CONTROL-SYSTEMS BASED ANALYSIS AND DESIGN METHODS FOR SCANNING PROBE MICROSCOPY

BY

CHI BUM LEE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mechanical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Assistant Professor Srinivasa M. Salapaka, Chair
Professor Placid M. Ferreira
Professor Geir E. Dullerud
Professor Narayana R. Aluru
Professor Petros G. Voulgaris

ABSTRACT

Recent demonstrations of nanoscience provide ample evidence indicating the feasibility of rational control, manipulation and interrogation of matter at the atomic scale. A class of devices, with micro-sized sensor probes, called Scanning Probe Microscopes (SPMs) are in the forefront of the technology that have demonstrated imaging and manipulation of sample properties. However, this technology is still far from realizing the promise of routinely tailoring matter at the atomic scale. Such an ability, once realized, will have far reaching impact and revolutionize every area of science and technology, especially the areas of material science, biology, medicine, and manufacturing.

This thesis presents control systems theoretic analysis and synthesis of new modes of operations that significantly expand the range of performance specifications and capabilities of SPMs. In particular, the focus of this thesis is on the two main requirements of SPMs, the precision positioning of the matter with respect to the probe and the obtaining of the surface topography from the probe data. A characterization of the inherent fundamental trade-offs between resolution, tracking-bandwidth, and reliability specifications on the positioning capability of these devices. A series of control designs which exploit these trade-offs appropriately to achieve pre-specified feasible performance objectives is discussed. These designs have two degrees of freedom (2DOF), that is, have the feedforward and the feedback components, and are obtained using the optimal control framework. Implementations and experimental results on the application of these designs show over 100-300% improvement over competing existing designs. For imaging, control systems tools have been used to model and analyze probe-sample (matter) interactions and design signals that estimate the sample topography. The central concept in this design is to view sample-topography signals as disturbance signals and use system theoretic tools to estimate them. Experiments using these

estimate signals show substantial improvements in imaging and detection bandwidths, and fewer artifacts and misinterpretations in SPM imaging.

To my wife, Seungyoon and my son, Daniel Jungwoo Lee,

for their love and support.

ACKNOWLEDGMENTS

Many people have contributed to this work. First, I would also like to thank my advisor Srinivasa M. Salapaka for his support, valuable advice, and constant encouragement. I was fortunate in having him as my advisor and mentor. I am indebted to him for much, more than advising and teaching me.

I give my gratitude to my doctoral committee: Professor Placid M. Ferreira, Professor Geir E. Dullerud, Professor Narayana R. Aluru and Professor Petros G. Voulgaris for teaching me several great courses, and giving me many valuable guidance and suggestions.

I would like to thank my friends, and office mates who spent this long time with me and made my life and work at UIUC a pleasant experience. My thanks go to all of these people along with a wish for the very best in their future endeavors.

I would like to thank my parents for all the love, concern and constant support. Finally, I thank my wife and son for encouraging me all the time and putting up with my long time of study. I cannot thank my wife enough and she is most precious to me. My son, who is the important meaning of my life, always gives me happiness. I dedicate this thesis to them.

TABLE OF CONTENTS

LIST OF FIGURES

ix

LISF OF ABBREVIATIONS

SPM   Scanning Probe Microscope

STM   Scanning Tunneling Microscopy

AFM   Atomic Force Microscopy

2DOF   Two Degree-Of-Freedom

NP   Nevanlinna-Pick

LMI   Linear Matrix Inequality

AM-AFM   Amplitude Modulation dynamic mode Atomic Force Microscopy

DMT   Derjaguin-Müller-Toporov

JKR   Johnson-Kendall-Roberts

KBM   Krylov-Bogoliubov-Mitropolskii

DSP   Digital Signal Processor

FPGA   Field Programmable Gate Array

CHAPTER 1

INTRODUCTION

The vision of studying the physics of nanoscale structures was first discussed by Richard Feynman in "There's Plenty of Room at the Bottom," a talk at the annual meeting of the American Physical Society at Caltech on December 29, 1959 [1] - It is said that this talk was the origin of "nanotechnology" though the words start being used later [2]. He asserted that laws of physics do not prevent manipulation and interrogation of material at the atomic scale. He provided a tantalizing glimpse of the potential impact of the ability to manipulate and interrogate matter at the atomic scale. He outlined a series of machines by which the ability to manipulate the individual atoms and molecules might be improved successively.

With the invention of the scanning tunneling microscope (STM) by G.Binnig and H. Roher in 1981 [3], Feynman's idea was embodied and the vision of atomic-scale interrogation and manipulation was realized. The STM was the first instrument to generate three dimensional images of surfaces with atomic resolution. The STM uses the quantum-mechanical phenomenon, tunneling current, which flows between two electrodes even through a thin insulator or a vacuum gap, as a sensing signal for the sample-topography. When a conducting tip is brought into proximity (less than 1 nm) of the *sample* surface, a bias voltage applied between the two can allow electrons to tunnel through the vacuum between them. The resulting tunneling current is exponentially dependent on a tip-surface separation with the atomic radius length scale and used to determine the topography of the surface.

The use of STM is limited to conductors and semiconductors. Imaging with atomic-scale resolution for insulators was unavailable until 1986, when G. Binnig, C. Quate, and C. Gerber invented the atomic force microscope (AFM) [4]. The AFM uses the interatomic force such as mechanical contact force, van der Waals forces, capillary forces, chemical bonding, electrostatic forces, magnetic forces to sense the sample-topography. When the tip

of a micro-cantilever probe is brought near to the sample surface, interatomic forces lead to a deflection of the cantilever. This deflection is used to determine the topography of the surface.

The STM and AFM heralded a new class of instruments, scanning probe microscopes (SPMs), based on various types of interaction between the tip and the sample surface. With an appropriate probe, SPMs have been developed that investigate and manipulate different physical variables such as magnetic, electrical, biological and chemical properties. All SPMs utilize the interaction between micro-probe and sample, and map the interaction to three dimensional images by the mechanical positioning for scanning. SPMs have the following advantages: a) The imaging resolution is not limited by diffraction, but by the probe-sample interaction volume. b) The probe-sample interaction can be used to create or modify the structure on sample. c) The operation does not necessarily require highly controlled (such as low temperature or high vacuum) environments, for example, the sample specimen can be in air or submerged in liquid. SPMs are in the forefront of nanoscale investigation, and form the basis of devices that promise realizing the potential of nanotechnology.

The concept of feedback played a crucial role in the development of SPMs. Before G.Binnig and H.Rohrer invented STM, earlier attempts to use tunneling current to image material with atomic-scale resolution were not successful due to extraneous vibrations that made it difficult to maintain a small separation between the probe and the sample. They were aware of the role of feedback strategies, and they used proportional-integral (PI) controllers. The feedback loop that controlled the spacing between the probe and the sample surface was pivotal to the success of tunneling experiment.

The current demands imposed by the growing number of applications of SPM technology include faster imaging, higher resolution imaging, robust operation, and quantitative measures of image fidelity. Although classical and simple control strategies were used in early STMs and AFMs, modern control and systems tools are needed to meet the demands of future SPM technology.

1.1 Scope of this Dissertation

The aim of this thesis is introducing systems-based principles and viewpoints that aim at fundamental changes to the existing technology towards tapping full potential of nano-technology. This research focuses on AFM, one of the most versatile and widely used SPM that has already demonstrated atomic-scale imaging and manipulation of matter. However, most of the methods developed in this work can be applied to a significantly larger class of devices than AFMs.
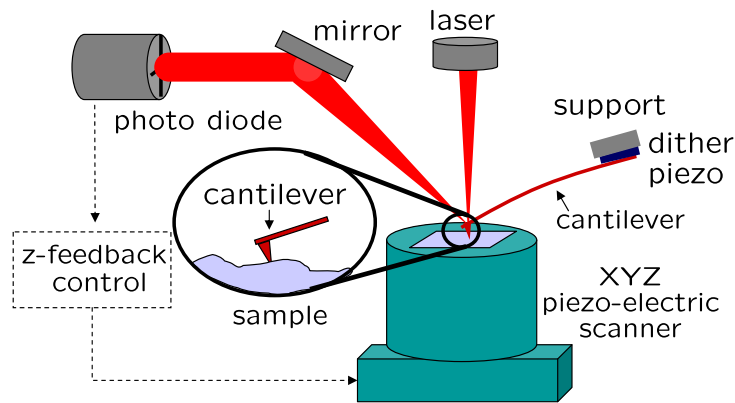
Figure 1.1: Atomic force microscope (AFM): A micro-cantilever deflects due to interactive forces between the atoms on the sample and the atoms on the tip. The deflection of the cantilever is registered by a laser incident on the cantilever, which reflects onto a split photodiode. The control signal, which regulates a reference set-point by adjusting the distance between the sample and the cantilever probe, gives a measure of the sample topography. A X-Y piezo scanner positions the sample relative to the cantilever in the lateral directions. A vertical piezo-actuator keeps the tip-sample interaction constant by positioning the sample or cantilever vertically.

The operating principle of an AFM is to employ a relatively soft micro-cantilever with very high resonance frequency capable of sensing extremely small interactive forces between the tip and the sample. The small stiffness facilitates detection of interatomic forces whereas high resonant frequency reduces the effect of various sources of vibrations. The imaging idea is to detect displacements of micro-cantilever tip as the sample surface moves under it (see Figure 1.1).

Considerable challenges need to be met to realize the full potential of SPMs. One of the

severe drawbacks of SPM is its imaging bandwidth. Existing methods of imaging that use SPMs are slow and therefore are rendered impractical for many applications that need high throughputs or track fast dynamic events. One of the primary reasons for this drawback is the low bandwidth of nanopositioning systems that are on the order of 1 kHz, which is two orders smaller than the resonance frequencies of the cantilevers. In this research, an optimal-control based framework for studying, analyzing, and designing control for nanopositioning systems is developed. This framework assesses fundamental limitations on employing specific design schemes, studies and quantifies trade-offs between different performance objectives, and designs and implements control laws for achieving feasible specifications for positioning resolution, bandwidth and robustness.

In view of nonlinear tip-sample interactions, sensor, electronic and thermal noise, and diverse operating conditions and requirements, the main challenge in design of the device and its control is preserving the high resolution sensing capability of the micro-cantilevers. Even though the robustness of devices to the above uncertainties is of primary importance for guaranteeing reliable resolution, it is typically ignored in existing methods. One of the main challenges in imaging arises from the complex tip-sample interaction, which makes it difficult to extract the sample-topography information. Even though, this complexity is diminished by use of feedback control for force regulation, existing methods impose their own constraints that prevent from utilizing the full information in the deflection signal for topography estimation.

To summarize, this research employs control and dynamics system theoretic tools to develop a framework for significantly improving AFM (and SPM) instrumentation. This framework results in understanding of certain observed phenomena, overcoming certain perceived limitations, spawning new operating modes, and obtaining substantial improvements in investigating bandwidths and precision.

1.2   Outline

I. Robust Broadband Nanopositioning

Part I of the thesis mainly presents analysis and design of 2DOF control methods for achieving robust high-precision high-bandwidth nanopositioning systems. In Chapter 2, a brief overview nanopositioning systems, a survey of nanopositioning related control systems literature, and analysis of theoretical and practical limitations on the control design for nanopositioning systems are presented. Chapter 3 develops the analysis and design of two degree-of-freedom (2DOF) control designs for four practical considerations in the context of nanopositioning systems. The controllers are designed in an optimal control framework, and its experimental implementation and verification are also presented. The advantages of these designs over one-degree of freedom (1DOF) designs, their limitations, and the relative roles of feedback and feedforward components in these designs, are discussed in Chapter 4.

II. Control Approach for Fast Imaging in Dynamic AFM

An analysis of probe-sample dynamics from a system theoretic viewpoint, and design and analysis of a new sample-topography estimate signal is presented in Part II. Chapter 5 describes working principles of different modes of operations of AFM and presents a brief literature survey of system theoretic approaches applied to them. The probe-sample interaction models in the context of *dynamic-mode* imaging and their simplifications are discussed in Chapter 6. Chapter 7 develops a sample-topography estimate signal that outperforms currently existing methods.

Observations, conclusions, and future directions for this research are presented in Chapter 8.

Part I

ROBUST BROADBAND NANOPOSITIONING

CHAPTER 2

ROBUST BROADBAND NANOPOSITIONING

Most of the applications in micro/nano-technology impose severe specifications on position-ing. This demand of ultra-high precision in positioning is one of the pivotal requirements in many applications. The ultra-high precision positioning systems are crucial in auto focus systems in optics [5], disk spin stands and vibration cancellation [6, 7, 8] in disk drives, wafer and mask positioning in microelectronics [9, 10, 11], piezo hammers in precision mechanics [12], and cell penetration and micro dispensing devices in medicine and biology [13]. As well as these applications, nanoscientific studies, such as in scanning probe microscopy, demand positioning systems with atomic scale resolutions. There is an added impetus on design of nanopositioning systems since they form the bottleneck in terms of speed and accuracy of most devices for nano-investigation, especially in SPMs. For instance in AFMs, the position-ing resolution and tracking bandwidth of positioning systems is typically few orders less than the imaging resolution and bandwidth that micro-cantilever probe provides. Besides high precision positioning, most nanoscientific studies and applications impose severe demands on the tracking bandwidth and reliability in terms of repeatability of experiments. High tracking bandwidth is required as many studies, especially in biology and material science, require assaying matter with nanoscale precision over areas with characteristic lengths that are typically three orders or more. Repeatability of experiments is essential for validation of the underlying studies. This requirement translates to robustness of positioning systems to modeling uncertainties and operating conditions. Devices that are insensitive to (robust to) diverse operating conditions give repeatable measurements, and are hence reliable.

Typical nanopositioning systems comprise of a flexure stage that provides frictionless motion through elastic deformation, an actuator, typically made from piezoelectric material that provides the required force to deform the flexure stage and/or sensing system along with

7

the control system. The main challenges to the design of robust broadband nanopositioning systems come from flexure-stage dynamics that limit the bandwidth of the positioning stage, from nonlinear effects of piezo-actuation such as hysteresis and creep that are difficult to model, and from sensor noise management issues in control feedback that can potentially hamper the tracking-resolution of the device.

The research on nanopositioning can largely be characterized as those that aim at re-designing the flexure stage mechanisms and those that propose new control methods for existing stages. The most common redesign approach is to design smaller and stiffer flexure stages which result in higher resonance frequencies, and therefore provide better tracking-bandwidths [14, 15]. However, the downside of this approach is that the resulting smaller stiffer devices have reduced traversal ranges.

In many nanopositioning systems that run in open loop or that apply feedforward control for improving performance, nonlinear piezoelectric effects, such as hysteresis and creep, form the main impediments to attaining high resolution and tracking bandwidth specifications. A feedforward scheme for improving the accuracy of a nanopositioning system is discussed in [16]. Energy-based models of piezoelectric behavior that incorporate the dependence of hysteresis loops on the scan rates are developed and applied to piezoelectric actuators in [17, 18]. These models, which are based on physical principles, predict a greater number of observed features in experiments than the phenomenological models, such as the Preisach model [19, 20]. Feedforward implementation schemes based on these models thus provide better positioning resolution [17]. The use of charge instead of voltage in driving piezoelectric tubes is another way for reducing hysteresis [21]. A feedforward control designed with $\mathcal{H}_\infty$ robust control framework based on experimentally identified tube type piezo-actuator model, compensate the lateral oscillation and provide better tracking accuracy in [22]. Preview based optimal inversion achieves precise output tracking when the desired output trajectory is known in advance for only a finite time interval [23].

Feedback control designs [24] with large gains at low frequencies have been demon-strated that make positioning resolution practically independent of piezoelectric nonlinear-ities, where nonlinear effects become negligible compared to measurement noise. Feedback control design that is robust to the modeling uncertainties are reported in [25, 26, 27, 28]. For

8

instance, a feedback control framework presented in [27] determines and quantifies trade-offs between performance objectives, assesses if desired specifications are feasible and provides a way to design controllers to achieve specifications when possible. In [29], a concerted effort of device and control design is presented, where a lighter (and therefore faster) device is designed by having a single moving mass and the responsibility of compensating mechanical coupling is delegated to the multi-input multi-output (MIMO) control design.

More recently, two degree-of-freedom designs (2DOF) that combine the feedforward and feedback strategies have been reported [30, 31, 32]. In [32], 2DOF design schemes are discussed where iterative learning control (ILC) schems are used along with the feedback control to make improvements on the feedback-only design. In [30], a polynomial-based feedback controller is designed to account for pole-location uncertainty along with inversion-based feedforward controller which provides improved tracking for raster-scan applications. However, a systematic study of the 2 DOF controllers with respect to nanopositioning, which aims at simultaneously achieving multi performance objectives in an optimal-control setting, is lacking. This thesis addresses this gap.

2.1 Performance Criteria and Limitations

A typical nanopositioning system used in SPM comprises a flexure stage, actuators (typically piezoelectric) and/or sensors along with the feedback system. One of the important contributions of control-systems theory to the design of nanopositioning systems is the quantification of performance objectives and fundamental design limitations. I present the analysis and design in terms of block diagrams as shown in Figure 2.1. In this figure, G is the transfer function of the *scanner* comprising of the actuator, the flexure stage, and the sensor. It represents the dynamical relationship between its output, the flexure stage displacement y (which is scaled by a sensor constant), and its input, the voltage u given to the actuator. The signal r represents the command signal that the positioning system needs to track, the disturbance signal d represents the *mechanical noise*- the effects of dynamics that are not incorporated in the model G, n represents the sensor noise, $y_m = y + n$ represents the noisy measurement signal, and the transfer function K represents the control transfer function.

The main objective for the design of the controller K is to make the *tracking error* small, that is to make the difference $r - y$ between the desired and actual motions small.
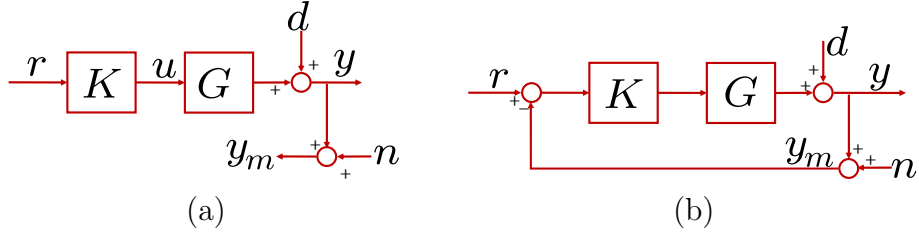
Figure 2.1: Block diagram schematics for nanopositioning systems: The controller acts only on the reference input signal in open-loop nanopositioning systems shown in (a) while it has access to the difference between the reference and the position y in the closed-loop (feedback-only) positioning system as shown in (b).

In open-loop positioning systems (Figure 2.1(a)), where the sensor signal is not fed back to the controller, the performance is severely limited by mechanical noise. The mechanical noise mainly consists of the slowly varying drift and creep, which are therefore prominent in slow scans, and the inertial lag at high frequencies, which is prominent in high speed scans. Hysteresis affects the systems at all frequencies and is particularly prominent in repetitive raster scanning. These nonlinear effects of drift, creep, and hysteresis are sensitive to changes in operating conditions such as ambient temperature, residual polarization in piezo-actuators, and most importantly the operating point, the reference value on the nonlinear input-output (input voltage versus stage displacement) graph about which stage motions are calibrated. The inclusion of their precise behavior in device models is practically infeasible. Feedback based schemes (Figure 2.1(b)) have demonstrated effective compensation for the creep, drift, hysteresis, and inertial lag problems without requiring their precise models [26]. They compensate for the mechanical noise but at the cost of feeding back the relatively smaller sensor noise. In the following, we exclude the consideration of open loop systems and analyze systems that have feedback along with feedforward components. An advantage of block diagram schematics, such as in Figure 2.1(b), is that objectives and limitations on the positioning systems can be quantified in terms of the transfer functions for closed-loop systems. We first present objectives and limitations in feedback-only designs and then present the 2DOF framework.

2.1.1 Performance characterization and control objectives

The performance of a nanopositioning system is characterized by its positioning resolution, tracking bandwidth, and robustness to modeling uncertainties. The resolution of the nanopositioning system is specified in terms of the standard deviation σ of the sensor output when there is no actuation of the positioning stage. The measurement noise typically exhibits a zero-mean Gaussian distribution. Thus, 3σ-resolution defined by the measurement noise gives over 99.7% confidence in any signal value that is greater than the resolution. The tracking bandwidth is the range of reference-signal frequencies that the nanopositioning system can track with a given precision. To characterize the robustness to modeling uncertainties, a metric is needed that quantifies how insensitive the closed-loop device is to the modeling errors and operating conditions.

In feedback-only (1DOF) configuration (Figure 2.1(b)), these performance specifications can be quantified by analyzing the tracking error. For a given controller K, the tracking error in this configuration is given by

$$e \;\; = \;\; r - y = S(r - d) + Tn, \tag{2.1}$$

where the *sensitivity transfer function* $S = (I + GK)^{-1}$ and the *complementary sensitivity transfer function* $T = I - S = (I + GK)^{-1}GK$. Thus, the tracking error, which is given by the superposition of the error contribution Sr from the reference signal r and the error contribution Tn from the sensor noise n, is adversely affected by large values of scan range of r and the amplitude of noise n. Thus, high resolution can be achieved by designing the feedback law $K(s)$ such that $S(s)$ and $T(s)$ are small in those frequency ranges where the frequency contents of r and n, respectively, are large. The resolution of the closed-loop positioning system is determined by the term Tn and therefore lower values of $T(s)$ over larger ranges of frequencies guarantee better resolutions. More specifically, the standard deviation σ of the zero-mean position signal when reference signal is identically zero, which

determines the resolution of the positioning system, is given by

$$\sigma = \int_0^\infty |T(j\omega)|^2 P_n(\omega)d\omega, \qquad (2.2)$$

where $P_n(\omega)$ denotes the power spectral density of the noise signal n. Thus smaller the bandwidth of T, which is characterized by the roll-off frequency ω_T (Figure 2.2(a)), smaller the standard deviation σ, and hence better the resolution of the closed-loop device.

(a)   (b)
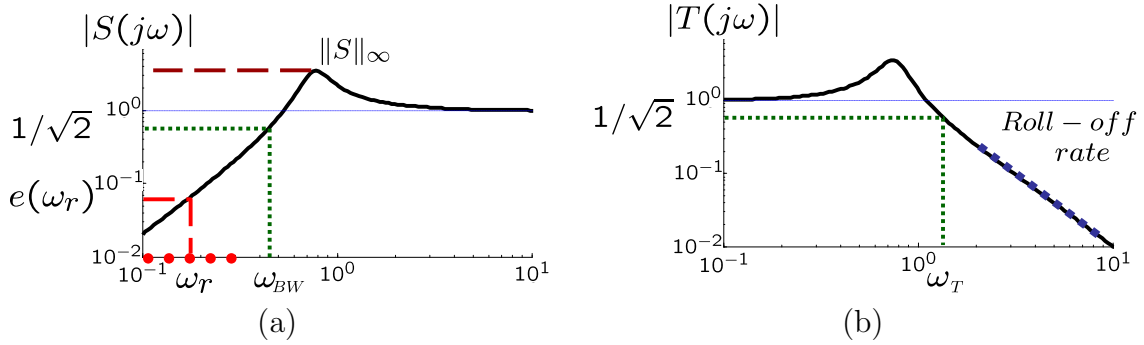
Figure 2.2: Objectives of feedback control: The objective of the control design is to achieve high positioning resolution (by achieving low values ω_T, high roll-off rates in (a) and small error at reference frequencies $e(\omega_r)$ in (b)), high tracking bandwidth (by designing for large ω_{BW}) and robustness (by achieving $\|S\|_\infty$ close to 1).

Similarly, in this configuration, the tracking bandwidth is determined by the range of frequencies over which the magnitude plot of the sensitivity transfer function $S(s)$ is below $1/\sqrt{2}$. It is characterized by the corner frequency ω_{BW} as shown in Figure 2.2(b). Note that unlike typical usage, we use ω_{BW} and not ω_T as a measure of bandwidth. Even though ω_T gives a higher numerical value for bandwidth for the same device, it is poorly related to actual time constants of the device [33].

The sensitivity function $S(s)$ is so called since it is equal to $\frac{dT/T}{dG/G}$, that is it gives the percentage change in the position signal for a given percentage change in the model; therefore it serves as a measure of 'sensitivity' or robustness of the closed-loop loop device to modeling uncertainties. We use the peak value of the magnitude of sensitivity function, $\|S\|_\infty$ to characterize the robustness of the system to modeling uncertainties and operating conditions (Figure 2.2(b)). $\|S\|_\infty$ corresponds to how stable the system is since it represent the distance

from -1 to open loop transfer function in Nyquist plot. Low (near 1) values of $\|S\|_\infty$ guarantee good gain margin and phase margin since $GM \geq \frac{\|S\|_\infty}{\|S\|_\infty - 1}$ and $PM \geq 2 \sin^{-1} \frac{1}{\|S\|_\infty}$ [33]. More on robustness will be discussed in Chapter 4.

The performance specifications translate to control design objectives of achieving high values of ω_{BW} for high tracking bandwidth, high roll-off rates of $|T|$ and smaller values of ω_T for better positioning resolution, and low values of $\|S\|_\infty$ for better robustness to modeling uncertainties.

2.1.2 Limitations

These objectives have to be achieved under some practical and fundamental algebraic limitations. For instance, it is impractical to have the sampling period faster than the time required by digital signal processor (DSP) to calculate steps in control logic. Thus, both high frequency sampling and high order control logic can not be implemented simultaneously. Similarly, to avoid errors due to digital implementation of control laws, sampling frequencies around thirty times the frequency range of interest are typically prescribed [34]. Another important practical limitation is that the actuation signal should be within saturation limits of the hardware.

The algebraic limitations on the control design of the positioning systems are *fundamental* since they are completely determined by scanner G and are independent of the controller K. For instance, in feedback-only systems $S(s) + T(s) \equiv 1$, which is evident from definitions of S and T. This algebraic limitation prevents error $e = r - y = Sr - Sd + Tn$ from becoming small in all frequencies since $|S|$ and $|T|$ can not be made small simultaneously. This motivates search for control designs that achieve a trade-off between the bandwidth and resolution requirements. In addition, for scanner G with phase margin less than 90 deg, which is true for most practical systems, it can be shown that the bandwidth ω_{BW} cannot be larger than ω_T [33]. This limitation prevents the feedback control to achieve noise attenuation over target reference frequency range. Another fundamental limitation that imposes a trade-off between the bandwidth, the resolution and the robustness requirements can be explained in terms of the Bode integral law [35, 36]. This law states that for any stable system G such

that the relative degree of the transfer function $K(s)G(s)$ is at least two,

$$\int_0^\infty \log |S(j\omega)| d\omega = 0. \tag{2.3}$$

The condition on the relative degree, which is the difference of the orders of the denominator and numerator polynomials in the transfer function [33], is typically satisfied. This is so since T needs a sufficiently fast roll-off rate at high frequencies for noise attenuation (and therefore better resolution), the open loop transfer function $K(s)G(s)$ is designed such that it has greater than or equal to relative degree of order 2. Also when a discrete system (or hybrid system-discrete control with analog system) is used, this relative degree condition is inherently satisfied [37].
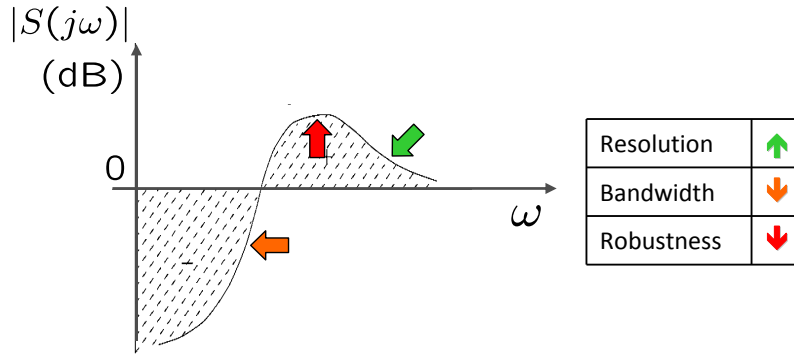
Figure 2.3: Trade-offs due to finite-waterbed effect: The Bode integral laws manifest themselves as waterbed effects, where decreasing the magnitude of sensitivity function at a certain frequency range results in its increase in some other frequency range. For instance, making sensitivity function small (near to 1) at high frequencies to ensure a high roll-off rate of the complementary sensitivity function (since $T = 1 - S$) for better resolution results in lower robustness to modeling uncertainties (due to higher values of the peak ($\|S\|_\infty$)) and lower values of tracking bandwidth (since ω_{BW} decreases). Similar trade-offs where one performance objective is sacrifices at the cost of others can be analyzed by studying the finite-water-bed effect.

The limitations from this algebraic law can be explained in terms of the *waterbed effect* - Since the area under the graph of $\log |S(j\omega)|$ over the entire frequency range is zero, $|S|$ made small at a frequency range has to be compensated by making it large at some other frequency ranges. One direct consequence of this law is that $|S|$ cannot be made less than 1 over all frequencies. Therefore $\|S\|_\infty$, the measure for robustness is at least 1. Moreover, if

the stable system G has real non-minimum phase zeros (that is real and positive roots), a stricter fundamental algebraic law holds

$$\int_0^\infty \log |S(j\omega)| \, W(z,\omega)d\omega = 0, \tag{2.4}$$

where $W(z,\omega) = \frac{2z}{z^2+\omega^2}$ for real positive zero z [36]. Typical scanner systems have non-collocated actuators and sensors that are separated by flexure stages. The transfer function models of such systems generally exhibit non-minimum phase zeros.

In this case it can be shown that the integral of $\log |S(j\omega)|$ over a *finite* frequency region can be bounded from below, thus manifesting a waterbed effect over a finite frequency range. This can be seen from the following analysis. Since the relative degree is greater than two, we have a $\omega_1 > 1$ such that

$$|G(j\omega)K(j\omega)| < \frac{1}{\omega^2} \text{ for all } \omega > \omega_1. \tag{2.5}$$

This is one way of saying that the loop bandwidth is less than ω_1. Then for $\omega > \omega_1$,

$$|(S(j\omega)| < \frac{1}{1 - |G(j\omega)K(j\omega)|} \leq \frac{\omega^2}{\omega^2 - 1}, \tag{2.6}$$

and hence

$$\int_{\omega_1}^\infty \log(|S(j\omega|)d\omega < \int_{\omega_1}^\infty \log(\frac{1}{1 - \omega^{-2}}) < \infty. \tag{2.7}$$

Thus, the simultaneous requirements of low $|S|$ over a large frequency for a high tracking bandwidth, high order roll-off rates of $|T|$ at high frequencies for high resolution and small peaks of the $|S|$ compete against each other under this limitation. For instance, small $|S|$ over a specified bandwidth might not leave out enough frequency range to zero out the area in the 'finite waterbed effect' even with $|S|$ at the allowed peak value for the remaining frequencies (see Figure 2.3).

2.2 Optimal Control Framework for Nanopositioning

The algebraic and practical limitations on the control design severely restrict the space of achievable performance specifications. The model-free based designs (such as proportional-integral-derivative (PID) designs) that are typically used in nanopositioning industry as well as designs based on loopshaping of the open-loop transfer functions further restrict the achievable space due to their inherent structural limitations. These techniques are inadequate to achieve *simultaneously* the multiple objectives of resolution, bandwidth, and robustness under the design challenges and fundamental limitations described above. The robust optimal-control theory provides an apt framework for control design of nanopositioning systems. In this framework, it is possible to determine if a set of design specifications are feasible, and when feasible the control law K is obtained by posing and solving an optimization problem. The main advantage of using this optimization framework is that it incorporates performance objectives directly into its cost function. This eliminates the tedious task of tuning gains (in trial-and-hit manner) as in the PID designs, where even the exhaustively tuned gains may fail to yield acceptable performance.
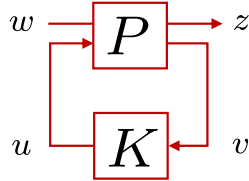
Figure 2.4: General framework for optimal control.

These optimization problems are of the form

$$\min_{K \in \mathcal{K}} \|\Phi(K)\|, \tag{2.8}$$

where $\mathcal{K}$ is the set of stabilizing controllers and Φ is a matrix transfer function whose elements are in terms of closed-loop transfer functions in (2.1) and $\|(\cdot)\|$ represents a metric on transfer functions. The design specifications are interpreted in terms of closed-loop signals z (such as tracking error e in (2.1) and then set Φ as the transfer function from external variables w (such as reference signal r and noise n) to signals z (see Figure 2.4). Demonstration of

this framework and discussion of its advantages are presented through experimental results in Chapter 3.

2.3 Device for Demonstration

For demonstration of the optimal-control framework, the two-dimensional flexure scanner of molecular force probe (MFP-3D) from Asylum Research Inc., Santa Barbara, CA. was used.

2.3.1 Device description
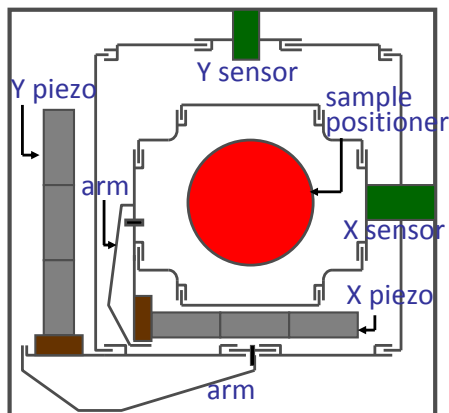
Figure 2.5: Schematic of flexure scanner: The sample is placed on the central block of the flexure stage which is driven by the x-piezo which in turn is driven by the y piezo. The x and y-sensors measure the stage position.

A schematic of the nanopositioning system (the scanner) is shown in Figure 2.5. It has two flexure components with component "X" stacked over the "Y" where the sample-holder is carried by the "X"-component. Each stage, by virtue of the serpentine spring design, deforms under the application of force, providing motion. These forces are generated by stack-piezos. There are three piezo-actuators in series for each axis. The motion of each flexure component is measured by the corresponding nanopositioning sensors which is modified from linear variable differential transformer (LVDT) and the associated demodulation circuit. The piezo-actuators lead to a travel range of 90 μm in close loop in both directions. The nanopositioing

sensors have noise less than 0.6 nm (deviation) over 0.1 to 1 kHz bandwidth.

The control law is discretized and implemented on a Texas instrument TMS320C6713 DSP using code composer studio with 16 bits A/D and 16 bits D/A channels. The setup of control system is shown in Figure 2.6.
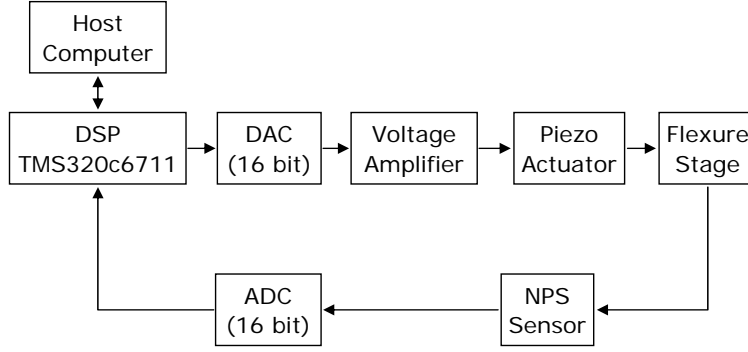
Figure 2.6: Control system setup for flexure scanner.

2.3.2 Identification

Physical modeling of the device is difficult due to its complicated structural design and poorly understood piezo-actuation phenomena and attempts in that direction typically result in complex dynamic models which are not easy to design control for. Therefore, system identification techniques were used to derive linear models about an operating point, where the sensor output gave a 'zero' reading corresponding to 'zero' input to the piezo-actuators. The device is viewed as a two-input two-output system in which the low-voltage signals to the X and Y amplifiers are the inputs and the motion of X and Y flexure-stages components measured by the corresponding sensors are the outputs. This results in four input-output transfer functions G_{ij}, i,j in $\{x,y\}$. Here G_{ij} represents the transfer function from input j to output i. The frequency-response based identification was done where a sine-sweep over a bandwidth from 1 Hz to 2 kHz with amplitude 10 mV was given to each axis using an HP 35670A signal analyzer. From the identification results, X and Y crosstalk represented by G_{xy} and G_{yx} are seen to be relatively small ($\|G_{xy}(j\omega)\|_\infty$ and $\|G_{xy}(j\omega)\|_\infty$ are less than -17.76 dB), which is expected since, by design, X and Y flexure components are decoupled

and are orthogonal to each other. Therefore, the nanopositioning system is modeled by two independent single input single output (SISO) units. The mode of operation of this device is such that higher bandwidth requirements are made on the smaller stage X whereas the Y stage is made to move relatively slow. Hence, there is a greater emphasis on the control designs for the X stage, which is presented in this work.
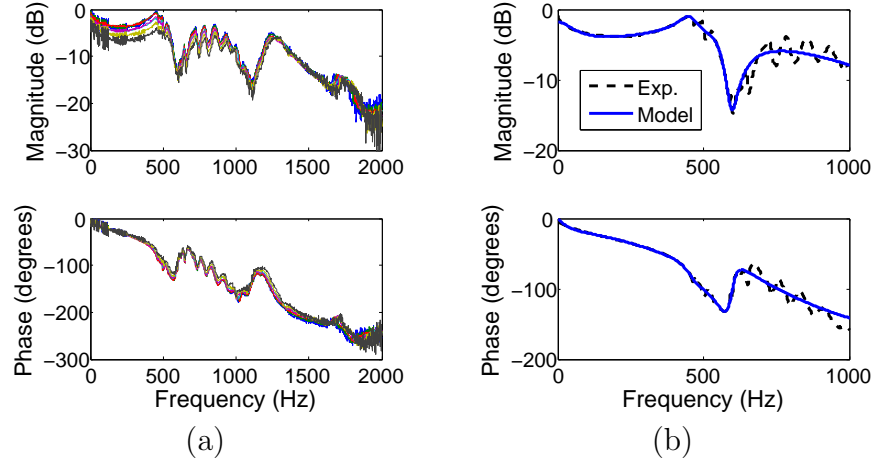
Figure 2.7: Identification of scanner: (a) Experimental frequency responses at various operating positions. (b) Nominal frequency response(dashed) and model frequency response(solid).

This process was repeated to obtain frequency responses of the system at different operating points (by giving various dc offsets) spanning the range of operation of the device (Figure 2.7(a)). The variation in these responses is indicative of the modeling errors (uncertainties) in our identification scheme. In addition, it was observed that the frequency response at the same operating point varies when obtained at different times. In view of these uncertainties, robustness of the closed-loop system is a critical requirement of control design. The nominal frequency response of the system is obtained from averaging on 5 experiments on the nominal operating point which is at dc offset corresponding to 0 V output value. Figure 2.7(b) shows the bode diagram of fitted mathematical model with nominal experimental result. Weighted iterative least square fitting was performed over $0 - 1$ kHz and

the reduction through balanced realization [38] resulted in the following 7th order model,

$$G_{xx}(s) = \frac{-1122.3157(s - 1.152 \times 10^4)(s + 543)}{(s + 390.3)(s^2 + 470.9s + 8.352 \times 10^6)}$$
$$\times \frac{(s^2 + 587.2s + 8.628 \times 10^6)(s^2 + 226.5s + 1.407 \times 10^7)}{(s^2 + 689.4s + 1.315 \times 10^7)(s^2 + 4950s + 2.44 \times 10^7)}. \tag{2.9}$$

This model did not capture dynamics beyond 500 Hz as shown in Figure 2.7(b). Its use is justified since the frequency range of interest is less than 500 Hz and larger models result in implementations of higher order control which cannot be accommodated by the processor with short sampling time. This modeling uncertainty from using low order model was accounted for by imposing the requirement of making the closed-loop system robust to it on the control design.

CHAPTER 3

2DOF CONTROL DESIGN AND EXPERIMENT

2DOF control is the most general feedback scheme where the plant input can be obtained by allowing the controller to process the reference-inputs and the measurements, independently. In control systems literature, the design of 2DOF systems has been topic of much research, which has been resolved using various techniques, each with their unique advantages. For instance, in [39], 2DOF control design is parameterized in terms of two independent stable rational parameters (each related to feedforward and feedback) and analyzed with this parameterization. In [40], 2DOF control design is developed based on parameterization and optimization of the cost functional which include the tracking performance and plant saturation. In [41], the 2DOF control design, with extra emphasis on robustness, is developed based on the integration of the Glover-McFarlane loopshaping [42] and model matching that makes the closed-loop responses close to a specified response.

In this research, we show that the feasible space of performance specifications, which are constrained by the limitations described above in feedback-only configuration, can be extended by using a 2DOF design scheme. In contrast to the feedback-only scheme described in Chapter 2, where the controller acts only on the difference between the reference r and the position-measurement y_m, in 2DOF scheme, the controller acts independently on them. This scheme is implemented in *equivalent* multiple architectures as shown in Figure 3.1.

The feedback-only control scheme is indeed a special case of the 2DOF scheme, where $K_{ff} = 0$ in (a), $K_r = -K_y$ in (b), and $K_{pre} = 1$ in (c) in Figure 3.1. Therefore, 2DOF control design should perform at least as well as the feedback-only design. In this research, we explore to what extent the performance becomes better. In the following, we mainly explain the design in feedforward-feedback structure (where $u = K_{ff}r + K_{fb}(r - y)$, (a) in Figure 3.1) for convenience. In 2DOF scheme, the robustness to modeling uncertainties as well as
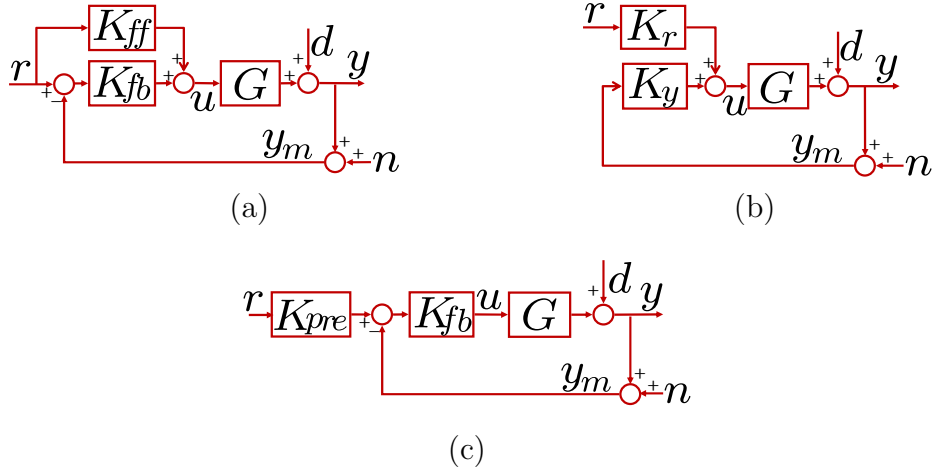
Figure 3.1: 2DOF control architectures: (a) The feedforward-feedback scheme where the actuation signal $u = K_{ff}r + K_{fb}(r - y_m)$, (b) Another scheme where $u = K_r r + K_y y_m$, and (c) Prefilter architecture where $u = K_{fb}(K_{pre}r - y_m)$. The schemes (a) and (b) are equivalent as control designs in that one can be retrieved exactly in terms of the other. Practical implementable designs for controllers in (a) and (b) can easily be derived from control design in (c), however the vice-versa may require certain factorization procedures.

resolution of the device are determined only by the feedback part of the controller, that is the transfer function from d to y that characterizes robustness to modeling uncertainties is still determined by the sensitivity function $S = (I + GK_{fb})^{-1}$, and the transfer function from n to y that characterizes resolution is still determined by the complementary sensitivity function $T = (I + GK_{fb})^{-1}GK_{fb}$. The main difference and advantage in 2DOF control design compared to the feedback-only design stems from the fact that the transfer functions from r to y and from n to y are different and can be designed independently. This difference gives greater independence in designing for better trade-offs between different performance objectives. We use T_{yr} and S_{er} to denote the transfer function from r to y and from r to e, respectively, that is $S_{er} = S(T - GK_{ff})$, $T_{yr} = SG(K_{ff} + K_{fb})$. In this notation, the relevant closed-loop signals are given by

$$
\begin{aligned}
\text{position:} \quad y &= T_{yr}r - Tn + Sd, \\
\text{tracking error:} \quad e &= S_{er}r + Tn - Sd, \\
\text{control (actuation) signal:} \quad u &= S(K_{ff} + K_{fb})r - SK_{fb}n - SK_{fb}d.
\end{aligned}
\tag{3.1}
$$

The performance objectives in tracking problems are characterized in terms of error e in (3.1). Small tracking error can be achieved by designing that make S_{er}, T and S small in those frequency ranges where the frequency contents of r, n and d are dominant. The control objectives translate to small roll-off frequency as well as high roll-off rates for $|T|$ to have good resolution, long range of frequencies for which $|S_{er}|$ is small to achieve large bandwidth and low (near 1) values of the peak in the magnitude plot of S for robustness to modeling uncertainties. Even though the 2DOF control design has greater flexibility than the feedback-only design, the main challenges to design still arise from practical and algebraic (albeit fewer) limitations. The constraints on hardware implementation in terms of sampling frequencies as well as saturation limits of actuation signals limit the scope of this design too. Similarly, the algebraic limitations constrain the control design in this setting too; for instance, the constraint $S(s) + T(s) \equiv 1$ and Bode integral law have the same ramifications on the trade-off between the resolution and robustness to modeling uncertainties as in the feedback-only design.

3.1 Optimal Prefilter Model Matching Design

3.1.1 Control design

In some positioning systems, there is a pre-designed feedback component K_{fb} which cannot be replaced or changed (for instance, some commercial scanners come with feedback components designed to accomplish specific tasks such as raster scanning). However, typically there are no such restrictions on the feedforward control design since it can be easily implemented as a prefilter on the reference signal. In the design presented here, the feedforward component K_{pre} is so chosen that the closed-loop positioning system mimics a *target* transfer function T_{ref} (Figure 3.2). This target transfer function T_{ref} is chosen so that it satisfies desired performance objectives. An advantage of using such model matching schemes is that desired transient characteristics (such as settling times and overshoots) can be incorporated by choosing appropriate model T_{ref}, and since the closed-loop device is designed to mimic the model, it inherits the transient characteristics too. After noting that the closed-loop

device transfer function from r to y is given by TK_{pre}, the feedforward component K_{pre} is chosen by solving an optimization problem such that the $\mathcal{H}_\infty$-norm of the mismatch transfer function $E = T_{ref} - TK_{pre}$ is minimized. Small values of $\|E\|_\infty$ guarantees small values for mismatch error signal (see Figure 3.2) given by

$$e = T_{ref}r - y = (T_{ref} - TK_{pre})r. \tag{3.2}$$

To ensure practical implementation, it is assumed that T_{ref} and T are stable, proper transfer functions.
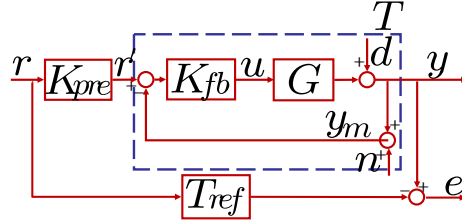
Figure 3.2: Model matching through prefilter problem.

Note that this optimization problem is trivial if T is a minimum phase transfer function, that is if it has only stable zeros. In this case T^{-1} is stable and the solution K_{pre} can be easily obtained as $T^{-1}T_{ref}$. However, typical nanopositioning systems are flexure-based with non-collocated actuators and sensors, which typically manifest as non-minimum phase zeros of T. In this case, the optimal solution can be found by applying Nevanlinna-Pick theory [43] as follows.

The model matching problem is equivalent to finding minimum γ such that $\|T_{ref} - TK_{pre}\|_\infty \leq \gamma$, where the minimum $\gamma = \gamma_{opt}$ is achieved for some stable K_{pre}. If we define $E_\gamma = \frac{1}{\gamma}(T_{ref} - TK_{pre})$ for $\gamma > 0$, then this problem can be restated as finding a stable K_{pre} which requires $\|E_\gamma\|_\infty \leq 1$. Note that, for stable K_{pre}, E_γ satisfies the interpolating conditions $E_\gamma(z_i) = \frac{1}{\gamma}T_{ref}(z_i)$ for every non-minimum phase zero z_i of the scanner G. Therefore, we can cast this as a Nevanlinna-Pick (NP) interpolation problem. The NP problem is to find a function E_γ in the space of stable, complex-rational functions satisfying the two

conditions

$$\|E_\gamma\|_\infty \leq 1, \tag{3.3}$$

$$\{(z_i, E_\gamma(z_i))\}_{i=1}^n. \tag{3.4}$$

Moreover, it can be shown that γ_{opt} is equal to the square root of the largest eigenvalue of the matrix $A^{-\frac{1}{2}}BA^{-\frac{1}{2}}$ where the elements of matrix A and B are respectively [43]

$$a_{ij} = \frac{1}{z_i + \bar{z}_j}, \quad b_{ij} = \frac{b_i \bar{b}_j}{z_i + \bar{z}_j}. \tag{3.5}$$

The prefilter is given by

$$K_{pre} = T^{-1}(T_{ref} - \gamma_{opt}E_\gamma). \tag{3.6}$$

Note that this method requires the additional steps for better performance. First, for better tracking, K_{pre} need to be scaled so that closed-loop function matches the target transfer function at 0 Hz (i.e. the dc gain match). Second, since the obtained K_{pre} may not be proper, a weight function W_0 which is a low pass filter needs to be multiplied so that the $W_0 K_{pre}$ becomes proper.

3.1.2 Implementation of the controller

For the purpose of illustration, we designed the feedback "pre-existing controller" by a feedback-only design using $\mathcal{H}_\infty$ optimal framework which achieves much better bandwidth (over sixty times) when compared to PI/PID designs [26]. Therefore, improvements resulting on application of our model matching designs are even more significant when applied to typical scanners that commonly use PI/PID based feedback controllers. A 9th order feedback controller K_{fb} was obtained as a result of a mixed sensitivity $\mathcal{H}_\infty$ optimization for feedback-

only designs given by

$$K_{fb} = \frac{872851.2498(s + 3.142 \times 10^5)(s^2 + 470.9s + 8.352 \times 10^6)}{(s + 1.123 \times 10^7)(s + 7.588 \times 10^4)(s + 1.261 \times 10^4)(s + 543.5)}$$
$$\times \frac{(s + 390.3)(s^2 + 689.4s + 1.315 \times 10^7)(s^2 + 4950s + 2.44 \times 10^7)}{(s + 0.03142)(s^2 + 587.1s + 8.628 \times 10^6)(s^2 + 226.7s + 1.407 \times 10^7)}. \qquad (3.7)$$

This feedback-only design yielded a bandwidth of 49.4 Hz, the roll-off frequency of 60.1 Hz, and $\|S\|_\infty$ of 1.15 for the closed-loop device (represented by solid lines in Figure 3.3).

The prefilter is designed in pursuit of matching $K_{pre}T$ to $T_{ref} = \frac{1}{0.0003s+1}$ by using the Nevalinna-Pick solution. The order of the prefilter was reduced and the prefilter is given in the following control law,

$$K_{pre} = \frac{2.0289 \times 10^{-10}(s + 1.123 \times 10^7)(s + 7.626 \times 10^4)}{(s + 3.142 \times 10^5)}$$
$$\times \frac{(s + 1.18 \times 10^4)(s + 544.6)(s + 349.9)}{(s + 2857)(s + 543)}. \qquad (3.8)$$

Before directly applying this control law, it is modified to account for steady state gain matching and multiplication of weighting function to ensure realizable transfer function as previously discussed. The control law from NP solution is improper and has relative order of degree -2, and therefore we multiplied by the weight function $W_0 = \frac{1}{(1 \times 10^{-4}s+1)^2}$ such that it becomes proper. The scaling factor 1.25 is also multiplied to $W_0 K_{pre}$ for the dc gain match.

Figure 3.3 shows the experimentally obtained transfer function from reference to error, i.e. S from feedback control and S_{er} from 2DOF control which represent the tracking performance (ω_{BW}=49.4 Hz (feedback), =214.5 Hz (2DOF)) and the transfer function from reference to output i.e. T of feedback control and T_{yr} of 2DOF control. This 2DOF design yields an improvement of over 330% in bandwidth over the feedback-only design. Since the feedback component of the two designs are the same and this component completely determines the robustness to modeling errors (characterized by $\|S\|_\infty$) as well as the positioning resolution (characterized by ω_T), the resolution and robustness remains the same for both the designs. An interesting observation is that the tracking-bandwidth ω_{BW} is greater than the roll-of
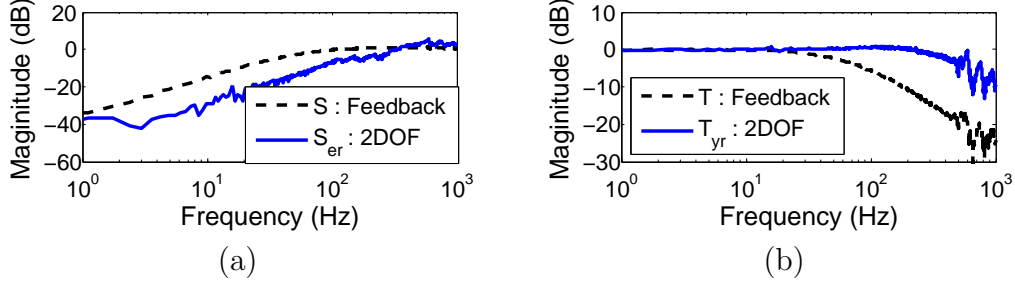
Figure 3.3: Comparison of experimentally obtained magnitude of $S(s)$ and $T(s)$ from $\mathcal{H}_\infty$ feedback-only control design(dashed) with $S_{er}(s)$ and $T_{yr}(s)$ from prefilter model matching 2DOF control(solid). The feedforward controller designed using prefilter model matching design achieves over 330% improvement in the tracking bandwidth of the closed-loop design. The robustness and resolution are determined by the feedback components S and T, and therefore remain the same for the two cases.

frequency $\omega_T = 60.1$ Hz which is impossible for feedback-only designs [33].

3.2 2DOF Optimal Robust Model Matching Design

3.2.1 Control design

In some nanopositioning systems, the issue of robustness to modeling uncertainties is critical. The systems with pre-designed feedback control have satisfactory performance (resolution and bandwidth) when experiments are conducted very carefully 'near nominal' operating conditions. However, there is rapid degradation of performance (sometimes even become unstable) when the operating conditions deviate from the nominal conditions. In [27], control designs based on Glover-McFarlane method [42, 44] that wrapped around pre-existing controllers were implemented that resulted in substantial improvements in robustness. Here, we present a control design where we pose and solve an optimal control problem that simultaneously designs a wrap-around feedback controller for robustness as well as the feedforward controller for better bandwidth. We use a 2DOF control design developed in [41, 45] that uses model matching where the transfer function of the closed-loop device is made to match a pre-specified target transfer function T_{ref}.

In this design formulation, we represent the modified scanner system by $G_s = GK_s$ where

27
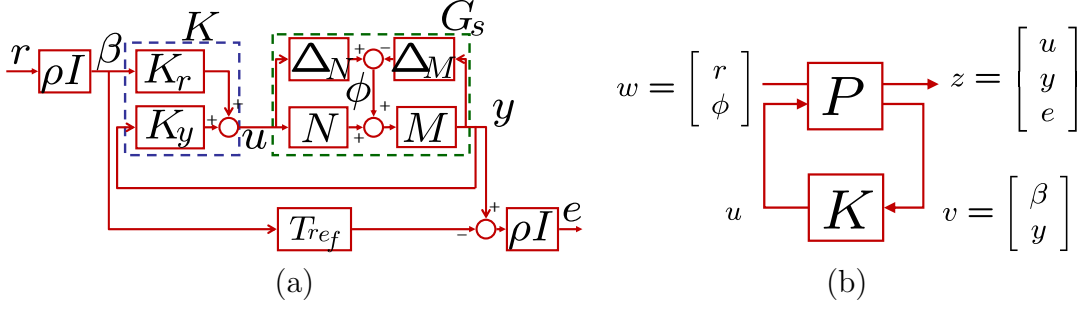
Figure 3.4: Control design for model matching and robustness to modeling errors: (a) The signals u, y and e that represent the control signal, the noise component in the position signal, and the mismatch error are chosen as regulated variables z. The transfer function from the reference signal r to these regulated variables z reflects the performance objectives of bounded control signals, noise attenuation, as well as the model matching. The transfer function from the effects of modeling error ϕ to z represents the effect of modeling errors (unmodeled dynamics) on performance objectives. (b) To achieve robust performance, a control design $K = [K_r \; K_y]$ which minimizes the $\mathcal{H}_\infty$-norm of the transfer function from w to z is sought through the optimal control problem.

G represents the scanner and K_s represents the pre-existing controller. The corresponding 2DOF optimal control problem can be explained in terms of Figure 3.4. The optimization routine seeks $K = [K_r \; K_y]$ such that the closed-loop system guarantees 'optimal' robustness to modeling uncertainties as well as minimizes the mismatch between the transfer function from r to y and a target transfer function T_{ref}. The condition for robustness is imposed by requiring the controller to guarantee stability for a set of transfer function models that are 'close' to the nominal model G_s. More specifically, the optimal control problem is cast in such a way that the resulting controller guarantees the stability of the closed-loop positioning system where the shaped scanner system is represented by *any* transfer function G_p in the following set

$$\{G_p | G_p = (M - \Delta_M)^{-1}(N + \Delta_N), \text{ where } \|[\Delta_M \; \Delta_N]\|_\infty \le \gamma^{-1}\}, \tag{3.9}$$

where $G_s = M^{-1}N$ is a coprime factorization [39], $[\Delta_M \; \Delta_N]$ represents the uncertain (un-modeled) dynamics, and γ specifies a bound on this uncertainty.

This characterization of robustness is specifically apt for many nanopositioning systems which typically have very low damping coefficients (< 0.01 are common). Since the set

described by G_p can have marginally stable/unstable systems, the critical robustness issue is adequately incorporated in this design. The model matching problem is incorporated by seeking the control design K such that the mismatch between the closed-loop transfer function and T_{ref} is minimized. The desired closed-loop transfer function T_{ref} is selected to satisfy desired response characteristics. The parameter ρ determines the emphasis between model matching and robustification in optimization. (If ρ is 0 then the objective becomes same as the Glover-McFarlane design problem).

The optimal control problem is cast in the $\mathcal{H}_\infty$ optimal control set up. More specifically, the signals $z = [u^T \ y^T \ e^T]^T$ are chosen as regulated variables and the controller K is obtained which minimizes the $\mathcal{H}_\infty$ norm of transfer function Φ_{zw} from $w = [r^T \ \phi^T]^T$ to z (see Figure 3.4) described by

$$
\begin{bmatrix} u \\ y \\ e \end{bmatrix} = \underbrace{\left[\begin{array}{c|c} \rho K_r S & K_y S M^{-1} \\ \rho G_s K_r S & S M^{-1} \\ \hline \rho^2 (G_s K_r S - M_0) & \rho S M^{-1} \end{array} \right]}_{=\Phi(K)} \begin{bmatrix} r \\ \phi \end{bmatrix}, \tag{3.10}
$$

where $S = (I - G_s K_y)^{-1}$. Here the exogenous signal ϕ represents a disturbance signal that is the effects of unmodeled dynamics. Thus by making the norm on matrix transfer function small, we make the closed-loop device insensitive to the effects of unmodeled dynamics; since smaller the $\|\Phi(K)\|_\infty$, the smaller is effect of the disturbance ϕ on the mismatch signal e. This implies that mismatch signal has been made insensitive to the modeling errors. The signals u and y are incorporated as regulated variables in the optimization problem to account for control-saturation constraint and for noise attenuation, respectively.

Note that an additional step is required to improve the tracking performance. As a final refinement for the tracking problem, K_r needs to be scaled so that closed-loop transfer function matches the reference transfer function at steady state problem as in optimal prefiter model matching design. A scale constant W_0 is defined as $W_0 = S(s)[G_s(s)K_y(s)]^{-1} T_{ref}|_{s=0}$ and the resulting controller becomes $K = [K_r W_0 \quad K_y]$.

3.2.2 Implementation of the controller

Proportional-integral (PI) and proportional-integral-integral (PII) controllers are the most common controllers currently used in commercial scanning-probe microscopes. Their popularity stems from the fact that they are simple to implement, and easy for users to develop a feel for tuning them. Moreover, PII controllers track ramp signals, which partially represent triangular-raster scan, with zero steady-state error. The PII controller has the structure $K_{PII} = k_p + \frac{k_i}{s} + \frac{k_{ii}}{s^2}$. After an exhaustive search over the space of controller parameters, $k_p = 0$ was chosen to get the roll-off of 2 for high frequency noise attenuation, and $k_i = 1.1 \times 10^3$ and $k_{ii} = 6.2 \times 10^4$ were chosen to meet the bandwidth and robustness requirements. Experiments with the resulting controller showed that the bandwidth of closed-loop device is 81 Hz and the roll-off frequency ω_T at 235 Hz, and $\|S\|_\infty$ is 1.54 (Figure 3.5 shows the experiment results for this PII-based feedback-only design). The design outlined in Section 3.2 was applied to G_{xx} with $K_s = K_{PII}$, $\rho = 3$, $T_{ref} = \frac{1}{0.0003s+1}$. The 2DOF controller was obtained as

$$
\begin{aligned}
K_r &= \frac{7.762 \times 10^9 (s + 56.47)(s + 56.36)(s^2 + 1116s + 3.445 \times 10^5)}{s^2(s + 8.286 \times 10^5)(s + 3333)(s + 541.4)(s + 56.36)} \\
&\times \frac{(s^2 + 479.2s + 8.373 \times 10^6)(s^2 + 678.4s + 1.318 \times 10^7)(s^2 + 4962s + 2.461 \times 10^7)}{(s^2 + 599.7s + 8.553 \times 10^6)(s^2 + 493.8s + 1.458 \times 10^7)(s^2 + 8440s + 4.552 \times 10^7)}, \\
K_y &= \frac{-1.992 \times 10^9 (s + 453.8)(s + 56.36)(s + 52.33)(s^2 + 454.9s + 8.264 \times 10^6)}{s^2(s + 8.286 \times 10^5)(s + 541.4)(s + 56.36)(s^2 + 599.7s + 8.553 \times 10^6)} \\
&\times \frac{(s^2 + 713.9s + 1.305 \times 10^7)(s^2 + 4920s + 2.411 \times 10^7)}{(s^2 + 493.8s + 1.458 \times 10^7)(s^2 + 8440s + 4.552 \times 10^7)}.
\end{aligned}
\tag{3.11}
$$

Figure 3.5(a) compares the experimentally obtained response from reference to error i.e. S from feedback-only control and S_{er} from 2DOF control which represent the tracking performance. There is over 64% improvement in tracking bandwidth ($\omega_{BW} = 81$ Hz (feedback), =133 Hz (2DOF)). Figure 3.5(b) compares the response from reference to output i.e. $T(s)$ from feedback control and $T_{yr}(s)$ from 2DOF control. Comparison of $S(j\omega)$ shows the improvement of robustness ($\|S\|_\infty = 1.52$ (feedback) = 1.21 (2DOF)). Note that this design process improves the bandwidth and robustness at the same time for fixed resolution which is generally impossible for feedback-only design due to algebraic fundamental limitations

discussed in Section 2.1.
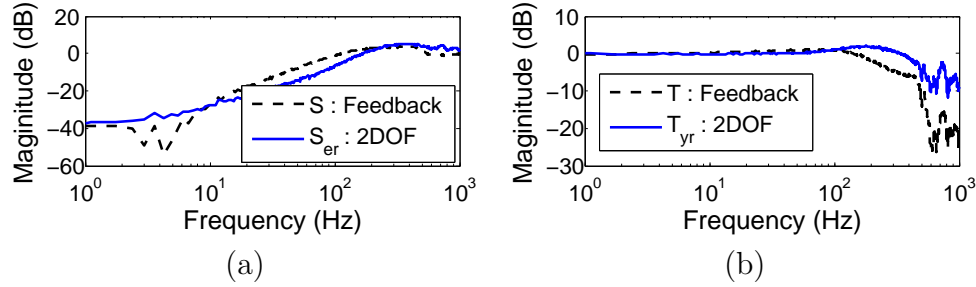
Figure 3.5: Comparison of experimentally obtained magnitude of $S(s)$ and $T(s)$ from PII feedback-only control design(dashed) with $S_{er}(s)$ and $T_{yr}(s)$ from 2DOF optimal robust model matching control(solid). The 2DOF optimal robust model matching controller achieves over 64% improvement in the tracking bandwidth.

The design achieves robustification and model matching simultaneously and wraps around the pre-existing control. Accordingly the closed-loop system with this design achieves better robustness to modeling errors and tracking bandwidth compared to the system with only pre-existing controller. However, if the pre-existing controller results in closed-loop device that is robust to modeling errors but has insufficient bandwidth, then this process increases the bandwidth without adversely affecting the robustness and vice versa [31].

3.3   2DOF Stacked Sensitivity Synthesis

3.3.1   Control design

In this control synthesis scheme, both the feedforward and the feedback control laws are solved in an optimal control setting. In order to reflect the performance objectives and physical constraints, the regulated outputs were chosen to be the weighted tracking error, $z_s = W_s e$, the weighted system output, $z_t = W_t y$, and the weighted control input, $z = W_u u$ (Figure 3.6).

Using (3.1), the closed-loop matrix transfer function from the exogenous variables $w =$
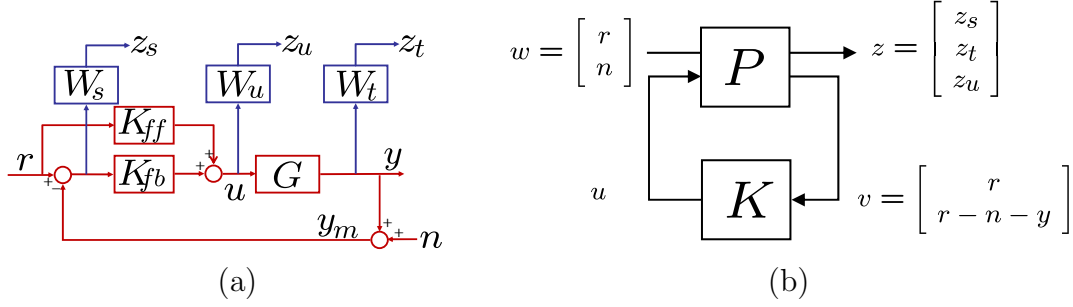
Figure 3.6: Mixed sensitivity problem for 2DOF control design: (a) The signals z_s, z_t, and z_u represent tracking error, the noise component in the position signal, and the control signal. The weights W_s, W_t and W_u are chosen to reflect the design specifications of tracking bandwidth, positioning resolution, and saturation limits on the control signal. (b) To achieve these objectives, a control design $K = [K_{ff}\ K_{fb}]$ which minimizes the $\mathcal{H}_\infty$-norm of the transfer function from $w = [r^T\ n^T]^T$ to $z = [z_s^T\ z_t^T\ z_u^T]^T$ is sought through the optimal control problem.

$[r^T\ n^T]^T$ to the regulated variables $z = [z_s^T\ z_t^T\ z_u^T]^T$ is given by

$$
\begin{bmatrix} z_s \\ z_t \\ z_u \end{bmatrix} = \underbrace{\begin{bmatrix} W_s S_{er} & -W_s S \\ W_t T_{yr} & -W_t T \\ W_u S(K_{ff} + K_{fb}) & -W_u S K_{fb} \end{bmatrix}}_{=\Phi(K)} \begin{bmatrix} r \\ n \end{bmatrix}.
\tag{3.12}
$$

Accordingly, the $\mathcal{H}_\infty$ optimal control problem that we pose is $\min_K \|\Phi(K)\|_\infty$. The minimization of z_s reflects the tracking-bandwidth requirement. If we design the *weight function* $W_s(j\omega)$ to be large in a frequency range of $[0\ \omega_{BW}]$ and ensure that z_s is small over the entire frequency range (through the above optimization problem) then the tracking-error e will be small in the frequency range $[0\ \omega_{BW}]$; that is the closed-loop positioning device has a bandwidth ω_{BW}. Alternatively, note that the transfer function from r to z_s is $W_s S_{er}$. The optimization problem along with our choice of W_s will ensure that the transfer function S_{er} is small in the frequency range $[0\ \omega_{BW}]$. Similarly the transfer function from n to z_t is the weighted complementary sensitivity function $W_t T$, whose minimization ensures better resolution as it forces low control gains at high frequencies, and the transfer function from r to z_u is $W_u S(K_{ff} + K_{fb})$, which measures the control effort. Its minimization reflects in

32

imposing the practical limitation of the control signals to be within saturation limits.

Note that the choice of *weight functions* is crucial in this design. The issue of conflicting requirements on the design of S and T transfer functions arising from competing performance objectives are resolved by achieving compromises (trade-offs) between them by designing appropriate weight transfer functions. The weight functions are used to scale closed-loop transfer functions to specify the frequency information of the performance objectives and system limitations. For instance, by including $\|W_t T\|_\infty$ instead of $\|T\|_\infty$ to be small (for achieving resolution objective) in the optimization problem, where the weight function $W_t(s)$ is designed as a high pass filter, forces the control solution such that $|T(s)|$ is small at high frequencies (and not necessarily at low frequencies). This weight function allows a compromise, for instance, between the resolution and bandwidth requirements. More specifically, if the condition for $\|T\|_\infty$ to be small is imposed instead (i.e. $|T|$ over all frequencies), the limitation from Bode integral law leaves little or no room for achieving robust performance objective.

In (3.12), weight functions $W_s(s)$ and $W_t(s)$ are to be so designed that the optimal control problem enforces simultaneously $|S_{er}|$ to be small in a frequency range that will determine the bandwidth of the positioning system, the peak value $\|S\|_\infty$ be small to ensure robustness to modeling uncertainties, as well as $|T|$ to roll-off at high frequencies to provide better resolution. Note that the optimization problem is severely constrained since we have only these *two* weight functions to shape *four* transfer functions S_{er}, T_{yr}, S and T, which in turn are limited by the algebraic constraints such as $S_{er} + T_{yr} = I$, $S + T = I$. The design of weight function W_u is relatively simpler as it has only one objective of control saturation to account for. In order to alleviate this problem, we introduce two new weight functions W_r and W_n that carry the information content of input signals r and n respectively. Thereby, the transfer matrix for optimization in (3.12) is replaced as

$$\begin{bmatrix} z_s \\ z_t \\ z_u \end{bmatrix} = \begin{bmatrix} W_s S_{er} W_r & -W_s S W_n \\ W_t T_{yr} W_r & -W_t T W_n \\ W_u S(K_{ff} + K_{fb}) W_r & -W_u K_{fb} S W_n \end{bmatrix} \begin{bmatrix} r \\ n \end{bmatrix}. \tag{3.13}$$

Note that inclusion of weights W_r and W_n does alleviate, to an extent, the severity of

constraints on the optimal control problem which requires shaping *four* transfer functions. Even with these additional weighting functions, the shaping of these transfer functions is still not unconstrained since they are not independent even though the input-weighting functions (W_r, W_n) and the output-weighting functions (W_s, W_t) can be chosen independently. The limitation stems from that these weight functions appear as products in (3.13) which reduces the flexibility in their choice.

3.3.2 Implementation of the controller
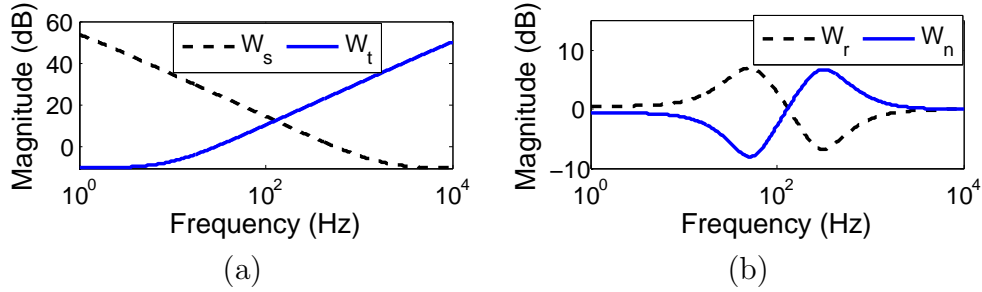
Figure 3.7: Choice of weight functions.

Four weight functions W_r, W_n, W_s, W_t are chosen to shape closed transfer functions: S_{er} with $W_s W_r$, S with $W_s W_n$, T_{yr} with $W_t W_r$, and T with $W_t W_n$. The performance objectives of high bandwidth, high resolution and robustness to modeling errors were reflected as follows. High resolution requires the roll-off frequency of ω_T to be small, that is $|T|$ to be small beyond ω_T. This is imposed by designing the weight frequency $W_t = \frac{1000s + 5.961 \times 10^4}{s + 1.885 \times 10^5}$ to be large at high frequencies (we chose ω_T around 75 Hz). The range of frequencies where $|S|$ is small (required for small tracking error) is restricted to small frequencies since small $|T|$ at high frequencies implies $|S|$ to be near 1 at high frequencies. Thus $W_s = \frac{0.3162s + 3456}{s + 3.456}$ ensures $|S|$ is small at low frequencies and allows for its cross-over frequency to be small enough to make designed ω_T feasible. Figure 3.7 shows the choice of the weight functions that reflect these objectives. The choice of W_r and W_n is made such that at the frequency the W_t starts increasing, W_r starts increasing and W_n starts decreasing (in fact W_n is chosen as inverse of W_r). Note that at the frequency the W_s stop rolls off, W_r and W_n converge to 1. This

choice of W_r and W_n ensures that $|S_{er}|$ is small even when $|S|$ is not small. This is done by exploiting that S is shaped by W_s while S_{er} is shaped by $W_s W_r$. The choice of weight function $W_u = 0.1$ restricted control signal values to be within saturation limits.

The result of $\mathcal{H}_\infty$ synthesis yielded the following control laws and the corresponding closed-loop transfer function are shown in Figure 3.8.

$$
\begin{aligned}
K_{ff} &= \frac{7.520 \times 10^9 (s + 1.885 \times 10^5)(s + 1.288 \times 10^4)(s + 2623)(s + 723.3)}{(s + 1.122 \times 10^9)(s + 3.041 \times 10^6)(s + 1.3 \times 10^4)(s + 3725)(s + 1001)} \\
&\times \frac{(s + 390.3)(s + 275.9)(s + 37.31)(s + 3.413)(s^2 + 470.9s + 8.352 \times 10^6)}{(s + 543)(s + 171)(s + 144)(s + 3.412)(s^2 + 1965s + 3.426 \times 10^6)} \\
&\times \frac{(s^2 + 689.4s + 1.315 \times 10^7)(s^2 + 4950s + 2.44 \times 10^7)}{(s^2 + 587.2s + 8.628 \times 10^6)(s^2 + 226.5s + 1.407 \times 10^7)}, \\
K_{fb} &= \frac{9.008 \times 10^9 (s + 1.885 \times 10^5)}{(s + 1.122 \times 10^9)(s + 3.041 \times 10^6)(s + 1.3 \times 10^4)(s + 3725)(s + 1001)(s + 543)} \\
&\times \frac{(s + 496.9)(s + 390.3)(s + 144)}{(s + 171)(s + 144)(s + 3.412)(s^2 + 1965s + 3.426 \times 10^6)(s^2 + 587.2s + 8.628 \times 10^6)} \\
&\times \frac{(s^2 + 1341s + 1.691 \times 10^6)(s^2 + 1965s + 3.426 \times 10^6)(s^2 + 470.9s + 8.352 \times 10^6)}{(s^2 + 226.5s + 1.407 \times 10^7)(s^2 + 689.4s + 1.315 \times 10^7)(s^2 + 4950s + 2.44 \times 10^7)}.
\end{aligned}
$$

$$(3.14)$$

The feedforward and feedback control laws obtained from $\mathcal{H}_\infty$ mixed sensitivity synthesis procedure were implemented. Figure 3.9(a) shows the experimentally obtained transfer function from reference to error i.e. $S_{er}(s)$ which represents the tracking performance ($\omega_{BW} = $ 148.2 Hz) and the transfer function from disturbance to error i.e. $S(s)$ and Figure 3.9(b) shows the transfer function from reference to output i.e. $T_{yr}(s)$ and the transfer function from noise to output i.e. $T(s)$. Comparing with the roll-off frequency $\omega_T = 75.4$ Hz, the tracking bandwidth ω_{BW} is higher as in optimal prefilter model matching design. There was in improvement of 290% in bandwidth for same values of resolution and robustness if compared to feedback-only design. Similarly, improvement in other performance objectives (resolution and robustness) can be obtained by appropriately designed the weight functions.
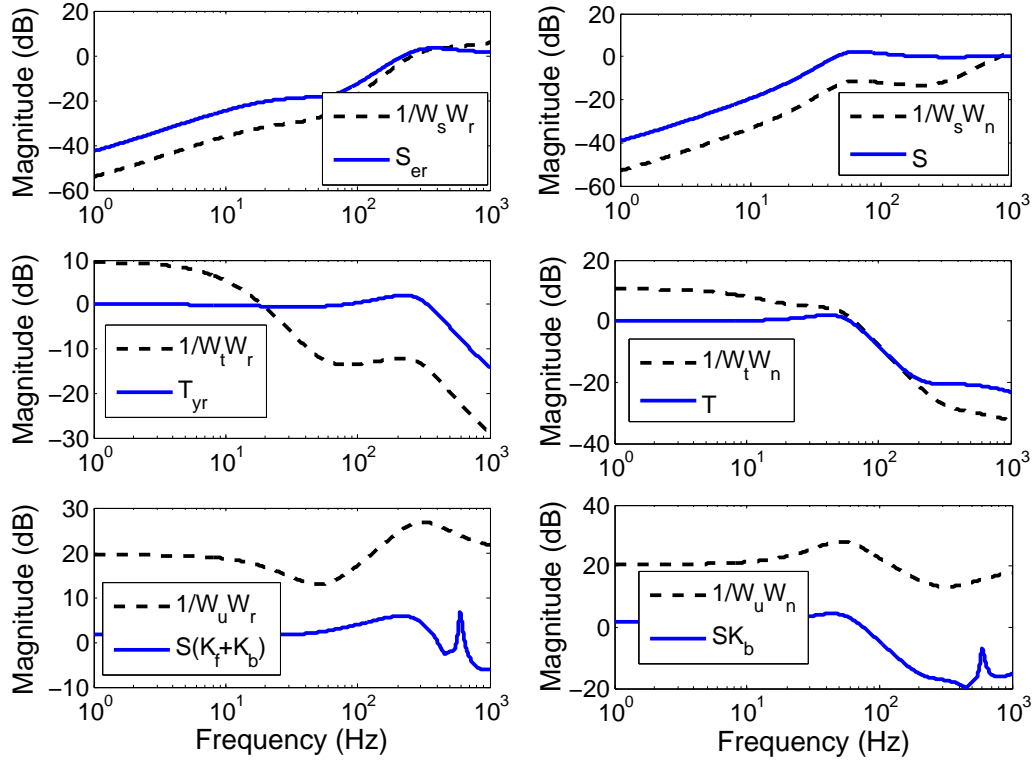
Figure 3.8: Mixed sensitivity synthesis of 2DOF control.
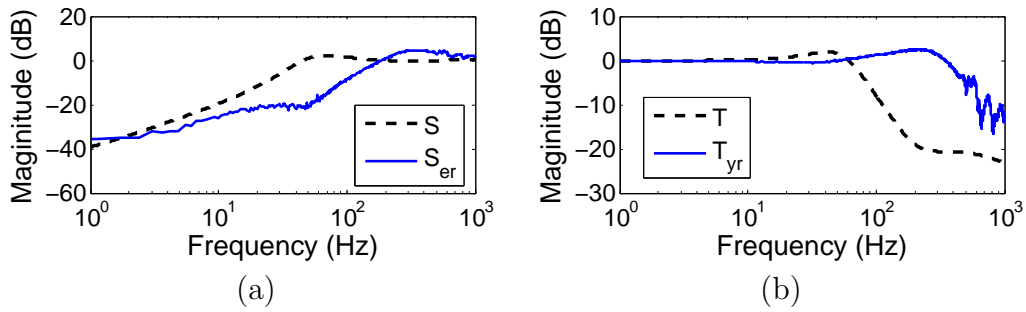
Figure 3.9: Magnitude of $S_{er}(s)$ and $T_{yr}(s)$(solid) obtained from experiment and compared to S and T(dashed).

3.4   2DOF Multi-objective Synthesis Design

3.4.1   Control design

In order to impose the performance objectives into optimal control setting, the configuration which is based on $\mathcal{H}_\infty$ stacked sensitivity framework [46] and model matching framework [43] is considered as shown in Figure 3.10. Main idea in this setup is to shape the closed-loop transfer functions S and T with weighting functions W_s and W_t to achieve robust stability, disturbance rejection and noise attenuation and to make the close-loop response close to a target reference response $T_{ref}r$. The regulated outputs are chosen to be $z_m = T_{ref}r - y$, the deviation from target response, $z_s = W_s(r - n - y)$, the weighted tracking error with noise, and $z_t = W_t y$, the weighted system output.
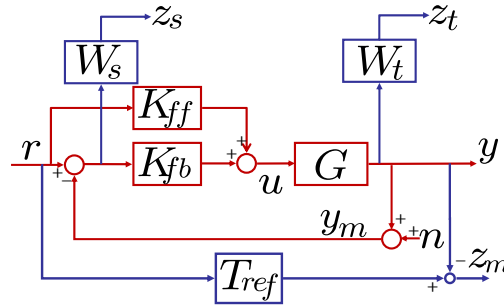
Figure 3.10: Control design for 2DOF stacked sensitivity with model matching design. The signals z_m, z_s, and z_t represent the deviation from target response, the weighted tracking error with noise, and the weighted system output. The weights W_s and W_t are chosen to reflect the design specifications of robust stability, disturbance rejection and noise attenuation. The target response function T_{ref} is chosen to specify the performance of close loop response to reference signal.

Using (3.1), the closed-loop matrix transfer function from the exogenous inputs $w = [r^T \ n^T]^T$ to the regulated outputs $z = [z_m^T \ z_s^T \ z_t^T]^T$ is given by

$$
\begin{bmatrix} z_m \\ z_s \\ z_t \end{bmatrix} = \underbrace{\begin{bmatrix} T_{ref} - T_{yr} & T \\ W_s S_{er} & -W_s S \\ W_t T_{yr} & -W_t T \end{bmatrix}}_{=\Phi(K)} \begin{bmatrix} r \\ n \end{bmatrix}. \tag{3.15}
$$

The minimization of matrix $\Phi(K)$ in (3.15) does not correspond to the design objective since $\Phi(K)$ includes the other redundant terms T, $W_s S_{er}$ and $W_t T_{yr}$ which are not related to design objectives. Also algebraic limitations of $S + T = I$ and $S_{er} + T_{yr} = I$ prevent from reaching satisfactory solution since these constraints severely restrict the feasible space of controllers that make $\|\Phi(K)\|$ small.

In our approach, specifications are directly imposed on certain transfer functions instead of posing the problem in terms of regulated variables. More specifically, we target the transfer functions $T_{ref} - T_{yr}$, $W_s S$ and $W_t T$. This is realizable by the multi-objective framework proposed in [47].

The cost objectives to be minimized are $T_{ref} - T_{yr}$ which is the transfer function from r to z_m and $-[W_s S \ W_t T]^T$ which is the transfer function from n to $[z_s^T \ z_t^T]^T$. The optimization problem is stated as

$$\min_{K \in \mathcal{K}} \rho \, \|T_{ref} - T_{yr}\|_{\alpha_1} + \left\| \begin{array}{c} W_s S \\ W_t T \end{array} \right\|_{\alpha_2}, \tag{3.16}$$

where $K = [K_{ff} \ K_{fb}]$, $\mathcal{K}$ is the set of stabilizing controllers and the parameter ρ reflects the user defined relative emphasis between model matching and robust performance and $\|\cdot\|_{\alpha_i, i \in \{1,2\}}$ are norms (generally 2-norm or ∞-norm) that can possibly be different from each other. Figure 3.11 shows the general framework of multi-objective optimization in (3.16) corresponding to Figure 3.10. This multi-objective problem can be formulated by linear matrix inequalities (LMIs).
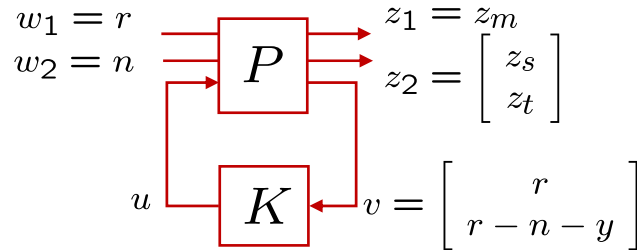
Figure 3.11: General framework for 2DOF control in (3.16): The multi-objective optimization problem is to design a controller that minimizes the sum of ρ times the transfer function from w_1 to z_1 and the transfer function from w_2 to z_2. This is in contrast to stacked sensitivity framework that minimizes the transfer function from $w = [w_1^T \ w_2^T]^T$ to $[z_1^T \ z_2^T]^T$.

3.4.2 Conversion to LMI optimization

This subsection describes the adaptation of the 2DOF control design in the framework developed in [47]. The generalized matrix for the system P in Figure 3.11 is given by

$$
\begin{bmatrix} z_m \\ z_s \\ z_t \\ \hline r \\ r-n-y \end{bmatrix} = \underbrace{\begin{bmatrix} T_{ref} & 0 & -G \\ W_s & -W_s & W_s G \\ 0 & 0 & W_t G \\ \hline I & 0 & 0 \\ I & -I & -G \end{bmatrix}}_{=P} \begin{bmatrix} r \\ n \\ u \end{bmatrix}.
\tag{3.17}
$$

We represent the state space realization of P by

$$
P \equiv \begin{bmatrix} A & B_w & B \\ \hline C_z & D_{zw} & D_z \\ C & D_w & 0 \end{bmatrix}
\tag{3.18}
$$

and the controller by

$$
K(s) = [K_{ff}(s) \quad K_{fb}(s)] \equiv \begin{bmatrix} A_k & B_k \\ \hline C_k & D_k \end{bmatrix}.
\tag{3.19}
$$

In terms of these realizations, the overall closed-loop transfer Φ in (3.15) is then derived as

$$
\Phi = LFT(P, K)
$$
$$
= \begin{bmatrix} A + BD_kC & BC_k & B_w + BD_kD_w \\ B_kC & A_k & B_kD_w \\ \hline C_z + D_zD_kC & D_zC_k & D_{zw} + D_zD_kDw \end{bmatrix}
$$
$$
\equiv \begin{bmatrix} \bar{A} & \bar{B} \\ \hline \bar{C} & \bar{D} \end{bmatrix}.
\tag{3.20}
$$

We represent the transfer functions from w_j to z_j in Figure 3.11 by Φ_j's $(j = 1, 2)$. (The transfer function $\Phi_1 = T_{ref} - T_{yr}$ represents the model mismatch and $\Phi_2 = -[W_s S \quad W_t T]^T$ reflects the robustness and performance objectives). Note that these are the transfer functions of interest to be minimized. Each of these transfer functions can be written in terms of the matrix transfer function Φ in (3.15) as $\Phi_j = L_j \Phi R_j$, where the matrices L_j and R_j are chosen as

$$L_1 = \begin{bmatrix} I & 0 & 0 \end{bmatrix}, \ R_1 = \begin{bmatrix} I \\ 0 \end{bmatrix}, \ L_2 = \begin{bmatrix} 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}, \ \text{and} \ R_2 = \begin{bmatrix} I \\ 0 \end{bmatrix}. \tag{3.21}$$

Let $B_j = B_w R_j$, $C_j = L_j C_z$, $D_j = L_j D_{zw} R_j$, $E_j = L_j D_z$, $H_j = D_w R_j$ $(j = 1, 2)$ then,

$$\Phi_j = \left[\begin{array}{c|c} \bar{A} & \bar{B} R_j \\ \hline L_j \bar{C} & L_j \bar{D} R_j \end{array} \right]$$

$$= \left[\begin{array}{cc|c} A + B D_k C & B C_k & B_j + B D_k H_j \\ B_k C & A_k & B_k H_j \\ \hline C_j + E_j D_k C & E_j C_k & D_j + E_j D_k H_j \end{array} \right]$$

$$\equiv \left[\begin{array}{c|c} \bar{A} & \bar{B}_j \\ \hline \bar{C}_j & \bar{D}_j \end{array} \right] \qquad \text{for } j = 1, 2. \tag{3.22}$$

Now, we can impose the performance condition (or design specification) for each Φ_j separately.

The condition for $\mathcal{H}_\infty$ performance $\|\Phi_j\|_\infty < \gamma_j$ can be imposed by the bounded-real lemma (See Appendix A.1). Equivalent problem to these performance conditions with Hurwitz $\bar{A}$ is determining $P_j > 0$ that satisfy the following matrix inequality

$$\begin{bmatrix} \bar{A}^T P_j + P_j \bar{A} & P_j \bar{B}_j & \bar{C}_j^T \\ \bar{B}_j^T P_j & -\gamma_j I & \bar{D}_j^T \\ \bar{C}_j & \bar{D}_j & -\gamma_j I \end{bmatrix} < 0 \tag{3.23}$$

for $j \in \{1, 2\}$.

The condition for $\mathcal{H}_2$ performance $\|\Phi_j\|_2 < \gamma_j$ is more complicated. The condition $\|\Phi_j\|_2^2 < \gamma_j$ is equivalent to find $P_j > 0$ and auxiliary parameter Q that satisfy the following matrix inequalities

$$\begin{bmatrix} \bar{A}^T P_j + P_j \bar{A} & P_j \bar{B}_j \\ \bar{B}_j^{\,T} P_j & -I \end{bmatrix} < 0, \quad \begin{bmatrix} P_j & C_j^T \\ C_j & Q \end{bmatrix} > 0, \quad Tr(Q) < \gamma_j, \quad D_j = 0 \tag{3.24}$$

for $j \in \{1, 2\}$ (See Appendix A.1).

However, the matrix inequalities in (3.23) or (3.24) are not linear in terms of the actual design variables $(A_k,\ B_k,\ C_k,\ D_k)$. They can be converted into LMIs if we impose the conservatism

$$P_1 = P_2 \equiv P \tag{3.25}$$

and change the variables through an appropriate transformation. This assumption brings the conservatism in design but it recovers linearity of variables.

A short sketch of construction of the transformation of variables that makes (3.23) and (3.24) LMIs in design variables is given below. We decompose the unknown positive definite matrix P and its inverse as

$$P = \begin{bmatrix} S & N \\ N^T & ? \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} R & M^T \\ M^T & ? \end{bmatrix} \tag{3.26}$$

where the terms shown as ? are unimportant.

The transformation is defined in terms of the following matrices

$$\Pi_1 = \begin{bmatrix} R & I \\ M^T & 0 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} I & S \\ 0 & N^T \end{bmatrix}. \tag{3.27}$$

Note that the decomposition of P (3.26) and that $PP^{-1} = P^{-1}P = I$ imply that

$$MN^T = I - RS, \quad P\Pi_1 = \Pi_2. \tag{3.28}$$

The new (transformed) variables are defined as

$$
\begin{aligned}
\hat{A} &= NA_kM^T + NB_kCR + SBC_kM^T + S(A + BD_kC)R, \\
\hat{B} &= NB_k + SBD_k, \\
\hat{C} &= C_kM^T + D_kCR, \\
\hat{D} &= D_k.
\end{aligned}
\tag{3.29}
$$

By congruence transformation of block diagonal matrix which consists of Π_1 and I, (3.23) and (3.24) can be transformed into the LMIs with variables of $\hat{A}, \hat{B}, \hat{C}, \hat{D}, R, S$.

For instance, if we impose $\mathcal{H}_\infty$ performance on Φ_2 which is natural choice for robustness, then congruence transform of $diag(\Pi_1, I, I)$ will change (3.23) into

$$
\begin{bmatrix}
\mathcal{Q}(AR + B\hat{C}) & (*) & (*) & (*) \\
\hat{A} + (A + B\hat{D}C)^T & \mathcal{Q}(SA + \hat{B}C) & (*) & (*) \\
(B_2 + B\hat{D}H_2)^T & B_2^T S + H_2^T \hat{B}^T & -\gamma_2 I & (*) \\
C_2 R + E_2 \hat{C} & C_2 + E_2 \hat{D}C & D_2 & -\gamma_2 I
\end{bmatrix} < 0
\tag{3.30}
$$

and $\Pi_1^T P\Pi_1$ becomes

$$
\begin{bmatrix}
R & (*) \\
I & S
\end{bmatrix} > 0.
\tag{3.31}
$$

Here, $(*)$ can be inferred by symmetry and the operation $\mathcal{Q}(Z)$ denotes $Z + Z^T$. Note that (3.30) and (3.31) are linear in $\hat{A}, \hat{B}, \hat{C}, \hat{D}, R, S$. Similar results can be derived for $\mathcal{H}_2$ performance case.

Now, we can solve the multi-objective optimization problem in (3.16). It has a solution if the following minimization problem has a solution $(\hat{A}, \hat{B}, \hat{C}, \hat{D}, R, S, \gamma_1, \gamma_2)$.

$$
\min \quad \rho\gamma_1 + \gamma_2
\tag{3.32}
$$

subject to LMIs corresponding to the performancespecifications on Φ_1 and Φ_2.

For example, if we impose $\mathcal{H}_\infty$ performance on both Φ_1 and Φ_2 (i.e. $\alpha_1 = \alpha_2 = \infty$) then

the minimization problem is

$$\min \quad \rho\gamma_1 + \gamma_2$$

subject to

$$
\begin{bmatrix}
\mathcal{Q}(AR + B\hat{C}) & (*) & (*) & (*) \\
\hat{A} + (A + B\hat{D}C)^T & \mathcal{Q}(SA + \hat{B}C) & (*) & (*) \\
(B_1 + B\hat{D}H_1)^T & B_1^T S + H_1^T \hat{B}^T & -\gamma_1 I & (*) \\
C_1 R + E_1 \hat{C} & C_1 + E_1 \hat{D}C & D_1 & -\gamma_1 I
\end{bmatrix} < 0,
$$

$$
\begin{bmatrix}
\mathcal{Q}(AR + B\hat{C}) & (*) & (*) & (*) \\
\hat{A} + (A + B\hat{D}C)^T & \mathcal{Q}(SA + \hat{B}C) & (*) & (*) \\
(B_2 + B\hat{D}H_2)^T & B_2^T S + H_2^T \hat{B}^T & -\gamma_2 I & (*) \\
C_2 R + E_2 \hat{C} & C_2 + E_2 \hat{D}C & D_2 & -\gamma_2 I
\end{bmatrix} < 0,
$$

$$
\begin{bmatrix}
R & (*) \\
I & S
\end{bmatrix} > 0.
\tag{3.33}
$$

Equation (3.33) is well defined convex optimization problem with LMIs in parameters $(\hat{A}, \hat{B}, \hat{C}, \hat{D}, R, S, \gamma_1, \gamma_2)$ and can be solved by standard convex optimization problem solving tools.

The final step is reconstruction of (A_k, B_k, C_k, D_k) from $(\hat{A}, \hat{B}, \hat{C}, \hat{D}, R, S)$ obtained from the previous convex problem. This is done using the equations in (3.29). For full order control design, M and N are invertible and (A_k, B_k, C_k, D_k) can be uniquely determined as

$$
\begin{aligned}
A_k &= N^{-1}(\hat{A} - (\hat{B} - SB\hat{D})CR - SB(\hat{C} - \hat{D}CR) - S(A + B\hat{D}C)R)M^{-T}, \\
B_k &= N^{-1}(\hat{B} - SB\hat{D}), \\
C_k &= (\hat{C} - \hat{D}CR)M^{-T}, \\
D_k &= \hat{D}.
\end{aligned}
\tag{3.34}
$$

3.4.3 Implementation of the controller

The controller resulting from the design outlined above is applied to G_{xx} with $T_{ref} = \frac{1}{0.0003s+1}$, $W_s = \frac{0.5s+394.8}{s+0.3948}$ and $W_t = \frac{100s+9.475\times10^4}{s+1.184\times10^5}$ which reflect the performance objectives of high bandwidth, high resolution and robustness to modeling errors. High bandwidth requires that $|S_{er}|$ is small for wide operating frequency. It is imposed by T_{ref} function since matching $T_{yr} = 1 - S_{er}$ with T_{ref} shapes S_{er} indirectly. High resolution requires the roll-off frequency of ω_T to be small, which is reflected in W_t. The robustness requirement is reflected in W_s which enforces $\|S\|_\infty \leq 2$. ∞-norm is chosen for both model matching and robust objectives and the weight of $\rho = 20$ is chosen for the multi-objective minimization,

$$\min_{K \in \mathcal{K}} \rho \left\| T_{ref} - T_{yr} \right\|_\infty + \left\| \begin{array}{c} W_s S \\ W_t T \end{array} \right\|_\infty. \tag{3.35}$$

The resulting 2DOF controller from this design is given by

$$
K_{ff} = \frac{2.746 \times 10^{14}(s+1.184\times10^5)(s+0.7883)(s+390.3)(s^2+470.9s+8.352\times10^6)}{(s+3.131\times10^{11})(s+1.296\times10^8)(s+7817)(s+3333)(s+543)(s+0.789)}
$$
$$
\times \frac{(s^2+689.4s+1.315\times10^7)(s^2+4950s+2.44\times10^7)}{(s^2+587.2s+8.628\times10^6)(s^2+226.5s+1.407\times10^6)},
$$
$$
K_{fb} = \frac{6.9497 \times 10^{13}(s+1.184\times10^5)(s+3333)(s+390.3)(s^2+470.9s+8.352\times10^6)}{(s+3.131\times10^{11})(s+1.296\times10^8)(s+7817)(s+3333)(s+543)(s+0.789)}
$$
$$
\times \frac{(s^2+689.4s+1.315\times10^7)(s^2+4950s+2.44\times10^7)}{(s^2+587.2s+8.628\times10^6)(s^2+226.5s+1.407\times10^6)}. \tag{3.36}
$$

The corresponding closed-loop transfer functions are shown in Figure 3.12.

The 2DOF control laws obtained from multi-objective synthesis were implemented. Figure 3.13(a) shows the experimentally obtained $\|S_{er}\|$ which gives the tracking performance (ω_{BW}= 161 Hz) with good robustness to modeling uncertainties ($\|S\|_\infty$=1.06) and Figure 3.13(b) shows the experimentally obtained $T_{yr}(s)$ with $\|T\|$ which gives the roll-off frequency ω_T=57.5 Hz.
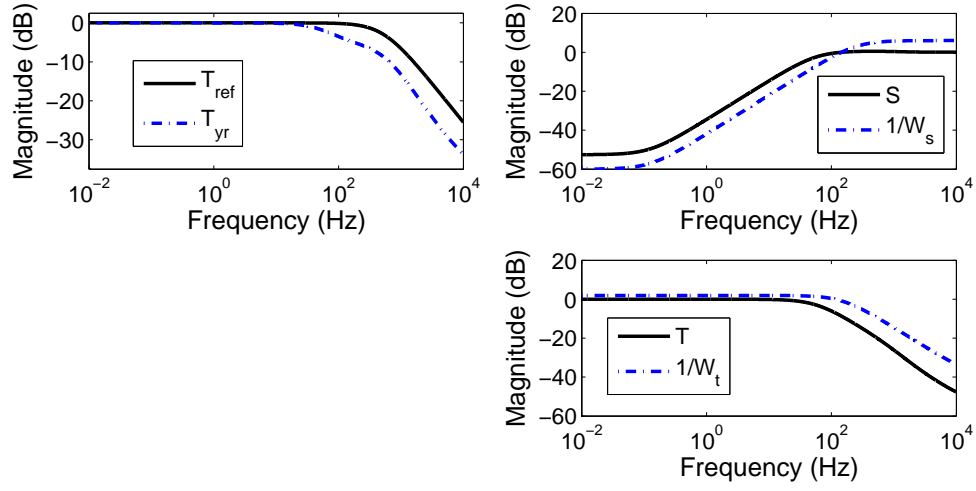
Figure 3.12: Multi-objective synthesis of 2DOF control using LMI. Minimization results corresponding weight $\rho = 20$ are $\|T_{ref} - T_{yr}\|_\infty = 0.526$ and $\left\|[W_s S \ W_t T]^T\right\|_\infty = 2.48$.
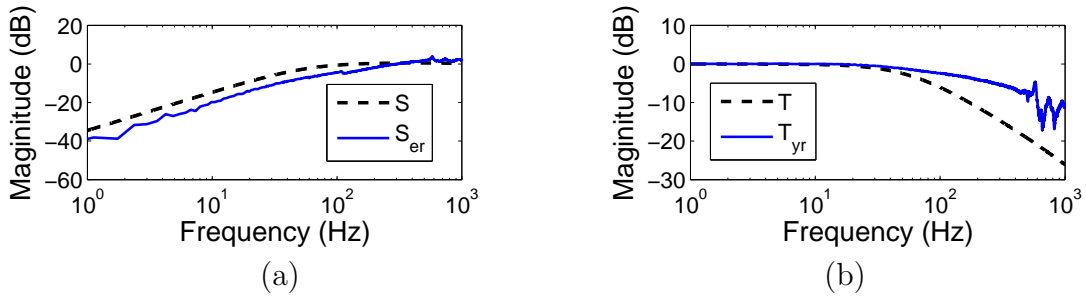
Figure 3.13: Magnitude of $S_{er}(s)$ and $T_{yr}(s)$(solid) obtained from experiment with S and T(dashed).

3.4.4 Comparison with feedback-only design

For comparison with usual feedback-only 1DOF design, the feedback-only controller was designed by S/T stacked sensitivity optimal control synthesis such that the closed-loop function has similar $|T|$ and $|S|$ as shown in Figure 3.14 which means almost identical resolution and robustness to model uncertainty.

The result of $\mathcal{H}_\infty$ synthesis yielded the 9th order following control laws,

$$
\begin{aligned}
K_{fb} = {} & \frac{6.0151 \times 10^{13}(s + 2.531 \times 10^5)(s + 390.3)(s^2 + 470.9s + 8.352 \times 10^6)}{(s + 1.471 \times 10^{13})(s + 3.292 \times 10^6)(s + 1.203 \times 10^4)(s + 543)} \\
& \times \frac{(s^2 + 689.4s + 1.315 \times 10^7)(s^2 + 4950s + 2.44 \times 10^7)}{(s + 0.8733)(s^2 + 587.2s + 8.628 \times 10^6)(s^2 + 226.6s + 1.407 \times 10^6)}.
\end{aligned} \tag{3.37}
$$
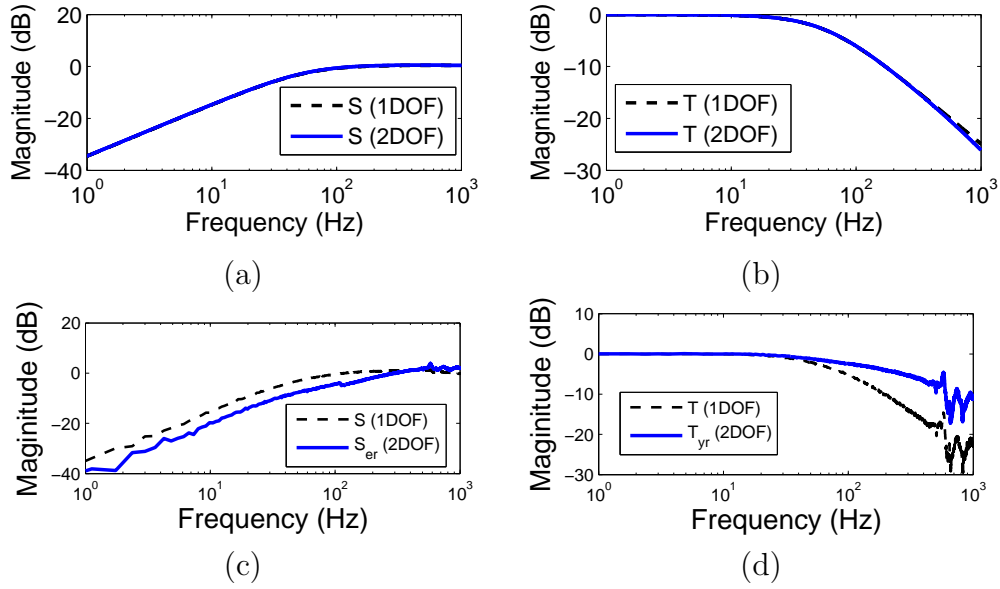
Figure 3.14: Comparison of $\mathcal{H}_\infty$ feedback-only stacked synthesis and 2DOF multi-objective synthesis control design. (a),(b) $|S|$ and $|T|$ from 1DOF control design(dashed) with $|S|$ and $|T|$ from 2DOF control design(solid). (c),(d) $|S|$ and $|T|$ from 1DOF control design(dashed) with $|S_{er}|$ and $|T_{yr}|$ from 2DOF control design(solid) obtained from experiment. The 2DOF multi-objective synthesis control achieves about 216% improvement in the tracking bandwidth.

The experiment results in $\|S\|_\infty$=1.17, ω_T=63.7 Hz and ω_{BW}= 51.0 Hz. There is an improvement of 216% in bandwidth for nearly the same values of resolution and robustness when compared to the optimal feedback-only design.

3.4.5 Discussion of this methodology

One of the principal advantages of the LMI-based approach is that it allows *simultaneous* incorporation of different metrics (such as $\mathcal{H}_\infty$, $\mathcal{H}_2$) of performance constraints as well as including design requirements such as specifications on passivity, asymptotic tracking and pole locations [47, 48]. This generality of approach, especially in incorporation of mixed-norm optimization problems, proves very useful in the context of nanopositioning. The measure on robustness is well characterized by infinity norm, but the natural metrics on resolution and bandwidth can vary depending on the applications and specific requirements. For instance, the metric on model mismatch can be given in terms of $\mathcal{H}_\infty$ norm as we have considered in this work or $\mathcal{H}_2$ norm depending on whether we are interested in the 'worst-case' or average performance. Thus, various specifications impose different optimization problems, say in the form

$$\min_{K \in \mathcal{K}} \rho \|T_{ref} - T_{yr}\|_\alpha + \|T\|_\beta \ \text{ subject to } \|W_s S\|_\gamma < 1,$$

where α, β, and γ can be two or ∞ norms.

This approach also allows for interchanging cost and constraints in an optimization problem. For instance, one can solve a problem where the cost function is maximization of robustness for given specifications of model mismatch and noise attenuation, that is, one can solve for problems of the form

$$\min_{K \in \mathcal{K}} \|S\|_\infty \text{ subject to} \|T_{ref} - T_{yr}\|_\infty < m_1 \text{ and } \|T\|_\infty < m_2.$$

Thus, this approach makes possible addressing a variety of nanopositioning applications with diverse objectives.

It should be emphasized that this approach imposes a technical condition (3.25) to solve the optimization problem. However, this condition adds conservatism to the solution in the sense that it minimizes over a subset of $\mathcal{K}$ instead of the entire $\mathcal{K}$. An area of active research is on relaxing this technical condition [49, 50].

CHAPTER 4

ANALYSIS AND DISCUSSION

4.1 Robustness to the Modeling Uncertainty

As emphasized in previous chapters, the complex structures of nanopositioning systems, lack of precise physical understanding of piezoelectric effects, as well as effects of environmental conditions (such as temperature and humidity), floor disturbances, and the need for simple models motivate control designs that make the closed-loop systems *robust* to these modeling uncertainties. That is, designs that guarantee some metric of performance for all plants that can be modeled as a perturbation of the *nominal* plant; the perturbation characterized by an uncertainty set. However, in the control designs in preceding chapters, an explicit characterization of the uncertainty set is not made. This omission can be justified as follows. The robustness analysis starts from identifying the uncertainty set. Based on the identification experiments in Chapter 2, the scanner described in this study can be analyzed as a system with multiplicative uncertainty in Figure 4.1(a).

The scanner system is represented by any transfer function G_p in the following set

$$\{G_p(s)|G_p(s) = (1 + w_i(s)\Delta(s))G_{xx}(s), \text{ where } \|\Delta\|_\infty \leq 1\}, \tag{4.1}$$

where G_{xx} is a nominal scanner transfer function (2.9), and $\Delta(s)$ is any stable transfer function. Figure 4.1(b) shows the nominal response of G_{xx}(solid) and the worst-case response of G_p(dashed) with simple choice of $w_i = \frac{s+1.026\times10^5}{s+5.132\times10^5}$. The uncertainty (the deviation from the nominal response) can be captured as shown. The sufficient and necessary condition for
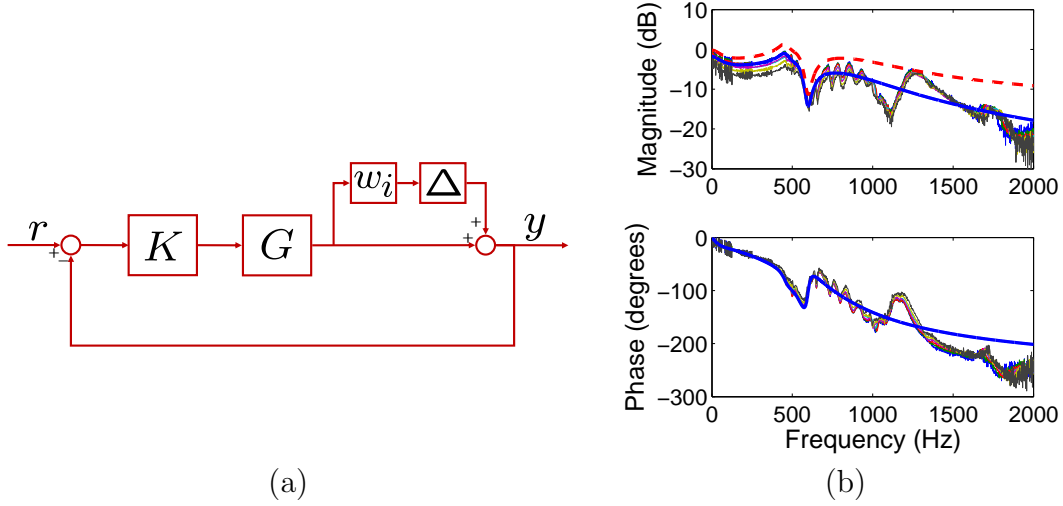
Figure 4.1: (a) Control system with multiplicative uncertainty (b) Nominal(solid) and worst-case(dashed) frequency response with experimental frequency responses.

robust stability for the systems with multiplicative uncertainty [33] is given as

$$|T| < \frac{1}{|w_i|}. \tag{4.2}$$

The constraints on T from our choices of W_T in Chapter 3 are stricter (i.e. $|w_i| > |W_T|$), and therefore this robust stability requirement is satisfied without difficulty.

4.2 Relative Roles of Feedforward and Feedback

One of the main advantages of the formulation of the optimal design in Chapter 3, besides providing an apt framework for incorporating design specifications and achieving them, is that their analysis gives significant insights into the trade-offs between difference objectives and relative roles of feedforward and feedback components. For instance, let us analyze the first two rows in (3.12) which reflects the tracking bandwidth and positioning resolution objectives. The term Sn in $z_s = W_s S_{er} r - W_s Sn$ and the term Tn in $z_t = W_t T_{yr} r - W_t Tn$ quantify the deteriorating effects of noise on tracking bandwidth and positioning resolution respectively. They cannot be made small in the same frequencies since S and T cannot be made small simultaneously. This problem is resolved by designing weight functions W_s and

49

W_t that separate these objectives in the frequency domain. An inappropriate choice of these weight functions can lead to infeasible optimization problem.

Since $S_{er} = S(1 - GK_{ff})$ and S cannot be made small over the entire bandwidth range (in order to allow for noise attenuation), the feedforward control K_{ff} plays an 'active' role in making S_{er} small beyond the frequency where S is not small (say greater than $1/\sqrt{2}$). Also since $S = (1 + GK_{fb})^{-1}$ is completely determined by K_{fb}, the feedback component is dominant in frequencies where S is small. This in turn implies that robustness to modeling errors is solely determined by K_{fb} and therefore feedforward component has a very little role in low frequencies (where typically robustness requirements are specified). For instance, Figure 4.2 shows the bode diagram for $K(s) = [K_r(s)\ K_y(s)]$ in terms of architecture in Figure 3.1(b). Figure 4.2(a) shows the controller from optimal prefilter model matching design. Note that, in low frequencies, the control law is essentially feedback-only since K_r and K_y have almost same magnitude and are 180 deg out of phase, i.e. $K_{tr} \approx -K_{ty}$, and therefore $u \approx K_y(r - y)$. The 2DOF nature of the control becomes active at high frequencies (> 50 Hz), that is near the bandwidth achieved by the pre-existing control design. Similarly in the controller from 2DOF optimal robust model matching design, the feedforward part is dominant at high frequencies (> 100 Hz) (Figure 4.2(b)).
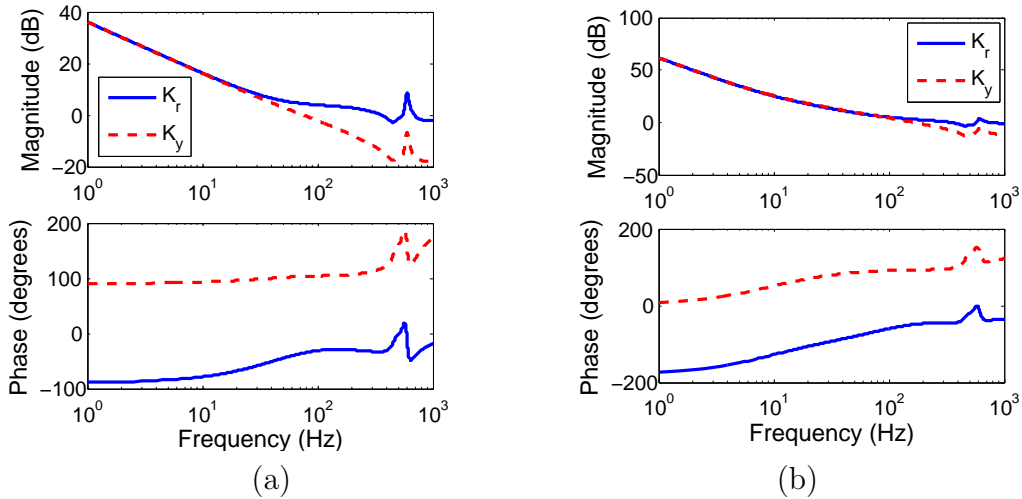

Figure 4.2: Bode plot $K(s) = [K_r(s)\ K_y(s)]$: (a) Optimal prefilter model matching controller (b) 2DOF optimal robust model matching controller.

From the above discussion, the main contribution of feedforward component is in the fre-

quency range where S is no longer small. However this frequency range is limited. Typically nanopositioning devices have very poor capabilities to provide motions beyond their flexure resonance frequencies. This is characterized by high roll-off rates of G. Therefore, very high control inputs (high voltage inputs) are needed to make the positioning systems beyond their flexure resonances. The saturation limits on control signals form the main constraints on attaining bandwidths beyond flexure resonances. Thus the feedforward components provide performance *enhancements* over feedback-only designs in the frequency range from corner frequency of S to the flexure resonance frequency. This perspective is further justified from the fact that the 2DOF mixed sensitivity optimal control designs (where the optimization problem formulation does not discriminate between feedforward and feedback components with respect to performance objectives) do not provide characteristically better performances than the 2DOF optimal robust model matching designs (where the optimization problem formulation places robustness requirement on feedback and the bandwidth enhancement objective on the feedforward component).

4.3   Breaking Barriers of Feedback-only Design

The 2DOF design is not bound by some fundamental limitations that constrain the feedback-only designs. For instance, the results in Chapter 3 show that in 2DOF design, the tracking bandwidth ω_{BW} of the closed-loop device can be made larger than the roll-off frequency ω_T which determines the resolution. The corner frequency ω_{BW} can *never* be made larger than ω_T in feedback-only design [33], which suffers from a stricter trade-off between the resolution and the bandwidth. This limitation is restricting as it is contrary to the performance objectives since high values of ω_{BW} and low values of ω_T are required for high tracking bandwidths and noise attenuation respectively. For optimal prefilter model matching (2DOF) control based on $\mathcal{H}_\infty$ controller has ω_{BW} of 214.5 Hz and ω_T of 60.1 Hz while 1DOF $\mathcal{H}_\infty$ controller has the bandwidth of 49.4 Hz and the same roll-off frequency. Also, the improvement of the bandwidth and robustness at the same time for fixed resolution is generally impossible due to fundamental limitation in feedback-only design but in 2DOF design, it is achieved as in 2DOF optimal robust model design example. These examples give a case where 2DOF

design achieves specifications that are *impossible* to attain with feedback-only design.

4.4   Control Design Extensions

The 2DOF designs presented here can be easily extended to achieve a larger class of design specifications than presented in Chapter 3 with trivial modifications to the designs. For instance, we have not exploited the frequency content of the reference signals in our designs. This can be easily incorporated by writing the reference signal as $r = H(j\omega)r'$, where r' is arbitrary signal (as used in our analysis) and the transfer function $H(j\omega)$ reflects the frequency content of the signal. Alternatively we can incorporate the frequency content of the reference signal through the model transfer functions T_{ref} or through weight transfer functions.

In this thesis, we have demonstrated experiments that achieve typical trade-offs between performance objectives. We can achieve extreme trade-offs by suitably designing the reference transfer functions T_{ref} or weight transfer functions. For instance, very high resolution can be achieved by trading-off severely on the tracking bandwidth. In [51], this trade-off (albeit in feedback-only design) yielded sub-nanometer resolution for the nanopositioning system described here. In fact, in this design the scanning speed for repetitive scans was not sacrificed by making the closed-loop device behave as a small-bandwidth device around the scanning frequency.

4.5   Integration of Device and Control Design

In this paper, we have presented control designs assuming that the scanner G has already been designed. The integration of control design and device design can lead to further increase in the closed-loop device performance. For instance, in [52] lighter positioning systems with single moving mass for multiple directions albeit with high cross-coupling effects are presented. The lighter mass results in higher flexure resonance frequencies and the control design decouples the cross interactions which has resulted in an overall superior device. The fundamental limitations developed in this paper can be used in the analysis

of device design that will enable better closed-loop devices. For instance the device design can be altered and resulting changes in non-minimum phase zeros (or their locations) be monitored that will result in better closed-loop devices.

Part II

CONTROL APPROACH FOR FAST IMAGING IN DYNAMIC AFM

CHAPTER 5

IMAGING IN AFM

5.1   Basic Principle of AFM

AFM is a device that functions as a scaled down sample topography profiler [53]. A typical AFM setup is shown in Figure 5.1. The laser beam is focused on the cantilever and is reflected from the cantilever surface into a split photodiode. The cantilever deflection changes the angle that the incident laser beam makes with the cantilever surface, which, in turn, changes the incidence position on the photodiode, which is registered as a change in the photodiode voltage. The difference between the readings from the top and the bottom cells gives a measurement proportional to the cantilever's normal deflection. The length of the reflected laser path amplifies the cantilever deflection. The fixed end of the cantilever is supported by the base, which is typically attached to a dither piezo which is required for dynamic mode operation. The dither piezo provides a means for oscillating the base of the cantilever. A scanner, which is actuated by piezoelectric material, positions the sample laterally and vertically during imaging. Except in very special cases such as constant height operational mode[1], a feedback control mechanism regulates the tip-sample interaction force at a constant value. More specifically, the measured deflection signal is used to design a feedback control that moves the piezo-positioner vertically in order to compensate for the effect of topographical features of the sample on the cantilever tip.

---

[1]Constant height mode is open loop operation without feedback control where the cantilever deflections are directly used for estimation of the sample topography
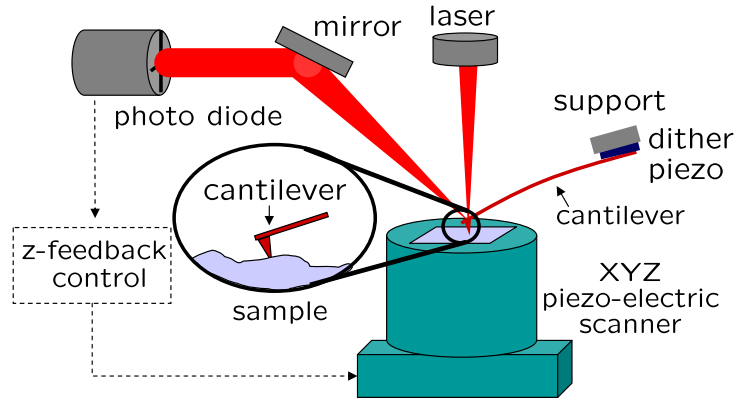
Figure 5.1: Atomic force microscope (AFM): AFM consists of a micro-cantilever, a detection system for measuring cantilever deflections, a sample positioner that are typically piezo-actuated to provide fine nanoscale displacements, and a control system for maintaining a desired operating condition.

5.1.1 Components of AFM

Cantilever and tip

A cantilever with a sharp tip needs to sense forces exerted on the tip by the atoms in the sample. To register small interatomic forces, the cantilever probe must be insensitive to extraneous disturbances from the surrounding environment. These disturbances are caused by sources such as building vibrations, whose power spectral density is significant in the frequency range 0-2 kHz. Therefore, to avoid the effects of these disturbances, the resonance frequency of the cantilever must be greater than 2 kHz. At the same time, the cantilevers need to be sufficiently compliant to sense the interatomic forces. The forces between the tip of the cantilever and the sample are in the range 10^{-7}-10^{-12} N.

Typical length, width, and thickness of micro-cantilevers used in AFM are $100 - 300$, $30 - 40$, and $2 - 5$ μm respectively. The stiffness of the micro-cantilever can vary from 0.06 to 100 N/m. The tetrahedral or cone shaped tip at the cantilever end has radius of 5-10 nm typically.

The tip-sample interaction forces are dominated by attractive and strong repulsive interatomic forces for long and short separations respectively (see Figure 5.2). Operating

modes in atomic force microscopy are characterized as contact, non-contact, or intermittent depending on the range of traversal of the cantilever-tip as described in Figure 5.2.
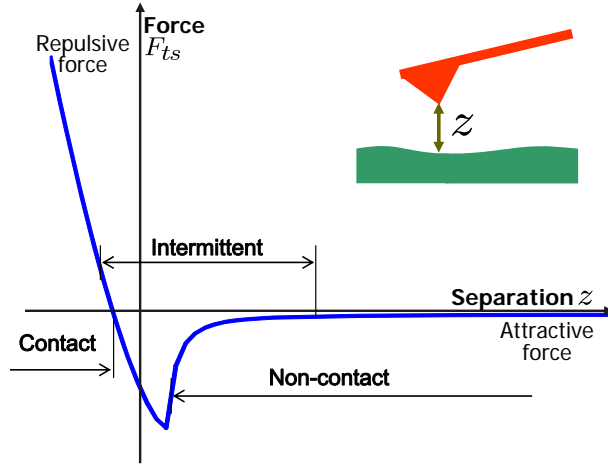
Figure 5.2: A typical tip sample interaction-force profile with long range attractive and short range strong repulsive forces.

Detection of cantilever deflection

The deflection of cantilever was measured by tunneling effect between the cantilever and the tunneling tip in the original AFM invented by G. Binnig, C. Quate, and C. Gerber [4]. Nowadays, it is commonly measured by a optical lever method (or laser beam bounce method) [54]. The components related with optical lever are generally composed of a micro-cantilever with reflective coating, a laser source, a photodiode for sensing laser incident.

Vertical actuator for feedback regulation

For AFM imaging, all three spatial direction actuations are required, two planar direction (typically X-Y) for scanning and one vertical direction (typically Z) for regulating the tip-sample interaction. The positioning system addressed in Part I is the planar direction scanner.

Most AFMs (and SPMs) use piezoelectric actuators, and there are two main designs, tube piezo and independent axis piezo-actuation. Tube piezo-actuators generate X-Y-Z motion by

applying voltages to the electrodes embedded on piezoelectric cylinders. Thus, the planar scanning and the feedback regulation are incorporated together and tube piezo-actuators usually have high crosstalk between planar motion and vertical motion. It requires the removal of cross coupling in imaging, typically by post processing, and the range of scanning is limited when compared to the independent type. Independent axis actuator separate the planar actuation and vertical actuation. Its independence feature removes the cross coupling and achieves large range in planar scanner.

In most available AFMs, the vertical actuation to control the separation distance between the cantilever tip and the sample is achieved either by actuating the entire cantilever and optical detection system while providing no mechanism for moving the sample or the vice versa. Typically, AFMs with separate lateral scanner and independent vertical scanner typically belongs to the former type while AFMs with tube actuators belongs to the latter type.

Feedback controller

The controller generates the input signal for vertical piezo-actuator to regulate the tip-sample interaction force, i.e. to track the constant level in force field (in general, the height profile of sample). Moreover it is also responsible for estimation of topography. The control signal regulates a constant force between the sample and the tip by compensating the topographical features on the sample. Therefore, this signal is also used as a measure of sample topography. In most commercial AFMs, Proportional-Integral (PI) or Proportional-Integral-double Integral (PII) control designs are used to regulate the tip-sample interaction.

5.1.2   Imaging modes

Depending on the sample and intended applications, various imaging modes are employed as through corresponding actuation strategies. The operating modes can be broadly classified as static or dynamic.

Static mode

In the static mode operation, the base of the cantilever is not forced and the tip scans the sample in contact with the surface. Since the measurement is broadband, the $1/f$ noise affects the detection scheme and plays a critical role in determining the imaging resolution [55]. For meaningful imaging, the deflection of the cantilever has to be larger than the root mean square deflection of the cantilever due to the dominant noise source. The deflection of the cantilever caused by the force on the cantilever tip due to the sample has to overcome the large measurement noise. Since the interaction forces between the tip and sample are small, the cantilever stiffness must be small (typically less than 1 N/m) to facilitate a sufficiently large deflection to overcome the measurement noise. The equilibrium points on the attractive portion of the tip-sample interaction in Figure 5.2 are unstable for soft cantilevers. Thus, the operating condition of static mode methods is restricted to the repulsive portion of the force curve, where the forces on the tip and sample are larger.

The most prevalent static mode operation is the constant force mode where the static tip deflection is used as a feedback signal. If the measured deflection is different from the set-point value, the feedback controller applies a voltage signal to extract or retract the vertical piezo to keep the deflection constant. A constant cantilever deflection means that a constant force is applied to the cantilever. The control signal is usually used as a measure of the height of topographic features on the sample. The advantages of this mode are the high scan speed, the atomic resolution imaging, the ease of data interpretation and the ease of implementation. Problems with this mode are caused by excessive tip-sample force. Lateral forces can distort features in the image. Soft samples such as biological tissues are not eligible for this mode since they can be easily damaged or lead to images with many artifacts.

Another mode used in static mode is the constant height mode where the static tip deflection is used as the direct measure of the sample topography. This mode is an open loop operation thus, the feedback controller is disabled. It has the advantage of fast scan and high resolution since the sensor noise is not fed back. However, since a sample is typically tilted, it is difficult to move the sample laterally without either losing proximity between

the tip and the sample or impacting the tip into the sample. Therefore scan sizes using this mode are limited. Another practical difficulty in using the constant height mode to image the topography of the sample is the nonlinear dependence of the cantilever deflection on the sample height. Only the qualitative behavior of the relation between them is known, which limits the ability to determine the topography of the sample from the deflection signal. Therefore this mode, when used, is applied on smooth surfaces (that is the heights on sample features are on the order of nm) which allows for linear approximations of the nonlinearity that are practical.

Dynamic mode

In the dynamic mode operation, the cantilever is externally oscillated at frequency close to its resonance frequency or a harmonic by forcing the base that supports the cantilever with a dither piezo (See Figure 5.1). A detailed review of this mode can be found in [56]. The cantilever oscillations vary when it interacts with the features on the sample. Thus, the changes in amplitude, phase and frequency of the cantilever oscillations are indicative of the effects of the tip-sample interaction forces and can be used to infer sample properties (the sample topography being one of them).

The single-mode cantilever model, which is useful to explain principle of dynamic mode operation, and the deflection measured by the photodiode are given by

$$\ddot{p} + \frac{\omega_0}{Q}\dot{p} + \omega_0^2 p = \frac{1}{m}F_{ts} + g(t) + \eta \tag{5.1}$$

$$p_m = p + n \tag{5.2}$$

where p, p_m, and n denote the deflection of the tip, measured deflection, and measurement noise, respectively. F_{ts}, g and η represent the tip-sample interaction force, the specific dither piezo excitation force and the specific thermal force. The parameters ω_0, m and Q are the first modal frequency (resonance frequency), the effective mass and the quality factor of the cantilever. (A detailed development of this model is presented in Chapter 6).

Figure 5.3 shows a power spectral density of the measured deflection signal p_m when the

60

force on the cantilever is the thermal noise force η. It is evident that the response of the cantilever are visible near the resonance of the cantilever, whereas, away from the resonance frequency, the measurement noise n clouds the thermal noise response of the cantilever. We note that, at a given temperature, thermally limited resolution is the best achievable resolution. When the signal of interest, for example, topography of the sample, is modulated such that required information is present near the resonance frequency of the cantilever, the measurement benefits from the high signal-to-noise ratio (SNR).
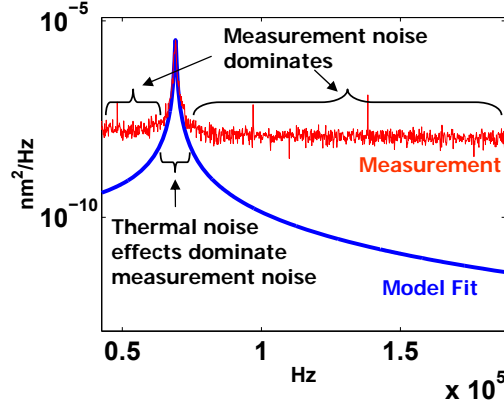

Figure 5.3: Effect of thermal noise on the cantilever deflection: The experimental thermal noise plot shows that, at and near resonance, the thermal noise response dominates the measurement noise. The second-order model fit provides an estimate of the spring constant of the cantilever [57].

Dynamic mode operations are explained by comparing the frequency responses of the cantilever with and without the influence of the sample in Figure 5.4. When the cantilever is not under the influence of the sample (dashed), a sinusoidal forcing with frequency ω results in a deflection signal that is sinusoidal with amplitude A_{sp}. In the absence of the sample, the resonance frequency of the cantilever is ω_R. Under the presence of a sample, with the intermittent contact causing an (average) repulsive force on the cantilever, the amplitude frequency curve (solid) shifts to the right with a increased resonance frequency ω_{R_e} and an decreased amplitude $A_{sp} - \Delta A$. (With the non-contact oscillation causing an (average) attractive force on the cantilever, the amplitude frequency curve shifts to the left.) The shift ΔA in the amplitude is exploited by the amplitude modulation AFM (AM-AFM)

method to measure the sample interaction. In frequency modulation AFM (FM-AFM), the difference $\Delta\omega_R$ between the *equivalent resonance frequency* ω_{Re} of the cantilever under the sample influence and the *free resonance frequency* ω_R is used to map the characteristics.
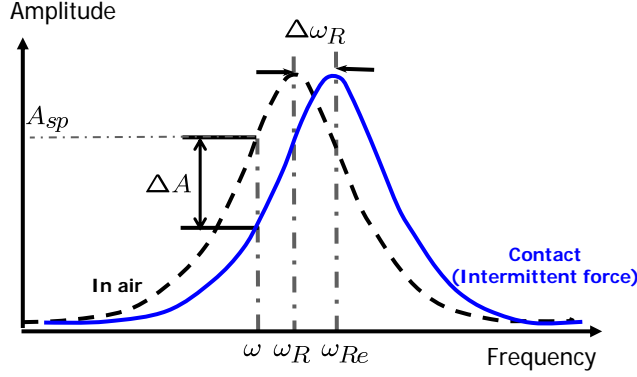

Figure 5.4: Effect of sample on the frequency response of the cantilever: The cantilever under the influence of the sample is modeled as an single-mode cantilever model with a changed resonance frequency ω_{Re}, with the original resonance frequency of the cantilever given by ω_R. The cantilever is forced with a drive frequency ω.

Dynamic mode of operation leads to smaller lateral and vertical forces. Because of its gentle nature, dynamic-mode operation is the preferred mode. Dynamic mode operations where the tip oscillate in the attractive regime in Figure 5.2 are called non-contact mode[58], and dynamic mode operations where the tip probes both the attractive and the repulse regime are called intermittent mode or tapping mode. The amplitude modulation intermittent mode is the most used mode for the characterization and modification of various materials in ambient condition.

Role of control in AFM imaging

The role of control design cannot be overemphasized in the imaging operation. Even in the very first AFM, which adopted static constant force mode operation, the controller played an essential role that enabled its operation. A controller is used to maintain constant contact force by positioning the sample in the vertical direction to compensate for the features on the sample topography as the lateral positioning system moves the sample in the lat-

eral directions. The force regulation enables the cantilever based sensing, otherwise it is difficult to move the sample laterally without either losing proximity between the tip and the sample or impacting the tip into the sample. The image of the sample topography is obtained by plotting the vertical control signal against the lateral coordinates. Similarly in amplitude-modulation AFM (AM-AFM) interaction force averaged over an oscillation cycle is maintained constant, and the resulting control required to compensate for the features on the sample topography provides for the imaging signal.

5.2   A Control Systems Perspective of AFM Imaging

To interpret the operating principle of imaging in control perspective, the block diagram schematic of an imaging in typical AFM is introduced in Figure 5.5, which excludes the x-y positioning system. The controller, the vertical piezo-positioner, the cantilever dynamics model which includes the tip-sample interaction force, and the signal conditioner are represented by K, G_p, $\mathcal{F}$, and $\mathcal{Q}$ respectively.

The various modes of operation differ in their designs of the dither control input g, the output y derived from the deflection signal p, and the way the feature height h is interpreted from the measurements. For instance, in contact-mode constant force microscopy, the dither is not excited ($g = 0$) and the measured deflection signal p_m is regulated at a constant ($\mathcal{Q}(\cdot)$ is the identity operator) value by appropriately designing the control signal u. If the effects of thermal η and sensor noise n are neglected, the tip-sample interaction force, which is a function of $p - h + v$, is approximately a constant since the deflection p is regulated at a constant value, which in turn implies that topography $h = v$. Since $v = G_p u$ and G_p is approximately a constant at low frequencies, the control signal u (the input to the vertical piezo-actuator) gives a measure proportional to h for low speed scans. Similarly in AM-AFM, the dither piezo is oscillated at a frequency ω close to the cantilever natural frequency ω_0, (i.e., $g(t) = g_0 \cos(\omega t)$), and the amplitude of the cantilever deflection p_m is regulated at a constant value. Again, the control signal u gives a measure of the topography h since the sample position v compensates for the effects of h to regulate the amplitude of the deflection signal. In fact, the control signal u from force regulation technique forms the topography
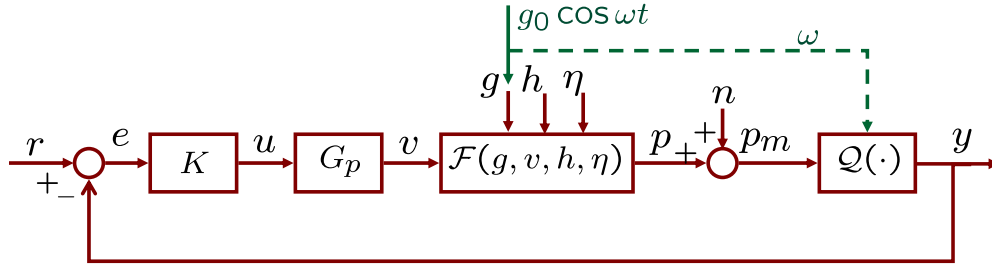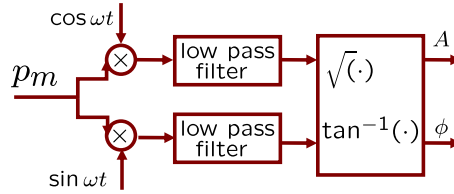
Figure 5.5: A block diagram schematic of an AFM: The controller K is designed to regulate the difference e between a derivative y of the deflection signal p and the set-point r to zero to compensate the effects of the sample topography h. The deflection p is described by a nonlinear dynamic model $\mathcal{F}$ which is explained in detail in Chapter 6. The deflection p is due to the forcing of the cantilever-transfer function $\mathcal{F}$, the dither piezo excitation g, the thermal noise η, and the tip-sample interaction force F_{ts} that depends on the sample-position v by vertical piezo-actuator and the sample height h. In contact-mode operation, $\mathcal{Q}$ is identity (i.e. $y = p_m$), while in the dynamic-mode operation, $\mathcal{Q}$ denotes lock-in amplifier or rms-to-dc converter when y is the amplitude of the oscillatory signal p for amplitude-modulation. The lock-in amplifier is described in (b). The deflection measurements p_m are corrupted by sensor noise, that is, $p_m = p + n$.

signal in most existing imaging modes. However, this signal yields distorted (or no) images for high speed scans (or rough samples) since G_p is not constant and u is not proportional to sample position v at high frequencies.

5.3 Related Literature on Imaging in AFMs

Compared to the research on nanopositioning, the problem of estimation of the sample topography has not been researched much in systems and control communities. Most of the studies on fast imaging have focused on the improvement of dynamic performance but there are some research with system theoretical approaches [59, 60, 61, 60, 62].

The followings are the previous studies classified by the component of AFMs [14, 63, 57].

5.3.1 Cantilever related studies

In dynamic mode operations, the cantilever response takes long time to settle down, which can be explained by the low damping in cantilever dynamics (5.1). Typically the dither piezo excitation to the cantilever support is given as $g(t) = g_0 \cos(\omega t)$, where ω is approximately equal or close to the first modal frequency of the cantilever. When the cantilever tip is far from the sample, when F_{ts} can be neglected, and if we ignore the effects of the thermal noise, the solution to (5.1) is given by

$$p(t) = \underbrace{A_0 \sin(\omega t + \phi)}_{\text{steady state}} + \underbrace{A' e^{-\frac{\omega_0 t}{2Q}} \sin(\omega' t + \phi')}_{\text{transient}}, \tag{5.3}$$

where $\omega' = \omega_0 \sqrt{1 - 1/4Q^2}$. Since the transient term decays as $e^{-\frac{\omega_0 t}{2Q}}$, the response time for transient signal to settle down is $\tau = \frac{2Q}{\omega_0}$. This implies that to achieve fast response, either decreasing the quality factor Q or increasing the first modal frequency or natural frequency ω_0 of the cantilever is needed. In this context, the following ideas have been proposed in the literature.

Active Q control

In [64, 65], the quality factor is actively controlled by feeding back the cantilever velocity $\dot{p}$ to the dither excitation $g(t)$. With $g(t) = g_0 \cos(\omega t) + k\dot{p}$, Equation (5.1) changes to

$$\ddot{p} + (\frac{\omega_0}{Q} - k)\dot{p} + \omega_0^2 p = \frac{1}{m}F_{ts}(p, v, h) + g_0 \cos(\omega t) + \eta \tag{5.4}$$

and the new system has the effective quality factor $Q' = \frac{\omega_0 Q}{\omega_0 - kQ}$. However, feeding back $\dot{p}$ is not practical since the measured deflection signal is noisy. This issue is easily addressed by noting that a shifted version of the deflection signal is proportional to its time derivative, that is, by noting that when $p(t) = A_0 \sin(\omega t + \phi)$, then time shift of $\frac{\pi}{2\omega}$ generates $p(t - \frac{\pi}{2\omega}) = -A\cos(\omega t + \phi) = \dot{p}/\omega$. Actual feedback through the dither piezo is achieved by implementing $g(t) = g_0 \cos(\omega t) + k'p(t - \frac{\pi}{2\omega})$. This active Q control is used to decrease the quality factor for the high speed of imaging in the intermittent mode [66] as well as to increase the quality factor for imaging in liquid where the original Q is very small [67].

Dual Q control

In active Q control, the feedbacked signal is not exactly $p(t)$ but $p_m(t - \frac{\pi}{2\omega})$, which is the noisy measured signal. The study in [62] used Kalman filter to obtain $\dot{p}$ and p from the measurement p_m. The Kalman filter utilizes (5.1) with noise characteristics. The new feedback of $g(t) = g_0 \cos(\omega t) + k_1\dot{\hat{p}} + k_2\hat{p}$ can change the effective stiffness as well as the effective damping. This requires very fast electrical circuit such as field programmable gate array (FPGA) to implement the Kalman filter which runs at time scales much faster than the excitation frequency ω.

Small cantilever

To increase the natural frequency, small cantilevers are used [68, 69]. The natural frequency is given as $\omega_0 = \sqrt{\frac{k}{m}}$. Since increasing the stiffness k is constrained by the requirement of reacting to small tip-sample interaction, decreasing the effective mass m by down-sizing cantilever is a possible solution. The cantilever used in [69] is approximately 6 μm long, 2 μm wide, and has the natural frequency of 3.5 MHz.

5.3.2   Detection of cantilever displacement related studies

Optical lever detector for small cantilever

The use of small cantilever requires the new optical devices consistent with small cantilever size. The incident and reflected laser beams are separated using a quarter wavelength plate and polarization splitter in [68]. In [70], a radio frequency power modulator is used for reducing the laser beam coherence [71].

Cantilever with integrated sensors

Cantilevers with integrated sensors are used in that they do not require the complicated optical devices such laser source and photo-detector [72, 73].

Amplitude calculation for amplitude modulation mode

For amplitude calculation from cantilever deflection signal ($Q(\cdot)$ in Figure 5.5), typically, the rms-to-dc converter or the lock-in amplifier is used. The rms-to-dc converter uses a rectifier circuit and a low-pass filter, and takes several oscillation to output the amplitude. The lock-in amplifier has also the large response time. These elements make the feedback loop of the amplitude modulation mode slow.

Use of the peak picking circuit generates the peak to peak value as the amplitude value in half cycle [71] and contribute to fast imaging of amplitude modulation mode.

Observer based methods

In [74], by using Kalman filter and a likelihood ratio test, the sample features are detected in transient signal and achieves two order higher detection rates. This methods aims the binary detection of feature, where the sample-height is either high or low. In [75], a real-time methodology for detecting probe loss in dynamic mode operation is developed. Reliability index, based on the observer method can give better detection of probe-loss than conventional amplitude signal.

5.3.3 Vertical Actuator related studies

Active cantilever

In [76], the cantilever with an PZT thin film is used to control the tip-sample spacing by applying the feedback actuation signal to the PZT layer and sense the deflection by itself. The excitation ac signal is also superimposed on the feedback actuation signal. In [77], the cantilever with an integrated piezoelectric actuator and an integrated thermal actuator is used as a substitute for the vertical piezo-actuator and has better bandwidth than the conventional setup. In [78], for fast imaging, the cantilever with a thin zinc oxide (ZnO) film to both actuate and sense deflection that contains both an integrated piezoelectric, is used instead of separated dither piezo-actuator and optical sensing mechanism and it can be used in liquid.

Counter balancing of the vertical actuator

In [71], the force that generates the structural resonance due to fast vertical actuation, is suppressed by using two identical vertical actuators in opposite directions.

Active damping to vertical actuator

In [79], active damping is used to minimize the vibration of the vertical scanner with counter balancing and contribute to achieve fast imaging.

5.3.4 Feedback control and topography estimation

Detecting the sample from thermal response

In [80], the thermal response of a cantilever is exploited and used to detect sample motions in attractive regime of tip-sample interaction. Without externally exciting the cantilever i.e. $g(t) = 0$, the shift in resonance frequency in thermal noise response is monitored to maintain the cantilever in the attractive portion. This method has been demonstrated to sense feature heights that smaller than an angstrom albeit with very low bandwidth.

Switching control depending on the probe-loss

In imaging samples with sudden changes, especially falls due to troughs in the sample topography, the cantilever loses contact with the sample (called as probe-loss) and hence the corresponding deflection measurements are spurious and lead to artifacts in images. In [81, 82], a switching of PID controller based on the probe-loss detection in dynamic mode operation, is used to reduce probe-loss affected regions in an image. In [81], the tip-loss is directly detected from oscillation amplitude, while in [82], the reliability index based on observer method [75] is used for a better tip-loss detection.

Exploting previous scan line information

In Multimode AFM by Veeco Instruments Inc., Santa Babara, CA and [83], based on the assumption that two adjacent scan lines are similar, information from previous scan lines is added as the feedfoward signal and improves the bandwidth and the accuracy in imaging. In [83], the feedforward controller for previous scan line, as well as the feedback controller in contact mode operation is designed in $\mathcal{H}_\infty$ robust optimal framework for having better bandwidth and robustness.

Separate signal for estimation

In [84], the separate estimator besides the controller is used to obtain better imaging bandwidth, where the control signal is used only for contact force regulation while a separate signal is derived to estimate the sample topography. It adopts $\mathcal{H}_\infty$ robust optimal control framework in the design of these signals.

5.4   Challenges in Dynamic Mode Imaging

One of the foremost requirements in many applications, especially in imaging of soft bio samples such as cells, tissues and proteins, is that the cantilever is gentle on the sample and does not damage the sample. These applications therefore preclude contact-mode AFM, since the cantilever can tear through the sample surface. The dynamic-mode AM-AFM is more commonly used for imaging, since they come in contact only intermittently with the sample and the cantilever does not drag through the sample.

Since the sample-topography data is interpreted from the steady state amplitude values of the deflection signal in existing methods, the imaging is slow. Another challenge for fast-imaging in AM-AFM stems from the high frequency oscillation of the cantilever which poses practical as well as analytical complications. The deflection signal $p(t)$ oscillates on the order of 100 kHz while the scanning systems are typically two orders slower, that is the control bandwidth is in $0.1 - 3$ kHz range. This forms the main motivation for using slow derivative signal (such as the amplitude signal) instead of the deflection signal itself. Obtaining these derivative signals adds further complexities in themodel which makes the analysis of the AM-AFM even more difficult. Although the resulting dynamics and their simplifications have been modeled and analyzed using various tools [56, 85, 86, 87, 88, 89], the models have not been used in designing control for force regulation. In current methodologies, the sample-topography is estimated either in an open-loop approach where sample-heights are determined from amplitude changes or through set-point regulation using proportional-integral-double integral (PII) feedback laws. The open-loop approach is sensitive to errors in calibration experiments and gives distorted images as the average force over an oscillation cycle between the tip and sample is not guaranteed to be constant. The PII methods are not model-based and typically designed by trial-error tuning which does not exploit the dynamics of the underlying process. Typically, the sample-topography affects both the amplitude as well as the phase of the cantilever deflection but only one of them is used for regulation. This procedure does not guarantee true compensation of sample-topography by control and thus can lead to spurious images.

CHAPTER 6

MODELING OF CANTILEVER DYNAMICS FOR DYNAMIC AFM

This chapter is devoted to the discussion on modeling of cantilever dynamics which is represented $\mathcal{F}(g, v, h, \eta)$ in Figure 5.5.

6.1   Tip-sample Interaction Model

The interaction force between the tip and the sample can be explained qualitatively by Lennard-Jones potential, which describes the interaction between a pair of neutral atoms or molecules. It is given by

$$V(r) = 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6}] \qquad (6.1)$$

where ϵ is the depth of the potential well, σ is the distance at which the inter-particle potential is zero, r is the separation between two atoms. The interaction force is given by $F(r) = -\frac{\partial V}{\partial r}$. The example of the tip-sample interaction force in the Lennard-Jones potential model is shown in Figure 6.1.
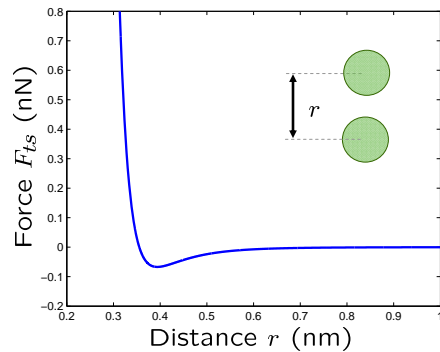

Figure 6.1: Tip-sample interaction force of Lennard-Jones potential model.

The r^{-6} term describes the Van der Waals interaction that results in an attractive force between the atoms. The effect of these long-range forces is perceptible when two particles are separated by distances in the order of 10 nm. When the atoms are too close, the overlap of electron clouds and ionic interactions cause strong repulsive forces between the atom on the cantilever and the atom on the sample surface. The effect of this Pauli exclusion principle is captured by the r^{-12} term in the Lennard-Jones force but the exponent 12 does not have theoretical justification. However, the Lennard-Jones model provides a good qualitative characterization of the tip-sample interaction but does not explain the diverse factors.

Since the tip and the sample interact not as the atom to the atom and the different forces are dominant at different tip-sample separations, generally other contact models are used for analysis of the contact behavior. These models typically consider both the Van der Waals force for long range interaction and repulsion force as Hertzian contact force for short range interaction. The representative models for the interaction between tip and sample include Hertz, Derjaguin-Müller-Toporov(DMT), Johnson-Kendall-Roberts(JKR), Maugis models [90, 91].

The Hertz model specifies the local stresses that develop as two spheres come in contact and deform slightly under the imposed loads. The amount of deformation is dependent on Young's moduli E, and Poisson ratios ν in contact. It gives the contact stress as a function of the normal contact force, the radii of curvature of both bodies and the modulus of elasticity of both bodies.

Using Hertz model, the interaction force can be derived as

$$F_{ts}(z) = \begin{cases} 0, & \text{if } z > 0 \\ \frac{4}{3}E^*\sqrt{R}(-z)^{\frac{3}{2}}, & \text{if } z \leq 0 \end{cases} \tag{6.2}$$

where R is the radius of the tip and $E^* = [\frac{1-\nu_{tip}^2}{E_{tip}} + \frac{1-\nu_{sample}^2}{E_{sample}}]^{-1}$ is the effective Young's modulus. This model does not include any forces except the repulsive elastic force. The example of the tip-sample interaction force in the Hertz model is shown in Figure 6.2.

The DMT model includes a Van der Waals force along with Hertz contact force. It is generally applied to tips with small curvature radius and high stiffness under dry conditions.
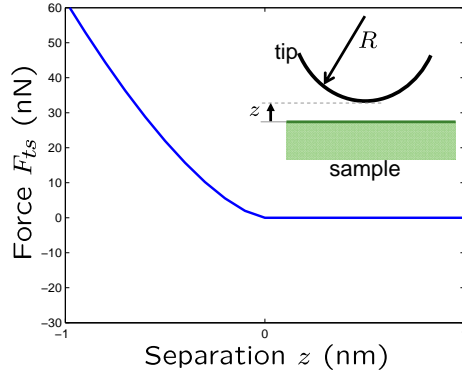
72

Figure 6.2: Tip-sample interaction force of Hertz model.

It is assumed that deformed surface geometries do not differ much from that given by the Hertz model solution. Consideration of the Van der Waals forces outside of the elastic contact area results in an additional probe-sample attraction which weakens forces of elastic repulsion.

$$
F_{ts}(z) = \begin{cases} -\frac{HR}{6z^2}, & \text{if } z > a_0 \\ -\frac{HR}{6a_0^2} + \frac{4}{3}E^*\sqrt{R}(a_0 - z)^{\frac{3}{2}}, & \text{if } z \le a_0 \end{cases} \tag{6.3}
$$

where H is the Hamaker constant (a material constant), and a_0 is the intermolecular distance. An example of tip-sample interaction force of the DMT model is shown in Figure 6.3.
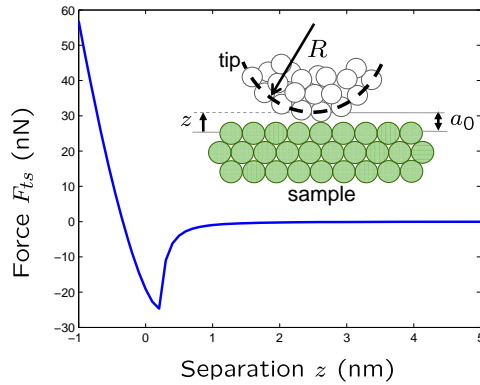

Figure 6.3: Tip-sample interaction force of DMT model.

73

The JKR model also includes a Van der Waals force with Hertz contact force. It is applied to tips with large curvature radius and small stiffness. In this model, the contact is considered to be strongly adhesive. It correlates the contact area to the elastic material properties plus the interfacial interaction strength. The model accounts for the influence of Van der Waals forces within the contact zone. Due to the adhesive contact, contacts can be formed during the unloading cycle also in the negative loading (pulling) regime. Thus, this model has hysteresis.

The Maugis model is in between DMT and JKR model which includes the adhesion related parameter λ. The DMT model is extreme case when $\lambda \to 0$ (stiff contact) and JKR model when $\lambda \to \infty$ (soft contact).

Besides these models, depending on the condition and cases, the forces are added to this model. For example, the model is modified with the capillary force terms for the moist air experiment [92].

6.2   Cantilever Model

A continuous beam equation can be obtained from classical Euler-Bernoulli beam theory with assumption of small deflection of slender, uniform beam. It is described by

$$EI\frac{\partial^4}{\partial x^4}[w(x,t) + \gamma\frac{\partial w(x,t)}{\partial t}] + \mu\frac{\partial^2 w(x,t)}{\partial t^2} = F(x,t), \qquad (6.4)$$

where x is the coordinate along the beam, $w(x,t)$ is the dynamic deflection at (coordinate x, time t), E is the Young's module, I is the moment of inertia, μ is the linear mass density, and γ is the internal damping coefficient of the beam. $F(x,t)$ is the force acting on the cantilever which may include the tip-sample interaction F_{ts}, the damping with the medium F_{damp}, the thermal excitation $F_{thermal}$, and the excitation from dither piezo F_{ext} for dynamic mode operation. The boundary conditions are assumed that one end ($x = 0$) of the cantilever is clamped and the other is free ($x = L$).

The single-mode cantilever model in Figure 6.4 is sufficient for most applications in conventional dynamic mode operation in air or vacuum where the excitation frequency is

near the resonance of the first eigenmode [93]. This point mass model is approximated by lumped mass and simplified from (6.4) and given by

$$\ddot{p} + \frac{\omega_0}{Q}\dot{p} + \omega_0^2 p = \frac{1}{m}[F_{ts} + F_{ext} + F_{thermal}], \tag{6.5}$$

where p is the deflection of the cantilever tip, ω_0 is the first modal frequency, and m is the effective mass of the cantilever and the tip. Q represents the associated quality factor, which is relevant with the damping with the medium F_{damp} and can be approximated by $Q = \frac{\sqrt{mk_c}}{c} =: \frac{1}{2\zeta}$ for the damping coefficient ζ. The quality factors Q can range up to 10 in liquid, 100 to 500 in air, and $10,000$ or higher under vacuum.
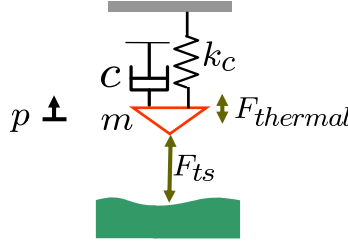

Figure 6.4: Single-mode point mass cantilever model.

In this research, the photodiode is assumed to register the position of the effective mass, while it measures the bending angle of the cantilever end.

The important main concern required for modeling $\mathcal{F}(g, v, h, \eta)$ is the location of vertical piezo-actuator. That is, whether the piezo-actuator is attached to the probe-head or attached to the sample, where the former actuation mechanism moves the cantilever (along with the optical-detection head) to regulate the distance between the sample and the cantilever-tip; while in the latter mechanism the distance is regulated by moving the sample. Closed-loop operation as seen in Figure 5.5 moves the vertical piezo-actuator to maintain tip-sample interaction at a constant value. Depending on whether the actuation is *on the cantilever* or *on the sample*, the model for $\mathcal{F}(g, v, h.\eta)$ is different.

In the vertical positioning on sample case in Figure 6.5(a), the tip-sample interaction
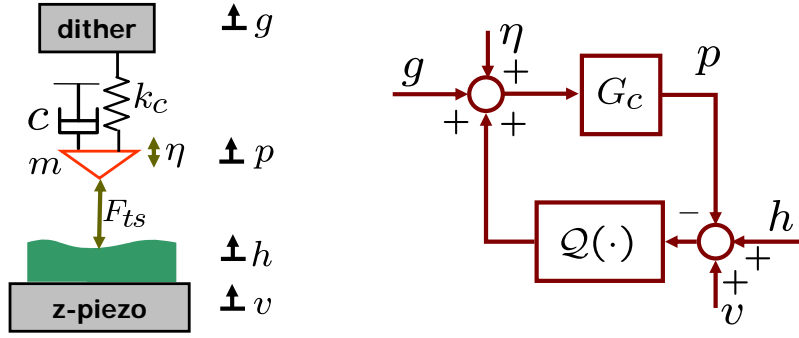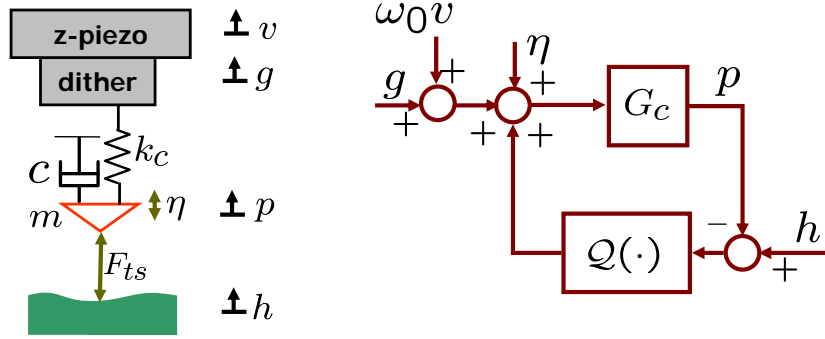
Figure 6.5: Classification by the location of vertical piezo-actuator: (a)vertical positioning on sample (b)vertical positioning on cantilever. The force terms are represented as $g = \frac{1}{m}F_{ext}$, $\eta = \frac{1}{m}F_{th}$, $\Phi(\cdot) = \frac{1}{m}F_{ts}(\cdot)$.

force F_{ts} is affected by vertical-piezo motion and (6.5) becomes

$$\ddot{p} + \frac{\omega_0}{Q}\dot{p} + \omega_0^2 p = \frac{1}{m}[F_{ts}(p - h - v) + F_{ext} + F_{thermal}]. \tag{6.6}$$

In the vertical positioning on cantilever case in Figure 6.5(b), the vertical-piezo movement changes the cantilever base position and (6.5) becomes

$$\ddot{p} + \frac{\omega_0}{Q}\dot{p} + \omega_0^2(p - v) = \frac{1}{m}[F_{ts}(p - h) + F_{ext} + F_{thermal}]. \tag{6.7}$$

Note that the photodiode is located in the same inertial frame with base of cantilever and the measurement corresponds to $p^* = p - v$ not p.

6.3   Asymptotic Perturbation Methods

Since its invention in late 1980s, the dynamic-mode operation has been used prevalently for imaging the soft samples such as biological matters. However, the analysis of it is particularly challenging because the tip traverses a long range of highly nonlinear tip sample interaction potential. The complexity of the dynamics is demonstrated by the experimental and theoretical studies that confirm the existence of chaotic behavior under certain operating conditions [94, 95]. Another indication is the experimentally observed feature of *bistability*, where the cantilever can settle into two different stable trajectories for the same cantilever-sample offset. Besides the analysis of numerical simulation in early 1990s [96], an asymptotic perturbation theory has been used to understand the behavior of this nonlinear system [97, 98, 99].

In dynamic mode operation of the AFM, the attractive forces act on the tip is small and the repulsive forces act only during a very small part of the vibration cycle, this vibration system can be considered as a weak nonlinear system. To approximate analytical solution of the dynamic behavior of the cantilever, the asymptotic perturbation methods is very appropriate since it can explain the motion as the amplitude and phase dynamics with high accuracy. The asymptotic perturbation methods was introduced and used for the open-loop

operation of the AFM where the vertical piezo-actuation does not exist i.e. $v = 0$.

In this study, we evaluate the perturbation methods on the closed-loop systems given on (6.6), (6.7).

The Krylov-Bogoliubov (KB) method [100, 101, 102] starts from the non-conservative second order equation

$$\ddot{x}(t) + \omega_n^2 x(t) = \epsilon f(x, \frac{dx}{dt}), \tag{6.8}$$

for small ϵ. It assumes the solution,

$$x(t) = A \cos \psi + \sum_{i=1}^{\infty} \epsilon^i u_i(A, \psi) \tag{6.9}$$

where $u_i(A, \psi)$ is a periodic function of ψ with the period of $T = 2\pi$. The dynamics for $A(t)$ and $\psi(t)$ are expressed in series expansion with respect to ϵ,

$$\frac{dA}{dt} = \sum_{i=1}^{\infty} \epsilon^i \mathcal{A}_i(A) \tag{6.10}$$

$$\frac{d\psi}{dt} = \omega_n + \sum_{i=1}^{\infty} \epsilon^i \mathcal{B}_i(A). \tag{6.11}$$

The first order approximation (i.e. i=1), which is generally used in analysis, is given by

$$\frac{dA}{dt} = -\frac{\epsilon}{2\pi\omega_n} \int_0^{2\pi} f(A \cos \theta, -A\omega \sin \theta) \sin \theta d\theta \tag{6.12}$$

$$\frac{d\phi}{dt} = -\frac{\epsilon}{2\pi\omega_n A} \int_0^{2\pi} f(A \cos \theta, -A\omega \sin \theta) \cos \theta d\theta \tag{6.13}$$

where $\phi := \psi - \omega_n t$.

Bogoliubov and Mitropolskii extended it for the case of non-autonomous system with periodic excitation

$$\ddot{x}(t) + \omega_n^2 x(t) = \epsilon \tilde{f}(x, \frac{dx}{dt}, \omega t). \tag{6.14}$$

The assumptions are modified by

$$x(t) = A \cos \psi + \sum_{i=1}^{\infty} \epsilon^i u_i(A, \psi, \omega t) \tag{6.15}$$

78

where $u_i(A, \psi, \omega t)$ is a periodic function of ψ and ωt with the period of $T = 2\pi$. The same series expansions on A and ψ are used as (6.10), (6.11).

For the harmonic excitation $\tilde{f}(x, \frac{dx}{dt}, \omega t) = f(x, \frac{dx}{dt}) + E\cos(\omega t)$ and the main resonance (i.e. $\omega \approx \omega_n$) case , the first order approximation (i.e. i=1) is given by

$$\frac{dA}{dt} = -\frac{\epsilon}{2\pi\omega} \int_0^{2\pi} f(A\cos\theta, -A\omega\sin\theta)\sin\theta d\theta - \frac{\epsilon E}{2\omega}\sin\phi \tag{6.16}$$

$$\frac{d\phi}{dt} = \omega_n - \omega - \frac{\epsilon}{2\pi\omega A} \int_0^{2\pi} f(A\cos\theta, -A\omega\sin\theta)\cos\theta d\theta - \frac{\epsilon E}{2\omega A}\cos\phi. \tag{6.17}$$

The Krylov-Bogoliubov-Mitropolskii (KBM) method can be applied for (6.6), (6.7). Consider the vertical positioning on sample type in Figure 6.5(a), first. For non-dimensionalization, substitute $\omega_0 t =: \tau$ and $\frac{\omega}{\omega_0} =: \Omega$. Then, (6.6) becomes

$$\frac{d^2p}{d\tau^2} + 2\zeta\frac{dp}{d\tau} + p = \frac{1}{m\omega_0^2}F_{ts}(p - h - v) + g_0\cos\Omega\tau + \eta. \tag{6.18}$$

where $\frac{1}{m}F_{ext} =: g_0\cos\omega t$ and $\frac{1}{m}F_{thermal} =: \eta$.

Rearranging (6.18) gives

$$\frac{d^2p}{d\tau^2} + p = \frac{1}{m\omega_0^2}F_{ts}(p - h - v) - 2\zeta\frac{dp}{d\tau} + \eta + g_0\cos\Omega\tau$$

$$= \epsilon f(p, \frac{dp}{d\tau}) + \epsilon E\cos\Omega\tau \tag{6.19}$$

where $f(p, \frac{dp}{d\tau}) := \frac{1}{\epsilon}[\frac{1}{\omega_0^2 m}F_{ts}(p-h-v) - 2\zeta\frac{dp}{d\tau} + \eta]$ and $E := \frac{g_0}{\epsilon}$. The assumption of small ϵ holds due to the small excitation g_0 in the cantilever support compared to the amplitude of the tip, the large resonance frequency ω_0, the very small damping ζ, and the small thermal noise η. However, an implicit assumption here is that h and v are constant parameters, that is their dynamics can be ignored (the time-scale separation between the fast dither dynamics and relatively slow actuation dynamics (on the order of 5% of the cantilever resonance frequency) partially justifies this assumption.

The KBM method applied to (6.19) results in the following equations.

$$\frac{dA}{d\tau} = -\zeta A - \frac{\epsilon E}{2\Omega} \sin\phi \tag{6.20}$$

$$\frac{d\phi}{d\tau} = 1 - \Omega - \frac{1}{2\pi\Omega A} \int_0^{2\pi} \frac{1}{\omega_0^2 m} F_{ts}(A\cos\theta - h - v)\cos\theta d\theta - \frac{\epsilon E}{2\Omega A}\cos\phi. \tag{6.21}$$

With original parameters and time scale,

$$\frac{dA}{dt} = -\zeta\omega_0 A - \frac{\omega_0^2 g_0}{2\omega} \sin\phi \tag{6.22}$$

$$\frac{d\phi}{dt} = \omega_0 - \omega - \frac{1}{2\pi\omega A} \int_0^{2\pi} \frac{1}{m} F_{ts}(A\cos\theta - h - v)\cos\theta d\theta - \frac{\omega_0^2 g_0}{2\omega A}\cos\phi. \tag{6.23}$$

The vertical positioning on tip type in Figure 6.5(b) is more complicated. The dynamics of global displacement p (6.7) is given by

$$\ddot{p} + 2\zeta\omega_0\dot{p} + \omega_0^2(p - v) = \frac{1}{m}F_{ts}(p - h) + g_0\cos\omega t + \eta \tag{6.24}$$

and by substitution of $p^* = p - v$, which is the displacement with respect to the head displacement v since the photodiode moves along with the head,

$$\ddot{p}^* + \omega_0^2 p^* = \frac{1}{m}F_{ts}(p^* - h + v) - 2\zeta\omega_0\dot{p}^* - (\ddot{v} + 2\zeta\omega_0\dot{v}) + \eta + g_0\cos\omega t \tag{6.25}$$

With the same non-dimensionalization and $f(p, \frac{dp}{d\tau}) := \frac{1}{\epsilon}[\frac{1}{\omega_0^2 m}F_{ts}(p - h + v) - 2\zeta\frac{dp}{d\tau} + \eta - (\frac{d^2 v}{d\tau^2} + 2\zeta\frac{dv}{d\tau})]$, the assumption of small ϵ holds again. Here, the displacement v of the vertical positioner is the output of the vertical piezo-actuator G_p. Since G_p can transmit the low frequency signal, typically less than 2 kHz, it can be assumed as a slow varying variable. Thus, its acceleration $\ddot{v}$, and the velocity $\dot{v}$ are small enough to be neglected. (Details are provided in Appendix A.2.)

The KBM method when applied to the case when the actuation is on the cantilever yields

the same equations as (6.22), (6.23) except that v is replaced by $-v$,

$$\frac{dA}{dt} = -\zeta\omega_0 A - \frac{\omega_0^2 g_0}{2\omega}\sin\phi \tag{6.26}$$

$$\frac{d\phi}{dt} = \omega_0 - \omega - \frac{1}{2\pi\omega A}\int_0^{2\pi}\frac{1}{m}F_{ts}(A\cos\theta - h + v)\cos\theta d\theta - \frac{\omega_0^2 g_0}{2\omega A}\cos\phi. \tag{6.27}$$

6.4 Validation

We verified this model through numerical simulations. This was done by using DMT (Derjaguin-Müller-Toporov) model model to evolve the dynamics of the cantilever deflection p; and then overlaying the amplitude a dynamics from the KBM model and checking if it matches with (envelops) the p dynamics from the DMT data. The DMT model forms a good benchmark since it is experimentally well verified and extensively corroborated in the literature [103, 96, 56, 86]. We also show that the KBM model predicts the *bistability* phenomenon which is observed in experiments.

Bistability

The obtained KBM equation is used in open loop simulation. Figure 6.6 shows steady-state amplitude and phase of the cantilever oscillation for different tip-sample separations h, as derived from KBM equation. This behavior can be interpreted as the saddle node bifurcation. Note that for large values of tip-sample separation, there is only one equilibrium point (stable), while there is a range of separation distances for which there two stable equilibrium points (bistability) and one unstable equilibrium.These results match the single mode dynamic model numerical simulation and experiment results shown in [96, 104, 86]. (More details about stability of cantilever oscillation can be found in these references.)

Dynamics of cantilever depending on varying sample topography

Amplitude and phase dynamics obtained from KBM methods show a good match with dynamics of the cantilever deflection signal p for slow varying sample height as shown in
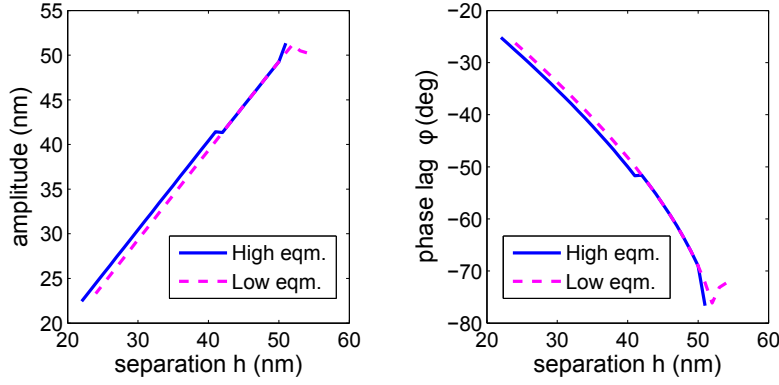
Figure 6.6: Equilibrium amplitude and phase solution of cantilever oscillation: Depending on the separation between tip and sample, the dynamics of cantilever can be mono-stable or bistable.

Figure 6.7. An arbitrary sample-profile h was given and the deflection signal $p(t)$ determined using (6.5), and compared the amplitude signal $A(t)$ and phase signal $\phi(t)$ as determined from the KBM method in (6.22) and (6.23) with $v = 0$. The results show the good agreement between the amplitude signal derived from $p(t)$ and $A(t)$. However, for sudden change in
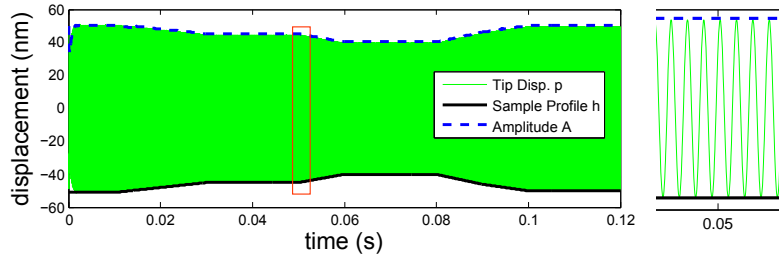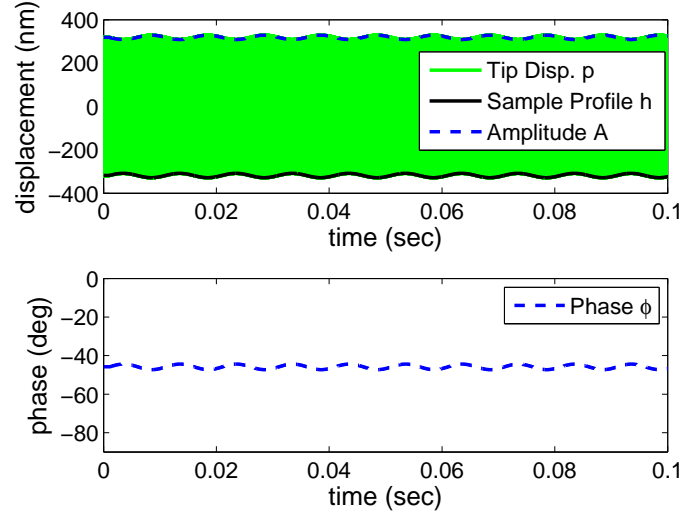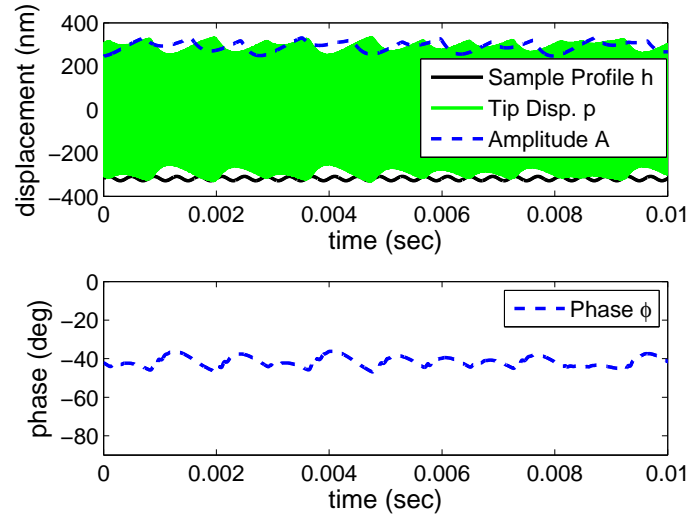

Figure 6.7: Comparison of original and KBM dynamics for slow varying height: The high-frequency signal (solid green) is the deflection $p(t)$ simulated using (6.5) and the dotted signal is amplitude as derived from the KBM methods (6.22), (6.23) with $v = 0$. The amplitude envelops the defection $p(t)$.

sample height (such as step input), it takes time to converge to match with the deflection $p(t)$. It suggests that the KBM method predicts well when the sample profile signal h is slow-varying (in fact, the sample height h was assumed to be a constant parameter in the development of KBM method). However, for fast varying $h(t)$, the KBM method does not approximate well. In order to see the range of frequencies, we simulated the KBM and

82

DMT models for sinusoidal sample-profile signals with different frequencies (for small height changes). In Figure 6.8, an noiseless sinusoidal sample-profile h with amplitude of 10 nm is chosen as test height signal and the simulations were applied for various frequencies. For low frequency (100 Hz) case in Figure 6.8(a), the amplitude signal derived from $p(t)$ is in the good agreement with $A(t)$ from KBM method, and the relation from sample height h to amplitude A is linear. However, the fast change case of 2500 Hz in Figure 6.8(b), the amplitude signal derived from $p(t)$ does not match with $A(t)$ from KBM method, and the amplitude A and oscillation $p(t)$ do not have the same frequency as the sample input h. The difference between two methods become significant for height signal frequencies over 2 kHz. However, it also depends on the amplitude of height change, which is obvious since the higher force from higher amplitude is the nonlinear terms in (6.5). These simulations conform that the KBM method is valid only for slow ($\ll$ 2 KHz) sample profile changes.

Figure 6.8: Comparison of original and KBM dynamics for sinusoidal height: Sample height is modeled as sinusoidal signal with 10 nm amplitude and frequency of 100 Hz in (a) and 2.5 kHz in (b). The cantilever oscillation has an amplitude of 330 nm with 70 kHz.

CHAPTER 7

ALTERNATIVE SIGNALS FOR SAMPLE-TOPOGRAPHY ESTIMATION IN DYNAMIC AFM

In this chapter, we address fast and accurate imaging in dynamic mode operation. In current amplitude modulation dynamic mode, the control signal u for amplitude regulation is typically used as the estimate for the sample topography. Therefore, the sample profile estimation signal is accurate as long as the control signal achieves good amplitude regulation. In the design scheme proposed in this chapter, the limitations of maintaining a constant amplitude while scanning at high bandwidth does not impose limitations on the reconstruction of the sample topography. In fact, we analytically prove that the sample-profile signal estimation problem can be solved independently of the control design scheme for amplitude regulation. Therefore, accurate sample-profile estimations can be obtained even at frequencies near and beyond the closed-loop control bandwidths. However, we show that the robustness of estimation does depend on the control design for regulation and in fact, the robustness of estimation is described by the closed-loop sensitivity transfer function. The new estimation method does not require any modification on control loop and guarantees more accurate image at higher speeds. The independence of the profile-estimation problem from the control design is another salient distinguishing characteristic. The estimation bandwidths by this new scheme can be increased by several times but is limited by noise corruption. Comparison with the existing methods of using the control signal as the image is provided. The experimental results corroborate the theoretical development.
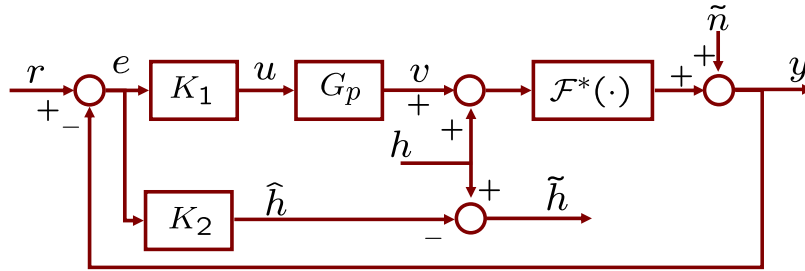
7.1   New Estimation Signal

In block diagram modeling of dynamic mode operation in Figure 7.1(a), the controller K_1 is designed to regulate the amplitude of cantilever deflection p to compensate the effects

of the sample topography h. In conventional method (without K_2 and $\hat{h}$ in Figure 7.1(a)), since the controller signal u is also used as the estimate for the sample topography (obtained by multiplying the control u by pre-calibrated vertical piezo-positioner sensitivity), the controller K_1 needs to achieve both good set-point regulation as well as good estimation of the sample topography.

The new approach proposes the use of a separate estimator K_2 to fully utilize the information in the system as seen in Figure 7.1(a). It is assumed that the set-point regulation controller K_1 is already given or fixed. In this design, a new estimator K_2 and a sample-topography estimate signal $\hat{h}$ are introduced, where a norm on error $h - \hat{h}$ serves a metric of the accuracy of estimation.
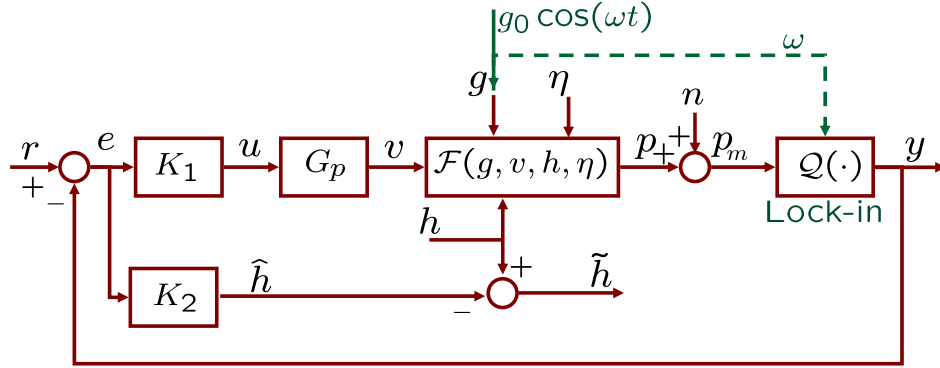
Figure 7.1: Block diagram for the model-based scheme for sample-topography estimation: K_2 is a separate estimator and $\hat{h}$ is a new estimate signal. The multi-input cantilever dynamics model $\mathcal{F}(g, v, h, \eta)$ and lock-in amplifier $\mathcal{Q}(\cdot)$ in (a) is approximated by the single input model $\mathcal{F}^*(\cdot)$.

In Figure 7.1(b), we present a block diagram that represents a model for AM-AFM. Here $\mathcal{F}^*(\cdot)$ represents the map whose output is the amplitude of the deflection signal when its input is the sum of the sample-topography and the piezo-actuation signals. This nonlinear map is obtained using asymptotic perturbation methods described in Chapter 6. This model can be thought of representing the combination of the cantilever dynamics along with the lock-in amplifier in Figure 7.1(a). Here, $\tilde{n}$ represents the uncertainties for using $\mathcal{F}^*(\cdot)$ as well as the effect of the noise n. Under assumptions that the given controller K_1 and the vertical piezo model G_p are linear maps, the input u is given as $u = K_1(r - n - \mathcal{F}^*(h + G_p(u)))$. Also, we represent the nonlinear map $\mathcal{F}^*$ as linear transfer function with multiplicative uncertainty $\mathcal{F}^* = F(1 + w_i \Delta)$ as in Figure 7.2, where the stable weight transfer function w_i reflects frequency dependence on uncertainty. Based on our identification experiments, the cantilever dynamics $\mathcal{F}(g, v, h, \eta)$ is nearly linear at low frequencies (≤ 2 kHz) and hence the weight w_i can be chosen to be small at this frequency range. In this framework, input u can be written as

$$u = \begin{bmatrix} SK_1 & -SK_1F \end{bmatrix} \begin{bmatrix} r - \tilde{n} \\ h \end{bmatrix} \tag{7.1}$$

where $S = (1 + K_1FG_p)^{-1}$ represents the sensitivity function. In conventional estimation,

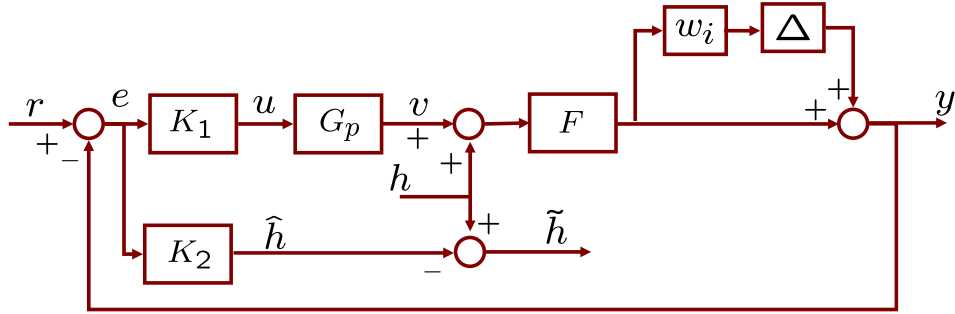Figure 7.2: Block diagram with multiplicative model uncertainty: $\mathcal{F}^* = F(1 + w_i \Delta)$ with any stable function with $\|\Delta\| < 1$.

the transfer function from the sample profile h to the control signal u is given by $SK_1F = K_1F/(1 + G_pK_1F)$. For achieving amplitude regulation, K_1 is required to be high at low frequencies, which implies that the transfer function from topography h to the control u can

be approximated by $1/G_p$ at low frequencies. Since the frequency response of the piezo-actuator is approximately a constant at low frequencies (up to its bandwidth), $|G_p(0)|u$ is used as an estimate of the sample topography h. However, at high frequencies, K_1 can not be designed to be large as it can make the closed-loop unstable especially in the view of modeling uncertainties $w_i\Delta$. In amplitude modulation dynamic mode, the large K_1 typically induces a *chatter-phenomenon* in imaging even though it does not make system unstable and the transfer function from the topography signal to the control signal u is not a constant. Therefore, this control signal (that is, the input to the piezo-actuator), which is typically used as an estimate for sample topography, gives low fidelity images during fast scanning. Note that the temporal frequency content of h depends on the spatial frequency content of the sample, i.e. how rough the sample is, and the scanning rate of the lateral positioners, i.e. how fast the sample is scanned. Thus, easy solution for good imaging is to use a slow lateral scanning rate (which will make h the low frequency signal) with moderate K_1 controller. However, this solution comes by sacrificing bandwidth, which is not tenable in many applications.

Now we analyze the proposed signal $\hat{h}$ for sample-topography estimation (see Figure 7.2). The estimate signal $\hat{h}$ can be written as

$$\hat{h} = \begin{bmatrix} K_2S & -K_2FS \end{bmatrix} \begin{bmatrix} r - \tilde{n} \\ h \end{bmatrix}, \tag{7.2}$$

where the sensitivity transfer function S depends only on K_1 and not on K_2.

An estimate $\hat{h}$ can be obtained by designing $K_2 = -S^{-1}F^{-1}$ since it minimizes the estimation error $\tilde{h}$ from h given by

$$\tilde{h} = \begin{bmatrix} -K_2S & 1 + K_2FS \end{bmatrix} \begin{bmatrix} r - \tilde{n} \\ h \end{bmatrix}. \tag{7.3}$$

Note that S is easily invertible since it is bi-proper and has no non-minimum zero when K and G are stable, which is typically true. If F is strictly proper, low pass filter weight function W of the order equal to the relative degree of F can be used to make $K_2 = -S^{-1}F^{-1}W$. If

F has a non-minimum zero, then the best K_2 can be obtained through a Nevanlinna-Pick solution to a model matching problem $arg\min_{/k_2}\|1 + K_2FS\|$ (e.g. see Section 3.1).

One important consequence of this design is that it decouples the objectives of regulation and sample-topography estimation. The design of controller K_1 for regulation can be made without any consideration towards estimation of sample topography. In this design, the estimation bandwidth is not limited by the regulation bandwidth. Another advantage of this separation of designs is that the improvement in the estimation bandwidth is achieved without increasing the range of frequencies where the magnitude of control $|K_1|$ has high values, which means that it avoids the stability or chattering issues discussed earlier.

Equation (7.3) clearly shows the trade-off between image estimation bandwidth and noise attenuation. If we design $K_2 = -S^{-1}F^{-1} = (1 + K_1FG_p)F^{-1}$ to achieve high estimation bandwidth, then the estimate signal $\hat{h}$ will be corrupted by the effect of noise $K_2 S\tilde{n} = F^{-1}\tilde{n}$. Thus, augmenting the estimator with a weight function W_h that carries the information content of height signal h results in

$$K_2 = -S^{-1}F^{-1}W_h \tag{7.4}$$

and the corresponding estimation error is given by $\tilde{h} = -F^{-1}W_h(r - \tilde{n}) + (1 - W_h)h$. If we choose the weight function W_h as low pass filter, then the effect of noise $F^{-1}W_h\tilde{n}$ rolls off at high frequencies.

In summary, new method separates the goals of force regulation and sample-topography estimation by designing two signals - u for force regulation and $\hat{h}$ for sample topography estimation. The control signal bandwidth is limited due to practical requirements such as stability and robustness to modeling uncertainties. Therefore, the control signal serves as a poor estimate of sample topography near and beyond the control-bandwidth frequency. The separation of goals allows the estimation signal $\hat{h}$ to depend directly on the whole frequency range of the amplitude signal (both below and beyond the control-bandwidth, thus more sample-topography information) and give accurate estimates of the sample topography.

7.2 Device Description

We demonstrated this design for sample-topography estimation on MFP-3D, an AFM developed by Asylum Research Inc. A schematic of this device is shown in Figure 7.3. Note that in this device, the vertical piezo-actuation is on the cantilever and not the sample.
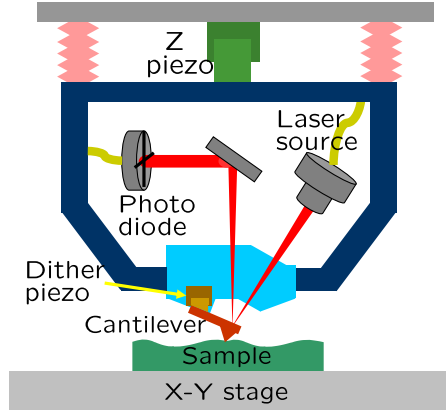

Figure 7.3: A schematic of imaging system of MFP-3D.

The voltage range of vertical piezo is -10 to 150 V and travel range is 25 μm. The cantilever deflection p is sampled at 5 MHz with 16 bit resolution for amplitude and phase calculations using a lock-in amplifier that is built on Field Programmable Gate Array (FPGA) and Digital Signal Processor (DSP). Amplitude signal is down-sampled to 100 kHz. The controller K_1 and the estimator K_2 which run at 100 kHz are implemented on DSP. This DSP code takes the amplitude signal and the set-point value as input and generates the control effort, which is converted by 20 bit 100 kHz DAC and applied with the vertical piezo after amplification.

The cantilever used for the experiment is AC240TS by Olympus co., which is approximately 240 μm long, 30 μm wide and its tip is approximately < 10 nm in diameter. The nominal value of resonance frequency is 70 kHz.

Most commercial AFMs offer the proportional-integral (PI) or proportional-integral-integral (PII) controllers as a default controller. The controller K_1 needs to be decided depending on the cantilever used since it determines the dynamics $\mathcal{F}(g, v, h, \eta)$ for the same

system with the vertical piezo-actuator G_p and the lock-in amplifier Q. The controller K_1 is tuned as $K_1(s) = -\frac{6000}{s}$, which did not induce the chattering and gave a good quality image at slow scan.

7.2.1 Vertical piezo identification

Similar to the X-Y piezo scanner in which physical modeling of the device is difficult, identification techniques were used to derive the transfer function from the vertical-piezo input u to the reading of z position sensor. At various points in the operating range, we obtained the frequency response of the vertical piezo over $0 - 10$ kHz (Figure 7.4(a)). The nominal frequency response of the device was chosen, at which the amplitude modulation would be set. Figure 7.4(b) shows the bode diagram of fitted mathematical model with nominal experimental result. Weighted iterative least square fitting was performed over $0 - 7.5$ kHz and resulted in following model:

$$
\begin{aligned}
G_z(s) = & \frac{-2.3808 \times 10^7 (s - 2.244 \times 10^4)(s^2 + 352.1s + 1.212 \times 10^8)}{(s + 5.455 \times 10^4)(s^2 + 523.4s + 1.164 \times 10^8)(s^2 + 7252s + 1.893 \times 10^8)} \\
& \times \frac{(s^2 + 3961s + 2.079 \times 10^8)(s^2 + 257.3s + 2.595 \times 10^8)(s^2 - 73.47s + 5.126 \times 10^8)}{(s^2 + 1056s + 1.993 \times 10^8)(s^2 + 240.9s + 2.637 \times 10^8)(s^2 + 1163s + 5.119 \times 10^8)} \\
& \times \frac{(s^2 - 2020s + 1.157 \times 10^9)(s^2 + 546s + 1.536 \times 10^9)(s^2 + 701.9s + 2.183 \times 10^9)}{(s^2 + 509.7s + 1.326 \times 10^9)(s^2 + 3090s + 1.678 \times 10^9)(s^2 + 2643s + 2.136 \times 10^9)}.
\end{aligned}
\tag{7.5}
$$

The z-position sensor has the sensitivity of 1.3×10^{-6} m/V. The transfer function between the vertical-piezo input u to its displacement v is given by

$$
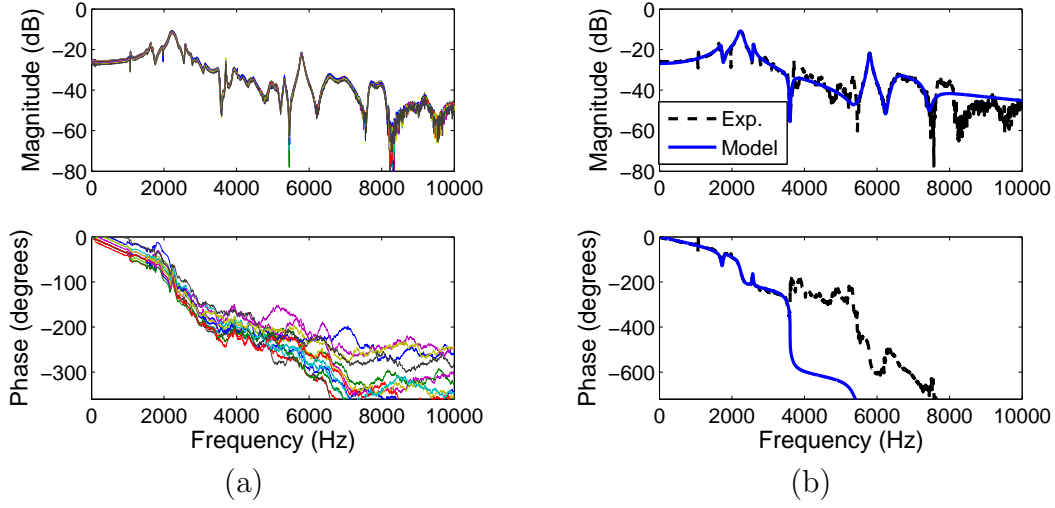G_p(s) = 1.3 \times 10^{-6} \, G_z(s).
\tag{7.6}
$$

Figure 7.4: Identification of vertical piezo: (a) Experimental frequency responses at various operating positions (b) Nominal frequency response(dashed) and model frequency response(solid).

7.3 Estimator Design and Validation

7.3.1 Estimator design

The cantilever dynamics model $\mathcal{F}^*$ is assumed as the KBM approximation described in Chapter 6. The DMT model is used the tip sample contact force model with parameters of the Hamaker constant of $H = 6.4 \times 10^{-20}$ J, the tip radius of $R = 10.0 \times 10^{-9}$ m, the effective Young's modulus of $E^* = 50.0 \times 10^{-9}$ Pa, and the intermolecular distance of $a_0 = 0.16 \times 10^{-9}$ m. The cantilever parameters were measured as the quality factor $Q = 158.5$, the stiffness $k = 1.99$ N/m, and the first modal frequency $f_0 = 75.731$ kHz - These parameters varied slightly on each attachment.

Initially, F was tried with the linearized model of the KBM approximation, but the estimator K_2 did not give a good estimation. It can be explained by the fact that the dynamics of the height h and the vertical piezo-position v is not negligible at high frequencies. In the asymptotic perturbation methods such as the one used to derive the KBM model, the error between the original dynamics (6.6), (6.7) and the approximated dynamics (6.22), (6.23) depends on the order of ϵ expansion and decays to zero in steady state. Thus, it has

good fidelity in slow varying inputs, where the decay is faster than the input changes, but not in fast varying inputs. For fast varying inputs or high frequency inputs, higher order approximation is required. For these reasons, F was chosen as the dc gain in the linearization of the KBM approximation ($F = -0.998 \approx -1$). W_h chosen as the low pass filter with 2 kHz bandwidth, $W_h = \frac{1}{7.958 \times 10^{-5} s + 1}$. Since the output of F is the amplitude of the displacement and required to converted to the voltage signal, F used in in the estimator design is scaled by the optical lever sensitivity, 359.7×10^{-9} m/V, which depends on each attachment of cantilever.
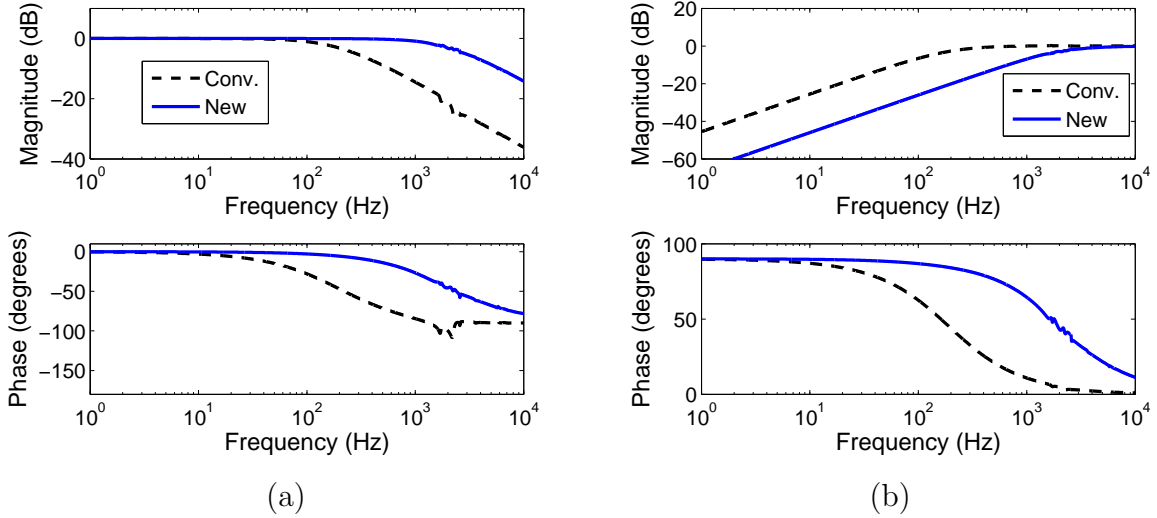

Figure 7.5: Comparison of transfer function of conventional(dashed) and new(solid) estimation: (a) transfer function from the sample height h to the conventional height estimate $\hat{u}$ (i.e. $-SK_1 F \times$ (Piezo Sensitivity))(dashed) and to the new height estimate $\hat{h}$ (i.e. $K_2 FS$)(solid) (b) transfer function from the sample height h to the conventional estimation error $h - \hat{u}$(dashed) and to the new height estimate $\tilde{h}$(solid).

The resulting estimator K_2 is obtained and the balanced model reduction results in the following 6th order model:

$$K_2 = \frac{5.4639(s + 4.639 \times 10^5)(s + 1211)(s^2 + 1357s + 1.186 \times 10^8)(s^2 + 2554s + 1.84 \times 10^8)}{(s + 0.001)(s + 6609)(s^2 + 1109s + 1.2 \times 10^8)(s^2 + 2116s + 1.96 \times 10^8)}.$$
(7.7)

Figure 7.5(a) shows the transfer function from the sample height h to the conventional height estimate $\hat{u}$ that scaled by vertical piezo sensitivity from the control input u (i.e.

93

$-SK_1F\times$ (Piezo Sensitivity)) and to the new height estimate $\hat{h}$ (i.e. K_2FS). Figure 7.5(b) shows the transfer function from the sample height h to the conventional estimation error $h - \hat{u}$ and to the new height estimate $\tilde{h}$. The imaging bandwidth can be defined by the range of frequencies over which the magnitude of this estimation error transfer function is below $1/\sqrt{2}$. According to this definition, the conventional estimation has the bandwidth of 187 Hz and the new estimation has the bandwidth of 1876 Hz. However, this calculation is based on the simplified F without modeling uncertainty, thus the real imaging bandwidth of the experiment may be different. (See the experimental results from using the calibration sample TDG01 in the next section.)

To validate this new estimation method, several samples are imaged. Every experiment was performed in raster scan with various scanning rate. The conventional height estimate $\hat{u}$ and the new height estimate $\hat{h}$ was obtained simultaneously while scanning.

7.3.2   1D imaging results

The calibration grating TDG01 from NT-MDT co. was imaged. It has 1 dimensional parallel ridge pattern with the height of approximately 50 nm and the period of 278 nm (3600 lines/mm). The structure was formed on the glass wafer. Figure 7.6 shows slow scan (4 μm/sec) image constructed from the conventional estimate signal $\hat{u}$.
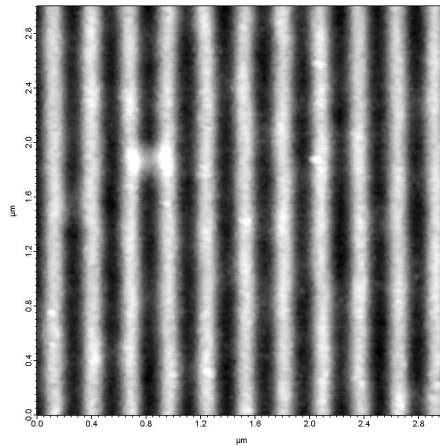

Figure 7.6: Slow scan (4 μm/sec) image of TDG01 constructed from the conventional estimate signal $\hat{u}$.
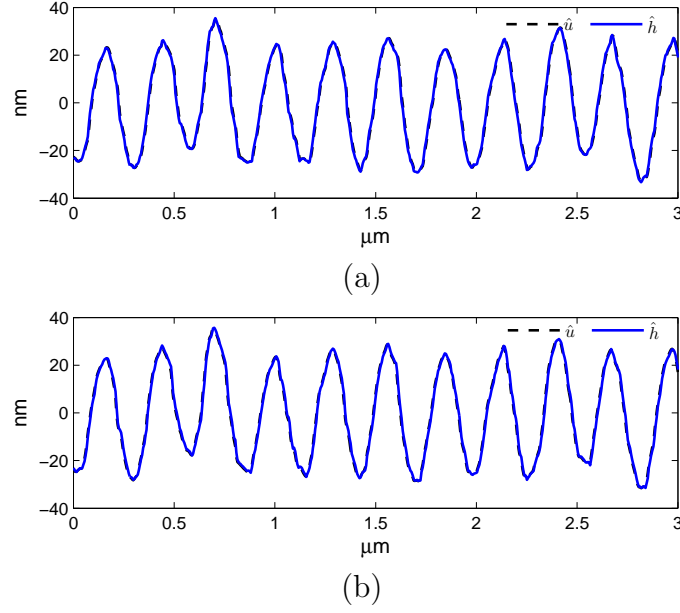
Figure 7.7: The conventional estimate signal $\hat{u}$(dashed) and the new estimate signal $\hat{h}$(solid) of slow scan of TDG01: (a) trace (left to right) direction (b)retrace (right to left) direction.

The conventional estimate signal $\hat{u}$ and the new estimate signal $\hat{h}$ of slow scan (4 μm/sec) are compared in trace (left to right) direction in Figure 7.7(a) and in retrace (right to left) direction in Figure 7.7(b). Since the two signals are almost identical in both trace and retrace, the average of these four signals is considered as the real feature and used as a reference for the calculation of error in fast scan. The deviations from the reference are 2.3013 nm ($\hat{u}$-reference in trace direction), 1.6855 nm ($\hat{u}$-reference in retrace direction), 1.1104 nm ($\hat{h}$-reference in trace direction), and 0.7641 nm ($\hat{h}$-reference in retrace direction) in root mean square (RMS) value.

Figure 7.8 compares the conventional estimate signal $\hat{u}$ and the new estimate signal $\hat{h}$ of trace scan lines in left column, and the error of conventional estimate $\hat{u} - $ reference and the error of the new estimate $\hat{h} - $ reference in right column.

Since the calibration grating TDG01 is manufactured for X-Y positioner calibration, its spacial pitch of 278 nm is quite accurate. Thus, we can calculate the temporal excitation frequency from this pitch and the scan speed, thus a scan speed of 10 μm/sec (shown in the scan in (a,b)) corresponds to a temporal sinusoidal frequency 36 Hz, and similarly 20
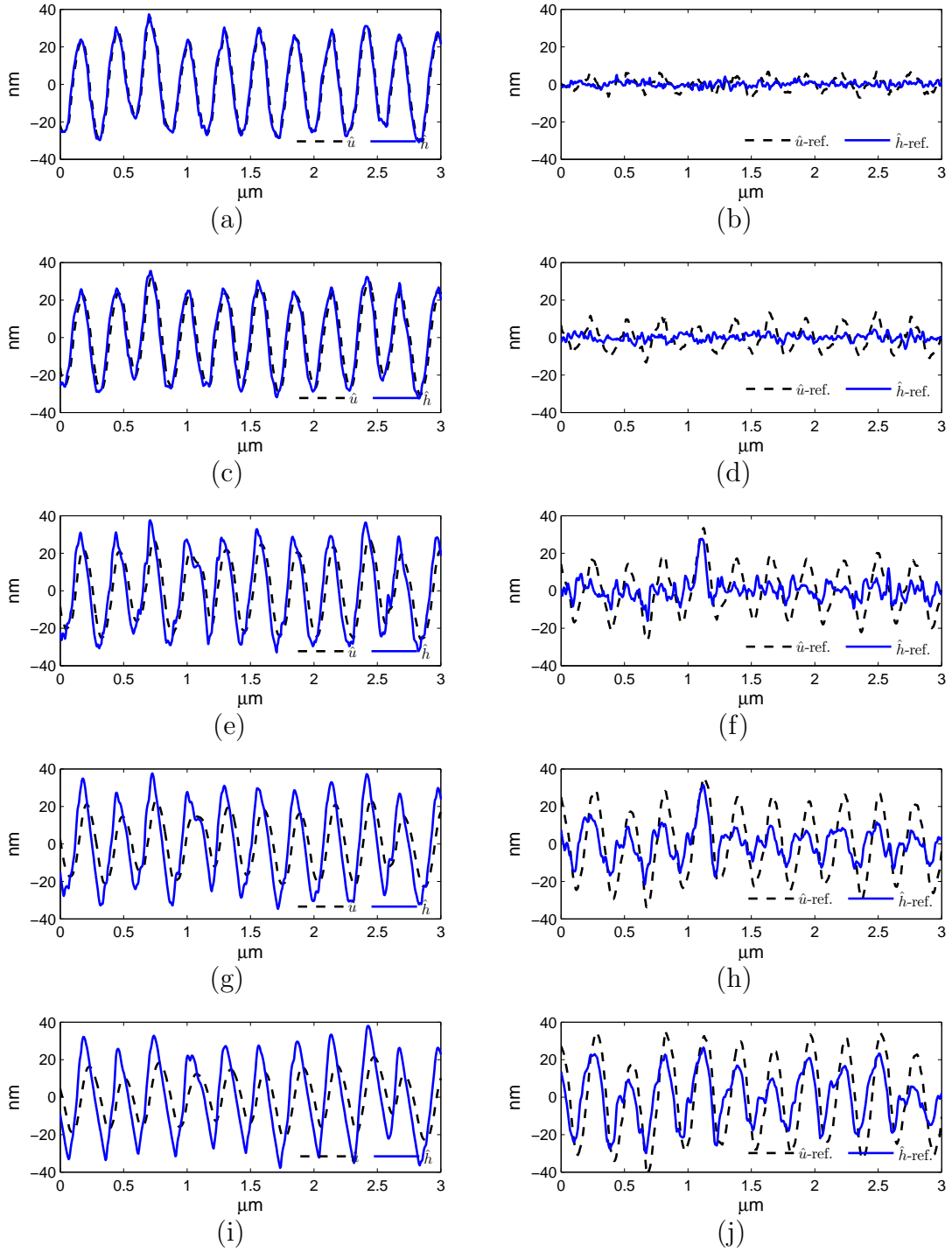
Figure 7.8: Comparison of the conventional $\hat{u}$(dashed) and new estimate signal $\hat{h}$(solid): The left graphs(a,c,e,g,i) compare $\hat{u}$(dashed) and the $\hat{h}$(solid). The right graphs(b,d,f,h,j) compare $\hat{u}$-reference(dashed) and $\hat{h}$-reference(solid). (a,b) scan rate of 10 μm/sec (c,d) scan rate of 20 μm/sec (e,f) scan rate of 40 μm/sec (g,h) scan rate of 60 μm/sec (i,j) scan rate of 80 μm/sec.

μm/sec (scan in (c,d)) to 72 Hz, 40 μm/sec (scan in (e,f)) to 144 Hz, 60 μm/sec (scan in (g,h)) to 216 Hz, and 80 μm/sec (scan in (i,j)) to 284 Hz. In addition to these scan speeds, the scan speeds of 110 μm/sec (396 Hz), 140 μm/sec (504 Hz), and 180 μm/sec (647 Hz) were also used.

Using Fourier transform on the above *input* sinusoid time signals, the transfer function from the sample height h to the conventional height estimate $\hat{u}$ and to the new height estimate $\hat{h}$ is obtained as shown in Figure 7.9(a). Figure 7.9(b) shows the transfer function from the sample height h to the conventional estimation error reference $-\hat{u}$ and to the new height estimate reference $-\hat{h}$. The bandwidth of conventional estimation is approximately 177 Hz and the new estimation approximately 301 Hz.

In Figure 7.9(b), low magnitude ($< -40dB$) values at low frequencies is not accurate since -40 dB corresponds to the 1% of the amplitude, which is around 0.25 nm and close the resolution limit of the intermittent mode in the machine. In addition, it assumed that the input is same as the calculated reference which has the maximum RMS error of 2.3 nm, which corresponds to -21 dB. At even higher frequencies, the approximation errors due to linearization dominate (as seen in Figure 7.8(e,f), the input-output response is not exactly linear). Therefore, the comparison of measured images might be an appropriate method to validate the new design.
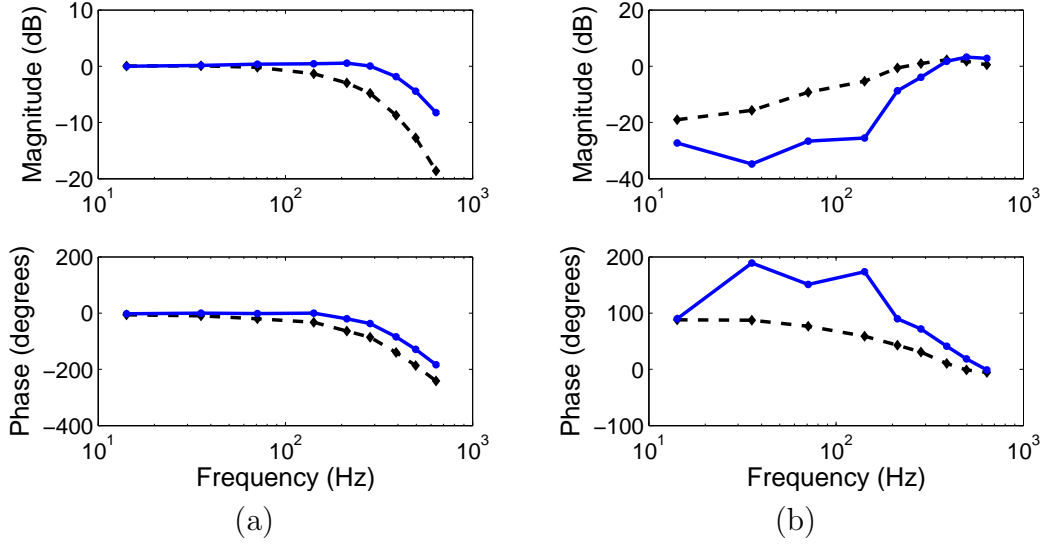
Figure 7.9: Experimentally obtained transfer function of conventional(dashed) and new(solid) estimation: (a) transfer function from the sample height h to the conventional height estimate $\hat{u}$(dashed) and to the new height estimate $\hat{h}$(solid) (b) transfer function from the sample height h to the conventional estimation error $h - \hat{u}$(dashed) and to the new height estimate $\tilde{h}$(solid).

7.3.3 2D imaging results

The 2D calibration grating, 607-AFM from Ted-pella inc. was imaged. It is a cellulose acetate replica of a waffle pattern diffraction grating. It has the height of approximately 31 nm and the line spacing of 463 nm (2160 lines/mm).

Figure 7.10 shows 2D image constructed from the conventional estimation $\hat{u}$ in left column and the new estimation $\hat{h}$ in right column. The difference are not noticeable in slow scan images in Figure 7.10(a,b). However, fast scan images show that the new estimation images in Figure 7.10(d,f) is more clear than the images in figure 7.10(c,e) and closes to the slow scan images in figure 7.10(a,b).
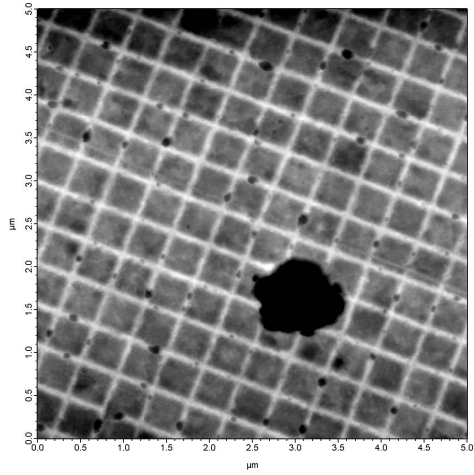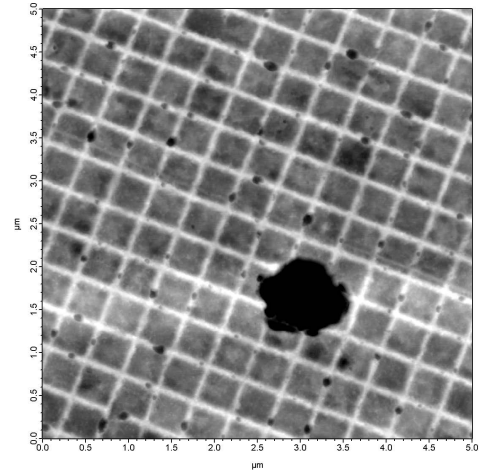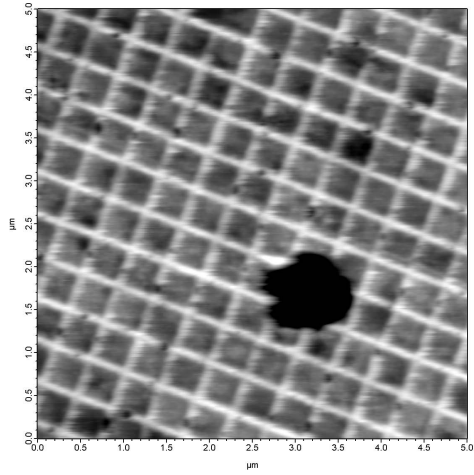
Figure 7.10: Comparison of 2D image of conventional(left) and new(right) estimation: (a,b) scan rate of 5 μm/sec. (c,d) scan rate of 100 μm/sec. (e,f) scan rate of 200 μm/sec.

7.4 Analysis and Discussion

7.4.1 Robustness of the new estimation

The transfer function $G_{\hat{h}h}$ from h to $\hat{h}$ is given by

$$G_{\hat{h}h} = -\frac{K_2 F}{1 + K_1 F G_p} \tag{7.8}$$

from (7.2) if we assume all transfer functions are linear.

The main source of the uncertainty in the dynamic model used in the estimation scheme is the cantilever model $\mathcal{F}_l^*$. The robustness of the new estimation signal $\hat{h}$ to the model $\mathcal{F}_l^*$ can be measured by the sensitivity of $G_{\hat{h}h}$ to the changes in the model $\mathcal{F}_l^*$. It is given by

$$\frac{dG_{\hat{h}h}}{G_{\hat{h}h}} = \frac{1}{1 + K_1 F G_p} \frac{dF}{F} = S \frac{dF}{F} \tag{7.9}$$

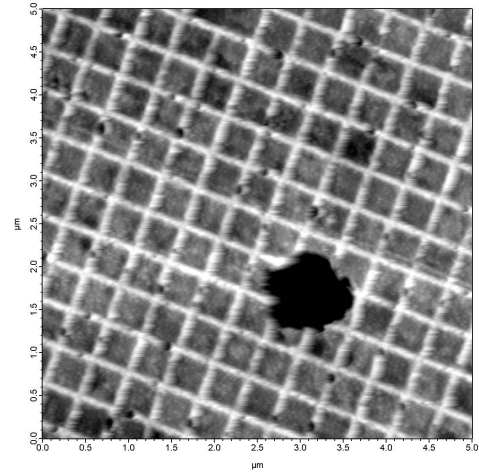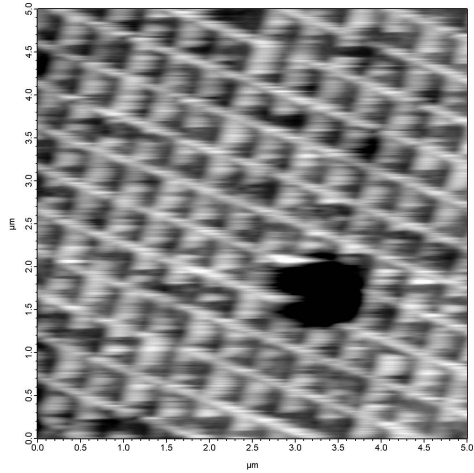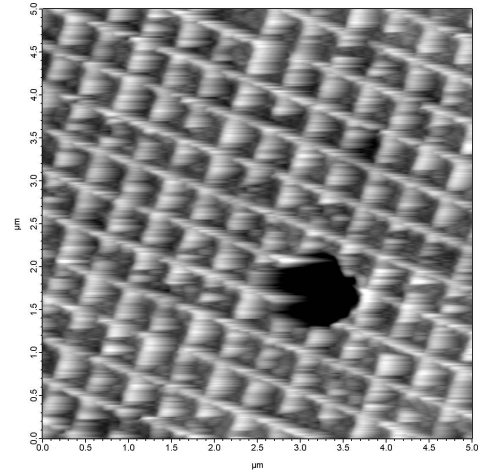and the robustness of the conventional estimation $\hat{u}$ to the model uncertainty $\frac{dG_{\hat{u}h}}{G_{\hat{u}h}}$ is the same. Thus, the conventional estimation method and the new estimation method has the same robustness to the cantilever model uncertainty.

7.4.2 Comparison with previous studies

Some estimation methods that do not directly use the controller u or the measured piezo displacement v have been proposed earlier. The *contact error mode* can be considered as intermediate between the mode of constant force mode and constant height mode [105]. This mode recognizes the fact that the control regulation bandwidth is limited and the error of regulation e has more information in Figure 7.1. Based on the assumption that the error e contains the sample topography information beyond control bandwidth, the sample topography estimation is obtained by summation of the scaled control u and error e signals, that is, estimate signal $\hat{h}$ is obtained as $\hat{h} = \alpha u + \beta e$, where α and β are constants. The main disadvantages of these mode are: 1) the prescription of α and β is ad-hoc and typically these estimates do not result in significant improvement of the imaging bandwidth. The primary

reason is that these scaling factors need to be dynamic as reflected in our design.; and 2) unlike our design, this design does not consider robustness to the noise and the disturbance signals. The effects of the noise is clearly seen from the equation for the estimate signal given by there are no consideration on the noise and disturbance signal. The error signal e is contaminated by the effect of the noise and disturbance, especially in dynamic mode operation in atmosphere. The estimate used in this mode can be given in

$$\hat{h} = (\alpha K_1 + \beta)S(r - \tilde{n}) - (\alpha + \beta K_1)S\mathcal{F}_l^*(h). \tag{7.10}$$

The proposed design can be viewed as a generalization of the method described in [84] for contact mode AFM. Contact mode AFM lends to simpler linear analysis for the following two reasons: since the tip-sample interaction is always regulated in the repulsive region, the tip-sample interaction force $F_{ts}(p-h-v)$ is well modeled by the linear relationship $F_{ts}(p-h-v) = k_o(p-h-v)$ for some constant k_o (see Figure 5.2); and also, the cantilever transfer function is assumed a constant G_c, which is a good approximation since micro-cantilevers have much larger bandwidths than the vertical positioning systems. In the research, the controller K_1 is designed by stacked sensitivity synthesis and the estimator K_2 design utilized the state estimator aspect of $\mathcal{H}_\infty$ solution of the controller K_1. However, this model based research has the same issue of model uncertainty in dynamic-mode.

7.4.3 Modeling uncertainty

The cantilever dynamics model $\mathcal{F}^*$ in this study is not accurate enough to make huge improvement in estimation bandwidth. The controller K_1 is chosen heuristically not by model based methods also for the same reason.

The static model F is the most extreme case and does not reflect the real cantilever dynamics $\mathcal{F}(g,v,h,\eta)$. In dynamic mode operation, $\mathcal{F}(g,v,h,\eta)$ also include the lock-in amplifier or the rms-to-dc converter that is nonlinear and the time lag component.

As shown in Figure 7.5 and 7.9, the analysis results from model and the experimentally obtained results have significant difference. If we consider one of the uncertainty in the model,

the low pass filter used in the lock-in amplifier of the given AFM, the transfer function will change as the following Figure 7.11.
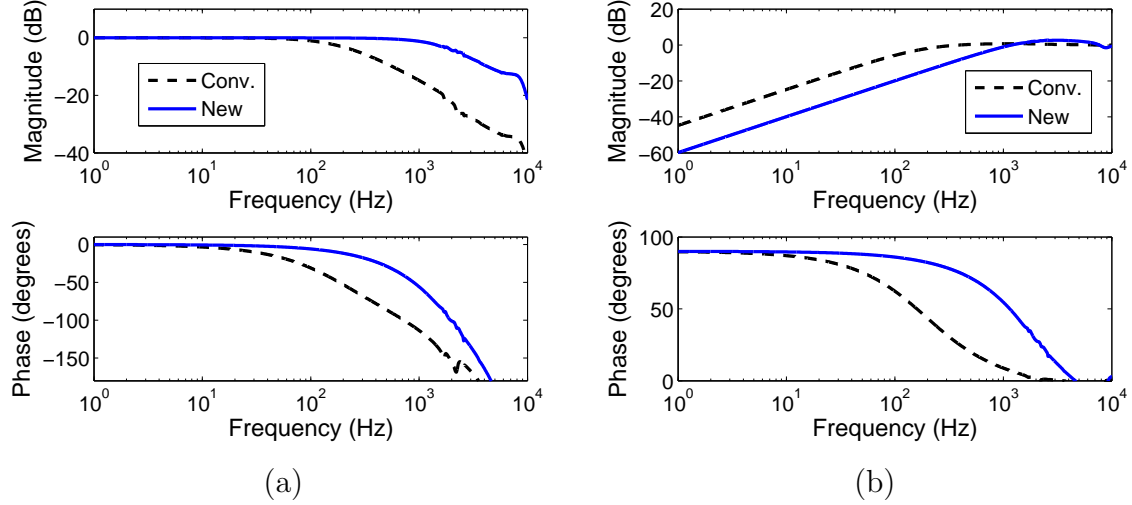
Figure 7.11: Transfer function of conventional(dashed) and new(solid) estimation with delay: (a) transfer function from the sample height h to the conventional height estimate $\hat{u}$ (i.e. $-SK_1\mathcal{F}_l^* LPF(s)\times$ (Piezo Sensitivity)))(dashed) and to the new height estimate $\hat{h}$ (i.e. $K_2\mathcal{F}_l^* LPF(s)S$)(solid) (b) transfer function from the sample height h to the conventional estimation error $h-\hat{u}$(dashed) and to the new height estimate $\tilde{h}$(solid).

In MFP-3D, the low pass filter used in the lock in amplifier, is elliptic type that has large delay. In the sense that the low pass filter is considered as the separate cascade block to F, this model is not accurate as well. (In real, it is integrated in lock-in amplifier, see Figure 5.5(b).) However, this low pass filter with cutoff frequency of 10 kHz decreases the bandwidth of imaging to 158 Hz in the conventional method and 761 Hz in the new method. The slowness and delay in lock-in amplifier or rms-to-dc converter for calculating the amplitude has been the issue in dynamic mode. The research in [71] addressed this problem and used peak picking circuit for amplitude detection.

7.4.4 Best sample-topography estimate for a given feedback control

The main contribution of this work is a design scheme that separates the regulation and the sample-topography estimation. Since design of K_1 does not depend on the design of K_2 (as

is clear from the block diagram in Figure 7.2), the regulation as well as disturbance rejection objectives can be achieved through an appropriate design of K_1 without giving any regard for the estimation objective. Also note that, for any linear design scheme, our design of K_2 achieves the maximum estimation bandwidth (through our model matching formulation) for a *given* controller K_1. Thus improvements in the design of K_1 for larger disturbance rejection bandwidths (better robustness) will lead to larger (and the maximum possible bandwidths through linear designs) estimation bandwidth. For designing K_1, PID-based methods are most commonly used. Therefore, it may be expected that better bandwidths (for disturbance rejection) can be obtained by using robust control theoretic tools. Our recent efforts using $\mathcal{H}_\infty$ synthesis design methods did give about 20% improvement in the bandwidth over the PID designs (exhaustively searched over the parameter space); however these improvements are restricted since the linearization errors are large which limit the disturbance-rejection bandwidth of any linear robust control design. Still larger improvements in the rejection bandwidths may require nonlinear control designs and better identification of the cantilever model $\mathcal{F}(g, v, h, \eta)$. However, our scheme, which separates the objectives of disturbance rejection and sample-topography estimation, will still give large improvements in sample-estimation bandwidths for a given design K_1 for the disturbance rejection. In addition, since the estimator K_2 is not in the feedback loop, it can be used either in real-time or post-processing modes.

CHAPTER 8

CONCLUSIONS AND FUTURE DIRECTIONS

For past decades, significant advances have seen in the fields of nanosciences and nano-technology. For exploring the full potential of these fields, there are increasing demands for faster and higher resolution imaging, robust operation, and quantitative measures of estimation fidelity in SPM technology. Beyond classical and simple control strategies in early SPMs, new systems-based principles and viewpoints are needed to meet the demands of future SPM technology. This thesis demonstrate them by addressing two of the primary enabling tools of nanotechnology: nanoscale positioning of materials at high bandwidth with robustness and obtaining the surface topography at the nanoscale.

Part I of the thesis addresses 2 DOF control methods for nanopositioning system. High bandwidth, high resolution, and good robustness are identified to be the primary needs of nanopositioning. The commonly employed open loop schemes and non model based controllers do not meet the current requirements for bandwidth and resolution, and also lack in robustness required to tackle the nonlinearities associated with piezo-actuation and the changing flexure dynamics. In this thesis, a new paradigm of 2DOF control is presented for the systematic design and implementation of controllers for nanopositioning system. In Chapter 2, the design goals of robustness, bandwidth and resolution can be quantified in the framework of modern robust control. 2DOF control design which deals the reference and the measured output signal separately has better capability to satisfy robustness and performance objectives is posed in Chapter 3. Four types of 2DOF control design based on different scenarios are presented. The experimental results clearly show the efficacy of systems tools and ideas in addressing the challenges. The implemented control designs have given substantial improvements in bandwidth (as high as 330%) for the same resolution and robustness. Other performance objectives can be improved by appropriately designing the

104

weight function or target transfer function. In Chapter 4, the analysis and discussion on 2DOF control framework are presented. The trade-off imposed by fundamental limitations in feedback control can be relaxed through 2DOF control and thus the 2DOF control can achieve the performance that were analytically (and therefore practically) impossible for feedback design.

In Part II, the dynamics of AFM is investigated and the new methods for fast imaging is exploited. The control systems perspective in Chapter 5 characterizes the principle of imaging in various mode and study the imaging performance in dynamic mode AFM. In Chapter 6, the control systems idea is further exploited to develop analytical tools for practically accessible models to make model-based control possible and achieve better imaging performance. The asymptotic perturbation method for weakly nonlinear oscillators can explain the distinctly nonlinear features observed in dynamic mode operation. In Chapter7, a design scheme is proposed that separates the issues of sample-profile estimation and amplitude regulation in dynamic-mode operation. Maintaining a constant amplitude while scanning does not impose limitations on the reconstruction of the sample topography. Therefore, accurate sample-profile estimations can be obtained even at frequencies near and beyond the closed-loop control bandwidths. The imaging bandwidths by new estimate signal are improved considerably over the conventional method.

These results clearly show the efficacy of control systems approaches in addressing the challenges posed by SPM technology and more widely nanotechnology. Future and ongoing work include a nonlinear model based control for regulation of interaction forces and estimation sample topography, and estimation of sample properties.

APPENDIX A

MATHEMATICAL BACKGROUND

A.1 Linear Matrix Inequalities Conditions

Lemma 1 (Kalman-Yakubovich-Popov lemma) *For* $\Phi(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$, *the matrix A is Hurwitz and $\|\Phi\|_\infty < \gamma$ if and only if there exist a matrix $P > 0$ such that*

$$
\begin{bmatrix} A^T P + PA & PB & C^T \\ B^T P & -\gamma I & D^T \\ C & D & -\gamma I \end{bmatrix} < 0. \tag{A.1}
$$

(See [38] for the proof.)

Lemma 2 *For* $\Phi(s) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$, *the matrix A is Hurwitz and $\|\Phi\|_2^2 < \gamma$ if and only if there exist a matrix $P > 0$ auxiliary parameter Q such that*

$$
\begin{bmatrix} A^T P + PA & PB_j \\ B^T P & -I \end{bmatrix} < 0, \tag{A.2}
$$

$$
\begin{bmatrix} P & C^T \\ C & Q \end{bmatrix} > 0, \tag{A.3}
$$

$$
Tr(Q) < \gamma \tag{A.4}
$$

(See [38] for the proof.)

A.2 KBM Derivation of the Vertical Positioning on Tip

Non-dimensionalized version of the photodiode measurement p^* (6.25) is given by

$$\frac{d^2 p^*}{d\tau^2} + p^* = \frac{1}{m\omega_0^2} F_{ts}(p^* - h - v) - 2\zeta\frac{dp^*}{d\tau} - (\frac{d^2 v}{d\tau^2} + 2\zeta\frac{dv}{d\tau}) + \eta + g_0\cos\Omega\tau \qquad \text{(A.5)}$$

Assume the vertical piezo transfer function G_p is a low pass filter of $G_p = \frac{1}{\frac{1}{\omega_p}s+1}$, then $\dot{v}$, and $\ddot{v}$ can be explained as

$$\dot{v} = -\omega_p v + \omega_p q \qquad \text{(A.6)}$$

$$\ddot{v} = -\omega_p^2 v - \omega_p^2 q + \omega_p \dot{q} \qquad \text{(A.7)}$$

where the signal q is the input to G_p. Again, non-dimensionalization results in

$$\frac{dv}{d\tau} = -\frac{\omega_p}{\omega_0}v + \frac{\omega_p}{\omega_0}q \qquad \text{(A.8)}$$

$$\frac{d^2 v}{d\tau^2} = -\frac{\omega_p^2}{\omega_0^2}v - \frac{\omega_p^2}{\omega_0^2}q + \frac{\omega_p}{\omega_0^2}\dot{q}. \qquad \text{(A.9)}$$

The cantilever natural frequency ω_0 (typically, order of 100 kHz) is much higher than the cutoff frequency of vertical piezo function ω_p (typically, order of $1-2$ kHz). It can be written as $\frac{\omega_p}{\omega_0} = \epsilon$ and (A.5) becomes

$$\frac{d^2 p^*}{d\tau^2} + p^* = \frac{1}{m\omega_0^2} F_{ts}(p^* - h - v) - 2\zeta\frac{dp^*}{d\tau} - (-\epsilon^2(v + q) + \epsilon(\frac{1}{\omega_0}\dot{q} - 2\zeta(v - q)))$$

$$+ \eta + g_0\cos\Omega\tau$$

$$= \epsilon f(p, \frac{dp}{d\tau}) + \epsilon E\cos\Omega\tau. \qquad \text{(A.10)}$$

where $f(p, \frac{dp}{d\tau}) := \frac{1}{\epsilon}[\frac{1}{\omega_0^2 m}F_{ts}(p - h + v) - 2\zeta\frac{dp}{d\tau} + \eta - (-\epsilon^2(v + q) + \epsilon(\frac{1}{\omega_0}\dot{q} - 2\zeta(v - q)))]$. The term $-\epsilon^2(v + q) + \epsilon(\frac{1}{\omega_0}\dot{q} - 2\zeta(v - q))$ is as order of ϵ and it will be neglected for the first order approximation.

REFERENCES

[1] Feynman, "There's plenty of room at the bottom: An invitation to enter a new field of physics," *Engineering and Science*, vol. 23, pp. 22–36, 1960.

[2] N. Taniguchi, "On the basic concept of 'Nanotechnology'," in *Proc. ICPE Tokyo. 2*, 1974, pp. 18–23.

[3] G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, "Tunnelling through a controllable vacuum gap," *Applied Physics Letters*, vol. 40(2), pp. 178–180, 1982.

[4] G. Binnig, C. F. Quate, and C. Gerber, "Atomic force microscope," *Physical Review Letters*, vol. 56, no. 9, pp. 930–933, March 1986.

[5] Z. Rihong, D. X., Y. Zhixing, and C. Jinbang, "Research on systems for measurements of CCD parameters," in *Proceedings of the SPIE- the International Society for Optical Engineering*, vol. 3553, September 1998, pp. 297–301.

[6] M. Jianxu and M. Ang Jr., "High-bandwidth macro/microactuation for hard disk drive," in *Proceedings of the SPIE- the International Society for Optical Engineering, Boston*, vol. 4194, November 2000, pp. 94–102.

[7] D. Shim, H. S. Lee, and L. Guo, "Mixed-objective optimization of a track-following controller using linear matrix inequalities," *Mechatronics, IEEE/ASME Transactions on*, vol. 9, no. 4, pp. 636–643, Dec. 2004.

[8] C. Du, L. Xie, J. N. Teoh, and G. Guo, "An improved mixed $\mathcal{H}_2/ \mathcal{H}_\infty$ control design for hard disk drives," *Control Systems Technology, IEEE Transactions on*, vol. 13, no. 5, pp. 832–839, Sept. 2005.

[9] D. White and O. Wood, "Novel alignment system for imprint lithography," *Journal of Vacuum Science & Technology B (microelectronics and Nanometer structures), Rancho Mirage*, vol. 18(6), pp. 3552–3556, May-June 2000.

[10] M. J. Zandvliet, C. W. Scherer, C. W. J. Hol, and M. M. J. van de Wal, "Multi-objective $\mathcal{H}_\infty$ control applied to a wafer stage model," in *Proc. CDC Decision and Control 43rd IEEE Conference on*, vol. 1, 17–17 Dec. 2004, pp. 796–802.

[11] S. Verma, W. jong Kim, and H. Shakir, "Multi-axis maglev nanopositioner for precision manufacturing and manipulation applications," *Industry Applications, IEEE Transactions on*, vol. 41, no. 5, pp. 1159–1167, Sept.–Oct. 2005.

[12] C. Lee, C. Lin, C. Hsia, and W. Liaw, "New tools for structural testing: piezoelectric impact hammers and acceleration rate sensors," *Journal of Guidance, Control and Dynamics*, vol. 21, no. 5, pp. 692–697, September-October 1998.

[13] D. R. Meldrum, W. H. Pence, S. E. Moody, D. L. Cunningham, M. Holl, P. J. Wiktor, M. Saini, M. P. Moore, L.-S. Jang, M. Kidd, C. Fisher, and A. Cookson, "Automated, integrated modules for fluid handling, thermal cycling and purification of dna samples for high throughput sequencing and analysis," in *Proc. IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, vol. 2, 8–12 July 2001, pp. 1211–1219.

[14] G. Schitter, K. Astrom, B. DeMartini, P. Thurner, K. Turner, and P. Hansma, "Design and modeling of a high-speed AFM-scanner," *Control Systems Technology, IEEE Transactions on*, vol. 15, no. 5, pp. 906–915, Sept. 2007.

[15] K. Leang and A. Fleming, "High-speed serial-kinematic AFM scanner: Design and drive considerations," *American Control Conference, 2008*, pp. 3188–3193, June 2008.

[16] D. Croft, G. Shedd, and S. Devasia, "Creep, hysteresis and vibration compensation for piezoactuators: Atomic force microscopy application," *Journal of Dynamic Systems, Measurement and Control*, vol. 123, pp. 35–43, 2001.

[17] A. G. Hatch, R. C. Smith, T. De, and M. V. Salapaka, "Construction and experimental implementation of a model-based inverse filter to attenuate hysteresis in ferroelectric transducers," *IEEE Transactions on Control Systems Technology*, vol. 14, no. 6, pp. 1058 – 1069, 2006.

[18] R. C. Smith, A. G. Hatch, T. De, M. V. Salapaka, R. C. H. del Rosario, and J. K. Raye, "Model development for atomic force microscope stage mechanisms," *SIAM Journal on Applied Mathematics*, vol. 66, no. 6, pp. 1998–2026, 2006.

[19] P. Ge and M. Jouaneh, "Tracking control of a piezoceramic acuator," *IEEE Transactions on Control Systems Technology*, vol. 4, no. 3, pp. 209–216, May 1996.

[20] D. Croft, G. Shedd, and S. Devasia, "Creep, hysteresis and vibration compensation for piezoactuators: Atomic force microscopy application," in *Proceedings of the American Control Conference, Chicago, Illinois*, June 2000, pp. 2123–2128.

[21] A. J. Fleming and S. O. R. Moheimani, "A grounded-load charge amplifier for reducing hysteresis in piezoelectric tube scanners," *Review of Scientific Instruments*, vol. 76, no. 7, p. 073707, 2005.

[22] G. Schitter and A. Stemmer, "Identification and open-loop tracking control of a piezoelectric tube scanner for high-speed scanning-probe microscopy," *Control Systems Technology, IEEE Transactions on*, vol. 12, no. 3, pp. 449–454, May 2004.

[23] Q. Zou and S. Devasia, "Preview-based optimal inversion for output tracking: application to scanning tunneling microscopy," *IEEE Transactions on Control Systems Technology*, vol. 12, no. 3, pp. 375–386, 2004.

[24] A. Daniele, M. V. Salapaka, S. Salapaka, and M. Dahleh, "Piezoelectric scanners for atomic force microscopes: design of lateral sensors, identification and control," in *Proceedings of the American Control Conference, San Diego, California*, June 1999, pp. 253–257.

[25] G. Schitter, P. Menold, H. F. Knapp, F. Allgower, and A. Stemmer, "High performance feedback for fast scanning atomic force microscopes," *Review of Scientific Instruments*, vol. 72, no. 8, pp. 3320–3327, August 2001.

[26] S. Salapaka, A. Sebastian, J. P. Cleveland, and M. V. Salapaka, "High bandwidth nano-positioner: A robust control approach," *Review of Scientific Instruments*, vol. 73, no. 9, pp. 3232–3241, 2002.

[27] A. Sebastian and S. Salapaka, "Design methodologies for robust nano-positioning," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 6, pp. 868–876, 2005.

[28] S. Bashash and N. Jalili, "Robust adaptive control of coupled parallel piezo-flexural nanopositioning stages," *Mechatronics, IEEE/ASME Transactions on*, vol. 14, no. 1, pp. 11–20, Feb. 2009.

[29] J. Dong, S. Salapaka, and P. Ferreira, "Robust control of a parallel-kinematic nanopositioner," *Journal of Dynamic Systems, Measurement, and Control*, vol. 130, pp. 041 007(1–15), July 2008.

[30] S. S. Aphale, S. Devasia, and S. O. R. Moheimani, "High-bandwidth control of a piezoelectric nanopositioning stage in the presence of plant uncertainties," *Nanotechnology*, vol. 19, no. 12, p. 125503 (9pp), Mar. 2008.

[31] C. Lee and S. Salapaka, "Optimal model matching design for high bandwidth, high resolution positioning in AFM," in *Proceedings of 17th IFAC Conference, Seoul, Korea*, July 2008, pp. 9230–9235.

[32] B. E. Helfrich, C. Lee, D. A. Bristow, X. Xiaohui, J. Dong, A. G. Alleyne, and S. Salapaka, "Combined $\mathcal{H}_\infty$-feedback and iterative learning control design with application to nanopositioning systems," in *Proceedings of American Control Conference, Seattle, WA*, June 2008, pp. 3893– 3900.

[33] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control, Analysis and Design*, 2nd ed. John Wiley and Sons, 2005.

[34] G. F. Franklin, J. D. Powell, and M. Workman, *Digital Control of Dynamic Systems*, 3rd ed. Pearson Education, 2005.

[35] H. Bode, *Network Analysis and Feedback Amplifier Design*. New York:Van Nostrand Reinhold, 1945.

[36] J. S. Freudenberg and D. P. Looze, "Right half-plane poles and zeros and design tradeoffs in feedback systems," *IEEE Transactions on Automatic Control*, vol. 30, no. 6, pp. 555–565, 1985.

[37] C. Mohtadi, "Bode's integral theorem for discrete-time systems," *Control Theory and Applications, IEE Proceedings D*, vol. 137, pp. 57–66, 1990.

[38] G. E. Dullerud and F. Paganini, *A Course in Robust Control Theory : A Convex Approach*, ser. Texts in Applied Mathematics, Vol. 36. Berlin, Heidelberg, New York: Springer, 2000.

[39] M. Vidyasagar, *Control System Synthesis: A Factorization Approach.* M. I. T. Press, 1985.

[40] D. Youla and J. Bongiorno, J., "A feedback theory of two-degree-of-freedom optimal wiener-hopf design," *Automatic Control, IEEE Transactions on*, vol. 30, no. 7, pp. 652–665, Jul 1985.

[41] D. J. Hoyle, R. A. Hyde, and D. J. N. Limebeer, "An $\mathcal{H}_\infty$ approach to two degree of freedom design," in *Proceedings of the IEEE Conference on Dicision and Control*, December 1991, pp. 1581–1585.

[42] K. Glover and D. McFarlane, "Robust stabilization of normalized coprime factor plant descriptions with $\mathcal{H}_\infty$-bounded uncertainty," *IEEE Transactions on Automatic Control*, vol. 34, no. 8, pp. 821–830, August 1989.

[43] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback control theory.* New York: MacMillan, 1992.

[44] D. McFarlane and K. Glover, "A loop shaping design procedure using $\mathcal{H}_\infty$ synthesis," *IEEE Transactions on Automatic Control*, vol. 37, no. 6, pp. 759–769, June 1992.

[45] D. Limebeer, E. Kasenally, and J. Perkins, "On the design of robust two degree of freedom controllers," *Automatica*, vol. 29(1), pp. 157–168, 1993.

[46] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control.* Prentice Hall, Upper Saddle River, NJ 07458, 1996.

[47] C. Scherer, P. Gahinet, and M. Chilali, "Multiobjective output-feedback control via LMI optimization," *Automatic Control, IEEE Transactions on*, vol. 42, no. 7, pp. 896–911, Jul 1997.

[48] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, ser. Studies in Applied Mathematics. Philadelphia, PA: SIAM, June 1994, vol. 15.

[49] M. de Oliveira, J. Geromel, and J. Bernussou, "An LMI optimization approach to multiobjective controller design for discrete-time systems," *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, vol. 4, pp. 3611–3616 vol.4, 1999.

[50] C. Schere, "Lower bounds in multi-objective $\mathcal{H}_2/\mathcal{H}_\infty$ problems," *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, vol. 4, pp. 3605–3610 vol.4, 1999.

[51] A. Shegaonkar and S. Salapaka, "Making high resolution positioning independent of scan rates: A feedback approach," *Applied Physics Letters*, vol. 91, p. 203513, 2007.

[52] J. Dong, S. Salapaka, and P. Ferreira, "Robust mimo control of a parallel kinematics nano-positioner forhigh resolution high bandwidth tracking and repetitive tasks," in *Proceedings of Control and Decision Conference, New Orleans, LA*, December 2007, pp. 4495–4500.

[53] R. C. Barrett and C. F. Quate, "High-speed, large-scale imaging with the atomic force microscope," *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures*, vol. 9, no. 2, pp. 302–306, 1991.

[54] G. Meyer and N. M. Amer, "Novel optical approach to atomic force microscopy," *Appl. Phys. Lett.*, vol. 53, no. 24, p. 2400, 1988.

[55] D. Sarid, *Scanning Force Microscopy with applications to Electric, Magnetic and Atomic Forces.* Oxford University Press, 1991.

[56] R. García and R. Pérez, "Dynamic atomic force microscopy methods," *Surf. Sci. Rep.*, vol. 47, p. 197, 2002.

[57] S. M. Salapaka and M. V. Salapaka, "Scanning probe microscopy," *IEEE Control Systems Magazine*, vol. 28, no. 2, pp. 65–83, April 2008.

[58] F. J. Giessibl, *Noncontact Atomic Force Microscopy*, ser. NanoScience and Technology. Berlin: Springer, 2002, ch. Principle of NC-AFM, p. 11.

[59] G. Schitter, "Model-based control of atomic force microscopes," Ph. D. Thesis, Swiss Federal Institute of Technology Zurich, Zurich, Switzerland, 2004.

[60] A. Sebastian, "Nanotechnology: A systems and control approach," Ph. D. Thesis, Iowa State University, Ames, Iowa, 2004.

[61] A. Gannepalli, "Theoretical and experimental explorations in atomic force microscopy," Ph. D. Thesis, Iowa State University, Ames, Iowa, 2004.

[62] D. R. Sahoo, "Transient force atomic force microscopy, dual Q control in atomic force microscopy, constructive control of quantum-mechanical systems, and feedback control for real-time scheduling," Ph. D. Thesis, Iowa State University, Ames, Iowa, 2006.

[63] T. Ando, T. Uchihashi, N. Kodera, D. Yamamoto, A. Miyagi, M. Taniguchi, and H. Yamashita, "High-speed AFM and nano-visualization of biomolecular processes," *Pflugers Archiv European Journal of Physiology*, vol. 456, pp. 211–225(15), April 2008.

[64] B. Anczykowski, J. P. Cleveland, D. Krüger, V. Elings, and H. Fuchs, "Analysis of the interaction mechanisms in dynamic mode SFM by means of experimental data and computer simulation," *Applied Physics A: Materials Science & Processing*, vol. 66, pp. 885–889, 1998.

[65] T. R. Rodríguez and R. García, "Theory of Q control in atomic force microscopy," *Applied Physics Ltters*, vol. 82, no. 26, p. 4821, 2003.

[66] T. Sulchek, R. Hsieh, J. Adams, G. Yaralioglu, S. Minne, C. Quate, J. Cleveland, A. Atalar, and D. Adderton, "High-speed tapping mode imaging with active Q control for atomic force microscopy," *Applied Physics Letters*, vol. 76, no. 11, p. 1473, 2000.

[67] J. Tamayo, A. Humphris, and M. Miles, "Piconewton regime dynamic force microscopy in liquid," *Appl. Phys. Lett.*, vol. 77, no. 4, p. 582, 2000.

[68] T. E. Schäffer, J. P. Cleveland, F. Ohnesorge, D. A. Walters, and P. K. Hansma, "Studies of vibrating atomic force microscope cantilevers in liquid," *Journal of Applied Physics*, vol. 80, no. 7, pp. 3622–3627, 1996.

[69] M. Kitazawa, K. Shiotani, and A. Toda, "Batch fabrication of sharpened silicon nitride tips," *Japanese Journal of Applied Physics*, vol. 42, no. Part 1, No. 7B, pp. 4844–4847, 2003.

[70] T. Fukuma and S. P. Jarvis, "Development of liquid-environment frequency modulation atomic force microscope with low noise deflection sensor for cantilevers of various dimensions," *Review of Scientific Instruments*, vol. 77, no. 4, p. 043701, 2006.

[71] 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 of the United States of America*, vol. 98, no. 22, pp. 12 468–12 472, 2001.

[72] T. Itoh and T. Suga, "Development of a force sensor for atomic force microscopy using piezoelectric thin films," *Nanotechnology*, vol. 4, no. 4, pp. 218–224, 1993.

[73] M. Tortonese, R. C. Barrett, and C. F. Quate, "Atomic resolution with an atomic force microscope using piezoresistive detection," *Applied Physics Letters*, vol. 62, no. 8, pp. 834–836, 1993.

[74] D. R. Sahoo, A. Sebastian, and M. V. Salapaka, "Harnessing the transient signals in atomic force microscopy," *International Journal of Robust and Nonlinear Control*, vol. 15, pp. 805–820, 2005.

[75] T. De, P. Agarwal, D. R. Sahoo, and M. V. Salapaka, "Real-time detection of probe loss in atomic force microscopy," *Applied Physics Letters*, vol. 89, no. 13, p. 133119, 2006.

[76] T. Itoh, C. Lee, and T. Suga, "Deflection detection and feedback actuation using a self-excited piezoelectric Pb(Zr,Ti)O3 microcantilever for dynamic scanning force microscopy," *Applied Physics Letters*, vol. 69, no. 14, pp. 2036–2038, 1996.

[77] T. Sulchek, S. C. Minne, J. D. Adams, D. A. Fletcher, A. Atalar, C. F. Quate, and D. M. Adderton, "Dual integrated actuators for extended range high speed atomic force microscopy," *Applied Physics Letters*, vol. 75, no. 11, pp. 1637–1639, 1999.

[78] B. Rogers, L. Manning, T. Sulchek, and J. D. Adams, "Improving tapping mode atomic force microscopy with piezoelectric cantilevers," *Ultramicroscopy*, vol. 100, no. 3-4, pp. 267 – 276, 2004.

[79] N. Kodera, H. Yamashita, and T. Ando, "Active damping of the scanner for high-speed atomic force microscopy," *Review of Scientific Instruments*, vol. 76, no. 5, p. 053708, 2005.

[80] A. Gannepalli, A. Sebastian, J. Cleveland, and M. Salapaka, "Thermally driven non-contact atomic force microscopy," *Applied Physics Letters*, vol. 87, no. 11, p. 111901, 2005.

[81] N. Kodera, M. Sakashita, and T. Ando, "Dynamic proportional-integral-differential controller for high-speed atomic force microscopy," *Review of Scientific Instruments*, vol. 77, no. 8, p. 083704, 2006.

[82] P. Agarwal, T. De, and M. V. Salapaka, "Real time reduction of probe-loss using switching gain controller for high speed atomic force microscopy," *Review of Scientific Instruments*, vol. 80, no. 10, p. 103701, 2009.

[83] G. Schitter, A. Stemmer, and F. Allgower, "Robust 2 dof-control of a piezoelectric tube scanner for high speed atomic force microscopy," *Proceeding of the American Control Conference, Denver, CO*, pp. 3720–3725, June 2003.

[84] S. Salapaka, T. De, and A. Sebastian, "A robust control based solution to the sample-profile estimation problem in fast atomic force microscopy," *International Journal of Robust and Nonlinear Control*, vol. 15, pp. 821–837, 2005.

[85] A. Sebastian, M. V. Salapaka, D. Chen, and J. P. Cleveland, "Harmonic and power balance tools for tapping-mode atomic force microscope," *Journal of Applied Physics*, vol. 89 (11), pp. 6473–6480, 2001.

[86] S. I. Lee, S. W. Howell, A. Raman, and R. Reifenberger, "Nonlinear dynamic perspectives on dynamic force microscopy," *Ultramicroscopy*, vol. 97, no. 1-4, pp. 185 – 198, 2003, proceedings of the Fourth International Conference on Scanning Probe Microscopy, Sensors and Nanostructures.

[87] F. Giessibl, "Advances in atomic force microscopy," *Reviews of Modern Physics*, vol. 75, pp. 949–983, July 2003.

[88] T. Sulchek, G. G. Yaralioglu, C. F. Quate, and S. C. Minne, "Characterization and optimization of scan speed for tapping-mode atomic force microscopy," *Review of Scientific Instruments*, vol. 73, pp. 2928–2936, Aug. 2002.

[89] M. Gauthier, N. Sasaki, and M. Tsukada, "Dynamics of the cantilever in noncontact dynamic force microscopy: The steady-state approximation and beyond," *Phy. Rev. B.*, vol. 64, no. 8, p. 085409, Aug. 2001.

[90] K. L. Johnson, *Contact Mechanics.* Cambridge: Cambridge University Press, 1985.

[91] B. Bhushan, *Handbook of Micro/Nano Tribology*, 2nd ed. CRC Press, 1999.

[92] J. N. Israelachvili, *Intermolecular and Surface Forces.* Academic Press, 1985.

[93] T. R. Rodríguez and R. García, "Tip motion in amplitude modulation (tapping-mode) atomic-force microscopy: Comparison between continuous and point-mass models," *Applied Physics Letters*, vol. 80, no. 9, pp. 1646–1648, 2002.

[94] S. Salapaka, M. Dahleh, and I. Mezic, "On the dynamics of a harmonic oscillator undergoing impacts with a vibrating platform," *Nonlinear Dynamics*, vol. 24, pp. 333–358, 2001.

[95] S. Hu and A. Raman, "Chaos in atomic force microscopy," *Physics Review Letters*, pp. 036 107:1–3, January 2006.

[96] R. García and A. San Paulo, "Attractive and repulsive tip-sample interaction regimes in tapping-mode atomic force microscopy," *Phys. Rev. B*, vol. 60, no. 7, pp. 4961–4967, Aug 1999.

[97] L. Wang, "Analytical descriptions of the tapping-mode atomic force microscopy response," *Applied Physics Letters*, vol. 73, no. 25, pp. 3781–3783, 1998.

[98] N. Sasaki and M. Tsukada, "Theory for the effect of the tip-surface interaction potential on atomic resolution in forced vibration system of noncontact AFM," *Applied Surface Science*, vol. 140, no. 3-4, pp. 339–343, 1999.

[99] A. Sebastian and M. Salapaka, "Amplitude phase dynamics and fixed points in tapping-mode atomic force microscopy," in *Proceedings of the American Control Conference, Boston, Massachusetts*, June 30 - July 2 2004, pp. 2499–2504.

[100] N. N. Bogoliubov and Y. A. Mitropolskii, *Asymptotic methods in the theory of nonlinear oscillations.* New York: Gordan and Breach, 1961.

[101] A. Nayfeh, *Perturbation Methods.* New York, U.S.A.: Wiley, 1973.

[102] J. Awrejcewicz, *Asymptotic Approaches in Nonlinear Dynamics: New Trends and Applications.* Berlin: Springer, 1998.

[103] N. A. Burnham, O. P. Behrend, F. Oulevey, G. Gremaud, P.-J. Gallo, D. Gourdon, E. Dupas, A. J. Kulik, H. M. Pollock, and G. A. D. Briggs, "How does a tip tap?" *Nanotechnology*, vol. 8, no. 2, pp. 67–75, 1997.

[104] R. García and A. San Paulo, "Dynamics of a vibrating tip near or in intermittent contact with a surface," *Phys. Rev. B*, vol. 61, no. 20, pp. R13 381–R13 384, May 2000.

[105] M.-H. W. Sergei N. Magonov, *Surface Analysis with STM and AFM: Experimental and Theoretical Aspects of Image Analysis.* Weinheim, New York, 1996.

AUTHOR'S BIOGRAPHY

Chibum Lee was born on May 19, 1975 in Seoul, Korea. He received his Bachelor's degree in February 1998 and Master's degree in February 2000, both from the Department of Mechanical Design and Production Engineering at Seoul National University, Korea. During his master's degree, he conducted the research on dynamics and vibration with special focus on the automotive applications. After his master's degree, he joined Hyundai Mobis, Korea, where he designed and developed the suspensions and their control logics, and analyzed the performance of vehicles as a member of Computer Aided Engineering team.

In the fall of 2005, he began his doctoral research in the System Engineering Lab for Nanoscale Investigation and Control (SENSIC) at the University of Illinois at Urbana-Champaign, under the guidance of Professor Srinivasa M. Salapaka. His primary research interests are the application of control system approaches on nano/micro-scale electromechanical systems, with a particular focus on the nanopositioning systems and Scanning Probe Microscopes (SPMs). His research endeavors at University of Illinois have led to four journal articles, and seven conference papers.