

1 **Alpacas in Space - An Autobiography**
2

3 By
4

5 JOSEPH LEVINE
6 DISSERTATION
7

8 Submitted in partial satisfaction of the requirements for the degree of
9
10 DOCTOR OF PHILOSOPHY

11 in
12

13 Physics
14

15 in the
16

17 OFFICE OF GRADUATE STUDIES
18
19 of the
20

21 UNIVERSITY OF CALIFORNIA
22

23 DAVIS
24

25 Approved:
26

27 J. Anthony Tyson, Chair
28

30 S. Mani Tripathi
31

33 Brian Kolner
34

36 Committee in Charge
37

38 2025

39

Abstract

40 comment: This is my qual write up and is here only as a place holder. Don't edit this
41 yet It is believed that a large majority of the mass in the universe comes from an as yet
42 undetermined source. This claim stems originally from work performed in 1933 by Fritz
43 Zwicky who noticed a discrepancy between measured velocities of galaxies within the Coma
44 Cluster and velocities predicted by applying the virial theorem. He wrote

45 *If this [experimental result] would be confirmed we would get the surprising result
46 that dark matter is present in much greater amount than luminous matter [1].*

47 Vera Rubin furthered this work by measuring velocities of stars rotating in the M31 galaxy
48 and concluded there must be additional non-luminous mass (originally published in [2],
49 and shown in Fig.0.1). Studies of other phenomena, such as lensing (see [**lensing2**]) have
50 confirmed the existence of this dark matter, and it is currently one of the major mysteries
51 in modern physics.

52 Since the 1980's, the leading theories have consisted of an unknown species of elementary
53 particle. The search for weakly interacting massive particles (WIMPs) have dominated
54 the budgets and schedules of the dark matter search effort but have yet to provide any
55 experimental evidence.

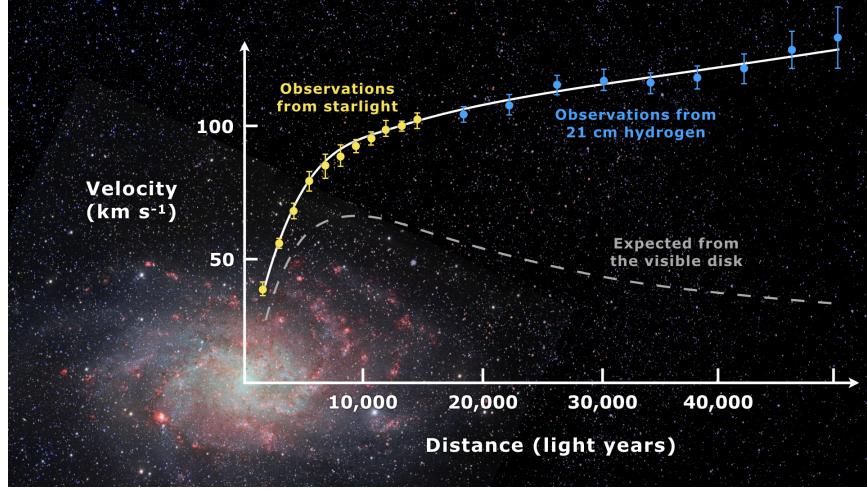


Figure 0.1: Expected vs observed velocity distributions of M33

56 In light of this, the 2017 community report on dark matter [3] highlights a need for a multi-
 57 experiment program in which many small scale experiments (< \$10M) split up to cover the
 58 vast landscape of potential dark matter candidates (see Fig. 0.2). Since very little is known
 59 about the dark matter, it is a playground for theoretical physicists to invent candidates.

60 This overwhelming search should be narrowed down.

61 The enormous mass range splits nicely into two regimes; waves and particles. At a mass of
 62 order 1eV the inter-particle spacing \approx wavelength. Lighter than this it is more convenient
 63 to think of dark matter as a wave. Alternatively, dark matter candidates heavier than this
 64 are more conveniently modeled as exhibiting particle-like behavior. The Dark E-Field Radio
 65 experiment searches for dark photons in the nano- to milli-eV mass range where dark matter
 66 is best described as a wave. This property means one would search for a dark photon using
 67 wave-like detectors, e.g. antennas.

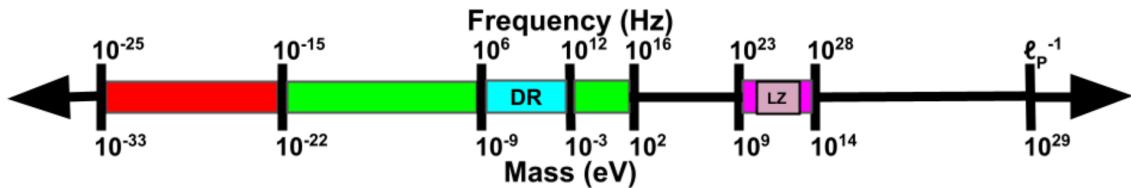


Figure 0.2: Cartoon depicting the mass scales over which dark matter may be found. Corresponding frequencies shown above. At heavy mass scales $> \mathcal{O}(1\text{ eV})$ the dark matter would behave like a particle, while on the lighter end it would behave like a wave. The Dark Radio Experiment searches at radio/microwave frequencies (blue) for a hidden photon using an antenna and spectrum analyzer. LUX-ZEPLIN Experiment (LZ) also shown.

68 comment: Discussion from paper. Good to pull from

69 This experiment extends the earlier results of our pilot experiment [4], which was designed to
 70 demonstrate feasibility of the Dark E-field Radio technique. The pilot experiment was run
 71 over the same frequency range as the experiment reported here, but did not make use of the
 72 calibration techniques to approximate statistical uniformity, nor did it fully account for the
 73 resonant enhancement of the cavity. In this paper we describe how we randomize antenna
 74 positions by moving it many times during the run. In addition, we detail EM simulations
 75 which give the average relation between the E-field at the antenna and the voltage into
 76 the LNA, accounting for resonant enhancement of the cavity. A 2^{24} -point FFT produces a
 77 spectrum dominated by background thermal noise which varies gradually with frequency.

78 We then searched over the full 50-300 MHz frequency span for any narrow-band dark photon
 79 signal of at least 5% global significance. Optimally filtering the resulting spectrum, we detect
 80 a single candidate which we are able to identify as interference, likely from our electronics.

81 Rejecting this candidate, we obtain a null result for any signal which could be attributed
82 to the dark photon in our frequency range. The resulting 95% exclusion limit for the dark
83 photon kinetic coupling ϵ is then obtained over this mass range of 0.2-1.2 μeV . Our null result
84 is a factor of ≈ 100 more sensitive than current astrophysical limits.

85 Ultimately, we can apply this detection technique at higher frequencies, ultimately going up
86 to the sub-THz band. This will require new antennas and microwave electronics. Cryogenic
87 cavities and LNAs could improve our sensitivity by an order of magnitude.

88 **Contents**

89	Contents	vi
90	List of Figures	ix
91	List of Tables	xii
92	1 The Dark Matter Story	5
93	1.1 The History of Dark Matter	6
94	1.1.1 Early ideas in dark matter	6
95	1.1.2 Into the modern era	7
96	1.1.2.1 Galactic rotation curves	11
97	1.1.3 Local dark matter density measurements	11
98	1.2 From “History” to “Current Events”	11
99	1.2.1 Modern evidence and observations of dark matter	12
100	1.2.1.1 The cosmic microwave background	12
101	1.2.1.2 Gravitational lensing	12
102	1.2.1.3 MOND and the bullet cluster	12
103	1.3 MACHOs, WIMPs and WISPs, oh my!	12
104	1.4 Dark Photon Physics	13
105	1.5 Detection Strategy: a Resource Allocation Problem	13
106	2 Experiment Overview and Design	14
107	2.1 Sources of Power in Measured Spectrum	16
108	2.1.1 Thermal Noise	17
109	2.1.1.1 Blackbody Electric Field Density	17
110	2.1.1.2 Antenna Noise	18
111	2.1.1.3 Dicke radiometer equation	22
112	2.1.2 Dark Photon Signal	27
113	2.1.3 Radio Frequency Interference	29
114	2.1.4 Amplifier Chain Noise	30
115	2.1.5 ADC effects	34
116	2.1.5.1 Spurious signals	35
117	2.1.5.2 ADC noise	35

118	2.2	Toy Analysis	36
119	2.2.1	Signal significance	37
120	2.2.1.1	Computing an exclusion limit	40
121	2.2.2	Predicted time to detection	43
122	2.3	Thermal Noise in A Cavity	44
123	2.4	Reverberation Chambers and Statistical Uniformity	44
124	2.5	System Design	44
125	2.5.1	Shielded room	47
126	2.5.2	Antenna	47
127	2.5.3	Terminator and fiber-optic switch control	52
128	2.5.4	Signal conditioning	52
129	2.5.4.1	Low noise amplifier	53
130	2.5.4.2	Secondary amplifier	53
131	2.5.4.3	Band pass filter	54
132	2.5.4.4	Putting together a signal conditioning chain	55
133	2.5.5	Veto antenna	58
134	2.5.6	12 V power system	58
135	2.5.7	GPU-Based Real-Time Spectrum Analyzer	61
136	3	System Characterization and Data Acquisition System	65
137	3.1	Measurement of Amplifier Chain Performance	66
138	3.1.1	Y-factor method	66
139	3.1.2	Confirming gain with tracking generator measurement	75
140	3.1.3	System stability over a run period	76
141	3.2	Measurement of shielding effectiveness and radio frequency interference	80
142	3.3	Spectrum Analyzer Characterization	86
143	3.3.1	Spectrum analyzer calibration	86
144	3.3.2	Spurious signal performance of the ADC	87
145	3.3.3	ADC clock performance	89
146	3.3.4	Real-time data collection efficacy	90
147	4	Data Acquisition, Data Analysis and Calibration	92
148	4.1	Data Acquisition	94
149	4.1.1	Raw data, S_o	95
150	4.2	Data Analysis	98
151	4.2.1	Fit background, $\hat{B}(\nu)$	100
152	4.2.2	Normalized spectrum, S_o^{norm}	102
153	4.2.3	Signal-matched filter	103
154	4.2.4	Monte carlo: pseudo experiments	106
155	4.2.5	Rejection of a single candidate	111
156	4.3	Calibration	112
157	4.3.1	Average effective aperture, $\langle A_e(\nu) \rangle$	113
158	4.3.2	Simulation of $\langle \tilde{Q} A_e \rangle$	114

159	4.3.3 Correction and uncertainty of $\langle \tilde{Q} A_e \rangle$	117
160	4.4 Hardware Injection Test	120
161	4.4.1 Injection test prerequisites	121
162	4.4.1.1 Determination of required injected power	121
163	4.4.1.2 Proxy dark photon signal injection	124
164	4.4.2 Performing the injection test	126
165	4.4.3 Inspection of Data	130
166	4.5 Results	134
167	4.5.1 Discussion of uncertainties	134
168	5 Beyond 300 MHz	137
169	A Overview of RTSA code base	138
170	A Tips for using the RTSA system	140
171	B Data structure and processing	141
172	B.1 Writing data	141
173	B.2 Reading and averaging data	145
174	Bibliography	148

¹⁷⁵ List of Figures

176	0.1	Expected vs observed velocity distributions of M33	iii
177	0.2	Dark matter mass regimes	iv
178	2.1	Blackbody electric field spectral density from radio to UV frequencies	19
179	2.2	An antenna and matched resistor in cavities which are in thermal equilibrium.	22
180	2.3	Simulated antenna noise voltage at room temperature in the time domain	23
181	2.4	Simulated antenna noise power spectral density at room temperature in the frequency domain	24
183	2.5	Effect of averaging on simulated antenna noise power spectral density at room temperature in the frequency domain	26
185	2.6	Schematic of LNA adding input referred noise	30
186	2.7	Schematic of cascade of amplifiers and their added noise	32
187	2.8	Cascaded noise temperature for system with two amplifiers	33
188	2.9	Input-referred power spectral density from simplified simulation	38
189	2.10	Noise spectrum containing sub-detection-threshold signal.	41
190	2.11	Noise spectrum containing sub-detection-threshold signal, various number of averages	42
192	2.12	XXXX	43
193	2.13	Schematic of the RF receiver system	45
194	2.14	Photo of dark radio lab	46
195	2.15	AB-900A biconical antenna free space effective aperture, simulated, measured and theoretical	49
197	2.16	simulated AB-900A biconical antenna free space complex input impedance, vs measured biconical data	50
199	2.18	Through gain (S_{21}) of bandpass filter	57
200	2.19	Schematic of Interlock board	59
201	2.20	Photo of lead-acid battery and slow turn on circuit	60
202	2.21	Real time DAQ data stream.	63
203	3.1	Cartoon showing Y-factor data	67
204	3.2	Hardware set up of Y-factor measurement.	69
205	3.3	Raw output power spectra used in computing gain and noise temperature for run 1.4 amp chain using the Y-factor method	70
207	3.4	Y-factor data from run 1.4 amp chain at a single frequency (625 MHz)	71

208	3.5	Frequency dependant noise temperature of run 1.4 amplifier chain, measured with Y-factor method	72
210	3.6	Frequency dependant gain of run 1.4 amplifier chain, measured with Y-factor method	73
212	3.7	Run 1.4 terminator, input referred power spectral density	74
213	3.8	Gain vs. frequency of run 1.4 amplifier chain	75
214	3.9	Gain vs. LNA voltage of run 1.4 amplifier chain	77
215	3.10	Voltage at LNA vs time.	78
216	3.11	Frequency-averaged power vs time of the terminator from run 1.4	79
217	3.12	Photo of set up to measure Shielding effectiveness of shielded room	82
218	3.13	Shielding effectiveness of shielded room, 50-300 MHz	83
219	3.14	Shielding effectiveness of shielded room, 125-3000 MHz	84
220	3.15	Spectrum from veto antenna during 300MHz data run	85
221	3.16	Scan of terminated input of Teledyne spectrum analyzer system to measure spur performance	88
223	3.17	Acquisition efficiency for GPU-based Real-time spectrum analyzer	91
224	4.1	Run 1A averaged, output-referred antenna power spectrum S_o	96
225	4.2	Run 1A averaged, output-referred, antenna and terminator power spectra	97
226	4.3	Flow chart outlining the logic of signal processing in the detection algorithm of sections 4.2.1 through 4.2.3.	99
228	4.4	Fitting background \hat{B} in the presence of a synthetic signal.	102
229	4.5	Normalized, output-referred power spectrum S_o^{norm}	104
230	4.6	S_o^{norm} after it has been passed through a matched filter.	106
231	4.7	Result of Monte Carlo pseudo experiments on signal detection.	109
232	4.8	Limit on output-referred total integrated signal power, P_o^{lim}	110
233	4.9	Screenshot of COMSOL simulation GUI for two-antenna validation	116
234	4.10	Simulated vs measured $\langle S_{21} \rangle$	117
235	4.11	Corrected average effective aperture.	119
236	4.12	Schematic of hardware injection test	122
237	4.13	Average S parameters of hardware injection test	123
238	4.14	Histogram of frequencies used for hardware injection test	126
239	4.15	Output-referred power spectral density from the hardware injection test	128
240	4.16	Standard deviation of output-referred power spectral density from the hardware injection test, computed with median absolute deviation	129
242	4.17	Output-referred power spectrum from hardware injection test, various levels of zoom	131
244	4.18	Normalized, output-referred power spectrum from hardware injection test, various levels of zoom	132
246	4.19	Matched filtered, output-referred power spectrum from hardware injection test, various levels of zoom	133
248	4.20	Zoom of 95 % exclusion limit on ϵ from Dark E-field Radio run 1A	135
249	4.21	Dark photon limits of various experiments with this work shown in red	136

²⁵¹ List of Tables

252	2.1	Caption	44
253	2.2	Specifications for the Pasternack PE15A1012-E. The voltage is regulated internally, so the exact voltage supplied is not critical	53
254	2.3	Specifications for the miniCircuits ZKL-1R5+ as measured with 9.0 V bias. There is no internal regulator, so the voltage is set using an external regulator (built around a TI LM317 [45]). comment: finish this table	54
255	2.4	Specifications for the spectrum analyzer used for run 1A.	62
259	4.1	Run 1.4 Details. Many specifications are related and can be computed from each other but are listed for reference. Efficacy differs from calculation in Fig. 3.17 mainly because of switching to terminator and brief daily pauses to move the antenna.	94
260	4.2	Threshold parameters that are part of the detection algorithm and Monte Carlo. X is the significance of the analysis. It is a parameter passed to the detection algorithm which specifies the significance threshold. The quantity $100\% - Y$ is the statistical power of the analysis. It is a parameter in the MC, which specifies a threshold on signal power where a given signal is detected in $100\% - Y$ of the MC iterations. We choose both X and Y = 5%.	108
269	A.1	Example database file.	142

270

Acknowledgments

271 Thank you to: Ben

272 Dave Hemmer

273 Nathan IT

274 Monda

275 emiljia

276 Bill Tuck

277 Buchholtz/Ayars/Nelson/Kagan/Clare/Patrova/Rick Danner/pechkis

278 parents

279 janet/joel

280 baba/papa

281 Belva

282 chatGPT

283 Mr Keye

284 Bill Nye

285 pechkis

286 Self plagiarism statement <https://journals.aps.org/copyrightFAQ.html>

287 As the author of an APS-published article, may I include my article or a portion
288 of my article in my thesis or dissertation? Yes, the author has the right to use
289 the article or a portion of the article in a thesis or dissertation without requesting
290 permission from APS, provided the bibliographic citation and the APS copyright
291 credit line are given on the appropriate pages.

292 <https://grad.ucdavis.edu/preparing-filing-your-thesis-or-dissertation>

293 Published or publishing pending material may be used with permission from the
294 copyright owner, if not you, and the Graduate Chair, if you are not the first-
295 author. See the Copyright FAQ for information about permissions when using
296 published material.

297 Material you have authored, and which has been previously published or is pend-
298 ing publishing, may be used in its published format with three exceptions: 1)
299 Margins - You must maintain 1" margins on all sides throughout the paper. 2)
300 Pagination - page numbers must follow the UC Davis standards. 3) The title
301 page must follow UC Davis standards and you must include an overall abstract.

302 comment: Need to add note about github and where code comes from

303

304 File: thesis.tex

305 Encoding: utf8

306 Sum count: 13512

307 Words in text: 10214

308 Words in headers: 221

309 Words outside text (captions, etc.): 2674

310 Number of headers: 50

311 Number of floats/tables/figures: 32

312 Number of math inlines: 382

313 Number of math displayed: 21

314 Subcounts:

315 text+headers+captions (#headers/#floats/#inlines/#displayed)

316 865+6+71 (2/2/7/0) _top_

317 147+4+0 (1/0/0/0) Chapter: The Dark Matter Story

318 0+5+0 (1/0/0/0) Section: The History of Dark Matter

319 154+5+0 (1/0/0/0) Subsection: Early ideas in dark matter

320 839+7+24 (2/0/21/7) Subsection: Into the modern era

321 20+5+0 (1/0/0/0) Subsection: Local dark matter density measurements

322 148+5+0 (1/0/0/0) Section: From ‘‘History’’ to ‘‘Current Events’’

323 61+18+0 (4/0/0/0) Subsection: Modern evidence and observations of dark matter

324 18+6+0 (1/0/0/0) Section: MACHOs, WIMPs and WISPs, oh my!

325 2+3+0 (1/0/0/1) Section: Dark Photon Physics

326 4+6+0 (1/0/0/0) Section: Detection Strategy: a Resource Allocation Problem

327 445+4+0 (2/0/0/0) Chapter: Experiment Overview and Design

328 37+6+0 (1/0/0/1) Section: Sources of Power in Measured Spectrum

329 280+8+148 (3/1/48/6) Subsection: Thermal Noise

330 49+0+0 (0/0/24/0) Subsection: Dark Photon Signal

331 6+0+0 (0/0/3/0) Subsection: Radio Frequency Interference

332 30+0+0 (0/0/16/0) Subsection: Amplifier Chain Noise \codedir{thesis/ch2/CH2.ipynb}

333 6+0+0 (0/0/6/0) Subsection: ADC effects

334 18+0+0 (0/0/4/0) Section: Toy Analysis \codedir{thesis/ch2/toyAnalysis.ipynb}

335 24+0+0 (0/0/14/0) Subsection: Signal significance

336 4+0+0 (0/0/2/0) Subsection: Shielded room

337 2+0+0 (0/0/2/0) Subsection: Antenna

338 3+0+0 (0/0/1/0) Subsection: Terminator and fiber-optic switch control

339 42+0+0 (0/0/20/0) Subsection: Signal conditioning

340 11+0+0 (0/0/7/0) Subsection: GPU-Based Real-Time Spectrum Analyzer \codedir{daqAnalysis}

341 139+6+0 (1/0/0/0) Chapter: System Characterization and Data Acquisition System

342 86+5+0 (1/0/0/0) Section: Measurement of Amplifier Chain Performance

343 449+6+445 (1/7/6/1) Subsection: Y-factor method \codedir{daqAnalysisAndExperiments/run1}

344 47+6+61 (1/1/0/0) Subsection: Confirming gain with tracking generator measurement

345 176+17+298 (1/3/0/0) Subsection: System stability over a run period \codedir{daqAnalys}

346 527+8+319 (1/4/3/0) Section: Measurement of shielding effectiveness and radio frequenc

347 33+3+0 (1/0/0/0) Section: Spectrum Analyzer Characterization
348 257+3+47 (1/0/0/0) Subsection: Spectrum analyzer calibration
349 372+6+63 (1/1/5/0) Subsection: Spurious signal performance of the ADC
350 298+3+0 (1/0/6/0) Subsection: ADC clock performance
351 64+4+80 (1/1/0/0) Subsection: Real-time data collection efficacy
352 354+6+0 (1/0/2/0) Chapter: Data Acquisition, Data Analysis and Calibration
353 171+2+39 (1/1/2/0) Section: Data Acquisition
354 289+2+114 (1/2/6/0) Subsection: Raw data, S_o {So}
355 374+2+59 (1/1/12/0) Section: Data Analysis
356 310+7+111 (1/1/13/0) Subsection: Fit background, \hat{B} {\hat{B}}
357 223+2+99 (1/1/24/1) Subsection: Normalized spectrum, $\boldsymbol{\hat{A}}$ {\hat{A}}
358 346+2+85 (1/1/8/0) Subsection: Signal-matched filter
359 430+4+185 (1/3/16/1) Subsection: Monte carlo: pseudo experiments
360 229+5+0 (1/0/2/0) Subsection: Rejection of a single candidate
361 264+1+0 (1/0/22/3) Section: Calibration
362 25+3+0 (1/0/14/0) Subsection: Average effective aperture, $\langle A_e \rangle$ {\langle A_e \rangle}
363 18+0+0 (0/0/8/0) Subsection: Simulation of \tilde{Q} , A
364 45+0+0 (0/0/25/0) Subsection: Correction and uncertainty of $\langle \tilde{Q} \rangle$ {\langle \tilde{Q} \rangle}
365 53+0+0 (0/0/21/0) Subsection: Injection test prerequisites
366 2+0+0 (0/0/1/0) Subsection: Inspection of Data
367 1+0+0 (0/0/1/0) Section: Results \codeline{COMPUTE LIMIT}
368 2+0+0 (0/0/4/0) Subsection: Discussion of uncertainties

369 6+3+0 (1/0/0/0) Chapter: Beyond 300 MHz
370 523+5+176 (1/1/2/0) Chapter: Overview of RTSA code base
371 62+6+0 (1/0/0/0) Section: Tips for using the RTSA system
372 94+4+0 (1/0/0/0) Section: Data structure and processing
373 226+2+216 (1/1/0/0) Subsection: Writing data
374 504+10+34 (1/0/4/0) Subsection: Reading and averaging data \codedir\daqAnalysisAndExpe
375
376 (errors:3)

³⁷⁷ Chapter 1

³⁷⁸ The Dark Matter Story

379

If we start making a list of things that aren't here, we could be here all night. You know, pens for instance. Let's stick with things we can see.

Wheatley

380 While there are many ways to begin a thesis on a dark matter search (of which many
381 hundreds are written every year), I have opted to prioritize narrative over completeness. I
382 will aim to answer the following questions:

- 383 • Why do we believe there to be some mysterious “dark matter” which we can’t even
384 see?
- 385 • Can we back up this prediction?
- 386 • What is the best way to balance answering this question with the economic reality of
387 finite money?

388 There will be a few digressions along the way, however the goal of this chapter is simply
389 to motivate what brought society to the point of paying young scientists to put antennas in
390 metal boxes and measure noise.

391 **1.1 The History of Dark Matter**

392 **1.1.1 Early ideas in dark matter**

393 The fundamental idea that things exist which can not be easily observed by human senses is
394 not a new one. Following a long history of discovery of previously unobserved phenomena,
395 the study of dark matter began to take shape around the end of the 19th century with the
396 discovery of dark regions among areas with a high density of stars [5]. It was Lord Kelvin
397 who began a dynamical study of the motions of stars in order to tease out the weight of

398 the luminiferous aether (who's "existence is a fact that can not be questioned" as stated
399 in his 1901 lecture which is transcribed on page 260 of [6]). Eventually the theory of the
400 aether would give way to special relativity, which ruled out a potential candidate for the
401 dark regions. This proposing of dark matter candidates and their rejection is a cycle that
402 continues to this day, and is the topic of this thesis.

403 1.1.2 Into the modern era

404 1905, known as Einstein's miracle year, ushered in a new era of "modern physics". The
405 aether that Lord Kelvin was trying to weigh was accepted as non-existent, but that didn't
406 answer the questions of galactic dynamics which were posed by those who believed in it.
407 Fritz Zwicky is credited with the first discovery of dark matter in 1933 [1], though it was
408 not widely accepted at this time. This lack of acceptance was in part due to his technique
409 of "morphological analysis" which is similar to arranging refrigerator magnets to arrive at
410 creative solutions¹. The many "creative" solutions are simultaneously to the benefit and
411 detriment of this technique. To quote Stephen Maurer[7],

412 *When researchers talk about neutron stars, dark matter, and gravitational lenses,*
413 *they all start the same way: "Zwicky noticed this problem in the 1930s. Back then,*
414 *nobody listened..."*.

415 It even seems if Zwicky himself didn't believe his own result[1],

¹This analogy is from a class I took from Andrew Wetzel at U.C. Davis. In researching it for this thesis, it is surprisingly accurate

416 *If this [experimental result] would be confirmed we would get the surprising result*
 417 *that dark matter is present in much greater amount than luminous matter [1].*

418 comment: Tony summarizes virial therm this in his '97 physics today article [8] In retrospect
 419 however, this idea is seemingly on firm theoretical footing. The viral theorem is a well known
 420 from statistical mechanics. It was formalized 1870 by Rudolf Clausius (English translation
 421 can be found here [9]). The name *virial* was coined by Clausius. Kinetic energy was, at
 422 the time, referred to as *vis viva* latin for “living force”. The plural of vis is virias, and
 423 since the theorem is concerned with many particles each with their own vis via the name
 424 *virial* theorem was chosen. The theorem is derived briefly here (following chapter 3 of
 425 Goldstein[10]) since it is informative, however the reader can skip to Eq. 1.6 for the result
 426 as it applies to galacite dynamics.

427 The “virial” the system is defined as

$$G \equiv \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{p}_i \quad (1.1)$$

428 Where \mathbf{r}_i is the position of the ith particle and \mathbf{p}_i is it's momentum, $m_i \mathbf{v}_i$.

429 By the product rule, the time derivative of G is

$$\dot{G} = \sum_{i=1}^N (\dot{\mathbf{r}}_i \cdot \mathbf{p}_i + \mathbf{r}_i \cdot \dot{\mathbf{p}}_i) . \quad (1.2)$$

430 Since the first term is really $m_i \mathbf{v}_i \cdot \mathbf{v}_i$ it can be seen as twice the kinetic energy, $2T$. The
 431 second term can be simplified by Newton's second law, $\mathbf{F} = \dot{\mathbf{p}}$. So,

$$\dot{G} = 2T + \sum_{i=1}^N (\mathbf{r}_i \cdot \mathbf{F}_i) . \quad (1.3)$$

Solving for the time-averaged, time-derivative of G ,

$$\begin{aligned}\bar{\dot{G}} &= \frac{1}{T} \int_0^\tau dt \dot{G} \\ &= \frac{G(\tau) - G(0)}{\tau}\end{aligned}$$

432 Where T is the period the average is taken over, not to be confused with the Kinetic energy.

433 In the case that the system is bound together (i.e. $\mathbf{r}_i, \mathbf{p}_i > \infty$), G is finite and as τ
434 approaches ∞ , $\bar{\dot{G}}$ approaches 0

435 Returning to Eq. 1.2, we can now say

$$\overline{T} = -\frac{1}{2} \overline{\sum_{i=1}^N (\mathbf{r}_i \cdot \mathbf{F}_i)}, \quad (1.4)$$

436 A familiar statement of the virial theorem, where the term on the right hand side is
437 known as *the virial of Clausius*.

438 While incredibly general (one can pull the ideal gas law out of this in just a few steps,
439 see again Ch. 3 of Goldstein [10]), we are concerned not with a jar filled with gas in some
440 lab, but with “gas” of stars (among which all labs are contained!).

441 To specify the equation to that of galactic dynamics, recognize for a conservative central
442 force, $\mathbf{F} = -\nabla U$ and $\mathbf{r} \cdot \mathbf{F} = rF$. If U is in a power law, i.e. of the form $k r^{n+1}$, it can quickly
443 be seen that the virial of Clausius of Eq. 1.4 can be written $\frac{n+1}{2} \overline{U}$.

444 For a problem involving an inverse square force as we are concerned with, $n = -2$, and we
445 arrive at the well known result which is usually just called “the virial theorem”, though as we
446 have seen, it reaches much deeper than a simple statement of kinetic vs potential energies
447 for galaxies:

$$\overline{T} = -\frac{1}{2}\overline{U}. \quad (1.5)$$

448 Finally following Edington's 1916 paper [11] we can form a useful formula which ul-
 449 timately gave a clue to the existence of dark matter. By setting $T = 1/2M\overline{v^2}$ and $U =$
 450 $GM^2/2R$ where M is the total mass of a cluster or galaxy, v is it's velocity, and R is it's
 451 radius, we arrive at

$$M \approx \frac{2R\overline{v^2}}{G} \quad (1.6)$$

452 As pointed out by Bertone [12], one of the earliest "clean" arguments for the existence
 453 of dark matter is known as the timing argument. It was derived by Kahn and Woltjer [13].
 454 The basic idea is given the negative red-shift (i.e. blue-shift) of the Andromeda galaxy, they
 455 are approaching (at 125 [km/s]), an indication that they are bound system. Given that the
 456 period of this system must be less than the age of the universe (assumed to be 10^{10} [yr]), a
 457 lower bound can be set on the reduced mass of the system (since orbital period T is inversely
 458 proportional to the reduced mass).

459 In the years following these arguments, WWII brought unprecedeted destruction, the
 460 reconstruction of which slowed the progress of all science that was not essential to the war
 461 effort, especially in Europe. The focus astronomy and astrophysics largely shifted to stellar
 462 structure and evolution, in large part due to the work and understanding of nuclear reactions
 463 around this time. The war also brought with it advances in radar technology. At the end of
 464 the war, the German occupation forces left large amounts of radar equipment strewn about

465 Europe. This included a large number of 7.5 [m] Würzburg antennas designed to 54[cm]
466 aircraft radar, but were also sensitive to the famous Hydrogen 21[cm] line [14]. This would
467 play a major roll in what was to come.

468 **1.1.2.1 Galactic rotation curves**

469 **1.1.3 Local dark matter density measurements**

470 comment: Add plots of local dark matter density estimate. <https://arxiv.org/pdf/1404.1938>
471 <https://arxiv.org/pdf/2201.01822> <https://arxiv.org/pdf/2305.13358>

472 **1.2 From “History” to “Current Events”**

473 As pointed out by Trimble, “Practicing scientists will normally put the cut between history
474 and current events at the time when they started reading the literature for themselves,
475 probably early in graduate school.” [15]. While that puts my personal cut off around 2020,
476 the cut off of the Dark E-field Radio group is closer to the mid-1960’s; a fruitful time for
477 research into dark matter.

478 A characteristic shift in the dark matter problem takes place in this era of current events.
479 Instead of asking “is there dark matter (and if so how much)”, we ask “what is this stuff?!”.
480 To be sure, the first question is still relevant today and the ever more precise answer was
481 outlined in Sec. 1.1.3. This section is concerned with more modern evidence for dark matter,
482 which tend to point to the modern non-baryonic particle models, which the Dark E-Field
483 Radio Experiment is searching for.

484 **1.2.1 Modern evidence and observations of dark matter**

485 **1.2.1.1 The cosmic microwave background**

486 Billy and Ben's thesis

487 **1.2.1.2 Gravitational lensing**

488 -ben thesis

489 -tony 92 article// -J. A. Tyson, G. P. Kochanski, and I. P. Dell'Antonio, Detailed mass map
490 of CL0024+1654 from strong lensing

491 **1.2.1.3 MOND and the bullet cluster**

492 -history of dark matter bertone

493 - billy and ben

494 - D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, et al., A direct
495 empirical proof of the existence of dark matter,

496 **1.3 MACHOs, WIMPs and WISPs, oh my!**

497 -Billy thesis sec 1.2

498 - G. Bertone, D. Hooper, and J. Silk, Particle dark matter: Evidence, candidates and con-
499 straints

500 **1.4 Dark Photon Physics**

501 - Ben's thesis

502

$$|\mathbf{E}_{\text{ant}}| \approx \epsilon \sqrt{\frac{2}{\varepsilon_0} \rho_{\text{DM}}}, \quad (1.7)$$

503 **1.5 Detection Strategy: a Resource Allocation**

504 **Problem**

505 - US Cosmic Visions 2017

506

507 Chapter 2

508 Experiment Overview and Design

509

A month in the laboratory can often save an hour in the library.

Frank Westheimer

510 Veljko Radeka said of detectors “One would imagine that in each particular case the best
511 solution is arrived at by 1) the detector design to maximize the significant signal, 2) reduction
512 of noise at its physical source, and 3) optimum filtering of signal and noise.” [16]. While
513 he was referring to position sensitive particle detectors, the same three principles apply to
514 this experiment. Put more directly, the goal is to maximize the signal to noise ratio. The
515 detector in this case is a low-noise, wide-band radio receiver system searching not for discrete
516 instances of particle-like interactions, but for coherent waves which are constant over long
517 periods. The signal is a small excess of narrow-band, radio frequency power received by an
518 antenna in a cavity. The noise is the white, thermal background due to the 300 K walls. The
519 system is shown schematically in Fig. 2.13.

520 This chapter begins with Sec. 2.1, devoted to exploring the sources of power in the mea-
521 sured spectrum. It walks through several back-of-the-envelope calculations to follow the
522 signal and noises as they progress from fields in free space through a simplified detector.
523 Section 2.2 takes the next step; it simulates these signals and noises and shows a frame-
524 work for statistical data analysis. These two sections serve to build up intuition about the
525 experiment which will explain design choices discussed in the later chapters. Furthermore,
526 this section is useful because the actual experiment will report a null result, i.e. a lack of
527 detection of a signal on a background. By following a signal *forward* through the system
528 and toy data analysis, it will be more clear how to infer an exclusion limit from a power
529 spectrum and working *backward* through the experiment. In the following two sections, 2.3
530 and 2.4, effects will be introduced that were not apparent from the simplistic analysis of
531 the first section. The final section, 2.5, details each piece of the system. Measurements of

532 system performance are put off until Ch. 3).

533 A note on experiment “run numbers”. Run 1A is the 50-300 MHz run. There have been
534 four attempts at collecting run 1A data. The final, published data run was collected between
535 May 10 and May 19, 2023 and is called run 1.4 or run 1p4 throughout my power points and
536 file names. I will refer to this run interchangeably as run 1A or run 1.4. Strictly speaking, it
537 should be run 1A.4, but if I change it now it will be inconsistent with file names. When you
538 forget about this note, you will have a hard time finding data in my file structure, especially
539 as future runs are completed.

540 **2.1 Sources of Power in Measured Spectrum**

541 This section will break down each term in the following equation for the input-referred power

$$P_i = \text{Thermal Noise} + \text{Dark Photon Signal} + \text{RFI} + \text{Amp Chain Noise} + \text{ADC Effects}/G, \quad (2.1)$$

542 where G is a gain factor. While convenient, this word equation is not rigorous, and relies
543 on the following subsections 2.1.1 - 2.1.5 for definition.

544 **2.1.1 Thermal Noise**

545 **2.1.1.1 Blackbody Electric Field Density**

546 This subsection estimates the noise-like¹ electric field in free space due to black body radi-
 547 ation. It assumes to be in *some* enclosure in that the ambient temperature is known to be
 548 300 K and not, for example, the 3.6 K of the sky.

549 Planck's law gives the black body spectral energy density as

$$u_\nu(\nu, T) d\nu = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/kT} - 1} d\nu \quad \left[\frac{J}{m^3} \right]. \quad (2.2)$$

550 This is frequently written in terms of spectral radiance,

$$B_\nu(\nu, T) d\nu = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} d\nu d\Omega \quad \left[\frac{W}{m^2} \right]. \quad (2.3)$$

551 Integrating this isotropic radiance over a solid angle 4π sr as well as a small frequency
 552 band $\Delta\nu$ gives the flux density $|\mathbf{S}|$ (AKA, the magnitude of the Poynting vector),

$$\begin{aligned} |\mathbf{S}| &= \int_0^{4\pi} \int_\nu^{\nu+\Delta\nu} B_\nu(\nu, T) d\nu d\Omega \\ &= \int_0^{4\pi} \int_\nu^{\nu+\Delta\nu} \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} d\nu d\Omega \quad \left[\frac{W}{m^2} \right]. \\ &\approx \frac{8\pi h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \Delta\nu \end{aligned} \quad (2.4)$$

553 The flux density can be related to the rms electric field from Poynting's theorem

$$|\mathbf{S}| = \frac{|E_{\text{rms}}|^2}{\eta} \quad \left[\frac{W}{m^2} \right], \quad (2.5)$$

¹ “Noise-like” simply means that the power contained in a signal is proportional to the bandwidth measured. Noise-like signals are more conveniently described as a power spectral density (PSD) which we will describe with the symbol S . As we will see, coherent signals which have finite width in frequency space $\Delta\nu_{\text{sig}}$ can share this property if the measurement bandwidth $\Delta\nu_{\text{RF}} \lesssim \Delta\nu_{\text{sig}}$, even though one wouldn't think of a coherent signal as being “noise-like”

554 where η is the impedance of free space. Equating Eqs. 2.4 and 2.5 and solving for the
 555 electric field gives

$$\frac{|E_{\text{rms}}|}{\sqrt{\Delta\nu}} = \sqrt{\eta \frac{8\pi h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1}} \quad \left[\frac{V}{\text{m} \cdot \sqrt{\text{Hz}}} \right], \quad (2.6)$$

556 and is plotted in Fig. 2.1.

557 Eq. 2.6 breaks up nicely into two regimes,

$$\frac{|E_{\text{rms}}|}{\sqrt{\Delta\nu}} = \begin{cases} \sqrt{\eta \frac{8\pi kT\nu^2}{c^2}} & \text{Rayleigh-Jeans regime } (h\nu \ll kT) \\ \sqrt{\eta \frac{8\pi h\nu^3}{c^2} e^{-h\nu/kT}} & \text{Wien approximation } (h\nu \gg kT) \end{cases} \quad \left[\frac{V}{\text{m} \cdot \sqrt{\text{Hz}}} \right]. \quad (2.7)$$

558 At frequencies and temperatures where the experiment is operated ($< 300 \text{ MHz}$ and
 559 300 K), $h\nu/kT \lesssim 5 \times 10^{-5}$ suggesting the Rayleigh-Jeans approximation is valid. At 300 K ,
 560 this yields electric field spectral densities of 1 and $6 \text{ nV}/(\text{m}\sqrt{\text{Hz}})$ at 50 and 300 MHz respectively.
 561

562 It is interesting to note, however, that for cryogenic experiments operating at a few GHz
 563 and in the sub K range (A common technique; see for example [17–19]), $h\nu/kT \approx 1$ and
 564 the full form of Eq. 2.6 must be used. This is shown at frequencies and temperatures to the
 565 right of the red dashed line in Fig. 2.1.

566 2.1.1.2 Antenna Noise

567 An antenna's effective aperture, $A_e [\text{m}^2]$, represents the effective area that it has to collect
 568 power density or irradiance [W/m^2] from an incident Poynting vector,

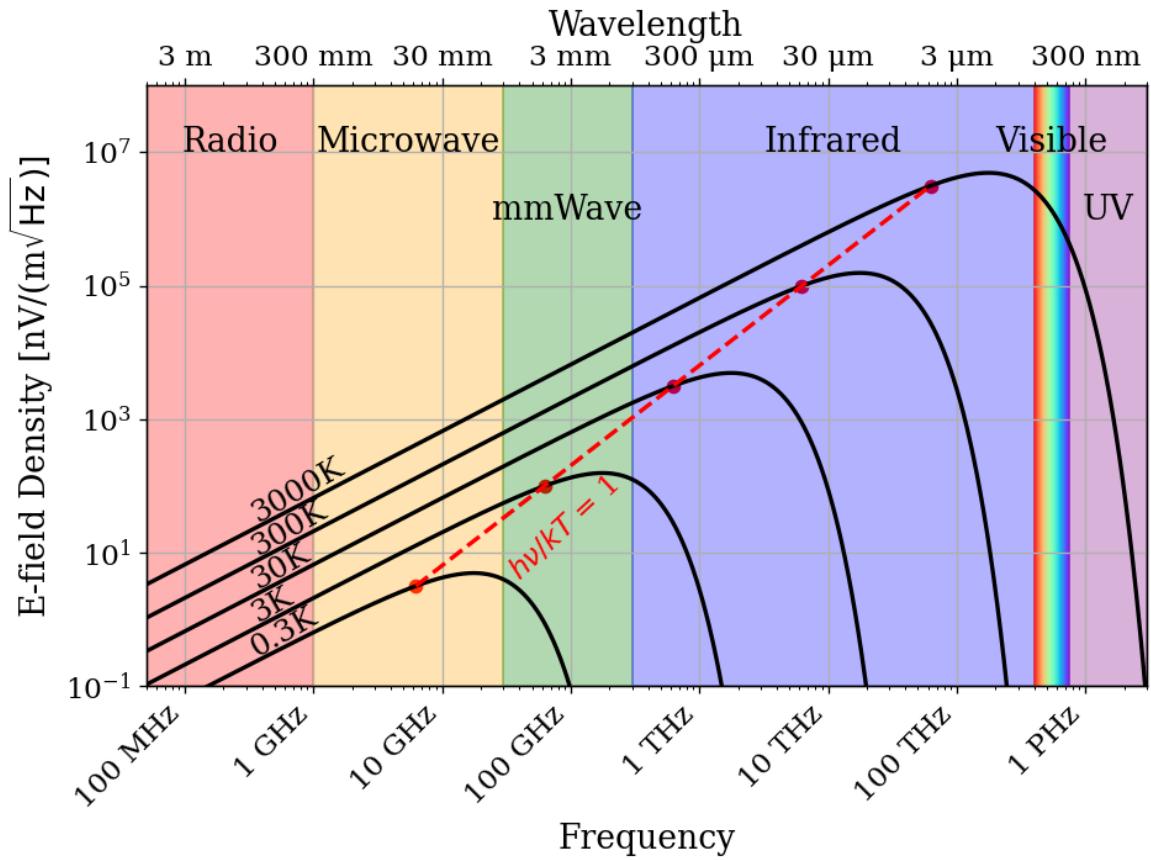


Figure 2.1: Blackbody electric field spectral density from radio to UV frequencies. A variety of temperatures are shown in black. The dashed red curve and red points indicate where along the black curves $h\nu = kT$, i.e. where Eq. 2.7 breaks up between the Rayleigh-Jeans regime to the Wien approximation. Note that these points are only a function of ν and T ; their vertical placement is chosen to lie on their respective curve but has no physical significance. The dark radio experiment is firmly in the Rayleigh-Jeans regime.

$$P_A = |\mathbf{S}| A_e, \quad (2.8)$$

569 Where $|\mathbf{S}|$ is the magnitude of the incident Poynting vector and P_r is the power received at

570 the antenna which is available at its terminals.

571 A_e is a directional quantity which varies with the antenna's directivity $D(\Omega)$, where Ω
 572 represents solid angle around the antenna. It varies with frequency ν , though it is generally
 573 discussed in terms of wavelength λ . Three matching parameters are introduced to describe
 574 how much actual power the antenna is able to deliver to a transmission line; p the polarization
 575 match of the wave to the antenna, m the impedance match of the antenna to the transmission
 576 line and η_a the efficiency of the antenna which represents how much power is absorbed
 577 compared to that lost to Joule heating of the antenna. p , m and η_a are all real, dimensionless
 578 and vary between 0 and 1.

$$A_e \equiv \frac{\lambda^2}{4\pi} D(\Omega) p m \eta_a. \quad (2.9)$$

579 This definition follows [20], though some authors do not include p in the definition [21,
 580 22].

581 A simple derivation of the direction-averaged effective aperture based on thermodynamics
 582 will provide intuition. An isotropic antenna placed in a cavity at temperature T will be
 583 illuminated by randomly polarized, isotropic radiation of the form given by the Rayleigh-
 584 Jeans limit of Eq. 2.4, $|\mathbf{S}| = 8\pi kT\Delta\nu\nu^2/c^2$. The power received by the antenna can be found
 585 by Eq. 2.8,

$$\begin{aligned} P_A &= \langle A_e \rangle \frac{1}{2} |\mathbf{S}| \\ &= \langle A_e \rangle \frac{4\pi kT\nu^2}{c^2} \Delta\nu, \end{aligned} \quad (2.10)$$

586 Where the factor of $1/2$ is introduced to account for the random polarization and the $\langle \rangle$
 587 indicate an average aperture over all angles around the antenna. If an resistor is placed in

588 a second cavity, also at temperature T, it will deliver its Johnson-Nyquest noise power [23,
 589 24]

$$P_R = kT\Delta\nu \quad (2.11)$$

590 into a matched transmission line. This is shown in Fig. 2.2. The second law of thermo-
 591 dynamics makes a very powerful statement here; the net power flow between antenna and
 592 terminator must equal 0 if the two temperatures are indeed equivalent. This means $P_R = P_A$
 593 or Eq. 2.10 = Eq. 2.12,

$$\langle A_e \rangle \frac{4\pi kT\nu^2}{c^2} \Delta\nu = kT\Delta\nu \quad (2.12)$$

594 or,

$$\begin{aligned} \langle A_e \rangle &= \frac{c^2}{4\pi\nu^2} \quad [\text{m}^2] \\ &= \frac{\lambda^2}{4\pi} \end{aligned} \quad (2.13)$$

595 This allows us to conclude that the power spectral density S_{ant} received by an antenna
 596 surrounded by an isotropic temperature is simply $kT\Delta\nu$ in the Rayleigh-Jeans limit of room
 597 temperatures and standard electronic frequencies. This result is independant of frequency,
 598 which can be understood by the reciprocal frequency dependence of the blackbody radiation
 599 (Eq. 2.4) and the average effective aperature (Eq. 2.13). The result is an antenna power
 600 which is equivalent to the well known result for a resistor at 290 K,

$$S_{\text{ant}} = 3.9 \times 10^{-22} [\text{W/Hz}] = -174 [\text{dBm/Hz}]. \quad (2.14)$$

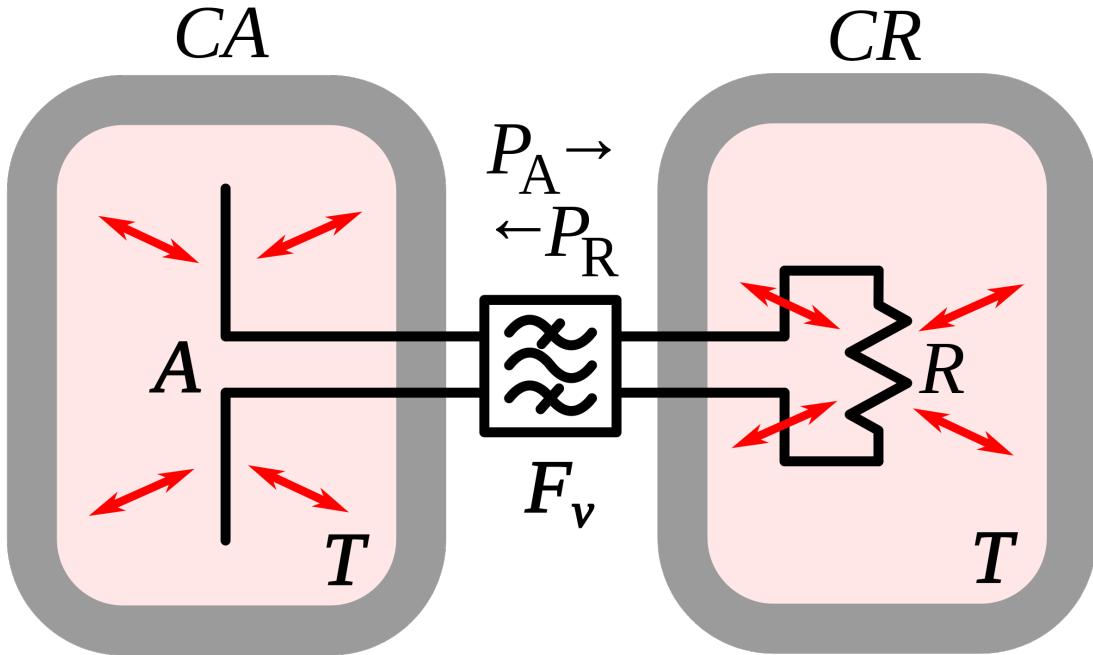


Figure 2.2: An antenna and matched resistor in cavities which are in thermal equilibrium. They are connected by a narrow filter permitting a narrow frequency band $\Delta\nu$. Image from Wikipedia, and is similar to the thought experiment proposed by Dicke [25].

601 Note that S_{ant} indicates power spectral density and should not be confused with S which

602 indicates a Poynting vector.

603 2.1.1.3 Dicke radiometer equation

604 Equation 2.14 gave the mean of a power spectrum which is inherently noisy. We will now

605 show the origin of this spectrum.

606 An enclosure who's electrically-lossy walls contain free charge carriers at finite temper-

607 ature will radiate incoherently by the fluctuation dissipation theorem. This theorem is

608 the underlying principal of phenomena such as Brownian motion [26] and Johnson-Nyquist

noise [23, 24], but was not generally proven until 1951 by Callen and Welton [27]. The random thermal fluctuation of the charge carriers will radiate a black-body spectrum. Observing the electric field in the time domain, one can imagine the radiation arriving at a detector at a wide variety of random frequencies and phases. This is incoherent noise in that at each time domain sample is independent of the one proceeding it ². The detector will produce a voltage which can be modeled as Gaussian with zero mean and standard deviation $\sigma = \sqrt{S_{\text{ant}} \Delta\nu_{\text{RF}} |Z|}$ where Z is the system impedance (here 50Ω). This is shown for a room temperature antenna (or equivalently a room temperature resistor, see Fig. 2.2) in Fig. 2.3.

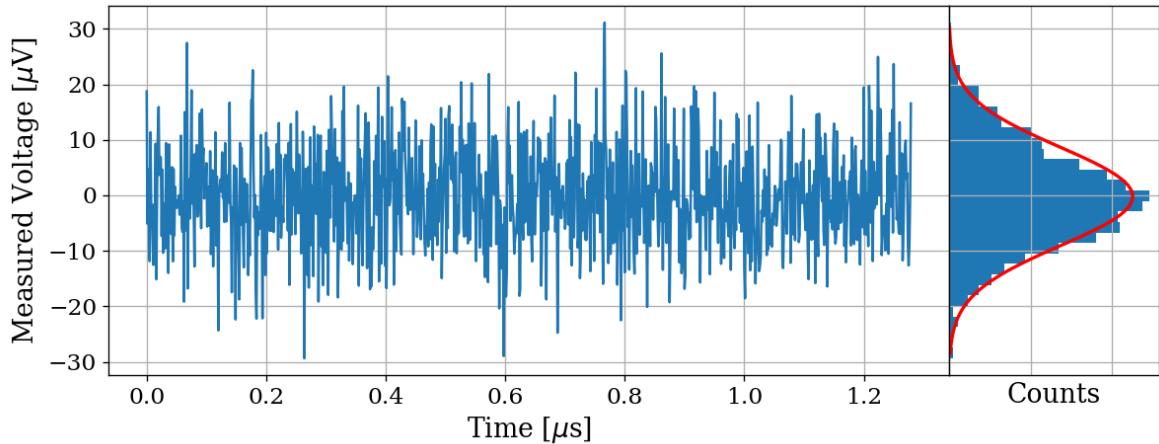


Figure 2.3: Simulated antenna noise voltage at room temperature in the time domain. Sample rate $\nu_s = 800\text{MHz}$ and number of samples $N = 2^{10} = 1024$. Bin width $\Delta\nu_{\text{RF}} = \nu_s/N \approx 800 \text{ kHz}$. Data are binned and plotted as a histogram to the right. Best fit Gaussian is shown on the histogram in red with $\mu = -0.199 \mu\text{V}$ and $\sigma = 8.86 \mu\text{V}$. Counts have been normalized such that the bins add up to unity.

The next step in converting this time domain voltage signal to a frequency domain

²The hidden assumption here is that blackbody radiation is totally incoherent. It actually has a coherence time $\tau_c \approx 2 \times 10^{-14} \text{ s}$ at 300 K [28]. For this statement to hold, the sample time $\tau_s >> \tau_c$. For run 1A, $\tau_s = 1/800 \text{ MHz} = 1.25^{-9} \text{ s}$, 5 orders of magnitude more than τ_c

618 power spectral denisty (PSD). The first step is taking a discrete Fourier transform. This is
 619 usually implemented with an algorithm known as a fast Fourier transform (FFT), so that
 620 $\tilde{V} = \text{FFT}(V)$. In order to convert to a power spectrum, a non-trivial normalization prefactor
 621 must be included;

$$S = \frac{2}{N^2 |Z|} |\Re(\tilde{V})[:N//2]|^2, \quad (2.15)$$

622 where N is the number of samples, Z is the system impedance, and [:N//2] is python notation
 623 for the second half of the samples contained in the \tilde{V} array. Performing this operation on
 624 the data in Fig. 2.3 yealds the data in Fig. 2.4

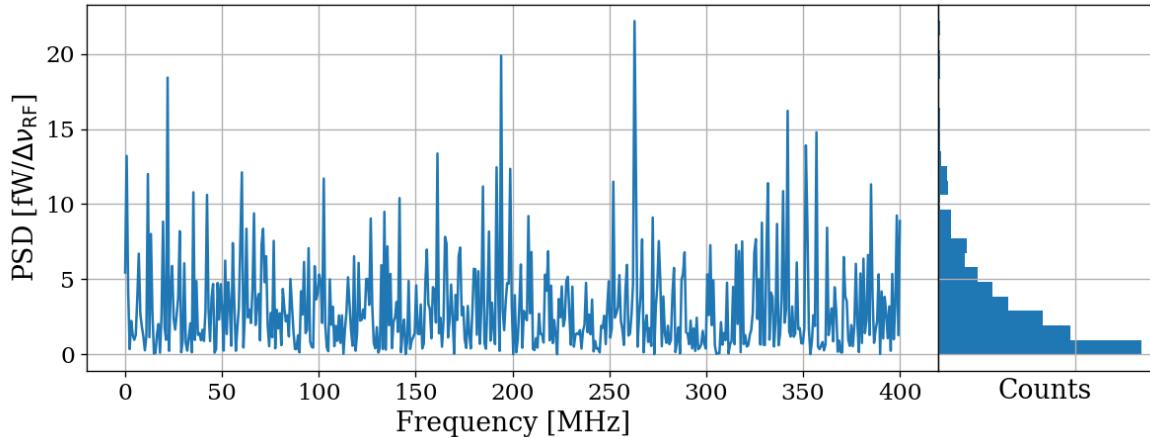


Figure 2.4: Simulated antenna noise power spectral density (PSD) at room temperature in the frequency domain. Sample rate $\nu_s = 800\text{MHz}$ and number of samples $N = 2^{10} = 1024$. Bin width $\Delta\nu_{\text{RF}} = \nu_s/N \approx 800\text{ kHz}$. Data are binned and plotted as a histogram to the right. Counts have been normalized such that the bins add up to unity.

625 The peculiar PDF of the histogram shown in Fig. 2.4 is known as a χ^2 distribution with
 626 1 degree of freedom and comes about because power is a positive-definite quantity and the

627 standard deviation of the PSD is greater than it's mean.

628 However, by averaging many of the these power spectra together the central limit theorem
629 dictates that we can expect the resulting PDF to be Gaussian. The transition from χ^2 to
630 Gaussian distributed spectra is shown in Fig. 2.5.

631 Finally, this averaged power spectrum can be modeled with the Dicke radiometer equa-
632 tion. The measured power (assuming only thermal noise) is given by

$$P_{\text{ant}} = kT\Delta\nu \left(1 \pm \frac{1}{\sqrt{\Delta\nu\tau}}\right) \quad [\text{W}]. \quad (2.16)$$

633 Here τ is the total acquisition time and so $\Delta\nu\tau$ is equivalent to the number of spectra
634 that are averaged together. This can be nondimensionalized and written

$$\frac{P_{\text{ant}}}{kT\Delta\nu_{\text{RF}}} = 1 \pm \frac{1}{\sqrt{N}} \quad [\text{W}], \quad (2.17)$$

635 which will become important during data analysis which is the topic of Sec. 4.2.

636 For the remainder of this thesis, unless otherwise stated, it will be assumed enough
637 spectra have been averaged together that a PSD is Gaussian and scales with the square root
638 of total acquisition time according to Eq. 2.16.

639 Thus far the analysis has focused only on thermal noise, however there are other sources
640 of noise and interfering signals which must be considered, not to mention the actual dark
641 photon signal.

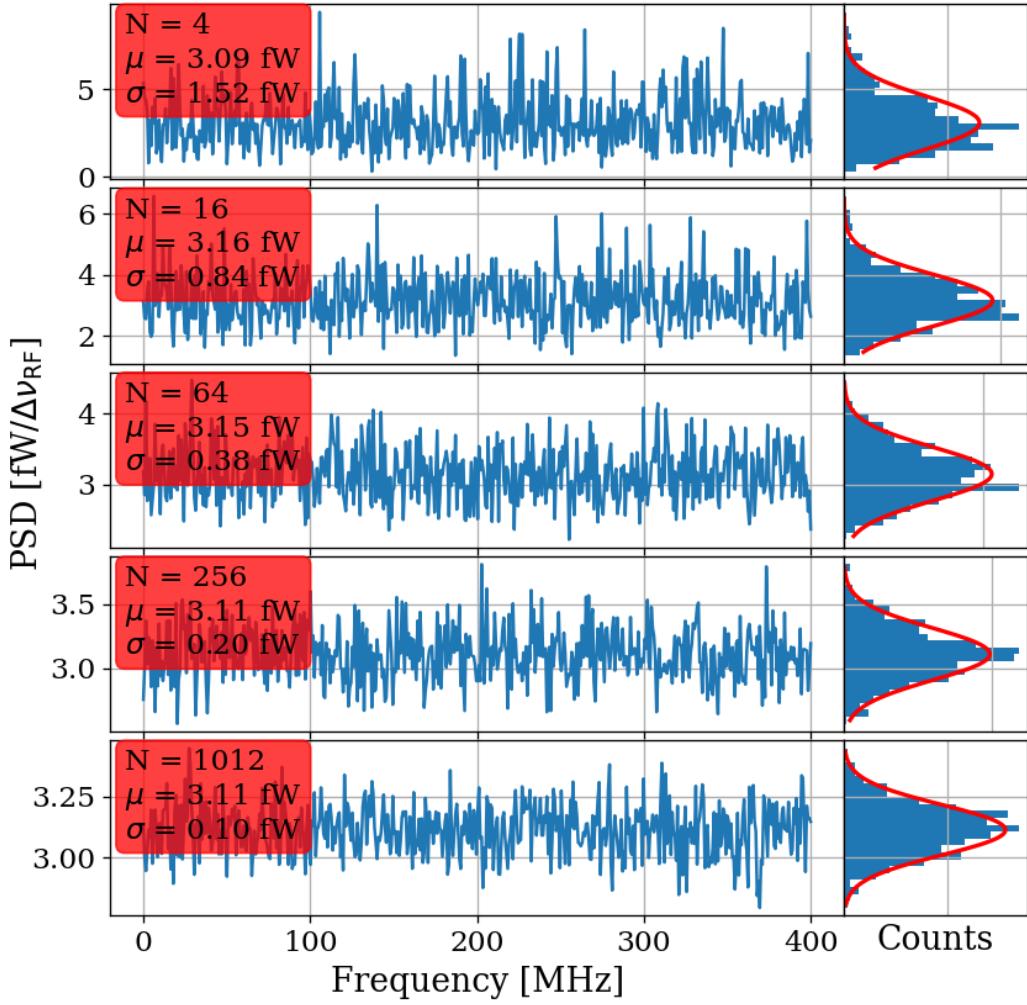


Figure 2.5: Effect of averaging on simulated antenna noise power spectral density (PSD) at room temperature in the frequency domain. Subplots show different number of averaged power spectra from $N = 4$ to 1024. Each subplot shows a factor of 4 times more averaging than the previous one. The Dicke radiometer equation predicts the standard deviation σ will scale like $N^{-1/2}$, once a large number of averages have been taken such that $\sigma \ll \mu$. Sample rate $\nu_s = 800\text{MHz}$ and number of samples $N = 2^{10} = 1024$. Bin width $\Delta\nu_{\text{RF}} = \nu_s/N \approx 800 \text{ kHz}$. Data are binned and plotted as a histogram to the right with a best fit Gaussian shown in red. Counts have been normalized such that the bins add up to unity. Note that the Y-axis scale changes between sub plots to show the improvement of the Gaussian fit. The standard deviation is indeed decreasing between plots.

642 **2.1.2 Dark Photon Signal**

643 The electric field of a kinetically mixed dark photon in free space $\mathbf{E}_{\text{ant}}^{\text{free space}}$ was shown in
 644 Eq. 1.7. In a cavity, the E-field will be enhanced by the quality factor Q of the cavity. This
 645 Q must be measured or simulated, but for this toy analysis we will assume it to be known
 646 ³. The E-field inside the cavity then is

$$\mathbf{E}_{\text{ant}} = \mathbf{E}_{\text{ant}}^{\text{free space}} \sqrt{Q}, \quad (2.18)$$

647 since Q is proportional to power, i.e. \mathbf{E}^2 . This E-field will then need to be converted from
 648 a wave in the cavity to a wave in a 50Ω transmission line by an antenna. Similar to the
 649 thermal noise of the previous section, this electric field will be converted via the effective
 650 aperture of the antenna. Similar to Q , aperture will be assumed to be known.

651 The total received power from a coherent signal inside the room then is

$$P = \frac{\mathbf{E}_{\text{ant}}^{\text{free space}^2}}{\eta} Q < A_e >, \quad (2.19)$$

652 where η is the impedance of free space.

653 The observed spread of the frequencies of the dark photon are in important effect which
 654 determine system design. In the following paragraphs of this section, several sections of
 655 Gramolin et al. [30] are summarized. Also note that the original calculation for the predicted
 656 line shape this appears to be Michael Turner in 1990 [31].

657 A simple model of the dark photon line assumes it is monochromatic, i.e. it's line shape
 658 is a delta function in frequency domain,

$$\nu_{\text{obs}} = \delta(\nu - \nu_{\text{DP}}). \quad (2.20)$$

³Typical values are in the ball park of 100. Some experiments have ultra-high Q cavities $\approx 10^{10}$ [29]

659 This is consistent with it's production add reference: misalignment mechanism/ch 1.

660 However, when observed in a frame other than its rest frame, the frequency of a (single)

661 dark photon will shift by an amount proportional to it's kinetic energy

$$\nu_{\text{obs}} = \left(1 + \frac{v_n^2}{2c^2}\right) \nu_{\text{DP}}, \quad (2.21)$$

662 where ν_{obs} is the observed frequency of the n th dark photon, v_n is its velocity, c is the speed

663 of light, and ν_{DP} is its rest frequency. The end result will be a signal that has some spread

664 in frequency, $\nu_{\text{DP}}/(\Delta\nu) \equiv Q_{\text{DP}} \approx 10^6$, with a line shape given by 2.24.

665 By summing over an infinity of dark photons of random phases and velocities (sampled

666 from the relative velocity of the dark matter halo), each with a frequency given by Eq. 2.21,

667 one can construct a power spectral density (PSD) of the dark photon signal as measured on

668 earth, S [W/Hz].

669 When performing a measurement, one records the voltage V emerging from a detector

670 for a period of time greater than the coherence of the dark photon $\tau_{\text{FFT}} \gg \tau_c$. The Fourier

671 transform of $V(t)$ is denoted $\tilde{V}(\nu)$

672 The signal will have a total power

$$P_0 = \frac{1}{\tau_{\text{FFT}}} \int_0^{\tau_{\text{FFT}}} \frac{V(t)^2}{|Z|} dt = \int_0^{1/\tau_{\text{FFT}}} S(\nu)^2 d\nu, \quad (2.22)$$

673 which is a statement of Parseval's theorem.

674 The normalized line shape is defined by dividing by P_0 ; $\lambda(\nu) \equiv S(\nu)/P_0$. This has the

675 property of being normalized to unity,

$$\int_0^\infty \lambda(\nu) d\nu = 1. \quad (2.23)$$

676 Finally, the result for this normalized line shape is

$$\lambda(\nu) = \frac{2 c^2}{\sqrt{\pi} v_0 v_{\text{lab}} \nu_{\text{DP}}} \exp \left(-\frac{\beta^2 v_0^2}{4 v_{\text{lab}}^2} - \frac{v_{\text{lab}}^2}{v_0^2} \right) \sinh \beta \quad \left[\frac{1}{\text{Hz}} \right], \quad (2.24)$$

677 where $|v_0| \approx 220$ km/s is the circular rotation speed of the Galaxy at the radius of the sun

678 (approximately 8 kpc), $v_{\text{lab}} \approx 233$ km/s is the relative velocity of the Sun to the rest frame

679 of the Galaxy and

$$\beta \equiv \frac{2 c v_{\text{lab}}}{v_0^2} \sqrt{\frac{2 (\nu - \nu_{\text{DP}})}{\nu_{\text{DP}}}}.$$

680 Equation 2.24 is used to generate Fig. 4.14. Note the quality factor $Q_{\text{DP}} \approx 10^6$ as

681 mentioned above.

682 2.1.3 Radio Frequency Interference

683 Radio Frequency Interference (RFI) includes any coherent interfering signals which can be

684 detected by the experiment. While noise is better described as a power spectral *density*

685 [W/Hz] or electric field *density* [V/(m $\sqrt{\text{Hz}}$)], RFI is made up of more narrow lines and is

686 discussed in terms of a power [W] or electric field [V/m]. In this experiment, RFI is mitigated

687 through the shielding effectiveness (SE) of the cavity. SE measurements and more details

688 about local RFI are discussed further in Sec. 2.5.1 and Sec. 3.2, and a plot of the local RFI

689 spectrum is shown in Fig. 3.15.

690 The peak RFI spike is at 186 MHz and approximately 100 $\mu\text{V}/\text{m}$, an energy density

691 of roughly 10^{-11} W/m². This will be reduced by the SE of the room (roughly 120 dB at

692 200MHz, see Fig. 3.12), but just like a coherent dark photon, it will be enhanced by the

693 Q/effective aperture. This will be right on the edge of detection, but in the actual data run
 694 it was not detected.

695 **2.1.4 Amplifier Chain Noise⁴**

696 Any amplifier will have some noise which it adds to an incoming signal which will degrade
 697 the signal to noise ratio (SNR) of the measurement ⁵. A low noise amplifier (LNA) is an
 698 amplifier which has been specifically designed to minimize the noise contribution. This
 699 process is shown schematically in Fig. 2.6.

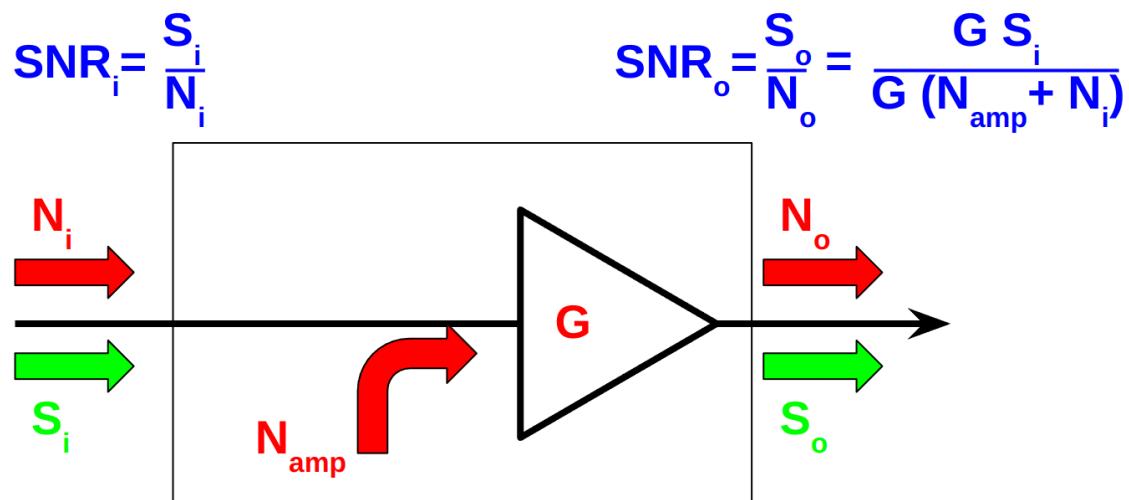


Figure 2.6: Schematic of LNA (with power gain G) adding input referred noise N_{amp} . Since N_{amp} is referred to the input, it can be directly added to the input noise which is itself input-referred. The ideal, noisless, amplifier is represented by the triangle, while the physical amplifier including noise is contained in the rectangle.

⁴Code for this section can be found at: <https://github.com/josephmlev/darkRadio/tree/master/thesis/ch2/CH2.ipynb>

⁵A great lecture on the subject by Prof. Greg Durgin can be found at [32]

700 The performance of an LNA is generally evaluated by its noise factor (F). F is defined
 701 to be the ratio the SNR at the input of an LNA to that at its output.

$$\begin{aligned}
 F &\equiv \frac{\text{SNR}_i}{\text{SNR}_o} \\
 &= \frac{S/N}{[S G]/[(N + N_{\text{amp}})G]} \\
 &= \frac{1}{1/[1 + N_{\text{amp}}/N]} \\
 &= 1 + \frac{N_{\text{amp}}}{N},
 \end{aligned} \tag{2.25}$$

702 where S and N are the signal and noise [W] presented to the LNA respectively, N_{amp} is
 703 the input-referred noise added by the LNA and G is the power gain. By factoring out the
 704 implicit $k \Delta\nu$ from $N = kT\Delta\nu$, we find

$$F = 1 + \frac{T_e}{T_0}, \tag{2.26}$$

705 where T_e is the noise temperature of a device and T_0 the temperature of the system being
 706 measured by the LNA.

707 Note that following the same derivation as Eq. 2.25, it is simple to show that the noise
 708 figure of an attenuator at temperature T with loss L is given by

$$F_{\text{att}} = 1 + \frac{(L - 1)T}{T_0}, \tag{2.27}$$

709 where T_0 is the reference temperature defined above. If $T = T_0$, Eq. 2.27 simplifies to
 710 $F_{\text{att}} = L$

711 In order to standardize device specifications for across system applications, it is common
 712 to choose a reference temperature T_0 of 290K. If not specified, it is generally safe to assume
 713 this has been done.

714 Noise factor is simply defined from noise figure,

$$\text{NF} \equiv 10\log_{10}(F). \quad (2.28)$$

715 When working with LNAs, all three measurements (T_e , F and NF) are frequently used
 716 and one must use Eqs. 2.26 and 2.28 to convert between them.

717 One important generalization is that of a cascaded series of amplifiers, shown schemati-
 718 cally in Fig. 2.7.

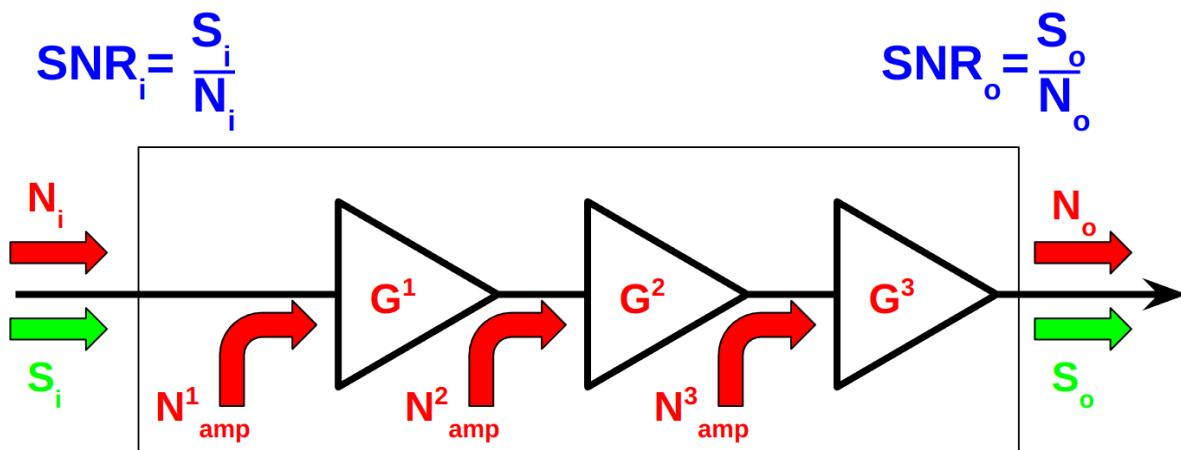


Figure 2.7: Schematic of cascade of $n = 3$ amplifiers and their added noise N_{amp}^n . Each amplifier has a gain of G^n . The SNR at the output is derived in Eq. 2.29. Note that superscripts in the figure and caption refer to index of each component and are not exponents.

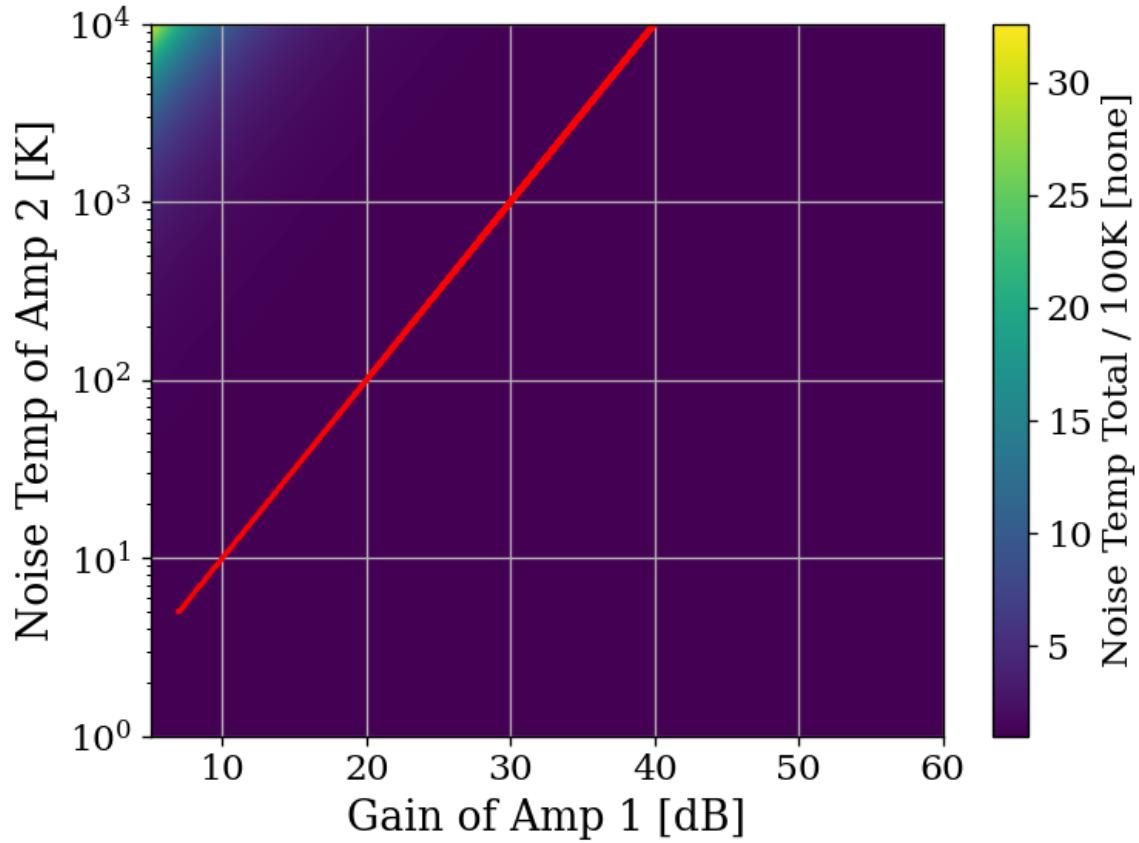


Figure 2.8: Cascaded noise temperature for system with $n = 2$ amplifiers, normalized to noise temperature of amplifier 1 (the so called LNA) = 100 K. Shown schematically (for $n = 3$) in Fig. 2.7. First amplifier's noise temperature and gain (40 dB and 100 K respectively) roughly equivalent to the LNA used in phase 1A of the experiment[33]. Red curve shows where the system's noise temperature is 1% higher than LNA noise temperature. This shows that for a 40 dB, 100 K LNA, in order to change the system noise temperature by 1 K, a second stage amplifier with a noise temperature of 10^4 would be required. Secondary amplifiers with noise temperatures closer to 500 K are realistic and inexpensive. Note that red curve should continue, but is cut off as a plotting artifact.

719 Here the total noise figure of n amplifiers can be shown to be

$$F_{\text{total}} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \cdots + \frac{F_n - 1}{G_1 G_2 \cdots G_{n-1}}, \quad (2.29)$$

720 following the same derivation as Eq. 2.25. Since the noise temperature of a system depends

721 on the noise temperature a given amplifier divided by he gain which precedes it, a front-end
722 LNA with modest gain ensures the total noise figure of the the system is equal to it's noise
723 figure to very good approximation. This is shown in Fig. 2.8. We will use this approximation
724 and assert

$$\text{Amp Chain Noise} = \text{LNA Noise}, \quad (2.30)$$

725 Where LNA here is taken to mean the first gain stage in the amplifier chain
726 Noise figures are typically frequency dependant, though they vary slowly over frequency
727 and can be approximated as constant over narrow frequency bands. Noise figures are typically
728 given on the data sheet of the LNA [33], but can also be measured. Measurement of LNA
729 noise is covered in Sec. 3.1.1 and is shown in Fig. 3.5 (which is in good agreement with the
730 LNA's data sheet [33]).

731 The power contributed by the LNA's noise is simply given by

$$P_{\text{LNA}} = kT_e\Delta\nu \quad (2.31)$$

732 This is again the mean of a fluctuating power, just as 2.16.

733 2.1.5 ADC effects

734 “ADC effects” is a catch all term which refers to power introduced by an analog-to-digital
735 converter. It contains are a three components, listed in order of importance;

$$\text{ADC Effects} = \text{Spurious Signals} + \text{ADC Noise Floor} \quad (2.32)$$

736 Equation 2.29 shows that gain G introduced before a noisy element in the RF chain, will
737 reduce the relative contribution of that noise by a factor of G . The same idea applies to
738 ADC effects, though one must be careful with the language used to describe this; spurious
739 signals are not noise, and the experiment's output *is* mostly noise.

740 **2.1.5.1 Spurious signals**

741 Spurious signals (also known as spurs) are coherent signals which are introduced into the
742 signal path at the ADC⁶. They are likely caused by candidates caused by RFI due to various
743 clocks in the PC in close proximity to the ADC. The coherence of spurs means they will pop
744 up above the noise with more averaging. Spurs don't degrade the SNR of the experiment in
745 the same way a noisy amplifier chain would; instead, they produce false positive candidates
746 which must be excluded, similar to RFI discussed in Sec. 2.1.3. Similar to ADC noise,
747 they can be measured easily by terminating the input of the ADC and scanning. They are
748 investigated in Sec. 3.3.2 and shown in Fig. 3.16, where they are shown to be nearly negligible,
749 having been mitigated by the gain of the system. There is a single spur detected after a few
750 days (see Sec. 4.2.5), but for this simple analysis we will assume spurs are negligible.

751 **2.1.5.2 ADC noise**

752 ADC noise can simply be measured by terminating the input and taking a scan. This is is
753 the same procedure as with spurs, and can be seen in Fig. 3.16. The result is that ADC

⁶Note that these spurs described here are not the same as the spurs that are described by the ADC specification *spur free dynamic range* (SFDR). SFDR is measured in dBc, i.e. *relative* to a carrier. Since our "carrier" is broadband noise, each bin produces some spurs which are -66 dBc [34] relative to itself. The aggregate of these spurs are also broadband, and averages down with the experiment's noise. The SFDR spurs are negligible for an experiment which looks at a noise-like background.

754 noise is $\approx -130\text{dBm}/47.7\text{Hz}$; a factor of 10^5 lower than the thermal noise of the experiment
 755 $\approx -81\text{dBm}/47.7\text{Hz}$, both output-referred⁷, which agrees with the ADC's data sheet [34].
 756 Since ADC noise follows the same scaling as the experiment's thermal noise (Eq. 2.16), this
 757 factor of 10^5 is independent of averaging, and ADC noise is totally negligible.

758 2.2 Toy Analysis⁸

759 With each of the terms of Eq. 2.1 defined in the previous section, we will now perform and
 760 view several simulations of a simplified dark photon signal on a simplified background. This
 761 section should provide intuition about the process of detecting a weak, narrow signal on a
 762 background PSD of thermal noise. It is assumed the noise has been averaged a sufficient
 763 number of times such that it's PDF is Gaussian (discussed in Sec. 2.1.1.3). Furthermore,
 764 following the discussion of Sec. 2.1, Eq. 2.1 can be simplified by setting RFI and ADC
 765 Effects to zero and combining Thermal Noise (300 K) with Amp Chain Noise (100 K) into a
 766 single term which represents the total of the noise in the whole system, $S_{\text{sys}} = P_{\text{sys}}/\Delta\nu_{\text{RF}} =$
 767 $k(T_{\text{ant}} + T_{\text{LNA}})$. With these simplifications, the input-referred measured power of Eq. 2.1
 768 reduces to

$$P_i = P_{\text{DP}} + kT_{\text{sys}}\Delta\nu_{\text{RF}} \left[1 \pm \frac{1}{\sqrt{N}} \right], \quad (2.33)$$

769 where $T_{\text{sys}} \equiv 400$ K, realistic for the experiment that is being simulated. Also note that this
 770 equation assumes the dark photon's line shape is much more narrow than $\Delta\nu_{\text{RF}}$ such that the

⁷Technically it doesn't matter where they are referred since they are taken in ratio. As long as they are referred to the same point!

⁸Code for this section can be found at: <https://github.com/josephmlev/darkRadio/tree/master/thesis/ch2/toyAnalysis.ipynb>

771 measured input-referred dark photon power is independent of $\Delta\nu_{\text{RF}}$. In the simulations the
772 line shape will be modeled as a delta function as in Eq. 2.20. The signal will be introduced
773 simply by adding some power in a single bin to a Gaussian background in frequency domain.
774 Performing an FFT on a perfect (discretized) sine wave can cause it's power to be split among
775 adjacent bins depending on the ratio of the sample rate to the sine wave's frequency. This
776 can be minimized by windowing the time-domain data as is discussed in Ch. 2 Sec. 2
777 of Ben Godfrey's thesis [35]. This effect is avoided by working in the frequency domain and
778 adding power to a single bin which is the method used here. A reminder that throughout
779 this thesis, code is available at github and is linked in the footnote of each section title.

780 With a simple simulation framework in place, we can now begin to generate signals
781 containing-spectra. Figure 2.9 qualitatively shows the 400 K system noise (input-referred
782 antenna plus LNA) averaging down, leaving a small signal visible. Further subsections in
783 this section will quantify this.

784 With a basic conceptual framework in place, it is now simpler to quantify signal detection
785 and the uncertainty on how many averages are required to detect a signal, and have that
786 detection be significant in that it isn't a random fluctuation (false positive).

787 **2.2.1 Signal significance**

788 The problem of the extraction of signal from noise is fundamentally a statistical one since
789 in general both the signal and noise are random variables. A method for computing a
790 *significance threshold* (ST) must be established, such that any bin containing more power

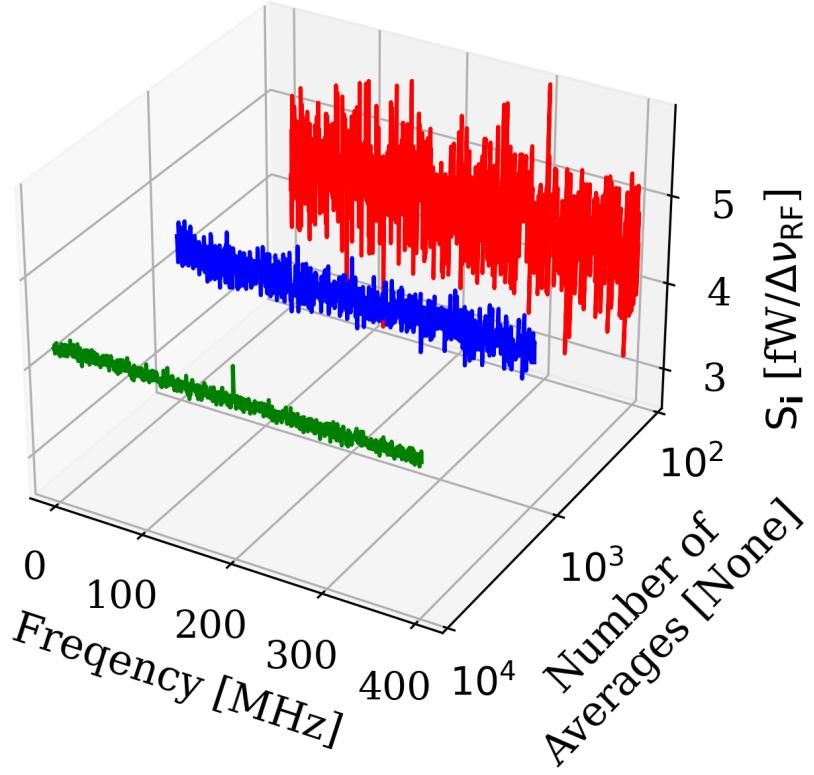


Figure 2.9: Input-referred power spectral density from simplified simulation illustrating noise averaging down to reveal a persistent, software-injected, dark photon proxy signal. The red, blue and green spectra represent 100, 1,000 and 10,000 averages respectively. The power of the signal P_{DP} is set to 0.4 fW and it only occupies a single bin. The mean of the noise is $\approx 4.3 \text{ fW}/\Delta\nu_{RF}$ where $\Delta\nu_{RF} \approx 0.78 \text{ MHz}$. The standard deviation ranges between $0.41 \text{ fW}/\Delta\nu_{RF}$ and $0.041 \text{ fW}/\Delta\nu_{RF}$ for 100 and 10,000 averages respectively. Note that this factor of 10 reduction in noise is predicted by the Dicke radiometer equation Eq. 2.16 for a factor of 100 times more averaging, as is shown in the red and green curves. Noise represents a 300 K antenna into a 100 K LNA for a total system temperature of 400 K. Signal is in a single bin at 200 MHz with a delta function line shape, defined in Eq. 2.20.

791 than this threshold is X % significant. In this way, it is possible have some known confidence

792 a given signal was not just a random fluctuation.

793 The probability that all N bins are less than z standard deviations $z\sigma$ for a standard

794 Gaussian distribution is given by

$$P(\max < z\sigma) = \left\{ \frac{1}{2} [1 + \operatorname{erf}(z/\sqrt{2})] \right\}^N, \quad (2.34)$$

795 where P is the probability, $\operatorname{erf}(z)$ is the standard error function and z is real. Setting
796 this equal to $100\% - X$ (where X is the *significance* or the desired probability a fluctua-
797 tion crosses the $z\sigma$ threshold assuming no signal), and inverting $\operatorname{erf}(z)$ yields a significance
798 threshold (ST). A convenient significance which was used in [36] is $X = 5\%$ corresponding to
799 a 5% probability that an observed fluctuation above this ST is due to chance rather than a
800 significant effect (i.e. a signal). A 5% ST for $2^{10}/2 = 512$ frequency bins⁹ works out to 3.9σ .

801 It should be noted that it is common in physics to discuss “ 5σ significance”. This means
802 that a given experiment has a $1 - \operatorname{erf}(5/\sqrt{2})$ probability (about 1 in 3×10^6) of a false positive.

803 The analysis of these normal spectra involves testing many independent frequency bins to see
804 if any one of them exceeds some threshold. It is helpful to view these bins as “independent
805 experiments”, each involving a random draw from the same parent Gaussian distribution.

806 In this context, we discuss global significance (all of the bins) in contrast to local significance
807 (a single bin). Setting a global 5% significance threshold is equivalent to setting a local
808 threshold of 3.9σ given 512 bins.

809 One concept which can assist in choosing the significance is known as the *cost* of a
810 decision. If an experiment requires a facility which charges by the hour and where the
811 schedule is set years in advance, a false positive is quite expensive since it will lead to

⁹Note that a real FFT produces half the number of frequency bins as an output compared to the time domain sample it received, hence the factor of 2 established in Fig. 2.3

812 publicity and ultimately humiliation. A follow up experiment will have to take out more
813 expensive time to verify the results, and until that happens theorists will spend their time
814 rewriting physics to explain the result of a random fluctuation. In this case, the global
815 significance should be quite low to avoid these high cost outcomes, hence the 5σ discussed
816 in the previous paragraph.

817 In the case of the dark radio experiment, a false positive is quite inexpensive. If a signal
818 is detected, just repeat the experiment. For run 1A, this is 9 days of averaging which is
819 mostly passive and is little more than an annoyance. If a signal is detected at the same
820 frequency, things become interesting. This concept of cost is discussed formally in Appendix
821 1 of Extraction of Signals From Noise by Wainstein and Zubakov [37].

822 For this reason, a significance of 5 % (i.e. 5 % chance of a false positive) is acceptable
823 for this experiment, where it is certainly not acceptable at the LHC.

824 **2.2.1.1 Computing an exclusion limit**

825 With the significance threshold (ST) defined in for a spectra containing a signal in the
826 previous subsection, we turn briefly to the concept of an exclusion limit. In the actual
827 experiment, no signals were detected. In this case, the null result must be translated into
828 an exclusion limit, as in Fig. 4.21. While this section on toy analysis looks at small signals
829 riding on noise spectra, it can be easy to loose site of the fact that no signal was observed.

830 When setting a exclusion limit on power

831 One common point of confusion can be cleared up with an example. Frequently people
832 think that the significance threshold [W] sets the limit above which one is 95 % confident

833 there

834 Really the detection limit is the threshold - the mean μ . This is usually significantly
835 below the significance threshold. This is demonstrated in Figs. 2.10 and 2.11.

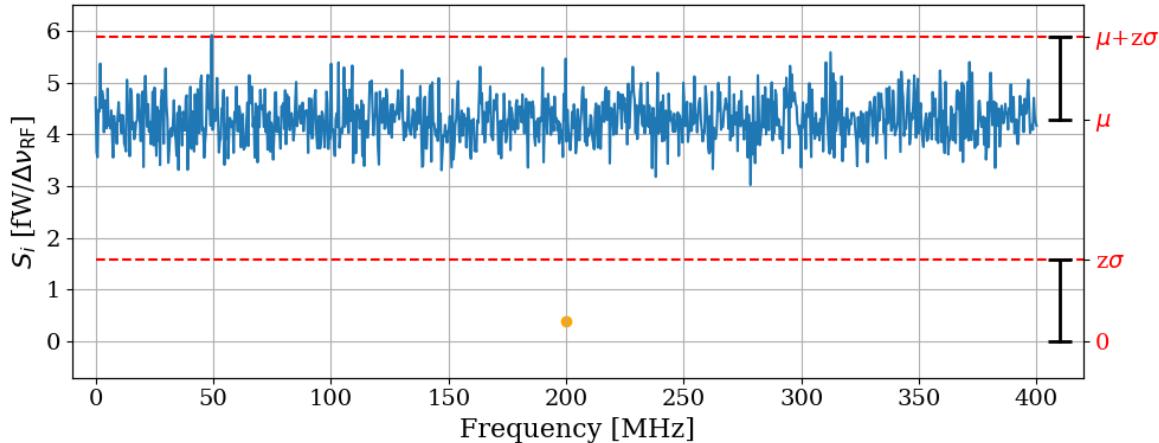


Figure 2.10: Simulated, input-referred noise spectrum containing small signal at 200MHz. The signal power = .4 fW, system temperature = 400 K and $N_{avg} = 100$; the same as is shown in the red spectrum of Fig. 2.9. The 0.4 fW signal is shown as an orange point. This power is added to the random background, so the measured power in the bin at 200 MHz is a Gaussian random variable given by Eq. 2.33. The significance threshold (upper dashed red line) is given by $\mu + z\sigma$ and was derived in section 2.2.1. The detection threshold (lower dashed red line) is the significance threshold minus the mean, $(\mu + z\sigma) - \mu = z\sigma$. This shows that the detection threshold is set by fluctuations of the measured power. While this is related to the ST, it is different. The detection threshold is below the mean, thanks to averaging. Note the detection at approximately 50 MHz; this a random fluctuation and is expected to occur in 5 out of 100 simulations of these spectra since the significance used to calculate the ST was set at 5%. **comment: make detection thresh 'lime'**

836 With the derivation of the significance threshold complete, the next subsection will focus
837 on predicting the amount of time.

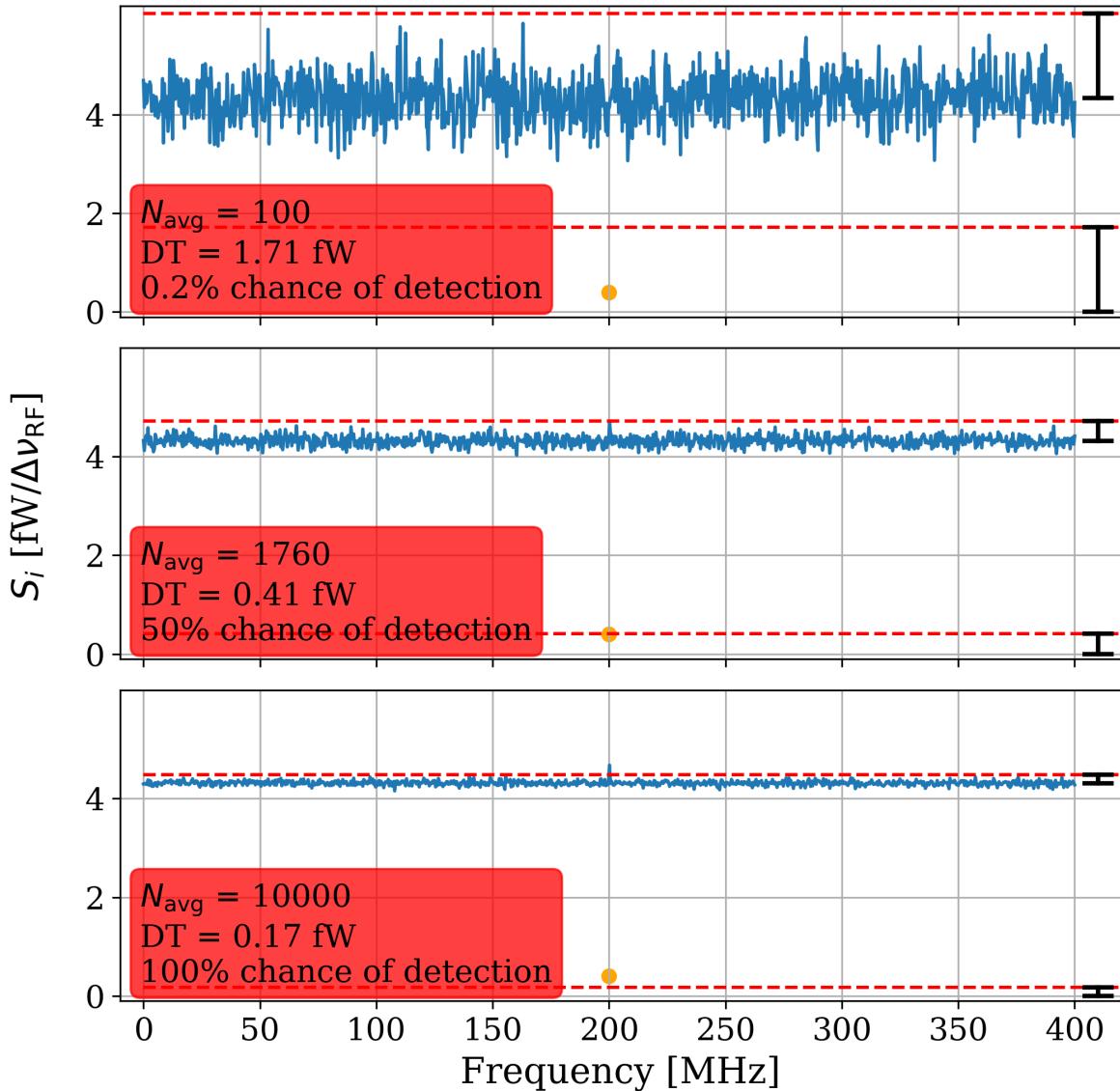


Figure 2.11: Simulated, input-referred noise spectrum containing small signal at 200MHz. The signal power = .4 fW, system temperature = 400 K and $N_{avg} = 100, 1,760$ and $10,000$; the first and last are shown as the red and green spectra respectively in Fig. 2.9. The 0.4 fW signal is shown as an orange point. The chance of detection is computed via a simple Monte Carlo where 100,000 of similar spectra containing noise and signal are generated, and the number of times the measured power at 200 MHz is greater than the ST is counted. The full set of statistics from this Monte Carlo is shown in Table 2.1. comment: make detection thresh 'lime'

838 2.2.2 Predicted time to detection

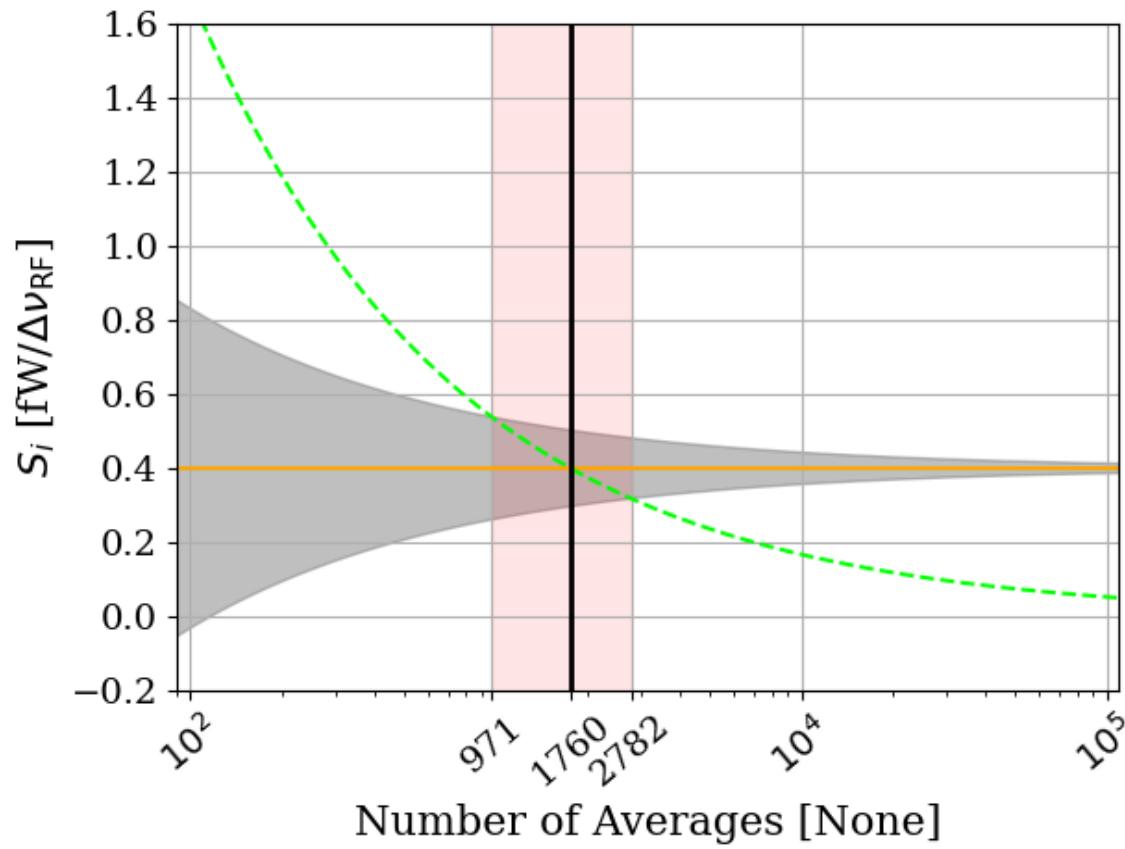


Figure 2.12: XXXX

Number of Averages	True Pos. [%]	False Neg. [%]	False Pos. [%]	True Neg. [%]
100	0.2	99.8	5.7	94.3
971	16.3	83.7	5.5	94.5
1,760	49.8	50.2	5.7	94.3
2,782	83.8	16.2	5.7	94.3
10,000	100.0	0.0	5.6	94.4

Table 2.1: Caption

839 **2.3 Thermal Noise in A Cavity**

840 **2.4 Reverberation Chambers and Statistical**

841 **Uniformity**

842 Generally it is convenient to think of electromagnetic cavities as containing a single mode

843 **2.5 System Design**

844 This section outlines the subsystems which make up the experiment. While specifics and

845 basic calculations are provided as they apply to design choices of subsystems, testing and

846 characterisation of the system as a whole is covered in Ch. 3 and calibration of the system

847 is discussed in Ch. 4. A simplified schematic of the entire experiment is shown in Fig. 2.13

848 and a photo of the lab is shown in Fig. 2.14.

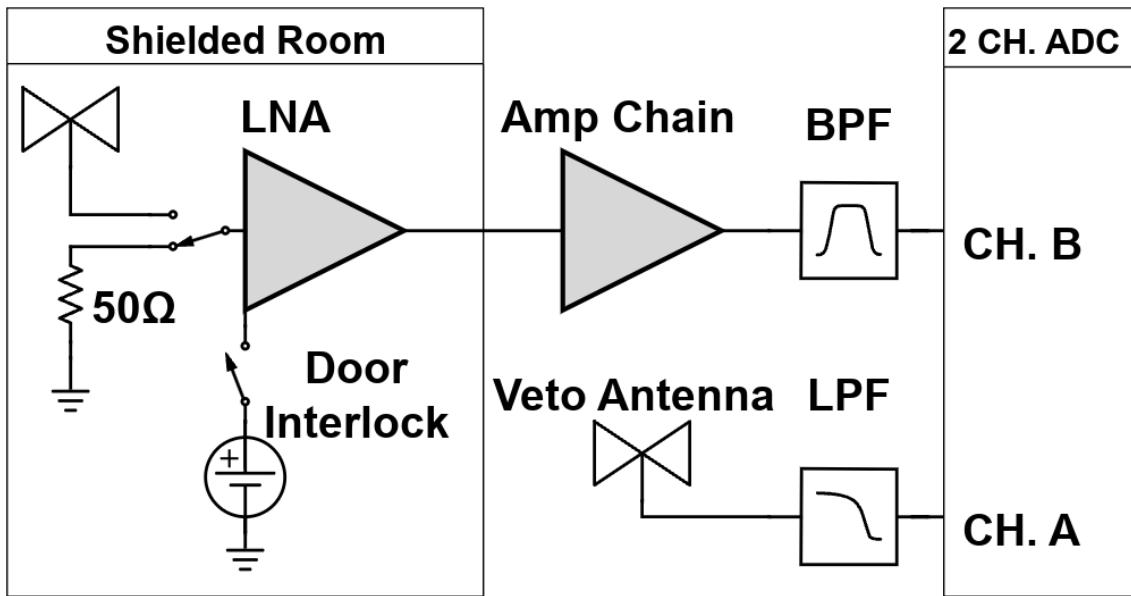


Figure 2.13: Schematic of the RF receiver system. An RF coax switch allows the PC to control the source (antenna or terminator). The switch is controlled by optical fiber to maintain isolation of the room. The LNA (Pasternack PE15A1012) has a nominal gain and noise temperature of 40dB and 100K (Figs. XXX and XXX). It is interlocked (Fig. 2.19) to the door to protect amp B and the ADC from large signals when the door is open. The secondary amplifier is a miniCircuits (MC) ZKL-1R5+ and has a nominal gain of 38dB. Not pictured after this amplifier is a fixed 4dB of attenuation. The band pass filter (BPF) is made up of a MC SLP-50+ high pass filter (HPF) and a MC ZX75LP-288-S+ low pass filter. The HPF reduces the bandwidth and therefore the total power of the signal before entering the ADC allowing for more gain before clipping. The LPF serves the same purpose while also acting as an anti-aliasing filter. These two filters define the experiment's bandwidth, $-3\text{ dB} \approx 40 - 320\text{ MHz}$. The veto antenna is outside of the room and interference is not reduced by the ≈ 100 dB SE of the room, so no gain is required. The LPF on the veto is for anti-aliasing.



Figure 2.14: Photo of dark radio lab. Shielded room contains the main antenna as well as the LNA and power supply (not visible). Veto antenna can be seen hanging out side of shielded room.

849 **2.5.1 Shielded room**

850 The shielding room [38] serves two purposes. The first is straightforward; to shield the
851 antenna, keeping radio frequency interference (RFI) *out*. The second purpose is a bit more
852 subtle; to keep any converted dark photons *in*. This second point is addressed further as an
853 aspect of system calibration in Ch. 4, but roughly can be described by the loaded quality
854 factor [39–41] of the antenna/room system. Namely, a more resonant system will be more
855 sensitive to coherent signals. This subsection will focus on the first point, keeping RFI out.

856 Shielding effectiveness SE is a measurement of a shielding enclosure’s ability to attenuate
857 electromagnetic waves from entering,

$$\text{SE} \equiv 10 \log_{10} \left(\frac{P_{\text{open}}}{P_{\text{closed}}} \right) = P_{\text{open}, dB} - P_{\text{closed}, dB} \quad (2.35)$$

858 where $P_{\text{open}}/P_{\text{closed}}$ are powers received with the door open/closed. The ratio of powers
859 allows all the specifics of antenna matching to cancel allowing for a very simple differential
860 measurement. The results of this are described in Sec. [add reference: system cheratization](#)

861 **2.5.2 Antenna**

862 The antenna plays an important roll in the experiment as the matching device between elec-
863 tromagnetic waves in the cavity and the receiver system. For a broadband search such as the
864 50-300 MHz run (a 6:1 bandwidth), a broadband antenna must be used. The chosen antenna
865 must provide a good impedance match and high efficiency since an inefficient antenna would
866 convert a substantial amount of the converted dark photon’s power into heat in the antenna’s

867 structure. For the 50-300 MHz run, a \approx 131 cm biconical antenna (bicon) was chosen. The
868 selected model is manufactured by COMPOWER, model AB-900A [42]. In a phone call with
869 the manufacturer as well as testing of the isolated balon, it was determined that the balun
870 used in the antenna was 1:1. This allows for simple simulation. In COMSOL, the lumped
871 port option allows for a balanced drive of an antenna. A match to a 50Ω transmission line
872 through a 1:1 balun is simply modeled as a lumped port, a very simple object in COMSOL
873 featured in nearly all of the antenna tutorials ¹⁰

¹⁰See for example the dipole antenna tutorial, available at <https://www.comsol.com/model/dipole-antenna-8715>

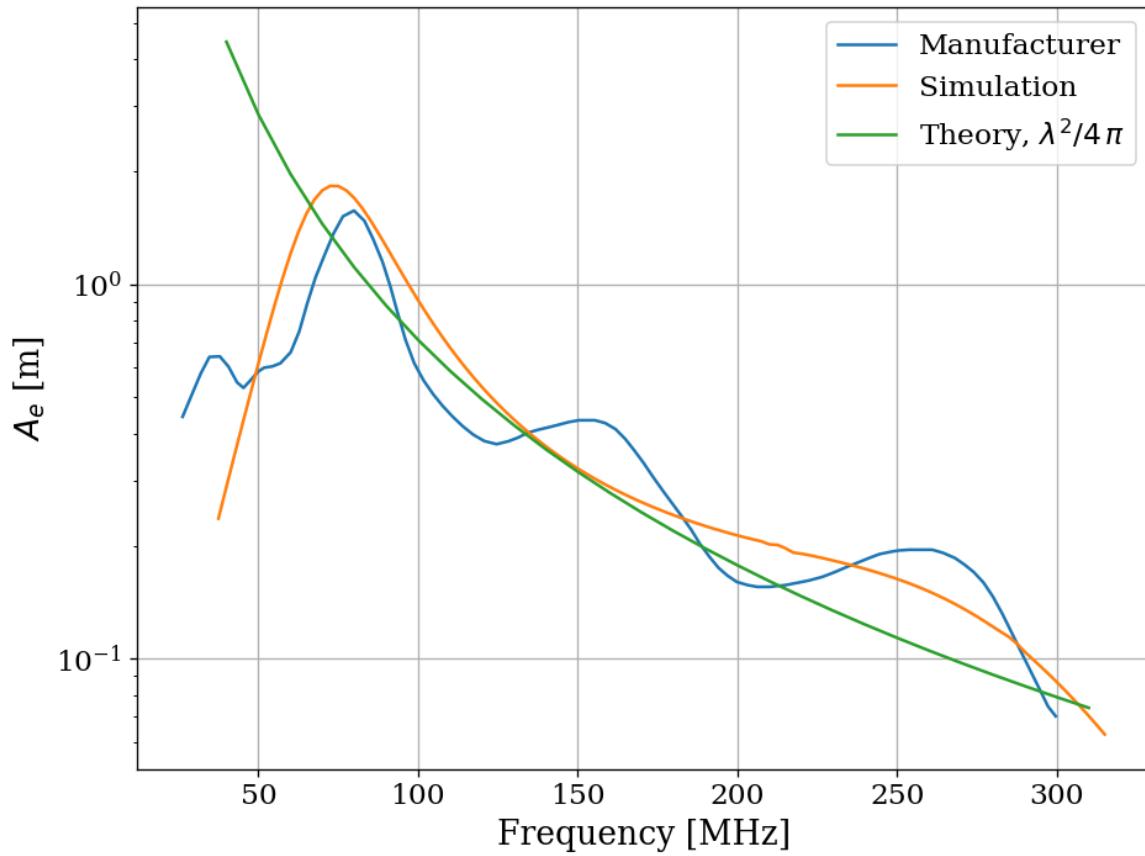


Figure 2.15: AB-900A biconical antenna effective aperture, simulated, measured and theoretical in free space. Simulation performed in COMSOL. The measurement was provided by manufacture [42] as an antenna factor and was converted using Eq. xxxx.

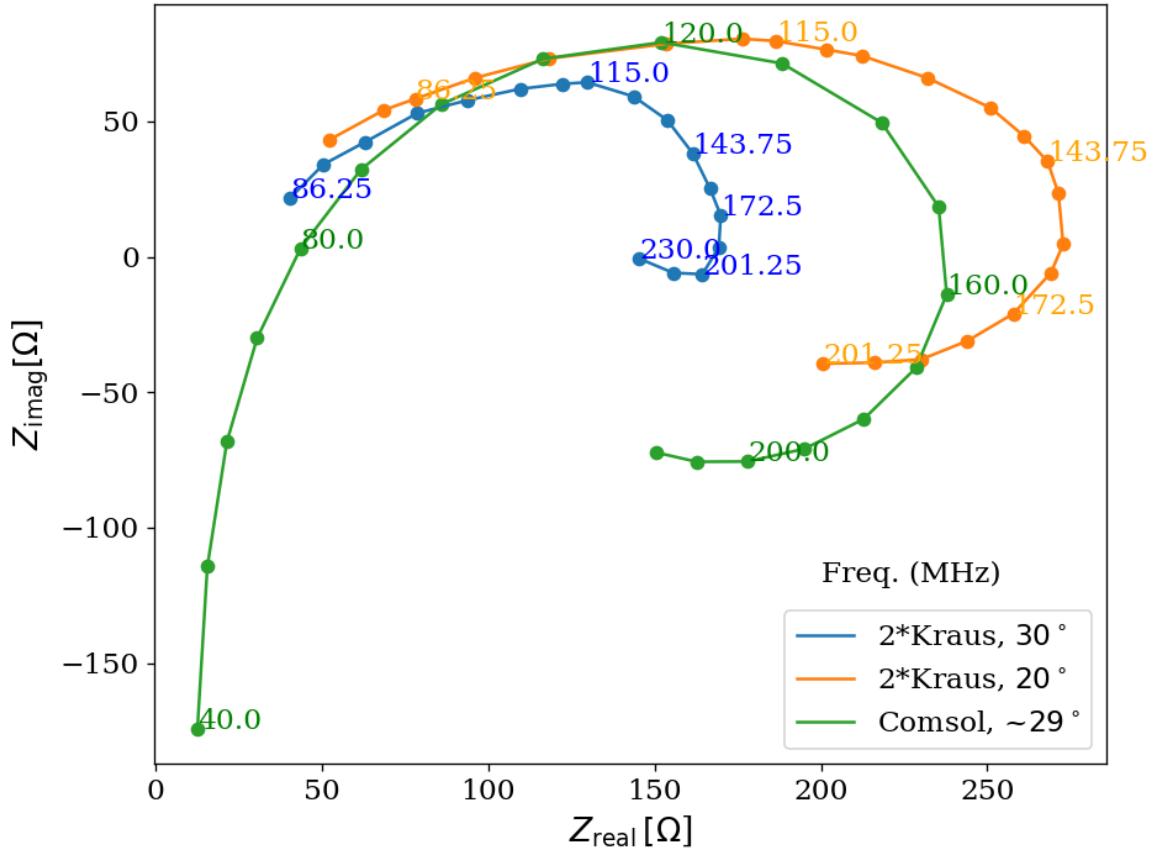


Figure 2.16: Simulated AB-900A biconical antenna free space complex input impedance shown in blue. Simulation was performed in COMSOL. Measurement from Kraus, second edition, Fig 8-13 [43], reproduced here in Fig. 2.17. Measured data is for a monoconical antenna and must be multiplied by 2 to compare to a biconical antenna. Numbers along curve indicate frequency in MHz.

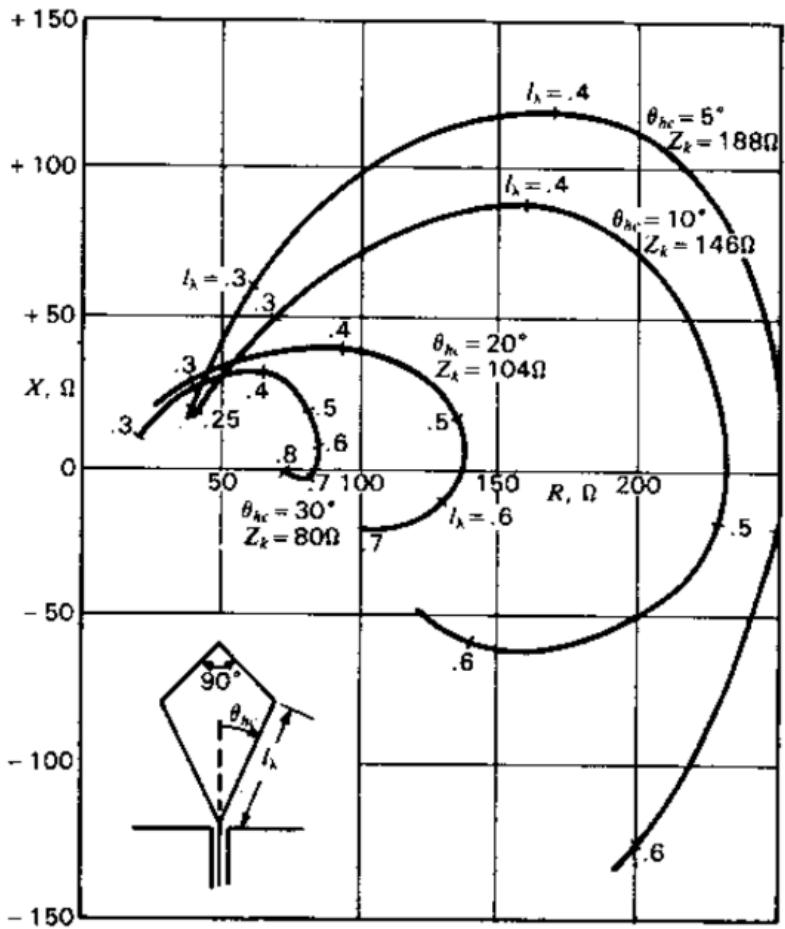


Figure 8-13 Measured input impedance of single cones with top hat as a function of cone length in wavelengths (l_λ). Impedance curves are presented for cones with half-angles of 5, 10, 20 and 30°.

Figure 2.17

874 **2.5.3 Terminator and fiber-optic switch control**

875 Experience has shown that it is advantageous to have a stable noise reference at the beginning
876 of the signal chain to measure periodically to monitor system function. One may think the
877 antenna can provide this following Sec. 2.3. In practice, it measures slightly less than this,
878 and has some frequency dependant variations. Furthermore, while RFI was not significant in
879 the actual run, it very well could have been, introducing some uncontrolled, coherent, power
880 into the spectrum which would have had an unknown origin. A terminator will give a very
881 constant $S_{\text{term}} = kT$ regardless of RFI and antenna position. For this reason one is included.
882 It is at the same temperature as the walls, and is a nice passive solution. The RF switch
883 shown in Fig. 2.13 is actuated periodically during the run to measure the terminator's noise
884 power. The switch is controlled via a fiber-optic link in order to reduce RFI. The fiber optic
885 is an extremely important feature which was overlooked for several years at the beginning
886 of the experiment leading to lots of RFI.

887 **2.5.4 Signal conditioning**

888 As discussed in Sec. 2.1.5, the ADC has internal signals which are mitigated by introducing
889 gain before the ADC. The amount of gain must be carefully chosen since too much will cause
890 the ADC to clip. Additionally, high frequencies must be limited before digitizing to prevent
891 aliasing. To accomplish, several RF components must be selected to condition the analog
892 signal: an LNA, a secondary amplifier, band pass filter and and several attenuators.

893 **2.5.4.1 Low noise amplifier**

894 The important concepts of the LNA have been introduced in Sec. 2.1.4. Relevant specifica-

895 tions for the Pasternack PE15A1012 [33] are summarized in table 2.2.

Specification	Value	Uncertainty (50-300MHz)	Units
Frequency Range	50-1,000	-	MHz
Gain	40	± 1	dB
Noise Temperature	110	± 10	K
Input Return Loss	< 15	-	dB
Output Return Loss	< 15	-	dB
Price	500	-	USD
Voltage	9-15		V
Current	100	10	mA

Table 2.2: Specifications for the Pasternack PE15A1012-E. The voltage is regulated internally, so the exact voltage supplied is not critical

896 We have several identical amplifier which are labeled with letters. At the time of writing,

897 amplifiers A-D are out of commission. Amplifier E was used for run 1.4.

898 **2.5.4.2 Secondary amplifier**

899 As discussed in 2.1.4, the noise temperature of a secondary amplifier has negligible impact

900 on the total system noise temperature. For this reason a cheaper secondary amplifier is

901 used. The specifications for the miniCircuits ZKL-1R5+ [44] are outlined in Table 2.3. This

902 amplifier has no internal regulator, so gain and noise temperature depend on the bias voltage.

903 This is set is set using an external regulator (built around a TI LM317 [45]).

Specification	Value	Uncertainty (50-300MHz)	Units
Frequency Range	50-1,000	-	MHz
Gain			
Noise Temperature			K
Input Return Loss	<	-	dB
Output Return Loss	<	-	dB
Price		-	USD
Voltage	9-15		V

Table 2.3: Specifications for the miniCircuits ZKL-1R5+ as measured with 9.0 V bias. There is no internal regulator, so the voltage is set using an external regulator (built around a TI LM317 [45]). **comment: finish this table**

904 2.5.4.3 Band pass filter

905 The band pass filter has two purposes. The first is to minimize the bandwidth entering
 906 the ADC so as much gain as possible can be introduced without wasting power amplifying
 907 frequencies where the antenna doesn't offer a good match. This will be computed below in
 908 Sec, The second is to prevent aliasing¹¹. Aliasing occurs when the analog signal contains
 909 frequency components at frequencies greater than half the sample rate, in otherwords when
 910 the signal and ADC don't obey a condition called the Nyquist criterion, $\nu_s/2 \nu$, where ν_s is
 911 the sample frequency and ν is the frequency of the analog signal.

912 When this criterion is not met, higher frequencies are mapped back down to a lower
 913 frequency, described by the aliasing formula:

$$\nu_{\text{alias}} = |\nu_{\text{signal}} - n \times \nu_s|, \quad (2.36)$$

¹¹Aliasing is rather complex topic which is greatly simplified here. Wikipedia's aliasing page is an excellent reference. For A more rigorous treatment, see Ch. 7 Sec. 3 of the second edition of Signals and Systems by Oppenheim et. al [46]

914 where ν_{alias} is the aliased frequency, ν_{signal} is the original signal frequency, ν_s is the
915 sampling rate, and n is an integer which specifies the Nyquist zone. In the simple case
916 with bandpass filters that roll off well within the first Nyquist zone, only $n = 1$ must be
917 considered.

918 As an example, a $\nu = 500$ MHz signal sampled at $\nu_s = 800$ MHz (the run 1.4 sample
919 rate), it would alias to $\nu_{\text{alias}} = 300$ MHz. This also means that 1,300 MHz (residing in the
920 second Nyquist zone) would alias down to 300 MHz, but again this and higher frequencies
921 won't be considered. This means that by choosing only to analyze frequencies less than 300
922 MHz, only frequencies greater than 500 MHz will alias into this analysis span.

923 The miniCircuits ZX75LP-288-S+ low pass filter was chosen [47]. It has an insertion loss
924 of 68 dB at 500 MHz, while being flat to within 2dB from 50 to 300 MHz. Additionally,
925 the miniCircuits SHP-50+ [48] was chosen as the high pass filter. When conected in series,
926 these two filters constitute the band pass filter.

927 2.5.4.4 Putting together a signal conditioning chain

928 The total gain required can be estimated by setting the output referred power of bandlimited,
929 400 K noise source times the gain equal to the maximum power the ADC can handle,
930 $V_{\text{RMS}}^2/Z = 0.63$ mW for $Z = 50\Omega$. Assuming a perfect filter from 50-300MHz (a 250 MHz
931 bandwidth), the gain required is

$$G = \frac{0.63 \text{ mW}}{k 400\text{K} 250 \text{ MHz}} = 4.5 \times 10^8 = 87 \text{ dB.} \quad (2.37)$$

932 Noise is a Gaussian random process however. This much gain ensures 1σ of the time

933 domain samples are below clipping. Since many samples are collected ($2^{24} \approx 1.6 \times 10^7$ in
934 run 1.4), and each has a probability of about 16% of clipping, many samples will clip with
935 87 dB of gain. However, it gives a good estimation for what to expect.

936 The band pass filter is not a brick wall from 50 - 300 MHz as was assumed in 2.37. The
937 effective bandwidth of a filter is defined here as the integral of the square magnitude of it's
938 through gain,

$$B_{\text{eff}} \equiv \int_{-\infty}^{\infty} d\nu 10^{S_{21}/10}. \quad (2.38)$$

939 Where S_{21} is measured in dB and we are interested in integrating a quantity that is
940 proportional to linear power, so it is divided by 10 rather than 20. The linear S parameters
941 of components connected in series multiply, but since dB are logarithmic, this is equivalent
942 to adding their S-parameters (in dB). Also note that $S_{21} \equiv \text{IL}$, the insertion loss. Insertion
943 loss is frequently given on data sheets.

944 With all this in mind, we can compute $B_{\text{eff}} = 237.06$ Hz for the actual filters from their
945 data sheets. Curves of S_{21} for the real band pass filter and an ideal brick wall filter are
946 shown in Fig. 2.18. Note that it is actually slightly less than the ideal 250 MHz bandwidth
947 since there is some loss in band.

948 Setting the gain is ultimately done by trial and error; taking a scan, adding some atten-
949 uation if it clips, taking a scan, etc. Experience and preliminary simulations have shown
950 that a few dB of clipping for an noise-dominated signal actually doesn't matter much, but
951 this was not fully explored. If future runs are to scan for longer, they may have a significant

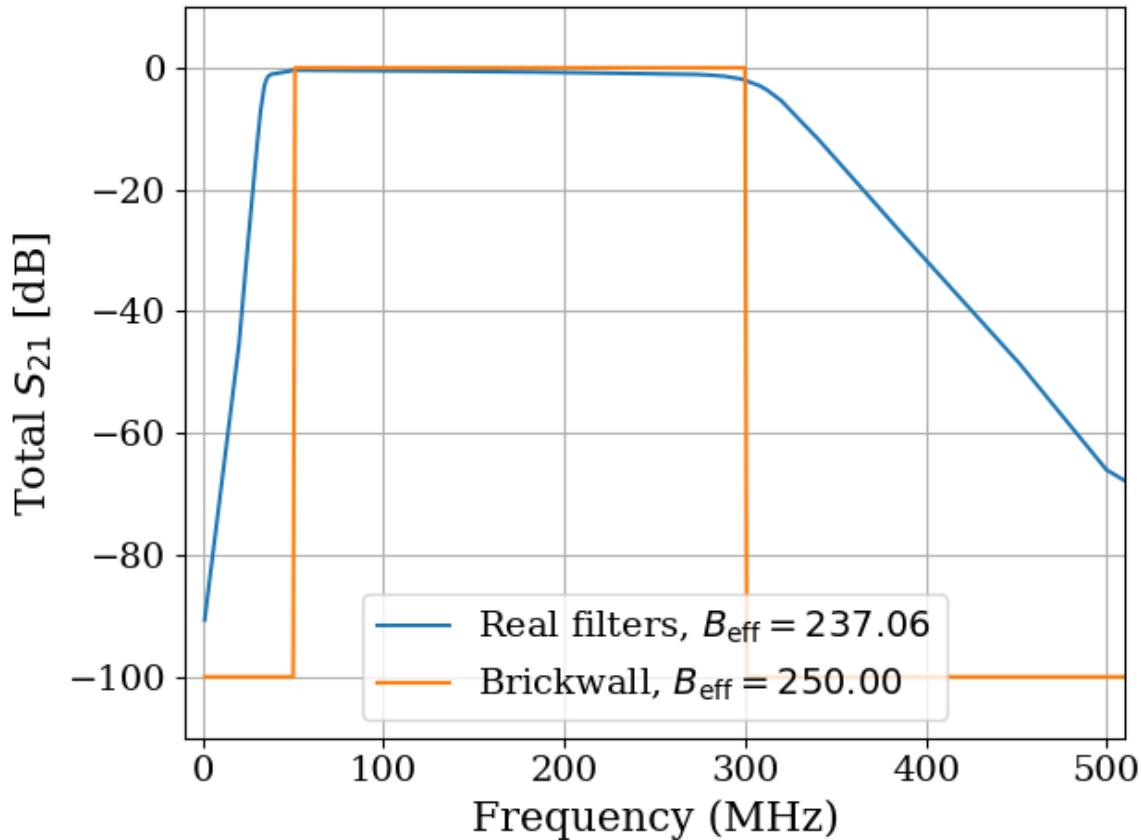


Figure 2.18: Through gain (S_{21}) of bandpass filter. Shown in dB, however B_{eff} is computed with linear S_{21} as in Eq. 2.38. Ideal brickwall filter shown for reference.
comment: Add Hz in legend

952 contamination from ADC effects, so additional gain driving into clipping may be fruitfully
 953 explored. Every dB of gain added allows for $10^{1 \text{dB}/5} \approx 44\%$ more averaging before ADC
 954 effects are at the same relative level ¹², so a little extra gain goes a long way.

¹²Calculations in dB like this are handy once they are understood, but can seemingly come out of nowhere. This can be worked out by a careful reading of section 2.2, using properties of logarithms and definition of the dB.

955 **2.5.5 Veto antenna**

956 The veto antenna is an identical Compower AB900 bicon antenna [49] as is used in the
957 shielded room. It is connected to channel A of the ADC, with no amplification. However, it
958 has a miniCircuits ZX75LP-288-S+ low pass filter [47] to prevent aliasing. This is the same
959 model low pass filter which is used for the main channel. The antenna is hung a few feet in
960 front of the door from some pipes on the ceiling. The proximity to these pipes likely give
961 the antenna a strange response, but it's purpose is simply to look for large RFI signals, so
962 this isn't that important.

963 **2.5.6 12 V power system**

964 **comment: Slow turn on circuit, plot of voltage vs time** The LNA and switch are active
965 component which require power to operate. The experiment is incredibly sensitive to RFI,
966 so while there is 120 VAC in the room, it is simpler to provide the power from a 12 V
967 battery than use a AC/DC regulator. Originally a 12 V lithium-ion (LiFePO₄) battery was
968 used. Lithium-ion batteries contain several cells and a controller to regulate charge/discharge
969 between the cells. In order to remove the possibility of this controller emitting RFI in the
970 room (which would create candidates that would be extremely difficult to veto), the lithium-
971 ion was replaced with a 12 V lead acid golf cart battery ¹³.

972 The LNA is interlocked to the door such that power is cut when the door is opened. This

¹³This was done between run 1.2 and 1.3 due to interference resulting in many candidate detections. Most of these were likely external RFI which were eliminated by cleaning the door (discussed in Sec. 3.2. Two variables were changed (cleaning of the door and replacement of the battery) so it's unclear if a lead acid is necessary. It works however, and if it aint broke don't fix it.

973 prevents the large radio signals being amplified once they enter the room and protects the
 974 ADC. The circuit that controls this is shown in Fig. 2.19. Note that it takes a few seconds
 975 for the slow turn on circuit to discharge, so the door should be opened slowly

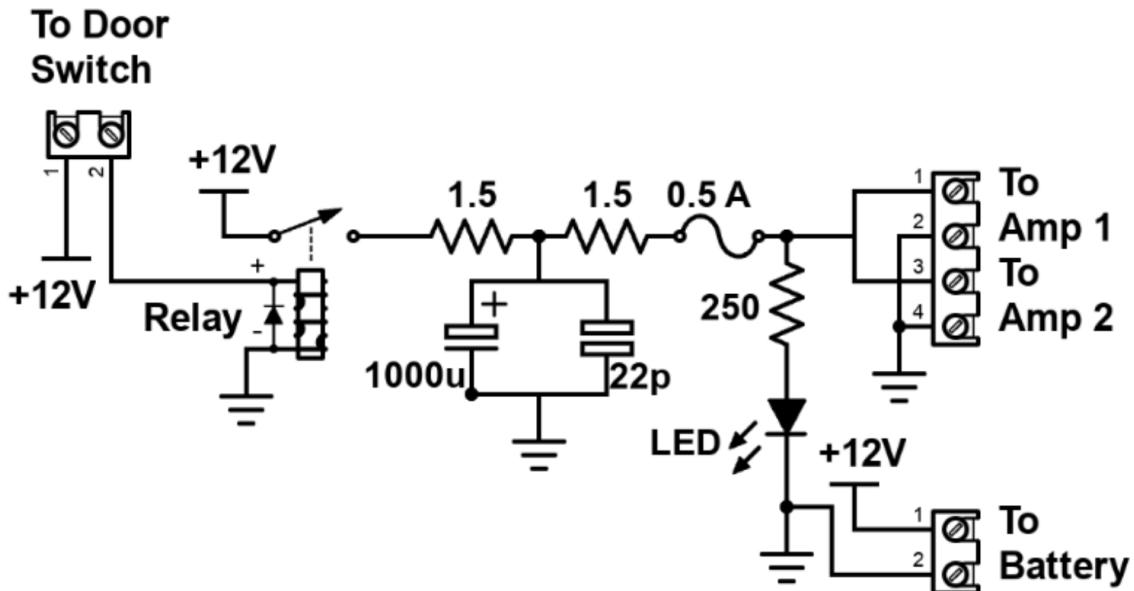


Figure 2.19: Schematic of Interlock board. No voltage regulation is provided because it is designed to work with amplifiers containing internal regulation (Pasternack PE15A1012 [33]). Not shown is a simple "slow turn on circuit" consisting of a 0.68 F capacitor and a 8Ω , 10 W resistor (time constant = 5.4 seconds) which was installed to protect the amplifier from transient voltages when the door is closed (can be seen in Fig. 2.20). Experience has shown the liberal use of fuses to be prudent when working with car batteries in a metal room.

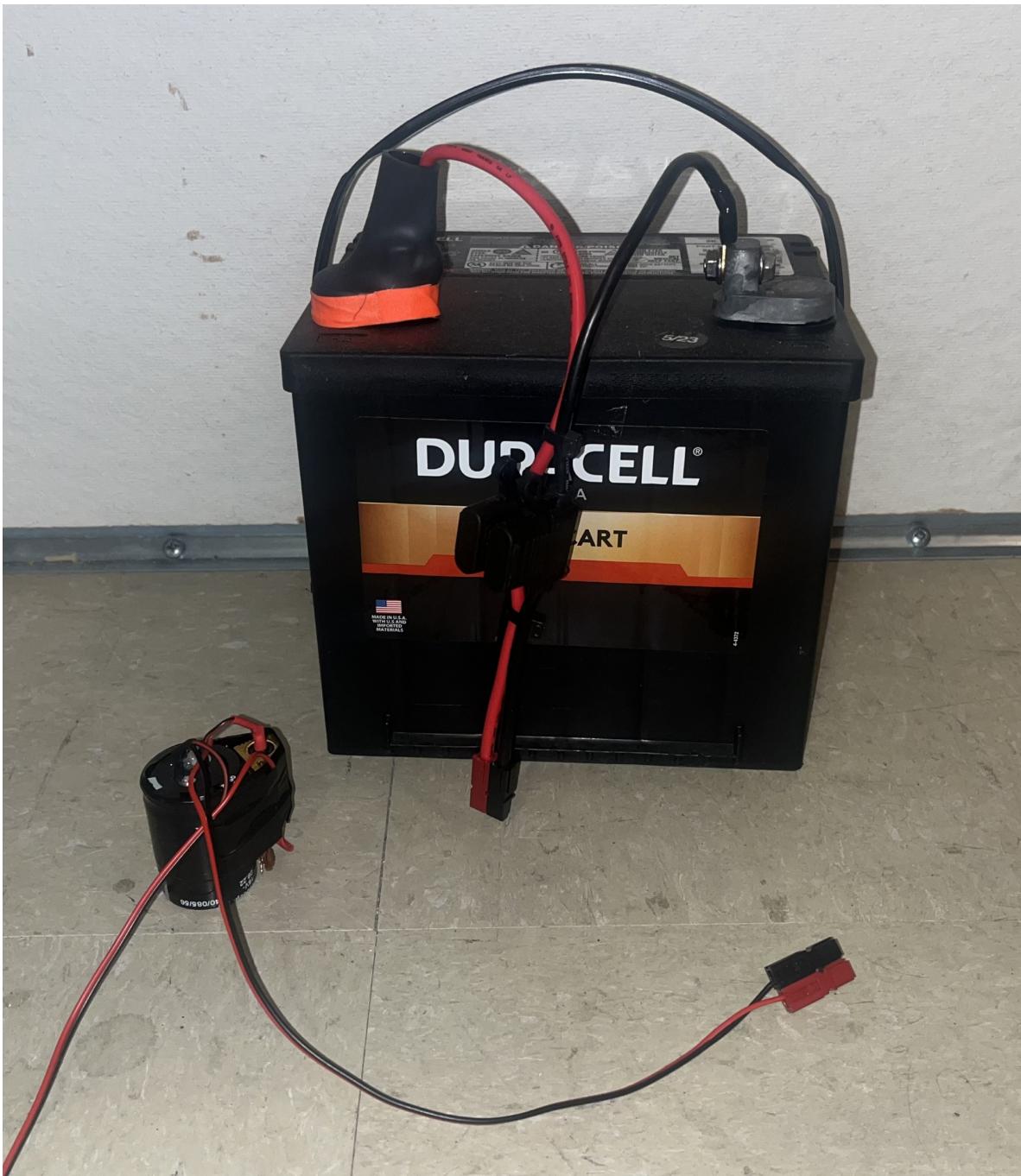


Figure 2.20: Photo of 12 V lead-acid battery and slow turn on circuit. Both connect directly to the interlock board shown in Fig. 2.19. Slow turn on circuit is connected between interlock board and amplifier. Battery includes a 2A fast fuse installed in commercial Anderson power pole assembly.

976 **2.5.7 GPU-Based Real-Time Spectrum Analyzer**¹⁴

977 The use of commercial Spectrum Analyzers (SAs) which feature so-called real time spectrum
978 analyzer (RTSA) mode come with several restrictions which limit the efficiency with which
979 they are able to perform wide-band scans with narrow frequency resolution, as pointed
980 out the dark radio pilot run [4]. The number of frequency bins output by a real discrete
981 Fourier transform (DFT) is equal to half of the number of time domain samples, while the
982 bandwidth is given by half of the sample rate. Furthermore, the ability to acquire data in real
983 time requires a DFT algorithm (generally implemented as a fast Fourier transform, FFT)
984 and computational resources which can operate on time domain data at least as fast as it is
985 acquired. In practice, real-time DFTs with high frequency resolution and wide bandwidth
986 require modest memory, transfer rates and processing resources. Commercial “real time
987 spectrum analyzers” tend to cheat a bit to reduce hardware requirements. Frequency mixers
988 reduce required sample rate (and therefore span), limits on FFT lengths reduce either span
989 or resolution, and limits on rate of scans make it so that these SAs aren’t real-time (at least
990 in the sense that we require them to be).

991 For this reason, we have constructed our own SA based on the Teledyne ADQ32 PCIE
992 digitizer[34], which is wide-bandwidth (up to 1.25 GHz frequency span), high resolution (2^{24}
993 point FFT), and nearly 100% real-time (see Fig. 3.17. We have been unable to find a
994 commercial SA with comparable capabilities. Specifications are shown in table 2.4.

¹⁴Code for this section can be found at: <https://github.com/josephmlev/darkRadio/tree/master/daqAnalysisAndExperiments/teledyne>

Bit depth	12 bits
Sample rate	800MHz
DFT input length	2^{24} samples
FFT compute time	2 ms
Channel count	2
Efficiency	99.765%

Table 2.4: Specifications for the spectrum analyzer used for run 1A.

995 After passing through and amplifier and filter chain outlined in Sec. 2.5.4, both the main
 996 and veto antenna's RF signals are digitized by the ADQ32's two independent ADCs. This
 997 raw, digitized time series is sampled at the digitizer's clock rate. Since the discrete Fourier
 998 transform (DFT) of a perfect sinusoid sampled by an unstable clock will have a finite spectral
 999 width, clock stability must be better than the expected spectral width of candidate signals,
 1000 which in our case is set by the expected $Q_{DP} \approx 10^6$. To achieve the required stability, we
 1001 synchronize the sample clock (Valon 5009 RF synthesizer) of our ADC to a 10 MHz rubidium
 1002 frequency standard (Stanford Research Systems FS725) which is further steered by the one
 1003 pulse-per-second (pps) signal from a GPS receiver.

1004 We utilize a GPU direct write in order to minimize CPU-GPU copies which in our
 1005 experience tend to be significantly slower than the FFT itself. This GPU direct write is
 1006 implemented by the digitizer's C++ API which we call from python. This happens in
 1007 the `teledyneTemplate/drDaq.py` script (within the directory linked as a footnote at the
 1008 beginning of this subsection). The process is shown graphically in Fig. 2.21

1009 First, several buffers are allocated in GPU memory. A record is acquired (2^{24} time
 1010 domain samples for run 1.4). This record is written to a pre-allocated, time-domain buffer

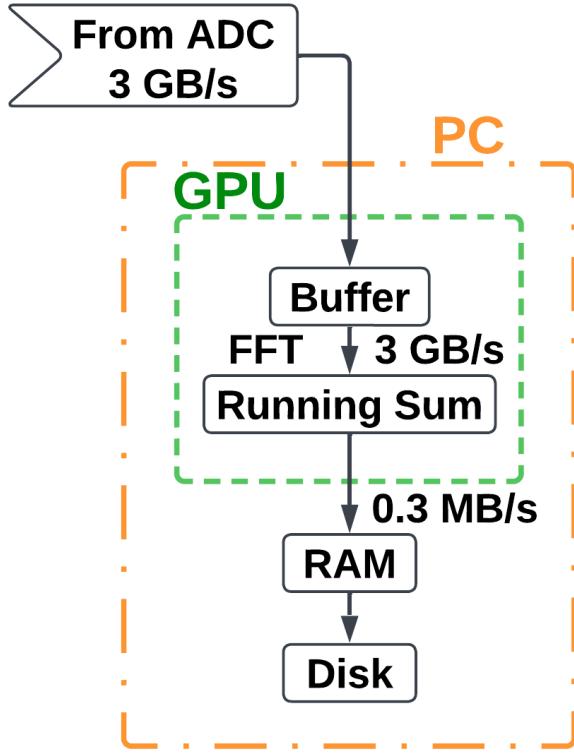


Figure 2.21: Real time DAQ data stream. Approximately 10^4 time series records (about 3 minutes of real time data) are written from the ADC directly to GPU buffers. FFTs are performed on these records resulting in a pre-averaged spectrum which can be saved to disk. This set up is duplicated for channels A and B, though the data rates indicate the sum of both channels.

1011 in the GPU's memory. Next, an FFT is performed using Pytorch which we found to be
 1012 the fastest across all the GPUs we tested (all Nvidia). [50]. Approximately 10^4 FFTs are
 1013 performed and added to a cumulative sum on the GPU (representing about 3 minutes of
 1014 real time data). Dividing by the number of FFTs provides an averaged spectrum that is
 1015 saved for offline processing. This *pre-averaging* reduces the raw ≈ 1.5 GB/s/channel data
 1016 stream to ≈ 0.15 MB/s/channel, which greatly reduces storage requirements. However, this
 1017 comes at the cost of temporal resolution of transient candidates. Since we are interested

1018 in constant (or at least very slowly varying) signals, this is not a problem, but in other
1019 radio astronomy applications, this step should be avoided. Pre-averaging is set using the
1020 NOF_BUFFERS_TO_RECEIVE variable, see Appendix ?? .

1021 **Chapter 3**

1022 **System Characterization and Data**

1023 **Acquisition System**

1024

“When you can measure what you are speaking about and express it in numbers you know something about it; but when you cannot measure it, when you cannot express it in numbers your knowledge is of meagre and unsatisfactory kind; it may be the beginning of knowledge but you have scarcely progressed in your thoughts to the stage of science whatever the matter may be.”

Lord Kelvin

1025 This chapter covers the characterization of the system as a whole, including the shielding
1026 room (Introduced in Sec. 2.5.1), amplifier chain (Sec. 2.5.4) and real time spectrum analyzer
1027 system (RTSA, Sec. 2.5.7). This section provides data and information about how it was
1028 collected from tests that were performed, but that distract from the narratives of the pre-
1029 ceding and following chapters. This chapter can be skimmed and referenced with further
1030 care during the reading of Ch. 4.

1031 **3.1 Measurement of Amplifier Chain Performance**

1032 The amp chain provides conditioning of the analog signal (described in Sec. 2.5.4). The
1033 following data show some measurements that are useful in confirming the system is operating
1034 as expected. Before, during and after a new data run, similar data should be collected and
1035 inspected so that any degradation of performance can be monitored. At the very least, one
1036 should take note of the power spectrum from a room temperature terminator (Fig. 3.11) and
1037 monitor it for any changes. The heads up display (Fig. A.1) makes this very easy.

1038 **3.1.1 Y-factor method¹**

1039 The Y-factor method [51] [52] is a common technique for measuring the gain and noise tem-
1040 perature of an amplifier chain (details in Sec. 2.1.4). A matched terminator is placed on
1041 the end of a transmission line which is connected to the input of an amp chain under test.
1042 The output of the amp chain is measured with the terminator at two different tempera-

¹Code for this section can be found at: <https://github.com/josephmlev/darkRadio/tree/master/daqAnalysisAndExperiments/run1p4/yFactorTest/yFactorTest.ipynb>

1043 tures². These temperatures must be known, and the larger the difference the better the
1044 measurement.

1045 There are a few different ways to handle the algebra, but the simplest is to fit a line of
1046 the form

$$T_{\text{out}}(\nu) = G T_{\text{term}} + B, \quad (3.1)$$

1047 where all terms are frequency dependant. The X-intercept (and therefore the negative
1048 amplifier temperature) than is simply G/B . This is shown as a cartoon for a single frequency
1049 in Fig. 3.1.

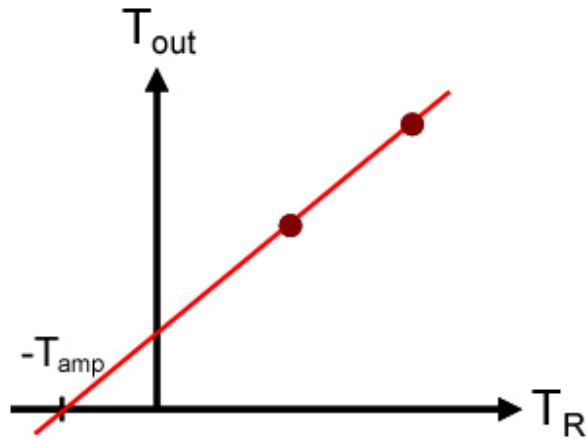


Figure 3.1: Cartoon showing Y-factor data. The (negative of) amplifier temperature is read off of the X-intercept. The gain is given by the slope. Figure from Wikipedia [53].

1050 To make this measurement, I wanted to use liquid nitrogen for the cold temperature
1051 (77 K) and a calibrated noise source for the hot temperature (12,000 K). The test was set

²alternatively, a calibrated noise source can be used to simulate a very hot terminator, usually thousands of degrees Kelvin.

1052 up for use of the noise source and used a 30 dB attenuator which was calibrated out later.
1053 However, while the test was set up to use this noise source, it gave unreliable results. I was
1054 able to confirm it became uncalibrated using a spectrum analyzer so I didn't use that data
1055 point. Additionally, I made the measurements with a much wider bandpass filter (0-1 GHz)
1056 so that the same data would also give the gain and noise temperature up to 1 GHz (though
1057 this is not shown in this thesis). Both the 30 dB attenuator and the different bandpass filter
1058 will be calibrated out. As a check this is done correctly, the gain is measured again using a
1059 tracking generator with the actual set up, avoiding corrections. The two gain curves agree
1060 quite well.

1061 The cryogenic set up is shown in Fig. 3.2. The raw data (including only liquid nitrogen
1062 and room temperature) are shown in Fig. 3.3.

1063 The spectra of Fig. 3.3 contain hot and cold measurements at 2^{18} frequency points. At
1064 each point, a fit is performed according to Eq. 3.1. To demonstrate, this is shown for a single
1065 frequency (625 MHz) in Fig. 3.4.

1066 Finally, fitting the raw spectra shown in Fig. 3.3 with Eq. 3.1 at each of the frequency
1067 points, the fully frequency dependant gain and noise temperature of the amp chain can be
1068 extracted. These are shown (after correcting for the 30 dB attenuator) in Figs. 3.5 and 3.6.

1069 As a final check, I took the run 1.4 terminator data and corrected it for the gain and
1070 noise figure found using the Y-factor method of this section, expecting to recover the famous
1071 -174 dBm/Hz as predicted by Johnson's formula for a terminator at room temperature. The
1072 agreement was quite good and is shown in Fig. 3.7.

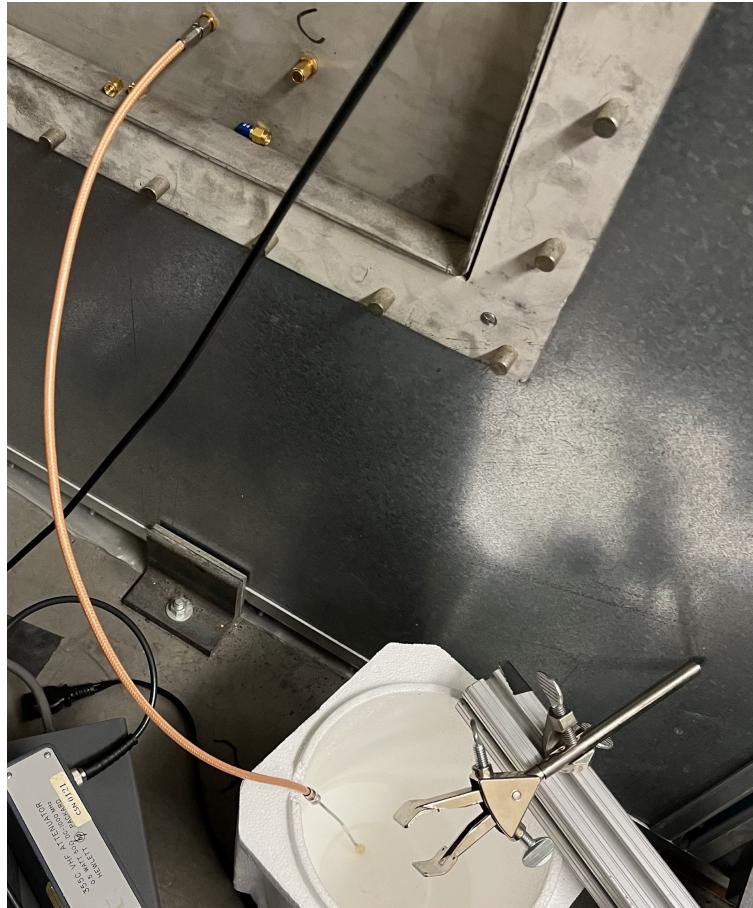


Figure 3.2: Hardware set up of Y-factor measurement. Note that actual data, cryogenic cable and terminator were inserted directly in place of the antenna, as depicted in Fig. 2.13. Also note that terminator is a standard Pasternak 50Ω (with the blue rubber case removed) and cable is semi-rigid cryogenic-specific. This test load (semi-rigid cable plus terminator) were confirmed to have an $S_{11} < -35 \text{ dB}$ at both room temperature and after a few minutes in liquid nitrogen, confirming performance at cryogenic temperatures.

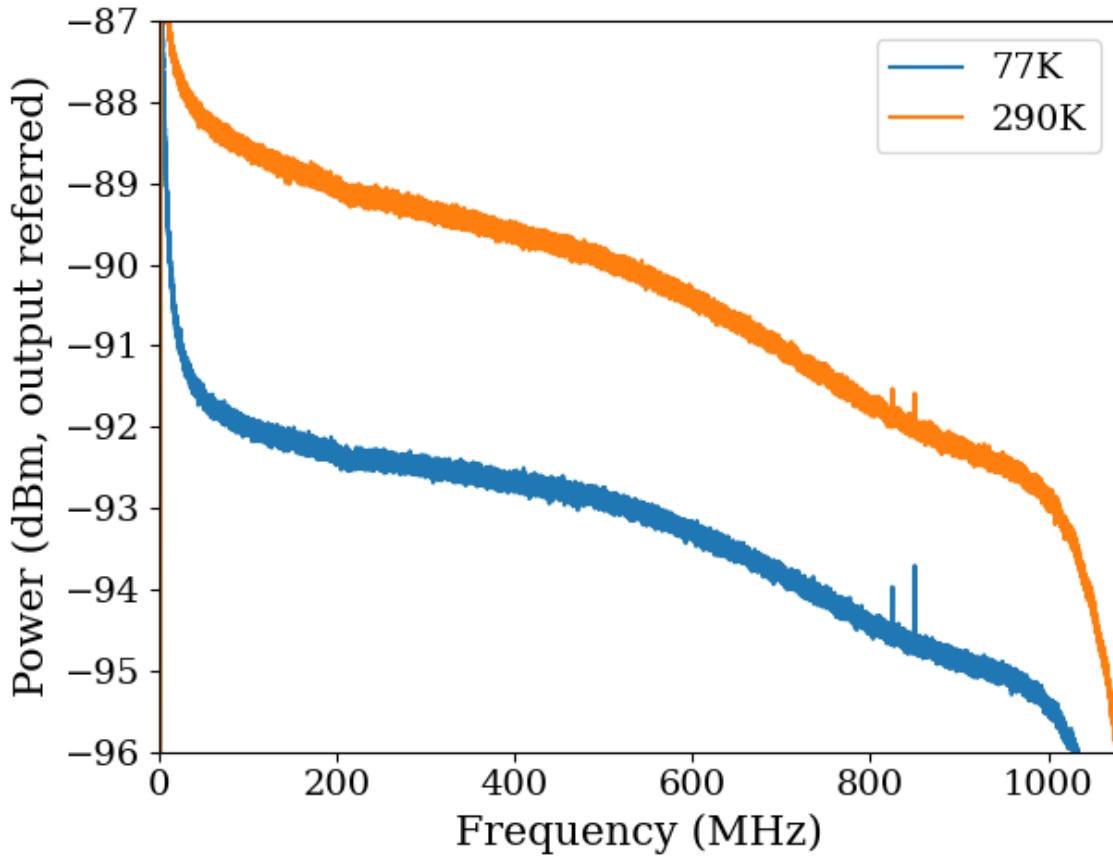


Figure 3.3: Raw output power spectra used in computing gain and noise temperature for run 1.4 amp chain using the Y-factor method. $\Delta\nu_{\text{RF}} \approx 9.5 \text{ kHz}$, and 10k averages were taken. These data are only used to characterize the system between 50 and 300 MHz, so the low frequency ADC effects and RFI around 800 MHz do not affect the calculation. Note that the amp chain under test has a 30 dB attenuator directly before the ADC which is discussed in the text. Obviously this had an impact on the measured gain and it will be corrected for in the following analysis. It has negligible impact on the noise figure however, as shown in Eq. 2.29.

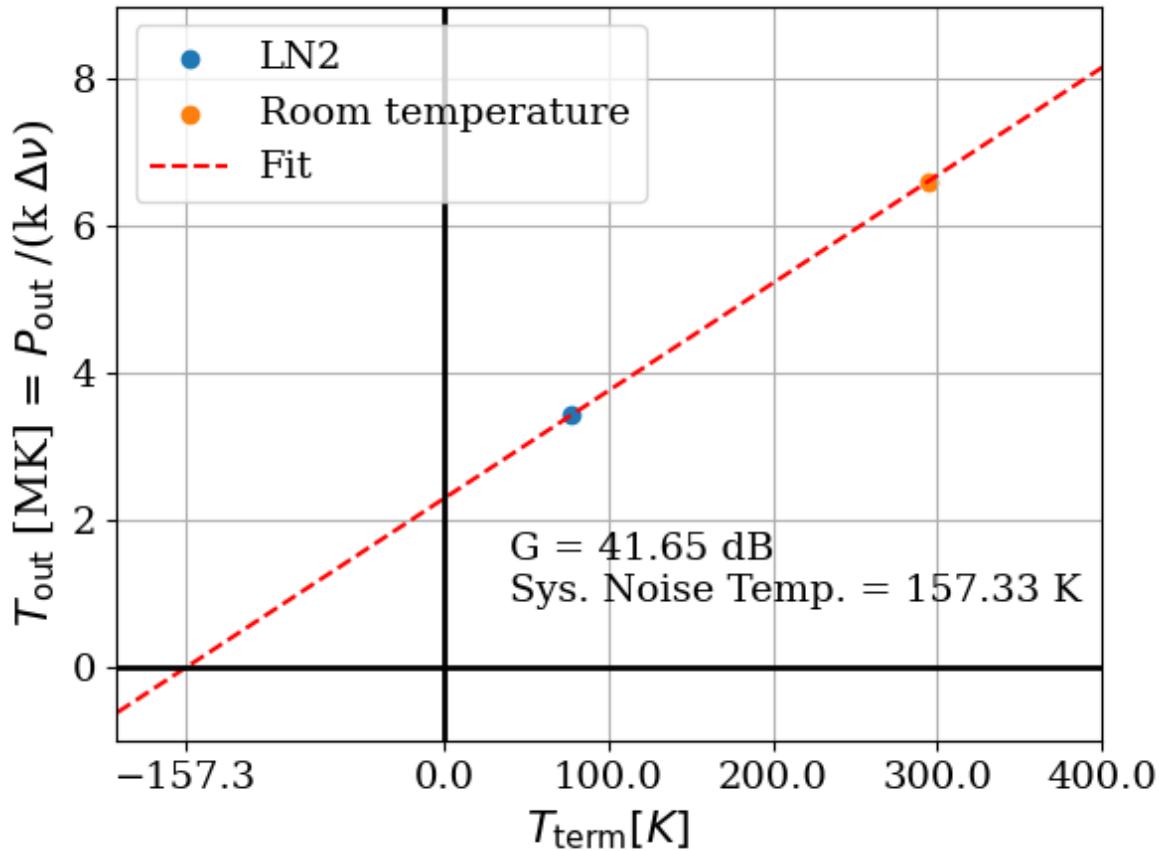


Figure 3.4: Y-factor data from run 1.4 amp chain at a single frequency (625 MHz). Note that T_{out} is in MK or millions of degrees Kelvin. This is expected since there is approximately 41 dB of gain on an input temperature of a few hundred Kelvin. Reminder, there is a 30 dB attenuator which has not yet been corrected for.

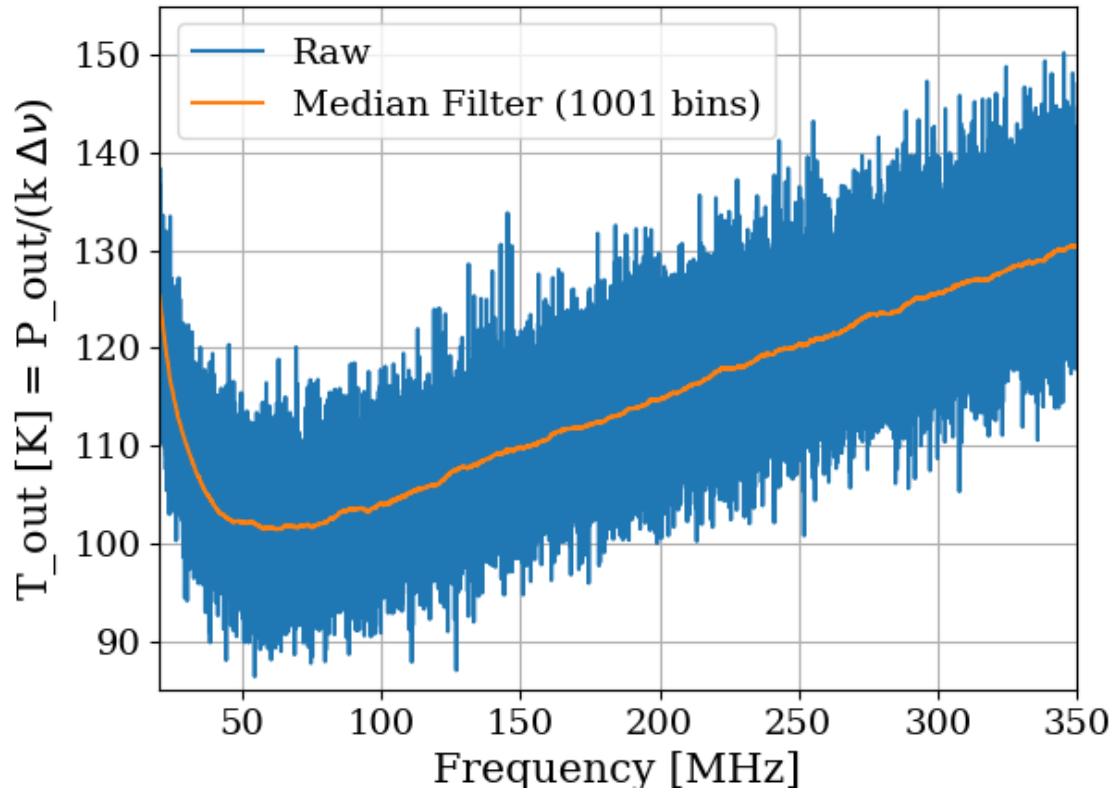


Figure 3.5: Frequency dependant noise temperature of run 1.4 amplifier chain, measured with Y-factor method. Low frequency behavior is consistent with the data sheet of the LNA [33]. Increase at high frequency due to attenuation of the cable before the LNA. Future runs may wish to place the LNA directly on the antenna to minimize this issue, though other complications will be introduced.

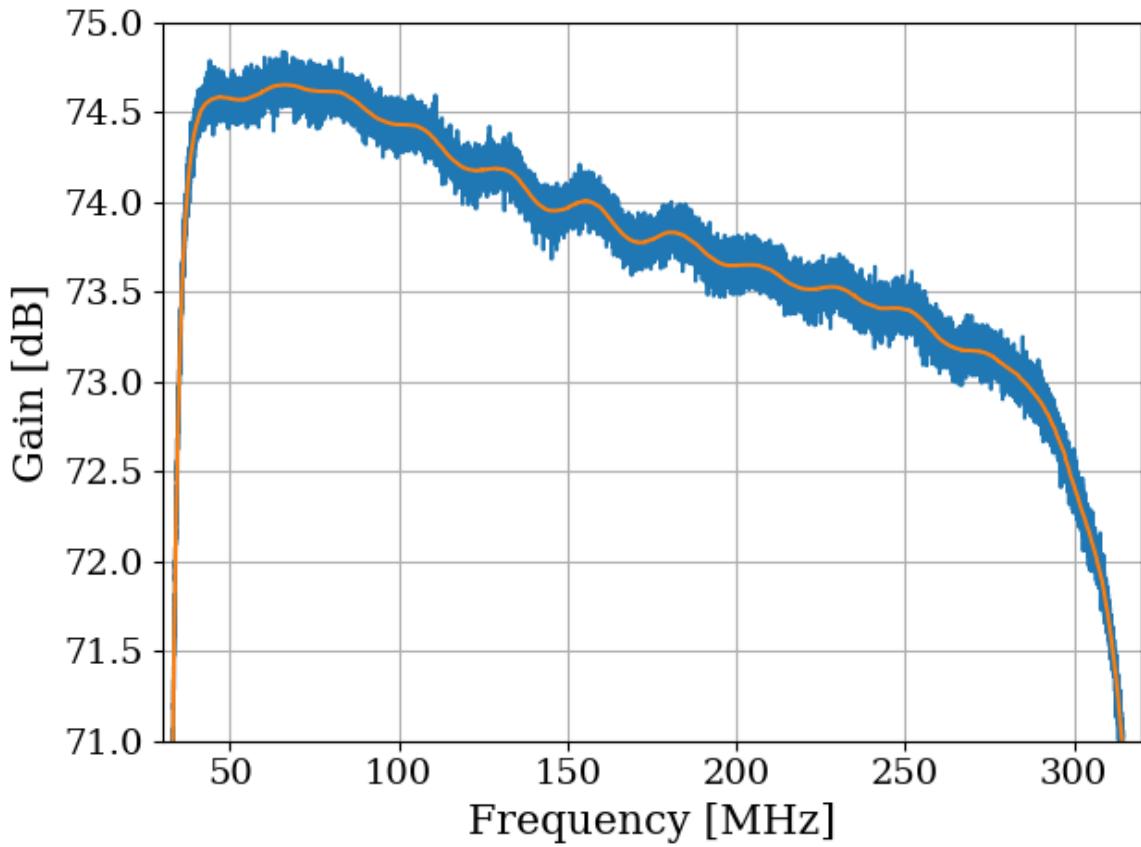


Figure 3.6: Frequency dependant gain of run 1.4 amplifier chain, measured with Y-factor method. The 30 dB attenuator and band pass filter have been corrected for in this plot. Slow decrease of gain with frequency is consistent with the data sheets of the LNA [33] and secondary amplifier [44]. Roll off at high and low frequencies is due to bandpass filter (see 2.5.4.3). Bumpy behavior is due to small impedance mismatch of the terminator.

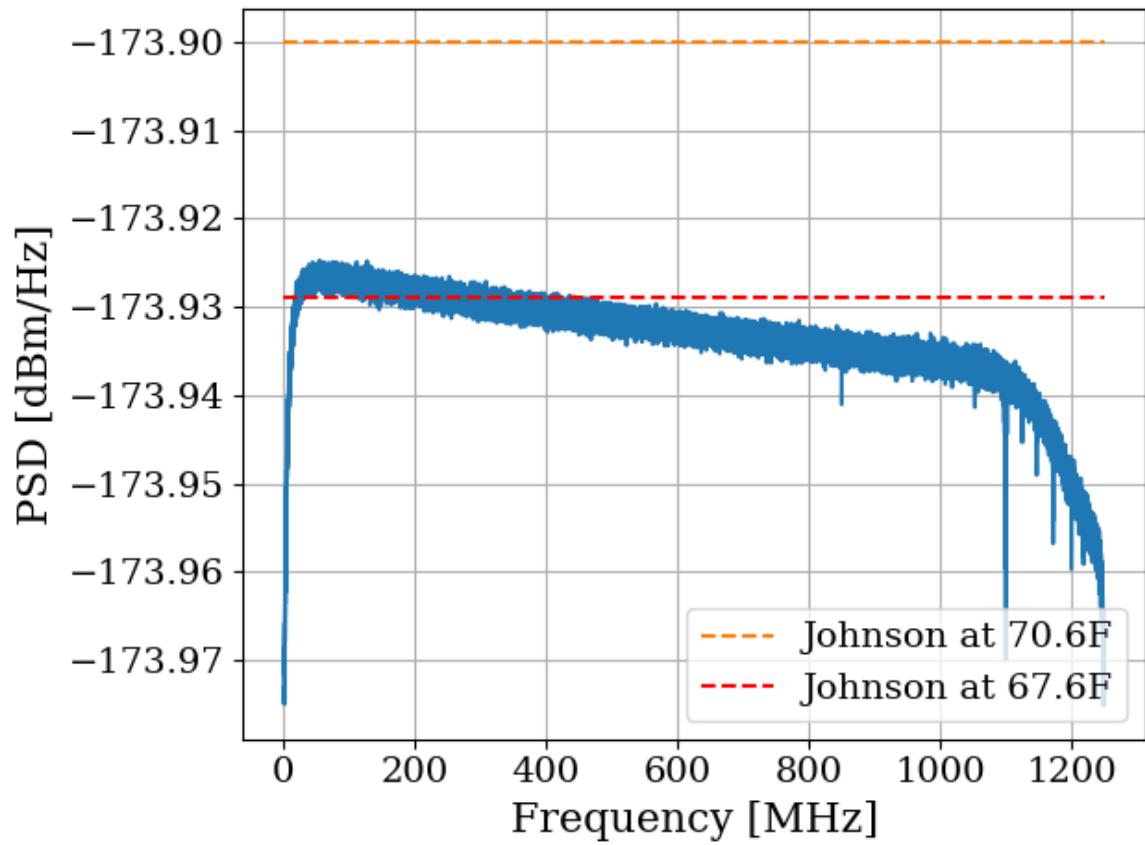


Figure 3.7: Run 1.4 terminator, input referred power spectral density. Excellent agreement with the predicted value from the formula for Johnson noise. Some reference temperatures are given with the dashed lines.

1073 **3.1.2 Confirming gain with tracking generator measurement**

1074 As a sanity check for the Y-factor method, the gain was measured again using the tracking
1075 generator of the Rigol RSA 5065TG [54], with good agreement. This measurement was
1076 useful since there were some corrections involved in the Y-factor measurement (the 30 dB
1077 attenuator and band pass filter).

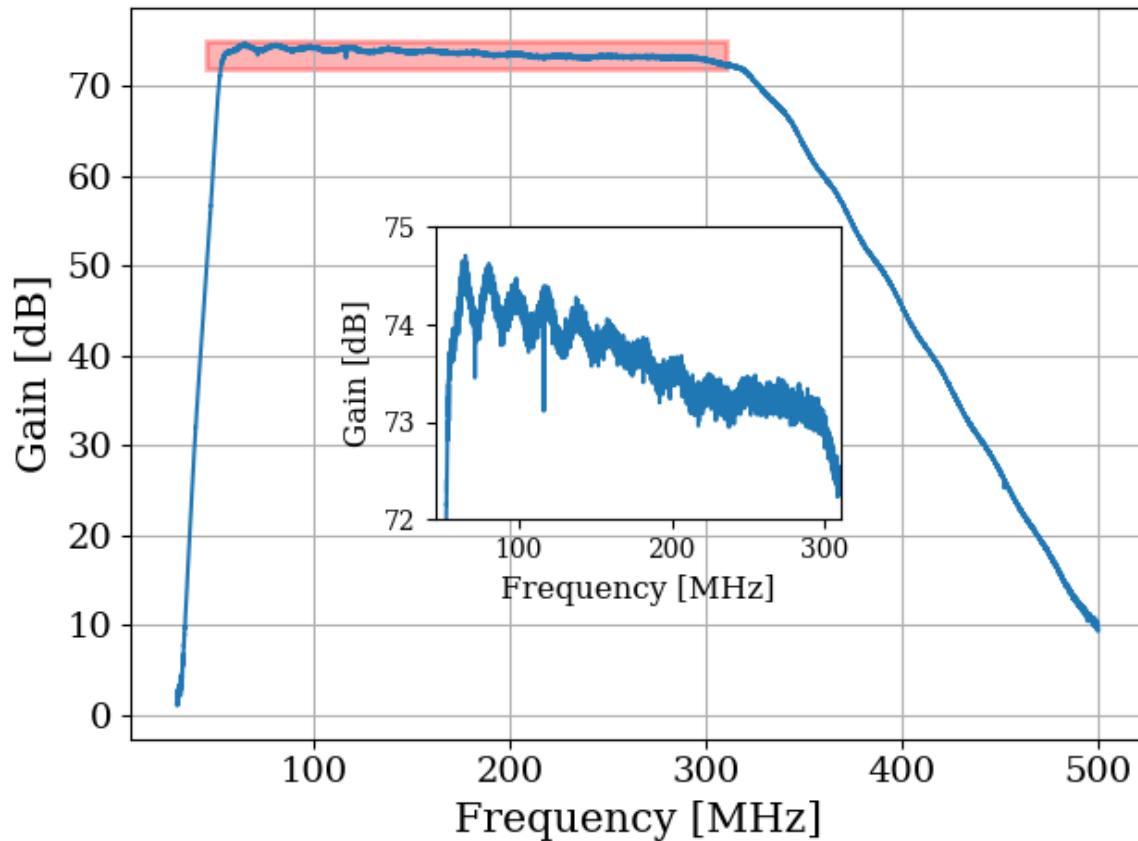


Figure 3.8: Gain vs. frequency of run 1.4 amplifier chain, including LNA, secondary amplifier band pass filter and attenuators. Inset shows zoom of region enclosed in red box. Measured using tracking generator on Rigol RSA 5065TG [54]. Wiggles visible in inset are a known artifact due to an impedance match of the Rigol's tracking generator. Gain was confirmed by comparison to Y-factor method (Sec. 3.1.1)

1078 **3.1.3 System stability over a run period**³

1079 The system's gain will likely vary slightly over the course of a run. An obvious culprit is
1080 the battery voltage decreasing with time⁴. At first glance, one may conclude that since the
1081 LNA is equipped with an internal voltage regulator, it is insensitive to the battery voltage
1082 as long as it's above some minimum value. However, at higher voltage, the regulator must
1083 dissipate more power and will make the LNA warmer which may have an effect. Tests of
1084 gain vs LNA voltage show that this is a minor issue. This is shown in Fig. 3.9.

1085 Knowing the gain vs voltage is only useful if one understands how the battery voltage
1086 will behave over the course of the run. These data is shown in Fig. 3.10.

1087 As a final check of system performance during the run, Fig. 3.11 shows the output referred
1088 power (averaged from 50-300 MHz) over the course of the 9 day run 1.4. The gain is shown to
1089 vary by about 0.7%, an insignificant amount given the other uncertainties involved (especially
1090 that of simulation and calibration, Sec. 4.3).

³Code for this section can be found at: [https://github.com/josephmlev/darkRadio/tree/master/daqAnalysisAndExperiments/run1p4/run1p4_timeDependence.ipynb](https://github.com/josephmlev/darkRadio/tree/master/daqAnalysisAndExperiments/run1p4/run1p4_analysis/run1p4_timeDependence.ipynb)
https://github.com/josephmlev/darkRadio/tree/master/daqAnalysisAndExperiments/run1p4/gainAndNf/gainAndNF_analysis.ipynb

⁴only the LNA is powered by battery. The secondary amplifier is outside of the room and powered with a voltage regulator.

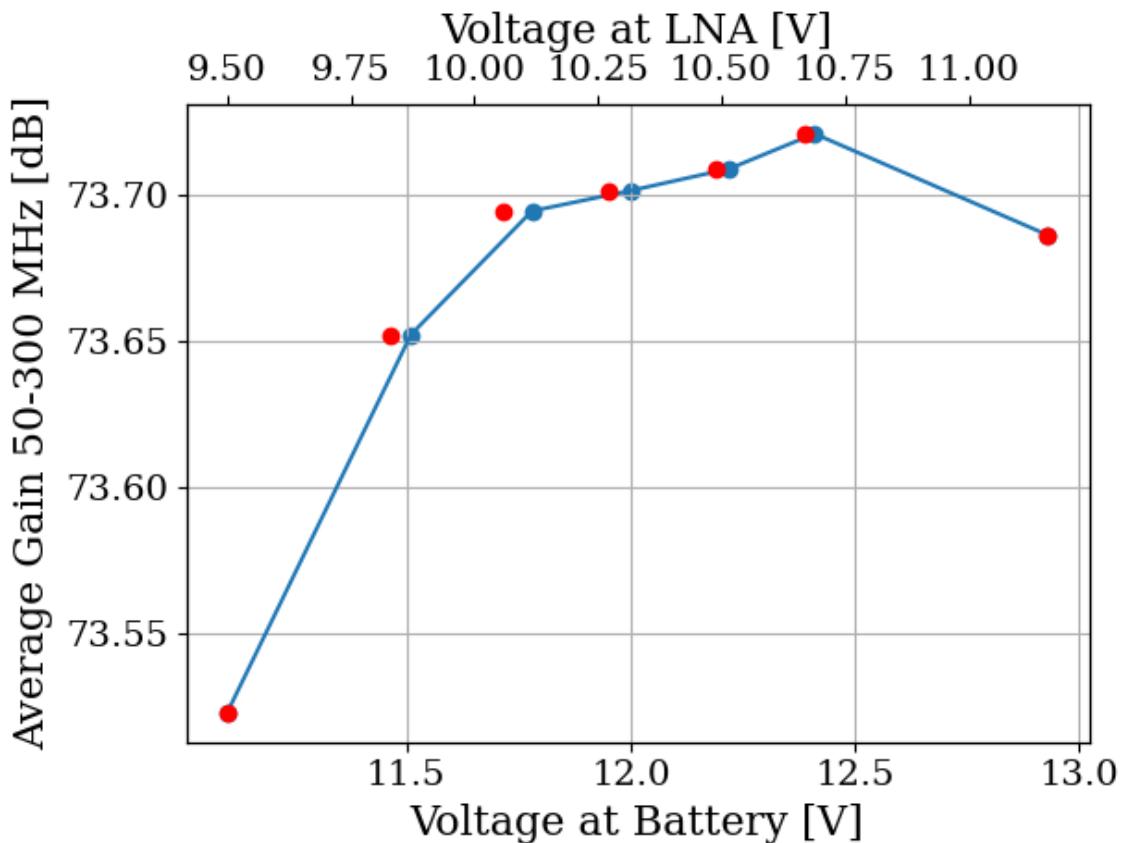


Figure 3.9: Gain vs. LNA voltage of run 1.4 amplifier chain, including LNA, secondary amplifier, band bass filter and attenuators. Measured using tracking generator on Rigol RSA 5065TG [54]. Gain is frequency dependant; its average value between 50 and 300 MHz is shown here. Voltage at amplifier was measured directly at terminals, after significant voltage drop due to interlock and slow turn on circuits (described in Sec. 2.5.6). Gain is in good agreement with the previous two subsections, 3.1.1 and 3.1.2. Voltage range specified on LNA data sheet [33] is 10-15 V.

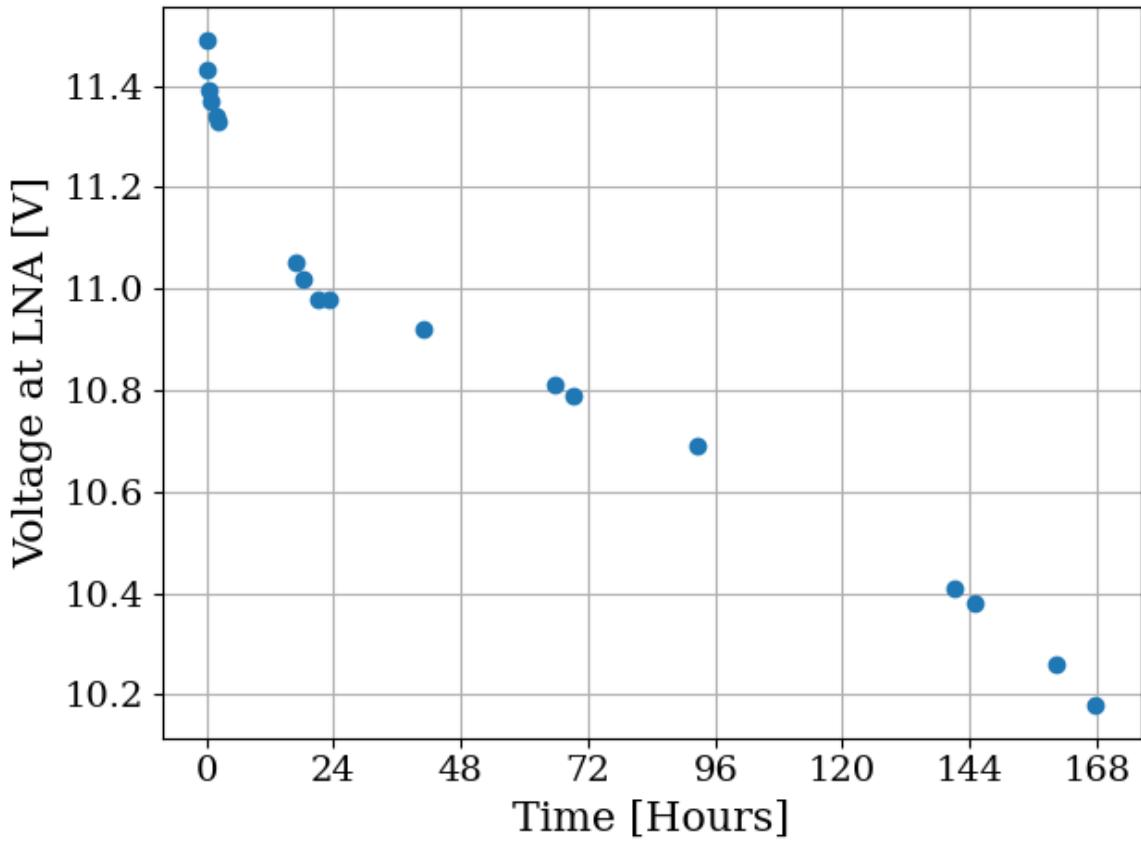


Figure 3.10: Voltage at LNA vs time. Data acquired 7/15/2024 through 7/22/2024 using battery number 2 immediately after being charged. The system was set up as it was for run 1.4, including the switching cycle (on for 3 minutes, off for 45 minutes). The battery supplies 177 mA when the switch is off, and 370 mA when the switch is on. There is an additional data point (not shown) around 196 hours where the voltage has dropped to 6.2 V, well below the minimum required voltage. The battery had approximately 20 charge/recharge cycles at the time of this test and was about 14 months old; a repeat of this test may be required to use the battery for longer periods between charges.

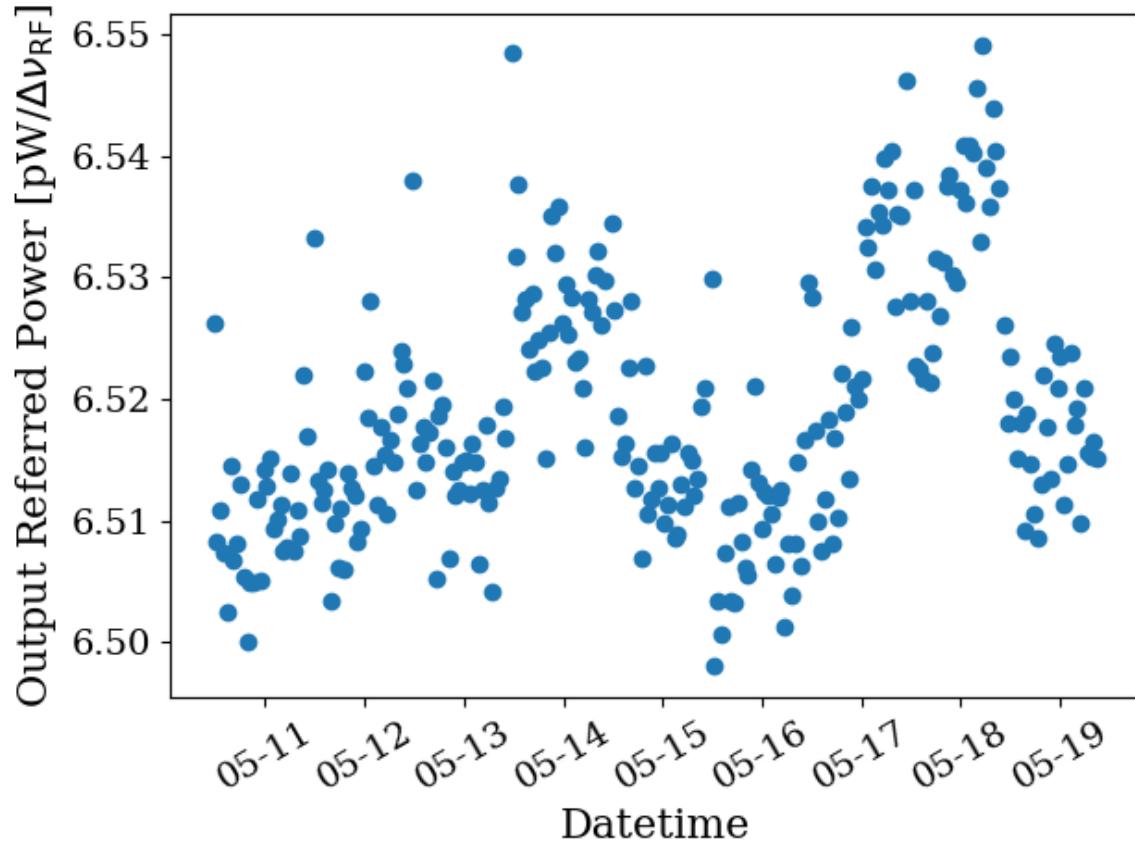


Figure 3.11: Frequency-averaged power vs time of the terminator from run 1.4 (schematic shown in Fig. 2.13). Each data point represents the mean power between 50 and 300 MHz of the terminator. In run 1.4, an approximately 3 minute long scan of the terminator took place after approximately 45 minutes of antenna scans. Dates shown are all May 2023 (month-day format, year suppressed) and X-axis ticks are shown at midnight.

1091 **3.2 Measurement of shielding effectiveness and radio**
1092 **frequency interference**

1093 Shielding effectiveness (SE) and radio frequency interference (RFI) were introduced in Secs.
1094 2.1.3 and 2.5.1. Their measurement is introduced here.

1095 Similar to the warning given in the previous section, SE should be measured before
1096 starting a data run or after entering a new frequency range/modifying anything the patch
1097 panel. In addition, it is probably worthwhile to clean the door before beginning a run (the
1098 procedure is described in the text of this section).

1099 Due to reciprocity between antennas, measuring the shielded room's ability to keep ra-
1100 diation *out* is equivalent to measuring the how well the room can keep it *in*. The latter
1101 method is simpler and was performed. The IEEE standard prescribing a very detailed SE
1102 measurement procedure [55] was used as a rough guideline but the simple results presented
1103 here should be viewed as an estimate. While simplified, the results do tend to agree with pre-
1104 dictions of when RFI should become detectable, based on comparisons to the veto antenna
1105 spectrum ⁵.

1106 The SE was measured by placing a Rigol DSG-830 signal generator inside the room,
1107 powered by the filtered 12 VAC Edison outlets inside the room, shown in Fig. 3.12. An
1108 antenna (bicon or Vivaldi) was connected to the signal generator Outside of the room, an
1109 identical antenna (again, bicon or Vivaldi) was connected to a Rigol RSA-5065-TG spectrum

⁵Or at least they did agree before cleaning the door. With the excellent isolation values, no RFI was detected in run 1.4.

1110 analyzer [54]. The spectrum is scanned to find a frequency without local interference which
1111 would confuse the results. The spectrum analyzer was set to attenuate its input to prevent
1112 clipping ⁶, and the signal generator was set to output a sine wave at maximum power.
1113 The power measured on the spectrum analyzer is P_{open} of Eq. 2.35. The door was closed,
1114 attenuation removed and P_{closed} was measured. In some cases the SE was so high a signal
1115 was buried below the noise floor of the spectrum analyzer. These data are shown in red in
1116 Figs. 3.13 3.14 and should be interpreted as an upper limit on isolation, since the true value
1117 is lower (better).

1118 SE can drop off to less than 80dB if the RF gasket around the door is not clean. Previous
1119 day-long test runs detected several hundred signals which originated from RFI emitted from
1120 the PC and several local radio stations. Cleaning was performed by scrubbing the copper
1121 gasket and steel mating surface with red scotch-brite using denatured alcohol as a lubricant
1122 ⁷. After two passes of polishing were complete, a layer of Deoxit D100L liquid was added.
1123 Maintenance cleanings were performed using only Deoxit D5 spray. These signals were not
1124 detected after the gasket was cleaned, which is consistent with calculations of Sec. 2.1.3

1125 The primary RFI in the 50-300 MHz span are from local radio broadcasts. The strongest
1126 of this is around 100 $\mu\text{V}/\text{m}$ as measured in the lab. There are also many lower level peaks
1127 which span the entire frequency range, though are more pronounced in a few frequency bands
1128 (60-75MHz, 130-140MHz, 270-290MHz). These peaks come from the PC which serves as the
1129 spectrum analyzer. Before the door was cleaned, many of these signals were detectable, but

⁶The spectrum analyzer automatically calibrates the displayed spectrum to its internal attenuator and pre-amplifier. If external gain/attenuation are used, they must be accounted for manually.

⁷Alcohol fumes are nasty. Be careful to open the door and use a fan to blow in air from the hall. Take frequent breaks. Don't get dizzy while standing on a chair!



Figure 3.12: Photo showing set up to measure SE of 314. Photo taken using Vivaldi antennas for data shown in Fig. 3.14. The same set up was used with bicon antennas for the 50-300 MHz data shown in Fig. 3.13. Bicon antenna shown in background of photo was not in use for this test.

1130 cleaning the door resulted in a clean spectrum for the 9 day data run. The spectrum from
1131 the veto antenna (discussed in Sec. 2.5.5) is shown in Fig. 3.15

314 Shielding Effectiveness

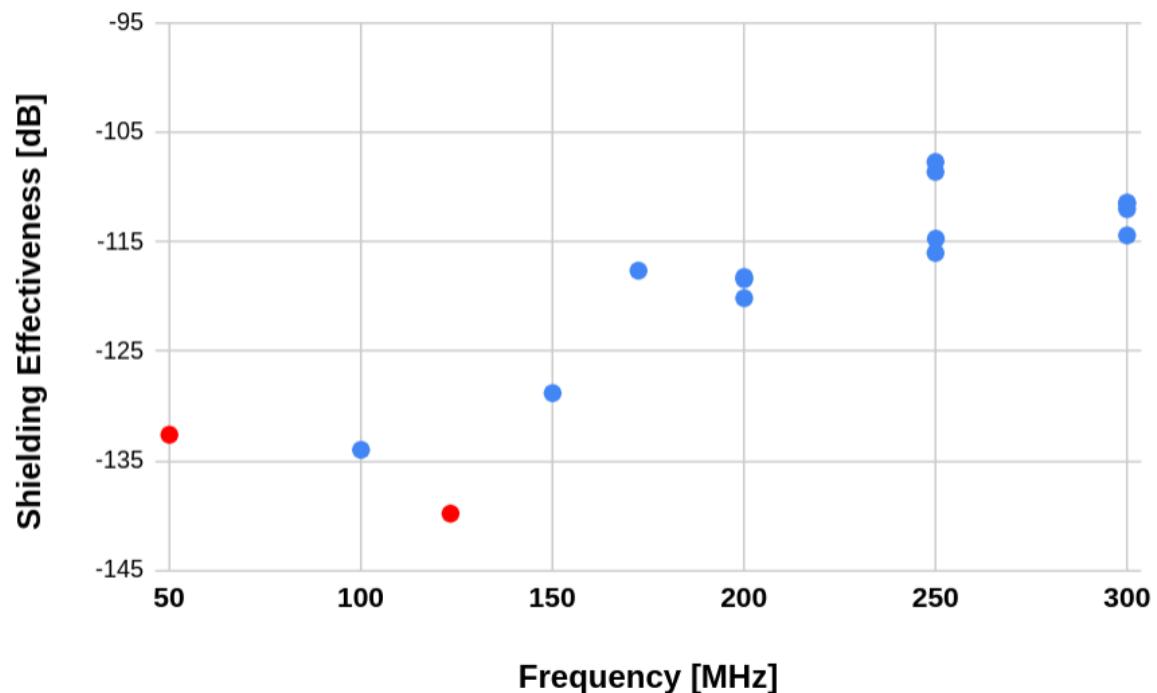


Figure 3.13: Shielding effectiveness of shielded room measured with Bicon antenna. Red points indicate measurements limited by the noise floor of the spectrum analyzer located outside and are an upper limit; true SE is lower (better). Measurement would require a stronger amplifier inside the room. At a few frequencies the antenna inside the room was moved to get a sense of the uncertainty of the measurement.

314 Shielding Effectiveness

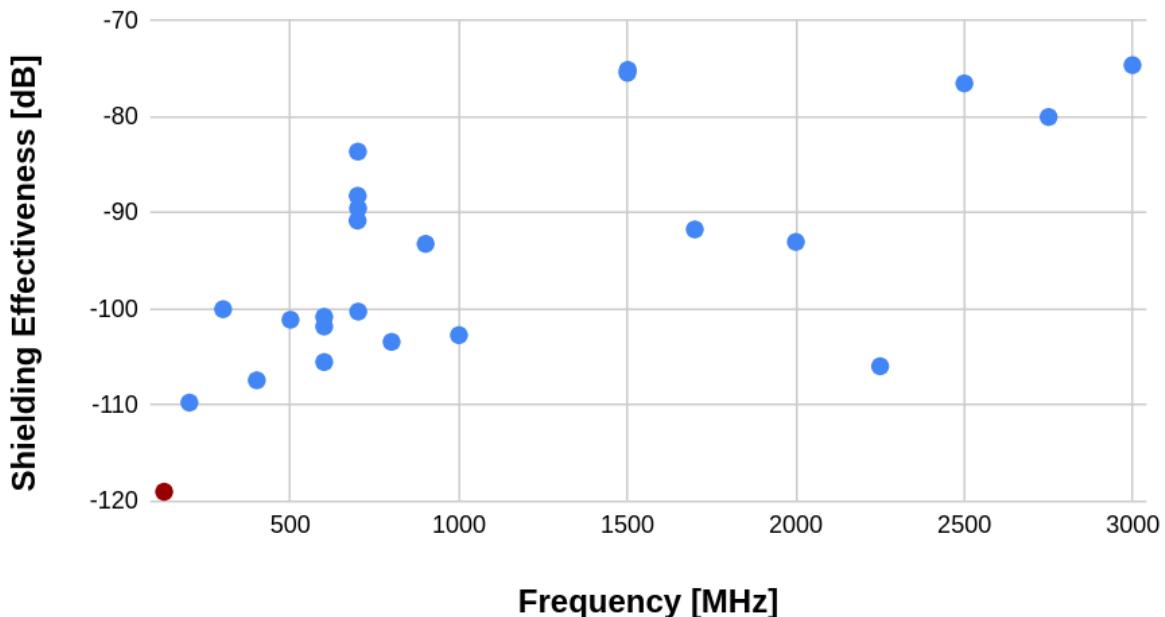


Figure 3.14: Shielding effectiveness of shielded room measured with Vivaldi antenna. Red points indicate measurements limited by the noise floor of the spectrum analyzer located outside and are an upper limit; true SE is lower (better). Measurement would require a stronger amplifier inside the room. At a few frequencies the antenna inside the room was moved to get a sense of the uncertainty of the measurement.

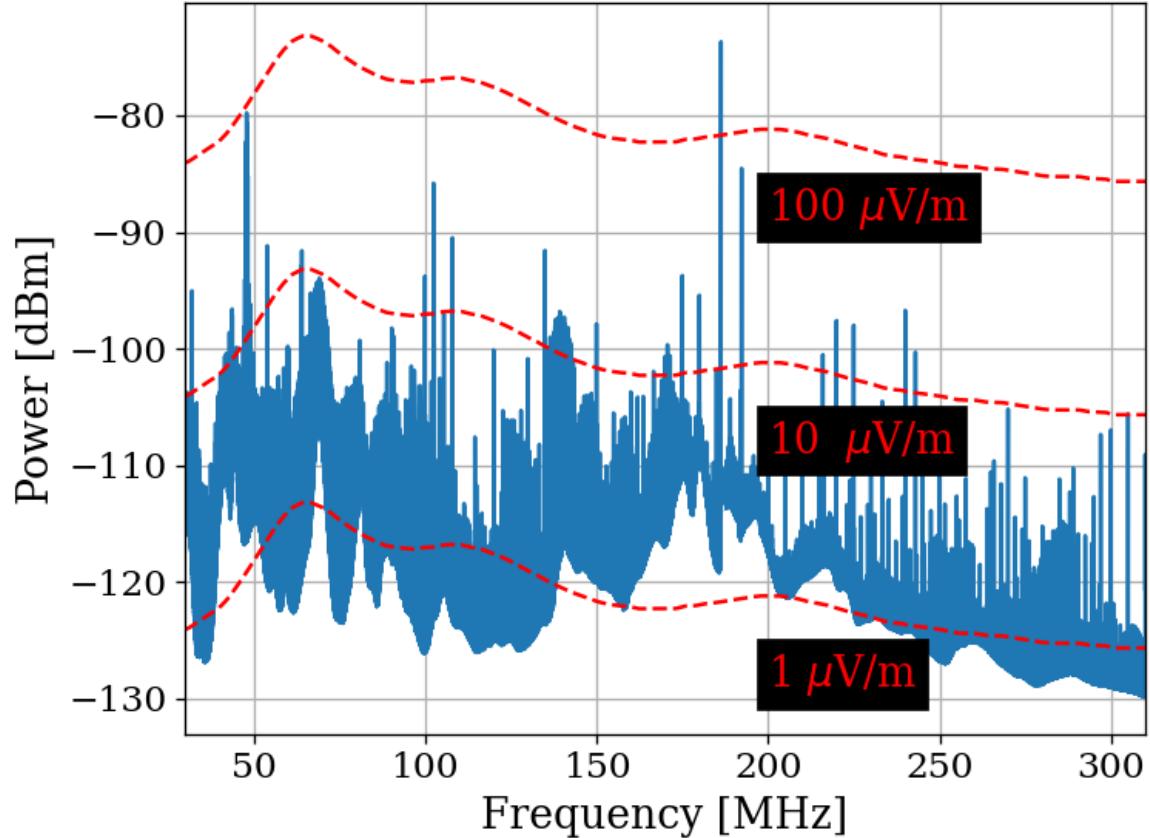


Figure 3.15: Spectrum from veto antenna during 300MHz data run. Spectrum plotted as measured power in dBm. Red curves indicate the equivalent field strength. These curves track the free-space antenna factor with frequency for the COMPOWER AB900 [42] bicon. Two identical AB900 antennas are used, one to search for dark photons in the shielded room, and a second to monitor the local RFI background.

1132 3.3 Spectrum Analyzer Characterization

1133 This section outlines several tests performed to ensure the GPU based real time spectrum
1134 analyzer (RTSA) system performs as expected. Lots of background information that relates
1135 to this system is described in Sec. 2.5.7.

1136 3.3.1 Spectrum analyzer calibration

1137 The first step in determining if the RTSA system is working adequately is to make sure
1138 it is able to correctly identify known, injected signals and noise. This calibration step is
1139 important since the FFT is normalized manually (see Eq. 2.15), and it is rather easy to
1140 mess up. The test signals were generated with a Rigol DSG-830 signal generator. The
1141 time-domain peak-to-peak voltage measured by the RTSA system agreed with the injected
1142 voltage to about 2%. The signal generator's output frequency and power was verified in
1143 both the frequency and time domain using a separate oscilloscope and spectrum analyzer.
1144 This is an important step since cable losses are significant at this level of precision. They
1145 can be taken into account by consulting the data sheet for the cable⁸.

1146 Next, it is important to understand how the system handles noise spectra, since this is
1147 the primary use of the system. This is simply tested by amplifying the Johnson noise of a
1148 terminator (see Sec. 2.1.1.2) and measuring the power spectral density. Knowing the gain
1149 and noise figure of the amp chain allows for calibration. This is the procedure outlined in
1150 detail earlier in this chapter, see Sec. 3.1, specifically Fig. 3.7. These results were confirmed

⁸There are several cheap black SMA cables which are great for testing, but not for calibration. They have more than 5dB of loss above 1GHz. Use a nice cable for these tests, and blow out the SMA connector with compressed air while you are at it.

1151 several times along the way leading up to the data run. Note that due to windowing, other
1152 spectrum analyzers will disagree with the theoretical calculation by a few dB if you don't
1153 include the equivalent noise bandwidth (ENBW) of the window function. Also remember to
1154 take cable effects into account as discussed in the previous paragraph.

1155 3.3.2 Spurious signal performance of the ADC

1156 Any practical spectrum analyzer will have small, coherent, signals which sneak into the the
1157 analog signal path and, after being digitized and Fourier transformed, manifest as spurious
1158 signals or *spurs* (see Sec. 2.1.5.1). These spurs likely come from clock signals within the same
1159 enclosure as the ADC. They can also come from local oscillator bleeding through mixers in
1160 a mixed, high frequency system such as the Rigol RSA 5065TG [54] or the future, upgraded
1161 dark radio system discussed in Ch. 5.

1162 Regardless of their origin, the spurs must be characterized such that if a signal is detected
1163 it can be classified as a spur and rejected as a false positive. Ideally, the system has enough
1164 gain that any spurs are subdominant to the amplified thermal noise, as discussed in Eq. 2.1.
1165 However, since there is a maximum amount of gain that can be added before clipping the
1166 ADC (see Sec. 2.5.4.3, this is not always possible.

1167 Experience with the Teledyne RTSA system has shown that spurs are rather simple to
1168 classify because they don't change much with time. They drift by about 1 kHz, and their
1169 amplitude changes by a few dB, but it still turns out this doesn't much matter.

1170 To classify them, one simply needs to terminate the input of the Teledyne and scan for

1171 a few minutes. It is important (and easy to overlook) that the settings are the same as will
 1172 be used for the run. For example, if an external clock is used, that must be set up the same
 1173 for the test. Length of FFT should be the same. Et cetera. A test scan with the same set
 1174 up as run 1.4 (2²⁴ point FFT, external 1.6 GHz clock, 2x sample skip. See Appendix A for
 1175 details.). It is shown in Fig. 3.16.

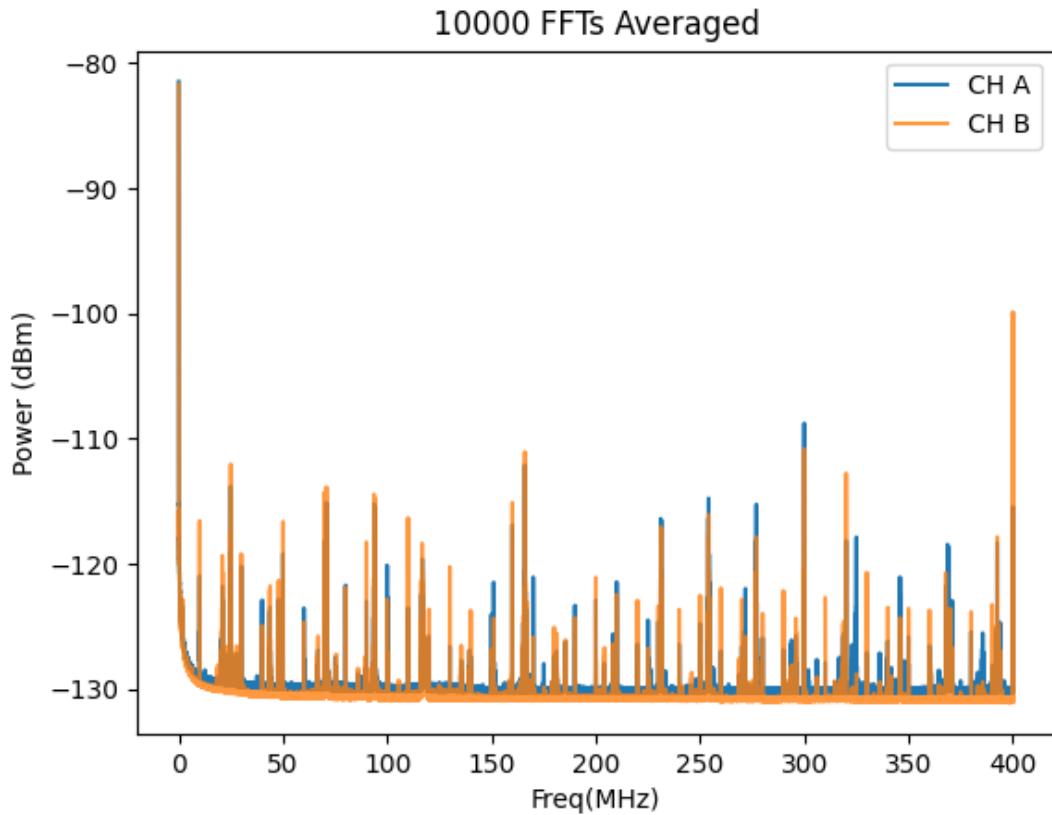


Figure 3.16: Scan of terminated input of Teledyne spectrum analyzer system to measure spur performance. Spectrum acquired with run 1A parameters ($\Delta\nu_{RF} = 47.7$ Hz) with 10,000 averages (about 3.5 minutes). Plot taken directly from test mode of the system (not post-processed in any way). The largest spur is at 299.97 MHz and turns out to be the single significant signal in run 1.4 (see Sec. 4.2.5).

1176 Once a test scan is acquired, comparing the largest spurs (≈ -110 dB) to the expected

1177 output-referred background ([Johnson noise + LNA noise] * Gain \approx -81 dBm/ $\Delta\nu_{\text{RF}}$) allows
1178 us to compute how many averages are required before the spurs are significant. Working
1179 this out is a good test of one's understanding of Sec. 2.2. We find that this spur should be
1180 detectable with 95% confidence after about 4 days, which is exactly what happened in run
1181 1.4 (this false positive candidate is described in Sec. 4.2.5).

1182 3.3.3 ADC clock performance

1183 Any ADC requires a clock which dictates the precise time a sample should be acquired. This
1184 subsection deals with the evaluation of the performance of that clock.

1185 Determining the performance of a clock generally relies on having a better (read: more
1186 expensive) clock known as a *standard* and measuring the clock under test against the stan-
1187 dard. The Valon 5009 RF synthesizer was tested against a borrowed rubidium frequency
1188 standard and was found to be just on the threshold of stable enough to work (this threshold
1189 is explained in Sec. 2.5.7). To ensure we are not limited by clock instability, we synchronize
1190 the sample clock of the ADC to a 10 MHz rubidium frequency standard (Stanford Research
1191 Systems FS725) which is further steered by the one pulse-per-second (pps) signal from a
1192 GPS receiver. This system has medium and long term fractional frequency stability (Allan
1193 deviation) of $\sigma_y(\tau) < 3 \times 10^{-12}$ (where τ is the averaging time) and phase noise of less than
1194 -65 dBc/Hz at offset frequencies > 50 Hz from the carrier [56]. This means that over the
1195 course of a single acquisition, the power contained in a bin will spread to an adjacent bin by
1196 less than 1 part in 10^6 which is more than sufficient for our experiment.

1197 This is a rare example of something in this experiment where we trust the data sheet.
1198 We don't have access to a clock which is better than this, so confirming it's performance is
1199 difficult. Furthermore, even if the specs are an order of magnitude worse than advertised
1200 there is still more than enough stability for the experiment. This statement holds at higher
1201 frequencies; since the Q of the signal is fixed, higher frequencies have wider bins so the
1202 relevant offset frequency at which the phase noise is measured increases with frequency. The
1203 dimensionless Allan deviation is also not affected.

1204 **3.3.4 Real-time data collection efficacy**

1205 As a final test, one would like to know how efficient data scans can be acquired and averaged.
1206 We call 100% efficient system "real time" as described in Sec. 2.5.7. Figure 3.17 illustrates
1207 this efficacy. For run 1.4, it is shown to be close to 99.765%. However, in practice run
1208 logistics such as moving the antenna prevent the full efficiency of the system from being
1209 reached.

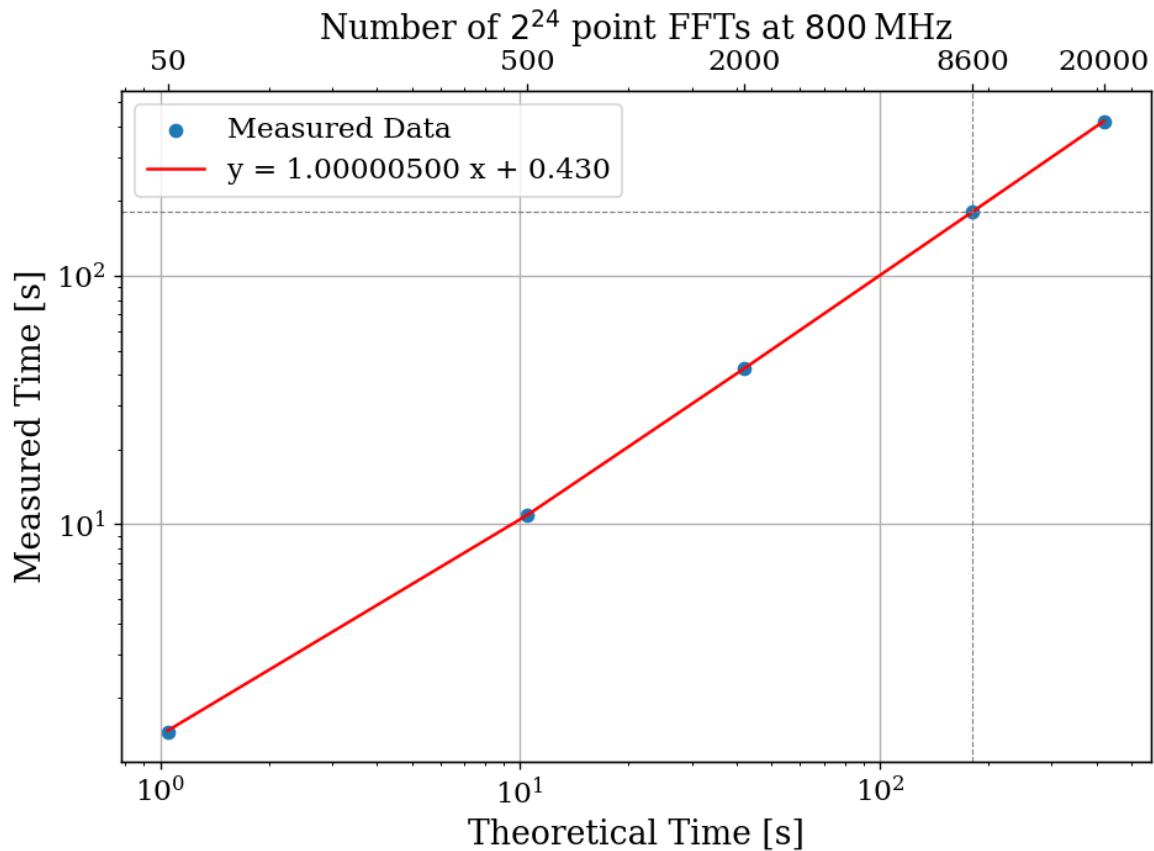


Figure 3.17: Acquisition efficiency for GPU-based real-time spectrum analyzer computed from measured vs theoretical times. The constant offset of 0.43 seconds corresponds to a small set up period when starting an acquisition containing, in the case of run 1A, 8600 FFTs. In the limit of an infinite length acquisition, the system's efficiency can be read off from the slope and is 99.9995%. The more realistic efficiency is a function of number of FFTs per acquisition and for run 1.4 = $180.3551 \text{ [s]} / 180.7782 \text{ [s]} = 99.765\%$.

₁₂₁₀ Chapter 4

₁₂₁₁ Data Acquisition, Data Analysis and
₁₂₁₂ Calibration

1213 Thus far this thesis has focused on building up a background on the dark radio technique
1214 as well as the design and testing of the experiment and it's sub-systems. This chapter outlines
1215 the process of acquiring actual data, analyzing that data to search for a small power excess,
1216 and in the absence of that excess, converting the null result into a exclusion limit on the dark
1217 photon mass/epsilon parameter space. Section 4.4 covers an injection test which confirms
1218 the experiment's ability to integrate all of the machinery developed and detect a realistic,
1219 hardware-injected dark photon proxy-signal. This chapter borrows heavily from my paper,
1220 New Limit on Dark Photon Kinetic Mixing in the 0.2-1.2 μeV Mass Range From the Dark
1221 E-Field Radio Experiment[36], which was published August 2024.

1222 I feel the cleanest way to mentally model the contents of this chapter is by compartmentalizing the first steps outlined above. For this reason, the chapter is organized as follows.
1223
1224 In Sec. 4.1 the data are acquired over a 9-day period. The procedure is outlined, and where
1225 appropriate references to previous sections are provided which are helpful in understanding
1226 how different procedures were developed. This section concluded with the output-referred
1227 power spectral density, S_o which must be searched for a power excess. We call this search
1228 *analysis* and it is covered in Sec. 4.2. The task of analysis is to extract a dark photon signal
1229 from this spectrum if it is present. Otherwise, in its absence, we would like to set a limit on
1230 the amount of output-referred power one would be able to detect *most of the time* were a
1231 narrow signal to be present in this averaged dataset. We quantify the meaning of “most of
1232 the time” by conducting a series of Monte Carlo “pseudo-experiments” on artificial signal-
1233 containing spectra for synthetic signals of varying powers and frequencies. With a limit on
1234 dark photon power extracted, Sec. 4.3 works back through the system to revolve a limit

1235 on epsilon above which we have some confidence we would have observed a signal. This
1236 produces the exclusion limit, which is ultimately the deliverable of this experiment. This
1237 limit is shown in Sec. 4.5.

1238 4.1 Data Acquisition

1239 Data were collected during a 9-day run from May 10 to May 19, 2023. This data run is
1240 referred to as run 1.4 throughout this thesis, and its details are shown in Table ??.

Specification	Value
Analysis span	50-300 MHz
Antenna	AB-900A biconical antenna
RTSA	Teledyne GPU system (Sec. 2.5.7)
Frequency resolution ($\Delta\nu_{RF}$)	47.7 Hz
Length of record	2^{24}
Sample rate	800 MHz
Window type	None (flattop)
Acquisition time per spectrum	20.96 ms
Run start time	2023-05-10 11:29:48
Total time of run	8 days, 21.13 hours
Efficacy (Time spent scanning antenna/total time)	92.03%

Table 4.1: Run 1.4 Details. Many specifications are related and can be computed from each other but are listed for reference. Efficacy differs from calculation in Fig. 3.17 mainly because of switching to terminator and brief daily pauses to move the antenna.

1241 Each day was subdivided into data-collection (23 hours 15 minutes) and setup (45 min-
1242 utes) periods. The setup period includes moving the antenna, changing a 12 V battery for
1243 the LNA (Sec. 2.5.6), file management and documentation. In order to reduce the data
1244 rate and storage requirements, all data were pre-averaged into 3-minute chunks and then

1245 saved. This pre-averaging is shown in Fig. 2.21. Additionally, an RF switch (see Fig. 2.13)
1246 is actuated for a 3 minute scan for every 15 antenna scans in order to monitor the status
1247 of the amplifier chain. For the data analysis, all 9 days of data were averaged together to
1248 create a single spectrum S_o (Fig. 4.1). If candidates are found, their time dependence can
1249 be observed by looking at the 3-minute pre-averages. All further analysis is performed on
1250 the full 9-day S_o spectrum and is described below (Sec. 4.2).

1251 4.1.1 Raw data, S_o

1252 All 9 days of pre-averaged data from the run are averaged together. The stability of the
1253 sample clock (Sec. 3.3.3) ensures that this is a simple process; frequency bins ($\Delta\nu_{RF} =$
1254 47.7 Hz) corresponding to a given frequency are added and normalized to the total number
1255 of pre-averaged spectra. This process produces the raw spectrum, S_o (Fig. 4.1), on which
1256 we will perform a search for power excess.

1257 Inspection of S_o reveals small power variations over spans of tens of kHz. The origin of
1258 these variations is explored in Sec. 2.3, but it is summarized here. Given an antenna in a
1259 cavity in thermal equilibrium with the input of an LNA, whose input is assumed to be real
1260 and matched, one would expect an output PSD which is constant with respect to frequency
1261 up to small variations in LNA gain. The theory for this is outlined by Dicke in [57]. These
1262 variations are not noise; for a given antenna position we repeatedly measure the same shape
1263 (though the noise riding on these variations *is* random). The origin of the observed small
1264 variations lies in the effective temperature difference between the room and LNA causing a

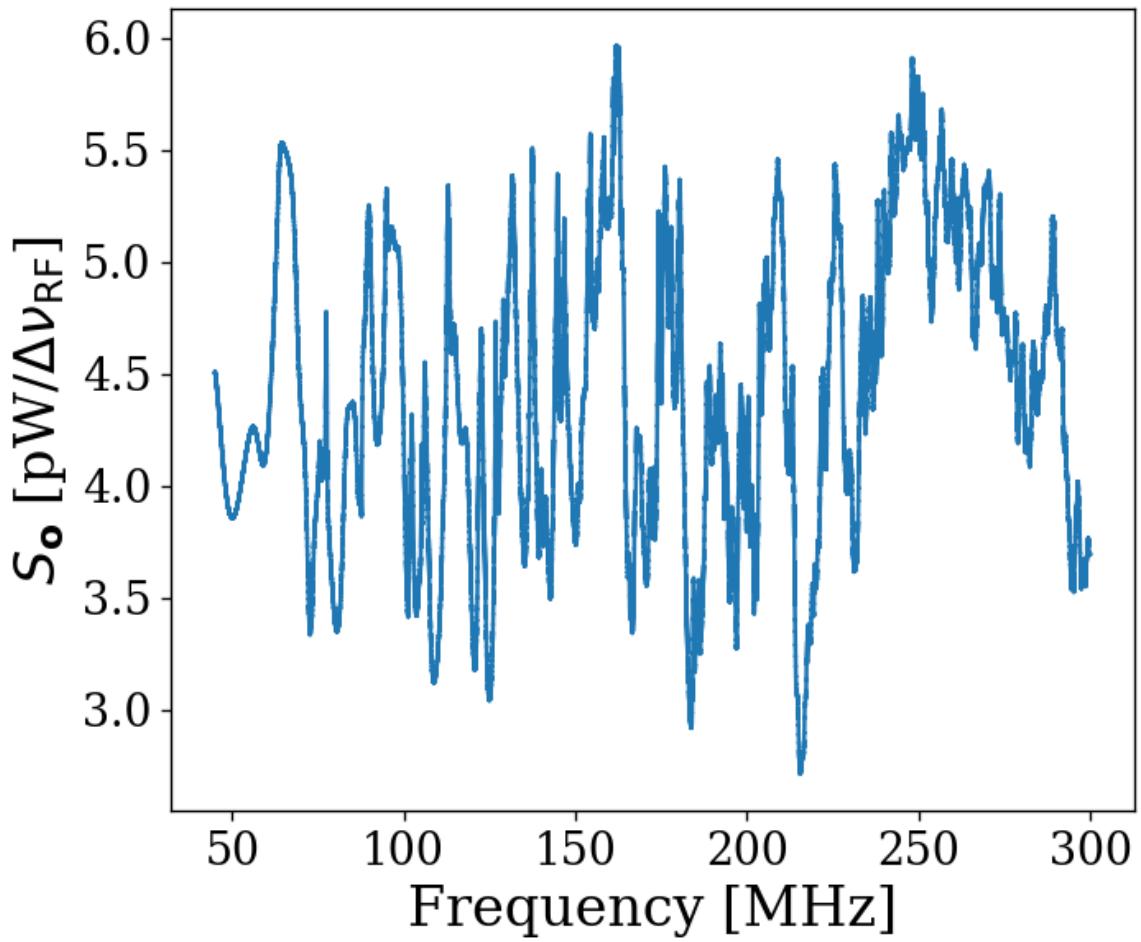


Figure 4.1: Run 1A averaged, output-referred antenna power spectrum S_o . Data were taken over a 9 day period at 9 antenna positions. The narrow variations are mainly due to the effective temperature difference between the room and LNA (Sec. 2.3), though there is a small contribution due to amplifier gain and noise temperature variations (Sec. 4.1.1). The variations seen here are not noise; their shape is repeatable for a given antenna position. The noise on this background is not visible at this level of zoom, but can be seen in Fig. 4.4, which shows a zoomed-in view of the spectrum at 240 MHz. The noise can also be seen nicely in Fig. 4.17



Figure 4.2: Run 1A averaged, output-referred, antenna and terminator power spectra

1265 net power flow from the antenna into the LNA. This effective temperature difference partially
1266 excites modes of the antenna/cavity system, causing the observed variations. We suspect this
1267 effect originates from a small reactive component of the LNA's input causing the electronic
1268 cooling described originally by Radeka [16]. This effect can be eliminated by adding an

1269 isolator between the antenna and LNA [29, 58] though for our experiment, it is impractical
1270 to get an isolator that covers such a wide band at relatively low frequency. Furthermore, the
1271 relatively wide (tens of kHz) variations can be handled by fitting to them which we discuss
1272 in Sec. 4.1.

1273 4.2 Data Analysis

1274 At this point, we have compiled a single, averaged, output-referred power spectrum, S_o
1275 (Fig. 4.1). The task of *analysis* is to extract a dark photon signal from this spectrum if
1276 it is present. Otherwise, in its absence, we would like to set a limit on the amount of
1277 output-referred power one would be able to detect *most of the time* were a narrow signal
1278 to be present in this averaged dataset. We quantify the meaning of “most of the time”
1279 by conducting a series of Monte Carlo “pseudo-experiments” on artificial signal-containing
1280 spectra for synthetic signals of varying powers and frequencies. The following subsections
1281 are organized as follows:

1282 4.2.1: Fit S_o to extract an estimate of the background B (which we call \hat{B}) whose origin was
1283 discussed in Sec. 4.1.1. See Fig. 4.4.

1284 4.2.2 Divide the spectrum by \hat{B} to generate the *normalized spectrum*, which very nearly
1285 follows a Gaussian distribution. Discuss statistics of the normalized spectrum and
1286 choose a global significance level and its associated *significance threshold*. See Fig. 4.5.
1287 Also note this was derived in Sec. 2.2.

1288 4.2.3 Apply a matched filter to the normalized spectrum and establish a significance thresh-

1289 old on its output using the same method defined in the previous section. See Fig. 4.6.

1290 The previous three steps comprise our *detection algorithm* which is shown in Fig. 4.3.

1291 4.2.4 Perform a Monte Carlo analysis to simulate the required power of a signal that can be

1292 detected above the significance threshold 95% of the time. We use this to recover a

1293 95% exclusion limit on the output referred power spectrum.

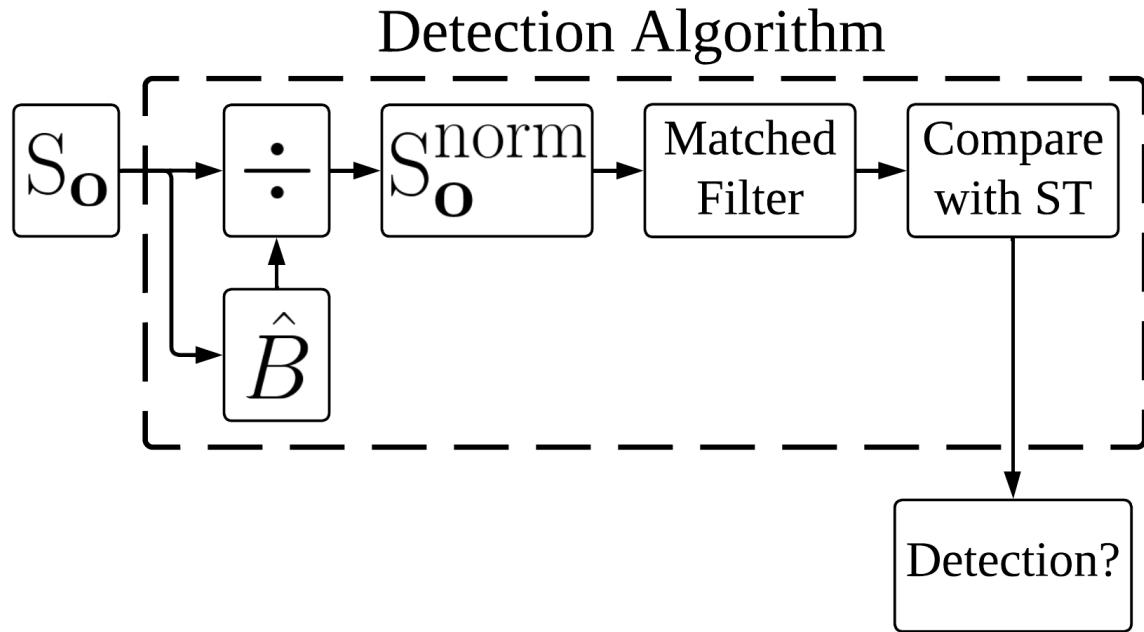


Figure 4.3: Flow chart outlining the logic of signal processing in the detection algorithm of sections 4.2.1 through 4.2.3. \hat{B} is the smoothed fit to S_o generated by low pass filtering. The output, *Detection?*, is a Boolean array which signifies a detection or lack thereof at each frequency bin. We detect a candidate if a bin contains more power than a significance threshold (ST) (Sec. 4.2.2).

1294 In Sec. 4.3 we convert this threshold on S_o into an actual limit on ϵ .

1295 Throughout the figures of this section we will follow a relatively large (40 fW, output-
1296 referred) software-injected, synthetic dark photon signal at 240 MHz to illustrate what a
1297 candidate would look like as it passes through the analysis procedure. This signal is added
1298 to S_o . For clarity, a single interfering candidate had been removed. This is discussed in
1299 Sec. 4.2.5.

1300 As a final note, at this point if you are unfamiliar with the data analysis, it would be
1301 advantageous to spend a few minutes looking through Figs. 4.17 through 4.19 to get a sense
1302 of the task at hand. It can be disorienting dealing with a spectrum that contains 8 million
1303 frequency points, zooming in on a signal can be very informative.

1304 **4.2.1 Fit background, $\hat{B}(\nu)$** ¹

1305 As shown in Fig. 4.1, the measured power spectrum looks like flat thermal noise *multiplied* by
1306 some frequency dependent background, $B(\nu)$ ². However, for this section we will not concern
1307 ourselves with the origin of B or any details of the experiment aside from two assumptions:

- 1308 1. The measured background is the product of a normally distributed spectrum and some
1309 background. This is enforced by the central limit theorem due to the large number of
1310 averaged spectra, independent of any experimental specifics.
- 1311 2. The line shape of the signal is known and the width of this signal is much narrower
1312 than the width of features on the background, viz. $\Delta\nu_{DP} \ll \Delta\nu_B$

¹Code for this section can be found at: https://github.com/josephmlev/darkRadio/tree/master/daqAnalysisAndExperiments/run1p4/run1p4_analysis/analysisClass.ipynb

²When I was first working on this, I used $H(\nu)$ to represent this background. This naming convention persists in the analysis code.

1313 The first assumption (1) implies that if we were able to extract the background, dividing
1314 S_o by this extracted background would yield a *dimensionless*, normally distributed power
1315 spectral density on which to perform a search for a dimensionless signal. The second as-
1316 sumption (2) will be critical in both performing the fit to the background (this section), and
1317 performing matched filtering (Sec. 4.2.3).

1318 In light of these assumptions, we attempt to fit for the background power spectrum. Since
1319 this fit estimates B , we use the symbol \hat{B} to refer to it. As discussed in [17], a particularly
1320 effective fitting technique that can discriminate between the wide bumps of S_o and a narrow
1321 signal is to use a low pass filter. We implement this filter in two stages:

- 1322 1. A median pre-filter (51 bins or about 2.4 kHz wide) attenuates any very narrow, very
1323 large excursions which would interfere with any following filters, causing them to
1324 “ring”³
- 1325 2. A 6th-order Butterworth low pass filter (corner frequency of 210 bins or 10 kHz)

1326 These bin widths/frequencies should be interpreted as the width of spectral features on
1327 S_o that are attenuated and will, therefore, not show up in the background fit. A narrow
1328 zoom of this fit with a synthetic signal is shown in orange in Fig. 4.4.

³I recommend that this pre-filtering step is omitted if the spectrum analyzer in use doesn’t produce large spurs. It is the slowest part of the analysis chain. It also causes the filtered spectrum to deviate slightly from Gaussian.

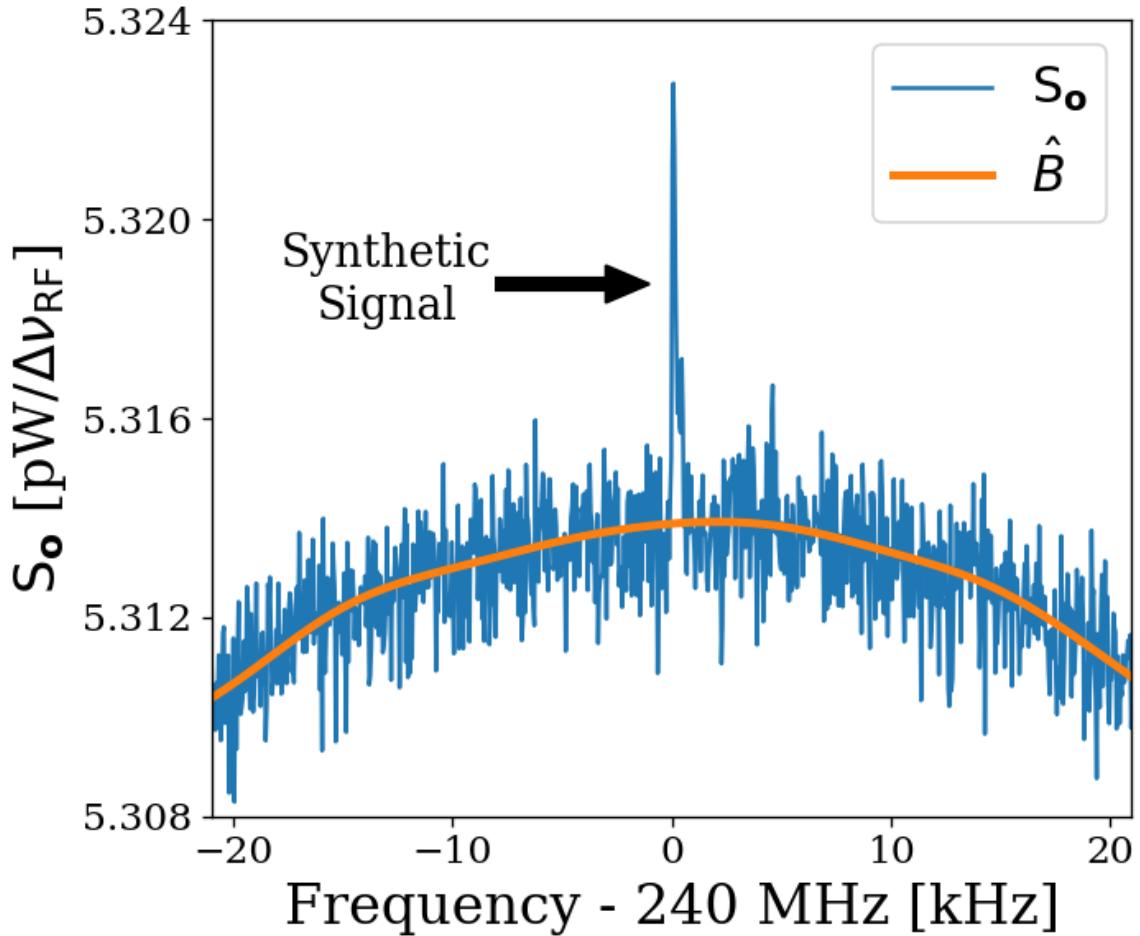


Figure 4.4: Fitting background \hat{B} in the presence of a synthetic signal injected at 240 MHz. Starting from the averaged, output-referred spectrum (S_o), we fit the background using a series of filters (section 4.2.1, and Fig. 4.3). This figure is a highly zoomed in view (240 MHz \pm 20 kHz) in order to show the noisy Rayleigh signal shape.

₁₃₂₉ **4.2.2 Normalized spectrum, S_o^{norm}**

₁₃₃₀ Once we have a fit to the background, \hat{B} , division of S_o by this fit yields a dimensionless,

₁₃₃₁ Gaussian distributed spectrum

$$S_o^{\text{norm}} \equiv \frac{S_o}{\hat{B}}. \quad (4.1)$$

1332 As discussed in Sec. 2.1.1.3, this normalized spectrum (Fig. 4.5) should have a mean $\mu_{\text{norm}} = 1$
 1333 and a standard deviation given by the Dicke radiometer equation $\sigma_{\text{norm}} = (\tau \Delta\nu_{\text{RF}})^{-1/2}$ where
 1334 τ is the total integration time (≈ 9 days) and $\Delta\nu_{\text{RF}}$ is the width of a bin (47.7 Hz). This
 1335 works out to a predicted σ_{norm} of 1.727×10^{-4} . μ_{norm} and σ_{norm} calculated from the data
 1336 are $1 - 1.2 \times 10^{-5}$ and 1.741×10^{-4} respectively, which agree with the predicted values
 1337 to better than 1%. Knowing the statistics of the background allow us to set a threshold
 1338 above which we have some confidence that a candidate is not a random fluctuation. This
 1339 significance threshold was derived in Sec. 2.2.1. As a reminder, the probability that all N
 1340 bins are less than z standard deviations, $z\sigma$, for a standard normal distribution is given
 1341 by $\left\{ \frac{1}{2} [1 + \text{erf}(z/\sqrt{2})] \right\}^N$, where $\text{erf}(z)$ is the standard error function and z is real. A 5%
 1342 ST for 5.2×10^6 bins (our 50-300 MHz analysis span) works out to 5.6σ . This is shown in
 1343 Fig. 4.5.

1344 It is possible to set a simple limit using this significance threshold on the normalized
 1345 spectrum, which was our method in [4]. However, knowing the line shape of the dark photon
 1346 signal provides additional information that improves sensitivity (up to a factor of ≈ 2) at
 1347 the higher frequency end of the spectrum, as shown in Fig. 4.8.

1348 4.2.3 Signal-matched filter

1349 As discussed in 4.2.2, one simple method to set a limit is to look for single-bin excursions
 1350 above some threshold. However, galactic dynamics impart a dark photon candidate with
 1351 a Rayleigh-distributed, spectral signature, which has a dimensionless width $Q_{\text{DP}} \approx 10^6$

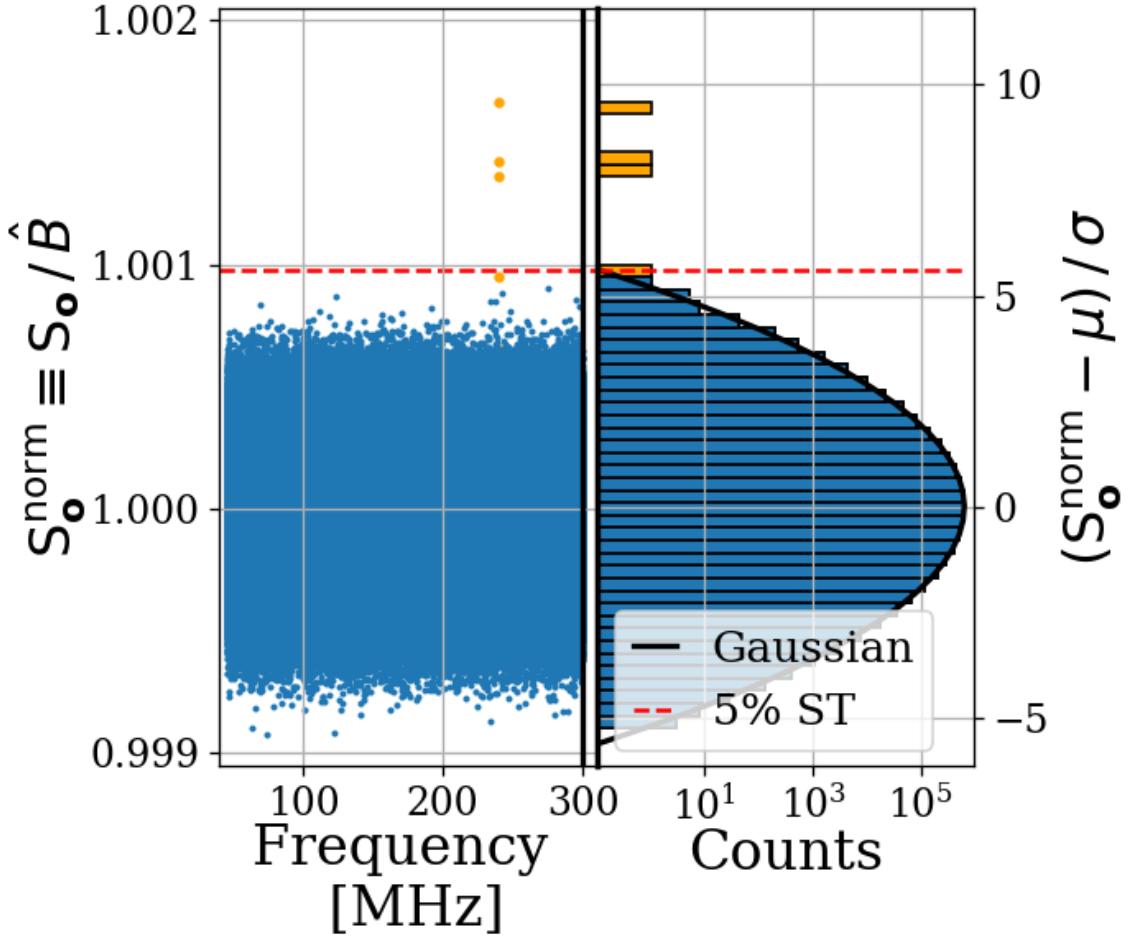


Figure 4.5: Dividing S_o by \hat{B} yields a dimensionless, normally distributed power spectrum that we define as S_o^{norm} . We show S_o^{norm} in two ways: a normalized power/frequency spectrum (*left*) and rescaled into Z-score units and collapsed into a histogram (*right*). The histogram shows power excess and Gaussian fit, but frequency information is lost. We compute a 5% significance threshold ST (*dashed red*), above which we will detect a candidate by chance 5% of the time. Bins adjacent to the 240MHz synthetic signal show up in orange on both plots. A single interfering signal has been removed for clarity. We discuss this further in Sec. 4.2.5

1352 [30]. This means that the expected width of a candidate signal over our analysis span (50-
 1353 300 MHz) varies between 50-300 Hz. We set $\Delta\nu_{\text{RF}} = 47.7 \text{ Hz}$ to maximize SNR for the lowest
 1354 expected signal width. However, this divides signal power between adjacent bins, an effect

1355 that becomes more pronounced at higher frequencies, leading to a decrease in sensitivity.
1356 By using a signal processing technique known as *signal-matched filtering* [59, 60], we restore
1357 some of the sensitivity lost due to the splitting of signal between the fixed-width frequency
1358 bins of the FFT. A similar “optimal weighting” procedure has been well established in axion
1359 haloscope experiments, notably by [17, 61]. Below, the signal-matched filter is referred to
1360 simply as a matched filter.

1361 For a known signal shape, the detection technique which optimizes SNR is the matched
1362 filter. This is implemented on the normalized power spectrum using the Rayleigh spectral
1363 line shape of [30] as a template. Since we have a constant $\Delta\nu_{\text{RF}}$ and expect the width of
1364 the signal to vary across our span, we must calculate several templates of varying width to
1365 match the expected line shape. Every 10% of fractional frequency change, a new template
1366 is generated and used to search that small sub-span of the normalized spectrum, each of
1367 which is also normally distributed though with its own standard deviation. This results in
1368 20 subspans (50-55 MHz, 55-60.5 MHz etc.). The normalized spectra of all 20 subspans and
1369 the histogram of the 227-250 MHz subspan are shown in Fig. 4.6.

1370 As the width of the templates increase, the standard deviation of the output decreases,
1371 resulting in the $\nu^{-1/2}$ shape of the 5% significance threshold shown in Fig. 4.6. It should be
1372 noted that since the total number of bins remains 5.2 million, the 5% significance threshold
1373 still corresponds to 5.6σ ; the shaping in Fig. 4.6 is due to the variation in σ for different
1374 templates, not a change in the $z = 5.6$ pre-factor.

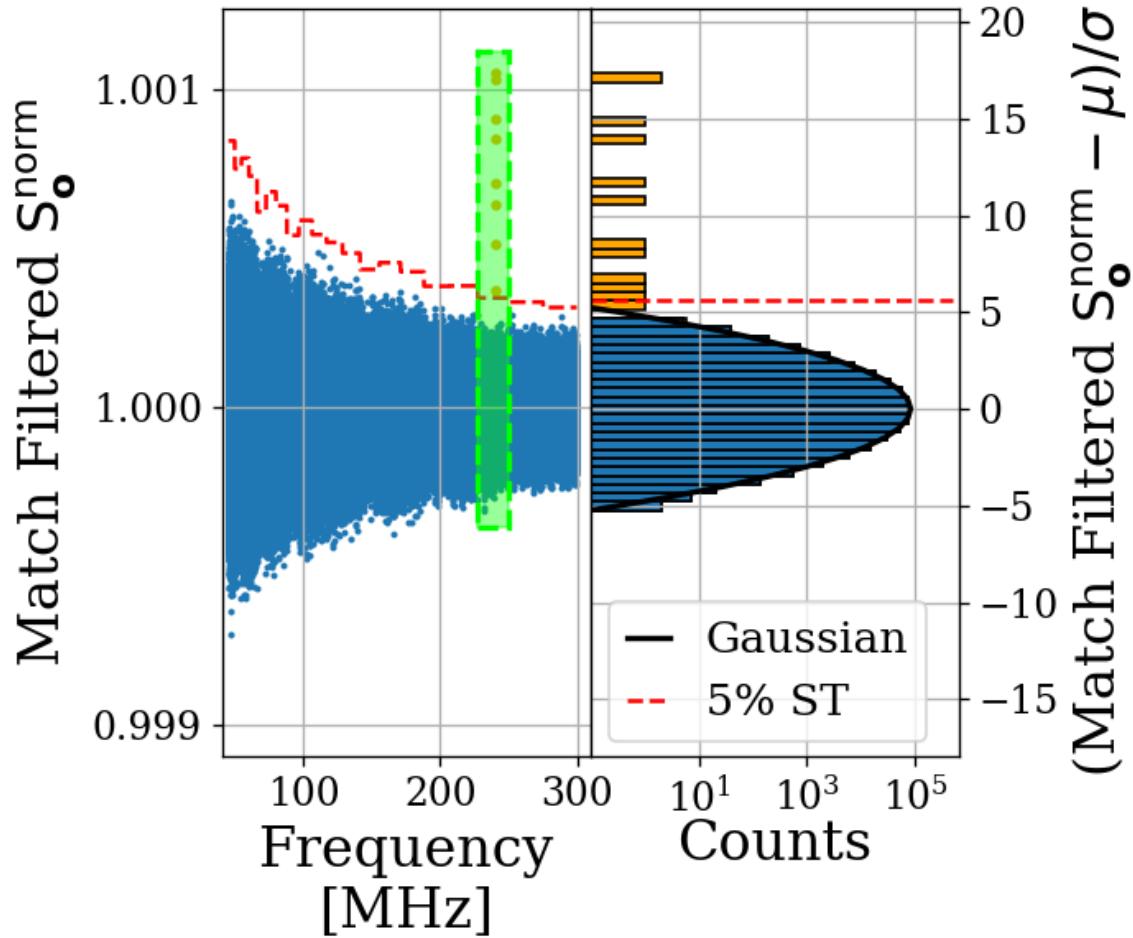


Figure 4.6: S_o^{norm} after it has been passed through a matched filter. The template varies in width throughout the frequency span resulting in 20 subspans, each with a constant 5% significance threshold ST (*dashed red*). Histogram only includes 227-250 MHz subspan (enclosed in the green box). The signal-to-threshold ratio of the synthetic signal (orange) improves by a factor of about 1.8 as compared to Fig. 4.5 without a matched filter. The frequency dependence of this effect is shown in Fig. 4.8. A single interfering signal has been removed for clarity.

¹³⁷⁵ 4.2.4 Monte carlo: pseudo experiments

¹³⁷⁶ The previous three sub-sections outline the procedure for detecting the presence of a signal

¹³⁷⁷ of known spectral line shape embedded in wide-band noise. We refer to this procedure as a

1378 *detection algorithm* (see Fig. 4.3) which we now calibrate through a Monte Carlo method.

1379 A synthetic spectrum is constructed by multiplying some B by randomly generated Gaussian white noise characterized by μ_{norm} and σ_{norm} , as discussed in section 4.2.2. A signal of known, total integrated, output-referred power and frequency, $\lambda(p, \nu)$, can now be added to this spectrum to create a test spectrum which can be passed through the detection algorithm. The frequencies of the synthetic signals are evenly spaced (approximately every 10 MHz). However because the signal spans a limited number of bins (one to six), the shape of the discretized signal is very sensitive to where its peak lands relative to the bins. To compensate for the fact we don't know where a dark photon's peak would land relative to the frequency bins, the frequency of the synthetic signal is randomly jittered by $\pm\Delta\nu_{\text{RF}}/2$, which is drawn from a uniform probability distribution at each iteration of the Monte Carlo. By repeating this with randomly generated Gaussian noise and various synthetic signals (including a small jittering of signal frequency outlined above), statistics are built up about how much total integrated power is required to detect a signal as a function of frequency *most of the time*. We quantify this as the statistical power of the detection algorithm and denote it $100\% - Y = 95\%$ following the standard convention of hypothesis testing.

1394 This Monte Carlo allows us to treat the detection algorithm as a black box which can be calibrated by passing it a known input (a synthetic S_0 containing a synthetic signal, both software-generated) and looking at its output; a Boolean array of frequency bins representing signal detection. These data along with a simple linear fit to the 95% confidence interval are shown in Fig. 4.7.

1399 The output MC_{lim} , normalized to standard deviation can simply be converted back to the

	Only Noise	Noise + Signal
Detection	X	100% – Y
No Detection	100% – X	Y

Table 4.2: Threshold parameters that are part of the detection algorithm and Monte Carlo. X is the significance of the analysis. It is a parameter passed to the detection algorithm which specifies the significance threshold. The quantity 100% – Y is the statistical power of the analysis. It is a parameter in the MC, which specifies a threshold on signal power where a given signal is detected in 100% – Y of the MC iterations. We choose both X and Y = 5%.

1400 limit on the total output-referred power contained in injected signals which can be detected
 1401 95% of the time,

$$P_o^{\lim} = MC_{\lim} \sigma_{\text{norm}} \hat{B}. \quad (4.2)$$

1402 MC_{\lim} is then a measure of how much the analysis' efficacy is reduced compared to an
 1403 ideal analysis where a signal with 5.6σ of power is detected at 95% significance half of the
 1404 time (remember, it's still a random process. Refer back to Sec 2.2).

1405 P_o^{\lim} is shown in Fig. 4.8 in blue. Also shown in Fig. 4.8 is a limit that does not include
 1406 any matched filtering (orange) to highlight the frequency dependent improvement of the
 1407 matched filter; this limit is only for illustration and not used in the following sections.

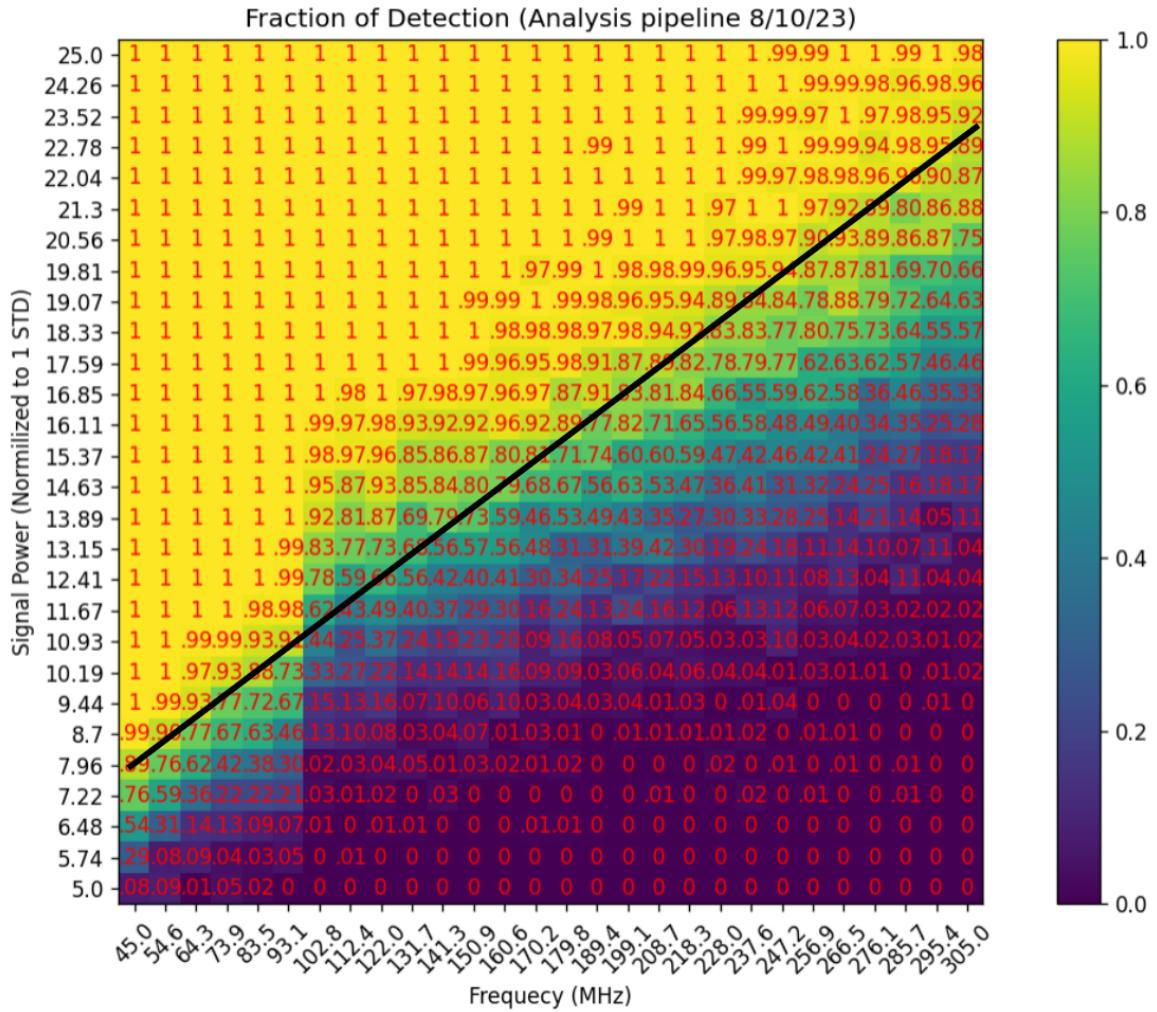


Figure 4.7: Result of Monte Carlo pseudo experiments on signal detection. Color/red numbering show fraction of detection for each frequency/injected signal power of the 784 combinations tested. Black line inserted by eye and gives an approximate fit to the frequency vs. injected signal power which results in a detection 95% of the time. Approximate form of back line is $y = 0.0686x + 2.411$

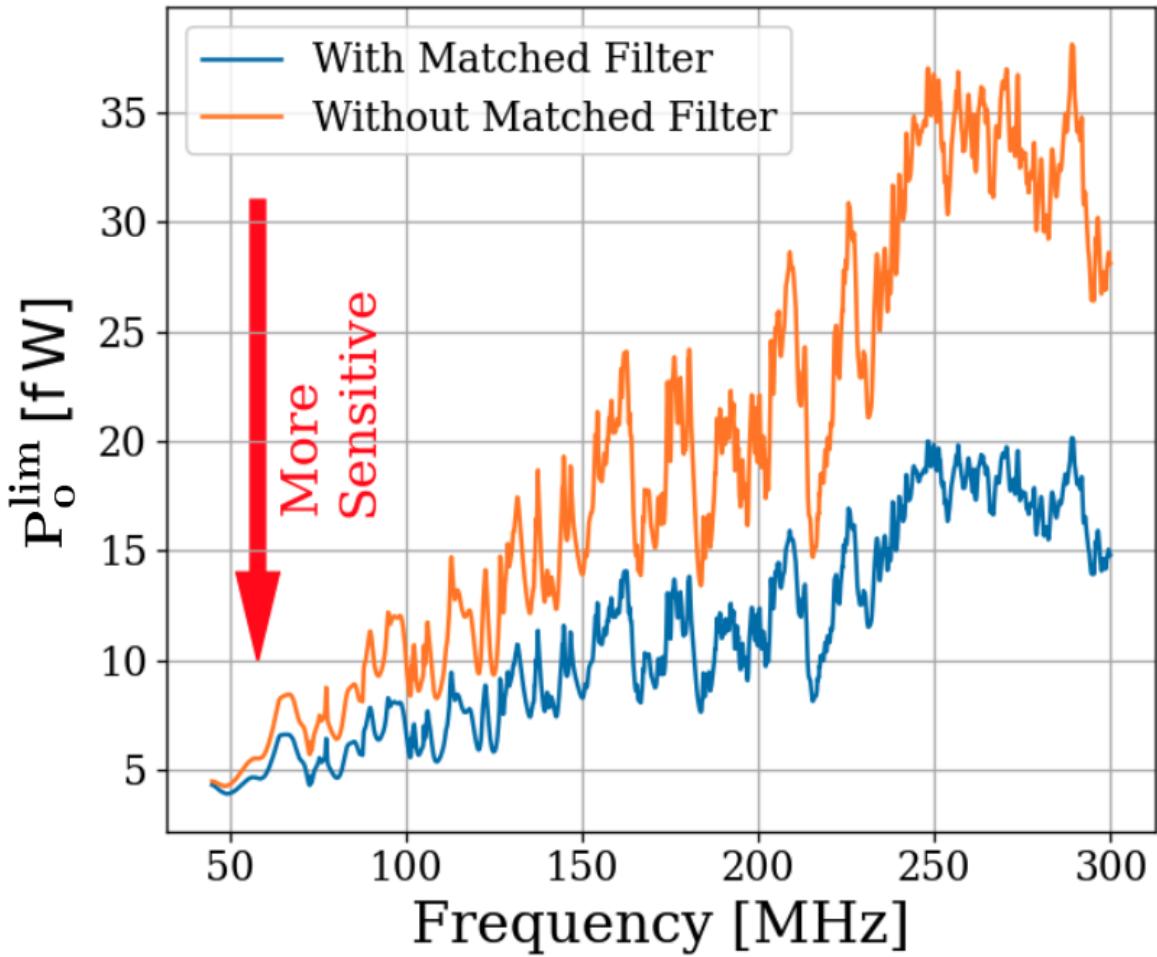


Figure 4.8: Limit on output-referred total integrated signal power, P_o^{\lim} . Limits computed with (blue) and without (orange) a matched filter (Sec. 4.2.3). The limits are similar at lower frequency but the matched filter improves sensitivity at higher frequencies where the signal power is split among more bins. The blue curve is used in the following sections.

1408 **4.2.5 Rejection of a single candidate**

1409 Passing S_o through the detection algorithm diagrammed in Fig. 4.3 yields a single candidate
1410 at 299.97 MHz which is approximately 1 kHz wide. This candidate first became detectable
1411 above the noise after about 4 days of averaging, indicating it is just on the threshold of what
1412 we are able to detect. Four factors cause us to conclude the candidate is an interfering signal
1413 originating from within the PC or ADC, allowing us to remove it:

- 1414 • The candidate is present not only in the main spectrum, but also the veto and termi-
1415 nator spectrum.
- 1416 • Inspection of the time evolution of this signal shows a narrow signal (about two bins,
1417 or 100 Hz wide) which seems to wander in frequency periodically over the course of a
1418 day and therefore with temperature. This is expected behavior for a quartz oscillator.
- 1419 • Reducing the gain of the system causes the SNR of the candidate to *increase*, indicating
1420 it enters the signal path after the gain stages.
- 1421 • Changing the clock rate causes the frequency of the candidate to change.

1422 The limit set in this section is referred to the output of the amplifier chain. A single
1423 significant candidate was found, but the method of ruling it out was outlined above. The
1424 topic of the next section will be to work back through the amp chain, to an E-field limit in
1425 the cavity and ultimately to a limit on ϵ .

1426 4.3 Calibration

1427 In this section we describe the calibration of our experiment and estimate our uncertainty.

1428 The previous section concluded with a limit on the output-referred power P_o^{\lim} (Fig. 4.8),

1429 which we now must convert into a frequency dependent limit on ϵ .

1430 We begin by inverting Eq. 1.7,

$$\epsilon(\nu) < \sqrt{\frac{|\mathbf{E}_{\text{ant}}^{\lim}|^2 \varepsilon_0}{2 \rho_{DM}}}, \quad (4.3)$$

1431 where the *lim* superscript indicates a limit, below which a detectable electric field may be

1432 hiding. The $<$ should be taken to mean that in setting a limit on $|\mathbf{E}_{\text{ant}}^{\lim}|$, ϵ is constrained to

1433 be less than the right hand side (if it exists at all).

1434 The first step of calibration is to convert from output-referred power to *antenna referred*

1435 *power*. This represents the signal power presented to the LNA by the antenna via a matched

1436 transmission line and is given by

$$P_{\text{ant}}(\nu) = \frac{P_o}{G} - T_{\text{amp}} k_B \Delta\nu_{\text{RF}}, \quad (4.4)$$

1437 where G and T_{amp} are the frequency-dependent amplifier gain and noise temperature (74–75 dB

1438 and 100–120 K respectively, measured via the Y-factor method, see Sec. 3.1.1) and k_B is

1439 Boltzmann’s constant.

1440 Ultimately, the exclusion limit is set by fluctuations on this baseline described by

$$\begin{aligned}
P_{\text{ant}}^{\lim}(\nu) &= \frac{P_{\text{o}}^{\lim}}{G} - \left(\frac{2}{n} \right)^{1/2} T_{\text{amp}} k_B \Delta\nu_{\text{RF}} \\
&= \frac{P_{\text{o}}^{\lim}}{G} - \left(\frac{2 \Delta\nu_{\text{RF}}}{\tau} \right)^{1/2} T_{\text{amp}} k_B,
\end{aligned} \tag{4.5}$$

1441 where the *lim* superscript indicates an exclusion limit, n is the total number of spectra
 1442 averaged together, and τ is the total integration time. In the second line we have used
 1443 $n = \Delta\nu_{\text{RF}} \tau$. In practice, the LNA correction is small; the first term divided by the second
 1444 varies with frequency between 7 and 50. The $\tau^{-1/2}$ dependence of P_{o}^{\lim} is implicit because it
 1445 was calculated from S_{o} which is itself an averaged spectrum. As mentioned above, this $\tau^{-1/2}$
 1446 dependence implies that the limit on ϵ scales as $\tau^{-1/4}$.

1447 In the remainder of this section we explore the relationship between P_{ant}^{\lim} and $|\mathbf{E}_{\text{ant}}^{\lim}|$
 1448 allowing us to use our experimental data to set a constraining limit on ϵ by employing
 1449 Eq. 4.3.

1450 4.3.1 Average effective aperture, $\langle A_e(\nu) \rangle$

1451 An antenna's effective aperture, $A_e [\text{m}^2]$, represents the effective area that it has to collect
 1452 power density or irradiance [W/m^2] from an incident Poynting vector. It was defined in
 1453 Eq. 2.9.

1454 A_e is useful for an antenna in free space, however some modifications must be made to
 1455 construct an analogous quantity for an antenna in a cavity.

1456 The first modification is to average over many configurations of the system. The back-
 1457 ground for this is given in Sec. 2.4. As discussed, we denote this averaging with $\langle \rangle$ so that the

1458 average, effective aperture is denoted $\langle A_e \rangle$. It is interesting to note that by averaging over
 1459 configurations (namely antenna direction), $\langle A_e \rangle$ simplifies since $\langle D(\Omega) \rangle = 1$ by construction
 1460 [20].

1461 The second modification is to introduce a resonant enhancement factor which corresponds
 1462 to the system's tendency to "ring up" in the same way any resonator will. We refer to
 1463 this as *composite Q* and represent it as \tilde{Q} . It is analogous to the standard quality factor
 1464 of a resonator with one important modification; we operate our experiment across a wide
 1465 frequency range so we define \tilde{Q} across the continuum of these resonances, not only on classical
 1466 eigenmodes of the system.

1467 These modifications allow us to construct a relationship between an observable E-field
 1468 (\mathbf{E}_{ant} in Eq. 4.3) and the power available at the port of an antenna for a given aperture

$$\langle P_{\text{ant}} \rangle = \frac{|\mathbf{E}_{\text{ant}}|^2}{\eta_0} \langle \tilde{Q} A_e \rangle, \quad (4.6)$$

1469 where η_0 is the impedance of free space. With this in mind, we perform an RF simulation
 1470 to compute $\langle \tilde{Q} A_e \rangle$.

1471 4.3.2 Simulation of $\langle \tilde{Q} A_e \rangle$

1472 It is difficult to make claims about statistical uniformity in the "undermoded" regime where
 1473 modes are not sufficiently mixed [62], so we have employed a commercial electromagnetic
 1474 finite-element modeling software package (COMSOL Multiphysics RF module [63]). Within
 1475 the simulation, a model of the antenna (with a 50Ω feed) is placed in a simplified room with

1476 wall features removed. Spot testing at various frequencies has shown that averaging results
1477 from various antenna positions using this simplified simulation behaves very similarly to one
1478 with the room features included at a fraction of computational complexity.

1479 Two similar simulations are run; driving an E-field while measuring the antenna's re-
1480 sponse and driving a second small monopole antenna and measuring the response of the
1481 primary antenna.

1482 In the first simulation, we drive currents on the walls which correspond to a surface E-
1483 field magnitude of 1 V/m (made up of equal components in the x, y and z directions) using
1484 COMSOL's source electric field option. This field takes the place of \mathbf{E}_{ant} in Eq. 4.6. The
1485 antenna/cavity system resonates and causes an enhancement by \tilde{Q} . The power received at
1486 the antenna's port is measured, allowing the calculation of $\tilde{Q} A_e$, again from Eq. 4.6. By
1487 repeating this simulation for several positions, averaging allows us to compute $\langle \tilde{Q} A_e \rangle$.

1488 The second simulation shares the same geometry, but is used to compute a correction
1489 factor to account for differences between simulation and measurement and to estimate un-
1490 certainty on the first simulation through comparison to physical measurement. Rather than
1491 driving the system through currents on the walls, power is injected into the system with
1492 a 40 cm monopole. From this simulation, two port scattering parameters (S parameters,
1493 defined in 4.3.3) are computed. A similar test is performed on the physical system using a
1494 vector network analyzer (VNA) which provides a physical measurement of the S parameters
1495 to compare with the simulation. The processing of the simulated and measured S parameter
1496 datasets are discussed in the following sub-section. A screenshot of the COMSOL model GUI
1497 is shown in Fig. 4.9. The resulting S parameters from the simulation are plotted against the

1498 measured S parameters in Fig. 4.10

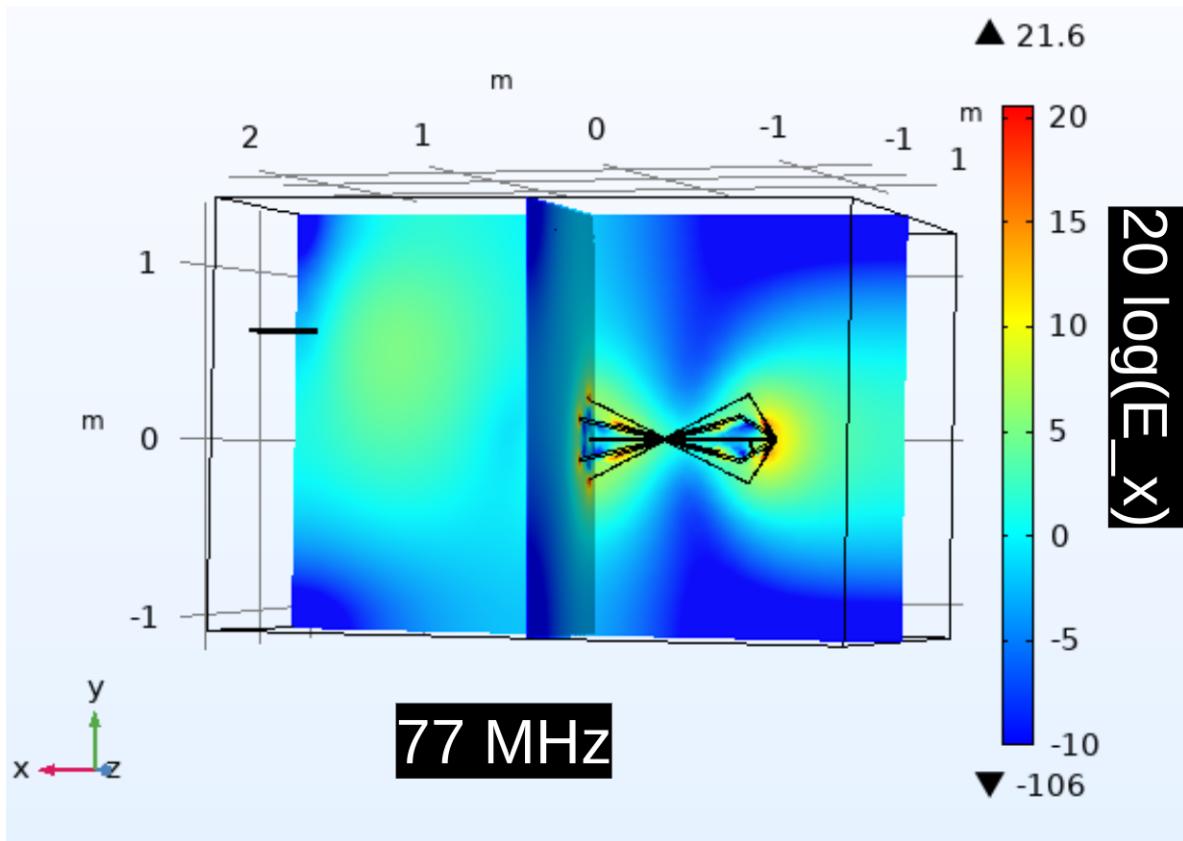


Figure 4.9: Screenshot of COMSOL simulation GUI for two-antenna validation. Shown only at a single position and single frequency (77 MHz). There are 18 antenna positions, 3 E-field components, and ≈ 1000 frequency points, so there are many similar figures to this one. X component of electric field shown as color in dB compared to 1 V/m.

1499 Both simulations are run at the same 18 positions; 9 of which are approximately equiv-
1500 alent to the physical antenna positions while the other 9 are different in order to estimate
1501 how many positions are required for decent convergence of $\langle \tilde{Q} A_e \rangle$. Repeatedly averaging 9
1502 different, random positions (with replacement) results in about 20% variation on their aver-
1503 aged S_{12} coefficients at each frequency, allowing us to conclude 9 positions and polarizations
1504 provides acceptable convergence.

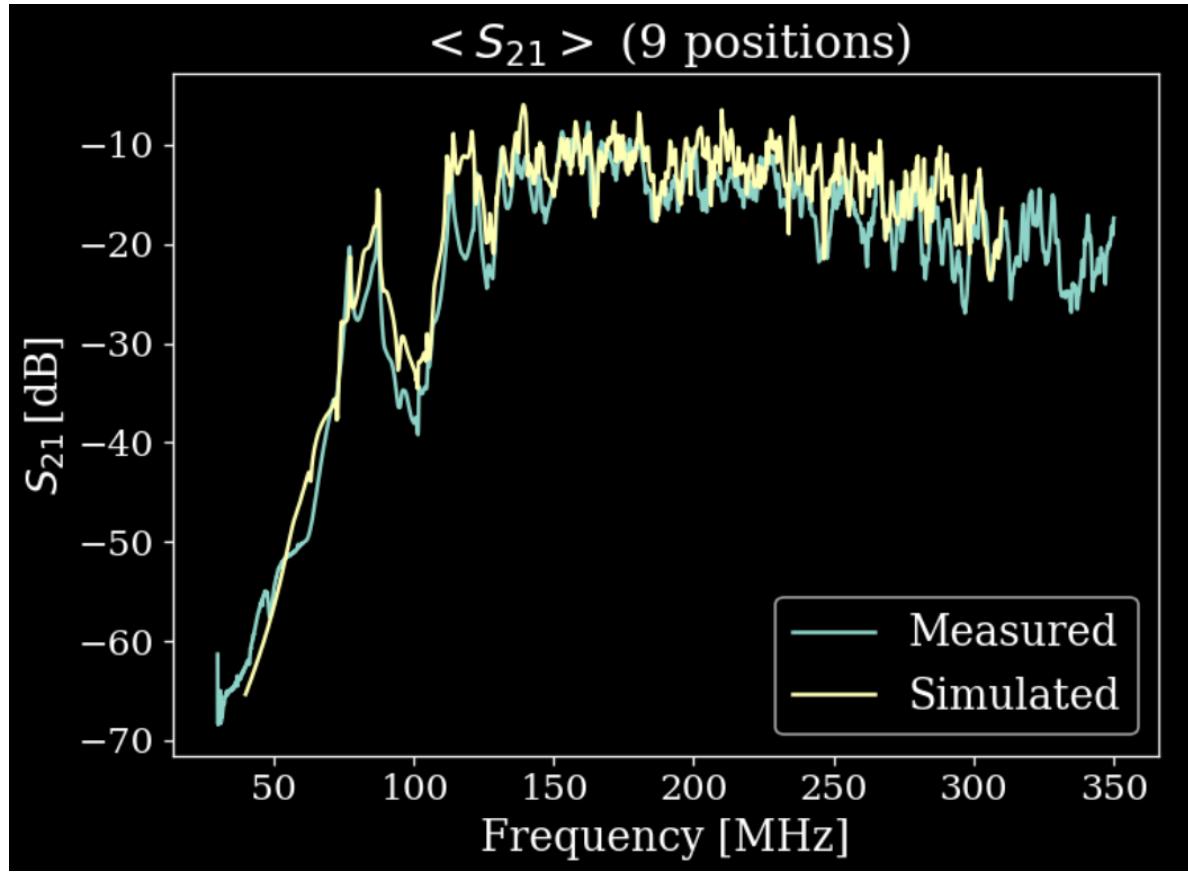


Figure 4.10: Simulated vs measured $\langle S_{21} \rangle$ comment: Replot without dark theme.

1505 4.3.3 Correction and uncertainty of $\langle \tilde{Q} A_e \rangle$

1506 As outlined above, we approximate the uncertainty of the simulation by injecting power into
 1507 the system via a second antenna and comparing the results to simulation.

1508 For a two port microwave device, the ratio between the voltage presented at port one
 1509 and the voltage measured at port two is known as S_{21} . For our system, S_{21} is a measurable
 1510 quantity which is similar to a dark photon detection in that it requires the antenna to convert
 1511 an electric field (which has interacted with the room) into a port voltage. Having frequency

1512 dependent measurements of S_{21} for simulation and measurement give us a correction to
 1513 the simulation (to account for discrepancies in geometry) and estimate the uncertainty on
 1514 $\langle \tilde{Q} A_e \rangle$.

1515 The difference between the measured and simulated values of $\langle |S_{21}| \rangle$ can be described by

$$\langle |S_{21}^{\text{meas}}|^2 \rangle = \alpha \langle |S_{21}^{\text{sim}}|^2 \rangle, \quad (4.7)$$

1516 where meas/sim indicates measured/simulated and the average is over all 18 measured/sim-
 1517 ulated positions and orientations of the antenna. We have taken the square since we are
 1518 interested in the aperture, which is proportional to the square of the voltage. This equa-
 1519 tion implies α is a frequency dependent, multiplicative correction factor which results in a
 1520 corrected $\langle |S_{21}^{\text{sim}}|^2 \rangle$. We find α to have a mean of 0.6, a minimum of 0.1 and a maximum of
 1521 2.

1522 To determine uncertainty on effective aperture, we define the following test statistic

$$\Delta = \frac{\langle |S_{21,n}^{\text{meas}}|^2 \rangle - \alpha \langle |S_{21,n}^{\text{sim}}|^2 \rangle}{\langle |S_{21}^{\text{meas}}|^2 \rangle}, \quad (4.8)$$

1523 where n refers to the subset of n measured/simulated positions sampled randomly with re-
 1524 placement. Δ defines the fractional difference between corrected, simulated S_{21} and measured
 1525 S_{21} . The test statistic, Δ , is calculated 1000 times, providing a distribution of frequency
 1526 dependent Δ s. The curves bounding 63% of these curves are taken to be the uncertainty on
 1527 Δ . Thus we can calculate the corrected $\langle \tilde{Q} A_e \rangle$ as well as its uncertainty. This is shown as
 1528 a function of frequency in Fig. 4.11. The uncertainty on it is shown in gray, and is simply

$$\delta \langle \tilde{Q} A_e \rangle = \langle \tilde{Q} A_e \rangle \delta \Delta. \quad (4.9)$$

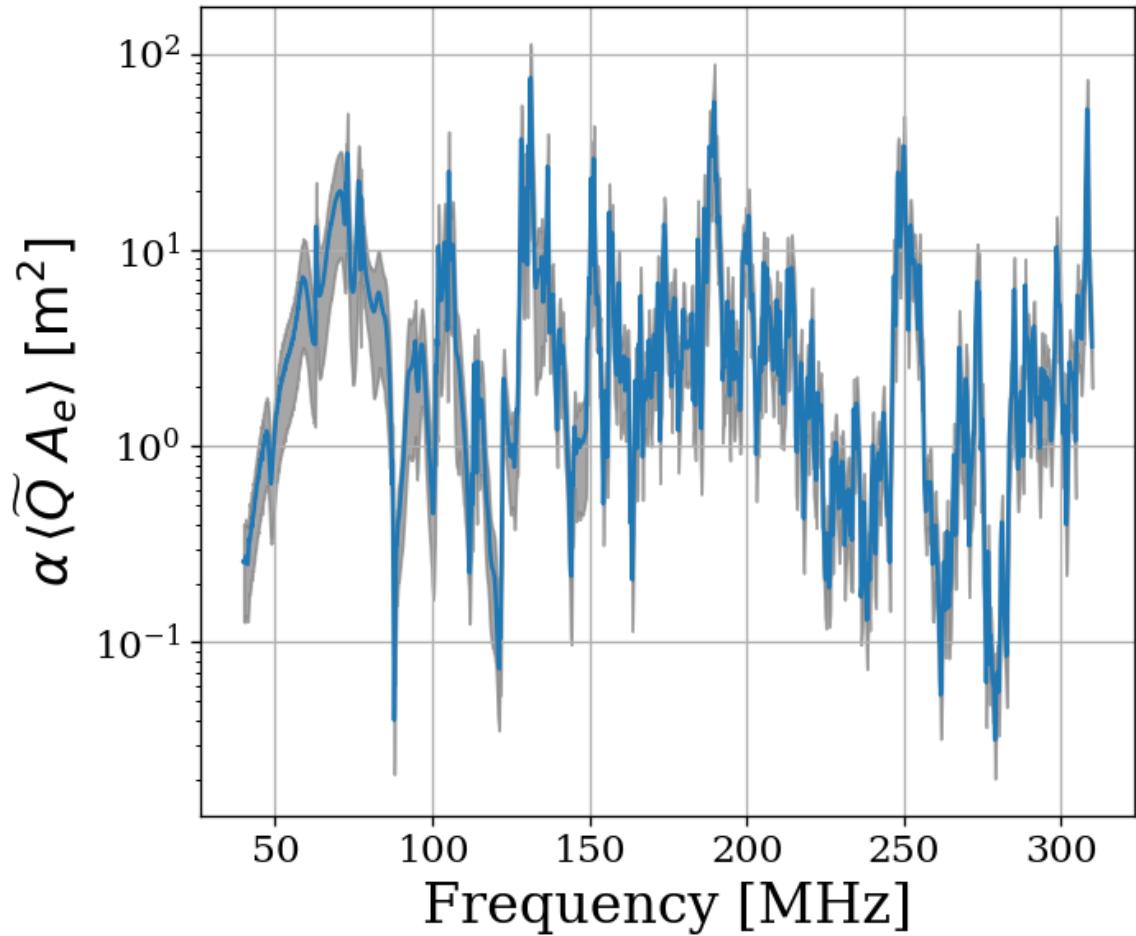


Figure 4.11: Corrected average effective aperture. Calculated with COMSOL RF. The aperture correction α (Eq. 4.7) and its uncertainty (gray) are estimated by comparing simulations to measured S parameters.

1529 A brief summary of the system's aperture is in order. In free space an antenna's ability
 1530 to couple an incoming wave's power density into a transmission line is given by it's effective
 1531 aperture, Eq. 2.9. An antenna in a cavity acts as a coupled oscillator which exhibits very

1532 complex resonances above the first few modes (around 100 MHz for our system). Attempts
 1533 to simulate an aperture for the antenna-cavity system are difficult because of the system's
 1534 extreme dependence on placement of any conductor in the room, especially the antenna.
 1535 Averaging over system configurations (antenna positions and polarizations in our case) allows
 1536 for a significantly more repeatable *statistical* treatment of the aperture/quality factor, which
 1537 we call $\langle \tilde{Q} A_e \rangle$. Comparison of simulated and measured S_{21} gives a small, dimensionless
 1538 correction factor α , Eq. 4.7.

1539 Armed with $\alpha \langle \tilde{Q} A_e \rangle$ we are now able to compute a limit on epsilon using measured and
 1540 simulated quantities via Eqs. 4.5 and 4.6,

$$\epsilon(\nu) < \sqrt{\frac{1}{2c\rho_{\text{DM}}} \frac{P_{\text{ant}}^{\text{lim}}}{\alpha \langle \tilde{Q} A_e \rangle}}, \quad (4.10)$$

1541 where c is the speed of light, ρ_{DM} is the local dark matter density and $P_{\text{ant}}^{\text{lim}}$ is defined in
 1542 Eq. 4.5. We have separated the equation into constants (or in the case of ρ_{DM} , values which
 1543 we fix) and values which we measure or simulate.

1544 In order to validate our entire detection system, we inject sub-threshold signals into the
 1545 shielded room to verify we are able to detect them.

1546 4.4 Hardware Injection Test ⁴

1547 To validate detection methodology, a separate, proof-of-concept run with an a proxy dark
 1548 photon signal injected into the shielded room was performed. Apart from the injection

⁴Code for this section can be found at: <https://github.com/josephmlev/darkRadio/tree/master/daqAnalysisAndExperiments/run1p4/injectionTesting/injectionTesting.ipynb>

1549 antenna (a 40 cm monopole, see Sec. 4.3.2), the setup was equivalent to run 1.4, including
1550 the data analysis. The proxy dark photon signal (detailed in Sec. 4.4.1.2) was injected at a
1551 frequency set by a colleague and was unknown to me at the time of analysis, constituting a
1552 “blind” analysis.

1553 4.4.1 Injection test prerequisites

1554 4.4.1.1 Determination of required injected power

1555 comment: S parameters need to be squared. Check on this. See Besnier [64] To accomplish
1556 the test, a minimum detectable power required for injection P_{inject} must be computed. P_{inject}
1557 should correspond to a signal that can be detected in a predictable amount of time (with
1558 some uncertainty, discussed in detail in Sec. 2.2.2). A simple way to begin is to read off the
1559 detectable, total integrated, power from Fig. 4.8. In other words the power contained in a
1560 dark photon line, integrated over the few bins spanned by the line ($Q_{\text{DP}} \approx 10^6$, discussed in
1561 Sec. 2.1.2). This gives the amount of output-referred power that would be detectable 95%
1562 of the time after 9 days of integration. Since we don’t want to wait 9 days for this test, it is
1563 simple to convert this detection limit into one which would be produced in a shorter time by
1564 the Dicke radiometer equation, Eq. 2.16. Namely, the limit on power scales like the square
1565 root of time ⁵, so a one hour integration will require a factor $\sqrt{9 \text{ days}/1 \text{ hour}} = 14.7$ more
1566 power than is shown in Fig. 4.8.

⁵It is important to point out that one needs to test whether or not the system in question actually behaves as predicted by the Dicke equation for the amount of averaging in question. After lots of averaging, one may encounter non-thermal backgrounds which do not scale properly. It is shown in Fig. 4.16 that the dark radio system follows the Dicke radiometer equation at least for 9 days.

1567 At this point the simplest way to proceed is to measure the the average through-power
1568 of the monopole to the bicon in several antenna positions $\langle |S_{21}|^2 \rangle$. This is the same as the
1569 set up described in Sec. 4.3.2. The bicon was moved to 9 positions and the resulting S
1570 parameters were measured at the reference planes shown in Fig. 4.12. They are shown in
1571 Fig. 4.13 after being averaged together.

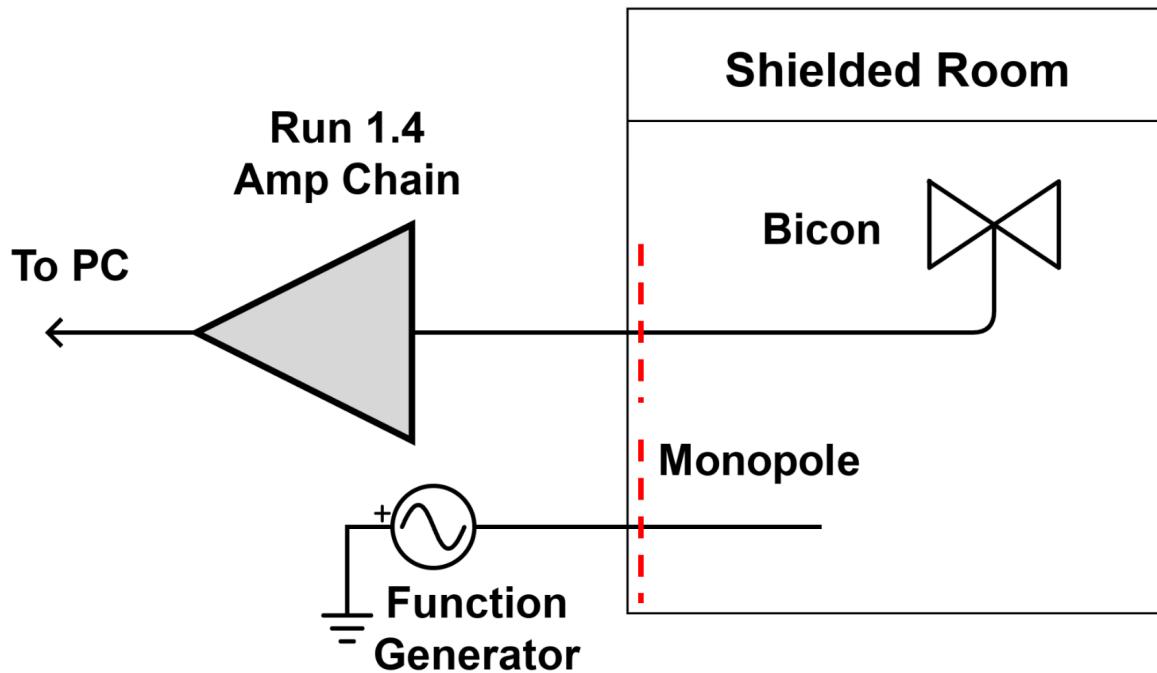


Figure 4.12: Schematic of hardware injection test. Dashed red lines indicate reference planes used to measure S parameters, shown in Fig. 4.13. “Amp Chain” includes amplifiers, attenuators and filters as described in Sec. 2.5.4.

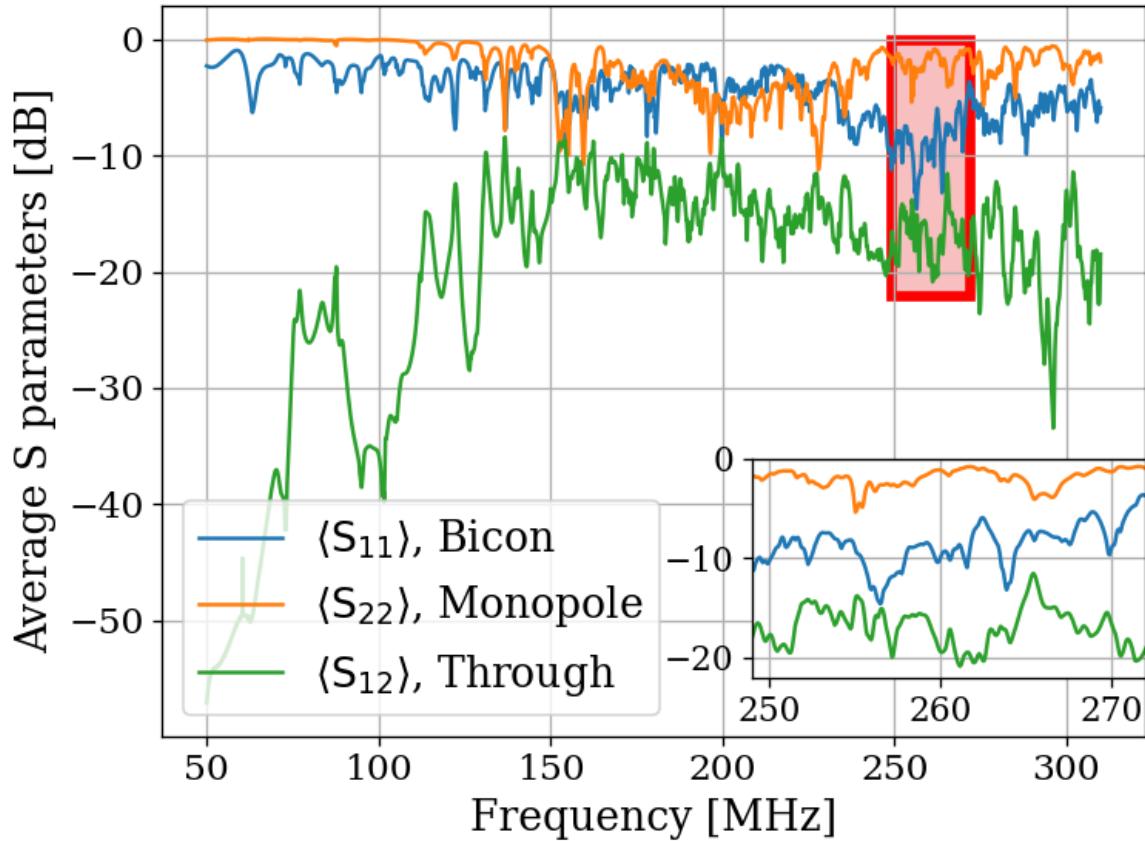


Figure 4.13: Average S parameters of hardware injection test. Taken at 9 positions of the bicon antenna and averaged together (linearly). Ports 1 and 2 are defined to be for the bicon and monopole respectively. Note that $\langle S_{12} \rangle = \langle S_{21} \rangle$ so only $\langle S_{12} \rangle$ is shown. Inset shows zoom on area enclosed by red box.

1572 Knowing the output-referred limit, the system's average $\langle |S_{21}|^2 \rangle$, and the gain G, allows
 1573 for a calculation of the required signal strength (as a function of frequency) which will become
 1574 detectable after a known amount of time. In order to simplify the test, a smaller 1 MHz
 1575 band is chosen between 268 and 269 MHz for the blind injection, where $\langle |S_{21}|^2 \rangle$ is constant
 1576 to about 1dB. At first glance, this seems to be cheating, however there are still approximately
 1577 21,000 frequency bins in this span, so a detection is very unlikely to be random. Furthermore,

1578 the entire 50-300 MHz span is sent to the detection algorithm (Discussed in Sec. 4.2 and
1579 illustrated in Fig. 4.3) which produces an output without knowing about this frequency
1580 restriction.

With all this in mind, the signal power required is simply

$$P_{\text{inject}} = \frac{P_{\text{o}}^{\text{lim}}}{G \langle S_{21} \rangle},$$

1581 where each term is a function of frequency. For the parameters described in this sec-
1582 tion (including the increase in the power limit $\sqrt{9 \text{ days}/1 \text{ hour}} = 14.7$), this works out to
1583 $\approx 6 \times 10^{-19} \text{ W}$ or -152.5 dBm. The Rigol DSG830 signal generator is not calibrated to
1584 such low levels, so this was achieved through attenuation ⁶.

1585 4.4.1.2 Proxy dark photon signal injection

1586 Now that the power for signal injection has been established, the finite-width proxy-signal
1587 ($Q_{\text{DP}} \approx 10^6$) can be generated. This is discussed in Sec. 2.1.2. An intuitive way to accomplish
1588 this would be with an arbitrary waveform generator injecting a time domain signal which is
1589 the Fourier transform of the the expected Rayleigh line shape, Eq. 2.24. This is the method
1590 of the ADMX experiment (see for example [65]). Another option would be to frequency-
1591 modulate a sine wave such that it slowly sweeps out the line shape, spending am amount of
1592 time at each frequency weighted by Eq. 2.24. While intuitive, I was unable to get this to
1593 work. Zhu et al.'s method of frequency hopping [66] is the discretized version of this and
1594 it was very simple to implement. The signal generator is set to change frequencies at some

⁶Experience with this signal generator has shown it exhibits the best performance is when it is set around -30 dBm. Higher than this, large non-harmonic distortions appear contaminating the run. Smaller, and the signal-to-spurious-noise-floor of the generator is poor, also contaminating the run.

1595 interval (discussed below). The frequency which is set is randomly drawn from the PDF of
1596 Eq. 2.24.

1597 There are two considerations that determine the frequency hop period τ_{FH} that the
1598 frequency is changed⁷. First, τ_{FH} should be much longer than the acquisition time of a single
1599 buffer τ_{FFT} ⁸. On the Rigol signal generator, the power is briefly shut off while the frequency
1600 is changed. $\tau_{\text{FFT}} \ll \tau_{\text{FH}}$ ensures that most FFTs of data don't contain a frequency-hop.
1601 Second, τ_{FH} should be small compared to the total time of integration τ , so that there are
1602 many frequencies represented in the entire run. In the limiting case, $\tau_{\text{FH}} = \tau$ will yield an
1603 averaged spectrum containing a single injected frequency; the proxy-signal will be a delta
1604 function in the frequency domain.

1605 Testing has shown that $\tau_{\text{FH}}/\tau_{\text{FFT}} \approx 10$ is more than adequate to address the first consider-
1606 ation. For run 1A (and therefore, this test which shares settings with run 1A), $\tau_{\text{FFT}} = 2^{24}/800$
1607 MHz = 21 ms, so τ_{FH} was set to 250 ms. This means that over 1 hour, the frequency will be
1608 set to $\approx 1.4 \times 10^4$, which addresses the second concern. A histogram of this signal is shown
1609 in Fig. 4.14.

⁷Zhu et al. randomized this period to prevent any unintentional periodic signals entering. I didn't find this to be necessary.

⁸One can likely bypass this restriction by coordinating the signal generator and ADC such that there is some dead time between collection of buffers, in which the frequency is set. Testing has shown that this restriction is adequate to avoid this extra programming step

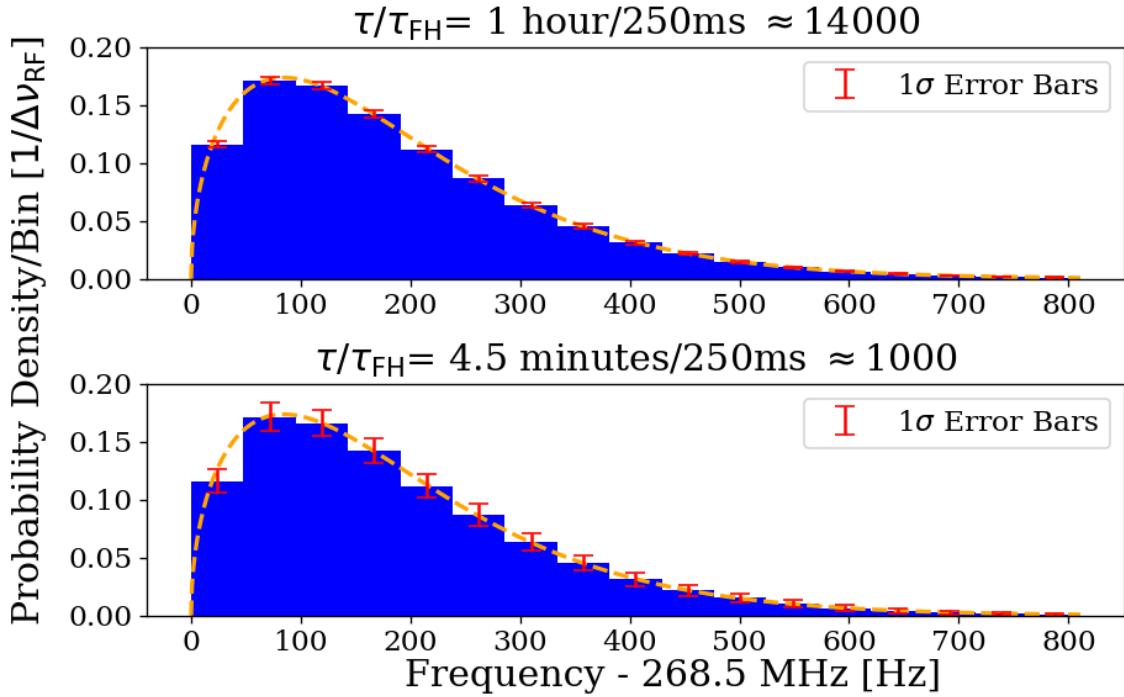


Figure 4.14: Histogram of frequencies used for hardware injection test, with realistic $\Delta\nu_{RF} = 47.7$ Hz. Orange dashed curve is expected line shape from 2.24, and is the PDF frequencies are drawn from. Blue histogram and error bars generated from Monte Carlo simulation and shows the mean value per bin, with 1σ error bars in red. This involves generating 1000 lists of random frequencies (each of length τ/τ_{FH}), binning the data and calculating the standard deviation of each bin. τ is the total acquisition time and τ_{FH} is the amount of time spent on each frequency before “hopping” to the next. Their ratio, τ/τ_{FH} is the number of frequencies which are injected in a given injection test, and was approximately 1.4×10^4 for the one hour test outlined in this section. Two plots give an idea of how error scales with τ/τ_{FH} .

1610 4.4.2 Performing the injection test

1611 Due to the uncertainties involved, more data were taken than the required 1 hour. This also
 1612 helped produce the pretty plot in Fig. 4 of Levine et al. [36]. 3.6 hours of data were collected,
 1613 and saved in 30 second pre-averages so that progressively more data could be averaged if
 1614 the signal was not detected at the predicated time. As mentioned above, the signal injected

1615 was at a relatively high frequency withing the span so that it would be split up into 5 or 6
1616 bins, testing the matched filter's effectiveness. The bicon was moved to 9 positions. Spectra
1617 resulting from 9 antenna positions and 30 seconds of pre-averaging at each position were
1618 averaged together giving 4.5 minute time resolution. Three of these spectra are shown in
1619 Fig. 4.15. The standard deviation of these spectra average down with the square root of time,
1620 closely following the Dicke radiometer equation (discussed in Sec. 2.1.1.3). This scaling is
1621 shown in Fig. 4.16.

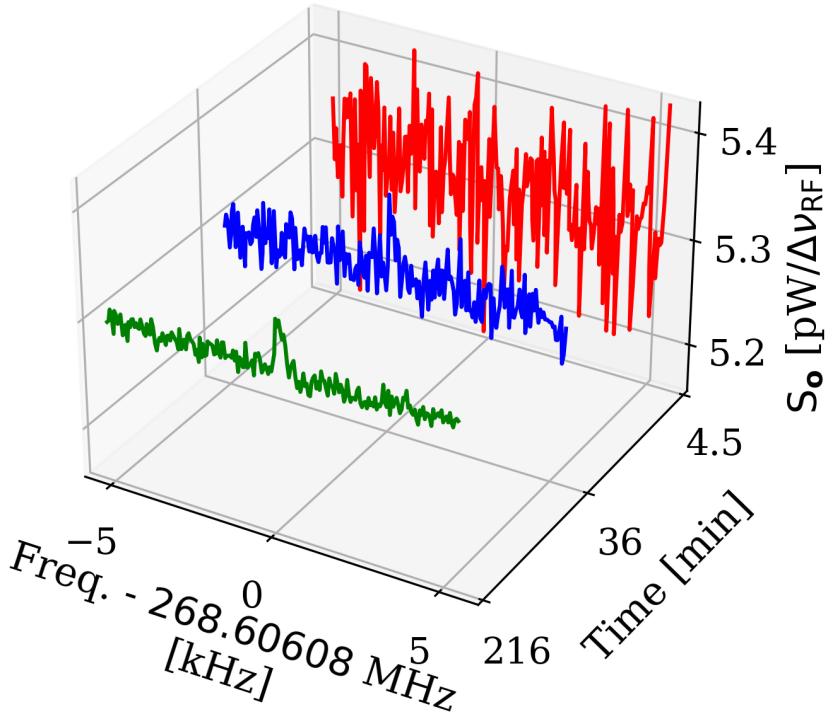


Figure 4.15: Output-referred power spectral density from the hardware injection test illustrating noise averaging down to reveal a persistent, hardware-injected, dark photon proxy signal. Spectra shown are highly zoomed around the injected frequency, 268.60608 MHz. The red, blue and green spectra represent 4.5, 36 and 216 minutes of integration time respectively. The standard deviation of these spectra (excluding the bins containing the injected signal) average down with the square root of time as expected. The blue spectrum shows the amount of averaging required for the signal to be detected by the detection algorithm (including the matched filter) at 5% significance. The tight zoom shown here is less than 1 part in 10^4 of the full 50-300 MHz spectrum analyzed.

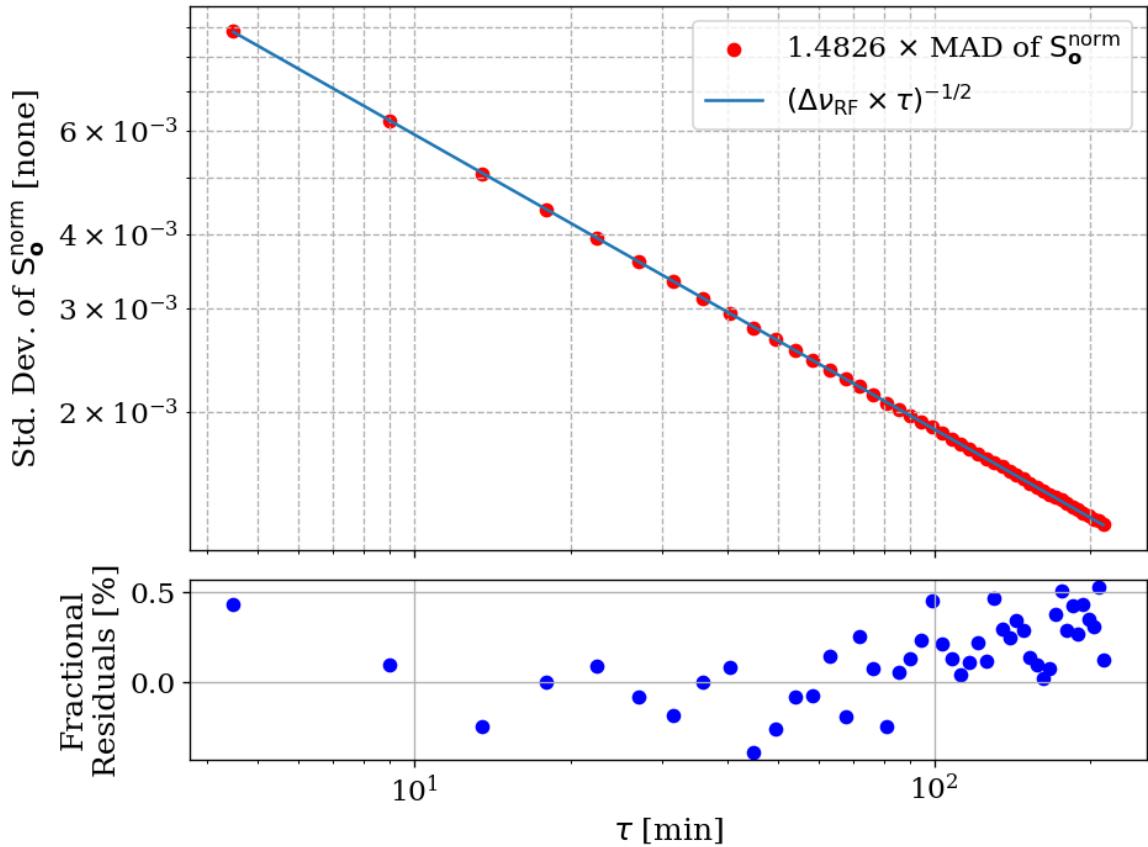


Figure 4.16: Standard deviation of output-referred power spectral density from the hardware injection test, computed with median absolute deviation (MAD). The blue curve represents the predicted standard deviation from the Dicke radiometer equation, Eq. 2.16. Each point corresponds to 9 antenna positions with an additional 4.5 minutes of data averaged (see Sec. 4.4.2). MAD provides a more robust measure of variability, reducing the influence of outliers and offering a better fit than direct standard deviation calculations.

1622 These spectra were generated one at a time and passed through the detection algorithm.
1623 The first spectrum where a signal was detected was at 36 minutes, shown in blue in Fig. 4.15.
1624 Although hardly detectable to the eye, the matched filter detects the signal with 5% signifi-
1625 cance. At the point the signal was detected (i.e. before all data were averaged together), the
1626 injection frequency was confirmed to have been correctly identified, resulting in a success-
1627 ful, blind, hardware injection test. Only after this confirmation were all the date averaged
1628 together to make Fig. 4.15.

1629 **4.4.3 Inspection of Data**

1630 This final subsection simply contains some full page figures which show data from the injec-
1631 tion test. They are all the same 34 minutes of data, but at different stages of processing,
1632 closely following the three main steps of Analysis, Sec. 4.2. They are meant to simulate the
1633 experience of inspecting a 2^{24} point FFT's power spectrum in a matplotlib widget window.
1634 The zoom is seemingly unending, a feature that is difficult to appreciate in a printed docu-
1635 ment. These figures should provide some context for how futile a manual search of unfiltered
1636 data would be. Try and pick out the signal in the top left panel of Fig. 4.17!

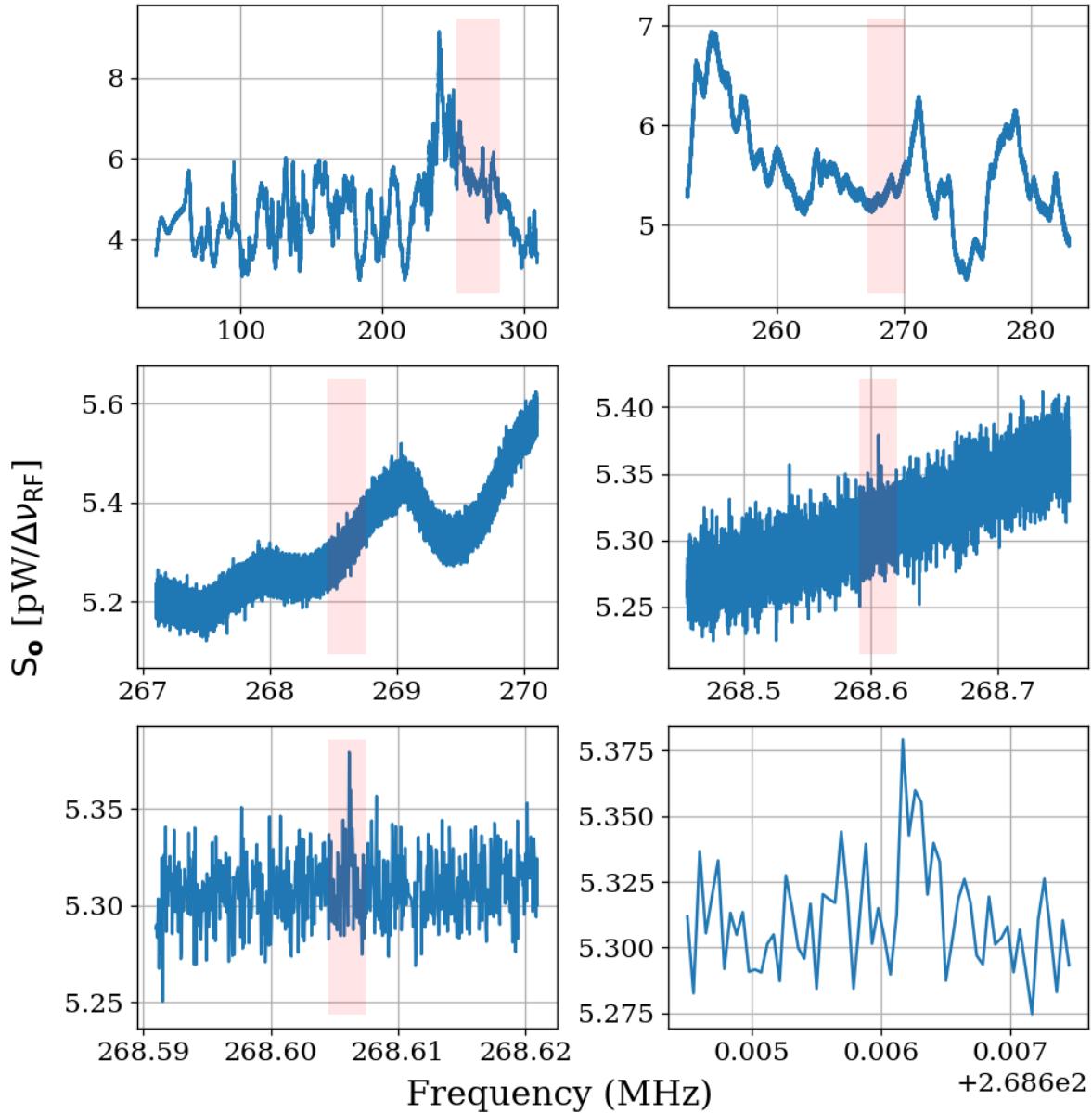


Figure 4.17: Output-referred power spectrum from hardware injection test. Injected signal at 268.60608 MHz. All spectra correspond to a total of 36 minutes of averaging, split evenly between 9 antenna positions. Full 50-300 MHz span contains $\approx 5.2 \times 10^6$ bins. Light pink boxes show zoom level on following plot.

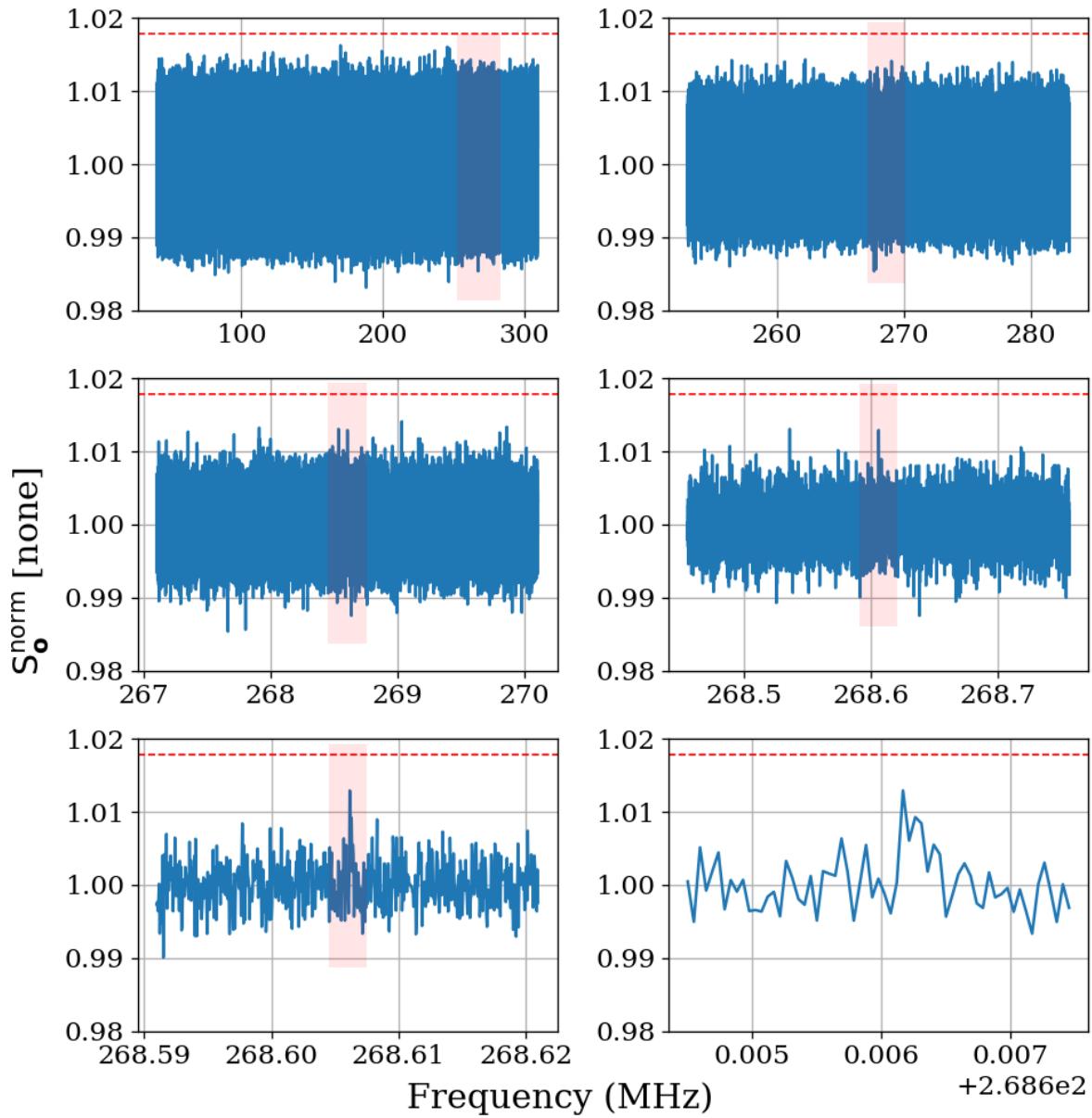


Figure 4.18: Normalized, output-referred power spectrum from hardware injection test. Injected signal at 268.60608MHz. Light pink boxes show zoom level on following plot. Red dashed line indicates the 5% significance threshold, derived in Sec. 2.2.1. Signal is not detectable above threshold.

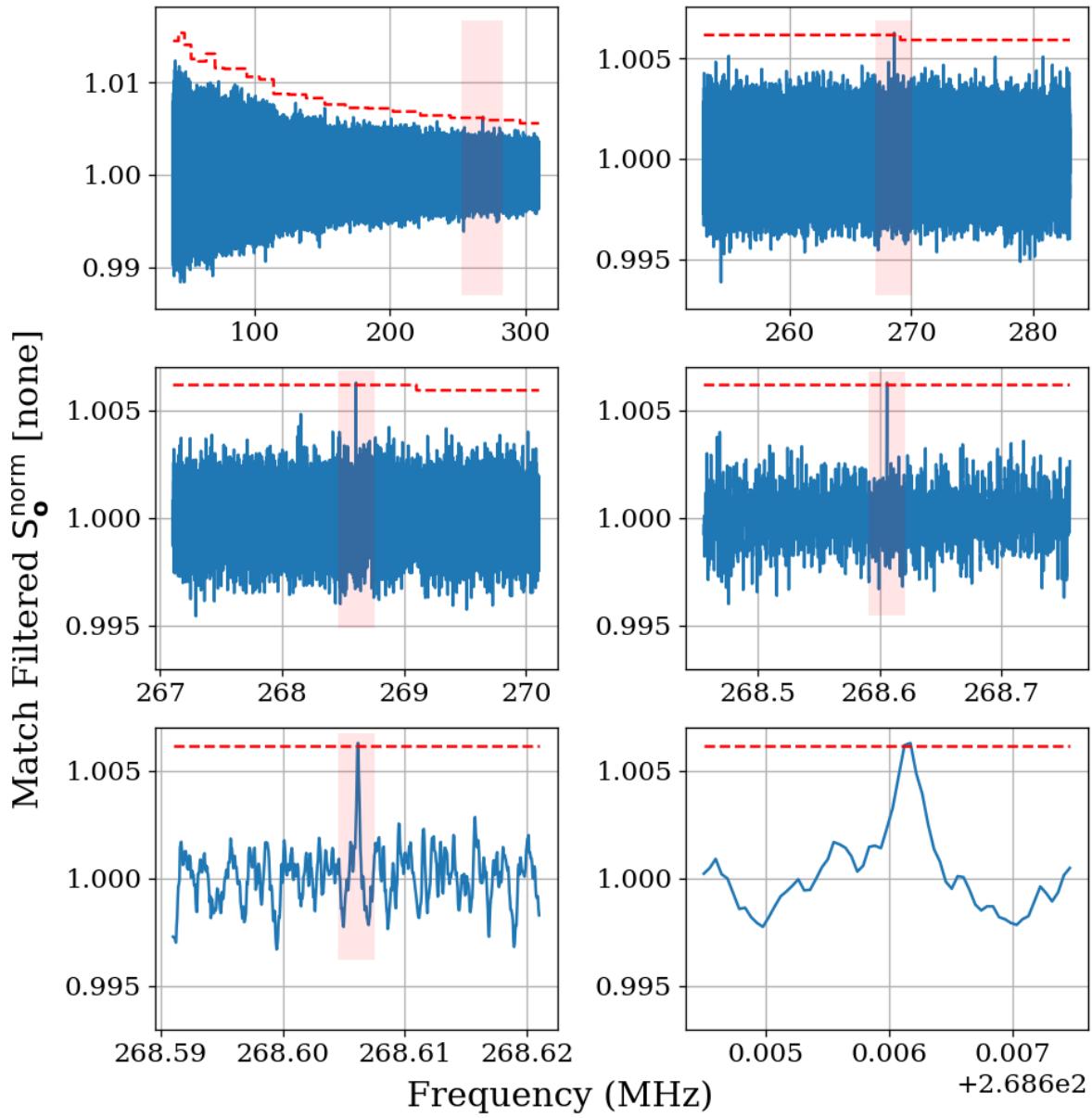


Figure 4.19: Matched filtered, output-referred power spectrum from hardware injection test. Injected signal at 268.60608MHz. Light pink boxes show zoom level on following plot. Red dashed line indicates the 5% significance threshold, derived in Sec. 2.2.1. Introducing the matched filter pushed signal above detectable above threshold compared to Fig. 4.18.

1637 **4.5 Results**⁹

1638 In this section, we report a 95%, frequency-dependent, exclusion limit on the kinetic mixing
1639 strength ϵ of the dark photon (Fig. 4.20). We discuss uncertainties on measured data, identi-
1640 fication of a candidate signal and our process to exclude it. Finally, we display our results in
1641 context by plotting these new limits on top of an aggregation of existing limits in Fig. 4.21.
1642 Future runs of this experiment from 0.3-14 GHz in similar room temperature RF enclosures
1643 and 100 K noise temperature LNAs are indicated (the foundation for such a system is out-
1644 lined in Ch. 5). We have only indicated planned runs, however at microwave frequencies,
1645 highly resonant cryogenic cavities and cryogenic LNAs as well as sub-THz instrumentation
1646 are feasible and could result in an order of magnitude improvement in the limit over the
1647 indicated frequency range and beyond.

1648 **4.5.1 Discussion of uncertainties**

1649 The systematic uncertainty in this experiment comes primarily from three sources, listed in
1650 order of their contribution from greatest to least:

- 1651 1. Fractional uncertainty on the simulated antenna aperture, which is discussed in Sec. 4.3.3,
1652 $\approx 60\%$
- 1653 2. Fractional uncertainty on the first-stage amplifier noise temperature, $\approx 10\%$
- 1654 3. Fractional uncertainty on the gain of the amplifier chain, $\approx 5\%$

⁹Code for this section can be found at: <https://github.com/josephmlev/darkRadio/tree/master/COMPUTELIMIT>

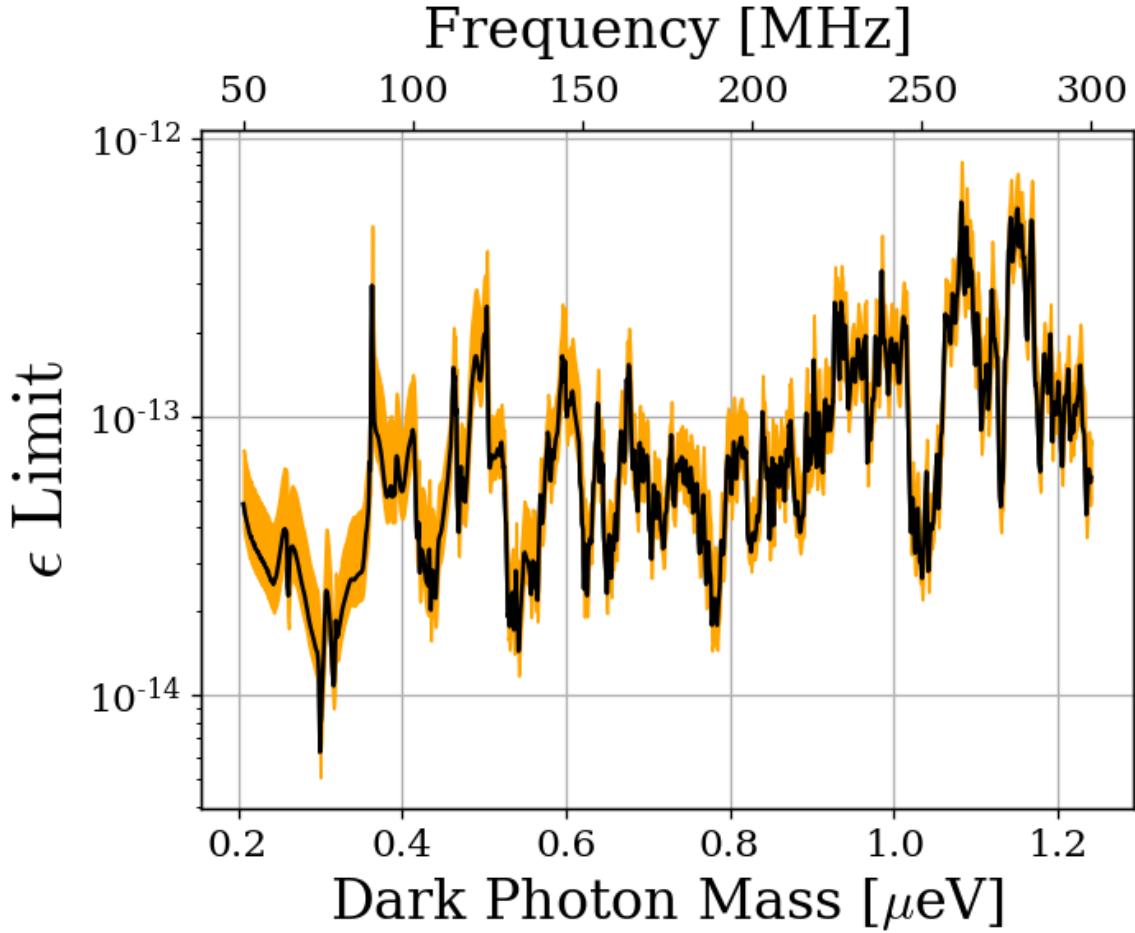


Figure 4.20: 95 % exclusion limit on ϵ with uncertainty shown in orange shaded region. This is based on a local dark matter density of $\rho_{\text{DM}} = 0.45 \text{ GeV/cm}^3$. The error estimate does not take the comparatively small gain and amplifier noise temperature errors into account.

1655 The uncertainty on the simulated antenna aperture is significantly larger than the other

1656 two, and so we neglect them in the uncertainty in the ϵ limit.

1657 We follow the convention of similar experiments where we fix the value of ρ_{DM} and

1658 solve for an ϵ limit given this value. Therefore we treat ρ_{DM} as a known constant with no

1659 uncertainty.

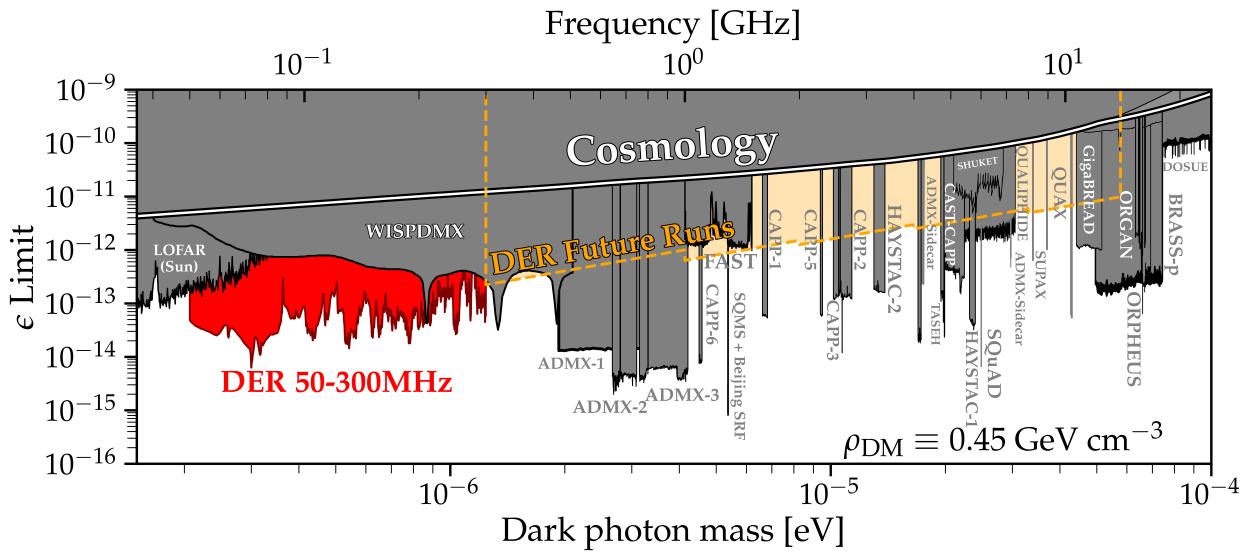


Figure 4.21: Dark photon limits of various experiments circa August 2024, with this work shown in red. The ragged lower bound is due to the complex structure of the resonant modes of the shielded room. Plot adapted by Ben Godfrey from [67] using code found at [68] and includes limit projections of various axion experiments. Astrophysical limits such as CMB interactions with the dark photon are in the region labeled *Cosmology*. Planned wideband extensions of our experiment search from 0.3-14 GHz in similar room temperature RF enclosures are indicated (*yellow*).

₁₆₆₀ Chapter 5

₁₆₆₁ Beyond 300 MHz

To infinity and beyond!

₁₆₆₂

Buzz Lightyear

1663 **Appendix A**

1664 **Overview of RTSA code base**

1665 A basic overview of the code which are used to acquire and process data are outlined here.

1666 The version control is very simple and each experiment has it's own directory containing

1667 several key files. The general usage template from which other experiments can be developed

1668 is in the `teledyneTemplate` directory. The important files, in order of importance, are

1669 • `settings.py`: Settings are controlled from this file. If the DAQ code isn't to be
1670 modified, this is all the user must interact with in normal usage. While it is a python
1671 script, it functions more like a text file. Descriptions and notes about allowed values
1672 are included as comments. Read them carefully, settings can conflict.

1673 • `drDaq*.py`: Main script which calls all the helper functions. To take data, run this
1674 script after modifying (and saving) `settings.py`. There is usually a suffix indicating
1675 the date and information about version.

1676 • `avgFftModule.py`: This is modified code from teledyne. It handles all the heavy

1677 lifting: Interfacing with the C++ API, pinning GPU memory, transfer of data from
1678 PCIE card to the GPU, computation of the FFT on the GPU. This is all wrapped in
1679 a class called `avgFft`. An instance of this class is called `avgSpec` and is the workhorse
1680 of `drDaq.py`

1681 • `daqHelpers.py`: Lots of helper functions which are separated here to keep other code
1682 clean. It is imported as a module in other files. Lots of useful code lives in here,
1683 including the code that converts time series to power spectra (normalization is non-
1684 trivial, see Eq. 2.15), writes info to `database.txt`, saves the pre-averaged spectra
1685 including metadata in an HDF5 file.

1686 • `plotTesting.py`: This uses `dash` to host a web app which allows interactive `plotly`
1687 graphs for simple visualisation of run data. This is extremely useful as it can tell you
1688 if amplifiers die. Without this, all data would have to be averaged and more fully
1689 analyzed, but this is a good light-weight option. There is some creative use of data
1690 down sampling so it runs quickly while not removing any candidates. An example
1691 window is shown in Fig. A.1.

1692 • `backup.sh`: Simple shell script which backs up run data to locations of your choosing.
1693 In the `teledyneTemplate` directory, it is set up to back up to the secondary hard disk
1694 drive in the DR2 machine, and to peloton, but this can be easily modified. When taking
1695 real data, this script should not be run at the same time as `drDaq.py`. `backup.sh`
1696 should be run first, so it completes while the antenna is moved and batteries changed.
1697 This ensures there is not a heavy load on the hard drive due to back ups while data

1698 acquisition is ongoing. `valonInit.py`: sets up the valon signal generator to work as a
1699 clock. This should probably be a function inside of `daqHelpers.py`, but there may have
1700 been a reason I kept it separate.

1701 • `gdrapi.py` Comes from Teledyne. Defines functions for the api. I have not modified
1702 it at all.

1703 • `helperCupy.py` Comes from Teledyne. Defines functions for the GPU. I have not
1704 modified it at all.

1705 • `streamingHelpers.py` comes from Teledyne. Defines functions for streaming from
1706 PCIE card to GPU. I have not modified it at all.

1707 A Tips for using the RTSA system

1708 • Make sure to run `insmod.sh` in the `teledyneInstall/gdrcopy` directory after restart-
1709 ing the machine.

1710 – look at `teledyneInstall/installGuide.txt` on how to handle common errors
1711 involving this process.

1712 • Read through the `settings.txt` file carefully! Some settings will conflict with others,
1713 but they are mostly noted. There are also some notes about possible upgrades which
1714 could be made, some of which without much effort.

1715 B Data structure and processing

1716 Figure 2.21 shows the handling of data as it comes in as an RF time series and is converted
1717 to pre-averaged spectra. While significantly less cumbersome than the raw data, processing
1718 these spectra still represents a challenge. This subsection outlines how I have attempted to
1719 handle it. While it is a little convoluted, this is the third iteration of how to handle this
1720 data processing and is likely simpler than it seems on first blush. In other words, there is
1721 probably a better way to do this, but don't knock it til you try it.

1722 B.1 Writing data

1723 Once a pre-averaged spectrum is computed by dividing the running sum by `NOF_BUFFERS_TO_RECEIVE`,
1724 this can be written to an HDF5 file. Whether or not it is, can be controlled by the `SAVE_H5`
1725 boolean variable¹. Each pre-averaged spectrum is uniquely specified by `ACQ_NUM` in a given
1726 data run. In order to simplify backups and avoid placing all our eggs in one HDF5 basket,
1727 `NUM_SPEC_PER_FILE`² pre-averaged spectra are saved into a single HDF5 file, before starting
1728 a new file. These files are simply named as a zero-indexed number followed by their `.hdf5`
1729 extension. These files are saved in `SAVE_DIRECTORY`³. I usually make this save directory in
1730 a secondary SSD named `drBiggerBoy` in order to preserve the main drive. I think the stress

¹This variable exists because when testing things or taking miscellaneous measurements, you will frequently want to acquire a single spectrum without engaging the complex machinery of the HDF5 saving procedure. You can

²I have kept this around 16 and not experimented much outside of this range, but it's probably fine. This keeps the files around 1 GB. For run 1.4, this is about 45 minutes of antenna data and 3 minutes of terminator data per file, so if a file is corrupted it's not a big deal. I have never had a problem, this is paranoia inherited from Ben.

³A reminder that this is specified, like all other variables, in `settings.py`. Note you must create this directory ahead of time and include a sub-directory called `data`. It says this in the comments of `settings.py` which you are reading, right?

1731 of continuous reads and writes will probably kill this drive, so I prefer to keep it separate
1732 from the main boot drive, drBigBoy.

1733 When `SAVE_H5 == 1`, a `database.txt` file is created in `SAVE_DIRECTORY`, shown in Table

1734 A.1.

Attribute	Pre-averaged spectrum 0	Pre-averaged spectrum 1
ACQ_NUM	0	1
DATETIME	2023-05-10 11:32:48.365	2023-05-10 11:35:49.193
SWITCH POS	0	1
ANT POS IDX	0	0
TEMP	295.64	295.54
LEN FFT LOG2	24	24
SAMPLE RATE MHZ	800.0	800.0
NOF BUFFERS	8600	8600
AMP1	1012_E_PbAcid	1012_E_PbAcid
AMP2	ZKL_9p05VReg	ZKL_9p05VReg
LPF	HSP50+	HSP50+
HPF	288S+	288S+
ATTENUATOR	4dB_FIXED	4dB_FIXED
ADC	ADQ32	ADQ32
CLOCK	SRS_VIA_VALON	SRS_VIA_VALON
File Number	0	0

Table A.1: Example database file from run 1.4. In this run, `ACQ_NUM` counts up to 4175, and these spectra are saved between 261 HDF5 files. Some of the values are auto-generated (`temp`, `DATETIME`, etc.) while others are manually entered into `settings.txt` (`AMP1`, `LPF`, etc.) Note that this table has been transposed in order to fit on the page.

1735 When setting up a data run, you must test that the data are saved how you expect. I have
1736 had success by reducing `NOF_BUFFERS_TO_RECEIVE` to a small number⁴, and taking a simu-
1737 lated data run. The process of switching is hacked together and can give you unpredictable

⁴`NOF_BUFFERS_TO_RECEIVE` times the time per buffer must be larger than around 2 seconds or it will crash

1738 results. Did I mention to read the comments in `settings.py`?

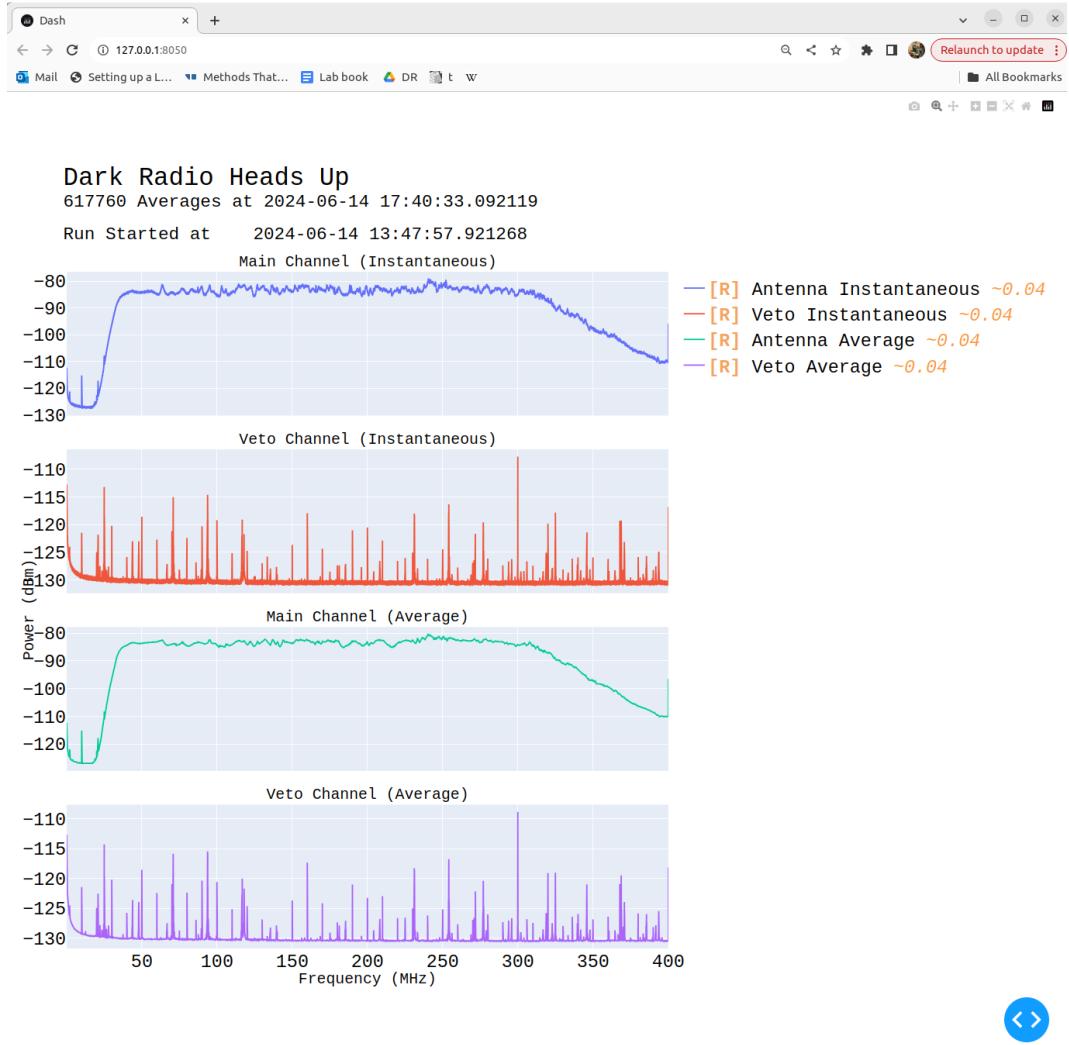


Figure A.1: Real time heads up window for data acquisition. Built with Plotly and Dash, and displays in web browser, in this case Google Chrome. Spectra contain 2^{23} frequency bins, so are down-sampled so as to display 10^4 bins at any given level of zoom. This allows for interaction with the plot in real time with very little lag. The resolution currently displayed (in MHz) is shown in yellow next to the legend (right of first plot). This down-sampling is "smart" in that it will show any excursions, and is implemented with the `plotlyResampler` package in `MinMaxAggregator` mode. Mousing over a curve causes a "Tooltip" to appear, indicating the frequency and power at that point in the curve. The frequency axes are all linked, so zooming on one plot will cause all plots to display the same frequency range. Icons on upper right allow for navigation of plot. Frustratingly there is no back button, so be careful when zooming on a very narrow feature; a wrong move means you have to go all the way back out and start over.

1739 **B.2 Reading and averaging data**⁵

1740 After taking a data run, you should have a `SAVE_DIRECTORY` containing a directory full of
1741 many HDF5 files and a `database.txt` file (and possibly a directory of plotting spectra if
1742 that is chosen in `settings.txt`. The basic idea is to “pre-process” the HDF5 files into a
1743 single, large HDF5 file which can be more quickly accessed to compute a single, averaged
1744 S_o spectrum for further analysis (see next chapter). The `database.txt` file is loaded into
1745 python as a pandas dataframe in order to find specific spectra in this large data structure.

1746 I tried (for a while) to keep the files separate to allow for multi-processing to speed up
1747 the averaging, but I ran into issues. The most serious issue is that all the files live on the
1748 same drive, so you can’t get much of a speed up since the drive is read-limited to around
1749 500 MB/s. I found multiprocessing is more effort than it’s worth. It takes about 20 minutes
1750 to pre-average run 1.4 (9 days of data, 3 minute pre-averages yields 4175 spectra which take
1751 about 280 GB), and you only have to do this once after a run.

1752 Following the example of `run1p4_packAvgAllAnalysis.ipynb`, the first step is to load
1753 `database.txt` as a pandas dataframe which can be used to index the pre-processed file.
1754 Next, a list of which `ACQ_NUMs` correspond to an antenna and terminator must be computed,
1755 after removing any which are known to be contaminated⁶. The list of indices which are
1756 to be averaged together is called either `antIdx` or `termIdx`. Note that you can devise any
1757 number of ways to generate these lists, or even hard code them. The code given is just an

⁵Code for this section can be found at: https://github.com/josephmlev/darkRadio/tree/master/daqAnalysisAndExperiments/run1p4/run1p4_analysis/run1p4_packAvgAllAnalysis.ipynb

⁶In run 1.4, the first antenna and terminator spectra are thrown out, since the computer monitor was on to verify the run started successfully. This was out of paranoia and probably didn’t do anything.

1758 example.

1759 Next, info in the dataframe (no power spectra data yet) are analyzed as a sanity check.

1760 These are not super important but a good check.

1761 The cell beginning with the comment `#pack pre proc dataset` is where the heavy lifting
1762 happens. A large HDF5 file is initialized. It contains two datasets, one for each channel
1763 (main experiment and veto in run 1.4). Note that in this example, it is hard coded to expect
1764 8388609 frequency bins ($2^{23} + 1$), and this must be changed if doing something other than a
1765 2^{24} -point FFT. The HDF5 sub-files are iterated over, their data is read, and written into the
1766 pre-processed HDF5 file. Make sure not to remove the error handling for closing the HDF5
1767 file, or you will corrupt it and have to rerun the pre-processing.

1768 Now that the pre-processed HDF5 file has been packed with data and a list of indices
1769 to be averaged has been generated, averaging is trivial (though still takes a few minutes, so
1770 maybe test it with a sub set of `antIdx`. The code is provided below:

```
1771
17721 def avgSpecFromPacked(avgIdxs, specStr):
17732     sum = np.zeros(2**23+1)
17743     for i in avgIdxs:
17754         sum += f[specStr][:,i]
17765     avgSpec = sum/(len(avgIdxs))
17776     return avgSpec
1778
```

Listing A.1: Python function for averaging spectra from pre-processed data

1779 Note that this assumes a single HDF5 file, `f`, is in scope. `specStr` is a string which
1780 specifies which channel to average, for example `spec_W_chA`.

1781 The output of this function is S_o which can be passed to following functions for further
1782 analysis.

¹⁷⁸³ Bibliography

- 1784 [1] Fritz Zwicky. “The Redshift of Extragalactic Nebulae”. In: (1933).
- 1785 [2] Vera C. Rubin and Jr. Ford W. Kent. “Rotation of the Andromeda Nebula from a
1786 Spectroscopic Survey of Emission Regions”. In: (1970).
- 1787 [3] Marco Battaglieri *et al.* *US Cosmic Visions: New Ideas in Dark Matter 2017: Com-*
1788 *munity Report*. 2017. DOI: 10.48550/ARXIV.1707.04591. URL: <https://arxiv.org/abs/1707.04591>.
- 1790 [4] Benjamin Godfrey et al. “Search for dark photon dark matter: Dark E field radio pilot
1791 experiment”. In: *Phys. Rev. D* 104 (1 July 2021), p. 012013. DOI: 10.1103/PhysRevD.
1792 104.012013. URL: <https://link.aps.org/doi/10.1103/PhysRevD.104.012013>.
- 1793 [5] E. E. Barnard and A. C. Ranyard. “Structure of the Milky Way”. In: *Knowledge: An*
1794 *Illustrated Magazine of Science* 17 (Nov. 1894), p. 253.
- 1795 [6] Kelvin, William Thomson, Baron. *Baltimore Lectures on Molecular Dynamics and the*
1796 *Wave Theory of Light*. English. Collection: University of California Libraries. London,
1797 Baltimore: C. J. Clay and sons; Publication agency of the Johns Hopkins University,
1798 1904, pp. xxi, 703. URL: <https://archive.org/details/baltimorelecture00kelvrich>.

- 1799 [7] Richard H. Maurer. “The Idea Man”. In: *SLAC Beam Line* 31.1 (2001). Accessed:
1800 2024-06-23. URL: <https://www.slac.stanford.edu/pubs/beamline/31/1/31-1-maurer.pdf>.
- 1801
1802 [8] Anthony Tyson. “Mapping Dark Matter with Gravitational Lenses”. In: *Physics Today*
1803 45.6 (June 1992), pp. 24–32. ISSN: 0031-9228. DOI: 10.1063/1.881309. eprint: https://pubs.aip.org/physicstoday/article-pdf/45/6/24/11176778/24_1_online.pdf. URL: <https://doi.org/10.1063/1.881309>.
- 1804
1805
1806 [9] R. Clausius. “XVI. On a mechanical theorem applicable to heat”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 40.265 (1870),
1807 pp. 122–127. DOI: 10.1080/14786447008640370. eprint: <https://doi.org/10.1080/14786447008640370>. URL: <https://doi.org/10.1080/14786447008640370>.
- 1808
1809
1810 [10] Herbert Goldstein, Charles P. Poole, and John L. Safko. *Classical Mechanics*. 3rd. San
1811 Francisco: Addison-Wesley, 2002. ISBN: 9780201657029.
- 1812 [11] A. S. Eddington. “The Kinetic Energy of a Star Clusters”. In: *Monthly Notices of*
1813 *the Royal Astronomical Society* 76.6 (Apr. 1916), pp. 525–528. ISSN: 0035-8711. DOI:
1814 10.1093/mnras/76.6.525. eprint: <https://academic.oup.com/mnras/article-pdf/76/6/525/18494199/mnras76-0525.pdf>. URL: <https://doi.org/10.1093/mnras/76.6.525>.
- 1815
1816
1817 [12] Gianfranco Bertone and Dan Hooper. “History of dark matter”. In: *Reviews of Modern*
1818 *Physics* 90.4 (2018), p. 045002.

- 1819 [13] F. D. Kahn and L. Woltjer. “Intergalactic Matter and the Galaxy.” In: *American*
1820 *Journal of Physics* 130 (Nov. 1959), p. 705. DOI: 10.1086/146762.
- 1821 [14] Hugo van Woerden and Richard G. Strom. “The beginnings of radio astronomy in
1822 the Netherlands”. In: *Journal of Astronomical History and Heritage* 9.1 (June 2006),
1823 pp. 3–20.
- 1824 [15] Virginia Trimble. “History of Dark Matter in Galaxies”. In: *Planets, Stars and Stel-*
1825 *lar Systems: Volume 5: Galactic Structure*. Ed. by Gerard Gilmore. New York, NY:
1826 Springer, 2013, pp. 1093–1116.
- 1827 [16] V. Radeka. “Signal, Noise and Resolution in Position-Sensitive Detectors”. In: *IEEE*
1828 *Transactions on Nuclear Science* 21.1 (1974), pp. 51–64. DOI: 10.1109/TNS.1974.
1829 4327444.
- 1830 [17] B. M. Brubaker et al. “HAYSTAC axion search analysis procedure”. In: *Physical Re-*
1831 *view D* 96.12 (Dec. 2017). DOI: 10.1103/physrevd.96.123008. URL: <https://doi.org/10.1103%2Fphysrevd.96.123008>.
- 1833 [18] C. Boutan et al. *Axion Dark Matter eXperiment: Run 1A Analysis Details*. 2023. arXiv:
1834 2312.16668.
- 1835 [19] Maximiliano Silva-Feaver et al. “Design Overview of DM Radio Pathfinder Experi-
1836 *ment*”. In: *IEEE Transactions on Applied Superconductivity* 27.4, 2631425 (June 2017),
1837 p. 2631425. DOI: 10.1109/TASC.2016.2631425.
- 1838 [20] David Hill. “Plane Wave Integral Representation for Fields in Reverberation Cham-
1839 bers”. en. In: 40 (1998-08-01 00:08:00 1998).

- 1840 [21] W. L. Stutzman and G. A. Thiele. *Antenna Theory and Design*. John Wiley Sons,
1841 1998.
- 1842 [22] J.D. Kraus. *Antennas*. McGraw-Hill, 1950. ISBN: 07-035410-3.
- 1843 [23] H. Nyquist. “Thermal Agitation of Electric Charge in Conductors”. In: *Phys. Rev.* 32
1844 (1 July 1928), pp. 110–113. DOI: 10.1103/PhysRev.32.110. URL: <https://link.aps.org/doi/10.1103/PhysRev.32.110>.
- 1845 [24] J. B. Johnson. “Thermal Agitation of Electricity in Conductors”. In: *Phys. Rev.* 32 (1
1847 July 1928), pp. 97–109. DOI: 10.1103/PhysRev.32.97. URL: <https://link.aps.org/doi/10.1103/PhysRev.32.97>.
- 1849 [25] Robert H Dicke. “The measurement of thermal radiation at microwave frequencies”.
1850 In: *Review of Scientific Instruments* 17.7 (1946), pp. 268–275.
- 1851 [26] Albert Einstein. “Über die von der molekularkinetischen Theorie der Wärme geforderte
1852 Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen”. In: *Annalen der
1853 Physik* 322.8 (1905), pp. 549–560.
- 1854 [27] Herbert B. Callen and Theodore A. Welton. “Irreversibility and Generalized Noise”.
1855 In: *Phys. Rev.* 83 (1 July 1951), pp. 34–40. DOI: 10.1103/PhysRev.83.34. URL:
1856 <https://link.aps.org/doi/10.1103/PhysRev.83.34>.
- 1857 [28] C. L. Mehta. “Coherence-time and effective bandwidth of blackbody radiation”. In:
1858 *Il Nuovo Cimento (1955-1965)* 28.2 (1963), pp. 401–408. DOI: 10.1007/BF02828589.
1859 URL: <https://doi.org/10.1007/BF02828589>.

- 1860 [29] Raphael Cervantes et al. *Deepest Sensitivity to Wavelike Dark Photon Dark Matter*
1861 *with SRF Cavities*. 2022. arXiv: 2208.03183 [hep-ex].
- 1862 [30] Alexander V. Gramolin et al. “Spectral signatures of axionlike dark matter”. In: *Physi-*
1863 *cal Review D* 105.3 (Feb. 2022). ISSN: 2470-0029. DOI: 10.1103/physrevd.105.035029.
1864 URL: <http://dx.doi.org/10.1103/PhysRevD.105.035029>.
- 1865 [31] Michael S. Turner. “Periodic signatures for the detection of cosmic axions”. In: *Phys.*
1866 *Rev. D* 42 (10 Nov. 1990), pp. 3572–3575. DOI: 10.1103/PhysRevD.42.3572. URL:
1867 <https://link.aps.org/doi/10.1103/PhysRevD.42.3572>.
- 1868 [32] Greg Durgin. *Lecture on Noise Figure*. Accessed: 2024-06-23. 2017. URL: https://www.youtube.com/watch?v=M_SRUF4TQgA.
- 1869 [33] Pasternack Enterprises. *PE15A1012: 1 Watt Amplifier Operating From 0.01 GHz to*
1870 *6 GHz*. Accessed: 2024-05-17. 2024. URL: <https://www.pasternack.com/images/>
1871 ProductPDF/PE15A1012.pdf.
- 1872 [34] SP Devices. *ADQ32 Datasheet*. Accessed: 2024-06-23. 2024. URL: https://www.spdevices.com/en-us/Products_/Documents/ADQ32/20-2378%20ADQ32%20datasheet.pdf.
- 1873 [35] Benjamin Godfrey. “A Novel Search for Dark Photons”. Ph.D. dissertation. UC Davis,
1874 2022. URL: <https://escholarship.org/uc/item/72b776v4>.
- 1875 [36] Joseph Levine et al. *New Limit on Dark Photon Kinetic Mixing in the 0.2-1.2 μeV*
1876 *Mass Range From the Dark E-Field Radio Experiment*. 2024. arXiv: 2405.20444
1877 [hep-ex]. URL: <https://arxiv.org/abs/2405.20444>.

- 1881 [37] L.A. Wainstein and V.D. Zubakov. *Extraction of Signals from Noise*. Englewood Cliffs,
1882 NJ: Prentice-Hall, 1962. ISBN: 9780132981330.
- 1883 [38] *Series 83 RF Shielded Enclosure*. Lindgren R.F. Enclosures, Inc. 2021. URL: <https://www.ets-lindgren.com/datasheet/shielding/rf-shielding-and-accessories/11003/1100307>.
- 1886 [39] David A. Hill. *Electromagnetic Theory of Reverberation Chambers*. Technical Note
1887 1506. Gaithersburg, MD: National Institute of Standards and Technology, 1998. URL:
1888 https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=24427.
- 1889 [40] D.A. Hill. “A reflection coefficient derivation for the Q of a reverberation chamber”. In:
1890 *IEEE Transactions on Electromagnetic Compatibility* 38.4 (1996), pp. 591–592. DOI:
1891 [10.1109/15.544314](https://doi.org/10.1109/15.544314).
- 1892 [41] Lawrence Krauss et al. “Calculations for cosmic axion detection”. In: *Physical Review
1893 Letters* 55.17 (1985), pp. 1797–1800. DOI: [10.1103/PhysRevLett.55.1797](https://doi.org/10.1103/PhysRevLett.55.1797). URL:
1894 <https://doi.org/10.1103/PhysRevLett.55.1797>.
- 1895 [42] *Biconical Antenna AB-900A*. Rev. D05.19. Com-Power Corporation. May 2019. URL:
1896 <https://documentation.com-power.com/pdf/AB-900A-1.pdf>.
- 1897 [43] John D. Kraus. *Antennas*. 2nd. New York: McGraw-Hill, 1988, p. 892. ISBN: 9780070354227.
- 1898 [44] Mini-Circuits. *ZKL-1R5: 50 Ω Gain Block Amplifier*. Accessed: 2024-06-23. 2024. URL:
1899 <https://www.minicircuits.com/pdfs/ZKL-1R5.pdf>.

- 1900 [45] Texas Instruments. *LM317: 3-Terminal Adjustable Regulator*. Accessed: 2024-06-23.
1901 2024. URL: <https://www.ti.com/lit/ds/symlink/lm317.pdf?ts=1720749307506>.
- 1902 [46] Alan V. Oppenheim, Alan S. Willsky, and S. Hamid Nawab. *Signals and Systems*. 2nd.
1903 Upper Saddle River, NJ: Prentice Hall, 1996. ISBN: 978-0138147570.
- 1904 [47] Mini-Circuits. *ZX75LP-288-S+: Low Pass Filter*. Accessed: 2024-06-23. 2024. URL:
1905 <https://www.minicircuits.com/pdfs/ZX75LP-288-S+.pdf>.
- 1906 [48] Mini-Circuits. *SHP-50+: High Pass Filter*. Accessed: 2024-06-23. 2024. URL: <https://www.minicircuits.com/pdfs/SHP-50+.pdf>.
- 1908 [49] *Biconical Antenna*. AB-900A. Rev. D. Com-Power Corporation. May 2019.
- 1909 [50] CuPy Developers. *CuPy: NumPy-compatible array library accelerating machine learn-*
1910 *ing with CUDA*. https://docs.cupy.dev/en/stable/user_guide/interoperability.html. Accessed on: May 11, 2023. 2021.
- 1912 [51] *The Y Factor Technique For Noise Figure Measurements*. 1MA178. Version 5e. Rhode
1913 & Schwarz. Oct. 2021.
- 1914 [52] *Noise Figure Measurement Accuracy: The Y-Factor Method*. 1MA178. Version 5e.
1915 Keysight Technologies. Oct. 2021.
- 1916 [53] Wikipedia contributors. *Y-factor Sketch*. Accessed: 2024-06-23. 2024. URL: https://en.wikipedia.org/wiki/Y-factor#/media/File:Y-factor_sketch.jpg.

- 1918 [54] Rigol Technologies. *RSA5000 Series: Real-Time Spectrum Analyzer Datasheet*. Ac-
1919 cessed: 2024-06-23. 2024. URL: https://beyondmeasure.rigoltech.com/acton/attachment/1579/f-0816/1/-/-/-/RSA5_datasheet.pdf.
- 1920
1921 [55] “IEEE Standard Method for Measuring the Effectiveness of Electromagnetic Shielding
1922 Enclosures”. In: *IEEE Std 299-1997* (1998), pp. 1–48. DOI: 10.1109/IEEESTD.1998.
1923 88115.
- 1924 [56] Stanford Research Systems. *FS725 Rubidium Frequency Standard Manual*. Accessed:
1925 2024-06-23. 2024. URL: <https://www.thinksrs.com/downloads/PDFs/Manuals/FS725m.pdf>.
- 1926
1927 [57] R. H. Dicke. “The Measurement of Thermal Radiation at Microwave Frequencies”.
1928 In: *Rev. Sci. Instrum.* 17 (1946), p. 268. DOI: 10.1063/1.1770483. URL: <https://aip.scitation.org/doi/10.1063/1.1770483>.
- 1929
1930 [58] S. Al Kenany et al. “Design and operational experience of a microwave cavity axion de-
1931 tector for the 20-100 micro-eV Range”. In: *Nuclear Instruments and Methods in Physics*
1932 *Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*
1933 854 (May 2017), pp. 11–24. ISSN: 0168-9002. DOI: 10.1016/j.nima.2017.02.012.
1934 URL: <http://dx.doi.org/10.1016/j.nima.2017.02.012>.
- 1935 [59] G. Turin. “An introduction to matched filters”. In: *IRE Transactions on Information*
1936 *Theory* 6.3 (1960), pp. 311–329. DOI: 10.1109/TIT.1960.1057571.
- 1937 [60] Robert N McDonough and Anthony D Whalen. *Detection of signals in noise*. Academic
1938 Press, 1995.

- 1939 [61] S. Asztalos et al. “Large-scale microwave cavity search for dark-matter axions”. In:
1940 *Phys. Rev. D* 64 (9 Oct. 2001), p. 092003. DOI: 10.1103/PhysRevD.64.092003. URL:
1941 <https://link.aps.org/doi/10.1103/PhysRevD.64.092003>.
- 1942 [62] G. Orjubin et al. “Statistical model of an undermoded reverberation chamber”. In:
1943 *IEEE Transactions on Electromagnetic Compatibility* 48.1 (2006), pp. 248–251. DOI:
1944 10.1109/TEMC.2006.870705.
- 1945 [63] *COMSOL Multiphysics*. 5.4. Available at <https://www.comsol.com/>. COMSOL, Inc.
1946 One First Street, Suite 4 Los Altos, CA 94022.
- 1947 [64] Philippe Besnier, Christophe Lemoine, and Jerome Sol. “Various estimations of com-
1948 posite Q-factor with antennas in a reverberation chamber”. In: *2015 IEEE Interna-*
1949 *tional Symposium on Electromagnetic Compatibility (EMC)*. 2015, pp. 1223–1227. DOI:
1950 10.1109/ISEMC.2015.7256344.
- 1951 [65] R. Khatiwada et al. “Axion Dark Matter Experiment: Detailed design and operations”.
1952 In: *Review of Scientific Instruments* 92.12 (Dec. 2021), p. 124502. ISSN: 0034-6748.
1953 DOI: 10.1063/5.0037857. eprint: https://pubs.aip.org/aip/rsi/article-pdf/doi/10.1063/5.0037857/16136265/124502\1_online.pdf. URL: <https://doi.org/10.1063/5.0037857>.
- 1955 [66] Yuqi Zhu et al. “An improved synthetic signal injection routine for the Haloscope
At Yale Sensitive To Axion Cold dark matter (HAYSTAC)”. In: *Review of Scientific*
1958 *Instruments* 94.5 (May 2023). ISSN: 1089-7623. DOI: 10.1063/5.0137870. URL: <http://dx.doi.org/10.1063/5.0137870>.

- 1960 [67] Andrea Caputo et al. “Dark photon limits: A handbook”. In: *Phys. Rev. D* 104 (9
1961 Nov. 2021), p. 095029. DOI: 10.1103/PhysRevD.104.095029. URL: <https://link.aps.org/doi/10.1103/PhysRevD.104.095029>.
- 1962
1963 [68] Ciaran O’Hare. *cajohare/AxionLimits: AxionLimits*. <https://cajohare.github.io/AxionLimits/>. Version v1.0. July 2020. DOI: 10.5281/zenodo.3932430.