63e3)	1.38e2(1.58e3)	1.35e2(1.53e3)	1.32e2(1.50e3)	1.30e2(1.47e3)
2×10^{16}	900	1.03e2(3.29e2)	9.49e1(3.22e2)	8.92e1(3.16e2)	8.47e1(3.12e2)	8.10e1(3.08e2)	7.79e1(3.05e2)	7.54e1(3.02e2)	7.31e1(3.00e2)
	50	5.12e1(2.36e2)	4.80e1(2.30e2)	4.61e1(2.25e2)	4.47e1(2.21e2)	4.38e1(2.18e2)	4.30e1(2.14e2)	4.23e1(2.12e2)	4.18e1(2.09e2)
	1100	3.89e1(2.19e2)	3.74e1(2.09e2)	3.64e1(2.01e2)	3.57e1(1.95e2)	3.51e1(1.89e2)	3.46e1(1.85e2)	3.42e1(1.81e2)	3.38e1(1.78e2)
	900	8.80e1(1.87e2)	8.24e1(1.81e2)	7.87e1(1.76e2)	7.59e1(1.72e2)	7.36e1(1.69e2)	7.15e1(1.66e2)	6.98e1(1.64e2)	6.82e1(1.62e2)
	80	5.17e1(1.24e2)	4.88e1(1.21e2)	4.71e1(1.19e2)	4.53e1(1.17e2)	4.49e1(1.15e2)	4.41e1(1.14e2)	4.35e1(1.13e2)	4.30e1(1.12e2)
	1100	4.01e1(1.26e2)	3.85e1(1.21e2)	3.75e1(1.18e2)	3.68e1(1.15e2)	3.62e1(1.12e2)	3.57e1(1.10e2)	3.52e1(1.08e2)	3.48e1(1.07e2)
2×10^{17}	900	7.37e1(4.35e2)	7.06e1(4.24e2)	6.77e1(4.15e2)	6.51e1(4.08e2)	6.28e1(4.03e2)	6.07e1(3.98e2)	5.88e1(3.94e2)	5.73e1(3.90e2)
	20	3.78e1(4.47e2)	3.44e1(4.32e2)	3.24e1(4.19e2)	3.10e1(4.09e2)	3.00e1(4.01e2)	2.92e1(3.93e2)	2.86e1(3.87e2)	2.81e1(3.81e2)
	1100	2.53e1(3.92e2)	2.42e1(3.69e2)	2.35e1(3.52e2)	2.30e1(3.39e2)	2.26e1(3.28e2)	2.23e1(3.19e2)	2.20e1(3.12e2)	2.18e1(3.05e2)
	900	4.00e1(5.45e1)	3.92e1(5.36e1)	3.82e1(5.24e1)	3.72e1(5.11e1)	3.62e1(4.98e1)	3.53e1(4.87e1)	3.45e1(4.77e1)	3.37e1(4.67e1)
	50	2.18e1(3.14e1)	1.90e1(2.90e1)	1.71e1(2.73e1)	1.58e1(2.62e1)	1.48e1(2.52e1)	1.40e1(2.46e1)	1.33e1(2.40e1)	1.28e1(2.35e1)
	1100	9.03e0(2.98e1)	7.90e0(2.89e1)	7.24e0(2.83e1)	6.82e0(2.78e1)	6.50e0(2.74e1)	6.27e0(2.70e1)	6.08e0(2.67e1)	5.92e0(2.65e1)
80	900	4.43e1(4.05e1)	4.20e1(4.04e1)	4.02e1(3.98e1)	3.88e1(3.91e1)	3.75e1(3.84e1)	3.64e1(3.77e1)	3.54e1(3.70e1)	3.45e1(3.64e1)
	1000	2.17e1(2.56e1)	1.89e1(2.33e1)	1.71e1(2.16e1)	1.58e1(2.05e1)	1.48e1(1.95e1)	1.40e1(1.88e1)	1.33e1(1.82e1)	1.28e1(1.77e1)
1100	9.09e0(1.83e1)	7.96e0(1.78e1)	7.33e0(1.73e1)	6.90e0(1.70e1)	6.59e0(1.67e1)	6.35e0(1.65e1)	6.16e0(1.63e1)	6.01e0(1.61e1)	5.90e0(1.60e1)

Table G.7: Look-up table for N_z (μm^{-2}) for boron ion implantation.

		Time (min)							
Dose (cm^{-2})	Energy (keV)	15	30	45	60	75	90	105	120
2×10^{14}	900	8.87e5(4.44e5)	9.05e5(4.43e5)	9.20e5(4.41e5)	9.35e5(4.40e5)	9.45e5(4.39e5)	9.56e5(4.38e5)	9.65e5(4.37e5)	9.74e5(4.36e5)
	20	1.22e6(4.49e5)	1.27e6(4.47e5)	1.31e6(4.46e5)	1.35e6(4.44e5)	1.38e6(4.42e5)	1.38e6(4.41e5)	1.39e6(4.40e5)	1.41e6(4.39e5)
	1100	1.53e6(5.00e5)	1.57e6(5.08e5)	1.60e6(5.11e5)	1.62e6(5.14e5)	1.63e6(5.17e5)	1.65e6(5.20e5)	1.66e6(5.22e5)	1.66e6(5.24e5)
	900	1.37e6(9.01e5)	1.37e6(8.99e5)	1.37e6(8.98e5)	1.37e6(8.96e5)	1.37e6(8.95e5)	1.37e6(8.93e5)	1.37e6(8.91e5)	1.37e6(8.90e5)
	50	1.48e6(8.49e5)	1.49e6(8.45e5)	1.50e6(8.38e5)	1.51e6(8.38e5)	1.52e6(8.36e5)	1.52e6(8.33e5)	1.53e6(8.30e5)	1.53e6(8.28e5)
	1100	1.60e6(8.26e5)	1.62e6(8.30e5)	1.64e6(8.31e5)	1.65e6(8.34e5)	1.66e6(8.35e5)	1.67e6(8.38e5)	1.68e6(8.41e5)	1.69e6(8.43e5)
2×10^{15}	900	1.58e6(1.15e6)	1.58e6(1.14e6)	1.57e6(1.14e6)	1.57e6(1.14e6)	1.57e6(1.14e6)	1.57e6(1.14e6)	1.57e6(1.14e6)	1.57e6(1.14e6)
	80	1.000	1.62e6(1.03e6)	1.62e6(1.08e6)	1.62e6(1.07e6)	1.62e6(1.07e6)	1.63e6(1.07e6)	1.63e6(1.06e6)	1.63e6(1.06e6)
	1100	1.67e6(1.02e6)	1.68e6(1.02e6)	1.68e6(1.02e6)	1.69e6(1.02e6)	1.69e6(1.03e6)	1.70e6(1.03e6)	1.71e6(1.03e6)	1.71e6(1.03e6)
	900	1.08e7(3.39e6)	1.16e7(3.38e6)	1.20e7(3.36e6)	1.21e7(3.35e6)	1.19e7(3.33e6)	1.19e7(3.32e6)	1.19e7(3.31e6)	1.19e7(3.29e6)
	20	1.54e7(6.00e6)	1.59e7(5.98e6)	1.62e7(5.95e6)	1.64e7(5.93e6)	1.65e7(5.92e6)	1.67e7(5.91e6)	1.68e7(5.91e6)	1.69e7(5.90e6)
	1100	1.73e7(6.97e6)	1.75e7(6.98e6)	1.76e7(6.99e6)	1.76e7(7.00e6)	1.77e7(7.02e6)	1.77e7(7.04e6)	1.77e7(7.06e6)	1.77e7(7.07e6)
2×10^{15}	900	1.39e7(7.38e6)	1.38e7(7.36e6)	1.38e7(7.33e6)	1.37e7(7.31e6)	1.36e7(7.29e6)	1.36e7(7.27e6)	1.36e7(7.25e6)	1.36e7(7.23e6)
	50	1.000	1.50e7(8.02e6)	1.53e7(8.00e6)	1.56e7(7.98e6)	1.58e7(7.96e6)	1.59e7(7.94e6)	1.61e7(7.93e6)	1.62e7(7.94e6)
	1100	1.69e7(8.65e6)	1.71e7(8.66e6)	1.72e7(8.67e6)	1.73e7(8.69e6)	1.73e7(8.70e6)	1.73e7(8.71e6)	1.74e7(8.72e6)	1.74e7(8.73e6)
	900	1.42e7(9.41e6)	1.43e7(9.39e6)	1.43e7(9.37e6)	1.43e7(9.35e6)	1.43e7(9.33e6)	1.43e7(9.31e6)	1.44e7(9.29e6)	1.44e7(9.27e6)
	80	1.000	1.56e7(9.64e6)	1.58e7(9.60e6)	1.60e7(9.57e6)	1.61e7(9.55e6)	1.62e7(9.53e6)	1.63e7(9.52e6)	1.64e7(9.50e6)
	1100	1.69e7(9.90e6)	1.71e7(9.91e6)	1.73e7(9.92e6)	1.74e7(9.93e6)	1.75e7(9.95e6)	1.75e7(9.96e6)	1.75e7(9.96e6)	1.75e7(9.96e6)
2×10^{16}	900	1.71e7(1.74e7)	1.83e7(1.81e7)	1.93e7(1.88e7)	2.03e7(1.94e7)	2.11e7(2.00e7)	2.19e7(2.04e7)	2.27e7(2.07e7)	2.34e7(2.10e7)
	20	5.59e7(5.45e7)	6.65e7(5.64e7)	7.53e7(5.78e7)	8.31e7(5.90e7)	9.02e7(5.98e7)	9.67e7(6.05e7)	1.03e8(6.10e7)	1.07e8(6.15e7)
	1100	1.70e8(9.20e7)	1.76e8(9.34e7)	1.77e8(9.39e7)	1.78e8(9.43e7)	1.78e8(9.46e7)	1.78e8(9.49e7)	1.79e8(9.51e7)	1.79e8(9.53e7)
	900	2.86e7(2.63e7)	2.97e7(2.67e7)	3.05e7(2.70e7)	3.12e7(2.74e7)	3.19e7(2.78e7)	3.26e7(2.82e7)	3.33e7(2.85e7)	3.39e7(2.89e7)
	50	1000	7.42e7(7.79e7)	8.44e7(8.54e7)	9.30e7(9.24e7)	1.00e8(9.84e7)	1.07e8(1.04e8)	1.13e8(1.10e8)	1.19e8(1.15e8)
	1100	1.83e8(1.14e8)	1.84e8(1.15e8)	1.85e8(1.16e8)	1.86e8(1.17e8)	1.86e8(1.17e8)	1.86e8(1.17e8)	1.87e8(1.18e8)	1.87e8(1.18e8)
80	900	4.19e7(3.77e7)	4.30e7(3.71e7)	4.39e7(3.69e7)	4.47e7(3.70e7)	4.54e7(3.72e7)	4.59e7(3.74e7)	4.65e7(3.76e7)	4.71e7(3.79e7)
	1000	9.19e7(8.97e7)	1.02e8(9.86e7)	1.10e8(1.06e8)	1.17e8(1.14e8)	1.24e8(1.19e8)	1.31e8(1.24e8)	1.36e8(1.29e8)	1.41e8(1.32e8)
	1100	1.86e8(1.24e8)	1.87e8(1.25e8)	1.88e8(1.26e8)	1.88e8(1.27e8)	1.89e8(1.27e8)	1.89e8(1.27e8)	1.89e8(1.28e8)	1.89e8(1.28e8)

APPENDIX G. ION IMPLANTATION LOOKUP TABLES

Table G.8: Look-up table for N_z (μm^{-2}) for phosphorus ion implantation.

		Time (min)							
Dose (cm^{-2})	Energy (keV)	15	30	45	60	75	90	105	120
2×10^{14}	900	1.11e6(3.01e5)	1.14e6(3.04e5)	1.16e6(3.07e5)	1.18e6(3.08e5)	1.19e6(3.09e5)	1.20e6(3.10e5)	1.21e6(3.11e5)	1.21e6(3.12e5)
	20	1.25e6(1.63e5)	1.26e6(1.64e5)	1.26e6(1.65e5)	1.26e6(1.65e5)	1.26e6(1.66e5)	1.25e6(1.66e5)	1.25e6(1.67e5)	1.26e6(1.67e5)
	1000	1.25e6(1.02e5)	1.25e6(1.02e5)	1.26e6(1.02e5)	1.26e6(1.03e5)	1.26e6(1.03e5)	1.27e6(1.03e5)	1.27e6(1.03e5)	1.27e6(1.03e5)
	900	1.40e6(9.41e5)	1.45e6(9.46e5)	1.48e6(9.50e5)	1.50e6(9.53e5)	1.51e6(9.55e5)	1.52e6(9.58e5)	1.53e6(9.60e5)	1.54e6(9.61e5)
	50	1.63e6(7.84e5)	1.66e6(7.87e5)	1.68e6(7.90e5)	1.69e6(7.91e5)	1.70e6(7.93e5)	1.71e6(7.95e5)	1.71e6(7.97e5)	1.71e6(7.98e5)
	1100	1.71e6(5.14e5)	1.71e6(5.17e5)	1.72e6(5.20e5)	1.72e6(5.23e5)	1.72e6(5.25e5)	1.73e6(5.27e5)	1.73e6(5.28e5)	1.73e6(5.29e5)
2×10^{15}	900	1.42e6(1.11e6)	1.46e6(1.11e6)	1.48e6(1.12e6)	1.49e6(1.12e6)	1.49e6(1.12e6)	1.50e6(1.12e6)	1.50e6(1.13e6)	1.50e6(1.13e6)
	80	1.55e6(1.03e6)	1.57e6(1.09e6)	1.59e6(1.09e6)	1.60e6(1.09e6)	1.61e6(1.09e6)	1.62e6(1.10e6)	1.63e6(1.10e6)	1.63e6(1.10e6)
	1100	1.65e6(8.07e5)	1.66e6(8.11e5)	1.66e6(8.15e5)	1.66e6(8.19e5)	1.67e6(8.22e5)	1.67e6(8.24e5)	1.67e6(8.26e5)	1.68e6(8.28e5)
	900	1.24e7(2.52e6)	1.28e7(2.53e6)	1.31e7(2.54e6)	1.31e7(2.55e6)	1.32e7(2.55e6)	1.32e7(2.55e6)	1.33e7(2.56e6)	1.33e7(2.56e6)
	20	1.37e7(1.45e6)	1.38e7(1.45e6)	1.38e7(1.46e6)	1.39e7(1.46e6)	1.39e7(1.46e6)	1.39e7(1.47e6)	1.39e7(1.47e6)	1.39e7(1.47e6)
	1100	1.38e7(8.44e5)	1.37e7(8.49e5)	1.36e7(8.54e5)	1.35e7(8.57e5)	1.35e7(8.60e5)	1.34e7(8.64e5)	1.34e7(8.67e5)	1.33e7(8.69e5)
2×10^{16}	900	1.63e7(8.26e6)	1.72e7(8.28e6)	1.76e7(8.29e6)	1.78e7(8.30e6)	1.79e7(8.31e6)	1.80e7(8.32e6)	1.81e7(8.32e6)	1.81e7(8.33e6)
	50	1.89e7(8.34e6)	1.89e7(8.34e6)	1.89e7(8.34e6)	1.89e7(8.34e6)	1.89e7(8.33e6)	1.89e7(8.33e6)	1.89e7(8.32e6)	1.88e7(8.32e6)
	1100	1.87e7(4.40e6)	1.85e7(4.40e6)	1.83e7(4.40e6)	1.82e7(4.40e6)	1.82e7(4.40e6)	1.81e7(4.41e6)	1.80e7(4.41e6)	1.80e7(4.41e6)
	900	1.52e7(1.07e7)	1.61e7(1.07e7)	1.65e7(1.08e7)	1.67e7(1.08e7)	1.69e7(1.08e7)	1.70e7(1.08e7)	1.71e7(1.08e7)	1.71e7(1.08e7)
	80	1.79e7(1.19e7)	1.80e7(1.19e7)	1.81e7(1.19e7)	1.81e7(1.19e7)	1.82e7(1.19e7)	1.82e7(1.19e7)	1.82e7(1.19e7)	1.82e7(1.19e7)
	1100	1.80e7(7.10e6)	1.78e7(7.10e6)	1.77e7(7.08e6)	1.76e7(7.08e6)	1.75e7(7.07e6)	1.75e7(7.07e6)	1.74e7(7.07e6)	1.73e7(7.06e6)
2×10^{17}	900	6.33e7(2.64e7)	6.49e7(2.65e7)	6.68e7(2.65e7)	6.88e7(2.66e7)	7.09e7(2.66e7)	7.30e7(2.66e7)	7.51e7(2.66e7)	7.73e7(2.67e7)
	20	1.32e8(2.71e7)	1.34e8(2.71e7)	1.35e8(2.71e7)	1.36e8(2.72e7)	1.37e8(2.72e7)	1.37e8(2.72e7)	1.38e8(2.72e7)	1.38e8(2.72e7)
	1100	1.41e8(1.71e7)	1.42e8(1.70e7)	1.42e8(1.69e7)	1.43e8(1.69e7)	1.43e8(1.69e7)	1.43e8(1.68e7)	1.44e8(1.68e7)	1.44e8(1.68e7)
	900	4.87e7(8.78e7)	5.30e7(8.87e7)	5.64e7(8.93e7)	5.95e7(9.00e7)	6.25e7(9.00e7)	6.54e7(9.00e7)	6.82e7(9.12e7)	7.09e7(9.14e7)
	50	1.71e8(1.18e8)	1.81e8(1.19e8)	1.85e8(1.19e8)	1.88e8(1.19e8)	1.89e8(1.20e8)	1.90e8(1.20e8)	1.91e8(1.20e8)	1.92e8(1.20e8)
	1100	1.97e8(6.90e7)	1.98e8(6.88e7)	1.98e8(6.86e7)	1.98e8(6.84e7)	1.98e8(6.83e7)	1.98e8(6.81e7)	1.98e8(6.80e7)	1.98e8(6.78e7)
2×10^{18}	900	6.27e7(1.10e8)	6.80e7(1.11e8)	7.15e7(1.12e8)	7.35e7(1.13e8)	7.50e7(1.14e8)	7.64e7(1.15e8)	7.81e7(1.16e8)	7.97e7(1.17e8)
	80	1.55e8(1.33e8)	1.73e8(1.39e8)	1.79e8(1.40e8)	1.83e8(1.40e8)	1.84e8(1.40e8)	1.86e8(1.41e8)	1.86e8(1.41e8)	1.87e8(1.41e8)
		1100	1.92e8(8.81e7)	1.92e8(8.78e7)	1.92e8(8.76e7)	1.92e8(8.74e7)	1.92e8(8.72e7)	1.91e8(8.70e7)	1.91e8(8.68e7)

Table G.9: Look-up table for N_z (μm^{-2}) for arsenic ion implantation.

		Time (min)							
Dose (cm^{-2})	Energy (keV)	15	30	45	60	75	90	105	120
2×10^{14}	900	2.36e5(1.38e4)	2.45e5(1.46e4)	2.52e5(1.52e4)	2.57e5(1.57e4)	2.61e5(1.62e4)	2.65e5(1.65e4)	2.68e5(1.69e4)	2.70e5(1.72e4)
	20	2.90e5(1.03e4)	2.94e5(1.06e4)	2.98e5(1.09e4)	2.98e5(1.10e4)	2.99e5(1.10e4)	3.01e5(1.11e4)	3.02e5(1.11e4)	3.03e5(1.12e4)
	1100	3.09e5(1.00e4)	3.12e5(1.00e4)	3.14e5(1.00e4)	3.16e5(9.98e3)	3.17e5(9.98e3)	3.18e5(9.98e3)	3.19e5(9.91e3)	3.20e5(9.87e3)
	900	1.36e6(2.09e5)	1.44e6(2.19e5)	1.48e6(2.27e5)	1.51e6(2.34e5)	1.53e6(2.40e5)	1.55e6(2.45e5)	1.56e6(2.50e5)	1.58e6(2.54e5)
	20	1.75e6(2.93e5)	1.81e6(3.05e5)	1.84e6(3.13e5)	1.87e6(3.19e5)	1.88e6(3.23e5)	1.90e6(3.27e5)	1.91e6(3.30e5)	1.91e6(3.32e5)
	1100	1.90e6(2.51e5)	1.89e6(2.55e5)	1.88e6(2.57e5)	1.87e6(2.59e5)	1.87e6(2.60e5)	1.87e6(2.61e5)	1.87e6(2.62e5)	1.86e6(2.63e5)
80	900	1.30e6(3.48e5)	1.36e6(3.66e5)	1.39e6(3.78e5)	1.40e6(3.87e5)	1.41e6(3.93e5)	1.42e6(3.98e5)	1.43e6(4.02e5)	1.43e6(4.05e5)
	20	1.54e6(5.92e5)	1.59e6(6.13e5)	1.62e6(6.28e5)	1.65e6(6.40e5)	1.67e6(6.50e5)	1.69e6(6.59e5)	1.71e6(6.66e5)	1.72e6(6.73e5)
	1100	1.78e6(5.6e5)	1.80e6(5.74e5)	1.80e6(5.80e5)	1.81e6(5.84e5)	1.80e6(5.86e5)	1.81e6(5.89e5)	1.80e6(5.91e5)	1.80e6(5.93e5)
	900	2.82e6(1.44e5)	3.01e6(1.50e5)	3.12e6(1.53e5)	3.21e6(1.55e5)	3.27e6(1.57e5)	3.31e6(1.58e5)	3.34e6(1.60e5)	3.37e6(1.61e5)
	20	3.52e6(8.0e4)	3.55e6(8.37e4)	3.55e6(8.57e4)	3.55e6(8.70e4)	3.54e6(8.80e4)	3.53e6(8.87e4)	3.52e6(8.93e4)	3.51e6(8.98e4)
	1100	3.41e6(6.74e4)	3.37e6(6.82e4)	3.35e6(6.87e4)	3.33e6(6.90e4)	3.32e6(6.94e4)	3.30e6(6.96e4)	3.29e6(6.97e4)	3.29e6(6.99e4)
2×10^{15}	900	8.69e6(2.13e6)	9.60e6(2.18e6)	1.04e7(2.21e6)	1.10e7(2.24e6)	1.16e7(2.28e6)	1.22e7(2.27e6)	1.26e7(2.29e6)	1.31e7(2.30e6)
	20	1.87e7(2.67e6)	1.96e7(2.70e6)	2.01e7(2.71e6)	2.04e7(2.72e6)	2.06e7(2.72e6)	2.07e7(2.72e6)	2.08e7(2.72e6)	2.09e7(2.72e6)
	1100	2.11e7(2.02e6)	2.10e7(2.01e6)	2.10e7(2.01e6)	2.09e7(2.01e6)	2.09e7(2.00e6)	2.08e7(2.00e6)	2.08e7(2.00e6)	2.07e7(2.00e6)
	900	9.54e6(4.01e6)	1.04e7(4.15e6)	1.11e7(4.26e6)	1.16e7(4.36e6)	1.20e7(4.45e6)	1.25e7(4.52e6)	1.28e7(4.58e6)	1.32e7(4.63e6)
	80	1.76e7(5.85e6)	1.85e7(5.93e6)	1.90e7(5.97e6)	1.93e7(6.00e6)	1.95e7(6.03e6)	1.97e7(6.05e6)	1.98e7(6.06e6)	1.99e7(6.06e6)
	1100	2.02e7(4.38e6)	2.02e7(4.36e6)	2.02e7(4.35e6)	2.01e7(4.33e6)	2.01e7(4.32e6)	2.00e7(4.30e6)	1.99e7(4.30e6)	1.99e7(4.28e6)
2×10^{16}	900	1.53e7(1.73e6)	1.58e7(1.75e6)	1.64e7(1.76e6)	1.70e7(1.77e6)	1.76e7(1.77e6)	1.82e7(1.78e6)	1.87e7(1.78e6)	1.92e7(1.78e6)
	20	2.83e7(1.21e6)	3.02e7(1.21e6)	3.14e7(1.21e6)	3.22e7(1.20e6)	3.28e7(1.20e6)	3.33e7(1.20e6)	3.37e7(1.20e6)	3.40e7(1.20e6)
	1100	3.55e7(8.40e5)	3.58e7(8.50e5)	3.59e7(8.51e5)	3.60e7(8.53e5)	3.60e7(8.54e5)	3.59e7(8.55e5)	3.59e7(8.56e5)	3.59e7(8.57e5)
	900	3.27e7(2.18e7)	3.26e7(2.16e7)	3.32e7(2.20e7)	3.41e7(2.24e7)	3.50e7(2.29e7)	3.59e7(2.34e7)	3.68e7(2.38e7)	3.76e7(2.43e7)
	20	6.14e7(3.48e7)	7.00e7(3.68e7)	7.69e7(3.84e7)	8.24e7(3.94e7)	8.73e7(4.03e7)	9.16e7(4.09e7)	9.51e7(4.15e7)	9.86e7(4.19e7)
	1100	1.34e8(2.81e7)	1.47e8(2.81e7)	1.57e8(2.81e7)	1.63e8(2.81e7)	1.68e8(2.80e7)	1.72e8(2.80e7)	1.76e8(2.80e7)	1.79e8(2.79e7)
80	900	2.61e7(3.03e7)	2.79e7(2.95e7)	2.93e7(2.96e7)	3.07e7(3.00e7)	3.19e7(3.06e7)	3.30e7(3.11e7)	3.40e7(3.16e7)	3.50e7(3.22e7)
	20	6.01e7(4.51e7)	6.89e7(4.90e7)	7.57e7(5.20e7)	8.15e7(5.44e7)	8.64e7(5.63e7)	9.06e7(5.79e7)	9.44e7(5.93e7)	9.79e7(6.04e7)
	1100	1.32e8(5.00e7)	1.46e8(5.06e7)	1.54e8(5.09e7)	1.61e8(5.12e7)	1.66e8(5.13e7)	1.70e8(5.14e7)	1.73e8(5.14e7)	1.75e8(5.15e7)

Table G.10: Look-up table for t_j (μm) for boron ion implantation.

		Time (min)							
		15	30	45	60	75	90	105	120
Dose (cm^{-2})	Energy (keV)	(Temp. (°C))							
2×10^{14}	900	0.58(0.74)	0.59(0.74)	0.59(0.74)	0.60(0.75)	0.60(0.75)	0.60(0.75)	0.61(0.75)	0.61(0.75)
	20	1000	0.66(0.94)	0.70(0.96)	0.74(0.99)	0.77(1.01)	0.80(1.03)	0.83(1.05)	0.86(1.07)
	1100	1.10(1.57)	1.29(1.69)	1.46(1.80)	1.60(1.90)	1.73(2.00)	1.86(2.08)	1.97(2.17)	2.07(2.25)
	900	0.99(1.08)	0.99(1.08)	1.00(1.08)	1.00(1.08)	1.00(1.09)	1.00(1.09)	1.00(1.09)	1.01(1.09)
	50	1000	1.01(1.22)	1.04(1.24)	1.06(1.26)	1.09(1.28)	1.12(1.30)	1.14(1.31)	1.16(1.33)
	1100	1.37(1.78)	1.53(1.89)	1.67(2.00)	1.80(2.10)	1.92(2.19)	2.03(2.28)	2.13(2.37)	2.23(2.45)
2×10^{15}	900	1.21(1.28)	1.21(1.29)	1.21(1.29)	1.22(1.29)	1.22(1.29)	1.22(1.29)	1.22(1.29)	1.22(1.29)
	80	1000	1.20(1.39)	1.23(1.41)	1.25(1.43)	1.27(1.45)	1.30(1.46)	1.32(1.48)	1.34(1.50)
	1100	1.52(1.92)	1.68(2.03)	1.82(2.13)	1.94(2.23)	2.06(2.32)	2.16(2.41)	2.26(2.49)	2.36(2.57)
	900	0.61(0.75)	0.62(0.75)	0.64(0.76)	0.66(0.76)	0.68(0.76)	0.70(0.76)	0.71(0.77)	0.73(0.77)
	20	1000	0.80(1.09)	0.88(1.12)	0.96(1.16)	1.02(1.19)	1.08(1.22)	1.14(1.25)	1.19(1.28)
	1100	1.45(1.92)	1.70(2.07)	1.92(2.21)	2.10(2.34)	2.27(2.46)	2.42(2.58)	2.56(2.68)	2.69(2.79)
2×10^{16}	900	1.46(1.51)	1.46(1.51)	1.47(1.51)	1.47(1.51)	1.47(1.52)	1.48(1.52)	1.48(1.52)	1.49(1.52)
	50	1000	1.45(1.58)	1.49(1.60)	1.52(1.62)	1.55(1.64)	1.58(1.66)	1.60(1.67)	1.63(1.69)
	1100	1.77(2.13)	1.94(2.26)	2.10(2.39)	2.24(2.51)	2.39(2.62)	2.52(2.73)	2.65(2.84)	2.77(2.94)
	900	1.92(1.95)	1.92(1.95)	1.92(1.95)	1.93(1.95)	1.93(1.95)	1.93(1.95)	1.94(1.95)	1.94(1.96)
	80	1000	1.84(1.95)	1.87(1.96)	1.89(1.98)	1.92(2.00)	1.94(2.01)	1.96(2.03)	1.99(2.05)
	1100	2.10(2.41)	2.25(2.53)	2.39(2.64)	2.52(2.75)	2.63(2.86)	2.75(2.96)	2.86(3.05)	2.96(3.14)
2×10^{17}	900	1.06(1.16)	1.09(1.17)	1.11(1.19)	1.13(1.21)	1.14(1.22)	1.16(1.24)	1.18(1.25)	1.20(1.27)
	20	1000	1.34(1.59)	1.52(1.69)	1.70(1.77)	1.86(1.85)	2.01(1.92)	2.15(1.98)	2.28(2.04)
	1100	2.66(2.67)	3.05(2.92)	3.35(3.14)	3.61(3.34)	3.84(3.52)	4.06(3.68)	4.25(3.84)	4.43(3.98)
	900	1.83(1.86)	1.84(1.86)	1.85(1.87)	1.87(1.88)	1.88(1.88)	1.89(1.89)	1.90(1.90)	1.92(1.90)
	50	1000	1.94(2.04)	2.03(2.13)	2.11(2.21)	2.20(2.30)	2.28(2.39)	2.37(2.49)	2.47(2.60)
	1100	2.78(2.85)	3.14(3.11)	3.44(3.34)	3.70(3.54)	3.94(3.73)	4.15(3.90)	4.35(4.06)	4.53(4.21)
2×10^{18}	900	2.77(2.76)	2.78(2.76)	2.79(2.77)	2.80(2.77)	2.80(2.77)	2.81(2.77)	2.82(2.78)	2.83(2.78)
	80	1000	2.68(2.72)	2.76(2.79)	2.83(2.86)	2.89(2.92)	2.96(2.98)	3.02(3.04)	3.08(3.10)
2×10^{19}	900	3.10(3.14)	3.33(3.32)	3.57(3.52)	3.80(3.70)	4.03(3.88)	4.23(4.05)	4.42(4.20)	4.60(4.36)

Table G.11: Look-up table for t_j (μm) for phosphorus ion implantation.

		Time (min)							
		15	30	45	60	75	90	105	120
		Dose (cm^{-2})	Energy (keV)	Temp. ($^{\circ}\text{C}$)					
2×10^{14}	900	0.28(0.45)	0.29(0.45)	0.29(0.45)	0.29(0.45)	0.30(0.46)	0.30(0.46)	0.31(0.46)	0.31(0.46)
	20	1000	0.38(0.67)	0.43(0.68)	0.47(0.70)	0.51(0.72)	0.55(0.73)	0.59(0.75)	0.62(0.76)
	1100	0.93(1.17)	1.12(1.26)	1.29(1.33)	1.43(1.41)	1.55(1.48)	1.67(1.54)	1.77(1.61)	1.88(1.67)
	900	0.50(0.58)	0.51(0.58)	0.51(0.58)	0.51(0.58)	0.51(0.59)	0.52(0.59)	0.52(0.59)	0.52(0.59)
	20	1000	0.55(0.82)	0.58(0.84)	0.60(0.86)	0.63(0.88)	0.66(0.90)	0.68(0.91)	0.71(0.93)
	1100	0.99(1.42)	1.18(1.52)	1.34(1.62)	1.48(1.71)	1.61(1.79)	1.73(1.87)	1.84(1.95)	1.94(2.02)
2×10^{15}	900	0.66(0.71)	0.66(0.71)	0.66(0.71)	0.67(0.71)	0.67(0.71)	0.67(0.71)	0.67(0.72)	0.67(0.72)
	80	1000	0.69(0.91)	0.71(0.93)	0.73(0.94)	0.75(0.96)	0.77(0.98)	0.79(0.99)	0.81(1.01)
	1100	1.05(1.52)	1.22(1.62)	1.38(1.72)	1.51(1.81)	1.64(1.90)	1.75(1.98)	1.86(2.06)	1.96(2.14)
	900	0.88(0.58)	0.90(0.59)	0.92(0.59)	0.93(0.59)	0.94(0.59)	0.95(0.60)	0.96(0.60)	0.96(0.60)
	20	1000	0.84(0.82)	0.87(0.84)	0.90(0.86)	0.93(0.88)	0.96(0.90)	0.99(0.92)	1.02(0.94)
	1100	1.27(1.48)	1.49(1.59)	1.68(1.69)	1.84(1.78)	1.99(1.87)	2.13(1.96)	2.25(2.04)	2.37(2.12)
2×10^{16}	900	0.69(0.96)	0.73(0.97)	0.77(0.97)	0.81(0.97)	0.84(0.97)	0.86(0.98)	0.88(0.98)	0.90(0.98)
	50	1000	0.76(0.91)	0.80(0.94)	0.85(0.97)	0.89(0.99)	0.93(1.02)	0.97(1.04)	1.00(1.07)
	1100	1.29(1.73)	1.52(1.85)	1.72(1.97)	1.90(2.08)	2.05(2.18)	2.19(2.28)	2.32(2.37)	2.44(2.46)
	900	0.82(1.07)	0.82(1.08)	0.83(1.08)	0.84(1.09)	0.84(1.09)	0.85(1.09)	0.86(1.09)	0.87(1.10)
	80	1000	0.87(0.98)	0.89(1.01)	0.91(1.03)	0.94(1.06)	0.97(1.08)	0.99(1.11)	1.02(1.13)
	1100	1.29(1.80)	1.52(1.93)	1.72(2.05)	1.89(2.16)	2.04(2.27)	2.18(2.37)	2.31(2.47)	2.43(2.56)
2×10^{16}	900	4.98(2.53)	4.98(2.54)	4.98(2.54)	4.98(2.55)	4.98(2.55)	4.98(2.56)	4.98(2.56)	4.98(2.56)
	20	1000	4.61(1.19)	4.59(1.22)	4.60(1.24)	4.61(1.27)	4.63(1.30)	4.65(1.32)	4.66(1.35)
	1100	4.15(1.89)	4.24(2.04)	4.32(2.17)	4.40(2.30)	4.49(2.42)	4.58(2.54)	4.69(2.64)	4.85(2.75)
	900	2.09(4.23)	2.20(4.22)	2.31(4.24)	2.42(4.22)	2.52(4.19)	2.62(4.19)	2.71(4.19)	2.80(4.20)
	50	1000	2.95(2.94)	2.93(2.95)	2.94(2.97)	2.97(2.99)	2.99(3.00)	3.01(3.01)	3.03(3.03)
	1100	2.59(2.03)	2.74(2.23)	2.90(2.41)	3.08(2.58)	3.24(2.73)	3.41(2.88)	3.56(3.01)	3.71(3.13)
80	900	1.27(4.19)	1.42(4.19)	1.56(4.21)	1.69(4.17)	1.81(4.22)	1.93(4.22)	2.03(4.20)	2.13(4.19)
	1000	2.21(3.09)	2.31(3.11)	2.35(3.12)	2.39(3.14)	2.43(3.15)	2.46(3.17)	2.49(3.18)	2.52(3.20)
	1100	2.21(2.09)	2.47(2.31)	2.72(2.50)	2.94(2.67)	3.13(2.83)	3.31(2.98)	3.48(3.12)	3.64(3.25)

Table G.12: Look-up table for t_1 (μm) for arsenic ion implantation.

		Time (min)							
		15	30	45	60	75	90	105	120
Dose (cm ⁻²)	Energy (keV)	Temp. (°C)							
2 × 10 ¹⁴	900	0.15(0.10)	0.16(0.10)	0.16(0.10)	0.16(0.10)	0.16(0.10)	0.16(0.10)	0.16(0.10)	0.16(0.10)
	20	0.16(0.12)	0.17(0.13)	0.18(0.14)	0.19(0.14)	0.20(0.15)	0.21(0.15)	0.22(0.16)	0.23(0.17)
	1100	0.34(0.27)	0.41(0.30)	0.46(0.34)	0.51(0.37)	0.56(0.39)	0.60(0.42)	0.64(0.44)	0.67(0.46)
2 × 10 ¹⁵	900	0.35(0.25)	0.35(0.25)	0.35(0.25)	0.35(0.25)	0.35(0.25)	0.35(0.25)	0.35(0.25)	0.35(0.25)
	50	0.35(0.29)	0.35(0.29)	0.36(0.29)	0.36(0.30)	0.36(0.30)	0.36(0.30)	0.37(0.30)	0.37(0.31)
	1100	0.43(0.40)	0.51(0.45)	0.57(0.50)	0.63(0.55)	0.68(0.58)	0.73(0.62)	0.77(0.65)	0.81(0.69)
2 × 10 ¹⁶	900	0.51(0.41)	0.51(0.41)	0.51(0.41)	0.51(0.41)	0.51(0.41)	0.51(0.41)	0.51(0.41)	0.51(0.41)
	80	0.52(0.44)	0.52(0.44)	0.52(0.45)	0.52(0.45)	0.52(0.45)	0.52(0.45)	0.53(0.45)	0.53(0.45)
	1100	0.55(0.51)	0.58(0.54)	0.62(0.58)	0.66(0.62)	0.71(0.66)	0.75(0.69)	0.79(0.73)	0.83(0.76)
2 × 10 ¹⁵	900	0.19(0.11)	0.19(0.11)	0.19(0.11)	0.19(0.11)	0.19(0.11)	0.19(0.11)	0.19(0.11)	0.19(0.11)
	20	0.20(0.15)	0.22(0.16)	0.24(0.17)	0.26(0.18)	0.27(0.18)	0.29(0.19)	0.30(0.20)	0.31(0.21)
	1100	0.44(0.33)	0.53(0.38)	0.60(0.43)	0.67(0.47)	0.72(0.50)	0.77(0.53)	0.82(0.56)	0.86(0.59)
2 × 10 ¹⁶	900	0.39(0.23)	0.39(0.23)	0.39(0.23)	0.39(0.23)	0.39(0.23)	0.39(0.23)	0.39(0.23)	0.39(0.23)
	50	1000	0.39(0.31)	0.40(0.31)	0.40(0.31)	0.40(0.31)	0.41(0.32)	0.42(0.32)	0.44(0.32)
	1100	0.61(0.49)	0.72(0.56)	0.81(0.62)	0.89(0.67)	0.96(0.72)	1.02(0.76)	1.08(0.80)	1.13(0.84)
2 × 10 ¹⁶	900	0.55(0.38)	0.55(0.38)	0.55(0.38)	0.55(0.38)	0.55(0.38)	0.55(0.38)	0.55(0.38)	0.55(0.38)
	80	1000	0.55(0.44)	0.55(0.44)	0.55(0.45)	0.55(0.45)	0.56(0.45)	0.56(0.45)	0.56(0.46)
	1100	0.63(0.57)	0.73(0.63)	0.82(0.69)	0.90(0.75)	0.96(0.80)	1.02(0.85)	1.08(0.89)	1.13(0.93)
2 × 10 ¹⁶	900	0.20(0.08)	0.20(0.09)	0.20(0.09)	0.20(0.10)	0.20(0.10)	0.20(0.10)	0.20(0.10)	0.21(0.11)
	20	0.29(0.17)	0.33(0.19)	0.37(0.20)	0.40(0.22)	0.43(0.23)	0.45(0.24)	0.48(0.25)	0.50(0.26)
	1100	0.67(0.43)	0.80(0.50)	0.90(0.55)	0.99(0.60)	1.07(0.65)	1.13(0.69)	1.20(0.73)	1.26(0.77)
2 × 10 ¹⁶	900	0.46(0.29)	0.46(0.29)	0.46(0.29)	0.46(0.29)	0.46(0.29)	0.46(0.29)	0.46(0.29)	0.46(0.29)
	50	1000	0.47(0.35)	0.47(0.38)	0.48(0.41)	0.51(0.44)	0.55(0.48)	0.59(0.50)	0.63(0.53)
	1100	0.92(0.74)	1.11(0.84)	1.27(0.93)	1.40(1.00)	1.51(1.07)	1.61(1.13)	1.71(1.18)	1.79(1.24)
2 × 10 ¹⁶	900	0.65(0.48)	0.65(0.48)	0.65(0.48)	0.65(0.48)	0.65(0.48)	0.65(0.48)	0.65(0.48)	0.65(0.48)
	80	1000	0.65(0.53)	0.66(0.53)	0.66(0.54)	0.66(0.54)	0.66(0.55)	0.67(0.57)	0.68(0.61)
	1100	0.94(0.85)	1.12(0.96)	1.26(1.05)	1.39(1.14)	1.51(1.21)	1.61(1.28)	1.70(1.34)	1.79(1.40)

Appendix H

Code

A great deal of the work presented here depends on numerical modeling and design optimization code. The most important components of the code are reproduced here for posterity. However, all of the code is also available online (<http://github.com/jcdoll>, <http://microsystems.stanford.edu/piezod> or by contacting me directly). Other useful but less crucial pieces of code (e.g. analyzing thermomechanical noise data) are also available online.

H.1 TSuprem simulations

The following three files were used to generate the ion implantation lookup tables used for the case studies. The first two files (`simulation.template` and `simulationControl.py`) were used to run the TSuprem simulations. Simulations were performed by substituting appropriate parameter values into the template and then running the simulation from the command line. Both steps were performed by `simulationControl.py`. Performing all 1296 simulations (3 dopants, 3 energies, 3 doses, 3 temperatures, 8 times, and 2 anneal conditions) took about 12 hours on the standard Stanford computing cluster where TSuprem4 was installed.

The simulations generated 2592 output text files containing the dopant concentration before and after the anneal. These text files were postprocessed (`postProcessTables.m` and `calculateProfileProperties.m`) in order to generate a `.mat` file that could be used for design

optimization. The output .mat file, including the raw post-anneal dopant concentration profiles ($n(z)$), was about 4MB. The lookup table was used to calculate β_1^* , β_2^* , R_s , N_z , x_j and $n(z)$ through linear interpolation.

As discussed in Section 7.3, the simulations are currently one-dimensional. A possible direction for future work would be extending the simulations to two dimensions in order to account for lateral dopant diffusion.

H.1.1 simulation.template

```
# Set up the 1D mesh
# Assumes a 10um device layer and 1um BOX
line x loc=0.0 tag=left
line x loc=1.0 tag=right
line y loc=0.0 spacing=0.01 tag=top
line y loc=5.0 spacing=0.01 tag=bot

region silicon xlo=left xhi=right ylo=top yhi=bot
initialize <100> impurity=${dopantBackground} i.resistivity=10
10 method vertical ddc.full pd.5str

# 1) Grow the 250A protection oxide
diffusion time=20 temperature=800 inert
diffusion time=20 temperature=800 t.final=900 inert
diffusion time=10 temperature=900 dry02
diffusion time=9 temperature=900 wet02
diffusion time=10 temperature=900 dry02
diffusion time=20 temperature=900 t.final=800 inert

20 # List the layers
select z=doping
print.1d x.val=0 layers

# 2) Ion implant
implant ${dopantShort} dose=${dose} energy=${energy} +
tilt=7 impl.tab=tr.${dopantLong} damage

# Output the total dopant concentration (unactivated)
select z=${dopantLong}
30 print.1d out.file=simulationOutputs/${outputFileName}_preAnneal.out

# List the layers
select z=doping
print.1d x.val=0 layers
```

```

# 4) Oxidize (optional) and anneal
${oxideStripCmd}

diffusion time=45 temperature=800 t.final=${temp} inert
diffusion time=${dryOxidationTime} temperature=${temp} dryO2
40 diffusion time=${wetOxidationTime} temperature=${temp} wetO2
diffusion time=${dryOxidationTime} temperature=${temp} dryO2
diffusion time=${time} temperature=${temp} inert
diffusion time=45 temperature=${temp} t.final=800 inert

# Output the post-anneal profile
# Specify the maximum temperature for correct solid solubility handling
select z=abs(doping) temperat=${temp}
print.1d out.file=simulationOutputs/${outputFileName}.out

50 # List the layers
select z=doping temperat=${temp}
print.1d x.val=0 layers

# 6) Analysis (optional)
# Sheet resistance
electric resistance x.val=0

# Peak doping level
select z=abs(doping) temperat=${temp}
60 extract silicon x.val=0 maximum val.extr assign name=Np0

# Junction depth
select z=doping temperat=${temp}
extract silicon x.val=0 value=0 d.extrac assign name=Dj

# Total carriers
select z=abs(doping) temperat=${temp}
extract silicon p1.y=0 p2.y=@Dj integral assign name=Ncm2

```

H.1.2 simulationControl.py

```
import sys, string, subprocess

# Parameter values
dopantValues = ['arsenic', 'boron', 'phosphorus']
doseValues = [2e14, 2e15, 2e16]
energyValues = [20, 50, 80]
timeValues = [15, 30, 45, 60, 75, 90, 105, 120]
tempValues = [900, 1000, 1100]

10 # Clear out all of the old simulation files
    subprocess.call('rm -rf *.out', shell=True)
    subprocess.call('rm -rf simulationInputs/*', shell=True)
    subprocess.call('rm -rf simulationOutputs/*', shell=True)

# Generate the tsuprem input files and run tsuprem on them
ii = 1
for dopantLong in dopantValues:
    if dopantLong == 'boron':
        dopantShort = 'B'
        dopantBackground = 'phosphorus'
    20 elif dopantLong == 'phosphorus':
        dopantShort = 'P'
        dopantBackground = 'boron'
    elif dopantLong == 'arsenic':
        dopantShort = 'As'
        dopantBackground = 'boron'
    else:
        print 'Unknown dopant type!'
        sys.exit(0)

30 for dose in doseValues:
    for energy in energyValues:
        for time in timeValues:
            for temp in tempValues:
```

```

print 'Iteration %d of %d' % (ii, 2*len(dopantValues)*
    len(doseValues)*len(energyValues)*len(timeValues)*len(tempValues))
ii += 1

# First simulate without passivation oxide
dryOxidationTime = 0
wetOxidationTime = 0
oxideStripCmd = ''

# Build the simulation input file through substitution
fileName = '%s_%s_%s_%s_noOxide' % (dopantShort,
    dose, energy, temp, time)
templateString = open('simulation.template', 'r').read()
s = string.Template(templateString)
fileContents = s.substitute(outputFileName=fileName, dose=dose,
    energy=energy, time=time, temp=temp, dopantShort=dopantShort,
    dopantLong=dopantLong, dopantBackground=dopantBackground,
    dryOxidationTime=dryOxidationTime,
    wetOxidationTime=wetOxidationTime,
    oxideStripCmd = oxideStripCmd)

# Save the file to disk
simulationFileName = ''.join(['simulationInputs/', fileName, '.input'])
f = open(simulationFileName, 'w')
f.write(fileContents)
f.close()

# Run the simulation
cmd = ''.join(['tsuprem4 ', simulationFileName])
subprocess.call(cmd, shell=True)

# Resimulate with an oxidation step immediately before the anneal
dryOxidationTime = 5
70 if temp == 900:
    wetOxidationTime = 66
elif temp == 1000:

```

```
    wetOxidationTime = 15
elif temp == 1100:
    wetOxidationTime = 5
oxideStripCmd = 'etch oxide all'

# Build the simulation input file through substitution
80   fileName = '%s_%s_%s_%s_%s_oxide' % (dopantShort, dose,
                                             energy, temp, time)
templateString = open('simulation.template', 'r').read()
s = string.Template(templateString)
fileContents = s.substitute(outputFileName=fileName, dose=dose,
                            energy=energy, time=time, temp=temp, dopantShort=dopantShort,
                            dopantLong=dopantLong, dopantBackground=dopantBackground,
                            dryOxidationTime=dryOxidationTime, wetOxidationTime=wetOxidationTime,
                            oxideStripCmd = oxideStripCmd)

# Save the file to disk
90   simulationFileName = ''.join(['simulationInputs/', fileName, '.input'])
f = open(simulationFileName, 'w')
f.write(fileContents)
f.close()

# Run the simulation
cmd = ''.join(['tsuprem4 ', simulationFileName])
subprocess.call(cmd, shell=True)
```

H.1.3 postProcessTables.m

```

function postProcessTables()
clear all
close all
clc

% size = (3, 3, 3, 3, 8, 2) -> (dopant, dose, energy, temp, time, oxidation)
dopants = {'B', 'P', 'As'};
dopantsLong = {'boron', 'phosphorus', 'arsenic'};
doses = [2e14, 2e15, 2e16];
10 energies = [20 50 80];
temp = [900 1000 1100];
times = [15 30 45 60 75 90 105 120];

% Setup the z interpolation grid
% All data will be calculated over this grid
z = 0:0.01:5;

% Allocate space
Beta1 = zeros(3, 3, 3, 3, 8, 2);
20 Beta2 = zeros(3, 3, 3, 3, 8, 2);
Rs = zeros(3, 3, 3, 3, 8, 2);
Nz_total = zeros(3, 3, 3, 3, 8, 2);
Nz = zeros(3, 3, 3, 3, 8, 2);
Xj = zeros(3, 3, 3, 3, 8, 2);
n = zeros(length(z), 3, 3, 3, 3, 8, 2);

ImplantDopants = [1 2 3];
ImplantEnergies = energies;
ImplantDoses = doses;
30 AnnealTemps = temp;
AnnealTimes = times;
AnnealOxidation = [1 2];

currentCount = 1;
for ii = 1:length(dopants)

```

```

for jj = 1:length(doses)
    for kk = 1:length(energies)
        for ll = 1:length(temp)
            for mm = 1:length(times)
                40      dopant = dopants{ii};
                dopantLong = dopantsLong{ii};
                dose = doses(jj);
                energy = energies(kk);
                temp = temps(ll);
                time = times(mm);

                % Output a status update
                fprintf('Processing %s, %0.1g/sq cm, %dkeV, %dC, %dmin\n', ...
                    dopant, dose, energy, temp, time);
                fprintf('%d of %d\n\n', currentCount, 2*length(dopants)* ...
                    length(doses)*length(energies)*length(temp)*length(times));
                50      % Generate the filenames
                doseExponent = floor(log10(dose));
                dosePrefactor = round(10^(rem(log10(dose), doseExponent)));
                name = sprintf('%s_%de+%d_%d_%d', dopant, ...
                    dosePrefactor, doseExponent, energy, temp, time);
                filename_noOxide = ['simulationOutputs/' name '_noOxide.out'];
                filename_oxide = ['simulationOutputs/' name '_oxide.out'];

                60      % No oxide case
                fprintf('%s\n', filename_noOxide);
                nInterp = loadAndCleanupData(filename_noOxide);
                [beta1tmp, Beta1tmp, Beta2tmp, Nztmp, Nz_totaltmp, Rstmp, Xjtmp] = ...
                    calculateProfileProperties(dopantLong, z*1e-6, nInterp);
                Rs(ii, jj, kk, ll, mm, 1) = Rstmp;
                Xj(ii, jj, kk, ll, mm, 1) = Xjtmp;
                Beta1(ii, jj, kk, ll, mm, 1) = Beta1tmp;
                Beta2(ii, jj, kk, ll, mm, 1) = Beta2tmp;
                Nz(ii, jj, kk, ll, mm, 1) = Nztmp;
                Nz_total(ii, jj, kk, ll, mm, 1) = Nz_totaltmp;
                70      n(:, ii, jj, kk, ll, mm, 1) = nInterp;

```

```

% Oxide case
fprintf('%s\n', filename_oxide);
nInterp = loadAndCleanupData(filename_oxide);
[betatmp, Beta1tmp, Beta2tmp, Nztmp, Nz_totaltmp, Rstmp, Xjtmp] = ...
    calculateProfileProperties(dopantLong, z*1e-6, nInterp);
Rs(ii, jj, kk, ll, mm, 2) = Rstmp;
Xj(ii, jj, kk, ll, mm, 2) = Xjtmp;
Beta1(ii, jj, kk, ll, mm, 2) = Beta1tmp;
Beta2(ii, jj, kk, ll, mm, 2) = Beta2tmp;
Nz(ii, jj, kk, ll, mm, 2) = Nztmp;
Nz_total(ii, jj, kk, ll, mm, 2) = Nz_totaltmp;
n(:, ii, jj, kk, ll, mm, 2) = nInterp;

% Move to the next iteration
currentCount = currentCount + 2;
end
90 end
end
end

ratio = Nz./Nz_total;
ratio = reshape(ratio, 1, 3*3*3*3*8*2);
fprintf('Nz/Nz_total: %.2f to %.2f (mean/median/std = %.2f/.2f/.2f)\n', ...
    min(ratio), max(ratio), mean(ratio), median(ratio), std(ratio));

100 % Output data
save lookupTable z Rs Xj Beta1 Beta2 Nz Nz_total n ImplantDopants ...
    ImplantEnergies ImplantDoses AnnealTemps AnnealTimes AnnealOxidation;

function nInterp = loadAndCleanupData(filename)
[zRaw, nRaw, material] = textread(filename, '%f %f %s', 'headerlines', 1);
siliconIndices = find(strcmp(material, 'silicon'));
zRaw = zRaw(siliconIndices);
nRaw = nRaw(siliconIndices);
zRaw = zRaw - zRaw(1);

```

```
110 nInterp = interp1(zRaw, nRaw, z, 'pchip'); % Arrange data on a standard grid
end
end
```

H.1.4 calculateProfileProperties.m

```

function [beta, beta1, beta2, Nz, Nz_total, Rsheet, Xj] = ...
    calculateProfileProperties(dopantType, z, n)
% Inputs:
% dopantType - 'phosphorus', 'boron', or 'arsenic'
% z - depth from surface (m)
% n - dopant concentration (per cc)
T = 300;

% Find the junction depth (concentration dips below end value)
10 junctionIndex = find(n <= n(end), 1);

% If the doping is constant...
if junctionIndex == 1
    junctionIndex = length(n);
end
Xj = z(junctionIndex);
t = max(z);

% Adjust the carrier concentration beyond the junction
20 ni = 1e10;
n(junctionIndex+1:end) = ni^2./n(junctionIndex+1:end);

% Calculate the mobility
[rho, mu] = resistivity(dopantType, T, n);
sigma = 1./rho;
P = richterModel(T, n);

% Calculate Beta*, beta1*, beta2*
zBeta = (t/2 - z)*1e2; % integrated from -t/2 to +t/2, output in cm
30 beta = 2/(t*1e2)*trapz(zBeta, sigma.*P.*zBeta)/trapz(zBeta, sigma);
calcRange = 1:junctionIndex;
beta1 = trapz(z(calcRange), P(calcRange).*sigma(calcRange)) ...
    /trapz(z(calcRange), sigma(calcRange));
beta2 = 1e6*trapz(z(calcRange), P(calcRange).*sigma(calcRange)) ...
    .*z(calcRange))/trapz(z(calcRange), sigma(calcRange));

```

```
% Calculate the sheet resistance and use to calculate the current density
Rsheet = 1/trapz(z*1e2, sigma); % convert z from m to cm
R = Rsheet;
40 I = 1/Rsheet;
J = sigma/1e2; % J = sigma*Em = A/cm^2

% Calculate the total and effective number of carriers
Nz_total = trapz(z, n*1e6);
Nz = 1e4*trapz(z*1e2, n.*mu)^2/trapz(z*1e2, n.*mu.^2);

end
```

H.1.5 richterModel.m

```
function P = richterModel(T, N)

% Transform from nx1 and 1xm to nxm matrices
T = T*ones(1, length(N));
N = ones(size(T,1), 1)*N;

Tn = T/300;
Nb = 6e19;
Nc = 7e20;
10 theta = 0.9;
alpha = 0.43;
gamma = 1.6;
beta = 0.1;
eta = 3;
P = Tn.^-theta.* (1 + Tn.^-beta.* (N./Nb).^alpha + Tn.^-eta.* (N./Nc).^gamma).^-1;
end
```

H.1.6 resistivity.m

```

function [rho, mu] = resistivity(dopantType, T, dopantConc)
k_b_eV = 8.617343e-5; % eV/K
q = 1.60218e-19; % Coulombs

Tnorm = T/300;
Eg = 1.170-(4.730e-4*T.^2)./(T+636);

ni = 2.9135e15.*T.^1.6.*exp(-Eg./(2*k_b_eV.*T));

10 switch dopantType
    case 'phosphorus'
        mumax = 1441;
        c = 0.07;
        gamma = 2.45;
        mu0d = 62.2.*Tnorm.^-0.7;
        mu0a = 132.*Tnorm.^-1.3;
        mu1d = 48.6.*Tnorm.^-0.7;
        mu1a = 73.5.*Tnorm.^-1.25;
        Cr1 = 8.5e16.*Tnorm.^3.65;
        Cr2 = 1.22e17.*Tnorm.^2.65;
        Cs1 = 4e20;
        Cs2 = 7e20;
        alpha1 = .68;
        alpha2 = .72;

        n = (dopantConc./2)+((dopantConc./2).^2 + ni.^2).^0.5;
        p = ni.^2./n;

        ND = dopantConc;
        NA = 0;
    case 'arsenic'
        mumax = 1441;
        c = 0.07;
        gamma = 2.45;
        mu0d = 55.*Tnorm.^-0.6;

```

```

mu0a = 132.*Tnorm.^-1.3;
mu1d = 42.4.*Tnorm.^-0.5;
mu1a = 73.5.*Tnorm.^-1.25;
Cr1 = 8.9e16.*Tnorm.^3.65;
40    Cr2 = 1.22e17.*Tnorm.^2.65;
Cs1 = 2.9e20;
Cs2 = 7e20;
alpha1 = .68;
alpha2 = .72;

n = (dopantConc./2)+((dopantConc./2).^2 + ni.^2).^.5;
p = ni.^2./n;

ND = dopantConc;
50    NA = 0;
case 'boron'
  mumax=470.5;
  c=0.0;
  gamma=2.16;
  mu0d=90.0*Tnorm.^-1.3;
  mu0a=44.0*Tnorm.^-0.7;
  mu1d=28.2*Tnorm.^-2.0;
  mu1a=28.2*Tnorm.^-0.8;
  Cr1=1.3e18*Tnorm.^2.2;
60    Cr2=2.45e17*Tnorm.^3.1;
  Cs1=1.1e18*Tnorm.^6.2;
  Cs2=6.1e20;
  alpha1=.77;
  alpha2=.719;

p=(dopantConc./2)+((dopantConc./2).^2+ni.^2).^.5;
n=ni.*ni./p;

NA=dopantConc;
70    ND = 0;
end

```

```
mu0 = (mu0d.*ND+mu0a.*NA)./(ND+NA);
mu1 = (mu1d.*ND+mu1a.*NA)./(ND+NA);
muL = mumax.*Tnorm.^(-gamma+c.*Tnorm);
second = (muL-mu0)./(1+(ND./Cr1).^alpha1 + (NA./Cr2).^alpha2);
third = mu1./(1+(ND./Cs1+NA./Cs2).^2);

mu = mu0+second-third;
80 sigma = q.*(mu.*n + mu.*p);
rho = 1./sigma;
end
```

H.2 Optimization code

The Matlab code used to model and optimize the cantilever designs (piezoD) is reproduced here for completeness. The latest version can always be found at either <http://microsystems.stanford.edu/piezod>, <http://github.com/jcdoll/PiezoD>, or if all else fails, by emailing me.

The code is written in object oriented Matlab and requires at least Matlab R2008a with the optimization toolbox. Cantilevers based upon silicon piezoresistors (diffusion, epitaxy, ion implantation and polysilicon), metal strain gauges and piezoelectric films are supported in both end loaded and surface stress loaded configurations. The thermal and multilayered beam models are integrated into the code for both sensor and actuator design.

The code consists of an abstract base class (`cantilever.m`) that is inherited by several cantilever subclasses (e.g. `cantileverDiffusion.m`). The user instantiates a subclass based upon the type of cantilever that they are designing. Each subclass has several additional instantiation parameters on top of the base class parameters (e.g. diffusion temperature and time), implements anything that is specific to the particular cantilever type (e.g. calculating the dopant concentration profile) and overrides several of the optimization functions so that everything fits together nicely. The cantilever base class includes all of the thermal modeling, noise and sensitivity functions and can model the effect of an actuator at the base of the sensor (i.e. the configuration used in the self-sensing and self-actuating probes that I developed). The code theoretically supports the optimization of the actuator, but the actuator design is typically so simple that it is quicker and easier to simple test several possible actuator lengths.

The optimization is a straightforward application of `fmincon`, Matlab's built-in generation optimization routine. The code supports arbitrary optimization constraints that are passed in as parameters to the optimization routine (see `sampleCode.m`) and additional nonlinear constraints can be added to the cantilever base class. Depending on the speed of your computer and the degree of complexity in the optimization (which can be controlled using flags, e.g. whether to account for variation in the thermal conductivity with temperature) an optimized cantilever design can be generated in as little as 30 seconds.

The code supports diffused (`cantileverDiffusion.m`), epitaxial (`cantileverEpitaxy.m`) and

ion implanted (`cantileverImplantation.m`) single crystal silicon piezoresistive cantilevers. Several other cases (polysilicon, metal strain gauges, piezoelectric films) are supported as well but the code (`cantileverPoly.m` and `cantileverPiezoelectric.m`) is not reproduced here because the results are not essential to the thesis and is not particularly well integrated into the rest of the code base. The polysilicon and piezoelectric design code can be found online. A significant amount of additional code (e.g. calculating $f_0/Q/k_c$ from thermomechanical noise data) that is not essential to the thesis but could potentially be quite useful to the reader can similarly be found online at <http://github.com/jcdoll>.

H.2.1 sampleCode.m

```
% sample_code.m
% This file demonstrates some of the basic features of piezoD
% and can also be used to verify that you have properly installed
% the code on your system.
%
% Check out cantilever.m for the many additional features included
% in the code. Examples include:
% - arbitrary parameter constraints (e.g. max temperature, diffusion temp)
% - piezoelectric and thermal actuator modeling
10 % - initial device deflections from intrinsic stress (Monte Carlo)
% - alternative optimization goals (e.g. surface stress)
% - alternative instrumentation amplifier noise characteristics

clear all
close all
clc

% Turn on multiprocessor support (if available)
if matlabpool('size') == 0
20 matlabpool
end

% Compare the model with Harley's 89 nm thick epitaxial cantilever (1999)
freq_min = 10;
freq_max = 1000;
l = 300e-6;
w = 44e-6;
t = 89e-9;
l_pr_ratio = 45/300;
30 v_bridge = 5;
doping_type = 'boron';
concentration_initial = 4e19;
t_pr_ratio = 30/89;
exampleCantilever = cantileverEpitaxy(freq_min, freq_max, l, w, t, l_pr_ratio, ...
v_bridge, doping_type, concentration_initial, t_pr_ratio);
```

```

exampleCantilever.fluid = 'vacuum'; % other options: 'air', 'water', 'arbitrary'
exampleCantilever.thermal_modeling = 'approx';
exampleCantilever.number_of_piezoresistors = 1;
exampleCantilever.print_performance();
40 exampleCantilever.print_performance_for_excel(); % tab delimited output
[TMax, TTip] = exampleCantilever.calculateMaxAndTipTemp();

% Setup optimization constraints
parameter_constraints = {{'min_t', 'max_t', 'min_w', 'max_v_bridge'}, ...
    {1e-6, 1e-6, 5e-6, 10}};
nonlinear_constraints = {{'omega_min_hz', 'max_power', 'min_k', 'max_k'}, ...
    {freq_max, 2e-3, 1e-3, 1e1}};
goal = cantilever.goalForceResolution;

50 % Optimize once from the current cantilever design
% Use optimize_performance() to randomize the optimization starting point
exampleCantilever = exampleCantilever.optimize_performance_from_current(parameter_constraints, ...
    nonlinear_constraints, goal);
exampleCantilever.print_performance();
exampleCantilever.plot_noise_spectrum();

% Diffusion Example
diffusion_time = 20*60; % 20 minutes
diffusion_temp = 800 + 273; % at 800C (1073K)
60 exampleCantilever = cantileverDiffusion(freq_min, freq_max, ...
    l, w, t, l_pr_ratio, v_bridge, doping_type, ...
    diffusion_time, diffusion_temp);
exampleCantilever.fluid = 'vacuum';
exampleCantilever.number_of_piezoresistors = 4;
exampleCantilever.print_performance();

% Ion implantation example
annealing_time = 20*60; % 20 minutes
annealing_temp = 950 + 273; % 950C
70 implantation_energy = 30; % keV
implantation_dose = 5e15; % 5e15/sq. cm
annealing_type = 'inert'; % or 'oxide'

```

```
exampleCantilever = cantileverImplantation(freq_min, freq_max, ...
    l, w, t, l_pr_ratio, v_bridge, doping_type, ...
    annealing_time, annealing_temp, annealing_type, ...
    implantation_energy, implantation_dose);
exampleCantilever.fluid = 'water';
exampleCantilever.number_of_piezoresistors = 4;
exampleCantilever.print_performance();
80 exampleCantilever.plotTempProfile();
```

H.2.2 cantilever.m

```
% Abstract cantilever base class
% Used by subclasses to implement common cantilever features
% The code is formatted to be < 85 columns wide for printing reproduction
%
% Subclasses must implement the following methods at a minimum:
% doping_profile
% doping_optimization_scaling
% doping_cantilever_from_state
% doping_current_state
10 % doping_initial_conditions_random
% doping_optimization_bounds
classdef cantilever
properties
    % Basic parameters that are absolutely required
    freq_min; % Minimum measured frequency, Hz
    freq_max; % Maximum measured frequency, Hz
    l; % overall cantilever/sensor length, m
    w; % overall cantilever width, m
    t; % overall cantilever thickness, m
    l_pr_ratio; % piezoresistor length ratio
    v_bridge; % bridge bias, V
    doping_type; % Dopant species: 'phosphorus', 'boron', or 'arsenic'

    % Optional parameters for general operation
    fluid; % 'vacuum', 'air', 'water', or 'arbitrary'. Default = 'air'
    number_of_piezoresistors; % # of PRs in the circuit. Default = 2.
    rms_actuator_displacement_noise = 1e-12; % Displacement noise from mounting, m
    default_alpha = 1e-5; % Hooge noise parameter, unitless
    amplifier; % Amplifier used: 'INA103' or 'AD8221'. Default = 'INA103'.
20    T; % Ambient temp, K. Default = 300.
    T_ref; % Reference temp for calculating thermal expansion, K. Default = 300
    tip_mass; % Tip mass loading (e.g. bead or sharp tip). Default = 0.
    R_base; % Thermal resistance at the cantilever base, K/W. Default = 10e3
    R_contact; % Excess resistance per contact, default = 170 ohm
```

```

% Include temp dependence in calculating electrical resistivity
% and thermal conductivity of silicon. 'yes' or 'no'. Default = 'no'
temperature_dependent_properties;

40   % Degree of accuracy to use in thermal modeling (e.g. see johnson_PSD())
% Options: 'none', 'approx' or 'exact'. Default = 'none'
thermal_modeling;

% Optional parameters for modeling arbitrary fluids (damping and heat transfer)
rho_arb; % Density of fluid
eta_arb; % Viscosity of fluid
k_arb; % Thermal conductivity of fluid
h_method; % How to calculate h_eff: 'fixed' or 'calculate'

50   % Effective convection coefficient of fluid, W/m^2-K
% Used in case of h_method = 'fixed' and fluid = 'arbitrary'
h_arb;

% Optional parameters for modeling a stiffener/actuator at the
% base of the device. Mechanical/thermal impact is included in the code.

% Options = 'none', 'step', 'thermal', 'piezoelectric'. Default = 'None'
cantilever_type;

60   % 'step' stack = silicon/oxide/metal (t/t_oxide/t_a)
% 'thermal' stack = silicon/oxide/metal (t/t_oxide/t_a)
% 'piezoelectric' stack = silicon/oxide/seed/metal/piezoelectric/metal
%           (t/t_oxide/t_a_seed/t_electrode_bottom/t_a/t_electrode_top)

l_a; % Overall length of the step/actuator, m
w_a; % Overall width of the step/actuator, m
w_a_active; % Active width of the step/actuator, m
t_a; % Thickness of the step/actuator, m
v_actuator; % Bias voltage for either thermal or PE actuation, V
70   R_heater; % Resistance of the thermal actuator heater, ohm
t_oxide; % Oxide layer thickness, m
t_a_seed; % Thickness of PE seed layer, m

```

```
t_electrode_bottom; % Thickness of the bottom PE metal electrode, m
t_electrode_top; % Thickness of the top PE metal electrode, m
d31_manual; % Allows manual specification of the PE d31 coefficient, pm/V

% Gap at the end of the actuator before the start of the sensor, m
% Typically on the order of 20 microns to allow for contacts, etc.
l_a_gap;
80
% Type of the metal used in the actuator/step
% Options: 'aluminum', 'titanium' or 'molybdenum'
metal_type;

% Probe deflection due to residual stress is modeled in the code
% Options for modeling: 'nominal' or 'random'
% The latter option performs a Monte Carlo simulation to predict
% the tip range from the input stress range (e.g. see sigma_si_range)
film_stress;
90 end

% Constants - can be referred to with cantilever.variableName
properties (Constant)

% Physical constants
k_b = 1.38e-23; % J/K
k_b_eV = 8.617343e-5; % eV/K
q = 1.60218e-19; % Coulombs
h_bar = 1.055e-34; % J-sec
100
% Define the number of points to use in discretized calculations
numFrequencyPoints = 1000; % For noise spectra plotting

% Number of points along the cantilever for calculating
% deflections and temperature profiles
numXPoints = 800;

% Number of points through the cantilever depth for calculating the
% dopant profile, sheet resistance, number of carriers, thermal conductivity
```

```

110    numZPoints = 200;

    % Number of Monte Carlo iterations for calculating the tip
    % deflection distribution
    numRandomStressIterations = 10;

    % Number of optimization iterations to perform without
    % convergence (to within 0.1%) before accepting
    % the best answer so far
    numOptimizationIterations = 20;

120

    % The width between the two cantilever legs. Affects the PR
    % resistance and sensitivity, but only if l_pr is small
    air_gap_width = 2e-6;

    % Standard fluid properties
    k_water = 0.610; % W/m-K
    rho_water = 996.6; % kg/m^3
    eta_water = 7.98e-4; % Pa-sec
130    h_water = 49218; % W/m^2-K

    k_air = 0.0262; % W/m-K
    rho_air = 1.164; % kg/m^3
    eta_air = 17e-6; % Pa-sec
    h_air = 2098; % W/m^2-K

    % For vacuum, use small but finite values for numerical stability
    k_vacuum = 1e-6; % W/m-K
    rho_vacuum = 1e-6; % kg/m^3
    eta_vacuum = 1e-6; % Pa-sec
140    h_vacuum = 1e-6; % W/m^2-K

    % Thermal conductivities, W/m-K
    % AlN: "Process-dependent thin-film thermal conductivities
    % for thermal CMOS MEMS"
    % Al: "MEMS test structure for measuring thermal conductivity of thin films"

```

```
k_si = 148;
k_al = 200;
k_sio2 = 1.4;
150   k_ti = 21.9;
k_aln = 60;
k_mo = 138;

% Coefficients of thermal expansion, 1/K
alpha_al = 23.1e-6;
alpha_sio2 = 0.5e-6;
alpha_si = 2.6e-6;
alpha_ti = 8.6e-6;
alpha_aln = 4.5e-6;
160   alpha_mo = 4.8e-6;

% Silicon temperature coefficient of resistance (TCR), 1/K
% Corresponds to a peak doping of about 1e20/cc. Used to predict
% resistance change with self-heating in the advanced thermal models
TCR = 1372e-6;

% Intrinsic stress (Pa)
% For modeling tip deflection from film stress
% Doping stress is treated separately elsewhere in the code
170   % film_stress == 'nominal' uses the average
% film_stress == 'random' uses a normal distribution
sigma_si_range = 1e6*[0 0];
sigma_sio2_range = 1e6*[-200 -300];
sigma_al_range = 1e6*[160 200];
sigma_aln_range = 1e6*[-300 300];
sigma_ti_range = 1e6*[-25 25];
sigma_mo_range = 1e6*[-25 25];

% Specific heat (J/kg-K) for calculating thermal time constants
180   Cv_si = 700;
Cv_al = 910;
Cv_sio2 = 700;
```

```

% Mechanical material properties (Pa)
nu_Si = 0.28; % <100> direction
nu_Al = 0.3;
nu_Ti = 0.3;
nu_AlN = 0.24;
nu_SiO2 = 0.17;
nu_Mo = 0.3;

190

% Assume plane strain conditions
% (i.e. cantilever much wider than it is thick)
E_si = 130e9/(1 - cantilever.nu_Si^2);
E_al = 70e9/(1 - cantilever.nu_Al^2);
E_ti = 90e9/(1 - cantilever.nu_Ti^2);
E_aln = 320e9/(1 - cantilever.nu_AlN^2); % or 345 GPa
E_sio2 = 75e9/(1 - cantilever.nu_SiO2^2);
E_mo = 329e9/(1 - cantilever.nu_Mo^2); % or 270 GPa

200

% Densities (kg/m^3)
rho_si = 2330;
rho_al = 2700;
rho_ti = 4506;
rho_aln = 3260;
rho_sio2 = 2200;
rho_mo = 10280;

% Transverse piezoelectric coefficient of AlN, pm/V or pC/N
210
% d31 varies with thickness, so interpolate from literature values
% Values are from papers written by the Piazza and Roukes groups
d31_t = 1e-9*[50 100 500 3000]; % nm
d31_aln = 1e-12*[1.9 2.3 2.5 2.6]; % pm/V

% Minimum and maximum quality factors for the cantilever
% Q is calculated from fluid damping. Results at atmospheric pressure
% are accurate, but other damping mechanisms that become important
% in vacuum (e.g. thermoelastic damping) are not included
maxQ = 5e3;
minQ = 1e-6;
220

```

```

% Define the possible optimization goals.
% Can be referred to as cantilever.goalForceResolution, etc.
goalForceResolution = 0;
goalDisplacementResolution = 1;
goalForceNoiseDensity = 2;
goalSurfaceStress = 3;

% Lookup table for calculating resonant frequency and quality factor in liquid
230 % Source: "Oscillations of cylinders..." by Brumley, Wilcox and Sader (2010)
% Values are stored as constants for calculation speed
A_lookup = [0 1/50 1/20 1/10 1/5 1/2 1 2 5 10 20 50 1000]; % t/w ratio
Beta_lookup = [-3 -2.5 -2 -1.5 -1 -.5 0 0.5 1 1.5 2 2.5 3 100]; % log(Re)

% Hydrodynamic function
% Includes fixed bottom row from published erratum
gamma_lookup_real = ...
[212.184 213.310 214.977 217.701 222.978 237.780 260.256 ...
240 207.210 169.667 154.616 145.909 139.855 134.720;
91.6984 92.2467 93.0601 94.3924 96.9808 104.295 115.542 ...
88.9011 70.8173 63.7655 59.7404 56.9653 54.6258;
41.6417 41.9209 42.3363 43.0185 44.3487 48.1380 54.0391 ...
39.8564 30.6996 27.2460 25.3060 23.9817 22.8730;
20.1196 20.2683 20.4907 20.8572 21.5753 23.6370 26.8847 ...
18.8235 13.9212 12.1457 11.1673 10.5072 9.95883;
10.4849 10.5677 10.6926 10.8998 11.3080 12.4883 14.3601 ...
9.43536 6.64606 5.68511 5.16801 4.82411 4.54093;
5.96655 6.01467 6.08871 6.21279 6.45897 7.17328 8.30052 ...
5.04739 3.35215 2.80394 2.51794 2.33126 2.17927;
250 3.73387 3.76344 3.81063 3.89099 4.05154 4.51368 5.22220 ...
2.89030 1.78322 1.45306 1.28807 1.18327 1.09943;
2.56548 2.58563 2.61959 2.67832 2.79515 3.11907 3.58531 ...
1.77617 0.994540 0.783333 0.684003 0.623512 0.576619;
1.91834 1.93509 1.96437 2.01450 2.11058 2.35665 2.68270 ...
1.17779 0.580514 0.435349 0.372208 0.336075 0.309503;
1.54554 1.56285 1.59247 1.64069 1.72687 1.92785 2.17551 ...
0.848104 0.357549 0.249659 0.206674 0.184001 0.168601;

```

```

1.32633 1.34658 1.37882 1.42757 1.50844 1.68437 1.88862 ...
    0.663505 0.235193 0.148772 0.117201 0.102069 0.0928779;
260   1.19577 1.2202 1.2555 1.3051 1.3833 1.5459 1.7259 ...
    0.55939 0.16703 0.093131 0.068128 0.057273 0.0515648;
1.11746 1.1465 1.1843 1.2346 1.3117 1.4670 1.6336 ...
    0.50051 0.12874 0.062098 0.040918 0.032516 0.0287745;
1 1.04551 1.08816 1.14064 1.21703 1.36368 1.51317 ...
    0.423881 0.0792129 0.0222121 0.00619303 0.00113212 0 ];
gamma_lookup_imag = ...
[1018.72 1021.37 1025.29 1031.66 1043.88 1077.39 1126.32 ...
    1008.65 915.159 874.583 850.149 832.704 817.599;
270   374.276 375.392 377.040 379.721 384.873 399.079 420.012 ...
    370.057 331.318 314.778 304.899 297.884 291.835;
140.659 141.144 141.862 143.031 145.284 151.534 160.848 ...
    138.825 122.228 115.278 111.167 108.266 105.776;
54.4049 54.6253 54.9508 55.4818 56.5079 59.3754 63.7087 ...
    53.5749 46.1812 43.1534 41.3825 40.1420 39.0836;
21.8269 21.9314 22.0855 22.3371 22.8247 24.2002 26.3169 ...
    21.4324 17.9905 16.6153 15.8210 15.2692 14.8012;
9.16870 9.22024 9.29587 9.41936 9.65973 10.3480 11.4345 ...
    8.96804 7.28929 6.63516 6.26219 6.00523 5.78862;
280   4.07467 4.10043 4.13779 4.19895 4.31957 4.67605 5.25977 ...
    3.95920 3.10274 2.77671 2.59298 2.46733 2.36186;
1.93366 1.94552 1.96256 1.99130 2.05107 2.24127 2.56535 ...
    1.85252 1.39790 1.22868 1.13429 1.07013 1.01639;
0.981710 0.985312 0.990956 1.00255 1.03157 1.13634 1.31768 ...
    0.915797 0.666095 0.575374 0.525354 0.491568 0.463359;
0.527773 0.526433 0.526077 0.529479 0.543868 0.602276 0.703142 ...
    0.474037 0.333253 0.283225 0.256021 0.237799 0.222666;
0.296143 0.291987 0.289093 0.289338 0.296683 0.328687 0.384789 ...
    0.253907 0.173548 0.145302 0.130165 0.120135 0.111868;
290   0.171115 0.16564 0.16234 0.16171 0.16525 0.18260 0.21384 ...
    0.13910 0.093151 0.076988 0.068405 0.062790 0.0582134;
0.100688 0.095021 0.092307 0.091476 0.093044 0.10247 0.11987 ...
    0.077266 0.051022 0.041760 0.036840 0.033652 0.0310905;
0 0 0 0 0 0 0 ...

```

```

    0 0 0 0 0 0];
end

% Abstract methods that MUST be defined in subclasses
methods (Abstract)
300    doping_profile(self)
    doping_optimization_scaling(self)
    doping_cantilever_from_state(self)
    doping_current_state(self)
    doping_initial_conditions_random(self)
    doping_optimization_bounds(self, parameter_constraints)

% Abstract functions for handling ion implantation
Nz(self)
alpha(self)
310    sheet_resistance(self)
end

methods
function self = cantilever(freq_min, freq_max, l, w, t, ...
                           l_pr_ratio, v_bridge, doping_type)
    % Define the cantilever using the required parameters
    self.freq_min = freq_min;
    self.freq_max = freq_max;
    self.l = l;
320    self.w = w;
    self.t = t;
    self.l_pr_ratio = l_pr_ratio;
    self.v_bridge = v_bridge;
    self.doping_type = doping_type;

    % Set default values for optional parameters
    % These can be overridden once the cantilever is created
    self.fluid = 'air';
    self.h_method = 'fixed';
330    self.metal_type = 'aluminum';
    self.film_stress = 'nominal';

```

```

    self.cantilever_type = 'none';
    self.l_a = 0;
    self.t_a = 0;
    self.w_a = 0;
    self.w_a_active = 0;
    self.d31_manual = 0;
    self.l_a_gap = 0;
    self.tip_mass = 0;
340   self.t_oxide = 100e-9;
    self.t_electrode_bottom = 50e-9;
    self.t_electrode_top = 50e-9;
    self.t_a_seed = 20e-9;
    self.number_of_piezoresistors = 2;
    self.amplifier = 'INA103';
    self.T = 273.15 + 23; % Assume operation around 23C
    self.T_ref = 273.15 + 23;
    self.R_base = 10e3; % Assume R_base is about 10e3 K/W by default
    self.R_contact = 170; % Assume 340 ohm total excess resistance
350   self.temperature_dependent_properties = 'no';
    self.thermal_modeling = 'none';
end

% Helpful getter function to calculate the absolute piezoresistor length
% Units: m
function l_pr = l_pr(self)
    l_pr = self.l * self.l_pr_ratio;
end

360   % Getter function for piezoresistor width
% Units: m
function w_pr = w_pr(self)
    w_pr = self.w/2;
end

% Determine the ion implantation table index from the dopant type
function dopantNumber = dopantNumber(self)
    switch self.doping_type

```

```

    case 'boron'
370     dopantNumber = 1;
    case 'phosphorus'
        dopantNumber = 2;
    case 'arsenic'
        dopantNumber = 3;
    otherwise
        fprintf('ERROR: Unknown dopant type: %s\n', self.doping_type);
        pause
    end
end

380
% Check if the cantilever is self-consistent
function check_valid_cantilever(self)
    validity_checks(1) = 1;
    validity_checks(2) = ~(strcmp(self.cantilever_type, 'none') ...
        && (self.l_a > 0));
    [valid, failed_index] = min(validity_checks);
    if ~valid
        fprintf('ERROR: Invalid cantilever - failed check #%d\n', failed_index);
        pause
    end
end

390

% Print the cantilever performance to the Matlab prompt
function print_performance(self)
    self.check_valid_cantilever();

    [omega_damped_hz, Q] = self.omega_damped_hz_and_Q();
    [x, active_doping, total_doping] = self.doping_profile();
    [TMax_approx TTip_approx] = self.approxTempRise();
    [TMax, TTip] = self.calculateMaxAndTipTemp();
    thermoLimit = self.thermo_integrated()/self.force_sensitivity();

    fprintf('=====\\n')
    fprintf('Freq range: %f to %f \\n', self.freq_min, self.freq_max)
    fprintf('Operating fluid: %s \\n', self.fluid);

```

```

        fprintf('Cantilever L/W/T: %f %f %f \n', ...
        self.l*1e6, self.w*1e6, self.t*1e6)
        fprintf('PR L/W: %f %f %f \n', self.l_pr()*1e6, self.w_pr()*1e6)
        fprintf('PR Length Ratio: %g \n', self.l_pr_ratio)
410      fprintf('\n')
        fprintf('Force resolution (N): %g \n', self.force_resolution())
        fprintf('Force noise at 1 kHz (fN): %g \n', ...
        self.force_noise_density(1e3))
        fprintf('Displacement resolution (m): %g \n', ...
        self.displacement_resolution())
        fprintf('Force sensitivity (V/N) %g \n', self.force_sensitivity())
        fprintf('Displacement sensitivity (V/m) %g \n', ...
        self.displacement_sensitivity())
        fprintf('Beta %g \n', self.beta())
420      fprintf('Thermomechanical force noise limit: %g \n', thermoLimit);
        fprintf('\n')
        fprintf('Stiffness (N/m): %g \n', self.stiffness())
        fprintf('Vacuum freq: %f \n', self.omega_vacuum_hz())
        fprintf('Damped freq: %f \n', omega_damped_hz)
        fprintf('Quality factor: %f \n', Q)
        fprintf('\n')
        fprintf('Wheatstone bridge bias voltage: %f \n', self.v_bridge)
        fprintf('Resistance: %f \n', self.resistance())
        fprintf('Sheet Resistance: %f \n', self.sheet_resistance())
430      fprintf('Power dissipation (mW): %g \n', self.power_dissipation()*1e3)
        fprintf('Approx. Temp Rises (C) - Tip: %f Max: %f\n', ...
        TTip_approx, TMax_approx)
        fprintf('F-D Temp Rises (C) - Tip: %f Max: %f\n', ...
        TTip, TMax)
        fprintf('\n')
        fprintf('Integrated noise (V): %g \n', self.integrated_noise())
        fprintf('Integrated johnson noise (V): %g \n', self.johnson_integrated())
        fprintf('Integrated 1/f noise (V): %g \n', self.hooge_integrated())
        fprintf('Amplifier noise (V): %g \n', self.amplifier_integrated())
440      fprintf('Thermomechanical noise (V): %g \n', self.thermo_integrated());
        fprintf('\n')
        fprintf('Johnson/Hooge: %g \n', ...

```

```

        self.johnson_integrated()/self.hooge_integrated())
450   fprintf('Knee frequency (Hz): %g \n', self.knee_frequency())
     fprintf('Number of Carriers: %g \n', self.number_of_carriers());
     fprintf('Nz: %g \n', self.Nz())
     fprintf('\n')
     fprintf('Number of silicon resistors: %f \n', ...
             self.number_of_piezoresistors)
     fprintf('Si Thermal Conductivity (W/m-K): %f \n', self.k_base())
     fprintf('E (GPa): %f \n', self.modulus()*1e-9)
     fprintf('Alpha: %g \n', self.alpha)

switch self.cantilever_type
  case 'none'
    % Do nothing special
  case 'step'
    fprintf('=====\\n')
    fprintf('Step at base (um): %f thick x %f long \n', ...
460       1e6*self.t_a, 1e6*self.l_a)
  case 'thermal'
    fprintf('=====\\n')
    [tau, freq] = self.heaterTimeConstant();
    fprintf('Actuator l/W/T: %f %f %f \n', ...
            1e6*self.l_a, 1e6*self.w_a, 1e6*self.t_a)
    fprintf('Neutral axis (um): %f \n', 1e6*self.actuatorNeutralAxis())
    fprintf('Actuator Voltage (): %f \n', self.v_actuator)
    fprintf('Heater resistance (kOhm): %f \n', 1e-3*self.R_heater)
    fprintf('Actuator Power (mW): %f \n', 1e3*self.heaterPower())
    fprintf('Tip Deflection (nm): %f \n', 1e9*self.tipDeflection())
    fprintf('Time Constant (microseconds): %f \n', tau*1e6)
    fprintf('‐3dB frequency (kHz): %f \n', freq*1e-3)
  case 'piezoelectric'
    fprintf('Actuator l/W/T: %f %f %f \n', ...
            1e6*self.l_a, 1e6*self.w_a, 1e6*self.t_a)
    fprintf('Neutral axis (um): %f \n', 1e6*self.actuatorNeutralAxis())
    fprintf('Actuator Voltage (): %f \n', self.v_actuator)
    fprintf('Tip Deflection (nm): %f \n', 1e9*self.tipDeflectipon())
  end

```

```

480     fprintf('=====\\n')
     end

     % Tab delimited output (for Excel, etc). Optionally provide a
     % file id (fid) to print the output there. print_performance_for_excel
     % is overridden in subclasses to provide detailed doping output
function print_performance_for_excel(self, varargin)
    varargin = varargin{1};
    optargin = size(varargin, 2);

490     % Determine where to print the output
    if optargin == 1
        fid = varargin{1};
    elseif optargin == 0
        fid = 1; % Print to the stdout
    else
        fprintf('ERROR: Extra optional arguments')
    end

    % Calculate intermediate quantities
500    [omega_damped_hz, Q] = self.omega_damped_hz_and_Q();
    [TMax, TTip] = self.calculateMaxAndTipTemp();
    [TMax_approx TTip_approx] = self.approxTempRise();
    thermoLimit = self.thermo_integrated()/self.force_sensitivity();

    variables_to_print = [self.freq_min, self.freq_max*1e-3, ...
        1e6*self.l 1e6*self.w 1e9*self.t 1e6*self.l_pr() ...
        1e6*self.l_a 1e6*self.w_a 1e9*self.t_a, ...
        self.force_resolution()*1e12, thermoLimit*1e12, ...
        self.displacement_resolution()*1e9, self.omega_vacuum_hz()*1e-3, ...
        omega_damped_hz*1e-3, Q, self.stiffness()*1e3, self.v_bridge, ...
        self.resistance()*1e-3, self.sheet_resistance(), ...
        self.power_dissipation()*1e6, TMax, TTip, TMax_approx, TTip_approx, ...
        self.number_of_piezoresistors, self.effective_mass(), self.tip_mass, ...
        self.force_sensitivity(), self.beta(), self.gamma(), ...
        self.integrated_noise()*1e6, self.johnson_integrated()*1e6, ...
        self.hooge_integrated()*1e6, self.amplifier_integrated()*1e6, ...

```

```

    self.thermo_integrated()*1e6, self.knee_frequency(), ...
    self.number_of_carriers(), 1e6*self.calculateDopantTipDeflection(), ...
    self.k_x(), self.modulus()*1e-9];

520
    fprintf(fid, '%s \t', self.doping_type);
    fprintf(fid, '%s\t', self.fluid);
    fprintf(fid, '%s\t', self.cantilever_type);
    for print_index = 1:length(variables_to_print)
        fprintf(fid, '%.4g \t', variables_to_print(print_index));
    end
end

% Calculate total resistance of piezoresistor.

530 % Units: ohms
function resistance = resistance(self)
    resistance = self.number_of_squares()*self.sheet_resistance() + ...
        2*self.R_contact;
end

% Calculate the number of resistor squares
% 1) Longitudinal region (2*l_pr/w_pr)
% 2) Transverse at end (air_gap_width/2*w_pr)
% 3) the connecting corners
540 % Units: squares
function number_of_squares = number_of_squares(self)
    number_of_squares = 2*self.l_pr()/self.w_pr() + ...
        self.air_gap_width/(2*self.w_pr()) + 2;
end

% Calculate R-R0/R0 for the cantilever including the TCR
% Used for calculating h_eff from experimental results
% Units: -
function dR = dR_with_temp_rise(self)
    dR = self.TCR*(self.averagePRTemp() - self.T);
end

550 function dR = approx_dR_with_temp_rise(self)

```

```

dR = self.TCR*(self.approxPRTemp() - self.T);
end

% Calculate conductivity for a given dopant concentration.
% Use vectors rather than looping for speed.
% Units: C/V-sec-cm
560 function sigma = conductivity(self, dopant_concentration)
switch self.temperature_dependent_properties
    case 'yes'
        [mu, sigma] = self.mobility(dopant_concentration, self.approxPRTemp());
    case 'no'
        [mu, sigma] = self.mobility(dopant_concentration, self.T);
    end
end

% Calculate the temp dependent carrier density, mobility, conductivity
570 % Current: "Electron and Hole Mobility ...", Reggiani et al. (2002)
% Previously: "Modeling of Carrier Mobility ...", Masetti et al. (1983)
function [mu, sigma] = mobility(self, dopantConc, T)
Tnorm = T/300;
Eg = 1.170-(4.730e-4*T.^2)./(T+636);
ni = (2.4e31.*T.^3.*exp(-Eg./(self.k_b_eV.*T))).^.5;

switch self.doping_type
    case 'boron'
        mumax=470.5;
        c=0.0;
        gamma=2.16;
580        mu0d=90.0*Tnorm^-1.3;
        mu0a=44.0*Tnorm^-0.7;
        mu1d=28.2*Tnorm^-2.0;
        mu1a=28.2*Tnorm^-0.8;
        Cr1=1.3e18*Tnorm^2.2;
        Cr2=2.45e17*Tnorm^3.1;
        Cs1=1.1e18*Tnorm^6.2;
        Cs2=6.1e20;
        alpha1=.77;

```

```

alpha2=.719;
ni=(2.4e31*T^3*exp(-1*Eg/8.617e-5/T))^5;
p=(dopantConc./2)+((dopantConc./2).^2+ni.^2).^5;
n=ni.*ni./p;
ND=n;
NA=p;
case 'phosphorus'
mumax = 1441;
c = 0.07;
600 gamma = 2.45;
mu0d = 62.2.*Tnorm.^-0.7;
mu0a = 132.*Tnorm.^-1.3;
mu1d = 48.6.*Tnorm.^-0.7;
mu1a = 73.5.*Tnorm.^-1.25;
Cr1 = 8.5e16.*Tnorm.^3.65;
Cr2 = 1.22e17.*Tnorm.^2.65;
Cs1 = 4e20;
Cs2 = 7e20;
alpha1 = .68;
610 alpha2 = .72;
n = (dopantConc./2)+((dopantConc./2).^2 + ni.^2).^5;
p = ni.^2./n;
ND = n;
NA = p;
case 'arsenic'
mumax = 1441;
c = 0.07;
gamma = 2.45;
mu0d = 55.*Tnorm.^-0.6;
620 mu0a = 132.*Tnorm.^-1.3;
mu1d = 42.4.*Tnorm.^-0.5;
mu1a = 73.5.*Tnorm.^-1.25;
Cr1 = 8.9e16.*Tnorm.^3.65;
Cr2 = 1.22e17.*Tnorm.^2.65;
Cs1 = 2.9e20;
Cs2 = 7e20;
alpha1 = .68;

```

```

alpha2 = .72;
n = (dopantConc./2)+((dopantConc./2).^2 + ni.^2).^5;
630 p = ni.^2./n;
ND = n;
NA = p;
end

mu0 = (mu0d.*ND+mu0a.*NA)./(ND+NA);
mu1 = (mu1d.*ND+mu1a.*NA)./(ND+NA);
muL = mumax.*Tnorm.^(-gamma+c.*Tnorm);
second = (muL-mu0)./(1+(ND./Cr1).^alpha1 + (NA./Cr2).^alpha2);
third = mu1./(1+(ND/Cs1+NA/Cs2).^2);

640 mu = mu0+second-third;
sigma = self.q.* (mu.*n + mu.*p);
end

% Calculate Rsheet(x) assuming temperature dependent cantilever properties
% Units: ohm/square
function [Rsheet_x] = RSheetProfile(self, x, T_x)
[z, active_doping, total_doping] = self.doping_profile();
% Units: z -> m, doping -> N/cm^3
650 n_z_x = total_doping'*ones(1, self.numXPoints);

% Generate a numZPoints x numXPoints matrix
if length(T_x) == 1
    T_z_x = ones(self.numZPoints, 1)*T_x'*ones(1, self.numXPoints);
else
    T_z_x = ones(self.numZPoints, 1)*T_x';
end

[mu_z_x, sigma_z_x] = self.mobility_temp_dependent(n_z_x, T_z_x + self.T);
660 Rsheet_x = 1./trapz(z*1e2, sigma_z_x); % Convert z from m to cm
end

% The number of current carriers in the piezoresistor
% The cantilever subclass implements Nz()

```

```

% Units: -
function number_of_carriers = number_of_carriers(self)
    resistor_area = self.w_pr()*(2*self.l_pr() + self.w + self.air_gap_width);
    number_of_carriers = self.Nz()*resistor_area;
end

670
% 1/f voltage power spectral density for the entire Wheatstone bridge
% Units: V^2/Hz
function hooge_PSD = hooge_PSD(self, freq)
    hooge_PSD = self.alpha()*self.v_bridge^2* ...
        self.number_of_piezoresistors./(4*self.number_of_carriers()*freq);
end

% Integrated 1/f noise density for the entire Wheatstone bridge
% Unit: V
680 function hooge_integrated = hooge_integrated(self)
    hooge_integrated = sqrt(self.alpha()*self.v_bridge^2* ...
        self.number_of_piezoresistors./(4*self.number_of_carriers())* ...
        log(self.freq_max/self.freq_min));
end

% Johnson noise PSD from the entire Wheatstone bridge
% Units: V^2/Hz
function johnson_PSD = johnson_PSD(self, freq)
    resistance = self.resistance(); % resistance() includes contacts
690    switch self.thermal_modeling
        case 'none'
            TPR = self.T; % the ambient temperature
        case 'approx'
            TPR = self.approxPRTemp();
        case 'exact'
            TPR = self.averagePRTemp();
        end
    resistance = resistance*(1 + TPR*self.TCR);

    700 if self.number_of_piezoresistors == 4
        R_external = resistance;

```

```

else
    % If we're using 2 piezoresistors, assume ideal 1 kOhm resistors
    R_external = 1e3;
end

johnson_PSD = 4*self.k_b*TPR*(resistance/2 + ...
    R_external/2) * ones(1, length(freq));
end

710
% Integrated Johnson noise
% Unit: V
function johnson_integrated = johnson_integrated(self)
    resistance = self.resistance();
    TPR = self.approxPRTemp();
    resistance = resistance*(1 + TPR*self.TCR);

    if self.number_of_piezoresistors == 4
        R_external = resistance;
    else
        720
        R_external = 700;
    end
    johnson_integrated = sqrt(4*self.k_b*TPR*(resistance/2 + ...
        R_external/2)*(self.freq_max - self.freq_min));
end

% Thermomechanical noise PSD
% Units: V^2/Hz
function thermo_PSD = thermo_PSD(self, freq)
    730
    [omega_damped_hz, Q_M] = self.omega_damped_hz_and_Q();
    switch self.thermal_modeling
        case 'none'
            TPR = self.T; % the ambient temperature
        case 'approx'
            TPR = self.approxPRTemp();
        case 'exact'
            TPR = self.averagePRTemp();
    end

```

```

thermo_PSD = (self.force_sensitivity())^2*4*self.stiffness()* ...
740      self.k_b*TPR/(2*pi*omega_damped_hz*Q_M) * ones(1, length(freq));
end

% Integrated thermomechanical noise
% Unit: V
function thermo_integrated = thermo_integrated(self)
[omega_damped_hz, Q_M] = self.omega_damped_hz_and_Q();
switch self.thermal_modeling
case 'none'
    TPR = self.T; % the ambient temperature
750 case 'approx'
    TPR = self.approxPRTemp();
case 'exact'
    TPR = self.averagePRTemp();
end
thermo_integrated = sqrt((self.force_sensitivity())^2* ...
4*self.stiffness()*self.k_b*TPR/(2*pi*omega_damped_hz*Q_M) ...
*(self.freq_max - self.freq_min));
end

760 % Accounts for displacement noise (vibrations, etc)
% Units: V
function actuator_noise_integrated = actuator_noise_integrated(self)
actuator_noise_integrated = self.rms_actuator_displacement_noise* ...
self.stiffness()*self.force_sensitivity(); % V
end

% Amplifier noise PSD
% Units: V^2/Hz
function amplifier_PSD = amplifier_PSD(self, freq)
770 switch self.amplifier
case 'INA103'
    A_VJ = 1.2e-9; % 1.2 nV/rtHz noise floor
    A_IJ = 2e-12; % 2 pA/rtHz noise floor
    A_VF = 6e-9; % 6 nV/rtHz @ 1 Hz
    A_IF = 25e-12; % 25 pA/rtHz @ 1 Hz

```

```

case 'AD8221'
A_VJ = 8e-9;
A_IJ = 40e-15;
A_VF = 12e-9;
A_IF = 550e-15;
otherwise
    fprintf('ERROR: UNKNOWN AMPLIFIER')
    pause
end

R = self.resistance()/2; % Resistance seen at the amp inputs
switch self.thermal_modeling
case 'none'
    TPR = self.T; % the ambient temperature
case 'approx'
    TPR = self.approxPRTemp();
case 'exact'
    TPR = self.averagePRTemp();
end
R = R*(1 + TPR*self.TCR);

amplifier_PSD = (A_VJ^2 + 2*(R*A_IJ)^2) + (A_VF^2 + 2*(R*A_IF)^2)./freq;
end

800 % Integrated amplifier noise
% Units: V
function amplifier_integrated = amplifier_integrated(self)

switch self.amplifier
case 'INA103'
    A_VJ = 1.2e-9; % 1.2 nV/rtHz noise floor
    A_IJ = 2e-12; % 2 pA/rtHz noise floor
    A_VF = 6e-9; % 6 nV/rtHz @ 1 Hz
    A_IF = 25e-12; % 25 pA/rtHz @ 1 Hz
case 'AD8221'
    A_VJ = 8e-9;
    A_IJ = 40e-15;

```

```

A_VF = 12e-9;
A_IF = 550e-15;
otherwise
    fprintf('ERROR: UNKNOWN AMPLIFIER')
    pause
end

820    R = self.resistance()/2; % Resistance seen at the amp inputs
switch self.thermal_modeling
case 'none'
    TPR = self.T; % the ambient temperature
case 'approx'
    TPR = self.approxPRTemp();
case 'exact'
    TPR = self.averagePRTemp();
end
R = R*(1 + TPR*self.TCR);

830    amplifier_integrated = sqrt(A_VJ^2*(self.freq_max - self.freq_min) ...
+ A_VF^2*log(self.freq_max/self.freq_min) + ...
2*(R*A_IJ)^2*(self.freq_max - self.freq_min) + ...
2*(R*A_IF)^2*log(self.freq_max/self.freq_min));
end

% Calculate the knee frequency (Hooge PSD == Johnson PSD)
% Units: Hz
function knee_frequency = knee_frequency(self)
    knee_frequency = self.number_of_piezoresistors*self.alpha()* ...
        self.v_bridge^2/(16*self.number_of_carriers()*self.k_b* ...
        self.approxPRTemp()*self.resistance());
end

840    % Integrated cantilever noise for given bandwidth
    % Pull the calculations into this function for speed (i.e. don't
    % calculate self.resistance() five separate times
    % Units: V
    function integrated_noise = integrated_noise(self)

```

```

850 [omega_damped_hz, Q_M] = self.omega_damped_hz_and_Q();
resistance = self.resistance()/self.gamma();
force_sensitivity = self.force_sensitivity();
spring_constant = self.stiffness();
switch self.thermal_modeling
    case 'none'
        TPR = self.T; % the ambient temperature
    case 'approx'
        TPR = self.approxPRTemp();
    case 'exact'
        TPR = self.averagePRTemp();
860 end
resistance = resistance*(1 + TPR*self.TCR);
switch self.amplifier
    case 'INA103'
        A_VJ = 1.2e-9; % 1.2 nV/rtHz noise floor
        A_IJ = 2e-12; % 2 pA/rtHz noise floor
        A_VF = 6e-9; % 6 nV/rtHz @ 1 Hz
        A_IF = 25e-12; % 25 pA/rtHz @ 1 Hz
    case 'AD8221'
        A_VJ = 8e-9;
        A_IJ = 40e-15;
        A_VF = 12e-9;
        A_IF = 550e-15;
    otherwise
        fprintf('ERROR: UNKNOWN AMPLIFIER')
        pause
870 end

actuator_noise_integrated = max(0, self.rms_actuator_displacement_noise* ...
880     spring_constant*force_sensitivity); % V
johnson_integrated = sqrt(4*self.k_b*TPR*resistance* ...
    (self.freq_max - self.freq_min));
hooge_integrated = sqrt(self.alpha()*self.v_bridge^2* ...
    self.number_of_piezoresistors./(4*self.number_of_carriers()) ...
    *log(self.freq_max/self.freq_min));
thermo_integrated = sqrt(force_sensitivity^2*4*spring_constant* ...

```

```

        self.k_b*TPR/(2*pi*omega_damped_hz*Q_M)*(self.freq_max - self.freq_min));
amplifier_integrated = sqrt(A_VJ^2*(self.freq_max - self.freq_min) + ...
A_VF^2*log(self.freq_max/self.freq_min) + 2*(resistance/2*A_IJ)^2* ...
(self.freq_max - self.freq_min) + 2*(resistance/2*A_IF)^2* ...
log(self.freq_max/self.freq_min));
integrated_noise = sqrt(actuator_noise_integrated^2 + ...
johnson_integrated^2 + hooge_integrated^2 + thermo_integrated^2 ...
+ amplifier_integrated^2);
end

% Calculate the noise in V/rtHz at a given frequency
function voltage_noise = voltage_noise(self, freq)
voltage_noise = sqrt(self.johnson_PSD(freq) + self.hooge_PSD(freq) + ...
self.thermo_PSD(freq) + self.amplifier_PSD(freq));
end

function f_min_cumulative = f_min_cumulative(self)
frequency = logspace(log10(self.freq_min), log10(self.freq_max), ...
cantilever.numFrequencyPoints);
noise = self.voltage_noise(frequency);
sensitivity = self.force_sensitivity();
force_noise_density = noise./sensitivity;
f_min_cumulative = sqrt(cumtrapz(frequency, force_noise_density.^2));
910 end

% Piezoresistance factor
% Accounts for dopant concentration dependent piezoresistivity in silicon
% Source: "Piezoresistance in p-type silicon revisited", Richter et al.
function P = piezoresistance_factor(self, dopant_concentration)
Nb = 6e19;
Nc = 7e20;
richter_alpha = 0.43;
richter_gamma = 1.6;
richter_beta = 0.1;
richter_eta = 3;
richter_theta = 0.9;
920

```

```

T0 = 300;
switch self.thermal_modeling
    case 'none'
        average_PR_temp = self.T; % the ambient temperature
    case 'approx'
        average_PR_temp = self.approxPRTemp();
    case 'exact'
        average_PR_temp = self.averagePRTemp();
end
Theta = average_PR_temp/T0;

P = Theta^-richter_theta*(1 + Theta^-richter_beta* ...
    (dopant_concentration/Nb).^richter_alpha + Theta^-richter_eta* ...
    (dopant_concentration/Nc).^richter_gamma).^ -1;
end
940
% Low concentration longitudinal piezoresistance coefficient
% Units: 1/Pa
function max_factor = max_piezoresistance_factor(self)
switch self.doping_type
    case 'boron'
        max_factor = 72e-11; % 110 direction
    case 'phosphorus'
        max_factor = 103e-11; % 100 direction
    case 'arsenic'
        max_factor = 103e-11; % 100 direction
950
    end
end

% Calculate the sensitivity factor (beta*)
% Accounts for finite piezoresistor thickness
% Units: None
function beta = beta(self)
[z, active_doping, total_doping] = self.doping_profile();
960
% Shift axis so that z varies from [-t/2, t/2] and m -> cm

```

```

z = (self.t/2 - z)*1e2;

switch self.temperature_dependent_properties
    case 'yes'
        [mu, sigma] = self.mobility(active_doping, self.approxPRTemp());
    case 'no'
        [mu, sigma] = self.mobility(active_doping, self.T);
    end
P = self.piezoresistance_factor(active_doping);
970 numerator = trapz(z, sigma.*P.*z);
denominator = trapz(z, sigma);
beta = 2*numerator/(self.t*1e2*denominator); % t: m -> cm

% For optimization, ensure that beta doesn't become negative
beta = max(beta, 1e-6);
end

% Ratio of piezoresistor resistance to total resistance (< 1)
function gamma = gamma(self)
    R = self.resistance();
    980 gamma = R/(R + 2*self.R_contact);
end

% Calculate the force sensitivity (V/N) for the 1/4-active Wheatstone bridge
% Includes the transverse portion at the end of the piezoresistive loop
% Units: V/N
function force_sensitivity = force_sensitivity(self)

    % For speed: precompute these parameters
    990 betaStar = self.beta();
    Rs = self.sheet_resistance();
    wheatstone_bridge_sensitivity = self.v_bridge/4;
    piMax = self.max_piezoresistance_factor();
    gamma = self.gamma();
    R = self.resistance();
    l_pr = self.l_pr();
    w_pr = self.w_pr();

```

```

% Pick the relative transverse PR coefficient for the doping type
1000 switch self.doping_type
    case 'boron'
        transverse_factor = -1;
    case {'phosphorus', 'arsenic'}
        transverse_factor = -.5;
    end

% Average length from the base to the piezoresistor centroid
longitudinal_l_avg = self.l - l_pr/2;
transverse_l_avg = self.l - l_pr;

1010 % Stress prefactor
stress_prefactor = 6*piMax/(self.w*self.t^2)*betaStar*gamma;

% Calculate the longitudinal and transverse resistances
% Assume that the transverse width is 2x the PR width
R_longitudinal = 2*Rs*l_pr/w_pr;
R_transverse = Rs*self.air_gap_width/(2*w_pr);

% Calculate deltaR values
1020 longitudinal_deltaR = stress_prefactor*longitudinal_l_avg*R_longitudinal;
transverse_deltaR = stress_prefactor*transverse_factor* ...
    transverse_l_avg*R_transverse;
deltaR_R = (longitudinal_deltaR + transverse_deltaR)/R;

force_sensitivity = deltaR_R*wheatstone_bridge_sensitivity;

% Compare with the sensitivity excluding the transverse effect
% (i.e. the simple formula that we usually include in papers)
force_sensitivity_no_transverse_effect = 3*(self.l - l_pr/2)* ...
    piMax/(2*self.w*self.t^2)*betaStar*gamma*self.v_bridge;

1030 end

% Calculate the input referred surface stress sensitivity
% Units: V/Pa

```

```

function stress_sensitivity = surface_stress_sensitivity(self)

    % Pick the relative transverse PR coefficient for the doping type
    switch self.doping_type
        case 'boron'
            longitudinal_factor = 1;
            transverse_factor = -1;
        case {'phosphorus', 'arsenic'}
            longitudinal_factor = 1;
            transverse_factor = -.5;
    end

    % The longitudinal and transverse stress is equal everywhere
    sensitivity_factor = abs(longitudinal_factor + transverse_factor);
    stress_sensitivity = 9*sensitivity_factor* ...
        1040 self.max_piezoresistance_factor()*self.beta()*self.gamma()* ...
        self.v_bridge/(16*self.t);
    end

    % Calculate the input referred displacement sensitivity
    % Units: V/m
    function displacement_sensitivity = displacement_sensitivity(self)
        displacement_sensitivity = self.force_sensitivity()*self.stiffness();
    end

    1060 % Power dissipation in the cantilever
    % Units: W
    function power_dissipation = power_dissipation(self)
        power_dissipation = (self.v_bridge/2)^2/self.resistance();
    end

    % Tip and maximum temperatures from the F-D model
    % Units: K
    function [TMax, TTip] = calculateMaxAndTipTemp(self)
        [tmp, Q, temp] = self.calculateTempProfile();
        1070 TMax = max(temp);
        TTip = temp(end);
    end

```

```

end

% Calculate the approximate PR temperature
% Units: K
function TPR = approxPRTemp(self)
    [TMax TTip] = self.approxTempRise();
    TPR = self.T + TMax/2;
end

1080
% Calculate the exact average PR temperature
% Units: K
function TPR = averagePRTemp(self)
    switch self.temperature_dependent_properties
        case 'yes'
            [x, Q, temp] = self.calculateTempProfileTempDependent();
        case 'no'
            [x, Q, temp] = self.calculateTempProfile();
    end
pr_indices = intersect(find(x >= self.l_a), ...
    find(x <= (self.l_a + self.l_pr())));
TPR = self.T + mean(temp(pr_indices));
end

% Calculate the temperature at the base of the PR
% Useful for the designing combined sensors/actuators
% Units: K
function TBase = tempRiseAtPRBase(self)
    [x, Q, temp] = self.calculateTempProfile();
    base_index = find(x >= self.l_a, 1);
    TBase = temp(base_index);
end

1100
% Calculate the maximum piezoresistor temperature
% Units: K
function TPR = maxPRTemp(self)
    [x, Q, temp] = self.calculateTempProfile();
    pr_indices = intersect(find(x >= self.l_a), ...

```

```

    find(x <= (self.l_a + self.l_pr())));
1110 TPR = max(temp(pr_indices));
      end

% Calculate the average temperature increase of the actuator
% Use for calculating A_XK (nm/K) for the thermal actuators
% Units: K
function TActuator = averageActuatorDeltaTemp(self)
  [x, Q, temp] = self.calculateTempProfile();
  actuator_indices = x <= (self.l_a);
  TActuator = mean(temp(actuator_indices));
1120 end

% Calculate the temp change of the PR in response to thermal actuation
% Units: K
function PRTempDelta = thermalCrosstalk(self)
  temp_hot = self.averagePRTemp();
  v_actuator_temp = self.v_actuator;
  self.v_actuator = 0;
  temp_cold = self.averagePRTemp();
  self.v_actuator = v_actuator_temp; %#ok<*MCHV2>
1130 PRTempDelta = temp_hot - temp_cold;
end

% Calculate the approx max and tip temperatures using lumped modeling
% Useful for quickly approximating the important temperatures
% Units: K
function [TMax TTip] = approxTempRiseAnalytical(self)
  k_c = self.k_base();
  R_conduction_pr = self.l_pr()/(2*self.w*self.t*k_c);
  TMax = self.power_dissipation()*R_total;
  [l_healing, tmp] = thermalHealingLengths(self);
  TTip = TMax_analytical*exp(-(self.l - 2/3*self.l_pr())/l_healing);
1140 end

% Approximate temperature modeling via a circuit model
% Much faster than the F-D model and slightly more accurate than the

```

```

% lumped parameter model
% Units: K

function [TMax TTip] = approxTempRise(self)
    h = self.lookupHeff();
    k_c = self.k_base();
    l_pr = self.l_pr();
    W = self.power_dissipation();

1150

% Model the system as current sources (PR or heater) and resistors
[E_metal, rho_metal, k_metal, alpha_metal] = self.lookup_metal_properties();

switch self.cantilever_type
    case 'none'
        switch self.fluid
            case 'vacuum'
                R_conduction_pr = l_pr/(2*self.w*self.t*k_c) + self.R_base;
                TMax = W*R_conduction_pr;
                TTip = TMax;
            otherwise
                R_conduction_pr = l_pr/(2*self.w*self.t*k_c) + self.R_base;
                R_convection_pr = 1/(2*h*l_pr*(self.w + self.t));
                R_conduction_tip = (self.l - l_pr)/(2*self.w*self.t*k_c);
                R_convection_tip = 1/(2*h*(self.l-l_pr)*(self.w + self.t));
                R_total = 1/(1/R_conduction_pr + 1/R_convection_pr + ...
                1/(R_conduction_tip + R_convection_tip));

1160
                TMax = W*R_total;
                TTip = W*R_total/(R_conduction_tip + ...
                R_convection_tip)*R_convection_tip;
            end
        case 'step'
            R_conduction_pr = l_pr/(2*self.w*self.t*k_c) + ...
            self.l_a/(self.w_a*(self.t*k_c + self.t_a*k_metal)) + ...
            self.R_base;
            R_convection_pr = 1/(2*h*(l_pr+self.l_a)*(self.w + self.t));
            R_conduction_tip = (self.l - l_pr)/(2*self.w*self.t*k_c);
            R_convection_tip = 1/(2*h*(self.l-l_pr)*(self.w + self.t));
        end
    end
end

```

```

R_total = 1/(1/R_conduction_pr + 1/R_convection_pr + ...
             1/(R_conduction_tip + R_convection_tip));

TMax = W*R_total;
TTip = W*R_total/(R_conduction_tip + ...
                  R_convection_tip)*R_convection_tip;
case 'piezoelectric'
1190   R_conduction_pr = l_pr/(2*self.w*self.t*k_c) + ...
              self.l_a/(self.w_a*(k_c*self.t + self.k_aln*(self.t_a + ...
              self.t_a_seed) + k_metal*(self.t_electrode_bottom + ...
              self.t_electrode_top))) + self.R_base;
          R_convection_pr = 1/(2*h*l_pr*(self.w + self.t));
          R_conduction_tip = (self.l - l_pr)/(2*self.w*self.t*k_c);
          R_convection_tip = 1/(2*h*(self.l-l_pr)*(self.w + self.t));
          R_total = 1/(1/R_conduction_pr + 1/R_convection_pr + ...
                         1/(R_conduction_tip + R_convection_tip));

TMax = W*R_total;
TTip = W*R_total/(R_conduction_tip + ...
                  R_convection_tip)*R_convection_tip;
case 'thermal'
1200   R_conduction_pr = l_pr/(2*self.w*self.t*k_c) + ...
              self.l_a/(self.w_a*(self.t*k_c + self.t_a*k_metal));
          R_convection_pr = 1/(2*h*l_pr*(self.w + self.t));
          R_conduction_tip = (self.l - l_pr)/(2*self.w*self.t*k_c);
          R_convection_tip = 1/(2*h*(self.l-l_pr)*(self.w + self.t));
          R_conduction_heater = self.l_a/(2*self.w_a*(self.t*k_c + ...
              self.t_a*k_metal));
          R_convection_heater = 1/(2*h*self.l_a*(self.w_a + self.t_a));
          R_total = 1/(1/(R_conduction_pr + 1/(1/R_convection_heater + ...
              1/R_conduction_heater)) + 1/R_convection_pr + ...
                         1/(R_conduction_tip + R_convection_tip));

T_heater = W/(1/R_convection_heater + ...
              1/R_conduction_heater);
TMaxDivider = 1/(1/R_convection_pr + 1/(R_conduction_tip + ...
              R_convection_tip)) / (R_conduction_pr + 1/(1/R_convection_pr + ...
              1/R_convection_heater));

```

```

1220           1/(R_conduction_tip + R_convective_tip)));
TTipDivide = R_convective_tip/(R_convective_tip + R_conduction_tip);
TMax = T_heater*TMaxDivide + W*R_total;
TTip = T_heater*TMaxDivide*TTipDivide + ...
W*R_total/(R_conduction_tip + ...
R_convective_tip)*R_convective_tip;
end
end

% Calculate the approx thermal healing lengths for the cantilever
1230 % Units: m
function [l_healing_cantilever, l_healing_step] = thermalHealingLengths(self)
A = self.w*self.t;
P = 2*(self.w + self.t);
h = self.lookupHeff();

k_c = self.k_base();
l_healing_cantilever = sqrt(k_c*A/h/P);

[E_metal, rho_metal, k_metal, alpha_metal] = self.lookup_metal_properties();
1240 l_healing_step = 0;
switch self.cantilever_type
case {'step', 'thermal'}
    l_healing_step = sqrt(self.w_a*(k_c*self.t + ...
self.k_sio2*self.t_oxide + k_metal*self.t_a)/ ...
(2*(self.w_a+self.t+self.t_a)*h));
case 'piezoelectric'
    l_healing_step = sqrt(self.w_a*(k_c*self.t + ...
self.k_sio2*self.t_oxide + self.k_aln*(self.t_a + self.t_a_seed) + ...
k_metal*(self.t_electrode_bottom + self.t_electrode_top)) / ...
(2*(self.w_a+self.t+self.t_a)*h));
1250
end
end

% Model the temp profile from Joule heating via finite differences
% Assumes convection to ambient, adiabatic tip, and R_base to the
% silicon die which is clamped at the ambient temperature

```

```

function [x, Q, T_increase] = calculateTempProfile(self, varargin)

% There are several ways to call calculateTempProfile()
1260 % No arguments: temperature independent solution
% One argument (legacy): temp-dependent thermal conductivity
% Two arguments: temp-dependent sheet resistance and conductivity
switch length(varargin)
    case 0
        k_c = self.k_base();
        Rsheet = self.sheet_resistance();
        k_x = ones(1, self.numXPoints)*k_c;
        Rsheet_x = ones(1, self.numXPoints)*Rsheet;
    case 1
1270     k_x = varargin{1};
        Rsheet = self.sheet_resistance();
        Rsheet_x = ones(1, self.numXPoints)*Rsheet;
    case 2
        k_x = varargin{1};
        Rsheet_x = varargin{2};
end

% Discretize the length of the cantilever
n_points = self.numXPoints;
1280 totalLength = self.l + self.l_a;
dx = totalLength/(n_points - 1);
x = 0:dx:totalLength;

% Determine the step and PR indices
step_indices = find(x <= self.l_a);
actuator_indices = find(x <= (self.l_a - self.l_a_gap));
cantilever_indices = find(x > self.l_a);
pr_indices = intersect(cantilever_indices, ...
    find(x < (self.l_a + self.l_pr())));
1290

% Calculate Qgen_x differently depending on our temperature range
switch self.temperature_dependent_properties
    % Calculate Qgen(x) considering temperature dependent sheet resistance

```

```

% This method does not converge quickly for design optimization,
% so is best used for modeling
case 'yes'
    index_range = x <= self.l_pr();
    R_x = 2*Rsheet_x(index_range)/self.w_pr();
    R_calc = 2*trapz(x(index_range), Rsheet_x(index_range)/self.w_pr());
    I_calc = (self.v_bridge/2)/R_calc;
    Qgen_x = I_calc^2*R_x;

    % Assume power/length is constant along the piezoresistor length
    % This method works well for modeling near the ambient
    % temperature and for design optimization
    % Note: calculate Qgen_x based upon the actual lengths
    % to avoid discretization errors (line 2 here is very important)
    case 'no'
        power = (self.v_bridge/2)^2/self.resistance();
        Qgen_x = power/(x(pr_indices(end)) - x(pr_indices(1)))*ones(length(x),1);
    end

    % Setup other variables
    tempAmbient = self.T;
    h = self.lookupHeff();
    [E_metal, rho_metal, k_metal, alpha_metal] = self.lookup_metal_properties();
    K = self.w.*k_x'.*self.t.*ones(n_points, 1);
    perimeter = 2*(self.w + self.t)*ones(n_points, 1);
    Q = zeros(n_points, 1);
    1320 Q(pr_indices) = Qgen_x(pr_indices);

    % Build K (area*k_c) and P
    switch self.cantilever_type
        case 'none'
        case 'step'
            K(step_indices) = self.w_a.* (k_x(step_indices)*self.t ...
                + k_metal*self.t_a);
            perimeter(step_indices) = 2*(self.w_a + self.t_a);
        case 'thermal'
            Qheater = self.heaterPower()/(x(actuator_indices(end)) - ...

```

```

        x(actuator_indices(1)));
        Q(actuator_indices) = Qheater;
        K(step_indices) = self.w_a*(k_x(step_indices)*self.t + ...
            self.t_oxide*self.k_sio2 + k_metal*self.t_a);
        perimeter(step_indices) = 2*(self.w_a + self.t_a);
    case 'piezoelectric'
        K(step_indices) = self.w_a*(k_x(step_indices)*self.t + ...
            self.k_aln*(self.t_a + self.t_a_seed) + self.t_oxide*self.k_sio2 + ...
            k_metal*(self.t_electrode_top + self.t_electrode_bottom));
        perimeter(step_indices) = 2*(self.w_a + self.t_a);
    end

    % Build A and RHS matrices
    A = zeros(n_points, n_points);
    rhs = zeros(n_points, 1);
    for ii = 2:n_points-1
        A(ii, ii-1) = -K(ii-1)/dx^2;
        A(ii, ii) = (K(ii-1) + K(ii+1))/dx^2 + h*perimeter(ii);
        A(ii, ii+1) = -K(ii+1)/dx^2;
        rhs(ii, 1) = Q(ii) + h*perimeter(ii)*tempAmbient;
    end
    A(n_points, n_points-1:n_points) = [1 -1]; % Adiabatic at tip
    A = sparse(A); % Leads to a significant speed improvement

    % Properly handle R_base
    A(1, 1) = -1-self.R_base*K(1)/dx;
    A(1, 2) = self.R_base*K(1)/dx;

    rhs(1, 1) = -self.T;
    T_absolute = A \ rhs;
    T_increase = T_absolute - tempAmbient;
end

% Model the temperature profile of a tip-heated cantilever
% via finite differences. Intended for modeling cantilever heating
% during laser doppler vibrometer or AFM measurements.
% For optical heating applications the user needs to account for

```

```

% reflections and limited silicon absorption.

% Assumptions:
1370 % 1) Piezoresistor is turned off
% 2) Temperature-independent material properties
% 3) Beam intensity is constant over the beam diameter
% 4) Accounts for the spot size being smaller/larger than the beam
% 5) Spot size = spot diameter
% Assumes temperature-independent properties for simplicity
% Input parameters:
% inputPower (W): power input to the cantilever
% spotSize (m): length back from the tip over which the power is spread
%
1380 % Note: most of the code is reused from above
function [x, Q, T_increase] = calculateTempProfileTipHeat(self, ...
    inputPower, spotSize)

    k_c = self.k_base();
    k_x = ones(1, self.numXPoints)*k_c;
    n_points = self.numXPoints;
    totalLength = self.l + self.l_a;
    dx = totalLength/(n_points - 1);
    x = 0:dx:totalLength;
    step_indices = find(x <= self.l_a);
    actuator_indices = find(x <= (self.l_a - self.l_a_gap));
    cantilever_indices = find(x > self.l_a);
    pr_indices = intersect(cantilever_indices, ...
        find(x < (self.l_a + self.l_pr())));
1390

    % Calculate the input power at the tip.
    % Accounts for some fraction of the beam missing the cantilever
    % if the cantilever width is comparable to the beam diameter
    % Assumes the center is aimed at the tip
    1400 Q = zeros(n_points, 1);
    fractionAbsorbed = min((spotSize/2)*self.w/(pi*(spotSize/2)^2), 1);
    Q(x > (self.l - spotSize)) = fractionAbsorbed*inputPower/spotSize;

    tempAmbient = self.T;

```

```

h = self.lookupHeff();
[E_metal, rho_metal, k_metal, alpha_metal] = self.lookup_metal_properties();
K = self.w.*k_x'.*self.t.*ones(n_points, 1);
perimeter = 2*(self.w + self.t)*ones(n_points, 1);

1410 switch self.cantilever_type
    case 'none'
    case 'step'
        K(step_indices) = self.w_a.* (k_x(step_indices)*self.t + ...
            k_metal*self.t_a);
        perimeter(step_indices) = 2*(self.w_a + self.t_a);
    case 'thermal'
        Qheater = self.heaterPower()/(x(actuator_indices(end)) - ...
            x(actuator_indices(1)));
        Q(actuator_indices) = Qheater;
        K(step_indices) = self.w_a_active*(k_x(step_indices)*self.t + ...
            self.t_oxide*self.k_sio2 + k_metal*self.t_a);
        perimeter(step_indices) = 2*(self.w_a + self.t_a);
    case 'piezoelectric'
        K(step_indices) = self.w_a*(k_x(step_indices)*self.t + ...
            self.k_aln*(self.t_a + self.t_a_seed) + self.t_oxide*self.k_sio2 + ...
            k_metal*(self.t_electrode_top + self.t_electrode_bottom));
        perimeter(step_indices) = 2*(self.w_a + self.t_a);
    end

1430 A = zeros(n_points, n_points);
rhs = zeros(n_points, 1);
for ii = 2:n_points-1
    A(ii, ii-1) = -K(ii-1)/dx^2;
    A(ii, ii) = (K(ii-1) + K(ii+1))/dx^2 + h*perimeter(ii);
    A(ii, ii+1) = -K(ii+1)/dx^2;
    rhs(ii, 1) = Q(ii) + h*perimeter(ii)*tempAmbient;
end
A(n_points, n_points-1:n_points) = [1 -1];
A = sparse(A);
1440 A(1, 1) = -1-self.R_base*K(1)/dx;
A(1, 2) = self.R_base*K(1)/dx;

```

```

rhs(1, 1) = -self.T;
T_absolute = A \ rhs;
T_increase = T_absolute - tempAmbient;
end

% Calculate the temp profile accounting for k(T) and R_s(T)
% If temperature_dependent_properties is set to 'yes' then this
1450 % function should be called correctly. The temperature dependent option
% is much slower than its temperature independent equivalent.
function [x, Q, T_final] = calculateTempProfileTempDependent(self)

% Iterative calculate the temperature distribution and k(x)
% Initialize k(x) solver with our initial guess
% This approach is required because the system is nonlinear
[x, Q, T_initial] = self.calculateTempProfile();
T_current = T_initial;
residual = 1;

1460 while residual > 1e-6
    absoluteTemp = T_current + self.T;

    % Calculate k_si and Rsheet from the current temperature guess
    [x, z, k_z_x] = self.thermal_conductivity_profile(absoluteTemp);
    k_x = mean(k_z_x, 1); % Average k for a cross-section
    Rsheet_x = self.RSheetProfile(x, T_current);

    % Scale the resistance so that it matches our nominal value
    % This choice is geared towards modeling experimental results
1470    R_nominal = self.resistance();
    x_resistor = x(x <= self.l_pr());
    Rsheet_resistor = Rsheet_x(x <= self.l_pr());
    R_calc = 2*trapz(x_resistor, Rsheet_resistor/self.w_pr()) + ...
        3.4*Rsheet_resistor(end) + 2*self.R_contact;
    Rsheet_x = Rsheet_x * (R_nominal/R_calc); % Scale

    % Update the temperature guess using k_x, Rsheet_x

```

```

[x, Q, T_new] = self.calculateTempProfile(k_x, Rsheet_x);

1480
    residual = sum((T_new - T_current).^2);
    T_current = T_new;
end
T_final = T_current;
end

% Calculate the approximate thermal conductivity of the cantilever
% Units: W/m-K
function k_base = k_base(self)
1490
    k_x = self.k_x();
    k_base = k_x(1);
end

% Calculate k(x) as a function of cantilever thickness
% Units: W/m-K
function k_effective = k_x(self) %#ok<*MANU>
switch self.thermal_modeling
    case {'none', 'approx'}
        % Model from Asheghi (1997)
        t0 = 120e-9;
        k_effective = self.k_si*self.t./(t0 + self.t);
    case 'exact'
        [x, z, k] = self.thermal_conductivity_profile(self.T);
        k_effective = mean(k);
    end
end

% Calculate k(x,z) using a temp/thickness dependent model
% Source: "Modeling and Data for Thermal Conductivity ...", Liu et al, 2006.
1510
% Units: W/m-K
function [x, z, k] = thermal_conductivity_profile(self, T)
    % Phonon Group velocities (m/sec)
    nu_T = 5860;
    nu_TU = 2000;
    nu_L = 8480;

```

```

% Density of state coefficients
C_T = self.k_b/(2*pi^2*nu_T)*(self.k_b/self.h_bar)^3;
C_TU = self.k_b/(2*pi^2*nu_TU)*(self.k_b/self.h_bar)^3;
1520 C_L = self.k_b/(2*pi^2*nu_L)*(self.k_b/self.h_bar)^3;

% Debye tempratures
theta_1 = 180; %K
theta_2 = 210; %K
theta_3 = 570; %K

% Phonon specific heat/volume
c_TU = 7.5e5; % J/m^3-K
c_L = 3.2e5; % J/m^3-K
1530

% Other Constants
B_T = 9.3e-13; % 1/K^3
B_TU = 5.5e-18; % s
B_L = 2e-24; % s/K
beta_TU = B_TU*(self.k_b/self.h_bar)^2; % 1/s-K^2
beta_L = B_L*(self.k_b/self.h_bar)^2; % 1/s-K^3
beta_T = B_T*(self.k_b/self.h_bar); % 1/s-K^4

% Phonon frequencies
1540 omega_2 = self.k_b/self.h_bar*theta_2;
omega_3 = self.k_b/self.h_bar*theta_3;

% Crystal lattice parameters
Gamma = 2.1e-4; % computed from isotopes
nu_s = (1/3*(nu_L^-1 + 2*nu_T^-1))^-1; % Speed of sound

R_si = 5.4e-10; %m - radius of the silicon host atom
V_si = (R_si/2)^3; % m^3 - silicon lattice volume
M_si = 4.59e-26; % kg - mass of silicon host atom
1550 A_isotopes = Gamma*V_si/(4*pi*nu_s^3);

% Other constants

```

```

Q = 4.2; % Constant scattering matrix parameter
gamma = 1.6; % Gruneisen parameter
psi = 3.5; % Modification factor to account for inconsistency of the Si

% Phonon mean free paths in nearly pure bulk silicon
function lambda_TU = lambda_TU(T)
    lambda_TU = C_TU*(theta_2^2-theta_1^2)./(2*nu_TU*c_TU*beta_TU.*T);
end

function lambda_L = lambda_L(T)
    lambda_L = C_L*theta_3./(nu_L*c_L*beta_L*T.^3);
end

% Phonon mean free paths including thermal vibrations
function lambda_TU_HT = lambda_TU_HT(T)
    lambda_TU_HT = (1./lambda_TU(T) + ...
        1/(nu_TU*(A_isotopes*omega_2^4)^-1)).^-1;
end

1570

function lambda_L_HT = lambda_L_HT(T)
    lambda_L_HT = (1./lambda_L(T) + ...
        1/(nu_L*(A_isotopes*omega_3^4)^-1)).^-1;
end

% Contribution to thermal conductivity from phonon modes
function k_D_layer_TU_HT = k_D_layer_TU_HT(T, d, n, dM, dR)
    A = 1./lambda_TU_HT(T) + ...
        1./(nu_TU*((n*V_si^2)/(4*pi*nu_s^3)).*(dM/M_si)^2 + ...
        2*n*V_si^2*Q^2*gamma^2./(pi*nu_s^3)*(dR/R_si)^2)*omega_2^4).^-1;
    k_D_layer_TU_HT = 2/3.*c_TU.*nu_TU.*A.^-1.*((1+3.)/(8*d/(A.^-1*psi))).^-1;
end

1580

function k_D_layer_L_HT = k_D_layer_L_HT(T, d, n, dM, dR)
    A = 1./lambda_L_HT(T) + ...
        1./(nu_L*((n*V_si^2)/(4*pi*nu_s^3)).*(dM/M_si)^2 + ...
        2.*n*V_si^2*Q^2*gamma^2./(pi*nu_s^3)*(dR/R_si)^2)*omega_3^4).^-1;
    k_D_layer_L_HT = 1/3.*c_L.*nu_L.*A.^-1.*((1+3.)/(8*d/(A.^-1*psi))).^-1;

```

```

1590      end

function k_D_layer_HT = k_D_layer_HT(T, d, n, dM, dR)
n = n*1e6; % Convert from 1/cc to 1/m^3
k_D_layer_HT = k_D_layer_L_HT(T, d, n, dM, dR) + ...
    k_D_layer_TU_HT(T, d, n, dM, dR);
end

% Doping constants
switch self.doping_type
    case 'arsenic'
        fprintf('ERROR: Arsenic dM and dR values not included');
        pause
    case 'boron'
        dM = 18*1.673e-27;
        dR = 0.175e-10;
    case 'phosphorus'
        dM = 3*1.673e-27;
        dR = 0.175e-10;
    end
1600
% Vector calculation of thermal conductivity
% Roughly 10x faster than iterative calculation!
[z, active_doping, total_doping] = self.doping_profile();
x = linspace(0, self.l, self.numXPoints);

% let dim(T) and dim(total_doping) be numZPoints x numXPoints
n_matrix = total_doping'*ones(1, self.numXPoints);
if length(T) == 1
    T_matrix = ones(self.numZPoints, 1)*T'*ones(1, self.numXPoints);
1620
else
    T_matrix = ones(self.numZPoints, 1)*T';
end

k = k_D_layer_HT(T_matrix, self.t, n_matrix, dM, dR);
end

```

```

% ===== Calculate resolution =====
function force_resolution = force_resolution(self)
    force_resolution = self.integrated_noise()/self.force_sensitivity();
1630    end

function force_resolution = thermomechanical_force_noise(self)
    force_resolution = self.thermo_integrated()/self.force_sensitivity();
end

function displacement_resolution = displacement_resolution(self)
    displacement_resolution = self.force_resolution()/self.stiffness();
end

1640 function force_noise_density = force_noise_density(self, freq)
    force_noise_density = self.voltage_noise(freq)/self.force_sensitivity();
end

function force_noise_density = resonant_force_noise_density(self)
    [omega_damped_hz, Q] = self.omega_damped_hz_and_Q(); %#ok<*NASGU>
    freq = omega_damped_hz;
    force_noise_density = self.force_noise_density(freq)*1e15;
end

1650 function surface_stress_resolution = surface_stress_resolution(self)
    surface_stress_resolution = self.integrated_noise()/
        self.surface_stress_sensitivity();
end

% ===== Multilayer beam mechanics and actuation =====

% Calculate the cantilever curvature per unit applied moment
1660 function Cm = calculateActuatorNormalizedCurvature(self)
    Zm = self.actuatorNeutralAxis();
    [z, E, A, I] = self.lookupActuatorMechanics();
    Z_offset = z - Zm;

```

```

Cm = 1./sum(E.*(I + A.*Z_offset.^2));
end

% Calculate the thickness of silicon required to have the same EI
% as the actuator/step at the base
function t_equivalent = calculateEquivalentThickness(self)
1670    function residual_squared = findEIResidual(t_equivalent_guess)
        EI_calculated = self.modulus()*self.w_a*t_equivalent_guess^3/12;
        residual_squared = (EI_effective/EI_calculated - 1)^2;
    end

    EI_effective = 1/self.calculateActuatorNormalizedCurvature();
    options = optimset('TolX', 1e-12, 'TolFun', 1e-12, 'Display', 'off');
    t_equivalent = fminbnd(@findEIResidual, self.t/100, self.t*100, options);
end

1680 function Zm = actuatorNeutralAxis(self)
    [z, E, I] = self.lookupActuatorMechanics();
    Zm = sum(z.*E.*A)/sum(E.*A);
end

function d31 = d31(self)
    if self.d31_manual == 0
        d31 = interp1(self.d31_t, self.d31_aln, self.t_a, 'spline');
    else
        d31 = self.d31_manual;
    end
1690 end

function [x, deflection] = calculateDeflection(self)
    n_points = self.numXPoints;
    totalLength = self.l + self.l_a;
    dx = totalLength/(n_points - 1);
    x = 0:dx:totalLength;

    M = 0; % external moment is zero
    P = 0; % external load is zero
1700

```

```

[z, E, A, I] = self.lookupActuatorMechanics();
stress = self.calculateActuatorStress();

% Calculate the curvature and neutral axis
% The curvature may vary as a function of position (e.g. thermal)
% so calculate the deflection by integrating the curvature twice
C = zeros(length(x), 1);

1710 % Calculate the curvature, C along the cantilever length
for ii = 1:length(x)
    C(ii) = -((M - sum(z.*A.*stress(ii,:)))*sum(E.*A) + ...
        (P + sum(A.*stress(ii,:)))*sum(E.*z.*A))/ ...
        (sum(E.*A)*sum(E.*(I+A.*z.^2)) - sum(z.*E.*A)^2);
end

% No curvature beyond the end of the actuator
% i.e. no dopant bending included in this calculation
C(x > self.l_a) = 0;

1720 % Calculate the deflection from the curvature by integrating
theta = cumtrapz(x, C);
deflection = cumtrapz(x, tan(theta));
end

% Calculate the stress as a function along the actuator length
function stress = calculateActuatorStress(self)
    [z, E, A, I] = self.lookupActuatorMechanics(); %#ok<*ASGLU>
    [E_metal, rho_metal, k_metal, alpha_metal] = self.lookup_metal_properties();
    [x_temp, Q, temp] = self.calculateTempProfile();
    temp = temp + self.T; % Use absolute temperature for these calculations

    intrinsic_stress = ones(self.numXPoints, 1)*self.film_intrinsic_stress();
    intrinsic_stress(x_temp > self.l_a, :) = 0;

    switch self.cantilever_type
        case {'step', 'thermal'}

```

```

    cte = [self.alpha_si self.alpha_sio2 alpha_metal];
    thermal_stress = (temp-self.T_ref).*ones(self.numXPoints,1)*(cte.*E);
1740   thermal_stress(x_temp > self.l_a, 2:3) = 0;
    piezoelectric_stress = zeros(self.numXPoints, 3);

    case 'piezoelectric'
    cte = [self.alpha_si self.alpha_sio2 self.alpha_aln ...
        alpha_metal self.alpha_aln alpha_metal];
    thermal_stress = (temp-self.T_ref).*ones(self.numXPoints,1)*(cte.*E);
    thermal_stress(x_temp > self.l_a, 2:6) = 0;

    % Calculate the piezoelectric stress
1750   % The seed electric field will vary depending on
    E_field = [0 0 0 0 self.v_actuator/self.t_a 0]';
    d31 = [0 0 0 0 self.d31() 0]';

    piezoelectric_stress = ones(self.numXPoints,1)*E*d31*E_field';
    piezoelectric_stress(x_temp > self.l_a - self.l_a_gap, :) = 0;
end

    % Scale the stress by the active width of the device
    stress = intrinsic_stress + thermal_stress + ...
1760   self.w_a_active/self.w_a*piezoelectric_stress;
end

function [mu_z sigma_z] = tip_deflection_distribution(self)
    v_actuator_temp = self.v_actuator;
    self.v_actuator = 0;
    self.film_stress = 'random';
    for ii = 1:cantilever.numRandomStressIterations
        [x_ii, z_ii] = self.calculateDeflection();

        x = x_ii;
        z(:, ii) = z_ii;

        z_tip(ii) = self.tipDeflection();
    end

```

```

    self.v_actuator = v_actuator_temp;
    self.film_stress = 'nominal';

    mu_z = mean(z_tip);
    sigma_z = std(z_tip);
1780 end

function plot_tip_deflection_distribution(self)
    v_actuator_temp = self.v_actuator;
    self.v_actuator = 0;
    self.film_stress = 'random';
    for ii = 1:cantilever.numRandomStressIterations
        [x_ii, z_ii] = self.calculateDeflection();

        x = x_ii;
        z(:, ii) = z_ii;
1790    end
    self.v_actuator = v_actuator_temp;
    self.film_stress = 'nominal';

    figure
    plot(x_ii*1e6, z*1e6, 'Color', [11 132 199]./255);
    xlabel('Distance from base (\mu m)');
    ylabel('Deflection (\mu m)');
    box off;
1800
end

function sigma_i = film_intrinsic_stress(self)
    % film_stress = 'nominal' => use average stress
    % film_stress = 'random' => use random, normally distributed value
    switch self.film_stress
        case 'nominal'
            sigma = 0;
        case 'random'
            sigma = 0.25; % 1/4 because 1/2 from diff and 1/2 from range = 2 STDevs
1810    end

```

```

switch self.cantilever_type
  case {'step', 'thermal'}
    sigma_i(1) = mean(self.sigma_si_range) + ...
      sigma*randn*diff(self.sigma_si_range);
    sigma_i(2) = mean(self.sigma_sio2_range) + ...
      sigma*randn*diff(self.sigma_sio2_range);
    sigma_i(3) = mean(self.sigma_al_range) + ...
      sigma*randn*diff(self.sigma_al_range);
1820
  case 'piezoelectric'
    sigma_i(1) = mean(self.sigma_si_range) + ...
      sigma*randn*diff(self.sigma_si_range);
    sigma_i(2) = mean(self.sigma_sio2_range) + ...
      sigma*randn*diff(self.sigma_sio2_range);
    sigma_i(3) = mean(self.sigma_aln_range) + ...
      sigma*randn*diff(self.sigma_aln_range);
    sigma_i(5) = mean(self.sigma_aln_range) + ...
      sigma*randn*diff(self.sigma_aln_range);
  switch self.metal_type
    case 'titanium'
      sigma_i(4) = mean(self.sigma_ti_range) + ...
        sigma*randn*diff(self.sigma_ti_range);
      sigma_i(6) = mean(self.sigma_ti_range) + ...
        sigma*randn*diff(self.sigma_ti_range);
    case 'molybdenum'
      sigma_i(4) = mean(self.sigma_mo_range) + ...
        sigma*randn*diff(self.sigma_mo_range);
      sigma_i(6) = mean(self.sigma_mo_range) + ...
        sigma*randn*diff(self.sigma_mo_range);
1840
    end
  end
end

% The thermal actuator power dissipation
% Units: W
function power = heaterPower(self)
  power = self.v_actuator^2/self.R_heater;

```

```

    end

1850
    % Calculate the time constant and -3 dB freq of the thermal actuator
    % Note that in practice this is roughly 4x too idealistic
    % Units: s and Hz
    function [tau, freq] = heaterTimeConstant(self)
        k_c = self.k_base();
        [E_metal, rho_metal, k_metal, alpha_metal] = self.lookup_metal_properties();
        h = self.lookupHeff();

        % Calculate the thermal resistances
1860
        R_conduction = (self.l_a-self.l_a_gap)/(2*self.w_a*(self.t*k_c + ...
            self.t_a*k_metal + self.t_oxide*self.k_sio2));
        R_convective = 1/(2*h*self.l_a*(self.w_a + self.t_a));
        R = 1/(R_conduction + self.R_base) + 1/R_convective; % K/W
        C = self.l_a*self.w_a*(self.t_a*self.rho_al*self.Cv_al + ...
            self.t*self.rho_si*self.Cv_si + ...
            self.t_oxide*self.rho_sio2*self.Cv_sio2); % J/K

        tau = R*C;
        freq = 1/(2*pi*tau);
1870
    end

    % Calculate the rise time using the Q-dependent formula
    % Source: "Control Systems Engineering", Nise
    function t_rise = mechanicalRiseTime(self)
        [omega_d, Q] = self.omega_damped_and_Q();
        zeta = 1/(2*Q);
        t_rise = 1/omega_d*(1.76*zeta^3-0.417*zeta^2+1.039*zeta+1);
    end

1880
    % Calculate dz_tip/dT_ambient -> temperature stability of the cantilever
    % Units: m/K
    function dzdT = tipTempDeflectionVsAmbientTemp(self)
        temp_delta_values = -5:5;

        T_original= self.T;

```

```
for ii = 1:length(temp_delta_values)
    self.T = T_original + temp_delta_values(ii);
    tipDeflection(ii) = self.tipDeflection();
end
1890 self.T = T_original;

p = polyfit(temp_delta_values, tipDeflection, 1);
dzdT = p(1);
end

% Calculate the electrical current in the heater
% Units: A
function current = heaterCurrent(self)
    current = self.v_actuator/self.R_heater;
1900 end

% Calculate the tip deflection including all of the various effects
% Units: m
function z_tip = tipDeflection(self)
    [x, z] = self.calculateDeflection();
    z_tip = z(end);
end

% Calculate the tip deflection with the actuator on vs. off
1910 % This approach cancels any static deflections (e.g. dopant bending)
% Units: m
function z_tip = actuatorTipDeflectionRange(self)
    v_actuator_temp = self.v_actuator;
    self.v_actuator = 0;
    [x, z_off] = self.calculateDeflection();
    z_initial = z_off(end);

    self.v_actuator = v_actuator_temp;
    [x, z_on] = self.calculateDeflection();
    z_final = z_on(end);

    z_tip = abs(z_final - z_initial);
1920
```

```

    end

    % Calculate dopant strain induced bending
    % The calculation method handles arbitrary strain gradients
    % Source: "Residual strain and resultant postrelease deflection ...", 1999
    function [x, C, theta, deflection] = calculateDopantBending(self)

1930      % Source: "Effects of phosphorus on stress...", JMM, 1999.
        delta = 4.5e-24;

        n_points = self.numXPoints;
        dx = self.l/(n_points - 1);
        x = 0:dx:self.l;

        [depth, active_doping, total_doping] = self.doping_profile();
        E = self.modulus();
        width = self.w; % But it doesn't matter
        thickness = self.t;

1940      % Shift coordinates so that strain is measured from the bottom
        y = linspace(0, thickness, length(depth))';
        dopant_concentration = flipud(active_doping);

        strain = dopant_concentration * delta;
        stress = strain*E;

        ybar = trapz(y, y.*stress') ./ trapz(y, stress');
        h1 = thickness - ybar;
        h2 = ybar;

1950      i1 = find(y > ybar);
        i2 = find(y <= ybar);
        y1 = y(i1);
        y2 = y(i2);
        stress1 = stress(i1);
        stress2 = stress(i2);

```

```

1960    sigma1 = 1/h1*trapz(y1, stress1);
    sigma2 = 1/h2*trapz(y2, stress2);

    ybar1 = trapz(y1, y1.*stress1')/trapz(y1, stress1');
    ybar2 = trapz(y2, y2.*stress2')/trapz(y2, stress2');

    I1 = width*h1^3/12;
    I2 = width*h2^3/12;
    I = width*thickness^3/12;

1970    sigmai = (sigma2 - sigma1)/(2 - h1*width*(ybar1 - ybar)^2/I1 + ...
    h2*width*(ybar - ybar2)^2/I2);

    M1 = sigmai*h1*width*(ybar1 - ybar);
    M2 = sigmai*h2*width*(ybar - ybar2);

    % Calculate the radius of curvature, assuming it is small
    R = (E*I)/(M1 + M2);
    C = zeros(length(x), 1);
    C_pr = 1/R;
1980    C(x <= self.l_pr()) = C_pr;

    theta = cumsum(C.*dx);
    deflection = cumsum(theta.*dx);
end

    % Calculate just the tip deflection from dopant bending
    % Units: m
    function tip_deflection = calculateDopantTipDeflection(self)
        [x, C, theta, deflection] = self.calculateDopantBending();
        tip_deflection = max(deflection);
1990    end

    % ===== Handy plotting functions =====
    function plotTempProfile(self)
        switch self.temperature_dependent_properties
            case 'yes'

```

```
[x, Q, temp] = self.calculateTempProfileTempDependent();
case 'no'
    [x, Q, temp] = self.calculateTempProfile();
end

2000 figure
hold all
plot(1e6*x, temp);
plot(1e6*x, Q, '--');
hold off
xlabel('X (um)');
ylabel('Temp Rise (K)');
legend('Temperature', 'Power/Unit Length');
end

function plotTempProfileActuatorComparison(self)

    [x, Q, temp] = self.calculateTempProfile();

    figure
    subplot(2,1,1)
    plot(1e6*x, Q, '--');
    box off;
2020 ylabel('Power (\muW/\mu m)');

    subplot(2,1,2)
    hold all
    plot(1e6*x, temp);

    v_actuator_temp = self.v_actuator;
    self.v_actuator = 0;
    [x, Q, temp] = self.calculateTempProfile();
    self.v_actuator = v_actuator_temp;

    plot(1e6*x, temp);
    box off;
2030
```

```

    hold off
    xlabel('Distance from base (\mu m)');
    ylabel('\Delta T (K)');
end

2040 function plot_thermal_conductivity_profile(self)
    [x, z, k] = self.thermal_conductivity_profile(self.T);

    figure
    plot(z*1e9, k);
    xlabel('Distance from Surface (nm)');
    ylabel('k (W/m-K)');
    ylim([0 150]);
    xlim([0 self.t*1e9]);
end

2050 function plotDopantBending(self)
    [x, C, theta, deflection] = self.calculateDopantBending();

    figure
    plot(1e6*x, 1e9*deflection);
    xlabel('Distance from Base (um)');
    ylabel('Cantilever Deflection (nm)');
    box off;
end

2060 function plotDeflectionAndTemp(self)

    v_actuator_temp = self.v_actuator;
    self.v_actuator = 0;
    [x, deflectionInitial] = self.calculateDeflection();

    self.v_actuator = v_actuator_temp; %#ok<*MCHV2>
    [x, deflectionFinal] = self.calculateDeflection();
    [x, Q, temp] = self.calculateTempProfile();
    deflection = deflectionFinal - deflectionInitial;

```

```

figure
subplot(2,1,1);
plot(1e6*x, temp);
ylabel('Temp Rise (K)');
box off;

subplot(2,1,2);
plot(1e6*x, 1e6*deflection);
2080 xlabel('Distance from Base (um)');
ylabel('Deflection (\mu m)');
box off;
end

function plot_noise_spectrum(self)
freq = logspace( log10(self.freq_min), log10(self.freq_max), ...
    cantilever.numFrequencyPoints);

figure
2090 hold all
plot(freq, sqrt(self.johnson_PSD(freq)), 'LineWidth', 2);
plot(freq, sqrt(self.hooge_PSD(freq)), 'LineWidth', 2);
plot(freq, sqrt(self.thermo_PSD(freq)), 'LineWidth', 2);
plot(freq, sqrt(self.amplifier_PSD(freq)), 'LineWidth', 2);
plot(freq, self.voltage_noise(freq), 'LineWidth', 2);
hold off
set(gca, 'xscale','log', 'yscale','log');
set(gca, 'LineWidth', 1.5, 'FontSize', 14);
ylabel('Noise Voltage Spectral Density (V/rtHz)', 'FontSize', 16);
xlabel('Frequency (Hz)', 'FontSize', 16);
legend('Johnson', 'Hooge', 'Thermo', 'Amp', 'Total')
2100 end

% ===== Handy lookup functions =====
% Calculate elastic modulus based upon dopant type
% Assume we're using the best piezoresistor orientation.
function elastic_modulus = modulus(self)

```

```

switch self.doping_type
  case 'boron'
    elastic_modulus = 169e9; % <110> direction
  case 'phosphorus'
    elastic_modulus = 130e9; % <100> direction
  case 'arsenic'
    elastic_modulus = 130e9; % <100> direction
end
end

% Lookup the appropriate fluid physical properties
function [rho_f, eta_f] = lookupFluidProperties(self)
2120 switch self.fluid
  case 'air'
    rho_f = self.rho_air;
    eta_f = self.eta_air;
  case 'water'
    rho_f = self.rho_water;
    eta_f = self.eta_water;
  case 'arbitrary'
    rho_f = self.rho_arb;
    eta_f = self.eta_arb;
  otherwise
    fprintf('ERROR - Unknown fluid: %s', self.fluid);
    pause
  end
end

% Lookup effective convection coefficient
% Or try calculating it using a shape factor (not recommended)
% Units: W/m^2-K
function h_eff = lookupHeff(self)
2140 switch self.h_method
  case 'fixed'
    switch self.fluid
      case 'vacuum'
        h_eff = self.h_vacuum;

```

```

        case 'air'
        h_eff = self.h_air;
    case 'water'
        h_eff = self.h_water;
    case 'arbitrary'
        h_eff = self.h_arb;
2150    end

    case 'calculate'
        switch self.fluid
            case 'vacuum'
                k_f = self.k_vacuum;
            case 'air'
                k_f = self.k_air;
            case 'water'
                k_f = self.k_water;
            case 'arbitrary'
                k_f = self.k_arb;
2160        end

        % Calculate h using the shape factor
        A = self.w*self.t;
        P = 2*(self.w + self.t);

        % Assume a thin rectangle to an infinite plane
        z = 450e-6; % cantilever die thickness, m
        S = 2*pi*self.l/log(2*pi*z/self.w);

2170    h_eff = k_f*S/(P*self.l); % calculate h from the Hu equation
    end
end

% Simple lookup function for the film stack properties
function [E, rho, k, alpha] = lookup_metal_properties(self)
    switch self.metal_type
        case 'aluminum'
            E = self.E_al;
2180            rho = self.rho_al;
            k = self.k_al;

```

```

alpha = self.alpha_al;
case 'titanium'
E = self.E_ti;
rho = self.rho_ti;
k = self.k_ti;
alpha = self.alpha_ti;
case 'molybdenum'
E = self.E_mo;
rho = self.rho_mo;
k = self.k_mo;
alpha = self.alpha_mo;
otherwise
fprintf('ERROR: Unknown metal_type');
pause
end
end

% Treat the actuator as encompassing the entire device width here
2200 % and scale the actuation later (to account for the metal traces
% running alongside
function [z_layers, E_layers, A, I] = lookupActuatorMechanics(self)
[E_metal, rho_metal, k_metal, alpha_metal] = self.lookup_metal_properties();
switch self.cantilever_type
case 'none';
error('ERROR in lookupActuatorMechanics() - cantilever_type = none');
case {'step', 'thermal'};
t_layers = [self.t self.t_oxide self.t_a];
w_layers = [self.w_a self.w_a self.w_a];
E_layers = [self.modulus() self.E_sio2 E_metal];
2210 case 'piezoelectric';
t_layers = [self.t self.t_oxide self.t_a_seed ...
self.t_electrode_bottom self.t_a self.t_electrode_top];
w_layers = [self.w_a self.w_a self.w_a ...
self.w_a self.w_a self.w_a];
E_layers = [self.modulus() self.E_sio2 self.E_aln ...
E_metal self.E_aln E_metal];
end

```

```

2220      z_layers = zeros(1, length(t_layers));
      for ii = 1:length(t_layers)
        % z(1) = t(1)/2, z(2) = t(1) + t(2)/2
        z_layers(ii) = sum(t_layers) - sum(t_layers(ii:end)) + t_layers(ii)/2;
      end
      A = w_layers.*t_layers;
      I = (w_layers.*t_layers.^3)/12;
    end

    % ===== Beam mechanics, resonant frequency and damping =====
2230
    % Calculate the cantilever spring constant
    % Units: N/m
    function stiffness = stiffness(self)
      k_tip = self.modulus() * self.w * self.t^3 / (4*self.l^3);

      switch self.cantilever_type
        case 'none'
          stiffness = k_tip;

2240      % For devices with an actuator/step at the base...
      % Apply a test force and calculate the resulting deflection
      % Treating the system as two springs in series is not valid because:
      % 1) the force is applied beyond the tip of the thick base,
      %   so the moment it sees is larger than would be expected
      % 2) the angle at the end of the thick base is integrated
      %   along the remaining length of the cantilever
      otherwise
        F = 1e-9; % Test force (N)
        EI_base = 1/self.calculateActuatorNormalizedCurvature();
2250      EI_tip = self.modulus()*self.w*self.t^3/12;

        % Deflection at the tip of the base
        z_base = F*self.l_a^2*(3*(self.l+self.l_a)-self.l_a)/(6*EI_base);

        % Angle at the tip of the base

```

```

theta_base = F*(2*(self.l_a+self.l)-self.l_a)*self.l_a/(2*EI_base);

% Deflection at the very tip
z_tip = z_base + tan(theta_base)*self.l + F*self.l^2*2*self.l/(6*EI_tip);
2260 stiffness = F/z_tip;
end
end

% Effective mass of the cantilever beam
% Does not include the effective mass of the base/actuator section
% f0 is calculated using Rayleigh-Ritz there
% Units: kg
function effective_mass = effective_mass(self)
cantilever_effective_mass = 0.243 * self.rho_si * self.w * self.t * self.l;
2270 effective_mass = cantilever_effective_mass + self.tip_mass;
end

% Resonant frequency for undamped vibration (first mode).
% For a simple cantilever, use Bernoulli beam theory.
% For step/actuator designs, use Rayleigh-Ritz (i.e. U_kin = U_strain)
% The R-R results agree with 2D and 3D FEA results to <3% when
% (1 < t_a/t < 6, 1/5 < l_a/l < 5)
% Units: radians/sec
function omega_vacuum = omega_vacuum(self)
2280 omega_bernoulli = sqrt( self.stiffness() / self.effective_mass());

switch self.cantilever_type
case 'none'
    omega_vacuum = omega_bernoulli;
otherwise
    options = optimset('TolX', 1e-6, 'TolFun', 1e-6, 'Display', 'off');
    options = optimset('TolX', 1e-12, 'TolFun', 1e-12, 'Display', 'off');
    omega_vacuum = fminbnd(@findEnergyResidual, 1, ...
        100*omega_bernoulli, options);
end

% Function for iteratively finding the resonant frequency

```

```

function residual_squared = findEnergyResidual(omega_guess)
    [U_elastic, U_kinetic] = calculateEnergies(omega_guess);
    residual_squared = (U_elastic/U_kinetic - 1)^2;
end

% Function for computing the elastic and kinetic energy of the beam
function [U_elastic, U_kinetic] = calculateEnergies(omega)

2300
    % Discretize the length of the cantilever
    totalLength = self.l + self.l_a;
    dx = totalLength/(self.numXPoints - 1);
    x = 0:dx:totalLength;
    base_indices = find(x <= self.l_a);
    tip_indices = find(x > self.l_a);

    % Define the multilayer mechanics
    EI_base = 1/self.calculateActuatorNormalizedCurvature();
    EI_tip = self.modulus()*self.w*self.t^3/12;

2310
    % Empirical correction factors that give better agreement with FEA results
    % Account for cases where t_a >> t, w_a >> w, l_a >> l
    EI_base = (self.t/self.t_a)^.25*EI_base;
    EI_base = (self.w/self.w_a)^.1*EI_base;
    EI_base = (self.l/self.l_a)^.1*EI_base;

    % Generate an approximate cantilever deflection profile assuming a
    % point load force at the tip of the beam. Stitch together the two
    % sections (the moment is constant despite the EI discontinuity)
    tip_deflection = 1e-6; % Apply a test force
    F = tip_deflection*3*EI_tip/self.l^3;
    moment = F*(totalLength-x);
    deflection(base_indices) = -F*x(base_indices).^2.* ...
        (3*totalLength-x(base_indices))/(6*EI_base);

    % x-coordinate from the end of l_a
    x_relative = x(tip_indices) - x(max(base_indices));

```

```

2330 % Continue with the slope that is at the end of the base section
if max(base_indices) > 1
    tip_slope = (deflection(max(base_indices)) - ...
        deflection(max(base_indices)-1))/dx;
% Catch any cases wher l_a = 0 (shouldn't normally happen)
else
    tip_slope = 0;
end
deflection(tip_indices) = deflection(max(base_indices)) - ...
F*x_relative.^2.* (3*self.l_a-x_relative)/(6*EI_tip) + tip_slope*x_relative;

2340 [E_metal, rho_metal, k_metal, alpha_metal] = ...
self.lookup_metal_properties();
dm_tip = self.w*a* self.t* self.rho_si;
switch self.cantilever_type
case {'step', 'thermal'}
    dm_base = self.w*a*(self.t* self.rho_si + ...
        self.t_oxide* self.rho_sio2 + self.t_a* rho_metal);
case 'piezoelectric'
    dm_base = self.w*a*(self.t* self.rho_si + ...
        self.t_oxide* self.rho_sio2 + ...
        self.rho_aln*(self.t_a + self.t_a_seed) + ...
        rho_metal*(self.t_electrode_bottom + self.t_electrode_top));
end

2350 % Piecewise kinetic and elastic energies
Udx_elastic(base_indices) = ...
    .5*moment(base_indices).^2*dx/EI_base;
Udx_kinetic(base_indices) = ...
    .5*(omega*deflection(base_indices)).^2*dx*dm_base;
Udx_elastic(tip_indices) = ...
    .5*moment(tip_indices).^2*dx/EI_tip;
Udx_kinetic(tip_indices) = ...
    .5*(omega*deflection(tip_indices)).^2*dx*dm_tip;

2360 U_elastic = trapz(x, Udx_elastic);
U_kinetic = trapz(x, Udx_kinetic);

```

```
    end
end

2370 % Resonant frequency for undamped vibration (first mode)
% Units: cycles/sec
function omega_vacuum_hz = omega_vacuum_hz(self)
    omega_vacuum_hz = self.omega_vacuum() / (2*pi);
end

% Calculate the damped natural frequency and Q (via Brumley/Sader)
function [omega_damped, Q] = omega_damped_and_Q(self)

    % If we're in vacuum, just return the vacuum frequency
2380 switch self.fluid
    case 'vacuum'
        omega_damped = self.omega_vacuum();
        Q = cantilever.maxQ;
        return;
    end

    % Inner function for solving the transcendental eqn to find
    % omega_damped. We're searching for a function minimum, so return
    % the residual squared (continuous and smooth)
2390 function residual_squared = find_natural_frequency(omega_damped)
    hydro = self.hydrodynamic_function(omega_damped, rho_f, eta_f);
    residual = omega_damped - omega_vacuum*(1 + pi * rho_f * ...
        self.w/(4 * self.rho_si * self.t) .* real(hydro)).^-0.5;
    residual_squared = residual^2;
end

% Lookup fluid properties once, then calculate omega_damped and Q
[rho_f eta_f] = self.lookupFluidProperties();
omega_vacuum = self.omega_vacuum();

2400 options = optimset('TolX', 1e-12, 'TolFun', 1e-12, 'Display', 'off');
omega_damped = fminbnd(@find_natural_frequency, 10, omega_vacuum, options);
```

```

hydro = self.hydrodynamic_function(omega_damped, rho_f, eta_f);
Q = (4*self.rho_si*self.t/(pi*rho_f*self.w)+real(hydro))/imag(hydro);

% Catch cases where Q is undefined, too large or too small
if Q < cantilever.minQ || isnan(Q)
    Q = cantilever.minQ;
2410 elseif Q > cantilever.maxQ
    Q = cantilever.maxQ;
end
end

function [omega_damped_hz, Q] = omega_damped_hz_and_Q(self)
[omega_damped, Q] = self.omega_damped_and_Q();
omega_damped_hz = omega_damped/(2*pi);
end

2420 % Calculate the Reynold's number - note that 'a = w/2' in the Brumley paper
function reynolds = reynolds(self, omega, rho_f, eta_f)
reynolds = (rho_f*omega*(self.w/2)^2)/eta_f;
end

% Calculate hydrodynamic function from the lookup table
function hydro = hydrodynamic_function(self, omega, rho_f, eta_f)

% A = aspect ratio, Beta = Reynold's number
A = self.t/self.w;
2430 Beta = self.reynolds(omega, rho_f, eta_f);
log_Beta = log10(Beta);

% If the Reynolds number gets too small, the calculation gets upset
% This only happens during random solution generation usually
if log_Beta < min(self.Beta_lookup)
    log_Beta = min(self.Beta_lookup);
end

gamma_real = interp2(cantilever.A_lookup, cantilever.Beta_lookup, ...
2440 cantilever.gamma_lookup_real, A, log_Beta, 'spline');

```

```
gamma_imag = interp2(cantilever.A_lookup, cantilever.Beta_lookup, ...
    cantilever.gamma_lookup_imag, A, log_Beta, 'spline');
hydro = complex(gamma_real, gamma_imag);
end

% ===== Optimization =====

% Calculate force resolution
% Units: pN
2450 function force_resolution = optimize_force_resolution(self, x0)
    self = self.cantilever_from_state(x0);
    force_resolution = self.force_resolution()*1e12;
end

% Calculate displacement resolution
% Units: nm
function displacement_resolution = optimize_displacement_resolution(self, x0)
    self = self.cantilever_from_state(x0);
    displacement_resolution = self.displacement_resolution()*1e9;
2460 end

% Calculate the force noise density on resonance
% Units: pN/rtHz
function force_noise_density = optimize_resonant_force_noise_density(self, x0)
    self = self.cantilever_from_state(x0);
    force_noise_density = self.resonant_force_noise_density()*1e12;
end

% Calculate the surface stress resolution
2470 % Units: Pa
function ss_resolution = optimize_surface_stress_resolution(self, x0)
    self = self.cantilever_from_state(x0);
    ss_resolution = self.surface_stress_resolution()*1e6;
end

% Used by optimization to bring all state variables to 0(10)
% If we didn't do that, we'd have 0(1e-9) and 0(1e20) variables
```

```

function scaling = optimization_scaling(self)
    l_scale = 1e5;
    w_scale = 1e7;
    t_scale = 1e8;
    l_pr_ratio_scale = 1e2;
    v_bridge_scale = 1e1;
    scaling = [l_scale, w_scale, t_scale, l_pr_ratio_scale, ...
               v_bridge_scale, self.doping_optimization_scaling()];

    % Actuator specific code
    switch self.cantilever_type
        case 'step'
            % Do nothing special
        case 'thermal'
            l_a_scale = 1e6;
            w_a_scale = 1e6;
            t_a_scale = 1e9;
            v_actuator_scale = 1;
            R_heater_scale = 1e-3;
            scaling = [scaling l_a_scale w_a_scale t_a_scale ...
                       v_actuator_scale R_heater_scale];
        case 'piezoelectric'
            l_a_scale = 1e6;
            w_a_scale = 1e6;
            t_a_scale = 1e9;
            v_actuator_scale = 1;
            scaling = [scaling l_a_scale w_a_scale t_a_scale v_actuator_scale];
        end
    end

    % Update the changed optimization parameters
    % All optimization takes place for the same object (i.e. we update 'self')
    % so that things like 'fluid' are maintained
function self = cantilever_from_state(self, x0)
    scaling = self.optimization_scaling();
    x0 = x0./scaling;

```

```

        self.l = x0(1);
        self.w = x0(2);
        self.t = x0(3);
        self.l_pr_ratio = x0(4);
        self.v_bridge = x0(5);
2520      self = self.doping_cantilever_from_state(x0);

        % Actuator specific code
        switch self.cantilever_type
            case 'step'
                % Do nothing special
            case 'thermal'
                self.l_a = x0(8);
                self.w_a = x0(9);
                self.t_a = x0(10);
2530            self.v_actuator = x0(11);
                self.R_heater = x0(12);
            case 'piezoelectric'
                self.l_a = x0(8);
                self.w_a = x0(9);
                self.t_a = x0(10);
                self.v_actuator = x0(11);
            end
        end

2540      % Return state vector for the current state
        function x = current_state(self)
            x(1) = self.l;
            x(2) = self.w;
            x(3) = self.t;
            x(4) = self.l_pr_ratio;
            x(5) = self.v_bridge;
            x = [x self.doping_current_state()];
        end

        % Actuator specific code
2550      switch self.cantilever_type
            case 'step'

```

```

% Do nothing special
case 'thermal'
    x(8) = self.l_a;
    x(9) = self.w_a;
    x(10) = self.t_a;
    x(11) = self.v_actuator;
    x(12) = self.R_heater;
case 'piezoelectric'
2560    x(8) = self.l_a;
    x(9) = self.w_a;
    x(10) = self.t_a;
    x(11) = self.v_actuator;
end
end

% Set the minimum and maximum bounds for the cantilever state variables.
% Bounds are written in terms of the initialization variables.
% Secondary constraints (e.g. power dissipation, resonant frequency)
2570 % are applied in optimization_constraints()
function [lb ub] = optimization_bounds(self, parameter_constraints)
    min_l = 10e-6;
    max_l = 3e-3;

    min_w = 2e-6;
    max_w = 100e-6;

    min_t = 1e-6;
    max_t = 100e-6;
2580    min_l_pr_ratio = 0.01;
    max_l_pr_ratio = 0.99;

    min_v_bridge = 0.1;
    max_v_bridge = 10;

    [doping_lb doping_ub] = ...
        self.doping_optimization_bounds(parameter_constraints);

```

```

2590    % Actuator specific code
    actuator_lb = [];
    actuator_ub = [];
    switch self.cantilever_type
        case {'step', 'none'}
            % Override the default values if any were provided
            % constraints = {{'min_l', 'max_v_bridge'}, {5e-6, 10}}
            if ~isempty(parameter_constraints)
                keys = parameter_constraints{1};
                values = parameter_constraints{2};
                for ii = 1:length(keys)
                    eval([keys{ii} '=' num2str(values{ii}) ';']);
                end
            end
        case 'thermal'
            min_l_a = 5e-6;
            max_l_a = 200e-6;

            min_w_a = 2e-6;
            max_w_a = 50e-6;

2600            min_t_a = 200e-9;
            max_t_a = 3e-6;

            min_v_actuator = .1;
            max_v_actuator = 10;

            min_R_heater = 200;
            max_R_heater = 5e3;

2620            % Override the default values if any were provided
            % constraints = {{'min_l', 'max_v_bridge'}, {5e-6, 10}}
            if ~isempty(parameter_constraints)
                keys = parameter_constraints{1};
                values = parameter_constraints{2};
                for ii = 1:length(keys)

```

```

    eval([keys{ii} '=' num2str(values{ii}) ';' ]);
end
end

2630 actuator_lb = [min_l_a min_w_a min_t_a min_v_actuator min_R_heater];
actuator_ub = [max_l_a max_w_a max_t_a max_v_actuator max_R_heater];
case 'piezoelectric'
min_l_a = 5e-6;
max_l_a = 200e-6;

min_w_a = 2e-6;
max_w_a = 30e-6;

min_t_a = 200e-9;
2640 max_t_a = 3e-6;

min_v_actuator = .1;
max_v_actuator = 10;

% Override the default values if any were provided
% constraints is a set of key value pairs, e.g.
% constraints = {{'min_l', 'max_v_bridge'}, {5e-6, 10}}
if ~isempty(parameter_constraints)
keys = parameter_constraints{1};
values = parameter_constraints{2};
for ii = 1:length(keys)
    eval([keys{ii} '=' num2str(values{ii}) ';' ]);
end
end

2650 actuator_lb = [min_l_a min_w_a min_t_a min_v_actuator];
actuator_ub = [max_l_a max_w_a max_t_a max_v_actuator];
end

lb = [min_l, min_w, min_t, min_l_pr_ratio, ...
       min_v_bridge, doping_lb, actuator_lb];
ub = [max_l, max_w, max_t, max_l_pr_ratio, ...

```

```

    max_v_bridge, doping_ub, actuator_ub];
end

% Generate a random cantilever design (x0) within our param bounds
function x0 = initial_conditions_random(self, parameter_constraints)
[lb, ub] = self.optimization_bounds(parameter_constraints);

2670    l_min = lb(1);
    l_max = ub(1);
    w_min = lb(2);
    w_max = ub(2);
    t_min = lb(3);
    t_max = ub(3);
    l_pr_ratio_min = lb(4);
    l_pr_ratio_max = ub(4);
    V_b_min = lb(5);
    V_b_max = ub(5);

2680    % Generate the random values
    l_random = l_min + rand*(l_max - l_min);
    w_random = w_min + rand*(w_max - w_min);
    t_random = t_min + rand*(t_max - t_min);
    l_pr_ratio_random = l_pr_ratio_min + rand*(l_pr_ratio_max - l_pr_ratio_min);
    v_bridge_random = V_b_min + rand*(V_b_max - V_b_min);

    x0_doping = self.doping_initial_conditions_random(parameter_constraints);

2690    % Actuator specific code
    x0_actuator = [];
    switch self.cantilever_type
        case 'thermal'
            l_a_min = lb(8);
            l_a_max = ub(8);
            w_a_min = lb(9);
            w_a_max = ub(9);
            t_a_min = lb(10);
            t_a_max = ub(10);

```

```

2700    v_actuator_min = lb(11);
    v_actuator_max = ub(11);
    R_heater_min = lb(12);
    R_heater_max = ub(12);

    l_a_random = l_a_min + rand*(l_a_max - l_a_min);
    w_a_random = w_a_min + rand*(w_a_max - w_a_min);
    t_a_random = t_a_min + rand*(t_a_max - t_a_min);
    v_actuator_random = v_actuator_min + ...
        rand*(v_actuator_max - v_actuator_min);
2710    R_heater_random = R_heater_min + rand*(R_heater_max - R_heater_min);
    x0_actuator = [l_a_random w_a_random t_a_random ...
        v_actuator_random R_heater_random];

case 'piezoelectric'
    l_a_min = lb(8);
    l_a_max = ub(8);
    w_a_min = lb(9);
    w_a_max = ub(9);
    t_a_min = lb(10);
    t_a_max = ub(10);
2720    v_actuator_min = lb(11);
    v_actuator_max = ub(11);

    l_a_random = l_a_min + rand*(l_a_max - l_a_min);
    w_a_random = w_a_min + rand*(w_a_max - w_a_min);
    t_a_random = t_a_min + rand*(t_a_max - t_a_min);
    v_actuator_random = v_actuator_min + ...
        rand*(v_actuator_max - v_actuator_min);
    x0_actuator = [l_a_random w_a_random t_a_random v_actuator_random];
end
2730
x0 = [l_random, w_random, t_random, l_pr_ratio_random, ...
    v_bridge_random, x0_doping, x0_actuator];
end

% Nonlinear optimization constraints.
% For a feasible design, all constraints are negative.

```

```

function [C, Ceq] = optimization_constraints(self, x0, nonlinear_constraints)
c_new = self.cantilever_from_state(x0);

2740    % Default aspect ratios that can be overriden
min_w_t_ratio = 2;
min_l_w_ratio = 2;
min_pr_l_w_ratio = 2;
min_pr_l = 2e-6;

% Read out the constraints as key-value pairs
% , e.g. {{'omega_min_hz', 'min_k'}, {1000, 10}}
if ~isempty(nonlinear_constraints)
    keys = nonlinear_constraints{1};
    values = nonlinear_constraints{2};
    for ii = 1:length(keys)
        eval([keys{ii} '=' num2str(values{ii}) ';']);
    end
end

% We start with this single element vector and then append
% any additional constraints that the user has provided.
% If a C(ii) > 0 then constraint ii is invalid

2750    % Force resolution must always be positive
    % Use a linear function so that it's smooth/continuous
resolution = c_new.force_resolution();
C(1) = -resolution*1e18;

% Resonant frequency
if exist('omega_min_hz', 'var')
    switch self.fluid
        case 'vacuum'
            freq_constraint = omega_min_hz - c_new.omega_vacuum_hz();
2770        otherwise
            [omega_damped_hz, tmp] = c_new.omega_damped_hz_and_Q();
            freq_constraint = omega_min_hz - omega_damped_hz;
        end

```

```

C = [C freq_constraint];
end

% Power dissipation
if exist('max_power', 'var')
    power_constraint = c_new.power_dissipation() - max_power;
2780    C = [C power_constraint];
end

% Temp constraints -- approximate lumped model
if exist('tip_temp', 'var') || exist('max_temp', 'var')
    [TMax, TTip] = c_new.approxTempRise();
end
if exist('tip_temp', 'var')
    temp_constraint = TTip - tip_temp;
    C = [C temp_constraint];
end
if exist('max_temp', 'var')
    temp_constraint = TMax - max_temp;
    C = [C temp_constraint];
end

% Temp constraints -- 1D finite differences
% Slower but good for refining designs
if exist('tip_temp_exact', 'var') || exist('max_temp_exact', 'var')
    [TMax, TTip] = c_new.calculateMaxAndTipTemp();
2800 end
if exist('tip_temp_exact', 'var')
    temp_constraint = TTip - tip_temp_exact;
    C = [C temp_constraint];
end
if exist('max_temp_exact', 'var')
    temp_constraint = TMax - max_temp_exact;
    C = [C temp_constraint];
end

% Min and maximum cantilever stiffness
2810

```

```

if exist('min_k', 'var')
    min_k_constraint = min_k - c_new.stiffness();
    C = [C min_k_constraint];
end

if exist('max_k', 'var')
    max_k_constraint = c_new.stiffness() - max_k;
    C = [C max_k_constraint];
end

2820
if exist('max_v_actuator', 'var')
    max_v_actuator_constraint = c_new.v_actuator - max_v_actuator;
    C = [C max_v_actuator_constraint];
end

if exist('min_tip_deflection', 'var')
    min_tip_deflection_constraint = min_tip_deflection - ...
        c_new.tipDeflection();
    C = [C min_tip_deflection_constraint];
end

2830
% Aspect ratio constraints. Default ratios can be changed.
length_width_ratio = min_l_w_ratio - c_new.l/c_new.w;
C = [C length_width_ratio];

width_thickness_ratio = min_w_t_ratio - c_new.w/c_new.t;
C = [C width_thickness_ratio];

pr_length_width_ratio = min_pr_l_w_ratio - c_new.l_pr()/c_new.w_pr();
C = [C pr_length_width_ratio];

pr_length_constraint = (min_pr_l - c_new.l_pr())*1e6;
C = [C pr_length_constraint];

2840
% Now for equality based constraints
Ceq = [];
```

```

% Fix the stiffness
2850 if exist('fixed_k', 'var')
    fixed_k_constraint = c_new.stiffness() - fixed_k;
    Ceq = [Ceq fixed_k_constraint];
end

if exist('fixed_v_bridge', 'var')
    fixed_v_bridge_constraint = c_new.v_bridge - fixed_v_bridge;
    Ceq = [Ceq fixed_v_bridge_constraint];
end

2860 % Fix the resonant frequency
if exist('fixed_f0', 'var')
    switch self.fluid
        case 'vacuum'
            fixed_f0_constraint = fixed_f0 - c_new.omega_vacuum_hz();
        otherwise
            [omega_damped_hz, tmp] = c_new.omega_damped_hz_and_Q();
            fixed_f0_constraint = fixed_f0 - omega_damped_hz;
        end
    Ceq = [Ceq fixed_f0_constraint];
end

2870

% Useful lines for debugging constraint failures
% (e.g. you made l_max too small for it to hit the desired f_0)
% C
% Ceq
% sprintf('Active Index: %d -- Value: %g\n', find(C==max(C)), max(C))
end

% The optimization isn't convex so isn't guaranteed to converge.
2880 % On average it converges 95-99% of the time.
% For this reason, we generally optimize from random initial guesses
% and keep going until several have converged to the same value
function optimized_cantilever = optimize_performance(self, ...
    parameter_constraints, nonlinear_constraints, goal)

```

```
percent_match = 0.01; % 1 percent match required between results
randomize_starting_conditions = 1;
converged = 0;
ii = 1;
2890 resolution = [];
while ~converged
    % Optimize another cantilever
    [c{ii}, exitflag] = ...
        self.optimize_performance_once(parameter_constraints, ...
            nonlinear_constraints, goal, randomize_starting_conditions);

    % If the optimization terminated abnormally, skip to the next iteration
    if ~(exitflag == 1 || exitflag == 2)
        continue
    end

    % Record the resolution for the latest cantilever
    if goal == cantilever.goalForceResolution
        resolution(ii) = c{ii}.force_resolution();
    elseif goal == cantilever.goalDisplacementResolution
        resolution(ii) = c{ii}.displacement_resolution();
    elseif goal == cantilever.goalForceNoiseDensity
        resolution(ii) = c{ii}.resonant_force_noise_density();
    elseif goal == cantilever.goalSurfaceStress
        resolution(ii) = c{ii}.surface_stress_resolution();
    end

    % If we have more than one result, consider stopping
    if ii > 1

        % Sort from smallest to largest, check if the two smallest values agree
        [temp_resolution, sortIndex] = sort(resolution);
        fprintf('Resolutions so far: %s\n', mat2str(temp_resolution, 3))
        resultsAgree = abs(1 - temp_resolution(1)/temp_resolution(2)) ...
2910             < percent_match;
```

```
% If the results agree, then stop the loop. Otherwise, continue
if resultsAgree
    fprintf('CONVERGED. Two best values: %s\n', ...
        mat2str(temp_resolution(1:2), 3))
    optimized_cantilever = c{sortIndex(1)};
    converged = 1;
else
    fprintf('NOT CONVERGED. Two best values: %s\n', ...
        mat2str(temp_resolution(1:2), 3))
end
end

if ii > cantilever.numOptimizationIterations
    [temp_resolution, sortIndex] = sort(resolution);
    optimized_cantilever = c{sortIndex(1)};
    converged = 1;
end

2940 ii = ii + 1; % Increment the loop counter
end
end

% Optimize, but don't randomize starting point
function optimized_cantilever = optimize_performance_from_current(self, ...
    parameter_constraints, nonlinear_constraints, goal)

    randomize_starting_conditions = 0;
    [optimized_cantilever, tmp] = ...
2950    self.optimize_performance_once(parameter_constraints, ...
        nonlinear_constraints, goal, randomize_starting_conditions);
end

function [optimized_cantilever, exitflag] = optimize_performance_once(self, ...
    parameter_constraints, nonlinear_constraints, ...
    goal, randomize_starting_conditions)

    scaling = self.optimization_scaling();
```

```
self.check_valid_cantilever();

2960
    % If random_flag = 1, start from random conditions
    % Else, start from the current cantilever state vector
    if randomize_starting_conditions == 1
        problem.x0 = ...
            scaling.*self.initial_conditions_random(parameter_constraints);
    else
        problem.x0 = scaling.*self.current_state();
    end

2970
    if goal == cantilever.goalForceResolution
        problem.objective = @self.optimize_force_resolution;
    elseif goal == cantilever.goalDisplacementResolution
        problem.objective = @self.optimize_displacement_resolution;
    elseif goal == cantilever.goalForceNoiseDensity
        problem.objective = @self.optimize_resonant_force_noise_density;
    elseif goal == cantilever.goalSurfaceStress
        problem.objective = @self.optimize_surface_stress_resolution;
    end

2980
    [lb ub] = self.optimization_bounds(parameter_constraints);
    problem.lb = scaling.*lb;
    problem.ub = scaling.*ub;

% problem.x0

    problem.options.TolFun = 1e-9;
    problem.options.TolCon = 1e-9;
    problem.options.TolX = 1e-9;

    problem.options.MaxFunEvals = 2000;
    problem.options.MaxIter = 2000;
    problem.options.Display = 'iter-detailed'; % iter, final, none, off
    problem.options.UseParallel = 'always'; % For multicore processors
    problem.options.TypicalX = ones(1, length(scaling));
    problem.options.InitBarrierParam = 10;
```

```
% interior-point does not do well if the design is already
% close to the optimal state (i.e. warm start). sqp is
% supposed to do better in this condition. try experimenting
3000    % with both, particular for implanted devices.

problem.options.Algorithm = 'sqp'; % sqp, interior-point
problem.solver = 'fmincon';

problem.nonlcon = @(x) self.optimization_constraints(x, ...
    nonlinear_constraints);

% These errors come up occasionally for ion implanted devices
warning off MATLAB:singularMatrix
warning off MATLAB:nearlySingularMatrix
3010

[x, tmp, exitflag] = fmincon(problem);
optimized_cantilever = self.cantilever_from_state(x);
end
end
end
```

H.2.3 cantileverDiffusion.m

```
% Model a diffused cantilever (particularly phosphorus/POCl3 diffusion)
classdef cantileverDiffusion < cantilever
    properties
        diffusion_time
        diffusion_temp
    end

    methods
        % Call superclass constructor
        function self = cantileverDiffusion(freq_min, freq_max, l, w, t, ...
10           l_pr_ratio, v_bridge, doping_type, diffusion_time, diffusion_temp)

            self = self@cantilever(freq_min, freq_max, l, w, t, l_pr_ratio, ...
                v_bridge, doping_type);

            self.diffusion_time = diffusion_time;
            self.diffusion_temp = diffusion_temp;
        end

        function print_performance(self)
            print_performance@cantilever(self); % print the base class stuff
            fprintf('Diffusion time (mins), temp (C): %f %f \n', ...
                self.diffusion_time/60, self.diffusion_temp-273);
            fprintf('Junction depth (nm): %f \n', self.junction_depth()*1e9);
            fprintf('=====\\n\\n')
        end

        function print_performance_for_excel(self, varargin)
30            % Call the superclass method first
            print_performance_for_excel@cantilever(self, varargin);

            % varargin gets another set of {} from the cantilever subclasses
            varargin = size(varargin, 2);
            if varargin == 1
```

```

        fid = varargin{1};
elseif optargin == 0
    fid = 1; % Print to the stdout
else
    fprintf('ERROR: Extra optional arguments')
end

% Then print out our additional variables
variables_to_print = [self.diffusion_time/60, self.diffusion_temp-273];
for print_index = 1:length(variables_to_print)
    fprintf(fid, '%4g\t', variables_to_print(print_index));
end
fprintf(fid, '\n');
end

50
% Calculate the diffusion profile for a constant surface source
% diffusion using self.diffusion_time and self.diffusion_temp as
% as well as self.doping_type
function [x, active_doping, total_doping] = doping_profile(self)
N_background = 1e15; % N/cm^3
N_surface = 1e20; % N/cm^3
n_points = self.numZPoints; % # of points of doping profile

switch self.doping_type
60
    case 'arsenic'
        D_0 = 9.17; % cm^2/sec
        E_a = 3.99; % eV

        % Simple diffusion model
        D = D_0*exp(-E_a/(self.k_b_eV*self.diffusion_temp));
        diffusion_length = sqrt(D*self.diffusion_time)*1e-2; % cm -> m

        junction_depth = 2*diffusion_length*erfcinv(N_background/N_surface);
        x = linspace(0, self.t, n_points);

70
        active_doping = N_surface*erfc(x/(2*diffusion_length));
        total_doping = active_doping;

```

```

case 'boron'
D_0 = 1.0;
E_a = 3.5;

% Simple diffusion model
D = D_0*exp(-E_a/(self.k_b_eV*self.diffusion_temp));
80   diffusion_length = sqrt(D*self.diffusion_time)*1e-2; % cm -> m

junction_depth = 2*diffusion_length*erfcinv(N_background/N_surface);
x = linspace(0, self.t, n_points);

active_doping = N_surface*erfc(x/(2*diffusion_length));
total_doping = active_doping;

% A model for phosphorus diffusion which accounts for the
% kink. Takes the solid solubility to be 2.1e20/cc, which
% is accurate for ~850C diffusion temp.
90   case 'phosphorus'
k_b_eV = 8.617343e-5;
x = linspace(0, self.t*1e2, n_points); % m -> cm

% Current (7/18/10)
if self.diffusion_temp < 780+273
    time_offset = 0;
else
    time_offset = 2;
end
100

% Alpha activation energy varies quadratically with temp
T_values = 273+[775 800 850 950];
Ea_values = [1.755 1.73 1.71 1.66];
p = polyfit(T_values, Ea_values, 2);
Ea_alpha = polyval(p, self.diffusion_temp);
alpha = .18*exp(-Ea_alpha./(k_b_eV*self.diffusion_temp));

Da = 100*exp(-3.77/k_b_eV/self.diffusion_temp);

```

```

110      Db = 2.3*exp(-1.95./(k_b_eV*self.diffusion_temp))*1e-5;
      Cb = 3*exp(-0.88./(k_b_eV*self.diffusion_temp))*1e23;

      % Add some offset time to account for the oxidation/purge steps
      t = self.diffusion_time + time_offset*60;

      % Temperature dependent electrically active concentration
      % Source: Solmi and Nobili papers (various)
      surface_concentration_total = 2.5e23* ...
          exp(-0.62./(k_b_eV*self.diffusion_temp));
      surface_concentration_active = 1.1e22* ...
          exp(-0.37./(k_b_eV*self.diffusion_temp));

      % Calculate the profile
      x0 = alpha*t;
      kappa = Cb/surface_concentration_total;

      F1 = erfc((x+alpha*t)/(2*sqrt(Da*t))) + ...
          erfc((x-3*alpha*t)/(2*sqrt(Da*t)));
      F2 = erfc((x+alpha*t)/(2*sqrt(Db*t))) + ...
          erfc((x-3*alpha*t)/(2*sqrt(Db*t)));

      130    Ca = (1-kappa)/2*surface_concentration_total* ...
          exp(-alpha/(2*Da)*(x-alpha*t)).*F1;
      Cb = kappa/2*surface_concentration_total* ...
          exp(-alpha/(2*Db)*(x-alpha*t)).*F2;

      C_total = Ca + Cb;
      C_total(find(C_total < 1e15)) = 1e15;
      C_total(x <= x0) = surface_concentration_total;
      total_doping = C_total;

      140    C_active = C_total;
      C_active(C_active >= surface_concentration_active) = ...
          surface_concentration_active;
      active_doping = C_active;

```

```

x = x*1e-2; % cm -> m
end
end
150
% Calculate sheet resistance.
% Units: ohms
function Rs = sheet_resistance(self)
    [x, active_doping, total_doping] = self.doping_profile();
    % Units: x -> m, doping -> N/cm^3
    conductivity = self.conductivity(active_doping); % ohm-cm
    Rs = 1/trapz(x*1e2, conductivity); % convert x to cm pre-integration
end

160 % Effective carrier density, N/m^2
function Nz = Nz(self)
    [z, N_active, N_total] = self.doping_profile();
    [mu, sigma] = self.mobility(N_active, self.T);
    % Nz_total = trapz(z, N_active*1e6) % doping: N/cm^3 -> N/m^3
    Nz = 1e4*trapz(z*1e2, N_active.*mu)^2/trapz(z*1e2, N_active.*mu.^2);
end

function plot_doping_profile(self)
    [x, active_doping, total_doping] = self.doping_profile();
170
    figure
    hold all
    plot(x, total_doping)
    plot(x, active_doping)
    hold off
    set(gca, 'yscale', 'log');
    xlim([0 self.t])
end

function x_j = junction_depth(self)
    [x, active_doping, total_doping] = self.doping_profile();
    x_j = x(find(active_doping == 1e15, 1));
end
180

```

```

function alpha = alpha(self)
    alpha = self.default_alpha; % use the alpha from the superclass
end

% =====
% ===== Optimization =====
% =====

190

function scaling = doping_optimization_scaling(self)
    diffusion_time_scale = 1e-3;
    diffusion_temp_scale = 1e-3;
    scaling = [diffusion_time_scale diffusion_temp_scale];
end

function self = doping_cantilever_from_state(self, x0)
    self.diffusion_time = x0(6);
    self.diffusion_temp = x0(7);
end

function x = doping_current_state(self)
    x(1) = self.diffusion_time;
    x(2) = self.diffusion_temp;
end

% Set the minimum and maximum bounds for the cantilever state
% variables. Bounds are written in terms of the initialization
210
% variables. Secondary constraints (e.g. power dissipation,
% piezoresistor thickness rather than ratio, resonant frequency)
% are applied in optimization_constraints()

function [lb ub] = doping_optimization_bounds(self, parameter_constraints)
    min_diffusion_time = 5*60; % seconds
    max_diffusion_time = 90*60;

    min_diffusion_temp = 273+800; % K
    max_diffusion_temp = 273+1000;

    % Override the default values if any were provided

```

```

% constraints is a set of key value pairs, e.g.
% constraints = {{'min_l', 'max_v_bridge'}, {5e-6, 10}}
if ~isempty(parameter_constraints)
    keys = parameter_constraints{1};
    values = parameter_constraints{2};
    for ii = 1:length(keys)
        eval([keys{ii} '=' num2str(values{ii}) ';' ]);
    end
end

230

lb = [min_diffusion_time, min_diffusion_temp];
ub = [max_diffusion_time, max_diffusion_temp];
end

function x0 = doping_initial_conditions_random(self, parameter_constraints)
    [lb, ub] = self.doping_optimization_bounds(parameter_constraints);

    diffusion_time_min = lb(1);
    diffusion_time_max = ub(1);

240

    diffusion_temp_min = lb(2);
    diffusion_temp_max = ub(2);

    % Generate the random values
    diffusion_time_random = diffusion_time_min + ...
        rand*(diffusion_time_max - diffusion_time_min);
    diffusion_temp_random = diffusion_temp_min + ...
        rand*(diffusion_temp_max - diffusion_temp_min);

250
    x0 = [diffusion_time_random, diffusion_temp_random];
end
end
end

```

H.2.4 cantileverEpitaxy.m

```
% Model an epitaxial cantilever (B, P, As well supported)
% Assumes negligible dopant diffusion
classdef cantileverEpitaxy < cantilever
    properties
        dopant_concentration
        t_pr_ratio
    end

    methods
        function self = cantileverEpitaxy(freq_min, freq_max, l, w, t, ...
10         l_pr_ratio, v_bridge, doping_type, dopant_concentration, t_pr_ratio)

            self = self@cantilever(freq_min, freq_max, l, w, t, ...
                l_pr_ratio, v_bridge, doping_type);
            self.dopant_concentration = dopant_concentration;
            self.t_pr_ratio = t_pr_ratio;
        end

        function print_performance(self)
20            print_performance@cantilever(self); % print the base stuff
            fprintf('Dopant concentration (per cc): %g \n', self.dopant_concentration);
            fprintf('PR Thickness Ratio: %g \n', self.t_pr_ratio);
            fprintf('PR Thickness: %g \n', self.junction_depth());
            fprintf('\n')
        end

        function x_j = junction_depth(self)
            x_j = self.t * self.t_pr_ratio;
        end
    end
30
    function print_performance_for_excel(self, varargin)
        % Call the superclass method first
        print_performance_for_excel@cantilever(self, varargin);

        % varargin gets another set of {} from the cantilever subclasses
```

```

optargin = size(varargin, 2);
if optargin == 1
    fid = varargin{1};
elseif optargin == 0
    fid = 1; % Print to the stdout
else
    fprintf('ERROR: Extra optional arguments')
end

% Then print out our additional variables
variables_to_print = [self.dopant_concentration, self.t_pr_ratio];
for print_index = 1:length(variables_to_print)
    fprintf(fid, '%4g\t', variables_to_print(print_index));
end
50    fprintf(fid, '\n');
end

% Return the electrically active dopant concentration profile
% Units: cm^-3
function [z, active_doping, total_doping] = doping_profile(self)
n_points = self.numZPoints; % # of points of doping profile
z = linspace(0, self.t, n_points);

% Initialize to the background concentration
60    background_concentration = 1e15;
active_doping = ones(1, n_points)*background_concentration;

% Fill in the epitaxial region, assume active = total
active_doping(z<=self.junction_depth()) = self.dopant_concentration;
total_doping = active_doping;
end

% Calculate sheet resistance
% Units: ohms
70    function Rs = sheet_resistance(self)
conductivity = self.conductivity(self.dopant_concentration); % ohm-cm
Rs = 1/(self.junction_depth()*1e2*conductivity); % t_j -> cm

```

```

    end

    % We assume a constant concentration so can just integrated to get N_z
    % Units: carries/m^2
    function Nz = Nz(self)
        Nz = self.junction_depth()*self.dopant_concentration*1e6;
    end

80
    function alpha = alpha(self)
        alpha = self.default_alpha; % use the alpha from the superclass
    end

    % ===== Optimization =====
    function scaling = doping_optimization_scaling(self)
        concentration_scale = 1e-19;
        t_pr_ratio_scale = 10;
        scaling = [concentration_scale t_pr_ratio_scale];
    end

90
    function self = doping_cantilever_from_state(self, x0)
        self.dopant_concentration = x0(6);
        self.t_pr_ratio = x0(7);
    end

    function x = doping_current_state(self)
        x(1) = self.dopant_concentration;
        x(2) = self.t_pr_ratio;
    end

100
    function [lb ub] = doping_optimization_bounds(self, parameter_constraints)

        min_dopant_concentration = 1e17;

        % Approximate solid solubilities at 1000C
        switch self.doping_type
            case 'boron'
                max_dopant_concentration = 2e20;

```

```

110      case 'phosphorus'
        max_dopant_concentration = 4e20;
    case 'arsenic'
        max_dopant_concentration = 8e20;
    end

    min_t_pr_ratio = 0.01;
    max_t_pr_ratio = 0.99;

    % Override the default values if any were provided
120    % constraints is a set of key value pairs, e.g.
    % constraints = {{'min_l', 'max_v_bridge'}, {5e-6, 10}}
    if ~isempty(parameter_constraints)
        keys = parameter_constraints{1};
        values = parameter_constraints{2};
        for ii = 1:length(keys)
            eval([keys{ii} '=' num2str(values{ii}) ';']);
        end
    end

130    lb = [min_dopant_concentration min_t_pr_ratio];
    ub = [max_dopant_concentration max_t_pr_ratio];
end

function x0 = doping_initial_conditions_random(self, parameter_constraints)
    [lb, ub] = self.doping_optimization_bounds(parameter_constraints);

    n_min = lb(1);
    n_max = ub(1);

140    t_pr_ratio_min = lb(2);
    t_pr_ratio_max = ub(2);

    % Generate the random values
    dopant_concentration_random = 10^(log10(n_min) + ...
        rand*(log10(n_max) - log10(n_min))); % logarithmically distributed
    t_pr_ratio_random = t_pr_ratio_min + rand*(t_pr_ratio_max - t_pr_ratio_min);

```

```
x0 = [dopant_concentration_random, t_pr_ratio_random];  
end  
150 end  
end
```

H.2.5 cantileverImplantation.m

```
% Model an ion implanted cantilever using lookup tables from TSUPREM
% B, P and As are supported with either inert or oxidizing anneal
% environments.
%
% For all conditions, a 250A protection oxide layer is grown
% before the ion implantation.
%
% After the implantation step...
%
% 'oxide' -> strip protection oxide, regrow 1500A oxide, inert anneal
10 % 'inert' -> leave protection oxide, inert anneal, strip oxide
%
% Lookup tables are linearly interpolated. Splines can result
% in negative values (e.g. sheet resistance). Gradient/Hessian
% discontinuities don't seem to pose a problem for optimization.
%
% The range of the simulation conditions (noted below) is equal to
% the limits of the highest quality data in TSUPREM
classdef cantileverImplantation < cantilever
    properties
        implantation_energy; % 20 - 80 keV
        implantation_dose; % 2e14 - 2e16 per sq cm
        annealing_temp; % 900 - 1100C
        annealing_time; % 15 - 120min
        annealing_type; % 'inert' or 'oxide'
        lookupTableData
    end

    methods
        % Call superclass constructor
30        function self = cantileverImplantation(freq_min, freq_max, ...
            l, w, t, l_pr_ratio, v_bridge, doping_type, ...
            annealing_time, annealing_temp, annealing_type, ...
            implantation_energy, implantation_dose)

            self = self@cantilever(freq_min, freq_max, l, w, t, l_pr_ratio, ...

```

```

v_bridge, doping_type);

self.implantation_energy = implantation_energy;
self.implantation_dose = implantation_dose;
40 self.annealing_type = annealing_type;
self.annealing_temp = annealing_temp;
self.annealing_time = annealing_time;

% Load the lookup table
try
    data = open('lookupTable.mat');
    self.lookupTableData = data;
catch
    fprintf('Dopant type: %s\n', self.doping_type);
    50 fprintf('Anneal type: %s\n', self.annealing_type);
    fprintf('Not available\n');
    return;
end
end

function print_performance(self)
print_performance@cantilever(self); % print the base stuff
fprintf('Implantation energy (keV), dose (cm-2): %.1f %1g \n', ...
60 self.implantation_energy, self.implantation_dose);
fprintf('Annealing time (mins), temp (C): %.1f %.1f \n', ...
    self.annealing_time/60, self.annealing_temp-273);
fprintf('Diffusion length (cm): %1g \n', self.diffusion_length());
fprintf('Junction depth (um): %.3f \n', self.junction_depth()*1e6);
fprintf('Alpha: %3g \n', self.alpha());
fprintf('\n')
end

function print_performance_for_excel(self, varargin)
70 % Call the superclass method first
print_performance_for_excel@cantilever(self, varargin);

```

```

% varargin gets another set of {} from the cantilever subclasses
optargin = size(varargin, 2);
if optargin == 1
    fid = varargin{1};
elseif optargin == 0
    fid = 1; % Print to the stdout
else
    fprintf('ERROR: Extra optional arguments')
end

% Then print out our additional variables
variables_to_print = [self.diffusion_time/60, self.diffusion_temp-273];
for print_index = 1:length(variables_to_print)
    fprintf(fid, '%4g\t', variables_to_print(print_index));
end
fprintf(fid, '\n');
end

90
=====
% ====== Doping Profile ======
% ====== =====
function annealNumber = annealNumber(self)
switch self.annealing_type
    case 'inert'
        annealNumber = 1;
    case 'oxide'
        annealNumber = 2;
    otherwise
        fprintf('Unknown anneal condition: %s\n', self.annealing_type);
        pause
    end
end

100
=====

% Lookup the concentration profile from the lookup table
function [x, active_doping, total_doping] = doping_profile(self)
    x = self.lookupTableData.z; % 10nm spacing from 0 to 5um

```

```

110 n = interpn(x, self.lookupTableData.ImplantDopants, ...
               self.lookupTableData.ImplantDoses, ...
               self.lookupTableData.ImplantEnergies, ...
               self.lookupTableData.AnnealTemps, ...
               self.lookupTableData.AnnealTimes, ...
               self.lookupTableData.AnnealOxidation, ...
               self.lookupTableData.n, ...
               x, self.dopantNumber(), self.implantation_dose, ...
               self.implantation_energy, self.annealing_temp-273, ...
               self.annealing_time/60, self.annealNumber(), 'linear');

120 % Remove data beyond the device thickness
% We don't need to remove data beyond the junction because beta/Rs/Nz
% are already just calculated to the junction in the lookup table
n(x > self.t) = [];
x(x > self.t) = [];

% Active = total unless the doping is higher than the solid solubility limit
% which is generally not the case in the ion implantation data
active_doping = n;
total_doping = n;

130
end

% Junction depth. Units = m
function Xj = junction_depth(self)
Xj = interpn(self.lookupTableData.ImplantDopants, ...
             self.lookupTableData.ImplantDoses, ...
             self.lookupTableData.ImplantEnergies, ...
             self.lookupTableData.AnnealTemps, ...
             self.lookupTableData.AnnealTimes, ...
             self.lookupTableData.AnnealOxidation, ...
             self.lookupTableData.Xj, ...
             self.dopantNumber(), self.implantation_dose, ...
             self.implantation_energy, self.annealing_temp-273, ...
             self.annealing_time/60, self.annealNumber(), 'linear');

140
end

```

```
function Rs = sheet_resistance(self)
    Rs = interpn(self.lookupTableData.ImplantDopants, ...
    self.lookupTableData.ImplantDoses, ...
    self.lookupTableData.ImplantEnergies, ...
    self.lookupTableData.AnnealTemps, ...
    self.lookupTableData.AnnealTimes, ...
    self.lookupTableData.AnnealOxidation, ...
    self.lookupTableData.Rs, ...
    self.dopantNumber(), self.implantation_dose, ...
    self.implantation_energy, self.annealing_temp-273, ...
    self.annealing_time/60, self.annealNumber(), 'linear');
end

160 % Nz is an abstract function specifically for this class
% In the other classes (epitaxy, diffusion) we calculate Nz from the profile
% Note that Nz != Nz_total due to current crowding
function Nz = Nz(self)
    Nz = interpn(self.lookupTableData.ImplantDopants, ...
    self.lookupTableData.ImplantDoses, ...
    self.lookupTableData.ImplantEnergies, ...
    self.lookupTableData.AnnealTemps, ...
    self.lookupTableData.AnnealTimes, ...
    self.lookupTableData.AnnealOxidation, ...
    self.lookupTableData.Nz, ...
    self.dopantNumber(), self.implantation_dose, ...
    self.implantation_energy, self.annealing_temp-273, ...
    self.annealing_time/60, self.annealNumber(), 'linear');
end

170
function beta = beta(self)
    Beta1 = interpn(self.lookupTableData.ImplantDopants, ...
    self.lookupTableData.ImplantDoses, ...
    self.lookupTableData.ImplantEnergies, ...
    self.lookupTableData.AnnealTemps, ...
    self.lookupTableData.AnnealTimes, ...
    self.lookupTableData.AnnealOxidation, ...
    self.lookupTableData.Beta1, ...
```

```

    self.dopantNumber(), self.implantation_dose, ...
    self.implantation_energy, self.annealing_temp-273, ...
    self.annealing_time/60, self.annealNumber(), 'linear');

Beta2 = interpn(self.lookupTableData.ImplantDopants, ...
self.lookupTableData.ImplantDoses, ...
190   self.lookupTableData.ImplantEnergies, ...
    self.lookupTableData.AnnealTemps, ...
    self.lookupTableData.AnnealTimes, ...
    self.lookupTableData.AnnealOxidation, ...
    self.lookupTableData.Beta2, ...
    self.dopantNumber(), self.implantation_dose, ...
    self.implantation_energy, self.annealing_temp-273, ...
    self.annealing_time/60, self.annealNumber(), 'linear');

% Beta2 is in units of microns, so convert t
200 beta = Beta1 - 2/(self.t*1e6)*Beta2;
end

% Calculated from the dopant diffusion length, not the Si diffusion length
% Based upon personal communications with L.K.J. Vandamme (Jan 2012)
% Note that most experimental data is from boron resistors
function diffusion_length = diffusion_length(self)
switch self.doping_type
    case 'arsenic'
        D0 = 22.9; %cm2/s
    210   Ea = 4.1; %eV
    case 'boron'
        D0 = 0.76; %cm2/s
        Ea = 3.46; %eV
    case 'phosphorus'
        D0 = 3.85; %cm2/s
        Ea = 3.66; %eV
    end
    diffusivity = D0*exp(-Ea/self.k_b_eV/self.annealing_temp);
    diffusion_length = (diffusivity*self.annealing_time)^0.5; %cm
220 end

```

```
% Calculate alpha from sqrt(Dt) using the data I compiled for the book
function alpha = alpha(self)
    alpha = 2.469e-10*self.diffusion_length()^-0.598;
end

% ===== Optimization =====
function scaling = doping_optimization_scaling(self)
    annealing_time_scale = 1e-2;
230    annealing_temp_scale = 1e-2;
    implantation_energy_scale = 1e0;
    implantation_dose_scale = 1e-13;
    scaling = [annealing_time_scale, annealing_temp_scale, ...
        implantation_energy_scale, implantation_dose_scale];
end

function self = doping_cantilever_from_state(self, x0)
    self.annealing_time = x0(6);
    self.annealing_temp = x0(7);
240    self.implantation_energy = x0(8);
    self.implantation_dose = x0(9);
end

function x = doping_current_state(self)
    x(1) = self.annealing_time;
    x(2) = self.annealing_temp;
    x(3) = self.implantation_energy;
    x(4) = self.implantation_dose;
end

250 % Set the minimum and maximum bounds for the cantilever state
% variables. Bounds are written in terms of the initialization
% variables. Secondary constraints (e.g. power dissipation,
% piezoresistor thickness rather than ratio, resonant frequency)
% are applied in optimization_constraints()
function [lb ub] = doping_optimization_bounds(self, parameter_constraints)
    min_annealing_time = 15*60; % seconds
```

```

max_annealing_time = 120*60;

260    min_annealing_temp = 273+900; % K
        max_annealing_temp = 273+1100;

        min_implantation_energy = 20; %keV
        max_implantation_energy = 80;

        min_implantation_dose = 2e14; %cm-2
        max_implantation_dose = 2e16;

        % Override the default values if any were provided
270    % constraints is a set of key value pairs, e.g.
        % constraints = {{'min_l', 'max_v_bridge'}, {5e-6, 10}}
        if ~isempty(parameter_constraints)
            keys = parameter_constraints{1};
            values = parameter_constraints{2};
            for ii = 1:length(keys)
                eval([keys{ii} '=' num2str(values{ii}) ';' ]);
            end
        end

        lb = [min_annealing_time, min_annealing_temp ...
               min_implantation_energy min_implantation_dose];
        ub = [max_annealing_time, max_annealing_temp ...
               max_implantation_energy max_implantation_dose];
    end

function x0 = doping_initial_conditions_random(self, parameter_constraints)
    [lb, ub] = self.doping_optimization_bounds(parameter_constraints);

    annealing_time_min = lb(1);
290    annealing_time_max = ub(1);

    annealing_temp_min = lb(2);
    annealing_temp_max = ub(2);

```

```
implantation_energy_min = lb(3);
implantation_energy_max = ub(3);

implantation_dose_min = lb(4);
implantation_dose_max = ub(4);

300
% Generate the random values
annealing_time_random = annealing_time_min + ...
    rand*(annealing_time_max - annealing_time_min);
annealing_temp_random = annealing_temp_min + ...
    rand*(annealing_temp_max - annealing_temp_min);
implantation_energy_random = implantation_energy_min + ...
    rand*(implantation_energy_max - implantation_energy_min);
implantation_dose_random = implantation_dose_min + ...
    rand*(implantation_dose_max - implantation_dose_min);

310
x0 = [annealing_time_random, annealing_temp_random, ...
    implantation_energy_random, implantation_dose_random];
end
end
end
```

Bibliography

- [1] D. R. and Carter, “Mechanical loading history and skeletal biology,” *Journal of Biomechanics*, vol. 20, no. 11-12, pp. 1095 – 1109, 1987.
- [2] A. S. Adhikari, J. Chai, and A. R. Dunn, “Mechanical load induces a 100-fold increase in the rate of collagen proteolysis by mmp-1,” *Journal of the American Chemical Society*, vol. 133, no. 6, pp. 1686–1689, 2011.
- [3] R. Fay, *Hearing in vertebrates: a psychophysics databook*. Hill-Fay Associates, 1988. [Online]. Available: <http://books.google.com/books?id=DoJwQgAACAAJ>
- [4] C. L. Degen, M. Poggio, H. J. Mamin, C. T. Rettner, and D. Rugar, “Nanoscale magnetic resonance imaging,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 5, pp. 1313–1317, 2009.
- [5] R. Eatock, R. Fay, and A. Popper, *Vertebrate hair cells*, ser. Springer handbook of auditory research. Springer, 2006. [Online]. Available: <http://books.google.com/books?id=Wvdk31NLSxoC>
- [6] R. Fettiplace and C. M. Hackney, “The sensory and motor roles of auditory hair cells,” *Nature Reviews Neuroscience*, vol. 7, pp. 19–29, Jan 2006.
- [7] R. Fettiplace, A. J. Ricci, and C. M. Hackney, “Clues to the cochlear amplifier from the turtle ear,” *Trends in Neurosciences*, vol. 24, pp. 169–175, 2001.
- [8] J. Holt and D. Corey, “Two mechanisms for transducer adaptation in vertebrate hair cells,” *Proceedings of the National Academy of Sciences*, Jan 2000.

- [9] R. Eatock, "Adaptation in hair cells," *Annual Reviews in Neuroscience*, Jan 2000.
- [10] H. J. Kennedy, A. C. Crawford, and R. Fettiplace, "Force generation by mammalian hair bundles supports a role in cochlear amplification," *Nature*, vol. 433, no. 7028, pp. 880–883, 2005.
- [11] H. Kennedy, M. Evans, A. Crawford, and R. Fettiplace, "Fast adaptation of mechanoelectrical transducer channels in mammalian cochlear hair cells," *Nature Neuroscience*, vol. 6, no. 8, pp. 832–836, 2003.
- [12] D. Corey and A. Hudspeth, "Kinetics of the receptor current in bullfrog saccular hair cells," *Journal of Neuroscience*, Jan 1983.
- [13] A. C. Crawford, M. G. Evans, and R. Fettiplace, "Activation and adaptation of transducer currents in turtle hair cells." *The Journal of Physiology*, vol. 419, no. 1, pp. 405–434, 1989.
- [14] N. Nise, *Control Systems Engineering*. John Wiley & Sons, 2010.
- [15] K. Oshima, K. Shin, M. Diensthuber, A. W. Peng, A. J. Ricci, and S. Heller, "Mechanosensitive hair cell-like cells from embryonic and induced pluripotent stem cells," *Cell*, vol. 141, no. 4, pp. 704 – 716, 2010.
- [16] J. H. Hoh, J. P. Cleveland, C. B. Prater, J. P. Revel, and P. K. Hansma, "Quantized adhesion detected with the atomic force microscope," *Journal of the American Chemical Society*, vol. 114, no. 12, pp. 4917–4918, 1992. [Online]. Available: <http://pubs.acs.org/doi/abs/10.1021/ja00038a075>
- [17] H. M. D Rugar, R. Budakian and B. Chui, "Single spin detection by magnetic resonance force microscopy," *Nature*, vol. 430, pp. 329–332, 2004.
- [18] K. Svoboda, C. F. Schmidt, B. J. Schnapp, and S. M. Block, "Direct observation of kinesin stepping by optical trapping interferometry," *Nature*, vol. 365, pp. 721–727, Oct. 1993.

- [19] C. Bustamante, J. C. Macosko, and G. J. Wuite, “Grabbing the cat by the tail: manipulating molecules one by one,” *Nature Reviews Molecular Cellular Biology*, vol. 1, pp. 130–136, 2000.
- [20] K. C. Neuman and A. Nagy, “Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy,” *Nat Meth*, vol. 5, no. 6, pp. 491–505, Jun 2008.
- [21] N. Malagnino, G. Pesce, A. Sasso, and E. Arimondo, “Measurements of trapping efficiency and stiffness in optical tweezers,” *Optics Communications*, vol. 214, no. 1-6, pp. 15 – 24, 2002. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0030401802021193>
- [22] J. Harley, “Advances in piezoresistive probes for atomic force microscopy,” Ph.D. dissertation, Stanford University, 2000.
- [23] D. R. Brumley, M. Willcox, and J. E. Sader, “Oscillation of cylinders of rectangular cross section immersed in fluid,” *Physics of Fluids*, vol. 22, no. 5, p. 052001, 2010.
- [24] A. A. Barlian, W.-T. Park, J. R. Mallon, A. J. Rastegar, and B. L. Pruitt, “Review: Semiconductor piezoresistance for microsystems,” *Proceedings of the IEEE*, vol. 97, no. 3, pp. 513–552, 2009.
- [25] S. Tadigadapa and K. Mateti, “Piezoelectric MEMS sensors: state-of-the-art and perspectives,” *Measurement Science and Technology*, vol. 20, no. 9, p. 092001, 2009.
- [26] D. J. Bell, T. J. Lu, N. A. Fleck, and S. M. Spearing, “MEMS actuators and sensors: observations on their performance and selection for purpose,” *Journal of Micromechanics and Microengineering*, vol. 15, no. 7, pp. S153–S164, 2005.
- [27] S. Trolier-McKinstry and P. Muralt, “Thin film piezoelectrics for MEMS,” *Journal of Electroceramics*, vol. 12, pp. 7–17, 2004.
- [28] T. Kenny, “Nanometer-scale force sensing with MEMS devices,” *Sensors Journal*, Jan 2001, testing 123.

- [29] G. Binnig, C. Quate, and C. Gerber, "Atomic force microscope," *Physical Review Letters*, vol. 56, no. 9, pp. 930–933, Mar 1986.
- [30] G. BINNING, H. ROHRER, C. GERBER, and E. WEIBEL, "Surface studies by scanning tunneling microscopy," *Physical Review Letters*, vol. 49, no. 1, pp. 57–61, 1982.
- [31] O. Solgaard, *Photonic microsystems: micro and nanotechnology applied to optical devices and systems*, ser. MEMS reference shelf. Springer, 2009. [Online]. Available: <http://books.google.com/books?id=7-ez-ZccV4QC>
- [32] T. Gabrielson, "Mechanical-thermal noise in micromachined acoustic and vibration sensors," *IEEE Transactions on Electron Devices*, vol. 40, no. 5, pp. 903–909, Jan 1993.
- [33] O. Chaudhuri, S. Parekh, W. Lam, and D. Fletcher, "Combined atomic force microscopy and side-view optical imaging for mechanical studies of cells," *Nature Methods*, vol. 6, pp. 383–387, 2009.
- [34] E. Canetta, A. Duperray, A. Leyrat, and C. Verdier, "Measuring cell viscoelastic properties using a force-spectrometer: Influence of protein-cytoplasm interactions," *Biorheology*, vol. 42, pp. 321–333, 2005.
- [35] M. Tortonese, R. Barrett, and C. Quate, "Atomic resolution with an atomic force microscope using piezoresistive detection," *Applied Physics Letters*, vol. 62, no. 8, pp. 834–836, Dec 1993.
- [36] J. Harley and T. Kenny, "High-sensitivity piezoresistive cantilevers under 1000 angstroms thick," *Applied Physics Letters*, vol. 75, no. 2, pp. 289–291, Jan 1999.
- [37] J. Tansock and C. Williams, "Force measurement with a piezoelectric cantilever in a scanning force microscope," *Ultramicroscopy*, vol. 42-44, no. Part 2, pp. 1464–1469, 1992.
- [38] T. Itoh and T. Suga, "Piezoelectric force sensor for scanning force microscopy," *Sensors and Actuators A: Physical*, vol. 43, no. 1-3, pp. 305–310, 1994.

- [39] N. Blanc, J. Brugger, N. F. de Rooij, and U. Durig, "Scanning force microscopy in the dynamic mode using microfabricated capacitive sensors," *J. Vac. Sci. Technol. B*, vol. 14, no. 2, pp. 901–905, 1996.
- [40] S. A. Miller, K. L. Turner, and N. C. MacDonald, "Microelectromechanical scanning probe instruments for array architectures," *Review of Scientific Instruments*, vol. 68, no. 11, pp. 4155–4162, 1997.
- [41] F. Beyeler, S. Muntwyler, and B. Nelson, "A six-axis MEMS force-torque sensor with micro-newton and nano-newtonmeter resolution," *Microelectromechanical Systems, Journal of*, vol. 18, no. 2, pp. 433 –441, april 2009.
- [42] S. Senturia, *Microsystem design*, 1st ed. New York: Springer, 2000.
- [43] J. C. Doll and B. L. Pruitt, "Design of piezoresistive versus piezoelectric contact mode scanning probes," *Journal of Micromechanics and Microengineering*, vol. 20, no. 9, p. 095023, 2010.
- [44] O. Brand and G. Fedder, *CMOS-MEMS*, ser. Advanced Micro & Nanosystems. Wiley-VCH, 2005, no. v. 2.
- [45] J. Adams, L. Manning, B. Rogers, M. Jones, and S. Minne, "Self-sensing tapping mode atomic force microscopy," *Sensors and Actuators A: Physical*, vol. 121, no. 1, pp. 262–266, 2005.
- [46] J. Adams, B. Rogers, L. Manning, Z. Hu, T. Thundat, H. Cavazos, and S. Minne, "Piezoelectric self-sensing of adsorption-induced microcantilever bending," *Sensors and Actuators A: Physical*, vol. 121, no. 2, pp. 457–461, 2005.
- [47] V. Mukundan and B. Pruitt, "Mems electrostatic actuation in conducting biological media," *Microelectromechanical Systems, Journal of*, vol. 18, no. 2, pp. 405 –413, april 2009.
- [48] V. Mukundan, P. Ponce, H. E. Butterfield, and B. L. Pruitt, "Modeling and characterization of electrostatic comb-drive actuators in conducting liquid media,"

- Journal of Micromechanics and Microengineering*, vol. 19, no. 6, p. 065008, 2009. [Online]. Available: <http://stacks.iop.org/0960-1317/19/i=6/a=065008>
- [49] J. Judy and N. Myung, “Magnetic materials for mems,” *Materials Research Society Symposium Proceedings*, pp. 23–26, 2002.
- [50] B. Lee, C. B. Prater, and W. P. King, “Lorentz force actuation of a heated atomic force microscope cantilever,” *Nanotechnology*, vol. 23, no. 5, p. 055709, 2012. [Online]. Available: <http://stacks.iop.org/0957-4484/23/i=5/a=055709>
- [51] C. Haber and D. Wirtz, “Magnetic tweezers for dna micromanipulation,” *Review of Scientific Instruments*, vol. 71, no. 12, pp. 4561–4570, 2000. [Online]. Available: <http://link.aip.org/link/?RSI/71/4561/1>
- [52] D. Lange, C. Hagleitner, C. Herzog, O. Brand, and H. Baltes, “Electromagnetic actuation and mos-transistor sensing for cmos-integrated micromechanical resonators,” *Sensors and Actuators A: Physical*, vol. 103, no. 1-2, pp. 150 – 155, 2003. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0924424702003072>
- [53] C. Vancura, M. Regg, Y. Li, C. Hagleitner, and A. Hierlemann, “Magnetically actuated complementary metal oxide semiconductor resonant cantilever gas sensor systems,” *Analytical Chemistry*, vol. 77, no. 9, pp. 2690–2699, 2005.
- [54] P. Kollmannsberger and B. Fabry, “High-force magnetic tweezers with force feedback for biological applications,” *Review of Scientific Instruments*, vol. 78, no. 11, p. 114301, 2007. [Online]. Available: <http://link.aip.org/link/?RSI/78/114301/1>
- [55] G. E. Fantner, W. Schumann, R. J. Barbero, A. Deutschinger, V. Todorov, D. S. Gray, A. M. Belcher, I. W. Rangelow, and K. Youcef-Toumi, “Use of self-actuating and self-sensing cantilevers for imaging biological samples in fluid,” *Nanotechnology*, vol. 20, no. 43, p. 434003, 2009. [Online]. Available: <http://stacks.iop.org/0957-4484/20/i=43/a=434003>

- [56] Y. Sarov, T. Ivanov, A. Frank, and I. Rangelow, “Thermally driven multi-layer actuator for 2d cantilever arrays,” *Applied Physics A: Materials Science and Processing*, vol. 102, pp. 61–68, 2011, 10.1007/s00339-010-6078-1. [Online]. Available: <http://dx.doi.org/10.1007/s00339-010-6078-1>
- [57] S. Manalis, S. Minne, and C. Quate, “Atomic force microscopy for high speed imaging using cantilevers with an integrated actuator and sensor,” *Applied Physics Letters*, vol. 68, no. 6, pp. 871–873, 1996.
- [58] S. Minne, S. Manalis, and C. Quate, “Parallel atomic force microscopy using cantilevers with integrated piezoresistive sensors and integrated piezoelectric actuators,” *Applied Physics Letters*, vol. 67, no. 26, pp. 3918–3920, Dec 1995.
- [59] S. C. Minne, S. R. Manalis, A. Atalar, and C. F. Quate, “Independent parallel lithography using the atomic force microscope,” *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 14, no. 4, pp. 2456–2461, July 1996.
- [60] S. C. Minne, G. Yaralioglu, S. R. Manalis, J. D. Adams, J. Zesch, A. Atalar, and C. F. Quate, “Automated parallel high-speed atomic force microscopy,” *Applied Physics Letters*, vol. 72, no. 18, pp. 2340–2342, May 1998.
- [61] S. Minne, J. Adams, G. Yaralioglu, S. Manalis, A. Atalar, and C. Quate, “Centimeter scale atomic force microscope imaging and lithography,” *Applied Physics Letters*, vol. 73, no. 12, pp. 1742–1744, Dec 1998.
- [62] S. R. Manalis, S. C. Minne, A. Atalar, and C. F. Quate, “Interdigital cantilevers for atomic force microscopy,” *Applied Physics Letters*, vol. 69, no. 25, pp. 3944–3946, Dec. 1996.
- [63] T. Sulcek, R. J. Grow, G. G. Yaralioglu, S. C. Minne, C. F. Quate, S. R. Manalis, A. Kiraz, A. Aydine, and A. Atalar, “Parallel atomic force microscopy with optical interferometric detection,” *Applied Physics Letters*, vol. 78, no. 12, pp. 1787–1789, Mar. 2001.

- [64] T. Sulcuk, R. Hsieh, J. D. Adams, S. C. Minne, C. F. Quate, and D. M. Adderton, "High-speed atomic force microscopy in liquid," *Review of Scientific Instruments*, vol. 71, no. 5, pp. 2097–2099, May 2000.
- [65] Y.-S. Kim, H.-J. Nam, S.-M. Cho, J.-W. Hong, D.-C. Kim, and J. U. Bu, "PZT cantilever array integrated with piezoresistor sensor for high speed parallel operation of AFM," *Sensors and Actuators A: Physical*, vol. 103, no. 1-2, pp. 122 – 129, 2003. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0924424702003114>
- [66] C. Shin, I. Jeon, Z. G. Khim, J. W. Hong, and H. Nam, "Study of sensitivity and noise in the piezoelectric self-sensing and self-actuating cantilever with an integrated wheatstone bridge circuit," *Review of Scientific Instruments*, vol. 81, no. 3, p. 035109, 2010. [Online]. Available: <http://link.aip.org/link/?RSI/81/035109/1>
- [67] T. Ivanov, T. Gotszalk, P. Grabiec, E. Tomerov, and I. Rangelow, "Thermally driven micromechanical beam with piezoresistive deflection readout," *Microelectronic Engineering*, vol. 67-68, no. 0, pp. 550 – 556, 2003, *[ce:title] Proceedings of the 28th International Conference on Micro- and Nano-Engineering [ce:title]*. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167931703001138>
- [68] R. Pedrak, T. Ivanov, K. Ivanova, T. Gotszalk, N. Abedinov, I. W. Rangelow, K. Edinger, E. Tomerov, T. Schenkel, and P. Hudek, "Micromachined atomic force microscopy sensor with integrated piezoresistive sensor and thermal bimorph actuator for high-speed tapping-mode atomic force microscopy phase-imaging in higher eigenmodes," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 21, no. 6, pp. 3102–3107, 2003. [Online]. Available: <http://link.aip.org/link/?JVB/21/3102/1>
- [69] R. Linnemann, T. Gotszalk, I. W. Rangelow, P. Dumania, and E. Oesterschulze, "Atomic force microscopy and lateral force microscopy using piezoresistive cantilevers," vol. 14, no. 2. AVS, 1996, pp. 856–860. [Online]. Available: <http://link.aip.org/link/?JVB/14/856/1>

- [70] S. Hafizovic, D. Barrettino, T. Volden, J. Sedivy, K. Kirstein, O. Brand, and A. Hierlemann, “Single-chip mechatronic microsystem for surface imaging and force response studies,” *PNAS*, vol. 101, no. 49, pp. 17 011–17 015, 2004.
- [71] S. Hafizovic, K.-U. Kirstein, and A. Hierlemann, “Integrated cantilevers and atomic force microscopes,” in *Applied Scanning Probe Methods V*, ser. NanoScience and Technology, B. Bhushan, S. Kawata, and H. Fuchs, Eds. Springer Berlin Heidelberg, 2007, pp. 1–22.
- [72] T. Ando, N. Kodera, E. Takai, D. Maruyama, K. Saito, and A. Toda, “A high-speed atomic force microscope for studying biological macromolecules,” *Proceedings of the National Academy of Sciences*, vol. 98, no. 22, pp. 12 468–12 472, 2001. [Online]. Available: <http://www.pnas.org/content/98/22/12468.abstract>
- [73] B. Chui, T. Stowe, T. Kenny, H. Mamin, and B. Terris, “Low-stiffness silicon cantilevers for thermal writing and piezoresistive readback with the atomic ɻdots,” *Applied Physics Letters*, Dec 1996.
- [74] R. Ried, H. Mamin, B. Terris, L. Fan, and D. Rugar, “6 mhz, 2 n/m piezoresistive atomic-force-microscope cantilevers with incisive tips,” *Journal Of Microelectromechanical Systems*, vol. 6, no. 12, pp. 294–302, 1997.
- [75] J. Harley and T. Kenny, “1/f noise considerations for the design and process optimization of piezoresistive cantilevers,” *Journal Of Microelectromechanical Systems*, vol. 9, no. 2, pp. 226–235, Jan 2000.
- [76] W. BRANTLEY, “Calculated elastic-constants for stress problems associated with semiconductor devices,” *Journal Of Applied Physics*, vol. 44, no. 1, pp. 534–535, 1973.
- [77] D. Franca and A. Blouin, “All-optical measurement of in-plane and out-of-plane young’s modulus and poisson’s ratio in silicon wafers by means of vibration modes,” *Meas Sci Technol*, vol. 15, no. 5, pp. 859–868, 2004.

- [78] I. S. Sokolnikoff, *Mathematical Theory of Elasticity*, 1st ed. Malabar, FL: R.E. Krieger Pub., 1983.
- [79] H. Rolnick, “Tension coefficient of resistance of metals,” *Physical Review*, vol. 36, no. 3, pp. 506–512, 1930.
- [80] C. Kittel, *Introduction to Solid State Physics*, 8th ed. New York: John Wiley & Sons, 2005.
- [81] J. F. Nye, *Physical Properties of Crystals: Their Representation by Tensors and Matrices*, 1st ed. Oxford: Oxford University Press, 1985.
- [82] W. C. Young and R. Budynas, *Roark's Formulas for Stress and Strain*, 7th ed. New York: McGraw-Hill, 2002.
- [83] K. PETERSEN, “Silicon as a mechanical material,” *Proceedings of the IEEE*, vol. 70, no. 5, pp. 420–457, 1982.
- [84] J. J. Wortman and R. A. Evans, “Young’s modulus, shear modulus, and poisson’s ratio in silicon and germanium,” *Journal of Applied Physics*, vol. 36, no. 1, pp. 153–156, 1965.
- [85] C. SMITH, “Piezoresistance effect in germanium and silicon,” *Physical Review*, vol. 94, no. 1, pp. 42–49, 1954.
- [86] P. Bridgman, “The effect of homogeneous mechanical stress on the electrical resistance of crystals,” *Physical Review*, vol. 42, no. 6, pp. 858–863, 1932.
- [87] W. MASON and R. THURSTON, “Use of piezoresistive materials in the measurement of displacement, force, and torque,” *J Acoust Soc Am*, vol. 29, no. 10, pp. 1096–1101, 1957.
- [88] Y. Kanda, “A graphical representation of the piezoresistance coefficients in silicon,” *Electron Devices*, vol. 29, no. 1, pp. 64–70, Jan 1982.

- [89] S. Suthram, J. C. Ziegert, T. Nishida, and S. E. Thompson, “Piezoresistance coefficients of (100) silicon nmosfets measured at low and high (similar to 1.5 gpa) channel stress,” *Ieee Electr Device L*, vol. 28, no. 1, pp. 58–61, 2007.
- [90] T. A. Rost, “Transistor design and layout for performance improvement with strain,” Patent US 0 095 988, 4 16, 2009.
- [91] E. Lund and T. Finstad, “Measurement of the temperature dependency of the piezoresistance coefficients in p-type silicon,” in *Proceedings of ASME Advances in Electronic Packaging*, Hawaii, USA, June 1999, pp. 215–218.
- [92] E. Lund and T. Finstad, “Design and construction of a four-point bending based set-up for measurement of piezoresistance in semiconductors,” *Review of Scientific Instruments*, vol. 75, no. 11, pp. 4960–4966, 2004.
- [93] C.-H. Cho, R. Jaeger, and J. Suhling, “Characterization of the temperature dependence of the piezoresistive coefficients of silicon from c to c,” *Sensors Journal, IEEE*, vol. 8, no. 8, pp. 1455 –1468, aug. 2008.
- [94] J. Richter, A. Hyldgard, K. Birkelund, M. B. Arnoldus, O. Hansen, and E. V. Thomsen, “CMOS-integrated silicon 3d force sensor system for micro component coordinate measurement machines,” in *Proceedings of IEEE 20th International Conference on Micro Electro Mechanical Systems (MEMS2007)*, Kobe, Japan, January 2007, pp. 69–72.
- [95] J. Richter, A. Hyldgard, K. Birkelund, O. Hansen, and E. V. Thomsen, “Realtime 3d stress measurement in curing epoxy packaging,” in *Proceedings of Transducers 2007: International Conference on Solid-State Sensors, Actuators and Microsystems*, Lyon, France, June 2007, pp. 2083–2086.
- [96] W. PFANN and R. THURSTON, “Semiconducting stress transducers utilizing transverse and shear piezoresistance effects,” *Journal of Applied Physics*, vol. 32, no. 10, pp. 2008–2019, 1961.

- [97] O. N. Tufte and E. L. Stelzer, “Piezoresistive properties of silicon diffused layers,” *Journal of Applied Physics*, vol. 34, no. 2, pp. 313–318, Jan 1963.
- [98] A. D. Kurtz and C. Gravel, “Semiconductor transducers using transverse and shear piezoresistance,” in *Proceedings of 22nd Annual ISA Conference*, Chicago, IL, USA, September 1967, pp. P4–PHYMMID–67 1–20.
- [99] S. Sze and K. Ng, *Physics of semiconductor devices*, ser. Wiley-Interscience publication. Wiley-Interscience, 2007.
- [100] M. A. Green, “Intrinsic concentration, effective densities of states, and effective mass in silicon,” *Journal of Applied Physics*, vol. 67, no. 6, pp. 2944–2954, 1990.
- [101] N. G. Nilsson, “An accurate approximation of the generalized einstein relation for degenerate semiconductors,” *physica status solidi (a)*, vol. 19, no. 1, pp. K75–K78, 1973.
- [102] J. Blakemore, “Approximations for fermi-dirac integrals, especially the function used to describe electron density in a semiconductor,” *Solid-State Electronics*, vol. 25, no. 11, pp. 1067 – 1076, 1982.
- [103] O. TUFTE and E. STELZER, “Piezoresistive properties of heavily doped n-ttpe silicon,” *Phys Rev A-Gen Phys*, vol. 133, no. 6A, pp. 1705–1716, 1964.
- [104] W. P. Mason, J. J. Forst, and L. M. Tornillo, “Recent developments in semiconductor strain transducers,” in *Proceedings of 15th Annual Conference of The Instrument Society of America*, New York, USA, October 1962, pp. 110–120.
- [105] D. KERR and A. MILNES, “Piezoresistance of diffused layers in cubic semiconductors,” *Journal of Applied Physics*, vol. 34, no. 4, pp. 727–731, 1963.
- [106] J. Richter, J. Pedersen, M. Brandbyge, E. Thomsen, and O. Hansen, “Piezoresistance in p-type silicon revisited,” *Journal of Applied Physics*, vol. 104, no. 2, p. 023715, July 2008.

- [107] F. MORIN, T. GEBALLE, and C. HERRING, “Temperature dependence of the piezoresistance of high-purity silicon and germanium,” *Physical Review*, vol. 105, no. 2, pp. 525–539, 1957.
- [108] J. Chen and N. MacDonald, “Measuring the nonlinearity of silicon piezoresistance by tensile loading of a submicron diameter fiber using a microinstrument,” *Review of Scientific Instruments*, vol. 75, no. 1, pp. 276–278, 2004.
- [109] K. MATSUDA, Y. Kanda, K. YAMAMURA, and K. Suzuki, “Nonlinearity of piezoresistance effect in p-type and n-type silicon,” *Sensors And Actuators A-Physical*, vol. 21, no. 1-3, pp. 45–48, 1990.
- [110] K. MATSUDA, K. Suzuki, K. YAMAMURA, and Y. Kanda, “Nonlinear piezoresistance effects in silicon,” *Journal of Applied Physics*, vol. 73, no. 4, pp. 1838–1847, 1993.
- [111] R. F. Pierret, *Semiconductor Fundamentals: Modular Series on Solid State Devices*, 1st ed. Reading, MA: Addison Wesley, 1983.
- [112] R. F. Pierret, *Semiconductor Device Fundamentals*, 1st ed. Reading, MA: Addison Wesley, 1996.
- [113] Y. Sun, S. E. Thompson, and T. Nishida, “Physics of strain effects in semiconductors and metal-oxide-semiconductor field-effect transistors,” *Journal of Applied Physics*, vol. 101, no. 10, p. 104503, 2007.
- [114] G. Kimball, “The electronic structure of diamond,” *J Chem Phys*, vol. 3, no. 9, pp. 560–564, 1935.
- [115] J. Mullaney, “Optical properties and electronic structure of solid silicon,” *Physical Review*, vol. 66, no. 11/12, pp. 326–339, 1944.
- [116] J. BARDEEN and W. SHOCKLEY, “Deformation potentials and mobilities in non-polar crystals,” *Physical Review*, vol. 80, no. 1, pp. 72–80, 1950.

- [117] C. HERRING, "Transport properties of a many-valley semiconductor," *Bell Syst Tech J*, vol. 34, no. 2, pp. 237–290, 1955.
- [118] C. HERRING and E. VOGT, "Transport and deformation-potential theory for many-valley semiconductors with anisotropic scattering," *Physical Review*, vol. 101, no. 3, pp. 944–961, 1956.
- [119] D. LONG, "Stress dependence of piezoresistance effect," *Journal of Applied Physics*, vol. 32, no. 10, pp. 2050–2051, 1961.
- [120] C. HERRING, "Theory of the thermoelectric power of semiconductors," *Physical Review*, vol. 96, no. 5, pp. 1163–1187, 1954.
- [121] R. W. Keyes, *The Effects of Elastic Deformation on the Electrical Conductivity of Semiconductors*, 1st ed. New York: Academic Press, 1960, vol. 11.
- [122] M. BECKER and H. FAN, "Optical properties of semiconductors .3. infra-red transmission of silicon," *Physical Review*, vol. 76, no. 10, pp. 1531–1532, 1949.
- [123] H. FAN, "Temperature dependence of the energy gap in monatomic semiconductors," *Physical Review*, vol. 78, no. 6, pp. 808–809, 1950.
- [124] B. ABELES and S. MEIBOOM, "Theory of the galvanomagnetic effects in germanium," *Physical Review*, vol. 95, no. 1, pp. 31–37, 1954.
- [125] M. SHIBUYA, "Magneto resistance effect in cubic semiconductors with spheroidal energy surfaces," *Physical Review*, vol. 95, no. 6, pp. 1385–1393, 1954.
- [126] S. MEIBOOM and B. ABELES, "Theory of the galvanomagnetic effects in n-germanium," *Physical Review*, vol. 93, no. 5, pp. 1121–1121, 1954.
- [127] R. Keyes, "Explaining strain," *Ieee Circuits Device*, vol. 18, no. 5, pp. 36–39, 2002.
- [128] M. LIN, "A better understanding of the channel mobility of si mosfets based on the physics of quantized subbands," *Ieee T Electron Dev*, vol. 35, no. 12, pp. 2406–2411, 1988.

- [129] K. Rim, K. Chan, L. Shi, D. Boyd, J. Ott, N. Klymko, F. Cardone, L. Tai, S. Koester, M. Cobb, D. Canaperi, B. To, E. Duch, I. Babich, R. Carruthers, P. Saunders, G. Walker, Y. Zhang, M. Steen, and M. Ieong, “Fabrication and mobility characteristics of ultra-thin strained Si directly on insulator (SSDOI) MOSFETs,” in *Proceedings of IEEE International Electron Devices Meeting (IEDM)*, Washington, DC, USA, December 2003, pp. 3.1.1–3.1.4.
- [130] S. Thompson, G. Sun, K. Wu, J. Lim, and T. Nishida, “Key differences for process-induced uniaxial vs. substrate-induced biaxial stressed Si and Ge channel MOSFETs,” in *Proceedings of IEEE International Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, December 2004, pp. 221–224.
- [131] E. Wang, P. Matagne, L. Shifren, B. Obradovic, R. Kotlyar, S. Cea, J. He, Z. Ma, R. Nagisetty, S. Tyagi, M. Stettler, and M. D. Giles, “Quantum mechanical calculation of hole mobility in silicon inversion layers under arbitrary stress,” in *Proceedings of IEEE International Electron Devices Meeting (IEDM)*, San Francisco, CA, USA, December 2004, pp. 147–150.
- [132] Y. OHMURA, “Role of the valence-band density-of-states in the piezoresistance of p-type semiconductors si and ge,” *J Phys Soc Jpn*, vol. 62, no. 10, pp. 3615–3620, 1993.
- [133] K. Matsuda, “Strain-dependent hole masses and piezoresistive properties of silicon,” in *IWCE-10 10th International Workshop on Computational Electronics*, Purdue University, Indiana, USA, October 2004, pp. 173–174.
- [134] M. Fischetti, Z. Ren, P. Solomon, M. Yang, and K. Rim, “Six-band k center dot p calculation of the hole mobility in silicon inversion layers: Dependence on surface orientation, strain, and silicon thickness,” *Journal Of Applied Physics*, vol. 94, no. 2, pp. 1079–1095, 2003.
- [135] R. Oberhuber, G. Zandler, and P. Vogl, “Subband structure and mobility of two-dimensional holes in strained si/sige mosfet’s,” *Phys Rev B*, vol. 58, no. 15, pp. 9941–9948, 1998.

- [136] T. Toriyama and S. Sugiyama, "Analysis of piezoresistance in p-type silicon for mechanical sensors," *Journal Of Microelectromechanical Systems*, vol. 11, no. 5, pp. 598–604, 2002.
- [137] G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors*, 1st ed. New York: John Wiley & Sons, 1974.
- [138] W. KLEINER and L. ROTH, "Deformation potential in germanium from optical absorption lines for exciton formation," *Physical Review Letters*, vol. 2, no. 8, pp. 334–336, 1959.
- [139] G. PIKUS and G. BIR, "Cyclotron and paramagnetic resonance in strained crystals," *Physical Review Letters*, vol. 6, no. 1, pp. 103–105, 1961.
- [140] J. HENSEL and G. FEHER, "Valence band parameters in silicon from cyclotron resonances in crystals subjected to uniaxial stress," *Physical Review Letters*, vol. 5, no. 7, pp. 307–309, 1960.
- [141] H. Hasegawa, "Theory of cyclotron resonance in strained silicon crystals," *Physical Review*, vol. 129, no. 3, pp. 1029–1040, 1963.
- [142] J. HENSEL and G. FEHER, "Cyclotron resonance experiments in uniaxially stressed silicon - valence band inverse mass parameters and deformation potentials," *Physical Review*, vol. 129, no. 3, pp. 1041–1062, 1963.
- [143] J. HENSEL, H. Hasegawa, and M. NAKAYAMA, "Cyclotron resonance in uniaxially stressed silicon .2. nature of covalent bond," *Physical Review*, vol. 138, no. 1A, pp. A225–A238, 1965.
- [144] J. HENSEL and K. Suzuki, "Quantum resonances in valence bands of germanium .2. cyclotron resonances in uniaxially stressed crystals," *Phys Rev B*, vol. 9, no. 10, pp. 4219–4257, 1974.
- [145] L. Smith, V. Moroz, G. Eneman, P. Verheyen, F. Nouri, L. Washington, M. Jurczak, O. Penzin, D. Pramanik, and K. D. Meyer, "Exploring the limits of stress-enhanced hole mobility," *Ieee Electr Device L*, vol. 26, no. 9, pp. 652–654, 2005.

- [146] S. E. Thompson, S. Suthram, Y. Sun, G. Sun, S. Parthasarathy, M. Chu, and T. Nishida, "Future of strained Si/semiconductors in nanoscale MOSFETs," in *Proceedings of International Electron Device Meeting*, San Francisco, CA, USA, December 2006, pp. 1–4.
- [147] V. P. Group, "Errors due to wheatstone bridge nonlinearity, tech note tn-507-1," Vishay Inc, Tech. Rep., 2005.
- [148] B. Konrad and M. Ashauer, "Demystifying Piezoresistive Pressure Sensors," *Sensors*, vol. 16, no. 7, pp. 12–25.
- [149] K. Anderson, "The new current loop: An instrumentation and measurement circuit topology," *Instrumentation and Measurement, IEEE Transactions on*, vol. 46, no. 5, pp. 1061–1067, 1997.
- [150] C. Lu, M. Lemkin, and B. Boser, "A monolithic surface micromachined accelerometer with digital output," *Ieee J Solid-St Circ*, vol. 30, no. 12, pp. 1367–1373, 1995.
- [151] T. Muller, M. Brandl, O. Brand, and H. Baltes, "An industrial cmos process family adapted for the fabrication of smart silicon sensors," *Sensors And Actuators A-Physical*, vol. 84, no. 1-2, pp. 126–133, 2000.
- [152] A. Bindra, "Integrated analog signal-conditioning ic handles harsh environments," *Electronic Design*, 2000.
- [153] A. J. Rastegar and J. Bryzek, "A high-performance CMOS processor for piezoresistive sensors," *Sensors*, vol. 14, no. 10, pp. 82–87, 1997.
- [154] S. Park, M. Goodman, and B. Pruitt, "Analysis of nematode mechanics by piezoresistive displacement clamp," *Proc Natl Acad Sci*, vol. 104, pp. 17376–17381, Oct 2007.
- [155] T. ISHIHARA, K. Suzuki, S. SUWAZONO, M. HIRATA, and H. TANIGAWA, "Cmos integrated silicon pressure sensor," *Ieee J Solid-St Circ*, vol. 22, no. 2, pp. 151–156, 1987.

- [156] T. Akiyama, U. Staufer, N. de Rooij, D. Lange, C. Hagleitner, O. Brand, H. Baltes, A. Tonin, and H. Hidber, “Integrated atomic force microscopy array probe with metal-oxide-semiconductor field effect transistor stress sensor, thermal bimorph actuator, and on-chip complementary metal-oxide-semiconductor electronics,” *J Vac Sci Technol B*, vol. 18, no. 6, pp. 2669–2675, 2000.
- [157] D. Lange, C. Hagleitner, O. Brand, and H. Baltes, “CMOS resonant beam gas sensing system with on-chip self excitation,” in *Proceedings of IEEE 14th International Conference on Micro Electro Mechanical Systems (MEMS2001)*, Interlaken, Switzerland, January 2001, pp. 547–552.
- [158] C. Hagleitner, D. Lange, T. Akiyama, A. Tonin, R. Vogt, and H. Baltes, “On-chip circuitry for a CMOS parallel scanning AFM,” in *Proceedings of SPIE’s Smart Structures and Materials 1999: Smart Electronics and MEMS*, Newport Beach, CA, USA, March 1999, pp. 240–248.
- [159] G. Villanueva, F. Perez-Murano, M. Zimmermann, J. Lichtenberg, and J. Bausells, “Piezoresistive cantilevers in a commercial cmos technology for intermolecular force detection,” *Microelectronic Engineering*, vol. 83, no. 4-9, pp. 1302–1305, Apr 2006.
- [160] W. Franks, D. Lange, S. Lee, A. Hierlemann, N. Spencer, and H. Baltes, “Nanochemical surface analyzer in cmos technology,” *Ultramicroscopy*, vol. 91, no. 1-4, pp. 21–27, 2002.
- [161] T. Volden, M. Zimmermann, D. Lange, O. Brand, and H. Baltes, “Dynamics of cmos-based thermally actuated cantilever arrays for force microscopy,” *Sensors And Actuators A-Physical*, vol. 115, no. 2-3, pp. 516–522, 2004.
- [162] J. Lu, T. Ikehara, Y. Zhang, T. Mihara, and T. Itoh, “Idots compatible single crystal silicon cantilever with separated on-chip piezoelectric actuator for ultra Idots,” *Micro Electro Mechanical Systems*, Dec 2008.
- [163] P. Gieschke, J. Richter, J. Joos, P. Ruther, and O. Paul, “Four-degree-of-freedom solid state MEMS joystick,” in *Proceedings of IEEE 21st International Conference*

- on Micro Electro Mechanical Systems (MEMS2008)*, Tucson, AZ, USA, January 2008, pp. 86–89.
- [164] M. Doelle, C. Peters, P. Ruther, and O. Paul, “Piezo-fet stress-sensor arrays for wire-bonding characterization,” *Journal Of Microelectromechanical Systems*, vol. 15, no. 1, pp. 120–130, 2006.
- [165] F. Fruett and G. C. M. Meijer, “The influence of the piezojunction effect at different temperatures on the accuracy of silicon temperature sensors,” in *Proceedings of the 15th International Conference on Microelectronics and Packaging, ICMP2000*, Manaus, Brazil, September 2000, pp. 153–157.
- [166] B. Kane, M. Cutkosky, and G. Kovacs, “A traction stress sensor array for use in high-resolution robotic tactile imaging,” *Journal Of Microelectromechanical Systems*, vol. 9, no. 4, pp. 425–434, 2000.
- [167] R. C. Jaeger, R. Ramani, J. C. Suhling, and Y. Kang, “CMOS stress sensor circuits using piezoresistive field-effect transistors (PIFETs),” in *Digest of Technical Papers 1995 Symposium on VLSI Circuits*, Kyoto, Japan, June 1995, pp. 43–44.
- [168] J. Bartholomeyczik, S. Brugger, P. Ruther, and O. Paul, “Multidimensional cmos in-plane stress sensor,” *Ieee Sens J*, vol. 5, no. 5, pp. 872–882, 2005.
- [169] G. Lin, K. S. J. Pister, and K. P. Roos, “Standard CMOS piezoresistive sensor to quantify heart cell contractile forces,” in *Proceedings of IEEE 9th International Conference on Micro Electro Mechanical Systems (MEMS1996)*, San Diego, CA, USA, February 1996, pp. 150–155.
- [170] B. Levey, P. Gieschke, M. Doelle, S. Spinner, A. Trautmann, P. Ruther, and O. Paul, “CMOS-integrated silicon 3d force sensor system for micro component coordinate measurement machines,” in *Proceedings of IEEE 20th International Conference on Micro Electro Mechanical Systems (MEMS2007)*, Kobe, Japan, January 2007, pp. 611–614.

- [171] G. Villanueva, J. A. Plaza, J. Montserrat, F. Perez-Murano, and J. Bausells, "Crystalline silicon cantilevers for piezoresistive detection of biomolecular forces," *Microelectronic Engineering*, vol. 85, no. 5-6, pp. 1120–1123, Feb 2008.
- [172] H. Baltes, O. Brand, A. Hierlemann, D. A. L. D. Lange, and C. A. H. C. Hagleitner, "CMOS MEMS—present and future," in *Proceedings of IEEE 15th International Conference on Micro Electro Mechanical Systems (MEMS2002)*, Las Vegas, NV, USA, January 2002, pp. 459–466.
- [173] H. Baltes and O. Brand, "Cmos-based microsensors and packaging," *Sensors And Actuators A-Physical*, vol. 92, no. 1-3, pp. 1–9, 2001.
- [174] B. W. Chui, L. Aeschimann, T. Akiyama, U. Staufer, N. F. de Rooij, J. Lee, F. Goericke, W. P. King, and P. Vettiger, "Advanced temperature compensation for piezoresistive sensors based on crystallographic orientation," *Review of Scientific Instruments*, vol. 78, no. 4, p. 043706, Apr. 2007.
- [175] J. Mallon, A. Rastegar, A. Barlian, M. Meyer, T. Fung, and B. Pruitt, "Low 1/f noise, full bridge, microcantilever with longitudinal and transverse piezoresistors," *Applied Physics Letters*, vol. 92, no. 3, p. 033508, Dec 2008.
- [176] P. Hoa, G. Suchaneck, and G. Gerlach, "Influence of polycrystalline silicon as electrical shield on reliability and stability of piezoresistive sensors," *Sensors And Actuators A-Physical*, vol. 120, no. 2, pp. 567–572, 2005.
- [177] P. Gray, P. Hurst, S. Lewis, and R. Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th ed. New York: John Wiley & Sons, 2001.
- [178] J. Johnson, "Thermal agitation of electricity in conductors," *Physical Review*, vol. 32, no. 1, p. 97, 1928.
- [179] H. Nyquist, "Thermal agitation of electric charge in conductors," *Physical Review*, vol. 32, no. 1, pp. 110–113, 1928.
- [180] M. Keshner, "1/f noise," *Proceedings of the IEEE*, vol. 70, no. 3, pp. 212–218, 1982.

- [181] R. VOSS and J. CLARKE, “Flicker (1-f) noise - equilibrium temperature and resistance fluctuations,” *Phys Rev B*, vol. 13, no. 2, pp. 556–573, 1976.
- [182] F. HOOGE and L. Vandamme, “Lattice scattering causes 1-f noise,” *Phys Lett A*, vol. 66, no. 4, pp. 315–316, 1978.
- [183] P. DUTTA and P. HORN, “Low-frequency fluctuations in solids - 1-f noise,” *Rev Mod Phys*, vol. 53, no. 3, pp. 497–516, 1981.
- [184] D. Fleetwood and N. Giordano, “Resistivity dependence of 1/f noise in metal films,” *Physical Review B*, vol. 27, no. 2, pp. 667–671, Dec 1983.
- [185] A. van der Ziel, “Unified presentation of 1/f noise in electron devices: fundamental 1/f noise sources,” *Proceedings of the IEEE*, vol. 76, no. 3, pp. 233–258, 1988.
- [186] L. Vandamme and G. Trefan, “1/f noise in homogeneous and inhomogeneous media,” *Iee P-Circ Dev Syst*, vol. 149, no. 1, pp. 3–12, Jan 2002.
- [187] L. VANDAMME, “Bulk and surface 1/f noise,” *Ieee T Electron Dev*, vol. 36, no. 5, pp. 987–992, 1989.
- [188] A. L. McWhorter, *1/f noise and related surface effects in germanium*, 1st ed. Cambridge, MA: MIT Lincoln Laboratory, 1955.
- [189] A. L. McWhorter, *Semiconductor Surface Physics*, 1st ed. Philadelphia, PA: University of Pennsylvania Press, 1957.
- [190] F. HOOGE, “1/f noise is no surface effect,” *Phys Lett A*, vol. A 29, no. 3, pp. 139–140, 1969.
- [191] F. Hooge, “1/f noise sources,” *Electron Devices*, vol. 41, no. 11, pp. 1926–1935, Jan 1994.
- [192] F. Hooge, T. Kleinpenning, and L. Vandamme, “Experimental studies on 1/f noise,” *Rep Prog Phys*, vol. 44, no. 5, pp. 479–532, 1981.

- [193] T. KLEINPENNING and A. DEKUIJPER, "Relation between variance and sample duration of 1/f noise signals," *Journal of Applied Physics*, vol. 63, no. 1, pp. 43–45, 1988.
- [194] X. Yu, J. Thaysen, O. Hansen, and A. Boisen, "Optimization of sensitivity and noise in piezoresistive cantilevers," *Journal of Applied Physics*, vol. 92, no. 10, pp. 6296–6301, Dec 2002.
- [195] M. Gel and I. Shimoyama, "Force sensing submicrometer thick cantilevers with ultra-thin piezoresistors by rapid thermal diffusion," *Journal of Micromechanics and Microengineering*, vol. 14, no. 3, pp. 423–428, Jan 2004.
- [196] J. C. Doll, B. C. Petzold, P. Ghale, M. B. Goodman, and B. L. Pruitt, "High frequency force sensing with piezoresistive cantilevers," in *Transducers International Conference*, june 2009.
- [197] A. Partridge, J. Reynolds, B. Chui, E. Chow, A. Fitzgerald, L. Zhang, N. Maluf, and T. Kenney, "A high-performance planar piezoresistive accelerometer," *Journal Of Microelectromechanical Systems*, vol. 9, no. 1, pp. 58–66, Mar 2000.
- [198] S.-J. Park, J. C. Doll, and B. L. Pruitt, "Piezoresistive cantilever performance - part I: Analytical model for sensitivity," *Journal Of Microelectromechanical Systems*, vol. 19, no. 1, pp. 149–161, feb. 2010.
- [199] A. A. Barlian, N. Harjee, V. Mukundan, T. H. Fung, S. J. Park, and B. L. Pruitt, "Sidewall epitaxial piezoresistor process for in-plane sensing applications," in *Proceedings of IEEE 21st International Conference on Micro Electro Mechanical Systems (MEMS2008)*, Tucson, AZ, USA, January 2008, pp. 331–334.
- [200] B. Neri, C. Ciofi, and V. Dattilo, "Noise and fluctuations in submicrometric al-si interconnect lines," *Ieee T Electron Dev*, vol. 44, no. 9, pp. 1454–1459, 1997.
- [201] E. Vandamme and L. Vandamme, "Current crowding and its effect on 1/f noise and third harmonic distortion - a case study for quality assessment of resistors," *Microelectron Reliab*, vol. 40, no. 11, pp. 1847–1853, 2000.

- [202] S.-J. Park, J. C. Doll, A. J. Rastegar, and B. L. Pruitt, “Piezoresistive cantilever performance - part II: Optimization,” *Microelectromechanical Systems, Journal of*, vol. 19, no. 1, pp. 149–161, feb. 2010.
- [203] F. N. Hooge, “Discussion of recent experiments on 1/[latin small letter f with hook] noise,” *Physica*, vol. 60, no. 1, pp. 130–144, July 1972.
- [204] L. Vandamme and W. van Bokhoven, “Conductance noise investigations with four arbitrarily shaped and placed electrodes,” *Applied Physics A: Materials Science & Processing*, vol. 14, no. 2, pp. 205–215, Oct. 1977.
- [205] R. H. M. Clevers, “1/f noise in ion-implanted resistors between 77 and 300 k,” *Journal of Applied Physics*, vol. 62, no. 5, pp. 1877–1881, 1987. [Online]. Available: <http://link.aip.org/link/?JAP/62/1877/1>
- [206] L. Vandamme and S. Oosterhoff, “Annealing of ionimplanted resistors reduces the 1/f noise,” *Journal of Applied Physics*, vol. 59, no. 9, pp. 3169–3174, Dec 1986.
- [207] L. Vandamme, “Noise as a diagnostic tool for quality and reliability of electronic devices,” *Ieee T Electron Dev*, vol. 41, no. 11, pp. 2176–2187, 1994.
- [208] M. Jevtic, “Impurity concentration dependence of the 1/f noise parameter [alpha] in silicon,” *Solid-State Electronics*, vol. 31, no. 6, pp. 1049 – 1052, 1988.
- [209] M. Tortonese, “Force sensors for scanning probe microscopy,” Ph.D. dissertation, Stanford University, 1993.
- [210] O. N. Tufte and D. Long, “Recent developments in semiconductor piezoresistive devices,” *Solid-State Electronics*, vol. 6, no. 4, pp. 323 – 338, 1963.
- [211] J. D. Plummer, M. D. Deal, and P. B. Griffin, *Silicon VLSI Technology Fundamentals, Practice, and Modeling*, 1st ed. New Jersey: Prentice Hall, 2000.
- [212] R. Pires, R. Dickstein, S. Titcomb, and R. Anderson, “Carrier freezeout in silicon,” *Cryogenics*, vol. 30, no. 12, pp. 1064 – 1068, 1990.

- [213] A. Volodin, K. Temst, C. V. Haesendonck, and Y. Bruynseraede, "Low temperature magnetic force microscopy with enhanced sensitivity based on piezoresistive detection," *Review of Scientific Instruments*, vol. 71, no. 12, pp. 4468–4473, 2000.
- [214] M. Roseman and P. Grütter, "Cryogenic magnetic force microscope," *Review of Scientific Instruments*, vol. 71, no. 10, pp. 3782–3787, 2000.
- [215] F. Trumbore, "Solid solubilities of impurity elements in germanium and silicon," *Bell Syst Tech J*, vol. 39, no. 1, pp. 205–233, 1960.
- [216] V. E. Borisenko and S. G. Yudin, "Steady-state solubility of substitutional impurities in silicon," *physica status solidi (a)*, vol. 101, no. 1, pp. 123–127, 1987.
- [217] A. Armigliato, D. Nobili, M. Servidori, and S. Solmi, "SiP precipitation within the doped silicon lattice, concomitant with phosphorus predeposition," *Journal of Applied Physics*, vol. 47, no. 12, pp. 5489–5491, 1976.
- [218] S. Solmi, A. Parisini, R. Angelucci, A. Armigliato, D. Nobili, and L. Moro, "Dopant and carrier concentration in si in equilibrium with monoclinic SiP precipitates," *Phys. Rev. B*, vol. 53, no. 12, pp. 7836–7841, Mar 1996.
- [219] G. Vick and K. Whittle, "Solid solubility and diffusion coefficients of boron in silicon," *Journal of The Electrochemical Society*, vol. 116, no. 8, pp. 1142–1144, 1969.
- [220] J. C. Wolfe and W. Zagódzon-Wosik, "Method and apparatus for doping silicon wafers using a solid dopant source and rapid thermal processing," Patent US 5 550 082, 8 27, 1996.
- [221] C. S. Yoo, *Semiconductor Manufacturing Technology*. World Scientific Pub Co Inc, 2008.
- [222] B. H. Justice and R. F. Aycock, "Spin-on dopant method," Patent US 4 514 440, 12 12, 1983.

- [223] S. Proctor, M. Crain, and K. Walsh, “P+ structural layers for microelectromechanical systems using spin-on dopants,” *Sens Lett*, vol. 2, no. 3-4, pp. 211–216, 2004.
- [224] Z. Zhu, E. Menard, K. Hurley, R. Nuzzo, and J. Rogers, “Spin on dopants for high-performance single-crystal silicon transistors on flexible plastic substrates,” *Applied Physics Letters*, vol. 86, no. 13, p. 133507, 2005.
- [225] S. Franssila and MyiLibrary, *Introduction to microfabrication*. Wiley Online Library, 2004.
- [226] R. Jaeger, *Introduction to microelectronic fabrication*, ser. Modular series on solid state devices. Prentice Hall, 2002.
- [227] S. Solmi and D. Nobili, “High concentration diffusivity and clustering of arsenic and phosphorus in silicon,” *Journal of Applied Physics*, vol. 83, no. 5, pp. 2484–2490, 1998.
- [228] J. Tsai, “Shallow phosphorus diffusion profiles in silicon,” *Proceedings of the IEEE*, vol. 57, no. 9, pp. 1499–1506, 1969.
- [229] Y. M. Haddara, B. T. Folmer, M. E. Law, and T. Buyuklimanli, “Accurate measurements of the intrinsic diffusivities of boron and phosphorus in silicon,” *Applied Physics Letters*, vol. 77, no. 13, pp. 1976–1978, 2000.
- [230] R. Duffy, V. C. Venezia, J. Loo, M. J. P. Hopstaken, M. A. Verheijen, J. G. M. van Berkum, G. C. J. Maas, Y. Tamminga, T. Dao, and C. Demeurisse, “Low-temperature diffusion of high-concentration phosphorus in silicon, a preferential movement toward the surface,” *Applied Physics Letters*, vol. 86, no. 8, pp. 081917–3, Feb. 2005.
- [231] A. N. Gorban’ and V. A. Gorodokin, “Low-temperature diffusion of phosphorus in silicon,” *Russian Physics Journal*, vol. 31, no. 9, pp. 737–740, 1988.
- [232] R. B. Fair and J. C. C. Tsai, “A quantitative model for the diffusion of phosphorus in silicon and the emitter dip effect,” *Journal of The Electrochemical Society*, vol. 124, no. 7, pp. 1107–1118, 1977.

- [233] S. M. Hu, P. Fahey, and R. W. Dutton, “On models of phosphorus diffusion in silicon,” *Journal of Applied Physics*, vol. 54, no. 12, pp. 6912–6922, 1983.
- [234] C. Sah, H. Sello, and D. Tremere, “Diffusion of phosphorus in silicon oxide film,” *Journal of Physics and Chemistry of Solids*, vol. 11, no. 3-4, pp. 288 – 298, 1959. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/002236975990229X>
- [235] M. Ghezzo and D. M. Brown, “Diffusivity summary of b, ga, p, as, and sb in sio[sub 2],” *Journal of The Electrochemical Society*, vol. 120, no. 1, pp. 146–148, 1973. [Online]. Available: <http://link.aip.org/link/?JES/120/146/1>
- [236] S. Jones, “Diffusion in silicon,” *IC Knowledge LLC*, 2008.
- [237] R. OHL, “Properties of ionic bombarded silicon,” *Bell Syst Tech J*, vol. 31, no. 1, pp. 104–121, 1952.
- [238] W. Shockley, “Forming semiconductive devices by ionic bombardment,” Patent US 2 787 564, 10 28, 1954.
- [239] W. CUSSINS, “Effects produced by the ionic bombardment of germanium,” *P Phys Soc Lond B*, vol. 68, no. 4, pp. 213–222, 1955.
- [240] J. MAYER, O. MARSH, G. SHIFRIN, and R. BARON, “Ion implantation of silicon: Ii. electrical evaluation using hall-effect measurements,” *Can J Phys*, vol. 45, no. 12, pp. 4073–4089, 1967.
- [241] J. GIBBONS, “Ion implantation in semiconductors: I. range distribution theory and experiments,” *Pr Inst Electr Elect*, vol. 56, no. 3, pp. 295–319, 1968.
- [242] L. LARGE, “Ion implantation - a new method of doping semiconductors,” *Contemp Phys*, vol. 10, no. 5, pp. 505–531, 1969.
- [243] T. SEIDEL and A. MACRAE, “Some properties of ion implanted boron in silicon,” *T Metall Soc Aime*, vol. 245, no. 3, pp. 491–498, 1969.

- [244] J. GIBBONS, “Ion implantation in semiconductors: II. damage production and annealing,” *Pr Inst Electr Elect*, vol. 60, no. 9, pp. 1062–1096, 1972.
- [245] J. W. Mayer, “Ion implantation in semiconductors,” in *Proceedings of International Electron Device Meeting*, Washington, DC, USA, December 1973, pp. 3–5.
- [246] J. DAVIES, “Early ion implantation history (or “how i met jim mayer and hogsted phim!”),” *Mater Chem Phys*, vol. 46, no. 2-3, pp. 111–117, 1996.
- [247] C. Yarling, “History of industrial and commercial ion implantation 1906-1978,” *J Vac Sci Technol A*, vol. 18, no. 4, pp. 1746–1750, 2000.
- [248] H. Chen, F. Sinclair, and M. Sugitani, “Dose control for use in an ion implanter,” 1998, uS Patent 5,760,409.
- [249] A. A. Barlian, S.-J. Park, V. Mukundan, and B. L. Pruitt, “Design and characterization of microfabricated piezoresistive floating element-based shear stress sensors,” *Sensors and Actuators A: Physical*, vol. 134, no. 1, pp. 77–87, 2007.
- [250] B. Chui, T. Kenny, H. Mamin, B. Terris, and D. Rugar, “Independent detection of vertical and lateral forces with a sidewall-implanted dual-axis piezoresistive cantilever,” *Applied Physics Letters*, vol. 72, no. 11, pp. 1388–1390, 1998.
- [251] M. C. Park, J. Y. Hur, K. W. Kwon, S.-H. Park, and K. Y. Suh, “Pumpless, selective docking of yeast cells inside a microfluidic channel induced by receding meniscus,” *Lab Chip*, vol. 6, no. 8, p. 988, Jan 2006.
- [252] Y. Li, T. Nishida, D. P. Arnold, and M. Sheplak, “Microfabrication of a wall shear stress sensor using side-implanted piezoresistive tethers,” in *Proceedings of SPIE’s International Symposium on Smart Structures and Materials*, San Diego, CA, USA, March 2007, pp. 65 290D 1–11.
- [253] C. Bergaud, E. Cocheteau, L. Bary, R. Plana, and B. Belier, “Formation of implanted piezoresistors under 100-nm thick for nanoelectromechanical systems,” in *Proceedings of IEEE 15th International Conference on Micro Electro Mechanical Systems (MEMS2002)*, Las Vegas, NV, USA, January 2002, pp. 360–363.

- [254] M. Gel and I. Shimoyama, "Sub-micron thick high sensitive piezoresistive cantilevers by boron etch stop and argon implantation," *Micro Electro Mechanical Systems*, Jan 2003.
- [255] F. N. Schwettmann, "Characterization of incomplete activation of high dose boron implants in silicon," vol. 45, no. 4, pp. 1918–1920, 1974.
- [256] K. Sakamoto, K. Nishi, F. Ichikawa, and S. Ushio, "Segregation and transport coefficients of impurities at the si/sio₂ interface," *Journal of Applied Physics*, vol. 61, no. 4, pp. 1553–1555, 1987. [Online]. Available: <http://link.aip.org/link/?JAP/61/1553/1>
- [257] B. Joyce and J. Baldrey, "Selective epitaxial deposition of silicon," *Nature*, vol. 195, no. 4840, pp. 485–486, Jan 1962.
- [258] L. Jastrzebski, "Soi by cvd: epitaxial lateral overgrowth (elo) process-review," *Journal of Crystal Growth*, Jan 1983.
- [259] R. Pagliaro, J. Corboy, L. Jastrzebski, and R. Soydan, "Uniformly thick selective epitaxial silicon," *Journal of The Electrochemical Society*, Jan 1987.
- [260] B. Ginsberg, J. Burghartz, G. Bronner, and S. Mader, "Selective epitaxial growth of silicon and some potential applications," *IBM Journal of Research and Development*, Jan 1990.
- [261] M. R. Goulding, "The selective epitaxial growth of silicon," *J. Phys. IV France*, vol. 02, no. C2, pp. C2–745–C2–778, Sep 1991.
- [262] M. Bartek, P. T. J. Gennissen, P. J. French, P. M. Sarro, and R. F. Wolffenbuttel, "Study of selective and non-selective deposition of single- and polycrystalline silicon layers in an epitaxial reactor," in *Proceedings of Transducers 1997: International Conference on Solid-State Sensors, Actuators and Microsystems*, Chicago, IL, USA, June 1997, pp. 1403–1406.
- [263] J. Pak, G. Neudeck, and A. Kabir, "Micromechanical sensors using merged epitaxial lateral overgrowth of silicon," Ph.D. dissertation, Purdue University, 1993.

- [264] J. Regolini, D. Bensahel, E. Scheid, and J. Mercier, "Selective epitaxial silicon growth in the 650–1100° c range in a reduced pressure chemical vapor dots," *Applied Physics Letters*, Jan 1989.
- [265] Y. Liang, S. Ueng, and T. Kenny, "Performance characterization of ultra-thin n-type piezoresistive cantilevers," in *Proceedings of Transducers 2001: International Conference on Solid-State Sensors, Actuators and Microsystems*, Munich, Germany, June 2001, pp. 998–1001.
- [266] Y. Onuma and K. Sekiya, "Piezoresistive properties of polycrystalline silicon thin-film," *Jpn J Appl Phys*, vol. 11, no. 1, pp. 20–23, 1972.
- [267] R. W. Gurtler and R. W. Zwernemann, "Polycrystalline silicon pressure sensor," Patent US 3 858 150, 12 31, 1974.
- [268] J. Seto, "The electrical properties of polycrystalline silicon films," *Journal of Applied Physics*, vol. 46, pp. 5247–5254, Jan 1975.
- [269] J. SETO, "Piezoresistive properties of polycrystalline silicon," *Journal of Applied Physics*, vol. 47, no. 11, pp. 4780–4783, 1976.
- [270] N. Lu, L. Gerzberg, C. Lu, and J. Meindl, "Modeling and optimization of monolithic polycrystalline silicon resistors," *IEEE T Electron Dev*, vol. 28, no. 7, pp. 818–830, 1981.
- [271] N. LU, L. GERZBERG, C. LU, and J. MEINDL, "A new conduction model for polycrystalline silicon films," *Electron Devic Lett*, vol. 2, no. 4, pp. 95–98, 1981.
- [272] M. Mandurah, K. Saraswat, and T. Kamins, "A model for conduction in polycrystalline silicon: I. theory," *Ieee T Electron Dev*, vol. 28, no. 10, pp. 1163–1171, 1981.
- [273] M. M. Mandurah, *The physical and electrical properties of polycrystalline-silicon*, 1st ed. Stanford, CA: Ph.D., Stanford University, 1981.

- [274] H. MIKOSHIBA, "Stress-sensitive properties of silicon-gate mos devices," *Solid State Electron*, vol. 24, no. 3, pp. 221–232, 1981.
- [275] W. GERMER and W. TODT, "Low-cost pressure force transducer with silicon thin-film strain-gauges," *Sensor Actuator*, vol. 4, no. 2, pp. 183–189, 1983.
- [276] J. ERSKINE, "Polycrystalline silicon-on-metal strain-gauge transducers," *Ieee T Electron Dev*, vol. 30, no. 7, pp. 796–801, 1983.
- [277] P. FRENCH and A. EVANS, "Piezoresistance in polysilicon," *Electron Lett*, vol. 20, no. 24, pp. 999–1000, 1984.
- [278] P. FRENCH and A. EVANS, "Piezoresistance in polysilicon and its applications to strain-gauges," *Solid State Electron*, vol. 32, no. 1, pp. 1–10, 1989.
- [279] T. Trupke, M. A. Green, P. Würfel, P. P. Altermatt, A. Wang, J. Zhao, and R. Corkish, "Temperature dependence of the radiative recombination coefficient of intrinsic crystalline silicon," *Journal of Applied Physics*, vol. 94, no. 8, pp. 4930–4937, 2003.
- [280] C. D. Thurmond, "The standard thermodynamic functions for the formation of electrons and holes in ge, si, gaas, and gap," *Journal of The Electrochemical Society*, vol. 122, no. 8, pp. 1133–1141, 1975.
- [281] S. Reggiani, M. Valdinoci, L. Colalongo, M. Rudan, G. Baccarani, A. Stricker, F. Illien, N. Felber, W. Fichtner, and L. Zullino, "Electron and hole mobility in silicon at large operating temperatures. i. bulk mobility," *Electron Devices, IEEE Trans. on*, vol. 49, no. 3, pp. 490–499, Mar. 2002.
- [282] A. Manara, A. Ostidich, G. Pedroli, and G. Restelli, "Anodic oxidation as sectioning technique for the analysis of impurity concentration profiles in silicon," *Thin Solid Films*, vol. 8, no. 5, pp. 359 – 375, 1971. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0040609071900848>
- [283] P. Gaworzewski, L. Kalman, H. Rausch, and M. Trapp, "Doping profile techniques for si epitaxial layers," *Journal of Radioanalytical and Nuclear Chemistry*, vol. 52, pp. 93–100, 1979, 10.1007/BF02517703.

- [284] R. Levy, *Microelectronic materials and processes*, ser. NATO Advanced Study Institutes series. Series E, Applied sciences. Kluwer Academic, 1989.
- [285] D. Schroder, *Semiconductor material and device characterization*. IEEE Press, 2006.
- [286] H. W. Werner, “Quantitative secondary ion mass spectrometry: A review,” *Surface and Interface Analysis*, vol. 2, no. 2, pp. 56–74, 1980.
- [287] P. Williams, “Secondary ion mass spectrometry,” *Annual Review of Materials Science*, vol. 15, no. 1, pp. 517–548, 1985.
- [288] B. A, F. G. Rdennauer, and H. W. Werner, *Secondary ion mass spectrometry : basic concepts, instrumental aspects, applications and trends*. John Wiley & Sons, 1987.
- [289] D. Dickey and J. Ehrestein, *Spreading resistance analysis for silicon layers with nonuniform resistivity: David H. Dickey and James R. Ehrstein*, ser. NBS special publication. U.S. Dept. of Commerce, 1979.
- [290] R. G. Mazur and D. H. Dickey, “A spreading resistance technique for resistivity measurements on silicon,” *Journal of The Electrochemical Society*, vol. 113, no. 3, pp. 255–259, 1966.
- [291] C. C. Williams, “Two-dimensional dopant profiling by scanning capacitance microscopy,” *Annual Review of Materials Science*, vol. 29, no. 1, pp. 471–504, 1999.
- [292] J. J. Kopanski, J. F. Marchiando, and J. R. Lowney, “Scanning capacitance microscopy measurements and modeling: Progress towards dopant profiling of silicon,” vol. 14, no. 1. AVS, 1996, pp. 242–247.
- [293] A. K. Henning, T. Hochwitz, J. Slinkman, J. Never, S. Hoffmann, P. Kaszuba, and C. Daglian, “Two-dimensional surface dopant profiling in silicon using scanning kelvin probe microscopy,” *Journal of Applied Physics*, vol. 77, no. 5, pp. 1888–1896, 1995.

- [294] R. Galloni and A. Sardo, “Fully automatic apparatus for the determination of doping profiles in si by electrical measurements and anodic stripping,” vol. 54, no. 3, pp. 369–373, 1983. [Online]. Available: <http://dx.doi.org/doi/10.1063/1.1137376>
- [295] A. Gharbi, B. Remaki, A. Halimaoui, D. Bensahel, and A. Souifi, “p-type silicon doping profiling using electrochemical anodization,” vol. 109, no. 2, p. 023715, 2011. [Online]. Available: <http://dx.doi.org/doi/10.1063/1.3534005>
- [296] A. D. Kurtz, *Adjusting crystal characteristics to minimize temperature dependency in Semiconductor and Conventional Strain Gages*, M. Dean, Ed., 1st ed. New York: Academic Press, 1962.
- [297] Q. Wang, J. Ding, and W. Wang, “Fabrication and temperature coefficient compensation technology of low cost high temperature pressure sensor,” *Sensors and Actuators A: Physical*, vol. 120, no. 2, pp. 468 – 473, 2005.
- [298] S.-C. Kim and K. Wise, “Temperature sensitivity in silicon piezoresistive pressure transducers,” *Electron Devices, IEEE Transactions on*, vol. 30, no. 7, pp. 802 – 810, jul 1983.
- [299] M. Olszacki, C. Maj, M. A. Bahri, J.-C. Marrot, A. Boukabache, P. Pons, and A. Napieralski, “Experimental verification of temperature coefficients of resistance for uniformly doped p-type resistors in soi,” *Journal of Micromechanics and Microengineering*, vol. 20, no. 6, p. 064008, 2010.
- [300] W. Bullis, F. Brewer, C. Kolstad, and L. Swartzendruber, “Temperature coefficient of resistivity of silicon and germanium near room temperature,” *Solid-State Electronics*, vol. 11, no. 7, pp. 639 – 646, 1968.
- [301] K. J. Kim and W. P. King, “Thermal conduction between a heated microcantilever and a surrounding air environment,” *Applied Thermal Engineering*, vol. 29, no. 8-9, pp. 1631–1641, 2009.
- [302] J. C. Doll, E. A. Corbin, W. P. King, and B. L. Pruitt, “Self-heating in piezoresistive cantilevers,” *Applied Physics Letters*, vol. 98, no. 22, p. 223103, 2011.

- [303] M. von Arx, O. Paul, and H. Baltes, “Process-dependent thin-film thermal conductivities for thermal CMOS MEMS,” *Microelectromechanical Systems, Journal of*, vol. 9, no. 1, pp. 136–145, mar 2000.
- [304] C. Liu, *Foundations of MEMS*, ser. Illinois ECE Series. Pearson Prentice Hall, 2006. [Online]. Available: <http://books.google.com/books?id=jLxZSgAACAAJ>
- [305] M. Asheghi, Y. K. Leung, S. S. Wong, and K. E. Goodson, “Phonon-boundary scattering in thin silicon layers,” *Applied Physics Letters*, vol. 71, no. 13, pp. 1798–1800, 1997. [Online]. Available: <http://link.aip.org/link/?APL/71/1798/1>
- [306] W. Liu, K. Etessam-Yazdani, R. Hussin, and M. Asheghi, “Modeling and data for thermal conductivity of ultrathin single-crystal SOI layers at high temperature,” *Electron Devices, IEEE Transactions on*, vol. 53, no. 8, pp. 1868–1876, aug. 2006.
- [307] F. Incropera, T. Bergman, A. Lavine, and D. DeWitt, *Fundamentals of Heat and Mass Transfer*. John Wiley & Sons, 2011. [Online]. Available: <http://books.google.com/books?id=vvyIoXEywMoC>
- [308] O. Ozsun, B. E. Alaca, A. D. Yalcinkaya, M. Yilmaz, M. Zervas, and Y. Leblebici, “On heat transfer at microscale with implications for microactuator design,” *Journal of Micromechanics and Microengineering*, vol. 19, no. 4, p. 045020, 2009.
- [309] X. J. Hu, A. Jain, and K. E. Goodson, “Investigation of the natural convection boundary condition in microfabricated structures,” *International Journal of Thermal Sciences*, vol. 47, no. 7, pp. 820–824, 2008.
- [310] K. Park, G. L. W. Cross, Z. M. Zhang, and W. P. King, “Experimental investigation on the heat transfer between a heated microcantilever and a substrate,” *Journal of Heat Transfer*, vol. 130, no. 10, p. 102401, 2008. [Online]. Available: <http://link.aip.org/link/?JHR/130/102401/1>
- [311] J. Lee, T. Beechem, T. Wright, B. Nelson, S. Graham, and W. King, “Electrical, thermal, and mechanical characterization of silicon microcantilever heaters,” *Microelectromechanical Systems, Journal of*, vol. 15, no. 6, pp. 1644–1655, 2006.

- [312] B. Nelson and W. King, "Temperature calibration of heated silicon atomic force microscope cantilevers," *Sensors and Actuators A: Physical*, vol. 140, no. 1, pp. 51–59, 2007.
- [313] M. Abel, T. Wright, W. King, and S. Graham, "Thermal metrology of silicon microstructures using raman spectroscopy," *Components and Packaging Technologies, IEEE Transactions on*, vol. 30, no. 2, pp. 200 –208, june 2007.
- [314] K. Goodson and M. Flik, "Effect of microscale thermal conduction on the packing limit of silicon-on-insulator electronic devices," *Components, Hybrids, and Manufacturing Technology, IEEE Transactions on*, vol. 15, no. 5, pp. 715–722, oct. 1992.
- [315] J. C. Doll, S.-J. Park, and B. L. Pruitt, "Design optimization of piezoresistive cantilevers for force sensing in air and water," *Journal of Applied Physics*, vol. 106, no. 6, p. 064310, 2009.
- [316] F. Goericke, J. Lee, and W. P. King, "Microcantilever hotplates with temperature-compensated piezoresistive strain sensors," *Sensors and Actuators A: Physical*, vol. 143, no. 2, pp. 181–190, 2008.
- [317] M. Engesser, A. Franke, M. Maute, D. Meisel, and J. Korvink, "A robust and flexible optimization technique for efficient shrinking of MEMS accelerometers," *Microsystem Technologies*, vol. 16, no. 4, pp. 647–654, 2010.
- [318] C. Zhu, R. Byrd, P. Lu, and J. Nocedal, "L-bfgs-b: Fortran subroutines for large-scale bound constrained optimization," *ACM Transactions on Mathematical Software*, vol. 23, no. 4, pp. 550–560, Jan 1997.
- [319] O. Hansen and A. Boisen, "Noise in piezoresistive atomic force microscopy," *Nanotechnology*, vol. 10, no. 1, pp. 51–60, 1999.
- [320] S. Bhardwaj, M. Sheplak, and T. Nishida, "S/N optimization and noise considerations for piezoresistive microphones," in *Proceedings of 16th International*

- Conference on Noise in Physical Systems and 1/f Fluctuations*, Gainesville, FL, USA, October 2001, pp. 549–552.
- [321] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004. [Online]. Available: <http://books.google.com/books?id=mYm0bLd3fc0C>
 - [322] J. Nocedal and S. Wright, *Numerical optimization*, ser. Springer series in operations research. Springer, 1999. [Online]. Available: <http://books.google.com/books?id=w1kJYpOkPykC>
 - [323] M. Engesser, A. Franke, M. Maute, D. Meisel, and J. Korvink, “Miniaturization limits of piezoresistive mems accelerometers,” *Microsystem Technologies*, vol. 15, pp. 1835–1844, 2009, 10.1007/s00542-009-0920-4. [Online]. Available: <http://dx.doi.org/10.1007/s00542-009-0920-4>
 - [324] S. Armbruster, F. Schafer, G. Lammel, H. Artmann, C. Schelling, H. Benzel, S. Finkbeiner, F. Larmer, P. Ruther, and O. Paul, “A novel micromachining process for the fabrication of monocrystalline si-membranes using porous silicon,” in *Proceedings of Transducers 2003: International Conference on Solid-State Sensors, Actuators and Microsystems*, Boston, MA, USA, June 2003, pp. 246–249.
 - [325] G. Lammel, S. Armbruster, C. Schelling, H. Benzel, J. Brasas, M. Illing, R. Gampp, V. Senz, F. Schafer, and S. Finkbeiner, “Next generation pressure sensors in surface micromachining technology,” in *Proceedings of Transducers 2005: International Conference on Solid-State Sensors, Actuators and Microsystems*, Seoul, Korea, June 2005, pp. 35–36.
 - [326] S. Iyer, A. Auberton-Hervé, and I. of Electrical Engineers, *Silicon wafer bonding technology: for VLSI and MEMS applications*, ser. EMIS datareviews series. Institution of Electrical Engineers, 2002. [Online]. Available: <http://books.google.com/books?id=HLaSUCsbvx4C>
 - [327] M. Anc, *SIMOX*, ser. EMIS processing series. Institution of Electrical Engineers, 2004. [Online]. Available: http://books.google.com/books?id=rzCv3-_WlzwC

- [328] A. Nafari, D. Karlen, C. Rusu, K. Svensson, and P. Enoksson, "Boron impurity at the si/sio₂ interface in soi wafers and consequences for piezoresistive mems devices," *Journal of Micromechanics and Microengineering*, vol. 19, no. 1, p. 015034, 2009.
- [329] N. Harjee, "Coaxial-tip piezoresistive cantilever probes for high-resolution scanning gate microscopy," Ph.D. dissertation, Stanford University, 2011.
- [330] T. Iida, T. Itoh, D. Noguchi, and Y. Takano, "Residual lattice strain in thin silicon-on-insulator bonded wafers: Thermal behavior and formation mechanisms," *Journal of Applied Physics*, vol. 87, no. 2, pp. 675–681, 2000. [Online]. Available: <http://link.aip.org/link/?JAP/87/675/1>
- [331] G. Celotti, D. Nobili, and P. Ostoja, "Lattice parameter study of silicon uniformly doped with boron and phosphorus," *Journal of Materials Science*, vol. 9, pp. 821–828, 1974, 10.1007/BF00761802.
- [332] C. S. Lee, J. H. Lee, C. A. Choi, K. No, and D. M. Wee, "Effects of phosphorus on stress of multi-stacked polysilicon film and single crystalline silicon," *Journal of Micromechanics and Microengineering*, vol. 9, no. 3, p. 252, 1999. [Online]. Available: <http://stacks.iop.org/0960-1317/9/i=3/a=306>
- [333] K. McQuhae and A. Brown, "The lattice contraction coefficient of boron and phosphorus in silicon," *Solid-State Electronics*, vol. 15, no. 3, pp. 259 – 264, 1972. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0038110172900792>
- [334] T. Hubbard and J. Wylde, "Residual strain and resultant postrelease deflection of surface micromachined structures," vol. 18, no. 2. AVS, 2000, pp. 734–737. [Online]. Available: <http://link.aip.org/link/?JVA/18/734/1>
- [335] J. Arlett, "Properties of piezoresistive silicon nano-scale cantilevers with applications to bionems," Ph.D. dissertation, Stanford University, 2011.
- [336] A. M. Lin, R. W. Dutton, and D. A. Antoniadis, "The lateral effect of oxidation on boron diffusion in $\langle 100 \rangle$ silicon," *Applied Physics Letters*, vol. 35, no. 10, pp. 799–801, 1979. [Online]. Available: <http://link.aip.org/link/?APL/35/799/1>

- [337] F. Hooge and L. Ren, “On generation-recombination noise,” *Physica B: Condensed Matter*, vol. 191, no. 34, pp. 220 – 226, 1993. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/092145269390080P>
- [338] H. Sadeghian, C.-K. Yang, J. F. L. Goosen, A. Bossche, U. Staufer, P. J. French, and F. van Keulen, “Effects of size and defects on the elasticity of silicon nanocantilevers,” *Journal of Micromechanics and Microengineering*, vol. 20, no. 6, p. 064012, 2010. [Online]. Available: <http://stacks.iop.org/0960-1317/20/i=6/a=064012>
- [339] X. Li, T. Ono, Y. Wang, and M. Esashi, “Ultrathin single-crystalline-silicon cantilever resonators: Fabrication technology and significant specimen size effect on young’s modulus,” *Applied Physics Letters*, vol. 83, no. 15, pp. 3081–3083, 2003. [Online]. Available: <http://link.aip.org/link/?APL/83/3081/1>
- [340] J. Sader, I. Larson, P. Mulvaney, and L. White, “Method for the calibration of atomic force microscope cantilevers,” *Review of Scientific Instruments*, Jan 1995.
- [341] M. Li, H. X. Tang, and M. L. Roukes, “Ultra-sensitive NEMS-based cantilevers for sensing, scanned probe and very high-frequency applications,” *Nat Nano*, vol. 2, no. 2, pp. 114–120, Dec 2007.
- [342] J. Sader, “Frequency response of cantilever beams immersed in viscous fluids with applications to the atomic dots,” *Journal of Applied Physics*, Jan 1998.
- [343] W. Chu, “Southwest research institute technical report no. 2, dtmb, contract nobs-86396(x),” 1963.
- [344] C. V. Eysden and J. Sader, “Frequency response of cantilever beams immersed in viscous fluids with applications to the atomic force microscope: Arbitrary mode order,” *Journal of Applied Physics*, vol. 101, no. 4, p. 044908, Jan 2007.
- [345] S. Boskovic, J. Chon, P. Mulvaney, and J. Sader, “Rheological measurements using microcantilevers,” *Journal of Rheology*, Jan 2002.

- [346] J. Arlett, J. Maloney, B. Gudlewski, M. Muluneh, and M. Roukes, "Self-sensing micro-and nanocantilevers with attonewton-scale force resolution," *Nano Letters*, vol. 6, no. 5, pp. 1000–1006, Jan 2006.
- [347] G. Villanueva, J. Montserrat, F. Prez-Murano, G. Rius, and J. Bausells, "Submicron piezoresistive cantilevers in a CMOS-compatible technology for intermolecular force detection," *Microelectronic Engineering*, vol. 73-74, pp. 480–486, Jun 2004.
- [348] P. Negrini, A. Ravaglia, and S. Solmi, "Boron predeposition in silicon using BBr₃," *Journal of The Electrochemical Society*, vol. 125, no. 4, pp. 609–613, 1978. [Online]. Available: <http://link.aip.org/link/?JES/125/609/1>
- [349] M. Weinberg, "Working equations for piezoelectric actuators and sensors," *J Microelectromech S*, vol. 8, no. 4, pp. 529–533, Dec 1999.
- [350] M. Akiyama, N. Ueno, H. Tateyama, K. Nagao, and T. Yamada, "Preparation of highly oriented aluminum nitride thin films on molybdenum bottom electrodes using metal interlayers," *Journal of Materials Science*, vol. 40, no. 5, pp. 1159–1162, 2005.
- [351] T. Kamohara, M. Akiyama, N. Ueno, K. Nonaka, and H. Tateyama, "Growth of highly -axis-oriented aluminum nitride thin films on molybdenum electrodes using aluminum nitride interlayers," *Journal of Crystal Growth*, vol. 275, no. 3-4, pp. 383–388, Mar 2005.
- [352] T. Kamohara, M. Akiyama, N. Ueno, K. Nonaka, and N. Kuwano, "Influence of aluminum nitride interlayers on crystal orientation and piezoelectric property of aluminum nitride thin films prepared on titanium electrodes," *Thin Solid Films*, vol. 515, no. 11, pp. 4565–4569, Apr. 2007.
- [353] T. Kamohara, M. Akiyama, N. Ueno, and N. Kuwano, "Improvement in crystal orientation of AlN thin films prepared on mo electrodes using AlN interlayers," *Ceramics International*, vol. 34, no. 4, pp. 985–989, May 2008.
- [354] S. H. Baek, J. Park, D. M. Kim, V. A. Aksyuk, R. R. Das, S. D. Bu, D. A. Felker, J. Lettieri, V. Vaithyanathan, S. S. N. Bharadwaja, N. Bassiri-Gharb, Y. B. Chen,

- H. P. Sun, C. M. Folkman, H. W. Jang, D. J. Kreft, S. K. Streiffer, R. Ramesh, X. Q. Pan, S. Trolier-McKinstry, D. G. Schlom, M. S. Rzchowski, R. H. Blick, and C. B. Eom, "Giant piezoelectricity on si for hyperactive mems," *Science*, vol. 334, no. 6058, pp. 958–961, 2011.
- [355] J. C. Doll, B. C. Petzold, B. Ninan, R. Mullapudi, and B. L. Pruitt, "Aluminum nitride on titanium for CMOS compatible piezoelectric transducers," *Journal of Micromechanics and Microengineering*, vol. 20, no. 2, p. 025008, 2010.
- [356] R. Ruby, J. Larson, and P. Bradley, "Manufacturing process for thin film bulk acoustic resonator (FBAR) filters," Patent US 0 088 257, 2005.
- [357] R. Ruby, P. Bradley, Y. Oshmyansky, A. Chien, and J. I. Larson, "Thin film bulk wave acoustic resonators (fbar) for wireless applications," vol. 1, 2001, pp. 813–821 vol.1.
- [358] V. V. Felmetser, P. N. Laptev, and R. J. Graham, "Deposition of ultrathin aln films for high frequency electroacoustic devices," *Journal of Vacuum Science and Technology A: Vacuum, Surfaces, and Films*, vol. 29, no. 2, p. 021014, 2011. [Online]. Available: <http://link.aip.org/link/?JVA/29/021014/1>
- [359] C. ming Lin, T. ta Yen, Y. ju Lai, V. Felmetser, M. Hopcroft, J. Kuypers, and A. Pisano, "Temperature-compensated aluminum nitride lamb wave resonators," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 57, no. 3, pp. 524 –532, march 2010.
- [360] F. Martin, P. Muralt, M. Dubois, and A. Pezous, "Thickness dependence of the properties of highly c-axis textured AlN thin films," *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 22, no. 2, pp. 361–365, Mar. 2004.
- [361] R. B. Karabalin, M. H. Matheny, X. L. Feng, E. Defaÿ, G. L. Rhun, C. Marcoux, S. Hentz, P. Andreucci, and M. L. Roukes, "Piezoelectric nanoelectromechanical resonators based on aluminum nitride thin films," *Applied Physics Letters*, vol. 95, no. 10, p. 103111, 2009. [Online]. Available: <http://link.aip.org/link/?APL/95/103111/1>

- [362] N. Sinha, G. E. Wabiszewski, R. Mahameed, V. V. Felmetsger, S. M. Tanner, R. W. Carpick, and G. Piazza, "Piezoelectric aluminum nitride nanoelectromechanical actuators," *Applied Physics Letters*, vol. 95, no. 5, p. 053106, 2009.
- [363] K. Ng and R. Liu, "On the calculation of specific contact resistivity on lang;100 rang; si," *Electron Devices, IEEE Transactions on*, vol. 37, no. 6, pp. 1535 –1537, jun 1990.
- [364] H. H. Berger, "Contact resistance and contact resistivity," *Journal of The Electrochemical Society*, vol. 119, no. 4, pp. 507–514, 1972. [Online]. Available: <http://link.aip.org/link/?JES/119/507/1>
- [365] D. Schroder and D. Meier, "Solar cell contact resistance - a review," *Electron Devices, IEEE Transactions on*, vol. 31, no. 5, pp. 637 – 647, may 1984.
- [366] C. Ting and M. Wittmer, "The use of titanium-based contact barrier layers in silicon technology," *Thin Solid Films*, vol. 96, no. 4, pp. 327 – 345, 1982. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0040609082905168>
- [367] S. P. Murarka, "Interactions in metallization systems for integrated circuits," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 2, no. 4, pp. 693–706, 1984. [Online]. Available: <http://link.aip.org/link/?JVB/2/693/1>
- [368] N. Einspruch, S. Cohen, and G. Gildenblat, *VLSI metallization*, ser. VLSI electronics. Academic Press, 1987. [Online]. Available: <http://books.google.com/books?id=3jRTAAAAMAAJ>
- [369] E. Chow, "Electrical through-wafer interconnects and microfabricated cantilever arrays using through-wafer silicon etching," Ph.D. dissertation, Stanford University, 2001.
- [370] A. J. Walton, "Microelectronic test structures," 1997.

- [371] W. Graff, M. Matson, T. Kellner, T. Pluym, and S. Nelson, “RF and microwave plasma for resist and post-etch polymer removal,” *Solid State Technology*, vol. 44, no. 12, pp. 37–+, 2001.
- [372] X. Xu, V. Kuryatkov, B. Borisov, M. Pandikunta, S. A. Nikishin, and M. Holtz, “The effect of bcl3 pretreatment on the etching of aln in cl2-based plasma,” *MRS Proceedings*, 2008.
- [373] D. Buttari, A. Chini, T. Palacios, R. Coffie, L. Shen, H. Xing, S. Heikman, L. McCarthy, A. Chakraborty, S. Keller, and U. K. Mishra, “Origin of etch delay time in cl₂ dry etching of algan/gan structures,” *Applied Physics Letters*, vol. 83, no. 23, pp. 4779–4781, 2003. [Online]. Available: <http://link.aip.org/link/?APL/83/4779/1>
- [374] F. Khan, L. Zhou, V. Kumar, I. Adesida, and R. Okojie, “High rate etching of aln using bcl3/cl2/ar inductively coupled plasma,” *Materials Science and Engineering: B*, vol. 95, no. 1, pp. 51 – 54, 2002. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0921510702001605>
- [375] R. Blanc, M. Borel, B. Gluck, M. Heitzmann, and E. Vermande, “Etching of molybdenum with slope control for mobile communication applications,” *Proceedings of PESM 2010*, 2010.
- [376] S. ju Park, C. ping Sun, and R. J. Purtell, “A mechanistic study of sf6/o₂ reactive ion etching of molybdenum,” *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 5, no. 5, pp. 1372–1373, 1987. [Online]. Available: <http://link.aip.org/link/?JVB/5/1372/1>
- [377] A. N. Cleland, M. Pophristic, and I. Ferguson, “Single-crystal aluminum nitride nanomechanical resonators,” *Applied Physics Letters*, vol. 79, no. 13, pp. 2070–2072, 2001. [Online]. Available: <http://link.aip.org/link/?APL/79/2070/1>
- [378] B. Morgan, X. Hua, T. Iguchi, T. Tomioka, G. S. Oehrlein, and R. Ghodssi, “Substrate interconnect technologies for 3-d mems packaging,” *Microelectronic*

- Engineering*, vol. 81, no. 1, pp. 106 – 116, 2005. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0167931705001425>
- [379] N. Roxhed, P. Griss, and G. Stemme, “A method for tapered deep reactive ion etching using a modified bosch process,” *Journal of Micromechanics and Microengineering*, vol. 17, no. 5, p. 1087, 2007. [Online]. Available: <http://stacks.iop.org/0960-1317/17/i=5/a=031>
- [380] P. Dixit and J. Miao, “High aspect ratio vertical through-vias for 3d mems packaging applications by optimized three-step deep rie,” *Journal of The Electrochemical Society*, vol. 155, no. 2, pp. H85–H91, 2008. [Online]. Available: <http://link.aip.org/link/?JES/155/H85/1>
- [381] G. C. Hill, J. I. Padovani, J. C. Doll, B. W. Chui, D. Rugar, H. J. Mamin, N. Harjee, and B. L. Pruitt, “Patterned cracks improve yield in the release of compliant microdevices from silicon-on-insulator wafers,” *Journal of Micromechanics and Microengineering*, vol. 21, no. 8, p. 087001, 2011. [Online]. Available: <http://stacks.iop.org/0960-1317/21/i=8/a=087001>
- [382] R. Popovic, “Metal-n-type semiconductor ohmic contact with a shallow n+ surface layer,” *Solid-State Electronics*, vol. 21, no. 9, pp. 1133 – 1138, 1978. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0038110178903490>
- [383] T. J. Faith, R. S. Irven, S. K. Plante, and J. J. J. O'Neill, “Contact resistance: Al and al-si to diffused n+ and p+ silicon,” *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 1, no. 2, pp. 443–448, 1983. [Online]. Available: <http://link.aip.org/link/?JVA/1/443/1>
- [384] J. Vossen, “In situ low damage techniques for cleaning silicon surfaces prior to metal deposition,” *Thin Solid Films*, vol. 126, no. 34, pp. 213 – 218, 1985. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/004060908590313X>
- [385] J. Osenbach, “Corrosion-induced degradation of microelectronic devices,” *Semiconductor science and technology*, vol. 11, p. 155, 1996. [Online]. Available: <http://iopscience.iop.org/0268-1242/11/2/002>

- [386] M. Schindler, S. K. Kim, C. S. Hwang, C. Schindler, a. Offenhäusser, and S. Ingebrandt, “Novel post-process for the passivation of a CMOS biosensor,” *physica status solidi (RRL) Rapid Research Letters*, vol. 2, no. 1, pp. 4–6, Jan. 2008. [Online]. Available: <http://doi.wiley.com/10.1002/pssr.200701242>
- [387] S. E. Potts, L. Schmalz, M. Fenker, B. Diaz, J. Swiatowska, V. Maurice, a. Seyeux, P. Marcus, G. Radnoci, L. Toth, and W. M. M. Kessels, “Ultra-Thin Aluminium Oxide Films Deposited by Plasma-Enhanced Atomic Layer Deposition for Corrosion Protection,” *Journal of The Electrochemical Society*, vol. 158, no. 5, p. C132, 2011. [Online]. Available: <http://link.aip.org/link/JESOAN/v158/i5/pC132/s1&Agg=doi>
- [388] W. Li, D. Rodger, P. Menon, and Y.-C. Tai, “Corrosion Behavior of Parylene-Metal-Parylene Thin Films in Saline,” *ECS Transactions*, vol. 11, no. 18, pp. 1–6, 2008. [Online]. Available: <http://link.aip.org/link/ECSTF8/v11/i18/p1/s1&Agg=doi>
- [389] G. Yang, D. Mathur, X. Wu, S. Dabral, J. Donald, T. Lu, and H. Bakhru, “Microstructures of parylene-n thin films and the effect on copper diffusion,” *Journal of Electronic Materials*, vol. 25, no. 11, pp. 1778–1783, 1996. [Online]. Available: <http://www.ingentaconnect.com/content/klu/jem/1996/00000025/00000011/00000035>
- [390] Z. X., M. J.F., W. B., D. S., and C. C., “Crystallinity properties of parylene-n affecting its use as an ILD in submicron integrated circuit technology,” *Thin Solid Films*, vol. 270, no. 1, pp. 508–511, 1995. [Online]. Available: <http://www.ingentaconnect.com/content/els/00406090/1995/00000270/00000001/art06842>
- [391] A. Offenhausser, J. Ruhe, and W. Knoll, “Neuronal cells cultured on modified microelectronic device surfaces,” *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, vol. 13, no. 5, pp. 2606–2612, Sept. 1995. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4930279

- [392] G. Schmitt, J. Schultze, F. Fassbender, G. Buss, H. Luth, and M. Schoning, “Passivation and corrosion of microelectrode arrays,” *Electrochim Acta*, vol. 44, no. 21-22, pp. 3865–3883, Jan 1999.
- [393] S. M. George, “Atomic layer deposition: An overview,” *Chemistry Review*, pp. 1–21, Jan 2010.
- [394] R. L. Puurunen, “Surface chemistry of atomic layer deposition: A case study for the trimethylaluminum/water process,” *J. Appl. Phys.*, vol. 97, no. 12, p. 121301, Jan 2005.
- [395] B. Ilic, S. Krylov, and H. G. Craighead, “Young’s modulus and density measurements of thin atomic layer deposited films using resonant nanomechanics,” *Journal of Applied Physics*, vol. 108, no. 4, p. 044317, 2010. [Online]. Available: <http://link.aip.org/link/?JAP/108/044317/1>
- [396] S. Lee, J. Licari, and I. Litant, “Electrical reliability of parylene films for device passivation,” *Metallurgical and Materials Transactions B*, vol. 1, pp. 701–711, 1970, 10.1007/BF02811598. [Online]. Available: <http://dx.doi.org/10.1007/BF02811598>
- [397] L. D. Hanley and J. H. Martin, “A test of parylene as a protective system for microcircuitry,” in *Reliability Physics Symposium, 1975. 13th Annual*, april 1975, pp. 53 –57.
- [398] C. Hassler, R. P. von Metzen, P. Ruther, and T. Stieglitz, “Characterization of parylene c as an encapsulation material for implanted neural prostheses,” *Journal of Biomedical Materials Research Part B: Applied Biomaterials*, vol. 93B, no. 1, pp. 266–274, 2010. [Online]. Available: <http://dx.doi.org/10.1002/jbm.b.31584>
- [399] T.-N. Chen, D.-S. Wuu, C.-C. Wu, C.-C. Chiang, Y.-P. Chen, and R.-H. Horng, “Improvements of permeation barrier coatings using encapsulated parylene interlayers for flexible electronic applications,” *Plasma Processes and Polymers*, vol. 4, no. 2, pp. 180–185, 2007. [Online]. Available: <http://dx.doi.org/10.1002/ppap.200600158>

- [400] P. Wu, G. Yang, J. McDonald, and T. Lu, “Surface reaction and stability of parylene n and f thin films at elevated temperatures,” *Journal of Electronic Materials*, vol. 24, pp. 53–58, 1995, 10.1007/BF02659727. [Online]. Available: <http://dx.doi.org/10.1007/BF02659727>
- [401] S. Dabral, J. Van Etten, X. Zhang, C. Apblett, G. Yang, P. Ficalora, and J. McDonald, “Stress in thermally annealed parylene films,” *Journal of Electronic Materials*, vol. 21, pp. 989–994, 1992, 10.1007/BF02684208. [Online]. Available: <http://dx.doi.org/10.1007/BF02684208>
- [402] R. Huang and Y. Tai, “Parylene to silicon adhesion enhancement,” in *Solid-State Sensors, Actuators and Microsystems Conference, 2009. TRANSDUCERS 2009. International*, june 2009, pp. 1027 –1030.
- [403] N. A. Burnham, X. Chen, C. S. Hodges, G. A. Matei, E. J. Thoreson, C. J. Roberts, M. C. Davies, and S. J. B. Tendler, “Comparison of calibration methods for atomic-force microscopy cantilevers,” *Nanotechnology*, vol. 14, no. 1, p. 1, 2003. [Online]. Available: <http://stacks.iop.org/0957-4484/14/i=1/a=301>
- [404] J. Cleveland, S. Manne, D. Bocek, and P. Hansma, “A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy,” *Review of Scientific Instruments*, 1993.
- [405] J. Sader, J. Chon, and P. Mulvaney, “Calibration of rectangular atomic force microscope cantilevers,” *Review of Scientific Instruments*, Jan 1999.
- [406] I. Behrens, L. Doering, and E. Peiner, “Piezoresistive cantilever as portable micro force calibration standard,” *Journal of Micromechanics and Microengineering*, vol. 13, no. 4, p. S171, 2003. [Online]. Available: <http://stacks.iop.org/0960-1317/13/i=4/a=325>
- [407] J. R. Pratt, J. A. Kramar, D. B. Newell, and D. T. Smith, “Review of si traceable force metrology for instrumented indentation and atomic force microscopy,” *Measurement Science and Technology*, vol. 16, no. 11, p. 2129, 2005. [Online]. Available: <http://stacks.iop.org/0957-0233/16/i=11/a=002>

- [408] R. S. Gates and J. R. Pratt, "Prototype cantilevers for si-traceable nanonewton force calibration," *Measurement Science and Technology*, vol. 17, no. 10, p. 2852, 2006. [Online]. Available: <http://stacks.iop.org/0957-0233/17/i=10/a=041>
- [409] R. B. Bhiladvala and Z. J. Wang, "Effect of fluids on the q factor and resonance frequency of oscillating micrometer and nanometer scale beams," *Phys. Rev. E*, vol. 69, p. 036307, Mar 2004. [Online]. Available: <http://link.aps.org/doi/10.1103/PhysRevE.69.036307>
- [410] L. Aeschimann, A. Meister, T. Akiyama, B. W. Chui, P. Niedermann, H. Heinzelmann, N. F. D. Rooij, U. Staufer, and P. Vettiger, "Scanning probe arrays for life sciences and nanobiology applications," *Microelectronic Engineering*, vol. 83, no. 4-9, 2006.
- [411] C.-P. Wu, S. S. Yang, S.-J. Hung, C.-T. Tu, and M. S.-C. Lu, "CMOS integrated cantilevers with sub-um tips for surface temperature measurement," *Journal of Micromechanics and Microengineering*, vol. 17, no. 12, p. 2432, 2007. [Online]. Available: <http://stacks.iop.org/0960-1317/17/i=12/a=007>
- [412] A. J. Ricci, H. J. Kennedy, A. C. Crawford, and R. Fettiplace, "The transduction channel filter in auditory hair cells," *Journal of Neuroscience*, vol. 25, no. 34, pp. 7831–7839, Aug 2005.
- [413] J. Linden and E. Oesterschulze, "Improving the quality factor of cantilevers in viscous fluids by the adaptation of their interface," *Applied Physics Letters*, vol. 100, no. 11, p. 113511, 2012. [Online]. Available: <http://link.aip.org/link/?APL/100/113511/1>
- [414] H. Cho, P. Kapur, and K. Saraswat, "Etch processes for nanoscale vertical mos devices," in *Proceedings of PEUG 2006*, 2006. [Online]. Available: http://www.avsusergroups.org/pag_pdfs/PEUG_2006_07_HoonCho.pdf
- [415] E. Rosenberg, *Trace Organic Analysis by Gas Chromatography with Selective Detectors*. John Wiley & Sons, 2006. [Online]. Available: <http://dx.doi.org/10.1002/9780470027318.a0875>

- [416] J. Davidow, M. Sarfaty, and D. Lymberopoulos, "Method and apparatus for monitoring the process state of a semiconductor device fabrication process," Patent US 6 455 437, 9 24, 2002.
- [417] K. Williams and R. Muller, "Etch rates for micromachining processing," *Microelectromechanical Systems*, Dec 1996.
- [418] K. Williams, K. Gupta, and M. Wasilik, "Etch rates for micromachining processing-part ii," *Microelectromechanical Systems, Journal of*, vol. 12, no. 6, pp. 761 – 778, dec. 2003.
- [419] S. Saravanan, E. Berenschot, G. Krijnen, and M. Elwenspoek, "A novel surface micromachining process to fabricate AlN unimorph suspensions and its application for RF resonators," *Sensors and Actuators A: Physical*, vol. 130-131, pp. 340–345, Aug. 2006.
- [420] J. Buhler, F.-P. Steiner, and H. Baltes, "Silicon dioxide sacrificial layer etching in surface micromachining," *Journal of Micromechanics and Microengineering*, vol. 7, no. 1, p. R1, 1997. [Online]. Available: <http://stacks.iop.org/0960-1317/7/i=1/a=001>
- [421] J. Doll, B. Petzold, B. Ninan, R. Mullapudi, and B. Pruitt, "A high d_{33} CMOS compatible process for aluminum nitride on titanium," in *Transducers International Conference*, june 2009.
- [422] J. Lee, J. Jung, M. Lee, and J. Park, "Effects of bottom electrodes on the orientation of AlN films and the frequency responses of resonators in AlN-based FBARs," *Thin Solid Films*, vol. 447-448, pp. 610–614, 2004.
- [423] M. Dubois and P. Muralt, "Stress and piezoelectric properties of aluminum nitride thin films deposited onto metal electrodes by pulsed direct current reactive sputtering," *Journal of Applied Physics*, vol. 89, no. 11, pp. 6389–6395, Dec 2001.
- [424] J. A. Ruffner, P. G. Clem, B. A. Tuttle, D. Dimos, and D. M. Gonzales, "Effect of substrate composition on the piezoelectric response of reactively sputtered AlN thin films," *Thin Solid Films*, vol. 354, no. 1-2, pp. 256–261, Oct. 1999.

- [425] W. Liu, S. Wu, C. Chen, Y. Lai, and C. Chuang, "Microstructural evolution and formation of highly c-axis-oriented aluminum nitride films by reactively magnetron sputtering deposition," *Journal of Crystal Growth*, vol. 276, no. 3-4, pp. 525–533, Apr. 2005.
- [426] J. Harley and T. Kenny, "A high-stiffness axial resonant probe for atomic force microscopy," *Microelectromechanical Systems*, Dec 2001.
- [427] T. Kamohara, M. Akiyama, and N. Kuwano, "Influence of polar distribution on piezoelectric response of aluminum nitride thin films," *Applied Physics Letters*, vol. 92, p. 3506, Feb 2008.
- [428] T. Kamohara, M. Akiyama, N. Ueno, M. Sakamoto, K. Kano, A. Teshigahara, N. Kawahara, and N. Kuwano, "Influence of sputtering pressure on polarity distribution of aluminum nitride thin films," *Applied Physics Letters*, vol. 89, no. 24, p. 243507, 2006.
- [429] A. Sanz-Hervas, L. Vergara, J. Olivares, E. Iborra, Y. Morilla, J. García-López, M. Clement, J. Sangrador, and M. Respaldiza, "Comparative study of c-axis AlN films sputtered on metallic surfaces," *Diamond and Related Materials*, vol. 14, no. 3-7, pp. 1198–1202, 2005.
- [430] K. Tonisch, V. Cimalla, C. Foerster, H. Romanus, O. Ambacher, and D. Dontsov, "Piezoelectric properties of polycrystalline AlN thin films for MEMS application," *Sensors and Actuators A: Physical*, vol. 132, no. 2, pp. 658–663, Nov. 2006.
- [431] M. Clement, L. Vergara, J. Olivares, E. Iborra, J. Sangrador, A. Sanz-Hervas, and C. Zinck, "SAW and BAW response of c-axis AlN thin films sputtered on platinum," in *Ultrasonics Symposium, 2004 IEEE*, vol. 2, 2004, pp. 1367–1370 Vol.2.
- [432] M. Akiyama, T. Kamohara, K. Kano, A. Teshigahara, and N. Kawahara, "Influence of oxygen concentration in sputtering gas on piezoelectric response of aluminum nitride thin films," *Applied Physics Letters*, vol. 93, no. 2, p. 021903, Jan 2008.

- [433] G. Iriarte, J. Bjurstrom, J. Westlinder, F. Engelmark, and I. Katardjiev, "Synthesis of c-axis-oriented AlN thin films on high-conducting layers: Al, mo, ti, TiN, and ni," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 52, no. 7, pp. 1170–1174, 2005.