

The Pennsylvania State University
The Graduate School
Eberly College of Science

4 DEVELOPMENT OF SCALABLE APPROACHES TO NEUTRINO MASS 5 MEASUREMENT WITH THE PROJECT 8 EXPERIMENT

A Thesis in
The Physics Department
by
Andrew Douglas Ziegler

10 © 2023 Andrew Douglas Ziegler

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

14 Doctor of Philosophy

15 December 2023

¹⁶ The thesis of Andrew Douglas Ziegler was reviewed and approved* by the following:

Luiz de Viveiros Souza
Assistant Professor of Physics
Thesis Advisor
Chair of Committee

Carmen Carmona
Assistant Professor of Physics

¹⁷ Doug Cowan
Professor of Physics and Astrophysics

Xingjie Ni
Associate Professor of Electrical Engineering

¹⁸ Stephanie Wissel
Associate Professor of Physics

¹⁹ *Signatures are on file in the Graduate School.

²⁰ Abstract

²¹ Neutrinos are fundamental particles in the standard model and play an important role
²² in the current understanding of the universe, however, the masses of the neutrinos, one
²³ of the most fundamental parameters for any particles, is currently unknown. This fact
²⁴ represents a gaping hole in our current knowledge of the universe that may provide clues
²⁵ to the energy scale of possible physics beyond the standard model. This dissertation
²⁶ summarizes research and development as a member of the Project 8 collaboration towards
²⁷ an experiment to measure the neutrino mass to a sensitivity below $50 \text{ meV}/c^2$, which
²⁸ is an order of magnitude below the most sensitive direct measurements of the neutrino
²⁹ mass to date. Project 8 will perform this measurement using Cyclotron Radiation
³⁰ Emission Spectroscopy (CRES) to measure the beta-decay endpoint spectrum of atomic
³¹ tritium. I present an analysis of the signal reconstruction performance of an antenna
³² array system designed to perform large-scale CRES measurements. Next, I discuss an
³³ approach to calibrating an antenna array CRES experiment using a unique probe antenna
³⁴ designed to mimic radiation from CRES events. Finally, I present design studies for a
³⁵ resonant cavity that could be used to perform a CRES experiment with atomic tritium
³⁶ at multi-cubic-meter scales.

Table of Contents

38	List of Figures	viii
39	List of Tables	xxvii
40	Chapter 1	
41	Introduction	1
42	1.1 Summary	1
43	1.2 Outline	4
44	Chapter 2	
45	Neutrinos and Neutrino Masses	6
46	2.1 Introduction	6
47	2.2 Neutrinos and Beta-decay	6
48	2.3 Neutrino Oscillations	8
49	2.4 Neutrino Masses in the Standard Model	11
50	2.5 Neutrino Absolute Mass Scale	13
51	2.5.1 Limits from Cosmology	13
52	2.5.2 Limits from Neutrinoless Double Beta-decay Searches	15
53	2.5.3 Limits from Beta-decay	17
54	Chapter 3	
55	Direct Measurement of the Neutrino Mass with Project 8	21
56	3.1 Introduction	21
57	3.2 Project 8 and Cyclotron Radiation Emission Spectroscopy	22
58	3.2.1 Cyclotron Radiation Emission Spectroscopy — CRES	22
59	3.2.2 The Project 8 Collaboration	26
60	3.2.3 The Project 8 Phased Development Plan	29
61	3.3 Phase II: First Tritium Beta Decay Spectrum and Neutrino Mass Mea- 62 surement with CRES	33
63	3.3.1 The Phase II CRES Apparatus	34
64	3.3.2 CRES Track and Event Reconstruction	38
65	3.3.3 Results from Phase II	42
66	3.4 Phase III R&D: Antenna Array CRES	46
67	3.4.1 The Basic Approach	46

68	3.4.2 The FSCD: Free-space CRES Demonstrator	48
69	3.5 Pilot-scale Experiments	53
70	3.5.1 Choice of Frequency	53
71	3.5.2 Pilot-scale Experiment Concepts	55
72	3.6 Phase IV	58

73	Chapter 4	
74	Signal Reconstruction Techniques for Antenna Array CRES and the	
75	FSCD	60
76	4.1 Introduction	60
77	4.2 FSCD Simulations	61
78	4.2.1 Kassiopeia	61
79	4.2.2 Locust	66
80	4.2.3 CRESana	70
81	4.3 Signal Detection and Reconstruction Techniques for Antenna Array CRES	70
82	4.3.1 Digital Beamforming	75
83	4.3.2 Matched Filtering	84
84	4.3.3 Machine Learning	96
85	4.4 Analysis of Signal Detection Algorithms for the FSCD	101
86	4.4.1 Introduction	102
87	4.4.2 Signal Detection with Antenna Array CRES	103
88	4.4.2.1 Antenna Array and Data Rate Estimates	103
89	4.4.2.2 Real-time Signal Detection	106
90	4.5 Signal Detection Algorithms	109
91	4.5.0.1 Power Threshold	110
92	4.5.0.2 Matched Filtering	112
93	4.5.0.3 Machine Learning	116
94	4.5.1 Methods	117
95	4.5.1.1 Data Generation	117
96	4.5.1.2 Template Number and Match Estimation	118
97	4.5.1.3 CNN Training and Data Augmentation	119
98	4.5.2 Results and Discussion	121
99	4.5.2.1 Trigger Classification Performance	121
100	4.5.2.2 Computational Cost and Hardware Requirements	122
101	4.5.3 Conclusion	125

102	Chapter 5	
103	Antenna and Antenna Measurement System Development for the	
104	Project 8 Experiment	127
105	5.1 Introduction	127
106	5.2 Antenna Measurements for CRES experiments	128
107	5.2.1 Antenna Parameters	128
108	5.2.1.1 Radiation Patterns	128
109	5.2.1.2 Directivity and Gain	129

110	5.2.1.3	Far-field and Near-field	130
111	5.2.1.4	Polarization	131
112	5.2.1.5	Antenna Factor and Effective Aperture	132
113	5.2.2	Antenna Measurement Fundamentals	134
114	5.2.2.1	Friis Transmission Equation	134
115	5.2.2.2	S-Parameters and Network Analyzers	135
116	5.2.2.3	Antenna Array Commissioning and Calibration Measurements	136
117	5.2.3	The Penn State Antenna Measurement System	137
118	5.3	Development of a Synthetic Cyclotron Antenna (SYNCA) for Antenna Array Calibration	140
119	5.3.1	Introduction	141
120	5.3.2	Cyclotron Radiation Phenomenology	143
121	5.3.3	SYNCA Simulations and Design	149
122	5.3.4	Characterization of the SYNCA	153
123	5.3.5	Beamforming Measurements with the SYNCA	157
124	5.3.6	Conclusions	160
125	5.4	SYNCA Development Discussion	161
126	5.5	FSCD Antenna Array Measurements with the SYNCA	164
127	5.5.1	Introduction	164
128	5.5.2	Measurement Setups	165
129	5.5.2.1	FSCD Array Setup	165
130	5.5.2.2	Synthetic Array Setup	167
131	5.5.3	Simulations, Analysis, and Results	168
132	5.5.3.1	Simulations	168
133	5.5.3.2	Phase Analysis	169
134	5.5.3.3	Magnitude Analysis	174
135	5.5.3.4	Beamforming Characterization	176
136	5.5.4	Conclusions	179

139	Chapter 6		
140	Development of Resonant Cavities for Large Volume CRES Measurements		181
141	6.1	Introduction	181
142	6.2	Cylindrical Resonant Cavities	182
143	6.2.1	General Field Solutions	182
144	6.2.2	TE and TM Modes	183
145	6.2.3	Resonant Frequencies of a Cylindrical Cavity	185
146	6.2.4	Cavity Q-factors	187
147	6.3	The Cavity Approach to CRES	191
148	6.3.1	A Sketch of a Molecular Tritium Cavity CRES Experiment	191
149	6.3.2	Magnetic Field, Cavity Geometry, and Resonant Modes	193
150	6.3.3	Trade-offs Between the Antenna and Cavity Approaches	196
151	6.4	Single-mode Resonant Cavity Design and Simulations	198

153	6.4.1	Open Cylindrical Cavities with Coaxial Terminations	198
154	6.4.2	Mode Filtering	201
155	6.4.3	Simulations of Open, Mode-filtered Cavities	203
156	6.5	Single-mode Resonant Cavity Measurements	206
157	6.5.1	Cavities and Setup	206
158	6.5.2	Results and Discussion	209
159	Chapter 7		
160	Conclusion and Future Prospects		213
161	Bibliography		215

List of Figures

162	2.1	A diagram of two different neutrino mass ordering scenarios. In the inverted hierarchy (inverted mass ordering) the lightest neutrino mass is m_3 , whereas, in the normal hierarchy (normal mass ordering) m_1 is the lightest neutrino. What cannot be measured by neutrino oscillations is the neutrino absolute mass scale, which is essentially the mass of the lightest neutrino mass eigenstate.	9
163	2.2	The masses of the neutrinos as a function of the lightest neutrino mass in both the normal (a) and inverted (b) mass ordering regimes.	10
164	2.3	The neutrino mass observable measured by cosmology as a function of the lightest neutrino mass eigenstate.	14
165	2.4	Feynman diagrams for double beta-decay (a) and $0\nu\beta\beta$ (b).	15
166	2.5	The discovery probabilities for the future generation of $0\nu\beta\beta$ experiments as a function of $m_{\beta\beta}$ and m_{least} . Figure from [1].	16
167	2.6	A Feynman diagram of beta decay	17
168	2.7	A figure from Fermi's 1934 paper on a theory of beta-decay depicting the kinetic energy spectrum of the emitted electron. The effect of the neutrino mass, written as μ , is to distort the shape of the spectrum near the endpoint from the zero-mass spectrum.	17
169	2.8	The tritium beta-decay spectrum. The effect of a massive neutrino on the spectrum is to change its shape near the endpoint by an amount proportional to the size of the neutrino mass. A sufficiently high-statistic and high-resolution measurement of the spectrum endpoint would be able to measure the neutrino mass.	19
170			
171			
172			
173			
174			
175			
176			
177			
178			
179			
180			
181			
182			
183			
184			
185			

186	3.1	A cartoon illustration of the CRES technique. An electron is contained in 187 a magnetic trap, which is a local minimum in the magnetic field, so that 188 it's cyclotron radiation can be detected by an array of antennas. Detecting 189 the cyclotron radiation allows us to measure its cyclotron frequency and 190 determine its kinetic energy.	23
191	3.2	An illustration of an electron in a bathtub magnetic trap generated by 192 two well-separated coils.	24
193	3.3	A plot of the main components of an electron's trajectory in a cylindrically 194 symmetric trap.	25
195	3.4	A plot of the final state distributions of atomic and molecular tritium. The 196 final state distribution provides the primary contribution to the width of 197 the molecular spectrum whereas thermal doppler broadening is responsible 198 for the width of the atomic spectrum.	27
199	3.5	Neutrino mass exclusion plot including limits from cosmological measure- 200 ments and the KATRIN experiment. Allowed ranges for neutrino masses 201 under the normal and inverted hierarchies are shown as the blue and 202 orange lines respectively. The black dashed line shows Project 8's goal 203 neutrino mass sensitivity for the Phase IV experiment.	30
204	3.6	Sensitivity calculations for a cavity based CRES experiment that demon- 205 strate the neutrino mass measurement goals of the Project 8 collaboration 206 throughout the phased development plan. The blue curves indicate molec- 207 ular tritium sources and the red curves indicate atomic tritium sources. 208 In the current plan, Phase III contains two tritium experiments. The first 209 is the Low-frequency Apparatus (LFA), which is a molecular tritium ex- 210 periment, and the second is the atomic tritium pilot-scale experiment that 211 officially ends Phase III. The sensitivity of these experiments is primarily 212 a function of statistics, however, there is a critical density beyond which 213 CRES electrons do not have enough time to radiate between collisions for 214 a high-resolution frequency measurement leading to worse sensitivity. . .	31
215	3.7	The Phase II CRES apparatus used to perform the first measurement of 216 the tritium beta-decay spectrum using CRES.	34
217	3.8	Diagram of the CRES cell portion of the Phase II apparatus.	36
218	3.9	RF system diagram for the Phase II apparatus.	37

219	3.10 The time-frequency spectrogram of a tritium CRES event in the Phase II 220 apparatus.	39
221	3.11 The sparse spectrogram obtained by placing a power cut on the raw 222 spectrogram shown in Figure 3.10.	41
223	3.12 Fits to the measured 17.8-keV ^{83m}Kr conversion line using the deep and 224 shallow trap configurations.	43
225	3.13 Measurements of the 17.8-keV ^{83m}Kr line using the deep trap configuration 226 for different values of the magnetic field from the field shifting solenoid.	44
227	3.14 The measured tritium spectrum from Phase II with Bayesian and frequen- 228 tist fits.	45
229	3.15 A cartoon illustration of the basics of the antenna array CRES technique.	47
230	3.16 An image of the MRI magnet installed in the Project 8 laboratory at the 231 University of Washington, Seattle.	49
232	3.17 (a) A model of the FSCD antenna array, magnetic trap, and tritium 233 containment vessel design.(b) A more detailed model of a prototype 234 design for the 5-slot waveguide antenna design.	50
235	3.18 (a) A plot of the decay rate for the two-body dipolar spin exchange 236 interaction for cc and dd state. (b) A plot of the decay rate of the dipolar 237 spin exchange interaction for d+d states as a function of magnetic field 238 magnitude. Lowering the magnetic field is key for reducing the losses 239 from this interaction.	54
240	3.19 A conceptual sketch of a large-volume antenna array based CRES experi- 241 ment to measure the neutrino mass.	56
242	3.20 A conceptual sketch of a pilot-scale cavity CRES experiment with an 243 atomic tritium beamline.	57

244	3.21 An illustration of a possible arrangement of ten pilot-scale cavity experiments for Phase IV. The experiments are arranged in a circle with an approximate diameter of 50 meters. Each atomic beamline connected to the bottom of each cavity is approximately 10 m in length. The cavities themselves are designed to operate at 325 MHz and are approximately 11 m tall. The circular arrangement of cavities has some advantages when it comes to cancellation of fringe fields from neighboring magnets, which is important due to the small magnetic field magnitudes consistent with these CRES frequencies. The advantage of ten independent atomic sources and cavities is that there is no single point of failure for the experiment. If an experiment goes down for repairs the other nine may continue running. Figure courtesy of Michael Huehn at UW-Seattle.	59
256	4.1 The geometry and parameters of the coils used to simulate the FSCD magnetic trap in Kassiopeia. Some axial profiles of the magnetic trap at different radial positions are show to demonstrate the shape of the magnetic field and trap depth as a function of position. Calulation of the magnetic field profiles was graciously done by René Reimann.	63
261	4.2 A map of the average ∇B -drift frequency for electrons trapped in the prototype FSCD trap shown in Figure 4.1. Negative drift frequencies indicate electrons that are drifting opposite to the standard direction, which means that they are close to escaping the magnetic trap.	65
265	4.3 The receiver chain used by Locust when simulating CRES events in the FSCD.	68
267	4.4 A high-level diagram depicting the process of CRES event reconstruction. The first step consists of identifying the presence of a signal in the data. This step is necessary to avoid the danger of performing a reconstruction of a false event, which would constitute a background contribution to the tritium spectrum measured by CRES.	72
272	4.5 An illustration of two PDFs associated a binary hypothesis test. The decision threshold is represented by the vertical line that partitions both distributions. The orange and red areas correspond to the true negative and false positive probabilities and the blue and green areas correspond to the false negative and true positive probabilities respectively. To decided between the two hypotheses we perform the likelihood ratio test specified by the Neyman-Pearson theorem. This approach achieves the highest true positive probability for a given false positive probability.	73

280	4.6 An example ROC curve formed by computing the P_{FP} and the P_{TP} for 281 a given likelihood ratio test. As the decision threshold is increased P_{FP} 282 decreases at the expense of a lower P_{TP} . The black dashed line indicates 283 the lower bound ROC curve obtained by randomly deciding between \mathcal{H}_0 284 and \mathcal{H}_1	74
285	4.7 An illustration of the constructive interference condition which is the 286 operating principle of digital beamforming using a uniform linear array 287 as an example.	75
288	4.8 A system level diagram of the laboratory setup used for beamforming 289 demonstrations at Penn State. For more information on this system see 290 Chapter 5. Signals near 26 GHz are fed to a dipole antenna using an 291 arbitrary waveform generator (AWG) and vector network analyzer (VNA), 292 which drive a mixer. The dipole radiation is collected by an array of 293 antennas connected to the digitizer data acquisition (DAQ) system.	77
294	4.9 Photographs of the beamforming demonstration setup. In (a) I show a 295 top-down view of the dipole antenna and the array of eight horn antennas. 296 Manual repositioning of the horn antennas allows one to synthesize a full- 297 circular antenna array. The dipole antenna is mounted on a camera tripod 298 mount that allows for manual position tuning. (b) is a close up image of 299 the dipole, which is manufactured from two segments of semi-rigid coaxial 300 cable. (c) is another image of the dipole and array.	78
301	4.10 An example of digital beamforming reconstruction of a dipole antenna 302 using a synthetic array of horn antennas. The beamforming image on 303 the right is constructed by computing the time-averaged power of the 304 summed signals for a two-dimensional grid of beamforming positions. In 305 the image, one can see a clear maximum that corresponds to the position 306 of the dipole antenna. On the left I show the frequency spectrum of the 307 time-series at the maximum power pixel. White Gaussian noise is added 308 to the signal to mimic a more realistic signal-to-noise-ratio. The signal 309 emitted by the dipole is clearly visible as the high power peak in the 310 frequency spectrum.	79
311	4.11 Beamforming images visualizing the reconstruction of an electron without 312 (a) and with (b) the cyclotron phase correction. The images were generated 313 using data from Locust simulations. The cyclotron phase refers to a phase 314 offset equal to the relative azimuthal position of an antenna in the array. 315 This phase offset is caused by the circular electron orbit and must be 316 corrected for during reconstruction.	80

317	4.12 Beamforming images visualizing the reconstruction of an electron located 318 off the central axis of the FSCD trap. In (a) beamforming is being 319 performed without the ∇B -drift correction, and in (b) we include the 320 ∇B -drift correction.	81
321	4.13 A plot of a typical frequency spectrum obtained by applying a Fourier 322 transform to the time-series obtained from beamforming. The frequency 323 spectra are plotted without noise on top of an example of a typical noise 324 spectrum to visualize a realistic signal-to-noise ratio. In the example we 325 see that without beamforming it would not be possible to detect anything 326 since the signal amplitudes would be reduced by a factor of sixty relative 327 to the noise. Additionally, we see that the ∇B -drift correction is needed 328 to detect this electron since it comes from a simulation of an electron with 329 a significant off-axis position.	82
330	4.14 A plot of the total signal power received by the FSCD array from trapped 331 electrons with different radial positions and pitch angles generated using 332 Locust simulations. The lines on the plot indicate a 10 dB detection 333 threshold above the mean value of the noise in the frequency spectrum. 334 With static beamforming electrons with radial positions larger than about 335 two centimeters are undetectable due to the change in the electron's 336 position over time causing losses from beamforming phase mismatch. This 337 is corrected by including ∇B -drift frequencies in the beamforming phases. 338 Both beamforming techniques fail to detect electrons below $\approx 88.0^\circ$, since 339 these signal are composed of several relatively weak sidebands that are 340 comparable to the noise.	83
341	4.15 Example of a convolution of a CRES signal template with a segment of 342 noisy data. A simulated CRES signal was simulated using Locust and 343 normalized to create a matched filter template. When this template is 344 convolved with noisy data the contains the matching signal the con- 345 volution output increases dramatically compared to data with only noise. 346 The decreasing convolution output as the time offset of the convolution 347 increases is caused by zero-padding of the data and template.	85

348	4.16 An example two-dimensional parameter distribution of a matched filter template bank and random test signals. θ refers to the pitch angle of the electron and E is the kinetic energy. The template bank forms a regular grid of in pitch angle and energy, whereas, the test signals are uniformly distributed in pitch angle and follow the tritium beta-decay kinetic energy distribution. This is why there are fewer test signals at higher energies. The need for high match across the full parameter space prevents one from reducing the density of templates in this low activity region. A zoomed in version of the template bank illustrates the relative density of templates and signals needed for match $> 90\%$	88
358	4.17 The matched filter scores of a dense grid of templates in pitch angle energy space. Dense template grids allow one to estimate the kinetic energy of the electron by identifying the best matching template. The uncertainty on this value is proportional to the space of templates that also match the test signal well. In the worst case matched filter templates can be completely degenerate where templates with different parameters match a signal with equal likelihood.	89
365	4.18 The mean match of the dense template grid shown in Figure 4.17 for different numbers of templates. Grids of different sizes were obtained by decimating a dense grid of templates and the average match for each grid was computed using the same set of randomly distributed test signals. Plotting the mean match against the size of the grid allows one to visualize the exponential relationship between match and template bank size. The noise in each curve is caused by sampling effects from the decimation algorithm. In general, longer templates are harder to match than shorter templates.	89
374	4.19 Matched filter template bank ROC curves as a function of mean match. One can see that for low match a matched filter is on average worse than the more straight forward beamforming detection approach.	90
377	4.20 Boundaries of detectable electrons in pitch angle kinetic energy space for a series of different signal detection algorithms. A detectable signal is defined as a signal that is above a consistent decision with at least 50% probability. This non-rigorous treatment of detection probability is primarily useful for the visualization the relative increases in detection performance provided by the different algorithms. The static beamforming (Static-BF) algorithm is the digital beamforming algorithm introduced above without the ∇B -drift correction. The DNN algorithm refers to a convolutional neural network classifier trained to detect CRES signals (see Section 4.3.3).	91

387	4.21 Two example illustrations of the correlation between kinetic energy and pitch angle imparted by the shape of the FSCD magnetic trap. The correlations manifest themselves as degeneracies in the matched filter score where multiple matched filter templates have the same matched filter for a particular signal. These degeneracies are a sign that the magnetic trap must be redesigned in order to break the correlation between pitch angle and kinetic energy.	95
394	4.22 A visualization of the correlation between energy and pitch angle in the FSCD magnetic trap. The image is formed by computing the match of the best template from a grid consisting of pitch angles from 84 to 90 degrees in steps of 0.05 degrees, kinetic energies from 17574 to 18574 eV, located at 2 cm from the central axis, and simulated for a length of three FSCD time-slices. The signals used to compute the best matching template consisted of a grid from 84 to 90 degrees in steps of 0.05 degrees, kinetic energies from 18550 to 18575 eV in steps of 0.25 eV, located 2 cm from the central axis, and simulated for three FSCD time-slices. The colored regions of the plot show how well signals with lower energy can match those of higher energy for the FSCD magnetic trap, which is proportional to the achievable energy resolution of the FSCD design.	96
406	4.23 A representation of a matched filter template bank as a convolutional neural network. The network has a single layer composed of the templates, which act as convolutional filters. The activation of the neural network is an absolute value followed by a max operator.	99
411	4.24 A graphical depiction of CRES signal detection using a CNN. A noisy segment of data is converted to a frequency series using digital beamforming and a FFT. The complex-valued frequency series is input into a trained CNN model that classifies the data as signal or noise using a decision threshold on the CNN output.	100
416	4.25 The detection efficiency and false alarm rate (false positive rate) as a function of the decision threshold for different values of the noise temperature. The model is trained to output a value close to one for data that contains a signal and outputs a value near zero when the data contains only noise. One sees that a lower decision threshold will have a high detection efficiency at the cost of a high rate of false alarms.	101
422	4.26 ROC curves for a CNN model classifying CRES signals. One can see that the area under the curve, which is a figure of merit that describes the performance of the classifier, is roughly linearly dependent with the noise temperature.	101

425	4.27 An illustration of the conceptual design for an antenna array CRES	
426	tritium beta-decay spectrum measurement. The antenna array geometry	
427	consists of a 20 cm interior diameter with 60 independent antenna channels	
428	arranged evenly around the circumference. The nominal antenna design	
429	is sensitive to radiation in the frequency range of 25-26 GHz, which	
430	corresponds to the cyclotron frequency of electrons emitted near the	
431	tritium beta-spectrum endpoint in a 0.96 T magnetic field. The array is	
432	located at the center of the magnetic trap produced by a set of current-	
433	carrying coils. The nominal magnetic trap design is capable of trapping	
434	electrons up to 5 cm away from the central axis of the array and traps	
435	electrons within an approximately 6 cm long axial region centered on the	
436	antenna array.	104
437	4.28 A block diagram illustration of the real-time triggering algorithm proposed	
438	for antenna array CRES reconstruction.	105
439	4.29 An illustration of the digital beamforming procedure. The blue lines	
440	indicate the distances from the beamforming position to each antenna.	
441	In the situation depicted the actual position of the electron matches the	
442	beamforming position, therefore, one expects constructive interference	
443	when the phase shifted signals are summed. To prevent the electron's	
444	∇B -motion from moving the electron off of the beamforming position,	
445	the beamforming phases include time-dependence to follow the trajectory	
446	of the electron in the magnetic trap.	106
447	4.30 Frequency spectra of simulated CRES events in the FSCD magnetic	
448	trap after beamforming. The signal of a 90° electron consists of a single	
449	frequency component that is clearly detectable using a power threshold	
450	on the frequency spectrum. This power threshold remains effective for	
451	signals with relatively large pitch angles such as 89.0° and 88.75°, which	
452	are composed of a main carrier and a few small sidebands. Signals with	
453	smaller pitch angles, below about 88.5°, are dominated by sidebands such	
454	that no single frequency component can be reliably distinguished from	
455	the noise with a power threshold.	108
456	4.31 PDFs of the power threshold test statistic for CRES signals with various	
457	pitch angles as well as the PDF for the noise-only signal case. The average	
458	PDF computed for pitch angles ranging from 85.5 to 88.5° is also shown.	
459	As the pitch angle is decreased the signal PDF converges towards the	
460	noise PDF which indicates that the power threshold trigger is unable to	
461	distinguish between small pitch angle signals and noise.	111

462	4.32 Plots of PDFs that describe the matched filter template bank test statistic for CRES signals with various pitch angles, as well as the estimated PDF for the noise only case. 10^5 matched filter templates are used and perfect match between signal and template i.e. $\Gamma_{\text{best}} = 1$ is assumed. The mean PDF includes signals ranging from $85.5 - 88.5^\circ$ in pitch angle. There is a larger distinction between the signal PDFs at small pitch angles compared to the power threshold, which indicates a higher detection efficiency for these signals.	115
470	4.33 The mean match of the matched filter template bank to a test set of randomly parameterized signals as a function of the number or density of templates. The parameter space includes pitch angles from $85.5 - 88.5^\circ$ and energies from $18575 - 18580$ eV.	119
474	4.34 Histograms of the trained CNN model output from the test dataset. The blue histogram shows the model outputs for signal data. The oddly shaped peak near the end is the result of the softmax function mapping the long tail of the raw output distribution to the range $[0, 1]$	120
478	4.35 ROC curves describing the detection efficiency or true positive rates for the three signal classification algorithms examined in this paper. The matched filter (MF) and Power Threshold curves are computed analytically using the distribution functions introduced in Section 4.5, and the CNN curve is computed numerically using the classification results on the test dataset. The percent match indicated in the legend refers to the mean match of the classifier.	121
485	5.1 An example radiation pattern generated using HFSS simulations. The color and radial distance of the surface from the origin indicate the relative magnitude of radiation power emitted by the antenna in that direction. The primary goal of most antenna measurements is typically to measure the antenna pattern, which is used to derive many useful antenna performance parameters.	128

491	5.2 An illustration of the three field regions important for the analysis of 492 an antenna system. Very close to the antenna the electric fields are 493 primarily reactive so there is no radiation. If a receiving antenna were 494 placed in this region most of the energy would be reflected back to the 495 transmitter. Outside of the reactive near-field is the radiative near field. 496 At these distances the antenna does radiate, but the radiation pattern is 497 not well-defined since it changes based on the distance of the receiving 498 antenna. It is only in the far-field region where the radiation pattern 499 becomes constant as a function of distance, which is where the majority 500 of antenna engineering is assumed to take place. The antenna arrays 501 developed by Project 8 for CRES measurements operate in the radiative 502 near-field due to the importance of limiting power loss from free-space 503 propagation, which complicates the design of the antenna system.	130
504	5.3 An illustration of the Friis measurement technique commonly used for 505 antenna characterization measurements.	134
506	5.4 Illustration of a two-port S-parameter measurement setup. S-parameters 507 characterize how incoming waves of voltage or power scatter off of the RF 508 device under test. This allows you to measure important properties of the 509 device. In particular, this framework can be used to model a two antenna 510 radiation pattern measurement, which can be automated using a VNA. . .	135
511	5.5 Two measurement approaches to characterizing an antenna array for CRES 512 measurements. The full-array approach (a) requires a complete antenna 513 array with all the associated hardware. The synthetic array approach 514 (b) utilizes a single antenna and a set of rotation/translation stages to 515 reposition the transmitter or the receiving antenna to synthesize the signals 516 that would be received by the full-array. This approach reduces the cost 517 and complexity of array measurements. A down-side of the synthetic 518 array approach is that multi-channel effects such as reflections cannot 519 be measured. Utilizing both the full-array and the synthetic array is a 520 powerful way to quantify the impact of errors from the multi-channel array.	136
521	5.6 Illustration of the antenna measurement system developed for the Project 522 8 Collaboration. The reference and test antennas can be connected to 523 different data acquisition configurations depending on the measurement 524 goals. The reference antenna is typically a standard horn antenna and 525 the test antenna is mounted on a set of translation stages for positioning. 526 Automated translation stages allow for relatively painless data-taking enabling 527 synthetic antenna array measurements using only a single receiving 528 antenna. Anechoic form designed to mitigate RF reflections surrounds 529 the setup.	138

530	5.7	Diagrams of two measurement system configurations. Configuration (a) utilizes a VNA and is more suited to antenna characterization. Configuration (b) utilizes an AWG and VNA as a signal generation system and digitizer to collect measurement data, which is more suited to simulating CRES measurements. The transmission chain utilizes a quadrature hybrid and a pair of baluns to drive the cross-dipole variant test antenna developed for synthetic CRES measurements.	138
531			
532			
533			
534			
535			
536			
537	5.8	A sketch of an antenna array large-volume CRES experiment. Electrons from β -decays are confined in a magnetic field using a set of trap coils. The cyclotron radiation produced by the motion of the trapped electrons can be detected by a surrounding antenna array to determine the electron energies. Measuring the energies of many electrons produces a β -decay spectrum.	141
538			
539			
540			
541			
542			
543	5.9	A schematic of the antenna array test stand. The circular antenna array has a radius of 10 cm with 60 independent channels (limited number shown for clarity). The test stand includes an arbitrary waveform generator (AWG), local oscillator (LO), and data acquisition (DAQ) hardware. Finally, a specialized Synthetic Cyclotron Antenna (SYNCA) is used to inject signals to test the antenna array.	142
544			
545			
546			
547			
548			
549	5.10	An electron (red dot) performing cyclotron motion in the x-y plane. The resulting cyclotron radiation is observed by an antenna located at the field point of interest.	143
550			
551			
552	5.11	A plot of the numeric solution to Equation 5.30. The time-domain representation of the signal (a) is composed of a zero frequency term and a series of harmonics separated by the main cyclotron frequency as shown in the plot of the frequency spectrum (b). We can see that the relative amplitude of the harmonics beyond $k = 7$ are smaller than the main carrier by a factor of about 10^{-5} and are completely negligible.	146
553			
554			
555			
556			
557			
558	5.12	The amplitude maxima of the cyclotron radiation form an Archimedean spiral as the radiation propagates outward from the cyclotron orbit center (a). A circular antenna array located at a fixed radius from the orbit center will receive electric fields with equal magnitude in each of its channels, but the phase of the electric field incident on each array channel will be linearly out of phase from its neighbor antennas by an amount equal to the angular separation of the two channels (b).	148
559			
560			
561			
562			
563			
564			

565	5.13 An idealized crossed-dipole antenna consists of two electric dipole antennas oriented perpendicular to each other and is fed with four signals with a quadrature phase relationship. An example antenna feed circuit is shown which is composed of a chained combination of a quadrature hybrid-coupler (Quad) and two baluns.	149
570	5.14 A model of the PCB crossed-dipole antenna with dimensions. The design has an inside diameter of 2.16 mm for the central circular trace, which is 0.13 mm wide. The dipole arms each have a width of 1.27 mm and protrude beyond the circular trace by 1.40 mm, which gives an overall width of 4.96 mm for the length of the antenna PCB trace from end-to-end. The overall size of the antenna is 20.0 mm the majority of which is the PCB dielectric material. This design was observed in simulation to maintain the field characteristics of the idealized crossed-dipole while being simpler to fabricate due to the increased size of the antenna.	150
579	5.15 A comparison of the electric field magnitudes, normalized by the maximum value of the electric field in each simulation, plotted on a 10 cm square to visualize the Archimedean spirals formed by the electron (a), the crossed-dipole antenna (b), and a PCB crossed-dipole antenna (c). The matching patterns indicate that the electric fields have similar phase characteristics. These images were generated using Locust simulations for the electron and ANSYS HFSS for both antennas.	151
586	5.16 A comparison of the normalized electric field magnitudes for the ideal crossed-dipole, PCB crossed-dipole, and a simulated electron as a function of the polar angle (θ). (a) Shows the total electric field, (b) shows the ϕ -polarized electric field component, and (c) shows the θ -polarized electric field component. These images were generated using Locust simulations for the electron and ANSYS HFSS for both antennas.	152
592	5.17 A comparison of the normalized electric field magnitudes for the crossed-dipole, PCB crossed-dipole, and a simulated electron as a function of the azimuthal angle (ϕ) evaluated at $\theta = 90^\circ$. This image was generated using Locust simulations for the electron and ANSYS HFSS for both antennas.	153

596	5.18 (a) A cartoon schematic which highlights the routing of the semi-rigid 597 coax transmission lines. (b) A photograph of a SYNCA constructed 598 using the modified crossed-dipole PCB antenna design. Visible in the 599 photograph of the SYNCA are four blobs of solder which are an artifact 600 of the SYNCA's hand-soldered construction. These solder blobs are the 601 most significant deviation from the SYNCA design shown in Figure 5.14 602 and are responsible for a significant fraction of the irregularities seen in 603 the antenna pattern.	154
604	605 5.19 A schematic of the VNA characterization measurements (a). This setup 606 allows for antenna gain and phase measurements across a full 360° of 607 azimuthal angles using a motorized rotation stage and control of the radial 608 position of the SYNCA using a translation stage. A photo of the setup in the lab is shown in (b).	155
609	610 5.20 Linear interpolations of the measured electric field magnitude (a) and 611 phase (b). The data was acquired using a VNA at 120 points spaced 612 by 3 degrees from 0 to 357 degrees of azimuthal angle. The different 613 color lines indicate the vertical offset of the horn antenna relative to the 614 SYNCA PCB and the dashed line shows the expected shape from electron 615 simulations. No significant difference in the antenna pattern is observed for the measured vertical offsets.	156
616	617 5.21 (a) A depiction of the relative phase differences for signals received by a 618 circular antenna array from an isotropic source. The phases correspond to a unique spatial position. (b) A schematic of the setup used to perform 619 digital beamforming.	157
620	621 5.22 Digital beamforming maps generated using a simulated 60 channel array 622 and electron simulated using the Locust package. (a) and (b) show the beamforming maps for simulated electrons without the cyclotron spiral 623 phases and with the cyclotron spiral phases respectively. (c) and (d) 624 show the beamforming maps produced from SYNCA measurements. We 625 observe good agreement between simulated electrons and the SYNCA 626 measurements.	159

627	5.23 A plot of the SYNCA's reconstructed position using the synthesized horn-	
628	antenna array and digital beamforming. (a) Shows the reconstructed	
629	position of the SYNCA compared with the target position indicated by	
630	the positioning system readout. (b) Shows the reconstruction error, which	
631	is the difference between the target and reconstructed positions. The error	
632	bars in (b) are the uncertainty in the mean position of the 2D Gaussian	
633	used to fit the digital beamforming reconstruction peak obtained from	
634	the fit covariance matrix. The mean fit position uncertainty of 0.02 mm	
635	is an order of magnitude smaller than the typical reconstruction error of	
636	0.3 mm obtained by calculating the standard deviation of the difference	
637	between the reconstructed and target position.	160
638	5.24 Images of an early prototype crossed-dipole antenna manufactured by	
639	hand and the first measurement setup. The antenna was constructed	
640	by hand using four stripped coaxial cables. The antenna was connected	
641	to one port of the VNA, and the remaining three ports on the VNA	
642	were connected to horn antenna arranged with 90 deg offsets around the	
643	crossed-dipole. The measured unwrapped S-parameter phases exhibit the	
644	desired relative phase behavior for a SYNCA. These early measurements	
645	were the first laboratory proof-of-principle for the crossed-dipole SYNCA.	162
646	5.25 An early iteration of a crossed-dipole SYNCA antenna simulated in HFSS.	
647	The antenna is electrically small at 26 GHz, which requires dipole arms on	
648	the order of 1 mm long. This design is limited by the minimum achievable	
649	distance between the dipole arms caused by the available diameters of	
650	coaxial cables. The assumed termination scheme for the coaxial cables to	
651	the antenna is hand-soldering, which introduces random variation in the	
652	antenna pattern from the inevitable blobs of solder left on the surface of	
653	the PCB.	163
654	5.26 A diagram of the array measurement system used to test the prototype	
655	FSCD antenna array. A VNA is used as the primary measurement tool,	
656	which is connected to the array through a series of switches. The other port	
657	of the VNA connects to the SYNCA through the quad-balun chain used	
658	to provide the SYNCA feed signals. During measurements the SYNCA	
659	is positioned inside the center of the antenna array and translated to	
660	different radial and axial positions using a 3-axis manual translation stage	
661	setup.	164
662	5.27 Photos of the prototype FSCD antenna (a), the FSCD array and SYNCA	
663	(b), and the translation stages and coordinate system used to position the	
664	SYNCA (c).	166

665	5.28 A photo of the FSCD antenna and the SYNCA in the synthetic array 666 measurement setup at Penn State.	168
667	5.29 The unwrapped phases of signals received by the FSCD antenna array 668 from an electron with a 90° pitch angle located in the plane of the antenna 669 array. The data points indicated the phases extracted from simulation 670 and the dashed lines show the model predictions.	169
671	5.30 Plots of the measured unwrapped phases from the FSCD array (a) and the 672 synthetic array (b) compared to the model predictions for a series of radial 673 positions. The different phases of the sinusoidal phase oscillations in the 674 two plots reflects differences in the coordinate systems of the measurements.	171
675	5.31 The phase errors between the measurement and model for the synthetic 676 array (blue) and the FSCD array (orange) for a series of radial positions. 677 The label JUGAAD refers to an alternative name for the FSCD array 678 setup. As the SYNCA is translated off-axis phase errors with progressively 679 higher oscillation frequency enter into the measurements.	172
680	5.32 Two dimensional plots of the phase errors for the synthetic array (a) 681 and the FSCD (JUGAAD) array (b). In both plots there is evidence 682 of a similar diffraction pattern with bilateral symmetry, but the FSCD 683 array measurments have an additional phase error contribution from the 684 different antennas and paths through the switch network.	173
685	5.33 The amplitude of the signals from CRESana for the FSCD array from a 686 90° electron. As the electron is moved from $R = 0$ the signals begin to 687 have unequal amplitudes depending on the distance from the electron to 688 the antenna.	174
689	5.34 The normalized magnitudes of the S21 parameters measured in the FSCD 690 (orange) and synthetic array (blue) setups. The dominant observed 691 behavior as a function of radius is the increase in the number of magnitude 692 peaks, which was noted in the phase error curves. There does not appear 693 to be a strong change in the relative amplitude of a group of antennas as 694 predicted by CRESana.	175
695	5.35	176

696	5.36 Beamforming images from the synthetic array (a) and FSCD array (b) 697 setups with the SYNCA positioned 15 mm off the central axis. In both 698 images, there is a clear maxima that corresponds to the true SYNCA 700 position. However, in the FSCD array there is an additional faint peak 701 located at the opposite position of the beamforming maximum. This 702 additional peak is the mirror of the true peak and is the result of reflections 703 between antennas in the FSCD array.	177
703	5.37 A comparison of the maximum signal amplitude obtained by beamforming 704 to the signal amplitude obtained with an ideal summation as a function of 705 the radial position of the SYNCA. The amplitudes for the synthetic array 706 are shown in (a) and the FSCD array are shown in (b). In both setups, 707 the signal amplitudes obtained from beamforming are smaller than the 708 signal amplitude that could be attained with the ideal summation without 709 phase mismatch.	178
710	5.38 The ratio of the beamforming signal amplitude to the ideal signal amplitude 711 for the FSCD and synthetic arrays. The FSCD array has a larger 712 power loss from phase error compared to the synthetic array which indi- 713 cates that calibration errors associated with the multiple channels as well 714 as reflections are impacting the signal reconstruction.	179
715	6.1 Geometry of a cylindrical waveguide with radius b	182
716	6.2 The geometry of a cylindrical cavity with length L and radius b	185
717	6.3 Relation of mode frequency to cavity length for a cylindrical cavity with 718 a radius of 18.32 cm.	186
719	6.4 A series RLC circuit.	187
720	6.5 Illustration of the behavior of the input impedance of the series RLC 721 circuit as a function of the driving frequency. The BW is proportion to 722 the width of the resonance, which is inversely proportional to Q.	189
723	6.6 A series RLC circuit coupled to an external circuit with input impedance 724 R_L	190

725	6.7 A cartoon depiction of a cavity CRES experiment. A metallic cavity filled 726 with tritium gas is inserted into a uniform background magnetic field 727 to perform CRES measurements. Electrons from beta-decays inside the 728 cavity can be trapped and used to excite a resonant mode(s). By coupling 729 to the cavity mode with a suitable probe one can measure the cyclotron 730 frequency of the electron and perform CRES.	191
731	6.8 Examples of the resonant mode frequencies of a cylindrical cavity. This 732 cavity has a radius of 18.32 cm and a length to diameter ratio of 4.55.	195
733	6.9 An image of an open cavity with coaxial terminations used for dielectric 734 constant measurements. Figure from [2].	199
735	6.10 The simplified geometry of an open cavity with coaxial terminations. 736 Figure from [3].	200
737	6.11 Electric field regions for the open cavity boundary value problem. Figure 738 from [3].	200
739	6.12 Two mode filtering concepts to break the degeneracy of TE_{01} and TM_{11} 740 modes. The resistive approach uses dielectric materials to impede currents 741 that travel vertically along the cavity while leaving azimuthal currents 742 unperturbed. An alternative approach is to impede the currents using 743 grooves cut into the cavity wall, which achieve the same effect with an 744 inductive impedance.	203
745	6.13 Four cavity design variations. (a) is a standard sealed cylindrical cavity, 746 (b) is an open cavity with smooth walls, (c) is an open cavity with resistive 747 walls, and (d) is an open cavity with grooved walls. The main cavity and 748 coaxial terminator parameter are identical for all four cavities.	204
749	6.14 The frequencies and Q-factors of the resonant modes identified by HFSS 750 for the cavity variations shown in Figure 6.13. The fully-sealed cavity with 751 smooth walls has several high-Q modes near the TE_{011} resonance. Intro- 752 ducing the open-termination preserves the Q-factors of the $TE_{01\ell}$ modes 753 and suppresses the Q-factors of the modes whose boundary conditions do 754 not match the cylindrical partition. Both the resistive and grooved wall 755 perturbations shift the resonant frequencies of the TM modes away from 756 the TE_{011} mode. By properly tuning the geometry of the grooves or the 757 resistive spacers several MHz of frequency separation can be achieved.	205

758	6.15 A cartoon depicting the design of the open-ended cavity prototype designed 759 to operate at approximately 6 GHz. The main cavity wall was composed of 760 a single copper pipe. A mode-filtered version of this cavity was constructed 761 by	207
762	6.16 Images depicting the measurement of the filtered and non-filtered open 763 cavities using the VNA. The coupling loop in the figure is shown in the 764 TE orientation.	208
765	6.17 HFSS simulation results for a the as-built cavity with the coaxial termi- 766 nators removed. The TE_{011}/TM_{111} frequency is approximately 5.78 GHz.	209
767	6.18 Measurements of the filtered and non-filtered prototype cavities acquired 768 with the VNA.	210

List of Tables

769	4.1	A summary of the CNN model layers and parameters. The output of each 1D-Convolution and Fully Connected layer is passed through a LeakyReLU activation function and re-normalized using batch normalization before being passed to the next layer in the model. The output of the final Fully Connected layer in the model is left without activation so that the model outputs can be directly passed to the Binary Cross-entropy loss function used during training. The first layer in the network has two input channels for the real and imaginary components of the spectrum. . .	117
770	6.1	A table of the values of p'_{nm}	184
771	6.2	A table of the values of p_{nm}	185
772	6.3	A table of cavity design parameters used for HFSS simulations.	204
773	6.4	A table of parameters describing the cavity prototypes. Certain values such as the cavity length and the distance between dielectric spacers are approximate due to variation in the machining of the copper. In particular, the filtered cavity was constructed from conducting copper segments that varied in size from 1.50" to 1.85".	207
774			
775			
776			
777			
778			
779			
780			
781			
782			
783			
784			
785			

786 **Chapter 1** |
787 **Introduction**

788 **1.1 Summary**

789 Neutrinos are one of the fundamental particles that comprise the standard model of
790 particle physics and account for a significant fraction of the matter in the universe.
791 Neutrinos are the most abundant fermions in the universe, but due to their weak
792 interactions neutrinos seldom interact with other particles. Regardless, neutrinos play a
793 unique role in the evolution of the early-universe, therefore, a detailed understanding of
794 the properties of the neutrino is key to understanding the universe at the cosmological
795 scale as well as the smallest particle physics regime.

796 It was uncertain that neutrinos had nonzero mass until vacuum neutrino flavor
797 oscillations were observed in the late 90's and early 00's. A simple relativistic argument
798 as to why oscillations are evidence for neutrino masses is that oscillations imply neutrinos
799 experience time, which means that they do not propagate at the speed of light, therefore
800 the masses of the neutrinos must be non-zero. Current neutrino oscillation data supports
801 that neutrino flavor states are actually a superposition of three separate neutrino states
802 with well-defined masses. Measurements of neutrino oscillations that have taken place
803 over the past couple of decades have measured the differences between neutrino mass
804 eigenstates with increasing precision. However, oscillation measurements cannot tell
805 us the mass scale of the neutrinos, which is required in order to measure the absolute
806 neutrino masses.

807 The neutrino mass scale remains an unknown quantity in the standard model of
808 particle physics. The value of the neutrino mass influences the evolution of the early
809 universe and is likely relevant to the energy-scale of new physics responsible for the factor
810 of 10^{-6} difference between the neutrino and electron masses. A model-independent way
811 to measure the neutrino mass is to measure the tritium beta-decay spectrum near its
812 endpoint. Energy conservation requires that the neutrino mass carry away some kinetic

813 energy from the beta-decay electron in the form of its mass, which causes a distortion in
814 the shape of the tritium beta-decay spectrum near the endpoint. The isotope tritium has
815 many advantages for this measurement, and has been used by the KATRIN collaboration
816 to perform the most sensitive direct neutrino mass measurement to date.

817 KATRIN represents the state-of-the-art in the current generation of neutrino mass
818 direct measurement experiments with a projected neutrino mass sensitivity of $m_\nu < 200$ meV.
819 This sensitivity does not fully exhaust the allowed parameter space of neutrino
820 masses under the normal and inverted neutrino mass ordering scenarios, which motivates
821 the development of a next generation of neutrino mass measurement experiments.

822 The Project 8 collaboration is developing a next-generation neutrino mass experiment
823 with a goal neutrino mass sensitivity of $m_\nu < 40$ meV. This sensitivity is sufficient to
824 exhaust the range of neutrino masses allowed under the inverted mass ordering regime.
825 Project 8 intends to achieve its sensitivity goal utilizing two technologies that are novel
826 to the space of direct neutrino mass measurements — atomic tritium and cyclotron
827 radiation emission spectroscopy (CRES). Atomic tritium is required in order to avoid
828 systematic broadening the tritium beta-decay spectrum caused by the final state of the
829 $^3\text{He}^+ - \text{T}$ molecule, and the CRES technique enables a differential measurement of the
830 tritium spectrum that is background-free and able to be directly integrated with the
831 atomic tritium source.

832 The Project 8 collaboration is currently engaged in a research and development
833 program intended to simultaneously develop the atomic tritium and CRES technologies
834 so that they can be combined in a next-generation experiment. This past year (2022)
835 Project 8 has used the CRES technique to measure the molecular tritium beta-decay
836 spectrum and place an upper limit on the neutrino mass: $m_\beta \leq 152$ eV. This measurement,
837 while not competitive scientifically, represents the first proof-of-principle that the CRES
838 technique can be used to measure the neutrino mass.

839 The future goals of the Project 8 collaboration are to develop the technologies
840 and techniques necessary to scale-up the volume in which CRES measurements can
841 be performed. Project 8's first neutrino mass measurement with CRES utilized a
842 measurement volume on the cubic-centimeter scale, however, sensitivity calculations
843 estimate that an experiment sensitive to neutrino masses of 40 meV will require several
844 tens of cubic-meters of experiment volume filled with atomic tritium. Developing a new
845 approach to performing CRES measurements that can be successfully scaled to these
846 volumes is a necessary step towards Project 8's neutrino mass measurement goal, and is
847 the primary topic of my dissertation research.

848 A parallel development is the technology necessary to produce, cool, trap, and
849 recirculate a supply of atomic tritium that is compatible with CRES measurements. The
850 atomic tritium system is equally important as the large-volume CRES measurement
851 technology, but will not be discussed at depth here.

852 The Project 8 collaboration has identified two scalable approaches to neutrino mass
853 measurement using the CRES technique. One approach is to use an array of antennas
854 that surrounds a volume of trapped atomic tritium that can perform CRES measurements
855 by collection the cyclotron radiation emitted by beta-decay electrons into free-space. The
856 other approach uses a resonant cavity filled with atomic tritium to perform CRES by
857 measuring the excitation of resonant cavity modes caused by the motion of electrons
858 trapped inside the cavity volume.

859 The cavity and antenna approaches to CRES have been studied in detail over the past
860 five years, and, while both approaches offer a physically viable path towards a 40 meV
861 neutrino mass measurement the collaboration has elected to pursue the cavity approach
862 for the foreseeable future. The major advantage of the cavity approach is a significant
863 reduction in the cost and complexity of the experiment design and data analysis, which
864 provides a lower risk path to Project 8’s scientific goals.

865 In this dissertation I summarize my most impactful contributions to the research and
866 development of antenna array and cavity CRES. In short these contributions are

- 867 • the development and analysis of signal reconstruction algorithms for antenna array
868 CRES, which provide key inputs to sensitivity analyses of antenna array CRES
869 experiments.
- 870 • The development of a specialized antenna, designed to synthesize fake CRES
871 radiation, which enables bench-top testing and validation of the antenna array
872 CRES technique.
- 873 • The development of an open-cavity design for CRES measurement, whose mode
874 structure can be tuned using perturbations that modify the impedance of the cavity
875 walls. The development of this cavity concept was one of many developments that
876 eventually lead to the adoption of cavities as the CRES technology of choice for
877 the future of Project 8.

878 1.2 Outline

879 The outline of this dissertation is as follows. In Chapter 2 I provide an introduction to
880 the basic physics of neutrinos and beta-decay, which provides context for a discussion of
881 various methods to measure the neutrino absolute mass scale.

882 Chapter 3 is an overview of the CRES technique and the Project 8 collaboration.
883 I highlight the Project 8 Phase II experiment, which was the first measurement of
884 the tritium beta-decay spectrum with CRES, and I discuss the planned research and
885 development for an antenna array CRES experiment in Phase III of the Project 8
886 collaboration’s experiment plan. I end Chapter 3 with a discussion of the pilot-scale and
887 Phase IV experiments, that will combine a scalable CRES measurement technology with
888 atomic tritium and measure the neutrino mass with 40 meV sensitivity.

889 Chapter 4 discusses the first of my contributions mentioned above, which is the
890 development of signal reconstruction techniques for antenna array CRES and an antenna
891 array demonstrator experiment called the FSCD. I discuss the key tools that Project 8
892 uses to simulate antenna array CRES before introducing signal reconstruction algorithms
893 that can be used to detect CRES signals using the array. I end Chapter 4 with a paper
894 that summarizes a detailed analysis and comparison of the signal detection performance
895 of each algorithm.

896 Chapter 5 describes my contributions to the development of antennas and an antenna
897 measurement system for Project 8, which is the second major contribution of this
898 dissertation. I begin with a general overview of basic principle of antennas and antenna
899 measurements, before including a paper that describes the development of unique antenna
900 designed to mimic the cyclotron radiation emitted by electrons in free-space when trapped
901 in a magnetic field. I call this antenna the synthetic cyclotron radiation antenna (SYNCA)
902 and its main purpose is to serve as a fake electron for laboratory validation measurements
903 of Project 8’s antenna array CRES simulations. Chapter 5 ends with an overview of
904 laboratory measurements of a prototype antenna array using the SYNCA, which were
905 compared with simulations to provide upper bounds on reconstruction errors caused by
906 imperfections in real-life measurements.

907 Chapter 6 discusses the cavity approach to CRES, which was adopted as the preferred
908 CRES technology for Phase IV late into my dissertation work. The chapter stars by
909 discussing resonant cavities in general before introducing the operating principles of the
910 cavity approach to CRES. I end the chapter by discussing a study of and open-cavity
911 design that could be used for CRES measurements and integrated with atomic tritium

912 and an electron gun calibration source for the pilot-scale and Phase IV experiments.

913 Finally, in Chapter 7 I conclude by briefly discussing the future directions of the
914 Project 8 collaboration as we continue towards a direct measurement of the neutrino
915 mass.

₉₁₆ **Chapter 2 |**

₉₁₇ **Neutrinos and Neutrino Masses**

₉₁₈ **2.1 Introduction**

₉₁₉ In this chapter I provide a cursory overview of background information relevant to
₉₂₀ neutrinos and neutrino mass measurements.

₉₂₁ In Section 2.2 I provide background information on the history of neutrinos and beta-
₉₂₂ decay. In Section 2.3 I describe the discovery of neutrino oscillations, which demonstrated
₉₂₃ unambiguously that neutrinos have non-zero masses. In Section 2.4 I discuss the current
₉₂₄ state of the theoretical understanding of neutrino masses in the standard model. Lastly,
₉₂₅ in Section 2.5 I discuss a few methods for measuring the absolute scale of the neutrino
₉₂₆ mass.

₉₂₇ **2.2 Neutrinos and Beta-decay**

₉₂₈ Late in the 19th century the phenomena of radioactivity was first observed in experiments
₉₂₉ performed by Henri Becquerel with uranium, and further studied using thorium and
₉₃₀ radium by Marie and Pierre Curie [4, 5]. Early work in radioactivity classified different
₉₃₁ forms of radiation based on it's ability to penetrate different materials. Rutherford was
₉₃₂ the first to separate radioactive emissions into two types, alpha and beta radiation [6].
₉₃₃ Alpha rays were easily stopped by a piece of paper or thin foil of metal, whereas beta
₉₃₄ radiation could penetrate metal several millimeters thick. Later a third form of radiation
₉₃₅ was identified by Villard [7], which was still more penetrating, later termed gamma
₉₃₆ radiation by Rutherford.

₉₃₇ When these forms of radioactivity were first discovered it was unclear what physically
₉₃₈ constituted an alpha, beta, or gamma particle. Experiments with radioactivity in magnetic
₉₃₉ fields were eventually able to identify the charge composition of the different forms of

radiation. In particular, experiments by Becquerel identified [8] that beta radiation had an identical charge-to-mass ratio to the electron. This was strongly suggestive that beta particles were indeed electrons.

Studies of beta radiation lead to the discovery that radioactivity resulted in the transmutation of elements [9] caused by the decay of a heavier nucleus to a lighter species. One feature of beta radiation, which we will refer to now as beta-decay, that differentiated it from alpha and gamma radiation is that the electrons produced by beta-decay have a continuous spectrum of kinetic energies, whereas, alpha and gamma particles are emitted with discrete energies. This feature of beta-decay was first observed by Chadwick in 1914 [10], and was extremely puzzling at the time, since the continuous spectrum apparently violates energy conservation [11].

Famously, in 1930 Pauli proposed the existence of a new neutral particle, which he termed the "neutron", that was also produced during beta-decay to resolve the missing energy problem posed by the beta-decay spectrum [12]. Because this particle carried no charge, it was hypothesized that it had simply not been observed in any previous experiments. This "neutron", which was initially estimated to have a mass no larger than that of an electron, was eventually renamed the "neutrino" by Fermi [13] after the discovery of the neutron by Chadwick in 1932 [14]. Later, in 1933, Fermi developed a quantum mechanical theory for beta-decay in which an electron and neutrino are produced by the decay of a neutron to a proton inside the radioactive nucleus [15].

Little more than a speculation when first introduced, indirect evidence for the existence of neutrinos was obtained in 1938 by the simultaneous observation of the electron and recoiling nucleus in cloud chambers by Crane and Halpern [16]. However, it wasn't until the Cowan-Reines experiment [17] in 1956 that direct evidence for the existence of neutrinos was observed through the observation of inverse beta-decays caused by neutrinos from a nuclear reactor interacting with protons contained in water molecules. The difficulty in detecting neutrinos is caused by their weak interactions with other particles. Later experiments revealed the existence of different types or flavors of neutrinos based on the nature of the leptons produced in neutrino charged-current interactions [18], but the existence of a neutrino mass remained an open question that would take more than 40 years to resolve.

971 2.3 Neutrino Oscillations

972 One of the first clues that neutrino flavor transitions or neutrino oscillations were occurring
 973 was the solar neutrino problem. The solar neutrino problem is a discrepancy between
 974 the measured and predicted flux of ν_e from the sum. The solar neutrino problem was
 975 famously observed by Ray Davis Jr. and collaborators in the 1960's [19] at the Homestake
 976 mine in South Dakota. In the early 2000's, the SNO experiment was able to resolve the
 977 solar neutrino problem by identifying neutrino oscillations as the cause of the observed
 978 deficit [20]. Furthermore, measurements of the atmospheric flux of neutrinos by the
 979 Super-Kamiokande experiment and others revealed that fewer muon-type neutrinos
 980 survived passage through the earth than expected providing strong evidence for neutrino
 981 oscillations for both flavors [21].

982 Neutrino oscillations occur because the weakly-interacting neutrino eigenstates are
 983 distinct from the mass eigenstates [22]. The neutrino mass eigenstates represent physical
 984 particles in that they are solutions to the free-particle Hamiltonian, whereas, the neutrino
 985 weak eigenstates correspond to the neutrino states that interact via the weak charged-
 986 current interaction. The neutrino weak eigenstates are a linear superposition of the
 987 neutrino mass eigenstates

$$988 \nu_\ell = \sum_i U_{\ell i} \nu_i, \quad (2.1)$$

988 where $\ell = e, \mu, \tau$ and $i = 1, 2, 3$. The matrix elements $U_{\ell i}$ are the elements of the
 989 Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix that describes the mixing between
 990 the neutrino flavor and mass states.

991 A standard parameterization [23] of the PMNS matrix is

$$992 U_{PMNS} = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix} \\ = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.2) \\ \times \begin{bmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

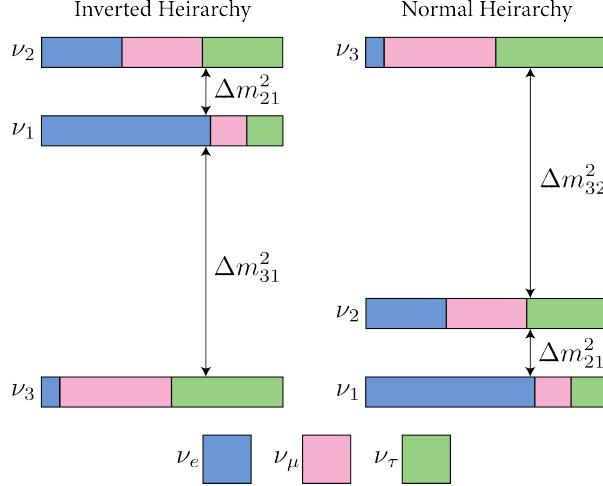


Figure 2.1. A diagram of two different neutrino mass ordering scenarios. In the inverted hierarchy (inverted mass ordering) the lightest neutrino mass is m_3 , whereas, in the normal hierarchy (normal mass ordering) m_1 is the lightest neutrino. What cannot be measured by neutrino oscillations is the neutrino absolute mass scale, which is essentially the mass of the lightest neutrino mass eigenstate.

where $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$. The parameters α_1 and α_2 are only included in the PNMS matrix if neutrinos are Majorana particles, something which represents a current area of research in neutrino physics. The phase δ quantifies the degree of CP-violation in the neutrino sector. Including the Majorana phases the PMNS matrix contains six independent parameters. Neutrino oscillation probabilities also depend on the squared mass differences between neutrino mass eigenstates

$$\Delta m_{ij}^2 = m_i^2 - m_j^2, \quad (2.3)$$

where $ij = 12, 32, 31$ respectively. Because $\Delta m_{32}^2 = \Delta m_{31}^2 - \Delta m_{21}^2$, this adds an additional two parameters that must be constrained by neutrino oscillations.

A large experimental effort over the past couple decades has greatly contained the majority of parameters in the PMNS matrix, many to relative uncertainties of only a few percent. However, certain ambiguities remain, which is the origin of the current uncertainty in the ordering of the neutrino masses (see Figure 2.1). The neutrino masses can be arranged by their relative masses. Current neutrino oscillation data supports that $m_2 > m_1$, however, the sign of Δm_{32}^2 is still unknown. Therefore, two mass-ordering scenarios are allowed, one where neutrino masses are arranged $m_3 > m_2 > m_1$, which is called the normal mass ordering (NMO), or alternatively neutrino masses may be ordered $m_2 > m_1 > m_3$, which is called the inverted mass ordering (IMO). Next-

1009 generation neutrino oscillation experiments such as JUNO [24], Hyper-Kamiokande [25],
 1010 and DUNE [26] are poised to resolve this ambiguity in the coming years.

1011 Neutrino oscillation probabilities are sensitive to the neutrino masses via the squared
 1012 mass differences. Therefore, oscillation probabilities are unaffected by the absolute scale
 1013 of the neutrino mass. However, oscillations can be used to obtain a lower bound on the
 1014 neutrino masses by setting the mass of the lightest neutrino mass state to zero. This
 1015 results in different lower limits depending on the ordering of the neutrino mass states.
 1016 Current best-fit values [23] with 1σ -uncertainties for the squared mass differences are

$$\Delta m_{21}^2 = (7.42^{+0.21}_{-0.20}) \times 10^{-5} \text{ eV}^2, \quad (2.4)$$

$$\Delta m_{31}^2 = (2.5176^{+0.026}_{-0.028}) \times 10^{-3} \text{ eV}^2 \text{ (NMO)}, \quad (2.5)$$

1017 for the normal mass ordering, and for the inverted ordering the limit is

$$\Delta m_{32}^2 = (-2.498^{+0.028}_{-0.028}) \times 10^{-3} \text{ eV}^2 \text{ (IMO).} \quad (2.6)$$

1018 The parameter Δm_{21}^2 is the same in the NMO and the IMO. Allowing the lightest neutrino
 1019 mass in each ordering scenario (m_{least}) to take on a range of values one can visualize the
 1020 relative masses of the neutrinos as a function of m_{least} (see Figure 2.2). The absolute
 1021 neutrino mass scale is effectively the value of this m_{least} parameter.

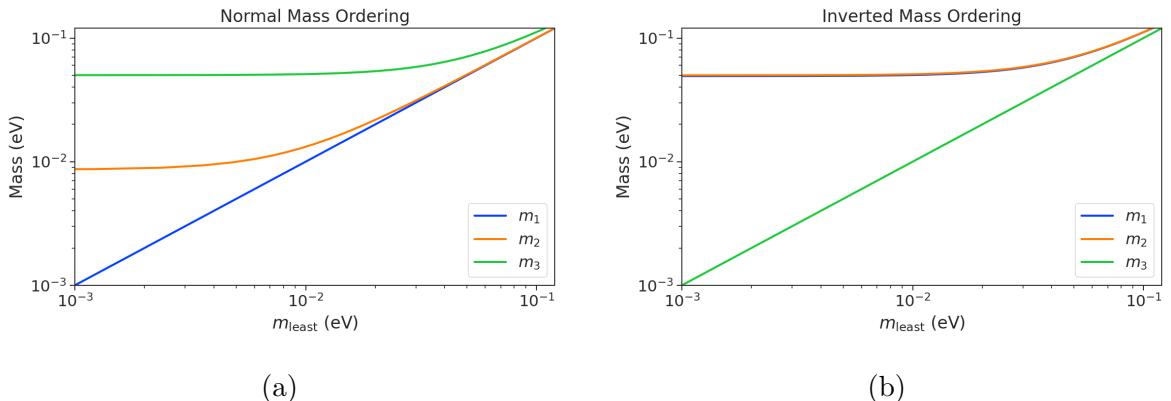


Figure 2.2. The masses of the neutrinos as a function of the lightest neutrino mass in both the normal (a) and inverted (b) mass ordering regimes.

2.4 Neutrino Masses in the Standard Model

In this section, I briefly summarize the current theoretical understanding of neutrino masses in the standard model [27–29]. Neutrinos are spin 1/2 particles, which are described using the Dirac equation.

$$(i\hbar\gamma^\mu\partial_\mu - mc)\psi(x) = 0, \quad (2.7)$$

where the field that describes the particle is denoted as $\psi(x)$. In the standard model fermions acquire mass through the Yukawa interaction, which add to the standard model Lagrangian terms of the form

$$\mathcal{L}_{\text{Yukawa}} = -Y_{ij}^\ell \bar{L}_{Li} \phi E_{Rj} + \text{h.c.}, \quad (2.8)$$

where Y_{ij}^ℓ is an element of the 3×3 Yukawa coupling matrix for leptons, L_{Li} is the left-handed lepton doublet for generation i , ϕ is the Higgs doublet, and E_{Rj} is the right-handed lepton field for generation j . Neutrinos are represented only as left-handed neutrinos and right-handed antineutrinos in the standard model, which is consistent with experimental observations. Since there are no right-handed neutrino singlet fields, there are no Yukawa interaction terms, thus neutrinos in the standard model are strictly massless. Therefore, non-zero neutrino mass is evidence for physics beyond the standard model.

For the charged leptons, the Yukawa interaction leads to masses of the form

$$m_{ij}^\ell = Y_{ij}^\ell \frac{v}{\sqrt{2}}, \quad (2.9)$$

where v is the Higgs vacuum expectation value. The observation of massive neutrinos motivates the extension of the standard model to explain the origin of neutrino masses, which can be approached in different ways, but all approaches add additional degrees of freedom to the standard model.

One approach is to introduce to the standard model a right-handed neutrino field that allows one to include Yukawa terms of the form

$$\mathcal{L}_{\nu \text{Yukawa}} = -Y_{ij}^\ell \bar{L}_{Li} \phi \nu_{Rj} + \text{h.c.} \quad (2.10)$$

where ν_{Rj} is the right-handed neutrino singlet. Because experimental evidence strongly

1045 predicts only three active neutrinos, these additional neutrinos are sterile and do not in-
1046 teract via the strong, weak, or electromagnetic interactions. After spontaneous symmetry
1047 breaking, the Yukawa interaction leads to mass terms given by

$$\mathcal{L}_D = -M_{Dij}\bar{\nu}_{Ri}\nu_{Lj} + \text{h.c.}, \quad (2.11)$$

1048 which is called a Dirac mass term. One of the issues with constructing neutrino masses
1049 in this way is that the required Yukawa couplings are at least a factor of 10^6 smaller than
1050 that of an electron, which begs the question: why are the Yukawa couplings so small for
1051 the neutrinos?

1052 An alternative approach is to allow the neutrinos to have a Majorana mass, which is
1053 possible because neutrinos are electrically neutral particles. The Majorana mass terms
1054 for neutrinos have the form

$$\mathcal{L}_M = -\frac{1}{2}(M_{Rij}\bar{\nu}_{Ri}\nu_{Rj}^c M_{Lij}\bar{\nu}_{Li}\nu_{Lj}^c) + \text{h.c.}, \quad (2.12)$$

1055 where M_{Rij} and M_{Lij} are right-handed and left-handed Majorana mass matrices. A
1056 consequence of neutrinos being Majorana particles is lepton number violation, which
1057 predicts the occurrence of neutrino-less double beta-decay at a rate proportional to the
1058 neutrino mass.

1059 In the most general case neutrinos have both Dirac and Majorana mass terms, which
1060 allows one to generate neutrino masses with Yukawa couplings similar to the rest of the
1061 standard model. Considering a single generation of neutrinos for demonstration, the
1062 combined neutrino mass Lagrangian can be written as

$$\mathcal{L}_{D+M} = -m_D\bar{\nu}_R\nu_L - \frac{1}{2}(m_L\bar{\nu}_L\nu_L^c + m_R\bar{\nu}_R\nu_R^c) + \text{h.c.}, \quad (2.13)$$

1063 or equivalently,

$$\mathcal{L}_{D+M} = -\frac{1}{2} \begin{bmatrix} \bar{\nu}_L & \bar{\nu}_R^c \end{bmatrix} \begin{bmatrix} m_L & m_D \\ m_D & m_R \end{bmatrix} \begin{bmatrix} \nu_L^c \\ \nu_R \end{bmatrix} + \text{h.c..} \quad (2.14)$$

1064 An example mass generation mechanism with this approach is the Type-I see-saw
1065 mechanism [30], in which we take $m_L = 0$ and $m_R \gg m_D$. By diagonalizing Equation
1066 2.14 one obtains the mass eigenvalues that represent the physical masses of the neutrinos.
1067 The light neutrino mass eigenstate, which represents the observed neutrino mass, has a
1068 mass given by

$$m_1 \approx \frac{m_D^2}{m_R}, \quad (2.15)$$

1069 and the heavy neutrino mass eigenstate, which represents the unobserved sterile neutrino,
1070 has a mass

$$m_2 \approx m_R. \quad (2.16)$$

1071 For m_D similar to the other quark or lepton masses, one obtains physical neutrino masses
1072 consistent with observations from sterile neutrino masses of $m_R \approx O(10^{15})$ GeV. This
1073 mass scale is well beyond the capabilities of modern particle accelerators to probe.

1074 2.5 Neutrino Absolute Mass Scale

1075 The neutrino absolute mass scale or simply "neutrino mass" cannot be probed with
1076 neutrino oscillations, since oscillation probabilities are determined by the squared mass
1077 differences between neutrino mass eigenstates, therefore, alternative techniques are needed
1078 to perform an effective measurement of the neutrino mass.

1079 2.5.1 Limits from Cosmology

1080 The Λ CDM model summarizes the current cosmological understanding of the universe [23].
1081 Λ CDM predicts that the universe originated from a single expansion event colloquially
1082 called the "Big Bang". During the Big Bang, the universe originated as a hot spacetime
1083 singularity, which abruptly experienced rapid expansion in a process known as inflation.
1084 After expansion the inflationary field eventually decayed into a population of quarks,
1085 gluons, leptons, and photons, which were kept in thermal equilibrium by the high-
1086 temperatures of the early universe.

1087 As the universe continued to expand it's density and temperature decreased until
1088 the formation of neutral atoms, primarily hydrogen, was possible. At which point the
1089 population of photons produced during the Big Bang decoupled from the primordial
1090 universe and began to freely propagate. A direct prediction of the Λ CDM model is that
1091 this population of photons is still present, but with a significantly reduced temperature
1092 due to the subsequent expansion of the universe. This is consistent with the observation of
1093 the CMB (cosmic microwave background), which is a population of microwave radiation
1094 with a blackbody temperature of 2.7 K. The CMB is extremely uniform in all directions
1095 with slight anisotropies that can be analyzed to study the evolution of the early universe.
1096 A series of experiments have measured the CMB with increasing levels of precision, which
1097 has lead to a significant increase in our current understanding of cosmology.

1098 In addition to the CMB, inflation predicts the existence of a $C\nu B$ (cosmic neutrino

background) [31], which are the remnant neutrinos produced during the Big Bang. Since neutrinos only interact via the weak force, they decouple from the Big Bang plasma at an earlier time than the CMB photons. The temperature at which the C ν B decouples depends on the neutrino rest mass. Neutrinos play a unique role in the Λ CDM model, due to the fact that neutrinos act as radiation early in the universe but as matter in the late universe. This leads to specific signatures that impact the expected anisotropies of the CMB as well as the distribution of matter in the universe [32]. By combining measurements of the CMB with measurements of the large-scale structure (LSS) of the universe one can constrain the neutrino mass scale by fitting these datasets with the Λ CDM model. This analysis results in some of the most stringent constraints on the neutrino mass. Recent analyses [23] have been able to constrain the neutrino mass scale to

$$\Sigma_{m_\nu} \equiv \sum_i m_i < 0.11 \text{ eV}, \quad (2.17)$$

where m_i are the neutrino mass eigenstates.

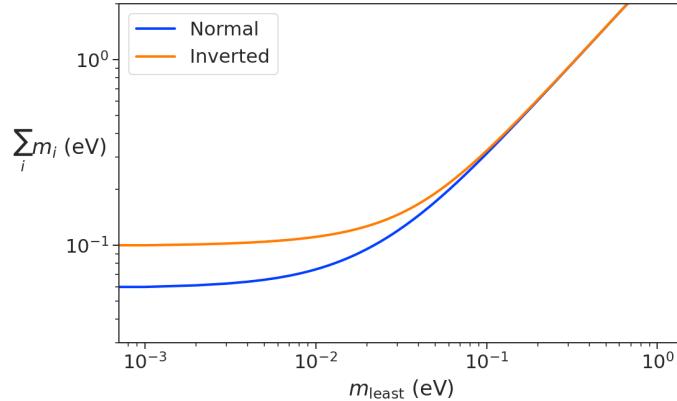


Figure 2.3. The neutrino mass observable measured by cosmology as a function of the lightest neutrino mass eigenstate.

The observable Σ_{m_ν} constrains the neutrino mass by setting the mass of the lightest neutrino mass eigenstate (m_{least}) (see Figure 2.3). In the normal mass ordering Σ_{m_ν} can be rewritten in the form

$$\Sigma_{m_\nu} = m_{\text{least}} + \sqrt{\Delta m_{21}^2 + m_{\text{least}}^2} + \sqrt{\Delta m_{32}^2 + m_{\text{least}}^2}, \quad (2.18)$$

where it is clear that a measurement of Σ_{m_ν} effectively sets the neutrino mass scale

1116 through m_{least} . The analogous formula for the inverted mass ordering is

$$\Sigma_{m_\nu} = m_{\text{least}} + \sqrt{-\Delta m_{32}^2 + m_{\text{least}}^2} + \sqrt{-\Delta m_{31}^2 + m_{\text{least}}^2}. \quad (2.19)$$

1117 Upcoming experiments [33] are planned to refine measurements of the CMB, LSS,
 1118 and other cosmological observables. With this additional data it is possible that in the
 1119 near future cosmological measurements will be able to positively constrain the neutrino
 1120 absolute mass scale. However, the strength of these limits strictly depend on the accuracy
 1121 of the Λ CDM model, which highlights the need for direct experimental measurements of
 1122 the neutrino mass to confirm the predictions of cosmology and to fix the neutrino mass
 1123 parameter in future cosmological analyses.

1124 2.5.2 Limits from Neutrinoless Double Beta-decay Searches

1125 If neutrinos are Majorana fermions, then the neutrino is equivalent to its own antiparticle
 1126 and lepton conservation is not an exact law of nature [34]. Limits on the rate of
 1127 neutrinoless double beta-decay ($0\nu\beta\beta$), are some of the most powerful current tests of
 1128 lepton number conservation [23]. If $0\nu\beta\beta$ were observed, it would direct evidence that
 1129 neutrinos are Majorana fermions and provide a method for measuring the neutrino mass
 1130 scale.

1131 Standard double beta-decay occurs when two neutrons in an unstable nucleus spon-
 1132 taneously decay into two protons, which results in the production of two electrons and
 1133 two neutrinos (see Figure 2.4). Whereas, during $0\nu\beta\beta$ the two neutrinos self-annihilate

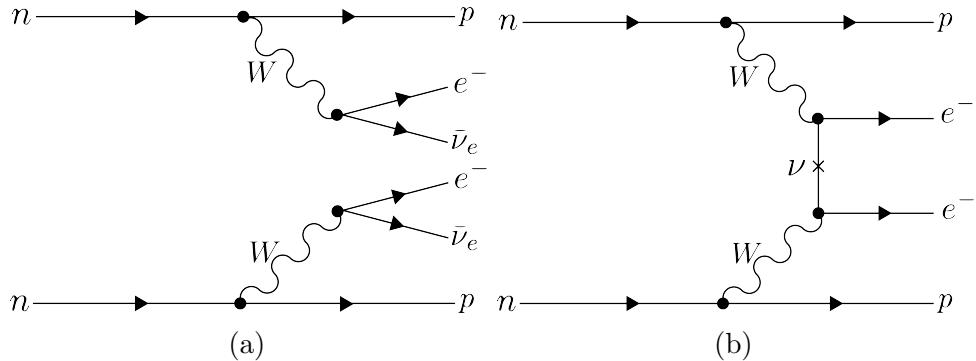


Figure 2.4. Feynman diagrams for double beta-decay (a) and $0\nu\beta\beta$ (b).

1133

1134 producing only two electrons, which violates lepton number by two.

1135 Assuming that the exchange of two Majorana neutrinos is the dominant channel for
 1136 $0\nu\beta\beta$, then a measurement of the $0\nu\beta\beta$ half-life for a particular isotope can be used to

1137 set the neutrino absolute mass scale [35]. The half-life is written in terms of the effective
 1138 neutrino mass for $0\nu\beta\beta$ ($m_{\beta\beta}$) using the equation

$$T_{1/2}^{0\nu} = \frac{1}{G|\mathcal{M}|^2 m_{\beta\beta}^2}, \quad (2.20)$$

1139 where G is the phase-space factor for the decay and \mathcal{M} is the relevant nuclear matrix
 1140 element. $m_{\beta\beta}$ is given by an incoherent sum of the neutrino mass eigenstates weighted
 1141 by the PMNS mixing matrix parameters,

$$m_{\beta\beta} = \left| \sum_i U_{ei}^2 m_i \right|. \quad (2.21)$$

1142 The information provided from $0\nu\beta\beta$ on the neutrino mass scale can be visualized by
 1143 expressing the value of $m_{\beta\beta}$ in terms of m_{least} and two relative Majorana phases [1]. The
 1144 allowed regions for $m_{\beta\beta}$ as a function of m_{least} are shown in Figure 2.5 as the regions
 1145 bounded by the black curves overlayed with the discovery probabilities of future $0\nu\beta\beta$
 decay experiments based on current neutrino data.

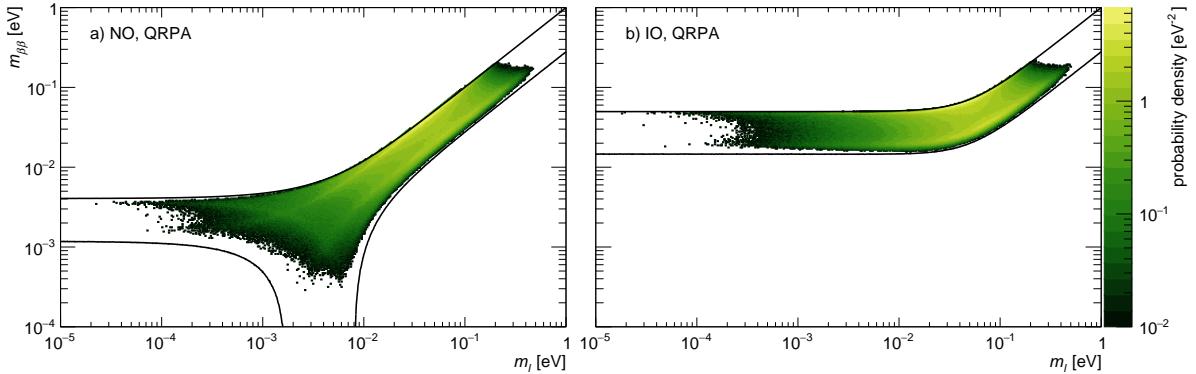


Figure 2.5. The discovery probabilities for the future generation of $0\nu\beta\beta$ experiments as a function of $m_{\beta\beta}$ and m_{least} . Figure from [1].

1146
 1147 Because of the possibility of cancellation due to the unknown Majorana phases included
 1148 in the sum specified by Equation 2.21, the neutrino mass information gained from $0\nu\beta\beta$
 1149 is necessarily imperfect. Additionally, theoretical uncertainties in the calculation of the
 1150 nuclear matrix elements complicates the calculation of $m_{\beta\beta}$ from a measurement of $0\nu\beta\beta$
 1151 half-life. Similar to cosmology, there is a high degree of complementarity between direct
 1152 measurements of the neutrino mass and $0\nu\beta\beta$. In particular, a measurement of m_{least} to
 1153 less than 0.1 eV sensitivity provides significant information for $0\nu\beta\beta$ searches based on
 1154 the discovery probabilities displayed in Figure 2.5.

2.5.3 Limits from Beta-decay

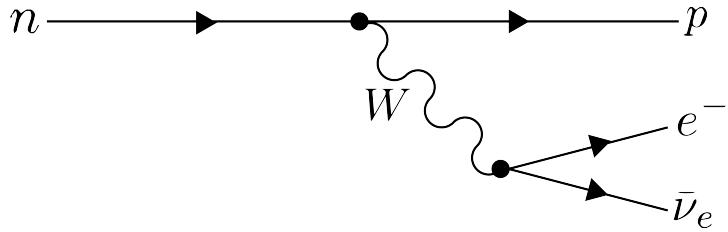


Figure 2.6. A Feynman diagram of beta decay

Certain processes involving neutrinos, in particular beta-decay (see Figure 2.6), have initial states with well-defined total energies and final states that can be measured with high accuracy and precision. Beta-decay involves the decay of an unstable isotope where a neutron spontaneously converts to a proton and emits an electron and anti-neutrino ("neutrino" for brevity) to conserve charge and lepton number [4]. Therefore, by applying the principles of energy and momentum conservation, a measurement of the kinematics of the final state can be used to constrain the neutrino mass [36].

Using beta-decay to measure the neutrino mass can be tied back to Fermi's original 1934 theory of nuclear beta-decay [15] (see Figure 2.7). Because the constraints on the

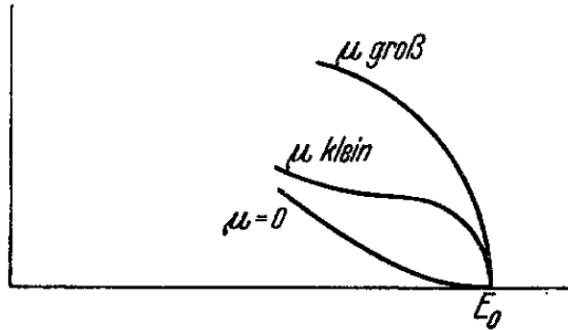


Figure 2.7. A figure from Fermi's 1934 paper on a theory of beta-decay depicting the kinetic energy spectrum of the emitted electron. The effect of the neutrino mass, written as μ , is to distort the shape of the spectrum near the endpoint from the zero-mass spectrum.

neutrino mass from beta-decay depend only on the final state measurement capabilities and the principles of energy and momentum conservation, neutrino mass measurements with beta-decay are called direct measurements. A direct measurement like beta-decay contrasts with other neutrino mass measurements approaches that are model-dependent such as cosmology and $0\nu\beta\beta$, which provide complementary ways to study the physics of massive neutrinos.

1171 The isotope of choice for direct neutrino mass measurements with beta-decay has
 1172 been tritium (3H_2) for many decades, because it conveniently fulfills many experimental
 1173 requirements. Of upmost importance is a decay with a low Q-value, which is the available
 1174 kinetic energy based on the mass difference between the initial and final states. The
 1175 effect of a massive neutrino on the shape of the spectrum is magnified for low Q-values
 1176 and tritium has an unusually low Q-value of 18.6 keV.

1177 Additionally, tritium beta-decay is super-allowed, which results in a relatively short
 1178 half-life of 12.3 years. Therefore, high source activity can be obtained with a relatively
 1179 small source mass. High-activity is desirable because of the low-activity near the tritium
 1180 spectrum endpoint. For tritium beta-decays, only a factor of 3×10^{-13} of the decays
 1181 occur in the last 1 eV of the spectrum. Isotopes with Q-values lower than tritium are
 1182 known [36], but this is outweighed by exceedingly long half-lives leading to unobtainable
 1183 source masses.

1184 The endpoint measurement approach involves quantifying the effect of the neutrino's
 1185 mass on shape of the electron's kinetic energy spectrum near the endpoint. The shape of
 1186 the kinetic energy spectrum (see Figure 2.8) is given by

$$\frac{d\Gamma}{dE} = \frac{G_F^2 |V_{ud}|^2}{2\pi^3} (G_V^2 + 3G_A^2) F(Z, \beta) \beta (E + m_e)^2 (E_0 - E) \\ \times \sum_{i=1,2,3} |U_{ei}|^2 [(E_0 - E)^2 - m_i^2]^{1/2} \Theta(E_0 - E - m_i), \quad (2.22)$$

1187 where G_F is the Fermi coupling constant, V_{ud} is an element of the CKM matrix, E
 1188 is the kinetic energy of the electron, β is the velocity of the electron divided by the
 1189 speed of light, E_0 is the endpoint energy assuming zero neutrino mass, $F(Z, \beta)$ is the
 1190 Fermi function, and $\Theta(E_0 - E - m_i)$ is the Heaviside function, which enforces energy
 1191 conservation. One can see that the decay spectrum is actually a combination of three
 1192 spectra with different endpoints based on the values of the neutrino mass eigenstates, m_i .
 1193 This produces "kinks" in the spectrum shape due to overlapping spectra with different
 1194 endpoint values, but such an effect would be nearly impossible to resolve given the finite
 1195 energy resolution of a real experiment.

1196 The neutrino mass scale variable measured by beta-decay is given by

$$m_\beta^2 = \sum_i |U_{ei}|^2 m_i^2, \quad (2.23)$$

1197 where m_β is the electron-weighted neutrino mass or simply "neutrino mass" for brevity.

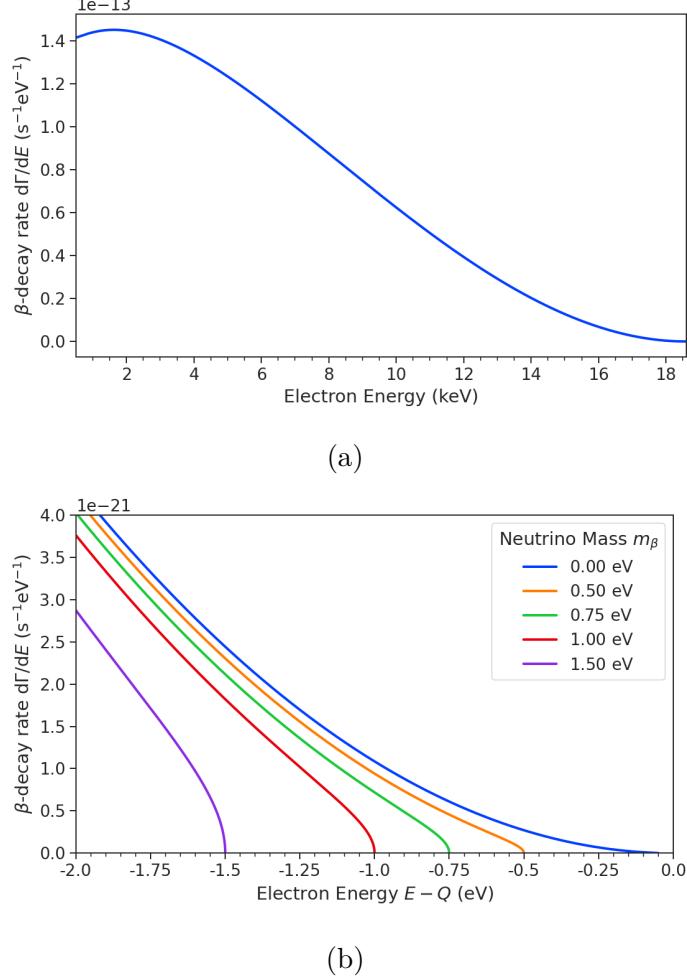


Figure 2.8. The tritium beta-decay spectrum. The effect of a massive neutrino on the spectrum is to change its shape near the endpoint by an amount proportional to the size of the neutrino mass. A sufficiently high-statistic and high-resolution measurement of the spectrum endpoint would be able to measure the neutrino mass.

1198 m_β corresponds to a particular weighted sum of the neutrino masses, which is distinct
 1199 from effective neutrino masses such as $m_{\beta\beta}$ [36]. Assuming unitarity, the neutrino mass
 1200 can be expressed in terms of the PMNS matrix elements, squared mass differences, and
 1201 the lightest neutrino mass eigenstate. For the normal mass ordering the equation is

$$m_\beta^2 = m_{\text{least}}^2 + |U_{e2}|^2 \Delta m_{21}^2 + |U_{e3}|^2 \Delta m_{31}^2, \quad (2.24)$$

1202 and for the inverted ordering the equation is

$$m_\beta^2 = m_{\text{least}}^2 + |U_{e1}|^2 (-\Delta m_{32}^2 - \Delta m_{21}^2) + |U_{e2}|^2 (-\Delta m_{32}^2). \quad (2.25)$$

1203 Therefore, a measurement of the neutrino mass in combination with neutrino mixing
1204 parameters is effectively a measurement of m_{least} .

1205 Since the neutrino mass is small (< 1 eV), it's effect on the spectrum is limited to
1206 the endpoint region. The affect of a non-zero neutrino mass on the endpoint spectrum is
1207 plotted for the reader in Figure 2.8. Resolving the small changes in the spectrum shape
1208 requires an experimental technique with high statistics, excellent energy resolution, and
1209 low background activity.

1210 **Chapter 3 |**

1211 **Direct Measurement of the Neutrino Mass**

1212 **with Project 8**

1213 **3.1 Introduction**

1214 A promising technique for direct measurements of the neutrino mass beyond the projected
1215 200 meV limit of the KATRIN experiment [37] is tritium beta-decay spectroscopy with
1216 an atomic tritium source [38]. Atomic tritium, combined with a large-volume, high-
1217 resolution energy measurement technique, is capable of measuring the neutrino mass
1218 with sensitivity below the 50 meV, which exhausts the range of neutrino masses allowed
1219 under the inverted hierarchy.

1220 Cyclotron Radiation Emission Spectroscopy (CRES) is a high-resolution energy
1221 measurement technique compatible with atomic tritium production and storage that can
1222 enable the next-generation of neutrino mass direct measurement experiments [39]. The
1223 Project 8 collaboration is currently engaged in a program of research and development
1224 (R&D) aimed at developing the technology necessary for a 40 meV sensitivity measurement
1225 of the neutrino mass using CRES and atomic tritium [40].

1226 In Section 3.2 I provide an introduction to the basics of the CRES technique as well as
1227 the goals of the Project 8 experiment. Additionally, I sketch out the phased experiment
1228 development plan being implemented by Project 8 to build towards a next-generation
1229 neutrino mass experiment.

1230 In Section 3.3 I give an overview of Phase II of the Project 8 experiment [41,42], which
1231 completed early in 2023. Although the bulk of the work presented in this dissertation is
1232 relevant to designs of future Project 8 experiments, a description of the work in Phase II
1233 provides useful context.

1234 In Section 3.4 I introduce a CRES measurement concept based on antenna arrays [43],
1235 which could be the basis for the ultimate Project 8 neutrino mass experiment. A

1236 significant portion of the R&D efforts of Project 8 in Phase III were directed towards
1237 simulating and modeling this experimental concept in order to understand the achievable
1238 sensitivity to the neutrino mass.

1239 Lastly, in Section 3.5 I introduce conceptual designs of pilot-scale experiments and
1240 Phase IV that combine atomic CRES with a large-volume CRES detection technique.
1241 This includes a design concept for an antenna array based experiment, but also a design
1242 for a resonant cavity based experiment. Resonant cavities are discussed in more depth in
1243 Chapter 6 and have become the default choice for the Phase IV experiment.

1244 **3.2 Project 8 and Cyclotron Radiation Emission Spec- 1245 troscopy**

1246 **3.2.1 Cyclotron Radiation Emission Spectroscopy — CRES**

1247 Time and frequency are two of the most precisely measured quantities in physics. Atomic
1248 clocks, which operate by measuring the frequencies of various atomic transitions, have
1249 been used to measure time with astounding relative uncertainties of 10^{-18} seconds [44].
1250 The extreme precision possible with frequency measurements is often summarized using
1251 the a quote from the Physicist Arthur Schawlow who said advise his students to "Never
1252 measure anything but frequency!" [45].

1253 Neutrino mass measurements using tritium beta-decay require the measurement of
1254 perturbations to the 18600 eV tritium endpoint with a precision as small as 0.1 eV,
1255 therefore, a spectroscopic technique with extremely high resolution is required. Frequency
1256 measurements are capable of such high-resolutions for the intuitive reason that they are
1257 essentially digital counting measurements, which average the number of oscillations of a
1258 physical system over time. By observing a rapidly oscillating system over a sufficient
1259 length of time one can obtain essentially arbitrary precision on a frequency limited only
1260 by the measurement time and signal-to-noise ratio (SNR) of the system.

1261 A method is required for translating an electron kinetic energy measurement into a
1262 frequency measurement. A straightforward way to accomplish this is to place a gaseous
1263 supply of tritium into a magnetic field, therefore, when a beta-decay occurs the resulting
1264 electron will immediately begin to orbit around a magnetic field line at the cyclotron
1265 frequency, proportional to its kinetic energy (see Figure 3.1). The acceleration caused
1266 by the orbit leads to the emission of cyclotron radiation that can be detected using an
1267 array of antennas or resonant cavity. The starting frequency of the radiation gives the

1268 electron's initial kinetic energy, which is used to build the beta-decay spectrum and
 1269 measure the neutrino mass. The name for this measurement technique is Cyclotron
 1270 Radiation Emission Spectroscopy or CRES [39].

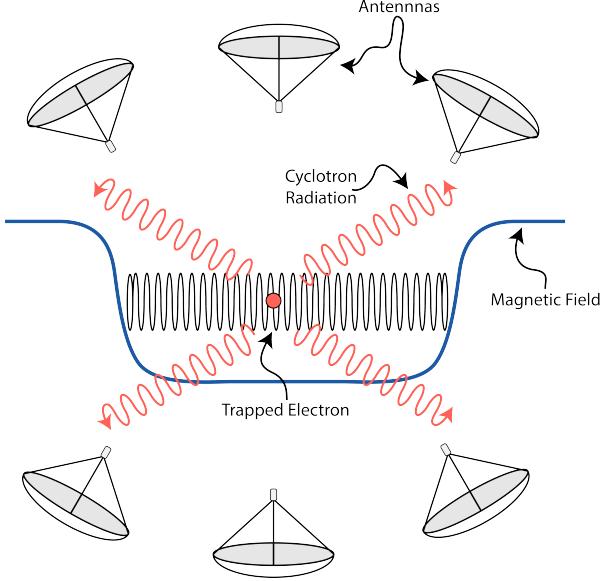


Figure 3.1. A cartoon illustration of the CRES technique. An electron is contained in a magnetic trap, which is a local minimum in the magnetic field, so that its cyclotron radiation can be detected by an array of antennas. Detecting the cyclotron radiation allows us to measure its cyclotron frequency and determine its kinetic energy.

1271 In the non-relativistic case, the cyclotron frequency is simply a function of the
 1272 charge-to-mass ratio of the particle, however, the relativistic correction to the cyclotron
 1273 frequency

$$f_c = \frac{qB}{2\pi m_e \gamma} = \frac{1}{2\pi} \frac{qB}{m_e + E_{\text{kin}}/c^2}, \quad (3.1)$$

1274 introduces a dependence of the kinetic energy (E_{kin}) to the inverse of the cyclotron
 1275 frequency (f_c). Electrons with kinetic energies of 18.6 keV are in the weakly relativistic
 1276 regime with $\beta = \frac{v}{c} = 0.263$ and $\gamma = 1.036$.

1277 The frequency resolution of a CRES measurement can be estimated by differentiating
 1278 Equation 3.1,

$$\frac{df_c}{dE_{\text{kin}}} = \frac{1}{2\pi} \frac{-qBc^2}{(m_e c^2 + E_{\text{kin}})^2}, \quad (3.2)$$

1279 from which we obtain the relationship between fractional differences in energy and
 1280 frequency,

$$\frac{df_c}{f_c} = \frac{1 - \gamma}{\gamma} \frac{dE_{\text{kin}}}{E_{\text{kin}}}. \quad (3.3)$$

1281 Therefore, an energy precision of 1 eV for an 18.6 keV electron requires a frequency
 1282 precision of approximately 2 ppm.

1283 The minimum observation time required to achieve this resolution can be estimated
 1284 using the uncertainty principle as formulated by Gabor [46]. Electrons from tritium
 1285 beta-decay experience random collisions with the background gas particles, which limits
 1286 the uninterrupted radiation lifetime. The time between collision events, referred to as
 1287 "track length", is an exponentially distributed variable. Differences in the track lengths
 1288 of a population of mono-energetic electrons leads to an uncertainty or broadening in the
 1289 distribution of measured frequencies, which is proportional to the mean track length, τ_λ .
 1290 The resulting frequency distribution has a Lorentzian profile, whose width is given by
 1291 the Gabor limit,

$$\tau_\lambda \Delta f_c = \frac{1}{2\pi} \implies \Delta f_c = \frac{1}{2\pi\tau_\lambda}. \quad (3.4)$$

1292 The cyclotron frequency for a 18.6-keV electron in a 1 T field is approximately
 1293 27 GHz, consequently, the minimum observation time for a 2 ppm frequency resolution
 1294 is approximately 3 μ sec. The Gabor limit is not the true lower bound on the frequency
 1295 resolution for a CRES signal, since it derives from the Fourier representation of a fixed
 1296 length time-series using a basis of infinite duration sinusoids. If one takes the approach of
 1297 fitting the CRES signal in the time-domain, then the lower limit on frequency precision
 1298 is given by the Cramér-Rao lower bound (CRLB) [47], which depends on the track length
 1299 and SNR. In general, the CRLB allows for better precision on the cyclotron frequency.

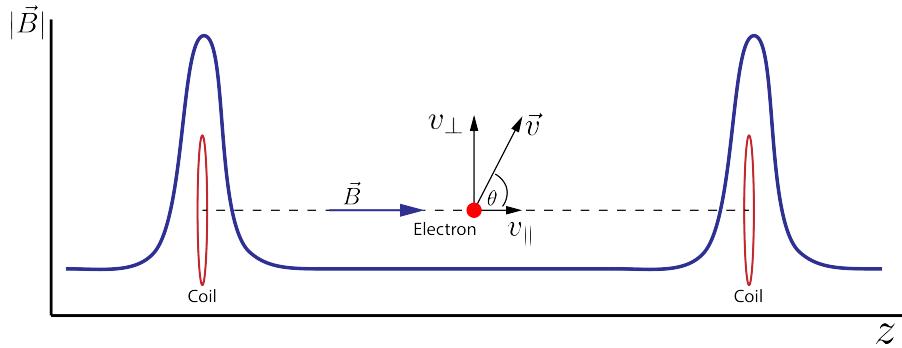


Figure 3.2. An illustration of an electron in a bathtub magnetic trap generated by two well-separated coils.

1300 Ensuring that an electron remains under observation long enough so that it's frequency
 1301 can be precisely measured requires a magnetic trap. A magnetic trap is a local minimum
 1302 in a background magnetic field generated an appropriate configuration of electromagnetic
 1303 coils. Since magnetic fields can do no work, there is no danger of the magnetic trap

affecting the kinetic energy electron after it is emitted from the beta-decay. One common approach to creating a magnetic trap is the "bathtub" trap configuration, which can be produced using two magnetic pinch coils aligned on a central axis that are separated by a distance that is large compared to the coil radius (see Figure 3.2). This configuration produces a trap with a uniform bottom and relatively steep walls, which is ideal for CRES measurements.

Electrons produced in the trap oscillate back and forth between the trap walls at a frequency that depends upon the pitch angle, unless they are produced with pitch angles too small to be contained in the trap. Pitch angle is defined as the angle between the component of the electron's velocity perpendicular to the magnetic field and the component parallel to the magnetic field

$$\tan \theta = \frac{v_{\perp}}{v_{\parallel}}. \quad (3.5)$$

The axial motion of the electron leads to variation in the cyclotron frequency caused by the changing value of the magnetic field. This leads to frequency modulation that generate sidebands in the cyclotron radiation spectrum. Resolving these sideband frequency components is necessary for a complete reconstruction of the CRES signal in the experiment.

Electrons trapped in a cylindrically symmetric trap have three primary components of motion (see Figure 3.3). The dominant component, typically with the highest frequency,

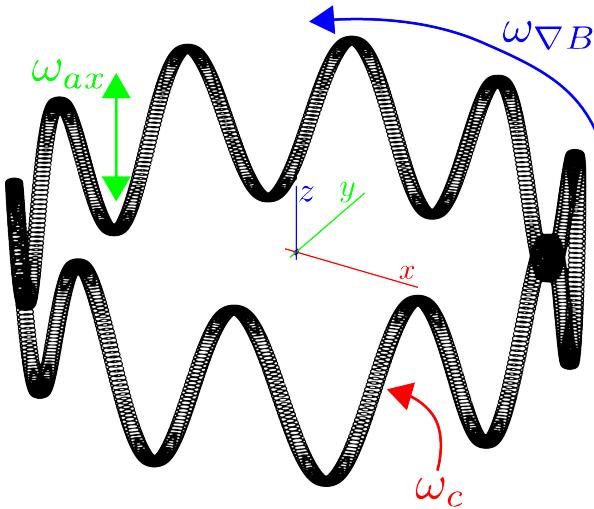


Figure 3.3. A plot of the main components of an electron's trajectory in a cylindrically symmetric trap.

is the electron's cyclotron orbit, which encodes information on the electron's kinetic

energy. Axial motion from the electron's pitch angle leads to frequency modulation, and a shift in the average magnetic field experienced by an electron. This leads to a correlation between the kinetic energy of the electron and the pitch angle depending on the particular shape of the magnetic trap, which can negatively impact energy resolution. To reduce this correlation one must engineer the trap to have a flat bottom with very steep walls, which is more easily achieved with a small aspect ratio bathtub trap. Radial gradients in the trap leads to a third component of motion called grad-B drift [48]. The equation for the drift velocity is

$$\mathbf{v}_{\nabla B} = \frac{m_e v_{\perp}^2}{2qB} \frac{\mathbf{B} \times \nabla B}{B^2}. \quad (3.6)$$

The total power of the radiation emitted by an electron in a free-space environment is given by the Larmor equation [49]

$$P(\gamma, \theta_p) = \frac{1}{4\pi\epsilon_0} \frac{2}{3} \frac{q^2 \omega_c^2}{c} (\gamma^2 - 1) \sin^2 \theta_p, \quad (3.7)$$

where ω_c is the cyclotron frequency multiplied by 2π and θ_p is the pitch angle to distinguish it from the spherical angle coordinate. A single electron with a 90° pitch angle and 18.6 keV of kinetic energy in a 1 T magnetic field emits a total radiation power of 1.2 fW, furthermore, one is typically only able to receive a fraction of this total power with an antenna or other detection system. Therefore, RF systems in CRES experiments must be operated at cryogenic temperatures to limit the noise power such that adequate SNR can be achieved for signal detection and reconstruction. Alternatively, longer tracks enable detection of weaker signals due to the increase in the total signal energy available for the detection algorithm.

3.2.2 The Project 8 Collaboration

The Project 8 collaboration¹ is a group of institutions in the United States and Germany aiming to measure the neutrino mass by developing a novel spectrometer technology based on CRES. In the ultimate Project 8 experiment, the CRES technique will be used to measure the beta-decay spectrum using a large source of atomic tritium sufficient to achieve the required statistics in the last $O(10)$ eV of the decay spectrum. Project 8 is targeting a neutrino mass sensitivity below 50 meV [50], which exhausts the range of possible neutrino masses under the inverted hierarchy and is a factor of four less than

¹<https://www.project8.org/>

1350 sensitivity projections for the ongoing KATRIN experiment.

1351 Project 8's proposed experiment requires the development of two novel technologies:
1352 the production and trapping of a source of atomic tritium on cubic-meter scales and
1353 technology to enable CRES measurements of individual electrons in the same volume.

1354 **Atomic Tritium**

1355 Previous measurements of the tritium beta-decay spectrum for neutrino mass measure-
1356 ments have relied on sources of molecular tritium for their measurements [37, 51, 52] due
1357 to the technical challenges associated with the production and storage of atomic tritium.

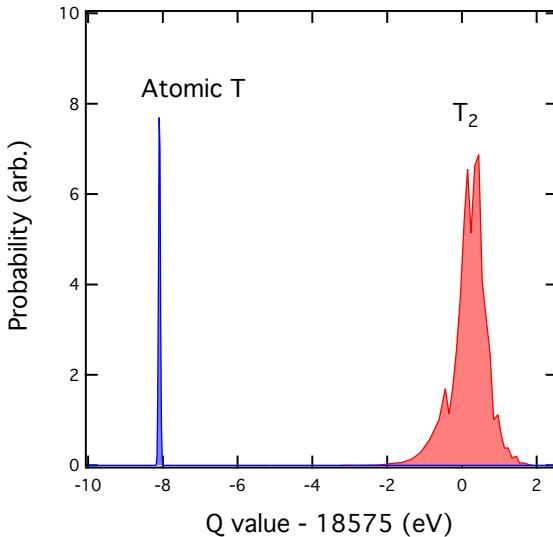


Figure 3.4. A plot of the final state distributions of atomic and molecular tritium. The final state distribution provides the primary contribution to the width of the molecular spectrum whereas thermal doppler broadening is responsible for the width of the atomic spectrum.

1358 One must supply sufficient energy to the tritium molecules to break the molecular
1359 bond and create atomic tritium. Common approaches include the use of hot coaxial
1360 filament atom crackers as well as plasma sources. Both involve heating the tritium atoms
1361 to temperatures of > 2500 K, which must then be cooled to temperatures on the order
1362 of a few mK so that the tritium atoms can be trapped. Cooling the atoms requires the
1363 construction of a large tritium infrastructure and cooling system that can supply a source
1364 of cold atoms to the trap.

1365 Once cold tritium atoms are produced they cannot make contact with any surfaces
1366 to avoid recombination of the atoms to molecules. Therefore, a magnetic trap is required
1367 to store the atoms for a sufficient length of time that they have a chance to decay before

1368 escaping the trap. Trapping the atoms requires the construction of a large and complex
1369 magnet system that must be cooled to cryogenic temperatures.

1370 The significant experimental complexity caused by atomic tritium makes a molecular
1371 source the obvious choice from practical considerations. However, the drawback of
1372 molecular tritium for neutrino mass measurement is the irreducible broadening in the
1373 electron's kinetic energy due to the final state spectrum of molecular tritium (see Figure
1374 3.4). The broadening of the final state spectra has a RMS amplitude of 436 meV [53, 54]
1375 caused by variation in the final vibrational state of the daughter molecule.

1376 For atomic tritium the primary sources of broadening in the final state spectrum are
1377 magnetic hyperfine splittings (magnitude of $O(10^{-5})$ eV) and thermal Doppler broadening
1378 caused by the motion of the trapped atom. Atomic tritium at a temperature of 1 mK
1379 has a broadening which is dominated by thermal Doppler broadening, providing about
1380 1 meV RMS of broadening to the electron's kinetic energy.

1381 The larger energy broadening with molecular tritium leads to an irreducible statistical
1382 uncertainty that limits the achievable sensitivity to approximately 100 meV at 90%
1383 confidence. For previous direct measurements of the neutrino mass this uncertainty is an
1384 insignificant contribution to the overall uncertainty budget, however, for experiments
1385 like Project 8 atomic tritium is a key component to the success of the experiment.

1386 CRES for Neutrino Mass Measurement

1387 Several features of the CRES technique make it an attractive choice for a next generation
1388 neutrino mass measurement experiment. Because CRES is a remote-sensing technique,
1389 it is possible to observe the kinetic energy of the electron without altering its trajectory
1390 or directly interacting with the particle, therefore, in a CRES experiment the source
1391 gas volume can be the same as the CRES spectrometer volume. Tritium gas is also
1392 transparent to cyclotron radiation, which means that the kinetic energies of electrons can
1393 be measured using a cavity or antenna array, located directly outside the atom trapping
1394 volume.

1395 Because source and spectrometer can be colocated, CRES experiments have an
1396 advantageous scaling law relative to the current state-of-the-art beta-decay spectroscopy
1397 experiment, KATRIN. KATRIN utilizes the magnetic adiabatic collimation with an
1398 electrostatic filter (MAC-E filter) technique to measure the beta-decay spectrum of
1399 molecular tritium. In this approach, a source of molecular tritium is located outside the
1400 spectrometer. When a beta-decay occurs the electron is guided out of the tritium source
1401 using a magnetic field and is transported through the MAC-E filter before it is detected

1402 on the other side of the filter using a charge sensor. The measurement statistics of the
1403 MAC-E filter are limited by the transverse area of the tritium source and filter due to the
1404 need to travel through the experiment without scattering. This scaling is less favorable
1405 than the volumetric scaling of CRES due to the ability to colocate source and detector.

1406 Another promising aspect of the CRES technique is the inherently high precision
1407 of frequency based measurements. The endpoint of the molecular tritium beta-decay
1408 spectrum is approximately 18.6 keV, which dwarfs the neutrino mass scale of $< 1 \text{ eV}/c^2$
1409 by at least a factor of 10^5 . Measuring the effect of such a small mass on a high energy
1410 electron requires excellent energy resolution. Since frequency measurements are essentially
1411 counting measurements they are intrinsically quite accurate due to the ability to measure
1412 the cyclotron frequency by effectively averaging over millions of cyclotron orbits. Using
1413 off-the-shelf RF components its is possible to achieve part-per-million accuracy on the
1414 kinetic energy with the CRES technique.

1415 CRES is also nearly immune to typical sources of backgrounds that can plague other
1416 experiments. Since CRES operates via a non-destructive measurement of the electron's
1417 cyclotron frequency, sources of background electrons are effectively filtered out by limiting
1418 the frequency bandwidth of the measurement. The fiducial volume of the experiment is
1419 free from any surfaces that could introduce stray electrons, and electrons from sources
1420 outside the fiducial volume can be prevented from entering the experiment.

1421 Neutrino Mass Sensitivity Goals

1422 Project 8's ultimate goal is to combine CRES with atomic tritium to measure the neutrino
1423 mass with 40 meV sensitivity at the 90% confidence level (see Figure 3.5). This sensitivity
1424 is sufficient to fully exhaust the range of allowable neutrino masses under the inverted
1425 neutrino mass ordering regime and is approximately an order of magnitude less than the
1426 projected final sensitivity of the KATRIN experiment. Excluding the full neutrino mass
1427 parameter space would require a sensitivity an order of magnitude lower than what is
1428 proposed by Project 8, which would require an experiment whose size and complexity
1429 are currently well beyond proposals for the next-generation of neutrino mass direct
1430 measurement experiments.

1431 3.2.3 The Project 8 Phased Development Plan

1432 Reaching 40 meV sensitivity requires the simultaneous development and eventually
1433 combination of CRES and atomic tritium. These technologies require a significant up-front

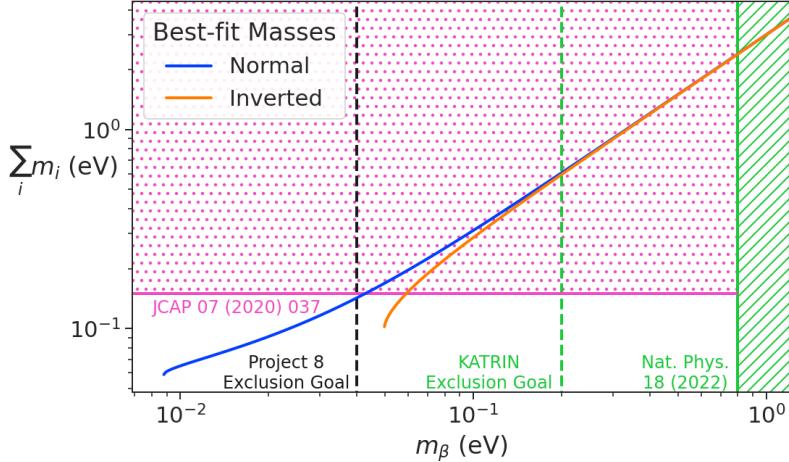


Figure 3.5. Neutrino mass exclusion plot including limits from cosmological measurements and the KATRIN experiment. Allowed ranges for neutrino masses under the normal and inverted hierarchies are shown as the blue and orange lines respectively. The black dashed line shows Project 8’s goal neutrino mass sensitivity for the Phase IV experiment.

1434 R&D investment to build-out the required capabilities for a 40 meV CRES experiment.
 1435 Therefore, Project 8 is following a phased experiment plan in which incremental progress
 1436 can be made towards the ultimate goal of a 40 meV neutrino mass measurement with
 1437 CRES.

1438 Phase I and II: Proof of Principle and First Tritium Measurements

1439 The earlier phases of the Project 8 experiment, Phase I and II, were focused on demon-
 1440 stration and development of the CRES technique itself as well as a proof-of-principle
 1441 measurement of the neutrino mass using the CRES technique.

1442 In Phase I, Project 8 performed a proof-of-principle measurement of the ^{83m}Kr
 1443 spectrum using CRES, which marked the first ever kinetic energy spectrum measurement
 1444 with CRES. The experiment included all the components of a basic CRES experiment.
 1445 An electron source consisting of a gas of ^{83m}Kr was supplied to a waveguide gas cell
 1446 constructed out of a segment of WR-42 waveguide and sealed with Kapton windows at
 1447 the top and bottom. A magnetic trapping region was created in the waveguide cell using
 1448 a single electromagnetic coil wrapped around the waveguide which provided a trapping
 1449 volume on the order of a few cubic-millimeters. Detection of the cyclotron radiation was
 1450 performed by connecting the waveguide cell to an additional segment of waveguide that
 1451 transmitted the radiation to a cryogenic amplifier.

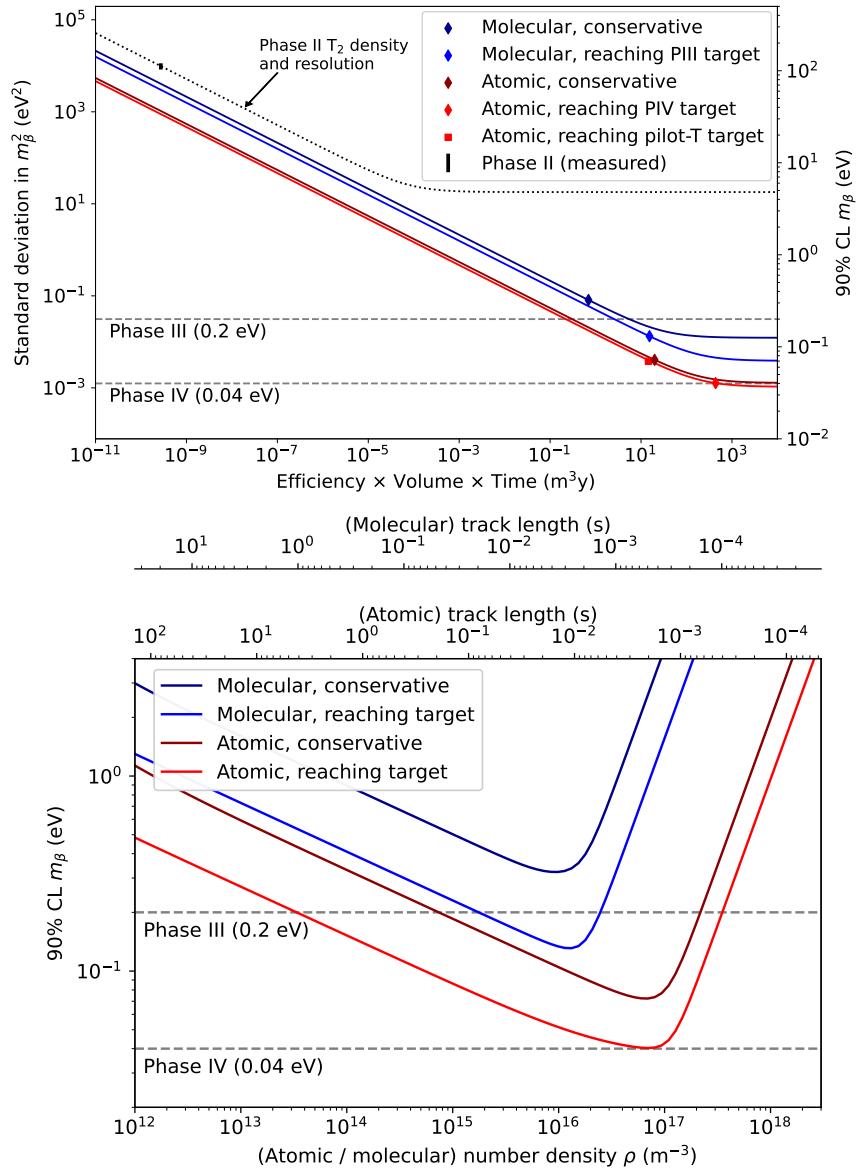


Figure 3.6. Sensitivity calculations for a cavity based CRES experiment that demonstrate the neutrino mass measurement goals of the Project 8 collaboration throughout the phased development plan. The blue curves indicate molecular tritium sources and the red curves indicate atomic tritium sources. In the current plan, Phase III contains two tritium experiments. The first is the Low-frequency Apparatus (LFA), which is a molecular tritium experiment, and the second is the atomic tritium pilot-scale experiment that officially ends Phase III. The sensitivity of these experiments is primarily a function of statistics, however, there is a critical density beyond which CRES electrons do not have enough time to radiate between collisions for a high-resolution frequency measurement leading to worse sensitivity.

1452 Success in Phase I was achieved with the 2014 publication of the measured ^{83m}Kr
1453 conversion spectrum [55], which contains a mono-energetic 17.8-keV as well as several
1454 other conversion lines at higher energies. Publication of this result marked the official
1455 end of Phase I and the start of Phase II, in which Project 8 shifted its focus to the
1456 demonstration of the first tritium beta-decay spectrum using CRES. For more information
1457 on Phase II please see Section 3.3.

1458 **Phase III: Research and Development and a Pilot-scale Experiment**

1459 After completing Phase II, Project 8 has shifted focus towards R&D aimed at the
1460 construction of an experiment that demonstrates all the technologies required for a
1461 40 meV measurement of the neutrino mass. The culmination of Phase III is a pilot-scale
1462 experiment that successfully retires all technological and engineering risks associated
1463 with the Phase IV experiment, while also being a scientifically interesting experiment in
1464 its own right. Sensitivity estimates of the pilot-scale experiment predict a neutrino mass
1465 sensitivity on par with the projected sensitivity of the KATRIN experiment.

1466 Phase III R&D is divided into two main efforts — atomic tritium and CRES detection
1467 techniques. Atomic tritium development in Phase III must retire all risks associated
1468 with the atomic tritium system. This includes the production of tritium atoms, atomic
1469 cooling and recirculation systems, purity and isotope concentration monitoring, and
1470 atom trapping. Currently, Project 8 is operating small scale atom cracking demonstrator
1471 systems to show that atom production at the estimated rates needed for Phase IV is
1472 achievable. Future efforts will continue the current developments on atom production
1473 and expand to include demonstrations of atomic cooling with an evaporative beam line
1474 as well as atom trapping using Halbach magnet arrays.

1475 The need for new CRES detection techniques is driven by the drastic increase in scale
1476 from Phase II to the pilot-scale experiments. The physical volume used for CRES in
1477 Phase II was on the order of a few cubic-centimeters, and achieving Project 8's sensitivity
1478 target of 40 meV requires an experiment volume on the multi-cubic meter scale. Therefore,
1479 the waveguide gas cell CRES detection technique used in Phase II is not a feasible option
1480 for the future of Project 8 due to its inability to scale to the required size.

1481 Two alternative CRES detection techniques have been proposed for the pilot-scale
1482 experiment — antenna arrays and resonant cavities (see Section 3.4 and Chapter 6).
1483 Both approaches have relative advantages and disadvantages, however, the improved
1484 understanding of the antenna array and cavity approaches to CRES in the recent years
1485 has led to cavities being the preferred technology for the pilot-scale experiment and

1486 Phase IV due to the estimated reduced cost and complexity of this approach. Since
1487 a large degree of the work presented in this dissertation is focused specifically on the
1488 development of the antenna array CRES technique as well as the design of demonstrator
1489 experiments, we described the proposed R&D plan for antenna array Section 3.4. A
1490 description of the cavity approach to CRES can be found in Chapter 6.

1491 Cavity CRES R&D consists of a series of demonstrator experiments intended to
1492 demonstrate cavity CRES at a variety of scales and magnetic fields. Radioactive sources
1493 gases include ^{83m}Kr and molecular tritium, as well as electrons produced by an electron-
1494 gun, which is a key calibration tool for future CRES experiments. The near-term cavity
1495 effort in Project 8 is the cavity CRES apparatus (CCA), which is a small-scale cavity
1496 experiment operating near 26 GHz. The CCA will perform the first CRES measurements
1497 using a small cavity, and will pave the way towards larger scale cavity experiments in
1498 preparation for the eventual pilot-scale tritium experiment.

1499 The pilot-scale experiment is the first experiment, which will combine atomic tritium
1500 and large-volume CRES detection in the same experiment. It will directly demonstrate
1501 all the technologies required for Phase IV such that no technical risks remain for scaling
1502 the experiment to required scale. A robust approach to scaling the pilot-scale experiment
1503 is to simply build multiple copies of it for the Phase IV experiment.

1504 **Phase IV: Project 8's Ultimate Neutrino Mass Experiment**

1505 The design of Phase IV should be a direct extension of the pilot-scale CRES experiment
1506 that marks the official end of Phase III (see Section 3.5). The Phase IV experiment
1507 represents the final experiment in the Project 8 neutrino mass measurement experiment
1508 plan and will have sensitivity to neutrino masses of 40 meV.

1509 **3.3 Phase II: First Tritium Beta Decay Spectrum and 1510 Neutrino Mass Measurement with CRES**

1511 In Phase II, Project 8 demonstrated the first ever measurement of the tritium beta-decay
1512 spectrum endpoint using the CRES technique, which lead to the first neutrino mass
1513 measurement by the Project 8 collaboration. This milestone was made possible by many
1514 improvements in the CRES technique and in the understanding of CRES systematics,
1515 which takes an important first step towards larger scale measurements of the tritium
1516 beta-decay spectrum with CRES. In this section, I briefly describe some important

elements of the Phase II experiment, with the goal of contextualizing the research and development efforts for Phases III and IV of Project 8. For more complete descriptions of the work that lead to Project 8’s Phase II results please refer to the relevant publications by the collaboration [41, 42].

3.3.1 The Phase II CRES Apparatus

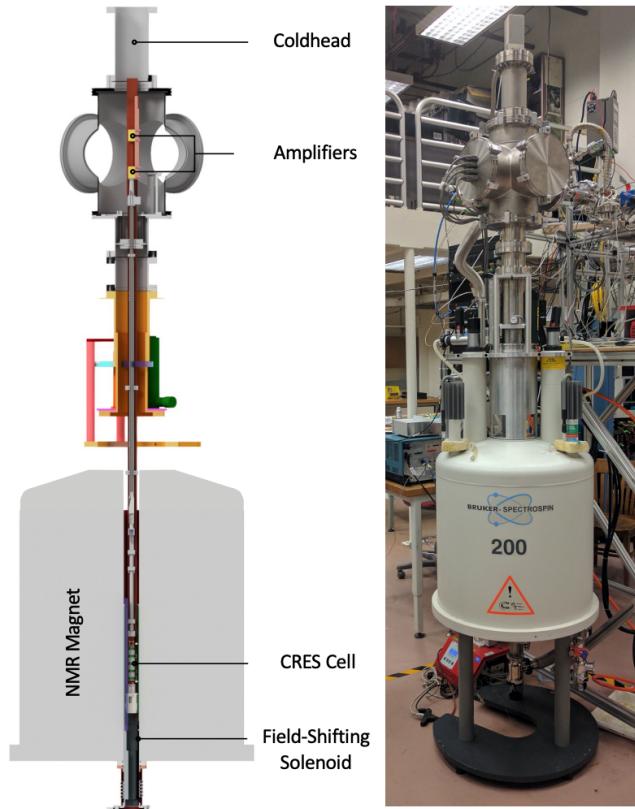


Figure 3.7. The Phase II CRES apparatus used to perform the first measurement of the tritium beta-decay spectrum using CRES.

Magnet and Cryogenics

The magnetic field for the Phase II experiment is provided by a nuclear magnetic resonance (NMR) spectroscopy magnet with a central bore diameter of 52 mm (see Figure 3.7). The magnet produces a background magnetic field with an average value of 0.959 T with a 10 ppm variation across the bore diameter achieved using several shim coils built into the magnet. Using an external NMR field probe, the variation of

1528 the magnetic field along the vertical axis of the magnet bore was measured to obtain
1529 an accurate model of the magnetic field so that the CRES cell could be positioned for
1530 optimal magnetic field uniformity.

1531 An external solenoid magnet was installed inside the magnet bore to provide the
1532 ability to shift the magnitude of the background magnetic field by a few mT. The solenoid
1533 has inside diameter of 46 mm and a length of 350 mm, which terminates in a vacuum
1534 flange that allows it to be inserted into the NMR magnet bore from the bottom. By
1535 shifting the value of the magnetic field by a few mT, the cyclotron frequencies of electrons
1536 produced by the 17.8 keV ^{83m}Kr internal-conversion line [56] can be shifted by frequencies
1537 of ± 100 MHz. This allows one to study the frequency dependent behavior of several
1538 CRES systematics such as detection efficiency that directly affect the measured shape of
1539 the tritium spectrum.

1540 The inside of the magnet bore diameter was pumped down to a vacuum of less than
1541 10 μtorr using a turbomolecular pump, which allows for cryogenic cooling of the CRES
1542 cell and RF system. Cooling power was supplied to the Phase II apparatus using a
1543 cryopump with its coldhead mounted above the primary magnet and CRES cell. This
1544 arrangement allowed for sufficient cooling power to be delivered to the amplifiers to cool
1545 them to a temperature of ≈ 40 K, while keeping the amplifiers far enough from the
1546 magnet so as not to be damaged by the large field strength. Thermal contact between
1547 the coldhead, amplifiers, RF system, and CRES cell is achieved using a copper bar that
1548 runs the full length of the apparatus. To prevent freeze-out of ^{83m}Kr on the walls of the
1549 CRES cell a separate heater was installed to keep the CRES cell near a temperature of
1550 85 K during the operation of the experiment.

1551 CRES Cell

1552 Located in the most uniform region of the magnetic field is the CRES cell, which is
1553 the region of the apparatus where radioactive decays of ^{83m}Kr and T_2 produce electrons
1554 that can be trapped and measured using CRES (see Figure 3.8). The CRES cell is
1555 manufactured from a segment of cylindrical waveguide designed to operate at K-band
1556 frequencies near 26 GHz. The diameter of the waveguide determines which resonant
1557 modes of the waveguide will couple to the electron and transmit its radiation to the
1558 amplifiers. For Phase II a waveguide diameter of 1 cm was selected, which allows electrons
1559 to couple to the TE_{11} and TM_{01} cylindrical waveguide modes. To reduce complexity in
1560 modeling and analyzing the CRES data, it is ideal to select a diameter that prevents
1561 electrons from coupling to higher-order waveguide modes beyond the fundamental TE

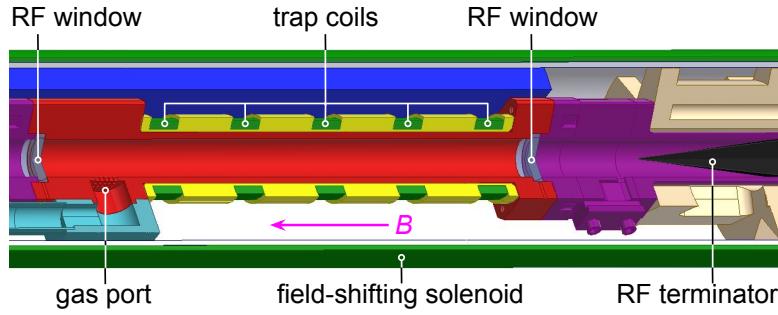


Figure 3.8. Diagram of the CRES cell portion of the Phase II apparatus.

1562 and TM modes.

1563 Around the exterior of the cylindrical waveguide are several magnetic coils used to
 1564 produce magnetic traps inside the CRES cell volume. Without a magnetic trap electrons
 1565 produced from decays inside the CRES cell quickly impact the cell wall, which prevents
 1566 a measurement of their cyclotron frequency using CRES. Each coil along the length of
 1567 the waveguide produces a separate trap that is approximately harmonic in shape. By
 1568 independently controlling the currents provided to each coil the traps can be configured
 1569 to have equal values of the magnetic field at the trap bottom despite a non-uniform field
 1570 from the NMR magnet.

1571 Two primary magnetic trap configurations were used during the Phase II experiment.
 1572 The first was a shallow trap configuration used primarily for it's high energy resolution to
 1573 study systematics using ^{83m}Kr decays, and the second was a deeper trap that could trap a
 1574 higher percentage of pitch angles. The trade-off with this trap is that the higher trapping
 1575 efficiency comes at the cost of lower energy resolution due to the greater variation in pitch
 1576 angle. The deep trap was the trap used to measure the tritium beta-decay spectrum in
 1577 Phase II.

1578 The source gases were delivered into the CRES cell through a gas port located near the
 1579 top end of the cylindrical waveguide. To prevent the gases from escaping the cell, vacuum
 1580 tight RF transparent windows are needed to contain the tritium and krypton source
 1581 gas across a 1 atm pressure differential, while still transmitting the cyclotron radiation
 1582 without distortion. The crystalline material, CaF_2 , which has a thermal expansion
 1583 coefficient similar to that of copper, was used for this purpose in the CRES cell. Two
 1584 windows, each 2.4 mm thick, were used to seal off the ends of the CRES cell. The
 1585 thickness of 2.4 mm corresponds to half of a cyclotron wavelength when one accounts for
 1586 the permittivity of CaF_2 .

1587 **RF System**

1588 The RF system in the Phase II apparatus propagates the cyclotron radiation from the
 1589 CRES cell to the receiver chain. The receiver chain performs the down-conversion and
 1590 digitization required to obtain signals that can be analyzed to determine the cyclotron
 frequencies of electrons in the CRES cell (see Figure 3.9).

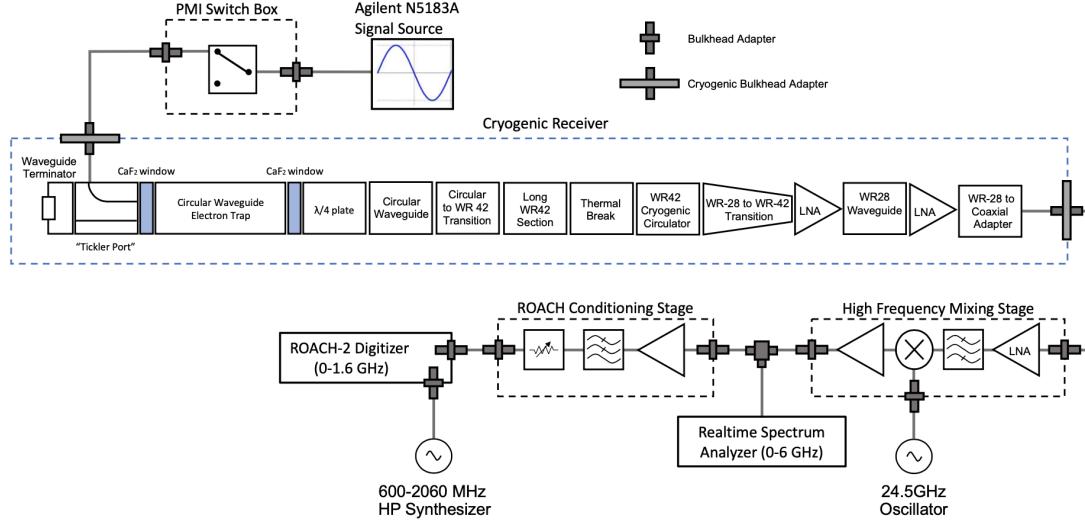


Figure 3.9. RF system diagram for the Phase II apparatus.

1591
 1592 Below the CRES cell, at the bottom of the Phase II apparatus, is a tickler port and
 1593 waveguide terminator. The tickler port is used to inject signals into the CRES cell and
 1594 RF system for testing and calibration purposes. The waveguide terminator is designed to
 1595 absorb cyclotron radiation emitted by electrons that transmits out of the bottom of the
 1596 CRES cell. This lowers the total power received from electrons in the CRES cell, since all
 1597 the energy radiated downwards is absorbed into the terminator. Earlier iterations of the
 1598 Phase II apparatus used an RF short in this location that reflected this power up towards
 1599 the amplifiers, however, interference between the upward traveling and reflected radiation
 1600 led to a disappearance in the signal carrier that made reconstruction impossible.

1601 Radiation traveling upward passes through the CaF_2 window passes through a $\lambda/4$
 1602 plate, which transforms the circularly polarized cyclotron radiation into linear polarization.
 1603 The linearly polarized fields next travel through a segment of circular waveguide that
 1604 transitions into a long segment of WR-42 waveguide that carries the fields out of the
 1605 high magnetic field region. A thermal break segment is included, which consists of a a
 1606 segment of gold-plated stainless steel WR-42 waveguide, to help thermally isolate the
 1607 relatively warm CRES cell from the colder amplifiers. The radiation then passes through

1608 a cryogenic circular, which prevents signals reflected from the amplifiers from interfering
1609 with the CRES cell before a WR-42 to WR-28 transition connects the waveguide to the
1610 first of the cryogenic amplifiers. The radiation passes through two cryogenic amplifiers
1611 before being coupled to a coaxial termination at the top of the Phase II apparatus.

1612 The coaxial cable transfers the cyclotron radiation signals to a high-frequency mixing
1613 stage that performs an analog frequency down-conversion using a 24.5 GHz LO. Two forms
1614 of digitization can be used at this stage to readout the CRES data. One is a real-time
1615 spectrum analyzer that digitizes the CRES signal data in time-domain and computes the
1616 frequency spectrum in real-time, which allows for direct visualization of CRES signal
1617 spectrograms as the experiment is running. The real-time spectrum analyzer is most
1618 useful for taking small amount of streamed data for debugging and analysis of the system.
1619 The other method, which was used to collect the majority of the CRES data in Phase II,
1620 is a ROACH-2 FPGA and digitizer system. The ROACH system consists of a fast ADC
1621 that samples the CRES signal data at 3.2 GSps. Internal digital down-conversion stages
1622 implemented in the FPGA perform a mixing operation that reduces the bandwidth of the
1623 CRES signals to 100 MHz. The FPGA implements a 8192 sample FFT and packetizes
1624 time and frequency domain records in parallel. The packetized data is then transferred
1625 from the ROACH to be analyzed by the data-processing pipeline.

1626 **3.3.2 CRES Track and Event Reconstruction**

1627 **Time-Frequency Spectrogram**

1628 The online data-processing software uses a real-time triggering algorithm that identifies
1629 interesting data that could contain CRES signals. Triggered data are collected into files
1630 that are transferred to a server for offline processing and analysis. The data files contain
1631 a continuous series of time-domain samples, broken into a set of records, which are 4096
1632 samples long. The time-series is made up of 8-bit IQ samples acquired at 100 MHz.

1633 Each time-series record is accompanied by an associated frequency spectrum consisting
1634 of 4096 frequency bins approximately 24.4 kHz wide, which is represented as a power
1635 spectral density. The individual frequency spectra can be organized temporally to create
1636 a time-frequency spectrogram that represents the evolution of the cyclotron frequency
1637 spectrum over the course of the CRES event (see Figure 3.10). The time-frequency
1638 spectrogram is represented as a two-dimensional image where the color of each pixel is
1639 proportional to the power spectral density. Each vertical slice of pixels in the image
1640 represents a frequency spectrum, therefore, each horizontal bin represents the data

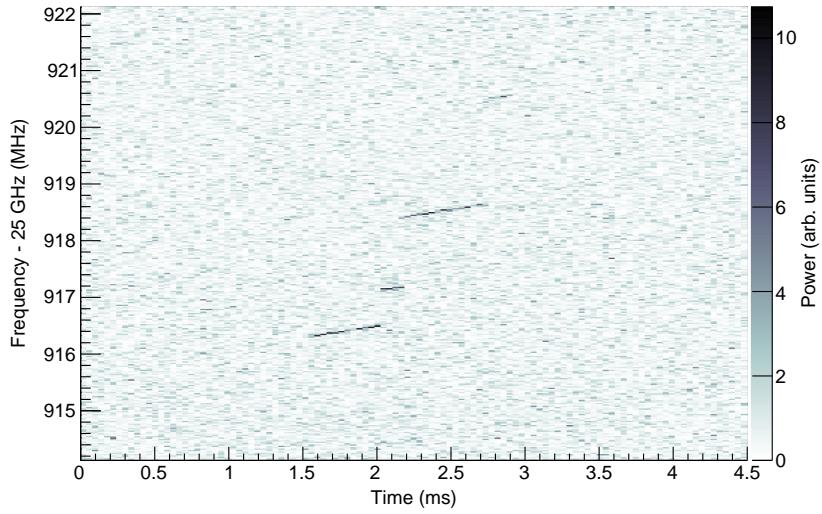


Figure 3.10. The time-frequency spectrogram of a tritium CRES event in the Phase II apparatus.

1641 obtained over a duration of $4096 \times 0.01 \text{ MHz}^{-1} = 40.96 \mu\text{sec}$.

1642 **CRES Event Data Features**

1643 Phenomenologically, a CRES signal appears as a sinusoidal signal whose frequency slow
 1644 increases over time in what is called a frequency "chirp". Axial motion of the electron in
 1645 the trap leads to the formation of frequency sidebands that surround the more powerful
 1646 carrier frequency. The critical piece of information that must be extracted from the track
 1647 and event reconstruction procedure is the carrier frequency, since it is this frequency that
 1648 gives the cyclotron frequency and thus the kinetic energy. Axial motion from non- 90°
 1649 pitch angles changes the average magnetic field experienced by an electron, which leads to
 1650 different cyclotron frequencies being measured for electrons with the same kinetic energy.
 1651 However, because of the low-SNR in Phase II sidebands were unable to be observed,
 1652 so no attempt to directly correct for this effect was attempted in the Phase II analysis.
 1653 The effect of different pitch angles is to broaden the peak of a monoenergetic electron
 1654 line, which can be quantified by measuring the instrumental resolution of the Phase II
 1655 apparatus.

1656 In the time-frequency spectrogram representation, the chirping carrier frequency
 1657 appears as a linear track of high-power frequency bins (see Figure 3.10). The vertical
 1658 slope of the tracks is caused by the emission of energy from the electron in the form of
 1659 cyclotron radiation, therefore, the size of the slope parameter is directly proportional

1660 to the Larmour power. The continuous track is periodically interrupted by random
1661 jumps to higher frequency and lower energy caused by random inelastic collisions with
1662 background gas molecules. The length of a track is an exponentially distributed variable
1663 whose mean value is inversely proportional to the gas density. The size of the frequency
1664 discontinuities is directly proportional to the energies of the rotational and vibrational
1665 states of background gas molecules.

1666 A CRES event refers to the collection of tracks produced by a trapped electron until
1667 it inevitably scatters into a pitch angle that can no longer be trapped. The goal of track
1668 and event reconstruction is to identify the set of tracks in a time-frequency spectrogram
1669 that represents a segment of data acquired in the Phase II apparatus. These tracks must
1670 be clustered into events, from which one can determine the first track produced by the
1671 electron and thus estimate it's starting cyclotron frequency and kinetic energy.

1672 Track Reconstruction

1673 The first step in CRES event reconstruction is the identification of tracks in the time-
1674 frequency spectrogram, which is essentially an image processing task. Track finding
1675 starts by normalizing the power spectral density based on the average noise power. Next
1676 a power threshold is applied to the normalized spectrogram where only bins that have a
1677 SNR ratio greater than five are selected to build tracks. In this case SNR is defined as the
1678 ratio between the normalized, unitless power of a bin divided by the average normalized
1679 power across the full frequency spectrum.

1680 The sparse spectrogram produced by this power cut consists only of a sparse collection
1681 of high-power frequency bins that could be part of a CRES signal track (see Figure
1682 3.11). In this form is it much easier to identify tracks "by eye", however, for the Phase II
1683 analysis Project 8 developed its own custom-made track finding algorithm, called the
1684 sequential track finder (STF).

1685 The STF algorithm processes the sparse spectrogram in sequential fashion, processing
1686 each time-slice one-by-one until the end of the spectrogram is reached. Tracks are found
1687 by searching for points in the sparse spectrogram that appear to fall on a straight line.
1688 Multiple configurable parameters are built into the STF algorithm that allow the user to
1689 tune the criteria for adding a point to an existing track or creating a new track. These
1690 include parameters such as maximum time and frequency differences between subsequent
1691 points in a track as well as minimum SNR values for the start and endpoints of the track.
1692 Additionally, tracks are required to have a minimum length and slope to be considered
1693 potential CRES tracks rather than random noise fluctuations.

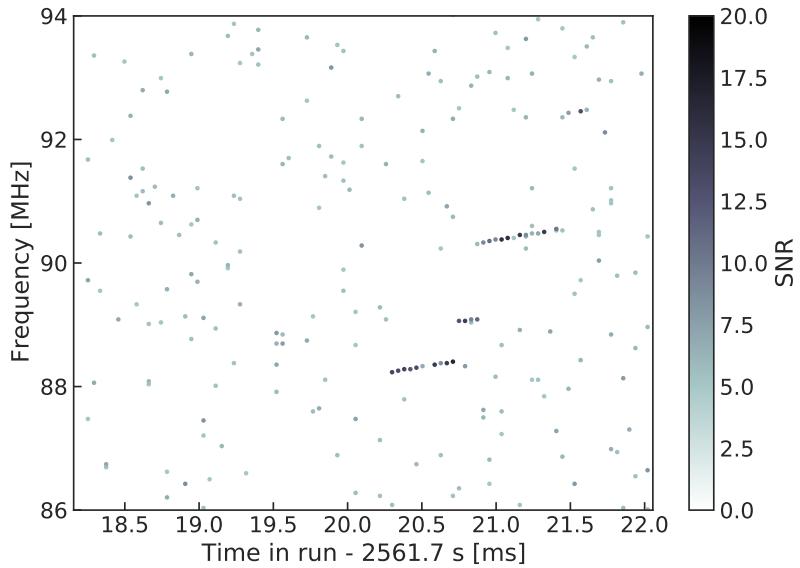


Figure 3.11. The sparse spectrogram obtained by placing a power cut on the raw spectrogram shown in Figure 3.10.

1694 The resulting output of the STF is a collection of track objects that consist of the track
 1695 point objects and their properties. The final step is to calculate track-level properties
 1696 and apply cuts to reject false tracks found by the STF. This involves the fitting of a
 1697 line to the collection of track points as well as the total and average power of the track
 1698 obtained by computing the sum and mean of the points powers. The starting frequency
 1699 of the track is determined by calculating the time coordinate that intersects with the
 1700 linear fit. A cut is performed to remove all tracks that do not have a specified average
 1701 power over their duration, which helps to remove the majority of noise fluctuations that
 1702 have passed all previous cuts up to this point.

1703 Event Reconstruction

1704 After track reconstruction comes event reconstruction where the identified tracks are
 1705 grouped into events that correspond to the trajectory of a single electron in the trap. This
 1706 procedure attempts to match tracks head to tail by checking if the start and end times
 1707 of a pair of tracks falls within a certain tolerance. This tolerance is a configurable
 1708 parameter that can be tuned to an optimal value using Monte Carlo simulations of events
 1709 in the Phase II apparatus.

1710 After the event building procedure has completed there remains a small likelihood
 1711 that false tracks have made it through to the event reconstruction stage. Typically, cuts

at the track level are able to remove 95% of the false tracks identified by the STF, which leads to a significant number of false tracks at the event building stage. However, the additional event-level information makes it possible to reject events that contain these false tracks with a high degree of confidence.

Two event level features are associated with events caused by real electrons — the duration of the first track as well as the number of tracks in the event. Real electrons tend to have event structures with longer first tracks and a higher number of total tracks. Based on the values of these two criteria, a minimum threshold on the average power in the first track was configured to reject false events. The average power in the first track was chosen due to the critical nature of the starting frequency of the first track in an event to the krypton and tritium spectrum analyses.

3.3.3 Results from Phase II

The main result from Phase II was the measurement of the tritium beta-decay spectrum using CRES, which lead to the first neutrino mass limit with CRES. However, Phase II also included a significant ^{83m}Kr measurement campaign to understand important systematics relevant to the tritium spectrum measurement, but also to understanding the fundamentals of the CRES technique itself. This required high-resolution measurements of the ^{83m}Kr internal-conversion spectrum [56], which is an interesting science result in its own right.

The results from Phase II represents a significant effort from the entire Project 8 collaboration over several years. Because the focus of my contributions to Project 8 is directed towards the research and development efforts for the Phase III experiments, the goal in this section is not to provide a detailed description of the analyses that lead to the Phase II results. Rather, I will provide brief descriptions of a few plots representative of the main results from Phase II.

Measurements with Krypton

Measurements with krypton were a key calibration tool for Phase II of the experiment and will continue to be useful in Phase III. In the context of Project 8 krypton measurements refers to CRES measurements of the internal-conversion spectrum of the metastable state of krypton-83, ^{83m}Kr , produced by electron capture decays of ^{83}Rb . A supply of ^{83}Rb was built into the Phase II apparatus gas system that supplied the CRES cell with ^{83m}Kr via emanation.

1744 The ^{83m}Kr internal-conversion spectrum consists of several lines based on the orbital
 1745 of the electron ejected during the decay. The conversion lines useful to Project 8 are
 1746 those that emit electrons with kinetic energies that fall inside the detectable frequency
 1747 bandwidth of the Phase II apparatus. These are the K; L2 and L3; M2 and M3; and N2
 1748 and N3 lines with kinetic energies of 17.8 keV, \approx 30.4 keV, \approx 31.9 keV, and \approx 32.1 keV,
 1749 respectively. The different energies of the lines allow a onw to test the linearity of the
 1750 relationship between kinetic energy and frequency across the range of frequencies covered
 1751 by the continuous tritium spectrum.

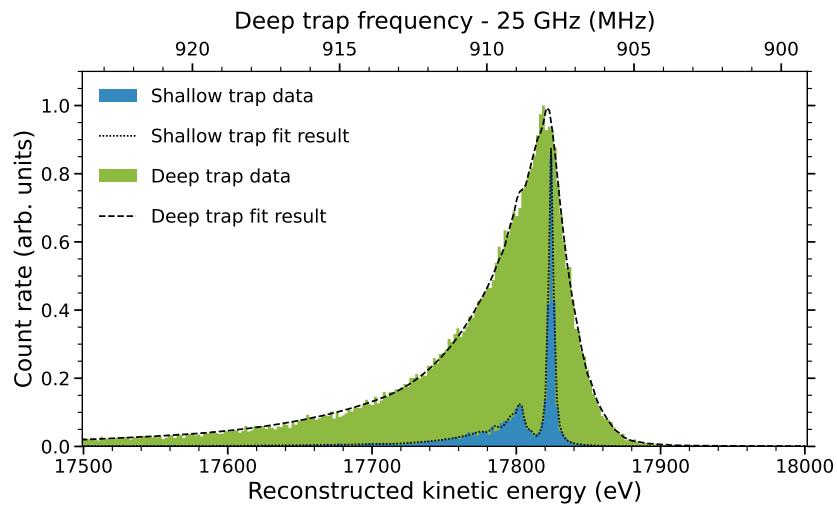


Figure 3.12. Fits to the measured 17.8-keV ^{83m}Kr conversion line using the deep and shallow trap configurations.

1752 Numerous detector related effects relevant to the tritium analysis can be characterized
 1753 by measuring the shape of the krypton spectrum. Specific examples include variations
 1754 in the magnetic field as a function of the radial position of the electron, variation in
 1755 the magnetic field caused by the trap shape, variation in the average magnetic field for
 1756 electrons with different pitch angles, and the effect of missing tracks due to scattering.
 1757 These spectrum shape measurements focused on the 17.8-keV krypton line and utilized
 1758 different trap geometries based on the particular goal of the dataset (see Figure 3.12).

1759 Krypton measurements with a shallow trap allow for high energy resolution, since
 1760 variation in frequency due to pitch angle differences is sharply reduced in the shallow
 1761 trap configuration. With this trap the main 17.8-keV peak of the conversion spectrum is
 1762 clearly visible along with additional satellite peaks at lower energy, which correspond to
 1763 the shakeup/shakeoff spectrum of the decay. The high accuracy of the fit demonstrates a
 1764 high degree of understanding of the CRES systematics.

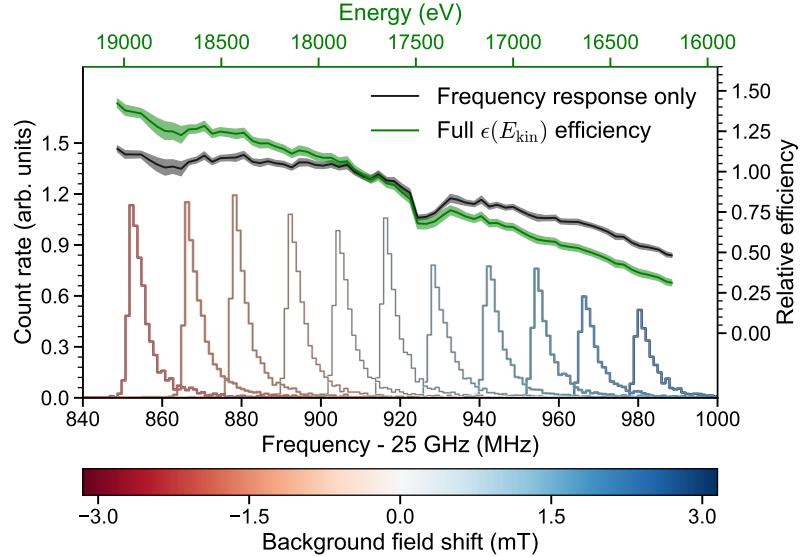


Figure 3.13. Measurements of the 17.8-keV ^{83m}Kr line using the deep trap configuration for different values of the magnetic field from the field shifting solenoid.

The broadening of the krypton spectrum seen for the deeper track is due to the large range of electron pitch angles that can be trapped. Furthermore, with a deeper trap there is a larger parameter space of electron that could be produced with pitch angles that are trappable but not visible in the time-frequency spectrogram. These electrons live in the trap and can scatter multiple times before randomly scattering to a visible pitch angle. This leads to one or more missing tracks earlier in the event, which leads to a misreconstruction of the true starting frequency. By measuring the krypton spectrum shape in the same trap used to detect tritium events, the effect this has on the spectrum shape can be characterized to mitigate its impact on the tritium measurements.

Changes in the Krypton spectrum shape as a function of CRES frequency were used to study the detection efficiency of the Phase II apparatus. Variations in the detection efficiency as a function of frequency directly influences the measured shape of the continuous tritium spectrum, which can lead to errors in the neutrino mass estimate if not modeled appropriately. Using the field shifting solenoid the cyclotron frequency of the krypton 17.83 keV line was shifted across the full frequency range of the tritium spectrum data (see Figure 3.13). Variations in the deep trap krypton spectrum shape can be used to infer the detection efficiency as a function of frequency and correct for this affect in the tritium measurements.

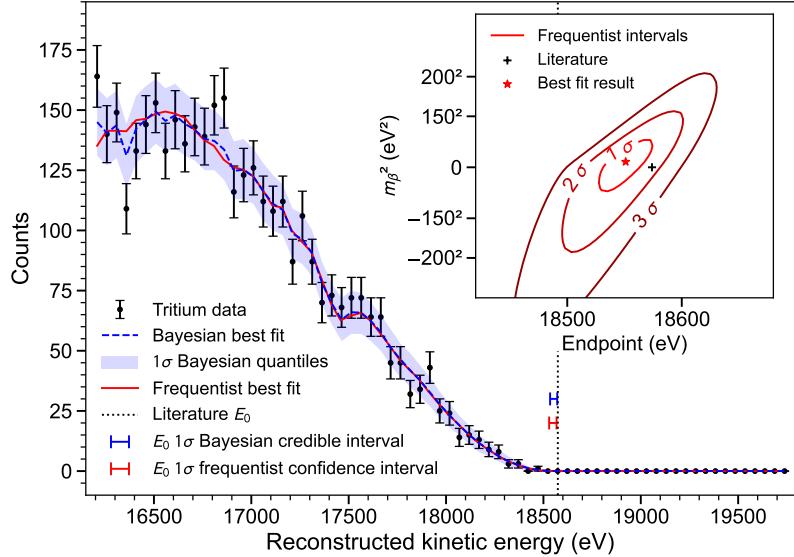


Figure 3.14. The measured tritium spectrum from Phase II with Bayesian and frequentist fits.

1783 Tritium Spectrum and Neutrino Mass Results

1784 The tritium measurement campaign resulted in the collection of 82 days of detector
 1785 live time during which 3770 total tritium events were detected. The track and event
 1786 reconstruction analysis extracted the starting frequencies of these tritium events, which
 1787 were used to build a frequency spectrum of tritium beta-decays. The resulting frequency
 1788 spectrum was then converted to an energy spectrum using the information gleaned from
 1789 the krypton measurement campaign to obtain the tritium beta-decay spectrum (see
 1790 Figure 3.14).

1791 CRES is inherently a very low background technique with the dominant source of
 1792 noise being random RF fluctuations. Monte Carlo simulations backed validated using
 1793 measurements of the RF noise background were used to set track and event cuts to
 1794 guarantee that zero false events would occur over the duration of the experiment with
 1795 90% confidence. Notably, the measured spectrum has zero events beyond the tritium
 1796 spectrum endpoint, which allows us to constrain the background rate in the Phase II
 1797 apparatus to less than 3×10^{-10} counts/ev/s. Achieving a low background is critical for
 1798 future neutrino mass experiments that seek to measure the neutrino mass with less than
 1799 100 meV sensitivity.

1800 Bayesian and frequentist based fits to the measured tritium spectrum, incorporating
 1801 information gained about CRES systematics from the krypton measurements, were
 1802 performed to extract upper limits on the tritium beta-decay spectrum endpoint as well as

1803 the neutrino mass. The estimated spectrum endpoints are 18553^{+18}_{-19} eV for the Bayesian
1804 analysis and 18548^{+19}_{-19} eV for the frequentist analysis. The quoted uncertainties are
1805 $1-\sigma$, and both results are within $2-\sigma$ of the literature endpoint value of 15574 eV. The
1806 estimated neutrino mass for both results is consistent with $m_\beta^2 = 0$. The 90% confidence
1807 upper limits for the Bayesian analysis is $m_\beta < 155$ eV/c² and $m_\beta < 152$ eV/c for the
1808 frequentist analysis.

1809 Though the neutrino mass results from Phase II are not competitive with KATRIN
1810 the experiment was a promising first step towards the development of more precise
1811 neutrino mass measurements using CRES. The low-background and high-resolution
1812 achievable with krypton measurements are promising features of the technique that were
1813 demonstrated with the Phase II apparatus. As new technologies are developed to enable
1814 CRES measurements in larger volume, many of the lessons learned from Phase II will
1815 continue to influence the operation and design of future experiments.

1816 **3.4 Phase III R&D: Antenna Array CRES**

1817 The goal of Phase III in the Project 8 experimental program is to develop the technologies
1818 and expertise required to build an experiment that uses CRES to measure the neutrino
1819 mass with a target sensitivity of 40 meV. One of the key technologies is a method for
1820 performing high resolution CRES measurements in a large volume, which allows one to
1821 observe a sufficient quantity of tritium to measure the low-activity endpoint region of
1822 the tritium spectrum.

1823 **3.4.1 The Basic Approach**

1824 One possible approach, suggested in the original CRES publication [39], is to use many
1825 antennas to surround a volume of tritium gas in a magnetic field (see Figure 3.15). When
1826 a decay occurs the electron will emit cyclotron radiation that can be collected by the array
1827 and used to perform CRES. Each antenna in the array collects only a small fraction of
1828 the electron's signal power, which is less than 1 fW for a 18.6 keV kinetic energy electron
1829 in a 1 T magnetic field. Scaling to large volumes with the antenna array approach is
1830 accomplished by increasing the number of antennas in the array, which increases the
1831 volume under observation proportionally.

1832 Several features of the antenna array approach make it an attractive candidate technol-
1833 ogy for a large volume experiment. One example is the accurate position reconstruction

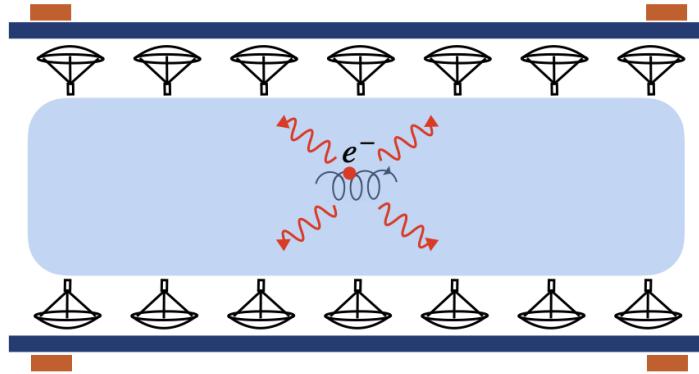


Figure 3.15. A cartoon illustration of the basics of the antenna array CRES technique.

1834 possible with a multichannel antenna array. Using techniques like digital beamforming,
 1835 it is possible to estimate the radial and azimuthal positions of the electron in the mag-
 1836 netic trap with a precision significantly less than the size of the cyclotron wavelength.
 1837 This capability allows one to perform event-by-event estimations of the magnetic fields
 1838 experienced by an electron, which helps achieve high energy resolution with the CRES
 1839 technique.

1840 The easy availability of position information with the antennas array approach
 1841 is potentially a unique advantage that provides significant flexibility in the magnetic
 1842 field uniformity requirements compared to other proposed approaches to large volume
 1843 CRES (see Chapter 6). Spatial discrimination using digital beamforming leads to pileup
 1844 reduction, which helps to reduce the potential of background events caused by missing
 1845 tracks or by incorrectly clustering a group of tracks into an event. Limits on the
 1846 background rate for a neutrino mass measurement with 40 meV sensitivity are stringent
 1847 and the total activity of the tritium source is gigantic relative to the activity near the
 1848 endpoint. Thus, pileup discrimination could be an important tool for a large scale CRES
 1849 experiment.

1850 Another beneficial quality of antenna arrays is that the volume of the experiment can
 1851 be scaled independent of frequency by simply adding more antennas to the array (see
 1852 Figure 3.19). Resonant cavities, the proposed alternative large volume CRES technology,
 1853 are ideally operated in magnetic fields that cause electrons to move with cyclotron
 1854 frequencies near the fundamental cavity resonance, to avoid complex coupling of the
 1855 electron to multiple cavity modes simultaneously. This leads to a coupling between the
 1856 cavity volume and the magnetic field magnitude, which forces one to lower the magnetic
 1857 field in order to increase the experiment scale. Whereas, for antenna arrays, in principle
 1858 there is no physical limitation on the size of the antenna array that can be used at a

particular magnetic field. However, this approach to scaling an antenna array experiment leads to rapidly increasing cost and complexity due to the large number of antennas, amplifiers, and data streams which require substantial computer processing power to effectively utilize.

3.4.2 The FSCD: Free-space CRES Demonstrator

The complexity of the antenna array CRES technique requires the construction of a small scale demonstration experiment to develop an understanding of technique itself and relevant systematics. Without a demonstrator experiment it is not possible to sufficiently retire the technical risks associated with the full-scale experiment. Therefore, Phase III of the Project 8 experimental program is primarily focused on the development and operation of demonstrator experiments to inform the design of the Phase IV experiment.

The Phase III demonstrator experiment for antenna array CRES is called the Free-space CRES Demonstrator or FSCD. The FSCD is also a capable neutrino mass measurement experiment in its own right, with a target neutrino mass sensitivity of a few eV using a molecular tritium source.

Magnetic Field

The background magnetic field for the FSCD is provided by a hospital-grade MRI magnet (see Figure 3.16). The magnet produces a magnetic field of approximately 0.958 T, which corresponds to a tritium spectrum endpoint frequency of approximately 25.86 GHz. The magnet is installed in the Project 8 laboratory located at the University of Washington, Seattle, and is shimmed to produce a uniform magnetic field with variations on the ppm-level. Measurements of the magnetic field non-uniformities are performed using a NMR probe and rotational gantry to capture measurements of the magnetic field around an elliptical surface in the center of the MRI magnet. During the operation of the FSCD an array of Hall or NMR magnetometers would be used to periodically measure the magnetic field to monitor its time stability.

Inside the field of the MRI magnet additional electromagnets would be installed that provide the capability to shift the value of the background magnetic field and produce a magnetic trap. Shifting the background magnetic field by a few μ T lets one control the cyclotron frequencies of electrons with a fixed kinetic energy, which is key to an effective calibration of the FSCD. The preferred calibration method for the FSCD is a mono-energetic electron gun that can inject electrons into the magnetic trap with a



Figure 3.16. An image of the MRI magnet installed in the Project 8 laboratory at the University of Washington, Seattle.

known kinetic energy. In combination with the field shifting magnet, one can vary the cyclotron frequencies of the electrons to measure the response of the antenna array as a function of the radiation frequency and electron position. This procedure characterizes the response of the antenna array and provides further information on magnetic field uniformity, which is important to achieving good energy resolution.

The design of the magnetic trap is absolutely critical to the success of a CRES experiment. The ideal shape is the perfect magnetic box, which has a flat bottom and step function walls. Any variation in the average magnetic field experienced by an electron leads to changes in the cyclotron frequency that can make determining the true starting kinetic energy more difficult. This includes changes in the magnetic field caused by the walls of the magnetic trap as well as radial magnetic field variations.

The ideal box trap is completely uniform and has infinitely steep walls that cause no change in the electron's cyclotron frequency as it is reflected from the trap wall, however, such a trap cannot be made from any combination of magnetic coils since it violates Maxwell's equations. One of the goals of magnetic trap design is to identify the configuration of coils that produces a trap that approximates the perfect box trap as closely as possible.

1908 **Antenna Array**

1909 The canonical antenna array design for CRES is a uniform cylindrical array of antennas
1910 that surrounds the magnetic trap volume. Since the FSCD is a demonstrator experiment,

1911 the antenna array design is the simplest form of the uniform cylindrical array, which is a single circular ring of antennas with a diameter of 20 cm (see Figure 3.17). Along this

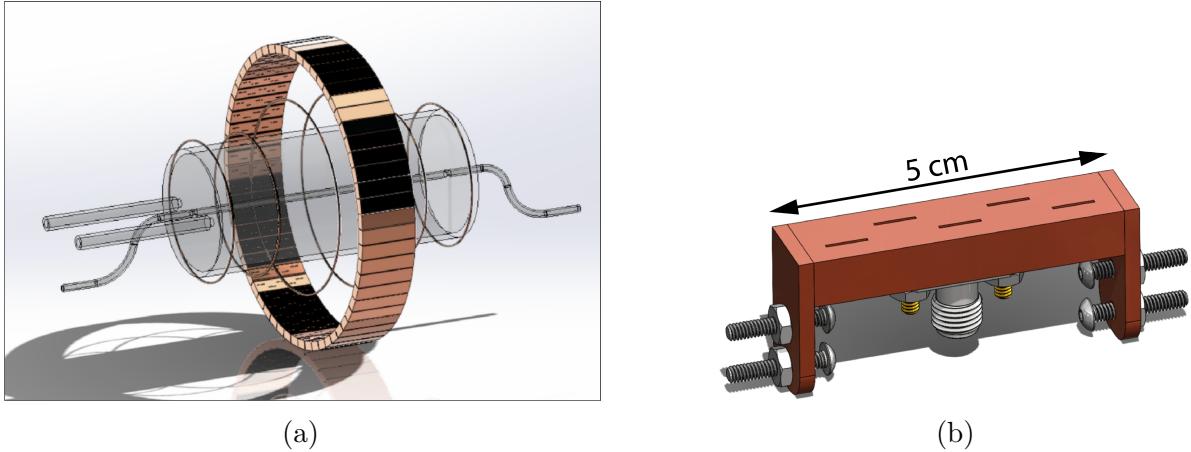


Figure 3.17. (a) A model of the FSCD antenna array, magnetic trap, and tritium containment vessel design.(b) A more detailed model of a prototype design for the 5-slot waveguide antenna design.

1912
1913 circle are sixty slotted waveguide antennas that fully populate the available space around
1914 the array circumference. In order to maximize the power collected from each electron
1915 it is optimal to cover as large a fraction of the solid angle around the magnetic trap as
1916 possible.

1917 The distance between antennas around the circumference of the array is proportional
1918 to the wavelength of the cyclotron radiation. Therefore, maximizing the solid angle
1919 coverage of the array, while minimizing channel count to keep the hardware and data
1920 acquisition costs manageable, biases one towards smaller array diameters. Antenna
1921 near-field effects limit the minimum diameter of the array for a given antenna design,
1922 since the radiation from electrons that are too close to the array cannot be detected due
1923 to destructive interference.

1924 Slotted waveguide antennas are used in the FSCD antenna array due to their high
1925 efficiency and low loss, which comes from the lack of dielectric materials in the antenna
1926 structure. Coupling to the waveguide is performed with a coaxial cable connected at the
1927 center of the antenna. One of the drawbacks of waveguide antennas is the large amount
1928 of space required to fit them inside the limited MRI magnet volume. Alternative antenna
1929 designs, constructed from microstrip printed circuit boards require significantly less space
1930 at the cost of slightly higher energy losses in the antenna structure.

1931 The FSCD antenna design is a 5 cm long segment of WR-34 waveguide with 5 vertical
1932 slots cut into the side. The distance between slots along the length of the waveguide is

1933 a half wavelength for optimal power combination between the individual antenna slots.
1934 Each slot is offset from the center of the antenna face a small distance in order to most
1935 effectively couple the slot to waveguide modes inside the antenna.

1936 The passive power combination achieved by placing 5 slots in a single waveguide is a
1937 compromise intended to reduce the cost and complexity of the antenna array system.
1938 Each additional channel in the array requires it's own cryogenic amplifier and also increase
1939 the required computer power to process the raw data collected by digitizing each channel.
1940 Passive summation, achieved by combining antennas into arrays axially, reduces the
1941 array channel count at the cost of losses from imperfect passive combination.

1942 Interference and re-radiation eventually limit the achievable the axial extent of passive
1943 power combination. The 5-slot designed developed for the FSCD is optimized to minimize
1944 the impact of these losses while achieving the maximum amount of axial coverage with a
1945 single ring of antennas. Scaling beyond the volume covered by a single ring of antennas is
1946 achieved by stacking additional rings of antennas together to cover a larger trap volume.
1947 A likely scenario for the FSCD experiment involves a staged experiment approach, where
1948 first a series of measurements is performed using only a single ring of antennas followed by
1949 experiments that add additional rings to the FSCD. The goal would be to first understand
1950 the principles of antenna array CRES using the simplest possible experiment, before
1951 attempting to scale the technique by expanding the antenna array size.

1952 **Tritium Source**

1953 While the primary purpose of the FSCD is as a technology demonstrator, it is impossible to
1954 retire all risks with the Phase IV experiment without an intermediate scale measurement
1955 of the neutrino mass. Therefore, the FSCD has the scientific goal of measuring the
1956 neutrino mass with a rough sensitivity goal in the range of a few eV. This level of precision
1957 is achievable using a molecular tritium source with a volume of approximately 1 L at a
1958 density comparable to potential Phase IV scenarios.

1959 Unlike previous CRES experiments, where the tritium source could be colocated
1960 with the receiving antenna inside a waveguide transmission line, the tritium source
1961 in the FSCD is thermally isolated from the antenna array to avoid freeze-out of the
1962 tritium molecules. The tiny radiation power emitted by electrons requires a system noise
1963 temperature of ≈ 10 K or less, in order to detect events at a high enough efficiency to
1964 reach the neutrino mass sensitivity goals of the experiment. Achieving a system noise of
1965 10 K requires that the antenna array and amplifiers operate at liquid helium temperatures
1966 of ≈ 4 K, which significantly lowers the vapor pressure of molecular tritium. By keeping

1967 the molecular tritium isolated in an RF-transparent vessel the tritium gas can be kept
1968 at a relatively warmer temperature in the range of 30 K to avoid the accumulation of
1969 tritium on the experiment surfaces.

1970 Data Acquisition and Reconstruction

1971 A fundamental change in the data acquisition system for the FSCD is the shift from
1972 single to multichannel reconstruction. This transition results in a significant increase in
1973 the data-generation rate, which is linearly related to the number of independent channels
1974 in the array. The larger data volume coincides with an increased demand for computer
1975 processing power based on the need for more precise signal reconstruction algorithms
1976 driven by the FSCD and Phase IV sensitivity goals. Therefore, the data acquisition
1977 system for the FSCD is likely to represent a significantly larger fraction of the experiment
1978 cost and complexity than in Phase II.

1979 Each antenna is connected to a cryogenic amplifier and down-converted from the
1980 26 GHz CRES frequency using an IQ-mixer to reduce the size of the analysis window.
1981 Using an LO with a frequency of approximately 25.80 GHz the antenna array signals can
1982 be digitized at a rate of 200 MHz, which is sufficient bandwidth to resolve the complete
1983 sideband spectrum produced by axial oscillations of electrons in the FSCD magnetic
1984 trap.

1985 Direct storage of the raw FSCD antenna array data is undesirable, since the estimated
1986 amount of raw data generated is $O(1)$ exabyte per year. The storage of such a large
1987 dataset is infeasible for a demonstrator experiment like the FSCD, since it would represent
1988 a disproportionate fraction of the total experiment budget in Phase III and Phase IV.
1989 Therefore, a goal of the FSCD experiment is the development of real-time reconstruction
1990 methods that could reduce the raw data volume by detecting and reconstructing CRES
1991 events in real-time. Ultimately, a real-time CRES reconstruction pipeline is desired, which
1992 takes raw voltages samples from the antenna array and converts them into measured
1993 starting kinetic energy values for electrons.

1994 The feasibility of a real-time reconstruction pipeline rests on the development of
1995 computationally efficient algorithms that can be implemented without the need for
1996 enormous computing resources. One challenge with the antenna array approach is that
1997 the small radiation power of a single electron is distributed among all channels in the
1998 array, such that reconstruction using only the information in a single channel is not
1999 possible. Therefore, the simply performing the initial step in reconstruction — signal
2000 detection — requires orders of magnitude more computational power than previous CRES

2001 experiments. This operation will then be followed by other, potentially more expensive,
2002 reconstruction steps that are required in order to determine the kinetic energy of the
2003 electron.

2004 **3.5 Pilot-scale Experiments**

2005 **3.5.1 Choice of Frequency**

2006 The optimal CRES frequency for Project 8 is that which allows us reach our target
2007 sensitivity of 40 meV, while minimizing the cost and complexity of the overall experiment.
2008 The magnitude of the background magnetic field determines the cyclotron frequency,
2009 which affects the entirety of the CRES detection system design, therefore, specifying the
2010 operating frequency of the CRES experiments is one of the first steps towards developing
2011 a full design.

2012 **Scaling Laws**

2013 The Phase I and II experiments utilized a background magnetic field of 0.959 T provided
2014 by an NMR magnet. This magnetic field was selected primarily for convenience, however,
2015 the cyclotron frequencies for electrons near the tritium endpoint in a 0.959 T field ranges
2016 from 25 to 26 GHz, which is within the standard RF Ka-band. Therefore, microwave
2017 electronics specialized for these frequencies are easily obtainable for relatively low cost.
2018 The operating frequency for the large-scale experiments must be selected in a more
2019 rigorous manner due to the increased scale and complexity of the systems as well as the
2020 requirements of the 40 meV neutrino mass science goal.

2021 There is a bias towards lower frequencies in a large-volume experiment, due to the
2022 direct relationship between wavelength and the physical size of the compatible RF
2023 components like antennas and cavities. With a longer wavelength more volume can
2024 be surrounded by an array with fewer antennas, which reduces hardware and data-
2025 processing costs. Additionally, the size of a cavity experiment is directly proportional to
2026 the wavelength, since this sets the physical dimensions of the cavity. It is also simpler to
2027 engineer a magnet that provides a uniform magnetic field across several cubic-meters of
2028 space at lower magnetic fields, which provides advantages in terms of cost-reduction.

2029 A concern with lower magnetic fields and frequencies is the scaling of the Larmour
2030 power equation, which is proportional to the square of the frequency. Naively, one would
2031 predict that the SNR would decrease with lower fields, however, two additional scaling

laws that affect the noise power also come into play. Noise power is directly proportional to the required bandwidth, which decreases linearly with the magnetic field. Furthermore, at lower frequencies it is possible to purchase amplifiers with lower noise temperatures until approximately 300 MHz at which point this relationship tends to flatten. Therefore, it is expected that the SNR remains approximately constant as the frequency decreases.

The SNR directly impacts the overall efficiency of the experiment through its effects on signal detection and energy resolution. Thus, the expectation that SNR remains the same at lower frequencies clearly biases large-scale experiments in this direction. One drawback of lower magnetic fields is the increased influence of external magnetic fields on the experiment. This includes magnetic fields from the building materials as well as variations in the earth's magnetic field. To deal with these affects a suitable magnetic field correction system will need to be devised, which includes constant monitoring of external fields.

Atomic Tritium Considerations

The pilot-scale experiments will be the first Project 8 experiments to combine CRES with atomic tritium, therefore, the optimal frequency should take into account the affect of the background magnetic field on the atom trap. The primary influence of the background

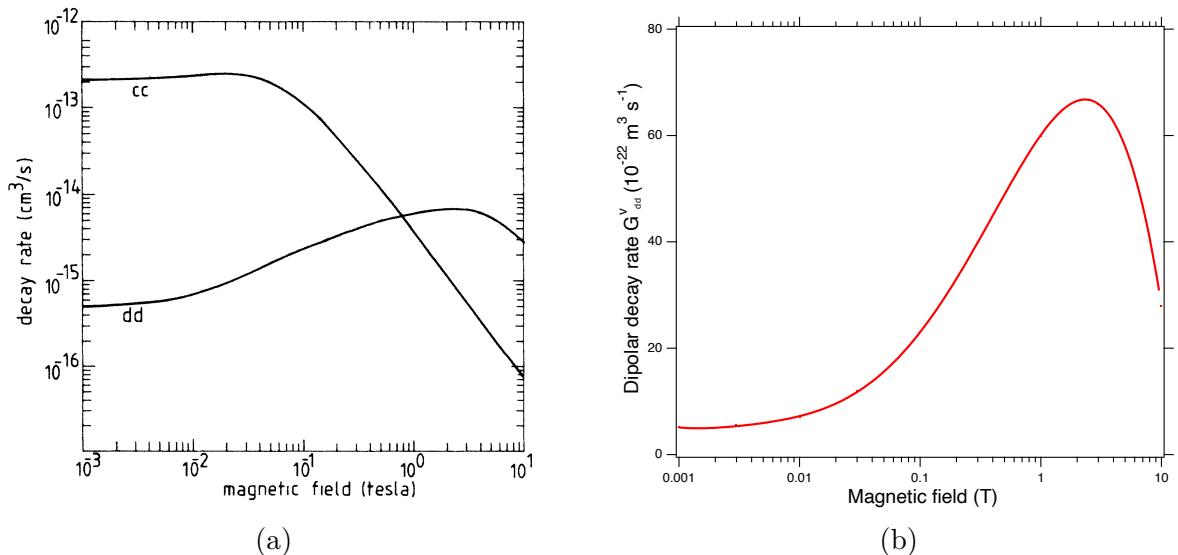


Figure 3.18. (a) A plot of the decay rate for the two-body dipolar spin exchange interaction for cc and dd state. (b) A plot of the decay rate of the dipolar spin exchange interaction for d+d states as a function of magnetic field magnitude. Lowering the magnetic field is key for reducing the losses from this interaction.

2049 field magnitude is through the rate of dipolar spin-flips caused by a spin exchange
2050 interaction between trapped atoms [57].

2051 Atomic tritium is a simple quantum system with a hyperfine structure given by the
2052 addition of the nuclear and atomic spins. The addition of two spins leads to a hyperfine
2053 structure with four states in the (m_s, m_I) basis [58]. The states with atomic spins directed
2054 anti-parallel to the magnetic field have $m_s = -1/2$ and are labeled as the a and b states.
2055 The a and b states are colloquially known as high-field seeking states, since their energy is
2056 minimized when in regions of higher magnetic field. This leads to losses in the magnetic
2057 trap as these atoms are drawn to higher fields away from the trap center. Alternatively,
2058 the c and d states, with atomic spin $m_s = +1/2$, minimize their energy in low magnetic
2059 fields because of the parallel alignment between spin and the magnetic field. Therefore,
2060 these low-field seeking states tend to stay trapped significantly longer than the high-field
2061 seeking states.

2062 It would be advantageous to prepare tritium atoms in purely c and d states before
2063 trapping, however, even in this case losses still occur due to dipolar interactions between
2064 pairs of c and d states leading to flipped atomic spins and subsequent losses from high-
2065 field seeking atoms. The rate of these interactions depends on the magnitude of the
2066 background magnetic field and is maximal for dd interactions around 1 T (see Figure
2067 3.18). The rate of losses from these interactions at 1 T requires atomic tritium production
2068 at a rate two orders of magnitude larger than at 0.1 T, thus, requirements on the whole
2069 atomic tritium system are significantly relaxed at lower magnetic fields, which provides
2070 powerful argument for moving to lower frequencies with the pilot-scale experiments and
2071 Phase IV.

2072 **3.5.2 Pilot-scale Experiment Concepts**

2073 While the pilot-scale experiments are still in the early stages, enough is known to sketch
2074 the general features of these experiments at the conceptual level.

2075 **Pilot-scale Antenna Array CRES Experiment Concept**

2076 A conceptual design for an antenna-based CRES experiment is shown in Figure 3.19.
2077 A large solenoid magnet provides a uniform background magnetic field less than 0.1 T
2078 in magnitude. Inside this region is the atom trapping magnet that generates a high
2079 magnetic field at the walls, which decays exponentially towards the central region. Known
2080 magnet designs that produce suitable atom trapping fields include Ioffe-Prichard traps,

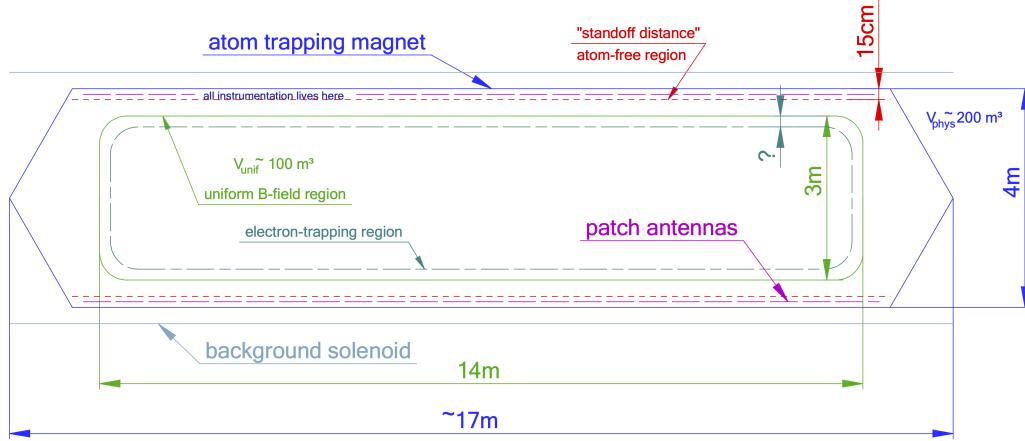


Figure 3.19. A conceptual sketch of a large-volume antenna array based CRES experiment to measure the neutrino mass.

which use conducting coils, as well as a Halbach array made from permanent magnets. Either magnet choice produces a region of high magnetic fields, which excludes atoms and allows for the placement of antennas inside the experiment.

Inside this region an array of microstrip patch antennas is inserted to collect the cyclotron radiation without providing a surface for atomic tritium recombination. Due to the lower frequency of cyclotron radiation antennas of a larger size can be used, which lowers the total number of antennas required to observe the experiment volume. Because of this scaling, the lower frequency experiment uses a similar number of antennas compared to a much smaller demonstrator experiment with a 1 T magnetic field.

The atomic tritium beamline that supplies fresh tritium atoms to the experiment is not shown in the figure. The general configuration would matches the one shown for the pilot-scale cavity experiment (see Figure 3.20).

2093 Pilot-scale Cavity CRES Experiment Concept

2094 The pilot-scale cavity experiment includes both an atomic tritium system and cavity
 2095 CRES system. The atomic system consists of a thermal atom cracker located at the
 2096 start of an evaporatively cooled atomic beamline. The atomic tritium system provides a
 2097 supply of tritium atoms to the trap with temperatures on the order of a few mK. Atoms
 2098 at this temperature can be trapped magneto-gravitationally, which is the reason for the
 2099 vertical orientation of the cavity. At these low magnetic fields the trapping requirements
 2100 for electrons and atoms differ enough such that it is advantageous to decouple the the
 2101 trapping potentials to avoid radioactive heating of the tritium atoms from excess trapped

2102 electrons. Electron trapping is provided by a set of magnetic pinch coils at the top and
2103 bottom of the cavity and a multi-pole Ioffe or Halbach magnet serves to contain the
2104 atoms.

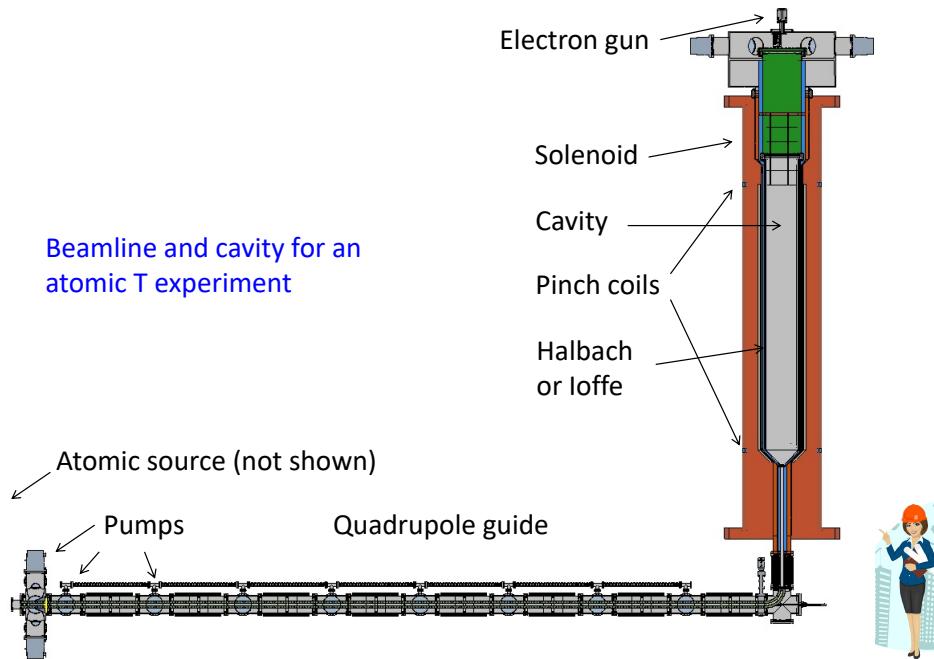


Figure 3.20. A conceptual sketch of a pilot-scale cavity CRES experiment with an atomic tritium beamline.

2105 The cavity design for the pilot-scale experiment consists of a large cylindrical cavity
2106 with a TE011 resonance of 325 MHz. Such a cavity is truly enormous, with a diameter
2107 of approximately 1.2 m and a height of 11 m. When an electron is produced inside
2108 the cavity with a cyclotron frequency that matches the TE011 resonant frequency it's
2109 cyclotron orbit couples the electron to the TE011, which drives a resonance in the cavity.
2110 These resonant fields can be read-out using an appropriate cavity coupling mechanism
2111 located at the center of the cavity. For more information on the cavity approach to
2112 CRES see Chapter 6.

2113 The bottom of the cavity has a cone termination to match the contour of the atom
2114 trapping magnet. This shape still allows for TE011 resonances with high internal Qs,
2115 which are required for good SNR in the cavity experiment. A small opening in the bottom
2116 of the cone serves as an entry point for the tritium atoms. To allow for calibration of
2117 the magnetic field inhomogeneities with an electron gun, the top of the cavity is left
2118 nearly completely open. Normally, this would drastically lower the Q-factor of the TE011
2119 mode, but a specially configured coaxial partition is inserted at the top. This termination

2120 scheme is designed to act as a perfect short for the TE011 mode since the circular shape
2121 of the partition matches the electric field boundary conditions for the TE011 mode.
2122 Simulations with HFSS have confirmed that this design results in a high quality TE011
2123 resonance despite the nearly completely open end.

2124 3.6 Phase IV

2125 The baseline CRES technology being pursued by the Project 8 collaboration are resonant
2126 cavities, which, due to their geometric properties, simple CRES signal structure, and low
2127 channel count, appear to be the better option for Phase IV. The current knowledge of the
2128 antenna array CRES approach reveals no technical obstacles that would preclude it as a
2129 baseline technology for Phase IV though it would most certainly be significantly more
2130 expensive. Therefore, antenna arrays represent a fallback approach if resonant cavities
2131 prove infeasible.

2132 The sensitivity of the pilot-scale atomic tritium experiment is estimated to be on
2133 the order of 0.1 eV, which means that increasing the sensitivity to reach the Phase IV
2134 goal will require an even larger experiment. Because of the direct coupling between the
2135 RF characteristics of a cavity and its geometry, the baseline plan is to build multiple
2136 copies of the pilot-scale experiment (see Figure 3.21) to obtain the required amount of
2137 volume rather than increase the size of the cavity beyond the pilot-scale. The built-in
2138 redundancy of this approach is useful in the sense that the experiment has no single
2139 point of failure, additionally, building several copies of the a pilot-scale experiment will
2140 minimize new engineering and design effort.

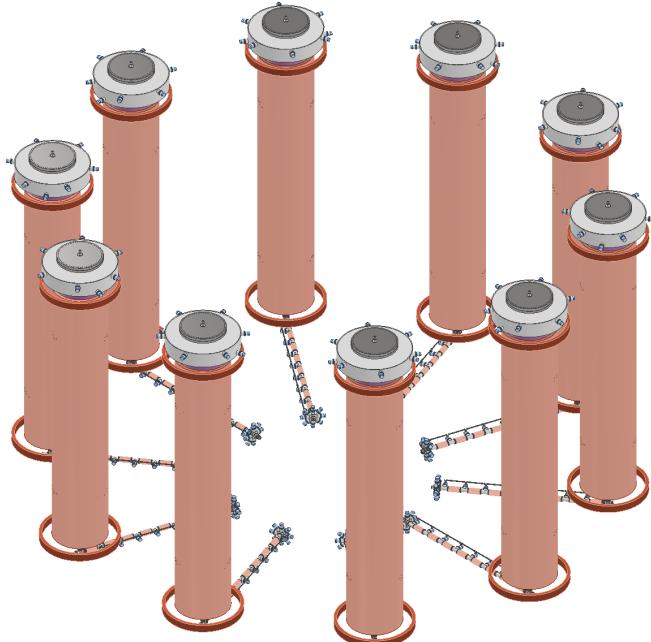


Figure 3.21. An illustration of a possible arrangement of ten pilot-scale cavity experiments for Phase IV. The experiments are arranged in a circle with an approximate diameter of 50 meters. Each atomic beamline connected to the bottom of each cavity is approximately 10 m in length. The cavities themselves are designed to operate at 325 MHz and are approximately 11 m tall. The circular arrangement of cavities has some advantages when it comes to cancellation of fringe fields from neighboring magnets, which is important due to the small magnetic field magnitudes consistent with these CRES frequencies. The advantage of ten independent atomic sources and cavities is that there is no single point of failure for the experiment. If an experiment goes down for repairs the other nine may continue running. Figure courtesy of Michael Huehn at UW-Seattle.

2141 **Chapter 4** |

2142 **Signal Reconstruction Techniques for An-**

2143 **tenna Array CRES and the FSCD**

2144 **4.1 Introduction**

2145 An antenna array CRES experiment introduces new challenges related to data acquisition,
2146 signal detection, and signal reconstruction caused by the multi-channel nature of the data.
2147 The development of signal reconstruction algorithms is crucial to the design of antenna
2148 array based experiments like the FSCD, because these algorithms directly influence the
2149 detection efficiency and energy resolution of the CRES experiment. In this Chapter I
2150 summarize my contributions to the development and analysis of signal reconstruction
2151 and detection algorithms for the FSCD experiment.

2152 In Section 4.2 I discuss the primary tool for this work, which is the Locust simulations
2153 package developed by the Project 8 experiment. Locust is used to simulate CRES events
2154 in the detector, which begins with calling a second software package — Kassiopeia — to
2155 calculate particle trajectory solutions for electrons in the magnetic trap. The trajectories
2156 are subsequently used to calculate the response of the antenna array to the cyclotron
2157 radiation produced by the electron, which results in signals that can be used to analyze
2158 the performance of different signal reconstruction algorithms. More recently, Project 8
2159 has developed CREsana, which is a new simulations package that takes a more analytical
2160 approach to CRES signal simulations for antenna arrays. Although CREsana signals
2161 were not used for the signal reconstruction algorithm development detailed here, we
2162 introduce the software as it is the simulation software used to model the antenna array
2163 measurements presented in Section 5.5.

2164 In Section 4.3 I discuss the signal reconstruction and detection approaches analyzed for
2165 the FSCD experiment. In general there are two steps to signal reconstruction — detection
2166 and parameter estimation. With signal detection one is concerned with distinguishing

2167 between data that contains a signal versus data that contains only noise, whereas, with
2168 parameter estimation one extracts the kinematic parameters of the electron encoded in
2169 the cyclotron radiation signal shape. Due to the low signal power of electrons near the
2170 spectrum endpoint in the FSCD experiment, signal detection is a non-trivial problem.
2171 This is magnified by the need to maximize the detection efficiency of the experiment
2172 in order to achieve the neutrino mass sensitivity goals. My contributions to signal
2173 reconstruction analyses for the FSCD are focused on the signal detection component of
2174 reconstruction.

2175 After discussing various signal detection approaches, in Section 4.4 I present a
2176 detailed analysis of the detection performance of three algorithms, which could be used
2177 to signal detection in the FSCD. This section was prepared for publication in JINST as
2178 a separate paper. The algorithms include a digital beamforming algorithm, a matched
2179 filter algorithm, and a neural network algorithm, which I analyze in terms of classification
2180 accuracy and estimated computational cost.

2181 **4.2 FSCD Simulations**

2182 Antenna array CRES and the FSCD require a combination of different capabilities
2183 not often found in a single simulation tool. In particular, accurate calculations of the
2184 magneto-static fields produced by current-carrying coils are needed to accurately model
2185 the magnetic trap and background magnets. The resulting magnetic fields must then be
2186 used to calculate the exact relativistic trajectory of electrons. The electron trajectories
2187 are required to calculate the electro-magnetic (EM) fields produced by the acceleration
2188 of the electron. Finally, the simulation must model the interaction of the antenna and
2189 RF receiver chain with the EM-fields in order to yield the simulated voltage signals from
2190 the antenna array. No available simulation tools adequately perform these combined
2191 functions, therefore, Project 8 developed a custom simulation framework to simulate the
2192 FSCD and CRES. This simulation framework includes custom simulation tools developed
2193 by Project 8, as well as open-source and proprietary software developed by third-parties.

2194 **4.2.1 Kassiopeia**

2195 Kassiopeia¹ is a particle tracking and static EM-field solver developed by the KATRIN
2196 collaboration for simulations of their spectrometer based on the MAC-E filter technique

¹<https://github.com/KATRIN-Experiment/Kassiopeia>

[59]. Unfortunately, Kassiopeia is not designed to solve for the EM-fields radiated by electrons in magnetic fields. However, it does provide efficient solvers for static electric and magnetic fields and charged particle trajectory solvers. Because of this, Project 8 has incorporated parts of Kassiopeia into the Locust simulation framework.

Magnetostatic Field Solutions

The solutions to the electric and magnetic fields generated by a static configuration of charges and currents is given by Maxwell's equations in the limit where the time-dependent terms go to zero. In their static form Maxwell's equations [48] are

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad (4.1)$$

$$\nabla \times \mathbf{E} = 0 \quad (4.2)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (4.3)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \quad (4.4)$$

where it can be seen that the electric and magnetic fields are completely decoupled from one another. The solution for the magnetic field in this boundary value problem is given by the Biot-Savart law

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int dr'^3 \frac{\mathbf{J}(\mathbf{r}') \times (\mathbf{r} - \mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|^3}, \quad (4.5)$$

which Kassiopeia can use a variety of numeric integration techniques to solve for a particular current distribution.

Kassiopeia Simulation of the FSCD Magnetic Trap

The trap developed for the FSCD experiment utilizes six current carrying coils, which surround a cylindrical tritium containment vessel (see Figure 4.1). Some critical aspects of the trap design include the total trapping volume, the maximum trap depth, the steepness of the trap walls, as well as the radial and azimuthal uniformity of the magnetic fields.

The volume of the FSCD trap is a cylindrically shaped region with a radius of 5 cm and a length of 15 cm resulting in a roughly 1 L total trap volume. The trap volume is an important design feature, because it sets the volume of the experiment that is potentially usable for CRES measurements. Trapping a larger volume allows one to observe a larger

number of tritium atoms, which increases the statistical power and sensitivity of the neutrino mass measurement. Due to the cost of constructing magnets with large and uniform magnetic fields it is important that the trap use as much of the available volume as possible to limit the overall cost of the experiment.

Coil	Radius (mm)	Z Pos. (mm)	Current (Amp.×Turns)
1	50.0	-92.3	750.0
2	50.1	-56.9	-220.3
3	68.5	-19.5	-250.0
4	68.5	19.5	-250.0
5	50.1	56.9	-220.3
6	50.0	92.3	750.0

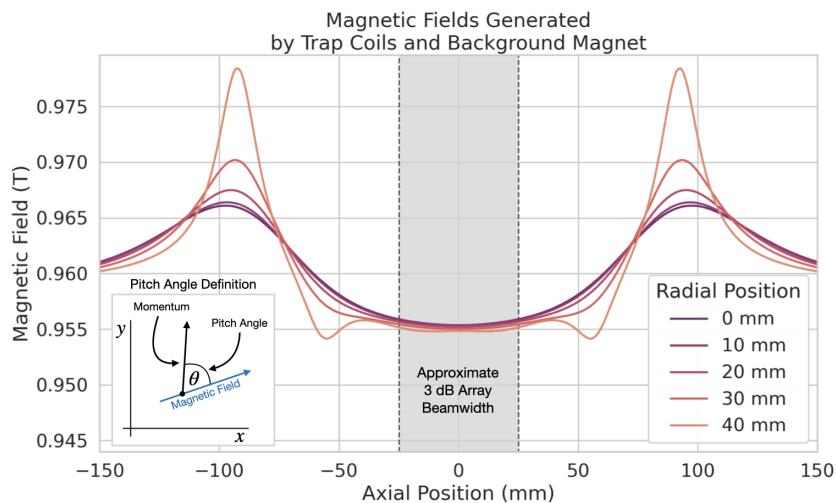
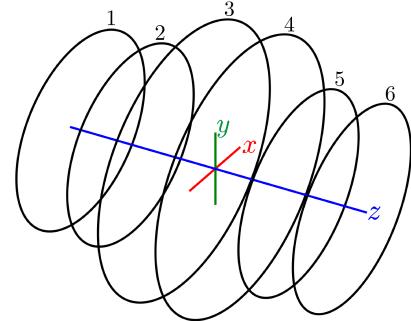


Figure 4.1. The geometry and parameters of the coils used to simulate the FSCD magnetic trap in Kassiopeia. Some axial profiles of the magnetic trap at different radial positions are shown to demonstrate the shape of the magnetic field and trap depth as a function of position. Calculation of the magnetic field profiles was graciously done by René Reimann.

The depth of the FSCD trap is approximately 10 mT when measured along the central axis, which is sufficient to trap electrons with pitch angles as small as 84° . The trap depth influences the efficiency of the experiment by directly controlling the range of electron pitch angles that can be trapped. If a higher fraction of pitch angles are trapped, in principle, more decay events can be observed. However, the signals from electrons with small pitch angles are significantly harder to detect in the FSCD than large pitch angles, which increases the likelihood of not detecting the first track of the CRES event and harms the energy resolution of the experiment.

The steepness of the trap walls as well as non-uniformities in the magnetic field

contribute to the total energy resolution of the CRES measurement by causing uncertainty in the relationship between an electron's kinetic energy and its cyclotron frequency. When an electron is trapped, it oscillates back and forth along the trap z-axis (see Figure 4.1) unless it has a pitch angle of exactly 90° [60]. As the electron is reflected from the trap walls it experiences a change in the total magnetic field, which causes a modulation in the cyclotron frequency. This change in magnetic field from the trap introduces a correlation between the pitch angle and kinetic energy parameters of the electron that can reduce energy resolution. In order to mitigate this effect it is important to make the trap walls as steep as possible.

Particle Trajectory Solutions

The magnetic fields solved by direct integration of the coil current densities are used to calculate the trajectories of electrons based on user specified initial conditions. Various statistical distributions are available, which can be sampled to replicate realistic event statistics. These include uniform, Gaussian, and Lorentzian distributions among others. In general, an electron has six kinematic parameters that define its trajectory, which are the three-dimensional coordinates of the initial position and the three components of the electron's momentum vector. However, when simulating CRES events it is common to parameterize the electron's trajectory in terms of the initial position, kinetic energy, pitch angle, and initial direction of the component of the electron's momentum perpendicular to the magnetic field. This parameterization is completely equivalent to specifying the starting position and momentum vectors.

From the initial parameters of the electron and the magnetic field, Kassiopeia solves for the trajectory of the electron. The direct approach proceeds by solving the motion of the electron using the Lorentz force equation, which takes the form of a set of differential equations

$$\frac{d\mathbf{r}}{dt} = \frac{\mathbf{p}}{\gamma m} \quad (4.6)$$

$$\frac{d\mathbf{p}}{dt} = e(\mathbf{E} + \frac{\mathbf{p} \times \mathbf{B}}{\gamma m}), \quad (4.7)$$

where \mathbf{r} is the position of the electron, \mathbf{p} is the electron's momentum, e is the charge of the electron, m is the electron's mass, and γ is the relativistic Lorentz term. Kassiopeia solves this pair of differential equations using numerical integration, however, the exact trajectory can be computationally intensive to solve. If the adiabatic approximation can

2262 be applied, then Kassiopeia can make use of a simpler set of equations that can be more
2263 readily solved numerically.

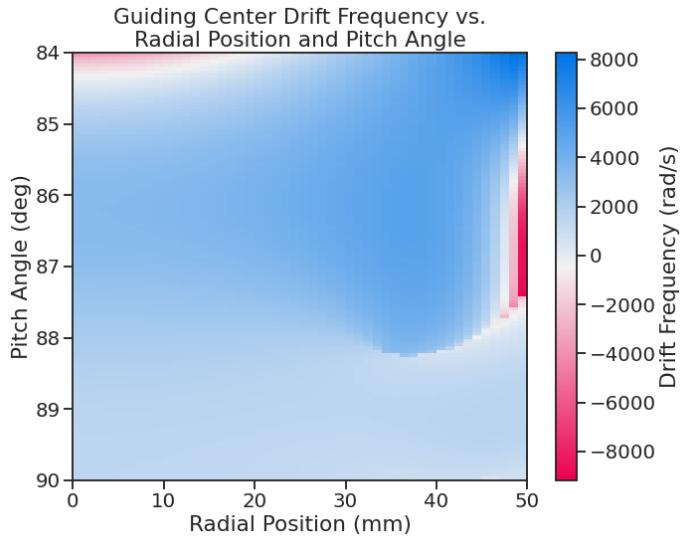


Figure 4.2. A map of the average ∇B -drift frequency for electrons trapped in the prototype FSCD trap shown in Figure 4.1. Negative drift frequencies indicate electrons that are drifting opposite to the standard direction, which means that they are close to escaping the magnetic trap.

2264 Though Kassiopeia is not directly capable of simulating the cyclotron radiation, it is
2265 an invaluable CRES simulation tool. With Kassiopeia it is possible to test the efficiency
2266 of a particular trap design, and analyze features of the electron trajectories that are
2267 important to the position, track, and event reconstruction (see Section 4.3). An example
2268 is the analysis of the average ∇B -drift frequency as a function of the electrons radial
2269 position and pitch angle in the FSCD trap (see Figure 4.2). Radial gradients in the trap
2270 cause the guiding center of the electron to drift around the center of the magnetic trap
2271 with an average frequency on the order of 10^3 rad/s. This frequency, while slow compared
2272 to the length of a typical CRES time-slice, is large enough to cause a significant loss in
2273 efficiency of certain signal reconstruction algorithms. Therefore, it is important to model
2274 the drift of the electron in the reconstruction algorithm in order to mitigate the effects
2275 of this motion on the reconstruction.

4.2.2 Locust

The Locust² software package [61] is the primary simulation tool developed and used by the Project 8 collaboration for CRES experiments. Locust simulates the responses of antennas and receiver electronics chain to rapidly time-varying electric fields using a flexible approach that allows one to choose from a variety of electric field sources and antennas. Similarly, one can simulate the receiver chain using a series of modular generators that include standard signal processing operations such as down-mixing and fast Fourier transforms (FFT). Since the primary focus of this chapter is the application of Locust to analyses of the FSCD, I shall describe only the most relevant aspects of the software rather than provide a comprehensive description.

Cyclotron Radiation Field Solutions

Simulating CRES events in the FSCD requires one to calculate the electric fields produced by the acceleration of the electron. In the general case, this can be a complicated computation, due to back-reaction forces on the electron. However, in the case of the FSCD it is possible to ignore such effects and approximate the electron as radiating into a free-space environment.

The equations that describe the EM fields from a relativistic moving point particle are the Liénard-Wiechert equations [62, 63], which are obtained by differentiating the Liénard-Wiechert potentials. In their full form, the Liénard-Wiechert field equations are

$$\mathbf{E} = e \left[\frac{\hat{n} - \boldsymbol{\beta}}{\gamma^2(1 - \boldsymbol{\beta} \cdot \hat{n})^3 |\mathbf{R}|^2} \right]_{t_r} + \frac{e}{c} \left[\frac{\hat{n} \times [(\hat{n} - \boldsymbol{\beta}) \times \dot{\boldsymbol{\beta}}]}{(1 - \boldsymbol{\beta} \cdot \hat{n})^3 |\mathbf{R}|} \right]_{t_r} \quad (4.8)$$

$$\mathbf{B} = [\hat{n} \times \mathbf{E}]_{t_r}, \quad (4.9)$$

where e is the charge of the particle, \hat{n} is the unit vector pointing from the particle to the position where the fields are calculated, $\boldsymbol{\beta}$ and $\dot{\boldsymbol{\beta}}$ are the velocity and acceleration of the particle divided by the speed of light (c), \mathbf{R} is the distance from the particle to the field calculation position, and γ is the relativistic Lorentz term. The subscript t_r indicates that the equations are evaluated at the retarded time so that the time-delay from the travel time of the electromagnetic radiation is taken into account.

The only required input to calculate the electric field at the position of an FSCD antenna is the velocity and acceleration of the electron, which can be obtained from Kassiopeia simulations. Therefore, when simulating a CRES event Locust first runs

²https://github.com/project8/locust_mc/tree/master

2304 a Kassiopeia simulation of the electron and subsequently calculates the electric field
 2305 incident on the antenna. This requires one to calculate the retarded time. The retarded
 2306 time corresponds to the time that a photon, which has just arrived at an antenna at
 2307 the space-time position (t, \mathbf{r}) , was actually emitted by the electron at the space-time
 2308 position of $(t_r, \mathbf{r}_e(t_r))$. To calculate the retarded time one solves

$$c(t - t_r) = |\mathbf{r} - \mathbf{r}_e(t_r)|, \quad (4.10)$$

2309 where the distance traveled by the photon between the measurement and retarded times
 2310 is equal to the distance between the antenna and the electron at the retarded time.
 2311 Locust solves Equation 4.10 using root finding algorithm to calculate the retarded time,
 2312 which yields the electric field emitted by the electron, at the position of each antenna in
 2313 the FSCD array.

2314 Antenna Response Modeling

2315 The electric field solutions are used to calculate the resulting voltages produced in the
 2316 antenna. However, direct simulation of the antenna itself is computationally expensive,
 2317 since it requires modeling the complex interactions of the electron's electric fields with
 2318 charge carriers in the antenna. Direct simulation of the antenna in Locust is avoided by
 2319 modeling the antenna response using the antenna factor, or antenna transfer function.
 2320 The antenna factor defines the voltage produced in the antenna terminal for an incident
 2321 electric field [64],

$$A_F = \frac{V}{|\mathbf{E}|}, \quad (4.11)$$

2322 where V is the voltage and $|\mathbf{E}|$ is the magnitude of the incident electric field. To obtain the
 2323 antenna factor for the antennas developed for the FSCD Project 8 employs Ansys HFSS.
 2324 HFSS is a commercially available finite element method electromagnetic solver widely
 2325 used throughout the antenna engineering industry [65]. HFSS is capable of calculating
 2326 the antenna factor and gain patterns for complex antenna designs and outputting the
 2327 resulting quantities in the form of a text file that can be used as a configuration input to
 2328 Locust.

2329 The antenna factor defines the steady-state response of the antenna to electromagnetic
 2330 plane waves in the frequency-domain. Since the antenna response is calculated in the
 2331 time-domain Locust models the antenna as a linear time-invariant system [66]. In this

2332 formalism the response of the system to the driving force is given by

$$y[n] = h * x = \sum_k h[k]x[n - k], \quad (4.12)$$

2333 where $y[n]$ is the discretely sampled response, x is the driving force stimulus, and h is
 2334 the finite impulse response (FIR) filter. When applied to the FSCD array, this formalism
 2335 calculates the voltage time-series produced in each antenna by convolving the electric
 2336 field time-series with the antenna FIR filter, which is obtained by performing an inverse
 2337 Fourier transform on the transfer function from HFSS.

2338 Radio-frequency Receiver and Signal Processing

2339 After obtaining the voltage time-series by computing the electron trajectory and antenna
 2340 response, Locust simulates the signal processing performed by the radio-frequency (RF)
 2341 receiver chain. The simulated Locust receiver chain includes all operations that would
 2342 be performed by the RF hardware (see Figure 4.3).

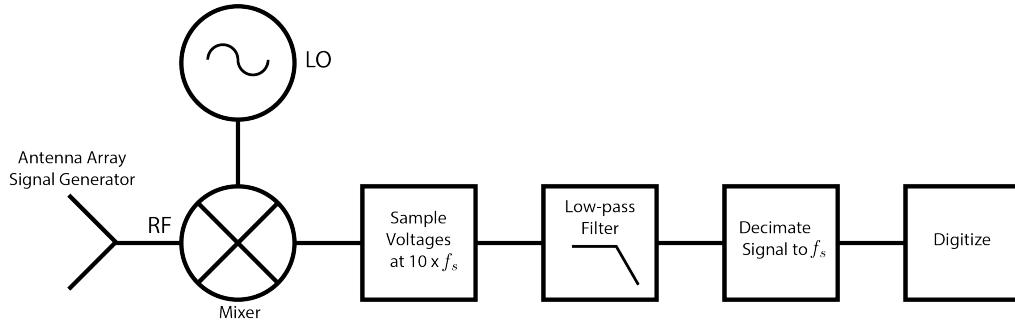


Figure 4.3. The receiver chain used by Locust when simulating CRES events in the FSCD.

2343 Frequency down-conversion reduces the digitization bandwidth required to read-out
 2344 CRES data. According to the Nyquist sampling theorem [67], the minimal sampling rate
 2345 that guarantees no information loss for a signal with a bandwidth Δf is given by

$$f_{\text{Nyq}} = 2\Delta f. \quad (4.13)$$

2346 The total bandwidth for CRES events ranges from 0 to 26 GHz in a 0.95 T magnetic field,
 2347 therefore, direct digitization of CRES signals from the FSCD would require sampling
 2348 frequencies greater than 50 GHz, which is infeasible for a real experiment. However, one
 2349 need only measure the shape of the spectrum in the last 100 eV, which corresponds to a
 2350 frequency bandwidth of 5 MHz, to effectively measure the neutrino mass.

2351 Down-conversion is a technique for reducing the base frequencies of signals in a
2352 bandwidth given by $[f_{\text{LO}}, f_{\text{LO}} + \Delta f]$ to the bandwidth $[0, \Delta f]$, by performing the following
2353 multiplication

$$x(t) \rightarrow x(t)e^{-2\pi f_{\text{LO}} t}. \quad (4.14)$$

2354 The signal, $(x(t))$, is multiplied by a sinusoidal signal with frequency f_{LO} to reduce the
2355 absolute frequencies of the signals in the bandwidth. In the FSCD, this allows one to
2356 detect events in the last 100 eV of the tritium spectrum, while sampling the data far
2357 below 50 GHz. The standard bandwidth used in the FSCD is 200 MHz, which allows for
2358 higher frequency resolution than the minimum sampling frequency for 100 eV of energy
2359 bandwidth.

2360 Directly simulating down-conversion with a frequency multiplication in Locust requires
2361 sampling the electric fields at each antenna in the FSCD array with a period of ≈ 20 ps,
2362 which is extremely slow computationally. To avoid this, Locust performs the down-
2363 conversion by intentionally under-sampling the electric fields with a frequency of 2 GHz.
2364 Sampling below the Nyquist limit causes the higher frequency components of the CRES
2365 signal to alias, however, Locust can remove these aliased frequency peaks using a
2366 combination of low-pass filtering and decimation to recreate frequency down-conversion.
2367 After filtering and decimation, Locust simulates digitization by an 8-bit digitizer at a
2368 sampling frequency of 200 MHz to recreate the conditions of the FSCD. The voltage
2369 offset and digitizer range must be configured by the user based on the characteristics of
2370 the simulation.

2371 Data

2372 The output of Locust simulations for the FSCD primarily consists of two data files. The
2373 first is the electron trajectory information calculated by Kassiopiea, which is output in
2374 the form of a `.root` file [68]. This file contains important kinematic information about
2375 the electron such as its position and pitch angle as a function of time. The other file
2376 is produced by Locust and contains the digitized signals acquired from each antenna
2377 in the array. The Locust output files conform to the Monarch specification developed
2378 by Project 8, which is based on the commonly used HDF5 file format, and matches the
2379 format of the files produced by the Project 8 data acquisition software. This makes it
2380 possible to use the same data analysis code to analyze both simulated and real data.

4.2.3 CRESana

Locust is the primary simulation tool used by Project 8 in the development and simulation of the FSCD. However, simulations of CRES events in larger antenna arrays (≥ 100 antennas) can take several hours to complete, which is prohibitively long when one is performing a sensitivity analysis and optimization. One reason for Locust's slow operation is that the electric fields from the electron must be solved numerically for each time-step for all antennas in the array. These numerical solutions allow Locust to accurately simulate the electric fields from arbitrarily complicated electron trajectories at the cost of more computations and slower simulations. Therefore, an additional simulation tool that sacrifices the accuracy of numerical approaches for computational efficiency is a useful tool for studying large antenna array experiments.

Recently, Project 8 has developed a new simulations package called CRESana³, specifically designed to perform analytical simulations of antenna array based CRES experiments. CRESana provides a significant increase in simulation speed by using well-justified analytical approximations of the electrons motion and electric fields in a magnetic trap. The electric fields and signals generated by CRESana are consistent with theoretical calculations of the electron's radiation, and are tested for accuracy using well-known test-case simulations and consistency checks.

4.3 Signal Detection and Reconstruction Techniques for Antenna Array CRES

Antenna Array CRES Signal Reconstruction

Antenna array CRES requires one to use the multichannel time-series obtained by digitizing the array to estimate the starting kinetic energies of electrons produced in the magnetic trap using CRES signal reconstruction algorithm. This procedure consists of a multi-stage process of detecting a CRES signal followed by an estimation of the electron's parameters.

Antenna array CRES requires a significantly different approach to signal reconstruction than previous Project 8 experiments. In Phases I and II, CRES was performed using a waveguide gas cell directly integrated into a waveguide transmission line. The transmission line efficiently propagates the cyclotron radiation along its length to an antenna at the

³<https://github.com/MCflowMace/CRESana>

2411 ends of the waveguide. However, with an antenna array the electron is radiating into
2412 free-space, therefore, the cyclotron radiation power collected by the array is directly
2413 proportional to the solid angle surrounding the electron that is covered with antennas.
2414 Because it is not practical to fully surround the magnetic trap with antennas, some of the
2415 cyclotron radiation power that would have been collected by the waveguide escapes into
2416 free-space. Furthermore, the power that is collected by the antenna array is split between
2417 every channel in the antenna array, which significantly lowers the signal-to-noise ratio
2418 (SNR) of CRES signals in a single antenna channel compared to a waveguide apparatus.
2419 Therefore, a suite of completely new signal reconstruction techniques are needed in order
2420 to perform CRES in the FSCD.

2421 Changes to the approach to CRES signal reconstruction are also motivated by the
2422 more ambitious scientific goals of the FSCD experiment. A measurement of the tritium
2423 beta-decay spectrum that is sensitive to neutrino masses as small as 40 meV requires that
2424 we measure the kinetic energies of individual electrons with a total energy broadening of
2425 115 meV [69]. This resolution includes all sources of uncertainty in the electron's kinetic
2426 energy such as magnetic field inhomogeneities. This precise energy resolution is only
2427 achieved by an event-by-event signal reconstruction approach where the kinetic energies,
2428 pitch angles, and other parameters of the CRES events are estimated for individual
2429 electrons before constructing the beta-decay spectrum.

2430 The event-by-event approach is distinct from the analysis done for the Phase I and
2431 Phase II experiments where only the starting cyclotron frequency of the event was
2432 measured by analyzing the tracks formed by the carrier frequency. These frequencies
2433 were then combined into a frequency spectrogram, which was converted to the beta-
2434 decay energy spectrum using an ensemble approach that averaged over all other event
2435 parameters. The ensemble approach to signal reconstruction results in poor energy
2436 resolution because other kinematic parameters such as pitch angle change the cyclotron
2437 carrier frequency due to changes in the average magnetic field experience by the electron.

2438 Components of Reconstruction: Signal Detection and Parameter Estimation

2439 CRES signal reconstruction is a two-step procedure consisting of signal detection followed
2440 by parameter estimation. In the former, one is concerned with identifying CRES signals
2441 in the data regardless of the signal parameters, whereas, in the latter one operates under
2442 the assumption that a signal is present and then estimates its parameters.

2443 More formally, signal detection can be posed as a binary hypothesis test between
2444 the signal and noise data classes, and parameter estimation is a process of fitting a

2445 signal model to the observed data. While both of these are required for a complete
 2446 reconstruction (see Figure 4.4), the focus of my work and this chapter is on the signal
 2447 detection aspect of antenna array CRES signal reconstruction.

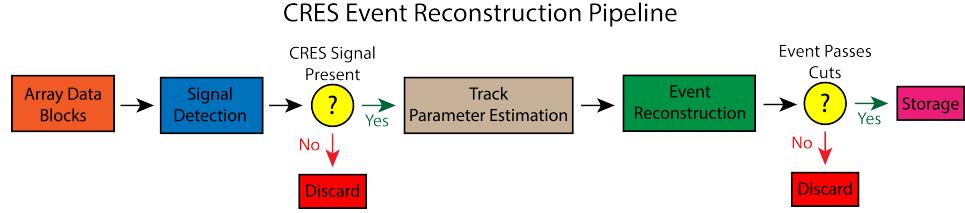


Figure 4.4. A high-level diagram depicting the process of CRES event reconstruction. The first step consists of identifying the presence of a signal in the data. This step is necessary to avoid the danger of performing a reconstruction of a false event, which would constitute a background contribution to the tritium spectrum measured by CRES.

2448 Detection Theory

2449 Signal detection is the process of deciding whether noisy data contains signal or noise,
 2450 which can be posed as a statistical hypothesis test [70]. For CRES signals, which are
 2451 essentially vectors with added white Gaussian noise (WGN), one needs to choose between

$$\mathcal{H}_0 : \mathbf{y} = \boldsymbol{\nu} \quad (4.15)$$

$$\mathcal{H}_1 : \mathbf{y} = \mathbf{x} + \boldsymbol{\nu}, \quad (4.16)$$

2452 where \mathbf{y} is the CRES data vector, $\boldsymbol{\nu}$ is a sample of WGN, and \mathbf{x} represents the CRES
 2453 signal. The hypothesis that the data contains only noise is labeled \mathcal{H}_0 and the hypothesis
 2454 that the data contains a signal is labeled \mathcal{H}_1 .

2455 For illustrative purposes, it is useful to study the case where only the first sample of
 2456 data is used to distinguish between \mathcal{H}_0 and \mathcal{H}_1 . The value of the first data sample is
 2457 distributed according to two possible Gaussian distributions (see Figure 4.5). By setting a
 2458 decision threshold on the value of this sample, one can choose the correct hypothesis with
 2459 a probability given by the area underneath the probability distribution curves. A true
 2460 positive corresponds to correctly identifying that the data contains signal, whereas, a true
 2461 negative means that one has correctly identified the data as noise. The rate at which the
 2462 detector performs a true positive classification is given by the green region underneath
 2463 $p(\mathbf{y}[0]; \mathcal{H}_0)$, and the rate at which the detector performs a true negative classification
 2464 is given by the orange region underneath $p(\mathbf{y}[0]; \mathcal{H}_1)$. Two types of misclassifications
 2465 are possible. Either one declares noise data as signal, which is called a false positive, or

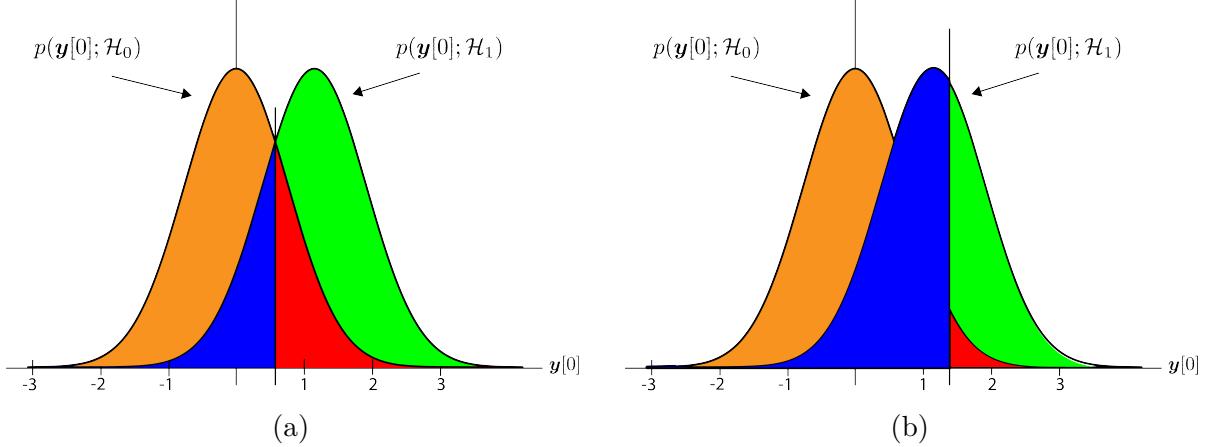


Figure 4.5. An illustration of two PDFs associated with a binary hypothesis test. The decision threshold is represented by the vertical line that partitions both distributions. The orange and red areas correspond to the true negative and false positive probabilities and the blue and green areas correspond to the false negative and true positive probabilities respectively. To decide between the two hypotheses we perform the likelihood ratio test specified by the Neyman-Pearson theorem. This approach achieves the highest true positive probability for a given false positive probability.

2466 one declares signal data as noise, which is a false negative. Note that it is only possible
 2467 to trade off these two types of errors by tuning the detection threshold. One cannot
 2468 simultaneously reduce the rate of false positives without also increasing the rate of false
 2469 negatives.

2470 The approach taken with CRES signals is to fix the rate of false positives by setting
 2471 a minimum decision threshold value. The rate of false positives that is acceptable at the
 2472 detection stage depends upon the total rate of background events compatible with the
 2473 sensitivity goals of the experiment. The ultimate goal of a neutrino mass measurement
 2474 with 40 meV sensitivity in general has strict requirements on the number of background
 2475 events, which requires a relatively high detection threshold to achieve. Consequently,
 2476 the ideal signal detection algorithm is the one that achieves the maximum rate of true
 2477 positives for a fixed rate of false positives, so that the detection efficiency of the experiment
 2478 is maximized and potential sources of background are kept to a minimum.

2479 According to the Neyman-Pearson theorem [71], the statistical hypothesis test that
 2480 maximizes the probability of detection for a fixed rate of false positives is the likelihood
 2481 ratio test, which is formed by computing the ratio of the signal likelihood to the noise
 2482 likelihood,

$$L(x) = \frac{P(\mathbf{y}; \mathcal{H}_1)}{P(\mathbf{y}; \mathcal{H}_0)} > \gamma. \quad (4.17)$$

2483 Here, the likelihood of the hypotheses \mathcal{H}_0 and \mathcal{H}_1 are described by the probability
2484 distributions $P(\mathbf{y}; \mathcal{H}_0)$ and $P(\mathbf{y}; \mathcal{H}_1)$ respectively, and γ is the threshold for deciding \mathcal{H}_1 .
2485 The decision threshold is determined by integrating $P(\mathbf{y}; \mathcal{H}_0)$ such that

$$P_{\text{FP}} = \int_{\gamma}^{\infty} P(\tilde{\mathbf{y}}; \mathcal{H}_0) d\tilde{\mathbf{y}} = \alpha, \quad (4.18)$$

2486 where α is the desired false positive detection rate given by the red colored areas shown
2487 in Figure 4.5. The true positive detection rate is given by the similar integral

$$P_{\text{TP}} = \int_{\gamma}^{\infty} P(\tilde{\mathbf{y}}; \mathcal{H}_1) d\tilde{\mathbf{y}}, \quad (4.19)$$

2488 which corresponds to the green areas in Figure 4.5.

2489 Changing the decision threshold allows one to trade-off between P_{TP} and P_{FP} as
2490 appropriate for the given situation. It is standard to summarize the relationship between
2491 P_{TP} and P_{FP} using the receiver operating characteristic (ROC) curve, which is obtained
2492 by evaluating the true positive and false positive probabilities as a function of the decision
threshold value (see Figure 4.6). The ROC curve provides a convenient way to compare

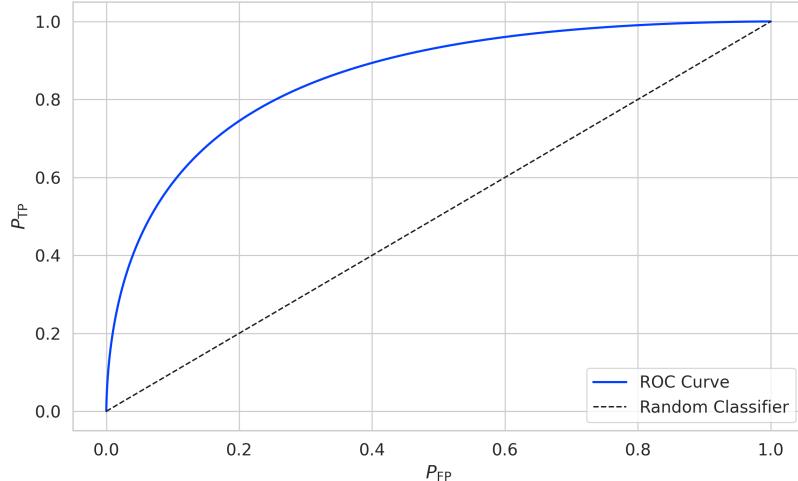


Figure 4.6. An example ROC curve formed by computing the P_{FP} and the P_{TP} for a given likelihood ratio test. As the decision threshold is increased P_{FP} decreases at the expense of a lower P_{TP} . The black dashed line indicates the lower bound ROC curve obtained by randomly deciding between \mathcal{H}_0 and \mathcal{H}_1 .

2493
2494 the performance of different signal detection algorithms. In general, a classifier with
2495 a higher the P_{TP} as a function of P_{FP} is desirable, which corresponds to a larger area
2496 underneath the respective ROC curve. A perfect classifier has an area underneath the

2497 curve of 1.0, however, such a classifier is never achieved in practice.

2498 4.3.1 Digital Beamforming

2499 Introduction to Beamforming

2500 Beamforming is an antenna array signal processing technique designed to enhance the
2501 radiation of the array in a particular direction and suppress it in other directions [64].
2502 Beamforming is of interest to Project 8 as a first level of signal reconstruction for the
2503 FSCD and other antenna array CRES experiments, which operates at the signal detection
2504 stage of reconstruction.

2505 Beamforming is performed using a phased summation of the signals received by the
2506 antenna array. The beamforming phases are selected such that the signals emitted by
2507 the array will constructively interfere at the point of interest (see Figure 4.7). As a
2508 consequence of the principle of reciprocity [72], when the array is operating in receive
2509 mode, the signals emitted from a source at the same point will constructively interfere
when summed. The origin of the phase delays in beamforming is the path-length difference

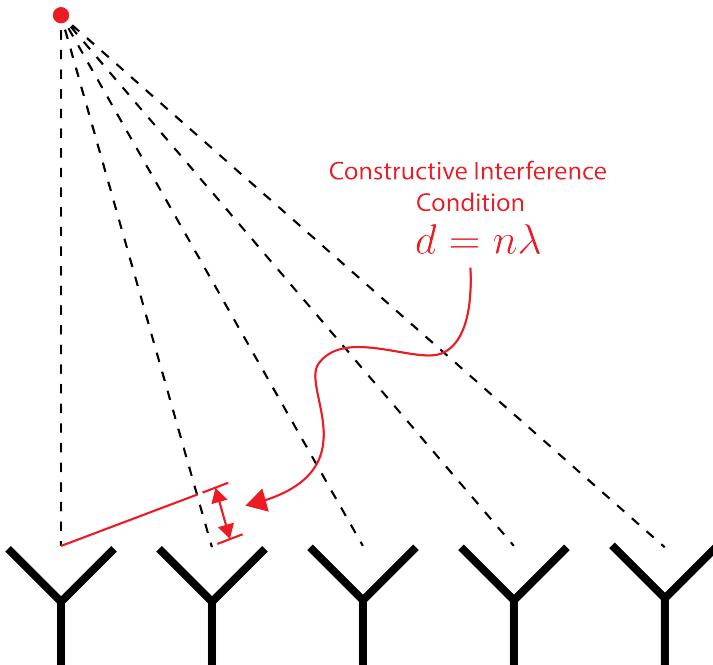


Figure 4.7. An illustration of the constructive interference condition which is the operating principle of digital beamforming using a uniform linear array as an example.

2510
2511 to the beamforming point between different antennas in the array. The relationship

2512 between the phase delay and the path-length difference is given by the familiar equation

$$\phi = \frac{2\pi d}{\lambda}, \quad (4.20)$$

2513 where ϕ is the phase delay, d is the path-length difference, and λ is the wavelength of
2514 the radiation. In practice, one chooses the values of d by specifying the beamforming
2515 positions of interest and then calculates the beamforming phases using Equation 4.20,
2516 which is guaranteed to follow the constructive interference condition shown in Figure 4.7.

2517 Beamforming can be neatly expressed mathematically using the vector equation

$$y[n] = \Phi^T[n] \mathbf{x}[n], \quad (4.21)$$

2518 where $\mathbf{x}[n]$ is the array snapshot vector, $\Phi[n]$ is a vector of beamforming shifts, and
2519 $y[n]$ is the resulting summed signal. The beamforming shifts consist of a set of complex
2520 numbers that contain the beamforming phase shift and an amplitude weighting factor,

$$\Phi[n] = [A_0[n]e^{-2\pi i\phi_0[n]}, A_1[n]e^{-2\pi i\phi_1[n]}, \dots, A_{N-1}[n]e^{-2\pi i\phi_{N-1}[n]}], \quad (4.22)$$

2521 where the set of magnitudes $A_i[n]$ are amplitude weighting factors and $\phi_i[n]$ are the
2522 phase shifts from the path-length differences. The index i is used to denote the antenna
2523 channel number. The amplitude weighting factor is the relative magnitude of the signal
2524 received by a particular antenna in the array. This factor properly accounts for antennas
2525 that are closer to the radiating source. In general, the beamforming phases can also be
2526 functions of time to track the motion of a non-stationary source.

2527 Digital beamforming specifically is the type of beamforming algorithm of interest to
2528 Project 8 for CRES. With digital beamforming, the phase shifts are applied to the array
2529 signals in software rather than employing fixed beamforming phase shifts in the receiver
2530 chain hardware. The advantage of digital beamforming is that for any given series of
2531 array data one can specify an arbitrarily large number of beamforming positions and
2532 search for electrons using a flexible and easily configurable beamforming grid.

2533 Digital beamforming can be viewed as the spatial filtering, which is a direct conse-
2534 quence of the constructive interference condition used to define the beamforming phases.
2535 Digital beamforming causes signals from multiple electrons at different positions in the
2536 trap to be separated, because the interference condition will cause the signals from
2537 electrons at other position to cancel out. This spatial filtering effect reduces pile-up that
2538 could become an issue for large scale CRES experiments using a dense tritium source.

Beamforming positions can be specified with arbitrary densities limited only by the available computational resources. This provides a very straight-forward way to estimate the position of the electron in the trap by using a dense grid of beamforming positions and maximizing the output power of the beamforming summation over this grid. This approach to position reconstruction is attractive due the requirements of an event-by-event signal reconstruction, which needs an accurate estimation of the exact magnetic field experienced by the electron in order to correctly estimate its kinetic energy. Combined with an accurate map of the magnetic field inhomogeneities of the trap obtained from calibrations, beamforming allows one to apply this magnetic field correction with a spatial resolution that is a fraction of the cyclotron wavelength.

Laboratory Beamforming Demonstrations

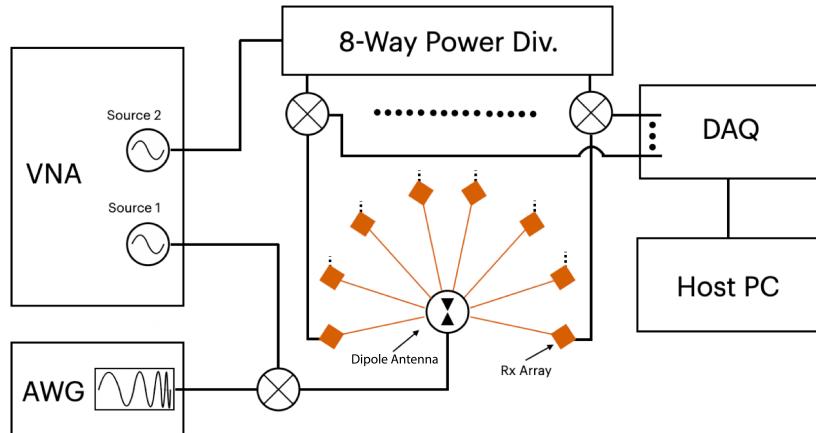


Figure 4.8. A system level diagram of the laboratory setup used for beamforming demonstrations at Penn State. For more information on this system see Chapter 5. Signals near 26 GHz are fed to a dipole antenna using and arbitrary waveform generator (AWG) and vector network analyzer (VNA), which drive a mixer. The dipole radiation is collected by an array of antennas connected to the digitizer data acquisition (DAQ) system.

An antenna measurement setup was constructed at Penn State to serve as a testbed for antenna prototypes and to perform laboratory validations of array simulations for the FSCD. This system is discussed in more detail in Chapter 5. Early versions of the antenna measurement system (see Figure 4.8 and Figure 4.9) were used to perform beamforming reconstruction studies of a simple probe antenna.

Signals from an arbitrary waveform generator were up-converted to 26 GHz using a mixer and a high-frequency source from a vector network analyzer and fed to a dipole

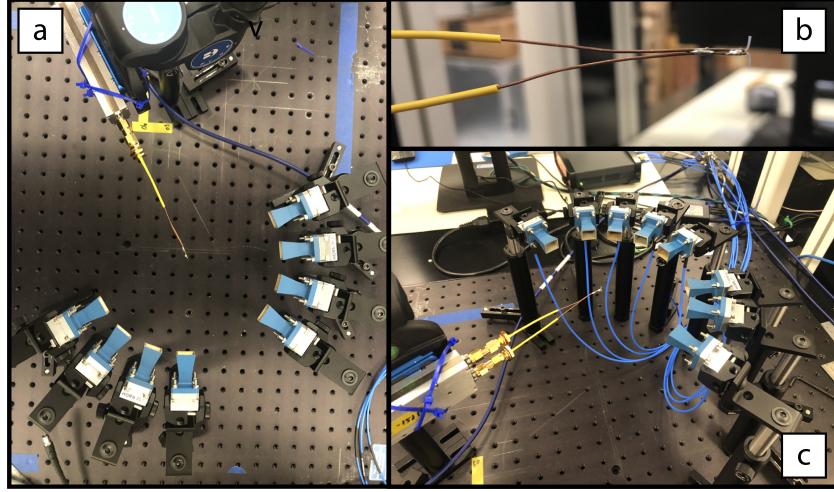


Figure 4.9. Photographs of the beamforming demonstration setup. In (a) I show a top-down view of the dipole antenna and the array of eight horn antennas. Manual repositioning of the horn antennas allows one to synthesize a full-circular antenna array. The dipole antenna is mounted on a camera tripod mount that allows for manual position tuning. (b) is a close up image of the dipole, which is manufactured from two segments of semi-rigid coaxial cable. (c) is another image of the dipole and array.

2557 antenna through a balun. The radiation from the dipole antenna was received by an
 2558 array of horn antennas. The signals from the horn antennas were down-converted to
 2559 baseband using a collection of mixers and an 8-way power divider. The signals were then
 2560 digitized and saved to a host computer for analysis.

2561 The data collected using the dipole and horn antenna array is reconstructed using the
 2562 beamforming reconstruction approach specified in Section 4.3.1. A two-dimensional grid
 2563 of xy-positions is defined and the beamforming phase shifts for each of these positions
 2564 is calculated. The phased summation can be visualized by plotting the time-averaged
 2565 power for each of the summations as a pixel in the resulting beamforming image (see
 2566 Figure 4.10). White Gaussian noise (WGN) can be added to the data at this stage
 2567 to simulate more realistic SNR if desired. The beamforming peak maxima is expected
 2568 to have a Bessel function shape due to the circular symmetry of the array, and by
 2569 analyzing the size of the beamforming maxima one can confirm that the beamforming
 2570 reconstruction measurement has similar position resolution as expected from Locust
 2571 simulations. Additionally, signal detection rates can be estimated from the data by
 2572 comparing the magnitude of the beamforming signal peak in the frequency spectra to
 2573 simulation.

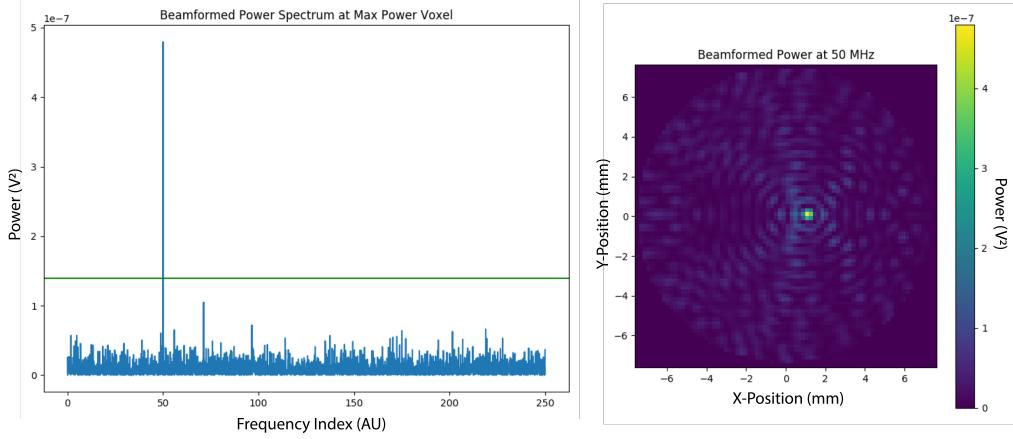


Figure 4.10. An example of digital beamforming reconstruction of a dipole antenna using a synthetic array of horn antennas. The beamforming image on the right is constructed by computing the time-averaged power of the summed signals for a two-dimensional grid of beamforming positions. In the image, one can see a clear maximum that corresponds to the position of the dipole antenna. On the left I show the frequency spectrum of the time-series at the maximum power pixel. White Gaussian noise is added to the signal to mimic a more realistic signal-to-noise-ratio. The signal emitted by the dipole is clearly visible as the high power peak in the frequency spectrum.

2574 FSCD Beamforming Simulations

2575 Locust simulations of the FSCD are used to generate simulated CRES signal data to
 2576 perform beamforming reconstruction studies. As mentioned in the previous section,
 2577 the beamforming procedure beings by specifying a set of beamforming positions and
 2578 corresponding beamforming shifts. The beamforming positions form a grid that covers
 2579 the region of interest. There are effectively an infinite number of ways to specify the
 2580 grid positions, however, uniform square grids are the most commonly used due to their
 2581 simplicity. In the actual experiment the number and pattern of beamforming positions
 2582 would be optimized to cover the most important regions of the trap volume, which
 2583 maximizes detection efficiency and minimizes superfluous calculations.

2584 The beamforming grids used for signal reconstruction with the FSCD consist of a set
 2585 of points that cover the two-dimensional plane formed by the perimeter of the antenna
 2586 array. The axial dimension is left out because electrons are treated as if they occupy only
 2587 their average axial position, which corresponds to the center of the magnetic trap. This
 2588 treatment is valid since it is impossible to resolve the axial position of the electron as a
 2589 function of time due to the rapid oscillation frequencies of trapped electrons.

2590 After beamforming, a summed time-series is obtained for each beamforming position
 2591 that can be check for a signal using a detection algorithm. A beamforming image is

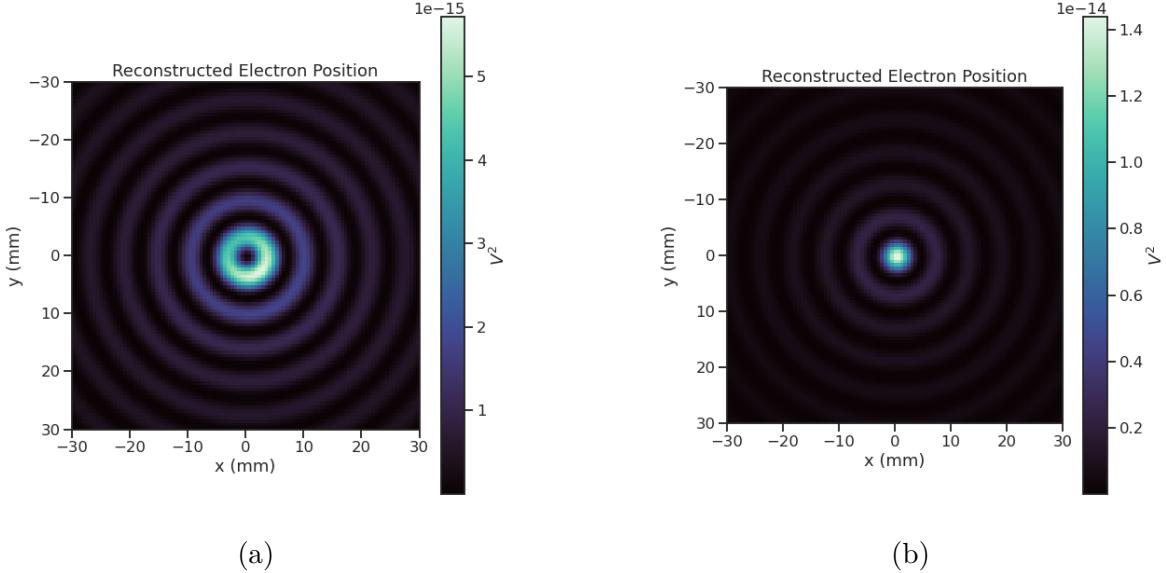


Figure 4.11. Beamforming images visualizing the reconstruction of an electron without (a) and with (b) the cyclotron phase correction. The images were generated using data from Locust simulations. The cyclotron phase refers to a phase offset equal to the relative azimuthal position of an antenna in the array. This phase offset is caused by the circular electron orbit and must be corrected for during reconstruction.

2592 a visualization method that is equivalent to arranging the beamforming grid points
 2593 according to their physical locations. Each pixel in the image corresponds to a summed
 2594 time-series obtained for a digital beamforming position, and the image is obtained taking
 2595 the time-averaged power at every pixel(see Figure 4.11).

2596 If only of the spatial beamforming phase component from Equation 4.20 is used, then
 2597 the resulting image contains a ring-shaped feature centered on the position of the electron
 2598 (see Figure 4.11a). The origin of this shape is an additional phase offset particular to
 2599 a cyclotron radiation source. The circular cyclotron orbitm introduces a relative phase
 2600 offset to the electric fields equal to the azimuthal position of the field measurement point.
 2601 Therefore, two antennas, one located at an azimuthal position of 0° and another located
 2602 at an azimuthal position of 90° , will recieve CRES signals out of phase by 90° , which is
 2603 the difference in their azimuthal positions. This phase offset can be corrected by adding
 2604 an additional term to the beamforming phase equation that is equal to the azimuthal
 2605 position of the antenna relative to the electron,

$$\phi_i[n] = \frac{2\pi d_i[n]}{\lambda} + \Delta\varphi_i[n], \quad (4.23)$$

2606 where $\Delta\varphi_i$ is difference between the azimuthal position of the electron and the i -th

2607 antenna channel. Using the updated beamforming phases changes the ring feature into
 2608 the expected Bessel peak whose maximum corresponds to the position of the electron.
 2609 Including this cyclotron phase correction significantly improves the signal detection and
 2610 reconstruction capabilities of beamforming by more than doubling the summed signal
 2611 power and shrinking the beamforming maxima feature size.

2612 The beamforming image examples in Figure 4.11 were produced using an electron
 2613 located on the central axis of the magnetic trap, which do not experience ∇B -drifts.
 2614 However, electrons produced at non-zero radial position the beamforming phases must
 2615 be made time-dependent to track the position of the electron's guiding center over
 2616 time. Without this correction the ∇B -drift causes the electron to move away from the
 2617 beamforming position, which effectively spreads the cyclotron radiation power over a
 wider area in the beamforming image (see Figure 4.12). This effect significantly reduces

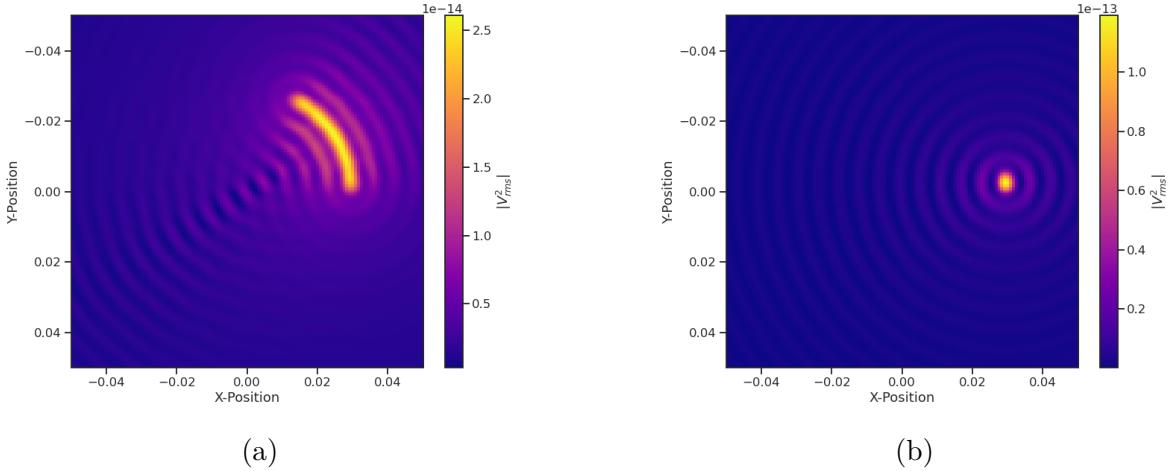


Figure 4.12. Beamforming images visualizing the reconstruction of an electron located off the central axis of the FSCD trap. In (a) beamforming is being performed without the ∇B -drift correction, and in (b) we include the ∇B -drift correction.

2618
 2619 the power of the beamforming maxima and increases the size of the beamforming features,
 2620 simultaneously harming detection efficiency and position reconstruction.

2621 The ∇B -drift correction simply adds a circular time-dependence to the beamforming
 2622 positions as a function of time,

$$r[n] = r_0 \quad (4.24)$$

$$\varphi[n] = \varphi_0 + \omega_{\nabla B} t[n], \quad (4.25)$$

2623 where $\omega_{\nabla B}$ is the drift frequency and $t[n]$ is the time vector. In the ideal case the ∇B -drift

frequencies from Figure 4.2 for the correct pitch angle and radial position would be used,
 however, it is not possible to know the electron's pitch angle a priori. In principle, one
 could perform multiple beamforming summations for a given beamforming position using
 different drift frequencies and choose the one that maximizes the summed power, but
 this approach leads to a huge computational burden that would be impractical for a
 real FSCD experiment. A compromise is to use an average value of $\omega_{\nabla B}$ obtained by
 averaging over the drift frequencies for electrons of different pitch angle at a particular
 radius. This approach keeps the computational cost of time-dependent beamforming to a
 minimum while still providing a significant increase in the detection efficiency of digital
 beamforming.

Signal Detection with Beamforming and a Power Threshold

Up to this point I have neglected a specific discussion of how digital beamforming is used
 for signal detection and reconstruction. Because, strictly speaking, digital beamforming
 consists only of the phased summation of the array signals and cannot be used alone for
 signal detection. The example beamforming images shown in Figure 4.11 and Figure 4.12
 were produced using simulated data that contained no noise, which significantly degrades
 the utility of analyzing the beamforming images for signal detection and reconstruction.

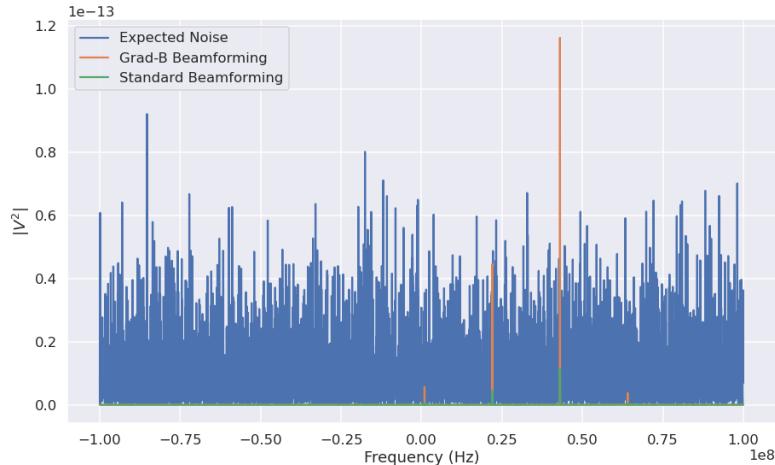


Figure 4.13. A plot of a typical frequency spectrum obtained by applying a Fourier transform to the time-series obtained from beamforming. The frequency spectra are plotted without noise on top of an example of a typical noise spectrum to visualize a realistic signal-to-noise ratio. In the example we see that without beamforming it would not be possible to detect anything since the signal amplitudes would be reduced by a factor of sixty relative to the noise. Additionally, we see that the ∇B -drift correction is needed to detect this electron since it comes from a simulation of an electron with a significant off-axis position.

2641 In Project 8, digital beamforming as a detection algorithm is understood to mean
 2642 digital beamforming plus a power or amplitude threshold placed on the frequency
 2643 spectrum obtained by applying a fast Fourier transform (FFT) to the summed time-series
 2644 (see Figure 4.13). This approach is similar to the time-frequency spectrogram analysis
 2645 employed in Phase I and II. However, it is possible to use any signal detection algorithm
 2646 after beamforming. In Section 4.4 I analyze the signal detection performance of the
 2647 power threshold approach in detail.

2648 Without a reconstruction technique that coherently combines the signals from the
 2649 full antenna the ability to detect CRES signals is drastically reduced (see Figure 4.13).
 2650 Because the CRES signals are in-phase at the correct beamforming position the summed
 2651 power increases as a function of N^2 compared to a single antenna channel, where N is
 2652 the number of antennas. It is true that the noise power is also increased by beamforming,
 2653 but, because the noise is incoherent, its power only increases linearly. Consequently, the
 2654 SNR of the CRES signal increases linearly with the number of antennas, which greatly
 2655 improves detection efficiency compared to using only the information in a single antenna.

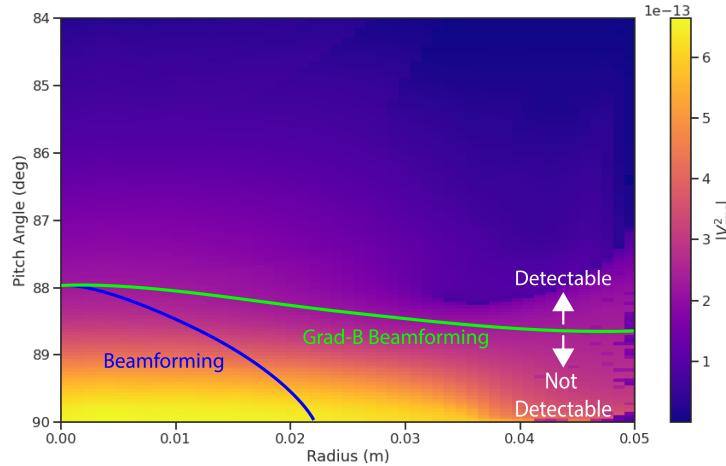


Figure 4.14. A plot of the total signal power received by the FSCD array from trapped electrons with different radial positions and pitch angles generated using Locust simulations. The lines on the plot indicate a 10 dB detection threshold above the mean value of the noise in the frequency spectrum. With static beamforming electrons with radial positions larger than about two centimeters are undetectable due to the change in the electron's position over time causing losses from beamforming phase mismatch. This is corrected by including ∇B -drift frequencies in the beamforming phases. Both beamforming techniques fail to detect electrons below $\approx 88.0^\circ$, since these signals are composed of several relatively weak sidebands that are comparable to the noise.

2656 The power threshold detection algorithm searches for high-power frequency bins that

2657 should correspond to a frequency component of the CRES signal. In order to prevent
 2658 random noise fluctuations from being mistaken as CRES signals the power threshold
 2659 must be set high enough so that it is unlikely that random noise could be responsible. A
 2660 consequence of this is that many electrons that can be trapped will go undetected because
 2661 the modulation caused by axial oscillations leads to the cyclotron carrier power to falling
 2662 below the decision threshold. The time-dependent beamforming used to correct for the
 2663 ∇B -drift increases the volume of the magnetic trap where electrons can be detected,
 2664 but it is ineffective at increasing the range of detectable pitch angles (see Figure 4.14).
 2665 Fundamentally, this is because the power threshold only uses a fraction of the signal
 2666 power to detect electrons and ignores the power present in the frequency sidebands. In
 2667 the subsequent sections I examine two other signal detection algorithms that seek to
 2668 improve the detection efficiency of the FSCD by utilizing the more of the signal shape to
 2669 compute the detection test statistics.

2670 **4.3.2 Matched Filtering**

2671 **Introduction to Matched Filtering**

2672 The problem of CRES signal detection is the problem of detecting a signal buried in
 2673 WGN, which has been examined at great depth in the signal processing literature [70].
 2674 For a fully known signal in WGN the optimal detector is the matched filter, which means
 2675 that it achieves the highest true positive rate for a fixed rate of false positives.

2676 The matched filter test statistic is calculated by taking the inner product of the data
 2677 with the matched filter template

$$\mathcal{T} = \left| \sum_n h^\dagger[n] y[n] \right|, \quad (4.26)$$

2678 where $h[n]$ is the matched filter template and $y[n]$ is the data. The matched filter test
 2679 statistic defines a binary hypothesis test in which the data vector is assumed to be an
 2680 instance of two possible data classes. By setting a decision threshold on the value of \mathcal{T} ,
 2681 one can classify a given data vector as belonging to two distinct hypotheses. Under the
 2682 first hypothesis the data is composed of pure WGN, and under the second hypothesis
 2683 the data is composed of the known signal with additive WGN.

2684 The matched filter template is obtained by rescaling the known signal in the following

2685 way

$$h[n] = \frac{x[n]}{\sqrt{\tau \sum_n x^\dagger[n]x[n]}}, \quad (4.27)$$

2686 where τ is the variance of the WGN and $x[n]$ is the known signal. Strictly speaking,
2687 Equation 4.27 is only true for noise with a diagonal covariance matrix, which is assumed
2688 to be true for the FSCD. Defining the matched filter templates in this way guarantees
2689 that the expectation value of \mathcal{T} is equal to one when the data contains only noise, which
2690 is the standard matched filter normalization.

2691 Although matched filters are canonically formulated in terms of a perfectly known
2692 signal, it is possible to apply the matched filter technique with imperfect information
2693 provided the signal is deterministic. From the discussion of CRES simulation tools (see
2694 Section 4.2) it was shown that the shape of CRES signals are completely determined
2695 by the initial parameters of the electron. The random collisions with background gas
2696 molecules, which cause the formation of signal tracks, are the only stochastic component
2697 of the CRES event after the initial beta-decay. Therefore, a matched filter can be used for
2698 the detection of CRES signal tracks between scattering events, which are fully determined
2699 by the initial parameters of the electron.

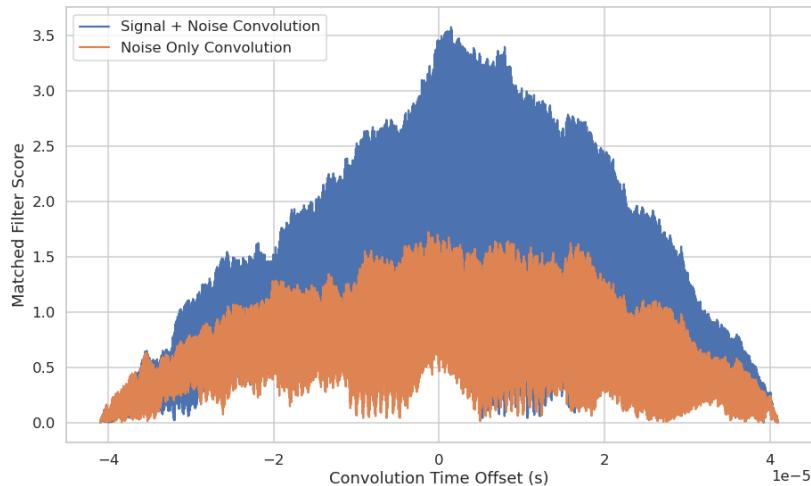


Figure 4.15. Example of a convolution of a CRES signal template with a segment of noisy data. A simulated CRES signal was simulated using Locust and normalized to create a matched filter template. When this template is convolved with noisy data the contains the matching signal the convolution output increases dramatically compared to data with only noise. The decreasing convolution output as the time offset of the convolution increases is caused by zero-padding of the data and template.

2700 The matched filter test statistic for CRES signals is a modified version of Equation

2701 4.26

$$\mathcal{T} = \max_{\mathbf{h}, m} |\mathbf{h} * \mathbf{y}| = \max_{\mathbf{h}, m} \left| \sum_k h^\dagger[k] x[m - k] \right|, \quad (4.28)$$

2702 where the matched filter inner product has been replaced with a convolution operation
2703 and a maximization over the template and convolution delay (m). Replacing the inner
2704 product with a convolution accounts for the fact that the start time of the CRES signal is
2705 now an unknown parameter, in addition, a maximization of the matched filter convolution
2706 is performed over a number of different templates. Because the shape of the signal is
2707 unknown, a range of different signal shapes, called a template bank, must be checked
2708 using an exhaustive search.

2709 The template bank approach, while powerful, can become computationally intractable.
2710 Specifically, the time-domain convolution specified by Equation 4.28 is particularly
2711 computationally intensive and is a major barrier towards the implementation of a
2712 matched filter for signal detection in an experiment like the FSCD. This can be avoided
2713 by using the convolution theorem to replace the time-domain convolution with an inner
2714 product in the frequency domain.

2715 The convolution theorem states that

$$\mathbf{f} * \mathbf{g} = \mathcal{F}^{-1}(\mathbf{F} \cdot \mathbf{G}) \quad (4.29)$$

2716 where \mathbf{f} and \mathbf{g} are discretely sampled time-series, \mathbf{F} and \mathbf{G} are the respective discrete
2717 Fourier transforms, and \mathcal{F}^{-1} is the inverse discrete Fourier transform operator. The
2718 convolution theorem allows us to perform the matched filter convolution by first com-
2719 puting the Fourier transform of the template and data, then performing a point-wise
2720 multiplication of the two frequency series, and finally performing the inverse Fourier
2721 transform to obtain the convolution output. Because discrete Fourier transforms can be
2722 performed extremely efficiently, the convolution theorem is almost always used in lieu of
2723 directly computing the convolution.

2724 One thing to note here is that the convolution theorem for discrete sequences shown
2725 here, is technically valid only for circular convolutions, which is not directly specified
2726 in Equation 4.28. However, because typical CRES track lengths are much longer than
2727 the Fourier analysis window and the frequency chirp rates are small compared to the
2728 time-slice duration, it is safe to use circular convolutions to evaluate matched filter scores
2729 for CRES signals, which allows one to apply the convolution theorem to compute matched
2730 filter scores for the FSCD.

2731 **Matched Filter Analysis of the FSCD**

2732 Since the matched filter is the optimal signal detection approach, it provides the ultimate
2733 upper bounds on signal detection. This makes it a useful algorithm for assessing the
2734 upper bounds on neutrino mass sensitivity for the FSCD, since it indicates the best
2735 possible detection efficiency achievable for that experiment configuration. The standard
2736 approach to performing these studies involves generating numerous simulated electron
2737 signals that span the kinematic parameter space of electrons.

2738 To limit the number of simulations required to evaluate the detection efficiency,
2739 the standard approach is to fix the starting axial position, starting azimuthal position,
2740 starting direction of the perpendicular component of the electron’s momentum, and event
2741 start time. This reduces the dimensionality of the simulated parameter space to three
2742 parameters — the starting radial position, starting kinetic energy, and starting pitch
2743 angle. The fixed variables are nuisance parameters, which do not affect the detection
2744 efficiency estimates for the FSCD design, because they simply introduce overall phase
2745 offsets that can be marginalized during the calculation of the matched filter score. Across
2746 radial position, kinetic energy, and pitch angle one defines a regular grid of parameters
2747 and uses Locust to simulate the corresponding signals (see Figure 4.16). This grid of
2748 simulated signals is used to estimate detection efficiency by calculating the detection
2749 probability of a randomly parameterized signal using the grid as a set of matched filter
2750 templates (see Section 4.4).

2751 The matched filter approach can also be used to estimate the achievable energy
2752 resolution of the experiment by using a dense grid of templates generated with parameters
2753 close to the unknown signal (see figure 4.17). Because matched filter templates with similar
2754 parameters have closely matching signal shapes, templates with incorrect parameters can
2755 have nearly identical matched filter scores as the correct template. Since only one sample
2756 of noise is included in a sample of real data, one cannot guarantee that the template
2757 with the maximum score corresponds to the ground truth parameters of the signal. This
2758 introduces uncertainty into the signal parameter estimation that manifests as an energy
2759 broadening. Dense grids of matched filter templates allow one to quantify this broadening
2760 by analyzing the parameter space of templates with matched filter scores close to the
2761 ground truth. This approach is analogous to maximum likelihood estimation and is one
2762 key component of a complete sensitivity analysis for an antenna array CRES experiment.

2763 A figure of merit that summarizes the performance of a matched filter template
2764 bank at signal detection is the mean match, which is defined as the average ratio of the
2765 highest matched filter score for a random signal to the matched filter score for a perfectly

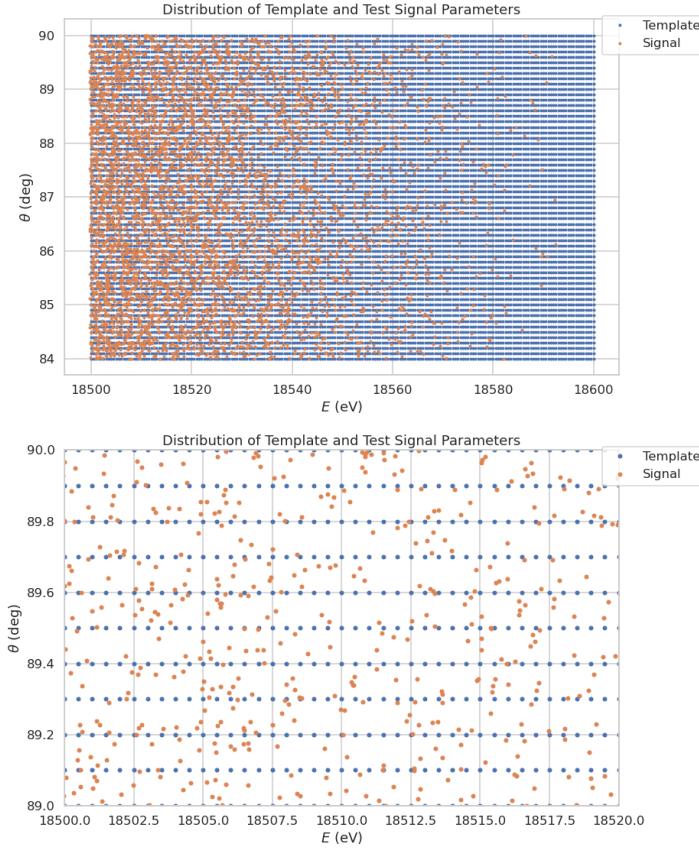


Figure 4.16. An example two-dimensional parameter distribution of a matched filter template bank and random test signals. θ refers to the pitch angle of the electron and E is the kinetic energy. The template bank forms a regular grid of in pitch angle and energy, whereas, the test signals are uniformly distributed in pitch angle and follow the tritium beta-decay kinetic energy distribution. This is why there are fewer test signals at higher energies. The need for high match across the full parameter space prevents one from reducing the density of templates in this low activity region. A zoomed in version of the template bank illustrates the relative density of templates and signals needed for match $> 90\%$.

2766 matching template. In equation form this is

$$\text{Match} \equiv \Gamma = \frac{\mathcal{T}_{\text{best}}}{\mathcal{T}_{\text{ideal}}}, \quad (4.30)$$

2767 where $\mathcal{T}_{\text{best}}$ is the matched filter score of the best fitting template in the bank and $\mathcal{T}_{\text{ideal}}$
 2768 is the hypothetical score one would measure if the signal perfectly matched the template.
 2769 Generally, one desires an average match as close to unity as possible. The mean match is
 2770 typically an exponential function of the number of templates in the template bank (see
 2771 Figure 4.18)..

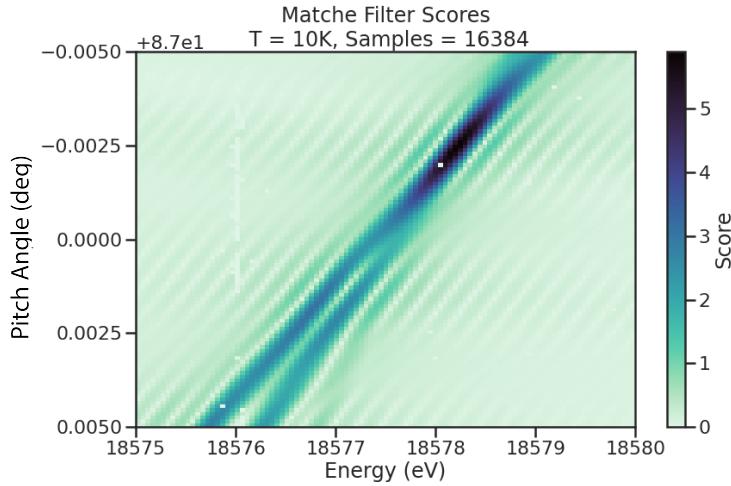


Figure 4.17. The matched filter scores of a dense grid of templates in pitch angle energy space. Dense template grids allow one to estimate the kinetic energy of the electron by identifying the best matching template. The uncertainty on this value is proportional to the space of templates that also match the test signal well. In the worst case matched filter templates can be completely degenerate where templates with different parameters match a signal with equal likelihood.

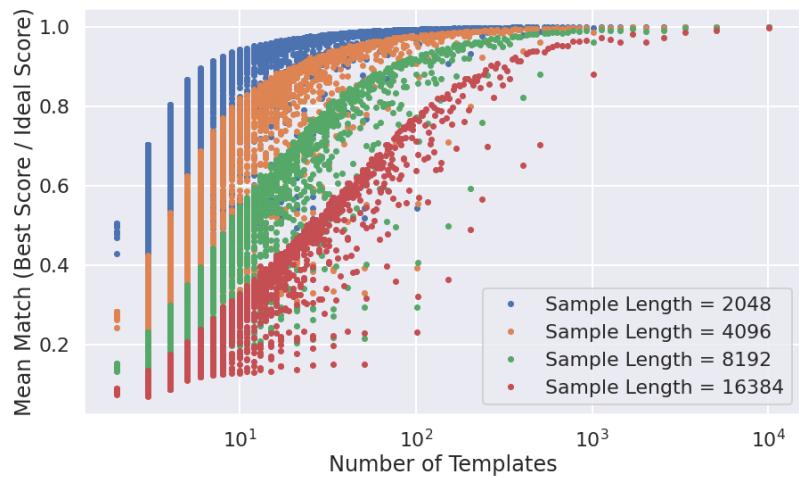


Figure 4.18. The mean match of the dense template grid shown in Figure 4.17 for different numbers of templates. Grids of different sizes were obtained by decimating a dense grid of templates and the average match for each grid was computed using the same set of randomly distributed test signals. Plotting the mean match against the size of the grid allows one to visualize the exponential relationship between match and template bank size. The noise in each curve is caused by sampling effects from the decimation algorithm. In general, longer templates are harder to match than shorter templates.

2772 The exponential relationship between match and template bank size manifests for
 2773 dense and sparse template grids. Sparse template grids are used for signal detection when
 2774 no prior information on the signal is available, whereas, dense templates grids are more
 2775 useful for parameter estimation. The mean match value directly influences the detection
 2776 efficiency of the template bank, but due to the exponential scaling, achieving a high
 2777 average match at the detection stage can easily overwhelm the available computational
 2778 resources.

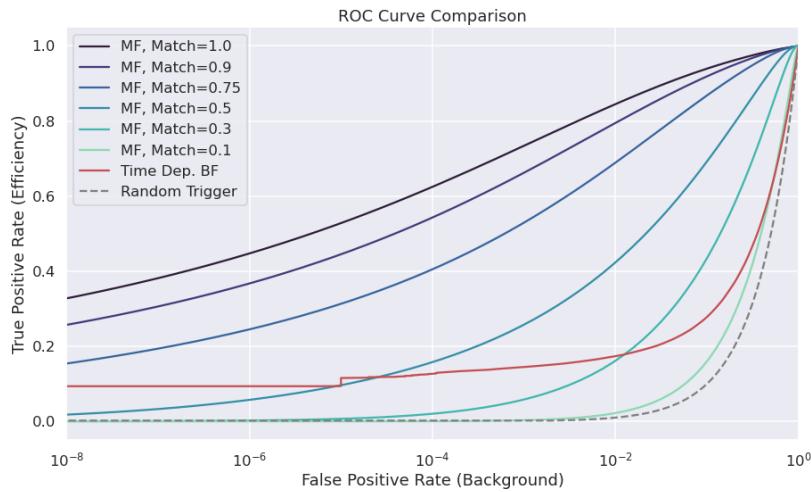


Figure 4.19. Matched filter template bank ROC curves as a function of mean match. One can see that for low match a matched filter is on average worse than the more straight forward beamforming detection approach.

2779 The effect of match on the detection efficiency of the matched filter template bank can
 2780 be summarized using the ROC curve (see Figure 4.19). A single ROC curve is obtained
 2781 by averaging over the PDFs that describe the detection probabilities of each individual
 2782 template.

2783 The distribution that describes the matched filter score under the signal hypothesis is
 2784 a Rician distribution, which has a mean value equal to the matched filter score multiplied
 2785 by the match ratio (see Section 4.4). Alternatively, the distribution of the matched
 2786 filter score when there is no signal in the data follows a Rayleigh distribution, which is
 2787 equivalent to a Rician distribution with zero mean. The matched filter score for each
 2788 template in the template bank is described by a separate Rician distribution. Therefore,
 2789 one way to model detection probability for a given signal is to average across all matched
 2790 filter distributions in the template bank to obtain a single distribution that describes the
 2791 statistical behavior of the matched filter score.

2792 A different way to visualize the detection performance for each algorithm is to specify
 2793 a minimum acceptable false positive rate at the trigger level. This is equivalent to
 2794 specifying a minimum threshold on the value of the matched filter score or the size of a
 2795 frequency peak for a beamforming power threshold trigger. One can then draw regions
 2796 of detectable signals as a function of the electron's pitch angle and radial position (see
 Figure 4.20). A kinetic energy shift is equivalent to an overall frequency shift of the

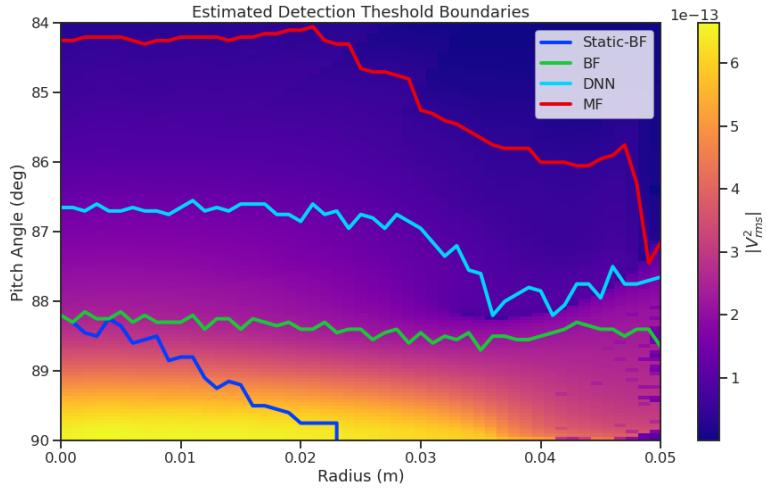


Figure 4.20. Boundaries of detectable electrons in pitch angle kinetic energy space for a series of different signal detection algorithms. A detectable signal is defined as a signal that is above a consistent decision with at least 50% probability. This non-rigorous treatment of detection probability is primarily useful for the visualization the relative increases in detection performance provided by the different algorithms. The static beamforming (Static-BF) algorithm is the digital beamforming algorithm introduced above without the ∇B -drift correction. The DNN algorithm refers to a convolutional neural network classifier trained to detect CRES signals (see Section 4.3.3).

2797
 2798 signal and should have no effect on the detection probability assuming sufficient density
 2799 of matched filter templates in the energy dimension. A electron is declared "detectable"
 2800 for the regions in Figure 4.20 if the signal has at least 50% probability of falling above the
 2801 decision threshold of the respective classifier. One can see that the parameter space of
 2802 detectable signals is greatly expanded beyond the beamforming power threshold trigger
 2803 with a matched filter (MF) or deep neural network (DNN) (see Section 4.3.3). Plots such
 2804 as Figure 4.20 are useful for visualization, but, since the handling of detection likelihood
 2805 is not sufficiently rigorous, the detection probability boundaries are not well-suited to
 2806 sensitivity estimates.

2807 **Optimized Matched Filtering Implementation for the FSCD**

2808 The biggest practical obstacle to the implementation of a matched filter template bank is
2809 the computational cost associated with exhaustively calculating the matched filter scores,
2810 therefore, one must employ several optimizations in a practical setting.

2811 Computing a matched filter score requires the convolution of two vectors, which can
2812 be performed very efficiently by computers if the convolution theorem and fast Fourier
2813 transforms (FFT) are utilized. Furthermore, one can apply digital beamforming as a
2814 pre-processing step to reduce the dimensionality of the data before the matched filter.
2815 In order to understand the relative gain in computational efficiency offered by these
2816 optimizations I analyze the total number of floating-point operations (FLOP) of several
2817 matched filter implementations in big O notation that utilize different combinations of
2818 optimizations.

2819 A direct implementation of a matched filter as specified by Equation 4.28 involves
2820 the convolution of N_{ch} signals of length N_s with template signals of length N_t . As a
2821 uniform metric we shall compare the FLOPs of the various matched filter implementations
2822 on a per-template basis, since each implementation scales linearly with the number of
2823 templates. The direct convolution approach to matched filtering costs

$$O(N_{\text{ch}}) \times O(N_s \times N_t) \quad (4.31)$$

2824 FLOP per-template, whose cost is dominated by the $O(M \times N)$ convolution operation.

2825 The computational cost of the direct matched filter approach can be significantly
2826 reduced by exploiting the convolution theorem and FFT algorithms. By restricting oneself
2827 to signals and templates that contain equal numbers of samples, the convolution can be
2828 calculated by Fourier transforming both vectors, performing the point-wise multiplication,
2829 and taking the inverse Fourier transform to obtain the convolution result. The FFT
2830 algorithm is able to compute the Fourier transform utilizing only $O(N \log N)$ operations.
2831 This optimization results in a computational cost per-template of

$$O(N_{\text{ch}}) \times O(N_s \log N_s) \quad (4.32)$$

2832 A typical signal vector in the FSCD contains $O(10^4)$ samples in which case the FFT
2833 reduces the computational cost of the matched filter by a factor of $O(10^3)$. This large
2834 reduction in computational cost implies that a direct implementation of a matched filter
2835 is completely infeasible in the FSCD due to resource constraints.

2836 Rather than relying solely on the matched filter it is tempting to consider using
 2837 digital beamforming as an initial step in the signal reconstruction for the purposes of
 2838 data reduction. The primary motivation is to reduce the dimensionality of the data by
 2839 a factor of N_{ch} by combining the array outputs coherently into a single channel. One
 2840 can view the beamforming operation as a partial matched filter, in the sense that the
 2841 matched filter convolution contains the beamforming phased summation along with a
 2842 prediction of the signal shape. By separating beamforming from the signal shape one
 2843 hopes to reduce the overall computational cost by effectively shrinking the number of
 2844 templates and reducing the number of operations required to check each one.

2845 The nature of this optimization requires that one account for the number of templates
 2846 used for pure matched filtering versus the hybrid approach. To first order, the total
 2847 number of templates at the trigger stage is a product of the number of guesses for each
 2848 of the electron's parameters

$$N_T = N_E \times N_\theta \times N_r \times N_\varphi, \quad (4.33)$$

2849 where N_E is the number of kinetic energies, N_θ is the number of pitch angles, N_r is the
 2850 number of starting radial positions, and N_φ is the number of starting azimuthal positions.
 2851 The starting axial position and cyclotron motion phase are not necessary to include in
 2852 the template bank, since these parameters manifest themselves as the starting phase of
 2853 the signal, which is effectively marginalized when using a FFT to compute the matched
 2854 filter convolution. Therefore, the total number of operations required by a matched filter
 2855 to detect a signal in a segment of array data is on the order of

$$O(N_T) \times O(N_{\text{ch}}) \times O(N_s \log N_s) \quad (4.34)$$

2856 With the hybrid approach one removes spatial parameters from the template bank
 2857 by using beamforming to combine the array signals into a single channel. Beamforming
 2858 explicitly assumes a starting position, which allows one to use matched filter templates
 2859 that span the two-dimensional space of kinetic energy and pitch angle. The total
 2860 computational cost of the hybrid method is directly proportional to the number of
 2861 beamforming positions. For the time-dependent beamforming defined in Section 4.3.1,
 2862 the number of beamforming positions is given by

$$N_{\text{BF}} = N_r \times N_\varphi \times N_{\omega_{\nabla B}}, \quad (4.35)$$

2863 where N_r and N_φ are the same spatial parameters encountered in the pure matched
 2864 filter template bank and $N_{\omega_{\nabla B}}$ is the number of ∇B -drift frequency assumptions. If a
 2865 unique drift frequency is used for each pitch angle then the hybrid approach is effectively
 2866 equivalent to a pure matched filter in the number of operations. The key efficiency gain
 2867 of the hybrid approach is to exploit the relatively small differences in $\omega_{\nabla B}$ for electrons
 2868 of different pitch angles by using only a few average drift frequencies.

2869 The total number of operations for the hybrid approach can be expressed as a sum of
 2870 the operations required by the beamforming and matched filtering steps,

$$O(N_{BF}) \times O(N_{ch}N_s) + O(N_{BF}) \times O(N_E N_\theta) \times O(N_s \log N_s). \quad (4.36)$$

2871 The first product in the sum is the number of operations required by beamforming,
 2872 which is simply the number of beamforming points times the computational cost of the
 2873 beamforming matrix multiplication, and the second product is the computational cost
 2874 of matched filtering the summed signal generated by each beamforming position. To
 2875 compare this to pure matched filtering we take the ratio of Equations 4.34 and 4.36 to
 2876 obtain

$$\Gamma_{BFMF} = \frac{O(N_{\omega_{\nabla B}})}{O(N_E N_\theta) \times O(\log N_s)} + \frac{O(N_{\omega_{\nabla B}})}{O(N_{ch})}. \quad (4.37)$$

2877 This expression can be simplified by observing that $O(N_E N_\theta) \times O(\log N_s) \gg O(N_{ch})$,
 2878 which means that the ratio of computational cost for the two methods can be reduced to

$$\Gamma_{BFMF} \approx \frac{O(N_{\omega_{\nabla B}})}{O(N_{ch})}. \quad (4.38)$$

2879 Limiting oneself to a number of estimated drift frequencies of $O(1)$, then it can be seen
 2880 that the estimated computational cost reduction of the hybrid approach is of $O(N_{ch})$.
 2881 This is a large reduction considering that the FSCD antenna array contains sixty antennas
 2882 in the baseline design.

2883 The main drawback of the hybrid approach is that the limited number of allowed
 2884 drift frequency guesses can lead to detection efficiency loss due to phase mismatch. The
 2885 degree of phase error from an incorrect drift frequency is proportional to the length of
 2886 the array data vector used by the signal detection algorithm. For signals with lengths
 2887 equal to the baseline FSCD Fourier analysis window of 8192 samples, typical phase errors
 2888 from using an average versus the exact ∇B -drift frequency are on the order of a few
 2889 percent in terms of the signal energy. This has a relatively small impact on the overall
 2890 detection efficiency, however, future experiments with antenna array CRES will want to

balance optimizations such as these during the design phase to keep experiment costs to a minimum while still achieving scientific goals.

Kinetic Energy and Pitch Angle Degeneracy

Accurate modeling of a matched filter requires one to consider the effects of mismatched signals and template, since this more accurately reflects the real-world usage of a matched filter. One way to study this is to use a signal grid to compute the matched filter scores between mismatched signals and templates and evaluate the matched filter scores under this scenario. What one finds when performing this analysis is that templates for signals with incorrect parameters can have matched filter scores that are indistinguishable from the matched filter score of the correct template (see Figure 4.21 and Figure 4.21).

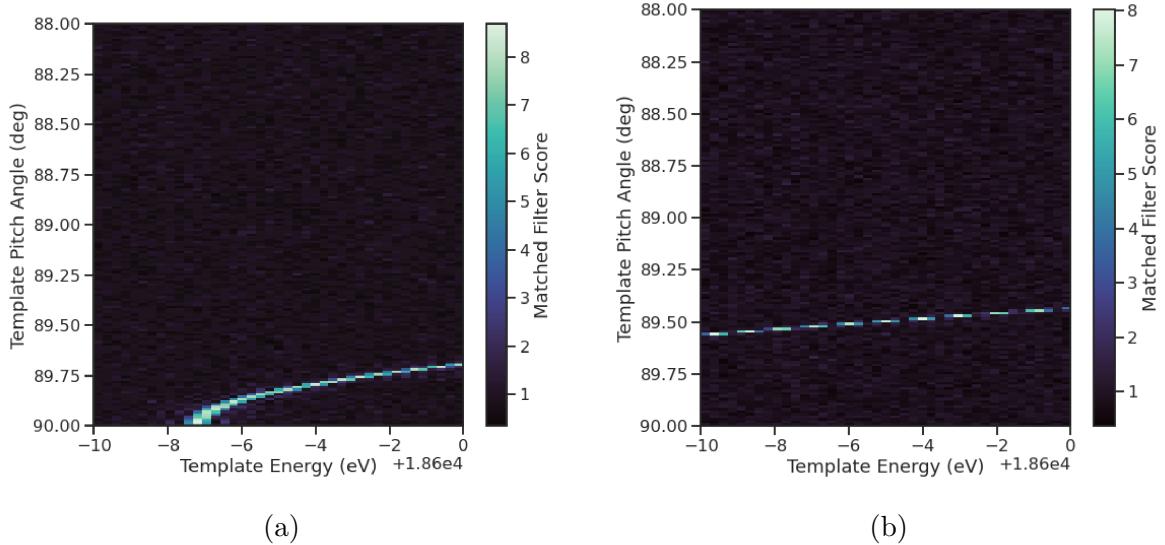


Figure 4.21. Two example illustrations of the correlation between kinetic energy and pitch angle imparted by the shape of the FSCD magnetic trap. The correlations manifest themselves as degeneracies in the matched filter score where multiple matched filter templates have the same matched filter for a particular signal. These degeneracies are a sign that the magnetic trap must be redesigned in order to break the correlation between pitch angle and kinetic energy.

This degeneracy in matched filter score is the result of correlations between the kinetic energy and pitch angle of the electron caused by the magnetic trap. These correlations are unacceptable since they greatly reduce the energy resolution of the experiment by causing electrons with specific kinetic energy to templates across a wide range of energies. It is important to emphasize that this degeneracy cannot be fixed by implementing a different signal reconstruction algorithm. As revealed by the matched filter scores the shapes of the signals for different parameters are identical. Resolving this degeneracy

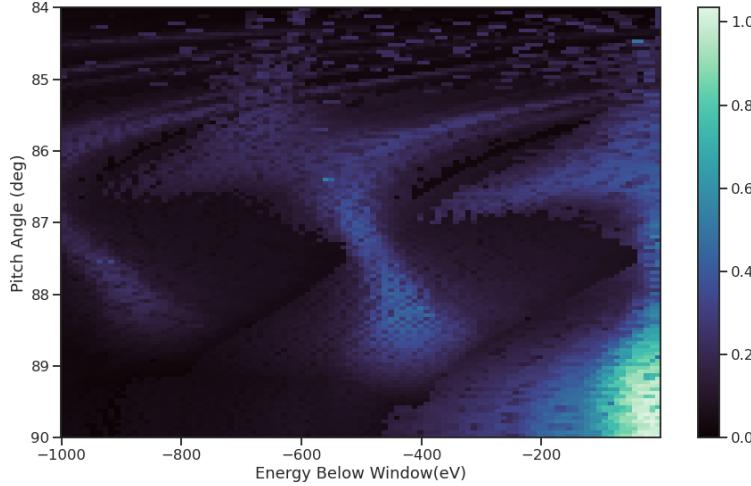


Figure 4.22. A visualization of the correlation between energy and pitch angle in the FSCD magnetic trap. The image is formed by computing the match of the best template from a grid consisting of pitch angles from 84 to 90 degrees in steps of 0.05 degrees, kinetic energies from 17574 to 18574 eV, located at 2 cm from the central axis, and simulated for a length of three FSCD time-slices. The signals used to compute the best matching template consisted of a grid from 84 to 90 degrees in steps of 0.05 degrees, kinetic energies from 18550 to 18575 eV in steps of 0.25 eV, located 2 cm from the central axis, and simulated for three FSCD time-slices. The colored regions of the plot show how well signals with lower energy can match those of higher energy for the FSCD magnetic trap, which is proportional to the achievable energy resolution of the FSCD design.

2908 between pitch angle and energy requires the design of a new magnetic trap with steeper
 2909 walls so that the average magnetic field experienced by an electron is less dependent on
 2910 pitch angle.

2911 4.3.3 Machine Learning

2912 Machine learning is a broad field of research [73] that has been particularly transformative
 2913 in the recent past. In this Section I provide a brief introduction to some concepts and
 2914 techniques of machine learning that were applied to CRES signal detection in my
 2915 dissertation.

2916 Introduction to Machine Learning

2917 Digitization of the FSCD antenna array generates large amounts of data that must
 2918 be rapidly processed for real-time signal detection and reconstruction. While digital
 2919 beamforming combined with a power threshold is relatively computationally inexpensive,

it is ineffective at detecting CRES signal with small pitch angles, since it relies on a visible frequency peak above the noise. On the other hand, a matched filter is able to detect signals with a significantly larger range of parameters, however, the exhaustive search of matched filter templates can be computationally expensive. Machine learning based triggering algorithms have been used successfully in many high-energy physics experiments [74], and recently have shown success in the detection of gravitational wave signals [75, 76] in place of more traditional matched filtering methods. The success of machine learning in these domains motivates the exploration of machine learning as a potential CRES signal detection algorithm.

Various approaches to machine learning are possible, but the one most important to our discussion here is the supervised learning approach. In supervised learning, one uses a differentiable model or function that is designed to map the input data to the appropriate label [73]. The data is represented as a multidimensional matrix of floating point values such as an image or a time-series, and the label is typically a class name such as signal or noise for classification problems, or a continuous value like kinetic energy for regression problems.

In supervised learning the model is trained to map from the data to the correct label by evaluating the output of the model using a training dataset consisting of a set of paired data and labels. To evaluate the difference between the model output and the correct label a loss function is used to quantify the error between the model prediction and the ground truth. For example, a common loss function in regression problems is the squared error loss function, which quantifies error using the squared difference between the model output and label.

Using the outputs of the loss function the next step in supervised learning is to compute the gradient of error with respect to the model parameters in a process called backpropagation. The gradients are used to update the model parameter values in order to minimize errors in the model predictions across the whole dataset. This loop is performed many times while randomly shuffling the dataset until the error converges to a minimum value at which point the training procedure has finished. It is standard practice to monitor the training procedure by evaluating the performance of the model using a separate validation dataset that matches the statistical distribution of the training data and to check the performance of the model after training using yet another dataset called the test dataset. These practices help to guard against overtraining which is a concern for models with many parameters.

2954 **Convolutional Neural Networks**

2955 A popular class of machine learning models are neural networks. A neural network is
2956 a function composed of a series of linear operations called layers, which take a piece of
2957 data typically represented as a matrix, multiply the elements of the data by a weight,
2958 and then sums these products to produce an output matrix. Neural networks composed
2959 of purely linear operations are unable to model complex non-linear behavior, therefore,
2960 non-linear activation functions are applied to the outputs of each of the layers to increase
2961 the ability of the neural network to model complex relationships between the data.

2962 Neural networks are typically composed of at least three layers, but with the present
2963 capabilities of computer hardware they typically contain much more than this. The first
2964 layer in a neural network is called the input layer, because it takes the data objects as
2965 input, and the last layer in a neural network is known as the output layer. The output
2966 layer is trained by machine learning to map the data an output label using the supervised
2967 learning procedure described in Section 4.3.3. Between the input and the output layer
2968 are typically several hidden layers that receive inputs from and transmit outputs to other
2969 layers in the neural network model. The term deep neural network (DNN) refers to those
2970 neural networks that have at least one hidden layer, which have proven to be extremely
2971 powerful tools for pattern recognition and function approximation.

2972 An important type of DNN are convolutional neural networks (CNN) that typically
2973 contain several layers which perform a convolution of the input with a set of filters. These
2974 convolution operations are typically accompanied by layers that attempt to down-sample
2975 the data along with the standard neural network activation functions. A standard CNN
2976 is composed of several convolutional layers at the beginning of the network and ends
2977 with a series of fully-connected neural network layers at the output. Intuitively, one
2978 can imagine that the convolutional layers are extracting features from the data that
2979 fully-connected layers use to perform the classification or regression task.

2980 **Deep Filtering for Signal Detection in the FSCD**

2981 CNNs have been extremely influential in the field of computer vision, particularly tasks
2982 such as image segmentation and classification, but have also been applied in numerous
2983 experimental physics contexts. Given the particular challenge posed by signal detection
2984 and reconstruction in the FSCD CNNs are an interesting choice for real-time signal
2985 detection, since this application requires both high efficiency and fast evaluation.

2986 In the machine learning paradigm, signal detection is a binary classification problem

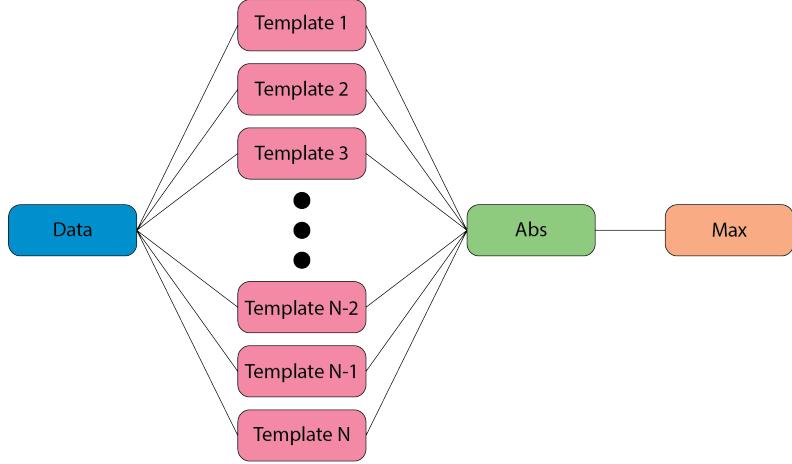


Figure 4.23. A representation of a matched filter template bank as a convolutional neural network. The network has a single layer composed of the templates, which act as convolutional filters. The activation of the neural network is an absolute value followed by a max operator.

2987 between the signal and noise data classes. My investigation focuses specifically on the
 2988 application of CNNs to signal detection in the FSCD, which is motivated by relatively
 2989 recent demonstrations of CNNs achieving classification accuracies for gravitational wave
 2990 time-series signals comparable to a matched filter template bank. In this framework
 2991 it is possible to interpret the matched filter as a type of CNN composed of a single
 2992 convolutional layer with the templates making up the layer filters (see Figure 4.23).
 2993 Since this neural network has no hidden layers, it is not a DNN, but one can attempt to
 2994 construct a proper CNN that attempts to reproduce the classification performance of the
 2995 matched filter network, which can be referred to as "deep filtering".

2996 The reason why deep filtering can be effective is that it may be possible to exploit
 2997 redundancies and correlations between templates, which allows one to perform signal
 2998 detection with similar accuracy but with fewer computations. This is relevant to real-time
 2999 detection scenarios like the FSCD experiment. In Section 4.4 I perform a detailed
 3000 comparison of the signal detection performance of a CNN to beamforming and a matched
 3001 filter template bank.

3002 Deep filtering is conceptually a simple technique. Similar to a matched filter template
 3003 bank, many simulated CRES signals are generated and used to train a model to distinguish
 3004 between signal and noise data (see Figure 4.24). To reduce the dimensionality of the
 3005 input FSCD data, a digital beamforming summation is applied to the raw time-series
 3006 data generated by Locust to compress the 60-channel data to a single time-series. CRES
 3007 signals have a sparse frequency representation and experiments training CNN's on time-
 3008 series and frequency-series data found that models trained on frequency spectrum data

3009 performed significantly better. Therefore, an FFT is applied to the summed time-series
 3010 before being normalized and fed to the classification model.

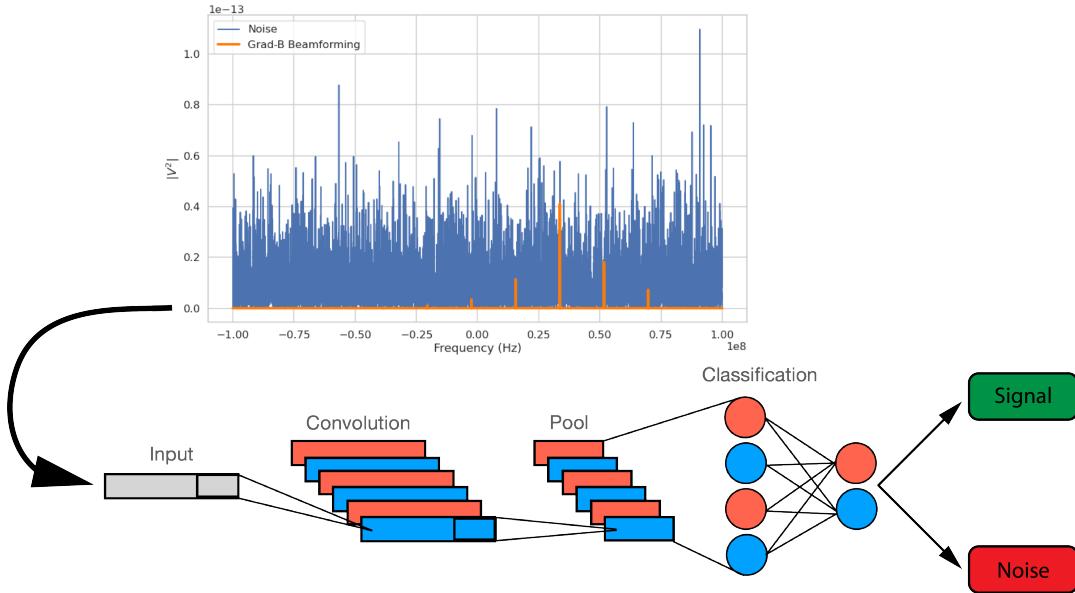


Figure 4.24. A graphical depiction of CRES signal detection using a CNN. A noisy segment of data is converted to a frequency series using digital beamforming and a FFT. The complex-valued frequency series is input into a trained CNN model that classifies the data as signal or noise using a decision threshold on the CNN output.

3011 The data used to train the model consists of an equal proportion of signal and noise
 3012 frequency spectra. Unique samples of WGN are generated and added to the signals during
 3013 training time to avoid have to pre-generate and store large samples of noise data. The
 3014 binary cross-entropy loss function combined with the ADAM optimizer proved effective
 3015 at training the models to classify CRES data. A simple hyperparameter optimization
 3016 was performed by manually tuning model, loss function, and optimizer parameters. The
 3017 model and training loops was implemented in python using the PyTorch deep learning
 3018 framework. Standard machine learning practices were followed when training the models,
 3019 such as overtraining monitoring using a validation dataset. Models were trained until the
 3020 training loss and accuracy converged and then evaluated using a separate test data set.

3021 The classification results of the test dataset are used to quantify the relationship
 3022 between the true positive rate and the false positive rate for the model. The true positive
 3023 rate is analogous to detection efficiency and the false positive rate is a potential source of
 3024 background in the detector. One can limit the rate of false positives using a sufficiently
 3025 high threshold on the model output at the cost of a lower detection efficiency (see Figure
 3026 4.25 and Figure 4.26). As expected, the performance of the model at signal classification

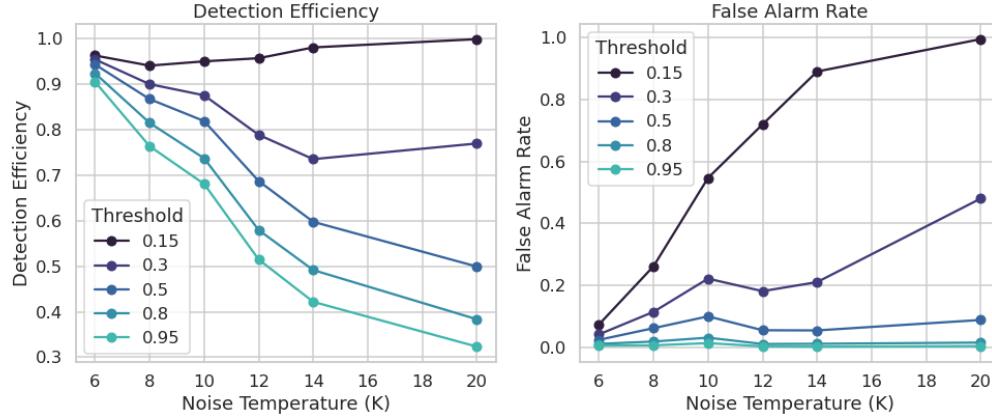


Figure 4.25. The detection efficiency and false alarm rate (false positive rate) as a function of the decision threshold for different values of the noise temperature. The model is trained to output a value close to one for data that contains a signal and outputs a value near zero when the data contains only noise. One sees that a lower decision threshold will have a high detection efficiency at the cost of a high rate of false alarms.

3027 is negatively effected the noise power, which is quantified by the noise temperature.

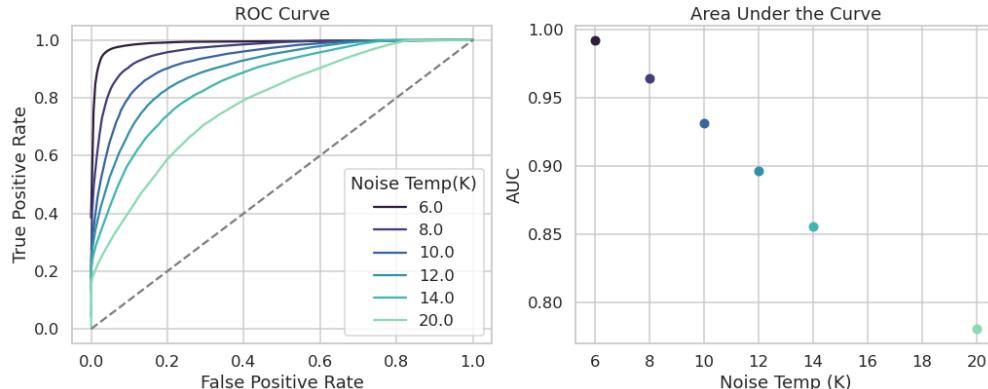


Figure 4.26. ROC curves for a CNN model classifying CRES signals. One can see that the area under the curve, which is a figure of merit that describes the performance of the classifier, is roughly linearly dependent with the noise temperature.

3028 4.4 Analysis of Signal Detection Algorithms for the FSCD

3029 This section contains an early version of the manuscript for the triggering paper prepared
 3030 for publication in JINST. I present a detailed analysis of the signal detection performance
 3031 of the three signal detection approaches discussed so far using a population of simulated

3032 CRES signals generated with Locust. The focus of the paper is on the performance of the
3033 signal detection algorithms for pitch angles below 88.5° where the beamforming power
3034 threshold is least effective.

3035 **4.4.1 Introduction**

3036 Cyclotron Radiation Emission Spectroscopy (CRES) is a technique for measuring the
3037 kinetic energies of charged particles by observing the frequency of the cyclotron radiation
3038 that is emitted as they travel through a magnetic field [39]. The Project 8 Collaboration
3039 is developing the CRES technique as a next-generation approach to tritium beta-decay
3040 endpoint spectroscopy for neutrino mass measurement. Recently, Project 8 has used
3041 CRES to perform the first ever tritium beta-decay energy spectrum and neutrino mass
3042 measurement [41, 42].

3043 Previous CRES measurements have utilized relatively small volumes of radiation source
3044 gas that are directly integrated with a waveguide transmission line, which propagates the
3045 cyclotron radiation emitted by magnetically trapped electrons to a cryogenic amplifier.
3046 While this technology has had demonstrable success, it is not a feasible option for scaling
3047 up to larger measurement volumes. In particular, the goal of the Project 8 Collaboration
3048 is to use CRES combined with atomic tritium to measure the neutrino mass with a
3049 40 meV sensitivity. Achieving this sensitivity goal will require a multi-cubic-meter scale
3050 measurement volume in order to obtain the required event statistics in the tritium
3051 beta-spectrum endpoint region; hence, there is a need for new techniques to enable large
3052 volume CRES measurements for future experiments.

3053 One approach is to surround a large volume with an array of antennas that together
3054 collect a portion of the cyclotron radiation emitted by trapped electrons [40, 77]. A
3055 promising design is an inward-facing uniform cylindrical array that surrounds the tritium
3056 containment volume. Increasing the size of the antenna array, by adding additional rings
3057 of antennas along the vertical axis, allows one to grow the experiment volume until a
3058 sufficient amount of tritium gas can be observed by the array. A challenging aspect of
3059 this approach is that the total radiated power emitted by an electron near the tritium
3060 spectrum endpoint is on the order of 1 fW or less in a 1 T magnetic field, which is then
3061 distributed among all antennas in the array. Because the CRES signal information is
3062 spread across the antenna array, detecting the presence of a CRES signal and determining
3063 the electron's kinetic energy requires reconstructing the entire array output over the
3064 course of the CRES event, posing a significant data acquisition and signal reconstruction
3065 challenge.

3066 Previous measurements with the CRES technique have utilized a threshold on the fre-
3067 quency spectrum formed from a segment of CRES time-series data. This algorithm relies
3068 on the detection of a frequency peak above the thermal noise background, which limits
3069 the kinematic parameter space of detectable electrons (see Section 4.4.2.2). Although a
3070 power threshold based classification was adequate for smaller detectors, improvements
3071 in detection efficiency are needed for better sensitivity to the neutrino mass. Better
3072 detection efficiency is possible by taking advantage of the deterministic CRES signal
3073 structure with a matched filter or machine learning based classifier [?]. In order to eval-
3074 uate the relative gains in detection efficiency that come from utilizing these algorithms for
3075 antennas, analytical models that describe the detection performance a power threshold
3076 and matched filter classifier are developed. In addition, a basic convolutional neural
3077 network (CNN) is implemented and tested as a first step towards the development of
3078 neural-network based classifiers for antenna array based CRES measurements. These
3079 results allow for a comparison between the estimated detection efficiencies of each of these
3080 methods, which are weighed against the associated computational costs for real-time
3081 applications.

3082 The outline of this paper is as follows. Section 4.4.2 is an overview of a prototyp-
3083 ical antenna array CRES experiment, and describes the approach to real-time signal
3084 identification. Section 4.5 develops models for the power threshold and matched filter
3085 algorithms and introduces the machine learning approach and CNN architecture. Section
3086 4.5.1 describes the process for generating simulated CRES signal data and the details of
3087 training the CNN. Finally, Section 4.5.2 compares the signal classification accuracy for
3088 the three approaches and discusses the relevant trade-offs in terms of detection efficiency
3089 and computational cost.

3090 **4.4.2 Signal Detection with Antenna Array CRES**

3091 **4.4.2.1 Antenna Array and Data Rate Estimates**

3092 In order to explore the potential of antenna array CRES for neutrino mass measurement,
3093 the Project 8 Collaboration has developed a conceptual design for a prototype antenna
3094 array CRES experiment [40, 77], called the Free-space CRES Demonstrator or FSCD (see
3095 Figure 4.27). The FSCD utilizes a single ring of antennas, which is the simplest form of a
3096 uniform cylindrical array configuration, to surround a radio-frequency (RF) transparent
3097 tritium gas vessel. A prototype version of this antenna array has been built and tested
3098 by the Project 8 collaboration [43] to validate simulations of the array radiation pattern

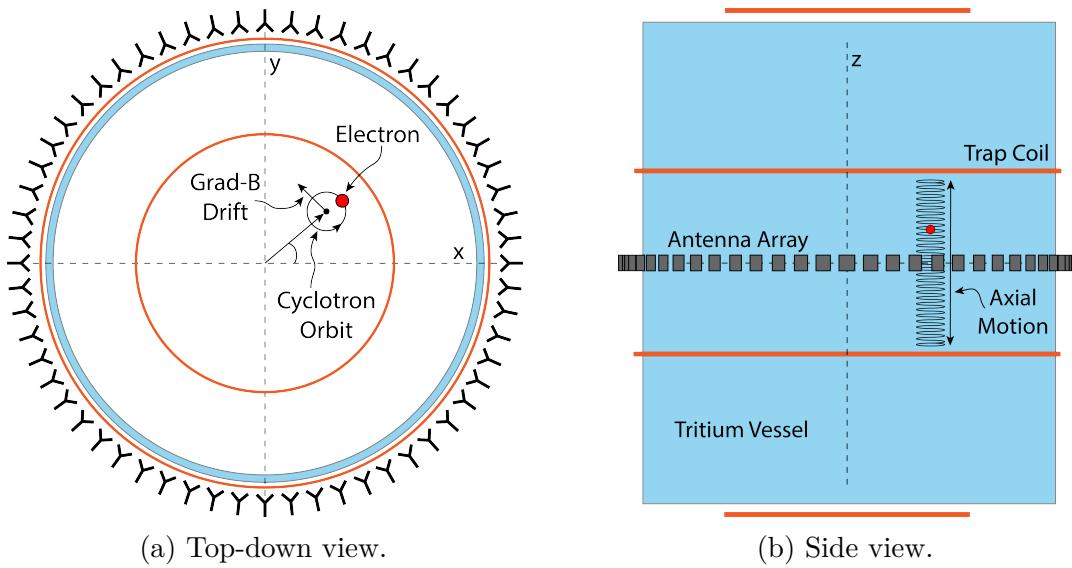


Figure 4.27. An illustration of the conceptual design for an antenna array CRES tritium beta-decay spectrum measurement. The antenna array geometry consists of a 20 cm interior diameter with 60 independent antenna channels arranged evenly around the circumference. The nominal antenna design is sensitive to radiation in the frequency range of 25-26 GHz, which corresponds to the cyclotron frequency of electrons emitted near the tritium beta-spectrum endpoint in a 0.96 T magnetic field. The array is located at the center of the magnetic trap produced by a set of current-carrying coils. The nominal magnetic trap design is capable of trapping electrons up to 5 cm away from the central axis of the array and traps electrons within an approximately 6 cm long axial region centered on the antenna array.

3099 and beamforming algorithms [?]. In the FSCD the antenna array is positioned at the
 3100 center of the magnetic trap formed by a set of electromagnetic coils, which create a local
 3101 minimum in the magnetic field with flat central region and steep walls in the radial and
 3102 axial directions.

3103 When an electron is trapped its motion consists of three primary components. The
 3104 component with the highest frequency is the cyclotron orbit whose frequency is determined
 3105 by the size of the background magnetic field. The FSCD design assumes a background
 3106 magnetic field value of approximately 0.96 T, which results in cyclotron frequencies
 3107 for electrons with kinetic energies near the tritium beta-spectrum endpoint of 26 GHz.
 3108 The component with the next highest frequency is the axial oscillation experienced by
 3109 electrons with pitch angles⁴ of less than 90° [60]. The flat region of the FSCD magnetic
 3110 trap extends approximately 3 cm above and below the antenna array plane causing
 3111 electrons to move back and forth as they are reflected from the trap walls. Typical

⁴Pitch angle is defined as the angle of the particle's total momentum with respect to the local magnetic field.

3112 oscillation frequencies are on the order of 10's of MHz, which results in an oscillation
 3113 period that is $O(10^3)$ smaller than the duration of a typical CRES event. Therefore, the
 3114 axial extent of the electron's motion is generally ignored for the purposes of reconstruction,
 3115 since the electron can be treated as if it is located in the average axial position at the
 3116 bottom of the magnetic trap. The component of motion with the smallest frequency
 3117 is the ∇B -drift caused by radial field gradients in the trap, producing an orbit of the
 3118 electron around the central axis of the trap with a frequency on the order of a few kHz,
 3119 dependent on the pitch angle and the radial position of the electron.

3120 Each component of motion influences the shape of the cyclotron radiation signals
 3121 received by the antenna array, therefore, the data acquisition (DAQ) system must be
 3122 properly designed in order to resolve the effects of the cyclotron motion, pitch angle, and
 3123 ∇B drift on the signal shape. Frequency down-conversion allows for intentional under-
 3124 sampling of the CRES signals at a nominal bandwidth of 200 MHz. The bandwidth is
 3125 required to be large enough to contain all sidebands produced by pitch angle modulation,
 3126 but must be limited to reduce the Nyquist-Johnson noise power for adequate signal-to-
 3127 noise ratio. The estimated noise temperature for the FSCD is ≈ 10 K, achievable with
 3128 low-noise HEMT amplifiers and cryogenic temperatures.

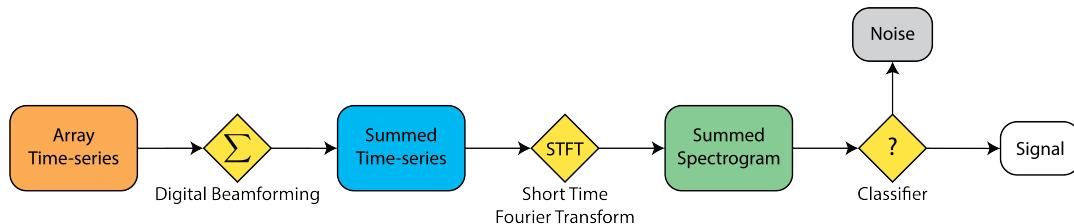


Figure 4.28. A block diagram illustration of the real-time triggering algorithm proposed for antenna array CRES reconstruction.

3129 A design goal for the FSCD DAQ system is to enable a significant portion of the
 3130 CRES event reconstruction to occur in real-time. The estimated data volume generated
 3131 by the FSCD is 1 exabyte of raw data per year of operation, with the nominal array size
 3132 of 60 antennas sampled at 200 MHz, which would be too expensive to store for offline
 3133 processing. Therefore, it is ideal to perform some CRES event reconstruction in real-time
 3134 so that it is possible to save a reduced form of the data for offline analysis. The first step
 3135 of the real-time reconstruction would be a real-time signal detection algorithm, which is
 3136 the focus of this paper. The basic approach consists of three operations performed on the
 3137 time-series data blocks including digital beamforming, a short time Fourier transform
 3138 (STFT), and a binary classification algorithm to distinguish between signal and noise

₃₁₃₉ data (see Figure 4.28).

₃₁₄₀ 4.4.2.2 Real-time Signal Detection

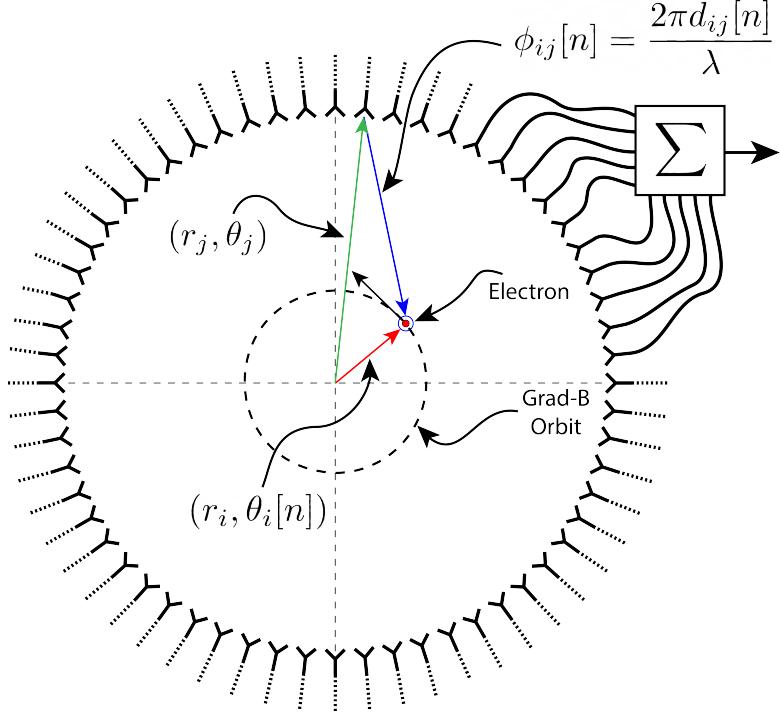


Figure 4.29. An illustration of the digital beamforming procedure. The blue lines indicate the distances from the beamforming position to each antenna. In the situation depicted the actual position of the electron matches the beamforming position, therefore, one expects constructive interference when the phase shifted signals are summed. To prevent the electron's ∇B -motion from moving the electron off of the beamforming position, the beamforming phases include time-dependence to follow the trajectory of the electron in the magnetic trap.

₃₁₄₁ The first step in the real-time detection algorithm is digital beamforming, which is a
₃₁₄₂ phased summation of the signals received by the array (see Figure 5.21). The phase shifts
₃₁₄₃ correspond to the path length differences between a spatial position and the antennas
₃₁₄₄ such that, when there is an electron located at the beamforming position, all the signals
₃₁₄₅ received by the array constructively interfere. Since one does not know a priori where an
₃₁₄₆ electron will be produced in the detector, a grid of beamforming positions is designed to
₃₁₄₇ cover the entire azimuthal plane where electrons can be trapped. A beamforming phased
₃₁₄₈ summation is performed for all points in the grid at each time-step. As shown in Section
₃₁₄₉ 4.4.2.1, the axial oscillation of the electrons prevents one from resolving its position along
₃₁₅₀ the z-axis, therefore, the beamforming grid need only cover the possible positions of the
₃₁₅₁ electron in the two-dimensional plane defined by the antenna array.

3152 Digital beamforming can be expressed as

$$\mathbf{y}[n] = \Phi^T[n]\mathbf{x}[n], \quad (4.39)$$

3153 where $\mathbf{x}[n]$ is the array snapshot vector at the sampled time n , $\Phi[n]$ is the matrix of
3154 beamforming phase shifts, and $\mathbf{y}[n]$ is the summed output vector that contains the
3155 voltages for each of the summed channels corresponding to a particular beamforming
3156 position. The elements of the beamforming phase shift matrix can be expressed as a
3157 weighted complex exponential

$$\Phi_{ij}[n] = A_{ij}[n] \exp(2\pi i \phi_{ij}[n]), \quad (4.40)$$

3158 where the indices i and j label the beamforming and antenna positions respectively. The
3159 weight A_{ij} accounts for the relative power increase for antennas that are closer to the
3160 position of the electron, and ϕ_{ij} is the total beamforming phase shift for the j -th antenna
3161 at the i -th beamforming position.

3162 The beamforming phase shift is a sum of two terms

$$\phi_{ij}[n] = \frac{2\pi d_{ij}[n]}{\lambda} + \theta_{ij}[n], \quad (4.41)$$

3163 where the first term is the phase shift originating from the path length difference ($d_{ij}[n]$)
3164 between the beamforming and antenna positions, which are represented by the vectors
3165 (r_j, θ_j) and $(r_i, \theta_i[n])$, and the second term is the angular separation ($\theta_{ij}[n]$) of the two
3166 positions. The angular separation enters into the beamforming phase due to an effect
3167 caused by the circular cyclotron orbit of the electron that produces radiation whose
3168 phase is linearly dependent on the relative azimuthal position of the antenna [78, 79].
3169 The time-dependence of the beamforming phases corrects for the effects of the ∇B -drift,
3170 which cause the guiding centers of electrons to orbit the center of the magnetic trap. The
3171 correction adds a linear time-dependence to the azimuthal beamforming position,

$$\theta_i[n] = \omega_{\nabla B} t[n] + \theta_{i,0}, \quad (4.42)$$

3172 where $\omega_{\nabla B}$ is the azimuthal grad-B drift frequency, $t[n]$ is the time vector and, $\theta_{i,0}$
3173 is the starting azimuthal position, which allows the beamforming phases to track the
3174 XY-position of the guiding center. Predicting accurate values of $\omega_{\nabla B}$ for a specific trap
3175 and set of kinematic parameters can be done with simulations, which are performed

3176 using the Locust software package [61] developed by Project 8.

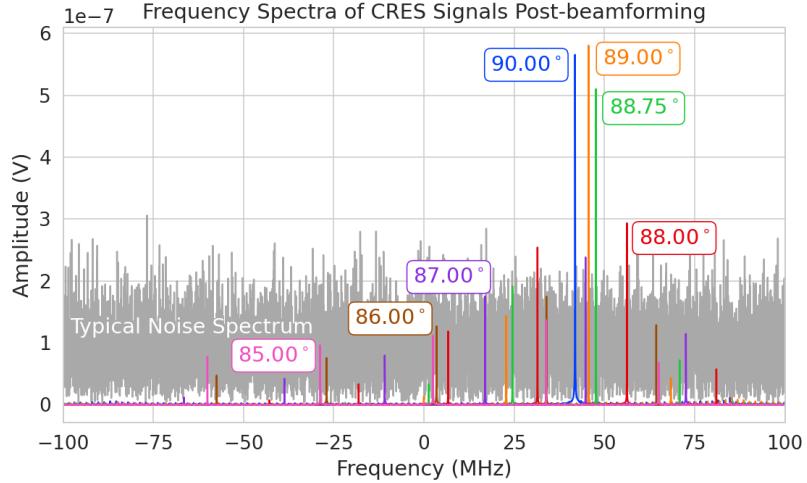


Figure 4.30. Frequency spectra of simulated CRES events in the FSCD magnetic trap after beamforming. The signal of a 90° electron consists of a single frequency component that is clearly detectable using a power threshold on the frequency spectrum. This power threshold remains effective for signals with relatively large pitch angles such as 89.0° and 88.75° , which are composed of a main carrier and a few small sidebands. Signals with smaller pitch angles, below about 88.5° , are dominated by sidebands such that no single frequency component can be reliably distinguished from the noise with a power threshold.

3177 After digital beamforming, a short-time Fourier transform (STFT) is applied to the
3178 summed time-series to obtain the signal frequency spectrum (see Figure 4.30). From the
3179 detection perspective, the frequency representation of the CRES data is advantageous
3180 compared to the time domain, due to the sparseness of CRES signals in the frequency
3181 domain. The frequency spectra of CRES signals are well-approximated by a frequency and
3182 amplitude modulated sinusoidal whose carrier frequency increases as a linear chirp [60].
3183 The modulation is caused by the axial oscillation of the electron in the magnetic trap,
3184 and the linear chirp is caused by the energy loss due to cyclotron radiation, which results
3185 in a relatively slow increase in the frequency components of the CRES signal over time.
3186 A typical CRES signal increases in frequency by approximately 15 kHz during the
3187 standard Fourier analysis window of 40.96 μ sec, which is smaller than the frequency
3188 bin width for a 200 MHz sample rate. Therefore, when considering a single frequency
3189 spectrum it is justifiable to neglect the effects of the linear frequency chirp.

3190 The majority of the CRES signal power for electrons in the FSCD trap is contained in
3191 a single frequency component when the electron has a pitch angle $\gtrsim 88.5^\circ$. The remain-
3192 ing signal power is distributed between a small number of sidebands with amplitudes

proportional to the electron's axial modulation (see Figure 4.30). Signal detection for these pitch angles is straightforward using a simple power threshold on the STFT, since the amplitude of the main signal peak is well above the thermal noise spectrum. However, as the pitch angle of the electron is decreased below 88.5° , the maximum amplitude of the frequency spectrum becomes comparable to typical noise fluctuations. At this point, the power threshold trigger is no longer able to distinguish between signal and noise leading to a reduction in detection efficiency, which is directly linked to the neutrino mass sensitivity of the FSCD. Because the distribution of electron pitch angles is effectively uniform, utilizing a signal detection algorithm that can improve efficiency for pitch angles less than 88.5° will lead to improvements in the neutrino mass sensitivity of the FSCD.

4.5 Signal Detection Algorithms

Modeling detection performance requires one to pose the signal detection problem in a consistent manner. The approach studied here uses the frequency spectra obtained from a STFT applied to the beamformed time-series from the FSCD to perform a binary hypothesis test. Mathematically, this is expressed as,

$$\mathcal{H}_0 : y[n] = \nu[n] \quad (4.43)$$

$$\mathcal{H}_1 : y[n] = x[n] + \nu[n]. \quad (4.44)$$

Under hypothesis \mathcal{H}_0 the vector representing the frequency spectrum ($y[n]$) is composed of complex white Gaussian noise (cWGN, $\nu[n]$) with total variance τ , and under hypothesis \mathcal{H}_1 the frequency spectrum is composed of a CRES signal ($x[n]$) with added cWGN. The dominant noise source for the FSCD is expected to be thermal Nyquist-Johnson noise, which is well approximated by a cWGN distribution. The hypothesis test is performed by calculating the ratio between the log-likelihood probability distributions for the classifier under \mathcal{H}_1 and \mathcal{H}_0 , which is the standard Neyman-Pearson approach to hypothesis testing [70]. The output of the log-likelihood ratio test is called the test statistic, which is used to assign the data as belonging to the noise or signal classes using a decision threshold.

In practice, the decision threshold is selected by finding the value of the test statistic that guarantees a tolerable rate of false positives. Given this false positive rate (FPR), one attempts to find a classifier that maximizes the true positive rate (TPR), which is the probability of correctly identifying if a piece of data contains signal or noise. Because

3222 FSCD signal classifiers will be used to evaluate the spectra of $O(10^2)$ beamforming
 3223 positions every 40.96 μ sec, there is a requirement that the signal classifiers with FPR
 3224 significantly smaller than 1% to reduce the burden placed on later stages of the CRES
 3225 reconstruction chain.

3226 4.5.0.1 Power Threshold

3227 The power threshold detection algorithm uses the maximum amplitude of the frequency
 3228 spectrum as the detection test statistic. Consider the \mathcal{H}_0 hypothesis where the signal is
 3229 pure cWGN. The performance of the power threshold can be modeled by first analyzing
 3230 a single bin in the frequency spectrum. The probability that the amplitude of a single
 3231 frequency bin falls below the decision threshold is given by the Rayleigh cumulative
 3232 distribution function (CDF),

$$\text{Ray}(x; \tau) = 1 - \exp(-|x|^2/\tau), \quad (4.45)$$

3233 where the complex value of the frequency bin is x , and τ is the cWGN variance. Because
 3234 the noise samples are independent and identically distributed (IID), the probability that
 3235 all bins in the frequency spectrum fall below the threshold is the joint CDF formed by
 3236 the product of each individual frequency bin CDF,

$$F_0(x; \tau, N_{\text{bin}}) = \text{Ray}(x; \tau)^{N_{\text{bin}}}. \quad (4.46)$$

3237 Finally, the PDF for the power threshold classifier can be obtained by differentiating
 3238 Equation 4.46.

3239 The noise variance of a beamformed frequency spectrum can be obtained directly
 3240 from the estimated noise power in a single antenna channel. The Nyquist-Johnson noise
 3241 power is given by $k_B T \Delta f$, where k_B is Boltzmann's constant, T is the system noise
 3242 temperature, and Δf is the sample rate. The beamformed noise variance is increased
 3243 by a factor of N_{ch} , where N_{ch} is the number of antennas, caused by the summation of
 3244 incoherent noise samples, however, the noise variance per frequency bin is decreased by a
 3245 factor equal to the number of samples in the STFT (N_{FFT}). The final expression for the
 3246 noise variance of the beamformed frequency spectrum is given by

$$\tau = k_B T \Delta f N_{\text{ch}} R / N_{\text{FFT}}, \quad (4.47)$$

3247 where the system impedance (R) has been used to convert from power to voltage-squared.

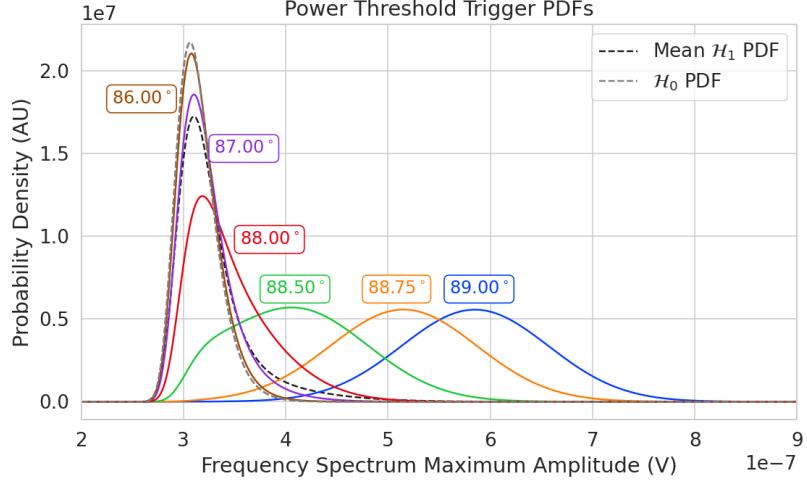


Figure 4.31. PDFs of the power threshold test statistic for CRES signals with various pitch angles as well as the PDF for the noise-only signal case. The average PDF computed for pitch angles ranging from 85.5 to 88.5° is also shown. As the pitch angle is decreased the signal PDF converges towards the noise PDF which indicates that the power threshold trigger is unable to distinguish between small pitch angle signals and noise.

3248 The probability distribution for the power threshold classifier under \mathcal{H}_1 is calculated
 3249 in a similar way, but the frequency bins that contain signal must be treated separately.
 3250 The probability that the amplitude of a frequency bin containing both signal and noise
 3251 bin falls below the decision threshold is described by a Rician CDF,

$$\text{Rice}(x; \tau, \alpha) = 1 - \int_x^\infty d|\tilde{x}| \frac{2|\tilde{x}|}{\tau} \exp\left(-\frac{|\tilde{x}|^2 + |\alpha|^2}{\tau}\right) \mathcal{I}_0\left(\frac{2|\tilde{x}||\alpha|}{\tau}\right), \quad (4.48)$$

3252 where the parameter $|\alpha|$ defines the noise-free amplitude of the signal. The CDF that
 3253 describes the probability that the entire spectrum falls below the decision threshold is
 3254 the product of both signal and noise CDFs,

$$F_1(x; \tau, \alpha, N_{\text{bin}}, N_s) = \text{Ray}(x; \tau)^{N_{\text{bin}} - N_s} \prod_{k=0}^{N_s} \text{Rice}(x; \tau, \alpha_k). \quad (4.49)$$

3255 The first half of Equation 4.49 is the contribution from the bins in the frequency spectrum
 3256 that contain only noise, and the second half is the product of the Rician CDFs for the
 3257 frequency bins that contain signal peaks with a noise-free amplitude of $|\alpha_k|$. Figure 4.31
 3258 shows plots of example PDFs under \mathcal{H}_1 and \mathcal{H}_0 .

3259 **4.5.0.2 Matched Filtering**

3260 The shape of a CRES signal in-between random scattering events with the background
 3261 gas is completely determined by the initial conditions of the electron, which implies that
 3262 it is possible to apply matched filtering as a signal detection algorithm. A matched filter
 3263 uses the shape of the known signal, which is called a template, to filter the incoming
 3264 data by computing the convolution between the signal and the data [70]. The matched
 3265 filter is the optimal detector, which means it achieves the maximum TPR for a particular
 3266 FPR, under the assumption that the signal is perfectly known and the noise is Gaussian
 3267 distributed. Since CRES signals have an unknown shape but are deterministic, the
 3268 matched filter can be applied by using simulations to generate a large number of signal
 3269 templates, called a "template bank", which spans the parameter space of possible signals.
 3270 Then at detection time, the template bank is used to identify signals by performing the
 3271 matched filter convolution for each template in an exhaustive search.

3272 CRES signals are highly periodic in nature. In such cases, it is advantageous to utilize
 3273 the convolution theorem to replace the matched filter convolution with an inner product
 3274 in the frequency-domain. Using the convolution theorem, the matched filter test statistic
 3275 is given by

$$\mathcal{T} = \max_h \left| \sum_{n=0}^{N_{\text{bin}}} h^\dagger[n] y[n] \right|, \quad (4.50)$$

3276 where $h^\dagger[n]$ is the complex conjugate of the signal template.

3277 The approach to deriving PDFs that describe the matched filter template bank will
 3278 be to first derive PDFs for \mathcal{H}_0 and \mathcal{H}_1 in the case of a single template and use these
 3279 solutions to create PDFs that describe the multi-template case. In the case when the
 3280 template bank consists of only a single template it is possible to derive an exact analytical
 3281 form for the PDF. Consider the \mathcal{H}_1 case, where the equation describing the matched
 3282 filter test statistic, also known as the matched filter score, becomes

$$\mathcal{T} = \left| \sum_{n=0}^{N_{\text{bin}}} h^\dagger[n] y[n] \right|. \quad (4.51)$$

3283 Each noisy frequency bin is a sum of signal and cWGN, which means $y[n]$ is also a
 3284 Gaussian distributed variable. Therefore, the value of the inner product between the
 3285 template and the data is also a complex Gaussian variable; and, since the matched filter
 3286 score is the magnitude of this inner product, it must follow a Rician distribution.

3287 The distribution that describes the matched filter score under \mathcal{H}_1 can be derived

3288 starting with the matched filter template equation. The matched filter template \mathbf{h} is a
 3289 simulated signal (\mathbf{x}_h) with a normalization factor

$$\mathbf{h} = \frac{\mathbf{x}_h}{\sqrt{\tau|\mathbf{x}_h|^2}}, \quad (4.52)$$

3290 where τ is the noise variance. Inserting this into Equation 4.50 and expressing the data
 3291 as a sum between a signal and a WGN vector yields,

$$\mathcal{T} = \frac{1}{\sqrt{\tau|\mathbf{x}_h|^2}} \left| \sum_{n=1}^{N_{\text{bin}}} x_h^\dagger[n]x[n] + \sum_{n=1}^{N_{\text{bin}}} x_h^\dagger[n]\nu[n] \right|. \quad (4.53)$$

3292 The first term is a scalar product between the signal and template vectors and the
 3293 second term is a complex Gaussian distributed variable with variance one. For the
 3294 purposes of identifying the statistical distribution, it is useful to rewrite the summation
 3295 describing an inner product

$$\sum_{n=1}^{N_{\text{bin}}} x_h^\dagger[n]x[n] = \mathbf{x}_h \cdot \mathbf{x} = |\mathbf{x}_h \cdot \mathbf{x}|e^{i\vartheta} \leq |\mathbf{x}_h||\mathbf{x}|e^{i\vartheta}, \quad (4.54)$$

3296 the last step utilizes the Cauchy-Schawrz inequality, where equality is guaranteed when
 3297 $\mathbf{x} = \mathbf{x}_h$. Instead of the inequality it is useful to define a quantity called "match" such that

$$|\mathbf{x}_h \cdot \mathbf{x}|e^{i\vartheta} = |\mathbf{x}_h||\mathbf{x}|\Gamma e^{i\vartheta}, \quad (4.55)$$

3298 where the match factor $\Gamma \in [0, 1]$. The match factor quantifies how well the template
 3299 matches the signal.

3300 The fact that the second term is a random complex Gaussian variable with unity
 3301 variance can be seen by noting that each of the noise samples are drawn from the complex
 3302 Gaussian distribution, $\mathcal{N}(0, \tau)$. Therefore,

$$\frac{x_h^\dagger[n]}{\sqrt{\tau|\mathbf{x}_h|^2}}\nu[n] \sim \mathcal{N}\left(0, \frac{x_h^\dagger[n]x_h[n]}{|\mathbf{x}_h|^2}\right), \quad (4.56)$$

$$n = \sum_{n=1}^{N_{\text{bin}}} \frac{x_h[n]}{\sqrt{\tau|\mathbf{x}_h|^2}}\nu[n] \sim \mathcal{N}\left(0, \frac{\sum_{n=1}^{N_{\text{bin}}} x_h^\dagger[n]x_h[n]}{|\mathbf{x}_h|^2}\right) = \mathcal{N}(0, 1). \quad (4.57)$$

3303 Equation 4.53 can now be simplified

$$\mathcal{T} = \left| |\mathbf{h}| |\mathbf{x}| \Gamma e^{i\vartheta} + n \right|, \quad (4.58)$$

3304 where Equation 4.52 has been used to redefine the inner product term. The quantity
3305 $|\mathbf{h}| |\mathbf{x}| \Gamma$ is a real number, which is the matched filter score that one would expect if the
3306 data contained no noise. The final simplification is to define $\mathcal{T}_{\text{ideal}} = |\mathbf{h}| |\mathbf{x}| \Gamma$, from which
3307 one obtains

$$\mathcal{T} = |\mathcal{T}_{\text{ideal}} e^{i\vartheta} + n|. \quad (4.59)$$

3308 From Equation 4.59 on can see that \mathcal{T} is simply the magnitude of a complex number
3309 with added cWGN of variance 1, which follows the Rician distribution, therefore the
3310 distribution that describes the matched filter score for a single template under \mathcal{H}_1 is

$$P_1(x; \mathcal{T}_{\text{ideal}}) = 2x \exp(- (x^2 + \mathcal{T}_{\text{ideal}}^2)) I_0(2x\mathcal{T}_{\text{ideal}}). \quad (4.60)$$

3311 The shape of the matched filter score distribution is controlled by the parameter $\mathcal{T}_{\text{ideal}}$,
3312 which is effectively the value of the matched filter score if the data contained no noise.
3313 Without noise, the data vector reduces to the signal, \mathbf{x} , in which case Equation 4.51
3314 becomes the magnitude of an inner product between two vectors. The magnitude of an
3315 inner product can be expressed in terms of the magnitudes of the vectors and a constant
3316 that describes the degree of orthogonality between them. Applying this to Equation 4.51,
3317 one obtains

$$\mathcal{T}_{\text{ideal}} = |\mathbf{h}^\dagger \cdot \mathbf{x}| = |\mathbf{h}| |\mathbf{x}| \Gamma \quad (4.61)$$

3318 where Γ describes the orthogonality between \mathbf{h} and \mathbf{x} . Γ effectively quantifies how well
3319 the template matches the unknown signal in the data.

3320 The matched filter score PDF under \mathcal{H}_0 is readily obtained from Equation 4.60 by
3321 setting the value of $\mathcal{T}_{\text{ideal}}$ to zero, since the data contains no signal in the noise case.
3322 Doing this, one obtains a Rayleigh distribution,

$$P_0(x) = 2x \exp(-x^2). \quad (4.62)$$

3323 Equations 4.60 and 4.62 describe the behavior of the matched filter test statistic
3324 under \mathcal{H}_0 and \mathcal{H}_1 for a single template. However, defining a PDF that describes the
3325 matched filter test statistic in the case of multiple templates is in general a mathematically
3326 intractable problem, since there is no guarantee of orthogonality between matched filter

3327 templates. This leads to correlations between the matched filter scores of different
 3328 templates, because only one sample of noise is used to compute the matched filter scores
 3329 of the template bank. In order to proceed, it is assumed that the matched filter scores for
 3330 all templates are IID variables, which allows one to ignore correlations between templates.
 3331 The overall effect of this will be an underestimate of the performance of the matched
 3332 filter by over-estimating the required number of templates and, therefore, the magnitude
 3333 of the statistical trials penalty.

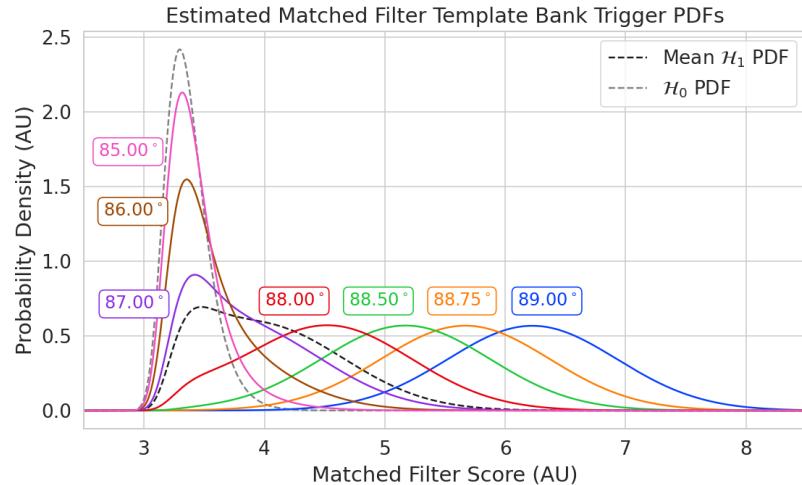


Figure 4.32. Plots of PDFs that describe the matched filter template bank test statistic for CRES signals with various pitch angles, as well as the estimated PDF for the noise only case. 10^5 matched filter templates are used and perfect match between signal and template i.e. $\Gamma_{\text{best}} = 1$ is assumed. The mean PDF includes signals ranging from $85.5 - 88.5^\circ$ in pitch angle. There is a larger distinction between the signal PDFs at small pitch angles compared to the power threshold, which indicates a higher detection efficiency for these signals.

3334 The probability that the matched filter score falls below the decision threshold under
 3335 \mathcal{H}_0 is again given by the CDF. Because of the assumption that matched filter scores from
 3336 different templates are independent, the probability that the matched filter score for all
 3337 templates falls below the threshold value is simply the joint CDF, which is

$$F_0(x) = \left(1 - e^{-x^2}\right)^{N_t}, \quad (4.63)$$

3338 where x is the matched filter score threshold and N_t is the number of templates. One
 3339 should expect that the distribution describing the maximum score of the matched filter
 3340 template bank depends on N_t , because with more templates there is a greater chance of
 3341 a random match between the template and data.

3342 The CDF that describes \mathcal{H}_1 is derived by starting with the CDF of the best matching
 3343 template, $F_{\text{best}}(x; \mathcal{T}_{\text{best}})$. Because of the orthogonality assumption, the matched filter
 3344 scores for all other templates are negligible ($\mathcal{T}_{\text{ideal}} \approx 0$). The joint CDF that describes
 3345 the total template bank is obtained by combining the distributions for all templates used
 3346 during detection. Therefore, the estimated CDF under \mathcal{H}_1 is

$$F_1(x; \mathcal{T}_{\text{best}}) = F_{\text{best}}(x; \mathcal{T}_{\text{best}}) \left(1 - e^{-x^2}\right)^{N_t}. \quad (4.64)$$

3347 Figure 4.32 shows plots of the matched filter template bank PDFs under \mathcal{H}_0 and \mathcal{H}_1 .

3348 4.5.0.3 Machine Learning

3349 The focus in this paper is on the potential of Convolutional Neural Networks (CNN)
 3350 as a machine learning based signal classifier at the trigger level. CNNs are constructed
 3351 using a series of convolutional layers, each composed of a set of filters that are convolved
 3352 with the input data. The individual convolutional filters can be viewed heuristically
 3353 as matched filter templates [?] that are learned from a set of simulated data rather
 3354 than being directly generated. This opens the possibility of finding a more efficient
 3355 representation of the matched filter templates during the training process that can
 3356 potentially reduce computational cost at inference time while retaining good classification
 3357 performance.

3358 The machine learning approach is distinct from the power threshold and matched
 3359 filtering in that there is no attempt to manually engineer a test statistic that can be
 3360 computed from the input data. Instead, a test statistic is calculated by constructing a
 3361 differentiable function that maps the complex frequency series to a binary classification
 3362 as signal or noise. The differentiable function is trained using supervised learning to
 3363 correctly perform this mapping. The test statistic for the machine learning classifier is
 3364 expressed mathematically as

$$\mathcal{T} = G(\mathbf{y}; \boldsymbol{\Omega}) \quad (4.65)$$

3365 where \mathbf{y} is the noisy data vector and $G(\mathbf{y}; \boldsymbol{\Omega})$ is the machine learning model parameterized
 3366 by the weights $\boldsymbol{\Omega}$.

3367 The CNN architecture used for this work is summarized by Table 4.1. No strategic
 3368 hyper-parameter optimization approach was implemented beyond the manual testing
 3369 of different CNN architecture variations, so this particular model is best viewed as a
 3370 proof-of-concept rather than a rigorously optimized design. Numerous model variations
 3371 were tested, some with significantly more layers and convolutions filters per layer, as

Table 4.1. A summary of the CNN model layers and parameters. The output of each 1D-Convolution and Fully Connected layer is passed through a LeakyReLU activation function and re-normalized using batch normalization before being passed to the next layer in the model. The output of the final Fully Connected layer in the model is left without activation so that the model outputs can be directly passed to the Binary Cross-entropy loss function used during training. The first layer in the network has two input channels for the real and imaginary components of the spectrum.

Layer	Type	Input Channels	Output Channels	Parameters
1	1D-Convolution	2	15	($N_{\text{kernel}} = 4$, $N_{\text{stride}} = 1$)
2	Maximum Pooling	15	15	($N_{\text{kernel}} = 4$, $N_{\text{stride}} = 4$)
3	1D-Convolution	15	20	($N_{\text{kernel}} = 4$, $N_{\text{stride}} = 1$)
4	Maximum Pooling	20	20	($N_{\text{kernel}} = 4$, $N_{\text{stride}} = 4$)
5	1D-Convolution	20	25	($N_{\text{kernel}} = 4$, $N_{\text{stride}} = 1$)
6	Maximum Pooling	25	25	($N_{\text{kernel}} = 4$, $N_{\text{stride}} = 4$)
7	Fully Connected	3200	512	NA
8	Fully Connected	512	64	NA
9	Fully Connected	64	2	NA

3372 well as others that were even smaller than the architecture in Table 4.1. Ultimately, the
 3373 model architecture choice was driven by the motivation to find the minimal model whose
 3374 classification performance was still comparable to the larger CNN’s tested, because of
 3375 the importance of minimizing computational cost in real-time applications. It is possible
 3376 that more sophisticated machine learning models could improve upon the classification
 3377 results achieved here, but this investigation is left for future work.

3378 4.5.1 Methods

3379 4.5.1.1 Data Generation

3380 Simulated CRES signals were generated using the Locust simulations package [61, 78].
 3381 Locust uses the separately developed Kassiopeia package [59] to calculate the magnetic
 3382 fields produced by a user defined set of current carrying coils along with any specified
 3383 background magnetic fields, resulting in a magnetic trap. Next, Kassiopeia calculates the
 3384 trajectory of an electron in this magnetic field starting from a set of user specified initial
 3385 conditions. The Locust software then uses the electron trajectories from Kassiopeia
 3386 to calculate the resulting electromagnetic fields using the Liénard-Wiechert equations,
 3387 and determines the voltages generated in the antenna array with the antenna transfer
 3388 function. Locust then simulates the down-conversion, filtering, and digitization steps
 3389 resulting in the simulated CRES signals for an electron.

3390 The shape of the received CRES signal is determined by the initial kinematic param-
3391 eters, including the starting position of the electron, the starting kinetic energy of the
3392 electron, and the pitch angle. The studies performed here are constrained to a single
3393 initial electron position located at $(x, y, z) = (5, 0, 0)$ mm. Two datasets are generated
3394 using this starting position by varying the initial kinetic energy and pitch angle. The
3395 first dataset consists of a two-dimensional square grid spanning an energy range from
3396 18575-18580 eV with a spacing of 0.1 eV, and pitch angles from 85.5-88.5° with a spacing
3397 of 0.001°, resulting in 153051 signals with a unique energy-pitch angle combination. This
3398 dataset is intended to represent a matched filter template bank. The upper range of pitch
3399 angles is limited because of the greater relative detection efficiency of the matched filter
3400 and neural network classifiers in this pitch angle range. The second dataset was generated
3401 by randomly sampling uniform probability distributions covering the same parameter
3402 space to produce approximately 50000 signals randomly parameterized in energy and
3403 pitch angle. This dataset provides the training and test data for the machine learning
3404 approach, and acts as a representative sample of signals to evaluate the performance of
3405 the matched filter template bank.

3406 Each signal was simulated for a duration of 40.96 μ s or 8192 samples starting at
3407 time $t = 0$ s for all simulations. This duration represents a single frequency spectrum
3408 generated by the STFT. The FSCD antenna array has sixty channels, and the output of
3409 the Locust simulations are a matrix of array snapshots with a size given by the number
3410 of channels times the event length ($N_{\text{ch}} \times N_{\text{sample}}$). The raw data from Locust is first
3411 summed using digital beamforming and converted to frequency spectra using a Fourier
3412 transform. The beamforming procedure uses the exact position and ∇B -drift correction
3413 to simplify the comparison between trigger algorithms. Many beamforming positions
3414 would be used in practice and potentially several estimates of a typical $\omega_{\nabla B}$ depending
3415 on the variation of the ∇B -drift frequency with pitch angle.

3416 **4.5.1.2 Template Number and Match Estimation**

3417 The estimated PDF for the matched filter template bank on the number of templates and
3418 the mean match (Γ_{best}). A given signal with random parameters will have a template in
3419 the filter bank that gives the highest matched filter score, therefore, the mean match
3420 ratio is obtained by averaging over the best matching templates for a representative
3421 population of test signals. Γ_{best} is a figure of merit that characterizes the performance of
3422 a template bank at signal detection. One expects that with more templates the value
3423 of Γ_{best} will increase, however, there is a point of diminishing returns at which more

3424 templates will not significantly increase match, but will still increase the likelihood of
3425 false positives. Therefore, it is desirable to use the minimum number of templates that provide an acceptable mean value of Γ_{best} .

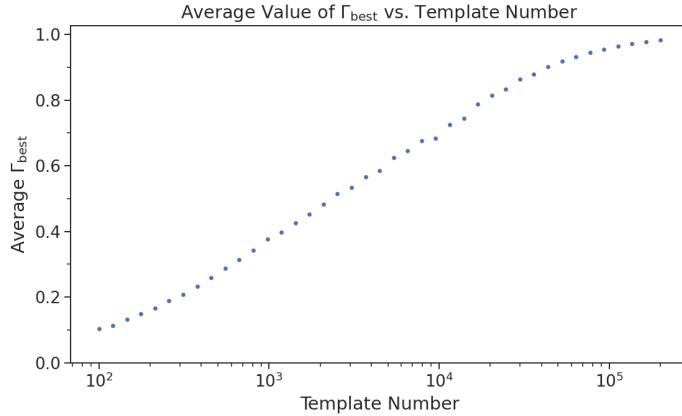


Figure 4.33. The mean match of the matched filter template bank to a test set of randomly parameterized signals as a function of the number or density of templates. The parameter space includes pitch angles from $85.5 - 88.5^\circ$ and energies from 18575 – 18580 eV.

3426
3427 To quantify the relationship between match and template number, the mean match
3428 of the random dataset to a selection of templates from the regularly spaced dataset was
3429 calculated. One sees that the average value of Γ_{best} is an exponential function of the
3430 number of templates (see Figure 4.33). Using this plot one can infer the required number
3431 of templates for the desired value of mean match.

3432 4.5.1.3 CNN Training and Data Augmentation

3433 The random dataset is split in half to create distinct training and test datasets for
3434 training the model. A randomly selected 20% of the training data is isolated for use as
3435 a validation set during the training loop. The size of the training, validation, and test
3436 datasets are tripled by appending two additional copies of the data to increase the sample
3437 size of the dataset after data augmentation. A different sample of noise is added to the
3438 simulation data during the training loop, which prevents the model from overtraining on
3439 noise features. The training and test datasets contain an equal split between signal and
3440 noise data, which are randomly shuffled after each training epoch.

3441 The Locust simulation data was augmented to make the datasets more representative
3442 of actual experiment data. As the signals are loaded for training a unique random phase
3443 shift is applied. Since the simulations are generated using the same initial axial position

3444 and cyclotron orbit phase, the randomization is an attempt to prevent overtraining on
 3445 these features. During each training epoch the data is randomly shuffled and split into
 3446 batches of 2500 signals. Each batch of signals is then circularly shifted by a random
 3447 number of frequency bins to simulate a kinetic energy shift from -75 to 20 eV, which
 3448 imitates a dataset with a larger energy range. Next, a sample of cWGN, consistent
 3449 with 10 K Nyquist-Johnson noise, is generated and added to the signal, which prevents
 3450 overtraining on noise features. As a final step, the data is renormalized by the standard
 3451 deviation of the noise so that the range of values in the data is close to $[-1, 1]$, which
 3452 ensures well-behaved back-propagation.

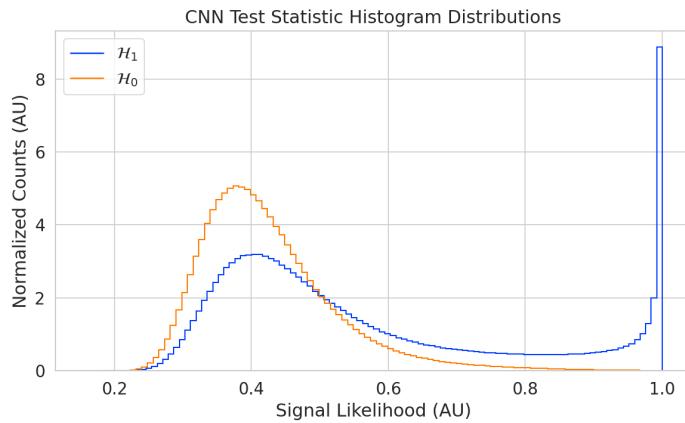


Figure 4.34. Histograms of the trained CNN model output from the test dataset. The blue histogram shows the model outputs for signal data. The oddly shaped peak near the end is the result of the softmax function mapping the long tail of the raw output distribution to the range $[0, 1]$.

3453 The Binary Cross-entropy loss function is used to compute the loss for each batch of
 3454 data, and the model weights are updated using the ADAM optimizer with a learning
 3455 rate of 5×10^{-3} . After each training epoch, the loss and classification accuracy of the
 3456 validation dataset are computed to monitor for overtraining. It was noticed that because
 3457 of the relatively high noise power and the fact that a new sample of noise was used for
 3458 each batch, it was nearly impossible to over-train the model. Typically, the loss and
 3459 classification accuracy of the model converged after a few hundred training epochs, but
 3460 the training loop was extended to 3000 epochs to attempt to achieve the best possible
 3461 performance. The training procedure generally took about 24 hrs using a single NVIDIA
 3462 V100 GPU [80].

3463 After training the model, it was used to classify the test dataset and generate histograms
 3464 of the model outputs for both classes of data. The data augmentation procedure for the

3465 evaluation of the test data mirrors the training procedure without the validation split.
 3466 Since a random circular shift and a new sample of WGN is added to each batch, the
 3467 testing evaluation loop is run for 100 epochs to get a representative sample of noise and
 3468 circular shifts. The model outputs are passed through a softmax activation and then
 3469 combined into histograms (see Figure 4.34).

3470 4.5.2 Results and Discussion

3471 4.5.2.1 Trigger Classification Performance

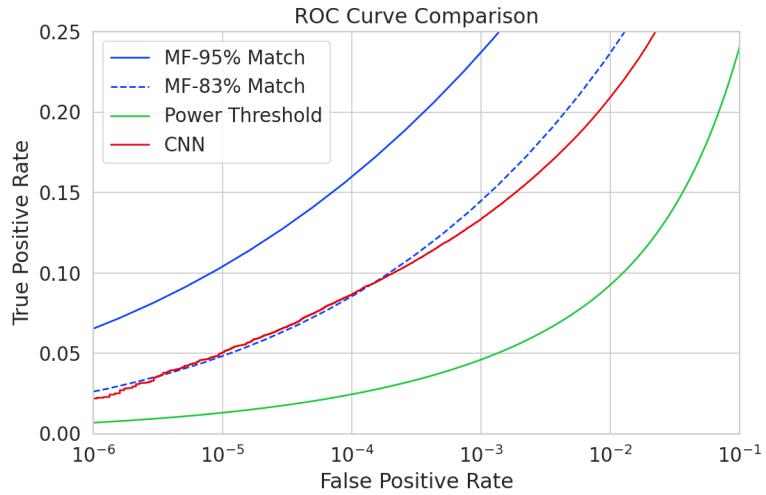


Figure 4.35. ROC curves describing the detection efficiency or true positive rates for the three signal classification algorithms examined in this paper. The matched filter (MF) and Power Threshold curves are computed analytically using the distribution functions introduced in Section 4.5, and the CNN curve is computed numerically using the classification results on the test dataset. The percent match indicated in the legend refers to the mean match of the classifier.

3472 The detection performance of the signal classifiers can be compared by computing
 3473 the receiver operating characteristic (ROC) curves (see Figure 4.35). A single ROC
 3474 curve is obtained for the matched filter and power threshold classifiers by averaging over
 3475 analytical ROC curves obtained from the distributions in Section 4.5. Two ROC curves
 3476 are calculated for the matched filter with different numbers of templates and mean match.
 3477 The ROC curve describing the CNN is obtained numerically from the histograms of the
 3478 model outputs for each signal class.

3479 The TPR of a signal classifier is equivalent to its detection efficiency, and one sees
 3480 that for the population of signals with pitch angles $< 88.5^\circ$ the power threshold has

3481 a consistently lower detection efficiency than the CNN and the matched filter. This
3482 result might have been predicted from the visualization of signal spectra in Figure 4.30,
3483 where it can be seen that a noise peak and a signal peak cannot be distinguished with
3484 high-confidence at small pitch angles. The CNN offers a significant and consistent increase
3485 in detection efficiency over the power threshold approach, with the relative improvement
3486 in detection efficiency increasing as the false positive rate decreases.

3487 If one compares the CNN to the matched filter, it can be seen that the performance of
3488 the tested network is roughly equivalent to a matched filter detector with a mean match
3489 of about 83%, which uses approximately 2×10^4 matched filter templates. The overall
3490 best detection efficiency is achieved by the matched filter classifier if a large enough
3491 template bank is used. The plot displays the ROC curve for a matched filter template
3492 bank with 95% mean match, which is achieved with approximately 10^5 templates. Since
3493 the matched filter is known to be statistically optimal for detecting a known signal in
3494 WGN, it is unsurprising that this algorithm has the highest detection efficiency.

3495 An important difference between the matched filter and CNN algorithms is that the
3496 CNN relies upon convolutions as its fundamental calculation mechanism, whereas our
3497 implementation of a matched filter utilizes an inner product. Since convolution is a
3498 translation invariant operation, the detection performance of CNN can be extended to
3499 a wider range of CRES event kinetic energies with less cost than the matched filter, a
3500 feature that is exploited during the CNN training by including circular translations of
3501 the CRES frequency spectra in the training loop. Increasing the range of detectable
3502 kinetic energies with a matched filter requires a proportional increase in the number of
3503 templates, which directly translates into increased computational and hardware costs.
3504 From a practical perspective, the detection algorithm is always limited by the available
3505 computational hardware, so estimating the relative costs is a key factor in determining
3506 their feasibility. A more detailed analysis of the relative costs of each of the detection
3507 algorithms is performed below.

3508 4.5.2.2 Computational Cost and Hardware Requirements

3509 The trade-off between better detection efficiency and computational cost is common
3510 to many signal detection problems and the FSCD is no exception. Computational
3511 costs can be related to actual hardware costs by calculating the theoretical amount of
3512 computer hardware required to implement the signal classifiers for real-time detection.
3513 The approach taken here utilizes order of magnitude estimates of the theoretical peak
3514 performance values for currently available Graphics Processing Units (GPUs) as a metric.

3515 This approach underestimates the amount of required hardware, since it is unlikely that
3516 any CRES detection algorithm could reach the theoretical peak performance of the
3517 hardware.

3518 Since the signal detection algorithms are designed to work using beamformed frequency
3519 spectra, the computational cost of beamforming combined with a fast Fourier transform
3520 (FFT) is constant for all classifiers. The beamforming grid is assumed to contain N_{bf}
3521 beamforming positions, each of which will produce a frequency spectrum containing N_{bin}
3522 after the FFT.

3523 Considering the power threshold classifier, this results in $N_{\text{bin}}N_b$ frequency bins
3524 that must be checked every N_{bin}/f_s seconds. The 20 cm diameter FSCD array requires
3525 $N_{\text{bf}} \approx O(10^2)$ for sufficient coverage and has a sampling frequency $f_s = 200$ MHz with a
3526 Fourier analysis window of $N_{\text{bin}} = 8192$ samples. Therefore the power threshold requires
3527 approximately $O(10^{10})$ FLOPS to check in real-time with these parameters

3528 Current generations of GPUs have peak theoretical performances in the range of
3529 $O(10^{13}) - O(10^{14})$ FLOPS [81], dependent on the required floating-point precision of
3530 the computation. Therefore, the entire computational needs of a real-time triggering
3531 system using a power threshold classifier, including digital beamforming and generation
3532 of the STFT, could be met by a single high-end GPU or a small number of less powerful
3533 GPUs. Since triggering is only one step of the full real-time signal reconstruction
3534 approach, limiting the computational cost of this stage is ideal. However, the power
3535 threshold classifier does not provide sufficient detection efficiency across the entire
3536 range of possible signals, which is the primary motivation for exploring more complicated
3537 triggering solutions.

3538 As discussed, the computational cost of the matched filter approach requires counting
3539 the number of templates that must be checked for each frequency spectra produced by
3540 the STFT. Computing the matched filter scores requires $O(N_{\text{bf}}N_tN_{\text{bin}})$ operations, since
3541 for each of the beamforming positions one must multiply N_t templates with a data vector
3542 that has length N_{bin} . The computation must be performed in a time less-than or equal
3543 to N_{bin}/f_s to keep up with the data generation rate. A 5 eV range of kinetic energies
3544 required 10^4 to 10^5 templates in order for the matched filter to exceed the performance
3545 of the CNN. The number of templates is expected to scale linearly with the total kinetic
3546 energy range of interest, therefore, 10^5 to 10^6 matched filter templates would be expected
3547 for the nominal 50 eV analysis window of the FSCD. Considering this, the estimated
3548 computational cost of the matched filter is between $O(10^{15})$ to $O(10^{16})$ FLOPS, which is
3549 $O(10^2)$ to $O(10^3)$ high-end GPUs.

3550 The computational cost of the CNN can be estimated by simply summing the compu-
3551 tational costs of the convolutions and matrix multiplications specified by the network
3552 architecture shown in Table 4.1. Each convolutional layer consists of $N_{\text{in}}N_{\text{out}}N_{\text{kernel}}L_{\text{input}}$
3553 floating-point operations, where N_{in} is the number of input channels, N_{out} is the number
3554 of output channels, N_{kernel} is the size of the convolutional kernel, and L_{input} is the length
3555 of the input vector, and the fully connected layers each contribute $N_{\text{in}}N_{\text{out}}$ operations.
3556 Summing all the neural network layers it is estimated that the CNN requires $O(10^6)$
3557 floating point operations to evaluate each frequency spectra; therefore, the total com-
3558 putational cost of the CNN trigger is value multiplied by the number of beamforming
3559 positions per the data acquisition time, which is $O(10^{13})$ FLOPS or $O(10^0)$ GPUs.

3560 Compared with the matched filter approach the CNN requires $O(100)$ to $O(1000)$
3561 fewer GPUs to implement, dependent on the exact number of templates used in the
3562 template bank. The 50 eV kinetic energy range is motivated by the application of these
3563 detection algorithms to an FSCD-like neutrino mass measurement experiment. However,
3564 if a significantly larger range of kinetic energies is required, a CNN may be the preferred
3565 detection approach despite the lower mean detection efficiency due to computational cost
3566 considerations.

3567 Additional experiments with larger CNNs, generated by increasing the depth and
3568 width of the neural network, were performed. It was observed that these changes
3569 provided minimal ($\lesssim 1\%$) improvement in the classification accuracy of the model. A
3570 potential reason for this could be the sparse nature of the signals in the frequency
3571 domain and the low SNR, which makes for a challenging dataset to learn from. Future
3572 work might investigate modifications to the neural network architecture such as sparse
3573 convolutions, which may improve the classification accuracy of the model or further
3574 reduce the computational costs of this approach. Alternatively, more complicated CNN
3575 architectures such as a ResNet [82] or VGG model [83] may provide improved classification
3576 performance over a basic CNN. An additional promising area of investigation are recurrent
3577 neural networks, which may be able to exploit the time-ordered features of the STFT for
3578 more accurate signal detection if the electron signals last for multiple Fourier transform
3579 windows.

3580 The estimate of the computational costs of the matched filter is somewhat naive if one
3581 notices that the majority of the values that make up a CRES frequency spectrum are zero
3582 (see Figure 4.30). Therefore, the majority of operations in the matched filter inner product
3583 are unnecessary, and one could instead evaluate the matched filter inner product using
3584 only the $\lesssim 10$ frequency peaks that make up the CRES signal. This optimization reduces

3585 the number of operations required to check each template by a factor of $O(100)$ to $O(1000)$,
3586 which brings the estimated computational cost of the matched filter in line with the
3587 CNN. Although this level of sparsity results in a multiplication with very low arithmetic
3588 complexity, the resulting sparse matched filter algorithm is still likely to be constrained
3589 by memory access speed rather than compute speed. Ultimately, the comparison of
3590 the relative computational and hardware costs between the matched filter and CNN
3591 will depend on the efficiency of the software implementation and hardware support for
3592 neural network and sparse matrix calculations, which will need to be determined using
3593 real-world benchmarks.

3594 **4.5.3 Conclusion**

3595 Increasing the detection efficiency and overall event rate of the CRES technique represents
3596 a key developmental path towards new scientific results and broader applications of the
3597 CRES technique. It is what motivates both the antenna array detection approach and
3598 the development of real-time signal reconstruction algorithms. The work presented here
3599 demonstrates that significant gains in the detection efficiency of the CRES technique
3600 are achievable by utilizing triggering algorithms that account for the specific shape of
3601 CRES signals in the detector. These algorithms emphasize the need for accurate and fast
3602 methods for CRES simulation, since they directly contribute to the success of matched
3603 filter methods by providing a way to generate expected signal templates and also serve
3604 as a source of training data for machine learning approaches.

3605 The down-side of these more advanced approaches to signal detection is the increase
3606 in computational resources required to implement them. However, it was shown that a
3607 CNN of minimal size was able to significantly improve detection performance above the
3608 baseline power threshold trigger algorithm with a theoretical computational cost of only
3609 $O(1)$ high-end GPU. This algorithm improves on detection performance while requiring
3610 at least a factor $O(10^2)$ less in computer relative to a matched filter template bank,
3611 which would be the classical approach to signal detection in Gaussian noise. Future work
3612 obtaining real-life benchmarks of the CNN and matched filter algorithms are required to
3613 support these conclusions, but this study has indicated that a real-time signal detection
3614 algorithm for an antenna array CRES experiment is computationally feasible without
3615 extraordinary compute power.

3616 While this work has focused on the real-time detection of CRES signals from antenna
3617 arrays, these same signal classifiers could be used in CRES experiments utilizing different
3618 detector technologies, since the same principles of signal detection will apply. For example,

3619 previous CRES measurements by the Project 8 collaboration that utilized a waveguide
3620 gas cell, could have improved their detection efficiency by employing a matched filter
3621 or neural network classifier to identify trapped electrons with pitch angles that are too
3622 small to be detected by the power threshold approach. Furthermore, alternative CRES
3623 detector technologies such as resonant cavities [40] could also see similar improvements
3624 in detection efficiency, which is of crucial importance to future efforts by the Project 8
3625 collaboration to utilize CRES to measure the neutrino mass.

3626 **Chapter 5 |**

3627 **Antenna and Antenna Measurement Sys-**

3628 **tem Development for the Project 8 Ex-**

3629 **periment**

3630 **5.1 Introduction**

3631 The FSCD and antenna array CRES represent an innovative approach to beta-decay
3632 spectroscopy. While much can be learned from simulations about the systematics
3633 of CRES with antenna arrays, laboratory measurements and demonstrations provide
3634 critical inputs to sensitivity and simulation models, and provide a means for calibration
3635 and commissioning of the experiment. Therefore, a robust program of antenna and
3636 antenna measurement hardware development is key to the success of the FSCD and the
3637 development of antenna array CRES more broadly.

3638 In this chapter I summarize the development of an antenna measurement system at
3639 Penn State to implement and test the techniques of antenna array CRES on the bench-top.
3640 In Section 5.2 I provide an introduction to some fundamental parameters and concepts
3641 related to antenna measurements as well as an overview of the Penn State antenna
3642 measurement system hardware. In Section 5.3 I include the manuscript of a paper [79]
3643 which details the design and characterization of a specialized antenna developed to mimic
3644 the electric fields emitted by an electron in a CRES experiment. This antenna, called
3645 the Synthetic Cyclotron Antenna (SYNCA), is intended as a calibration tool for antenna
3646 arrays developed for CRES measurements. Lastly, in Section 5.5 I summarize a set of
3647 prototype FSCD antenna array measurements with the SYNCA [43], which I use to
3648 validate the simulated performance of the antenna array and estimate systematic errors
3649 associated with the antenna array.

3650 **5.2 Antenna Measurements for CRES experiments**

3651 **5.2.1 Antenna Parameters**

3652 Antenna characterization measurements are intended to validate simulations of the
3653 antenna array performance, which ultimately informs the neutrino mass sensitivity of
3654 the experiment. In this section, I shall summarize a few fundamental concepts relating
3655 to antennas and antenna measurement, before introducing how Project 8 uses antenna
3656 measurements for the development of antenna array CRES.

3657 **5.2.1.1 Radiation Patterns**

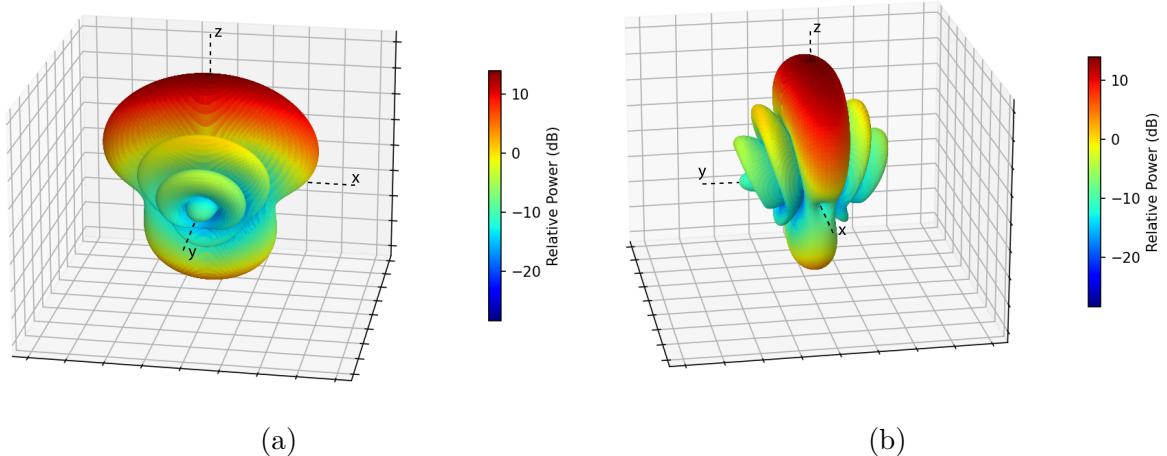


Figure 5.1. An example radiation pattern generated using HFSS simulations. The color and radial distance of the surface from the origin indicate the relative magnitude of radiation power emitted by the antenna in that direction. The primary goal of most antenna measurements is typically to measure the antenna pattern, which is used to derive many useful antenna performance parameters.

3658 Antennas are conductive structures designed to carry alternating electric currents
3659 to transmit energy in the form of EM waves [64]. Perhaps the most fundamental way
3660 to characterize an antenna, is to map out the radiated power density as a function of
3661 position, which is called the radiation pattern (see Figure 5.1). The radiation power
3662 density is obtained by calculating the time-averaged Poynting vector for all positions
3663 surrounding the antenna, which in equation form is

$$\mathbf{W}(x, y, z) = \langle \mathbf{E}(x, y, z, t) \times \mathbf{H}^*(x, y, z, t) \rangle_t, \quad (5.1)$$

3664 where $\mathbf{E}(x, y, z, t)$ and $\mathbf{H}(x, y, z, t)$ are the time-dependent electric and magnetic fields
 3665 produced by the antenna [48]. The radiation power density has units of W/m^2 and is
 3666 more typically called the energy flux density in physics applications, since it is a measure
 3667 of the amount of energy passing through a unit area over time.

3668 Because the radiation power density is a measure of power per unit area, its value
 3669 in a particular direction will depend on the distance from the antenna at which one is
 3670 measuring. This is undesirable for practical applications. A related quantity, which is
 3671 distance independent, is the energy flux per unit solid angle or radiation intensity, which
 3672 is computed directly from the radiation power density by multiplying by the squared
 3673 distance from the antenna. Specifically,

$$U = r^2 W(x, y, z), \quad (5.2)$$

3674 where r is the distance from the antenna to the field measurement point. The radiation
 3675 intensity is typically defined in regions where the Poynting vector consists only of a radial
 3676 component where it is safe to treat as a scalar quantity.

3677 5.2.1.2 Directivity and Gain

3678 Since the radiation intensity is a measure of average power per unit solid angle, it is
 3679 independent of distance and more useful as feature for antenna measurement. The
 3680 radiation intensity is directly related to antenna directivity and gain, which are common
 3681 antenna engineering figures-of-merit. Directivity is defined as the ratio between the
 3682 radiation intensity at particular point on the radiation pattern to the average radiation
 3683 intensity computed over all solid angles [64]. The equation that relates the radiation
 3684 intensity to directivity is

$$D = \frac{U}{U_0} = \frac{4\pi U}{P_{\text{rad}}}, \quad (5.3)$$

3685 where U_0 is the average radiation intensity over all solid angles, which simply the total
 3686 radiated power (P_{rad}) divided by 4π . Closely related to directivity is antenna gain, which
 3687 accounts for energy losses that occur inside then antenna when attempting to transmit
 3688 or receive a signal. The antenna gain is given by

$$G = \frac{4\pi U}{P_{\text{in}}}, \quad (5.4)$$

3689 where P_{in} is the total power delivered to the antenna. Gain can be thought of as the ratio
 3690 of the antenna's radiation intensity to that of a hypothetical isotropic, lossless radiator.

3691 The maximum values of gain and directivity exhibited by the main lobe of the antenna
 3692 pattern as well as the ratio between the gain of the main lobe and any side-lobes are
 3693 important figures-of-merit to evaluate antenna design performance.

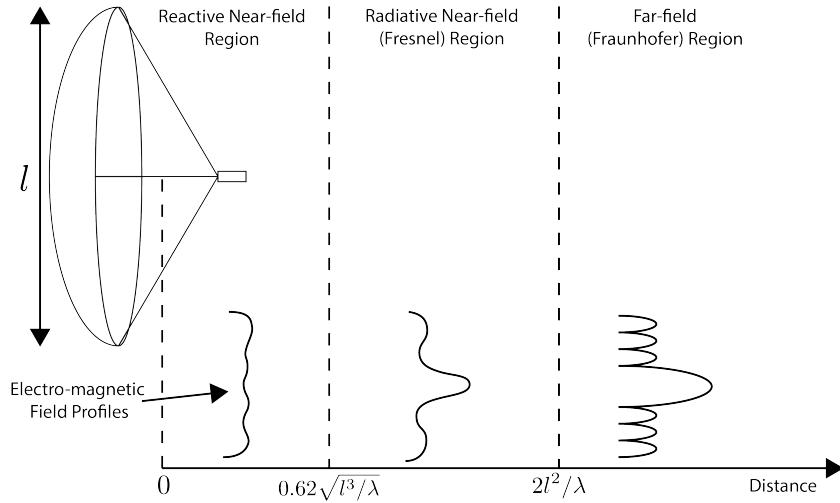


Figure 5.2. An illustration of the three field regions important for the analysis of an antenna system. Very close to the antenna the electric fields are primarily reactive so there is no radiation. If a receiving antenna were placed in this region most of the energy would be reflected back to the transmitter. Outside of the reactive near-field is the radiative near field. At these distances the antenna does radiate, but the radiation pattern is not well-defined since it changes based on the distance of the receiving antenna. It is only in the far-field region where the radiation pattern becomes constant as a function of distance, which is where the majority of antenna engineering is assumed to take place. The antenna arrays developed by Project 8 for CRES measurements operate in the radiative near-field due to the importance of limiting power loss from free-space propagation, which complicates the design of the antenna system.

3694 5.2.1.3 Far-field and Near-field

3695 Radiation patterns are well-defined only in regions where the shape of the radiation
 3696 pattern is independent of distance. The region where this approximation is valid is called
 3697 the "far-field", and in this region the EM fields from the antenna can be approximated as
 3698 spherical plane waves. A rule of thumb for antennas is that the far-field approximation
 3699 applies when the condition

$$R > \frac{2l^2}{\lambda} \quad (5.5)$$

3700 is true. In this expression, R is the distance from the antenna, l is the largest characteristic
 3701 dimension of the antenna, and λ is the wavelength of the radiation (see Figure 5.2).

3702 The region very close to the antenna is called the reactive near-field, because in this
 3703 region the reactive component of the EM field is dominant. Unlike radiative electric

3704 fields, the reactive electric and magnetic fields are out of phase from each other by 90° ,
 3705 since they are caused by electrostatic and magnetostatic effects from the self-capacitance
 3706 and self-inductance of the antenna. The reactive fields are unable to transfer energy a
 3707 significant distance from the antenna and are thus completely negligible for most antenna
 3708 applications. The limit of the reactive near-field for an electrically-large antenna is
 3709 typically taken to be

$$R < 0.62\sqrt{l^3/\lambda}. \quad (5.6)$$

3710 The unique application of antennas by Project 8 is limited by reactive near-field effects,
 3711 since it defines an absolute minimum distance for detectable electrons inside the uniform
 3712 cylindrical antenna array. If electrons are too close to the edge of the array than reactive
 3713 near-field effects leads to a large reduction in the received power and detection efficiency.
 3714 This leads to a significant volume inside the antenna array that is unsuitable for CRES
 3715 lowering the volumetric efficiency of the antenna array CRES technique.

3716 Between the reactive near-field and the far-field is the radiative near-field region. In
 3717 this region the fields are primarily radiative, however, it is too close to the antenna for
 3718 the spherical plane wave approximation to apply. Therefore, interference effects between
 3719 EM waves emitted from different points on the antenna occur causing the shape of the
 3720 radiation pattern to change as a function of distance from the antenna. Evaluating the
 3721 far-field distance limit for the FSCD antennas one finds an estimated far-field distance
 3722 of 43 cm, which is a factor of four larger than the radius of the antenna array designed
 3723 for the experiment. Consequently, it is expected that near-field effects will influence
 3724 the performance of the antenna array highlighting the importance of calibration and
 3725 characterization measurements to mitigate these effects.

3726 5.2.1.4 Polarization

3727 The polarization of an EM wave defines the spatial orientation of the electric field
 3728 oscillations. Conventionally, polarization vectors a defined in the plane perpendicular
 3729 to the direction of propagation for the EM wave. For radiation moving in the radial (\hat{r})
 3730 direction the electric field can be decomposed into the orthogonal basis

$$\mathbf{E}_{\text{tot}} = E_\theta \hat{\theta} + E_\phi \hat{\phi}, \quad (5.7)$$

3731 assuming a spherical coordinate system.

3732 In general, one defines partial radiation patterns, directivities, and gains so that the
 3733 performance of the antenna can be analyzed for the desired polarization. The radiation

³⁷³⁴ pattern defined in terms of partial patterns is

$$U_{\text{tot}} = U_\phi + U_\theta, \quad (5.8)$$

³⁷³⁵ where U_ϕ and U_θ are the radiation intensities in a particular direction for the respective
³⁷³⁶ polarization components. Similarly, a quantity such as gain can be written in terms of
³⁷³⁷ partial gains,

$$G_{\text{tot}} = G_\phi + G_\theta = \frac{2\pi U_\phi}{P_{\text{in}}} + \frac{2\pi U_\theta}{P_{\text{in}}}. \quad (5.9)$$

³⁷³⁸ An electron performing a circular orbit in the XY-plane from the side, viewed along
³⁷³⁹ the X or Y axes, would be seen as performing a linear oscillation perpendicular to the
³⁷⁴⁰ viewing axis. From this picture, one would predict that the primary polarization of
³⁷⁴¹ electric fields from CRES events is linearly polarization in the $\hat{\phi}$ direction in the XY-plane.

³⁷⁴² 5.2.1.5 Antenna Factor and Effective Aperture

³⁷⁴³ A useful way to characterize the performance of an antenna is to measure the electric
³⁷⁴⁴ field magnitude required to produce a signal with an amplitude of one volt in the antenna
³⁷⁴⁵ terminals. This ratio between the magnitude of the incoming electric field and the
³⁷⁴⁶ magnitude of the signal produced by the antenna is called the antenna factor, which is
³⁷⁴⁷ written as

$$A_F = \frac{|\mathbf{E}_{\text{in}}|}{V_{\text{ant}}}, \quad (5.10)$$

³⁷⁴⁸ where A_F is the antenna factor, E_{in} is the incoming electric field, and V_{ant} is the magnitude
³⁷⁴⁹ of the voltage produced by the antenna.

³⁷⁵⁰ The antenna factor can be expressed in terms of the antenna's gain through a related
³⁷⁵¹ quantity called the effective aperture. The effective aperture defines for a given incident
³⁷⁵² radiation power density (W/m^2) the power that is received by the antenna. Therefore,
³⁷⁵³ the effective aperture gives the equivalent area of the antenna,

$$A_{\text{eff}} = \frac{P_{\text{rec}}}{P_{\text{in}}} = \frac{\lambda^2}{4\pi} G, \quad (5.11)$$

³⁷⁵⁴ where the received power is P_r and the total incoming power is P_{in} .

³⁷⁵⁵ The magnitude of the Poynting vector can be written as

$$|\mathbf{S}_{\text{in}}| = |\mathbf{E}_{\text{in}}|^2 / \eta_0, \quad (5.12)$$

3756 where η_0 is the impedance of free-space, which relates the magnitudes of the electric and
 3757 magnetic fields in a vacuum, and is defined by

$$\eta_0 = \frac{|\mathbf{E}|}{|\mathbf{H}|} = \sqrt{\frac{\epsilon_0}{\mu_0}}. \quad (5.13)$$

3758 Therefore, the total received power by the antenna is

$$P_{\text{rec}} = |\mathbf{S}_{\text{in}}| A_{\text{eff}} = |\mathbf{S}_{\text{in}}| \frac{\lambda^2}{4\pi} G = \frac{|\mathbf{E}_{\text{in}}|^2 \lambda^2 G}{4\pi \eta_0}. \quad (5.14)$$

3759 To relate this to the antenna factor recall that the voltage produced by the antenna
 3760 is related to the received power by

$$P_{\text{rec}} = \frac{V_{\text{ant}}^2}{Z} = \frac{|\mathbf{E}_{\text{in}}|^2}{A_{\text{F}}^2 Z}, \quad (5.15)$$

3761 where Z is the system impedance. Setting Equations 5.14 and 5.15 equal to each other,
 3762 one obtains the following expression for antenna factor in terms of gain

$$A_{\text{F}} = \sqrt{\frac{4\pi\eta_0}{ZG\lambda^2}} = \frac{9.73}{\lambda\sqrt{G}}. \quad (5.16)$$

3763 The second expression in Equation 5.16 is obtained by evaluating the constant terms
 3764 assuming a system impedance of 50Ω .

3765 This exercise highlights that the majority of antenna parameters that one cares
 3766 to measure about an antenna can be obtained from the radiation or gain pattern of
 3767 the antenna. The antenna factor is a particularly important parameter for CRES
 3768 measurements due to its relevance to antenna array simulations with the Locust software
 3769 [61, 78].

3770 To compute the response of the antenna to the electric field, Locust relies upon
 3771 linear time-invariant system theory, which computes the response of the antenna (i.e. the
 3772 voltage time series generated by the antenna) using a convolution between the electric field
 3773 time-series and the antenna impulse response. This approach is necessary for correctly
 3774 modeling the antenna response to the electric field due to the broadband and non-
 3775 stationary nature of the electric fields from CRES events. Since antenna measurements
 3776 take place under steady-state conditions, parameters such as the radiation pattern, gain,
 3777 and antenna factor are defined in the frequency domain. However, by performing an
 3778 inverse Fourier transform on the antenna factor one obtains the antenna impulse response,

3779 which is used to calculate CRES signal voltages in Locust.

3780 5.2.2 Antenna Measurement Fundamentals

3781 5.2.2.1 Friis Transmission Equation

3782 The antenna factor or antenna transfer function is used to model how the antenna
3783 responds to electric fields emitted from a CRES event. Therefore, directly measuring the
3784 antenna transfer functions of the array is a key step in the commissioning and calibrating
3785 the FSCD experiment. A common approach to antenna characterization is to perform a
3786 two antenna transmit-receive measurement where an antenna with a known gain is used
3787 to characterize the unknown gain of the antenna under test (see Figure 5.3).

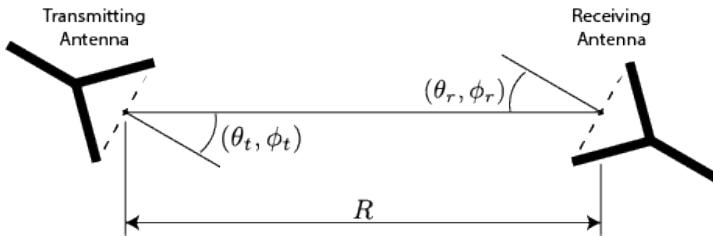


Figure 5.3. An illustration of the Friis measurement technique commonly used for antenna characterization measurements.

3787 Analyzing this two antenna setup involves calculating the power received from the
3788 transmitting antenna. The received power density is expressed as a function of the
3789 antenna gain in a direction (θ_t, ϕ_t) at frequency f and distance R

$$w_t = \frac{P_t}{4\pi R^2} G_t(\theta_t, \phi_t, f), \quad (5.17)$$

3791 where the subscript t denotes the transmitting antenna, and P_t is the total power delivered
3792 to the transmitting antenna. The power density is power per unit area, so the total
3793 power delivered to the receiving antenna is the transmitted power density multiplied by
3794 the effective area of the receiving antenna

$$P_r = w_t A_{\text{eff},r} = P_t \frac{G_t(\theta_t, \phi_t, f) G_r(\theta_r, \phi_r, f) c^2}{(4\pi R f)^2}, \quad (5.18)$$

3795 where $G_r(\theta_r, \phi_r, f)$ is the gain of the receiving antenna. Equation 5.18 is called the Friis
3796 transmission equation [84], which is of fundamental importance for antenna measurements,
3797 since it allows one to measure the gain of an unknown antenna by measuring the power

3798 received from an antenna with a known gain pattern. Alternatively, if an antenna with a
 3799 known gain pattern is unavailable, two identical antennas with unknown gain patterns
 3800 can be used.

3801 **5.2.2.2 S-Parameters and Network Analyzers**

3802 It is more common to measure the ratio of the received power to the transmitted power
 3803 instead of the absolute received power

$$\frac{P_r}{P_t} = \frac{G_t(\theta_t, \phi_t, f) G_r(\theta_r, \phi_r, f) c^2}{(4\pi R f)^2}. \quad (5.19)$$

3804 This power ratio can be easily measured using a vector network analyzer (VNA), which
 3805 automates a significant fraction of the measurement process. Network analyzers are used
 3806 to measure the scattering or S-parameters of a multi-port RF device [85], which describes
 3807 how waves are scattered between the device ports. Friis antenna measurements can be
 3808 modeled as a two-port microwave device that is characterized by measuring how incident
 3809 voltage waves are transmitted or reflected (see Figure 5.4). The scattered waves (V_1^-

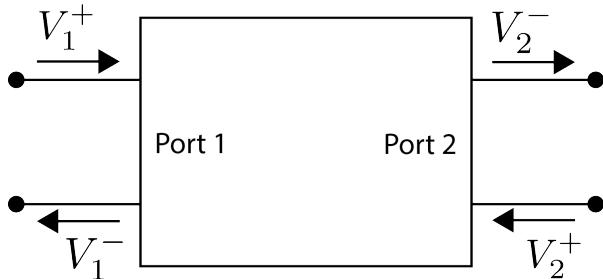


Figure 5.4. Illustration of a two-port S-parameter measurement setup. S-parameters characterize how incoming waves of voltage or power scatter off of the RF device under test. This allows you to measure important properties of the device. In particular, this framework can be used to model a two antenna radiation pattern measurement, which can be automated using a VNA.

3809
 3810 and V_2^-) can be written in terms of the incident (V_1^+ and V_2^+) waves using the scattering
 3811 matrix

$$\begin{pmatrix} V_1^- \\ V_2^- \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} V_1^+ \\ V_2^+ \end{pmatrix}, \quad (5.20)$$

3812 where the elements of the matrix are the device S-parameters. It is assumed that,
 3813 when exciting the device from a particular port, that all other ports in the network are
 3814 terminated at the system impedance. This ensures that the incident waves from other
 3815 ports in the network are zero. Therefore, the S-parameters are the ratios between the

3816 scattered and incident waves,

$$S_{ij} = \frac{V_i^-}{V_j^+}. \quad (5.21)$$

3817 Alternatively, S-parameters can be defined as the ratio of the scattered and incident
3818 power, which is proportional to the ratio of the squared voltage waves.

3819 Returning to the antenna measurement setup, it is clear that measuring the ratio of
3820 the received to the transmitted power is equivalent to measuring the ratio of power being
3821 scattered from port 1 to port 2 in a RF network. Therefore, measuring an antenna's gain
3822 can be accomplished quite easily using a VNA to perform a two port S_{21} measurement.

3823 **5.2.2.3 Antenna Array Commissioning and Calibration Measurements**

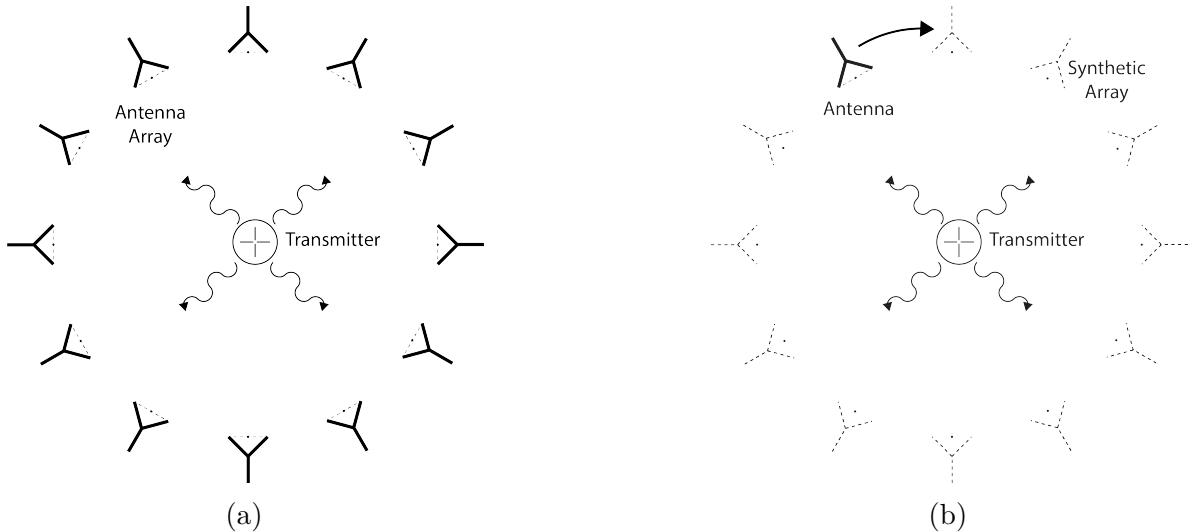


Figure 5.5. Two measurement approaches to characterizing an antenna array for CRES measurements. The full-array approach (a) requires a complete antenna array with all the associated hardware. The synthetic array approach (b) utilizes a single antenna and a set of rotation/translation stages to reposition the transmitter or the receiving antenna to synthesize the signals that would be received by the full-array. This approach reduces the cost and complexity of array measurements. A down-side of the synthetic array approach is that multi-channel effects such as reflections cannot be measured. Utilizing both the full-array and the synthetic array is a powerful way to quantify the impact of errors from the multi-channel array.

3824 Measuring the gain of each individual array element allows to predict the features of
3825 the signals received during a CRES event (see Section 5.2.1.5). However, unpredictable
3826 changes to the antenna performance can be introduced by the incorporation of the
3827 antennas into the circular array geometry, therefore, both individual antenna and full-

array characterization measurements are performed as part of the commissioning of the FSCD.

There are two main approaches to array measurements that could be used for characterization and calibration (see Figure 5.5). One approach is to construct the complete array and use a omni-directional transmitting antenna to measure the power received by each channel in the antenna array. In Section 5.3 I describe the development of an omni-directional transmitter that also mimics the radiation phase characteristics of a CRES event, which is useful because the entire array can be tested without repositioning. Alternatively, a full antenna array can be synthesized by repeatedly moving and measuring a single array element. This approach is ideal for identifying if different channels in the antenna array are affecting each other through multi-path interference by comparing the measurement results of the synthetic array to the real array.

5.2.3 The Penn State Antenna Measurement System

The development of antenna array based CRES requires the capability to test and calibrate different antenna array designs to validate the performance of the as-built antenna array before and during the experiment. With these aims in mind an antenna measurement system was developed at Penn State specifically designed to mimic the characteristics of the FSCD experiment.

The Penn State antenna measurement system utilizes a two antenna measurement configuration with a stationary reference antenna and a test antenna mounted on a set of motorized translation and rotation stages (see Figure 5.6). The antenna measurement system can be operated in two distinct modes, one focused on the characterization of the radiation patterns of prototype antennas, and the other focused on the validation of data-acquisition (DAQ) and CRES signal reconstruction techniques to bridge the gap between real measurements and simulation. In both measurement configurations, it is critical to isolate the antennas from the environment so that multi-path reflections do not negatively influence the measurement results. For this reason the measurement volume is surrounded with microwave absorber foam (AEMI AEC-1.5) specifically designed to attenuate microwave radiation near the 26 GHz measurement range of the system.

In the first measurement configuration, the reference antenna is a well-characterized horn antenna as pictured, since horn antennas have well-known and stable radiation patterns making them ideal as standard references. For characterization measurements, the test antenna represents the antenna-under-test whose pattern is being characterized. Mounting the test antenna on motorized rotation and translation stages allows for

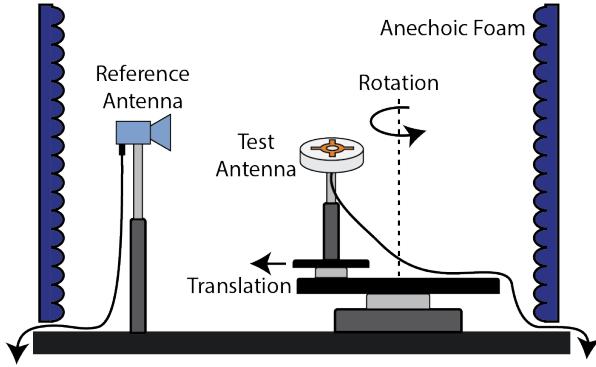


Figure 5.6. Illustration of the antenna measurement system developed for the Project 8 Collaboration. The reference and test antennas can be connected to different data acquisition configurations depending on the measurement goals. The reference antenna is typically a standard horn antenna and the test antenna is mounted on a set of translation stages for positioning. Automated translation stages allows for relatively painless data-taking enabling synthetic antenna array measurements using only a single receiving antenna. Anechoic form designed to mitigate RF reflections surrounds the setup.

3862 automation, which significantly speeds up the radiation pattern measurement process.

3863 The second measurement configuration mimics the conditions of the FSCD as it
 3864 concerns the antenna array and DAQ system. In this configuration, the reference antenna
 3865 is a prototype FSCD antenna, and the test antenna is a specially designed synthetic
 3866 cyclotron antenna (SYNCA) as picture in Figure 5.6. The SYNCA is designed such that
 3867 the radiation pattern mimics that of a CRES electron so that the signals received by the
 3868 prototype CRES array antenna mimic what is expected for a real CRES experiment.

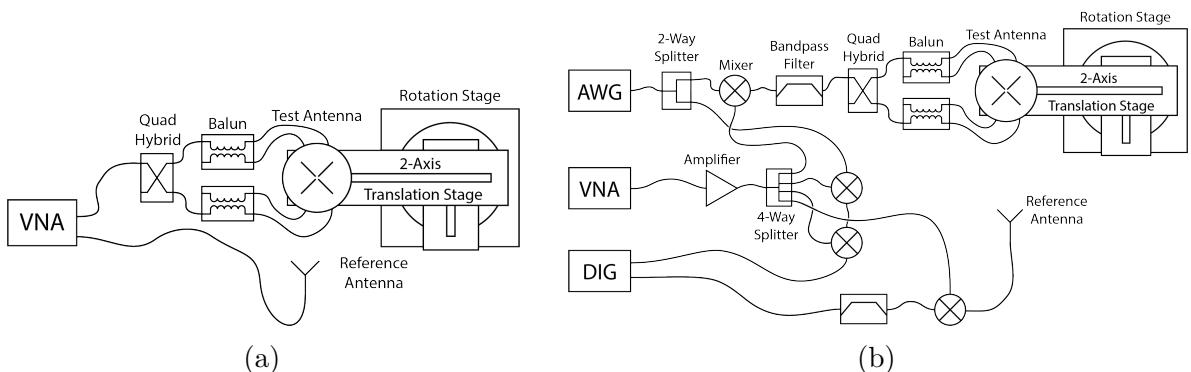


Figure 5.7. Diagrams of two measurement system configurations. Configuration (a) utilizes a VNA and is more suited to antenna characterization. Configuration (b) utilizes an AWG and VNA as a signal generation system and digitizer to collect measurement data, which is more suited to simulating CRES measurements. The transmission chain utilizes a quadrature hybrid and a pair of baluns to drive the cross-dipole variant test antenna developed for synthetic CRES measurements.

3869 Figure 5.7 shows two high-level system diagrams of the Penn State antenna measurement
3870 system that depict the important system components and the connections between
3871 them. The two configurations of the measurement system utilize different hardware. For
3872 characterization and radiation pattern measurements, the configuration shown in Figure
3873 5.7a is used. In this case a vector network analyzer (VNA) is used as the transmission
3874 source and data acquisition system, which is easy to calibrate over a wide range of
3875 frequencies. The configuration in 5.7b is used to mimic the FSCD experiment, since this
3876 system includes a more realistic receiver chain.

3877 The characterization configuration utilizes a network analyzer (Keysight N5222A)
3878 with two independent sources and four measurement ports as the primary measurement
3879 tool. A standard reference antenna is connected to one measurement port, and the test
3880 antenna is connected to a second port. The typical reference antenna used for these
3881 studies is a Pasternack PF9851 horn antenna. In the measurement shown, the test
3882 antenna represents a SYNCA antenna, which requires a transmission chain consisting of
3883 quadrature hybrid coupler (Marki QH-0226) connected to two baluns (Marki BAL-0026)
3884 to generate feed signals with the appropriate phases. The VNA measures the radiation
3885 pattern by performing a transmission S-parameter measurement, which can be used with
3886 the knowledge of the reference antenna's radiation pattern to determine the radiation
3887 pattern of the test antenna (see Section 5.2.1).

3888 The second configuration incorporates more hardware components to mimic the DAQ
3889 system envisioned for the FSCD experiment. The basic approach is to produce CRES-like
3890 radiation and use an antenna combined with a realistic RF receiver chain to acquire the
3891 signals. On the transmit side, an arbitrary waveform generator (AWG, RIGOL DG5252)
3892 is used to generate a waveform that mimics a CRES signal at a baseband frequency up
3893 to 250 MHz. This frequency is then up-converted to the CRES signal frequency band
3894 of 25.8 to 26.0 GHz using a mixer (Marki MM1-0832L) and a bandpass filter (K&L
3895 Microwave 3C62-25900/T200-K/K) to reject unwanted mixing components outside out
3896 of the 200 MHz CRES signal band. The local oscillator signal for mixing is provided by
3897 one of the VNA sources configured to run in a continuous wave setting. On the receive
3898 side, a prototype antenna is used to detect the radiation emitted by the test antenna,
3899 which is down-converted and filtered using the same mixer and bandpass filter as the
3900 transmission chain. Lastly, data acquisition is performed using a 14-bit ADC sampling
3901 at 500 MSa/s (CAEN DT530) to digitize the down-converted signals.

3902 In order to distribute the LO to all mixers a 4-way power splitter (MiniCircuits
3903 ZC4PD-18263-S+) along with an amplifier (Marki APM-6848) is used to drive the four

3904 mixers used in the measurement system. A limitation of using the VNA as an LO source
3905 is that there is no control of the LO phase when a measurement is triggered by the
3906 control script, which leads to a random phase offset between acquisitions. This makes it
3907 impossible to perform synthetic array measurements, which require strict control over
3908 the starting phase of the transmitted signal. In order to monitor the random phase of the
3909 LO, a 2-way power splitter (MiniCircuits Z99SC-62-S+) is used to split the signal from
3910 the AWG between the transmission path and a LO monitoring path. The LO monitoring
3911 path consists of an up-conversion and down conversion using two mixers connected by a
3912 coaxial cable, and monitors the relative phase of the LO using a channel on the digitizer
3913 to sample this path. A phase shift in the LO will lead to a proportional phase shift in
3914 the mixed signal, which is measured and removed from the received signals.

3915 The test antenna is mounted on a set of motorized stages, which are identical for
3916 both measurement configurations. A rotational stage (ThorLabs PRMTZ8) is used as
3917 the base layer with additional translation stages mounted on top. The rotational stage is
3918 ideal for measuring a complete azimuthal scan of the test antenna's radiation pattern
3919 as well as for moving a SYNCA antenna in circular motion to recreate the symmetry
3920 of the FSCD antenna array. On top of the rotational stage, are mounted two linear
3921 translation stages (ThorLabs MTS50-Z8 and MTS25-Z8) in a cross-wise manner so that
3922 the test antenna can be moved along two perpendicular axes. Using the linear stages in
3923 combination with the rotational stage allows one to fine-tune the positioning of the test
3924 antenna so that it can be perfectly aligned with the central axis of the array. A LabView
3925 script was developed to automate the measurement of a full 360° radiation pattern and
3926 control the measurement electronics. Data from these acquisitions is stored on university
3927 provided cloud storage.

3928 **5.3 Development of a Synthetic Cyclotron Antenna (SYNCA)** 3929 **for Antenna Array Calibration**

3930 This section is the manuscript of the publication [79] detailing the development of a
3931 Synthetic Cyclotron Antenna (SYNCA) for antenna array characterization measurements
3932 by the Project 8 collaboration.

3933 5.3.1 Introduction

3934 Neutrinos are the most abundant standard model fermions in our universe, but due to
3935 weak interaction cross-sections with other particles, neutrinos are particularly difficult
3936 to study. Consequently, many fundamental properties of neutrinos are still unknown
3937 including the absolute scale of the neutrino mass [23]. Direct, kinematic measurements of
3938 the neutrino mass are particularly valuable due to their model independent nature [36].
3939 To date the most sensitive direct neutrino mass measurements have been performed by
3940 the KATRIN collaboration [86], which measures the molecular tritium β -decay spectrum
3941 to infer the neutrino mass. Current data from neutrino oscillation measurements [23]
3942 allow for neutrino masses significantly smaller than the design sensitivity of the KATRIN
3943 experiment; therefore, there is a need for new technologies for performing direct neutrino
3944 mass measurements to probe lower neutrino masses.

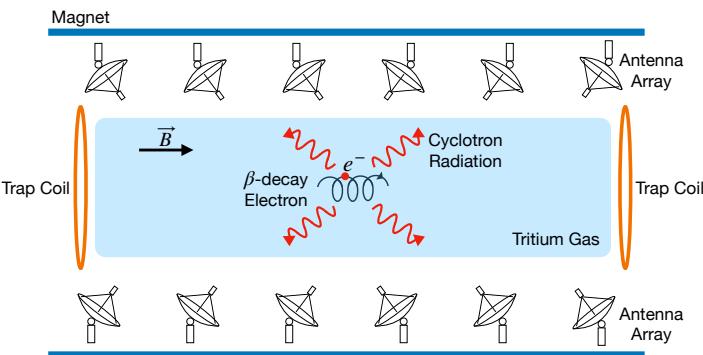


Figure 5.8. A sketch of an antenna array large-volume CRES experiment. Electrons from β -decays are confined in a magnetic field using a set of trap coils. The cyclotron radiation produced by the motion of the trapped electrons can be detected by a surrounding antenna array to determine the electron energies. Measuring the energies of many electrons produces a β -decay spectrum.

3945 The Project 8 collaboration is developing new methods for neutrino mass measurement
3946 based on Cyclotron Radiation Emission Spectroscopy (CRES) [55, 87–89], with the goal
3947 of measuring the absolute scale of the neutrino mass with a $40 \text{ meV}/c^2$ sensitivity [?, 36].
3948 This sensitivity goal will require the development of two separate technical capabilities.
3949 First is the development of an atomic tritium source, which avoids significant spectral
3950 broadening due to molecular final states [54]. Second is the technology for performing
3951 CRES in a multi-cubic-meter experimental volume with high combined detection and
3952 reconstruction efficiency, which is required in order to obtain sufficient event statistics
3953 near the tritium spectrum endpoint.

One approach for a large-volume CRES experiment is to use an array of antennas, which surrounds a volume of tritium gas, to detect the cyclotron radiation produced by the β -decay electrons when they are trapped in a background magnetic field using a set of magnetic trapping coils (see Figure 5.8). Project 8 has developed a conceptual experiment design to study the feasibility of this approach. The design consists of a single circular array of antennas with a radius of 10 cm and 60 independent channels positioned around the center of the magnetic trap. The motivation behind this antenna array design is to first develop an understanding of the antenna array approach to CRES with a small scale experiment before attempting to scale the technique to large volumes by using multiple antenna rings to construct the full cylindrical array. The development of the antenna array approach to CRES has largely proceeded through simulations using the Locust software package [78, 90], which is used to model the fields emitted by CRES events and predict the signals received by the surrounding antenna array. To validate these simulations, a dedicated test stand is being constructed to perform characterization measurements of the prototype antenna array developed by Project 8 (see Figure 5.9) and benchmark signal reconstruction methods using a specially designed transmitting calibration probe antenna.

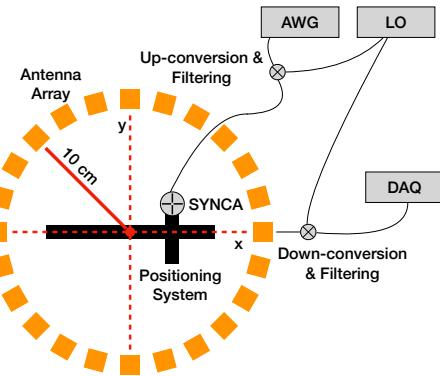


Figure 5.9. A schematic of the antenna array test stand. The circular antenna array has a radius of 10 cm with 60 independent channels (limited number shown for clarity). The test stand includes an arbitrary waveform generator (AWG), local oscillator (LO), and data acquisition (DAQ) hardware. Finally, a specialized Synthetic Cyclotron Antenna (SYNCA) is used to inject signals to test the antenna array.

We call this probe antenna the Synthetic Cyclotron Antenna or SYNCA. The SYNCA is a novel antenna design that mimics the cyclotron radiation generated by individual charged particles trapped in a magnetic field, which will be used in the antenna test stand to perform characterization measurements, simulation validation, and reconstruction benchmarking. This paper provides an overview of the design, construction, and

3976 characterization measurements of the SYNCA performed in preparation for its usage as
 3977 a transmitting calibration probe.

3978 In Section 5.3.2 we provide a description of the cyclotron radiation field characteristics
 3979 that we recreate with the SYNCA. In Section 5.3.3 we give an overview of the simulations
 3980 performed to develop an antenna design that mimics the characteristics of cyclotron
 3981 radiation. In Section 5.3.4 we outline characterization measurements to validate that
 3982 the fields generated by the SYNCA match simulation, and finally in Section 5.3.5 we
 3983 demonstrate an application of the SYNCA to test phased array reconstruction techniques
 3984 on the bench-top.

3985 5.3.2 Cyclotron Radiation Phenomenology

3986 To understand the cyclotron radiation phenomenology that the SYNCA should mimic,
 3987 we consider a charged particle moving at relativistic speed in the presence of an external
 3988 magnetic field (see Figure 5.10). In the special case we shall examine, the entirety of
 3989 the electron's momentum is directed perpendicular to the magnetic field; therefore, the
 3990 trajectory of the electron is confined to the cyclotron orbit plane. Because the momentum
 3991 vector is oriented perpendicular to the magnetic field, electrons with these trajectories
 3992 are said to have pitch angles of 90°.

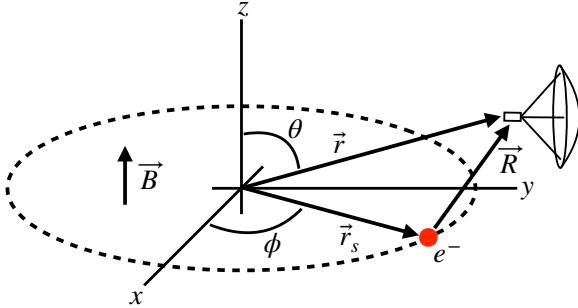


Figure 5.10. An electron (red dot) performing cyclotron motion in the x-y plane. The resulting cyclotron radiation is observed by an antenna located at the field point of interest.

3993 The cyclotron radiation fields generated by this circular trajectory are those which
 3994 we aim to reproduce with the SYNCA. We can describe the electromagnetic (EM) fields
 3995 using the Liénard-Wiechert equations [48, 78], which in non-covariant form express the
 3996 electric field as

$$\vec{E} = e \left[\frac{\hat{n} - \vec{\beta}}{\gamma^2(1 - \vec{\beta} \cdot \hat{n})^3 |\vec{R}|^2} \right]_{tr} + \frac{e}{c} \left[\frac{\hat{n} \times [(\hat{n} - \vec{\beta}) \times \dot{\vec{\beta}}]}{(1 - \vec{\beta} \cdot \hat{n})^3 |\vec{R}|} \right]_{tr}, \quad (5.22)$$

3997 where e is the particle's charge, $\hat{n} = (\vec{r} - \vec{r}_s)/|\vec{r} - \vec{r}_s|$ is the unit vector pointing from the
 3998 electron to the field measurement point, $\vec{\beta} = \dot{\vec{r}}_s/c$ is the velocity of the particle divided
 3999 by the speed of light, and γ is the relativistic Lorentz factor. The equation is meant to
 4000 be evaluated at the retarded time as indicated by $t_r = t - |\vec{R}|/c$, which accounts for the
 4001 time delay due to the finite speed of light between the point where the field was emitted
 4002 and the point where the field is detected.

4003 We would like to simplify Equation 5.22 it at all possible. As a first step we analyze
 4004 the relative magnitudes of the electric field polarization components. Consider an electron
 4005 following a circular cyclotron orbit in a uniform magnetic field whose guiding center
 4006 is positioned at the origin of the coordinate system. The equation of motion can be
 4007 expressed as

$$\vec{r}_s = (r_c \cos \omega_c t_r) \hat{x} + (r_c \sin \omega_c t_r) \hat{y}. \quad (5.23)$$

4008 For single antenna located along the y -axis at position $\vec{r} = r_a \hat{y}$ we are interested in the
 4009 incident electric fields from the electron. The electric field is given by Equation 5.22,
 4010 which we evaluate in the regime where $r_a \gg r_c$. This limit can be justified by comparing
 4011 the radius of the cyclotron orbit for an electron with the tritium beta-spectrum endpoint
 4012 energy of 18.6 keV in a 1 T magnetic field to the typical ($r_a \simeq 100$ mm) radial position
 4013 of the receiving antenna. We find that the cyclotron orbit has a radius of 0.46 mm which
 4014 is approximately a factor of 200 smaller than the typical antenna radial position. In this
 4015 regime we can make the approximation $\vec{R} \simeq r_a \hat{y}$ and the expression for the electric field
 4016 at the antenna's position becomes

$$\vec{E} = \frac{e}{\gamma^2 r_a^2} \frac{\hat{x} \left(\frac{r_c \omega_c}{c} \sin \omega_c t_r \right) + \hat{y} \left(1 - \frac{r_c \omega_c}{c} \cos \omega_c t_r \right)}{(1 - \frac{r_c \omega_c}{c} \cos \omega_c t_r)^3} - \frac{e}{c r_a} \frac{\hat{x} \left(\frac{r_c^2 \omega_c^3}{c^2} - \frac{r_c \omega_c^2}{c} \cos \omega_c t_r \right)}{(1 - \frac{r_c \omega_c}{c} \cos \omega_c t_r)^3}. \quad (5.24)$$

4017 Since the receiving antenna is part of a circular array of antennas, it is useful to rewrite
 4018 Equation 5.24 in terms of the azimuthal ($\hat{\phi}$) and radial (\hat{r}) polarizations. Making use of
 4019 the fact that for an antenna located at $R = r_a \hat{y}$ that $\hat{\phi} = -\hat{x}$ and $\hat{r} = \hat{y}$ we find

$$\vec{E} = \hat{\phi} E_\phi + \hat{r} E_r \quad (5.25)$$

$$E_\phi = \frac{e}{(1 - \frac{r_c \omega_c}{c} \cos \omega_c t_r)^3} \left[-\frac{\frac{r_c \omega_c}{c} \sin \omega_c t_r}{\gamma^2 r_a^2} + \frac{\omega_c \left(\frac{r_c^2 \omega_c^2}{c^2} - \frac{r_c \omega_c}{c} \cos \omega_c t_r \right)}{c r_a} \right] \quad (5.26)$$

$$E_r = \frac{e \left(1 - \frac{r_c \omega_c}{c} \sin \omega_c t_r \right)}{\gamma^2 r_a^2 (1 - \frac{r_c \omega_c}{c} \cos \omega_c t_r)^3}. \quad (5.27)$$

4020 For the purposes of designing a synthetic cyclotron radiation antenna we are interested
 4021 in the dominant electric field polarization emitted by the electron. The antenna is being
 4022 designed to mimic the cyclotron radiation produced by electrons with kinetic energies of
 4023 approximately 18.6 keV in a 1 T magnetic field [54]. Since the relativistic beta factor for
 4024 an electron with this kinetic energy is $|\vec{\beta}| \simeq \frac{1}{4}$, the approximations $\gamma \simeq 1$ and $\frac{r_c \omega_c}{c} \simeq \frac{1}{4}$ are
 4025 justified. Inserting these expressions into the equations for the electric field components
 4026 above simplifies the comparison of the magnitudes of the two components. Additionally,
 4027 we compare the time-averaged magnitudes to evaluate the root mean squared electric
 4028 field ratio. The time-averaged ratio of the radial and azimuthally polarized electric fields
 4029 with the above simplifications is given by

$$\frac{\langle |E_r| \rangle}{\langle |E_\phi| \rangle} = \frac{8 - \sqrt{2}}{\left| 1 - \frac{r_a}{r_c} \frac{1-2\sqrt{2}}{8} \right|} \simeq \frac{r_c}{r_a} \frac{8(8 - \sqrt{2})}{2\sqrt{2} - 1} = 0.13, \quad (5.28)$$

4030 where we have made use of the fact that for these magnetic fields and kinetic energies
 4031 the cyclotron radius is much smaller than the radius of the antenna array.

4032 From Equation 5.28 we see that the time-averaged azimuthal polarization is larger than
 4033 the radial polarization by about a factor of 8, which makes it the dominant contribution
 4034 to the electric fields at the position of the antenna. We must also consider the directivity
 4035 of the receiving antenna which can have a gain that is disproportionately large for a
 4036 specific polarization component. Because the E_ϕ component is dominant, the receiving
 4037 antenna array is designed with an azimuthal polarization, which negates the voltages
 4038 induced in the antenna from the radially polarized fields. Therefore, we conclude that
 4039 for the purpose of designing the SYNCA antenna it is acceptable to approximate the
 4040 electric fields from Equation 5.22 as purely azimuthally or ϕ -polarized. The simplified
 4041 expression for the electric field received by an antenna becomes

$$\vec{E} = E_\phi \hat{\phi} = \frac{e^{\frac{r_c \omega_c}{c}}}{4r_a r_c} \left[\frac{\frac{r_c \omega_c}{c} - \cos \omega_c t - \frac{4r_c}{r_a} \sin \omega_c t}{(1 - \frac{r_c \omega_c}{c} \cos \omega_c t)^3} \right]_{t_r} \hat{\phi}, \quad (5.29)$$

4042 where the radius of the cyclotron orbit is called r_c , the cyclotron frequency is called ω_c ,
 4043 and the radial position of the receiving antenna is called r_a . Equation 5.29 has been
 4044 evaluated in the non-relativistic limit where $\gamma \simeq 1$, which is justified by the fact that
 4045 $|\vec{\beta}| \simeq \frac{c}{4}$ for an electron with an 18.6 keV kinetic energy in a 1 T magnetic field.

4046 This rather complicated expression can be simplified using Fourier analysis. Assuming
 4047 a background magnetic field of 1 T and a kinetic energy of 18.6 keV we calculate

4048 numerically the electric field using Equation 5.29 and apply a discrete Fourier Transform
to visualize the frequency spectrum (see Figure 5.11).

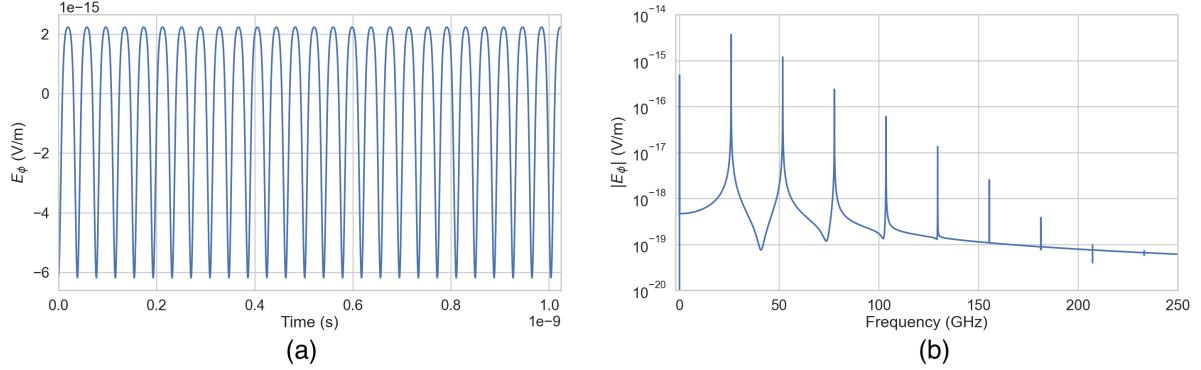


Figure 5.11. A plot of the numeric solution to Equation 5.30. The time-domain representation of the signal (a) is composed of a zero frequency term and a series of harmonics separated by the main cyclotron frequency as shown in the plot of the frequency spectrum (b). We can see that the relative amplitude of the harmonics beyond $k = 7$ are smaller than the main carrier by a factor of about 10^{-5} and are completely negligible.

4049
4050 We observe that the azimuthally polarized electric field is periodic with a base cyclotron
4051 frequency of 25.898 GHz corresponding to the highest power frequency component in
4052 Figure 5.11. The frequency spectrum reveals that the signal is composed of a constant
4053 term with zero frequency and a series of harmonics separated by 25.898 GHz. Therefore,
4054 we can represent the azimuthal electric fields from the electron as a linear combination
4055 of pure sinusoids with frequencies given by $\omega_k = k\omega_c$ ($k \in 0, 1, 2, \dots$) and amplitudes
4056 extracted from the Fourier representation. Using this representation we can transform
4057 the equation for the azimuthally polarized electric fields in Equation 5.29 into

$$E_\phi = \frac{e^{\frac{r_c \omega_c}{c}}}{4r_a r_c} \sum_{k=0}^7 A_k e^{i\omega_k t_r}, \quad (5.30)$$

4058 where we have truncated the sum over harmonics at the 7th order for completeness. The
4059 amplitudes A_k are dimensionless complex numbers, which encode the relative powers of
4060 the harmonics as well as the starting overall phase of the cyclotron radiation. Because
4061 magnitude of the relative amplitudes exponentially decreases for higher harmonics, it is
4062 usually sufficient to consider only the terms up to $k = 4$ where the relative amplitude
4063 of the harmonics has decreased from the main carrier by a factor of approximately 100.
4064 However, for completeness we include harmonics up to 7th order in Equation 5.30. The
4065 range of frequencies to which the receiving antenna array in the antenna test stand is
4066 sensitive is defined by the antenna's transfer function. The receptive bandwidth for

4067 the antennas used in the test stand is a range of frequencies with a bandwidth on the
 4068 order of a few GHz centered around the main cyclotron carrier frequency of 25.898 GHz.
 4069 Therefore, the higher order harmonics as well as the zero frequency term can be ignored
 4070 when considering only the signals that will be received by the antenna array.

4071 Considering only the 1st order harmonic term from Equation 5.30, which represents
 4072 the portion of the electric field that will be detected by the array, and evaluating this at
 4073 the retarded time we obtain the following for the ϕ -polarized electric fields

$$E_\phi \propto \cos \left(\omega_c \left(t - |\vec{R}|/c \right) - \Delta \right), \quad (5.31)$$

4074 where the arbitrary phase Δ is defined by $A_k = |A_k|e^{i\Delta}$. We are interested in the
 4075 characteristics of the amplitude of the electric field as a function of the radial distance
 4076 component ($|\vec{R}|$) of the retarded time. In particular, the maximum of E_ϕ occurs when
 4077 the argument of the cosine function is equal $n\pi$ where $n \in \{0, \pm 2, \pm 4, \dots\}$; however, the
 4078 solutions where n is negative can be discarded since they represent unphysical negative
 4079 overall phases. Applying this condition to Equation 5.31 gives a condition on the radial
 4080 position of the maximum of E_ϕ

$$\omega_c(t - |\vec{R}|/c) - \Delta = n\pi, \quad (5.32a)$$

$$|\vec{R}| = \frac{c}{\omega_c} ((\omega_c t - \Delta) - n\pi), \quad (5.32b)$$

4081 which is a function of time in the frame of the moving electron (t). Equation 5.32 can
 4082 be further simplified by noticing that the azimuthal position of the electron ($\phi_e(t)$) as a
 4083 function of time is defined by $\phi_e(t) = \omega_c t - \Delta$ which reduces Equation 5.32 to

$$|\vec{R}| = \frac{c}{\omega_c} (\phi_e(t) - n\pi). \quad (5.33)$$

4084 Equation 5.33 represents an archimedian spiral which is formed when plotting the
 4085 amplitude of E_ϕ in the x-y plane. The solution where $n = 0$ represents the leading edge
 4086 of the radiation spiral which propagates outward from the electron at the speed of light.
 4087 The additional solutions for $n > 0$ represent the persistent spiral at radii inside the
 4088 leading edge of the radiated fields that have not yet been detected by the receiver at the
 4089 current time. In Figure 5.12a we show the expected spiral pattern for the maxima of the
 4090 cyclotron radiation.

4091 In particular, we note that for the circular array geometry of the test stand, depicted
 4092 as the series of circles in Figure 5.12a, each antenna receives a linearly polarized wave

4093 with a phase offset that corresponds to the azimuthal angle for that antenna element.
 4094 Therefore, as we show in Figure 5.12b, when the relative phase of the received signal is
 4095 plotted as a function of the receiving antenna's azimuthal position the result is also an
 4096 Archimedean spiral.

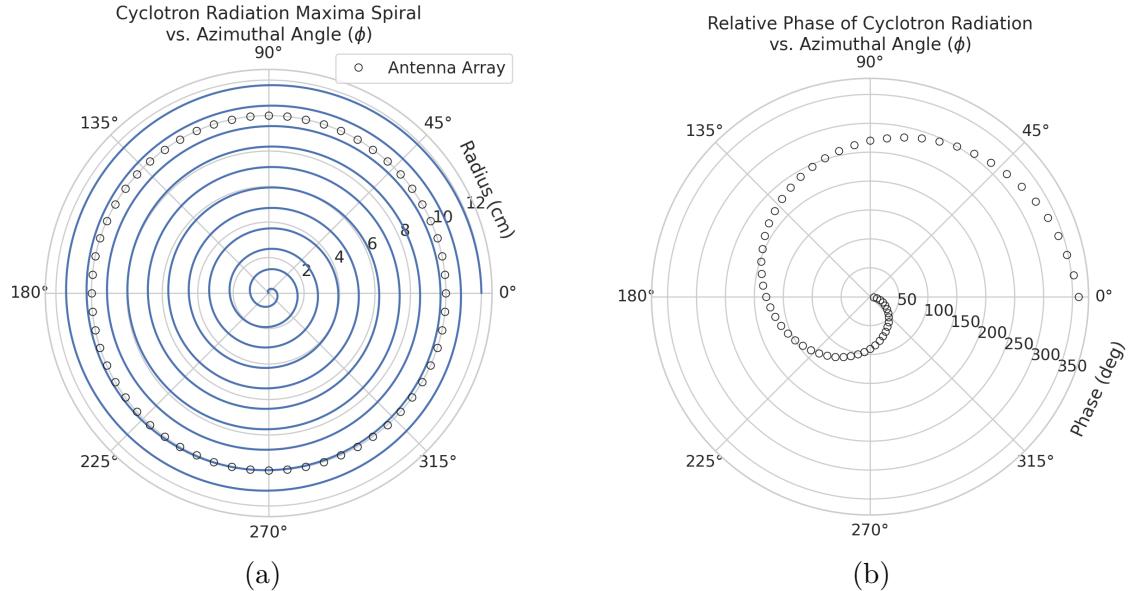


Figure 5.12. The amplitude maxima of the cyclotron radiation form an Archimedean spiral as the radiation propagates outward from the cyclotron orbit center (a). A circular antenna array located at a fixed radius from the orbit center will receive electric fields with equal magnitude in each of its channels, but the phase of the electric field incident on each array channel will be linearly out of phase from its neighbor antennas by an amount equal to the angular separation of the two channels (b).

4097 Based on these analytical calculations we can characterize the magnitude, polarization,
 4098 and phase of the signals received by the antenna array using three criteria. These criteria
 4099 are the basis of comparison for the radiation produced by the SYNCA and cyclotron
 4100 radiation emitted by electrons and will be used to evaluate the performance of antenna
 4101 designs. The criteria are:

- 4102 1. Electric fields that are ϕ -polarized near $\theta = 90^\circ$
- 4103 2. Uniform time-averaged electric field magnitudes around the circumference of a
 4104 circle centered on the antenna
- 4105 3. Electric fields whose phase is equal to the azimuthal angle at the point of measure-
 4106 ment plus a constant

4107 The Locust simulation package [90] can be used to directly simulate the EM fields
 4108 generated by electrons performing cyclotron motion to validate the analytical calculations.
 4109 Locust simulates the EM fields by first calculating the trajectory of the electrons in
 4110 the magnetic trap using the Kassiopeia software package [91]. The trajectory can then
 4111 be used to solve for the EM fields using the Liénard-Wiechert equations directly with
 4112 no approximations. The resulting electric field solutions drive a receiving antenna by
 4113 convolving the time-domain fields with the finite-impulse response filter of the antenna
 4114 or they can be examined directly to study the field characteristics that the SYNCA must
 4115 reproduce. In the next section we compare the radiation field patterns for electrons
 4116 simulated with Locust to patterns from a SYNCA antenna design.

4117 **5.3.3 SYNCA Simulations and Design**

4118 One potential SYNCA design is the crossed-dipole antenna [92]. A crossed-dipole antenna
 4119 consists of two dipole antennas, one of which is rotated 90° with respect to the other,
 4120 which are fed with signals that are out of phase from the opposite dipole by 90° (see
 Figure 5.13). This arrangement causes the signals fed to each arm of the dipole to be

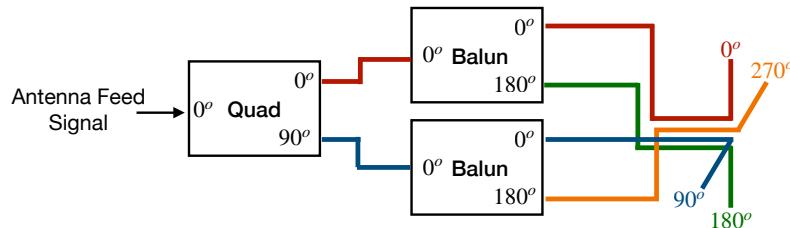


Figure 5.13. An idealized crossed-dipole antenna consists of two electric dipole antennas oriented perpendicular to each other and is fed with four signals with a quadrature phase relationship. An example antenna feed circuit is shown which is composed of a chained combination of a quadrature hybrid-coupler (Quad) and two baluns.

4121
 4122 out of phase from each of the neighboring arms by 90°, which mirrors the spatial phase
 4123 relationship of cyclotron radiation fields.

4124 A potential drawback of this design is that standard crossed-dipole antennas do not
 4125 radiate uniform electric fields near the $\theta = \pi/2$ plane. Typical crossed-dipole antennas
 4126 use dipole arm lengths equal to $\lambda/4$ or larger [92], where λ is the wavelength at the
 4127 desired operating frequency. Such large arm lengths cause the electric field magnitude
 4128 to vary significantly around the circumference of the antenna. However, making the

4129 antenna electrically small by shrinking the arm length can improve the antenna pattern
4130 uniformity.

4131 In general, the criterion for an electrically small antenna is that the largest dimension
4132 of the antenna (D) obey $D \lesssim \lambda/10$ [64]. In our application, we are attempting to mimic
4133 the cyclotron radiation emitted by electrons produced from tritium β -decay with energies
4134 near the spectrum endpoint. For a background magnetic field of 1 T, the corresponding
4135 cyclotron frequency of tritium endpoint electrons is approximately 26 GHz. Therefore, the
4136 electrically small condition would require that the largest dimension of the crossed-dipole
4137 antenna be smaller than 1.2 mm.

4138 A crossed-dipole antenna with an overall size of 1.2 mm is challenging to fabricate due
4139 to the small dimensions of the dipole arms that, in practice, are fragile and unsuitable
4140 for use as a calibration probe. To mitigate some of the challenges with the fabrication
4141 of such a small antenna, a variant crossed-dipole antenna design using printed circuit
4142 board (PCB) technology (see Figure 5.14) was developed in partnership with an antenna
prototyping company, Field Theory Consulting ¹.

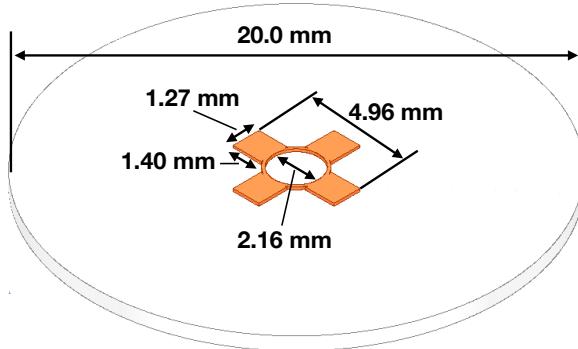


Figure 5.14. A model of the PCB crossed-dipole antenna with dimensions. The design has an inside diameter of 2.16 mm for the central circular trace, which is 0.13 mm wide. The dipole arms each have a width of 1.27 mm and protrude beyond the circular trace by 1.40 mm, which gives an overall width of 4.96 mm for the length of the antenna PCB trace from end-to-end. The overall size of the antenna is 20.0 mm the majority of which is the PCB dielectric material. This design was observed in simulation to maintain the field characteristics of the idealized crossed-dipole while being simpler to fabricate due to the increased size of the antenna.

4143
4144 The PCB crossed-dipole design uses four rectangular pads to represent the dipole arms,
4145 which are connected by a thin circular trace. The circular trace both adds mechanical
4146 stability to the antenna and improves the azimuthal uniformity of the electric fields
4147 compared to a more standard crossed-dipole geometry. Furthermore, the circular trace

¹<https://fieldtheoryinc.com/>

4148 allows for a greater separation between dipole arms than standard crossed-dipoles, which
 4149 is required to accommodate the coaxial connections to each pad. The pads each contain
 4150 a through-hole solder joint to connect coaxial transmission lines using hand soldering.
 4151 The antenna PCB has no ground plane on the bottom layer as this was observed in
 4152 simulation to significantly distort the radiation pattern in the plane of the PCB. The
 4153 only ground planes present in the model are the outer conductors of the four coaxial
 4154 transmission lines which feed the antenna. These are left unterminated on the bottom of
 4155 the PCB dielectric material.

4156 The antenna design development utilized a combination of Locust electron simula-
 4157 tions and antenna simulations using ANSYS HFSS [65], a commercial finite-element
 4158 electromagnetic simulation software. Two antenna designs were simulated: an idealized
 4159 electrically small crossed-dipole antenna with an arm length of 0.40 mm and an arm
 4160 separation of 0.05 mm, as well as a PCB crossed-dipole antenna with the dimensions
 4161 shown in Figure 5.14. Plotting the magnitude of the electric fields generated by the
 4162 antennas across a 10 cm square located in the same plane as the respective antennas
 4163 reveals the expected cyclotron spiral pattern (see Figure 5.15) which closely matches
 4164 the prediction for simulated electrons. The spiral pattern demonstrates that the electric
 4165 fields have the appropriate phases to mimic cyclotron radiation, which fulfills SYNCA
 criterion 3 identified in Section 5.3.2.

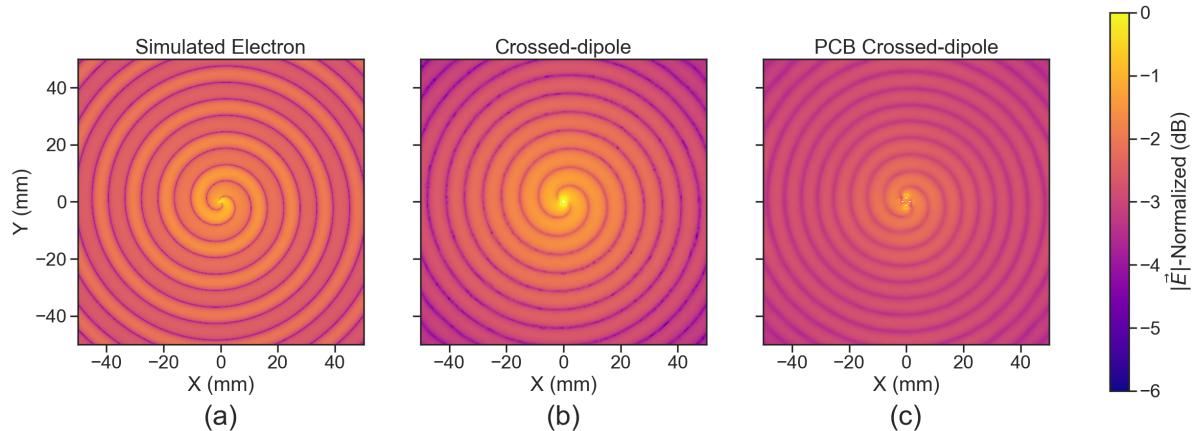


Figure 5.15. A comparison of the electric field magnitudes, normalized by the maximum value of the electric field in each simulation, plotted on a 10 cm square to visualize the Archimedean spirals formed by the electron (a), the crossed-dipole antenna (b), and a PCB crossed-dipole antenna (c). The matching patterns indicate that the electric fields have similar phase characteristics. These images were generated using Locust simulations for the electron and ANSYS HFSS for both antennas.

4166

4167 As we can see from Figure 5.16, the crossed-dipole antenna, which uses an idealized

4168 geometry, exhibits good agreement with simulation. The antenna has a maximum
 4169 deviation from a simulated electron of approximately 0.5 dB in the total electric field, 1
 4170 dB for the ϕ -polarized electric field and 1 dB for the θ -polarized electric field.

4171 In comparison, the pattern of the PCB crossed-dipole antenna, because the simulation
 4172 incorporates the geometry of the coax transmission lines, exhibits some distortion from
 4173 the idealized cross-dipole simulations. The vertically oriented ground planes of the coax
 4174 lines introduce more θ -polarized electric fields than are observed for simulated electrons
 4175 near $\theta = 90^\circ$. The significant θ -polarized field minimum is still present but shifted
 to approximately $\theta = 65^\circ$. The θ -polarized field deviations of the PCB crossed-dipole

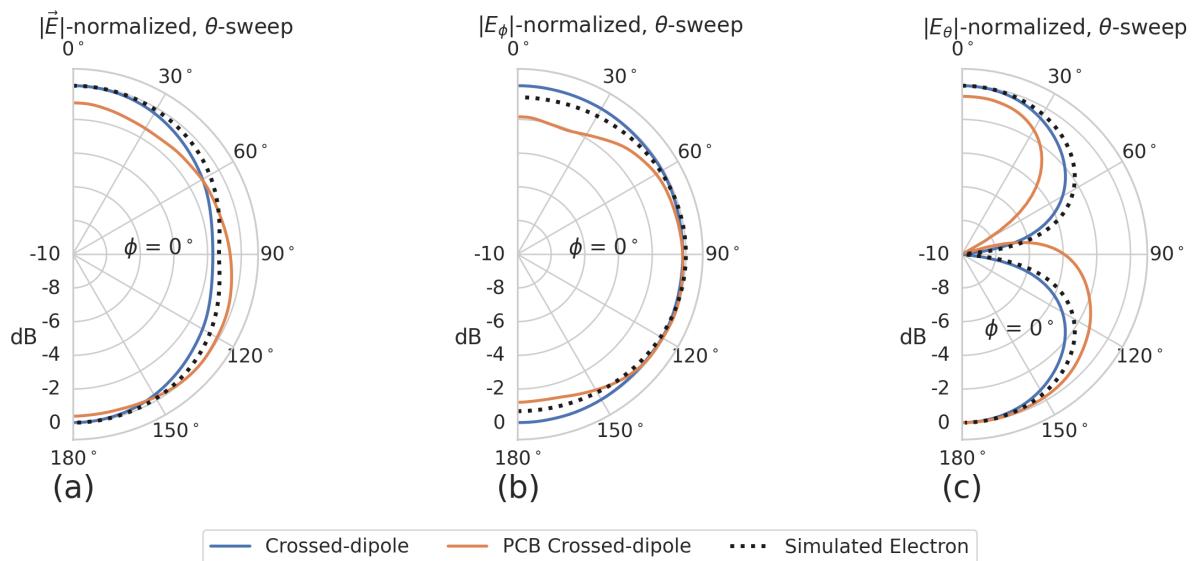


Figure 5.16. A comparison of the normalized electric field magnitudes for the ideal crossed-dipole, PCB crossed-dipole, and a simulated electron as a function of the polar angle (θ). (a) Shows the total electric field, (b) shows the ϕ -polarized electric field component, and (c) shows the θ -polarized electric field component. These images were generated using Locust simulations for the electron and ANSYS HFSS for both antennas.

4176
 4177 antenna should not greatly impact the performance of the antenna because the receiving
 4178 antenna array is primarily ϕ -polarized. Therefore deviations in the θ -polarized fields
 4179 will be suppressed due to the polarization mismatch. More importantly, the ϕ -polarized
 4180 electric field pattern generated by the PCB crossed-dipole closely matches simulated
 4181 electrons across the polar angle range of $50^\circ < \theta < 150^\circ$. In this region the PCB crossed-
 4182 dipole differs by less than 0.5 dB from simulated electrons. This range greatly exceeds
 4183 the beamwidth of the receiving antenna array which is designed to be most sensitive
 4184 to fields produced near $\theta = 90^\circ$. Therefore, we conclude that the PCB crossed-dipole
 4185 antenna generates a ϕ -polarized radiation pattern that fulfills SYNCA criterion 1 from

4186 Section 5.3.2.

4187 The final SYNCA criterion is related to the uniformity of the electric fields when
4188 measured azimuthally around the antenna. As we saw for real electrons in Section 5.3.2
4189 it is expected that the magnitude of the electric field be completely uniform as a function
4190 of the azimuthal angle due to the symmetry of the cyclotron orbit. In Figure 5.17 we plot
4191 the total electric field as a function of azimuthal angle for an electron, the crossed-dipole
antenna, and the PCB crossed-dipole antenna. The crossed-dipole antenna exhibits

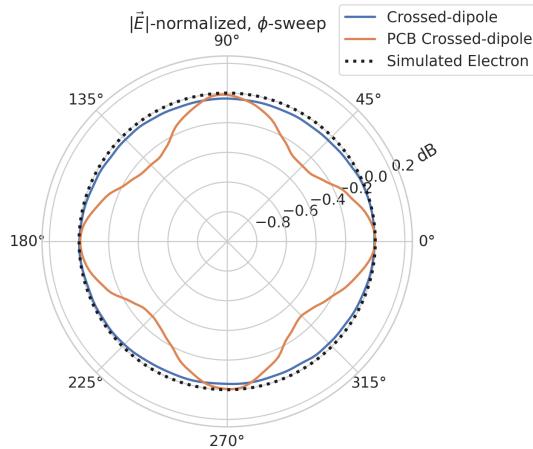


Figure 5.17. A comparison of the normalized electric field magnitudes for the crossed-dipole, PCB crossed-dipole, and a simulated electron as a function of the azimuthal angle (ϕ) evaluated at $\theta = 90^\circ$. This image was generated using Locust simulations for the electron and ANSYS HFSS for both antennas.

4192
4193 perfect uniformity around the azimuthal angle, whereas the PCB crossed-dipole has a
4194 small periodic deviation with a maximum difference of 0.3 dB caused by the coaxial
4195 transmission lines below the PCB. Such a small deviation from uniformity is acceptable
4196 since it is smaller than the expected variation in uniformity caused by imperfections in
4197 the antenna fabrication process, which modifies the antenna shape in an uncontrolled
4198 manner by introducing solder blobs with a typical size of a few tenths of a millimeter on
4199 the dipole arms (see Figure 5.18). Additionally, the SYNCA will be separately calibrated
4200 to account for azimuthal differences in the electric field magnitude. Therefore we see
4201 from the simulated performance of the PCB crossed-dipole antenna that this antenna
4202 design meets all three of the SYNCA criteria.

4203 **5.3.4 Characterization of the SYNCA**

4204 Two SYNCAs were manufactured using the PCB crossed-dipole design (see Figure 5.18).
4205 The antenna PCB (Matrix Circuit Board Materials, MEGTRON 6) is connected to

4206 four 2.92 mm coaxial connectors (Fairview Microwave, SC5843) using semi-rigid coax
 4207 (Fairview Microwave, FMBC002), which also physically support the antenna PCB. The
 4208 antenna PCB consists only of two layers which correspond to the copper antenna trace
 4209 and the PCB dielectric. Each coax line is connected to the associated dipole arm using
 4210 through-hole soldering and phase matched to ensure that the electrical length of each
 4211 of the transmission lines is identical at the operating frequency. The antenna PCB is
 4212 further reinforced using custom cut polystyrene foam blocks, which have an electrical
 4213 permittivity nearly identical to air. A custom 3D printed mount is included at the base
 4214 of the antenna to support the coax connectors and to provide a sturdy mounting base.

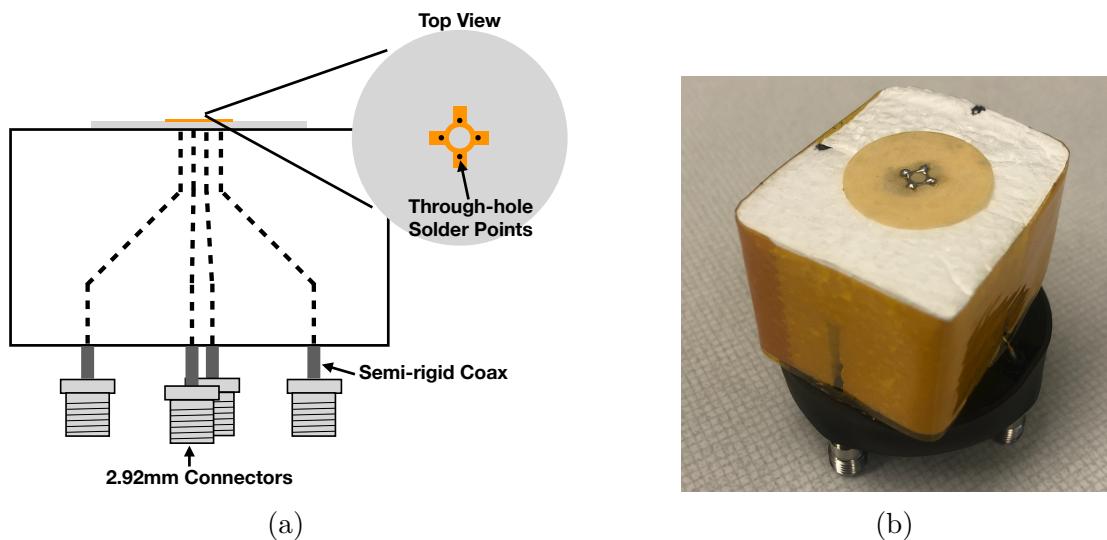


Figure 5.18. (a) A cartoon schematic which highlights the routing of the semi-rigid coax transmission lines. (b) A photograph of a SYNCA constructed using the modified crossed-dipole PCB antenna design. Visible in the photograph of the SYNCA are four blobs of solder which are an artifact of the SYNCA’s hand-soldered construction. These solder blobs are the most significant deviation from the SYNCA design shown in Figure 5.14 and are responsible for a significant fraction of the irregularities seen in the antenna pattern.

4215 Characterization measurements were performed using a Vector Network Analyzer
 4216 (VNA) to measure the electric field magnitude and phase radiated by the SYNCA to
 4217 verify the radiation pattern (see Figure 5.19). The VNA is connected to the SYNCA
 4218 at one port through a hybrid-coupler whose outputs are connected to two baluns to
 4219 generate the signals with the appropriate phases to feed the SYNCA (see Figure 5.13).
 4220 The other port of the VNA is connected to a single reference horn antenna that serves
 4221 as a field probe. To position the SYNCA, a combination of translation and rotation
 4222 stages are used to characterize the antenna’s fields across the entire radiation pattern

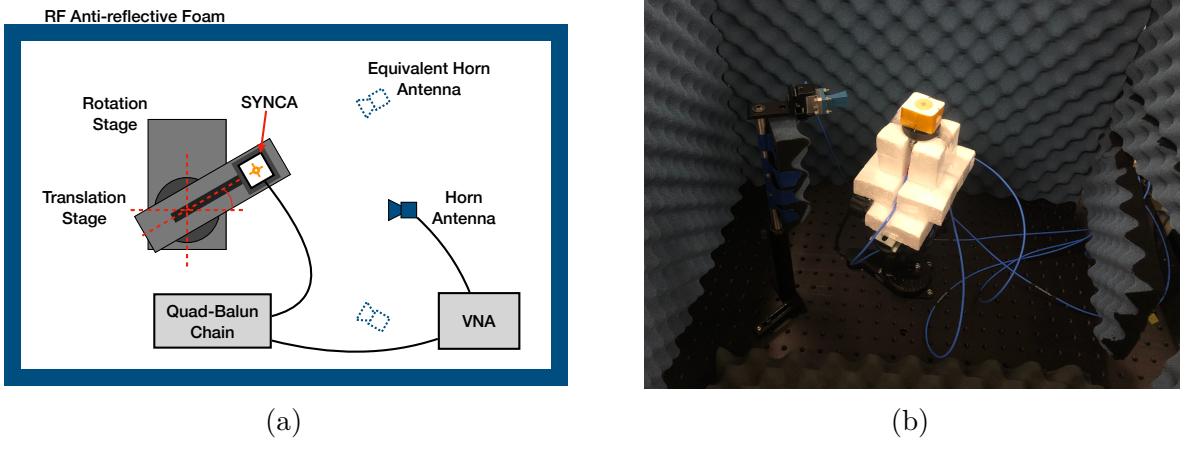


Figure 5.19. A schematic of the VNA characterization measurements (a). This setup allows for antenna gain and phase measurements across a full 360° of azimuthal angles using a motorized rotation stage and control of the radial position of the SYNCA using a translation stage. A photo of the setup in the lab is shown in (b).

4223 circumference. This measurement scheme is equivalent to measuring the fields generated
 4224 by the SYNCA using a full circular array of probe antennas.

4225 The antenna measurement space is surrounded by RF anti-reflective foam to isolate
 4226 the measurements from the lab environment (see Figure 5.19b) and remaining reflections
 4227 are removed using the VNA’s time-gating feature. The SYNCA is affixed to the stages
 4228 by a custom RF transparent mount made of polystyrene foam. The coaxial cables deliver
 4229 the antenna feed signals generated by the VNA to the SYNCA while still allowing
 4230 unrestricted rotation. The horn antenna probe is nominally positioned in the plane
 4231 formed by the antenna PCB ($\theta = 90^\circ$ or $z = 0$ mm) at a distance of 10 cm from the
 4232 SYNCA, to match the expected position of the antenna array relative to the SYNCA in
 4233 the antenna array test stand. The horn antenna can be manually raised or lowered to
 4234 different relative vertical positions to characterize the radiation pattern at different polar
 4235 angles.

4236 Several 360° scans were performed with probe vertical offsets of -10.0 mm, -5.0 mm,
 4237 0.0 mm, 5.0 mm, and 10.0 mm relative to the antenna PCB plane. These probe offsets
 4238 cover a 2 cm wide vertical region centered on the SYNCA PCB, approximately equal to
 4239 ± 6 degrees of polar angle. The measurements show that the SYNCA is generating fields
 4240 with nearly isotropic magnitude across the probed region. The standard deviation of the
 4241 electric field magnitude measured around the antenna circumference is approximately
 4242 2.9 dB for a typical rotational scan. The presence of a significant pattern null is noted

4243 near 45° (see Figure 5.20), which we attribute to small imperfections in the antenna
 4244 PCB that could be introduced from the hand soldered terminations connecting the coax
 4245 cables to the antenna. There is no significant difference in the radiation pattern when
 4246 measured across the 2 cm vertical range. The measured relative phases closely follow
 4247 the expectation for an electron, being linear with the measurement rotation angle and
 4248 forming the expected spiral pattern. Other than the small phase imperfections there is
 4249 a slight sinusoidal bias to the phase data, which we determined is the result of a small
 4250 ($\lesssim 1$ mm) offset of the antenna's phase center from the rotation axis of the automated
 4251 stages.

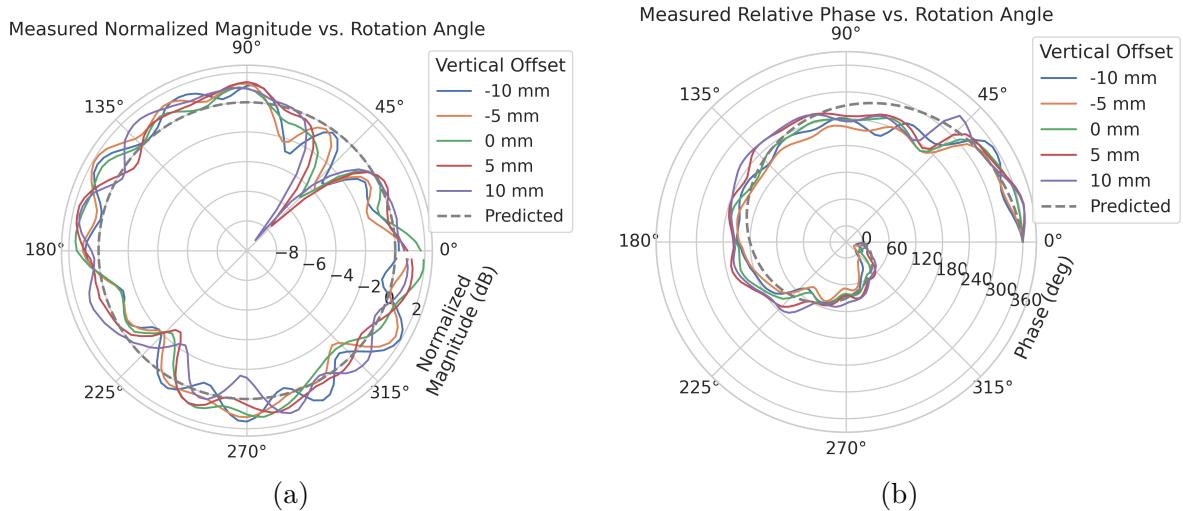


Figure 5.20. Linear interpolations of the measured electric field magnitude (a) and phase (b). The data was acquired using a VNA at 120 points spaced by 3 degrees from 0 to 357 degrees of azimuthal angle. The different color lines indicate the vertical offset of the horn antenna relative to the SYNCA PCB and the dashed line shows the expected shape from electron simulations. No significant difference in the antenna pattern is observed for the measured vertical offsets.

4252 The characterization measurements confirm the simulated performance of the SYNCA.
 4253 As expected the fields generated by the antenna are nearly isotropic in magnitude, ϕ -
 4254 polarized, and are linearly out of phase around the circumference of the antenna as
 4255 predicted for cyclotron radiation in Section 5.3.2. Small imperfections in the magnitude
 4256 and phase of the antenna are expected, particularly at the antenna's high operating
 4257 frequency of 26 GHz where small geometric changes can have significant impacts on
 4258 electrical properties. However, calibration through careful characterization measurements
 4259 can be used to remove the majority of these pattern imperfections, including the relatively
 4260 large pattern null near 45° , which will allow for the usage of the SYNCA as a test source

for free-space CRES experiments utilizing antenna arrays. In the next section we use the VNA measurements obtained here as a calibration for signal reconstruction using digital beamforming.

5.3.5 Beamforming Measurements with the SYNCA

Digital beamforming is a standard technique for signal reconstruction using a phased array [93]. The SYNCA, since it exhibits the same cyclotron phases as a trapped electron, can be used to perform simulated CRES digital beamforming reconstruction experiments on the bench-top without the need for the magnet, cryogenics, and vacuum systems required by a full CRES experiment. The fields received by the individual elements of the antenna array will have phases dependent on the spatial position of the source relative to the antennas. Therefore, a simple summation of the received signals will fail to reconstruct the signal due to destructive interference between the individual channels in the array. However, applying a phase shift associated with the source's spatial position removes phase differences and results in a constructive summation of the channel signals (see Figure 5.21). We can summarize the digital beamforming operation succinctly using the following equation

$$y[t_n] = \sum_{m=0}^{N-1} x_m[t_n] A_m e^{i\phi_m}, \quad (5.34)$$

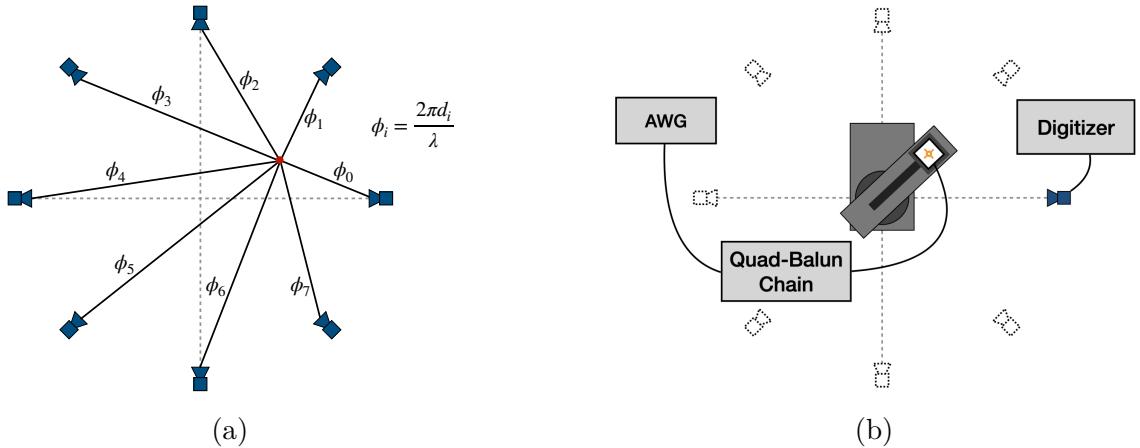


Figure 5.21. (a) A depiction of the relative phase differences for signals received by a circular antenna array from an isotropic source. The phases correspond to a unique spatial position. (b) A schematic of the setup used to perform digital beamforming.

4277 where $y[t_n]$ represents the summed array signal at time t_n , $x_m[t_n]$ is the signal received
 4278 by channel m at time t_n , ϕ_m is the phase shift applied to the signal received at channel
 4279 m , and A_m is an amplitude weighting factor that accounts for the different signal power
 4280 received by individual channels. By changing the digital beamforming phases, the point
 4281 of constructive interference can be scanned across the sensitive region of the array to
 4282 search for the location of a radiating source, which is identified as the point of maximum
 4283 summed signal power above a specified threshold. The digital beamforming phases consist
 4284 of two components,

$$\phi_m = 2\pi d_m / \lambda + \theta_m, \quad (5.35)$$

4285 where d_m is the distance from the m -th array element to the source, and θ_m is the
 4286 relative angle between the source position and the m -th antenna. The first component is
 4287 the standard digital beamforming phase that corresponds to the spatial position of the
 4288 source, and the second component is the cyclotron phase that corresponds to the relative
 4289 azimuthal phase offset.

4290 With a small modification to the hardware used to characterize the SYNCA (see
 4291 Figure 5.19), we can perform a digital beamforming reconstruction of a synthetic CRES
 4292 event. By replacing the VNA with an arbitrary waveform generator (AWG), the SYNCA
 4293 can be used to generate cyclotron radiation with an arbitrary signal structure, which
 4294 can then be detected by digitizing the signals received by the horn antenna. Rotational
 4295 symmetry allows us to use the rotational stage of the positioning system to rotate the
 4296 SYNCA to recreate the signals that would have been received by a complete circular
 4297 array of antennas.

4298 Using this setup, signals from a 60 channel circular array of equally spaced horn
 4299 antennas were generated with the SYNCA positioned 10 mm off the central array axis,
 4300 reconstructed using digital beamforming, and compared to Locust simulation (see Figure
 4301 5.22). When the cyclotron spiral phases are not used, which is equivalent to setting θ_m
 4302 in Equation 5.35 to zero, the SYNCA's position is reconstructed as a relatively faint ring
 4303 as predicted by simulation. However, when the appropriate cyclotron phases are used
 4304 during the beamforming procedure, both the simulated electron and the SYNCA appear
 4305 as a single peak of high relative power corresponding to the source position. Therefore,
 4306 we observe good agreement between the simulated and SYNCA reconstructions. While it
 4307 may seem that for the case with no cyclotron phase corrections the ring reconstructs the
 4308 position of the electron as effectively as beamforming with the cyclotron phase corrections,
 4309 it is important to note that the simulations and measurements were generated without a
 4310 realistic level of thermal noise. The larger maxima region and lower signal power, which

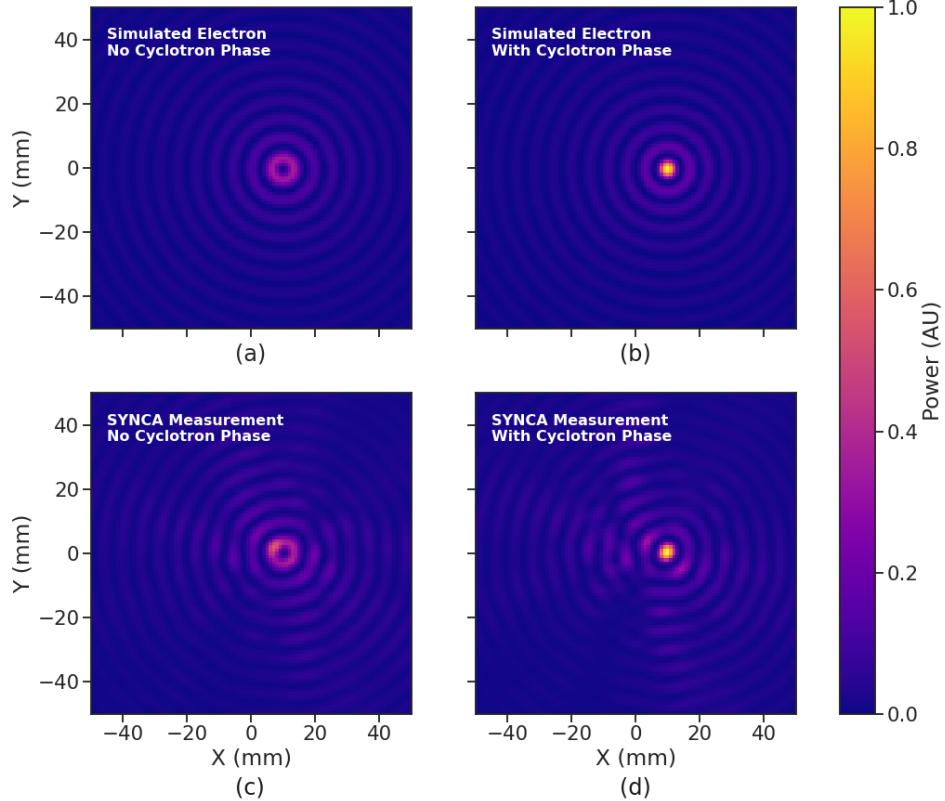


Figure 5.22. Digital beamforming maps generated using a simulated 60 channel array and electron simulated using the Locust package. (a) and (b) show the beamforming maps for simulated electrons without the cyclotron spiral phases and with the cyclotron spiral phases respectively. (c) and (d) show the beamforming maps produced from SYNCA measurements. We observe good agreement between simulated electrons and the SYNCA measurements.

4311 occurs without the cyclotron phase corrections, significantly reduce the probability of
 4312 detecting an electron in a realistic noise background.

4313 To bound the beamforming capabilities of the synthetic array of horn antennas, we
 4314 performed a series of beamforming reconstructions where the SYNCA was progressively
 4315 moved off the central axis of the array (see Figure 5.23). To extract an estimate of the
 4316 position of the SYNCA using the digital beamforming image we apply a 2-dimensional
 4317 (2D) Gaussian fit to the image data and extract the estimated centroid value. We find
 4318 that the synthetic horn antenna array reconstructs the position of the SYNCA with a
 4319 1σ -error of 0.3 mm with no apparent trend across the 30 mm measurement range. This
 4320 reconstruction error is an order of magnitude larger than mean fit position uncertainty
 4321 of 0.02 mm indicating that systematic effects related to the SYNCA positioning system
 4322 could be contributing additional uncertainty to the measurements. Note that the current

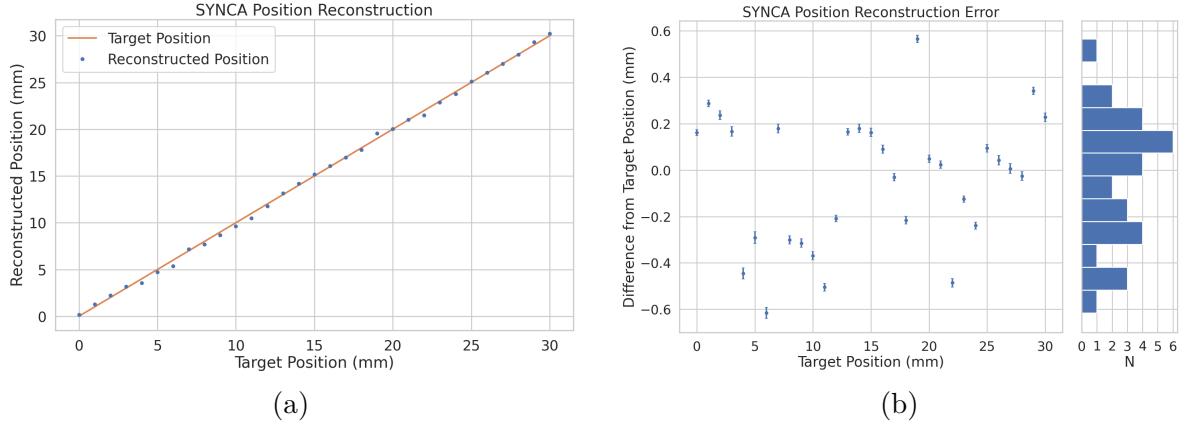


Figure 5.23. A plot of the SYNCA’s reconstructed position using the synthesized horn-antenna array and digital beamforming. (a) Shows the reconstructed position of the SYNCA compared with the target position indicated by the positioning system readout. (b) Shows the reconstruction error, which is the difference between the target and reconstructed positions. The error bars in (b) are the uncertainty in the mean position of the 2D Gaussian used to fit the digital beamforming reconstruction peak obtained from the fit covariance matrix. The mean fit position uncertainty of 0.02 mm is an order of magnitude smaller than the typical reconstruction error of 0.3 mm obtained by calculating the standard deviation of the difference between the reconstructed and target position.

4323 mean reconstruction error of 0.3 mm is a factor of 20 smaller than the full width at half
 4324 maximum of the digital beamforming peak (6 mm), which could be interpreted as a naive
 4325 estimate of the position reconstruction performance of this technique. Because these
 4326 measurements are intended as a proof-of-principle demonstration, we do not investigate
 4327 potential sources of systematic errors further; however, we expect that a similar and
 4328 more thorough investigation will be performed using the Project 8 antenna array test
 4329 stand, where typical reconstruction errors can be used to estimate the energy resolution
 4330 limits of antenna array designs.

4331 5.3.6 Conclusions

4332 In this paper we have introduced the SYNCA, which is a novel antenna design that
 4333 emits radiation that mimics the unique properties of the cyclotron radiation generated by
 4334 charged particles moving in a magnetic field. The characterization measurements of the
 4335 SYNCA validated the simulated performance of the PCB crossed-dipole antenna design.
 4336 Additionally, the SYNCA was used to estimate the position reconstruction capabilities
 4337 of a synthesized array of horn antennas and experimentally reproduced the simulated
 4338 digital beamforming reconstruction of electrons.

4339 While the SYNCA performs well, there exist discrepancies in the phase and magnitude
4340 of the radiation pattern compared to the simulated SYNCA design that are related to
4341 the small geometric differences in the soldered connections. Future design iterations that
4342 replace the soldered connections with a fully surface mount design could improve the
4343 radiation pattern at the cost of some complexity and expense. Furthermore, improving
4344 the design of the antenna PCB and mounting system would allow the antenna to be
4345 inserted into a cryogenic and vacuum environment where in-situ antenna measurement
4346 calibrations could be performed.

4347 The discrepancies in the radiation pattern and phases exhibited by the as-built
4348 SYNCA should not greatly impact its performance as a calibration probe. Both magni-
4349 tude and phase variations can be accounted by applying the SYNCA characterization
4350 measurements as a calibration to the data collected by the antenna array test stand. The
4351 separate calibration of the SYNCA radiation does not impact the primary goals for the
4352 antenna array test stand which are array calibration and signal reconstruction algorithm
4353 performance characterization, because it can be performed with standard reference horn
4354 antennas with well understood characteristics.

4355 The SYNCA antenna technology advances the CRES technique by providing a
4356 mechanism to characterize free-space antenna arrays for CRES measurements without
4357 the need for a magnet and cryogenics system, which would be required for calibration
4358 using electron sources. Both the Project 8 collaboration as well as future collaborations
4359 which are developing antenna array based CRES experiments can make use of SYNCA
4360 antennas as an important component of their calibration and commissioning phases.

4361 **5.4 SYNCA Development Discussion**

4362 A crossed-dipole antenna (see Figure 5.24) was identified early on as a candidate SYNCA
4363 design. The crossed-dipole is a circularly polarized antenna, consequently, the electric
4364 fields measured in the plane of the dipole antenna exhibit the same relative phase offsets
4365 as a 90° electron in a magnetic trap. This is explained in greater detail in Section 5.3.
4366 These phase offsets were measured with the first rudimentary crossed-dipole prototype
4367 manufactured from coaxial cables with the insulation and shield stripped away.

4368 Because the SYNCA is ultimately a calibration tool, it is desireable that the antenna
4369 have a well-characterized and robust antenna pattern. Therefore, manufacturing a
4370 SYNCA using the stripped wire method shown in Figure 5.24 is infeasible. Studies of
4371 crossed-dipole antennas manufactured out of printed circuit boards were performed using

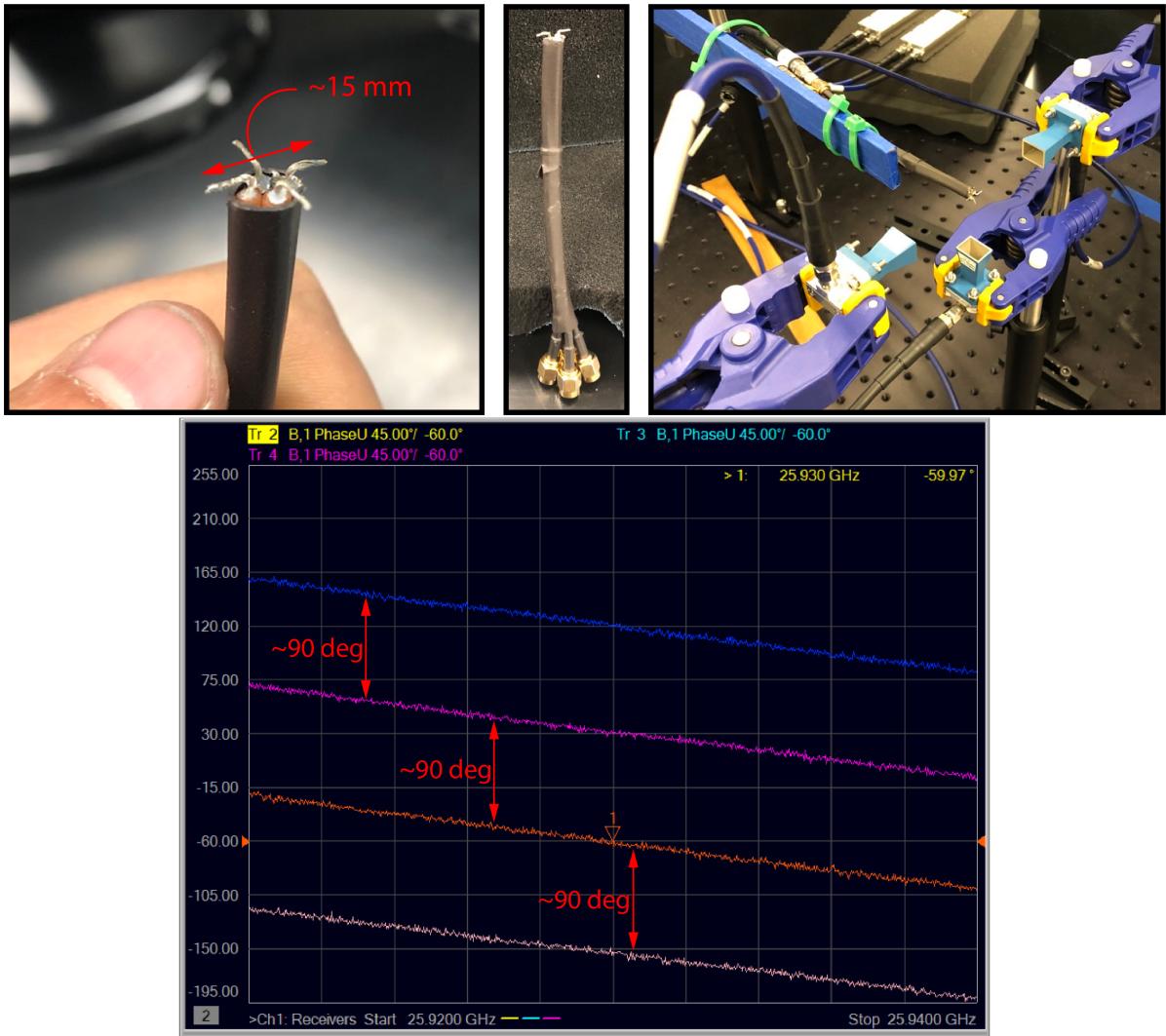


Figure 5.24. Images of an early prototype crossed-dipole antenna manufactured by hand and the first measurement setup. The antenna was constructed by hand using four stripped coaxial cables. The antenna was connected to one port of the VNA, and the remaining three ports on the VNA were connected to horn antenna arranged with 90 deg offsets around the crossed-dipole. The measured unwrapped S-parameter phases exhibit the desired relative phase behavior for a SYNCA. These early measurements were the first laboratory proof-of-principle for the crossed-dipole SYNCA.

4372 HFSS to identify an antenna design that imitated an electron, while being more robust
 4373 and simpler to manufacture (see Figure 5.25).

4374 Identifying a design that was robust, manufacturable, and matched the electric fields
 4375 of a trapped electron proved to be a non-trivial task. The primary factor driving the
 4376 difficulty was the high operating frequency of the antenna (26 GHz) combined with
 4377 the requirement that the antenna be electrically-small. An antenna that is electrically-

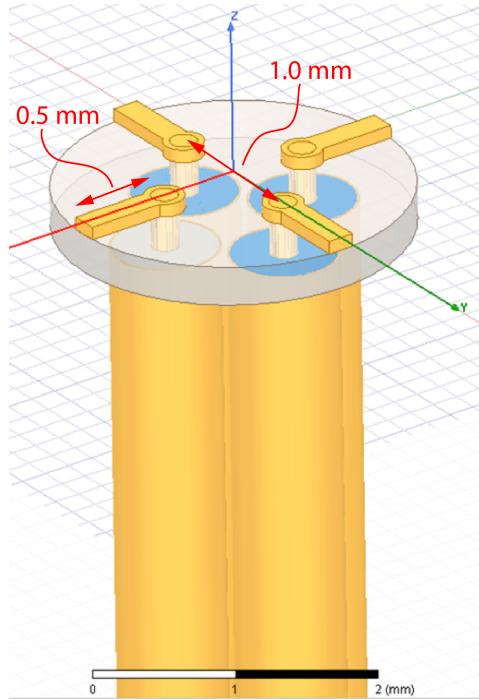


Figure 5.25. An early iteration of a crossed-dipole SYNCA antenna simulated in HFSS. The antenna is electrically small at 26 GHz, which requires dipole arms on the order of 1 mm long. This design is limited by the minimum achievable distance between the dipole arms caused by the available diameters of coaxial cables. The assumed termination scheme for the coaxial cables to the antenna is hand-soldering, which introduces random variation in the antenna pattern from the inevitable blobs of solder left on the surface of the PCB.

4378 small at 26 GHz has a largest dimension on the order of 1 mm, which poses significant
 4379 manufacturability challenges given the limited available budget for SYCNA fabrication.

4380 One of the key limitations with the small size requirements is the diameter of the
 4381 coaxial cables needed to feed the crossed-dipole antenna. The smallest commonly available
 4382 rigid coaxial cables available on the market have diameters of approximately 0.5 mm,
 4383 which limited the spacing between dipole arms to a minimum of about 1 mm. The
 4384 crossed-dipole antenna performs better as a SYNCA if the dipole arm separation is
 4385 significantly less than the operating wavelength. Therefore, the high operating frequency
 4386 ultimately limited how well the SYNCA could mimic an electron. If the desired cyclotron
 4387 frequency was lowered by an order of magnitude to approximately 3 GHz a significantly
 4388 higher quality SYNCA could be manufactured at lower cost.

4389 The decision to use coaxial transmission lines terminated on the antenna PCB with a
 4390 hand-soldered connection was driven primarily to limit the costs of SYNCA development
 4391 and contributed to the observable variations in the SYNCA's gain and phase patterns.

4392 A second iteration of the SYNCA design that minimized hand-soldering by using surface-
 4393 mount components could significantly reduce variations in the antenna pattern. The
 4394 major drawback in the development of a surface-mount SYNCA is the cost, and given the
 4395 transition to a cavity based design for Phase IV, such a design was never investigated.

4396 **5.5 FSCD Antenna Array Measurements with the SYNCA**

4397 **5.5.1 Introduction**

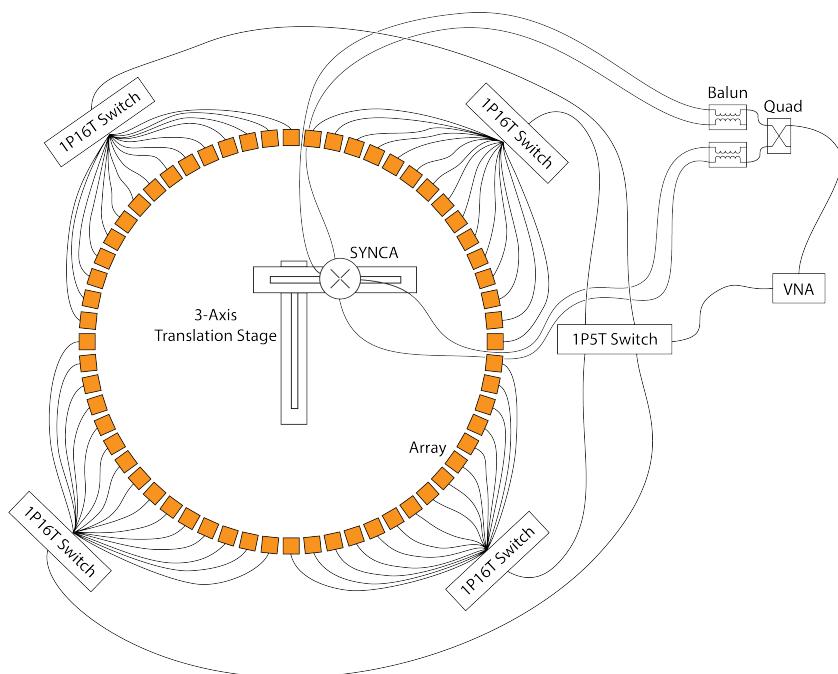


Figure 5.26. A diagram of the array measurement system used to test the prototype FSCD antenna array. A VNA is used as the primary measurement tool, which is connected to the array through a series of switches. The other port of the VNA connects to the SYNCA through the quad-balun chain used to provide the SYNCA feed signals. During measurements the SYNCA is positioned inside the center of the antenna array and translated to different radial and axial positions using a 3-axis manual translation stage setup.

4398 Using the SYNCA it is possible perform full-array measurements of prototype versions
 4399 of the FSCD antenna array with a realistic cyclotron radiation source (see Figure 5.26).
 4400 The goal is to compare the measured power received to FSCD simulations as a function
 4401 of the radial and axial position of the SYNCA source. These measurements are intended
 4402 to validate the antenna research and development by Project 8, which has been driven
 4403 primarily by simulations with Locust [61] and CRESana (see Section 4.2.3), and identify

4404 any discrepancies with these simulations tools. This knowledge will provide confidence
4405 in the simulations necessary for the analysis of the sensitivity of larger antenna array
4406 based CRES experiment designs to the neutrino mass.

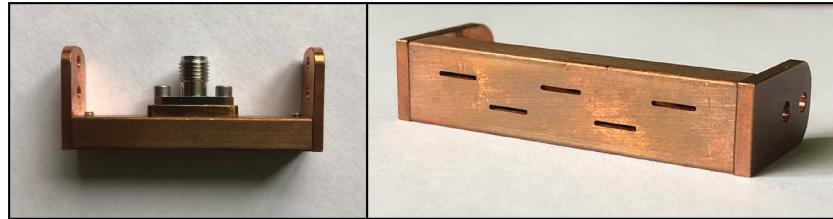
4407 As shown in Section 5.3, the SYNCA has some radiation pattern imperfections
4408 that complicate the comparison between measurement and simulation data. One way
4409 to disentangle the effects of these imperfections is to perform an additional set of
4410 measurements using a synthetic antenna array setup along with the SYNCA antenna.
4411 Since the synthetic array setup uses only a single array antenna, the data should be
4412 free of errors associated with individual antenna differences and multi-path interference,
4413 which are two error sources being tested with the full-array setup. By comparing the
4414 synthetic array data to the FSCD array data and to simulation data one can evaluate the
4415 significance of these effects relative to the errors introduced by SYNCA imperfections.

4416 **5.5.2 Measurement Setups**

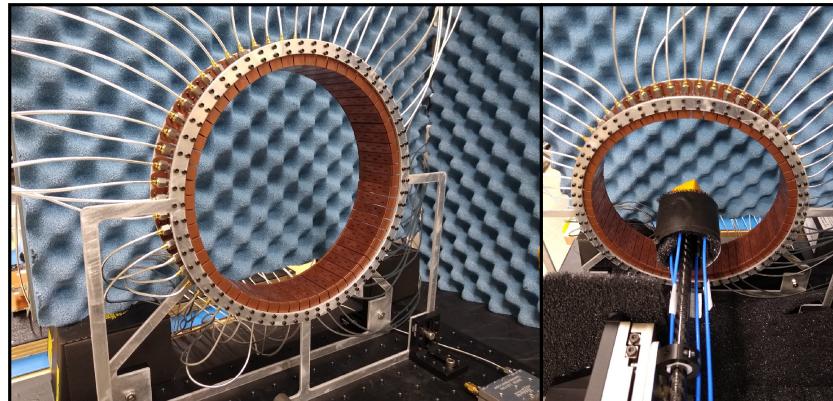
4417 **5.5.2.1 FSCD Array Setup**

4418 The antenna design that composes the array is the 5-slot waveguide antenna developed
4419 for the FSCD experiment (see Figure 5.27a). The antenna is 5 cm long and is constructed
4420 out of WR-34 waveguide with a 2.92 mm coax connector located at the center of the
4421 antenna. Copper flanges located on both ends of the antenna are used to mount the
4422 antenna in the array support structure. The antennas are supported by two circular steel
4423 brackets that can be bolted to both ends of the waveguide to construct the circular array
4424 (see Figure 5.27b). The antenna array consists of sixty identical waveguide antennas
4425 with a radius of 10 cm. The array is mounted perpendicular to an optical breadboard
4426 surface using a pair of the steel brackets, which provide sufficient space for the coaxial
4427 cable connections and allows for easy positioning of the SYNCA antenna. The SYNCA is
4428 mounted on the end of a carbon fiber rod attached to a set of manual translation stages,
4429 which are used to move the SYNCA antenna to different positions inside the array (see
4430 Figure 5.27c). The stages allow for independent motion in three different axes and can
4431 position the SYNCA at radial distances up to 5 cm from the center.

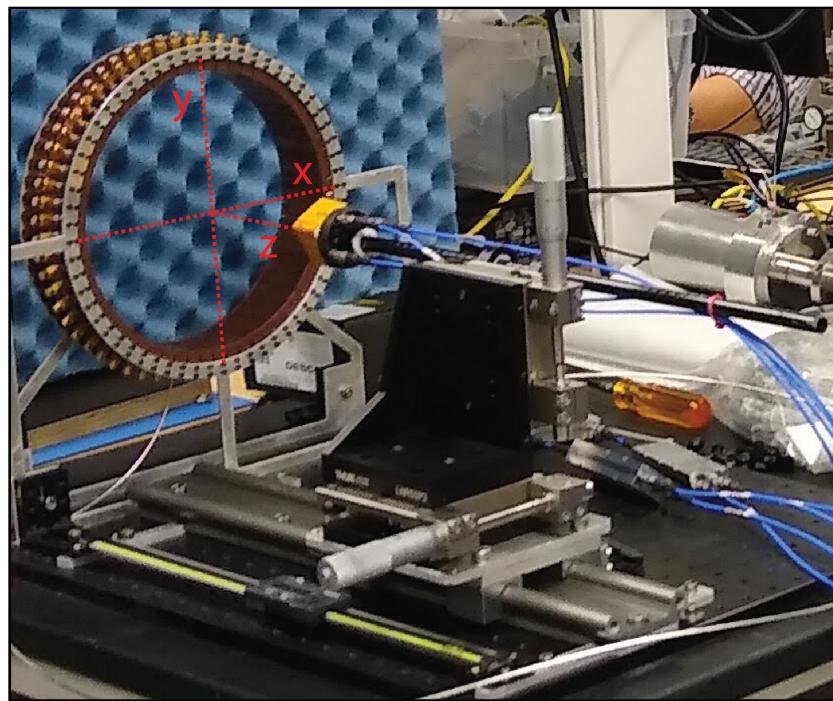
4432 Data acquisition is accomplished using a two-port VNA in combination with a series
4433 of microwave switches that allow the VNA to connect to each channel in the array . The
4434 first port of the VNA is connected to the quad-balun chain used to feed the SYNCA (see
4435 Section 5.3), and the second port of the VNA connects to a 1P5T microwave switch. The
4436 1P5T switch is connected to four separate 1P16T switch boards that connect directly



(a)



(b)



(c)

Figure 5.27. Photos of the prototype FSCD antenna (a), the FSCD array and SYNCA (b), and the translation stages and coordinate system used to position the SYNCA (c).

4437 to the array. The data acquisition is controlled by a python script running on a lab
4438 computer, which is connected to the VNA and an Arduino board programmed to control
4439 the microwave switches. The script uses the switches to iteratively connect each of the
4440 antennas in the array to the VNA. The VNA is configured to load a specific calibration
4441 file for each antenna channel and performs the measurements of all available S-parameters.
4442 The separate calibration files is an attempt to remove phase and magnitude errors caused
4443 by different propagation through the RF switches. Array measurements were performed
4444 for the set of SYNCA positions consisting of radial (x-axis) positions from 0 to 50 mm
4445 in 5 mm steps and axial (z-axis) positions from 0 to 50 mm in 5 mm steps resulting in
4446 121 array measurements. At each SYNCA position the two-port S-parameter matrix
4447 is measured using a linear frequency sweep from 25.1 to 26.1 GHz with 101 discrete
4448 frequencies.

4449 5.5.2.2 Synthetic Array Setup

4450 A photograph of the setup used to perform the synthetic array measurements is shown
4451 in Figure 5.28. A difference between this setup and the FSCD array setup is that the
4452 synthetic array measurements were performed with a waveform generator and digitizer
4453 instead of a VNA. The electronics configuration is identical to the diagram in Figure
4454 5.7b. Despite the differences, one is still able to compare the measured phases of the
4455 synthetic array and the relative magnitude of the power, since the digitized signal power
4456 is directly proportional to S21.

4457 The arbitrary waveform generator in the setup is configured to produce a 64 MHz
4458 sine wave signal that is up-converted to 25.864 GHz using a mixer and the VNA source.
4459 This signal is passed through a bandpass filter and fed to the SYNCA quad-balun chain.
4460 A single FSCD antenna is positioned 10 cm from the SYNCA and aligned vertically so
4461 that the center of the 5-slot waveguide is in the plane of the SYNCA PCB (see Figure
4462 5.28). This position corresponds to $z = 0$ in Figure 5.27c. The SYNCA is rotated
4463 in three degree steps to synthesize an antenna array with 120 channels. This channel
4464 count is more than could physically fit in a 10 cm radius array, but there is no cost to
4465 over-sampling. The signals from the FSCD antenna are down-converted using the second
4466 mixer connected to the VNA source before being digitized at 250 MHz and saved to
4467 disk. Several synthetic array measurement scans were performed by using the linear
4468 translation stage to change the radial position of the SYNCA. In total eight scans were
4469 taken from 0 to 35 mm using a radial position step size of 5 mm.

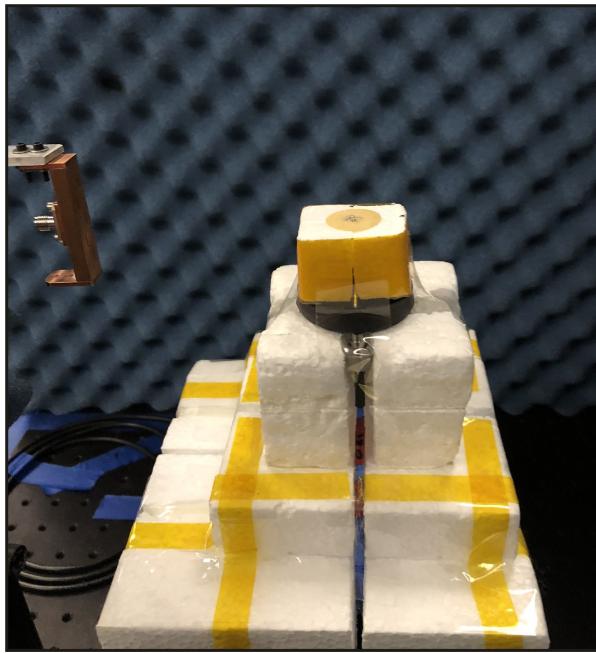


Figure 5.28. A photo of the FSCD antenna and the SYNCA in the synthetic array measurement setup at Penn State.

4470 5.5.3 Simulations, Analysis, and Results

4471 The Locust and CRESana simulation packages utilize the antenna transfer functions
4472 to calculate the power that would be received by each antenna from a CRES electron.
4473 The equivalent quantity in the measurement setup is the S21 matrix element, which
4474 indicates the ratio of the power received by an antenna in the array to the amount of
4475 power delivered to the SYNCA. Therefore, the analysis focuses on comparing the relative
4476 magnitudes and phase of the S21 parameters measured by the VNA as a function of the
4477 array channel and the SYNCA position. Additionally, a beamforming reconstruction
4478 using the S21 data is done to evaluate how the summed power and beamforming images
4479 change as a function of the position of the SYNCA.

4480 5.5.3.1 Simulations

4481 Simulations for the FSCD array measurements were performed using CRESana, which
4482 performs analytical calculations of the EM-fields produced by an electron at the position
4483 of the antennas. At each sampled time CRESana computes the electric field vector at the
4484 antenna positions, which is projected onto the antenna polarization axis to obtain the
4485 co-polar electric field. The magnitude of the co-polar electric field is then multiplied by
4486 a flat antenna transfer function to calculate the corresponding voltage signal. CRESana

4487 simulations exploit the flat transfer functions of the FSCD antennas, which allows the
 4488 electric field to be multiplied by the antenna transfer function rather than performing
 4489 the full FIR calculation. These calculations produce a voltage time-series for each of the
 4490 antennas in the array that can be compared to the laboratory measurements.

4491 CRESana was configured to simulate a 90° electron in a constant background magnetic
 4492 field of ≈ 0.958 T with a kinetic energy of 18.6 keV. These parameters were chosen
 4493 in order to mimic a CRES event near the tritium beta-decay spectrum endpoint in
 4494 the FSCD experiment. The constant background magnetic field guarantees that the
 4495 guiding center of the electron is stationary across the duration of the simulation which is
 4496 consistent with the SYNCA in the laboratory measurements. Simulations were performed
 4497 with the electron's guiding center at radial positions from 0 to 45 mm in steps of 1 mm
 4498 and axial positions from 0 to 30 mm in steps of 1 mm. The simulations generated time
 4499 series consisting of 8192 samples at 200 MHz for the sixty channel FSCD antenna array
 4500 geometry.

4501 5.5.3.2 Phase Analysis

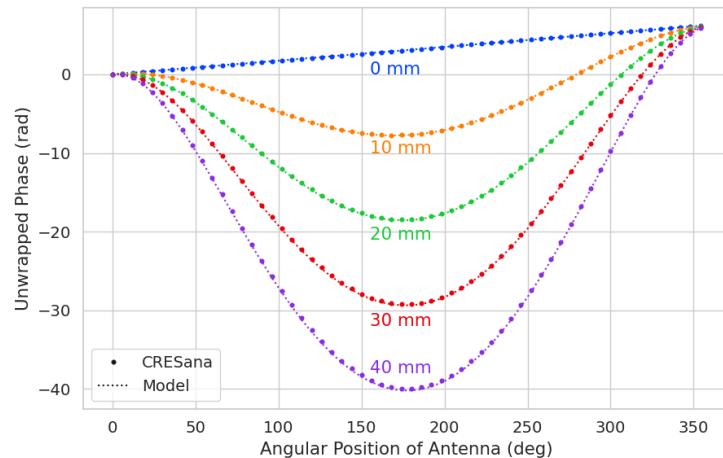


Figure 5.29. The unwrapped phases of signals received by the FSCD antenna array from an electron with a 90° pitch angle located in the plane of the antenna array. The data points indicated the phases extracted from simulation and the dashed lines show the model predictions.

4502 Correct modeling of the signal phases is fundamental to reconstruction for both
 4503 beamforming and matched filter approaches. The beamforming reconstruction relies on
 4504 a signal phase model developed from Locust simulations, which allows one to predict the
 4505 relative signal phases for a specific magnetic trap and electron position. The equation

4506 for the model is

$$\phi_{ij}(t) = \frac{2\pi d_{ij}(t)}{\lambda} + \theta_{ij}(t), \quad (5.36)$$

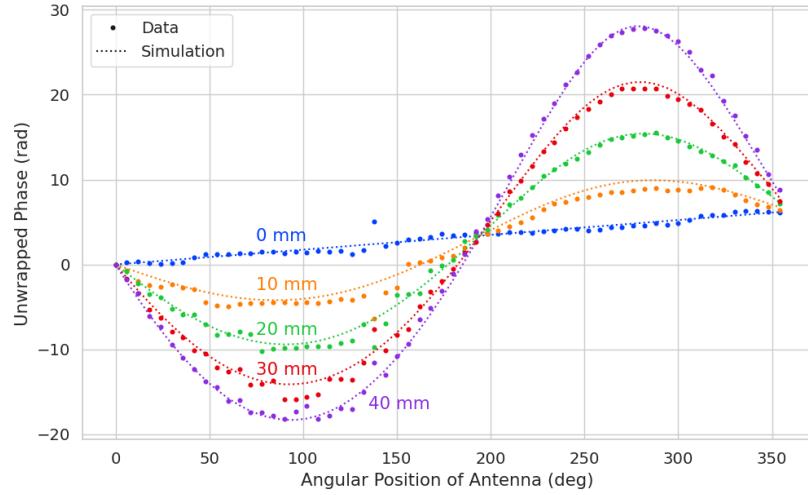
4507 where $d_{ij}(t)$ is distance between the assumed electron position and the antenna position,
4508 and $\theta_{ij}(t)$ is the angular separation between the electron and antenna positions. For
4509 details on the components of the phase model see Section 5.3.2. In Figure 5.29 I compare
4510 the phases predicted by Equation 5.36 to phases extracted from CRESana simulations of
4511 an electron located in the plane of the antenna array at a series of radial positions. One
4512 observes excellent agreement between the model and simulation.

4513 The measured signal phases from the FSCD array and synthetic array are shown
4514 in Figures 5.30a and 5.30b compared to the signal phase model. The axial position of
4515 the SYNCA in both plots is $z = 0$ mm, such that the plane of the PCB is aligned with
4516 the center of the FSCD antenna. The data shown in Figure 5.30a corresponds to the
4517 S-parameters measured at 25.80 GHz which is the frequency closest to the one used in
4518 the synthetic array setup. The different slope and sinusoidal phases exhibited by Figure
4519 5.30a and 5.30b reflects differences in the coordinate system for each setup. In general,
4520 the phase model predicts the large scale features of the phases well, but there are some
4521 small scale deviations or errors from the phase model that do not appear to be present
4522 in simulation.

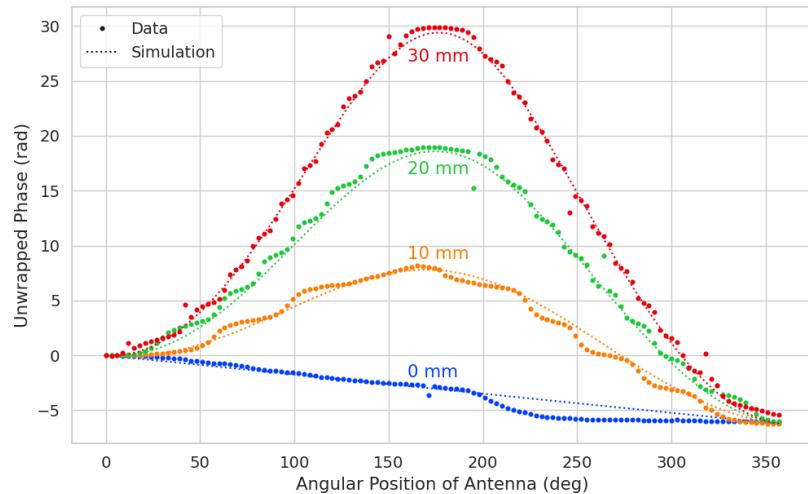
4523 A comparison of the phase errors, which are the difference between measurement and
4524 model is shown in Figure 5.31. The FSCD array data is referred to as the JUGAAD
4525 data in the plot legend, which is an alternative name for the FSCD array setup.

4526 The phase error at $R = 0$ in Figure 5.31 forms a smooth curve, with the exception of
4527 an outlier data point caused by a bug in the data acquisition script. One can attribute
4528 the observed phase error at this position to imperfections in the antenna pattern of the
4529 SYNCA. As the SYNCA is moved away from $R = 0$ mm one observes that the phase
4530 error exhibits oscillations whose frequency increases as a function of the radial position
4531 of the SYNCA. These oscillations have the appearance of a diffraction pattern, which
4532 is particularly clear for the radii ≥ 15 mm, due to the bilateral symmetry of the phase
4533 error peaks around 180° .

4534 One can observe a higher average variance in the phase errors measured for the FSCD
4535 array compared to the synthetic array. This is best seen by comparing the curves at
4536 $R \leq 15$ mm where the smooth synthetic array curves are distinct from the relatively
4537 noisy FSCD array errors. The extra noise in the FSCD array is most likely caused by
4538 differences in the radiation patterns of the antennas that make up the array as well as
4539 differences in the transmission lines through the switch network that introduce additional



(a)



(b)

Figure 5.30. Plots of the measured unwrapped phases from the FSCD array (a) and the synthetic array (b) compared to the model predictions for a series of radial positions. The different phases of the sinusoidal phase oscillations in the two plots reflects differences in the coordinate systems of the measurements.

phase errors into the measurement. Since the synthetic array measurements use only a single antenna, these extra error terms are not present, which explains the relatively smoother phase error curves. Despite the extra phase errors in the FSCD array, it is still possible to observe a similar phase error oscillation effect as the SYNCA is moved away from $R = 0$ mm.

The diffraction pattern exhibited by the phase error oscillations is more easily observed

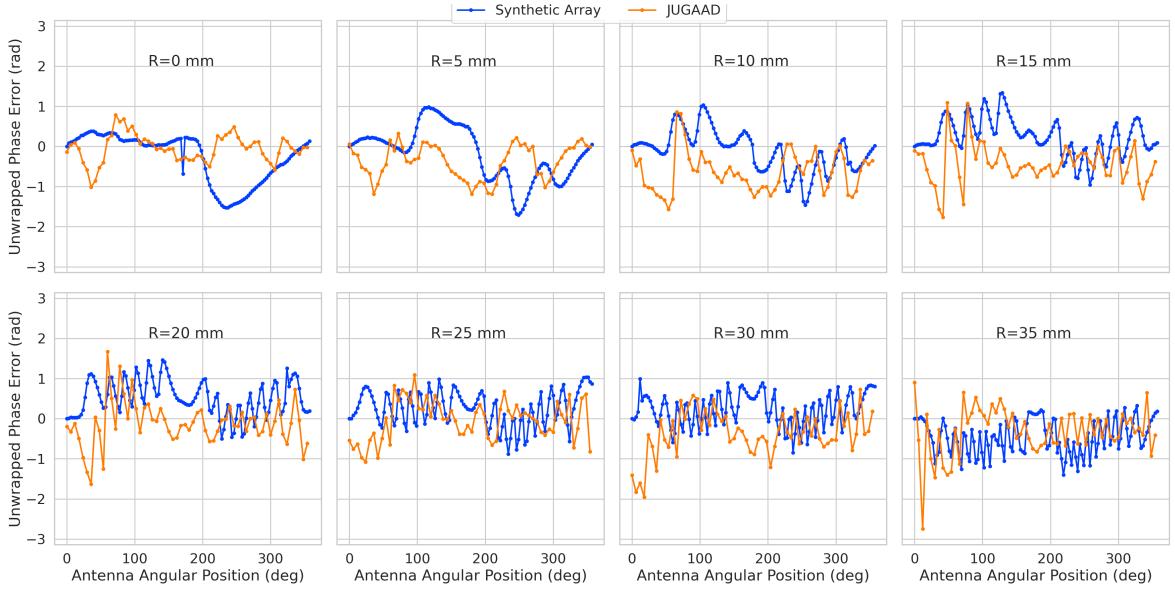
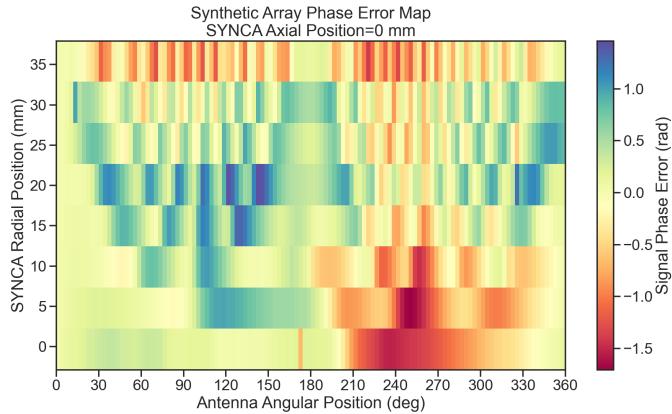


Figure 5.31. The phase errors between the measurement and model for the synthetic array (blue) and the FSCD array (orange) for a series of radial positions. The label JUGAAD refers to an alternative name for the FSCD array setup. As the SYNCA is translated off-axis phase errors with progressively higher oscillation frequency enter into the measurements.

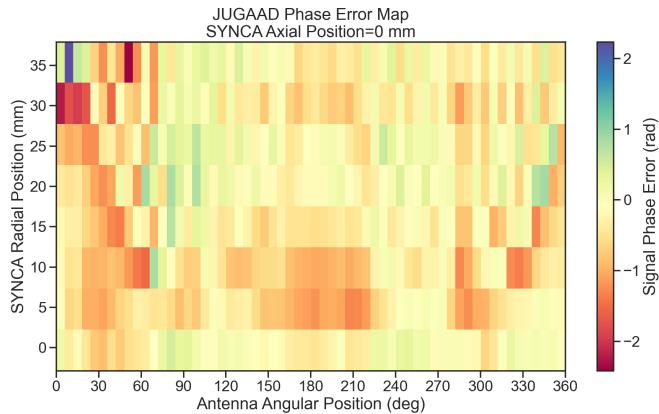
4546 by plotting the phase errors in a two-dimensional map, which is done in Figures 5.32a and
 4547 5.32b. For the synthetic array data ones observes a relatively clear diffraction pattern
 4548 that emerges as the SYNCA is moved radially. The bilateral symmetry of the diffraction
 4549 patterns is due to the bilateral symmetry of the circular synthetic array around the
 4550 translation axis of the SYNCA. A similar pattern is also visible in the FSCD array data,
 4551 although, it is obscured by the additional phase error that results from the multi-channel
 4552 array.

4553 The physical origin of the phase error diffraction pattern is attributed to interference
 4554 effects arising from path-length differences between the individual slots in the FSCD
 4555 antenna and the SYNCA transmitter. Since measurements are being performed in the
 4556 radiative near-field of the FSCD antenna, the path length differences between the slots
 4557 introduces a significant change in the summation of the signals that occurs inside the
 4558 waveguide, which causes the radiation pattern of the antenna to change as a function of
 4559 distance. Therefore, when the SYNCA is positioned off-axis the different path-lengths
 4560 from the SYNCA to each antenna results in different radiation patterns leading to the
 4561 observed diffraction pattern.

4562 This near-field effect is not present in simulations, because in order to simplify the
 4563 calculations it is assumed that the far-field approximation can be applied to the FSCD



(a)



(b)

Figure 5.32. Two dimensional plots of the phase errors for the synthetic array (a) and the FSCD (JUGAAD) array (b). In both plots there is evidence of a similar diffraction pattern with bilateral symmetry, but the FSCD array measurements have an additional phase error contribution from the different antennas and paths through the switch network.

antennas. This means that the radiation pattern and antenna transfer functions are independent of the distance between the transmitter and the receiving antenna. In principle, the near-field effects can be accounted for with a more detailed simulation of the FSCD antennas either in CRESana or Locust, which would result in an additional term in the beamforming phase model. However, this would increase the computational intensity of the simulation software. In the next section I briefly discuss the impact of these near-field effects on the measured magnitude of the signals.

4571 5.5.3.3 Magnitude Analysis

4572 Exactly as for the signal phase, one can use simulations to construct a model that
4573 describes the magnitude of the signals received by each channel in the antenna array.
4574 By examining the results of simulations or by analyzing the Liénard-Wiechert equation
4575 one can show that radiation pattern from a 90° pitch angle electron in a magnetic field
4576 is omni-directional. Therefore the relative magnitudes of the signals received by each
4577 channel will be determined by the free-space power loss, which is proportional to the
4578 inverse distance between the assumed electron position and the antenna.

4579 A consequence of this is that the signals produced in the array for electrons off the
4580 central axis will have larger amplitudes for the antennas closer to the electron compared
4581 to those which are further away. The amplitudes of the signals received by the array
from an electron located at a series of radial positions are shown in Figure 5.33.

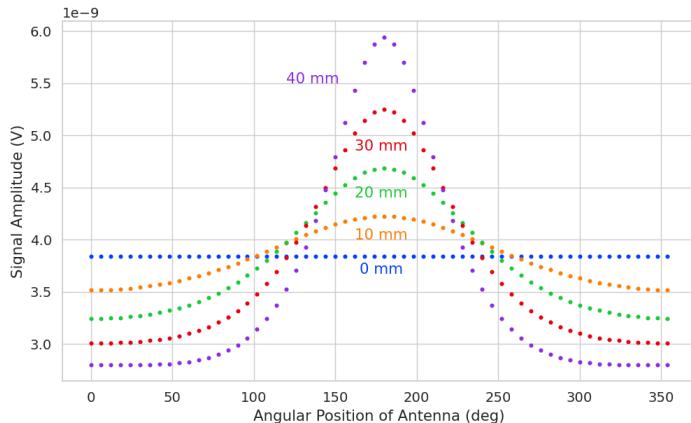


Figure 5.33. The amplitude of the signals from CRESana for the FSCD array from a 90° electron. As the electron is moved from $R = 0$ the signals begin to have unequal amplitudes depending on the distance from the electron to the antenna.

4582
4583 One expects to see a similar trend in the signal magnitudes in both the FSCD and
4584 synthetic arrays. The normalized signal magnitudes extracted from the full and synthetic
4585 array setups for a series of radial SYNCA positions are shown in Figure 5.34. The data
4586 corresponds to a SYNCA axial position of $z = 0$ mm and at a frequency 25.86 GHz. One
4587 complication is that the radiation pattern of the SYNCA is not perfectly omni-directional,
4588 which causes the measured magnitudes at $R = 0$ mm to diverge from the perfectly flat
4589 behavior exhibited by electrons.

4590 As the SYNCA is moved off-axis one observes a similar increase in the number of
4591 magnitude peaks in the synthetic array data that one would expect from a diffraction

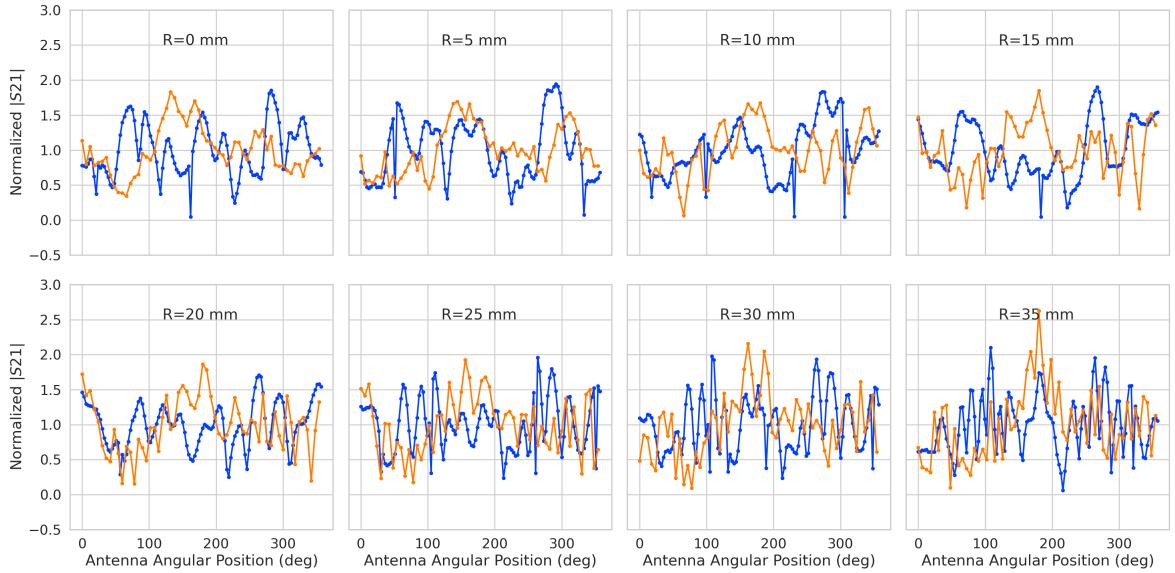
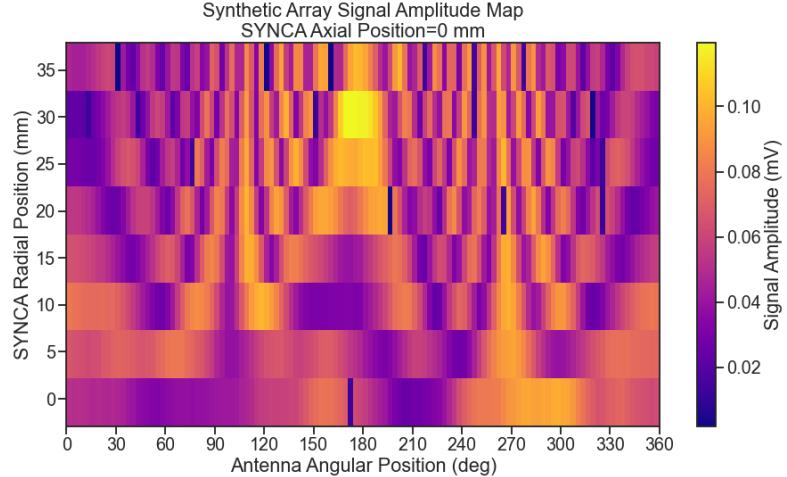


Figure 5.34. The normalized magnitudes of the S21 parameters measured in the FSCD (orange) and synthetic array (blue) setups. The dominant observed behavior as a function of radius is the increase in the number of magnitude peaks, which was noted in the phase error curves. There does not appear to be a strong change in the relative amplitude of a group of antennas as predicted by CRESana.

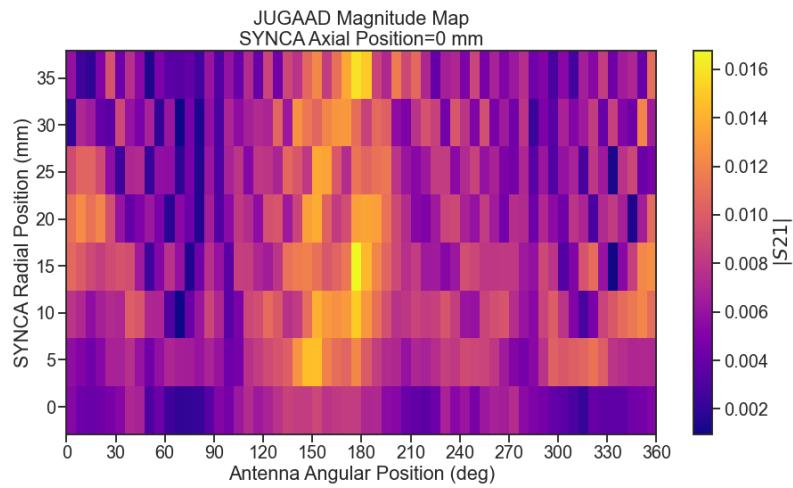
pattern, although this trend is not as stark compared to the phase data. Noticeably, there does not appear to be a set of channels with disproportionately larger amplitude at large R , which would be expected based on the trends from CRESana.

Comparing the magnitudes of the synthetic array to the FSCD array in Figure 5.34, one observes a similar amount of variability in the magnitudes at $R = 0$ mm, although there is potentially more small scale error in the magnitude curve caused by channel differences in the FSCD array. A similar trend is seen in the number of magnitude error peaks in the FSCD array data to the synthetic array data, which mirrors the diffraction effect observed in the phase data. The diffraction effect can be visualized more clearly by plotting a similar two-dimensional map of the magnitudes (see Figure 5.35).

The fact that one observes a similar diffraction pattern in the signal magnitudes as a function the SYNCA position reinforces the conclusions from the phase analysis that near-field effects are having a significant impact on the radiation pattern of the FSCD array. These near-field effects lead to changes in the magnitude and phase of the radiation pattern of the FSCD antenna as a function of distance. If left uncorrected these errors reduce detection efficiency by causing power loss in the beamforming or matched filter reconstruction due to phase mismatch. I explore the impact of these phase and



(a)



(b) The two-dimensional maps showing the diffractive pattern exhibited by the FSCD and synthetic array signal magnitudes.

Figure 5.35.

4609 magnitude errors on beamforming in the next section.

4610 5.5.3.4 Beamforming Characterization

4611 Errors in the signal magnitudes and phases lead to errors in signal reconstruction. For
 4612 example, a matched filter reconstruction requires accurate knowledge of the signals in
 4613 each channel to achieve optimal performance. Uncorrected errors leads to mismatches
 4614 between the template and signal, which reduces detection efficiency and introduces
 4615 uncertainty in the parameter estimation. In this section, I analyze the beamformed

4616 signal amplitude as a function of the position of the SYNCA to quantify the impact of
 4617 the phase and magnitude errors on signal reconstruction. Because of the imperfections
 4618 in the SYNCA source, it is inappropriate to directly compare the beamformed signal
 4619 amplitude of the FSCD array or synthetic array. Such a comparison would not allow
 4620 one to disentangle losses that occur because of the antenna array from those that occur
 4621 because of the source. Therefore, I focus on comparing the beamforming of the FSCD
 4622 array to the synthetic array.

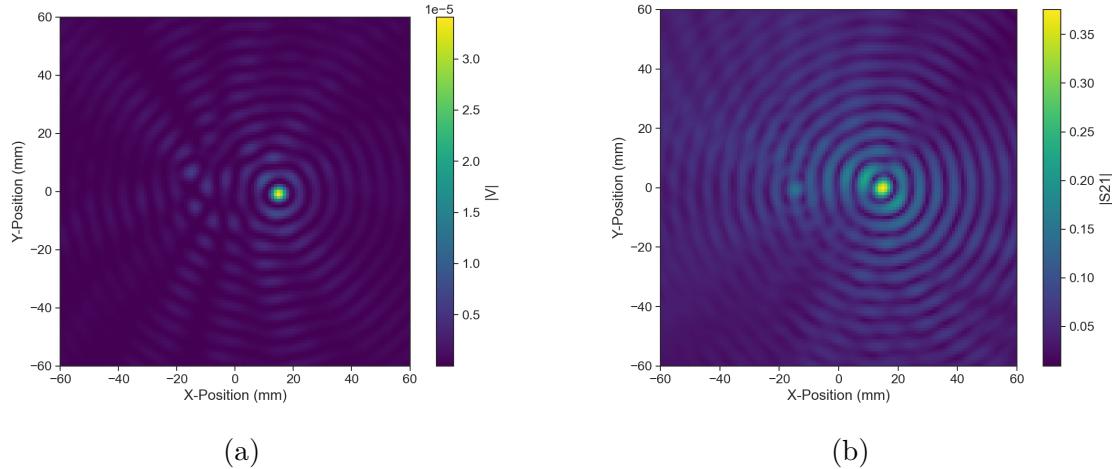


Figure 5.36. Beamforming images from the synthetic array (a) and FSCD array (b) setups with the SYNCA positioned 15 mm off the central axis. In both images, there is a clear maxima that corresponds to the true SYNCA position. However, in the FSCD array there is an additional faint peak located at the opposite position of the beamforming maximum. This additional peak is the mirror of the true peak and is the result of reflections between antennas in the FSCD array.

4623 The first method of comparison is to analyze the images generated by applying the
 4624 beamforming reconstruction specified in Section 4.3.1 to the FSCD and synthetic array
 4625 data (see Figure 5.36). The beamforming grid consisting of a square 121×121 grid
 4626 spanning a range of -60-mm to 60 mm in the x and y dimensions. The beamforming
 4627 images formed from the synthetic array produces a three-dimensional matrix where each
 4628 grid position contains a summed time series. A single beamforming image is formed from
 4629 this data matrix by taking the mean over the time dimension. In the case of the FSCD
 4630 array, the VNA generates frequency domain data such that each grid position contains a
 4631 summed frequency series produced by the VNA sweep. For this data a single image is
 4632 formed by averaging in the frequency domain.

4633 There is a clear difference between the synthetic and FSCD array beamforming images,
 4634 which is the additional faint beamforming maxima located directly opposite the maxima

4635 corresponding to the SYNCA position. The images in Figure 5.36 were generated with
 4636 data collected at a SYNCA radial position of 15 mm, which agrees well with the observed
 4637 beamforming maximum in both images. The faint beamforming peak is located directly
 4638 opposite of the true beamforming maximum similar to a mirror image. Therefore, the
 4639 origin of this additional feature appears to be reflections between the two sides of the
 4640 circular antenna array that are not present for the synthetic array since only a single
 4641 physical antenna is used.

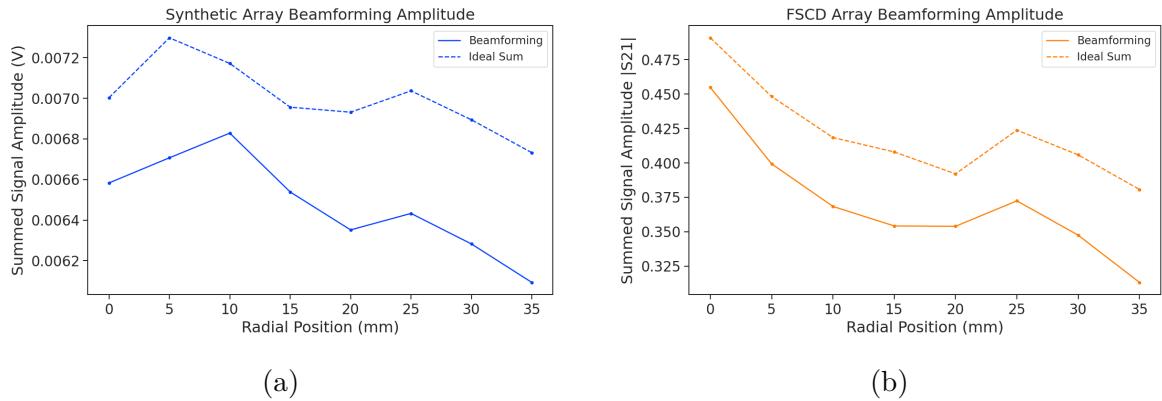


Figure 5.37. A comparison of the maximum signal amplitude obtained by beamforming to the signal amplitude obtained with an ideal summation as a function of the radial position of the SYNCA. The amplitudes for the synthetic array are shown in (a) and the FSCD array are shown in (b). In both setups, the signal amplitudes obtained from beamforming are smaller than the signal amplitude that could be attained with the ideal summation without phase mismatch.

4642 From the beamforming images the maximum amplitude is extracted, which can be
 4643 plotted as a function of the radial position of the SYNCA (see Figure 5.37). The phase
 4644 errors observed in the FSCD and synthetic arrays leads to power loss at the beamforming
 4645 stage due to phase mismatches between the signals at different channels. This power loss
 4646 can be quantified by comparing the signal amplitude obtained from beamforming to the
 4647 amplitude which would be obtained from an ideal summation. The ideal summation is
 4648 performed by phase shifting each array channel to an identical phase and then summing.
 4649 The comparison between the beamforming and ideal sums is shown in Figure 5.37,
 4650 where it is seen that the synthetic and FSCD arrays experience power losses from the
 4651 beamforming summation.

4652 The beamforming power loss can be quantified using the ratio of the beamforming to
 4653 ideal signal amplitudes. Computing this ratio as a function of SYNCA radial position
 4654 radius for the FSCD and synthetic arrays, it is found that the FSCD array has a uniformly
 4655 smaller beamforming amplitude ratio, which means that the FSCD array has a larger

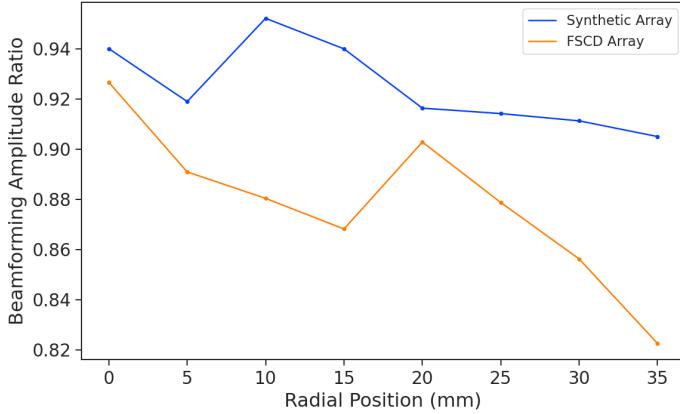


Figure 5.38. The ratio of the beamforming signal amplitude to the ideal signal amplitude for the FSCD and synthetic arrays. The FSCD array has a larger power loss from phase error compared to the synthetic array which indicates that calibration errors associated with the multiple channels as well as reflections are impacting the signal reconstruction.

beamforming power loss (see Figure 5.38). The primary contributions to the beamforming power loss in the synthetic array are phase errors from the SYNCA and phase errors from the FSCD antenna near-field. Both of these phase errors contribute to beamforming losses in the FSCD array, but there are clearly additional phase errors in the FSCD array measurements contributing to the smaller ratio. Two potential error sources include phase differences in the different antenna channels that could not be corrected by calibration as well as reflections between antennas in the array. The total effect of these additional phase errors is to reduce the beamforming amplitude ratio by about 5% from the beamforming ratio of the synthetic array. Therefore, it is estimated that if no effort is made to correct these phase errors in an FSCD-like experiment, then one would expect approximately a 10% total signal amplitude loss from a beamforming signal reconstruction.

5.5.4 Conclusions

The estimated power loss of a beamforming reconstruction obtained from this analysis provides valuable inputs to sensitivity calculations of a FSCD-like antenna array experiment to measure the neutrino mass, since it helps to bound systematic uncertainties from the antenna array and reconstruction pipeline. This power loss lowers the estimated detection efficiency of the experiment since some of the signal power is lost due to improper combining between channels and also increases the uncertainty in the electron's kinetic energy by contributing to errors in the estimation of the electron's cyclotron frequency.

⁴⁶⁷⁶ If these reconstruction losses prove unacceptable there are steps that can be taken
⁴⁶⁷⁷ to mitigate their effects. Some examples include the development of a more accurate
⁴⁶⁷⁸ antenna simulation approach that can reproduce the observed near-field interference
⁴⁶⁷⁹ patterns of the FSCD antennas and the implementation of a calibration approach that
⁴⁶⁸⁰ allows for the relative phase delays of the array to be measured without changing or
⁴⁶⁸¹ disconnecting the antenna array configuration.

Chapter 6

Development of Resonant Cavities for Large Volume CRES Measurements

6.1 Introduction

The cavity approach was originally an alternative CRES measurement technology under consideration by the Project 8 collaboration for the Phase IV experiment. After pursuing an antenna array based CRES demonstrator design for several years, the increasing costs and complexity of the antenna arrays led to a reconsideration of the baseline technology for the ultimate CRES experiment planned by Project 8. Currently, a cavity based CRES experiment is the preferred technology choice for future experiments by the Project 8 collaboration including the Phase IV experiment.

In this chapter I provide a brief summary of resonant cavities and sketch out the key features of a cavity based CRES experiment. In Section 6.2 I provide a brief introduction to cylindrical resonant cavities and the solutions for the electromagnetic fields in the cavity volume.

In Section 6.3 I describe the main components of a cavity based CRES experiment, including the background and trap magnets, cavity geometry and design, and cavity coupling considerations. I also discuss some relevant trade-offs between an antenna array and cavity CRES experiment, and highlight some reasons for the transition of Project 8 to the development of a cavity based experiment.

Finally, in Sections 6.4 and 6.5, I present the design and development of an open mode-filtered cavity that could be used in a cavity based CRES experiment with atomic tritium. The results of the cavity simulations are confirmed by laboratory measurements of a proof-of-principle prototype that demonstrates key features of the design.

4706 6.2 Cylindrical Resonant Cavities

4707 Resonant cavities are sealed conductive containers, which allows one to describe the
4708 electromagnetic (EM) fields contained in the cavity volume as a superposition of resonant
4709 modes [85]. The field shapes of the resonant modes are determined by Maxwell's equations
4710 and the boundary conditions enforced by the cavity geometry. Of interest to Project 8
4711 for CRES measurements are cylindrical cavities due to their ease of construction and
4712 integration with atom and electron trapping magnets.

4713 6.2.1 General Field Solutions

4714 Consider a long segment of conducting material with a cylindrical cross-section (see
4715 Figure 6.1). A geometry such as this can be used as a waveguide transmission line to
4716 transfer EM energy from point to point, or, if conducting shorts are inserted on both
4717 ends of the cylinder, the waveguide becomes a resonant cavity.

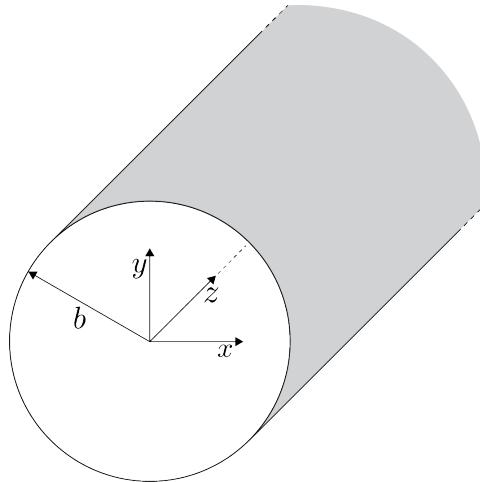


Figure 6.1. Geometry of a cylindrical waveguide with radius b .

4718 The fields allowed inside a cylindrical cavity are determined by the boundary conditions
4719 of the cylindrical geometry. The general approach to solving the fields begins by assuming
4720 solutions to Maxwell's equations of the form

$$\mathbf{E}(x, y, z) = (\mathbf{e}(x, y) + \hat{z}e_z(x, y))e^{-i\beta z}, \quad (6.1)$$

$$\mathbf{H}(x, y, z) = (\mathbf{h}(x, y) + \hat{z}h_z(x, y))e^{-i\beta z}. \quad (6.2)$$

4721 The solutions assume a harmonic time dependence of the form $e^{i\omega t}$ and propagation

⁴⁷²² along the positive z-axis. The functions $\mathbf{e}(x, y)$ and $\mathbf{h}(x, y)$ represent the transverse
⁴⁷²³ (\hat{x}, \hat{y}) components of the electric and magnetic fields respectively, and $e_z(x, y)$, $h_z(x, y)$
⁴⁷²⁴ represent the longitudinal components. The version of Maxwell's equations in the case
⁴⁷²⁵ where there are no source terms can be written as a pair of coupled differential equations,

$$\nabla \times \mathbf{E} = -i\omega\mu\mathbf{H}, \quad (6.3)$$

$$\nabla \times \mathbf{H} = i\omega\epsilon\mathbf{E}, \quad (6.4)$$

⁴⁷²⁶ where ϵ and μ are the permittivity and permeability of the material inside the waveguide
⁴⁷²⁷ or cavity. Using the field solutions from Equations 6.1 and 6.2 one can solve for the
⁴⁷²⁸ transverse components of the fields in terms of the longitudinal fields. Because cylindrical
⁴⁷²⁹ cavities are of interest, it is advantageous to write the field solutions in cylindrical
⁴⁷³⁰ coordinates. After performing this transformation, the set of four equations for the
⁴⁷³¹ transverse field components are

$$H_\rho = \frac{i}{k_c^2} \left(\frac{\omega\epsilon}{\rho} \frac{\partial E_z}{\partial\phi} - \beta \frac{\partial H_z}{\partial\rho} \right), \quad (6.5)$$

$$H_\phi = \frac{-i}{k_c^2} \left(\omega\epsilon \frac{\partial E_z}{\partial\rho} + \frac{\beta}{\rho} \frac{\partial H_z}{\partial\phi} \right), \quad (6.6)$$

$$E_\rho = \frac{-i}{k_c^2} \left(\beta \frac{\partial E_z}{\partial\rho} + \frac{\omega\mu}{\rho} \frac{\partial H_z}{\partial\phi} \right), \quad (6.7)$$

$$E_\phi = \frac{i}{k_c^2} \left(\frac{-\beta}{\rho} \frac{\partial E_z}{\partial\phi} + \omega\mu \frac{\partial H_z}{\partial\rho} \right), \quad (6.8)$$

⁴⁷³² where k_c is the cutoff wavenumber defined by $k_c^2 = k^2 - \beta^2$ with $k = \omega\sqrt{\mu\epsilon}$ being the
⁴⁷³³ wavenumber of the EM radiation.

⁴⁷³⁴ This set of equations can be used to solve for a variety of different modes, which can
⁴⁷³⁵ be obtained by setting conditions on E_z and H_z . For cylindrical cavities two types of
⁴⁷³⁶ modes are allowed, which correspond to solutions where $E_z = 0$ and $H_z = 0$ respectively.

⁴⁷³⁷ 6.2.2 TE and TM Modes

⁴⁷³⁸ The TE family of modes corresponds to the case where $E_z = 0$. This implies that H_z is
⁴⁷³⁹ a solution to the Helmholtz wave equation

$$(\nabla^2 + k^2)H_z = 0. \quad (6.9)$$

4740 For solutions of the form $H_z(\rho, \phi, z) = h_z(\rho, \phi)e^{-i\beta z}$, Equation 6.9 can be solved using
 4741 the standard technique of separation of variables. Rather than reproduce the derivation
 4742 here I shall simply quote the solutions for the transverse fields [85], which are

$$H_\rho = \frac{-i\beta}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.10)$$

$$H_\phi = \frac{-i\beta n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.11)$$

$$E_\rho = \frac{-i\omega\mu n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.12)$$

$$E_\phi = \frac{i\omega\mu}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}. \quad (6.13)$$

4743 One observes that the solutions have a periodic dependence on ϕ , and radial profiles
 4744 given by the Bessel functions of the first kind. The integer indices n and m arise from
 4745 continuity conditions on the EM fields in the azimuthal and radial directions. For the
 4746 TE modes, the indices range from $n \geq 0$ and $m \geq 1$. $k_{c_{nm}}$ is the cutoff wavenumber for
 4747 the TE_{nm} mode given by

$$k_{c_{nm}} = \frac{p'_{nm}}{b}, \quad (6.14)$$

4748 where b is the radius of the cavity or waveguide and p'_{nm} is the m -th root of the derivative
 4749 of the n -th order Bessel function (see Table 6.1).

Table 6.1. A table of the values of p'_{nm} .

n	p'_{n1}	p'_{n2}	p'_{n3}
0	3.832	7.016	10.174
1	1.841	5.331	8.536
2	3.054	6.706	9.970

4750 The TM mode family corresponds to the case where $H_z = 0$, and $(\nabla^2 + k^2)E_z = 0$.
 4751 Again, solutions are assumed of the form $E_z(\rho, \phi, z) = e_z(\rho, \phi)e^{-i\beta z}$, for which the general
 4752 form of the solutions is the same as for the TE modes. However, the different boundary
 4753 conditions for the TM modes results in particular solutions with a different form, which I
 4754 shall quote here without derivation. The transverse fields of the TM modes are given by

$$H_\rho = \frac{-i\omega\epsilon n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.15)$$

$$H_\phi = \frac{-i\omega\epsilon}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z} \quad (6.16)$$

$$E_\rho = \frac{-i\beta}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.17)$$

$$E_\phi = \frac{-i\beta n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.18)$$

4755 which one may notice are the same solutions as the TE modes with H and E flipped.

4756 The cutoff wavenumber for the TM modes is given by, $k_{c_{nm}} = p_{nm}/b$, where the values of

4757 p_{nm} correspond to the m -th zero of the n -th order Bessel function (see Table 6.2).

Table 6.2. A table of the values of p_{nm} .

n	p_{n1}	p_{n2}	p_{n3}
0	2.405	5.520	8.654
1	3.832	7.016	10.174
2	5.135	8.417	11.620

4758 6.2.3 Resonant Frequencies of a Cylindrical Cavity

4759 A cylindrical cavity is constructed by taking a section of cylindrical waveguide and
4760 shorting both ends with conductive material. This means that the electric fields inside a
4761 cylindrical cavity are exactly those derived in Section 6.2.2 with the additional condition
 that the electric fields must go to zero at $z = 0$ and $z = L$ (see Figure 6.2).

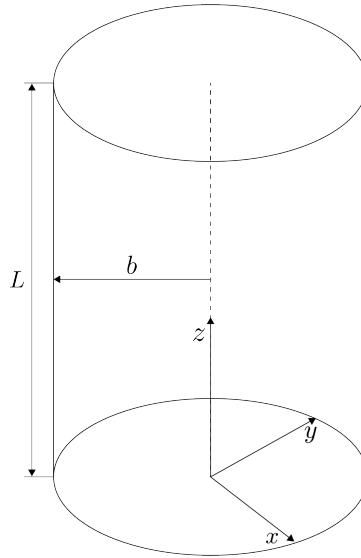


Figure 6.2. The geometry of a cylindrical cavity with length L and radius b .

⁴⁷⁶³ The transverse electric field solutions for a cylindrical waveguide are of the form

$$\mathbf{E}(\rho, \phi, z) = \mathbf{e}(\rho, \phi) (A_+ e^{-i\beta_{nm}z} + A_- e^{i\beta_{nm}z}), \quad (6.19)$$

⁴⁷⁶⁴ where A_+ and A_- are arbitrary amplitudes of forward and backward propagating waves.

⁴⁷⁶⁵ In order to enforce that \mathbf{E} is zero at both ends of the cavity it is required that

$$\beta_{nm}L = 2\pi\ell, \quad (6.20)$$

⁴⁷⁶⁶ where $\ell = 0, 1, 2, 3, \dots$. Using this constraint on the propagation constant one can solve

⁴⁷⁶⁷ for the resonant frequencies of the TE_{nml} and the TM_{nml} modes in a cylindrical cavity.

⁴⁷⁶⁸ For the TE modes the resonant frequencies are

$$f_{nml} = \frac{c}{2\pi\sqrt{\mu_r\epsilon_r}} \sqrt{\left(\frac{p'_{nm}}{b}\right)^2 + \left(\frac{\ell\pi}{L}\right)^2}, \quad (6.21)$$

⁴⁷⁶⁹ and the frequencies of the TM modes are

$$f_{nml} = \frac{c}{2\pi\sqrt{\mu_r\epsilon_r}} \sqrt{\left(\frac{p_{nm}}{b}\right)^2 + \left(\frac{\ell\pi}{L}\right)^2}. \quad (6.22)$$

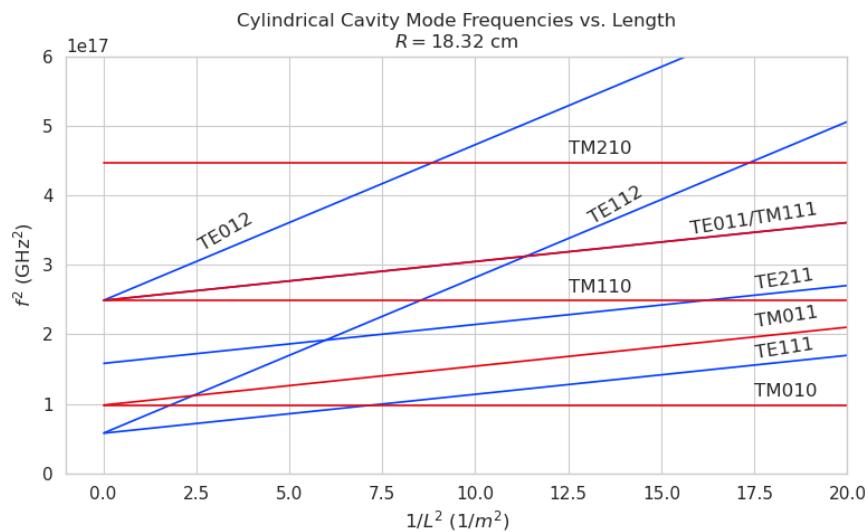


Figure 6.3. Relation of mode frequency to cavity length for a cylindrical cavity with a radius of 18.32 cm.

4770 6.2.4 Cavity Q-factors

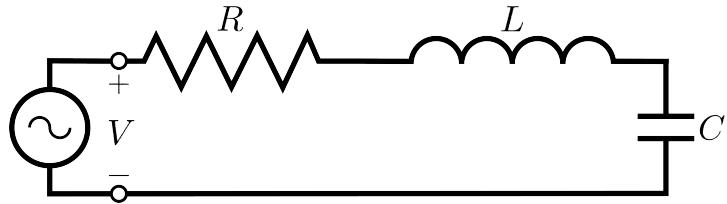


Figure 6.4. A series RLC circuit.

4771 The resonant behavior of cylindrical cavities can be modeled as a series RLC circuit
 4772 (see figure 6.4). The input impedance of the circuit can be obtained by applying
 4773 Kirchhoff's laws to calculate the impedance of the equivalent circuit. For a series RLC
 4774 circuit the input impedance is

$$Z_{\text{in}} = \left(\frac{1}{R} + \frac{1}{i\omega L} + i\omega C \right). \quad (6.23)$$

4775 The resistance in the circuit represents all sources of loss in the cavity, which is primarily
 4776 caused by the finite conductivity of the cavity walls. The inductor and capacitor represent
 4777 the energy stored in the cavity in the form of electric and magnetic fields. If the circuit
 4778 is being driven by an external power source the input power can be written in terms of
 4779 the circuit input impedance and the source voltage

$$P_{\text{in}} = \frac{1}{2} Z_{\text{in}} |I|^2 = \frac{1}{2} |I|^2 \left(\frac{1}{R} + \frac{1}{i\omega L} + i\omega C \right). \quad (6.24)$$

4780 The resistor introduces a loss into the system with a power given by

$$P_{\text{loss}} = \frac{1}{2} |I|^2 R, \quad (6.25)$$

4781 and the capacitor and inductor store energies given by

$$W_e = \frac{1}{4} \frac{|I|^2}{\omega^2 C}, \quad (6.26)$$

$$W_m = \frac{1}{4} |I|^2 L, \quad (6.27)$$

4782 respectively. Using these expressions the input power and input impedance can be written

⁴⁷⁸³ in terms of the lost power and stored energy

$$P_{\text{in}} = P_{\text{loss}} + 2i\omega(W_m - W_e), \quad (6.28)$$

$$Z_{\text{in}} = \frac{P_{\text{loss}} + 2i\omega(W_m - W_e)}{\frac{1}{2}|I|^2}. \quad (6.29)$$

⁴⁷⁸⁴ The condition for resonance in the RLC circuit is that the stored magnetic energy
⁴⁷⁸⁵ is equal to the stored electric energy ($W_e = W_m$). When this occurs $Z_{\text{in}} = R$, which is a
⁴⁷⁸⁶ purely real impedance, and $P_{\text{in}} = P_{\text{loss}}$. The resonant frequency of the circuit can be
⁴⁷⁸⁷ determined from the condition $W_e = W_m$ from which one finds that

$$\omega_0 = \frac{1}{\sqrt{LC}}. \quad (6.30)$$

⁴⁷⁸⁸ An important performance parameter for any resonant system is the Q-factor, which
⁴⁷⁸⁹ quantifies the quality of the resonator as the ratio of the stored energy multiplied by the
⁴⁷⁹⁰ resonant frequency to the average energy lost per second. For the series RLC circuit, the
⁴⁷⁹¹ Q-factor is given by the expression

$$Q_0 = \omega \frac{W_e + W_m}{P_{\text{loss}}} = \frac{1}{\omega_0 RC}, \quad (6.31)$$

⁴⁷⁹² from which one observes that as the resistance of the RLC circuit is decreased the quality
⁴⁷⁹³ factor of the resonator increases. From the perspective of cylindrical cavities this implies
⁴⁷⁹⁴ that as one decreases the resistance of the cavity walls it is expected that the Q-factor of
⁴⁷⁹⁵ the cavity should increase, which is indeed the case. In certain applications where a high
⁴⁷⁹⁶ Q is desireable it is possible to manufacture a cavity out of superconducting materials in
⁴⁷⁹⁷ order to minimize the power losses of the system.

⁴⁷⁹⁸ The Q-factor of the resonator also determines with bandwidth (BW) of the system. A
⁴⁷⁹⁹ cavity with a high Q-factor will resonant with a smaller range of frequencies than a cavity
⁴⁸⁰⁰ with a low Q-factor. To see this examine the behavior of the RLC circuit when driven by
⁴⁸⁰¹ frequencies near the resonance. For a frequency $\omega = \omega_0 + \Delta\omega$, where $\Delta\omega = \omega - \omega_0 \ll \omega_0$,
⁴⁸⁰² the input impedance can be written as

$$Z_{\text{in}} = R + i\omega L \left(\frac{\omega^2 - \omega_0^2}{\omega^2} \right), \quad (6.32)$$

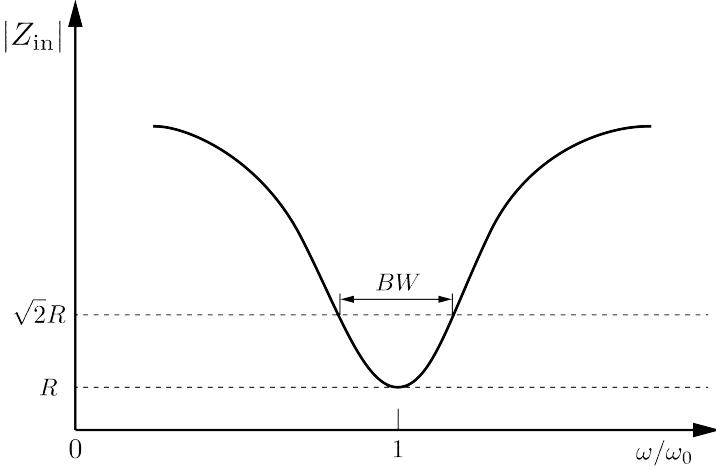


Figure 6.5. Illustration of the behavior of the input impedance of the series RLC circuit as a function of the driving frequency. The BW is proportion to the width of the resonance, which is inversely proportional to Q.

4803 and by expanding $(\omega^2 - \omega_0^2)/\omega^2$ to first order in $\Delta\omega$, one obtains

$$Z_{\text{in}} \approx R + i \frac{2RQ_0\Delta\omega}{\omega_0}. \quad (6.33)$$

4804 Therefore, the magnitude of the input impedance near the resonance is given by

$$|Z_{\text{in}}| = R \sqrt{1 + 4Q_0^2 \frac{\Delta\omega^2}{\omega^2}}, \quad (6.34)$$

4805 from which it is seen that for the series RLC circuit the input impedance is minimized
 4806 at the resonant frequency, which corresponds to the maximum input power (see Figure
 4807 6.5). The half-power BW is the range of frequencies over which the input power drops to
 4808 half the input power on resonance. This occurs when $|Z_{\text{in}}| = \sqrt{2}R$, which corresponds to
 4809 $\Delta\omega/\omega = \text{BW}/2$. Using Equation 6.34 one can find that

$$2R^2 = R^2(1 + Q_0^2\text{BW}^2), \quad (6.35)$$

4810 which implies

$$\text{BW} = \frac{1}{Q_0} \quad (6.36)$$

4811 It is important to emphasize that the Q-factor defined here, Q_0 , is technically the
 4812 unloaded Q. It reflects the quality of the cavity or resonant circuit without the influence
 4813 of any external circuitry. In practice, however, a cavity is invariably coupled to an

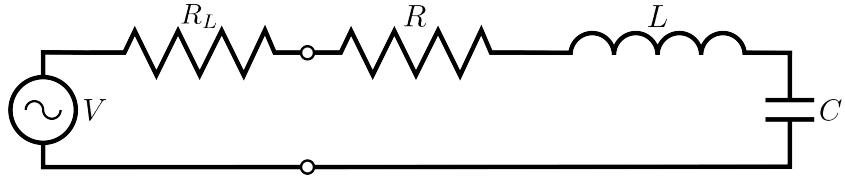


Figure 6.6. A series RLC circuit coupled to an external circuit with input impedance R_L .

4814 external circuit to drive a cavity resonance or to measure the energy of a resonant mode.
 4815 Coupling a cavity to an external circuit changes the Q by loading the equivalent cavity
 4816 RLC circuit (see Figure 6.6). The Q-factor of the cavity when it is loaded by an external
 4817 circuit is called the loaded Q, which is the quantity that one actually measures when
 4818 exciting a resonance in the cavity. Using the series RLC circuit model one can see that
 4819 the load resistor in Figure 6.6 will add in series with the resistor in the circuit for a total
 4820 equivalent resistance of $R + R_L$. Therefore, the loaded Q is given by

$$Q_L = \frac{1}{\omega_0(R + R_L)C}, \quad (6.37)$$

4821 from which one observes that the loaded Q is always less than the intrinsic Q of the
 4822 cavity.

4823 The amount of coupling that is desireable depends on the specific application of
 4824 the resonator. If one wants a resonator that is particular frequency selective than it
 4825 makes sense to limit the amount of coupling to the cavity to maintain a small BW,
 4826 alternatively, if a larger BW is need one can increase the cavity coupling by tuning the
 4827 input impedance of the external circuit. The critical point, where maximum power is
 4828 transferred between the cavity and the external circuit, occurs when the input impedance
 4829 of the cavity matches the input impedance of the external transmission line. For the
 4830 series RLC circuit on resonance, this matching condition corresponds to

$$Z_0 = Z_{in} = R, \quad (6.38)$$

4831 where Z_0 is the impedance of the transmission line. The loaded Q at this critical point
 4832 is, therefore,

$$Q_L = \frac{1}{2\omega_0 Z_0 C} = \frac{Q_0}{2}. \quad (6.39)$$

4833 One can described the degree of coupling between the cavity and an external circuit by

4834 defining a coupling factor, g , such that,

$$g = \frac{Q_0}{Q_L} - 1. \quad (6.40)$$

4835 When $g = 1$ then $Q_L = Q_0/2$, and the cavity is said to be critically coupled. If
4836 $Q_L < Q_0/2$, then the cavity is undercoupled to the transmission line, corresponding to
4837 $g < 1$. Alternatively, if $Q_L > Q_0/2$, then $g > 1$, and the cavity is overcoupled to the
4838 transmission line. Various specialized circuits can be used to tune the input impedance
4839 of the external circuit as seen by the cavity to achieve a wide range of different coupling
4840 factors based on the desired application of the cavity.

4841 6.3 The Cavity Approach to CRES

4842 6.3.1 A Sketch of a Molecular Tritium Cavity CRES Experiment

4843 Resonant cavities can be used to perform CRES measurements, and they represent the
4844 current preferred technology by the Project 8 collaboration. The basic approach to a
4845 neutrino mass measurement using a resonant cavity and molecular tritium beta-decay
source is illustrated by Figure 6.7.

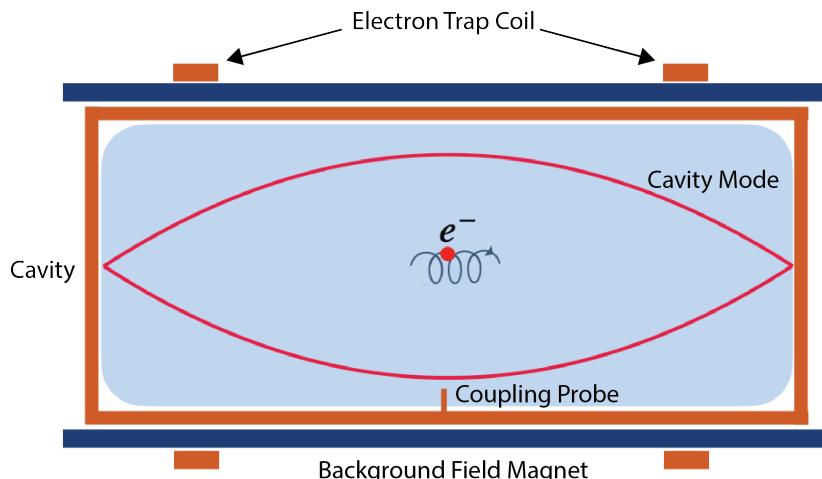


Figure 6.7. A cartoon depiction of a cavity CRES experiment. A metallic cavity filled with tritium gas is inserted into a uniform background magnetic field to perform CRES measurements. Electrons from beta-decays inside the cavity can be trapped and used to excite a resonant mode(s). By coupling to the cavity mode with a suitable probe one can measure the cyclotron frequency of the electron and perform CRES.

4846

4847 At the core of the experiment is a large resonant cavity filled with tritium gas. The
4848 filled cavity is then placed in a uniform magnetic field provided by a primary magnet,
4849 which provides the background magnetic field. The value of the background magnetic field
4850 sets the range of cyclotron frequencies for electrons emitted near the tritium spectrum
4851 endpoint. When a beta-decay electron is produced in the cavity it is trapped using a set
4852 of magnetic pinch coils that keep electrons inside the cavity volume.

4853 Electrons trapped inside the cavity do not radiate in the same way as electrons
4854 in free-space. Effectively, the same boundary conditions that were used to derive the
4855 resonant modes of a cylindrical cavity in Section 6.2 apply to the radiation of the electron
4856 as well. The coupling of an electron performing cyclotron motion in a cavity has been
4857 studied in detail for measurements of the electron’s magnetic moment [94–96]. If an
4858 electron is emitted with a kinetic energy that corresponds to a cyclotron frequency that
4859 matches a resonant frequency of the cavity, then energy radiated by the electron excites
4860 a corresponding resonance in the cavity. The strength of the electron’s coupling to the
4861 cavity is given to first order by the dot product between the electrons trajectory and
4862 the electric field vector of the resonant mode. Additional effects, such as the Purcell
4863 enhancement [97], alter the emitted power from the free-space Larmor equation [49]. If an
4864 electron is moving with a cyclotron frequency that is far from any resonant modes in the
4865 cavity, then radiation from the electron is suppressed. One can interpret this somewhat
4866 surprising effect as the metallic walls of the cavity reflecting the radiated energy back to
4867 the electron.

4868 Detecting an electron in the cavity is accomplished by coupling the cavity to an
4869 external transmission line that leads to an amplifier and RF receiver chain [98]. The
4870 coupling of the cavity resonance to the amplifier occurs through a coupling probe or
4871 aperture designed to read-out the excitation of the mode(s) excited by the electron. For
4872 CRES measurements, the placement of a wire antenna coupling probe inside the cavity
4873 volume leads to unacceptable losses of tritium atoms due to recombination to molecular
4874 tritium on the antenna surface, therefore, apertures are the preferred coupling method
4875 for cavity CRES experiments.

4876 One of the attractive features of the CRES technique for neutrino mass measurement
4877 is the gain in statistics that comes from the differential nature of the tritium spectrum
4878 measurement. Initially, this seems incompatible with cavities, due to the narrow reso-
4879 nances of cavity modes giving relatively small bandwidth. However, by intentionally
4880 over-coupling to a single cavity mode one can achieve bandwidths of a few 10’s of MHz
4881 (see Section 6.2), which is sufficient for a measurement of the tritium spectrum endpoint

4882 region.

4883 **6.3.2 Magnetic Field, Cavity Geometry, and Resonant Modes**

4884 **Magnetic Field and Volume Scaling**

4885 For a CRES experiment, cylindrical cavities are a natural choice since they match
4886 the geometry of standard solenoid magnets, which are needed in order to produce the
4887 background magnetic field for CRES measurements. Furthermore, the cylindrical shape is
4888 compatible with a Halbach array, which is the leading choice of atom trapping technology
4889 for future atomic tritium experiments by the Project 8 collaboration. Cylindrical
4890 cavities also benefit from well-established machining practices that are able to achieve
4891 high geometric precision at large lengths scales. More exotic cavity designs are under-
4892 consideration and there are on-going efforts to investigate the potential advantages these
4893 may have over the standard cylindrical geometry.

4894 As shown in Section 6.2, the physical dimensions of the cavity are directly coupled to
4895 the resonant frequencies of the cavity. This dependency links the size of the cavity to
4896 the magnitude of the background magnetic field, because the magnetic field determines
4897 the cyclotron frequencies of trapped electrons. Specifically, as the size of the cavity is
4898 increased to accommodate larger volumes of tritium gas, the frequencies of the resonant
4899 modes decrease proportionally. This requires that the magnetic field also decrease in
4900 order to maintain coupling between electrons and the desired cavity mode.

4901 The required cavity size is ultimately determined by the required statistics in the
4902 tritium spectrum endpoint region. Because the gas density must be kept below a certain
4903 level to ensure that electrons have sufficient time to radiate before scattering, larger
4904 volumes become the only way to achieve higher event statistics. To achieve the sensitivity
4905 goals of Phase III and IV cavity volumes on the order of several cubic-meters are required,
4906 which pushes one towards frequencies in the range of 100's of MHz.

4907 **Single-mode Cavity CRES**

4908 It is tempting to consider maintaining a high magnetic field, while still increasing the size
4909 of the cavity, in order to increase the radiated power from trapped electrons for better
4910 SNR. However, if one were to maintain the same magnetic field while increasing the
4911 size of the cavity, the electrons would begin to couple to higher order modes with more
4912 complicated transverse geometries. The danger with this approach is that a complicated
4913 mode structure could introduce systematic errors into the CRES signals. Example

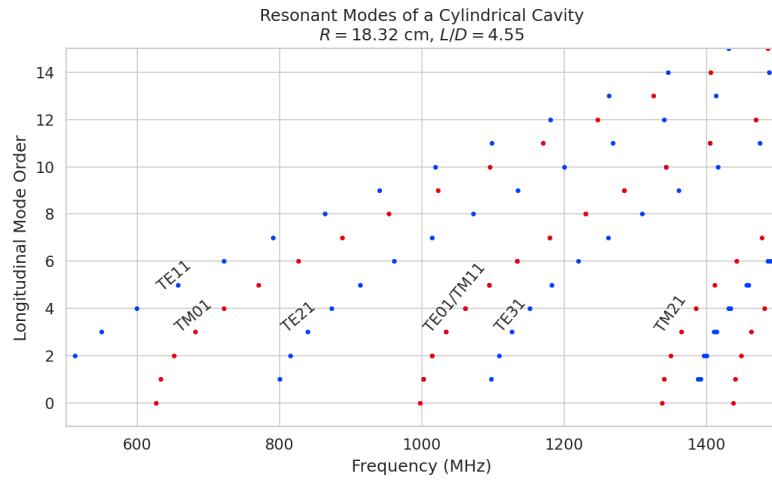
4914 systematics include unpredicted mode hybridization or changes in the mode shapes from
4915 imperfections in the cavity construction, which would prevent reconstruction of the
4916 electron's starting kinetic energies with adequate resolution. For this reason, it is ideal
4917 to operate with magnetic fields that give cyclotron frequencies near the fundamental
4918 frequency of the cavity, where the mode structure is relatively simple (see Figure 6.8).
4919 In this frequency region it is possible to perform CRES by coupling to only a single
4920 resonant mode, however, it is currently an open question if a single mode measurement
4921 will provide enough information about an individual electron's position to reconstruct
4922 the full event. Regardless, developing a solid understanding of the CRES phenomenology
4923 when an electron is coupling to a single mode will be a necessary step towards a future
4924 multi-mode cavity experiment.

4925 Considerations for Resonant Mode Selection

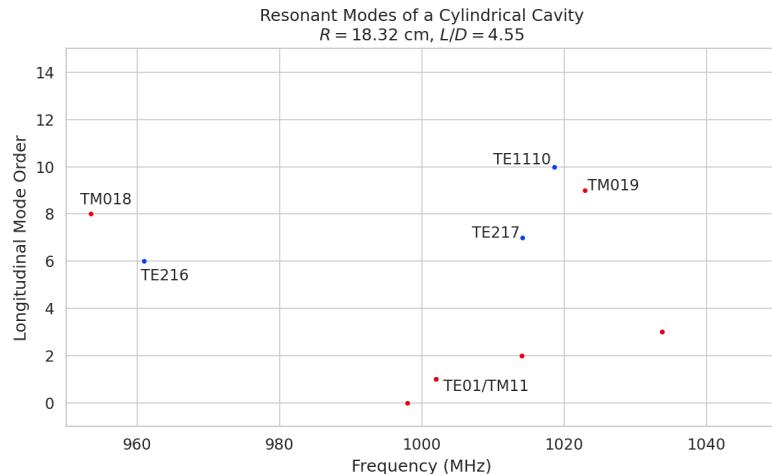
4926 A single-mode cavity experiment begs the question, which resonant mode is best for
4927 CRES measurements? There is an immediate bias towards low order TE_{nm} and TM_{nm}
4928 modes due to the multi-mode considerations discussed above. Additionally, there is a
4929 preference towards modes with longitudinal index $\ell = 1$ with a single antinode along the
4930 vertical axis of the cylindrical cavity. The reason for this is that there is a phase change
4931 in the electric fields between antinodes that leads to modulation effects that destroy the
4932 carrier frequency signal information.

4933 A second consideration for mode selection is the volumetric efficiency of the mode.
4934 Volumetric efficiency can be thought of as an integral over the volume of the cavity
4935 weighted by the relative amplitude of the mode. From the perspective of simply maximiz-
4936 ing the volume useable for CRES measurements this integral would be as close to unity
4937 as possible. However, there is a requirement to reconstruct the position of the electrons
4938 inside the cavity volume so that the local magnetic fields can be used to convert the
4939 measured cyclotron frequency to a kinetic energy. With a single mode this necessarily
4940 requires a variable transverse mode amplitude, which lowers the volumetric efficiency, so
4941 that position of the electron in the cavity can be estimated from the average amplitude
4942 of the CRES signal. Longitudinal indices of $\ell = 1$ have an advantage in volumetric
4943 efficiency over higher order ℓ modes, since there are only two longitudinal nodes, one at
4944 each end of the cavity. Therefore, the average coupling strength of trapped electrons as
4945 they oscillate axially is higher for $\ell = 1$ modes.

4946 The longitudinal variation in the mode strength is ultimately critical for achieving the
4947 energy resolution required for neutrino mass measurements. Correcting for the change in



(a)



(b)

Figure 6.8. Examples of the resonant mode frequencies of a cylindrical cavity. This cavity has a radius of 18.32 cm and a length to diameter ratio of 4.55.

the average magnetic fields experienced by electrons with different pitch angles requires that information on the axial motion of the electron be encoded into the CRES signal. The longitudinal variation in the mode amplitude leads to amplitude modulation of the CRES signal with a frequency proportional to the electron's pitch angle.

An additional factor for mode selection is the intrinsic or unloaded Q of the mode. In terms of SNR it is advantageous to use a mode with a very high Q_0 , which is then highly overcoupled to achieve the necessary bandwidth to cover the tritium endpoint spectrum. This scheme leads to a decoupling of the physical cavity temperature from the effective noise temperature after the amplifier, which allows us to achieve adequate SNR without

4957 the requirement of cooling the entire cavity to single Kelvin temperatures.

4958 An example of a resonant mode that exhibits these traits is the TE₀₁₁ mode. At present
4959 the TE₀₁₁ mode is the preferred resonance for a single-mode cavity CRES experiment
4960 by the Project 8 collaboration. TE₀₁₁ is a low order mode located in a region relatively
4961 far from other cavity modes. Furthermore, the separation of the TE₀₁₁ mode can be
4962 improved by various mode-filtering techniques discussed in Section 6.4.2 below. TE₀₁₁
4963 consists of a single longitudinal antinode that can provide pitch angle information in the
4964 form of amplitude modulation, and has an electric field with a radial profile given by the
4965 J'_0 Bessel function allowing for radial position estimation. Lastly, the TE₀₁₁ mode has a
4966 relatively high intrinsic Q compared to nearby modes, which helps with SNR. Unloaded
4967 Q's greater than 80000 are achievable for a 1 GHz TE₀₁₁ resonance using a copper walled
4968 cavity.

4969 **6.3.3 Trade-offs Between the Antenna and Cavity Approaches**

4970 The choice between cavities and antennas for large-scale CRES measurements is not
4971 without trade-offs. Both the antenna array and cavity approaches are relatively immature
4972 techniques, at present there are no known obstacles that would prevent either approach
4973 from being used for a large scale neutrino mass experiment. The preference for cavities
4974 is largely driven by important practical considerations that could make a cavity based
4975 experiment significantly cheaper than an antenna experiment of similar size and scope.
4976 However, the switch to cavities also introduces new challenges less relevant to the
4977 antenna array, which must be solved in order for Project 8 to achieve its neutrino mass
4978 measurement goals.

4979 One of the major relative drawbacks of the antenna array approach is the size and
4980 complexity of the data-acquisition system. A large-scale antenna array experiment
4981 requires $O(100)$ antennas independently digitized at rates of $O(10)$ to $O(100)$ MHz. Since
4982 there is insufficient information in a single antenna channel to detect or reconstruct the
4983 CRES signal, the entire array output must be processed during the signal reconstruction.
4984 Because data storage becomes an issue with these data volumes, there is a real-time
4985 signal reconstruction requirement that allows one to detect CRES signals buried in the
4986 thermal noise. As discussed in Section 4.4, the computational cost of these real-time
4987 detection algorithms are potentially quite large for even a small scale antenna array
4988 experiment. However, the operating principle of a cavity experiment allows the CRES
4989 signal to be detected using only a single read-out channel digitized at rates of $O(10)$ MHz,
4990 which reduces the cost of the data acquisition system by many orders of magnitude.

4991 From an engineering perspective, the simple geometry and thin-walls of a cylindrical
4992 cavity are simpler to interface with the cryogenic and magnetic subsystems needed for a
4993 CRES experiment. Whereas, the antenna array requires careful design and engineering
4994 to accommodate the antenna array and receiver electronics in proximity to the trapping
4995 magnets. Additionally, due to near-field interference effects, the antenna array is unable
4996 to reconstruct CRES events within the reactive near-field distance of the antennas.
4997 Because atom trapping requirements require magnetic fields which correspond to cyclotron
4998 frequencies for endpoint electrons less than 1 GHz, the required stand-off distance leads to
4999 a significant loss in useable experiment volume, necessitating larger and more expensive
5000 magnets.

5001 Another advantage to the cavity approach is the relatively compact sideband structure,
5002 which is a result of the low modulation index for cavity CRES signals. The axial motion
5003 in an antenna array experiment leads to frequency modulation and sidebands. The shape
5004 of the sideband structure is determined by the modulation index, $h = \frac{\Delta f}{f_a}$, where Δf
5005 is the size of the frequency deviation and f_a is the axial frequency. The large electron
5006 traps required for a cubic-meter-scale experiment leads to high modulation indices, which
5007 causes the signal spectrum to be made up of numerous low power sidebands that make
5008 reconstruction and detection challenging. This behavior was observed in simulations
5009 of the FSCD in which carrier power decreased with pitch angle due to the increase in
5010 modulation index (see Figure 4.30). For cavities, however, the modulation index remains
5011 near $h = 1$ even for very long magnetic traps due to the high phase velocity in cavities
5012 relative to the axial velocity of the electron. This results in an almost ideal spectrum
5013 shape that has a strong carrier frequency with a few sidebands whose relative amplitudes
5014 encode pitch angle information.

5015 A downside of the cavity approach is the apparent difficulty of estimating the position
5016 of the electron using only the coupling of the electron to a single mode. The amplitude of
5017 the TE₀₁₁ mode is completely independent of the azimuthal coordinate, therefore, position
5018 reconstruction using the TE₀₁₁ mode is only able to estimate the radial position of the
5019 electron. This position degeneracy may lead to magnetic field uniformity requirements
5020 that are too challenging to meet due to mechanical uncertainties in cavity and magnet
5021 construction, as well as uncertainties caused by nuisance external magnetic fields such
5022 as the Earth's field and magnetic fields from building materials. A multi-mode cavity
5023 experiment may provide a way to extract more precise information on the position of
5024 the electron by analyzing the coupling of the electron to several modes that overlap in
5025 different ways.

5026 **6.4 Single-mode Resonant Cavity Design and Simulations**

5027 The single-mode cylindrical cavities envisioned for the Phase III and IV experiments must
5028 be carefully engineered in order to measure the neutrino mass with the desired sensitivity.
5029 In this section I summarize some simulation studies performed to analyze early design
5030 concepts for a single-mode cavity. The primary tool for these investigations was Ansys
5031 HFSS, which was also used for the development of the SYNCA antenna described in
5032 Section 5.3.

5033 **6.4.1 Open Cylindrical Cavities with Coaxial Terminations**

5034 **Design Concept**

5035 A basic cavity design question relevant to Project 8's ultimate goal of an atomic tritium
5036 CRES experiment is how to build a cavity that can be efficiently filled with atomic
5037 tritium. To keep the rate of atom loss from recombination on surfaces it is ideal if the
5038 ends of the cylindrical cavity are as open as possible so that tritium atoms can flow
5039 inside unimpeded. Additionally, one of the primary calibration techniques planned for
5040 future CRES experiments involves CRES measurements using electrons injected from
5041 an electron gun source, which also requires an opening at the cavity end. Cylindrical
5042 cavities with open ends can be manufactured, however, the intrinsic Q-factors of these
5043 cavities are orders of magnitude less than their sealed counterparts, which reduces the
5044 signal-to-noise ratio when that cavity is used for CRES measurement.

5045 Cylindrical cavities with mostly open ends that also exhibit Q values for the $TE_{01\ell}$
5046 modes similar to sealed cavities can be built by using coaxial endcaps to terminate the
5047 cavity. Cavities of this type have been manufactured for specialized applications related
5048 to the measurements of the dielectric constants of liquefied gasses (see Figure 6.9) [2, 3].
5049 This cavity design leaves the ends of the cavity wide open, but retains high Q-values for
5050 the $TE_{01\ell}$ modes due to the coaxial endcap, which are designed to perfectly reflect the
5051 electric fields of $TE_{01\ell}$ modes. Coupling to the $TE_{01\ell}$ mode is achieved via an aperture
5052 located at the center of the cavity wall.

5053 A cavity similar to Figure 6.9 is a candidate design for the future CRES experiments
5054 by Project 8, since it appears to elegantly solve many practical issues that arise when
5055 combining cavity CRES and atomic tritium. The coaxial endcaps leave significant regions
5056 of the cavity ends completely open, which allows for the entrance of atomic tritium as
5057 well as the pumping away of molecular tritium that has recombined on the cavity walls.

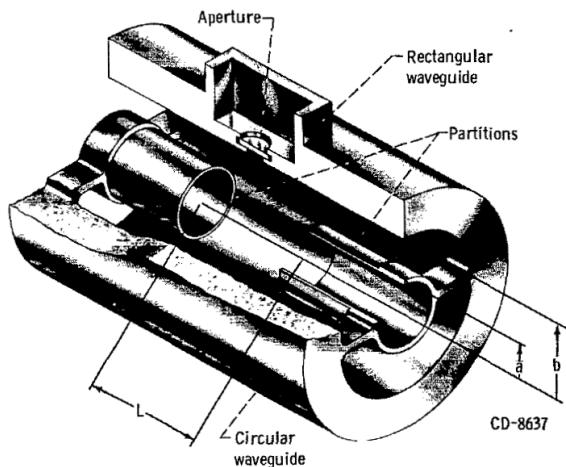


Figure 6.9. An image of an open cavity with coaxial terminations used for dielectric constant measurements. Figure from [2].

These open ends are achieved while preserving the high Q-values of the $\text{TE}_{01\ell}$ modes, which is important for extracting as much signal power from the electron as possible. In subsequent sections this cavity design will be analyzed in more detail, primarily by using HFSS simulations to analyze the resonant mode structure of this cavity geometry.

Coaxial Terminator Constraints

The reason that coaxial endcaps can be used to achieve high Q-values for the $\text{TE}_{01\ell}$ modes is that the electric fields for these modes are purely azimuthally polarized (see Equations 6.12 and 6.13). Therefore, the boundary conditions that require the electric field to go to zero at the cavity ends can be supplied using a coaxial partition of the correct radius (see Figure 6.10). Because the cylindrical shape enforced by the partition does not match the boundary conditions of other cavity modes, these terminations also significantly suppress the Q-factors of non- $\text{TE}_{01\ell}$ modes, which is potentially beneficial for a single-mode cavity CRES experiment.

The correct radius of the cylindrical partition is derived by setting up the boundary value problem in Figure 6.10, and analyzing the reflection and transmission coefficients for waves incident on the coaxial terminators. The basic problem is to identify the radius a where the reflection coefficient for the $\text{TE}_{01\ell}$ modes becomes equal to 1. One can show that if the coaxial partitions are made sufficiently long relative to the wavelength of the TE_{01} modes than perfect reflection can be achieved. This derivation is quite lengthy and complex and is presented in full in [3]. Here, I shall simply explain the resulting

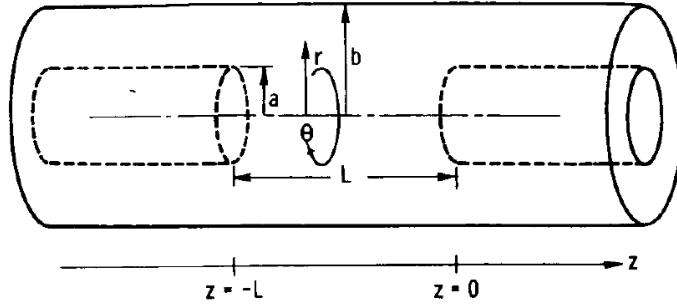


Figure 6.10. The simplified geometry of an open cavity with coaxial terminations. Figure from [3].

5078 conditions on the partition radius for perfect reflection.

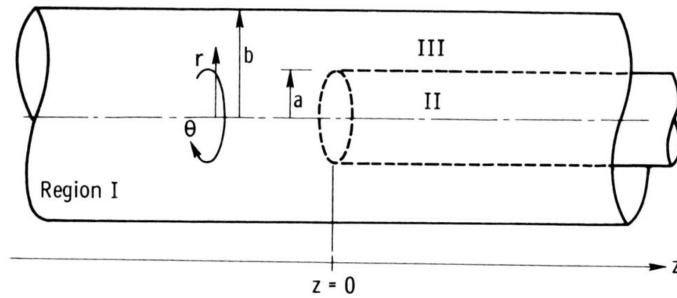


Figure 6.11. Electric field regions for the open cavity boundary value problem. Figure from [3].

5079 The open cavity boundary value problem is solved by expressing the forms of the
 5080 electric fields in the different regions of the cavity and requiring that the electric fields are
 5081 continuous. There are effectively three distinct regions in the open cavity corresponding
 5082 to the central cavity volume, the inner coaxial volume, and the outer coaxial volume (see
 5083 Figure 6.11).

5084 In Region I, the boundary conditions are those of a cylindrical waveguide, and it
 5085 is required that E_ϕ for the TE_{0m} modes go to zero at the cavity wall ($r = b$). This
 5086 necessitates $J'_{0m}(k_{c0m}b) = 0$. A solution for the radius a is desired such that the TE_{01}
 5087 mode propagates, but other TE_{0m} modes are below the cutoff frequency for the circular
 5088 waveguide. This is equivalent to requiring

$$3.832 < k_{c0m}b < 7.016, \quad (6.41)$$

5089 where the numbers 3.832 and 7.016 correspond to the first and second zeros of the Bessel
 5090 function (see Table 6.1).

5091 In Region II the boundary conditions are those of a cylindrical waveguide, but with
 5092 a smaller radius. The condition that $E_\phi = 0$ at the cylindrical partition radius is that
 5093 $J'_{0m}(k_{c0m}a) = 0$. To ensure perfect reflection, all modes in Region 1 of the cavity must be
 5094 below the cutoff frequency of the circular waveguide formed by the inner volume of the
 5095 coaxial terminator. Therefore, solutions where the condition

$$k_{c0m}a < 3.832, \quad (6.42)$$

5096 is true are required.

5097 Finally, in Region III the boundary condition are those of a coaxial waveguide. One
 5098 needs to guarantee that $E_\phi = 0$ at both $r = b$ and $r = a$, which involves finding the
 5099 eigenvalues of the following equation

$$J'_0(k_{c0m}a)Y'_0(k_{c0m}b) - J'_0(k_{c0m}b)Y'_0(k_{c0m}a) = 0, \quad (6.43)$$

5100 where Y'_0 the zeroth-order derivatives of the Bessel function of the second kind. The
 5101 solutions to this equation depend on the value of the ratio b/a . The approximate solution
 5102 is given by

$$\delta_n a \simeq \frac{n\pi}{b/a - 1}, \quad (6.44)$$

5103 where δ_n are eigenvalues of Equation 6.43. Similar to Region II, solutions for which
 5104 the TE₀₁ modes of Region I are below the cutoff frequency of Region III are needed.
 5105 Therefore, it is required that

$$k_{c0m} < \delta_1. \quad (6.45)$$

5106 In general, one has some freedom in specifying the value of b/a . A value typically used
 5107 in practice is $b/a = 2.082$, which corresponds to positioning the radius of the cylindrical
 5108 partition at the maxima of the TE₀₁ electrical fields.

5109 Using the constraints from the three field regions one can develop a coaxial terminator
 5110 that acts as a virtual perfectly conducting surface for the TE₀₁ modes. The only required
 5111 inputs are the desired frequency of the TE₀₁₁ mode and a choice for the value of b/a .

5112 **6.4.2 Mode Filtering**

5113 The general case of an electron coupling to a resonant cavity is complicated. This is
 5114 because cavities contain an infinite number of resonant modes, which for higher order
 5115 modes, have couplings to the electron with a complex spatial dependence. The danger is

5116 that improper modeling of the electron's coupling to the cavity can lead to systematic
5117 errors in the CRES measurements that prevent a high-resolution measurement of the
5118 electron's kinetic energy. This in part drives the preference for a single-mode cavity
5119 experiment that uses only the electron's coupling to the TE₀₁₁ mode to perform CRES,
5120 assuming that sufficient information on the electron's position can be obtained with a
5121 single mode.

5122 The TE₀₁₁ mode is in a region where there are relatively few other modes to which
5123 the electron could couple(see Figure 6.8). However, one can see that the frequency of
5124 the TE₀₁₁ is perfectly degenerate with the TM₁₁₁ mode, which means that electrons will
5125 inevitably couple to both modes if they have the correct cyclotron frequency.

5126 The magnitude of the impact of the electron coupling to both TE₀₁₁ and TM₁₁₁ is
5127 currently unknown. To first order an electron coupling to more both modes will lose more
5128 energy overtime, which can be measured by observing the frequency chirp rate of the
5129 signal. This effect may be small enough to be negligible or simple enough to model that
5130 the cavity can be treated as an effective single-mode cavity. Alternatively, the one could
5131 consider devising a coupling scheme that is sensitive to both the TE₀₁₁ and the TM₁₁₁
5132 modes. By measuring the coupling of the electron to both modes more information on
5133 the position of the electron could be obtained, which could improve the position and
5134 energy resolution of the CRES measurements.

5135 A different approach is the mode filtering approach, which seeks to obtain a single
5136 TE₀₁₁ mode cavity using perturbations to the cavity walls that selectively impede the
5137 TM modes, while leaving the TE modes mostly unperturbed. The type of perturbations
5138 required can be determined by visualizing the surface currents induced in the cavity
5139 walls by each type of mode (see Figure 6.12). By definition, all TM have electric fields
5140 directed along the vertical axis of the cylindrical cavity, which means that perturbations
5141 that impede currents in this direction will modify TM resonances. On the other hand,
5142 the TE₀₁ modes induce azimuthal currents in the cavity walls, therefore, it is possible to
5143 break the degeneracy between TE₀₁ and TM₁₁ using a cavity perturbation that impedes
5144 axial currents, but does not affect the flow of azimuthal currents.

5145 Figure 6.12 shows two cavity design concepts that achieve this selective current
5146 perturbation. The resistive approach inserts a series of thin dielectric rings into the walls
5147 of the cavity that introduces a resistive and capacitive impedance to the longitudinal
5148 currents, while leaving azimuthal current paths intact. Cavities of this type with high
5149 TE₀₁ Q's have also been constructed by tightly wrapping a thin, dielectric coated wire
5150 around a mold to form the cavity wall. An alternative method is to introduce an inductive

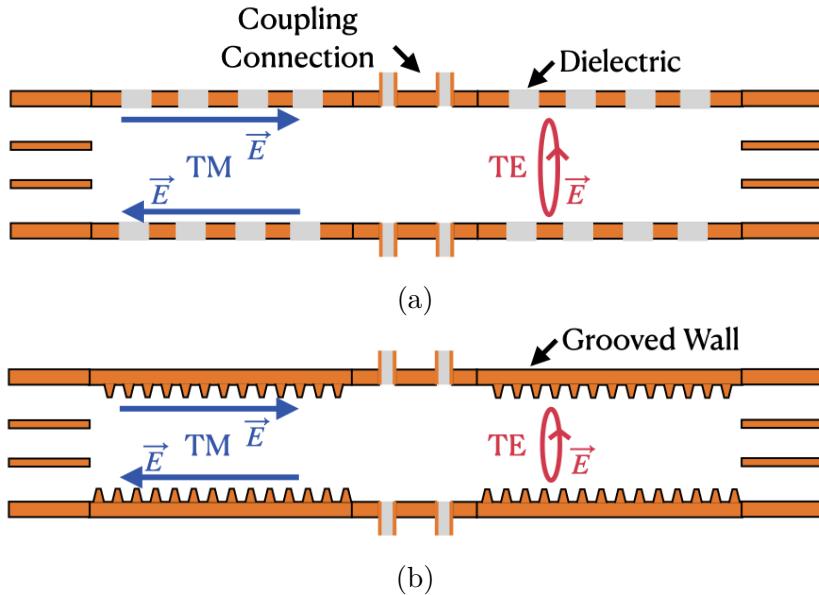


Figure 6.12. Two mode filtering concepts to break the degeneracy of TE_{01} and TM_{11} modes. The resistive approach uses dielectric materials to impede currents that travel vertically along the cavity while leaving azimuthal currents unperturbed. An alternative approach is to impede the currents using grooves cut into the cavity wall, which achieve the same effect with an inductive impedance.

5151 impedance by cutting grooves or a thread pattern on the inside wall of the cavity. For
 5152 reasons of manufacturability and compatibility with tritium the grooved cavity approach
 5153 is the preferred method for mode-filtered cavity construction by Project 8.

5154 **6.4.3 Simulations of Open, Mode-filtered Cavities**

5155 A candidate design for a single TE_{011} mode CRES experiment is a cavity that utilizes
 5156 the coaxial terminations combined with a mode-filtering wall. The first step towards
 5157 validating that a cavity that combines these two design features will operate as expected
 5158 is a thorough simulation effort for which finite element method (FEM) simulation software
 5159 is invaluable. The primary tool for electromagnetic FEM calculations inside Project 8 is
 5160 Ansys HFSS, which has a robust and well-established eigenmode solver that can identify
 5161 the resonant frequencies and associated Q-factors for given structure.

5162 Four variations of a cavity design with a ~ 1 GHz TE_{011} resonance were implemented
 5163 in HFSS (see Figure 6.13). The four designs include a standard cylindrical cavity, an
 5164 open cavity with smooth walls, an open cavity with resistive walls, and an open cavity
 5165 with grooved walls. The relevant design parameters are summarized in Table 6.3. All
 5166 cavities were simulated using copper walls and filled with a vacuum dielectric. The

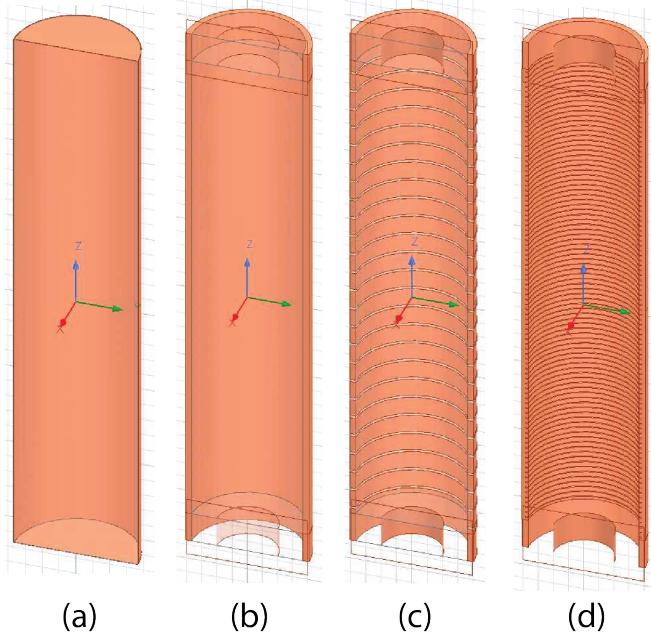


Figure 6.13. Four cavity design variations. (a) is a standard sealed cylindrical cavity, (b) is an open cavity with smooth walls, (c) is an open cavity with resistive walls, and (d) is an open cavity with grooved walls. The main cavity and coaxial terminator parameter are identical for all four cavities.

5167 identities of the resonant modes found by HFSS were validated by visual inspection of
 5168 the electric and magnetic field patterns and by comparison to analytical calculations of
 5169 the mode frequencies.

Table 6.3. A table of cavity design parameters used for HFSS simulations.

Name	Qty.	Unit	Description
D_{cav}	326.4	mm	Cavity diameter
L_{cav}	1668.0	mm	Cavity length
D_{term}	200.2	mm	Inner diameter of coaxial terminator
L_{term}	100.0	mm	Terminator length
l_{die}	8.3	mm	Dielectric spacer thickness
Δl_{die}	66.7	mm	Distance between dielectric spacers
l_{groove}	3.0	mm	Groove height
d_{groove}	9.0	mm	Groove depth
Δl_{groove}	18.3	mm	Distance between grooves

5170 The results of the HFSS simulations validate our predictions of the resonant behavior
 5171 of an open, mode-filtered cavity developed in the preceding sections (see Figure 6.14) One
 5172 can see that for a standard cavity the TE_{01} and the TM_{11} are degenerate in frequency
 5173 with relatively high Q-factors. The open-ended cavity preserves the high Q-factors of

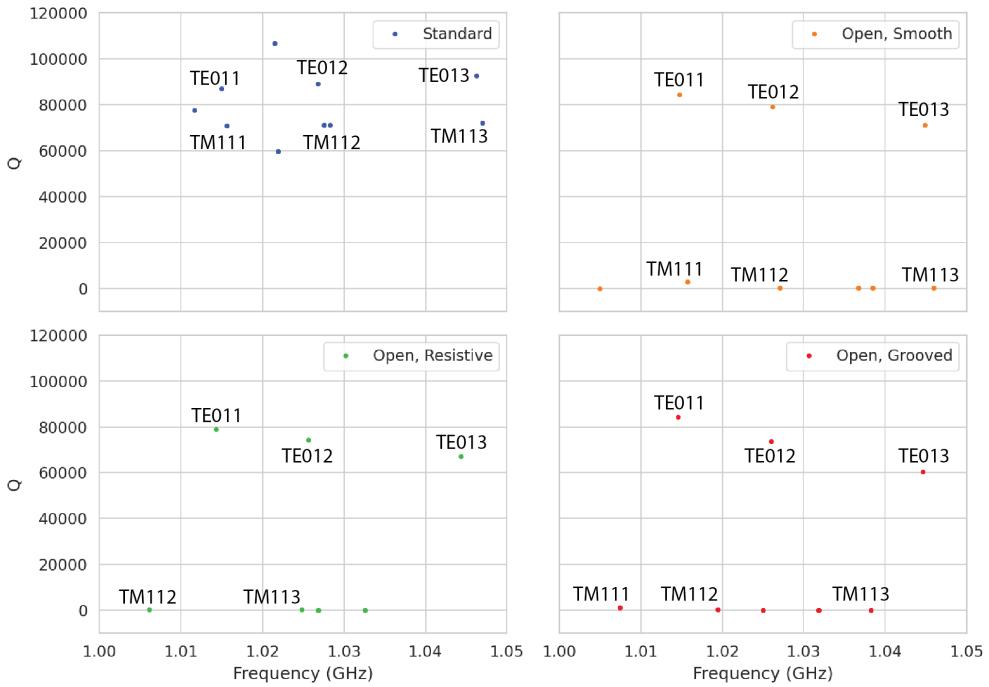


Figure 6.14. The frequencies and Q-factors of the resonant modes identified by HFSS for the cavity variations shown in Figure 6.13. The fully-sealed cavity with smooth walls has several high-Q modes near the TE_{011} resonance. Introducing the open-termination preserves the Q-factors of the $TE_{01\ell}$ modes and suppresses the Q-factors of the modes whose boundary conditions do not match the cylindrical partition. Both the resistive and grooved wall perturbations shift the resonant frequencies of the TM modes away from the TE_{011} mode. By properly tuning the geometry of the grooves or the resistive spacers several MHz of frequency separation can be achieved.

the TE_{01} modes, while the other modes, since their boundary conditions do not match the coaxial geometry, have their Q-factors suppressed. One can see that the effect of the resistive and inductive mode-filtering schemes is to effectively shift the resonant frequencies of the TM_{11} modes below those of the associated TE_{01} modes, which breaks the degeneracy. Optimization of the dielectric spacer or groove parameters can ensure that the TE_{011} mode is isolated from other modes by $O(10)$ MHz, which provides sufficient bandwidth for a measurement of the tritium spectrum endpoint.

Further optimization of the cavity design requires a more detailed cavity simulation that includes the cavity coupling mechanism as well as other geometry modifications required for integration into the magnetic and tritium gas subsystems. Perhaps more important is the development of the capability to simulate the interaction of electrons with the cavity so that simulated CRES signals can be generated using cavities designed for CRES measurements. Simulated CRES signals can then be used to estimate the

5187 neutrino mass sensitivity of the experiment, which allows for the optimization of the cavity
5188 design towards the configuration that provides the best measurement of the neutrino
5189 mass.

5190 **6.5 Single-mode Resonant Cavity Measurements**

5191 Measurement test stands play an important role in the research and development process
5192 that cannot be replaced by simulations. For example, constructing a prototype CRES
5193 cavity forces one to consider important practical issues such as manufacturability and
5194 machine tolerances that may require modifications to the design. Furthermore, by
5195 comparing laboratory measurements of a real cavity to simulations, one can quantify
5196 the impact of imperfections and real-life measurement systematics, which allows for
5197 more accurate sensitivity estimates of the experiment. Lastly, the development of these
5198 prototypes helps to build the necessary experience and expertise within the collaboration
5199 required for more complicated experiments to succeed.

5200 In this spirit a prototype cavity was constructed to demonstrate the open, mode-
5201 filtered cavity concept explored in the previous sections. The primary goal of the
5202 measurements was to validate that an open, mode-filtered cavity suppressed the TM_{11}
5203 modes as predicted by HFSS simulations.

5204 **6.5.1 Cavities and Setup**

5205 Two rudimentary, cavities were constructed using segments of copper pipe available from
5206 McMaster-Carr (see Figure 6.15). The design consists of copper pipes of two diameters.
5207 The larger diameter pipe forms the main cavity wall and the smaller diameter pipe is
5208 used to create a coaxial termination. The diameter of the outer pipe was chosen to
5209 produce a TE_{011} resonance of approximately 6 GHz, while the diameter of the smaller
5210 pipe was selected based on the open termination criteria introduced in Section 6.4.1. The
5211 approximate diameters and lengths of the copper pipe are summarized in Table 6.4.

5212 Coupling to the cavity was achieved using a hand-formable segment of coaxial cable
5213 stripped at one end to form a loop antenna. This was inserted into a small hole located
5214 at the center of the main cavity wall. The coaxial terminators were supported inside the
5215 main cavity by carving a spacer from polystyrene foam (styrofoam) so that they could
5216 be easily inserted into the cavity and repositioned. The dielectric constant of styrofoam
5217 is quite close to air at microwave frequencies so this is expected to have minimal impact

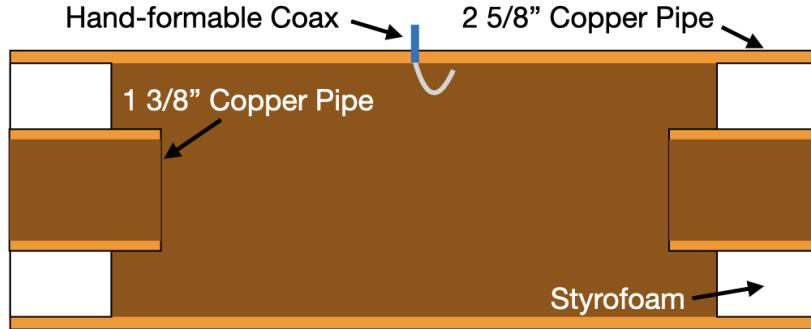


Figure 6.15. A cartoon depicting the design of the open-ended cavity prototype designed to operate at approximately 6 GHz. The main cavity wall was composed of a single copper pipe. A mode-filtered version of this cavity was constructed by

5218 on the resonant properties of the cavity.

Table 6.4. A table of parameters describing the cavity prototypes. Certain values such as the cavity length and the distance between dielectric spacers are approximate due to variation in the machining of the copper. In particular, the filtered cavity was constructed from conducting copper segments that varied in size from 1.50" to 1.85".

Name	Qty.	Unit	Description
D_{cav}	2.625	in	Cavity diameter
L_{cav}	≈ 13	in	Cavity length
D_{term}	1.375	in	Inner diameter of coaxial terminator
L_{term}	1.575	in	Terminator length
l_{die}	0.75	in	Dielectric spacer thickness
Δl_{die}	≈ 1.50 to 1.85	in	Distance between dielectric spacers

5219 The actual length of the cavity is given by the distance between the inner edges of the
 5220 coaxial terminations. The length of the outer section of pipe that forms the main wall of
 5221 the cavity is approximately 16" in length which leads to a cavity length of $\approx 13"$ when
 5222 both terminators are inserted in the cavity. Because the terminators were not rigidly
 5223 mounted this distance is only approximate, however, the uncertain length of the cavity
 5224 will not prevent us from validating the open cavity design.

5225 Along with the smooth-walled open cavity a resistively mode-filtered cavity was
 5226 constructed by creating dielectric spacers out of segments of clear PVC pipe (see Figure
 5227 6.16). The spacers were machined such that the conductive segments of the cavity would
 5228 be separated by 0.75" when the cavity was fully assembled. Due to variations in the
 5229 lengths of the copper segments that make up the cavity wall the distance between spacers
 5230 has significant variation with average value of about 1.7". Eight total spacers were used
 5231 to build the cavity, which when assembled was approximately 16" in total length similar

to the non-filtered cavity.

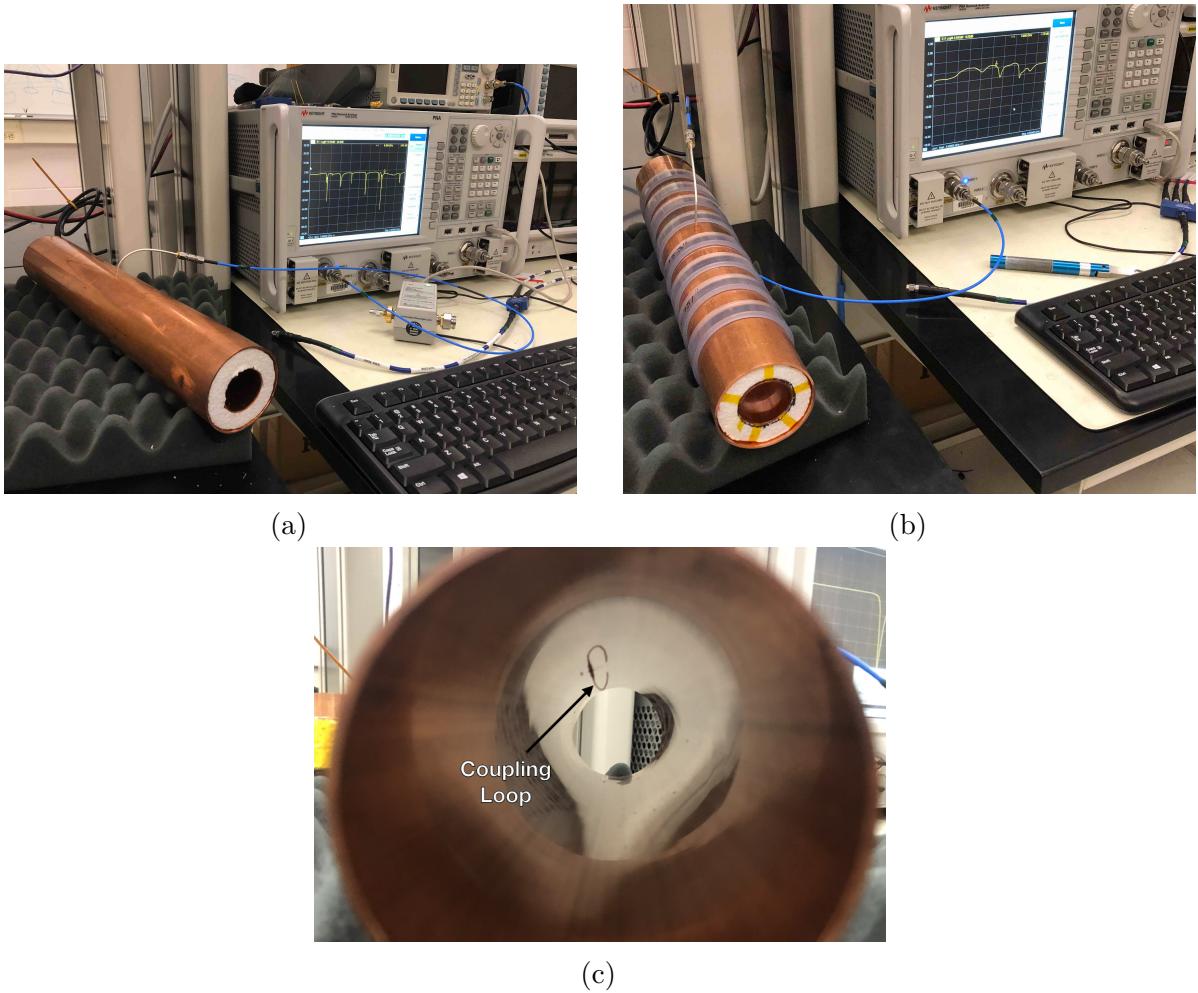


Figure 6.16. Images depicting the measurement of the filtered and non-filtered open cavities using the VNA. The coupling loop in the figure is shown in the TE orientation.

5232

5233

Measurements of both cavities were performed using a VNA connected to the cavity
5234 coupling probe (see Figure 6.16). By measuring the return loss over a range of frequencies
5235 one can measure the frequencies and relative Q-factors of the resonant modes in the
5236 cavity. Due to the opposite polarity of the electric fields for the TE and TM modes,
5237 the loop coupling probe must be rotated 90° to change the polarity of the loop antenna.
5238 When the antenna is oriented such that the loop opening faces the ends of the cavity, it
5239 couples primarily to the TE modes which have magnetic fields directed along the long
5240 axis of the cavity (see Figure 6.16). If the coupling loop is turned by 90° from where
5241 it is shown in the image then it will couple to the TM modes which have azimuthally
5242 directed magnetic fields. In this way both the TE and TM resonances can be measured

5243 independently.

5244 6.5.2 Results and Discussion

5245 The primary analysis for the prototype cavities involved a simple visualization of the
5246 return loss as measured by the VNA and a comparison between the filtered and non-
5247 filtered variations. Since the resonances measured by the VNA are not labeled, there is
5248 an uncertainty about the true identities of the modes measured by the VNA. To resolve
5249 this I performed a simulation of the simplest possible cavity that could be created from
5250 the prototype components, which is a fully open cavity created by removing the coaxial
5251 inserts. The fully-open cavity with the as-built dimensions was simulated in HFSS to get
5252 estimates on the positions of the TE₀₁₁ and TM₁₁₁ modes (see Figure 6.17).

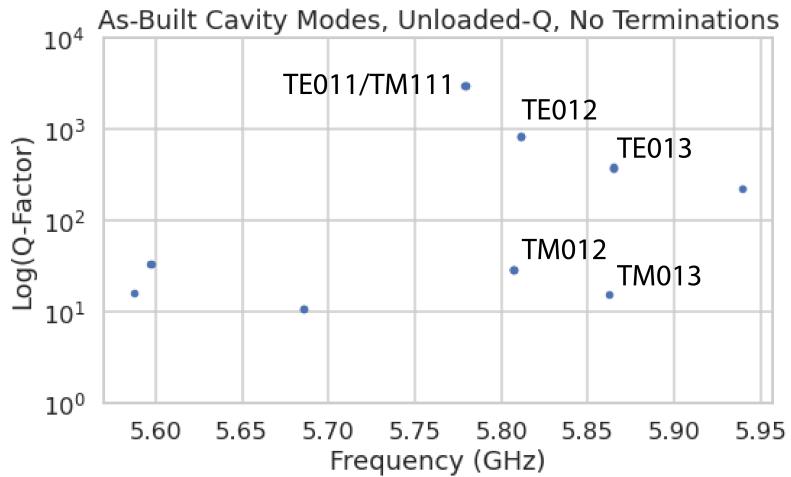


Figure 6.17. HFSS simulation results for a the as-built cavity with the coaxial terminators removed. The TE₀₁₁/TM₁₁₁ frequency is approximately 5.78 GHz.

5253 Simulation of the fully open cavity shows that the TE₀₁₁/TM₁₁₁ modes have a
5254 frequency of approximately 5.78 GHz in the fully open cavity. If the frequency of this
5255 mode is compared to the measurements of the fitered and non-filtered cavities with the
5256 terminators removed one can easily identify the TE₀₁₁ mode at approximately 5.75 GHz
5257 (see Figure 6.18).

5258 Both variations of the non-filtered cavities one sees that the TE₀₁₁ mode is degenerate
5259 in frequency with what appears to be a doublet of TM modes located at the TM₁₁₁
5260 frequency position. This doublet is actually the TM₁₁₁ mode, which has two polarizations
5261 with opposite polarizations. Because the pipe used to construct the cavity is not perfectly

round, the frequency degeneracy between the two polarizations is broken resulting in the doublet peaks.

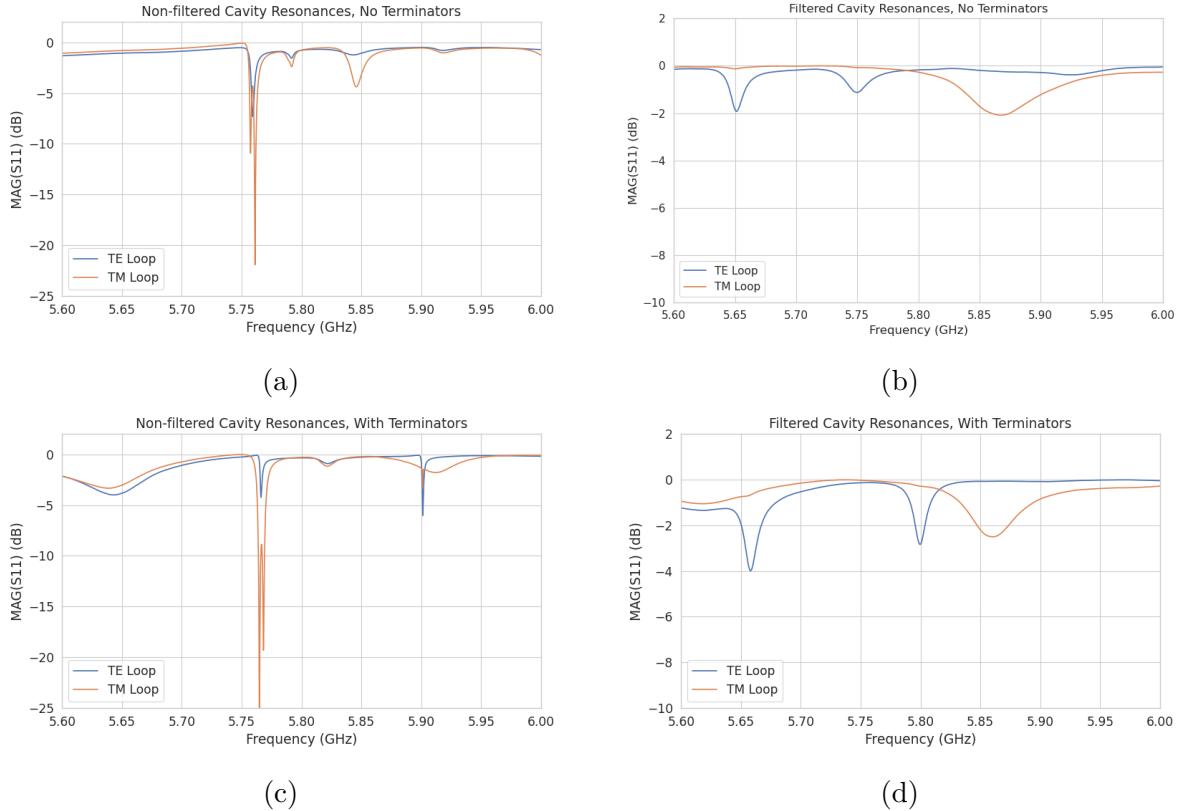


Figure 6.18. Measurements of the filtered and non-filtered prototype cavities acquired with the VNA.

The S-parameter plot for the filtered cavity without terminators has an isolated TE resonance at 5.65 GHz, associated with the TE_{011} mode. The frequency of this mode is lower than the non-filtered cavity due to a difference in the overall lengths of the cavities. An obvious difference between the filtered and non-filtered cavities is that there is no TM_{111} doublet at the TE_{011} frequency. This is what one would expect if the mode-filtering was suppressing the TM modes. There appears to be a noticeable difference in the Q of the TE_{011} resonance between non-filtered and filtered variations as indicated by the increased resonance depth for the filtered cavity. Overall, the Q-factors of the filtered cavity appear significantly smaller than the non-filtered cavity due to the increase in resonance width. This is likely caused by the relatively large widths of the dielectric spacers, which are partially impeding the TE modes.

One can see from these cavity measurements that, in principle, resistive mode-filtering can be used to separate the TE_{011} resonance from the degenerate TM_{111} modes in

5277 combination with the open cavity endcaps. This finding agrees with the expectations
5278 from HFSS, which should provide confidence that the eigenmode solver is correctly
5279 modeling the behavior of the cavity. Although I did not perform a similar study using
5280 a cavity with grooved walls it is expected that the resonant mode structure would be
5281 similar to the cavity studied here.

5282 While this prototype cavity is a good first step, several deficiencies prevent this setup
5283 from providing more than qualitative information to the design of cavities for CRES. This
5284 includes the rudimentary approach to cavity coupling using a stripped coax antenna and
5285 the inability to map the field density in the cavity volume. Improvements in these areas
5286 are required so that measurements from a real cavity can provide useful information to
5287 cavity CRES simulations that will ultimately inform neutrino mass sensitivity estimates.

5288 Future work with prototype cavities must include an improved cavity coupling scheme,
5289 which is robust and compatible with atomic tritium. Since the cavity will ultimately
5290 be filled with atomic tritium, a coupling antenna cannot be used due to the losses of
5291 atomic tritium caused by recombination on the antenna surfaces. Possible non-invasive
5292 coupling schemes include aperture coupling, where the cavity is coupled to an external
5293 waveguide structure through an aperture, or a split-ring coupling approach, where the
5294 center segment of the cylindrical cavity wall is replaced an isolated conductive ring with
5295 a small vertical slit. The aperture coupling approach is a standard coupling scheme [85]
5296 used in a wide range of applications, but at low frequencies the size of the external
5297 waveguide conflicts with design of the atom trapping magnet and cryogenics system.
5298 The split-ring approach could potentially be coupled to a small coaxial transmission line
5299 which is more compatible with the rest of the experiment design. A challenge is achieving
5300 adequate coupling through impedance tuning, which is a focus of current research.

5301 The robustness of the coupling mechanism is relevant due to the difficulty in modeling
5302 its effect on the cavity modes. Small changes in geometry can have a large influence on
5303 the coupling and hence the performance of the cavity, therefore, correctly modeling the
5304 cavity coupling is critical for accurate CRES simulations. Coupling schemes that rely
5305 on connections to coaxial lines are potentially at a disadvantage in this regard due to
5306 the affect of soldering imperfections or unintended bends in the coax on the coupling.
5307 Future work will identify a coupling scheme for the cavity compatible with the neutrino
5308 mass goals of Project 8.

5309 Imperfections in the geometry of a real cavity will necessarily distort the resonant
5310 modes away from simulation predictions. This will change the coupling of an electron
5311 to the cavity and thus change the expected signal structure. Ultimately, this effect will

5312 limit the achievable energy resolution of the experiment unless the differences between
5313 simulation and a real cavity can be sufficiently characterized and calibrated. One possible
5314 approach to this is to utilize a "bead puller" system [99] to strategically perturb the cavity
5315 by moving a conductive bead through the cavity volume. The small perturbation caused
5316 by the bead affects the phase of the cavity resonances proportional to the total magnitude
5317 of the electric field at that position, so by moving the bead through the cavity volume
5318 the total electric field can be mapped and compared to simulation. This information can
5319 provide bounds on the relative perturbations to the cavity mode structure from real-life
5320 imperfections compared to the idealized cavity in HFSS.

5321 **Chapter 7 |**

5322 **Conclusion and Future Prospects**

5323 In this dissertation we have discussed research and development efforts towards the
5324 development of a scalable CRES measurement technology that can be used to build a
5325 CRES experiment at cubic-meter scales with sensitivity to neutrino masses of 40 meV.
5326 The primary contributions of my dissertation are the development and analysis of signal
5327 reconstruction algorithms for an antenna array based CRES experiment [100], which leads
5328 to estimates of the neutrino mass sensitivity; the development of a synthetic cyclotron
5329 radiation antenna (SYNCA) [79], which allowed for laboratory validation of antenna
5330 array CRES simulation models [43]; and the development of an open-ended cavity design
5331 compatible with atomic tritium for a cavity based CRES experiment. A measurable
5332 impact of this work is the transition of the Project 8 collaboration’s experimental plan
5333 from an antenna array based approach to a cavity based approach, where my work played
5334 a key role in demonstrating the significantly higher cost and complexity of the antenna
5335 array experiment.

5336 The transition from antenna arrays to cavities requires a new set of demonstrator
5337 experiments to make incremental progress towards a 40 meV measurement of the neutrino
5338 mass. At the time of writing, the near-term plan of Project 8 is to design and construct a
5339 small-scale cavity CRES experiment utilizing the 1 T magnet installed in the UW-Seattle.
5340 This cavity is designed to have a TE011 resonance with a frequency of about 26 GHz with
5341 a length-to-diameter ratio that mimics the larger cavities intended for the pilot-scale and
5342 Phase IV experiments. The goal of this experiment is to demonstrate cavity CRES as
5343 well as validate models of CRES systematics using electrons from ^{83m}Kr and an electron
5344 gun. Though the primary goal is demonstration, near-term physics measurements are
5345 available in the form of high-resolution measurements of the ^{83m}Kr conversion spectrum
5346 of interest to the KATRIN collaboration.

5347 Furthermore, Project 8 is currently constructing a low-frequency CRES setup located
5348 at Yale University to better understand the principles of cavity based CRES at lower

5349 magnetic fields. The Low, UHF Cavity Krypton Experiment at Yale (LUCKEY) is
5350 a 1.5 GHz cavity CRES experiment the will use conversion electrons from ^{83m}Kr to
5351 perform CRES measurements at the lowest frequencies ever attempted with the technique.
5352 LUCKEY will validate frequency scaling models developed by Project 8 and will pave
5353 the way for the future Low-Frequency Apparatus (LFA), which will be a larger, 1 GHz
5354 cavity CRES experiment that includes a molecular tritium source. The target for the
5355 LFA is a measurement of the neutrino mass with a sensitivity of approximately 0.2 eV,
5356 which will build towards the atomic pilot-scale CRES experiment.

5357 In parallel to the development of cavity CRES is the development of the atomic
5358 tritium source. Recent demonstrations of the production of atomic hydrogen are excellent
5359 steps towards the atomic tritium production needed for the pilot-scale experiment. One
5360 area of future study includes the development of a more detailed understanding of the
5361 efficiency of atomic hydrogen production. Near-term plans include the development of a
5362 magnetic, evaporatively cooled beamline, as well as the prototyping of a Halbach array
5363 atoms trap. Nearly all of the components of the atomic tritium system will require
5364 demonstration before the complete system can be built. The long-term goal of the
5365 atomic tritium work is to construct a full atomic tritium prototype that demonstrates
5366 the production, cooling, trapping, and recycling of tritium at the rates needed for the
5367 pilot-scale experiment.

5368 More broadly, the long-term goal of the Project 8 collaboration is to fully develop
5369 both the atomic tritium and cavity CRES technologies so that both can be combined in
5370 a pilot-scale CRES experiment. It is envisioned that this process will take approximately
5371 10 years for both atomic tritium and cavity CRES. After these developments comes
5372 the pilot-scale experiment which will be the first CRES experiment that simultaneously
5373 demonstrates all the required technologies for Phase IV. Scaling to Phase IV with cavity
5374 CRES will require the construction of multiple copies (approximately 10) of the pilot-scale
5375 experiment to obtain sufficient statistics for 40 meV sensitivity.

5376 Development of the CRES experimental technique by Project 8 has led to new
5377 experiments utilizing the CRES technique for basic physics research, such as the ^6He -
5378 CRES collaboration [101], and has also found applications as a new approach to x-ray
5379 spectroscopy [102]. Recently, a new experimental effort called CRESDA has begun in
5380 the UK to develop new quantum technologies applied to CRES measurements for the
5381 neutrino mass [103]. This flourishing of new experimental efforts based on the CRES
5382 technique is likely to continue as Project 8 continues to develop the technique towards
5383 its neutrino mass measurement goal.

Bibliography

- [1] AGOSTINI, M., G. BENATO, and J. DETWILER (2017) “Discovery probability of next-generation neutrinoless double- β decay experiments,” *Phys. Rev. D*, **96**(5), p. 053001, 1705.02996.
- [2] WENGER, N. C. and J. SMETANA (1972) “Hydrogen Density Measurements Using an Open-Ended Microwave Cavity,” *IEEE Transactions on Instrumentation and Measurement*, **21**(2), pp. 105–114.
- [3] WENGER, N. (1967) “Resonant Frequency of Open-Ended Cylindrical Cavity,” *IEEE Transactions on Microwave Theory and Techniques*, **15**(6), pp. 334–340.
- [4] DAS, A. and T. FERBEL (2003) *Introduction to Nuclear and Particle Physics*, World Scientific.
- [5] CURIE, P. (1898) “Sur une nouvelle substance fortement radioactive, contenue dans la pechblende,” *Comptes rendus de l’Academie des Sciences, Paris*, **127**, pp. 1215–1217.
- [6] RUTHERFORD, E. (1899) “Uranium radiation and the electrical conduction produced by it,” *Philos. Mag.*, **47**, p. 109.
URL <https://cds.cern.ch/record/261896>
- [7] GERWARD, L. (1999) “Paul Villard and his Discovery of Gamma Rays,” *Physics in Perspective*, **1**, pp. 367–383.
- [8] BECQUEREL, H. (1900) “Deviation du rayonnement du radium dans un champ électrique,” *C.R.*, **130**, pp. 809–815.
- [9] D.Sc., E. R. M. and F. S. B.A. (1902) “XLI. The cause and nature of radioactivity.—Part I,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **4**(21), pp. 370–396, <https://doi.org/10.1080/14786440209462856>.
URL <https://doi.org/10.1080/14786440209462856>
- [10] CHADWICK, J. (1914) *Intensitätsverteilung im magnetischen Spektrum der β -Strahlen von Radium B+C*, Druck von Friedr. Vieweg und Sohn.
URL <https://books.google.com/books?id=8s8UmQEACAAJ>

- 5413 [11] BROWN, L. M. (1978) “The idea of the neutrino,” *Physics Today*, **31**(9).
- 5414 [12] PAULI, W. (1978) “Dear radioactive ladies and gentlemen,” *Phys. Today*, **31N9**,
5415 p. 27.
- 5416 [13] AMALDI, E. (1984) “From the discovery of the neutron to the discovery of nuclear
5417 fission,” *Physics Reports*, **111**(1), pp. 1–331.
5418 URL <https://www.sciencedirect.com/science/article/pii/037015738490214X>
- 5420 [14] CHADWICK, J. (1932) “The existence of a neutron,” *Proceedings of the Royal
5421 Society of London. Series A, Containing Papers of a Mathematical and Physical
5422 Character*, **136**(830), pp. 692–708, <https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1932.0112>.
5424 URL <https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1932.0112>
- 5426 [15] WILSON, F. L. (1968) “Fermi’s Theory of Beta Decay,” *American Journal of
5427 Physics*, **36**(12), pp. 1150–1160, https://pubs.aip.org/aapt/ajp/article-pdf/36/12/1150/11891052/1150_1_online.pdf.
5429 URL <https://doi.org/10.1119/1.1974382>
- 5430 [16] CRANE, H. R. and J. HALPERN (1938) “New Experimental Evidence for the
5431 Existence of a Neutrino,” *Phys. Rev.*, **53**, pp. 789–794.
5432 URL <https://link.aps.org/doi/10.1103/PhysRev.53.789>
- 5433 [17] COWAN, C. L., F. REINES, F. B. HARRISON, H. W. KRUSE, and A. D.
5434 MC GUIRE (1956) “Detection of the Free Neutrino: a Confirmation,” *Science*,
5435 **124**(3212), pp. 103–104, <https://www.science.org/doi/pdf/10.1126/science.124.3212.103>.
5437 URL <https://www.science.org/doi/abs/10.1126/science.124.3212.103>
- 5438 [18] ANICIN, I. V. (2005) “The Neutrino - Its Past, Present and Future,” *arXiv e-prints*,
5439 physics/0503172, [physics/0503172](https://arxiv.org/abs/physics/0503172).
- 5440 [19] DAVIS, R. (2003) “Nobel Lecture: A half-century with solar neutrinos,” *Rev. Mod.*
5441 *Phys.*, **75**, pp. 985–994.
- 5442 [20] McDONALD, A. B. ET AL. (2002) “Direct evidence for neutrino flavor transfor-
5443 mation from neutral-current interactions in SNO,” *AIP Conf. Proc.*, **646**(1), pp.
5444 43–58.
- 5445 [21] FUKUDA, Y. ET AL. (1998) “Evidence for oscillation of atmospheric neutrinos,”
5446 *Phys. Rev. Lett.*, **81**, pp. 1562–1567, [hep-ex/9807003](https://arxiv.org/abs/hep-ex/9807003).
- 5447 [22] GIGANTI, C., S. LAVIGNAC, and M. ZITO (2018) “Neutrino oscillations: The rise
5448 of the PMNS paradigm,” *Prog. Part. Nucl. Phys.*, **98**, pp. 1–54, [1710.00715](https://arxiv.org/abs/1710.00715).

- 5449 [23] WORKMAN, R. L. and OTHERS (2022) “Review of Particle Physics,” *PTEP*, **2022**,
 5450 p. 083C01.
- 5451 [24] AN, F. ET AL. (2016) “Neutrino Physics with JUNO,” *J. Phys. G*, **43**(3), p. 030401,
 5452 1507.05613.
- 5453 [25] BIAN, J. ET AL. (2022) “Hyper-Kamiokande Experiment: A Snowmass White
 5454 Paper,” in *Snowmass 2021*, 2203.02029.
- 5455 [26] ABI, B. ET AL. (2020) “Long-baseline neutrino oscillation physics potential of the
 5456 DUNE experiment,” *Eur. Phys. J. C*, **80**(10), p. 978, 2006.16043.
- 5457 [27] ZUBER, K. (2020) *Neutrino Physics*, CRC Press.
- 5458 [28] XING, Z. Z. and S. ZHOU (2011) *Neutrinos in Particle Physics, Astronomy and*
 5459 *Cosmology*, Springer.
- 5460 [29] MOHAPATRA, R. N. ET AL. (2007) “Theory of neutrinos: A White paper,” *Rept.*
 5461 *Prog. Phys.*, **70**, pp. 1757–1867, hep-ph/0510213.
- 5462 [30] DE GOUVÊA, A. (2016) “Neutrino Mass Models,” *Ann. Rev. Nucl. Part. Sci.*, **66**,
 5463 pp. 197–217.
- 5464 [31] WONG, Y. Y. Y. (2011) “Neutrino mass in cosmology: status and prospects,”
 5465 *Ann. Rev. Nucl. Part. Sci.*, **61**, pp. 69–98, 1111.1436.
- 5466 [32] ADE, P. A. R. ET AL. (2016) “Planck 2015 results. XIII. Cosmological parameters,”
 5467 *Astron. Astrophys.*, **594**, p. A13, 1502.01589.
- 5468 [33] CARLSTROM, J. ET AL. (2019) “CMB-S4,” in *Bulletin of the American Astronomical Society*, vol. 51, p. 209, 1908.01062.
- 5469 [34] ELLIOTT, S. R. and P. VOGEL (2002) “Double beta decay,” *Ann. Rev. Nucl. Part. Sci.*, **52**, pp. 115–151, hep-ph/0202264.
- 5470 [35] DOLINSKI, M. J., A. W. P. POON, and W. RODEJOHANN (2019) “Neutrinoless
 5471 Double-Beta Decay: Status and Prospects,” *Ann. Rev. Nucl. Part. Sci.*, **69**, pp.
 5472 219–251, 1902.04097.
- 5473 [36] FORMAGGIO, J. A., A. L. C. DE GOUVÊA, and R. G. H. ROBERTSON (2021)
 5474 “Direct measurements of neutrino mass,” *Physics Reports*, **914**, pp. 1–54, direct
 5475 measurements of neutrino mass.
 5476 URL [https://www.sciencedirect.com/science/article/pii/
 5477 S0370157321000636](https://www.sciencedirect.com/science/article/pii/S0370157321000636)
- 5478 [37] AKER, M. ET AL. (2022) “New Constraint on the Local Relic Neutrino Background
 5479 Overdensity with the First KATRIN Data Runs,” *Phys. Rev. Lett.*, **129**(1), p.
 5480 011806, 2202.04587.

- 5483 [38] FORMAGGIO, J. A., A. L. C. DE GOUVÊA, and R. G. H. ROBERTSON (2021)
 5484 “Direct Measurements of Neutrino Mass,” *Phys. Rept.*, **914**, pp. 1–54, 2102.00594.
- 5485 [39] MONREAL, B. and J. A. FORMAGGIO (2009) “Relativistic cyclotron radiation
 5486 detection of tritium decay electrons as a new technique for measuring the neutrino
 5487 mass,” *Phys. Rev. D*, **80**, p. 051301.
 5488 URL <https://link.aps.org/doi/10.1103/PhysRevD.80.051301>
- 5489 [40] ESFAHANI, A. A. ET AL. (2022) “The Project 8 Neutrino Mass Experiment,”
 5490 2203.07349.
- 5491 [41] ——— (2023) “Cyclotron Radiation Emission Spectroscopy of Electrons from
 5492 Tritium Beta Decay and ^{83m}Kr Internal Conversion,” 2303.12055.
- 5493 [42] ——— (2023) “Tritium Beta Spectrum and Neutrino Mass Limit from Cyclotron
 5494 Radiation Emission Spectroscopy,” 2212.05048.
- 5495 [43] “Antenna Arrays for Physics Measurements with Large-scale CRES Detectors,” *In
 5496 preparation*.
- 5497 [44] NICHOLSON, T. ET AL. (2015) “Systematic evaluation of an atomic clock at 2 x
 5498 10-18 total uncertainty,” *Nature Communications*, **6**(6896).
- 5499 [45] HÄNSCH, T. W. (2006) “Nobel Lecture: Passion for precision,” *Rev. Mod. Phys.*,
 5500 **78**, pp. 1297–1309.
 5501 URL <https://link.aps.org/doi/10.1103/RevModPhys.78.1297>
- 5502 [46] GABOR, D. (1946) “Theory of communication. Part 1: The analysis of information,”
 5503 *Journal of the Institution of Electrical Engineers-part III: radio and communication
 5504 engineering*, **93**(26), pp. 429–441.
- 5505 [47] ESFAHANI ASHTARI, A. ET AL. (2022) “Viterbi decoding of CRES signals in
 5506 Project 8,” *New J. Phys.*, **24**(5), p. 053013, 2112.05265.
- 5507 [48] JACKSON, J. D. (1999) *Classical electrodynamics*, 3rd ed., Wiley, New York, NY.
 5508 URL <http://cdsweb.cern.ch/record/490457>
- 5509 [49] F.R.S., J. L. D. (1897) “LXIII. On the theory of the magnetic influence on
 5510 spectra; and on the radiation from moving ions,” *The London, Edinburgh, and
 5511 Dublin Philosophical Magazine and Journal of Science*, **44**(271), pp. 503–512,
 5512 <https://doi.org/10.1080/14786449708621095>.
 5513 URL <https://doi.org/10.1080/14786449708621095>
- 5514 [50] ESFAHANI, A. A. ET AL. (2022) “The Project 8 Neutrino Mass Experiment,” in
 5515 *2022 Snowmass Summer Study*, 2203.07349.
- 5516 [51] KRAUS, C. ET AL. (2005) “Final results from phase II of the Mainz neutrino mass
 5517 search in tritium beta decay,” *Eur. Phys. J. C*, **40**, pp. 447–468, [hep-ex/0412056](https://arxiv.org/abs/hep-ex/0412056).

- 5518 [52] ASEEV, V. N. ET AL. (2011) “An upper limit on electron antineutrino mass from
 5519 Troitsk experiment,” *Phys. Rev. D*, **84**, p. 112003, 1108.5034.
- 5520 [53] SAENZ, A., S. JONSELL, and P. FROELICH (2000) “Improved Molecular Final-
 5521 State Distribution of HeT+ for the β -Decay Process of T2,” *Phys. Rev. Lett.*, **84**(2),
 5522 p. 242.
- 5523 [54] BODINE, L. I., D. S. PARNO, and R. G. H. ROBERTSON (2015) “Assessment
 5524 of molecular effects on neutrino mass measurements from tritium β decay,” *Phys.
 5525 Rev. C*, **91**(3), p. 035505, 1502.03497.
- 5526 [55] ASNER, D. M. ET AL. (2015) “Single electron detection and spectroscopy via
 5527 relativistic cyclotron radiation,” *Phys. Rev. Lett.*, **114**(16), p. 162501, 1408.5362.
- 5528 [56] VÉNOS, D., J. SENTKERESTIOVÁ, O. DRAGOUN, M. SLEZÁK, M. RYŠAVÝ, and
 5529 A. ŠPALEK (2018) “Properties of 83mKr conversion electrons and their use in the
 5530 KATRIN experiment,” *JINST*, **13**(02), p. T02012.
- 5531 [57] LAGENDIJK, A., I. F. SILVERA, and B. J. VERHAAR (1986) “Spin exchange and
 5532 dipolar relaxation rates in atomic hydrogen: Lifetimes in magnetic traps,” *Phys.
 5533 Rev. B*, **33**, pp. 626–628.
 5534 URL <https://link.aps.org/doi/10.1103/PhysRevB.33.626>
- 5535 [58] SILVERA, I. and J. WALRAVEN (1986) “Spin-polarized atomic hydrogen,” *Prog.
 5536 Low Temp. Phys.*, **X**, pp. 139–370.
- 5537 [59] FURSE, D. ET AL. (2017) “Kassiopeia: a modern, extensible C++ particle tracking
 5538 package,” *New Journal of Physics*, **19**(5), p. 053012.
 5539 URL <https://doi.org/10.1088/1367-2630/aa6950>
- 5540 [60] ESFAHANI, A. A. ET AL. (2019) “Electron radiated power in cyclotron radiation
 5541 emission spectroscopy experiments,” *Phys. Rev. C*, **99**, p. 055501.
 5542 URL <https://link.aps.org/doi/10.1103/PhysRevC.99.055501>
- 5543 [61] ASHTARI ESFAHANI, A. ET AL. (2019) “Locust: C++ software for simulation of
 5544 RF detection,” *New J. Phys.*, **21**, p. 113051, 1907.11124.
- 5545 [62] WIECHERT, E. (1901) “Elektrodynamische Elementargesetze,” .
 5546 URL <https://doi.org/10.1002/andp.19013090403>
- 5547 [63] LIÉARD, A. (1898) “Champ électrique et Magnétique,” *Léclairage électrique*, **16**(27-
 5548 29).
- 5549 [64] BALANIS, C. (2015) *Antenna Theory: Analysis and Design*, Wiley.
 5550 URL <https://books.google.com/books?id=PTFcCwAAQBAJ>
- 5551 [65] <https://www.ansys.com/products/electronics/ansys-hfss>.

- 5552 [66] https://en.wikipedia.org/wiki/Linear_time-invariant_system.
- 5553 [67] NYQUIST, H. (1928) “Certain Topics in Telegraph Transmission Theory,” *Transactions of the American Institute of Electrical Engineers*, **47**(2), pp. 617–644.
- 5555 [68] BRUN, R. and F. RADEMAKERS (1997) “ROOT: An object oriented data analysis
5556 framework,” *Nucl. Instrum. Meth. A*, **389**, pp. 81–86.
- 5557 [69] ASHTARI ESFAHANI, A. ET AL. (2021) “Bayesian analysis of a future β decay
5558 experiment’s sensitivity to neutrino mass scale and ordering,” *Phys. Rev. C*, **103**,
5559 p. 065501.
5560 URL <https://link.aps.org/doi/10.1103/PhysRevC.103.065501>
- 5561 [70] KAY, S. (1998) *Fundamentals of Statistical Signal Processing: Detection Theory,*
5562 *Volume II*, Pearson.
- 5563 [71] NEYMAN, J. and E. PEARSON (1933) “On the problem of the the most efficient
5564 tests of statistical hypotheses,” *Phil. Trans. R. Soc. Lond. A*, **231**.
- 5565 [72] STUMPF, M. (2018) *Electromagnetic Reciprocity in Antenna Theory*, Wiley.
- 5566 [73] BISHOP, C. (2016) *Pattern Recognition and Machine Learning*, Springer.
- 5567 [74] PLEHN, T., A. BUTTER, B. DILLON, and C. KRAUSE (2022) “Modern Machine
5568 Learning for LHC Physicists,” [2211.01421](#).
- 5569 [75] GEORGE, D. and E. A. HUERTA (2018) “Deep Learning for Real-time Gravitational
5570 Wave Detection and Parameter Estimation: Results with Advanced LIGO Data,”
5571 *Phys. Lett. B*, **778**, pp. 64–70, [1711.03121](#).
- 5572 [76] GABBARD, H., C. MESSENGER, I. S. HENG, F. TONOLINI, and R. MURRAY-
5573 SMITH (2022) “Bayesian parameter estimation using conditional variational au-
5574 toencoders for gravitational-wave astronomy,” *Nature Phys.*, **18**(1), pp. 112–117,
5575 [1909.06296](#).
- 5576 [77] REIMANN, R. (2022) “Project 8: R&D for a next-generation neutrino mass experi-
5577 ment,” *PoS*, **PANIC2021**, p. 283.
- 5578 [78] BUZINSKY, N. (2021) *Statistical Signal Processing and Detector Optimization in*
5579 *Project 8*, Ph.D. thesis, Massachusetts Institute of Technology.
- 5580 [79] ESFAHANI, A. A. ET AL. (2023) “SYNCA: A Synthetic Cyclotron Antenna for
5581 the Project 8 Collaboration,” *Journal of Instrumentation*, **18**(01), p. P01034.
5582 URL <https://dx.doi.org/10.1088/1748-0221/18/01/P01034>
- 5583 [80] <https://www.nvidia.com/en-us/data-center/v100/>.
- 5584 [81] <https://www.nvidia.com/en-us/data-center/h100/>.

- 5585 [82] HE, K., X. ZHANG, S. REN, and J. SUN (2016) “Deep Residual Learning for
 5586 Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern*
 5587 *Recognition (CVPR)*, pp. 770–778.
- 5588 [83] SIMONYAN, K. and A. ZISSERMAN (2015) “Very Deep Convolutional Networks
 5589 for Large-Scale Image Recognition,” in *3rd International Conference on Learning*
 5590 *Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference*
 5591 *Track Proceedings* (Y. Bengio and Y. LeCun, eds.).
 5592 URL <http://arxiv.org/abs/1409.1556>
- 5593 [84] FRIIS, H. (1946) “A Note on a Simple Transmission Formula,” *Proceedings of the*
 5594 *IRE*, **34**(5), pp. 254–256.
- 5595 [85] POZAR, D. M. (2005) *Microwave engineering; 3rd ed.*, Wiley, Hoboken, NJ.
 5596 URL <https://cds.cern.ch/record/882338>
- 5597 [86] AKER, M. ET AL. (2022) “Direct neutrino-mass measurement with sub-electronvolt
 5598 sensitivity,” *Nature Physics*, **18**(2), pp. 160–166.
 5599 URL <https://doi.org/10.1038/s41567-021-01463-1>
- 5600 [87] MONREAL, B. and J. A. FORMAGGIO (2009) “Relativistic Cyclotron Radiation
 5601 Detection of Tritium Decay Electrons as a New Technique for Measuring the
 5602 Neutrino Mass,” *Phys. Rev. D*, **80**, p. 051301, 0904.2860.
- 5603 [88] ASHTARI ESFAHANI, A. ET AL. (2017) “Determining the neutrino mass with
 5604 cyclotron radiation emission spectroscopy—Project 8,” *J. Phys. G*, **44**(5), p. 054004,
 5605 1703.02037.
- 5606 [89] ESFAHANI, A. A. ET AL. (2022) “The Project 8 Neutrino Mass Experiment,” in
 5607 *2022 Snowmass Summer Study*, 2203.07349.
- 5608 [90] ASHTARI ESFAHANI, A. ET AL. (2019) “Locust: C++ software for simulation of
 5609 RF detection,” *New Journal of Physics*, **21**(11), p. 113051.
 5610 URL <https://doi.org/10.1088/1367-2630/ab550d>
- 5611 [91] FURSE, D. ET AL. (2017) “Kassiopeia: a modern, extensible C++ particle tracking
 5612 package,” *New Journal of Physics*, **19**(5), p. 053012.
 5613 URL <https://doi.org/10.1088/1367-2630/aa6950>
- 5614 [92] BALANIS, C. (2011) *Modern Antenna Handbook*, Wiley.
 5615 URL <https://books.google.com/books?id=UYpV8L8GNcwc>
- 5616 [93] WIRTH, W. (2001) *Radar Techniques Using Array Antennas*, Institution of Engi-
 5617 neering and Technology.
 5618 URL <https://books.google.com/books?id=ALht42gkzLsC>

- 5619 [94] BROWN, L. S., G. GABRIELSE, K. HELMERSON, and J. TAN (1985) “Cyclotron
 5620 motion in a microwave cavity: Lifetime and frequency shifts,” *Phys. Rev. A*, **32**,
 5621 pp. 3204–3218.
 5622 URL <https://link.aps.org/doi/10.1103/PhysRevA.32.3204>
- 5623 [95] HANNEKE, D., S. FOGWELL HOOGERHEIDE, and G. GABRIELSE (2011) “Cavity
 5624 control of a single-electron quantum cyclotron: Measuring the electron magnetic
 5625 moment,” *Phys. Rev. A*, **83**, p. 052122.
 5626 URL <https://link.aps.org/doi/10.1103/PhysRevA.83.052122>
- 5627 [96] HANNEKE, D. A. (2007) *Cavity control in a single-electron quantum cyclotron: an
 5628 improved measurement of the electron magnetic moment*, Ph.D. thesis, Harvard U.
- 5629 [97] PURCELL, E. (1946) “Spontaneous Emission Probabilities at Radio Frequencies,”
 5630 *Phys. Rev.*, **69**, pp. 674–674.
 5631 URL <https://link.aps.org/doi/10.1103/PhysRev.69.674>
- 5632 [98] MOSKOWITZ, B. E. and J. ROGERS (1988) “ANALYSIS OF A MICROWAVE
 5633 CAVITY DETECTOR COUPLED TO A NOISY AMPLIFIER,” *Nucl. Instrum.
 5634 Meth. A*, **264**, pp. 445–452.
- 5635 [99] RAPIDIS, N. M. (2018) “Application of the Bead Perturbation Technique to a
 5636 Study of a Tunable 5 GHz Annular Cavity,” *Springer Proc. Phys.*, **211**, pp. 45–51,
 5637 1708.04276.
- 5638 [100] “Real-time Signal Detection for Cyclotron Radiation Emission Spectroscopy Mea-
 5639 surements using Antenna Arrays,” *In preparation*.
- 5640 [101] BYRON, W. ET AL. (2022) “First observation of cyclotron radiation from MeV-scale
 5641 e^{pm} following nuclear beta decay,” 2209.02870.
- 5642 [102] KAZKAZ, K. and N. WOOLLETT (2021) “Using Cyclotron Radiation Emission
 5643 for Ultra-high Resolution X-Ray Spectroscopy,” *New J. Phys.*, **23**(3), p. 033043,
 5644 1911.05869.
- 5645 [103] CANNING, J. A. L., F. F. DEPPISCH, and W. PEI (2023) “Sensitivity of future
 5646 tritium decay experiments to New Physics,” *JHEP*, **03**, p. 144, 2212.06106.

5647

Vita

5648

Andrew Douglas Ziegler

5649

Education

5650

- Doctor of Philosophy, Physics, The Pennsylvania State University, University Park, Pennsylvania, USA, 2023
- Bachelor of Science, Physics, The University of Minnesota, Minneapolis, Minnesota, USA, 2017

5654

Selected Publications

5655

- Astari Esfahani, A. et al. (2023) "Antenna Arrays for CRES-based Neutrino Mass Measurement", *Phys. Rev. C*, In preparation.
- Astari Esfahani, A. et al. (2023) "Real-time Signal Detection for Cyclotron Radiation Emission Spectroscopy Measurements using Antenna Arrays", *Journal of Instrumentation*, In preparation.
- Astari Esfahani, A. et al. (2023) "Tritium Beta Spectrum and Neutrino Mass Limit from cyclotron Radiation Emission Spectroscopy", *Phys. Rev. Lett.*, Accepted.
- Astari Esfahani, A. et al. (2022) "SYNCA: A Synthetic Cyclotron Antenna for the Project 8 Collaboration", *Journal of Instrumentation*, **18**(01).

5664

Selected Presentations

5665

- *New Developments in the CRES Technique for Neutrino Mass Measurement*, Invited Talk, Fall 2022 Meeting of the APS Division of Nuclear Physics, New Orleans, Louisiana, USA, 2022
- *Signal Detection Algorithms for Phase III of the Project 8 Experiment*, Contributed Talk, APS April Meeting 2022, New York, New York, 2022
- *Synthetic Electron Antenna for Calibrating the Project 8 Neutrino Mass Experiment*, Contributed Talk, Fall 2021 Meeting of the APS Division of Nuclear Physics, Virtual, 2021