

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

<sup>16</sup> 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

<sup>17</sup> Doug Cowan  
Professor of Physics and Astrophysics

Xingjie Ni  
Associate Professor of Electrical Engineering

<sup>18</sup> Stephanie Wissel  
Associate Professor of Physics

<sup>19</sup> \*Signatures are on file in the Graduate School.

# <sup>20</sup> 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  
25 clues to the energy scale of possible physics beyond the standard model. This thesis  
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	<b>List of Figures</b>	viii
39	<b>List of Tables</b>	xxvii
40	<b>Chapter 1</b>	
41	<b>Introduction</b>	1
42	1.1 Summary . . . . .	1
43	1.2 Outline . . . . .	4
44	<b>Chapter 2</b>	
45	<b>Neutrinos and Neutrino Masses</b>	6
46	2.1 Introduction . . . . .	6
47	2.2 Neutrinos and Beta-decay . . . . .	6
48	2.3 Neutrino Oscillations . . . . .	7
49	2.4 Neutrino Masses in the Standard Model . . . . .	10
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 . . . . .	16
54	<b>Chapter 3</b>	
55	<b>Direct Measurement of the Neutrino Mass with Project 8</b>	21
56	3.1 Introduction . . . . .	21
57	3.2 Cyclotron Radiation Emission Spectroscopy and Project 8 . . . . .	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 . . . . .	30
61	3.3 Phase II: First Tritium Beta Decay Spectrum and Neutrino Mass Mea- surement with CRES . . . . .	33
62	3.3.1 The Phase II CRES Apparatus . . . . .	34
63	3.3.2 CRES Track and Event Reconstruction . . . . .	38
64	3.3.3 Results from Phase II . . . . .	42
65	3.4 Phase III R&D: Antenna Array CRES . . . . .	46
66	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 . . . . .	56
72	3.6 Phase IV . . . . .	58

73 **Chapter 4**

74 **Signal Reconstruction Techniques for Antenna Array CRES and the**

75 **FSCD**

76	4.1 Introduction . . . . .	60
77	4.2 FSCD Simulations . . . . .	61
78	4.2.1 Kassiopeia . . . . .	62
79	4.2.2 Locust . . . . .	66
80	4.2.3 CRESana . . . . .	70
81	4.3 Signal Detection and Reconstruction Techniques for Antenna Array CRES	71
82	4.3.1 Digital Beamforming . . . . .	75
83	4.3.2 Matched Filtering . . . . .	85
84	4.3.3 Machine Learning . . . . .	97
85	4.4 Analysis of Signal Detection Algorithms for the Antenna Array Demonstrator	104
86	4.4.1 Introduction . . . . .	104
87	4.4.2 Signal Detection with Antenna Array CRES . . . . .	106
88	4.4.2.1 Antenna Array and DAQ System . . . . .	106
89	4.4.2.2 Real-time Signal Detection . . . . .	109
90	4.4.3 Signal Detection Algorithms . . . . .	113
91	4.4.3.1 Power Threshold . . . . .	113
92	4.4.3.2 Matched Filtering . . . . .	114
93	4.4.3.3 Machine Learning . . . . .	119
94	4.4.4 Methods . . . . .	120
95	4.4.4.1 Data Generation . . . . .	120
96	4.4.4.2 Template Number and Match Estimation . . . . .	121
97	4.4.4.3 CNN Training and Data Augmentation . . . . .	122
98	4.4.5 Results and Discussion . . . . .	124
99	4.4.5.1 Trigger Classification Performance . . . . .	124
100	4.4.5.2 Computational Cost and Hardware Requirements . . . . .	125
101	4.4.6 Conclusion . . . . .	128

102 **Chapter 5**

103 **Antenna and Antenna Measurement System Development for the**

104 **Project 8 Experiment**

105	5.1 Introduction . . . . .	129
106	5.2 Antenna Measurements for CRES experiments . . . . .	130
107	5.2.1 Antenna Parameters . . . . .	130
108	5.2.1.1 Radiation Patterns . . . . .	130
109	5.2.1.2 Directivity and Gain . . . . .	131

110	5.2.1.3	Far-field and Near-field . . . . .	132
111	5.2.1.4	Polarization . . . . .	133
112	5.2.1.5	Antenna Factor and Effective Aperture . . . . .	134
113	5.2.2	Antenna Measurement Fundamentals . . . . .	136
114	5.2.2.1	Friis Transmission Equation . . . . .	136
115	5.2.2.2	S-Parameters and Network Analyzers . . . . .	137
116	5.2.2.3	Antenna Array Commissioning and Calibration Measurements . . . . .	138
117	5.2.3	The Penn State Antenna Measurement System . . . . .	140
118	5.3	Development of a Synthetic Cyclotron Antenna (SYNCA) for Antenna Array Calibration . . . . .	143
119	5.3.1	Introduction . . . . .	143
120	5.3.2	Cyclotron Radiation Phenomenology . . . . .	146
121	5.3.3	SYNCA Simulations and Design . . . . .	152
122	5.3.4	Characterization of the SYNCA . . . . .	156
123	5.3.5	Beamforming Measurements with the SYNCA . . . . .	160
124	5.3.6	Conclusions . . . . .	163
125	5.4	SYNCA Development Discussion . . . . .	164
126	5.5	FSCD Antenna Array Measurements with the SYNCA . . . . .	167
127	5.5.1	Introduction . . . . .	167
128	5.5.2	Measurement Setups . . . . .	168
129	5.5.2.1	FSCD Array Setup . . . . .	168
130	5.5.2.2	Synthetic Array Setup . . . . .	170
131	5.5.3	Simulations, Analysis, and Results . . . . .	171
132	5.5.3.1	Simulations . . . . .	171
133	5.5.3.2	Phase Analysis . . . . .	172
134	5.5.3.3	Magnitude Analysis . . . . .	177
135	5.5.3.4	Beamforming Characterization . . . . .	179
136	5.5.4	Conclusions . . . . .	182

139	<b>Chapter 6</b>		
140	<b>Development of Resonant Cavities for Large Volume CRES Measurements</b>		184
141	6.1	Introduction . . . . .	184
142	6.2	Cylindrical Resonant Cavities . . . . .	185
143	6.2.1	General Field Solutions . . . . .	185
144	6.2.2	TE and TM Modes . . . . .	186
145	6.2.3	Resonant Frequencies of a Cylindrical Cavity . . . . .	188
146	6.2.4	Cavity Q-factors . . . . .	190
147	6.3	The Cavity Approach to CRES . . . . .	194
148	6.3.1	A Sketch of a Molecular Tritium Cavity CRES Experiment . . . . .	194
149	6.3.2	Magnetic Field, Cavity Geometry, and Resonant Modes . . . . .	196
150	6.3.3	Trade-offs Between the Antenna and Cavity Approaches . . . . .	199
151	6.4	Single-mode Resonant Cavity Design and Simulations . . . . .	201

153	6.4.1	Open Cylindrical Cavities with Coaxial Terminations . . . . .	201
154	6.4.2	Mode Filtering . . . . .	204
155	6.4.3	Simulations of Open, Mode-filtered Cavities . . . . .	206
156	6.5	Single-mode Resonant Cavity Measurements . . . . .	209
157	6.5.1	Cavities and Setup . . . . .	209
158	6.5.2	Results and Discussion . . . . .	212
159	<b>Chapter 7</b>		
160	<b>Conclusion and Future Prospects</b>		<b>216</b>
161	<b>Bibliography</b>		<b>218</b>

# <sup>162</sup> List of Figures

<sup>163</sup> 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. . . . .	8
<sup>169</sup> 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
<sup>171</sup> 2.3	The neutrino mass observable measured by cosmology as a function of the lightest neutrino mass eigenstate. . . . .	14
<sup>173</sup> 2.4	Feynman diagrams for double beta-decay (a) and $0\nu\beta\beta$ (b). . . . .	15
<sup>174</sup> 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
<sup>176</sup> 2.6	A Feynman diagram of beta decay . . . . .	17
<sup>177</sup> 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 $\mu$ , is to distort the shape of the spectrum near the endpoint from the zero-mass spectrum. . . . .	17
<sup>181</sup> 2.8	The tritium beta-decay spectrum. The affect of a massive neutrino on the spectrum is to change it's shape near the endpoint by an amount proportional to the size of the neutrino mass. This suggests that a sufficiently high-statistic and high-resolution measurement of the spectrum endpoint would be able to measure the neutrino mass. . . . .	19

186	3.1	A cartoon illustration of the CRES technique. An electron is contained in a magnetic trap so that it's 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. . . . .	23
190	3.2	An illustration of an electron in a bathtub magnetic trap generated by two well-separated coils. . . . .	24
192	3.3	A plot of the main components of an electron's trajectory in a cylindrically symmetric trap. . . . .	26
194	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. . . . .	27
198	3.5	Neutrino mass exclusion plot including limits from cosmological measure- ments 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. . . . .	30
203	3.6	Sensitivity calculations for a cavity based CRES experiment that demon- strate the neutrino mass measurement goals of the Project 8 collaboration throughout the phased development plan. The blue curves indicate molec- ular 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 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. . .	31
214	3.7	The Phase II CRES apparatus used to perform the first measurement of the tritium beta-decay spectrum using CRES. . . . .	34
216	3.8	Diagram of the CRES cell portion of the Phase II apparatus. . . . .	36
217	3.9	RF system diagram for the Phase II apparatus. . . . .	37
218	3.10	The time-frequency spectrogram of a tritium CRES event in the Phase II apparatus. . . . .	39

220	3.11 The sparse spectrogram obtained by placing a power cut on the raw 221 spectrogram shown in Figure 3.10. . . . .	41
222	3.12 Fits to the measured 17.8-keV $^{83m}\text{Kr}$ conversion line using the deep and 223 shallow trap configurations. . . . .	43
224	3.13 Measurements of the 17.8-keV $^{83m}\text{Kr}$ line using the deep trap configuration 225 for different values of the magnetic field from the field shifting solenoid. . . . .	44
226	3.14 The measured tritium spectrum from Phase II with Bayesian and frequen- 227 tist fits. . . . .	45
228	3.15 A cartoon illustration of the basics of the antenna array CREES technique. . . . .	47
229	3.16 An image of the MRI magnet installed in the Project 8 laboratory at the 230 University of Washington, Seattle. . . . .	49
231	3.17 (a) A model of the FSCD antenna array, magnetic trap, and tritium 232 containment vessel design.(b) A more detailed model of a prototype 233 design for the 5-slot waveguide antenna design. . . . .	50
234	3.18 (a) A plot of the decay rate for the two-body dipolar spin exchange 235 interaction for cc and dd state. (b) A plot of the decay rate of the dipolar 236 spin exchange interaction for d+d states as a function of magnetic field 237 magnitude. Lowering the magnetic field is key for reducing the losses 238 from this interaction. . . . .	55
239	3.19 A conceptual sketch of a large-volume antenna array based CREES experi- 240 ment to measure the neutrino mass. . . . .	56
241	3.20 A conceptual sketch of a pilot-scale cavity CREES experiment with an 242 atomic tritium beamline. . . . .	57

243	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
255	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
260	4.2 A map of the average $\nabla B$ -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
264	4.3 The receiver chain used by Locust when simulating CRES events in the FSCD. . . . .	68
266	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
271	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. . . . .	74

279	4.6 An example ROC curve formed by computing the $P_{\text{FP}}$ and the $P_{\text{TP}}$ for 280 a given likelihood ratio test. As the decision threshold is increased $P_{\text{FP}}$ 281 decreases at the expense of a lower $P_{\text{TP}}$ . The black dashed line indicates 282 the lower bound ROC curve obtained by randomly deciding between $\mathcal{H}_0$ 283 and $\mathcal{H}_1$ . . . . .	75
284	4.7 An illustration of the constructive interference condition which is the 285 operating principle of digital beamforming using a uniform linear array 286 as an example. . . . .	76
287	4.8 System level diagram of the laboratory setup used for beamforming 288 demonstrations at Penn State. For more information on this system see 289 Chapter 5. Signals near 26 GHz are fed to a dipole antenna using an 290 arbitrary waveform generator (AWG) and vector network analyzer (VNA), 291 which drive a mixer. The dipole radiation is collected by an array of 292 antennas connected to the digitizer data acquisition (DAQ) system. . . .	78
293	4.9 Photographs of the beamforming demonstration setup. In (a) I show a 294 top-down view of the dipole antenna and the array of eight horn antennas. 295 Manual repositioning of the horn antennas allows one to synthesize a full- 296 circular antenna array. The dipole antenna is mounted on a camera tripod 297 mount that allows for manual position tuning. (b) is a close up image of 298 the dipole, which is manufactured from two segments of semi-rigid coaxial 299 cable. (c) is another image of the dipole and array. . . . .	79
300	4.10 An example of digital beamforming reconstruction of a dipole antenna 301 using a synthetic array of horn antennas. The beamforming image on 302 the right is constructed by computing the time-averaged power of the 303 summed signals for a two-dimensional grid of beamforming positions. In 304 the image one can see a clear maximum that corresponds to the position 305 of the dipole antenna. On the left I show the frequency spectrum of the 306 time-series at the maximum power pixel. White gaussian noise is added 307 to the signal to mimic a more realistic signal-to-noise-ratio. The signal 308 emitted by the dipole is clearly visible as the high power peak in the 309 frequency spectrum. . . . .	79
310	4.11 Beamforming images visualizing the reconstruction of an electron without 311 (a) and with (b) the cyclotron phase correction. The images were generated 312 using data from Locust simulations. The cyclotron phase refers to a phase 313 offset equal to the relative azimuthal position of an antenna in the array. 314 This phase offset is caused by the circular electron orbit and must be 315 corrected for during reconstruction. . . . .	81

316	4.12 Beamforming images visualizing the reconstruction of an electron located 317 off the central axis of the FSCD trap. In (a) we performing beamforming 318 without the $\nabla B$ -drift correction, and in (b) we include the $\nabla B$ -drift 319 correction. . . . .	82
320	4.13 A plot of a typical frequency spectrum obtained by applying a Fourier 321 transform to the time-series obtained from beamforming. The frequency 322 spectra are plotted without noise on top of an example of a typical noise 323 spectrum to visualize a realistic signal-to-noise ratio. In the example we 324 see that without beamforming it would not be possible to detect anything 325 since the signal amplitudes would be reduced by a factor of sixty relative 326 to the noise. Additionally, we see that the $\nabla B$ -drift correction is needed 327 to detect this electron since it comes from a simulation of an electron with 328 a significant off-axis position. . . . .	83
329	4.14 A plot of the total signal power received by the FSCD array from trapped 330 electrons with different radial positions and pitch angles generated using 331 Locust simulations. The lines on the plot indicate a 10 dB detection 332 threshold above the mean value of the noise in the frequency spectrum. 333 With static beamforming electrons with radial positions larger than about 334 two centimeters are undetectable due to the change in the electron's 335 position over time causing losses from beamforming phase mismatch. This 336 is corrected by including $\nabla B$ -drift frequencies in the beamforming phases. 337 Both beamforming techniques fail to detect electrons below $\approx 88.0^\circ$ , since 338 these signal are composed of several relatively weak sidebands that are 339 comparable to the noise. . . . .	84
340	4.15 Example of a convolution of a CRES signal template with a segment of 341 noisy data. A simulated CRES signal was simulated using Locust and 342 normalized to create a matched filter template. When this template is 343 convolved with noisy data the contains the matching signal the con- 344 volution output increases dramatically compared to data with only noise. 345 The decreasing convolution output as the time offset of the convolution 346 increases is caused by zero-padding of the data and template. . . . .	87



385	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. . . . .	97
392	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. . . . .	98
404	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. . . . .	100
409	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. . . . .	102
415	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. . . . .	103
420	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. . . . .	103

423	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 1 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. . . . .	106
435	4.28 A high-level diagram of the DAQ system architecture envisioned for the FSCD. . . . .	107
437	4.29 A block diagram illustration of the real-time triggering algorithm proposed for antenna array CRES reconstruction. . . . .	108
439	4.30 An illustration of the digital beamforming procedure. The blue lines indicate the various distances from the beamforming position to the antenna. In the situation depicted the actual position of the electron matches the beamforming position, so we should expect constructive interference when the phase shifted signals are summed. To prevent the electron's $\nabla B$ -motion from moving the electron off of the beamforming position, the beamforming phase include a time-dependence to follow the trajectory of the electron in the magnetic trap. . . . .	109
447	4.31 Frequency spectra of simulated CRES signals post-beamforming. The signal of a 90° electron consists of a single frequency component that is easy to detect with a power threshold on the frequency spectrum. This power threshold is still 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°, tend to be dominated by sidebands such that no single frequency component can be reliably distinguished from the noise with a power threshold. . . . .	111

456	4.32 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. . . . .	114
462	4.33 Plots of the estimated PDFs for the matched filter template bank test 463 statistic for CRES signals with various pitch angles as well as the estimated 464 PDF for the noise only signal case. We assume an estimated number 465 of templates of $10^5$ and perfect match between signal and template i.e. 466 $\Gamma_{\text{best}} = 1$ . The mean PDF includes signals ranging from 85.5 – 88.5° in 467 pitch angle. There is a much larger distinction between the signal PDFs 468 at small pitch angle compared to the power threshold indicating a higher 469 detection efficiency for these signals. . . . .	118
470	4.34 The mean match of the matched filter template bank to a test set of 471 randomly parameterized signals as a function of the number or density of 472 templates. The parameter space includes pitch angles from 85.5 – 88.5° 473 and energies from 18575 – 18580 eV. . . . .	122
474	4.35 Histograms of the trained CNN model output from the test dataset. The 475 blue histogram shows the model outputs for signal data. The oddly shaped 476 peak near the end is the result of the softmax function mapping the long 477 tail of the raw output distribution to the range [0, 1]. . . . .	123
478	4.36 ROC curves describing the detection efficiency or true positive rates for 479 the three signal classification algorithms examined in this paper. . . . .	124
480	5.1 An example radiation pattern generated using HFSS simulations. The 481 color and radial distance of the surface from the origin indicate the 482 relative magnitude of radiation power emitted by the antenna in that 483 direction. The primary goal of most antenna measurements is typically to 484 measure the antenna pattern, which is used to derive many useful antenna 485 performance parameters. . . . .	130

486	5.2 An illustration of the three field regions important for the analysis of 487 an antenna system. Very close to the antenna the electric fields are 488 primarily reactive so there is no radiation. If a receiving antenna were 489 placed in this region most of the energy would be reflected back to the 490 transmitter. Outside of the reactive near-field is the radiative near field. 491 At these distances the antenna does radiate, but the radiation pattern is 492 not well-defined since it changes based on the distance of the receiving 493 antenna. It is only in the far-field region where the radiation pattern 494 becomes constant as a function of distance, which is where the majority 495 of antenna engineering is assumed to take place. The antenna arrays 496 developed by Project 8 for CRES measurements operate in the radiative 497 near-field due to the importance of limiting power loss from free-space 498 propagation, which complicates the design of the antenna system. . . . .	132
499	5.3 An illustration of the Friis measurement technique commonly used for 500 antenna characterization measurements. . . . .	136
501	5.4 Illustration of a two-port S-parameter measurement setup. S-parameters 502 characterize how incoming waves of voltage or power scatter off of the RF 503 device under test. This allows you to measure important properties of the 504 device. In particular, we can use this framework to model a two antenna 505 radiation pattern measurement, which we can then automate using a VNA.	138
506	5.5 Two measurement approaches to characterizing an antenna array for CRES 507 measurements. The full-array approach (a) requires a complete antenna 508 array with all the associated hardware. The synthetic array approach 509 (b) utilizes a single antenna and a set of rotation/translation stages to 510 reposition the transmitter or the receiving antenna to synthesize the signals 511 that would be received by the full-array. This approach reduces the cost 512 and complexity of array measurements. A down-side of the synthetic 513 array approach is that multi-channel effects such as reflections cannot 514 be measured. Utilizing both the full-array and the synthetic array is a 515 powerful way to quantify the impact of errors from the multi-channel array.	139
516	5.6 Illustration of the antenna measurement system developed for the Project 517 8 Collaboration. The reference and test antennas can be connected to 518 different data acquisition configurations depending on the measurement 519 goals. The reference antenna is typically a standard horn antenna and 520 the test antenna is mounted on a set of translation stages for positioning. 521 Automated translation stages allow for relatively painless data-taking enabling 522 synthetic antenna array measurements using only a single receiving 523 antenna. Anechoic form designed to mitigate RF reflections surrounds 524 the setup. . . . .	140

525	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. . . . .	141
532	5.8	A sketch of an antenna array large-volume CRES experiment. Electrons from $\beta$ -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 $\beta$ -decay spectrum. . . . .	144
538	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. . . . .	145
544	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. . . . .	146
547	5.11	A plot of the numeric solution to Equation 5.31. 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. . . . .	149
553	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). . . . .	151

560	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. . . . .	152
565	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. . . . .	153
574	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. . . . .	154
581	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 ( $\theta$ ). (a) Shows the total electric field, (b) shows the $\phi$ -polarized electric field component, and (c) shows the $\theta$ -polarized electric field component. These images were generated using Locust simulations for the electron and ANSYS HFSS for both antennas. . . . .	155
587	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 ( $\phi$ ) evaluated at $\theta = 90^\circ$ . This image was generated using Locust simulations for the electron and ANSYS HFSS for both antennas. . . . .	156

591	5.18 (a) A cartoon schematic which highlights the routing of the semi-rigid 592 coax transmission lines. (b) A photograph of a SYNCA constructed 593 using the modified crossed-dipole PCB antenna design. Visible in the 594 photograph of the SYNCA are four blobs of solder which are an artifact 595 of the SYNCA's hand-soldered construction. These solder blobs are the 596 most significant deviation from the SYNCA design shown in Figure 5.14 597 and are responsible for a significant fraction of the irregularities seen in 598 the antenna pattern. . . . .	157
599	5.19 A schematic of the VNA characterization measurements (a). This setup 600 allows for antenna gain and phase measurements across a full $360^\circ$ of 601 azimuthal angles using a motorized rotation stage and control of the radial 602 position of the SYNCA using a translation stage. A photo of the setup in 603 the lab is shown in (b). . . . .	158
604	5.20 Linear interpolations of the measured electric field magnitude (a) and 605 phase (b). The data was acquired using a VNA at 120 points spaced 606 by 3 degrees from 0 to 357 degrees of azimuthal angle. The different 607 color lines indicate the vertical offset of the horn antenna relative to the 608 SYNCA PCB and the dashed line shows the expected shape from electron 609 simulations. No significant difference in the antenna pattern is observed 610 for the measured vertical offsets. . . . .	159
611	5.21 (a) A depiction of the relative phase differences for signals received by a 612 circular antenna array from an isotropic source. The phases correspond to 613 a unique spatial position. (b) A schematic of the setup used to perform 614 digital beamforming. . . . .	160
615	5.22 Digital beamforming maps generated using a simulated 60 channel array 616 and electron simulated using the Locust package. (a) and (b) show the 617 beamforming maps for simulated electrons without the cyclotron spiral 618 phases and with the cyclotron spiral phases respectively. (c) and (d) 619 show the beamforming maps produced from SYNCA measurements. We 620 observe good agreement between simulated electrons and the SYNCA 621 measurements. . . . .	161

622	5.23 A plot of the SYNCA's reconstructed position using the synthesized horn-	
623	antenna array and digital beamforming. (a) Shows the reconstructed	
624	position of the SYNCA compared with the target position indicated by	
625	the positioning system readout. (b) Shows the reconstruction error, which	
626	is the difference between the target and reconstructed positions. The error	
627	bars in (b) are the uncertainty in the mean position of the 2D Gaussian	
628	used to fit the digital beamforming reconstruction peak obtained from	
629	the fit covariance matrix. The mean fit position uncertainty of 0.02 mm	
630	is an order of magnitude smaller than the typical reconstruction error of	
631	0.3 mm obtained by calculating the standard deviation of the difference	
632	between the reconstructed and target position. . . . .	163
633	5.24 Images of an early prototype crossed-dipole antenna manufactured by	
634	hand and the first measurement setup. The antenna was constructed	
635	by hand using four stripped coaxial cables. The antenna was connected	
636	to one port of the VNA, and the remaining three ports on the VNA	
637	were connected to horn antenna arranged with 90 deg offsets around the	
638	crossed-dipole. The measured unwrapped S-parameter phases exhibit the	
639	desired relative phase behavior for a SYNCA. These early measurements	
640	were the first laboratory proof-of-principle for the crossed-dipole SYNCA.	165
641	5.25 An early iteration of a crossed-dipole SYNCA antenna simulated in HFSS.	
642	The antenna is electrically small at 26 GHz, which requires dipole arms on	
643	the order of 1 mm long. This design is limited by the minimum achievable	
644	distance between the dipole arms caused by the available diameters of	
645	coaxial cables. The assumed termination scheme for the coaxial cables to	
646	the antenna is hand-soldering, which introduces random variation in the	
647	antenna pattern from the inevitable blobs of solder left on the surface of	
648	the PCB. . . . .	166
649	5.26 A diagram of the array measurement system used to test the prototype	
650	FSCD antenna array. A VNA is used as the primary measurement tool,	
651	which is connected to the array through a series of switches. The other port	
652	of the VNA connects to the SYNCA through the quad-balun chain used	
653	to provide the SYNCA feed signals. During measurements the SYNCA	
654	is positioned inside the center of the antenna array and translated to	
655	different radial and axial positions using a 3-axis manual translation stage	
656	setup. . . . .	167
657	5.27 Photos of the prototype FSCD antenna (a), the FSCD array and SYNCA	
658	(b), and the translation stages and coordinate system used to position the	
659	SYNCA (c). . . . .	169

660	5.28 A photo of the FSCD antenna and the SYNCA in the synthetic array 661 measurement setup at Penn State. . . . .	171
662	5.29 The unwrapped phases of signals received by the FSCD antenna array 663 from an electron with a 90° pitch angle located in the plane of the antenna 664 array. The data points indicated the phases extracted from simulation 665 and the dashed lines show the model predictions. . . . .	172
666	5.30 Plots of the measured unwrapped phases from the FSCD array (a) and the 667 synthetic array (b) compared to the model predictions for a series of radial 668 positions. The different phases of the sinusoidal phase oscillations in the 669 two plots reflects differences in the coordinate systems of the measurements.	174
670	5.31 The phase errors between the measurement and model for the synthetic 671 array (blue) and the FSCD array (orange) for a series of radial positions. 672 The label JUGAAD refers to an alternative name for the FSCD array 673 setup. As the SYNCA is translated off-axis phase errors with progressively 674 higher oscillation frequency enter into the measurements. . . . .	175
675	5.32 Two dimensional plots of the phase errors for the synthetic array (a) 676 and the FSCD (JUGAAD) array (b). In both plots we observe evidence 677 of a similar diffraction pattern with bilateral symmetry, but the FSCD 678 array measurements have an additional phase error contribution from the 679 different antennas and paths through the switch network. . . . .	176
680	5.33 The amplitude of the signals from CRESana for the FSCD array from a 681 90° electron. As the electron is moved from $R = 0$ the signals begin to 682 have unequal amplitudes depending on the distance from the electron to 683 the antenna. . . . .	177
684	5.34 The normalized magnitudes of the S21 parameters measured in the FSCD 685 (orange) and synthetic array (blue) setups. The dominant observed 686 behavior as a function of radius is the increase in the number of magnitude 687 peaks, which was noted in the phase error curves. There does not appear 688 to be a strong change in the relative amplitude of a group of antennas as 689 predicted by CRESana. . . . .	178
690	5.35 . . . . .	179

691	5.36 Beamforming images from the synthetic array (a) and FSCD array (b) 692      setups with the SYNCA positioned 15 mm off the central axis. In both 693      images we see a clear maxima that corresponds to the true SYNCA 694      position. However, in the FSCD array there is an additional faint peak 695      located at the opposite position of the beamforming maximum. This 696      additional peak is the mirror of the true peak and is the result of reflections 697      between antennas in the FSCD array. . . . .	180
698	5.37 A comparison of the maximum signal amplitude obtained by beamforming 699      to the signal amplitude obtained with an ideal summation as a function 700      of the radial position of the SYNCA. The amplitudes for the synthetic 701      array are shown in (a) and the FSCD array are shown in (b). In both 702      setups we observe that the signal amplitudes obtained from beamforming 703      are smaller than the signal amplitude that could be attained with the 704      ideal summation without phase mismatch. . . . .	181
705	5.38 The ratio of the beamforming signal amplitude to the ideal signal amplitude 706      for the FSCD and synthetic arrays. We see that the FSCD array has 707      a larger power loss from phase error compare to the synthetic array which 708      indicates that calibration errors associated with the multiple channels as 709      well as reflections are impacting the signal reconstruction. . . . .	182
710	6.1 Geometry of a cylindrical waveguide with radius $b$ . . . . .	185
711	6.2 The geometry of a cylindrical cavity with length $L$ and radius $b$ . . . . .	188
712	6.3 Relation of mode frequency to cavity length for a cylindrical cavity with 713      a radius of 18.32 cm. . . . .	189
714	6.4 A series RLC circuit. . . . .	190
715	6.5 Illustration of the behavior of the input impedance of the series RLC 716      circuit as a function of the driving frequency. The BW is proportion to 717      the width of the resonance, which is inversely proportional to Q. . . . .	192
718	6.6 A series RLC circuit coupled to an external circuit with input impedance 719 $R_L$ . . . . .	193

720	6.7	A cartoon depiction of a cavity CRES experiment. A metallic cavity filled 721 with tritium gas is inserted into a uniform background magnetic field 722 to perform CRES measurements. Electrons from beta-decays inside the 723 cavity can be trapped and used to excite a resonant mode(s). By coupling 724 to the cavity mode with a suitable probe one can measure the cyclotron 725 frequency of the electron and perform CRES. . . . .	194
726	6.8	Examples of the resonant mode frequencies of a cylindrical cavity. This 727 cavity has a radius of 18.32 cm and a length to diameter ratio of 4.55. . .	198
728	6.9	An image of an open cavity with coaxial terminations used for dielectric 729 constant measurements. Figure from [2]. . . . .	202
730	6.10	The simplified geometry of an open cavity with coaxial terminations. 731 Figure from [3]. . . . .	203
732	6.11	Electric field regions for the open cavity boundary value problem. Figure 733 from [3]. . . . .	203
734	6.12	Two mode filtering concepts to break the degeneracy of $TE_{01}$ and $TM_{11}$ 735 modes. The resistive approach uses dielectric materials to impede currents 736 that travel vertically along the cavity while leaving azimuthal currents 737 unperturbed. An alternative approach is to impede the currents using 738 grooves cut into the cavity wall, which achieve the same effect with an 739 inductive impedance. . . . .	206
740	6.13	Four cavity design variations. (a) is a standard sealed cylindrical cavity, 741 (b) is an open cavity with smooth walls, (c) is an open cavity with resistive 742 walls, and (d) is an open cavity with grooved walls. The main cavity and 743 coaxial terminator parameter are identical for all four cavities. . . . .	207
744	6.14	The frequencies and Q-factors of the resonant modes identified by HFSS 745 for the cavity variations shown in Figure 6.13. The fully-sealed cavity with 746 smooth walls has several high-Q modes near the $TE_{011}$ resonance. Intro- 747 ducing the open-termination preserves the Q-factors of the $TE_{01\ell}$ modes 748 and suppresses the Q-factors of the modes whose boundary conditions do 749 not match the cylindrical partition. Both the resistive and grooved wall 750 perturbations shift the resonant frequencies of the TM modes away from 751 the $TE_{011}$ mode. By properly tuning the geometry of the grooves or the 752 resistive spacers several MHz of frequency separation can be achieved. . .	208

753	6.15 A cartoon depicting the design of the open-ended cavity prototype designed 754 to operate at approximately 6 GHz. The main cavity wall was composed of 755 a single copper pipe. A mode-filtered version of this cavity was constructed 756 by . . . . .	210
757	6.16 Images depicting the measurement of the filtered and non-filtered open 758 cavities using the VNA. The coupling loop in the figure is shown in the 759 TE orientation. . . . .	211
760	6.17 HFSS simulation results for a the as-built cavity with the coaxial termi- 761 nators removed. The $TE_{011}/TM_{111}$ frequency is approximately 5.78 GHz. . . . .	212
762	6.18 Measurements of the filtered and non-filtered prototype cavities acquired 763 with the VNA. . . . .	213

# List of Tables

764	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. . . . .	120
765	6.1	A table of the values of $p'_{nm}$ . . . . .	187
766	6.2	A table of the values of $p_{nm}$ . . . . .	188
767	6.3	A table of cavity design parameters used for HFSS simulations. . . . .	207
768	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". . . . .	210
770			
771			
772			
773			
774			
775			
776			
777			
778			
779			

780 **Chapter 1** |  
781 **Introduction**

782 **1.1 Summary**

783 Neutrinos are one of the fundamental particles that comprise the standard model of  
784 particle physics and account for a significant fraction of the matter in the universe.  
785 Neutrinos are the most abundant fermions in the universe, but due to their weak  
786 interactions neutrinos seldom interact with other particles. Regardless, neutrinos play a  
787 unique role in the evolution of the early-universe, therefore, a detailed understanding  
788 of the properties of the neutrino is important to understanding the cosmology of the  
789 universe as well as understanding the universe at the fundamental particle physics scale.

790 Unlike other fermions it was unclear that neutrinos had nonzero mass until neutrino  
791 flavor oscillations were definitively observed in the late 90's and early 00's. Flavor  
792 oscillations require that neutrinos experience time so that when acted upon by the  
793 time-evolution operator the initial neutrino state can evolve to a new flavor state. This  
794 implies that the neutrino flavor states are really a superposition of at least three separate  
795 neutrino states with well-defined masses. Measurements of neutrino oscillations that have  
796 taken place over the past couple of decades have measured the differences between  
797 neutrino mass eigenstates with increasing precision. However, oscillation measurements  
798 cannot tell us the mass scale of the neutrinos, which is required in order to measure the  
799 absolute neutrino masses.

800 The neutrino mass scale remains an unknown quantity in the standard model of  
801 particle physics. The value of the neutrino mass influences the evolution of the early  
802 universe and is likely relevant to the energy-scale of new physics responsible for the factor  
803 of  $10^{-6}$  difference between the neutrino and electron masses. A model-independent way  
804 to measure the neutrino mass is to measure the tritium beta-decay spectrum near its  
805 endpoint. Energy conservation requires that the neutrino mass carry away some kinetic  
806 energy from the beta-decay electron in the form of its mass, which causes a distortion in

807 the shape of the tritium beta-decay spectrum near the endpoint. The isotope tritium has  
808 many advantages for this measurement, and has been used by the KATRIN collaboration  
809 to perform the most sensitive direct neutrino mass measurement to date.

810 KATRIN represents the state-of-the-art experiment in the current generation of  
811 neutrino mass direct measurement experiments and has a final projected sensitivity to  
812 neutrino masses  $m_\nu < 200$  meV. This sensitivity does not fully exhaust the allowed  
813 parameter space of neutrino masses under the normal and inverted neutrino mass  
814 ordering scenarios, which motivates the development of a next generation of neutrino  
815 mass measurement experiments.

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

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

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

842 A parallel development is the technology necessary to produce, cool, trap, and  
843 recirculate a supply of atomic tritium that is compatible with CRES measurements. The  
844 atomic tritium system is equally important as the large-volume CRES measurement  
845 technology, but it will not be the focus of this dissertation since I did not contribute  
846 significantly to this effort.

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

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

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

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

## 873 1.2 Outline

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

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

884 Chapter 4 discusses the first of the contributions mentioned above, which is the  
885 development of signal reconstruction techniques for antenna array CRES and an antenna  
886 array demonstrator experiment called the FSCD. I discuss the important tools that Project  
887 8 uses to simulate antenna array CRES before introducing three signal reconstruction  
888 algorithms that can be used to detect CRES signals using the array. I end Chapter 4  
889 with a paper that summarizes a detailed analysis and comparison of the signal detection  
890 performance of each algorithm.

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

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

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

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

<sub>911</sub> **Chapter 2 |**

<sub>912</sub> **Neutrinos and Neutrino Masses**

<sub>913</sub> **2.1 Introduction**

<sub>914</sub> In this chapter I provide a cursory overview of background information relevant to  
<sub>915</sub> neutrinos and neutrino mass measurements.

<sub>916</sub> In Section 2.2 I provide some background information on the history of neutrinos and  
<sub>917</sub> beta-decay. In Section 2.3 I describe the discover of neutrino oscillations, which proved  
<sub>918</sub> unambiguously that neutrinos have non-zero masses. In Section 2.4 I discuss the current  
<sub>919</sub> state of the theoretical understanding of neutrino masses in the standard model. Lastly,  
<sub>920</sub> in Section 2.5 I discuss methods for measuring the absolute scale of the neutrino mass.

<sub>921</sub> **2.2 Neutrinos and Beta-decay**

<sub>922</sub> Late in the 19th century the phenomena of radioactivity was first observed in experiments  
<sub>923</sub> performed by Henri Becquerel with uranium, and further studied using thorium and  
<sub>924</sub> radium by Marie and Pierre Currie [4,5]. Early work in radioactivity classified different  
<sub>925</sub> forms of radiation based on it's ability to penetrate different materials. Rutherford was  
<sub>926</sub> the first to separate radioactive emissions into two types, alpha and beta radiation [6].  
<sub>927</sub> Alpha rays can be easily stopped by a piece of paper or thin foil of metal, whereas beta  
<sub>928</sub> radiation could penetrate metals several millimeters thick. Later a third form of radiation  
<sub>929</sub> was identified by Villard [7], which was still more penetrating, and was eventually termed  
<sub>930</sub> gamma radiation by Rutherford.

<sub>931</sub> When these forms of radioactivity were first discovered it was unclear what physically  
<sub>932</sub> constituted an alpha, beta, or gamma particle. Experiments with radioactivity in  
<sub>933</sub> magnetic fields was eventually able to identify the charge composition of different forms  
<sub>934</sub> of radiation. In particular, experiments by Becquerel identified that beta radiation had

935 an identical charge-to-mass ratio to the electron discovered by Thompson in his work on  
936 cathode rays [8]. This was strongly suggestive that beta particles were indeed electrons.

937 Further studies of beta radiation lead to the discovery that radioactivity resulted in  
938 the transmutation of elements [9] caused by the decay of a heavier nucleus to a lighter  
939 species. One feature of beta radiation, which we now properly call beta-decay, that  
940 was different from alpha-decays and gamma radiation is that the electrons produced by  
941 beta-decay have a continuous spectrum of kinetic energies, whereas, alpha and gamma  
942 particles are emitted with discrete energies. This feature of beta-decay was first observed  
943 by Chadwick in 1914 [10], and was extremely puzzling at the time since the continuous  
944 spectrum apparently violates energy conservation [11].

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

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

## 965 **2.3 Neutrino Oscillations**

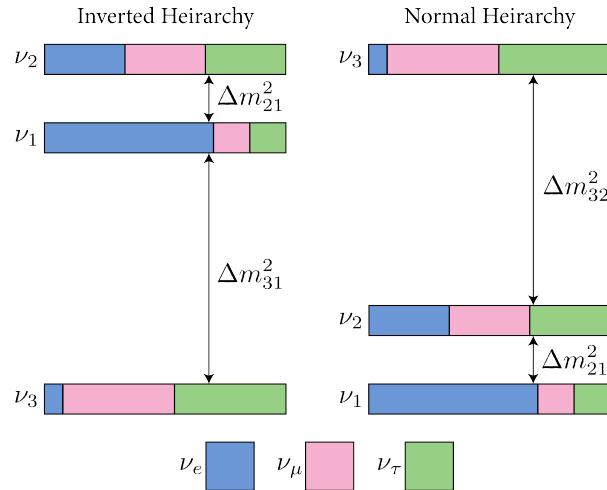
966 The first hint of neutrino flavor transitions or neutrino oscillations was indicated by  
967 the solar neutrino problem, which referred to discrepancies between the predicted flux

of  $\nu_e$  from the standard solar model and measurements of the solar neutrino flux such as the famous experiment at the Homestake mine by Ray Davis Jr. and collaborators in the 1960's [19]. Essentially, fewer electron-type neutrinos than expected were being observed from the sun. Finally, in the early 2000's the SNO experiment was able to resolve the solar neutrino problem by identifying neutrino oscillations as the cause of the observed deficit [20]. Furthermore, measurements of the atmospheric flux of neutrinos by the Super-Kamiokande experiment and others revealed that fewer muon-type neutrinos survived passage through the earth than expected providing strong evidence for neutrino oscillations for both flavors [21].

The origin of neutrino oscillations is that the weak eigenstates are distinct from the mass eigenstates [22]. The neutrino mass eigenstates represent physical particles in the sense that they are solutions to the free-particle Hamiltonian, whereas, the neutrino weak eigenstates correspond to the neutrino states that interact via the weak charged-current interaction. The neutrino weak eigenstates are a linear superposition of the neutrino mass eigenstates

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

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



**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.

986 One standard parameterization of the PMNS matrix is

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

987 where  $c_{ij} = \cos \theta_{ij}$  and  $s_{ij} = \sin \theta_{ij}$ . The parameters  $\alpha_1$  and  $\alpha_2$  are only included in the  
988 PNMS matrix if neutrinos are Majorana particles, something which represents a current  
989 area of research in neutrino physics. The phase  $\delta$  quantifies the degree of CP-violation  
990 in the neutrino sector. Including the Majorana phases the PMNS matrix contains six  
991 independent parameters. In addition, neutrino oscillation probabilities depend on the  
992 squared mass differences between neutrino mass eigenstates

$$\Delta m_{ij}^2 = m_i^2 - m_j^2, \tag{2.3}$$

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

995 A giant experimental effort over the past couple of decades has greatly contained the  
996 majority of parameters in the PMNS matrix, many to relative uncertainties of only a  
997 few percent. However, some parameters still remain relatively unconstrained, which is  
998 the origin of the current uncertainty in the ordering of the neutrino masses (see Figure  
999 2.1). The neutrino masses can be organized by their relative mass. The current neutrino  
1000 oscillation data can confirm that  $m_2 > m_1$ , however, the sign of  $\Delta m_{32}^2$  is still unknown.  
1001 This leads to two scenarios where neutrino masses follow the ordering  $m_3 > m_2 > m_1$ ,  
1002 which is called the normal mass ordering (NMO), or alternatively neutrino masses may  
1003 be ordered  $m_2 > m_1 > m_3$ , which is called the inverted mass ordering (IMO). Next-  
1004 generation neutrino oscillation experiments such as JUNO [23], Hyper-Kamiokande [24],  
1005 and DUNE [25] are poised to resolve this ambiguity in the coming years.

1006 Neutrino oscillation probabilities are only sensitive to the neutrino masses via the  
1007 squared mass differences. Therefore, oscillation probabilities are unaffected by the

absolute scale of the neutrino mass. However, oscillations can be used to obtain a lower bound on the neutrino masses by setting the mass of the lightest neutrino mass state to zero. This results in different lower limits depending on the ordering of the neutrino mass states. Current best-fit values [26] with  $1\sigma$ -uncertainties for the squared mass differences are

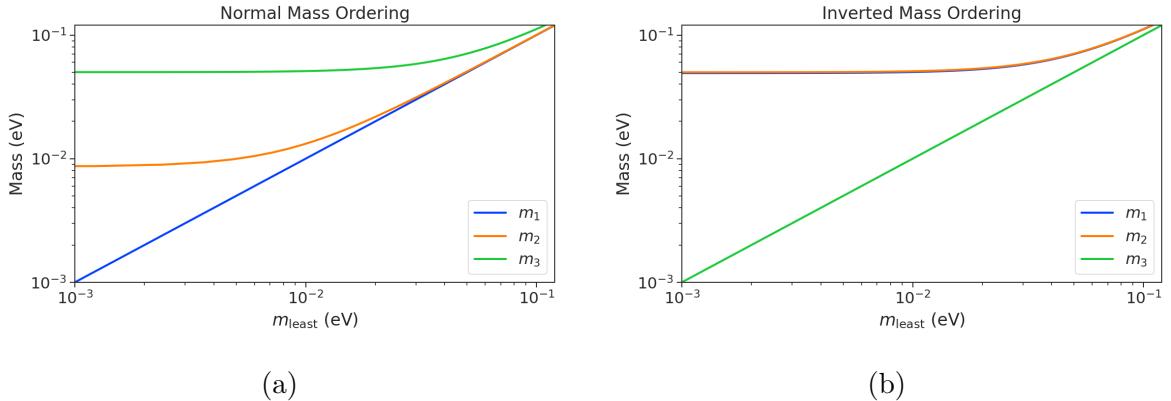
$$\Delta m_{21}^2 = (7.42^{+0.21}_{-0.20}) \times 10^{-5} \text{ eV}^2, \quad (2.4)$$

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

for the normal mass ordering, and in the case of the inverted ordering we have

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

By letting the lightest neutrino mass in each ordering scenario ( $m_{\text{least}}$ ) take on a range of values one can visualize the relative masses of the neutrinos as a function of  $m_{\text{least}}$  (see Figure 2.2).



**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.

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

1021 where the field that describes the particle is denoted as  $\psi(x)$ . In the standard model  
 1022 fermions acquire mass through the Yukawa interaction, which add to the standard model  
 1023 Lagrangian terms of the form

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

1024 where  $Y_{ij}^\ell$  is an element of the  $3 \times 3$  Yukawa coupling matrix for leptons,  $L_{Li}$  is the  
 1025 left-handed lepton doublet for generation  $i$ ,  $\phi$  is the Higgs doublet, and  $E_{Rj}$  is the  
 1026 right-handed lepton field for generation  $j$ . Neutrinos are represented only as left-handed  
 1027 neutrinos and right-handed antineutrinos in the standard model, which is consistent  
 1028 with experimental observations. Since there are no right-handed neutrino singlet fields,  
 1029 there are no Yukawa interaction terms, thus neutrinos in the standard model are strictly  
 1030 massless. Therefore, non-zero neutrino mass is evidence for physics beyond the standard  
 1031 model.

1032 For the charged leptons, the Yukawa interaction leads to masses of the form

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

1033 where  $v$  is the Higgs vacuum expectation value. The observation of massive neutrinos  
 1034 motivates the extension of the standard model to explain the origin of neutrino masses,  
 1035 which can be approached in different ways, but all approaches add additional degrees of  
 1036 freedom to the standard model. One approach is to introduce to the standard model a  
 1037 right-handed neutrino field that allows one to include Yukawa terms of the form

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

1038 where  $\nu_{Rj}$  is the right-handed neutrino singlet. Because experimental evidence strongly  
 1039 predicts only three active neutrinos, these additional neutrinos are sterile and do not in-  
 1040 teract via the strong, weak, or electromagnetic interactions. After spontaneous symmetry  
 1041 breaking, the Yukawa interaction leads to mass terms given by

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

1042 which is called a Dirac mass term. One of the issues with constructing neutrino masses  
 1043 in this way is that the required Yukawa couplings are at least a factor of  $10^6$  smaller than  
 1044 that of an electron, which begs the question: why are the Yukawa couplings so small for

1045 the neutrinos?

1046 An alternative approach is to allow the neutrinos to have a Majorana mass, which is  
1047 possible because neutrinos are electrically neutral particles. The Majorana mass terms  
1048 for the neutrino have the form

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

1049 where  $M_{Rij}$  and  $M_{Lij}$  are right-handed and left-handed Majorana mass matrices. A  
1050 consequence of neutrinos being Majorana particles is lepton number violation, which  
1051 predicts the occurrence of neutrino-less double beta-decay at a rate proportional to the  
1052 neutrino mass.

1053 In the most general case neutrinos have both Dirac and Majorana mass terms, which  
1054 allows one to generate neutrino masses with Yukawa couplings similar to the rest of  
1055 the standard model. Considering just one generation of neutrinos for illustration, the  
1056 combined Lagrangian can be written as

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

1057 or equivalently,

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

1058 An example mass generation mechanism with this approach is the Type-I see-saw  
1059 mechanism [30], in which we take  $m_L = 0$  and  $m_R \gg m_D$ . By diagonalizing Equation  
1060 2.14 one obtains the mass eigenvalues that represent the physical masses of the neutrinos.  
1061 The light neutrino mass eigenstate, which represents the observed neutrino mass, has a  
1062 mass given by

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

1063 and the heavy neutrino mass eigenstate, which represents the unobserved sterile neutrino,  
1064 has a mass

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

1065 For  $m_D$  similar to the other quark or lepton masses, one obtains physical neutrino masses  
1066 consistent with observations from sterile neutrino masses of  $m_R \approx O(10^{15})$  GeV. This  
1067 mass scale is well beyond the capabilities of modern particle accelerators.

## **2.5 Neutrino Absolute Mass Scale**

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

### **2.5.1 Limits from Cosmology**

The  $\Lambda$ CDM model summarizes our current cosmological understanding of our universe [31].  $\Lambda$ CDM predicts that the universe originated from a single expansion event colloquially called the "Big Bang". During the Big Bang, the universe originated as a hot spacetime singularity, which abruptly experienced rapid expansion in a process known as inflation. After expansion the inflationary field eventually decayed into a population of quarks, gluons, leptons, and photons, which were kept in thermal equilibrium by the high-temperatures of the early universe.

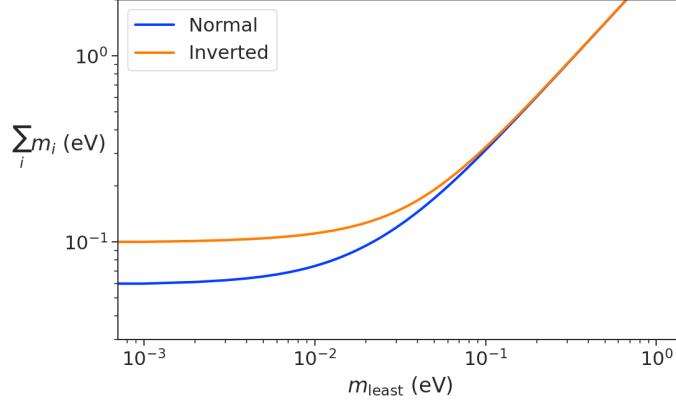
As the universe continued to expand its density and temperature decreased until the formation of neutral atoms, primarily hydrogen, was possible. At which point the population of photons produced during the Big Bang thermally decoupled. A direct prediction of the  $\Lambda$ CDM model is that this population of photons should still be present, but with a significantly reduced temperature due to the expansion of the universe. This is consistent with the observation of the CMB (cosmic microwave background), which is a population of microwave radiation with a blackbody temperature of 2.7 K. The CMB is extremely uniform in all directions with slight anisotropies that can be analyzed to study the evolution of the early universe. A series of experiments have measured the CMB with increasing levels of precision, which has lead to a significant increase in our current understanding of cosmology.

In addition to the CMB, inflation predicts the existence of a  $C\nu B$  (cosmic neutrino background) [32], which are the remnant neutrinos produced during the Big Bang. Since neutrinos only interact via the weak force, they decouple from the hot Big Bang plasma at an earlier time than the CMB radiation. The temperature at which the  $C\nu B$  decouples depends on the neutrino rest mass. Neutrinos play a unique role in the  $\Lambda$ 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 [33]. 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  $\Lambda$ CDM model. This analysis results in some of the most stringent constraints on the neutrino mass. Recent analyses [31] 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  $\Sigma_{m_\nu}$  constrains the neutrino mass by setting the mass of the lightest neutrino mass eigenstate ( $m_{\text{least}}$ ). In the normal mass ordering  $\Sigma_{m_\nu}$  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  $\Sigma_{m_\nu}$  effectively sets the neutrino mass scale through  $m_{\text{least}}$ . The analogous formula for the inverted mass ordering is

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

In figure 2.3 we plot the observable  $\Sigma_{m_\nu}$  as a function of  $m_{\text{least}}$ .

Upcoming experiments [34] are planned to refine measurements of the CMB, LSS, and other cosmological observables. With this additional data it is possible that in the near future cosmological measurements will be able to positively constrain the neutrino absolute mass scale. However, the strength of these limits strictly depend on the accuracy of the  $\Lambda$ CDM model, which highlights the need for direct experimental measurements of the neutrino mass to confirm the predictions of cosmology and to fix the neutrino mass

parameter in future cosmological analyses.

### 2.5.2 Limits from Neutrinoless Double Beta-decay Searches

If neutrinos are Majorana fermions, then the neutrino is equivalent to its own antiparticle and lepton conservation is not an exact law of nature [35]. Limits on the rate of neutrinoless double beta-decay ( $0\nu\beta\beta$ ), are some of the most powerful current tests of lepton number conservation [31]. If  $0\nu\beta\beta$  were observed it would direct evidence that neutrinos are Majorana fermions, and provide a method for measuring the neutrino mass scale.

Standard double beta-decay occurs when two neutrons contained in the nucleus spontaneously decay into two protons, which results in the production of two electrons and two neutrinos (see Figure 2.4). However, during  $0\nu\beta\beta$  the two neutrinos self-annihilate

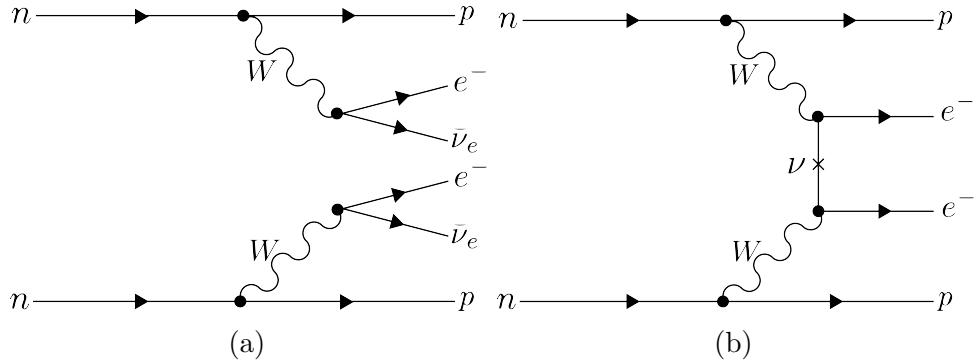


Figure 2.4. Feynman diagrams for double beta-decay (a) and  $0\nu\beta\beta$ (b).

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

Assuming that the exchange of two Majorana neutrinos is the dominant channel for  $0\nu\beta\beta$ , then a measurement of the  $0\nu\beta\beta$  half-life for a particular isotope can be used to set the neutrino absolute mass scale [36]. The half-life is written in terms of the effective neutrino mass for  $0\nu\beta\beta$  ( $m_{\beta\beta}$ ) using the equation

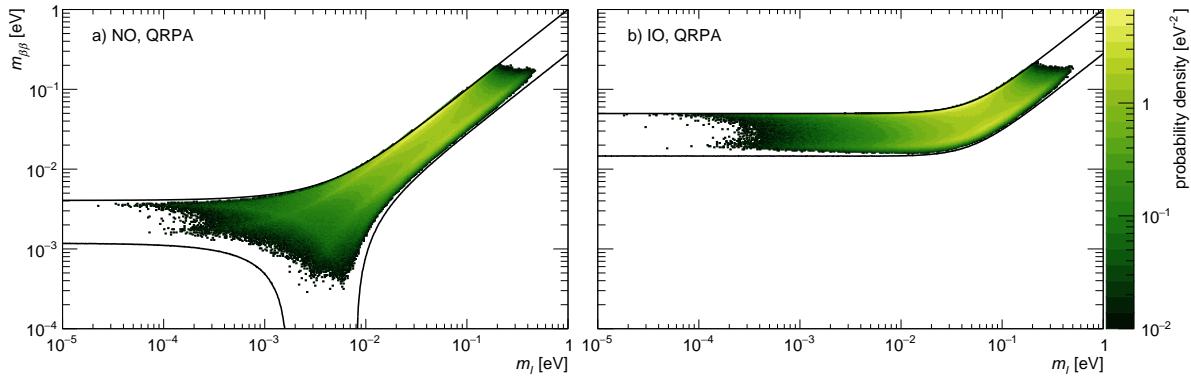
$$T_{1/2}^{0\nu} = \frac{1}{G|\mathcal{M}|^2 m_{\beta\beta}^2}, \quad (2.20)$$

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

1136 by the PMNS mixing matrix parameters,

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

1137 The information provided from  $0\nu\beta\beta$  on the neutrino mass scale can be visualized by  
 1138 expressing the value of  $m_{\beta\beta}$  in terms of  $m_{\text{least}}$  and two relative Majorana phases [1]. The  
 1139 allowed regions for  $m_{\beta\beta}$  as a function of  $m_{\text{least}}$  are shown in Figure 2.5 as the regions  
 1140 bounded by the black curves overlayed with the discovery probabilities of future  $0\nu\beta\beta$   
 decay experiments based on current neutrino data.

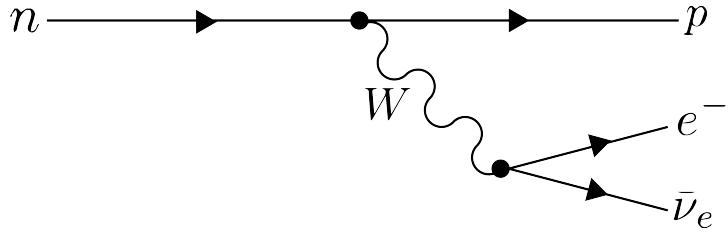


**Figure 2.5.** The discovery probabilities for the future generation of  $0\nu\beta\beta$  experiments as a function of  $m_{\beta\beta}$  and  $m_{\text{least}}$ . Figure from [1].

1141  
 1142 Because of the possibility of cancellation due to the unknown Majorana phases  
 1143 included in the sum specified by Equation 2.21, the information gained is necessarily  
 1144 imperfect. Additionally, theoretical uncertainties in the calculation of the nuclear matrix  
 1145 elements complicates the calculation of  $m_{\beta\beta}$  from a measurement of  $0\nu\beta\beta$  half-life. Similar  
 1146 to cosmology, there is a high degree of complementarity between direct measurements  
 1147 of the neutrino mass and  $0\nu\beta\beta$ . In particular, a measurement of  $m_{\text{least}}$  to less than  
 1148 than 0.1 eV sensitivity provides significant information for  $0\nu\beta\beta$  searches based of the  
 1149 discovery probabilities of Figure 2.5.

### 1150 2.5.3 Limits from Beta-decay

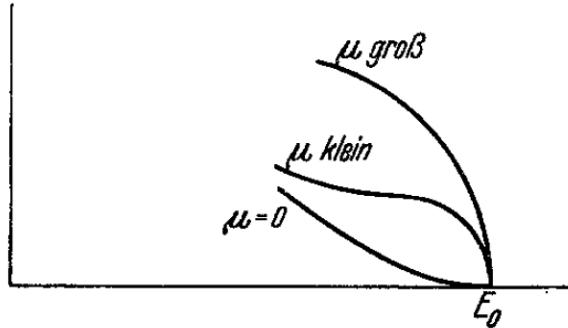
1151 Certain processes involving neutrinos, in particular beta-decay (see Figure 2.6), have  
 1152 initial states with well-defined total energies and final states that can be measured with  
 1153 high accuracy and precision. Beta-decay involves the decay of an unstable isotope where  
 1154 a neutron spontaneously converts to a proton and emits and electron and anti-neutrino



**Figure 2.6.** A Feynman diagram of beta decay

1155 ("neutrino" for brevity) to conserve charge and lepton number [4]. Therefore, by applying  
 1156 the principles of energy and momentum conservation, a measurement of the kinematics  
 1157 of the final state can be used to constrain the neutrino mass [37].

1158 Using beta-decay to measure the neutrino mass can be tied back to Fermi's original  
 1159 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  $\mu$ , is to distort the shape of the spectrum near the endpoint from the zero-mass spectrum.

1159  
 1160 neutrino mass from beta-decay depend only on the final state measurement capabilities  
 1161 and the principles of energy and momentum conservation, neutrino mass measurements  
 1162 with beta-decay are sometimes called direct measurements. A direct measurement like  
 1163 beta-decay contrasts with other neutrino mass measurements approaches that are model-  
 1164 dependent such as cosmology and  $0\nu\beta\beta$ , which provide complementary ways to study  
 1165 the physics of massive neutrinos.

1166 The isotope of choice for direct neutrino mass measurements with beta-decay has  
 1167 been tritium ( ${}^3H_2$ ) for many decades, because it conveniently fulfills many experimental  
 1168 requirements. Of upmost importance is a decay with a low Q-value, which is the available  
 1169 kinetic energy based on the mass difference between the initial and final states. The  
 1170 effect of a massive neutrino on the shape of the spectrum is magnified for low Q-values

1171 and tritium decays have an unusually low Q-value of 18.6 keV.

1172 Additionally, tritium beta-decay is a super-allowed decay, which results in a relatively  
1173 short half-life of 12.3 years. Therefore, it is relatively easy to obtain a high-activity  
1174 using a small source mass. High-activity is desirable because of the low-activity near  
1175 the tritium spectrum endpoint. For tritium beta-decays only a factor of  $3 \times 10^{-13}$  of  
1176 the decays occur in the last 1 eV of the spectrum. Isotopes with Q-values lower than  
1177 tritium are known [37], but this is outweighed by exceedingly long half-lives leading to  
1178 unobtainable source masses.

1179 The measurement involves quantifying the effect of the neutrino's mass on shape of  
1180 the electron's kinetic energy spectrum near the endpoint. The shape of the kinetic energy  
1181 spectrum (see Figure 2.8) is given by

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

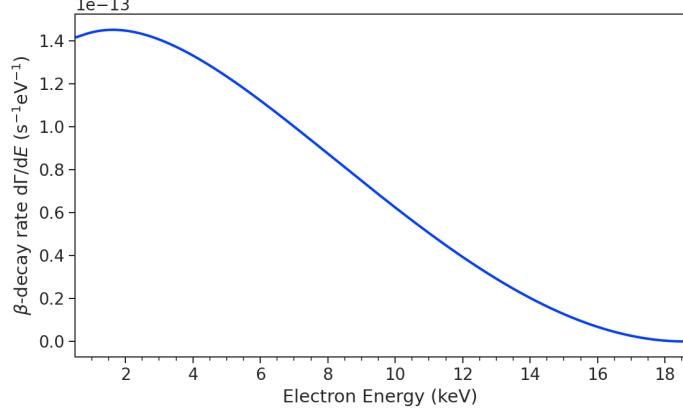
1182 where  $G_F$  is the Fermi coupling constant,  $V_{ud}$  is an element of the CKM matrix,  $E$  is  
1183 the kinetic energy of the electron,  $\beta$  is the velocity of the electron divided by the speed  
1184 of light,  $E_0$  is the endpoint energy assuming zero neutrino mass,  $F(Z, \beta)$  is the Fermi  
1185 function, and  $\Theta(E_0 - E - m_i)$  is the Heaviside function, which enforces energy conservation.  
1186 One can see that the decay spectrum is actually a combination of three spectra with  
1187 different endpoints based on the actual values of the neutrino mass eigenstates,  $m_i$ . This  
1188 results in "kinks" in the spectrum shape due to the overlapping spectra, but such an  
1189 effect would likely be impossible to resolve given the finite energy resolution of a real  
1190 experiment and low statistics.

1191 The neutrino mass scale variable measured by beta-decay is given by

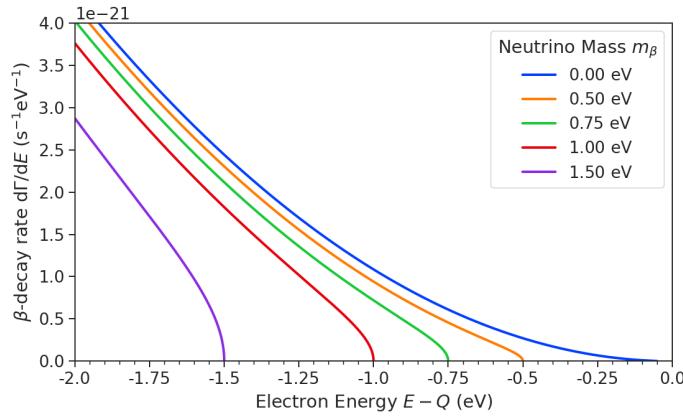
$$m_\beta^2 = \sum_i |U_{ei}|^2 m_i^2, \quad (2.23)$$

1192 where  $m_\beta$  is the electron-weighted neutrino mass or simply "neutrino mass" for brevity.  
1193  $m_\beta$  corresponds to a particular weighted sum of the neutrino masses, which is distinct  
1194 from effective neutrino masses such as  $m_{\beta\beta}$  [37]. Assuming unitarity, the neutrino mass  
1195 can be expressed in terms of the PMNS matrix elements, squared mass differences, and  
1196 the lightest neutrino mass eigenstate. For the normal mass ordering the equation is

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



(a)



(b)

**Figure 2.8.** The tritium beta-decay spectrum. The affect of a massive neutrino on the spectrum is to change it's shape near the endpoint by an amount proportional to the size of the neutrino mass. This suggests that a sufficiently high-statistic and high-resolution measurement of the spectrum endpoint would be able to measure the neutrino mass.

and for the inverted ordering the equation is

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

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

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

<sub>1204</sub> low background activity.

1205 **Chapter 3 |**

1206 **Direct Measurement of the Neutrino Mass**

1207 **with Project 8**

1208 **3.1 Introduction**

1209 A promising technique for direct measurements of the neutrino mass beyond the projected  
1210 limit of the ongoing KATRIN experiment [38] is tritium beta-decay spectroscopy with an  
1211 atomic tritium source [39]. Atomic tritium, combined with a large-volume, high-resolution  
1212 energy measurement technique, is capable of measuring the neutrino mass with sensitivity  
1213 below the 50 meV limit allowed by neutrino oscillations.

1214 Cyclotron Radiation Emission Spectroscopy or CRES is a high-resolution energy  
1215 measurement technique compatible with atomic tritium production and storage that can  
1216 enable the next-generation of neutrino mass direct measurement experiments [40]. The  
1217 Project 8 collaboration is currently engaged in a program of research and development  
1218 (R&D) aimed at developing the technology necessary for a 40 meV sensitivity measurement  
1219 of the neutrino mass using CRES and atomic tritium [41].

1220 In Section 3.2 I provide an introduction to the basics of the CRES technique as well as  
1221 the goals of the Project 8 experiment. Additionally, I sketch out the phased experiment  
1222 development plan being implemented by Project 8 to build towards a next-generation  
1223 neutrino mass experiment.

1224 In Section 3.3 I give a brief overview of Phase II of the Project 8 experiment [42, 43],  
1225 which completed early in 2023. Although the bulk of the work presented in this thesis is  
1226 relevant to designs of future Project 8 experiments, a description of the work in Phase II  
1227 provides useful context for the rest of the work.

1228 In Section 3.4 I introduce a CRES measurement concept based on antenna arrays [44],  
1229 which could be the basis for the ultimate Project 8 neutrino mass experiment. A  
1230 significant portion of the R&D efforts of Project 8 in Phase III were directed towards

1231 simulating and modeling this experimental concept in order to understand the achievable  
1232 sensitivity to the neutrino mass.

1233 Lastly, in Section 3.5 I introduce conceptual designs of pilot-scale experiments and  
1234 Phase IV that combine atomic CRES with a large-volume CRES detection technique.  
1235 This includes a design concept for an antenna array based experiment, but also a design  
1236 for a resonant cavity based experiment. Resonant cavities are discussed in more depth in  
1237 Chapter 6 and have become the default choice for the Phase IV experiment.

## 1238 **3.2 Cyclotron Radiation Emission Spectroscopy and Project** 1239 **8**

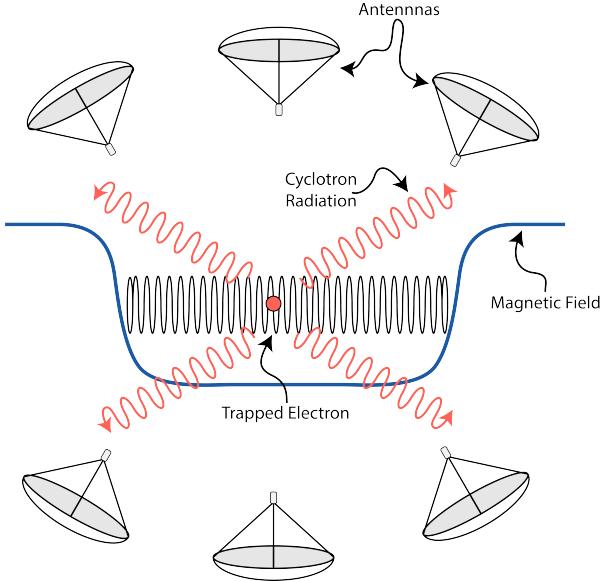
### 1240 **3.2.1 Cyclotron Radiation Emission Spectroscopy — CRES**

1241 Time and frequency are two of the most precisely measured quantities in physics. It is  
1242 often advantageous to convert measurements of other physical quantities like mass or  
1243 length into frequency measurements due to the digital nature of frequency measurements  
1244 that make them immune to many sources of noise. Atomic clocks, which operate by  
1245 measuring the frequencies of various atomic transitions, have been used to measure  
1246 time with astounding relative uncertainties of  $10^{-18}$  seconds [45]. The extreme precision  
1247 possible with frequency measurements is often summarized using the a quote from the  
1248 Physicist Arthur Schawlow who said advise his students to "Never measure anything but  
1249 frequency!" [46].

1250 Neutrino mass measurements using tritium beta-decay require us to measure pertur-  
1251 bations of the 18600 eV tritium endpoint to a precision as low as 0.1 eV, therefore, a  
1252 spectroscopic technique with extremely high resolution is required for this measurement.  
1253 The intuitive explanation for why frequency measurements are capable of such high reso-  
1254 lutions is that they are essentially counting measurements, which average the number of  
1255 oscillations of a physical system over time. By observing a rapidly oscillating system over  
1256 a sufficient length of time one can obtain essentially arbitrary precision on a frequency  
1257 limited only by the time available for measurement and the SNR of the system.

1258 What is required is that one translate the kinetic energy of the electron into a frequency,  
1259 and a straightforward way to accomplish this is to place a gaseous supply of tritium into  
1260 a magnetic field. When an atom decays the resulting electron will immediately begin  
1261 to orbit around a magnetic field line at the cyclotron frequency, which is proportional  
1262 to its kinetic energy (see Figure 3.1). The acceleration caused by the orbit leads to the

1263 emission of cyclotron radiation that can be detected using an array of antennas or a  
 1264 different RF sensor such as a resonant cavity. The frequency of the radiation gives the  
 1265 electron's kinetic energy, which is used to build the beta-decay spectrum and measure  
 1266 the neutrino mass. The name for this measurement technique is Cyclotron Radiation  
 1267 Emission Spectroscopy or CRES [40].



**Figure 3.1.** A cartoon illustration of the CRES technique. An electron is contained in a magnetic trap so that it's 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.

1268 For non-relativistic particles the cyclotron frequency is simply a function of the  
 1269 charge-to-mass ratio of the particle, however, from the relativistic form of the cyclotron  
 1270 frequency

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

1271 one can see that the kinetic energy ( $E_{\text{kin}}$ ) of the electron is directly proportional to the  
 1272 inverse of the cyclotron frequency ( $f_c$ ). Electrons with kinetic energies of 18.6 keV are in  
 1273 the weakly relativistic regime with  $\beta = \frac{v}{c} = 0.263$  and  $\gamma = 1.036$ .

1274 The required frequency resolution needed for neutrino mass measurement can be  
 1275 obtained by differentiating Equation 3.1,

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

1276 from which we can obtain the relationship between fractional differences in energy and

1277 frequency,

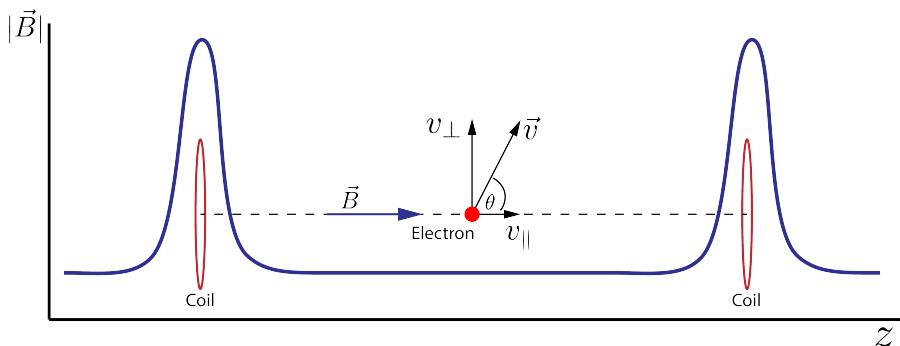
$$\frac{df_c}{f_c} = \frac{1 - \gamma}{\gamma} \frac{dE_{\text{kin}}}{E_{\text{kin}}}. \quad (3.3)$$

1278 Therefore, an energy precision of 1 eV for an 18.6 keV electron requires a frequency  
1279 precision of approximately 2 ppm.

1280 The minimum observation time required to achieve this resolution can be estimated  
1281 using the uncertainty principle as formulated by Gabor [47]. Electrons from tritium  
1282 beta-decay experience random collisions with the background gas particles, which limits  
1283 the uninterrupted radiation lifetime. The time between collision events, referred to  
1284 as track length in the context of CRES measurements, is an exponentially distributed  
1285 variable. Differences in the track lengths of a population of mono-energetic electrons leads  
1286 to uncertainty or broadening in the distribution of measured frequencies proportional to  
1287 the mean track length,  $\tau_\lambda$ . The resulting frequency distribution has a Lorentzian profile,  
1288 whose width is given by the Gabor limit,

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

1289 The cyclotron frequency for a 18.6-keV electron in a 1 T field is approximately  
1290 27 GHz, from which one can estimate the minimum observation time for 2 ppm frequency  
1291 resolution at approximately 3  $\mu$ sec. The Gabor limit is not the true lower bound on the  
1292 frequency resolution for a CRES signal, since it is based on the details of the Fourier  
1293 representation of a time-series with a fixed length. If one takes the approach of fitting  
1294 the CRES signal in the time-domain, then one finds that the limit on frequency precision  
1295 is given by the Cramér-Rao lower bound (CRLB) [48], which depends on the track length  
1296 and SNR. The CRLB allows for better precision on the cyclotron frequency, however,  
1297 the Gabor limit provides an intuitive limit with the correct order of magnitude.



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

Ensuring that an electron remains under observation long enough so that it's frequency can be precisely measured requires a magnetic trap. A magnetic trap is a local minimum in a background magnetic field generated an appropriate configuration of electromagnetic 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 in it's simplest form consists of two high magnetic field pinch coils aligned on a central axis that are well separated (see Figure 3.2). This configuration produces a trap with a flat uniform bottom and relatively steep walls, which is ideal for CRES measurements.

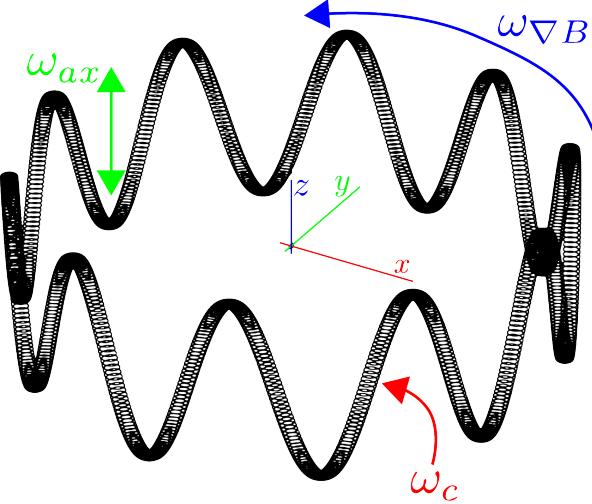
Electrons produced in the trap oscillate back and forth between the trap walls at a frequency that depends upon the pitch angle, unless they are produced with pitch angles too small to be contained in the trap. Pitch angle is defined as the angle between the component of the electron's velocity perpendicular 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 due to the changing value of the magnetic fields. 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, 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 but also 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 wall both of which are 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 [49]. The equation for the drift velocity is

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



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

1328 These additional components of motion all influence the shape of the CRES signal so  
 1329 modeling their effects is critical to proper measurement of the kinetic energy.

1330 The total power of the radiation emitted by an electron in a free-space environment  
 1331 is given by the Larmor equation [50]

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

1332 where  $\omega_c$  is the cyclotron frequency multiplied by  $2\pi$  and  $\theta_p$  is the pitch angle to distinguish  
 1333 it from the spherical angle coordinate. A single electron with a  $90^\circ$  pitch angle and  
 1334 18.6 keV of kinetic energy in a 1 T magnetic field emits a total radiation power of 1.2 fW,  
 1335 which is quite small compared with typical RF systems, furthermore, one is typically  
 1336 only able to receive a fraction of this total power with an antenna or other detection  
 1337 system. Therefore, RF systems in CRES experiments must be operated at cryogenic  
 1338 temperatures to limit the noise power such that adequate SNR can be achieved for signal  
 1339 detection and reconstruction. Alternatively, longer tracks enable detection of weaker  
 1340 signals due to the increase in the total signal energy available for the detection algorithm.

### 1341 3.2.2 The Project 8 Collaboration

1342 The Project 8 collaboration<sup>1</sup> is a group of institutions in the United States and Germany  
 1343 aiming to measure the neutrino mass by developing a novel spectrometer technology

---

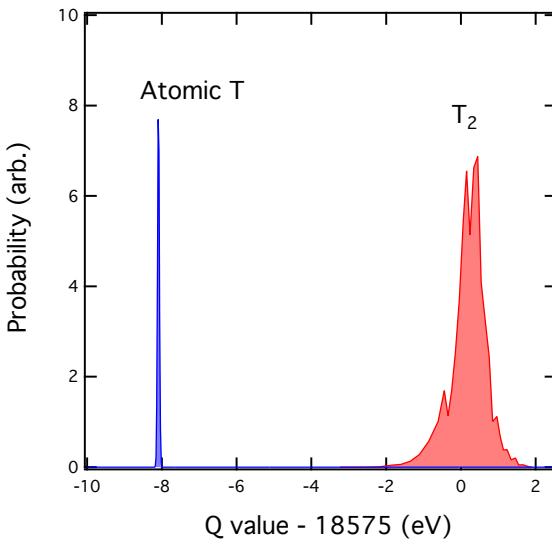
<sup>1</sup><https://www.project8.org/>

1344 based on CRES. In the ultimate Project 8 experiment the CRES technique will be used  
1345 to measure the beta-decay spectrum using a large source of atomic tritium sufficient to  
1346 achieve the required statistics in the last  $O(10)$  eV of the decay spectrum. Project 8 is  
1347 targeting a neutrino mass sensitivity below 50 meV [51], which exhausts the range of  
1348 possible neutrino masses under the inverted hierarchy and is a factor of four less than  
1349 sensitivity projections for the ongoing KATRIN experiment.

1350 Project 8's proposed experiment requires the development of two novel technologies:  
1351 the production and trapping of a source of atomic tritium on cubic-meter scales and  
1352 technology to enable CRES measurements of individual electrons in the same volume.

### 1353 Atomic Tritium

1354 Previous measurements of the tritium beta-decay spectrum for neutrino mass measure-  
1355 ments have relied on sources of molecular tritium for their measurements [38, 52, 53] due  
1356 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.

1357 One must supply sufficient energy to the tritium molecules to break the molecular  
1358 bond and create atomic tritium. Common approaches to this include the use of hot  
1359 coaxial filament atom crackers as well as plasma atom sources. Both involve heating the  
1360 tritium atoms to temperatures of  $> 2500$  K, which must then be cooled to temperatures  
1361 on the order of a few mK so that the tritium atoms can be trapped. Cooling the atoms

1362 requires the construction of a large tritium infrastructure and cooling system that can  
1363 supply a source of cold atoms to the trap.

1364 Once cold tritium atoms are produced they cannot make contact with any surfaces  
1365 to avoid recombination of the atoms to molecules. Therefore, a magnetic trap is required  
1366 to store the atoms for a sufficient length of time that they have a chance to decay before  
1367 escaping the trap. Trapping the atoms requires the construction of a large and complex  
1368 magnet system that must be cooled to cryogenic temperatures.

1369 The significant experimental complexity caused by atomic tritium makes a molecular  
1370 source the obvious choice from practical considerations. However, the drawback of  
1371 molecular tritium for neutrino mass measurement is the irreducible broadening in the  
1372 electron's kinetic energy due to the final state spectrum of molecular tritium (see Figure  
1373 3.4). The broadening of the final state spectra has a RMS amplitude of 436 meV [54, 55]  
1374 caused by variation in the final vibrational state of the daughter molecule. For atomic  
1375 tritium the primary sources of broadening in the final state spectrum are magnetic  
1376 hyperfine splittings (magnitude of  $O(10^{-5})$  eV) and thermal Doppler broadening caused  
1377 by the motion of the trapped atom. For atomic tritium at a temperature of 1 mK thermal  
1378 broadening is the dominant contribution, providing about 1 meV RMS of broadening to  
1379 the electron's kinetic energy.

1380 The larger energy broadening with molecular tritium leads to an irreducible statistical  
1381 uncertainty that limits the achievable sensitivity to approximately 100 meV at 90%  
1382 confidence. For previous direct measurements of the neutrino mass this uncertainty is an  
1383 insignificant contribution to the overall uncertainty budget, however, for experiments  
1384 like Project 8 atomic tritium is a key component to the success of the experiment.

### 1385 **CRES for Neutrino Mass Measurement**

1386 Several features of the CRES technique make it an attractive choice for a next generation  
1387 neutrino mass measurement experiment. For example, with a CRES experiment the  
1388 volume of the source gas can be the same as the volume of the CRES spectrometer.  
1389 This is due to the fact that CRES is a remote-sensing technique that can observe the  
1390 energy of the electron without altering its trajectory or directly interacting with the  
1391 electron. Given that tritium gas is transparent to cyclotron radiation the kinetic energies  
1392 of electrons can be measured with an appropriate sensing technology, such as a cavity or  
1393 antenna array, located directly outside the atom trapping volume.

1394 The current state-of-the-art tritium beta-decay spectroscopy experiment, KATRIN,  
1395 utilizes the magnetic adiabatic collimation with an electrostatic filter (MAC-E filter)

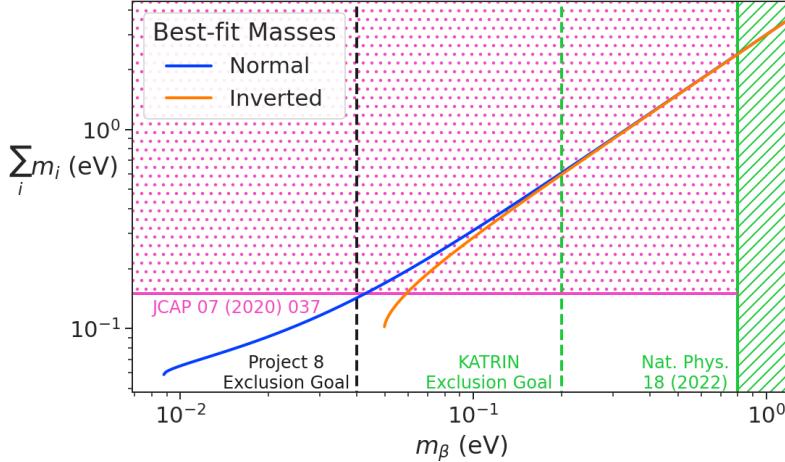
1396 technique to measure the beta-decay spectrum of molecular tritium. In this approach,  
1397 a source of molecular tritium is located outside the spectrometer. When a beta-decay  
1398 occurs the electron must exit the tritium source and travel through the MAC-E filter  
1399 before it can be detected on the other side of the filter using a charge sensor. The  
1400 measurement statistics of the MAC-E filter are limited by the transverse areas of the  
1401 tritium source and the filter due to the need to travel through the experiment without  
1402 scattering. This scaling is less favorable than the volumetric scaling of CRES due to the  
1403 ability to co-locate source and detector.

1404 Another promising aspect of the CRES technique is the inherently high precision  
1405 of frequency based measurements. The endpoint of the molecular tritium beta-decay  
1406 spectrum is approximately 18.6 keV, which dwarfs the neutrino mass scale of  $< 1 \text{ eV}/c^2$   
1407 by at least a factor of  $10^5$ . Measuring the effect of such a small mass on a high energy  
1408 electron requires excellent energy resolution. Since frequency measurements are essentially  
1409 counting measurements they are intrinsically quite accurate due to the ability to measure  
1410 the cyclotron frequency by effectively averaging over millions of cyclotron orbits. Using  
1411 off-the-shelf RF components its is possible to achieve part-per-million accuracy on the  
1412 kinetic energy with the CRES technique.

1413 CRES is also nearly immune to typical sources of backgrounds that plague other  
1414 experiments. Since CRES operates via non-destructive measurements of the electron's  
1415 cyclotron frequency potential sources of background electrons are effectively filtered out  
1416 by limiting the frequency bandwidth of the measurement. The fiducial volume of the  
1417 experiment is free from any surfaces that could introduce stray electrons and electrons  
1418 from sources outside the fiducial volume can be prevented from entering the experiment.

## 1419 Neutrino Mass Sensitivity Goals

1420 Project 8's ultimate goal is to combine CRES with atomic tritium to measure the neutrino  
1421 mass with 40 meV sensitivity at the 90% confidence level (see Figure 3.5). This sensitivity  
1422 is sufficient to fully exhaust the range of allowable neutrino masses under the inverted  
1423 neutrino mass ordering regime and is approximately an order of magnitude less than the  
1424 projected final sensitivity of the KATRIN experiment. Excluding the full neutrino mass  
1425 parameter space would require a sensitivity an order of magnitude lower than what is  
1426 proposed by Project 8, which would require an experiment whose size and complexity  
1427 are currently well beyond proposals for the next-generation of neutrino mass direct  
1428 measurement experiments.



**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.

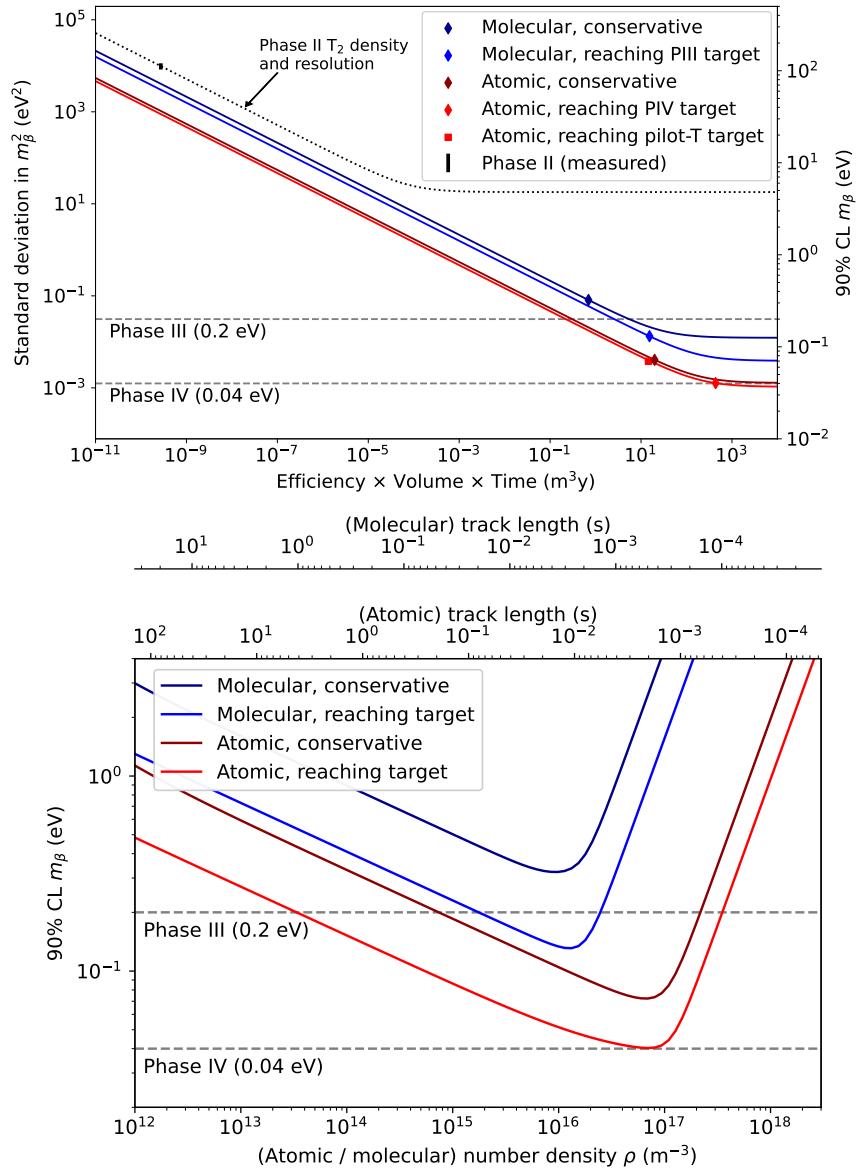
### 1429 3.2.3 The Project 8 Phased Development Plan

1430 Reaching 40 meV sensitivity requires the simultaneous development and eventually  
 1431 combination of CRES and atomic tritium. These technologies require a significant up-  
 1432 front research and development (R&D) investment to build-out the required capabilities  
 1433 for a 40 meV CRES experiment. Therefore, Project 8 is following a phased experiment  
 1434 plan in which incremental progress can be made towards the ultimate goal of a 40 meV  
 1435 neutrino mass measurement with CRES.

#### 1436 Phase I and II: Proof of Principle and First Tritium Measurements

1437 The earlier phases of the Project 8 experiment, Phase I and II, were focused on demon-  
 1438 stration and development of the CRES technique itself as well as a proof-of-principle  
 1439 measurement of the neutrino mass using the CRES technique.

1440 In Phase I, Project 8 performed a proof-of-principle measurement of the  $^{83m}\text{Kr}$   
 1441 spectrum using CRES, which marked the first ever energy spectrum measurement with  
 1442 CRES. The experiment included all of the main components expected for the full-scale  
 1443 version of the experiment. An electron source consisting of a gas of  $^{83m}\text{Kr}$  was supplied  
 1444 to a waveguide gas cell constructed out of a segment of WR-42 waveguide and sealed  
 1445 with Kapton windows at the top and bottom. A magnetic trapping region was created  
 1446 in the waveguide cell using a single electromagnetic coil wrapped around the waveguide



**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 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.

1447 which provided a trapping volume on the order of a few cubic-millimeters. Detection of  
1448 the cyclotron radiation was performed by connecting the waveguide cell to an additional  
1449 segment of waveguide that transmitted the radiation to a cryogenic amplifier.

1450 Success in Phase I was achieved with the 2014 publication of the measured  $^{83m}\text{Kr}$   
1451 conversion spectrum [56], which contains a mono-energetic 17.8-keV as well as several  
1452 other conversion lines at higher energies. Publication of this result marked the official  
1453 end of Phase I and the start of Phase II in which Project 8 shifted its focus to the  
1454 demonstration of the first tritium beta-decay spectrum using CRES. Phase II successfully  
1455 concluded in 2023 with the submission of the papers demonstrating the first tritium  
1456 beta-decay spectrum endpoint and neutrino mass measurement using CRES. For more  
1457 information on Phase II please see Section 3.3.

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

1459 After Phase II Project 8 has shifted focus to R&D towards the construction of an  
1460 experiment that demonstrates all the technologies required for a 40 meV measurement of  
1461 the neutrino mass. The goal for this pilot-scale experiment is to successfully retire all  
1462 technological and engineering risks associated with the Phase IV experiment, while being  
1463 a scientifically interesting experiment in its own right that has sensitivity to neutrino  
1464 masses on par with KATRIN’s final projected sensitivity.

1465 Phase III R&D is divided into two equally important efforts — atomic tritium  
1466 and CRES detection techniques. Atomic tritium development in Phase III includes  
1467 the development of all aspects of the tritium system. This includes the production of  
1468 tritium atoms, atomic cooling and recirculation systems, purity and isotope concentration  
1469 monitoring, and atom trapping. Currently, Project 8 is operating small scale demonstrator  
1470 systems developing atom crackers to show that atom production at the estimated rates  
1471 needed for Phase IV is achievable. Future efforts will continue the current developments  
1472 on atom production and expand to include demonstrations of atomic cooling with an  
1473 evaporative beam line as well as atom trapping using Halbach magnet arrays.

1474 The need for new CRES detection techniques is driven by the drastic increase in scale  
1475 from Phase II to the Phase IV and the pilot-scale experiments. The physical volume  
1476 used for CRES in Phase II was on the order of a few cubic-centimeters, and achieving  
1477 Project 8’s sensitivity target of 40 meV requires an experiment volume on the multi-cubic  
1478 meter scale. Therefore, the waveguide gas cell CRES detection technique used in Phase  
1479 II is not a feasible option for the future of Project 8 due to it’s inability to scale to the  
1480 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 a  
1487 large degree of the work presented in this thesis is focused on the development of the  
1488 antenna array CRES technique as well as the design of demonstrator experiments, we  
1489 described the proposed R&D plan for antenna array CRES in Phase III in Section 3.4.

1490 Cavity CRES R&D in Phase III consists of a series of demonstrator experiments  
1491 intended to demonstrate cavity CRES at a variety of scales and magnetic fields using  
1492 electrons from  $^{83m}\text{Kr}$ , an electron gun, and potentially molecular tritium sources. The  
1493 near-term cavity effort in Project 8 is the cavity CRES apparatus (CCA), which is a  
1494 small-scale cavity experiment operating near 26 GHz, that will perform the first CRES  
1495 measurements using a small cavity. This experiment will pave the way towards larger  
1496 scale cavity experiments in preparation for the eventual pilot-scale tritium experiment.

1497 The pilot-scale experiment is the first experiment, which will combine atomic tritium  
1498 and large-volume CRES detection in the same experiment. It will directly demonstrate  
1499 all the technologies required for Phase IV such that no technical risks remain for scaling  
1500 the experiment to required scale. A robust approach to scaling the pilot-scale experiment  
1501 is to simply build multiple copies of it for the Phase IV experiment.

## 1502 **Phase IV: Project 8’s Ultimate Neutrino Mass Experiment**

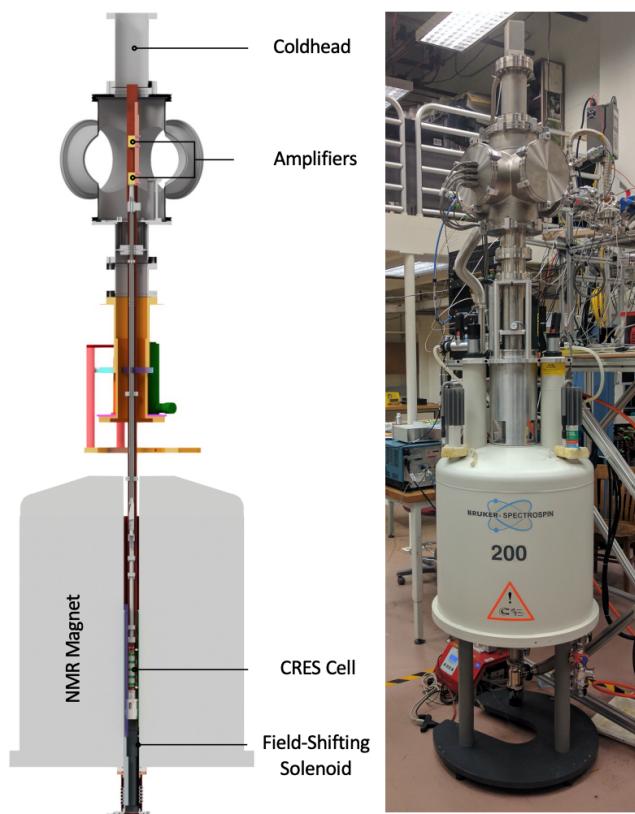
1503 The design of Phase IV should be a direct extension of the pilot-scale CRES experiment  
1504 that marks the official end of Phase III (see Section 3.5). The Phase IV experiment  
1505 represents the final experiment in the Project 8 neutrino mass measurement experiment  
1506 plan and will have sensitivity to neutrino masses of 40 meV.

## 1507 **3.3 Phase II: First Tritium Beta Decay Spectrum and** 1508 **Neutrino Mass Measurement with CRES**

1509 In Phase II Project 8 demonstrated the first ever measurement of the tritium beta-decay  
1510 spectrum endpoint using the CRES technique, which lead to the first neutrino mass  
1511 measurement by the Project 8 collaboration. This milestone was made possible by many

improvements in the CRES technique and in the understanding of CRES systematics, which takes an important first step towards larger scale measurements of the tritium 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 [42, 43].

### 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.

### 1520 Magnet and Cryogenics

1521 The magnetic field for the the Phase II experiment is provided by a nuclear magnetic  
1522 resonance (NMR) spectroscopy magnet with a central bore diameter of 52 mm (see Figure

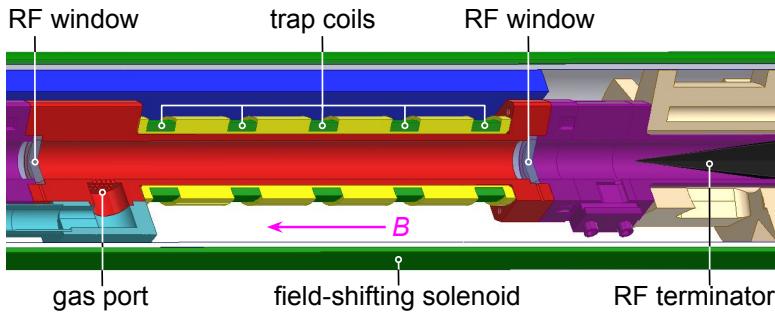
1523 3.7). The magnet produces a background magnetic field with an average value of 0.959 T  
1524 and a 10 ppm variation across the bore diameter achieved using several shim coils built  
1525 into the magnet. Using an external NMR field probe the variation of the magnetic field  
1526 along the vertical axis of the magnet bore was measured to obtain an accurate model of  
1527 the magnetic field so that the CRES cell could be positioned for optimal magnetic field  
1528 uniformity.

1529 An external solenoid magnet was installed inside the magnet bore to provide the  
1530 ability to shift the magnitude of the background magnetic field by values on the order of  
1531 a few mT. The solenoid has inside diameter of 46 mm and a length of 350 mm, which  
1532 terminates in a vacuum flange that allows it to be inserted into the NMR magnet bore  
1533 from the bottom. By shifting the value of the magnetic field by a few mT, the cyclotron  
1534 frequencies of electrons produced by the 17.8 keV  $^{83m}\text{Kr}$  internal-conversion line [57]  
1535 can be shifted over a range of frequencies on the order of 100 MHz. This allows one to  
1536 study the frequency dependent behavior of multiple CRES systematics such as detection  
1537 efficiency that directly affect the measured shape of the tritium spectrum.

1538 The inside of the magnet bore diameter was pumped down to a vacuum of less than  
1539 10  $\mu\text{torr}$  using a turbomolecular pump, which allows for cryogenic cooling of the CRES  
1540 cell and RF system. Cooling power was supplied to the Phase II apparatus using a  
1541 cryopump with its coldhead mounted above the primary magnet and CRES cell. This  
1542 arrangement allowed for sufficient cooling power to be delivered to the amplifiers to cool  
1543 them to a temperature of  $\approx$  40 K, while keeping the amplifiers far enough from the  
1544 magnet so as not to be damaged by the large field strength. Thermal contact between  
1545 the coldhead, amplifiers, RF system, and CRES cell is achieved using a copper bar that  
1546 runs the full length of the apparatus. To prevent freeze-out of  $^{83m}\text{Kr}$  on the walls of the  
1547 CRES cell a separate heater was installed to keep the CRES cell near a temperature of  
1548 85 K during the operation of the experiment.

## 1549 **CRES Cell**

1550 Located in the most uniform region of the magnetic field is the CRES cell, which is the  
1551 region of the apparatus where radioactive decays of  $^{83m}\text{Kr}$  and  $T_2$  emit electrons that can  
1552 be trapped and measured using CRES (see Figure 3.8). The CRES cell is manufactured  
1553 from a segment of cylindrical waveguide designed to operate at K-band frequencies  
1554 near 26 GHz. The diameter of the waveguide determines which resonant modes of the  
1555 waveguide will couple to the electron and transmit its radiation to the amplifiers. For  
1556 Phase II a waveguide diameter of 1 cm was selected, which allows electrons to couple to



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

1557 the TE<sub>11</sub> and TM<sub>01</sub> cylindrical waveguide modes. To reduce complexity in modeling and  
 1558 analyzing the CRES data, it is ideal to select a diameter that prevents electrons from  
 1559 coupling to higher-order waveguide modes beyond the fundamental TE and TM modes.

1560 Around the exterior of the cylindrical waveguide are several magnetic coils used to  
 1561 produce magnetic traps inside the CRES cell volume. Without a magnetic trap electrons  
 1562 produced from decays inside the CRES cell quickly impact the cell wall, which prevents  
 1563 a measurement of their cyclotron frequency using CRES. Each coil along the length of  
 1564 the waveguide produces a separate trap that is approximately harmonic in shape. By  
 1565 independently controlling the currents provided to each coil the traps can be configured  
 1566 to have equal values of the magnetic field at the trap bottom despite a non-uniform field  
 1567 from the NMR magnet.

1568 Two primary magnetic trap configurations were used during the Phase II experiment.  
 1569 The first was a shallow trap configuration used primarily for its high energy resolution to  
 1570 study systematics using <sup>83m</sup>Kr decays, and the second was a deeper trap that could trap a  
 1571 higher percentage of pitch angles. The trade-off with this trap is that the higher trapping  
 1572 efficiency comes at the cost of lower energy resolution due to the greater variation in pitch  
 1573 angle. The deep trap was the trap used to measure the tritium beta-decay spectrum in  
 1574 Phase II.

1575 The source gases were delivered into the CRES cell through a gas port located near the  
 1576 top end of the cylindrical waveguide. To prevent the gases from escaping the cell, vacuum  
 1577 tight RF transparent windows are needed to contain the tritium and krypton source  
 1578 gas across a 1 atm pressure differential, while still transmitting the cyclotron radiation  
 1579 without distortion. The crystalline material, CaF<sub>2</sub>, which has a thermal expansion  
 1580 coefficient similar to that of copper, was used for this purpose in the CRES cell. Two  
 1581 windows, each 2.4 mm thick, were used to seal off the ends of the CRES cell. The  
 1582 thickness of 2.4 mm corresponds to half of a cyclotron wavelength when one accounts for

1583 the permittivity of  $\text{CaF}_2$ .

1584 **RF System**

1585 The RF system in the Phase II apparatus transferred the cyclotron radiation from the  
1586 CRES cell to the receiver chain. The receiver chain performs the down-conversion and  
1587 digitization required to obtain signals that can be analyzed to determine the cyclotron  
frequencies of electrons in the CRES cell (see Figure 3.9).

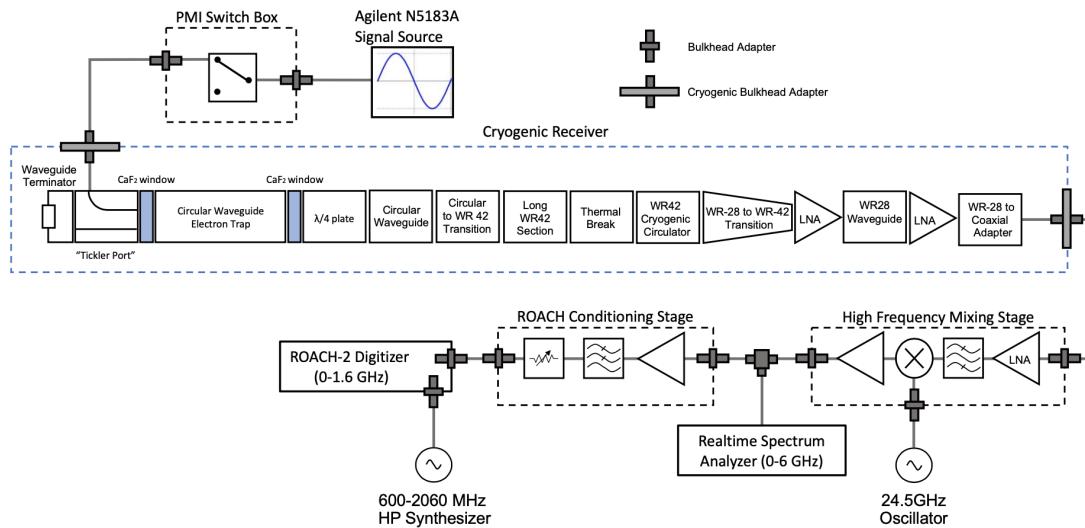


Figure 3.9. RF system diagram for the Phase II apparatus.

1588  
1589 Below the CRES cell, at the bottom of the Phase II apparatus, is a tickler port and  
1590 waveguide terminator. The tickler port is used to inject signals into the CRES cell and  
1591 RF system for testing and calibration purposes. The waveguide terminator is designed to  
1592 absorb cyclotron radiation emitted by electrons that transmits out of the bottom of the  
1593 CRES cell. This lowers the total power received from electrons in the CRES cell, since all  
1594 the energy radiated downwards is absorbed into the terminator. Earlier iterations of the  
1595 Phase II apparatus used an RF short in this location that reflected this power up towards  
1596 the amplifiers, however, interference between the upward traveling and reflected radiation  
1597 led to a disappearance in the signal carrier that made reconstruction impossible.

1598 Radiation traveling upward passes through the  $\text{CaF}_2$  window passes through a  $\lambda/4$   
1599 plate, which transforms the circularly polarized cyclotron radiation into linear polarization.  
1600 The linearly polarized fields next travel through a segment of circular waveguide that  
1601 transitions into a long segment of WR-42 waveguide that carries the fields out of the  
1602 high magnetic field region. A thermal break segment is included, which consists of a a

1603 segment of gold-plated stainless steel WR-42 waveguide, to help thermally isolate the  
1604 relatively warm CRES cell from the colder amplifiers. The radiation then passes through  
1605 a cryogenic circular, which prevents signals reflected from the amplifiers from interfering  
1606 with the CRES cell before a WR-42 to WR-28 transition connects the waveguide to the  
1607 first of the cryogenic amplifiers. The radiation passes through two cryogenic amplifiers  
1608 before being coupled to a coaxial termination at the top of the Phase II apparatus.

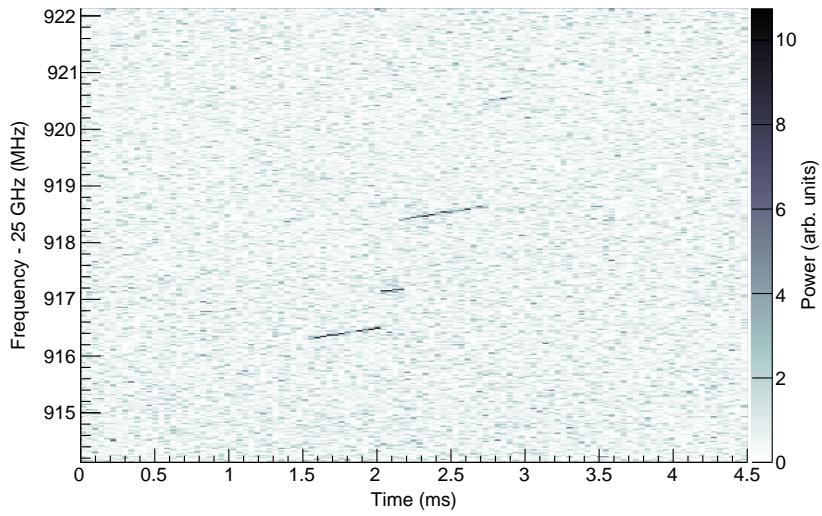
1609 The coaxial cable transfers the cyclotron radiation signals to a high-frequency mixing  
1610 stage that performs an analog frequency down-conversion using a 24.5 GHz LO. Two forms  
1611 of digitization can be used at this stage to readout the CRES data. One is a real-time  
1612 spectrum analyzer that digitizes the CRES signal data in time-domain and computes the  
1613 frequency spectrum in real-time, which allows for direct visualization of CRES signal  
1614 spectrograms as the experiment is running. The real-time spectrum analyzer is most  
1615 useful for taking small amount of streamed data for debugging and analysis of the system.  
1616 The other method, which was used to collect the majority of the CRES data in Phase II,  
1617 is a ROACH-2 FPGA and digitizer system. The ROACH system consists of a fast ADC  
1618 that samples the CRES signal data at 3.2 GSps. Internal digital down-conversion stages  
1619 implemented in the FPGA perform a mixing operation that reduces the bandwidth of the  
1620 CRES signals to 100 MHz. The FPGA implements a 8192 sample FFT and packetizes  
1621 time and frequency domain records in parallel. The packetized data is then transferred  
1622 from the ROACH to be analyzed by the data-processing pipeline.

### 1623 **3.3.2 CRES Track and Event Reconstruction**

#### 1624 **Time-Frequency Spectrogram**

1625 The online data-processing is intended to identify interesting data that could contain  
1626 CRES signals using a software real-time triggering algorithm. Interesting segments of  
1627 data identified by this algorithm are collected into files that are transferred to a server for  
1628 offline processing and analysis. The data files contain a continuous series of time-domain  
1629 samples, broken into a set of records, which are 4096 samples long. The time-series is  
1630 made up of 8-bit IQ samples acquired at 100 MHz.

1631 Each time-series record is accompanied by an associated frequency spectrum consisting  
1632 of 4096 frequency bins approximately 24.4 kHz wide, which is represented as a power  
1633 spectral density. The individual frequency spectra can be organized temporally to create  
1634 a time-frequency spectrogram that represents the evolution of the cyclotron frequency  
1635 spectrum over the course of the CRES event (see Figure 3.10). The time-frequency



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

1636 spectrogram is represented as a two-dimensional image where the color of each pixel is  
 1637 proportional to the power spectral density. Each vertical slice of pixels in the image  
 1638 represents a frequency spectrum, therefore, each horizontal bin represents the data  
 1639 obtained over a duration of  $4096 \times 0.01 \text{ MHz}^{-1} = 40.96 \mu\text{sec}$ .

#### 1640 **CRES Event Data Features**

1641 Phenomenologically, a CRES signal appears as a sinusoidal signal whose frequency slow  
 1642 increases ("chirps") over time. Axial motion of the electron in the trap leads to the  
 1643 formation of frequency sidebands that surround the more powerful carrier frequency, due  
 1644 to Doppler modulation of the electron's frequency as it bounces between the walls of the  
 1645 magnetic trap. The critical piece of information that must be extracted from the track  
 1646 and event reconstruction procedure is the carrier frequency, since it is this frequency  
 1647 that gives the cyclotron frequency and thus the kinetic energy. While axial motion from  
 1648 non- $90^\circ$  pitch angles does change the average magnetic field experienced by an electron  
 1649 and, therefore, changes the cyclotron frequency. Because of low-SNR sidebands were  
 1650 unable to be observed in Phase II, so a correction for the effect of the pitch angle on the  
 1651 cyclotron frequency was not possible.

1652 In the time-frequency spectrogram representation the chirping carrier frequency  
 1653 appears as a linear track of high-power frequency bins (see Figure 3.10). The vertical  
 1654 slope of the tracks is caused by the emission of energy from the electron in the form of

1655 cyclotron radiation, therefore, the size of the slope parameter is directly proportional  
1656 to the Larmour power. The continuous track is periodically interrupted by random  
1657 jumps to higher frequency and lower energy caused by random inelastic collisions with  
1658 background gas molecules. The length of a track is an exponentially distributed variable  
1659 whose mean value is inversely proportional to the gas density. The size of the frequency  
1660 discontinuities is directly proportional to the energies of the rotational and vibrational  
1661 states of background gas molecules.

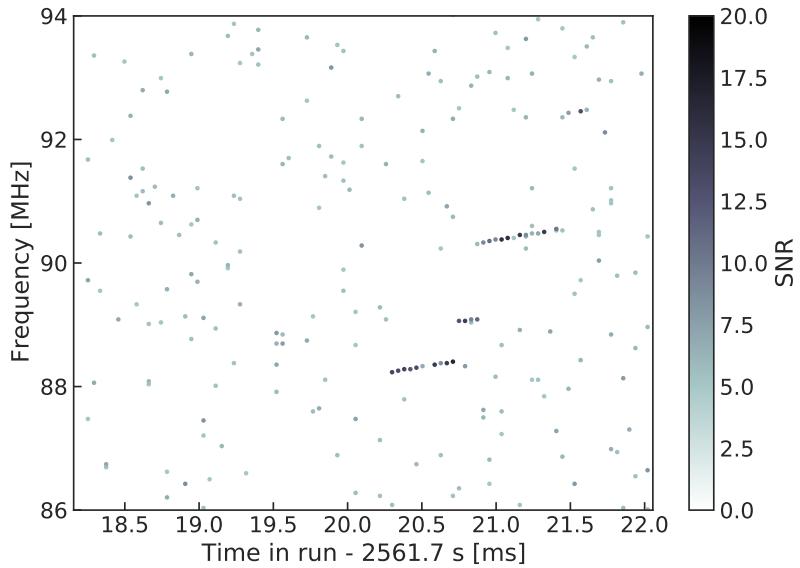
1662 A CRES event refers to the collection of tracks produced by a trapped electron until  
1663 it inevitably scatters into a pitch angle that can no longer be trapped. The goal of track  
1664 and event reconstruction is to first identify the set of tracks present in a time-frequency  
1665 spectrogram that represents a segment of data acquired in the Phase II apparatus. These  
1666 tracks must then be clustered into events from which we can determine the first track  
1667 produced by the electron and thus estimate it's starting cyclotron frequency and kinetic  
1668 energy.

## 1669 **Track Reconstruction**

1670 The first step in CRES event reconstruction is the identification of tracks in the time-  
1671 frequency spectrogram, which is essentially an image processing task. Track finding  
1672 starts by normalizing the power spectral density based on the average noise power.  
1673 Next a power threshold is applied to the normalized spectrogram where only bins that  
1674 have a signal-to-noise ratio greater than five are selected to build tracks. In this case  
1675 signal-to-noise ratio is defined as the ratio between the normalized, unitless power of a  
1676 bin divided by the average normalized power across the full frequency spectrum.

1677 The sparse spectrogram produced by this power cut consists only of a sparse collection  
1678 of high-power frequency bins that could be part of a CRES signal track (see Figure  
1679 3.11). In this form is it much easier to identify tracks "by eye", however, for the Phase II  
1680 analysis Project 8 developed its own custom-made track finding algorithm, called the  
1681 sequential track finder (STF).

1682 The STF algorithm processes the sparse spectrogram in sequential fashion, processing  
1683 each time-slice one-by-one until the end of the spectrogram is reached. Tracks are found  
1684 by searching for points in the sparse spectrogram that appear to fall on a straight line.  
1685 Multiple configurable parameters are built into the STF algorithm that allow the user to  
1686 tune the criteria for adding a point to an existing track or creating a new track. These  
1687 include parameters such as maximum time and frequency differences between subsequent  
1688 points in a track as well as minimum SNR values for the start and endpoints of the track.



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

1689 Additionally, tracks are required to have a minimum length and slope to be considered  
 1690 potential CRES tracks rather than random noise fluctuations.

1691 The resulting output of the STF is a collection of track objects that consist of the track  
 1692 point objects and their properties. The final step is to calculate track-level properties  
 1693 and apply cuts to reject false tracks found by the STF. This involves the fitting of a  
 1694 line to the collection of track points as well as the total and average power of the track  
 1695 obtained by computing the sum and mean of the points powers. The starting frequency  
 1696 of the track is determined by calculating the time coordinate that intersects with the  
 1697 linear fit. A cut is performed to remove all tracks that do not have a specified average  
 1698 power over their duration, which helps to remove the majority of noise fluctuations that  
 1699 have passed all previous cuts up to this point.

## 1700 Event Reconstruction

1701 After track reconstruction comes event reconstruction where the identified tracks are  
 1702 grouped into events that correspond to the trajectory of a single electron in the trap. This  
 1703 procedure attempts to match tracks head to tail by checking if the start and end times  
 1704 of a pair of tracks falls within a certain tolerance. This tolerance is a configurable  
 1705 parameter that can be tuned to an optimal value using Monte Carlo simulations of events  
 1706 in the Phase II apparatus.

1707 After the event building procedure has completed there remains a small likelihood  
1708 that false tracks have made it through to the event reconstruction stage. Typically, cuts  
1709 at the track level are able to remove 95% of the false tracks identified by the STF, which  
1710 leads to a significant number of false tracks at the event building stage. However, the  
1711 additional event-level information makes it possible to reject events that contain these  
1712 false tracks with a high degree of confidence.

1713 Two event level features are associated with events caused by real electrons — the  
1714 duration of the first track as well as the number of tracks in the event. Real electrons  
1715 tend to have event structures with longer first tracks and a higher number of total tracks.  
1716 Based on the values of these two criteria, a minimum threshold on the average power in  
1717 the first track was configured to reject false events. The average power in the first track  
1718 was chosen due to the critical nature of the starting frequency of the first track in an  
1719 event to the krypton and tritium spectrum analyses.

### 1720 **3.3.3 Results from Phase II**

1721 The main result from Phase II was the measurement of the tritium beta-decay spectrum  
1722 using CRES, which lead to the first neutrino mass limit with CRES. However, Phase  
1723 II also included a significant  $^{83m}\text{Kr}$  measurement campaign to understand important  
1724 systematics relevant to the tritium spectrum measurement, but also to understanding the  
1725 fundamentals of the CRES technique itself. This required high-resolution measurements  
1726 of the  $^{83m}\text{Kr}$  internal-conversion spectrum [57], which is an interesting science result in  
1727 its own right.

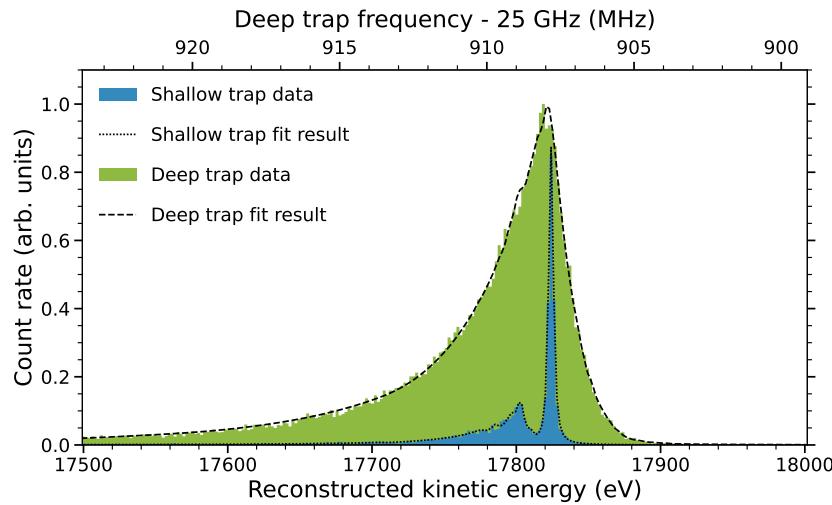
1728 The results from Phase II represents a significant effort from the entire Project 8  
1729 collaboration over several years. Because the focus of my contributions to Project 8 is  
1730 directed towards the research and development efforts for the Phase III experiments, the  
1731 goal in this section is not to provide a detailed description of the analyses that lead to  
1732 the Phase II results. Rather, I will provide brief descriptions of a few plots representative  
1733 of the main results from Phase II.

### 1734 **Measurements with Krypton**

1735 Measurements with krypton were a key calibration tool for Phase II of the experiment and  
1736 will continue to be useful in Phase III. In the context of Project 8 krypton measurements  
1737 refers to CRES measurements of the internal-conversion spectrum of the metastable state  
1738 of krypton-83,  $^{83m}\text{Kr}$ , produced by electron capture decays of  $^{83}\text{Rb}$ . A supply of  $^{83}\text{Rb}$

<sup>1739</sup> was built into the Phase II apparatus gas system that supplied the CRES cell with  $^{83m}\text{Kr}$   
<sup>1740</sup> via emanation.

<sup>1741</sup> The  $^{83m}\text{Kr}$  internal-conversion spectrum consists of several lines based on the orbital  
<sup>1742</sup> of the electron ejected during the decay. The conversion lines useful to Project 8 are  
<sup>1743</sup> those that emit electrons with kinetic energies that fall inside the detectable frequency  
<sup>1744</sup> bandwidth of the Phase II apparatus. These are the K; L2 and L3; M2 and M3; and N2  
<sup>1745</sup> and N3 lines with kinetic energies of 17.8 keV,  $\approx$  30.4 keV,  $\approx$  31.9 keV, and  $\approx$  32.1 keV,  
<sup>1746</sup> respectively. The different energies of the lines allow a onw to test the linearity of the  
<sup>1747</sup> relationship between kinetic energy and frequency across the range of frequencies covered  
<sup>1748</sup> by the continuous tritium spectrum.

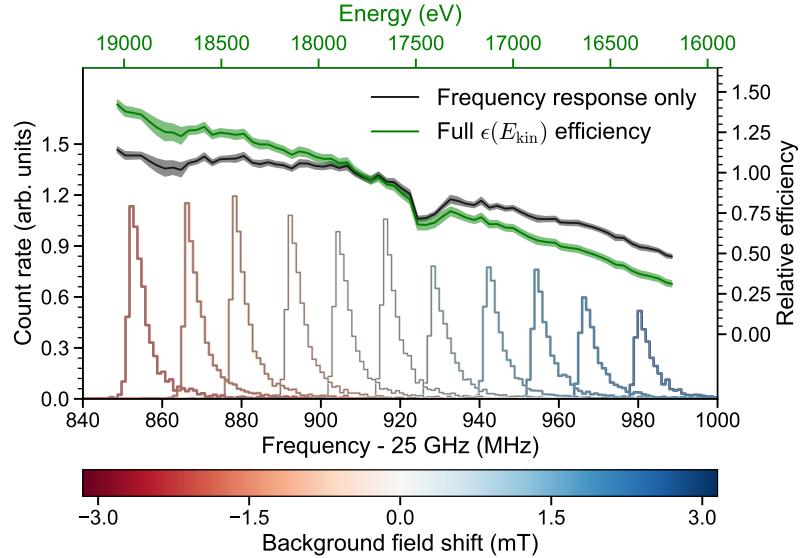


**Figure 3.12.** Fits to the measured 17.8-keV  $^{83m}\text{Kr}$  conversion line using the deep and shallow trap configurations.

<sup>1749</sup> Numerous detector related effects relevant to the tritium analysis can be characterized  
<sup>1750</sup> by measuring the shape of the krypton spectrum. Specific examples include variations  
<sup>1751</sup> in the magnetic field as a function of the radial position of the electron, variation in  
<sup>1752</sup> the magnetic field caused by the trap shape, variation in the average magnetic field for  
<sup>1753</sup> electrons with different pitch angles, and the effect of missing tracks due to scattering.  
<sup>1754</sup> These spectrum shape measurements focused on the 17.8-keV krypton line and utilized  
<sup>1755</sup> different trap geometries based on the particular goal of the dataset (see Figure 3.12).

<sup>1756</sup> Krypton measurements with a shallow trap allow for high energy resolution, since  
<sup>1757</sup> variation in frequency due to pitch angle differences is sharply reduced in the shallow  
<sup>1758</sup> trap configuration. With this trap the main 17.8-keV peak of the conversion spectrum is  
<sup>1759</sup> clearly visible along with additional satellite peaks at lower energy, which correspond to

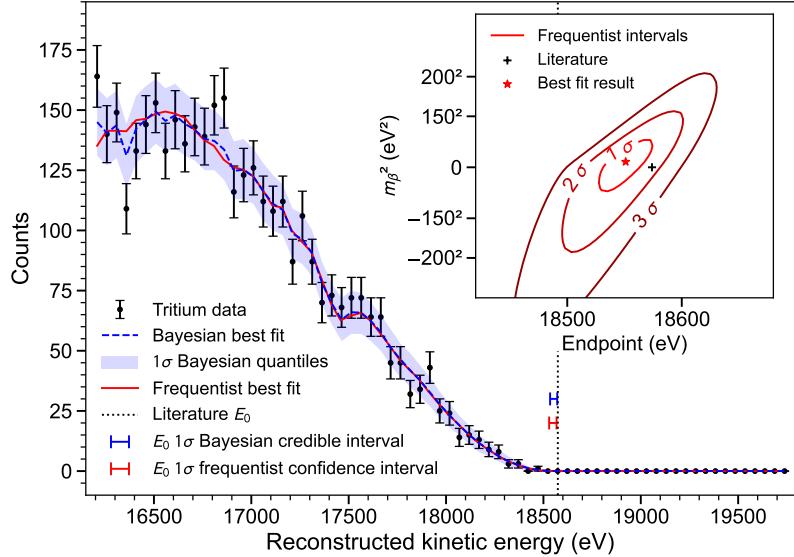
1760 the shakeup/shakeoff spectrum of the decay. The high accuracy of the fit demonstrates a  
 1761 high degree of understanding of the CRES systematics.



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

1762 The broadening of the krypton spectrum seen for the deeper track is due to the large  
 1763 range of electron pitch angles that can be trapped. Furthermore, with a deeper trap  
 1764 there is a larger parameter space of electron that could be produced with pitch angles  
 1765 that are trappable but not visible in the time-frequency spectrogram. These electrons  
 1766 live in the trap and can scatter multiple times before randomly scattering to a visible  
 1767 pitch angle. This leads to one or more missing tracks earlier in the event, which leads to  
 1768 a misreconstruction of the true starting frequency. By measuring the krypton spectrum  
 1769 shape in the same trap used to detect tritium events, the effect this has on the spectrum  
 1770 shape can be characterized to mitigate its impact on the tritium measurements.

1771 Changes in the Krypton spectrum shape as a function of CRES frequency were  
 1772 used to study the detection efficiency of the Phase II apparatus. Variations in the  
 1773 detection efficiency as a function of frequency directly influences the measured shape of  
 1774 the continuous tritium spectrum, which can lead to errors in the neutrino mass estimate  
 1775 if not modeled appropriately. Using the field shifting solenoid the cyclotron frequency  
 1776 of the krypton 17.83 keV line was shifted across the full frequency range of the tritium  
 1777 spectrum data (see Figure 3.13). Variations in the deep trap krypton spectrum shape  
 1778 can be used to infer the detection efficiency as a function of frequency and correct for



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

1779 this affect in the tritium measurements.

#### 1780 **Tritium Spectrum and Neutrino Mass Results**

1781 The tritium measurement campaign resulted in the collection of 82 days of detector  
 1782 live time during which 3770 total tritium events were detected. The track and event  
 1783 reconstruction analysis extracted the starting frequencies of these tritium events, which  
 1784 were used to build a frequency spectrum of tritium beta-decays. The resulting frequency  
 1785 spectrum was then converted to an energy spectrum using the information gleaned from  
 1786 the krypton measurement campaign to obtain the tritium beta-decay spectrum (see  
 1787 Figure 3.14).

1788 CRES is inherently a very low background technique with the dominant source of  
 1789 noise being random RF fluctuations. Monte Carlo simulations backed validated using  
 1790 measurements of the RF noise background were used to set track and event cuts to  
 1791 guarantee that zero false events would occur over the duration of the experiment with  
 1792 90% confidence. Notably, the measured spectrum has zero events beyond the tritium  
 1793 spectrum endpoint, which allows us to constrain the background rate in the Phase II  
 1794 apparatus to less than  $3 \times 10^{-10}$  counts/ev/s. Achieving a low background is critical for  
 1795 future neutrino mass experiments that seek to measure the neutrino mass with less than  
 1796 100 meV sensitivity.

1797 Bayesian and frequentist based fits to the measured tritium spectrum, incorporating

1798 information gained about CRES systematics from the krypton measurements, were  
1799 performed to extract upper limits on the tritium beta-decay spectrum endpoint as well as  
1800 the neutrino mass. The estimated spectrum endpoints are  $18553^{+18}_{-19}$  eV for the Bayesian  
1801 analysis and  $18548^{+19}_{-19}$  eV for the frequentist analysis. The quoted uncertainties are  
1802  $1-\sigma$ , and both results are within  $2-\sigma$  of the literature endpoint value of 15574 eV. The  
1803 estimated neutrino mass for both results is consistent with  $m_\beta^2 = 0$ . The 90% confidence  
1804 upper limits for the Bayesian analysis is  $m_\beta < 155$  eV/c<sup>2</sup> and  $m_\beta < 152$  eV/c for the  
1805 frequentist analysis.

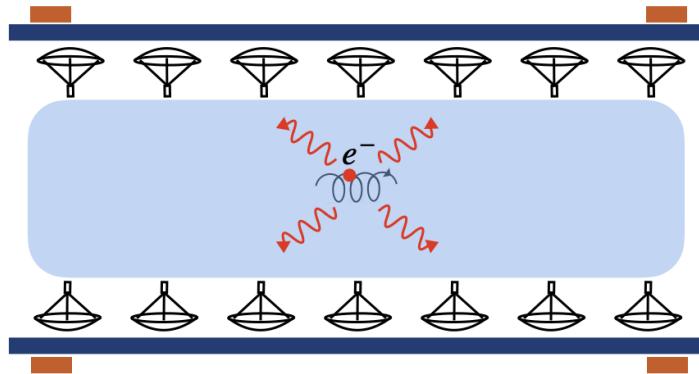
1806 Though the neutrino mass results from Phase II are not competitive with KATRIN  
1807 the experiment was a promising first step towards the development of more precise  
1808 neutrino mass measurements using CRES. The low-background and high-resolution  
1809 achievable with krypton measurements are promising features of the technique that were  
1810 demonstrated with the Phase II apparatus. As new technologies are developed to enable  
1811 CRES measurements in larger volume, many of the lessons learned from Phase II will  
1812 continue to influence the operation and design of future experiments.

## 1813 **3.4 Phase III R&D: Antenna Array CRES**

1814 The goal of Phase III in the Project 8 experimental program is to develop the technologies  
1815 and expertise required to build an experiment that uses CRES to measure the neutrino  
1816 mass with a target sensitivity of 40 meV. One of the key technologies is a method for  
1817 performing high resolution CRES measurements in a large volume, which allows one to  
1818 observe a sufficient quantity of tritium to measure the low-activity endpoint region of  
1819 the tritium spectrum.

### 1820 **3.4.1 The Basic Approach**

1821 One possible approach, suggested in the original CRES publication [40], is to use many  
1822 antennas to surround a volume of tritium gas in a magnetic field (see Figure 3.15). When  
1823 a decay occurs the electron will begin to emit cyclotron radiation that can be collected  
1824 by the array and used to perform CRES. Each antenna in the array collects only a small  
1825 fraction of the electron's signal power, which is less than 1 fW for a 18.6 keV kinetic  
1826 energy electron in a 1 T magnetic field. Scaling to large volumes with the antenna  
1827 array approach is accomplished by increasing the number of antennas in the array, which  
1828 increases the volume under observation proportionally, so that a sufficient population of



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

1829 tritium atoms can be observed to measure the tritium spectrum endpoint shape.

1830 Several features of the antenna array approach make it an attractive candidate technol-  
 1831 ogy for a large volume experiment. One example is the accurate position reconstruction  
 1832 made possible by the multichannel nature of the array. Using techniques like digital  
 1833 beamforming it is possible to estimate the radial and azimuthal positions of the electron  
 1834 in the magnetic trap with a precision significantly less than the size of the cyclotron  
 1835 wavelength. This capability allows one to perform event-by-event estimations of the  
 1836 magnetic field experienced by an electron, which is crucial to achieving high energy  
 1837 resolution with the CRES technique.

1838 The easy availability of position information with the antennas array approach  
 1839 is potentially a unique advantage that provides significant flexibility in the magnetic  
 1840 field uniformity requirements compared to other proposed approaches to large volume  
 1841 CRES (see Chapter 6). Spatial discrimination using digital beamforming leads to pileup  
 1842 reduction, which helps to reduce the potential of background events caused by missing  
 1843 tracks or by incorrectly clustering a group of tracks into an event. Limits on the  
 1844 background rate for a neutrino mass measurement with 40 meV sensitivity are stringent  
 1845 and the total activity of the tritium source for such an experiment is gigantic relative to  
 1846 the activity near the endpoint. Thus, pileup discrimination could be an important tool  
 1847 for a large scale CRES experiment.

1848 Another beneficial quality of the antenna array approach is that the volume of the  
 1849 experiment can be scaled independent of frequency by simply adding more antennas to  
 1850 the array (see Figure 3.19). Resonant cavities, the proposed alternative large volume  
 1851 CRES technology, are ideally operated in magnetic fields that cause electrons to move  
 1852 with cyclotron frequencies near the fundamental cavity resonance, to avoid complex  
 1853 coupling of the electron to many cavity modes simultaneously. This leads to a coupling

1854 between the cavity volume and the magnetic field magnitude, which forces one to lower  
1855 the magnetic field in order to increase the experiment scale. Whereas, for antenna arrays,  
1856 in principle there is no physical limitation on the size of the antenna array that can be  
1857 used at a particular magnetic field. However, the nature of scaling an antenna array  
1858 based experiment leads to rapidly increasing cost and complexity due to the large number  
1859 of antennas, amplifiers, and data streams that require substantial computer processing  
1860 power to effectively analyze.

### 1861 **3.4.2 The FSCD: Free-space CRES Demonstrator**

1862 The complexity of the antenna array CRES technique requires the construction of a  
1863 small scale demonstration experiment to develop an understanding of technique itself and  
1864 relevant systematics. Without a demonstrator experiment it is not possible to sufficiently  
1865 retire the technical risks associated with the full-scale experiment. Therefore, Phase  
1866 III of the Project 8 experimental program is primarily focused on the development and  
1867 operation of demonstrator experiments to inform the design of the Phase IV experiment.

1868 The demonstrator experiment developed for antenna array CRES in Phase III is called  
1869 the Free-space CRES Demonstrator or FSCD. The FSCD is intended as a demonstration  
1870 of antenna array CRES, but is also a capable neutrino mass measurement experiment  
1871 in its own right, with a target neutrino mass sensitivity of a few eV using a molecular  
1872 tritium source.

#### 1873 **Magnetic Field**

1874 The background magnetic field for the FSCD is provided by a hospital-grade MRI magnet  
1875 (see Figure 3.16). The magnet produces a magnetic field of approximately 0.958 T, which  
1876 corresponds to a tritium spectrum endpoint frequency of approximately 25.86 GHz. The  
1877 magnet is installed in the Project 8 laboratory located at the University of Washington,  
1878 Seattle, and is shimmed to produce a uniform magnetic field with variations on the  
1879 ppm-level. Measurements of the magnetic field non-uniformities are performed using a  
1880 NMR probe and rotational gantry to capture measurements of the magnetic field around  
1881 an elliptical surface in the center of the MRI magnet. During the operation of the FSCD  
1882 an array of Hall or NMR magnetometers would be used to periodically measure the  
1883 magnetic field to monitor its time stability.

1884 Inside the field of the MRI magnet additional electromagnets would be installed that  
1885 provide the capability to shift the value of the background magnetic field and produce



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

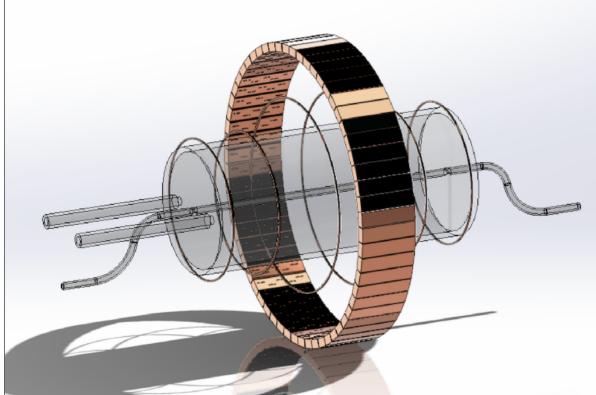
1886 a magnetic trap. Shifting the background magnetic field by a few  $\mu$  T lets one control  
 1887 the cyclotron frequencies of electrons with a fixed kinetic energy, which is key to an  
 1888 effective calibration of the FSCD. The preferred calibration method for the FSCD is  
 1889 a mono-energetic electron gun that can inject electrons into the magnetic trap with a  
 1890 known kinetic energy. In combination with the field shifting magnet, one can vary the  
 1891 cyclotron frequencies of the electrons to measure the response of the antenna array as a  
 1892 function of the radiation frequency and electron position. This procedure characterizes  
 1893 the response of the antenna array and provides further information on magnetic field  
 1894 uniformity, which is important to achieving good energy resolution.

1895 The design of the magnetic trap is absolutely critical to the success of a CRES  
 1896 experiment. The ideal shape is the perfect magnetic box, which has a flat bottom and  
 1897 step function walls. Any variation in the average magnetic field experienced by an  
 1898 electron leads to changes in the cyclotron frequency that can make determining the true  
 1899 starting kinetic energy more difficult. This includes changes in the magnetic field caused  
 1900 by the walls of the magnetic trap as well as radial magnetic field variations.

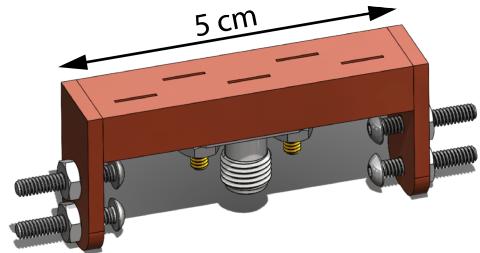
1901 The ideal box trap is completely uniform and has infinitely steep walls that cause  
 1902 no change in the electron's cyclotron frequency as it is reflected from the trap wall,  
 1903 however, such a trap cannot be made from any combination of magnetic coils since it  
 1904 violates Maxwell's equations. One of the goals of magnetic trap design is to identify the  
 1905 configuration of coils that produces a trap that approximates the perfect box trap as  
 1906 closely as possible.

1907 **Antenna Array**

1908 The canonical antenna array design for CRES is a uniform cylindrical array of antennas  
1909 that surrounds the magnetic trap volume. Since the FSCD is a demonstrator experiment,  
1910 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



(a)



(b)

**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.

1911  
1912 circle are sixty slotted waveguide antennas that fully populate the available space around  
1913 the array circumference. In order to maximize the power collected from each electron  
1914 it is optimal to cover as large a fraction of the solid angle around the magnetic trap as  
1915 possible.

1916 The distance between antennas around the circumference of the array is proportional  
1917 to the wavelength of the cyclotron radiation. Therefore, maximizing the solid angle  
1918 coverage of the array, while minimizing channel count to keep the hardware and data  
1919 acquisition costs manageable, biases one towards smaller array diameters. Antenna  
1920 near-field effects limit the minimum diameter of the array for a given antenna design  
1921 since the radiation from electrons that are too close to the array cannot be detected  
1922 due to destructive interference caused by path-length differences from the electron to  
1923 different points on the antenna surface.

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 can be performed with a coaxial cable connected  
1927 at the center or on either end of the waveguide. One of the drawbacks of waveguide

1928 antennas is the large amount of space required to fit them inside the limited MRI magnet  
1929 volume. Alternative antenna designs, constructed from microstrip printed circuit boards  
1930 require significantly less space at the cost of slightly higher energy loss in the antenna  
1931 structure.

1932 The FSCD antenna design is a 5 cm long segment of WR-34 waveguide with 5 vertical  
1933 slots cut into the side. The distance between slots along the length of the waveguide is  
1934 a half wavelength for optimal power combination between the individual antenna slots.  
1935 Each slot is offset from the center of the antenna face a small distance in order to most  
1936 effectively couple the slot to waveguide modes inside the antenna.

1937 The passive power combination achieved by placing 5 slots in a single waveguide is a  
1938 compromise intended to reduce the cost and complexity of the antenna array system.  
1939 Each additional channel in the array requires it's own cryogenic amplifier and also increase  
1940 the required computer power to process the raw data collected by digitizing each channel.  
1941 Passive summation, achieved by combining antennas into arrays axially, reduces the array  
1942 channel count at the cost of losses from imperfect passive combination. Imperfect passive  
1943 combination is caused by effects such as re-radiation of energy from and destructive  
1944 interference between slots in the waveguide antenna.

1945 Interference and re-radiation eventually limit the achievable the axial extent of passive  
1946 power combination. The 5-slot designed developed for the FSCD is optimized to minimize  
1947 the impact of these losses while achieving the maximum amount of axial coverage with a  
1948 single ring of antennas. Scaling beyond the volume covered by a single ring of antennas is  
1949 achieved by stacking additional rings of antennas together to cover a larger trap volume  
1950 for a higher statistics measurement of the tritium spectrum endpoint region. A likely  
1951 scenario for the FSCD experiment involves a staged experiment approach, where first  
1952 a series of measurements is performed using only a single ring of antennas followed by  
1953 experiments that add additional rings to the FSCD. The goal would be to first understand  
1954 the principles of antenna array CRES using the simplest possible experiment, before  
1955 attempting to scale the technique by expanding the antenna array size.

## 1956 **Tritium Source**

1957 While the primary purpose of the FSCD is as a technology demonstrator, it is unlikely  
1958 for the collaboration to gain the required confidence in the antenna array CRES tech-  
1959 nique to perform neutrino mass measurements at the 40 meV sensitivity level without  
1960 an intermediate scale measurement of the neutrino mass using antenna array CRES.  
1961 Therefore, the FSCD has an additional scientific goal of measuring the neutrino mass

1962 with a rough sensitivity goal of a few eV. This level of precision is achievable using a  
1963 source of molecular tritium with a volume of approximately 1 L at a density comparable  
1964 to potential Phase IV scenarios.

1965 Unlike previous CRES experiments, where the tritium source could be co-located  
1966 with the receiving antenna inside a waveguide transmission line, the tritium source  
1967 in the FSCD is thermally isolated from the antenna array to avoid freeze-out of the  
1968 tritium molecules. The tiny radiation power emitted by electrons requires a system noise  
1969 temperature of  $\approx 10$  K or less, in order to detect events at a high enough efficiency to  
1970 reach the neutrino mass sensitivity goals of the experiment. Achieving a system noise of  
1971 10 K requires that the antenna array and amplifiers operate at cryogenic, liquid helium  
1972 temperatures of  $\approx 4$  K, which significantly lowers the vapor pressure of molecular tritium.  
1973 By keeping the molecular tritium isolated in an RF-transparent vessel the tritium gas can  
1974 be kept at a relatively warmer temperature in the range of 30 K to avoid the accumulation  
1975 of tritium on the experiment surfaces.

## 1976 Data Acquisition and Reconstruction

1977 A fundamental change in the data acquisition system for the FSCD is the shift from  
1978 single to multi-channel reconstruction. This transition results in a significant increase in  
1979 the data-generation rate, which is linearly related to the number of independent channels  
1980 in the array. The larger data volume coincides with an increased demand for computer  
1981 processing power based on the need for more precise signal reconstruction algorithms  
1982 driven by the FSCD and Phase IV sensitivity goals. Therefore, the data acquisition  
1983 system for the FSCD is likely to represent a significantly larger fraction of the experiment  
1984 cost and complexity than previous CRES experiments.

1985 Each antenna in the array is connected to a cryogenic amplifier and down-converted  
1986 from the 26 GHz CRES frequency using an IQ-mixer to reduce the size of the analysis  
1987 window in which the tritium spectrum is measured. Using an LO with a frequency of  
1988 approximately 25.80 GHz the antenna array signals can be digitized at a rate of 200 MHz,  
1989 which is sufficient bandwidth to resolve the complete sideband spectrum produced by  
1990 axial oscillations of electrons in the FSCD magnetic trap.

1991 Direct storage of the raw FSCD antenna array data is undesirable, since the estimated  
1992 amount of raw data generated is  $O(1)$  exabyte per year. The management and storage  
1993 of such a large dataset is infeasible for a demonstrator experiment on the scale of the  
1994 FSCD and would represent a large fraction of the budget for a Phase IV scale antenna  
1995 array based CRES experiment. Therefore, a sub-goal of the FSCD experiment is the

1996 development of real-time reconstruction methods that could reduce the raw data volume  
1997 by detecting and reconstructing CRES events in real-time. The ultimate goal would be  
1998 a complete real-time reconstruction pipeline that takes raw voltages samples from the  
1999 antenna array and returns estimates for the starting kinetic energies of CRES events in  
2000 the data.

2001 The feasibility of a real-time reconstruction pipeline rests on the development of  
2002 computationally efficient algorithms that can be implemented without the need for  
2003 enormous computing resources. One challenge with the antenna array approach is that  
2004 the small radiation power of a single electron is distributed between each channel in  
2005 the array, such that reconstruction using only the information in a single channel is not  
2006 possible. Therefore, the simply performing the initial step in reconstruction — signal  
2007 detection — requires orders of magnitude more computational power than previous CRES  
2008 experiments. This operation will then be followed by other, potentially more expensive,  
2009 reconstruction steps that are required in order to determine the kinetic energy of the  
2010 electron.

## 2011 **3.5 Pilot-scale Experiments**

### 2012 **3.5.1 Choice of Frequency**

2013 The optimal CRES frequency for Project 8 is that which can reach our target sensitivity  
2014 of 40 meV, while minimizing the cost and complexity of the overall experiment. The  
2015 magnitude of the background magnetic field determines the cyclotron frequency, which  
2016 affects the entirety of the CRES detection system design, specifying the operating  
2017 frequency of the CRES experiments is one of the first steps towards developing a full  
2018 design.

### 2019 **Scaling Laws**

2020 The Phase I and II experiments utilized a background magnetic field of 0.959 T provided  
2021 by an NMR magnet. This magnetic field was selected primarily for convenience, however,  
2022 the cyclotron frequencies for electrons near the tritium endpoint in a 0.959 T field ranges  
2023 from 25 to 26 GHz, which is within the standard RF Ka-band. Therefore, microwave  
2024 electronics specialized for these frequencies are easily obtainable for relatively low cost.  
2025 The operating frequency for the large-scale experiments must be selected in a more  
2026 rigorous manner due to the increased scale and complexity of the systems as well as the

2027 requirements of the 40 meV neutrino mass science goal.

2028 There is a bias towards lower frequencies in a large-volume experiment, due to the  
2029 direct relationship between wavelength and the physical size of the compatible RF  
2030 components like antennas and cavities. With a longer wavelength more volume can  
2031 be surrounded by an array with fewer antennas, which reduces hardware and data-  
2032 processing costs. Additionally, the size of a cavity experiment is directly proportional  
2033 to the wavelength since this sets the physical dimensions of the cavity. Furthermore,  
2034 it is easier to engineer a magnet that provides a uniform magnetic field across several  
2035 cubic-meters of space at lower magnetic fields, which provides advantages in terms of  
2036 cost-reduction and field uniformity.

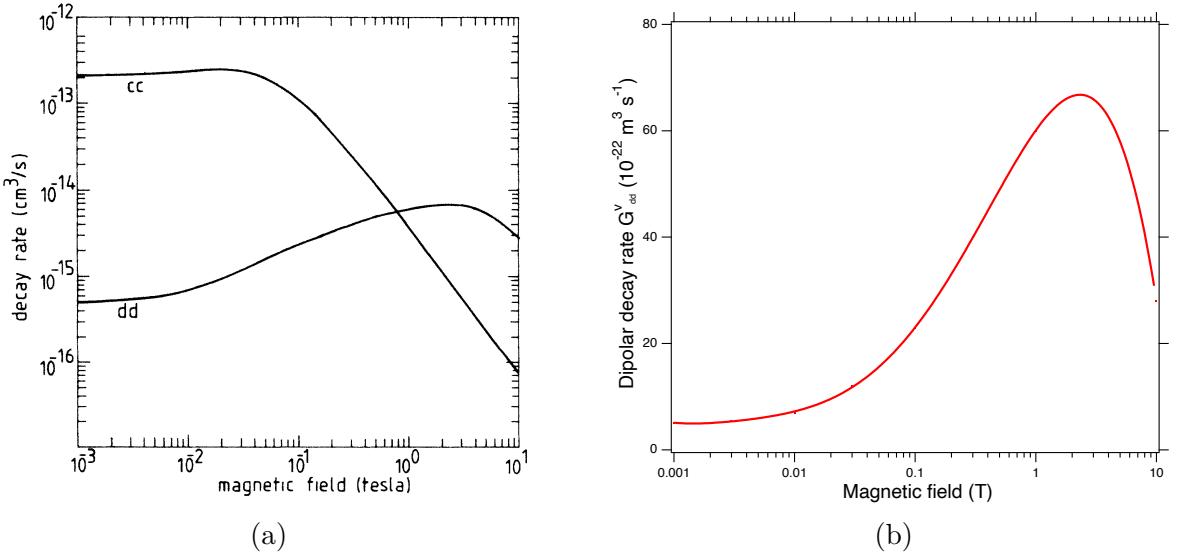
2037 A concern with lower magnetic fields and frequencies is the scaling of the Larmour  
2038 power equation, which is proportional to the square of the frequency. Naively, one would  
2039 predict that the SNR would decrease with lower fields, however, two additional scaling  
2040 laws that affect the noise power also come into play. Noise power is directly proportional  
2041 to the required bandwidth, which decreases linearly with the magnetic field. Furthermore,  
2042 at lower frequencies it is possible to purchase amplifiers with lower noise temperatures  
2043 until approximately 300 MHz at which point this relationship tends to flatten. Therefore,  
2044 it is expected that the SNR remains approximately constant as the frequency decreases.

2045 The SNR directly impacts the overall efficiency of the experiment through its effects  
2046 on signal detection and energy resolution. Thus, the expectation that SNR remains the  
2047 same at lower frequencies clearly biases large-scale experiments in this direction. One  
2048 drawback of lower magnetic fields is the increased influence of external magnetic fields  
2049 on the experiment. This includes magnetic fields from the building materials as well as  
2050 variations in the earth's magnetic field. To deal with these affects a suitable magnetic  
2051 field correction system will need to be devised, which includes constant monitoring of  
2052 external fields.

## 2053 **Atomic Tritium Considerations**

2054 The pilot-scale experiments will be the first Project 8 experiments to combine CRES with  
2055 atomic tritium, therefore, the optimal frequency should take into account the affect of the  
2056 background magnetic field on the atom trap. The primary influence of the background  
2057 field magnitude is through the rate of dipolar spin-flips caused by a spin exchange  
2058 interaction between trapped atoms [58].

2059 Atomic tritium is a simple quantum system with a hyperfine structure given by the  
2060 addition of the nuclear and atomic spins. The addition of two spins leads to a hyperfine

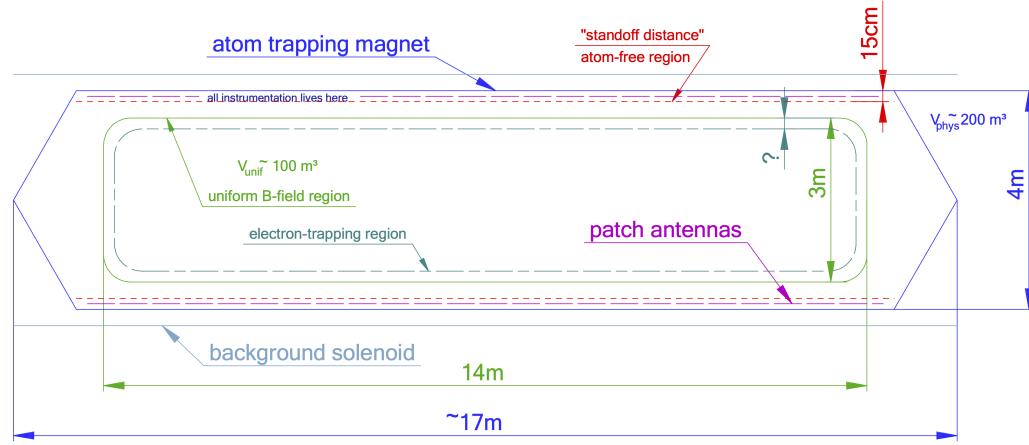


**Figure 3.18.** (a) A plot of the decay rate for the two-body dipolar spin exchange interaction for cc and dd state. (b) A plot of the decay rate of the dipolar spin exchange interaction for d+d states as a function of magnetic field magnitude. Lowering the magnetic field is key for reducing the losses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

2061 structure with four states in the  $(m_s, m_I)$  basis [59]. The states with atomic spins directed  
 2062 anti-parallel to the magnetic field have  $m_s = -1/2$  and are labeled as the a and b states.  
 2063 The a and b states are colloquially known as high-field seeking states, since their energy is  
 2064 minimized when in regions of higher magnetic field. This leads to losses in the magnetic  
 2065 trap as these atoms are drawn to higher fields away from the trap center. Alternatively,  
 2066 the c and d states, with atomic spin  $m_s = +1/2$ , minimize their energy in low magnetic  
 2067 fields because of the parallel alignment between spin and the magnetic field. Therefore,  
 2068 these low-field seeking states tend to stay trapped significantly longer than the high-field  
 2069 seeking states.

2070 It would be advantageous to prepare tritium atoms in purely c and d states before  
 2071 trapping, however, even in this case losses still occur due to dipolar interactions between  
 2072 pairs of c and d states leading to flipped atomic spins and subsequent losses from high-field  
 2073 seeking atoms. The rate of these interactions depends on the magnitude of the background  
 2074 magnetic field and is maximal for dd interactions around 1 T (see Figure 3.18). The rate  
 2075 of losses from these interactions at 1 T requires atomic tritium production at a rate two  
 2076 orders of magnitude larger than at 0.1 T, thus, requirements on the whole atomic tritium  
 2077 system are significantly relaxed at lower magnetic fields, which provides an additional  
 2078 argument for transitioning to lower frequencies with the pilot-scale experiments.

2079 **3.5.2 Pilot-scale Experiment Concepts**



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

2080 While the pilot-scale experiments are still in the early stages, enough is known to  
2081 sketch the general features of these experiments at the conceptual level.

2082 **Pilot-scale Antenna Array CRES Experiment Concept**

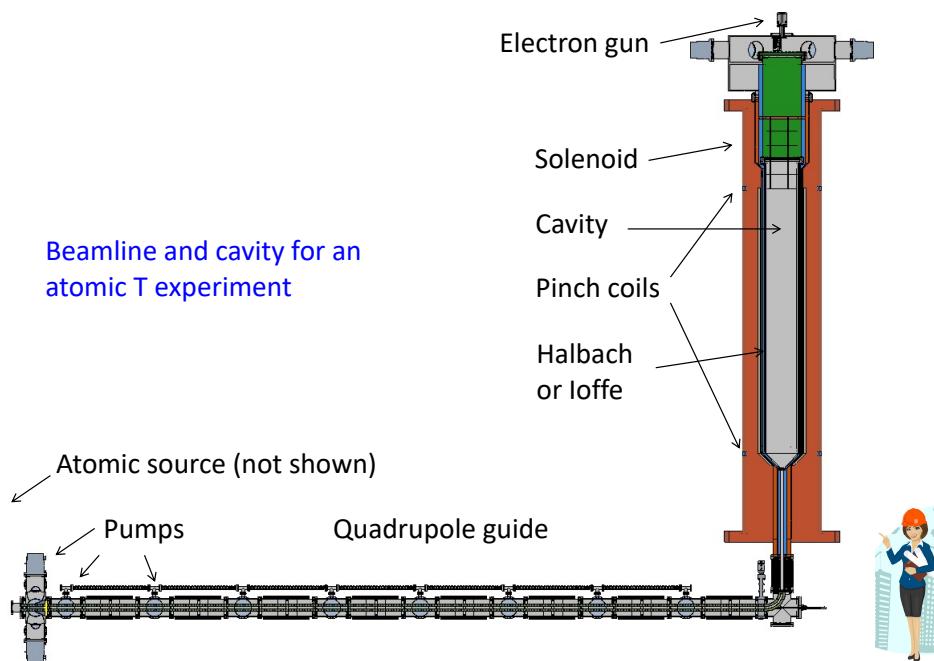
2083 A conceptual design for an antenna-based CRES experiment is shown in Figure 3.19.  
2084 A large solenoid magnet provides a uniform background magnetic field less than 0.1 T  
2085 in magnitude. Inside this region is the atom trapping magnet that generates a high  
2086 magnetic field at the walls, which decays exponentially towards the central region. Known  
2087 magnet designs that produce suitable atom trapping fields include Ioffe-Prichard traps,  
2088 which use conducting coils, as well as a Halbach array made from permanent magnets.  
2089 Either magnet choice produces a region of high magnetic fields, which excludes atoms  
2090 and allows for the placement of antennas inside the experiment.

2091 Inside this region an array of microstrip patch antennas is inserted to collect the  
2092 cyclotron radiation without providing a surface for atomic tritium recombination. Due  
2093 to the lower frequency of cyclotron radiation antennas of a larger size can be used,  
2094 which lowers the total number of antennas required to observe the experiment volume.  
2095 Because of this scaling, the lower frequency experiment uses a similar number of antennas  
2096 compared to a much smaller demonstrator experiment with a 1 T magnetic field.

2097 The atomic tritium beamline that supplies fresh tritium atoms to the experiment is  
2098 not shown in the figure. The general configuration would matches the one shown for the  
2099 pilot-scale cavity experiment (see Figure 3.20).

## 2100 Pilot-scale Cavity CRES Experiment Concept

2101 The pilot-scale cavity experiment includes both an atomic tritium system and cavity  
2102 CRES system. The atomic system consists of a thermal atom cracker located at the  
2103 start of an evaporatively cooled atomic beamline. The atomic tritium system provides a  
2104 supply of tritium atoms to the trap with temperatures on the order of a few mK. Atoms  
2105 at this temperature can be trapped magneto-gravitationally, which is the reason for the  
2106 vertical orientation of the cavity. At these low magnetic fields the trapping requirements  
2107 for electrons and atoms differ enough such that it is advantageous to decouple the the  
2108 trapping potentials to avoid radioactive heating of the tritium atoms from excess trapped  
2109 electrons. Electron trapping is provided by a set of magnetic pinch coils at the top and  
2110 bottom of the cavity and a multi-pole Ioffe or Halbach magnet serves to contain the  
2111 atoms.



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

2112 The cavity design for the pilot-scale experiment consists of a large cylindrical cavity  
2113 with a TE011 resonance of 325 MHz. Such a cavity is truly enormous, with a diameter  
2114 of approximately 1.2 m and a height of 11 m. When an electron is produced inside  
2115 the cavity with a cyclotron frequency that matches the TE011 resonant frequency it's  
2116 cyclotron orbit couples the electron to the TE011, which drives a resonance in the cavity.  
2117 These resonant fields can be read-out using an appropriate cavity coupling mechanism

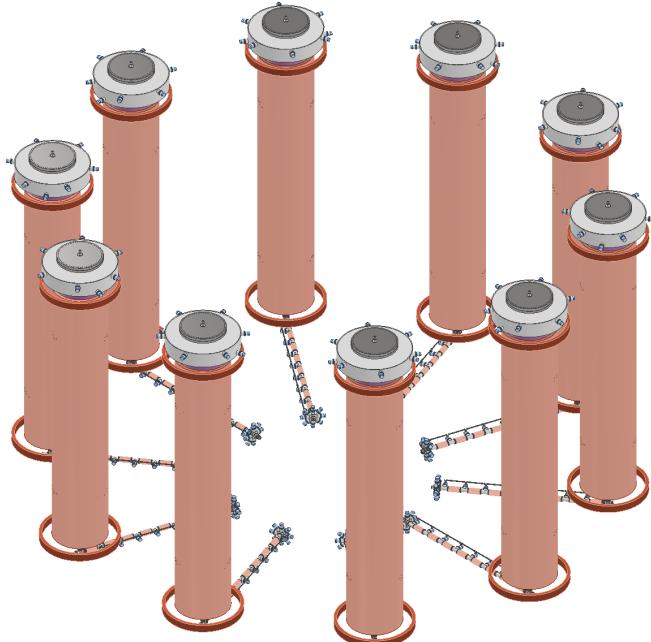
2118 located at the center of the cavity. For more information on the cavity approach to  
2119 CRES see Chapter 6.

2120 The bottom of the cavity has a cone termination to match the contour of the atom  
2121 trapping magnet. This shape still allows for TE011 resonances with high internal Qs,  
2122 which are required for good SNR in the cavity experiment. A small opening in the bottom  
2123 of the cone serves as an entry point for the tritium atoms. To allow for calibration of  
2124 the magnetic field inhomogeneities with an electron gun, the top of the cavity is left  
2125 nearly completely open. Normally, this would drastically lower the Q-factor of the TE011  
2126 mode, but a specially configured coaxial partition is inserted at the top. This termination  
2127 scheme is designed to act as a perfect short for the TE011 mode since the circular shape  
2128 of the partition matches the electric field boundary conditions for the TE011 mode.  
2129 Simulations with HFSS have confirmed that this design results in a high quality TE011  
2130 resonance despite the nearly completely open end.

## 2131 3.6 Phase IV

2132 The baseline CRES technology being pursued by the Project 8 collaboration are resonant  
2133 cavities, which, due to their geometric properties, simple CRES signal structure, and low  
2134 channel count, appear to be the better option for Phase IV. The current knowledge of the  
2135 antenna array CRES approach reveals no technical obstacles that would preclude it as a  
2136 baseline technology for Phase IV though it would most certainly be significantly more  
2137 expensive. Therefore, antenna arrays represent a fallback approach if resonant cavities  
2138 prove infeasible.

2139 The sensitivity of the pilot-scale atomic tritium experiment is estimated to be on  
2140 the order of 0.1 eV, which means that increasing the sensitivity to reach the Phase IV  
2141 goal will require an even larger experiment. Because of the direct coupling between the  
2142 RF characteristics of a cavity and its geometry, the baseline plan is to build multiple  
2143 copies of the pilot-scale experiment (see Figure 3.21) to obtain the required amount of  
2144 volume rather than increase the size of the cavity beyond the pilot-scale. The built-in  
2145 redundancy of this approach is useful in the sense that the experiment has no single  
2146 point of failure, additionally, building several copies of the a pilot-scale experiment will  
2147 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.

## **Chapter 4**

# **Signal Reconstruction Techniques for Antenna Array CRES and the FSCD**

### **4.1 Introduction**

The transition from a waveguide CRES experiment to an antenna array CRES experiment introduces new challenges related to data acquisition, signal detection, and signal reconstruction caused by the multi-channel nature of the data. The development of signal reconstruction algorithms is crucial to the design of antenna array based experiments like the FSCD, because these algorithms directly influence the detection efficiency and energy resolution of the CRES experiment. In this Chapter I summarize my contributions to the development and analysis of signal reconstruction and detection algorithms for the FSCD experiment.

In Section 4.2 I discuss the primary tool for this work, which is the Locust simulations package developed by the Project 8 experiment. Locust is used to simulate CRES events in the detector. Locust uses Kassiopeia to calculate particle trajectory solutions for electrons in the magnetic trap. The trajectories are then used to calculate the response of the antenna array to the cyclotron radiation produced by the electron, which results in signals that can be used to analyze the performance of different signal reconstruction algorithms. More recently, Project 8 has developed CREsana, which is a new simulations package that takes a more analytical approach to CRES signal simulations for antenna arrays. Although CREsana signals were not used for the signal reconstruction algorithm development detailed here, we introduce the software as it plays a role in the antenna array measurements presented in Section 5.5.

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

2174 only with distinguishing between data that contains a signal versus data that contains only  
2175 noise, whereas, with parameter estimation one is interested in extracting the kinematic  
2176 parameters of the electron encoded in the cyclotron radiation signal shape. Due to  
2177 the low signal power of electrons near the spectrum endpoint in the FSCD experiment,  
2178 signal detection is a non-trivial problem. This is magnified by the need to maximize the  
2179 detection efficiency of the experiment in order to achieve the neutrino mass sensitivity  
2180 goals. My contributions to signal reconstruction analysis for the FSCD are focused on  
2181 this signal detection component of reconstruction.

2182 After the discussion of various signal detection approaches, in Section 4.4 I present a  
2183 more detailed analysis of the detection performance of three algorithms, which could be  
2184 used to signal detection in the FSCD. This section was originally prepared for publication  
2185 in JINST as a separate paper. The algorithms include a digital beamforming algorithm,  
2186 a matched filter algorithm, and a neural network algorithm, which I analyze in terms of  
2187 classification accuracy and estimated computational cost.

## 2188 **4.2 FSCD Simulations**

2189 Antenna array CRES and the FSCD requires a combination of different capabilities  
2190 not often found in a single simulation tool. First of all, accurate calculations of the  
2191 magneto-static fields produced by current-carrying coils are required in order to accurately  
2192 model the magnetic trap and background magnets. The resulting magnetic fields must  
2193 then be used to calculate the exact relativistic trajectory of electrons, which is required  
2194 in order to calculate the electro-magnetic (EM) fields produced by the acceleration of  
2195 the electron. Finally, the simulation has to model the interaction of the antenna and  
2196 RF receiver chain with these EM-fields in order to produce the simulated voltage signals  
2197 produced by the antenna array during the CRES event. At the time when Project 8 was  
2198 developing this simulation capability, no single available simulation tool was known to  
2199 adequately perform this suite of calculations, which prompted the development of custom  
2200 simulation framework to simulate the FSCD. This simulation framework includes custom  
2201 simulation tools developed by Project 8 as well as other open-source and proprietary  
2202 software developed by third-parties.

2203 **4.2.1 Kassiopeia**

2204 Kassiopeia<sup>1</sup> is a particle tracking and static EM-field solver developed by the KATRIN  
2205 collaboration for simulations of their spectrometer based on magnetic adiabatic collimation  
2206 with an electrostatic filter [60]. Due to the measurement technique employed by the  
2207 KATRIN collaboration, Kassiopeia is not designed to solve for the EM-fields produced by  
2208 electrons in magnetic fields. However, it does provide efficient solvers for static electric  
2209 and magnetic fields and charged particle trajectory solvers. Because of this, Project 8  
2210 has incorporated parts of Kassiopeia into its own simulation framework.

2211 **Magnetostatic Field Solutions**

2212 The solutions to the electric and magnetic fields generated by a static configuration of  
2213 charges and currents is given by Maxwell's equations in the limit where the time-dependent  
2214 terms go to zero. In their static form Maxwell's equations [49] are

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

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

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

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

2215 where we can see that the electric and magnetic fields are now completely decoupled  
2216 from each other. The solution for the magnetic field in this boundary value problem is  
2217 given by the Biot-Savart law

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

2218 which Kassiopeia uses a variety of numeric integration techniques to solve for a user  
2219 defined current distribution.

2220 **Kassiopeia Simulation of the FSCD Magnetic Trap**

2221 The trap developed for the FSCD experiment utilizes six current carrying coils, which  
2222 surround a cylindrical tritium containment vessel (see Figure 4.1). Some critical aspects  
2223 of the trap design include the total trapping volume, the maximum trap depth, the

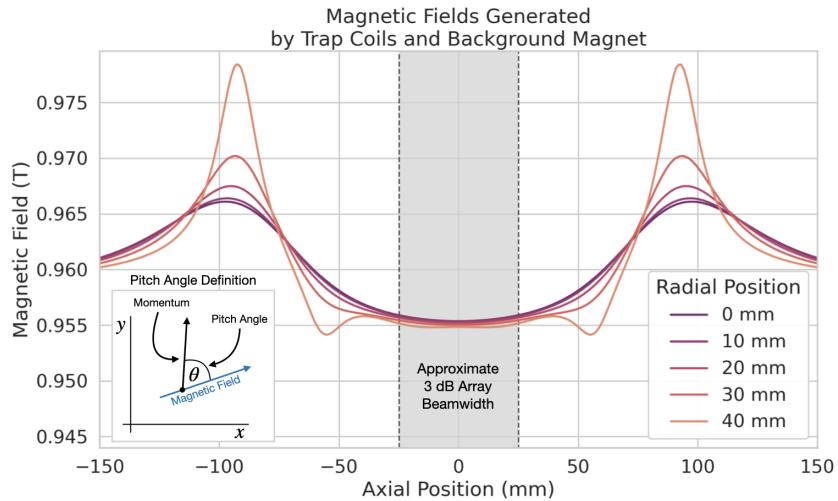
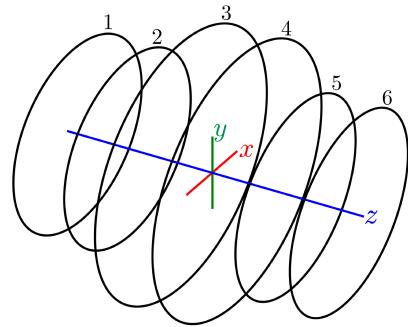
---

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

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 (A×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 Kassiopéia. 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  $84^\circ$ . The trap depth factors into 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 then, in principle, more decay events can be observed. However, the signals from electrons with small pitch angles are typically significantly harder to detect than larger pitch angles when using an antenna array, 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 any 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 it's cyclotron frequency. When an electron is trapped, it oscillates back and forth along the trap z-axis (see Figure 4.1) unless it is produced with a pitch angle of exactly 90° [61]. 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

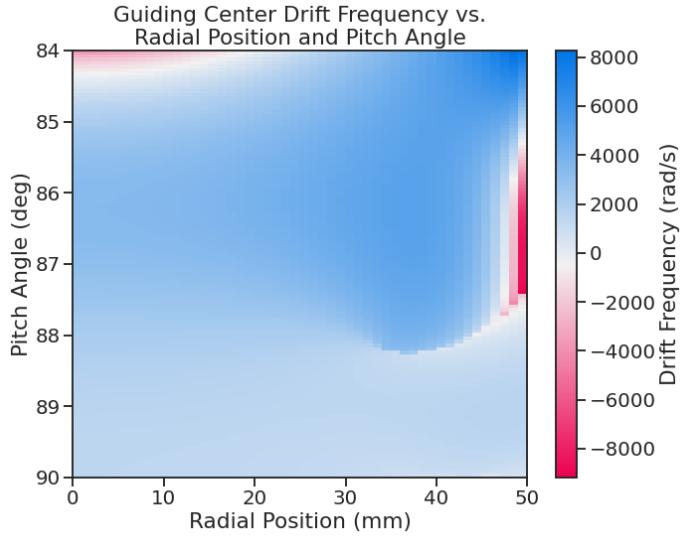
The magnetic fields solved by direct integration of the electron's current density can be used by Kassiopeia to solve for the trajectory of electrons based on user specified initial conditions. Various distributions are available within Kassiopeia that can be sampled in order to replicate realistic event statistics, including uniform, Gaussian, and Lorentzian 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 more common to parameterize the electron's trajectory in terms of it's initial position, the kinetic energy, the pitch angle, and the initial direction of the component of the electron's momentum perpendicular to the magnetic field. This parameterization is completely equivalent to specify each component of the electrons initial 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  $\gamma$  is the relativistic Lorentz term. To account for kinetic energy losses from radiation Kassiopeia includes an additional term in the momentum differential equation, which calculates the change in the electron's momentum induced by synchrotron radiation. 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 be applied, then Kassiopeia can make use of a simpler set of equations that can be more readily solved numerically.



**Figure 4.2.** A map of the average  $\nabla B$ -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.

Even though Kassiopeia is not directly capable of simulating the cyclotron radiation, it is still an invaluable CRES simulation tool, due to the accurate trajectory solutions for electrons in magnetic traps. With Kassiopeia it is possible to test the efficiency of a particular trap design and analyze features of the electron trajectories that are important to the position, track, and event reconstruction algorithms (see Section 4.3). One example of this for the FSCD is the analysis of the average  $\nabla B$ -drift frequency as a function of the electrons radial position and pitch angle in the magnetic trap (see Figure 4.2). Radial gradients in the trap cause the guiding center of the electron to drift around the center of

the magnetic trap with an average frequency on the order of  $10^3$  rad/s. This frequency, while slow compared to the length of a typical CRES time-slice, is large enough to cause a significant loss in efficiency of certain signal reconstruction algorithms. Therefore, it is important to model the drift of the electron in the reconstruction algorithm in order to mitigate the effects of this motion on the reconstruction.

### 4.2.2 Locust

The Locust<sup>2</sup> software package [62] 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, we 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 that we calculate the electric fields produced by the acceleration of the electron. In the general case, this can be a complicated question to answer, due to back-reaction forces on the electron from its own electric fields that occur when the electron is surrounded by conductive material such as a waveguide or cavity. 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 electromagnetic fields from a relativistic moving point particle are the Liénard-Wiechert field equations [63, 64], which are obtained by differentiating the Liénard-Wiechert potentials. In their full form the Liénard-Wiechert field equations are

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

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

---

<sup>2</sup>[https://github.com/project8/locust\\_mc/tree/master](https://github.com/project8/locust_mc/tree/master)

2311 where  $e$  is the charge of the particle,  $\hat{n}$  is the unit vector pointing from the particle to the  
 2312 position where the fields are calculated,  $\beta$  and  $\dot{\beta}$  are the velocity and acceleration of the  
 2313 particle divided by the speed of light ( $c$ ),  $\mathbf{R}$  is the distance from the particle to the field  
 2314 calculation position, and  $\gamma$  is the relativistic Lorentz term. The subscript  $t_r$  indicates  
 2315 that the equations must be evaluated at the retarded time so that the time-delay from  
 2316 the travel time of the electromagnetic radiation is correctly accounted for.

2317 The only required input to calculate the electric field at the position of an FSCD  
 2318 antenna is the velocity and acceleration of the electron, which can be obtained from  
 2319 Kassiopeia simulations. Therefore, when simulating a CRES event Locust first runs a  
 2320 Kassiopeia simulation of the electron and calculates the electric field incident on the  
 2321 antenna. The only difficulty with this approach is the determination of the retarded time.  
 2322 The retarded time corresponds to the time that a photon, which has just arrived at an  
 2323 antenna at the space-time position  $(t, \mathbf{r})$ , was actually emitted by the electron at the  
 2324 space-time position of  $(t_r, \mathbf{r}_e(t_r))$ . Defined in this way, finding the retarded time requires  
 2325 solving

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

2326 where the distance traveled by the photon between the measurement and retarded times  
 2327 is equal to the distance between the antenna and the electron at the retarded time.  
 2328 Locust solves Equation 4.10 using a built-in root finding algorithm to find the retarded  
 2329 time, and thus the electric field produced by the electron at the position of each antenna  
 2330 in the FSCD array.

### 2331 Antenna Response Modeling

2332 With the electric field it is possible, in principle, to calculate the resulting voltages pro-  
 2333 duced in the antenna. However, direct simulation of the antenna itself is computationally  
 2334 expensive since it would require the modeling of complex interactions of the electron's  
 2335 electric fields with charge carriers in the conductive elements of the antenna. Direct  
 2336 simulation of the antenna in Locust can be avoided by modeling the antenna response  
 2337 using the antenna factor, or antenna transfer function, approach. The antenna factor  
 2338 defines the voltage produced in the antenna terminal for an incident electric field [65],

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

2339 where  $V$  is the voltage and  $|\mathbf{E}|$  is the magnitude of the incident electric field. To obtain the  
 2340 antenna factor for the antennas developed for the FSCD Project 8 employs Ansys HFSS.  
 2341 HFSS is a commercially available finite element method electromagnetic solver widely  
 2342 used throughout the antenna engineering industry [66]. HFSS is capable of calculating  
 2343 the antenna factor and gain patterns for complex antenna designs and outputting the  
 2344 resulting quantities in the form of a text file that can be used as an input to the Locust  
 2345 simulation.

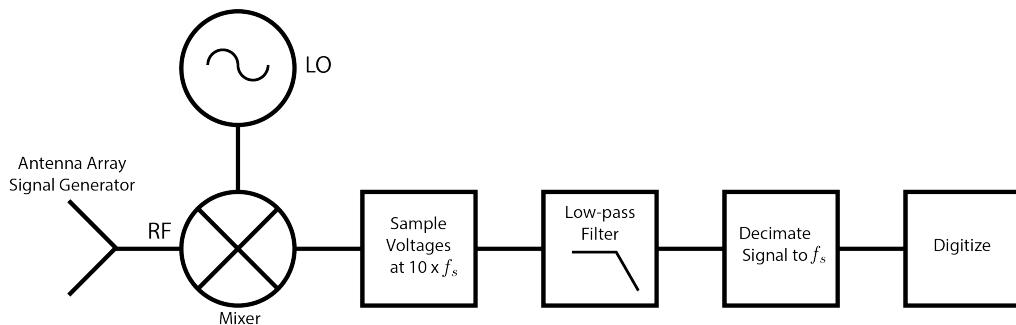
2346 The antenna factor defines the steady-state response of the antenna to electromagnetic  
 2347 plane waves and is a function of the frequency of the radiation. Therefore, in order to  
 2348 apply the transfer function for the calculation of the antenna voltage response in the  
 2349 time domain, Locust models the antenna as a linear time-invariant system [67]. In this  
 2350 formalism the response of the system to the driving force is given by

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

2351 where  $y[n]$  is the discretely sampled response,  $x$  is the driving force stimulus, and  $h$  is  
 2352 the finite impulse response (FIR) filter. When applied to the FSCD array, this formalism  
 2353 calculates the voltage time-series produced in each antenna by convolving the electric  
 2354 field time-series with the antenna FIR filter, which is obtained by performing a inverse  
 2355 Fourier transform on the transfer function from HFSS.

## 2356 Radio-frequency Receiver and Signal Processing

2357 After obtaining the voltage time-series by computing the electron trajectory and antenna  
 2358 response, Locust simulates the signal processing associated with the radio-frequency  
 2359 receiver chain. The standard receiver chain used in Locust simulations of the FSCD  
 2360 attempts to mimic the operations that would actually occur in hardware (see Figure 4.3).



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

2361 Frequency down-conversion is used in the FSCD to reduce the digitization bandwidth  
 2362 required to read-out CRES data. According to the Nyquist sampling theorem [68], the  
 2363 minimal sampling rate that guarantees no information loss for a signal with a bandwidth  
 2364  $\Delta f$  is given by

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

2365 The total bandwidth of CRES signal frequencies from tritium beta-decay ranges from 0  
 2366 to 26 GHz in a 0.95 T magnetic field, therefore, direct digitization of CRES signals from  
 2367 the FSCD would require sampling frequencies greater than 50 GHz, which is infeasible for  
 2368 a real experiment. However, for the purposes of neutrino mass measurement we are only  
 2369 interested in measuring the shape of the spectrum in the last 100 eV, which corresponds  
 2370 to a frequency bandwidth of 5 MHz. Down-conversion is a technique for reducing the  
 2371 base frequencies of signals in a bandwidth given by  $[f_{\text{LO}}, f_{\text{LO}} + \Delta f]$  to the bandwidth  
 2372  $[0, \Delta f]$ , by performing the following multiplication

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

2373 In down-conversion the signal ( $x(t)$ ) is multiplied by a sinusoidal signal with frequency  
 2374  $f_{\text{LO}}$  to reduce the absolute frequencies of the signals in the bandwidth. In the FSCD this  
 2375 allows us to detect events in the last 100 eV of the tritium spectrum while sampling the  
 2376 data far below 50 GHz. The standard bandwidth used in the FSCD is 200 MHz, which  
 2377 allows for higher frequency resolution than the minimum sampling frequency for 100 eV  
 2378 of energy bandwidth.

2379 Trying to directly simulate down-conversion with a frequency multiplication in Locust  
 2380 would require the sampling of the electric fields at each antenna in the FSCD array with  
 2381 a period of  $\approx 20$  ps, which is extremely slow computationally. To avoid this Locust  
 2382 performs the down-conversion by intentionally under-sampling the electric fields with  
 2383 a frequency of 2 GHz. Sampling below the Nyquist limit causes the higher frequency  
 2384 components of the CRES signal to alias, however, Locust can remove these aliased  
 2385 frequency peaks using a combination of low-pass filtering and decimation to recreate  
 2386 frequency down-conversion. After filtering and decimation, Locust simulates digitization  
 2387 by an 8-bit digitizer at a sampling frequency of 200 MHz to recreate the conditions of  
 2388 the FSCD. The voltage offset and the digitizer range must be configured by the user  
 2389 based on the characteristics of the simulation.

2390 **Data**

2391 The output of Locust simulations for the FSCD primarily consists of two data files. The  
2392 first is the electron trajectory information calculated by Kassiopiea, which is output in  
2393 the form of a `.root` file [69]. This file contains important kinematic information about  
2394 the electron such as it's position and pitch angle as a function of time. The other file is  
2395 produced by Locust and it contains the digitized signals acquired from each antenna in  
2396 the FSCD array. The Locust output files conform to the Monarch specification developed  
2397 by Project 8, which is based on the commonly used HDF5 file format, and matches the  
2398 format of the files produced by the Project 8 data acquisition software. This makes it  
2399 possible to use the same data analysis code to analyze both simulated and real data.

2400 **4.2.3 CRESana**

2401 Locust is the primary simulation tool used by Project 8 in the development and simulation  
2402 of the FSCD. However, simulations of CRES events in larger antenna arrays ( $\geq 100$   
2403 antennas) using Locust can take several hours to complete, which is prohibitively long  
2404 when one is performing a sensitivity analysis for a large scale antenna experiment. One  
2405 of the reasons for Locust's slow operation is that the electric fields from the electron  
2406 must be solved numerically for each time-step for each of the antennas in the array.  
2407 These numerical solutions allow Locust to accurately simulate the electric fields from  
2408 arbitrarily complicated electron trajectories at the cost of more computations and slower  
2409 simulations. Therefore, an additional simulation tool that sacrifices some accuracy for  
2410 computational efficiency would be extremely useful simulations and sensitivity analyses  
2411 of larger antenna array experiments.

2412 To fill this need, Project has developed a new simulations package called CRESana<sup>3</sup>,  
2413 specifically designed to perform analytical simulations of antenna array based CRES  
2414 experiments. CRESana is not as flexible as Locust, but it provides a significant increase  
2415 in simulation speed. It does this by using well-justified analytical approximations of the  
2416 electrons motion in the magnetic field and the resulting electric fields from the electron's  
2417 acceleration. The electric fields and signals generated by CRESana are consistent with  
2418 theoretical calculations of the electron's radiation, and are test for accuracy using  
2419 well-known test-case simulations and consistency checks.

---

<sup>3</sup><https://github.com/MCFlowMace/CRESana>

## **2420 4.3 Signal Detection and Reconstruction Techniques for 2421 Antenna Array CRES**

### **2422 Antenna Array CRES Signal Reconstruction**

**2423** A robust set of FSCD simulation tools are vital to the development of the analysis  
**2424** algorithms necessary for antenna array CRES to succeed. In order to perform CRES  
**2425** measurements using an antenna array, one must develop an algorithm that uses the  
**2426** multi-channel time-series obtained by digitizing the array to estimate the starting kinetic  
**2427** energies of electrons produced in the magnetic trap. This procedure consists of a multi-  
**2428** stage process of detecting a CRES signal then estimating the parameters of the electron  
**2429** that produced and is often referred to as simply CRES signal reconstruction.

**2430** Compared with the signal reconstruction approaches of the Phase I and II CRES  
**2431** experiments, antenna array CRES requires a significantly different approach to signal  
**2432** reconstruction. In Phase I and II, CRES was performed using a waveguide gas cell that  
**2433** could be directly connected to a waveguide transmission line. The transmission line  
**2434** efficiently transmits the cyclotron radiation along its length to an antenna at either end  
**2435** of the waveguide. However, with an antenna array the electron is essentially radiating  
**2436** into free-space, therefore, the cyclotron radiation power collected by the array is directly  
**2437** proportional to the solid angle surrounding the electron that is covered with antennas.  
**2438** Because it is not practical to fully surround the magnetic trap with antennas, some of the  
**2439** cyclotron radiation power that would have been collected by the waveguide escapes into  
**2440** free-space. Furthermore, the power that is collected by the antenna array is split between  
**2441** every channel in the antenna array, which significantly lowers the signal-to-noise ratio  
**2442** (SNR) of CRES signals in a single antenna channel compared to a waveguide apparatus.  
**2443** Therefore, a suite of completely new signal reconstruction techniques are needed in order  
**2444** to perform CRES in the FSCD.

**2445** Changes to the approach to CRES signal reconstruction are also motivated by the  
**2446** more ambitious scientific goals of the FSCD experiment. A measurement of the tritium  
**2447** beta-decay spectrum that is sensitive to neutrino masses as small as 40 meV requires that  
**2448** we measure the kinetic energies of individual electrons with a total energy broadening  
**2449** of 115 meV [70]. This resolution includes all sources of uncertainty in the electron's  
**2450** kinetic energy such as magnetic field inhomogeneities. This level of energy resolution is  
**2451** compatible only with an event-by-event signal reconstruction approach where the kinetic  
**2452** energies, pitch angles, and other parameters of the CRES events are estimated before

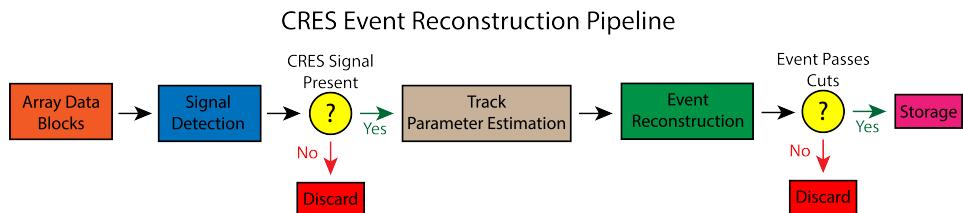
2453 constructing the beta-decay spectrum.

2454 The event-by-event approach is distinct from the analysis done for the Phase I and  
2455 Phase II experiments where only the starting cyclotron frequency of the event was  
2456 estimated by analyzing the tracks formed by the carrier frequency in the time-frequency  
2457 spectrogram. These frequencies were then combined into a frequency spectrogram, which  
2458 was converted to the beta-decay energy spectrum using an ensemble approach that  
2459 averaged over all other event parameters. The ensemble approach to signal reconstruction  
2460 results in poor energy resolution because other kinematic parameters such as pitch angle  
2461 change the cyclotron carrier frequency due to changes in the average magnetic field  
2462 experience by the electron, and it is therefore incompatible with the future goals of the  
2463 Project 8 collaboration.

#### 2464 Components of Reconstruction: Signal Detection and Parameter Estimation

2465 CRES signal reconstruction can be viewed as a two-step procedure consisting of signal  
2466 detection followed by parameter estimation. In the former, one is concerned with  
2467 identifying CRES signals in the data regardless of the signal parameters, whereas, in the  
2468 latter one operates under the assumption that a signal is present and then estimates its  
2469 parameters.

2470 More formally, signal detection is essentially a binary hypothesis test between the  
2471 signal and noise data classes and parameter estimation describes a procedure of fitting a  
2472 model to the observed data. While both of these processes are required for a complete  
2473 reconstruction (see Figure 4.4), the focus of my work and this chapter is on the signal  
2474 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.

2475 **Detection Theory**

2476 The problem of signal detection can be posed as a statistical hypothesis test [71]. For  
2477 CRES signals, which are essentially vectors with added white Gaussian noise (WGN),  
2478 one needs to choose between two hypotheses

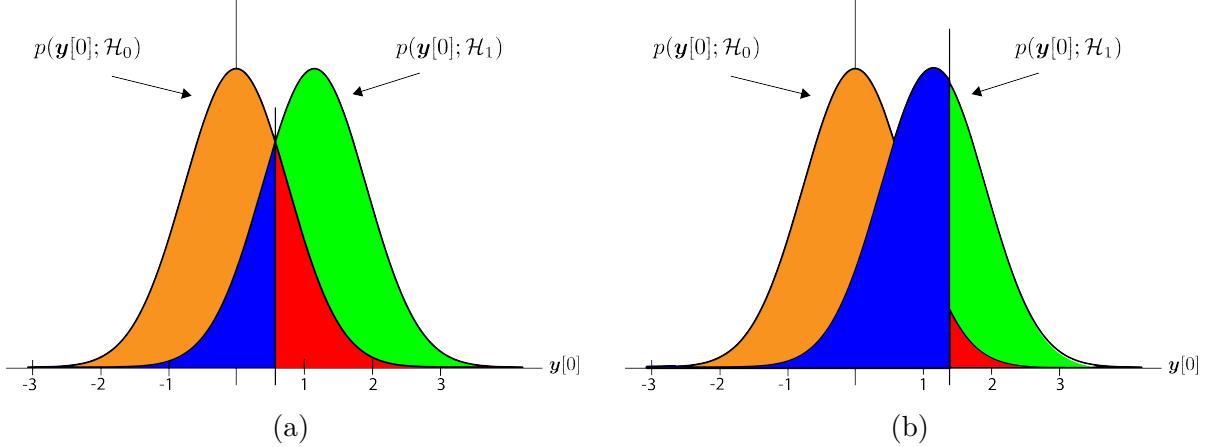
$$\mathcal{H}_0 : \mathbf{y} = \boldsymbol{\nu} \quad (4.15)$$

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

2479 where  $\mathbf{y}$  is the CRES data vector,  $\boldsymbol{\nu}$  is a sample of WGN, and  $\mathbf{x}$  represents the CRES  
2480 signal. The hypothesis that the data contains only noise is labeled  $\mathcal{H}_0$  and the hypothesis  
2481 that the data contains a signal is labeled  $\mathcal{H}_1$ .

2482 For illustrative purposes one can examine the case where one the first sample of  
2483 data is used to distinguish between  $\mathcal{H}_0$  and  $\mathcal{H}_1$ . The value of the first data sample is  
2484 distributed according to two gaussian distributions corresponding to  $\mathcal{H}_0$  and  $\mathcal{H}_1$  (see  
2485 Figure 4.5). By setting a decision threshold on the value of this sample, one can choose  
2486 the correct hypothesis with a probability given by the areas underneath the probability  
2487 distribution curves. A true positive corresponds to correctly identifying that the data  
2488 contains signal, whereas, a true negative means that one has correctly identified the data  
2489 as noise. The rate at which the detector performs a true positive classification is given  
2490 by the green region underneath  $p(\mathbf{y}[0]; \mathcal{H}_0)$ , and the rate at which the detector performs  
2491 a true negative classification is given by the orange region underneath  $p(\mathbf{y}[0]; \mathcal{H}_1)$ . Two  
2492 types of misclassifications are possible. Either we declare noise data as signal, which is  
2493 call a false positive, or we declare signal data as noise, which is a false negative. Note  
2494 that it is only possible to trade off these two types of errors by tuning the detection  
2495 threshold. One cannot simultaneously reduce the rate of false positives without also  
2496 increasing the rate of false negatives.

2497 The approach taken with CRES signals is to fix the rate of false positives by setting  
2498 a minimum value for a detection threshold. The rate of false positives that is acceptable  
2499 at the detection stage depends upon the rate of background events compatible with the  
2500 sensitivity goals of the experiment. The ultimate goal of a neutrino mass measurement  
2501 with 40 meV sensitivity in general has strict requirements on the number of background  
2502 events, which requires a relatively high detection threshold to achieve. Consequently,  
2503 the ideal signal detection algorithm is the one that achieves the maximum rate of true  
2504 positives for a fixed rate of false positives, so that the detection efficiency of the experiment  
2505 is maximized and potential sources of background are kept to a minimum.



**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.

According to the Neyman-Pearson theorem [72], the statistical hypothesis test that maximizes the probability of detection for a fixed rate of false positives is the likelihood ratio test, which is formed by computing the ratio of the signal likelihood to the noise likelihood,

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

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

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

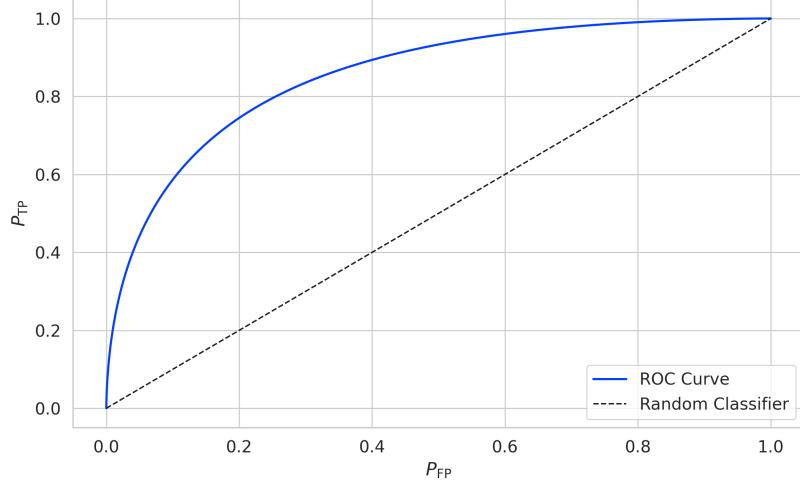
where  $\alpha$  is the desired false positive detection rate given by the red colored areas shown in Figure 4.5. The true positive detection rate is given by the similar integral

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

which corresponds to the green areas in Figure 4.5.

Changing the decision threshold allows one to trade-off between  $P_{\text{TP}}$  and  $P_{\text{FP}}$  as

appropriate for the given situation. It is common to summarize the relationship between  $P_{\text{TP}}$  and  $P_{\text{FP}}$  using the receiver operating characteristic (ROC) curve, which is obtained 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  $P_{\text{FP}}$  and the  $P_{\text{TP}}$  for a given likelihood ratio test. As the decision threshold is increased  $P_{\text{FP}}$  decreases at the expense of a lower  $P_{\text{TP}}$ . The black dashed line indicates the lower bound ROC curve obtained by randomly deciding between  $\mathcal{H}_0$  and  $\mathcal{H}_1$ .

the performance of different signal detection algorithms. In general, a classifier with a higher the  $P_{\text{TP}}$  as a function of  $P_{\text{FP}}$  is desirable, which corresponds to a larger area underneath the respective ROC curve. A perfect classifier has an area underneath the curve of 1.0, however, such a classifier is almost never achievable in practice.

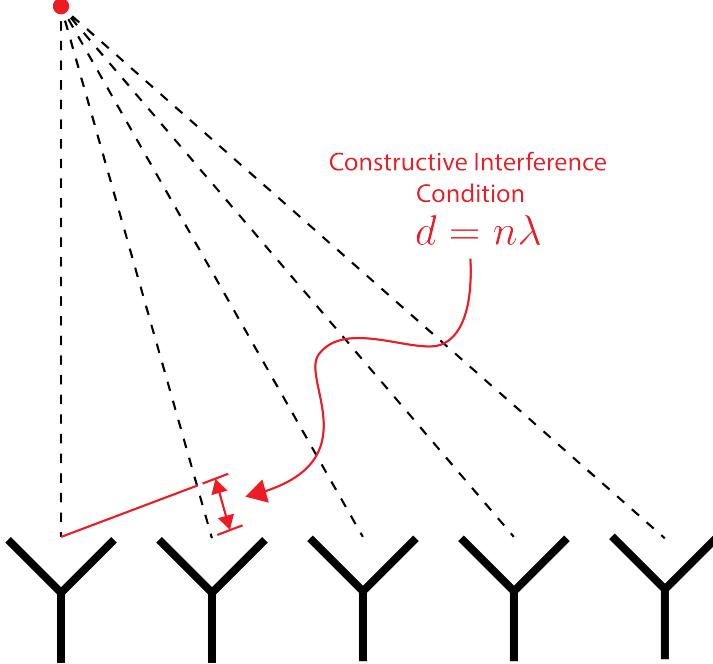
### 4.3.1 Digital Beamforming

#### Introduction to Beamforming

Beamforming refers to a suite of antenna array signal processing techniques that are designed to enhance the radiation or gain of the array in certain directions and suppress it in other direction [65]. Beamforming is of interest to Project 8 as a first level of signal reconstruction for the FSCD and other antenna array CRES experiments, which operates at the signal detection stage of reconstruction.

Beamforming is accomplished by performing a phased summation of the signals received by the antenna array. The beamforming phases are chosen such that the signals

2534 emitted by the array will constructively interfere at the point of interest (see Figure  
 2535 4.7). As a consequence of the principle of reciprocity [73], when the array is operating in  
 2536 receive 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-



**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.

2537  
 2538 length difference to the beamforming point between different antennas in the array. The  
 2539 relationship between the phase delay and the path-length difference is given by the  
 2540 familiar equation

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

2541 where  $\phi$  is the phase delay,  $d$  is the path-length difference, and  $\lambda$  is the wavelength of  
 2542 the radiation. In practice, one chooses the values of  $d$  by specifying the beamforming  
 2543 positions of interest and then calculates the beamforming phases using Equation 4.20,  
 2544 which is guaranteed to follow the constructive interference condition shown in Figure 4.7.

2545 Beamforming can be neatly expressed mathematically using the vector equation

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

2546 where  $\mathbf{x}[n]$  is the array snapshot vector,  $\Phi[n]$  is a vector of beamforming shifts, and  
 2547  $y[n]$  is the resulting summed signal. The beamforming shifts consist of a set of complex

2548 numbers that contain the beamforming phase shift and an amplitude weighting factor,

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

2549 where the set of magnitudes  $A_i[n]$  are amplitude weighting factors and  $\phi_i[n]$  are the phase  
2550 shifts from the path-length differences. The index  $i$  is used to denote the antenna channel  
2551 number. The amplitude weighting factor is the relative magnitude of the signal received  
2552 by a particular antenna to the other antennas in the array, such that the antennas that  
2553 receive signals with higher amplitude, due to being closer to the source, have more  
2554 weight in the beamforming summation. The input and outputs signals beamforming  
2555 are naturally expected to be functions of time as indicated by the index  $[n]$ , however, it  
2556 is also possible to use time dependent beamforming phases that shift the beamforming  
2557 position of the array over time.

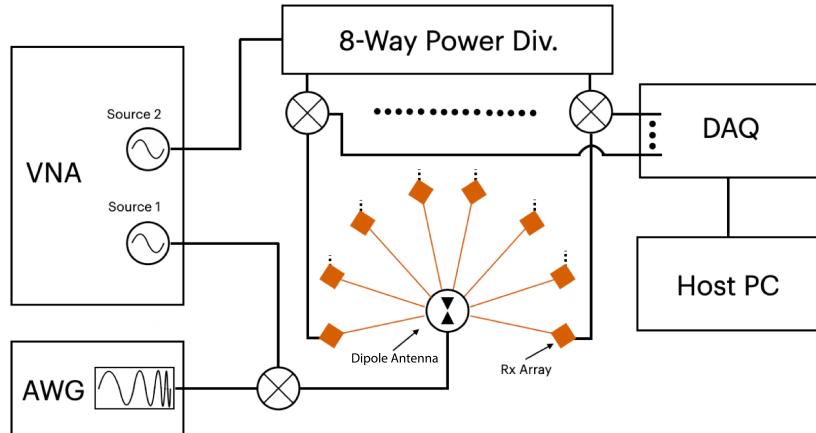
2558 Digital beamforming is the type of beamforming algorithm of interest to Project 8 for  
2559 CRES. Specifically, digital beamforming means that the beamforming phases are applied  
2560 to the array signals in software rather than employing fixed beamforming phase shifts in  
2561 the receiver chain hardware. The advantage of digital beamforming is that for a given  
2562 series of array snapshots one can specify a large number of beamforming positions and  
2563 effectively search for electrons by performing the beamforming summation associated  
2564 with each point and applying a signal detection algorithm to identify the presence of a  
2565 CRES signal.

2566 One of the most attractive features of digital beamforming is the spatial filtering  
2567 effect, which is a direct consequence of the constructive interference condition used to  
2568 define the beamforming phases. Spatial filtering allows for signals from multiple electrons  
2569 at different positions in the trap to be effectively separated, because the constructive  
2570 interference condition will force the signals from electrons at positions different from the  
2571 beamforming position to cancel. This helps to reduce signal pile-up that could become  
2572 an issue for large scale CRES experiments using a dense tritium source.

2573 The digital beamforming positions can be specified with arbitrary densities limited  
2574 only by the available computational resources. This provides a very straight-forward way  
2575 to estimate the position of the electron in the trap by using a dense grid of beamforming  
2576 positions and maximizing the output power of the beamforming summation over this  
2577 grid. This natural approach to position reconstruction is attractive due the requirements  
2578 of an event-by-event signal reconstruction, which needs an accurate estimation of the  
2579 exact magnetic field experienced by the electron in order to correctly estimate it's kinetic

2580 energy. Combined with an accurate map of the magnetic field inhomogeneities of the  
 2581 trap obtained from calibrations, beamforming allows one to apply this magnetic field  
 2582 correction with a spatial resolution that is a fraction of the cyclotron wavelength.

2583 **Laboratory Beamforming Demonstrations**

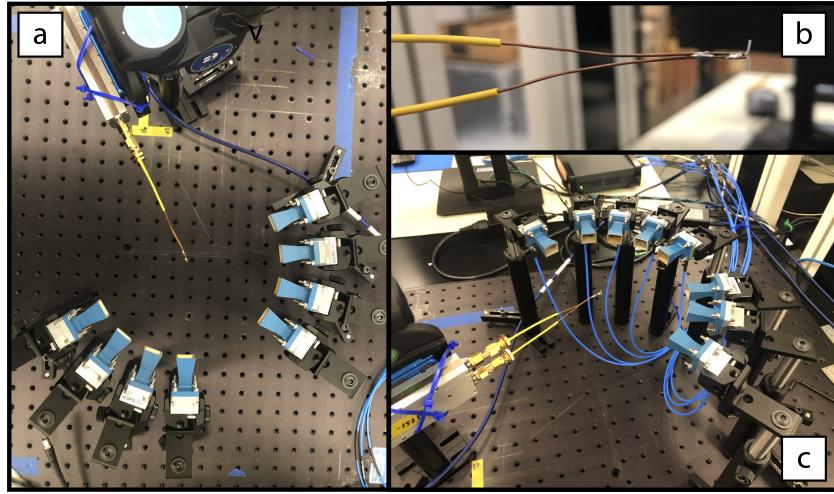


**Figure 4.8.** System level diagram of the laboratory setup used for beamforming demonstrations at Penn State. For more information on this system see Chapter 5. Signals near 26 GHz are fed to a dipole antenna using an arbitrary waveform generator (AWG) and vector network analyzer (VNA), which drive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

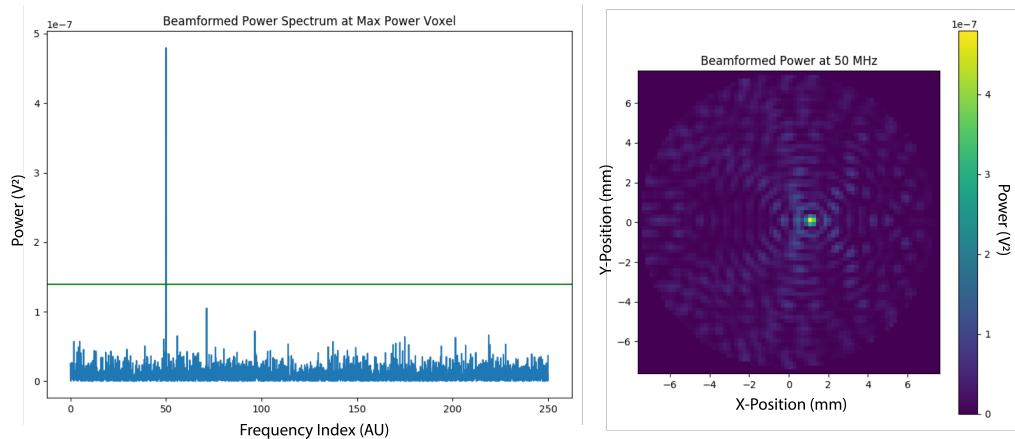
2584 As part of the development of antenna array CRES for the FSCD, an antenna  
 2585 measurement setup was constructed at Penn State to serve as a testbed for antenna  
 2586 prototypes and to perform laboratory validations of array simulations. This system  
 2587 is discussed in more detail in Chapter 5. Early versions of the antenna measurement  
 2588 system (see Figure 4.8 and Figure 4.9) were used to perform beamforming reconstruction  
 2589 studies of a simple probe antenna to better understand the principles of beamforming  
 2590 and confirm the estimated beamforming performance of Locust.

2591 Signals from an arbitrary waveform generator were up-converted to 26 GHz using a  
 2592 mixer and a high-frequency source from a vector network analyzer and fed to the dipole  
 2593 antenna through a balun. The radiation from the dipole antenna was received by an  
 2594 array of horn antennas. The signals from the horn antennas were then down-converted  
 2595 to baseband using a collection of mixers and an 8-way power divider. The signals were  
 2596 then digitized and saved to a host computer for analysis.

2597 The data collected using the dipole and horn antenna array is reconstructed using the  
 2598 beamforming reconstruction approach specified in Section 4.3.1. A two-dimensional grid



**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.



**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.

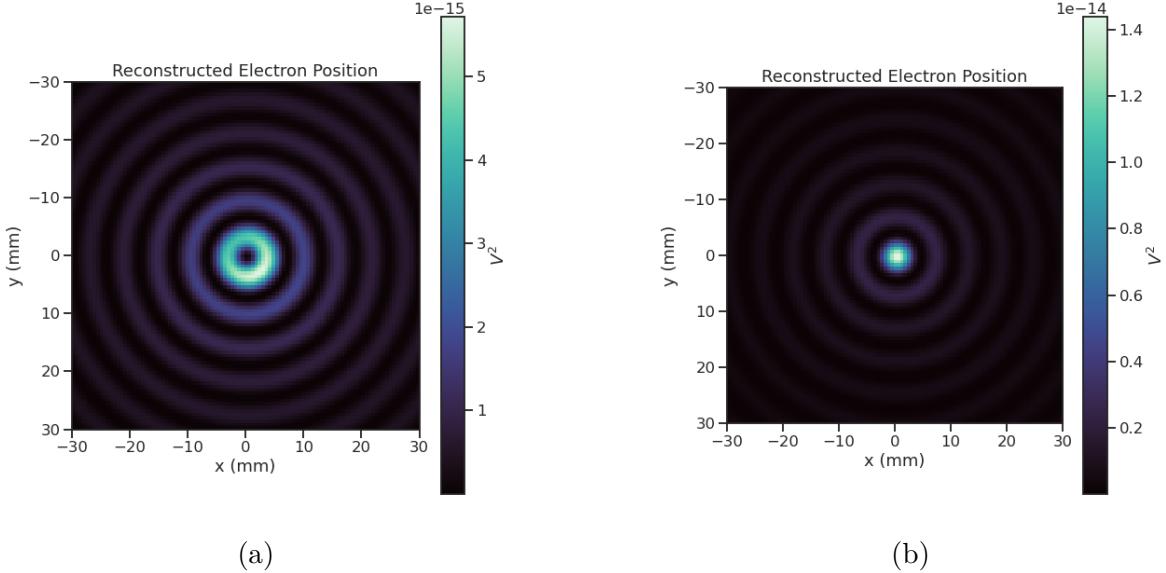
of xy-positions is defined and the beamforming phase shifts for each of these positions is calculated. The phased summation can be visualized by plotting the time-averaged power for each of the summations as a pixel in the resulting beamforming image (see Figure 4.10). White Gaussian noise (WGN) can be added to the data at this stage to simulate more realistic signal-to-noise ratios (SNR) if desired. The beamforming peak maxima is expected to have a Bessel function shape due to the circular symmetry of the array, and by analyzing the size of the beamforming maxima one can confirm that the beamforming reconstruction measurement has similar position resolution as expected from Locust simulations. Additionally, signal detection rates can be estimated from the data by comparing the magnitude of the beamforming signal peak in the frequency spectra to simulation.

## FSCD Beamforming Simulations

Using Locust simulations of the FSCD one can perform beamforming reconstruction studies using the simulated CRES signal data. As we mentioned in the previous section, the beamforming procedure begins by specifying a set of beamforming positions and corresponding beamforming shifts. The beamforming positions form a grid that covers the region of interest in the field of view of the antenna array. There are effectively an infinite number of ways to specify the grid positions, however, uniform square grids are the most commonly used due to their simplicity. In the FSCD experiment the number and pattern of the grid positions would be optimized to cover the most important regions of the trap volume to maximize detection efficiency while minimizing superfluous calculations.

The beamforming grids used for signal reconstruction with the FSCD consist of a set of points that cover a region of the two-dimensional plane formed by the perimeter of the antenna array. The axial dimension is left out of the beamforming grid because the electrons are assumed to occupy only an average axial position, which corresponds to the center of the magnetic trap. This is because it is impossible to resolve the axial position of the electron as a function of time due to the rapid axial oscillation frequencies of trapped electrons relative to the FSCD time-slice duration.

After beamforming, a summed time-series is obtained for each beamforming position that can be evaluated for the presence of a signal using a detection algorithm. A beamforming image is a visualization method that is equivalent to arranging the beamforming grid points according to their physical locations to form a three-dimensional matrix where the first two dimensions encode the XY-position of the beamforming point and the third dimension contains the summed time-series. The image is formed by taking the



**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.

2633 time-averaged power (see Figure 4.11). Beamforming images are purely for the purposes  
 2634 of visualization and are not particularly useful for signal detection or reconstruction.

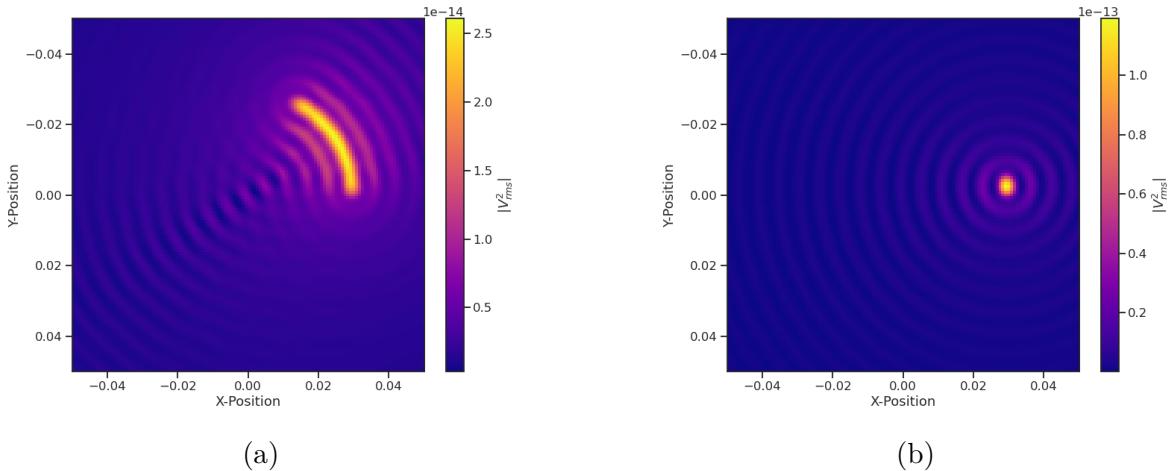
2635 If the beamforming phases consist only of the spatial phase component from Equation  
 2636 4.20, then the resulting beamforming image contains a relatively high-power ring-shaped  
 2637 region that is centered on the position of the electron (see Figure 4.11a). The origin  
 2638 of this shape is an additional phase offset particular to a cyclotron radiation source.  
 2639 Essentially, the circular motion that produces the cyclotron radiation introduces a relative  
 2640 phase offset to the electric fields that is equal to the azimuthal position of the field  
 2641 measurement point. For example, if we have two antennas, one located at an azimuthal  
 2642 position of  $0^\circ$  and another located at an azimuthal position of  $90^\circ$ , then the CRES signals  
 2643 received by these antennas will be out of phase by  $90^\circ$ , which is the difference in their  
 2644 azimuthal positions. This phase offset can be corrected by adding an additional term to  
 2645 the beamforming phase equation that is equal to the azimuthal position of the antenna  
 2646 relative to the electron,

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

2647 where  $\Delta\varphi_i$  is difference between the azimuthal position of the electron and the  $i$ -th  
 2648 antenna channel. Using the updated beamforming phases in the summation changes the

ring feature into a Bessel function peak whose maximum corresponds to the position of the electron. Including this cyclotron phase correction significantly improves the signal detection and reconstruction capabilities of beamforming by more than doubling the summed signal power and shrinking the beamforming maxima feature size.

The beamforming image examples in Figure 4.11 were produced using an electron located on the central axis of the magnetic trap, which do not experience  $\nabla B$ -drift. However, for electrons produced at non-zero radial position the beamforming phases must be made time-dependent in order to track the position of the electron's guiding center over time. Without this correction the  $\nabla B$ -drift causes the electron to move between beamforming positions, which effectively spreads the cyclotron radiation power over a wider area in the beamforming image (see Figure 4.12). This effect significantly



**Figure 4.12.** Beamforming images visualizing the reconstruction of an electron located off the central axis of the FSCD trap. In (a) we performing beamforming without the  $\nabla B$ -drift correction, and in (b) we include the  $\nabla B$ -drift correction.

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

The  $\nabla B$ -drift correction simply adds a circular time-dependence to the beamforming positions as a function of time,

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

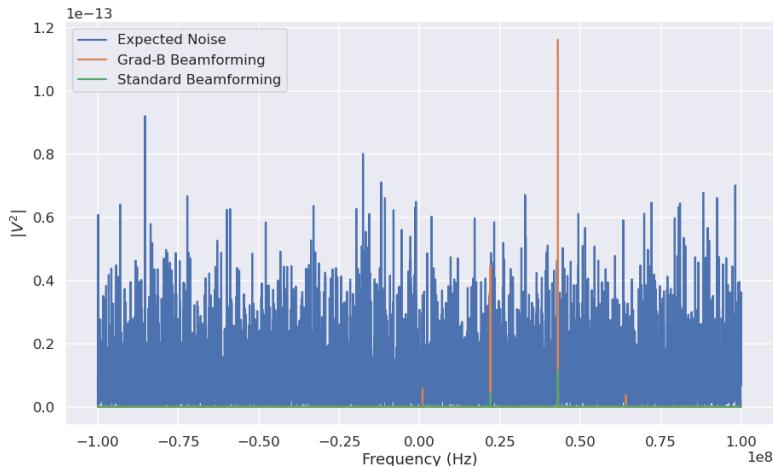
$$\varphi[n] = \varphi_0 + \omega_{\nabla B} t[n], \quad (4.25)$$

where  $\omega_{\nabla B}$  is the drift frequency and  $t[n]$  is the time vector. In the ideal case the  $\nabla B$ -drift frequencies from Figure 4.2 for the correct pitch angle and radial position would be used,

2666 however, it is not possible to know the electron's pitch angle a priori. In principle, one  
 2667 could perform multiple beamforming summations for a given beamforming position using  
 2668 different drift frequencies and choose the one that maximizes the summed power, but  
 2669 this approach leads to a huge computational burden that would be impractical for a  
 2670 real FSCD experiment. A compromise is to use an average value of  $\omega_{\nabla B}$  obtained by  
 2671 averaging over the drift frequencies for electrons of different pitch angle at a particular  
 2672 radius. This approach keeps the computational cost of time-dependent beamforming to a  
 2673 minimum while still providing a significant increase in the detection efficiency of digital  
 2674 beamforming.

## 2675 **Signal Detection with Beamforming and a Power Threshold**

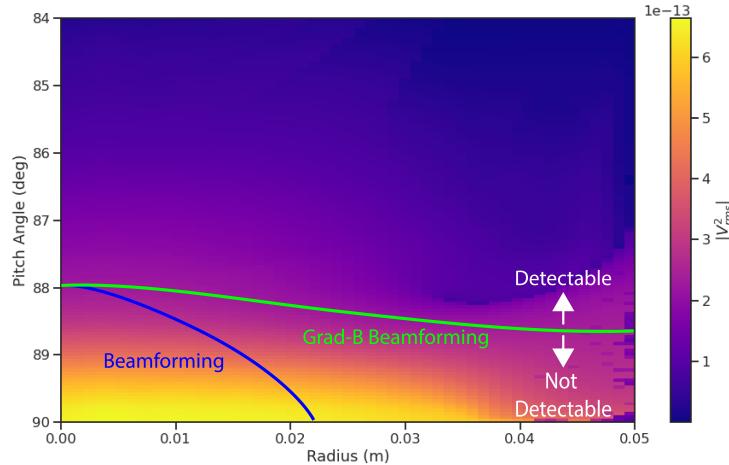
2676 Up to this point we have neglected any specific discussion of how digital beamforming is  
 2677 used for signal detection and reconstruction. This is because, strictly speaking, digital  
 2678 beamforming consists only of the phased summation of the array signals and cannot  
 2679 be used alone for signal detection. The example beamforming images shown in Figure  
 2680 4.11 and Figure 4.12 were produced using simulated data that contained no noise, which  
 2681 significantly degrades the utility of analyzing the beamforming images for signal detection  
 2682 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  $\nabla B$ -drift correction is needed to detect this electron since it comes from a simulation of an electron with a significant off-axis position.

2683     Digital beamforming as a detection algorithm is understood to mean digital beam-  
 2684     forming plus a detection threshold placed on the amplitude of the frequency spectrum  
 2685     obtained by applying a fast Fourier transform (FFT) to the summed time-series (see  
 2686     Figure 4.13). This approach is most similar to the time-frequency spectrogram analysis  
 2687     employed in previous CRES experiments, however, in principle any signal detection  
 2688     algorithm could be used after the beamforming procedure. In Section 4.4 I analyze the  
 2689     signal detection performance of the power threshold approach in detail.

2690     From the example frequency spectra in Figure 4.13 it is clear that without a re-  
 2691     construction technique that coherently combines the signals from the full antenna our  
 2692     ability to detect CRES signals will be drastically reduced. Because the CRES signals are  
 2693     in-phase at the correct beamforming position the summed power increases as a function  
 2694     of  $N^2$  compared to a single antenna channel, where  $N$  is the number of antennas. It  
 2695     is true that the noise power is also increased by beamforming, but, because the noise  
 2696     is incoherent, its power only increases linearly. Consequently, the signal-to-noise ratio  
 2697     (SNR) of the CRES signal increases linearly with the number of antennas, which greatly  
 2698     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  $\nabla B$ -drift frequencies in the beamforming phases. Both beamforming techniques fail to detect electrons below  $\approx 88.0^\circ$ , since these signal are composed of several relatively weak sidebands that are comparable to the noise.

2699     The power threshold detection algorithm searches for high-power frequency bins that  
 2700    should correspond to a frequency component of the CRES signal. In order to prevent  
 2701    random noise fluctuations from being mistaken as CRES signals the power threshold  
 2702    must be set high enough so that it is unlikely that random noise could be responsible. A  
 2703    consequence of this is that many electrons that can be trapped will go undetected because  
 2704    the modulation caused by axial oscillations leads to the cyclotron carrier power to falling  
 2705    below the decision threshold. The time-dependent beamforming used to correct for the  
 2706     $\nabla B$ -drift increases the volume of the magnetic trap where electrons can be detected,  
 2707    but it is ineffective at increasing the range of detectable pitch angles (see Figure 4.14).  
 2708    Fundamentally, this is because the power threshold only uses a fraction of the signal  
 2709    power to detect electrons and ignores the power present in the frequency sidebands. In  
 2710    the subsequent sections I examine two other signal detection algorithms that seek to  
 2711    improve the detection efficiency of the FSCD by utilizing the more of the signal shape to  
 2712    compute the detection test statistics.

### 2713    **4.3.2 Matched Filtering**

#### 2714    **Introduction to Matched Filtering**

2715    The problem of CRES signal detection is the problem of detecting a signal buried in  
 2716    WGN, which has been examined at great depth in the signal processing literature [71].  
 2717    For a fully known signal in WGN the optimal detector is the matched filter, which means  
 2718    that it achieves the highest true positive rate for a fixed rate of false positives. The  
 2719    matched filter test statistic is calculated by taking the inner product of the data with  
 2720    the matched filter template

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

2721    where  $h[n]$  is the matched filter template and  $y[n]$  is the data. The matched filter test  
 2722    statistic defines a binary hypothesis test in which the data vector is assumed to be an  
 2723    instance of two possible data classes. By setting a decision threshold on the value of  $\mathcal{T}$ ,  
 2724    one can classify a given data vector as belonging to two distinct hypotheses. Under the  
 2725    first hypothesis the data is composed of pure WGN, and under the second hypothesis the  
 2726    data is composed of the known signal with additive WGN. The matched filter template

2727 is obtained by rescaling the known signal in the following way

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

2728 where  $\tau$  is the variance of the WGN and  $x[n]$  is the known signal. Strictly speaking,  
 2729 Equation 4.27 is only true for noise with a diagonal covariance matrix, however, in the  
 2730 context of the FSCD we are justified in assuming this to be true. Defining the matched  
 2731 filter templates in this way guarantees that the expectation value of  $\mathcal{T}$  is equal to one  
 2732 when the data contains only noise, which is the standard matched filter normalization in  
 2733 the signal processing literature.

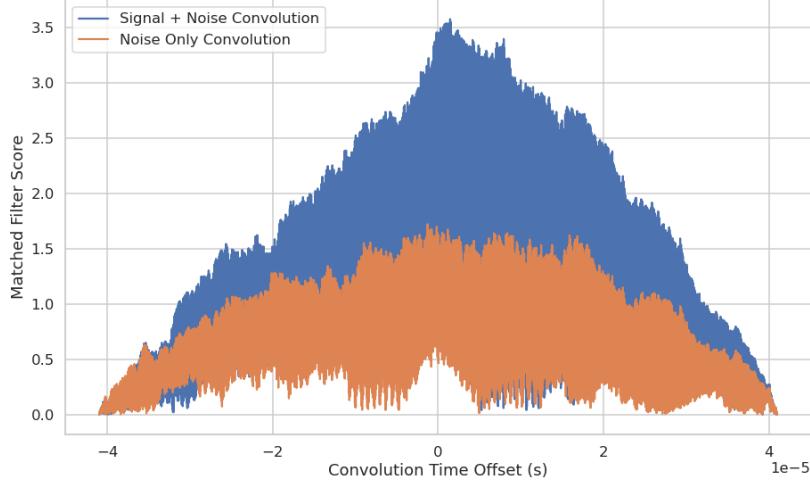
2734 Although matched filters are canonically formulated in terms of a perfectly known  
 2735 signal, it is still possible to apply the matched filter technique given imperfect information  
 2736 about the signal provided that the signal is deterministic. From our discussion of CRES  
 2737 simulation tools for the FSCD (see Section 4.2) we know that the shape of CRES signals  
 2738 are completely determined by the initial parameters of the electron. The random collisions  
 2739 with background gas molecules which cause the formation of signal tracks are the only  
 2740 stochastic component of the CRES event after the initial beta-decay, therefore, it is  
 2741 possible to develop a matched filter for the detection of CRES signal tracks which are fully  
 2742 determined by the parameters of the electron after the initial beta-decay or subsequent  
 2743 collision events.

2744 The matched filter test statistic for CRES signals is a modified version of Equation  
 2745 4.26

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

2746 where the matched filter inner product has been replaced with a convolution operation  
 2747 and a maximization over the template and convolution delay ( $m$ ). Replacing the inner  
 2748 product with a convolution accounts for the fact that the start time of the CRES signal is  
 2749 now an unknown parameter, in addition, we now perform a maximization of the matched  
 2750 filter convolution over a number of different templates. Because the shape of the signal is  
 2751 unknown we are forced to guess a number of different signal shapes to create a template  
 2752 bank with which we can identify unknown signals by performing an exhaustive search.

2753 The template bank approach to matched filtering, while quite powerful, can quickly  
 2754 become computationally intractable. This is especially true in the case of the FSCD  
 2755 because of the large amount of raw data produced by the array that must be analyzed.  
 2756 Specifically, the time-domain convolution specified by Equation 4.28 is particularly



**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.

2757 computationally intensive and is a major barrier towards the implementation of a  
 2758 matched filter for signal detection in an experiment like the FSCD. This can be avoided  
 2759 by using the convolution theorem to replace the time-domain convolution with an inner  
 2760 product in the frequency domain.

2761 The convolution theorem states that

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

2762 where  $\mathbf{f}$  and  $\mathbf{g}$  are discretely sampled time-series,  $\mathbf{F}$  and  $\mathbf{G}$  are the respective discrete  
 2763 Fourier transforms, and  $\mathcal{F}^{-1}$  is the inverse discrete Fourier transform operator. The  
 2764 convolution theorem allows us to perform the matched filter convolution by first com-  
 2765 puting the Fourier transform of the template and data, then performing a point-wise  
 2766 multiplication of the two frequency series, and finally performing the inverse Fourier  
 2767 transform to obtain the convolution output. Because discrete Fourier transforms can be  
 2768 performed extremely efficiently, the convolution theorem is almost always used in lieu of  
 2769 directly computing the convolution.

2770 One thing to note here is that the convolution theorem for discrete sequences shown  
 2771 here, is technically valid only for circular convolutions, which is not directly specified  
 2772 in Equation 4.28. However, because typical CRES track lengths are much longer than

2773 the Fourier analysis window and also that the frequency chirp rates are small compared  
2774 to the time-slice duration, it is relatively safe to use circular convolutions to evaluate  
2775 matched filter scores for CRES signals, which allows us to apply the convolution theorem  
2776 to compute matched filter scores using the frequency representation of the data and  
2777 matched filter template.

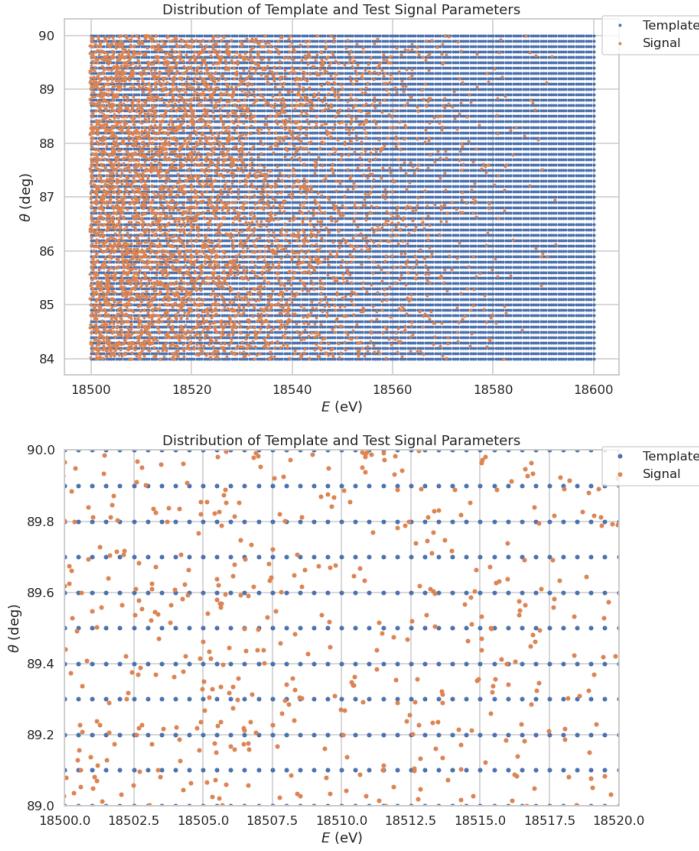
2778 **Matched Filter Analysis of the FSCD**

2779 The optimality provided by the matched filter makes it a useful algorithm for analysis  
2780 of CRES experiment designs for sensitivity analyses, since it indicates the best possible  
2781 detection efficiency achievable by an experiment configuration. The standard approach to  
2782 performing these studies involves generating a large number of simulated electron signals  
2783 that span the kinematic parameter space of electrons in the magnetic trap. In general,  
2784 electrons have six kinematic parameters along with an additional start time parameter.

2785 In order to limit the number of simulations required to evaluate the detection efficiency  
2786 the standard approach is to fix the starting axial position, starting azimuthal position,  
2787 starting direction of the perpendicular component of the electron's momentum, and event  
2788 start time to reduce the parameter space to starting radial position, starting kinetic  
2789 energy, and starting pitch angle. The fixed variables are true nuisance parameters that do  
2790 not affect the detection efficiency estimates for the FSCD design, because they manifest  
2791 as phases which are marginalized during the calculation of the matched filter score.

2792 Across radial position, kinetic energy, and pitch angle one defines a regular grid of  
2793 parameters and uses Locust to simulate the corresponding signals (see Figure 4.16). This  
2794 grid of simulated signals can be used to estimate the likelihood of detecting signals,  
2795 because the matched filter score specifies the shape of the PDF that defines the detection  
2796 probability and the size of the template bank influences the likelihood of a good match  
2797 between a template and a random signal.

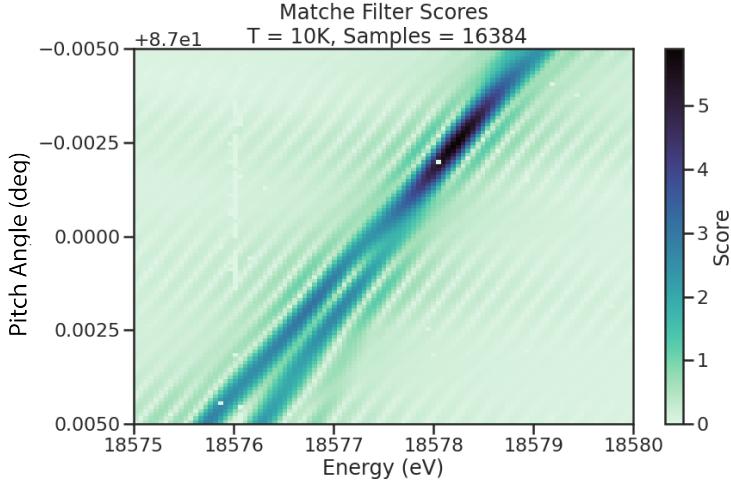
2798 The matched filter approach can also be used to estimate the achievable energy  
2799 resolution of the experiment by using a dense grid of templates generated with parameters  
2800 close to the unknown signal (see figure 4.17). Because matched filter templates with  
2801 similar parameters have signal shapes that are also similar, templates with incorrect  
2802 parameters can have nearly identical matched filter scores as the correct template. Since  
2803 only one sample of noise is included in a sample of real data, one cannot guarantee that  
2804 the best matching template corresponds to the ground truth parameters of the signal.  
2805 This introduces uncertainty into the signal parameter estimation that manifests as an  
2806 energy broadening. Dense grids of matched filter templates allows one to quantify this



**Figure 4.16.** An example two-dimensional parameter distribution of a matched filter template bank and random test signals.  $\theta$  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  $> 90\%$ .

broadening by analyzing the parameter space of templates with matched filter scores close to the ground truth. This approach is analogous to maximum likelihood estimation and is one key component of a complete sensitivity analysis for an antenna array CRES experiment.

A key parameter for describing the performance of a matched filter template bank at signal detection is match, which we define as the average ratio of the highest matched filter score for a random signal to the matched filter score for a perfectly matching



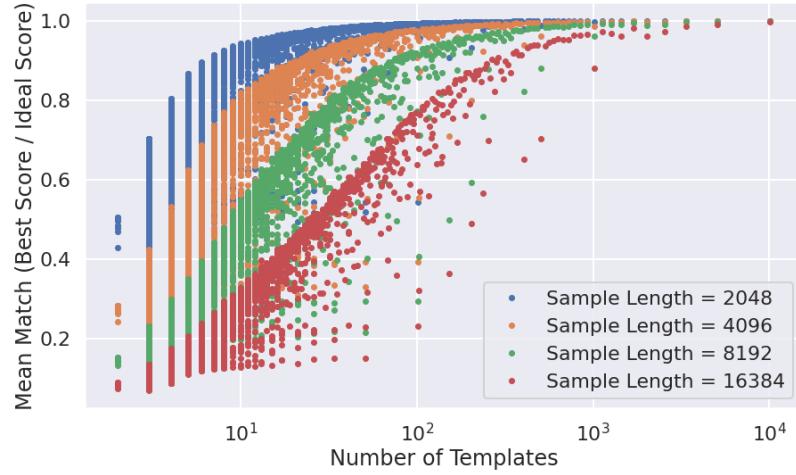
**Figure 4.17.** The matched filter scores of a dense grid of templates in pitch angle energy space. Dense template grids allow one to estimate the kinetic energy of the electron by identifying the best matching template. The uncertainty on this value is proportional to the space of templates that also 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.

2814 template. In equation form this is

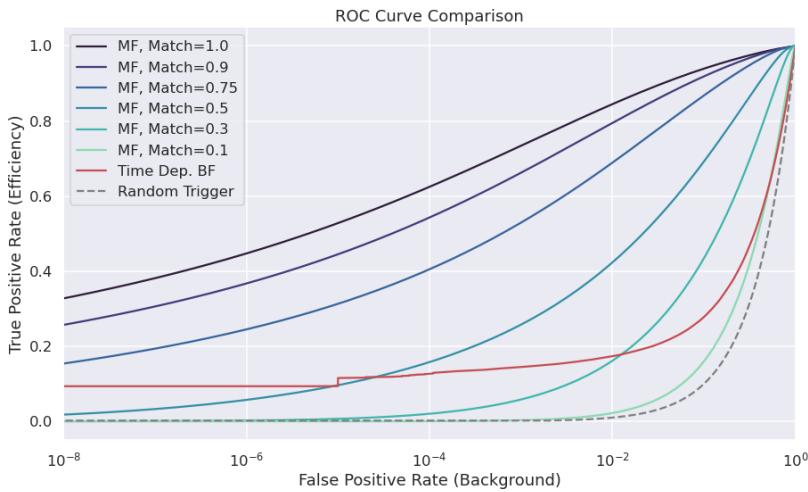
$$\text{Match} \equiv \Gamma = \frac{\mathcal{T}_{\text{best}}}{\mathcal{T}_{\text{ideal}}}, \quad (4.30)$$

2815 where  $\mathcal{T}_{\text{best}}$  is the matched filter score of the best fitting template in the bank and  $\mathcal{T}_{\text{ideal}}$  is  
 2816 the hypothetical matched filter score one would measure if the signal perfectly matched  
 2817 the template. Generally, one desires an average match as close to one as possible, however,  
 2818 the average match value is an exponential function of the number of templates in the  
 2819 template bank (see Figure 4.18). This behavior is observed for dense matched filter grids  
 2820 like the one in Figure 4.17. A dense grid was used to calculate the average value of match  
 2821 for different template bank sizes shown in Figure 4.18.

2822 The exponential relationship between match and template bank size is also evident  
 2823 for template banks that cover a wide range of parameters, such as the template bank  
 2824 visualized in Figure 4.16. Since no prior knowledge of the signal parameters is available,  
 2825 one has no choice but to use a template bank that covers a large range of parameters for  
 2826 signal detection. Achieving a high average match in this scenario can easily overwhelm  
 2827 the available computational resources, so in practice only a limited number of templates  
 2828 could be used at the detection stage. Therefore, accurately modeling the effects of match  
 2829 is key to correct sensitivity calculations.



**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 than shorter templates.



**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.

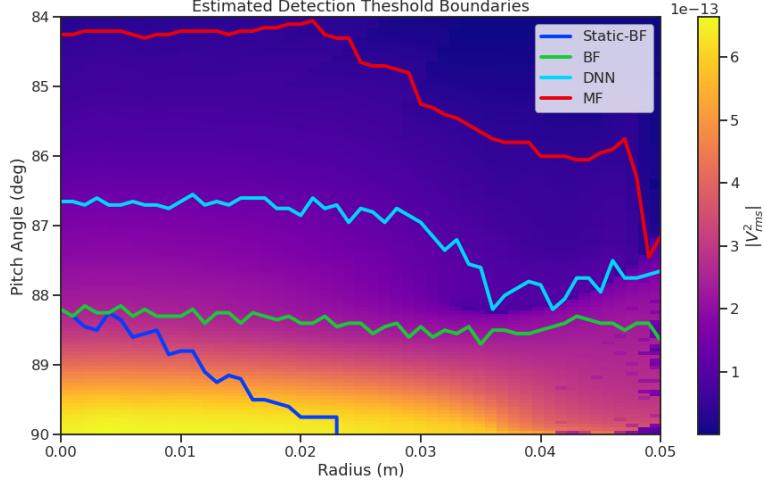
2830        The effect of match on the detection efficiency of the matched filter template bank can  
2831        be summarized using the ROC curve (see Figure 4.19). A single ROC curve is obtained  
2832        by averaging over the PDFs that describe the detection probabilities of each individual  
2833        template. The matched filter score for a template follows a Rician distribution with a  
2834        mean value equal to the matched filter score multiplied by the match ratio between the  
2835        template and signal. Therefore, the distribution that describes the average matched filter  
2836        score when there is a signal in the data is obtained by averaging over the distributions  
2837        for every template, whose expectation values are multiplied by the average match ratio.

2838        The distribution of the matched filter score when there is no signal in the data follows  
2839        a Rayleigh distribution. Therefore, a trials penalty, which is the statistical penalty one  
2840        pays for randomly checking many templates in order to avoid a random match between  
2841        noise and a template, is included by computing the joint distribution of  $N_{\text{template}}$  Rayleigh  
2842        distributions, where  $N_{\text{template}}$  is the size of the template bank. For more information on  
2843        the calculation of matched filter template bank ROC curves please refer to Section 4.4.

2844        An alternative way to visualize the detection performance for each algorithm is to  
2845        specify a minimum acceptable false positive rate at the trigger level. This is equivalent  
2846        to specifying a minimum threshold on the value of the matched filter score or the size of  
2847        a frequency peak for a beamforming power threshold trigger. One can then draw regions  
2848        of detectable signals as a function of the electron's pitch angle and radial position (see  
2849        Figure 4.20). A kinetic energy shift is equivalent to an overall frequency shift of the  
2850        signal and should have no effect on the detection probability assuming sufficient density  
2851        of matched filter templates in the energy dimension. A electron is declared "detectable"  
2852        for the regions in Figure 4.20 if the signal has at least 50% probability of falling above the  
2853        decision threshold of the respective classifier. One can see that the parameter space of  
2854        detectable signals is greatly expanded beyond the beamforming power threshold trigger  
2855        with a matched filter (MF) or deep neural network (DNN) (see Section 4.3.3). Plots such  
2856        as Figure 4.20 are useful for visualization, but, since the handling of detection likelihood  
2857        is not sufficiently rigorous, the detection probability boundaries are not well-suited to  
2858        sensitivity estimates.

## 2859        **Optimized Matched Filtering Implementation for the FSCD**

2860        The biggest practical obstacle to the implementation of a matched filter template bank  
2861        detection approach is oftentimes the computational cost associated with exhaustively  
2862        calculating the matched filter scores of the template bank, and the FSCD is no exception  
2863        in this regard. At a basic level computing a matched filter score requires the convolution



**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  $\nabla B$ -drift correction. The DNN algorithm refers to a convolutional neural network classifier trained to detect CRES signals (see Section 4.3.3).

of two vectors, which can be performed very efficiently by computers if the convolution theorem and fast Fourier transforms (FFT) are utilized. Furthermore, one can consider applying digital beamforming as a pre-processing step to reduce the dimensionality of the data before the matched filter is applied. In order to understand the relative gain in computational efficiency offered by these optimizations we analyze the total number of floating-point operations (FLOP) of several matched filter implementations in big  $O$  notation that utilize different combinations of optimizations.

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

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

FLOP per-template, whose cost is dominated by the  $O(M \times N)$  convolution operation. The computational cost of the direct matched filter approach can be significantly

2878 reduced by exploiting the convolution theorem and FFT algorithms. If we restrict  
 2879 ourselves to signals and templates that contain equal numbers of samples then the  
 2880 convolution can be calculated by Fourier transforming both vectors, performing the  
 2881 point-wise multiplication, and then performing the inverse Fourier transform to obtain  
 2882 the convolution result. The FFT algorithm is able to compute the Fourier transform  
 2883 utilizing only  $O(N \log N)$  operations compared to  $O(N^2)$  for a naive Fourier transform  
 2884 implementation. This optimization results in a computational cost per-template of

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

2885 A typical signal vector in the FSCD contains  $O(10^4)$  samples in which case the FFT  
 2886 reduces the computational cost of the matched filter by a factor of  $O(10^3)$ . This large  
 2887 reduction in computational cost implies that a direct implementation of a matched filter  
 2888 is completely infeasible in the FSCD due to resource constraints.

2889 Rather than relying solely on the matched filter it is tempting to consider using  
 2890 digital beamforming as an initial step in the signal reconstruction for the purposes of  
 2891 data reduction. The primary motivation is to reduce the dimensionality of the data by  
 2892 a factor of  $N_{\text{ch}}$  by combining the array outputs coherently into a single channel. One  
 2893 can view the beamforming operation as a partial matched filter, in the sense that the  
 2894 matched filter convolution contains the beamforming phased summation along with a  
 2895 prediction of the signal shape. By separating beamforming from the signal shape one  
 2896 hopes to reduce the overall computational cost by effectively shrinking the number of  
 2897 templates and reducing the number of operations required to check each one.

2898 The nature of this optimization requires that we account for the number of templates  
 2899 used for pure matched filtering versus the hybrid approach. To first order, the total  
 2900 number of templates at the trigger stage is a product of the number of guesses for each  
 2901 of the electron's parameters

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

2902 where  $N_E$  is the number of kinetic energies,  $N_\theta$  is the number of pitch angles,  $N_r$  is the  
 2903 number of starting radial positions, and  $N_\varphi$  is the number of starting azimuthal positions.  
 2904 The starting axial position and cyclotron motion phase are not necessary to include in  
 2905 the template bank since these parameters manifest themselves as the starting phase of  
 2906 the signal, which is effectively marginalized when using a FFT to compute the matched  
 2907 filter convolution. Therefore, the total number of operations required by a matched filter

2908 to detect a signal in a segment of array data is on the order of

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

2909 With the hybrid approach we attempt to remove the spatial parameters from the  
2910 template bank by using beamforming to combine the array signals into a single channel.  
2911 Beamforming explicitly assumes a starting position, which allows us to only use matched  
2912 filter templates that span the two-dimensional space of kinetic energy and pitch angle.  
2913 The total computational cost of the hybrid method is directly proportional to the number  
2914 of beamforming positions. For the time-dependent beamforming defined in Section 4.3.1,  
2915 the number of beamforming positions is given by

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

2916 where  $N_r$  and  $N_\varphi$  are the same spatial parameters encountered in the pure matched  
2917 filter template bank and  $N_{\omega_{\nabla B}}$  is the number of  $\nabla B$ -drift frequency assumptions. If a  
2918 unique drift frequency is used for each pitch angle then the hybrid approach is effectively  
2919 equivalent to a pure matched filter in the number of operations. The key efficiency gain  
2920 of the hybrid approach is to exploit the relatively small differences in  $\omega_{\nabla B}$  for electrons  
2921 of different pitch angles by using only a small number of average drift frequencies.

2922 The total number of operations for the hybrid approach can be expressed as a sum of  
2923 the operations required by the beamforming and matched filtering steps,

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

2924 The first product in the sum is the number of operations required by beamforming,  
2925 which is simply the number of beamforming points times the computational cost of the  
2926 beamforming matrix multiplication, and the second product is the computational cost  
2927 of matched filtering the summed signal generated by each beamforming position. To  
2928 compare this to pure matched filtering we take the ratio of Equations 4.34 and 4.36 to  
2929 obtain

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

2930 This expression can be simplified by observing that  $O(N_E N_\theta) \times O(\log N_s) \gg O(N_{\text{ch}})$ ,

2931 which means that the ratio of computational cost for the two methods can be reduced to

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

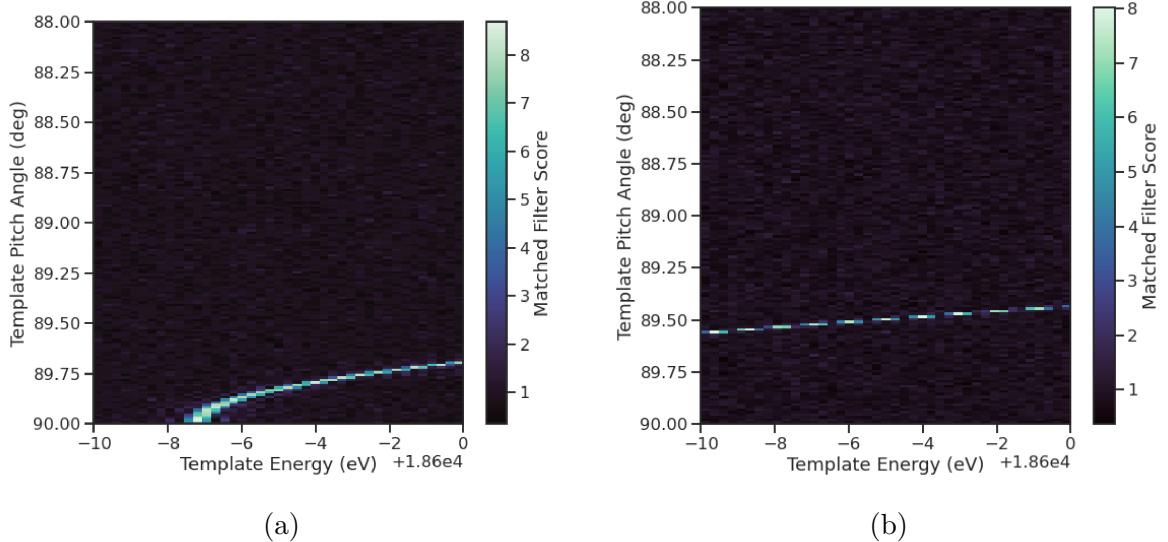
2932 If we limit ourselves to a number of estimated drift frequencies of  $O(1)$  then we see that  
2933 the estimated computational cost reduction of the hybrid approach is of  $O(N_{\text{ch}})$ . This is  
2934 quite a large reduction considering that the FSCD antenna array contains sixty antennas  
2935 in the baseline design.

2936 The main drawback of the hybrid approach is that the limited number of allowed  
2937 drift frequency guesses can lead to detection efficiency loss due to phase mismatch. The  
2938 degree of phase error from an incorrect drift frequency is proportional to the length of  
2939 the array data vector used by the signal detection algorithm. For signals with lengths  
2940 equal to the baseline FSCD Fourier analysis window of 8192 samples, typical phase errors  
2941 from using an average versus the exact  $\nabla B$ -drift frequency are on the order of a few  
2942 percent in terms of the signal energy. This has a relatively small impact on the overall  
2943 detection efficiency, however, future experiments with antenna array CRES will want to  
2944 balance optimizations such as these during the design phase to keep experiment costs to  
2945 a minimum while still achieving scientific goals.

## 2946 Kinetic Energy and Pitch Angle Degeneracy

2947 More accurate modeling of a matched filter requires that we consider the effects of  
2948 mismatched signals and template, since this more accurately reflects the real-world usage  
2949 of a matched filter where many incorrect templates are convolved with the data until the  
2950 matching template is found. One way to study this is to use the grid of simulated signals  
2951 to compute the matched filter scores between mismatched signals and templates and  
2952 evaluate the matched filter scores under this scenario. What one finds when performing  
2953 this analysis is that templates for kinetic energies and pitch angles that do not match  
2954 the underlying signal can have matched filter scores that are indistinguishable from the  
2955 matched filter score of the correct template (see Figure 4.21 and Figure 4.21).

2956 This degeneracy in matched filter score is the result of correlations between the kinetic  
2957 energy of the electron and the pitch angle caused by changes in the average magnetic field  
2958 experienced by an electron for different pitch angles. While in principle helpful for the  
2959 purposes of signal detection these correlations are unacceptable since they greatly reduce  
2960 the energy resolution of the experiment by causing electrons with specific kinetic energy  
2961 to templates across a wide range of energies. It is important to emphasize that this



**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.

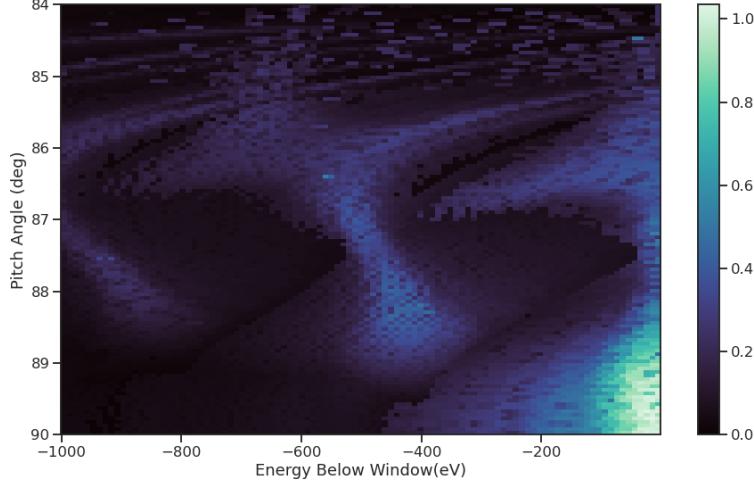
2962 degeneracy cannot be fixed by implementing a different signal reconstruction algorithm.  
 2963 As revealed by the matched filter scores the shapes of the signals for different parameters  
 2964 are identical. Resolving this degeneracy between pitch angle and energy requires the  
 2965 design of a new magnetic trap with steeper walls so that the average magnetic field  
 2966 experienced by an electron is less dependent on pitch angle.

### 2967 **4.3.3 Machine Learning**

2968 Machine learning is a vast and rapidly developing field of research [74]. In this Section  
 2969 we shall attempt to provided a brief introduction to some of the concepts and techniques  
 2970 of machine learning that were applied to CRES signal detection rather than attempt a  
 2971 comprehensive overview.

#### 2972 **Introduction to Machine Learning**

2973 Digitization of the FSCD antenna array generates large amounts of data that must be  
 2974 rapidly processed to enable real-time signal detection and reconstruction. While digital  
 2975 beamforming combined with a power threshold is relatively computationally inexpensive,  
 2976 it is relatively ineffective at detecting CRES signal with small pitch angles, since it relies



**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.

2977 on a visible frequency peak above the noise. On the other hand, a matched filter is able  
 2978 to detect signals with a significantly larger range of parameters, however, the exhaustive  
 2979 search of matched filter templates can be computationally expensive. Machine learning  
 2980 based triggering algorithms have been used successfully in many different high-energy  
 2981 physics experiments [75] and recent developments have shown success in the detection  
 2982 of gravitational wave signals using machine learning techniques [76, 77] in place of the  
 2983 more traditional matched filtering method. This motivates the exploration of machine  
 2984 learning as a potential CRES signal detection algorithm.

2985 There are several different approaches to machine learning, but the one most important  
 2986 to our discussion here is the supervised learning approach. In supervised machine learning  
 2987 one uses a differentiable model or function that is designed to map the input data to the  
 2988 appropriate label [74]. The data is represented as a multidimensional matrix of floating  
 2989 point values such as an image or a time-series, and the label is generally a class name  
 2990 such as signal or noise for classification problems or a continuous value like kinetic energy  
 2991 in the case of regression problems.

2992 In supervised learning the model is trained to map from the data to the correct label  
2993 by evaluating the output of the model using a training dataset consisting of a set of  
2994 paired data and labels. To evaluate the difference between the model output and the  
2995 correct label a loss function is used to quantify the error between the model prediction  
2996 and the ground truth. For example, a common loss function in regression problems is the  
2997 squared error loss function, which quantifies error using the squared difference between  
2998 the model output and label.

2999 Using the outputs of the loss function the next step in supervised learning is to  
3000 compute the gradient of error with respect to the model parameters in a process called  
3001 backpropagation. Using the model parameter gradients the last step in the supervised  
3002 learning process is to perform an update of the parameter values in order to minimize  
3003 the error in the model predictions across the whole dataset. This loop is performed many  
3004 times while randomly shuffling the dataset until the error converges to a minimum value  
3005 at which point the training procedure has finished. It is standard practice to monitor  
3006 the training procedure by evaluating the performance of the model using a separate  
3007 validation dataset that matches the statistical distribution of the training data and to  
3008 check the performance of the model after training using yet another dataset called the  
3009 test dataset. These practices help to guard against overtraining which is a concern for  
3010 models with many parameters.

### 3011 Convolutional Neural Networks

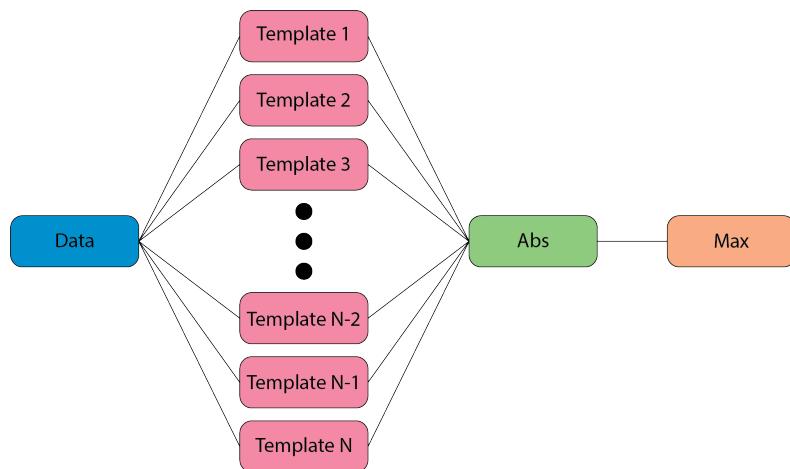
3012 A popular class of machine learning models are neural networks. A neural network is  
3013 essentially a function composed of a series of linear operations called layers which take a  
3014 piece of data typically represented as a matrix, multiplies the elements of the data by a  
3015 weight, and then sums these products to produce an output matrix. Neural networks  
3016 composed of purely linear operations are unable to model complex non-linear behavior,  
3017 therefore, non-linear activation functions are applied to the outputs of each of the layers  
3018 to increase the ability of the neural network to model complex relationships between the  
3019 data.

3020 Neural networks are typically composed of at least three layers, but with the present  
3021 capabilities of computer hardware they more often contain many more than this. The  
3022 first layer in a neural network is called the input layer, because it takes the data objects  
3023 as input, and the last layer in a neural network is known as the output layer. The  
3024 output layer is trained by machine learning to map the data to a desired output using  
3025 the supervised learning procedure described in Section 4.3.3. In between the input and

3026 the output layer are typically several hidden layers that receive inputs from and transmit  
3027 outputs to other layers in the neural network model. The term deep neural network  
3028 (DNN) refers to those neural networks that have at least one hidden layer, which have  
3029 proven to be extremely powerful tools for pattern recognition and function approximation.

3030 An important type of DNN are convolutional neural networks (CNN) that typically  
3031 contain several layers which perform a convolution of the input with a set of filters. These  
3032 convolution operations are typically accompanied by layers that attempt to down-sample  
3033 the data along with the standard neural network activation functions. A standard CNN  
3034 is composed of several convolutional layers at the beginning of the network and ends  
3035 with a series of fully-connected neural network layers at the output. Intuitively, one  
3036 can imagine that the convolutional layers are extracting features from the data that  
3037 fully-connected layers use to perform the classification or regression task.

3038 **Deep Filtering for Signal Detection in the FSCD**



**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.

3039 CNNs have been extremely influential in the field of computer vision, particularly tasks  
3040 such as image segmentation and classification, but have also been applied in numerous  
3041 experimental physics contexts. Given the particular challenge posed by signal detection  
3042 and reconstruction in the FSCD we are interested in exploring the potential of machine  
3043 learning as an effective algorithm for real-time signal detection, since this application  
3044 requires both high efficiency and fast evaluation.

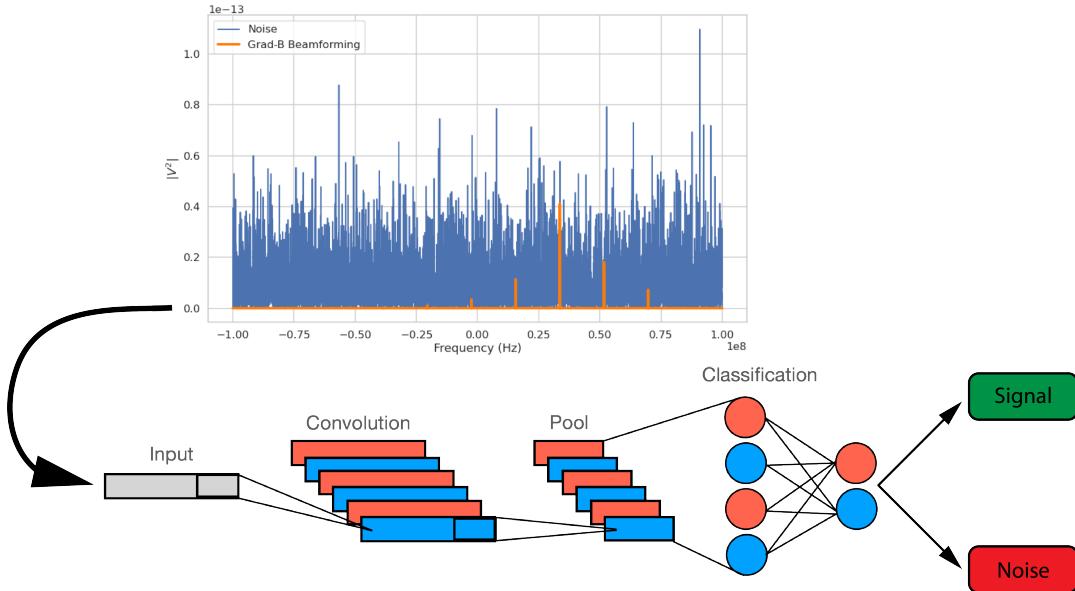
3045 In the machine learning paradigm signal detection is equivalent to a binary classifi-

3046 cation problem between the signal and noise data classes, and my investigation focuses  
3047 specifically on the application of CNNs to signal detection in the FSCD, which is moti-  
3048 vated by relatively recent demonstrations of CNNs achieving classification accuracies for  
3049 gravitational wave time-series signals comparable to a matched filter template bank. In  
3050 this framework it is possible to interpret the matched filter as a type of CNN composed  
3051 of a single convolutional layer with the templates making up the layer filters (see Figure  
3052 4.23). Since this neural network has no hidden layers, it is not a DNN like we have  
3053 been discussing so far, but we can attempt to construct a proper CNN that attempts to  
3054 reproduce the classification performance of the matched filter network.

3055 The name deep filtering refers to this scheme of replacing a matched filter template  
3056 bank with a DNN. The reason why one might want to do this is that it may be possible to  
3057 exploit redundancies and correlations between templates that may allow one to perform  
3058 signal detection with similar accuracy but with fewer computations, which is important  
3059 for real-time detection scenarios like the FSCD experiment. In Section 4.4 we perform a  
3060 detailed comparison of the signal detection performance of a CNN to beamforming and a  
3061 matched filter template bank.

3062 Deep filtering is conceptually a simple technique. Similar to a matched filter template  
3063 bank a large number of simulated CREs signals are generated and used to train a model  
3064 to distinguish between signal and noise data (see Figure 4.24). In order to reduce the  
3065 dimensionality of the input FSCD data a digital beamforming summation is applied  
3066 to the raw time-series data generated by Locust to compress the 60-channel data to a  
3067 single time-series. CREs signal have a sparse frequency representation and experiments  
3068 training CNN's on time-series and frequency series data found that models trained on  
3069 frequency spectrum data performed significantly better, therefore, an FFT is applied to  
3070 the summed time-series before being normalized and fed to the classification model.

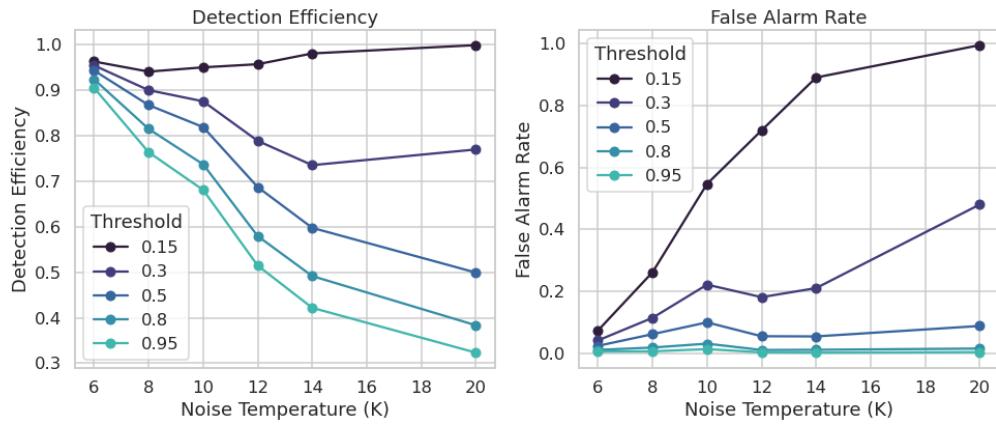
3071 The data used to train the model consists of an equal proportion of signal and noise  
3072 frequency spectra. Unique samples of WGN are generated and added to the signals during  
3073 training time to avoid having to pre-generate and store large samples of noise data. The  
3074 binary cross-entropy loss function combined with the ADAM optimizer proved effective  
3075 at training the models to classify CREs data. A simple hyperparameter optimization  
3076 was performed by manually tuning model, loss function, and optimizer parameters. The  
3077 model and training loops were implemented in python using the PyTorch deep learning  
3078 framework. Standard machine learning best practices were followed when training the  
3079 models, such as overtraining monitoring using a validation dataset. Models were trained  
3080 until the training loss and accuracy converged and then evaluated using a separate test



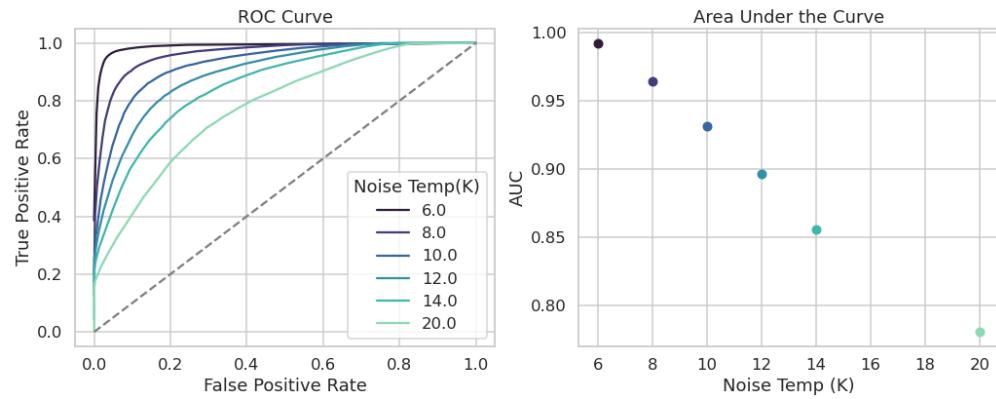
**Figure 4.24.** A graphical depiction of CRES signal detection using a CNN. A noisy segment of data is converted to a frequency series using digital beamforming and a FFT. The complex-valued frequency series is input into a trained CNN model that classifies the data as signal or noise using a decision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

3081 data set.

3082 The classification results of the test dataset are used to quantify the relationship  
 3083 between the true positive rate and the false positive rate for the model. The true positive  
 3084 rate is analogous to detection efficiency and the false positive rate is a potential source of  
 3085 background in the detector. One can limit the rate of false positives using a sufficiently  
 3086 high threshold on the model output at the cost of a lower detection efficiency (see Figure  
 3087 4.25 and Figure 4.26). As expected, the performance of the model at signal classification  
 3088 is negatively effected the noise power, which is quantified by the noise temperature.



**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.



**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.

## 3089   **4.4 Analysis of Signal Detection Algorithms for the An-** 3090   **tenna Array Demonstrator**

3091   This section contains an early version of the manuscript for the triggering paper prepared  
3092   for publication in JINST. In it I present a relatively detailed analysis of the signal  
3093   detection performance of the three signal detection approaches discussed so far using a  
3094   population of simulated CRES signals generated with Locust. The focus of the paper is  
3095   on the performance of the signal detection algorithms for pitch angles below 88.5° where  
3096   the beamforming power threshold begins to fail.

### 3097   **4.4.1 Introduction**

3098   Cyclotron Radiation Emission Spectroscopy (CRES) is a technique for measuring the  
3099   kinetic energies of charged particles by observing the frequency of the cyclotron radiation  
3100   that is emitted as they travel through a magnetic field [40]. The Project 8 Collaboration  
3101   is developing the CRES technique as a next-generation approach to tritium beta-decay  
3102   endpoint spectroscopy for neutrino mass measurement. Recently, Project 8 has used  
3103   CRES to perform the first ever tritium beta-decay energy spectrum and neutrino mass  
3104   measurement [42, 43].

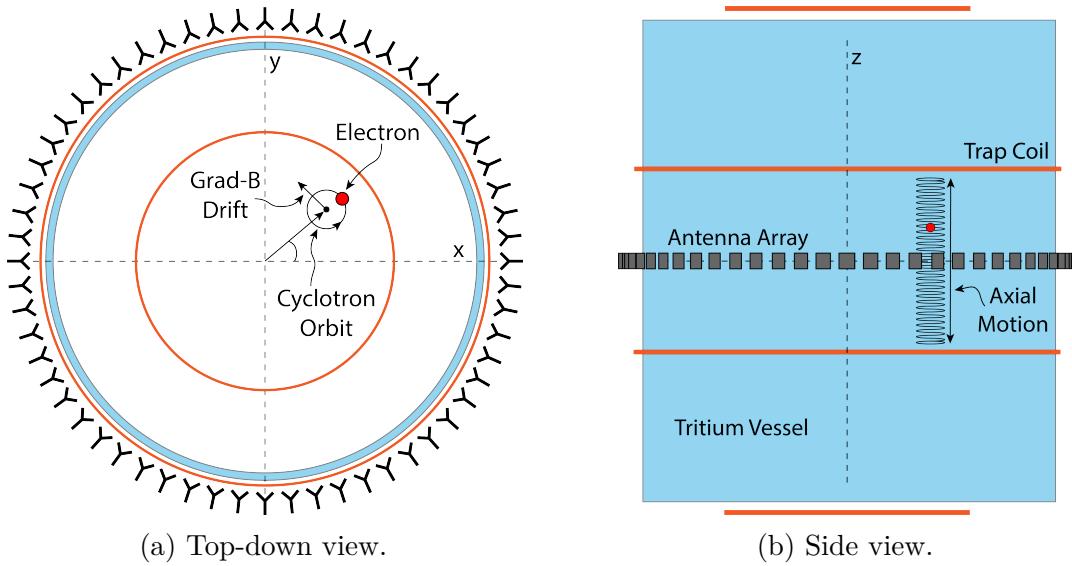
3105   Previous CRES measurements have utilized relatively small volumes of gas that are  
3106   directly integrated with a waveguide transmission line, which transmits the cyclotron  
3107   radiation emitted by the trapped electrons to a cryogenic amplifier. While this technology  
3108   has had demonstrable success, it is not a feasible option for scaling up to significantly  
3109   larger measurement volumes. In particular, the goal of the Project 8 Collaboration  
3110   is to use CRES combined with atomic tritium to measure the neutrino mass with a  
3111   40 meV sensitivity. Achieving this sensitivity goal will require a multi-cubic-meter scale  
3112   measurement volume in order to obtain the required event statistics in the tritium  
3113   beta-spectrum endpoint region; hence, there is a need for new techniques to enable large  
3114   volume CRES measurements for future experiments.

3115   One approach is to surround a large volume with an array of antennas that together  
3116   collect the cyclotron radiation emitted by trapped electrons [41, 78]. A promising  
3117   array design is an inward-facing uniform cylindrical array that surrounds the tritium  
3118   containment volume. Increasing the size of the antenna array, by adding additional  
3119   rings of antennas along vertical axis, allows one to grow the experimental volume until a  
3120   sufficient amount of tritium gas can be observed by the array. A challenging aspect of

3121 this approach is that the total radiated power emitted by an electron near the tritium  
3122 spectrum endpoint is on the order of 1 fW or less, which is then distributed between  
3123 all the antennas in the array. Consequently, detecting the presence of a CRES signal  
3124 and determining the electron's kinetic energy requires reconstructing the entire antenna  
3125 array output over the course of the CRES event, posing a significant data acquisition  
3126 and signal reconstruction challenge.

3127 Project 8 has developed a triggering system to enable real-time identification of CRES  
3128 events using an antenna array [79]. Previous measurements with the CRES technique  
3129 have utilized a threshold on the frequency spectrum formed from a segment of CRES  
3130 time-series data. This algorithm relies on the detection of a frequency peak above the  
3131 thermal noise background, which limits the kinematic parameter space of detectable  
3132 electrons. Due to the limitations of this power threshold, Project 8 has been investigating  
3133 alternative signal identification approaches, including both matched filtering and machine  
3134 learning based classifiers, to improve the detection efficiency of the experiment. In  
3135 order to evaluate the relative gains in detection efficiency that come from utilizing  
3136 these alternative algorithms, we develop analytical models for the power threshold and  
3137 matched filter signal classifier performance applicable to an antenna array based CRES  
3138 detector. In addition, we implement and test a basic convolutional neural network (CNN)  
3139 as a first step towards the development of neural-network based classifiers for CRES  
3140 measurements. These results allow us to compare the estimated detection efficiencies of  
3141 each of these methods, which we weigh against the associated computational costs for  
3142 real-time applications.

3143 The outline of this paper is as follows. In Section 4.4.2 we give an overview of a  
3144 prototypical antenna array CRES experiment, and describe the major steps involved  
3145 in the proposed approach to real-time signal identification. In Section 4.4.3 we develop  
3146 models for the power threshold and matched filter algorithms, and introduce the machine  
3147 learning approach and CNN architecture. In Section 4.4.4 we describe our process for  
3148 generating simulated CRES signal data and the details of training the CNN. Finally,  
3149 in Section 4.4.5 we perform a comparison of the signal classification accuracy of the  
3150 three approaches and discuss the relevant trade-offs in terms of detection efficiency and  
3151 computational cost.



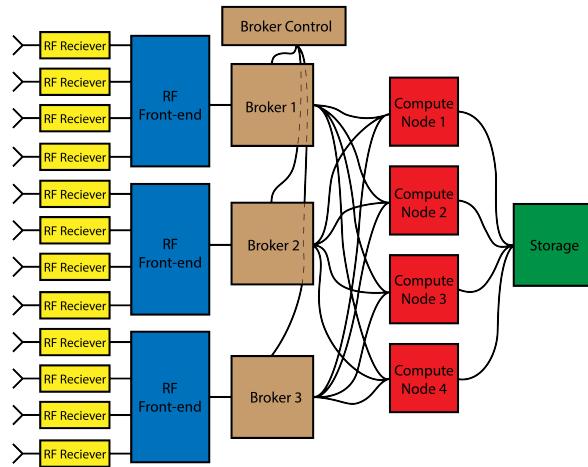
**Figure 4.27.** An illustration of the conceptual design for an antenna array CRES tritium beta-decay spectrum measurement. The antenna array geometry consists of a 20 cm interior diameter with 60 independent antenna channels arranged evenly around the circumference. The nominal antenna design is sensitive 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 1 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.

## 3152 4.4.2 Signal Detection with Antenna Array CRES

### 3153 4.4.2.1 Antenna Array and DAQ System

3154 In order to explore the potential of antenna array CRES for neutrino mass measurement,  
 3155 the Project 8 Collaboration has developed a conceptual design for a prototype antenna  
 3156 array CRES experiment [41, 78], called the Free-space CRES Demonstrator or FSCD,  
 3157 which could be used as a demonstration of the antenna array measurement technique  
 3158 (see Figure 4.27). The FSCD utilizes a single ring of antennas, which is the simplest  
 3159 form of a uniform cylindrical array configuration, to surround a radio-frequency (RF)  
 3160 transparent tritium gas vessel. A prototype version of this antenna array has been built  
 3161 and tested by the Project 8 collaboration to validate simulations of the array radiation  
 3162 pattern and beamforming algorithms [44]. In the FSCD the antenna array is positioned  
 3163 at the center of the magnetic trap formed by a set of electro-magnetic coils that are  
 3164 designed to produce a magnetic trap with flat central region and steep walls both radially  
 3165 and axially.

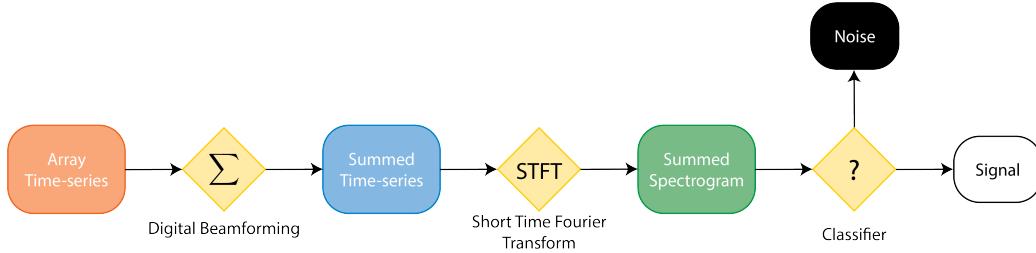
3166 When a beta-decay electron is trapped its motion consists of three primary components.  
 3167 The component with the highest frequency is the cyclotron orbit whose frequency is  
 3168 determined by the size of the background magnetic field. The FSCD design assumes  
 3169 a background magnetic field value of approximately 0.96 T, which results in cyclotron  
 3170 frequencies for electrons with kinetic energies near the tritium beta-spectrum endpoint  
 3171 from 25 to 26 GHz. The component with the next highest frequency is the axial oscillation  
 3172 experienced by electrons with pitch angles of less than 90° [61]. The flat region of the  
 3173 FSCD magnetic trap extends approximately 3 cm above and below the antenna array  
 3174 plane causing electrons to move back and forth as they are reflected from the trap walls.  
 3175 Typical oscillation frequencies are on the order of 10's of MHz, which results in an  
 3176 oscillation period that is  $O(10^3)$  smaller than the duration of a typical CRES event.  
 3177 Therefore, when reconstructing CRES events we treat the electron as occupying only an  
 3178 average axial position at the center of the magnetic trap, since we are not able to resolve  
 3179 the axial position as a function of time. The component of motion with the smallest  
 3180 frequency is  $\nabla B$ -drift caused by radial field gradients in the trap, producing an orbit of  
 3181 the electron around the central axis of the trap with a frequency on the order of a few  
 3182 kHz, dependent on the pitch angle and the radial position of the electron.



**Figure 4.28.** A high-level diagram of the DAQ system archctecture envisioned for the FSCD.

3183 The data acquisition (DAQ) system digitizes the signals from the antenna array and  
 3184 combines thee data streams into a time-ordered matrix of array snapshots that can be  
 3185 used by the reconstruction algorithms. The FSCD DAQ system design [79] is divided into  
 3186 three layers 4.28. The first layer is the RF front-end, which includes the antenna array,  
 3187 the RF receiver boards, and the digitization electronics. The receiver board contains an  
 3188 amplifier, RF mixer, and bandpass filter to enable down-conversion, and is followed by

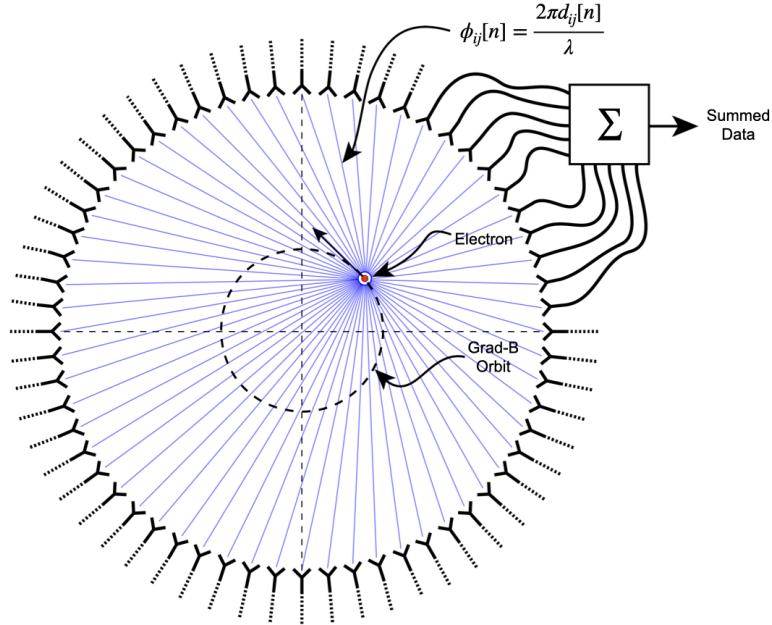
3189 the digitization electronics that sample the CRES signals at 200 MHz. In order to achieve  
 3190 an adequate signal-to-noise ratio to detect CRES events, the DAQ system for the antenna  
 3191 array demonstrator must have a total system noise temperature of  $\approx 10$  K, which we  
 3192 can achieve by using low-noise amplifiers and operating at cryogenic temperatures. After  
 3193 digitization, the array data must be reorganized from individual data streams sorted  
 3194 by channel into array snapshots sorted by time. In order to solve this data transfer  
 3195 and networking problem the second layer of the DAQ system consists of a set of broker  
 3196 computer nodes that reorganize the array data into time-ordered chunks. This approach  
 3197 allows us accommodate different data transfer requirements by scaling the number of  
 3198 broker nodes in this layer accordingly. Next, the broker layer distributes these chunks  
 3199 of array data to the final layer of the DAQ system, which consists of a set of identical  
 3200 reconstruction nodes that perform the calculations required for CRES reconstruction.  
 3201 Similar to the broker layer, the number of reconstruction nodes can be increased or  
 3202 decreased depending on the amount of computer power required for real-time CRES  
 3203 reconstruction.



**Figure 4.29.** A block diagram illustration of the real-time triggering algorithm proposed for antenna array CRES reconstruction.

3204 The design of the FSCD DAQ system is intended to enable a significant portion of  
 3205 the CRES event reconstruction to occur in real-time. The motivation for this comes from  
 3206 the fact that the FSCD antenna array generates approximately 1 exabyte of raw data  
 3207 per year of operation. Therefore, in order to reduce the data-storage requirements, it is  
 3208 ideal to perform at least some of the CRES event reconstruction in real-time so that it  
 3209 is possible to save a reduced form of the data for offline analysis. The first step of the  
 3210 real-time reconstruction would be a real-time signal detection algorithm, which is the  
 3211 focus of this paper. Our approach consists of three main operations performed on the  
 3212 time-series data blocks including digital beamforming, a short time Fourier transform  
 3213 (STFT), and a binary classification algorithm to distinguish between signal and noise  
 3214 data (see Figure 4.29).

3215 4.4.2.2 Real-time Signal Detection



**Figure 4.30.** An illustration of the digital beamforming procedure. The blue lines indicate the various distances from the beamforming position to the antenna. In the situation depicted the actual position of the electron matches the beamforming position, so we should expect constructive interference when the phase shifted signals are summed. To prevent the electron's  $\nabla B$ -motion from moving the electron off of the beamforming position, the beamforming phase include a time-dependence to follow the trajectory of the electron in the magnetic trap.

3216 The first step in the real-time detection algorithm is digital beamforming, which is  
 3217 a phased summation of the signals received by individual antennas in the array (see  
 3218 Figure 5.21). The phase shifts correspond to the path length differences between a spatial  
 3219 position and each individual antenna such that, when there is an electron located at  
 3220 the beamforming position, all the signals received by the array constructively interfere.  
 3221 Since we do not know ahead of time where an electron will be produced in the detector,  
 3222 we define a grid of beamforming positions that cover the entire region where electrons  
 3223 can be trapped and perform a phased summation for each of these points for every  
 3224 time-step in the array data block. As we saw in Section 4.4.2.1, the axial oscillation  
 3225 of the electrons prevents us from resolving it's position along the Z-axis of the trap,  
 3226 therefore our beamforming grid need only cover the possible positions of the electron in  
 3227 the two-dimensional plane defined by the antenna array.

3228 The equation defining digital beamforming can be expressed as

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

3229 where  $\mathbf{x}[n]$  is array snapshot vector at the sampled time  $n$ ,  $\Phi[n]$  is the matrix of  
 3230 beamforming phase shifts, and  $\mathbf{y}[n]$  is summed output vector that contains the voltages  
 3231 for each of the summed channels that correspond to a particular beamforming position.  
 3232 The elements of the beamforming phase shift matrix can be expressed as a weighted  
 3233 complex exponential

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

3234 where the indices  $i$  and  $j$  label the beamforming and antenna positions respectively. The  
 3235 weight  $A_{ij}$  accounts for the relative power increase for antennas that are closer to the  
 3236 position of the electron, and  $\phi_{ij}$  is the total beamforming phase shift for the  $j$ -th antenna  
 3237 at the  $i$ -th beamforming position.

3238 The beamforming phase shift is a sum of two terms

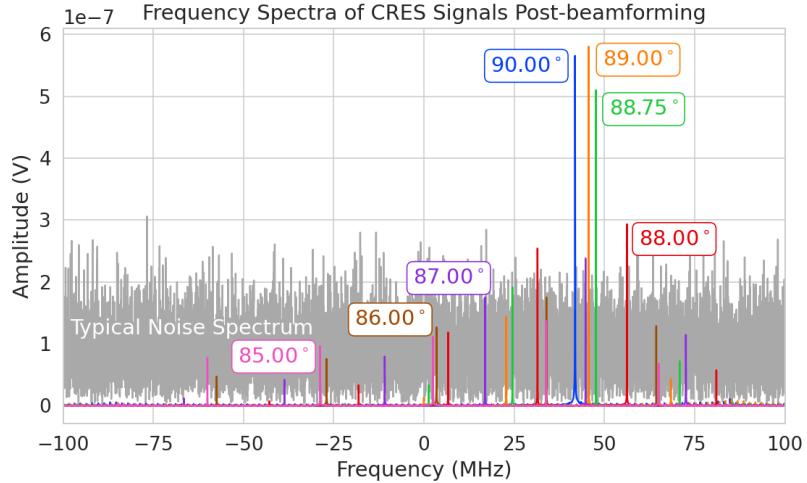
$$\phi_{ij}[n] = \frac{2\pi d_{ij}[n]}{\lambda} + \theta_{ij}[n], \quad (4.41)$$

3239 where the first term is the phase shift originating from the path length difference ( $d_{ij}[n]$ )  
 3240 between the beamforming and antenna positions, which are represented by the vectors  
 3241  $(r_j, \theta_j)$  and  $(r_i, \theta_i[n])$ , and the second term is the angular separation ( $\theta_{ij}[n]$ ) of the two  
 3242 positions. The angular separation enters into the beamforming phase due to an effect  
 3243 caused by the circular orbit of the electron that produces radiation whose phase is linearly  
 3244 dependent on the relative azimuthal position of the antenna [80,81]. The time-dependence  
 3245 of the beamforming phases is intended to correct for the effects of  $\nabla B$ -drift, which cause  
 3246 the guiding centers of electrons to orbit the center of the magnetic trap. By including a  
 3247 linear time-dependence in the azimuthal beamforming position,

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

3248 where  $\omega_{\nabla B}$  is the azimuthal grad-B drift frequency,  $t[n]$  is the time vector and,  $\theta_{i,0}$  is the  
 3249 starting azimuthal position, we can configure the beamforming phases to effectively track  
 3250 the XY-position of the guiding center over the event duration. Predicting accurate values  
 3251 of  $\omega_{\nabla B}$  for a specific trap and set of kinematic parameters will be done by simulations,  
 3252 which are performed using the Kassiopeia software package [60] by Project 8.

3253 After digital beamforming, we apply a short-time Fourier transform (STFT) to the



**Figure 4.31.** Frequency spectra of simulated CRES signals post-beamforming. The signal of a  $90^\circ$  electron consists of a single frequency component that is easy to detect with a power threshold on the frequency spectrum. This power threshold is still effective for signals with relatively large pitch angles such as  $89.0^\circ$  and  $88.75^\circ$ , which are composed of a main carrier and a few small sidebands. Signals with smaller pitch angles, below about  $88.5^\circ$ , tend to be dominated by sidebands such that no single frequency component can be reliably distinguished from the noise with a power threshold.

3254 summed time-series to obtain the frequency spectrum representation of the signals (see  
 3255 Figure 4.31). From the detection perspective, the frequency representation of the CRES  
 3256 data is advantageous compared to the time domain, because the frequency spectra of  
 3257 CRES signals are well-approximated by a frequency and amplitude modulated sinusoidal  
 3258 whose carrier frequency increases as a linear chirp. The modulation is caused by the axial  
 3259 oscillation of the electron in the magnetic trap and produce frequency spectra that are  
 3260 well-described by a small number of frequency components. The linear chirp is caused  
 3261 by the energy loss due to cyclotron radiation, which results in a relatively slow increase  
 3262 in the frequency components of the CRES signal over time. During the standard Fourier  
 3263 analysis window for the FSCD of  $40.96 \mu\text{sec}$ , we expect a typical CRES signal to increase  
 3264 in frequency by approximately 15 kHz, which is smaller than the frequency bin width  
 3265 given the 200 MHz sample rate. Therefore when considering a single frequency spectrum  
 3266 it is justifiable to neglect the effects of the linear frequency chirp.

3267 In the cases where the electron's pitch angle is  $\gtrsim 88.5^\circ$ , the majority of the signal  
 3268 power is contained in a single frequency component, with the remaining signal power  
 3269 contained in a small number of sidebands proportional to the electron's axial modulation  
 3270 (see Figure 4.31). In these cases detection is relatively straight-forward by implementing

3271 a power threshold on the STFT, since the amplitude of the main signal peak is distinct  
 3272 from the thermal noise spectrum. However, as the pitch angle of the electron is decreased  
 3273 below  $88.5^\circ$ , the modulation index of the signal increases causing the maximum amplitude  
 3274 of the frequency spectrum to be comparable to typical noise fluctuations. At this point,  
 3275 the power threshold trigger is no longer able to distinguish between signal and noise  
 3276 leading to a reduction in detection efficiency. The neutrino mass sensitivity of the FSCD  
 3277 is directly linked to the overall detection efficiency. And, because the distribution of  
 3278 electron pitch angles is effectively uniformly distributed across the range of pitch angles  
 3279 that can be trapped, the overall detection efficiency is directly influenced by the range of  
 3280 pitch angles that have detectable signals. Therefore, utilizing a signal detection algorithm  
 3281 that can more effectively identify signals with pitch angles less than  $88.5^\circ$  will improve  
 3282 both detection efficiency and ultimately the neutrino mass sensitivity of the FSCD and  
 3283 other CRES experiments.

3284 Modeling the detection performance of alternative signal detection algorithms for  
 3285 the FSCD requires that we pose the signal detection problem in a consistent manner.  
 3286 The approach we take is to perform a binary hypothesis test on the frequency spectra  
 3287 generated by the STFT. Mathematically, this is expressed as,

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

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

3288 Where under hypothesis  $\mathcal{H}_0$ , the vector representing the frequency spectrum ( $y[n]$ ) is  
 3289 composed of pure white Gaussian noise (WGN) represented by  $\nu[n]$ , and under hypothesis  
 3290  $\mathcal{H}_1$  the frequency spectrum is composed of a CRES signal ( $x[n]$ ) with added WGN. The  
 3291 dominant source of noise in a FSCD-like experiment is expected to be thermal Nyquist-  
 3292 Johnson noise, which is well approximated by a WGN distribution. In order to decide  
 3293 between these two hypotheses we follow the standard Neyman-Pearson approach by  
 3294 performing a log-likelihood ratio test between the probability distributions of the signal  
 3295 classifier output under  $\mathcal{H}_1$  and  $\mathcal{H}_0$  [71]. The output of the log-likelihood ratio test is  
 3296 called the test statistic, which is used to assign the data as belonging to the noise ( $\mathcal{H}_0$ )  
 3297 or signal ( $\mathcal{H}_1$ ) classes by setting a decision threshold on the value of the test statistic.

3298 In practice, we select the decision threshold by finding the value of the test statistic  
 3299 that guarantees an acceptable rate of false positives and then attempt to maximize  
 3300 the signal detection probability under that fixed false positive rate. Because the signal  
 3301 classifier will be used to evaluate the spectra of  $O(10^2)$  beamforming positions every

3302 40.96  $\mu$ sec, we will require the signal classifiers to operate with decision thresholds that  
 3303 provide false positive rates significantly smaller than 1%. This reduces the burden placed  
 3304 on later stages of the CRES reconstruction chain to reject these false positives and  
 3305 decreases the overall likelihood of reconstructing a false event. Below, we calculate the  
 3306 probability distributions that allow us characterize how different detection algorithms  
 3307 will perform for CRES signals in an FSCD experiment.

### 3308 **4.4.3 Signal Detection Algorithms**

#### 3309 **4.4.3.1 Power Threshold**

3310 The power threshold detection algorithm uses the maximum amplitude of the frequency  
 3311 spectra as the detection test statistic. To model the performance of this approach,  
 3312 consider first the case where the signal is pure WGN. For a single bin in the frequency  
 3313 spectrum, the probability that the amplitude falls below a specific threshold value is  
 3314 given by the Rayleigh cumulative distribution function (CDF),

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

3315 where the complex amplitude of the frequency bin is  $x$ , and  $\tau$  is the WGN variance.  
 3316 Because the noise samples for each frequency bin are independent and identically dis-  
 3317 tributed (IID), the probability that every bin in the frequency spectrum falls below the  
 3318 threshold is the joint CDF formed by the product of each individual frequency bin CDF,

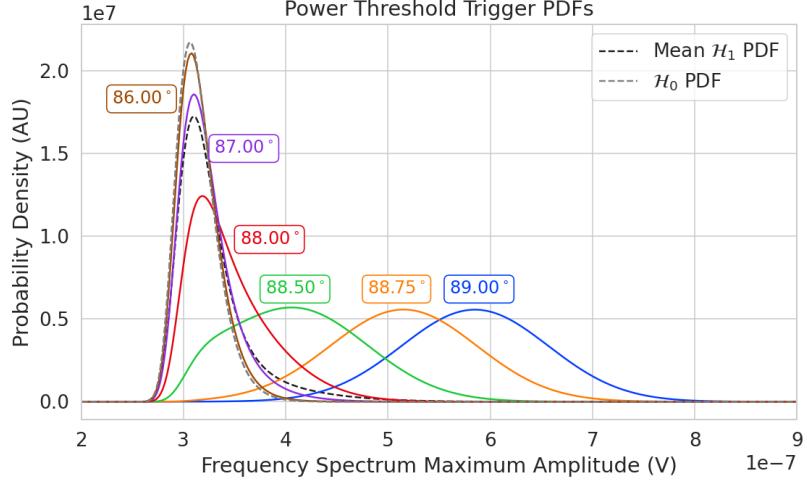
$$F_0(x; \tau, N_{\text{bin}}) = \text{Ray}(x; \tau)^{N_{\text{bin}}}. \quad (4.46)$$

3319 The PDF for the power threshold classifier can then be obtained by differentiating the  
 3320 CDF.

3321 The probability distribution for the power threshold classifier under  $\mathcal{H}_1$  is formed in  
 3322 a similar way, but the frequency bins that contain signal must be treated separately. For  
 3323 a frequency bin that contains both signal and noise we can describe the probability that  
 3324 the amplitude of the bin will fall below our threshold using the Rician CDF,

$$\text{Rice}(x; \tau, \nu) = 1 - Q_1 \left( \frac{|\nu|}{\sqrt{2\tau}}, \frac{|x|}{\sqrt{2\tau}} \right), \quad (4.47)$$

3325 where the parameter  $|\nu|$  defines the noise-free amplitude of the signal and  $Q_1$  is the  
 3326 Marcum Q-function. This time the CDF that describes the probability that the entire



**Figure 4.32.** 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.

3327 spectrum falls below the decision threshold is the product of both signal and noise CDFs,

$$F_1(x; \tau, \nu, N_{\text{bin}}, N_s) = \text{Ray}(x; \tau)^{N_{\text{bin}} - N_s} \prod_{k=0}^{N_s} \text{Rice}(x; \tau, \nu_k). \quad (4.48)$$

3328 The first half of Equation 4.48 is the contribution from the bins in the frequency spectrum  
 3329 that contain only noise, and the second half is the product of the Rician CDFs for the  
 3330 frequency bins that contain signal peaks with a noise-free amplitude of  $|\nu_k|$ . In Figure  
 3331 4.32 we show plots of example PDFs under  $\mathcal{H}_1$  and  $\mathcal{H}_0$ .

3332 **4.4.3.2 Matched Filtering**

3333 The shape of a CRES signal is completely determined by the initial conditions of the  
 3334 electron as it is emitted from beta-decay, which implies that it is possible to apply  
 3335 matched filtering as a signal detection algorithm. With a matched filter one uses the  
 3336 shape of the known signal, which is called a template, to filter the incoming data by  
 3337 computing the convolution between the signal and the data [71]. For cases where the  
 3338 signal is buried in WGN, the matched filter is the optimal detector in that it achieves  
 3339 the maximum probability of a true detection for a fixed false positive rate. Since CRES  
 3340 signals have an unknown shape but are deterministic, we can apply a matched filter by

3341 using simulations to generate a large number of signal templates called a template bank,  
 3342 which spans the parameter space of possible signals. Then at detection time, we use the  
 3343 template bank to identify signals by performing the matched filter convolution for each  
 3344 template in an exhaustive search.

3345 As we saw from the frequency spectra in Figure 4.31, CRES signals are highly periodic  
 3346 in nature. In such cases, it is advantageous to utilize the convolution theorem to replace  
 3347 the matched filter convolution with an inner product in the frequency-domain. With the  
 3348 convolution theorem, the matched filter test statistic that describes the detection of a  
 3349 signal buried in WGN using a matched filter template bank is given by

$$\mathcal{T} = \max_{\mathbf{h}} \left| \sum_{n=0}^{N_{\text{bin}}} h^\dagger[n] y[n] \right|, \quad (4.49)$$

3350 where  $h^\dagger[n]$  is the complex conjugate of the signal template. For the case when our  
 3351 template bank consists of only a single template it is possible to derive an exact analytical  
 3352 form for the PDF describing the matched filter test statistic. First, we derive PDF under  
 3353 the signal hypothesis, where the equation describing the matched filter test statistic, also  
 3354 known as the matched filter score, becomes

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

3355 Each noisy frequency bin represented by  $y[n]$  is the sum between value of the signal  
 3356 at that bin and complex WGN, which means that  $y[n]$  is itself Gaussian distributed.  
 3357 Therefore, the value of the inner product between the template and the data is also a  
 3358 complex Gaussian variable; and, since the matched filter score is the magnitude of this  
 3359 inner product, it must follow a Rician distribution.

3360 We can derive the equation for the Rician PDF by expressing the matched filter  
 3361 template  $\mathbf{h}$  in terms of the corresponding simulated signal, which we write as  $\mathbf{x}_h$  to  
 3362 distinguish from the signal in the data. Using the standard normalization and assuming  
 3363 uncorrelated WGN, the matched filter templates can be written as

$$\mathbf{h} = \frac{\mathbf{x}_h}{\sqrt{\tau |\mathbf{x}_h|^2}} \quad (4.51)$$

3364 where  $\tau$  is the noise variance. Inserting this into Equation 4.49 and expressing the data

3365 as a sum between a signal and a WGN vector yields,

$$\mathcal{T} = \frac{1}{\sqrt{\tau|\mathbf{x}_h|^2}} \left| \sum_{n=1}^{N_{\text{bin}}} x_h[n] (x[n] + \nu[n]) \right|. \quad (4.52)$$

3366 Next, we transform the expression by isolating the randomly distributed components  
 3367 giving

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

3368 The first term of 4.53 can be simplified by using the Cauchy-Schawrz inequality to express  
 3369 the magnitude of the inner product in terms of the magnitudes of the signal and template  
 3370 as well as an orthogonality constant which we call "match" ( $\Gamma$ ). Using this we obtain,

$$\mathcal{T} = |\mathbf{h}| |\mathbf{x}| \Gamma + \frac{1}{\sqrt{\tau|\mathbf{x}_h|^2}} \left| \sum_{n=1}^{N_{\text{bin}}} x_h[n] \nu[n] \right|. \quad (4.54)$$

3371 The second term is a sum of Gaussian distributed variables, which we should expect also  
 3372 follows a Gaussian distribution. Each of the samples  $\nu[n]$  is described by

$$\nu[n] \sim \mathcal{N}(0, \tau), \quad (4.55)$$

3373 where  $\mathcal{N}(0, \tau)$  is a complex Gaussian distribution with zero mean and variance  $\tau$ . There-  
 3374 fore,

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

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

$$|\mathbf{h}| |\mathbf{x}| \Gamma + \sum_{n=1}^{N_{\text{bin}}} \frac{x_h[n]}{\sqrt{\tau|\mathbf{x}_h|^2}} \nu[n] \sim \mathcal{N}(|\mathbf{h}| |\mathbf{x}| \Gamma, 1). \quad (4.58)$$

3375 We see that  $\mathcal{T}$  is magnitude of a complex variable with mean  $|\mathbf{h}| |\mathbf{x}| \Gamma$  and variance one. In  
 3376 order to simply the expression a bit further, we define the quantity  $\mathcal{T}_{\text{ideal}} = |\mathbf{h}| |\mathbf{x}| \Gamma$ , which  
 3377 we call the ideal matched filter score, because it represents the value of the matched  
 3378 filter inner product that we would expect if no noise was present in the signal. We can  
 3379 write the matched filter test statistic as the magnitude of a two-dimensional vector in

3380 the complex plane

$$\mathcal{T} = |(\mathcal{T}_{\text{ideal}} + n_r, n_i)|, \quad (4.59)$$

3381 where  $n_r$  and  $n_i$  are the real and imaginary components of the noise each with variance  
3382  $1/2$ , which is modeled by a Rician distribution with shape factor  $\mathcal{T}_{\text{ideal}}$ . Therefore, the  
3383 probability distribution of the matched filter test statistic is given by,

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

3384 where  $I_0$  is the zeroth-order modified Bessel function.

3385 The shape of the matched filter score distribution is controlled by the parameter  
3386  $\mathcal{T}_{\text{ideal}}$ , which is effectively the value of the matched filter score if the data contained no  
3387 noise. Without noise, the data vector reduces to the signal,  $\mathbf{x}$ , in which case Equation  
3388 4.50 becomes the magnitude of an inner product between two vectors. We can write  
3389 the magnitude of an inner product in terms of the lengths of the individual vectors and  
3390 a constant that describes the degree of orthogonality between them. Applying this to  
3391 Equation 4.50, we obtain

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

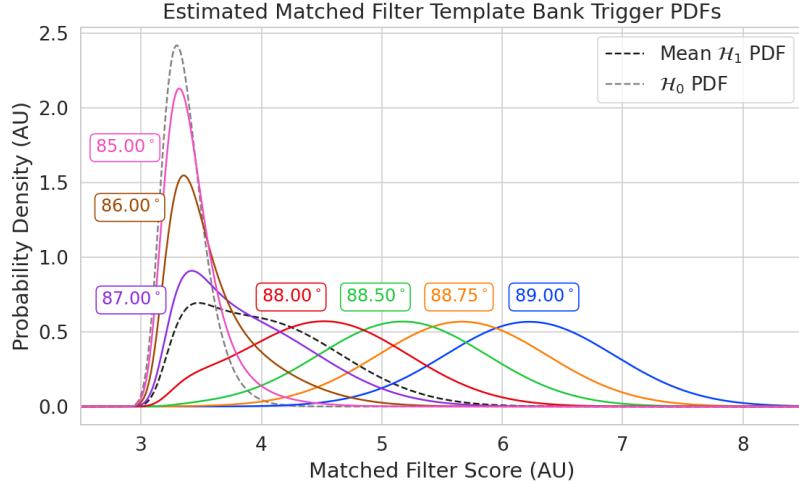
3392 where  $\Gamma$  describes the orthogonality between  $\mathbf{h}$  and  $\mathbf{x}$ . From the point of view of matched  
3393 filtering, we can interpret  $\Gamma$  as describing how well the template matches the underlying  
3394 signal in the data.

3395 The matched filter score PDF under the noise hypothesis can be readily obtained  
3396 from Equation 4.60 by setting the value of  $\mathcal{T}_{\text{ideal}}$  to zero, since the data contains no signal  
3397 in the noise case. Doing this, we obtain the Rayleigh distribution that describes the  
3398 matched filter score under  $\mathcal{H}_0$ ,

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

3399 Equations 4.60 and 4.62 describe the behavior of the matched filter test statistic  
3400 under  $\mathcal{H}_0$  and  $\mathcal{H}_1$  for a single template. However, defining a PDF that describes the  
3401 matched filter test statistic in the case of multiple templates is in general a mathematically  
3402 intractable problem, since there is no guarantee of orthogonality between matched filter  
3403 templates. This leads to correlations between the matched filter scores of different  
3404 templates because only one sample of noise is used to compute the matched filter scores  
3405 of the template bank. In order to proceed, we need to make the simplifying assumption  
3406 that we can treat the matched filter scores as IID variables, which allows to ignore

3407 correlations between templates. The overall effect of this will be an underestimate of the  
3408 performance of the matched filter, since we are under counting the number of templates  
3409 that could contribute a detectable score.



**Figure 4.33.** Plots of the estimated PDFs for the matched filter template bank test statistic for CRES signals with various pitch angles as well as the estimated PDF for the noise only signal case. We assume an estimated number of templates of  $10^5$  and perfect match between signal and template i.e.  $\Gamma_{\text{best}} = 1$ . The mean PDF includes signals ranging from  $85.5 - 88.5^\circ$  in pitch angle. There is a much larger distinction between the signal PDFs at small pitch angle compared to the power threshold indicating a higher detection efficiency for these signals.

3410 For  $\mathcal{H}_0$  we model the probability that the matched filter score falls below our threshold  
3411 using the CDF obtained by integrating Equation 4.62. Because we are assuming that  
3412 the matched filter scores using different templates are independent, the probability that  
3413 the matched filter score for all templates falls below a threshold value is the joint CDF  
3414 formed by multiplying the CDF for each template. Under  $\mathcal{H}_0$  this is

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

3415 where  $x$  is the matched filter score threshold and  $N_t$  is the number of templates. We  
3416 should expect that the distribution describing the matched filter template bank maximum  
3417 score depends on  $N_t$ , because with more templates there is a greater chance of a random  
3418 match between the template and data.

3419 For  $\mathcal{H}_1$ , we start by denoting the CDF of the best matching template as  $F_{\text{best}}(x; \mathcal{T}_{\text{best}})$ ,  
3420 and treat the matched filter scores for all other templates as negligible ( $\mathcal{T}_{\text{ideal}} \approx 0$ ). Then  
3421 we form the joint CDF by combining the distributions for all templates used during

3422 detection. Since we are exhaustively checking the matched filter scores, the number of  
 3423 templates checked will be a randomly distributed variable that ranges from zero to the  
 3424 total number of available templates. If we assume that signals are uniformly distributed  
 3425 across the parameter space spanned by the template bank then on average we check  
 3426  $(N_t - 1)/2 \approx N_t/2$  templates for each inference. Therefore, the estimated CDF under  $\mathcal{H}_1$   
 3427 is

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

3428 In Figure 4.33 we show plots of the estimated matched filter template bank classifier  
 3429 PDFs under both  $\mathcal{H}_0$  and  $\mathcal{H}_1$ .

#### 3430 4.4.3.3 Machine Learning

3431 In this paper we focus on Convolutional Neural Networks (CNN) as an example of  
 3432 a machine learning based signal classifier. CNNs are constructed using a series of  
 3433 convolutional layers, each composed of a set of filters that are convolved with the input  
 3434 data. The individual convolutional filters can be viewed as matched filter templates that  
 3435 are learned from a set of simulated data rather than being directly generated. This opens  
 3436 the possibility of finding a more efficient representation of the matched filter templates  
 3437 during the training process that can potentially reduce computational cost at inference  
 3438 time while still offering good classification performance.

3439 The machine learning approach is distinct from both the power threshold and matched  
 3440 filtering in that we do not attempt to manually engineer a test statistic that is computed  
 3441 from the data for classification. Instead, we attempt calculate the test statistic by  
 3442 constructing a differentiable function that maps the complex frequency series generated  
 3443 by the STFT to a binary classification as either signal or noise. The test statistic for the  
 3444 machine learning classifier can be expressed as

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

3445 where  $\mathbf{y}$  is the noisy data vector and  $G(\mathbf{y}; \boldsymbol{\Omega})$  is the machine learning model parameterized  
 3446 by the weights  $\boldsymbol{\Omega}$ . By using supervised learning on a labeled set of training signals, we  
 3447 can modify the function parameters to learn the mapping from the data to the likelihood  
 3448 of  $\mathbf{y}$  belonging to either  $\mathcal{H}_1$  or  $\mathcal{H}_0$ .

3449 The CNN architecture used for this work is summarized by Table 4.1. No strategic  
 3450 hyper-parameter optimization approach was implemented beyond the manual testing  
 3451 of different CNN architecture variations, so this particular model is best viewed as a

**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.

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

<sup>3452</sup> proof-of-concept rather than a rigorously optimized design. Numerous model variations  
<sup>3453</sup> were tested, some with significantly more layers and convolutions filters per layer, as  
<sup>3454</sup> well as others that were even smaller than the architecture in Table 4.1. Ultimately, the  
<sup>3455</sup> model architecture choice was driven by the motivation to find the minimal model whose  
<sup>3456</sup> classification performance was still comparable to the larger CNN’s tested, because of  
<sup>3457</sup> the importance of minimizing computational cost in real-time applications. It is possible  
<sup>3458</sup> that more sophisticated machine learning models could improve upon the classification  
<sup>3459</sup> results achieved here, but we leave this investigation for future work.

#### <sup>3460</sup> 4.4.4 Methods

##### <sup>3461</sup> 4.4.4.1 Data Generation

<sup>3462</sup> To test the triggering performance of the classifiers, simulated CRES signals were  
<sup>3463</sup> generated using the Locust simulations package [62, 80] developed by the Project 8  
<sup>3464</sup> collaboration. Locust uses the separately developed Kassiopeia package to calculate the  
<sup>3465</sup> magnetic fields produced by a user defined set of current carrying coils along with any  
<sup>3466</sup> specified background magnetic fields, resulting in a magnetic trap. Next, Kassiopeia  
<sup>3467</sup> calculates the trajectory of an electron in this magnetic field starting from a set of user  
<sup>3468</sup> specified initial conditions. The Locust software then uses the electron trajectories from  
<sup>3469</sup> Kassiopeia to calculate the resulting electromagnetic fields using the Liénard-Wiechert  
<sup>3470</sup> equations, and determine the voltages generated in the antenna array with the antenna

3471 transfer function. Locust then simulates the down-conversion, filtering, and digitization  
3472 steps resulting in the simulated CRES signals for an electron.

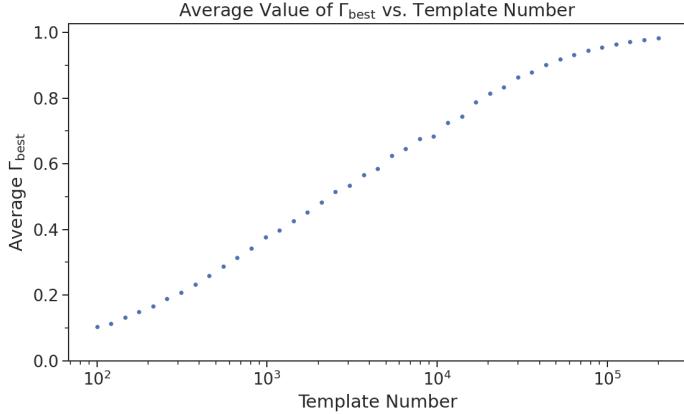
3473 The shape of the received CRES signal is determined by the initial kinematic param-  
3474 eters, including the starting position of the electron, the starting kinetic energy of the  
3475 electron, and the pitch angle. For the studies performed here we constrain ourselves to a  
3476 single initial electron position located at  $(x, y, z) = (5, 0, 0)$  mm, and using this starting  
3477 position we generate two datasets by varying the initial kinetic energy and the starting  
3478 pitch angle. The first dataset consists of a two-dimensional square grid of kinetic energy  
3479 and pitch angle spanning an energy range from 18575-18580 eV with a spacing of 0.1 eV,  
3480 and pitch angles from 85.5-88.5° with a spacing of 0.001°, resulting in 153051 signals with  
3481 a unique energy-pitch angle combination. This dataset is intended to represent a matched  
3482 filter template bank. The second dataset was generated by randomly sampling uniform  
3483 probability distributions covering the same parameter space to produce approximately  
3484 50000 signals randomly parameterized in energy and pitch angle. This dataset provides  
3485 the training and test data for the machine learning approach, and acts as a representative  
3486 sample of signals to evaluate the performance of the matched filter template bank.

3487 Each signal was simulated for a duration of 40.96  $\mu$ s, which is equivalent to 8192  
3488 samples at the FSCD digitization rate, and begins at time  $t = 0$  s for all simulations.  
3489 This duration represents a single frequency spectrum generated by the STFT. The output  
3490 of the Locust simulation is a matrix of array snapshots with size given by the number of  
3491 channels times the event length ( $N_{\text{ch}} \times N_{\text{sample}}$ ), which we pre-process using the digital  
3492 beamforming summation and STFT described in Section 4.4.2.2. The  $\nabla B$ -drift correction  
3493 uses the exact value of  $\omega_{\nabla B}$ , obtained from the Kassiopeia simulation of that electron.  
3494 In practice, an average value for  $\omega_{\nabla B}$  could be used, because there is limited variation in  
3495 drift frequency across this parameter space.

#### 3496 4.4.4.2 Template Number and Match Estimation

3497 The estimated PDF for the matched filter template bank depends on the score of the  
3498 best matching template or equivalently the match of the best template ( $\Gamma_{\text{best}}$ ) as well  
3499 as the number of templates. One expects that with a higher number of templates the  
3500 average value of  $\Gamma_{\text{best}}$  will increase, however, there is a point of diminishing returns at  
3501 which more templates will not significantly increase match, but will still increase the  
3502 likelihood of false positives. Therefore, it is desirable to use the minimum number of  
3503 templates that provide an acceptable mean value of  $\Gamma_{\text{best}}$ .

3504 To quantify the relationship between match and template number, we calculated



**Figure 4.34.** 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.

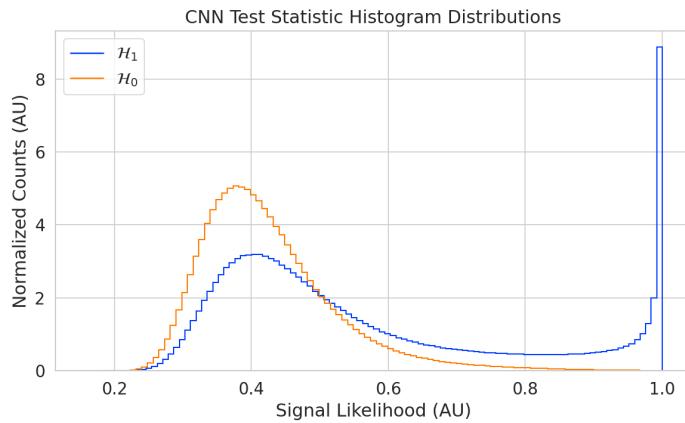
the mean match of the random dataset to a selection of templates obtained from the regularly spaced dataset. The results are shown in Figure 4.34, where we find that the average value of  $\Gamma_{\text{best}}$  is an exponential function of the number of templates. From this plot we select the desired value of mean match at which we would like to evaluate the matched filter PDF and can infer the required number of templates.

#### 4.4.4.3 CNN Training and Data Augmentation

To prepare the data for training the model, we split the random dataset in half to create distinct training and test datasets. Additionally, a randomly selected 20% of the training data is isolated for use as a validation set during the training loop. The size of the training, validation, and test datasets are then tripled by appending two additional copies of the data to increase the sample size of the dataset after data augmentation. The data is loaded with no noise, which is added to each data batch during the training phase by generating a new noise sample from a complex WGN distribution. In order to ensure an even split between signal and noise data we append to the noise-free signals an equal number of empty signals composed of all zeros. Therefore, as the data is randomly shuffled during training, on average an equal number of empty signals will be included with the training signals. After adding the sample of WGN to the data batch, the empty signals represent the noise-only data that the model must distinguish from signal data.

As the training signals are loaded we apply a unique random phase shift as the first form of data augmentation. Since the data is generated using the same initial axial position and cyclotron orbit phase, the randomization is an attempt to prevent

3526 overtraining on these features. During each training epoch the data is randomly shuffled  
 3527 and split into batches of 2500 signals. Each batch of signals is then circularly shifted  
 3528 by a random number of frequency bins to simulate a kinetic energy shift from  $-75$  to  
 3529  $20$  eV to simulate a training dataset with a larger energy range. Next, a sample of  
 3530 complex WGN, consistent with the expected  $10$  K Nyquist-Johnson noise expected for  
 3531 the FSCD, is generated and added to the signal, which prevents overtraining on noise  
 3532 features. As a final step, the data is renormalized by the standard deviation of the noise  
 3533 so that the range of values in the data is close to  $[-1, 1]$ , which helps ensure well-behaved  
 3534 back-propagation.



**Figure 4.35.** 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  $[0, 1]$ .

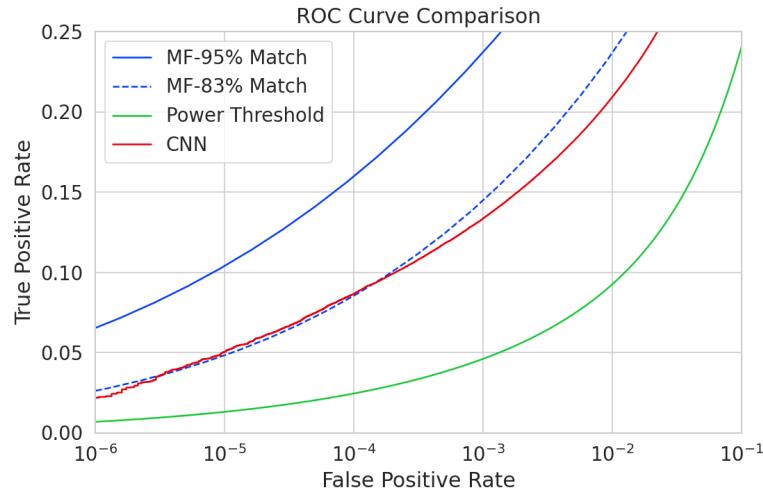
3535 The Binary Cross-entropy loss function is used to compute the loss for each batch of  
 3536 data and the model weights are updated using the ADAM optimizer with a learning rate  
 3537 of  $5 \times 10^{-3}$ . After each training epoch, the loss and classification accuracy of the validation  
 3538 dataset are computed to monitor for overtraining. It was noticed that the relatively high  
 3539 noise power and the fact that a new sample of noise was used for each batch together  
 3540 provided a strong form of regularization, since no evidence of over-training was observed  
 3541 even after several thousand epochs. Typically, the loss and classification accuracy of  
 3542 the model converged after a few hundred training epochs, but the training loop was  
 3543 extended to 3000 epochs to attempt to achieve the best possible performance. The  
 3544 training procedure generally took about 24 hrs using a single NVIDIA V100 GPU [82].

3545 After training the model, we use it to classify the test dataset and generate  
 3546 histograms of the model outputs for both classes of data. The data augmentation

procedure for the evaluation of the test data mirrors the training procedure without the validation split. Since a random circular shift and a new sample of WGN is added to each batch, the testing evaluation loop is run for 100 epochs to get a representative sample of noise and circular shifts. The model outputs for each batch are passed through a softmax activation and then combined into histograms, which we show in Figure 4.35.

## 3552 4.4.5 Results and Discussion

### 3553 4.4.5.1 Trigger Classification Performance



**Figure 4.36.** ROC curves describing the detection efficiency or true positive rates for the three signal classification algorithms examined in this paper.

Using the matched filter and power threshold CDFs, along with the classification results from the CNN, we compare detection performance by computing receiver operating characteristic (ROC) curves. Specifically, we compare the detection performance averaged over the full signal parameter space in order to get a measure of the overall detection efficiency achieved by each algorithm. For the power threshold and matched filter algorithms, we obtain the mean ROC curve by taking the average over all signals in the regularly spaced dataset. In the case of the matched filter, we examine two cases using different numbers of templates, which have different values of mean match. The ROC curve describing the CNN is obtained by forming a histogram of the network outputs for each class of signal and from this computing the estimated CDFs and ROC curve. In Figure 4.36, we show the ROC curves obtained for each of the detection algorithms, visualized in terms of true positive rate and false positive rate.

3566        The true positive rate of a signal classifier is equivalent to its detection efficiency, and  
3567        we see that for the population of signals with pitch angles  $< 88.5^\circ$  the power threshold  
3568        has a consistently lower detection efficiency than the CNN and the matched filter. This  
3569        result could have been predicted from the visualization of signal spectra in Figure 4.31,  
3570        where we see that there is no way to distinguish between a noise peak and a signal peak  
3571        with high confidence at small pitch angles. The CNN offers a significant and consistent  
3572        increase in detection efficiency over the power threshold approach, with the relative  
3573        improvement in detection efficiency increasing as the false positive rate decreases. If  
3574        we compare the CNN to the matched filter, we see that the performance of the tested  
3575        network is roughly equivalent to a matched filter detector with an average match of  
3576        about 83%, which uses approximately 20000 matched filter templates. The overall best  
3577        detection efficiency is achieved by the matched filter classifier if a large enough template  
3578        bank is used. We show in the plot the ROC curve for a matched filter template bank  
3579        with 95% average match, which is achieved with approximately 100000 templates. Since  
3580        the matched filter is known to be statistically optimal for detecting a known signal in  
3581        WGN, it is somewhat expected that this algorithm has the highest detection efficiency.

3582        A potentially impactful difference between the matched filter and CNN algorithms is  
3583        that the CNN relies upon convolutions as its fundamental calculation mechanism, whereas  
3584        our implementation of a matched filter utilizes an inner product. Since convolution is  
3585        a translation invariant operation, the detection performance of CNN can be extended  
3586        to a wider range of CRES event kinetic energies with less cost than the matched filter,  
3587        a feature that we exploited during the CNN training by including circular translations  
3588        of the CRES frequency spectra in the training loop. Increasing the range of kinetic  
3589        energies detectable by a matched filter requires a proportional increase in the number of  
3590        templates, which directly translates into increased computational and hardware costs.  
3591        From a practical perspective, the detection algorithm is always limited by the available  
3592        computational hardware, so estimating the relative costs is a key factor in determining  
3593        their feasibility. Below we perform a more detailed analysis of the relative costs of each  
3594        of the detection algorithms.

#### 3595        4.4.5.2 Computational Cost and Hardware Requirements

3596        In the process of investigating triggering approaches for an antenna array CRES experi-  
3597        ment, we have uncovered a strong tension between detection efficiency and computational  
3598        resources. To relate the computational cost estimates to actual costs, we compare the  
3599        theoretical amount of computer hardware required to implement the signal classifiers

3600 for real-time detection in an FSCD experiment. To do this we shall utilize order of  
3601 magnitude estimates of the theoretical peak performance values for currently available  
3602 Graphics Processing Units (GPUs) as a metric. This approach will underestimate the  
3603 amount of required hardware, since it is unlikely that any CRES detection algorithm  
3604 could reach the theoretical peak performance of the hardware.

3605 Of the three detection algorithms tested, the power threshold classifier is the least  
3606 expensive. It requires that we check whether the amplitude of each frequency bin in  
3607 the STFT is below or above our decision threshold. The STFT combined with digital  
3608 beamforming produces  $N_{\text{bin}}N_b$  frequency bins that must be checked every  $N_{\text{bin}}/f_s$  seconds.  
3609 This requires approximately  $O(10^{10})$  FLOPS to check in real-time. Current generations of  
3610 GPUs have peak theoretical performances in the range of  $O(10^{13}) - O(10^{14})$  FLOPS [83],  
3611 dependent on the required floating-point precision of the computation. Therefore, the  
3612 entire computational needs of a real-time triggering system using a power threshold  
3613 classifier, including digital beamforming and generation of the STFT, could be met by a  
3614 single high-end GPU or a small number of less powerful GPUs. Since triggering is only  
3615 one step of the full real-time signal reconstruction approach, limiting the computational  
3616 cost of this stage is ideal. However, we have seen that the power threshold classifier does  
3617 not provided sufficient detection efficiency across the entire range of possible signals,  
3618 which is the primary motivation for exploring more complicated triggering solutions.

3619 As discussed, the computational cost of the matched filter approach requires counting  
3620 the number of templates that must be checked for each frequency spectra produced by the  
3621 STFT. Computing the matched filter scores requires  $O(N_bN_tN_{\text{bin}})$  operations, since for  
3622 each of the  $N_b$  beamforming positions we must multiply  $N_t$  templates with a data vector  
3623 that has length  $N_{\text{bin}}$ . The time within which we must perform this calculation is equal  
3624 to  $N_{\text{bin}}/f_s$  to keep up with the data generation rate. To cover the 5 eV kinetic energy  
3625 range spanned by the template bank, we saw that  $10^4$  to  $10^5$  templates are required in  
3626 order to match or exceed the detection efficiency of the CNN. If the number of templates  
3627 scales linearly with then kinetic energy range of interest as expected, then we would  
3628 require  $10^5$  to  $10^6$  matched filter templates with this more realistic range of energies.  
3629 Considering this, the estimated computational cost of the matched filter is between  
3630  $O(10^{15})$  to  $O(10^{16})$  FLOPS, which is  $O(10^2)$  to  $O(10^3)$  high-end GPUs.

3631 Lastly, we have the CNN classifier. To estimate the computational cost we simply  
3632 sum the number of convolutions and matrix multiplications specified by the network  
3633 architecture shown in Table 4.1. Each convolutional layer consists of  $N_{\text{in}}N_{\text{out}}N_{\text{kernel}}L_{\text{input}}$   
3634 floating-point operations, where  $N_{\text{in}}$  is the number of input channels,  $N_{\text{out}}$  is the number

3635 of output channels,  $N_{\text{kernel}}$  is the size of the convolutional kernel, and  $L_{\text{input}}$  is the length  
3636 of the input vector, and the fully connected layers each contribute  $N_{\text{in}}N_{\text{out}}$  operations.  
3637 Summing all the neural network layers we estimate that the CNN would require  $O(10^6)$   
3638 floating point operations for each frequency spectra; therefore, the total computation  
3639 cost of the CNN trigger is this cost times the number of beamforming positions per the  
3640 data acquisition time, which is  $O(10^{13})$  FLOPS or  $O(10^0)$  GPUs.

3641 Compared with the matched filter approach the CNN requires  $O(100)$  to  $O(1000)$   
3642 fewer GPUs to implement, dependent on the exact number of templates used in the  
3643 template bank. The 100 eV kinetic energy range is motivated by the application of these  
3644 detection algorithms to an FSCD-like neutrino mass measurement experiment. However,  
3645 if a significantly larger range of kinetic energies is required, a CNN may be the preferred  
3646 detection approach despite the lower average detection efficiency due to computational  
3647 cost considerations. The low estimated computational cost of the CNN is directly related  
3648 to the small network size.

3649 Additional experiments with larger CNNs, generated by increasing the depth and  
3650 width of the neural network, and we observed that these changes provided minimal  
3651 ( $\lesssim 1\%$ ) improvement in the classification accuracy of the model. A potential reason  
3652 for this could be the sparse nature of the signals in the frequency domain and the low  
3653 SNR which makes for a challenging dataset to learn from. Future work could investigate  
3654 modifications to the neural network architecture such as sparse convolutions, which may  
3655 improve the classification accuracy of the model or further reduce the computational  
3656 costs of this approach. Alternatively, more complicated CNN architectures such as a  
3657 ResNet [84] or VGG model [85] may provide improved classification performance over a  
3658 basic CNN. An additional promising area of investigation are recurrent neural networks,  
3659 which may be able to exploit the time-ordered features of the STFT for more accurate  
3660 signal detection if the electron signals last for multiple Fourier transform windows.

3661 Our estimate of the computational cost of the matched filter is somewhat naive if  
3662 we notice that the majority of the values that make up a CRES frequency spectra are  
3663 zero (see Figure 4.31). Therefore, the majority of operations in the matched filter inner  
3664 product are unnecessary, and we could instead evaluate the matched filter inner product  
3665 using only the  $\lesssim 10$  frequency peaks that make up CRES signal. This optimization  
3666 reduces the number of operations required to check each template by a factor of  $O(100)$   
3667 to  $O(1000)$ , which brings the estimated computational cost of the matched filter in  
3668 line with the CNN. Although this level of sparsity results in a multiplication with very  
3669 low arithmetic complexity, the resulting sparse matched filter algorithm is still likely

3670 to be constrained by memory access speed rather than compute speed. Ultimately, the  
3671 comparison of the relative computational and hardware costs between the matched filter  
3672 and CNN will depend on the efficiency of the software implementation and hardware  
3673 support for neural network and sparse matrix calculations.

#### 3674 **4.4.6 Conclusion**

3675 Increasing the detection efficiency and overall event rate of the CRES technique represents  
3676 a key developmental path towards new scientific results and broader applications of the  
3677 CRES technique. It is what motivates both the antenna array detection approach and  
3678 the development of real-time signal reconstruction algorithms. We have demonstrated  
3679 that significant gains in the detection efficiency of the CRES technique are achievable  
3680 by utilizing triggering algorithms that account for the specific shape of CRES signals in  
3681 the detector. These algorithms emphasize the need for accurate and fast methods for  
3682 CRES simulation, since they directly contribute to the success of matched filter methods  
3683 by providing a way to generate expected signal templates and also serve as a source of  
3684 training data for machine learning approaches.

3685 The improvements in detection efficiency offered by these alternative approaches to  
3686 triggering are crucial to the success of efforts to develop scalable technologies for CRES  
3687 measurement, since they provide a significant increase in the detectable parameter space  
3688 of CRES events, which allows for a better utilization of the larger detection volume.  
3689 While we have focused on the real-time detection of CRES signals from antenna arrays,  
3690 these same signal classifiers could be used in CRES experiments utilizing a different  
3691 detector technologies, since the same principles of signal detection will apply. For example,  
3692 previous CRES measurements by the Project 8 collaboration that utilized a waveguide  
3693 gas cell, could have improved their detection efficiency by employing a matched filter  
3694 or neural network classifier to identify trapped electrons with pitch angles that are too  
3695 small to be detected by the power threshold approach. Furthermore, alternative CRES  
3696 detector technologies such as resonant cavities [41] could also see similar improvements  
3697 in detection efficiency, which is of crucial importance to future efforts by the Project 8  
3698 collaboration to utilize CRES to measure the neutrino mass.

# **Chapter 5**

## **Antenna and Antenna Measurement System Development for the Project 8 Experiment**

### **5.1 Introduction**

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

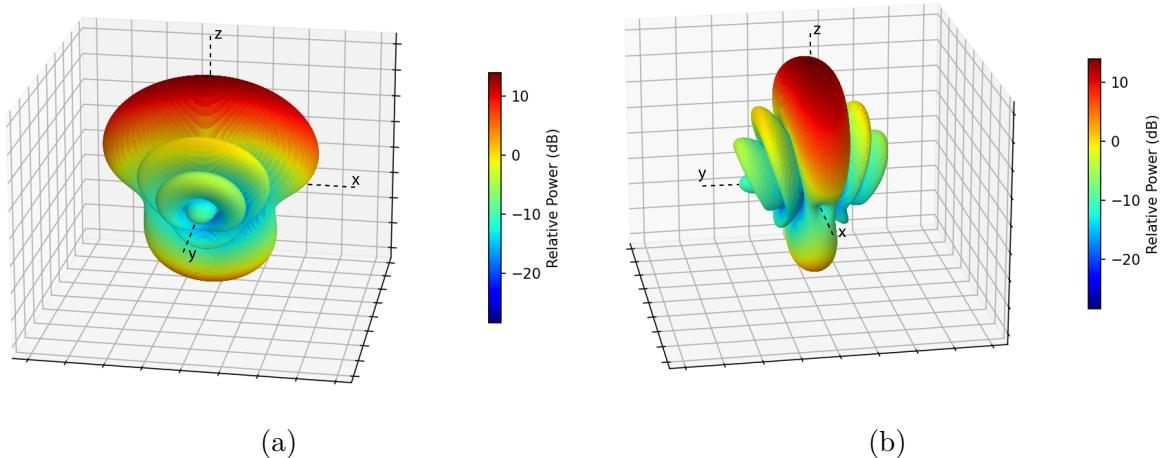
In this chapter we summarize the development of an antenna measurement system at Penn State to implement and test the techniques of antenna array CRES on the bench-top, in order to support the efforts of the Project 8 collaboration. In Section 5.2 we provide an introduction to some fundamental parameters and concepts related to antenna measurements as well as an overview of the Penn State antenna measurement system hardware. In Section 5.3 we include the manuscript of a paper [81] which details the design and characterization of a specialized antenna developed to mimic the electric fields emitted by an electron in a CRES experiment. This antenna, called the Synthetic Cyclotron Antenna (SYNCA), is intended as a calibration tool for antenna arrays developed for CRES measurements. Lastly, in Section 5.5 we summarize a set of prototype FSCD antenna array measurements with the SYNCA [44], which we use to validate the simulated performance of the antenna array and estimate systematic errors associated with the antenna array.

## 3724 5.2 Antenna Measurements for CRES experiments

### 3725 5.2.1 Antenna Parameters

3726 Antenna characterization measurements are intended to validate simulations of the  
3727 antenna array performance, which ultimately informs the neutrino mass sensitivity of  
3728 the experiment. In this section, I shall summarize a few fundamental concepts relating  
3729 to antennas and antenna measurement, before introducing how Project 8 uses antenna  
3730 measurement for the development of antenna array CRES.

#### 3731 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.

3732 Antennas are conductive structures designed to carry alternating electric currents  
3733 in order to transmit energy in the form of electro-magnetic (EM) waves [65]. Perhaps  
3734 the most fundamental way to characterize an antenna, is to map out the radiated power  
3735 density as a function of position, which is called the radiation pattern (see Figure 5.1).  
3736 We find the radiation power density by calculating the time-averaged Poynting vector for  
3737 all positions surrounding the antenna, which in equation form is

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

3738 where  $\mathbf{E}(x, y, z, t)$  and  $\mathbf{H}(x, y, z, t)$  are the time-dependent electric and magnetic fields  
 3739 produced by the antenna [49]. The radiation power density has units of  $\text{W/m}^2$  and is  
 3740 more typically called the energy flux density in physics applications, since it is a measure  
 3741 of the amount of energy passing through a unit area over time.

3742 Because the radiation power density is a measure of power per unit area, its value  
 3743 in a particular direction will depend on the distance from the antenna at which we are  
 3744 measuring. This is undesirable for practical applications A related quantity, which is  
 3745 distance independent, is the energy flux per unit solid angle or radiation intensity, which  
 3746 is computed directly from the radition power density by multiplying by the squared  
 3747 distance from the antenna. Specifically,

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

3748 where  $r$  is the distance from the antenna to the field measurement point. The radiation  
 3749 intensity is typically defined in regions where the Poynting vector consists only of a radial  
 3750 component where it is safe to treat as a scalar quantity.

### 3751 5.2.1.2 Directivity and Gain

3752 Since the radiation intensity is a measure of average power per unit solid angle, it is  
 3753 independent of distance and more useful as feature for antenna measurement. However,  
 3754 most antenna measurements are performed in terms of the directly related directivity  
 3755 and gain quantities. Directivity is defined as the ratio between the radiation intensity at  
 3756 particular point on the radiation pattern to the average radiation intensity computed  
 3757 over all solid angles [65]. The equation that relates the radiation intensity to directivity  
 3758 is

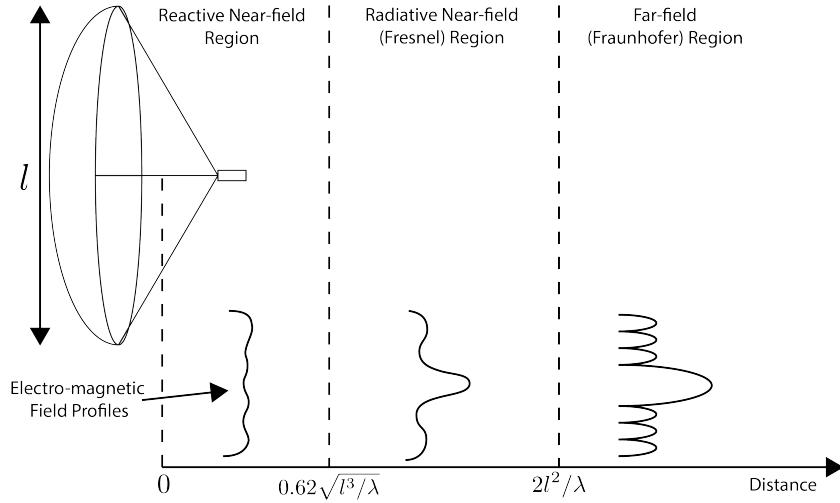
$$D = \frac{U}{U_0} = \frac{4\pi U}{P_{\text{rad}}}, \quad (5.3)$$

3759 where  $U_0$  is the average radiation intensity over all solid angles, which simply the total  
 3760 radiated power ( $P_{\text{rad}}$ ) divided by  $4\pi$ . Closely related to directivity is concept of gain,  
 3761 which accounts for energy losses that occur inside then antenna when attempting to  
 3762 transmit or receive a signal. The antenna gain is given by

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

3763 where  $P_{\text{in}}$  is the total power delivered to the antenna. Gain can be thought of as the ratio  
 3764 of the antenna's radiation intensity to that of a hypothetical isotropic, lossless radiator.

3765 The maximum values of gain and directivity exhibited by the main lobe of the antenna  
 3766 pattern as well as the ratio between the gain of the main lobe and any side-lobes are  
 3767 important figures of merit used to evaluate antenna designs.



**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.

### 3768 5.2.1.3 Far-field and Near-field

3769 Radiation patterns are only well-defined in regions where the shape of the radiation  
 3770 pattern is independent of distance. The region where this approximation is valid is called  
 3771 the "far-field", and in this region we can treat the EM fields from the antenna as spherical  
 3772 plane waves. A rule of thumb for antennas is that the far-field approximation is valid  
 3773 when the condition

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

3774 is met. In this expression,  $R$  is the distance from the antenna,  $l$  is the largest characteristic  
 3775 dimension of the antenna, and  $\lambda$  is the wavelength of the radiation (see Figure 5.2).

3776 The region very close to the antenna is called the reactive near-field, because in this  
 3777 region the reactive component of the EM field is dominant. Unlike radiative electric

3778 fields, the reactive electric and magnetic fields are out of phase from each other by  
 3779  $90^\circ$ , since they are the result of electrostatic and magnetostatic effects coming from the  
 3780 self-capacitance and self-inductance of the antenna. The reactive fields are unable to  
 3781 transfer energy a significant distance from the antenna and are thus completely negligible  
 3782 for most antenna applications. The limit of the reactive near-field for an electrically-large  
 3783 antenna is typically taken to be

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

3784 The unique application of antennas by Project 8 is somewhat limited by reactive near-  
 3785 field effects in the form of a maximum radial position for electrons inside the uniform  
 3786 cylindrical antenna array. If electrons are too close to the edge of the array than reactive  
 3787 near-field effects leads to a large reduction in the received power and consequently  
 3788 detection efficiency. This leads to a significant volume inside of the antenna array that  
 3789 is unsuitable for CRES lowering the volumetric efficiency of the antenna array CRES  
 3790 technique relative to a cavity experiment.

3791 In between the reactive near-field and the far-field is the radiative near-field region.  
 3792 In this region the fields are primarily radiative, however we are still too close to the  
 3793 antenna for the spherical plane wave approximation to apply. Therefore, interference  
 3794 effects between EM waves emitted from different points on the antenna occur causing the  
 3795 shape of the radiation pattern to change as a function of distance from the antenna. If we  
 3796 evaluate the far-field distance limit for the FSCD one finds an estimated far-field distance  
 3797 of 43 cm, which is a factor of four larger than the radius of the antenna array designed for  
 3798 the experiment. Consequently, we expect near-field effects to influence the performance  
 3799 of the antenna array highlighting the importance of calibration and characterization  
 3800 measurements.

### 3801 5.2.1.4 Polarization

3802 The polarization of an EM wave defines the spatial orientation of the electric field  
 3803 oscillations in the plane perpendicular to the direction of the propagation, and is defined  
 3804 in terms of orthogonal polarization components. In our application, one analyzes the  
 3805 properties of radiation propagating along the radial ( $\hat{r}$ ) direction away from the antenna,  
 3806 which implies that the electric fields can be described as a linear combination of orthogonal  
 3807 polarization components

$$\mathbf{E}_{\text{tot}} = E_x \hat{x} + E_y \hat{y} + E_z \hat{z}, \quad (5.7)$$

3808 in Cartesian coordinates, or

$$\mathbf{E}_{\text{tot}} = E_\theta \hat{\theta} + E_\phi \hat{\phi}, \quad (5.8)$$

3809 in spherical coordinates.

3810 In general, one defines partial radiation patterns, directivities, and gains so that the  
3811 performance of the antenna for the desired polarization can be analyzed. The radiation  
3812 pattern defined in terms of partial patterns is

$$U_{\text{tot}} = U_\phi + U_\theta, \quad (5.9)$$

3813 where  $U_\phi$  and  $U_\theta$  are the radiation intensities in a particular direction for the respective  
3814 polarization components. Similarly, a quantity such as gain can be written in terms of  
3815 partial gains,

$$G_{\text{tot}} = G_\phi + G_\theta = \frac{2\pi U_\phi}{P_{\text{in}}} + \frac{2\pi U_\theta}{P_{\text{in}}}. \quad (5.10)$$

3816 If we view an electron performing a circular orbit in the XY-plane from the side, that  
3817 is, along the X or Y axes, then we would observe the electron to be performing a linear  
3818 oscillation perpendicular to the viewing axis. From this intuitive picture, we can predict  
3819 that the primary polarization of electric fields from CRES events to be linearly polarized  
3820 in the  $\hat{\phi}$  direction when viewed with an antenna positioned in the XY-plane.

### 3821 5.2.1.5 Antenna Factor and Effective Aperture

3822 A useful way to characterize the performance of an antenna is to measure the electric  
3823 field magnitude required to produce a signal with an amplitude of one volt in the antenna  
3824 terminals. This ratio between the magnitude of the incoming electric field and the  
3825 magnitude of the signal produced by the antenna is called the antenna factor, which is  
3826 written as

$$A_F = \frac{|\mathbf{E}_{\text{in}}|}{V_{\text{ant}}}, \quad (5.11)$$

3827 where  $A_F$  is the antenna factor,  $E_{\text{in}}$  is the incoming electric field, and  $V_{\text{ant}}$  is the magnitude  
3828 of the voltage produced by the antenna.

3829 The antenna factor can be expressed in terms of the antenna's gain through a related  
3830 quantity called the effective aperture. The effective aperture defines for a given incident  
3831 radiation power density ( $\text{W/m}^2$ ) the power that is received by the antenna. Therefore,

3832 the effective aperture gives the equivalent area of the antenna,

$$A_{\text{eff}} = \frac{P_{\text{rec}}}{P_{\text{in}}} = \frac{\lambda^2}{4\pi} G, \quad (5.12)$$

3833 where the received power is  $P_r$  and the total incoming power is  $P_{\text{in}}$ .

3834 If we express the incident power in terms of the magnitude of the Poynting vector,  
3835 then

$$|\mathbf{S}_{\text{in}}| = |\mathbf{E}_{\text{in}}|^2 / \eta_0, \quad (5.13)$$

3836 where  $\eta_0$  is the impedance of free-space, which relates the magnitudes of the electric and  
3837 magnetic fields in a vacuum, and is defined by

$$\eta_0 = \frac{|\mathbf{E}|}{|\mathbf{H}|} = \sqrt{\frac{\epsilon_0}{\mu_0}}. \quad (5.14)$$

3838 The total received power by the antenna can therefore be expressed as

$$P_{\text{rec}} = |\mathbf{S}_{\text{in}}| A_{\text{eff}} = |\mathbf{S}_{\text{in}}| \frac{\lambda^2}{4\pi} G = \frac{|\mathbf{E}_{\text{in}}|^2 \lambda^2 G}{4\pi \eta_0}. \quad (5.15)$$

3839 To relate this to the antenna factor recall that we can relate the voltage produced by  
3840 the antenna to the received power with

$$P_{\text{rec}} = \frac{V_{\text{ant}}^2}{Z} = \frac{|\mathbf{E}_{\text{in}}|^2}{A_F^2 Z}, \quad (5.16)$$

3841 where  $Z$  is the system impedance. Setting Equations 5.15 and 5.16 equal to each other,  
3842 we obtain the following expression for antenna factor in terms of gain

$$A_F = \sqrt{\frac{4\pi\eta_0}{ZG\lambda^2}} = \frac{9.73}{\lambda\sqrt{G}}. \quad (5.17)$$

3843 The second expression in Equation 5.17 is obtained by evaluating the constant terms  
3844 assuming a system impedance of  $50 \Omega$ .

3845 We have gone through the effort of expressing the antenna factor in terms of gain  
3846 to highlight that the majority of antenna parameters that we care to measure for a  
3847 CRES experiment can be obtained from the radiation or gain pattern of the antenna.  
3848 The antenna factor is a particularly important parameter for CRES measurements  
3849 due to its relevance to antenna array simulations with the Locust software [62, 80].  
3850 Specifically, Locust simulates the trajectory of an electron in a magnetic trap by running

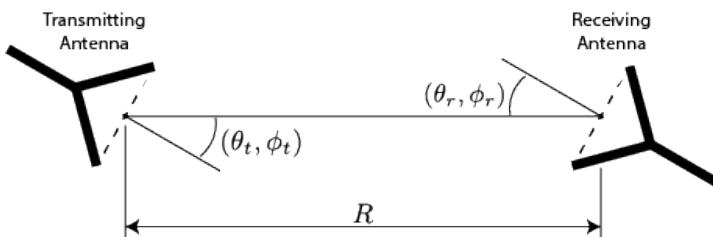
3851 the Kassiopeia software package [60] and then uses the Liénard-Wiechert equations [63, 64]  
3852 to calculate the electric fields that are incident on the antenna.

3853 To compute the response of the antenna to the electric field, Locust relies upon  
3854 linear time-invariant system theory, which computes the response of the antenna (i.e. the  
3855 voltage time series generated by the antenna) using a convolution between the electric field  
3856 time-series and the antenna impulse response. This approach is necessary for correctly  
3857 modeling the antenna response to the electric field due to the broadband and non-  
3858 stationary nature of the electric fields from CRES events. Since antenna measurements  
3859 take place under steady-state conditions, parameters such as the radiation pattern, gain,  
3860 and antenna factor are defined in the frequency domain. However, by performing an  
3861 inverse Fourier transform on the antenna factor we can obtain the antenna impulse  
3862 response, which allows us to simulate CRES events in the antenna array demonstrator  
3863 experiment.

### 3864 **5.2.2 Antenna Measurement Fundamentals**

#### 3865 **5.2.2.1 Friis Transmission Equation**

3866 The antenna factor, sometimes called the antenna transfer function, is used to model  
3867 how the antenna will respond to electric fields emitted from a CRES event. Therefore,  
3868 being able to measure the antenna transfer function of the antenna array is a key step  
3869 in the commissioning and calibration phases of an antenna array CRES experiment. A  
3870 common approach to antenna characterization is to perform a two antenna transmit-  
3871 receive measurement where an antenna with a known gain is used to characterize the  
3872 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.

3872  
3873 To analyze this two antenna setup we seek to calculate the amount of power from  
3874 the transmitting antenna that we will detect with the receiving antenna. Using our  
3875 understanding of antenna gain, we can calculate the power density transmitted by an

3876 antenna in a direction  $(\theta_t, \phi_t)$  at frequency  $f$  and distance  $R$ , which is given by

$$w_t = \frac{P_t}{4\pi R^2} G_t(\theta_t, \phi_t, f). \quad (5.18)$$

3877 Here,  $P_t$  is the total power delivered to the transmitting antenna and  $G_t(\theta_t, \phi_t, f)$  is  
 3878 the value of the transmitting antenna gain. The power density is the power per unit  
 3879 area, so to calculate the total power delivered to the receiving antenna we multiply the  
 3880 transmitted power density by the effective area of the receiving antenna,

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

3881 where  $G_r(\theta_r, \phi_r, f)$  is the gain of the receiving antenna. Equation 5.19 is called the Friis  
 3882 transmission equation [86], which is of fundamental importance for antenna measurements,  
 3883 since it allows one to measure the gain of an unknown antenna by measuring the power  
 3884 received from an antenna with a known gain pattern. Alternatively, if no antenna with a  
 3885 known gain pattern is available, two identical antennas with unknown gain patterns can  
 3886 be used.

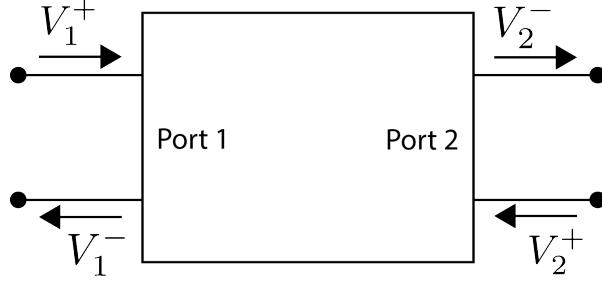
### 3887 5.2.2.2 S-Parameters and Network Analyzers

3888 Instead of directly measuring the power received by the antenna under test, it is more  
 3889 common to measure the ratio of the received power to the transmitted power,

$$\frac{P_r}{P_t} = \frac{G_t(\theta_t, \phi_t, f) G_r(\theta_r, \phi_r, f) c^2}{(4\pi R f)^2}. \quad (5.20)$$

3890 This power ratio can be easily measured using a vector network analyzer (VNA), which  
 3891 automates a significant fraction of the measurement process. Network analyzers are  
 3892 used to measure the scattering or S-parameters of a multi-port RF device [87], which  
 3893 describes how waves are scattered between the device ports. The antenna measurements  
 3894 we have been considering can be modeled as a two-port microwave device that we can  
 3895 characterize by measuring how incident voltage waves are transmitted or reflected (see  
 3896 Figure 5.4). We can write the scattered waves ( $V_1^-$  and  $V_2^-$ ) in terms of the incident ( $V_1^+$   
 3897 and  $V_2^+$ ) waves using the scattering matrix

$$\begin{pmatrix} V_1^- \\ V_2^- \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} V_1^+ \\ V_2^+ \end{pmatrix}, \quad (5.21)$$



**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, we can use this framework to model a two antenna radiation pattern measurement, which we can then automate using a VNA.

3898 where the elements of the matrix are the device S-parameters. It is assumed that,  
 3899 when exciting the device from a particular port, that all other ports in the network are  
 3900 terminated at the system impedance. This ensures that the incident waves from other  
 3901 ports in the network are zero. Therefore, the S-parameters are the ratios between the  
 3902 scattered and incident waves,

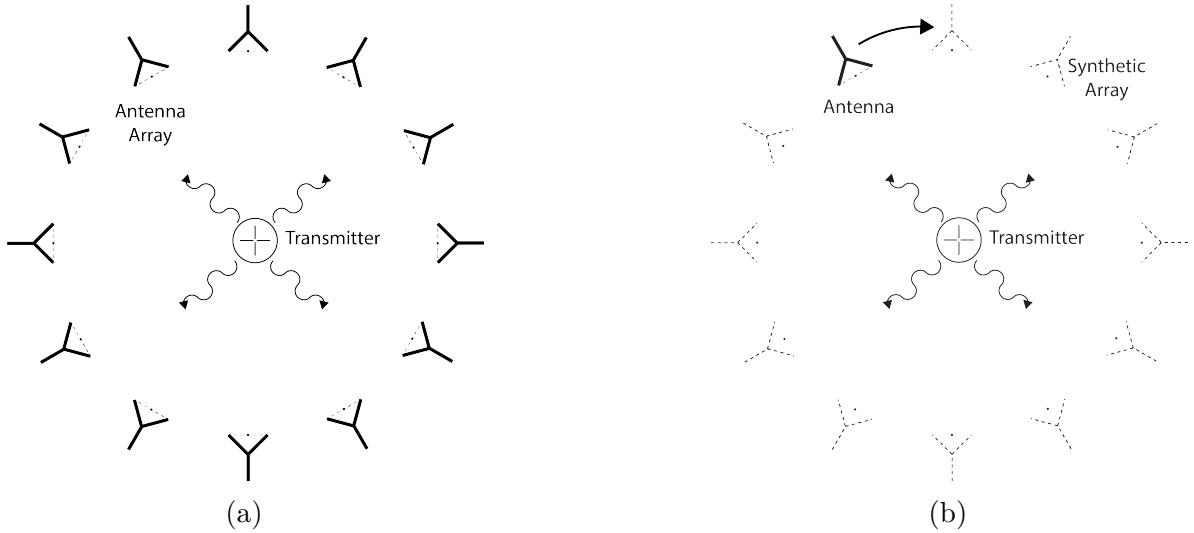
$$S_{ij} = \frac{V_i^-}{V_j^+}. \quad (5.22)$$

3903 Alternatively, S-parameters can be defined as the ratio of the scattered and incident  
 3904 power, which is proportional to the ratio of the squared voltage waves. Returning to  
 3905 our antenna measurement setup, we see that measuring the ratio of the received to the  
 3906 transmitted power is equivalent to measuring the ratio of power being scattered from port  
 3907 1 to port 2 in a RF network. Therefore, measuring an antenna's gain can be accomplished  
 3908 quite easily, by using a VNA to perform a two port  $S_{21}$  measurement.

### 3909 5.2.2.3 Antenna Array Commissioning and Calibration Measurements

3910 Up to this point we have been discussing calibration and commissioning measurements  
 3911 as they apply to a single antenna. While these measurements play an important role  
 3912 in validating the radiation patterns of the individual array elements, the ultimate goal  
 3913 is to use a phased array of these antennas. Therefore, we must also consider antenna  
 3914 measurement techniques that apply to the whole array system.

3915 By measuring the gain of each individual array element we can predict the features of  
 3916 the signals received during a CRES event using the antenna factor (see Section 5.2.1.5).  
 3917 However, unpredictable changes to the antenna performance can be introduced by the  
 3918 incorporation of the antennas into the circular array geometry, therefore, we employ both



**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.

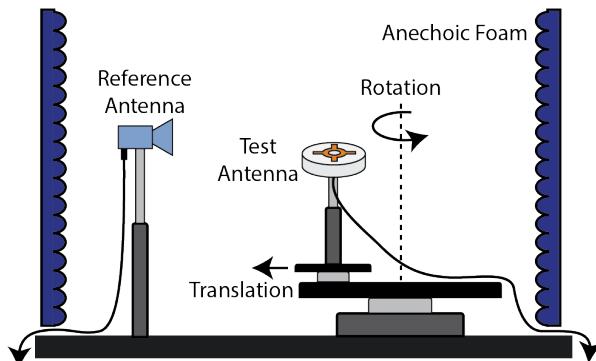
3919 individual antenna and full-array measurements in the commissioning of the FSCD to  
 3920 account for these effects.

3921 There are two main approaches to array measurements that could be used for  
 3922 characterization and calibration (see Figure 5.5). One approach is to construct the  
 3923 complete array and use an omni-directional transmitting antenna to measure the power  
 3924 received by each channel in the antenna array. In Section 5.3 we describe the development  
 3925 of an omni-directional transmitter that also mimics the radiation phase characteristics of a  
 3926 CRES event, which is useful because the entire array can be tested without repositioning.  
 3927 Alternatively, a full antenna array can be synthesized by repeatedly moving and measuring  
 3928 a single array element. This approach is ideal for identifying if different channels in the  
 3929 antenna array are affecting each other through multi-path interference by comparing the  
 3930 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 we developed an antenna measurement system at Penn State specifically designed to mimic the characteristics of the antenna experiment designed for demonstration of the antenna array CRES technique by the Project 8 collaboration.

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 we surround the measurement volume with microwave absorber foam (AEMI AEC-1.5) specifically designed to attenuate microwave radiation near the 26 GHz measurement range of the system.

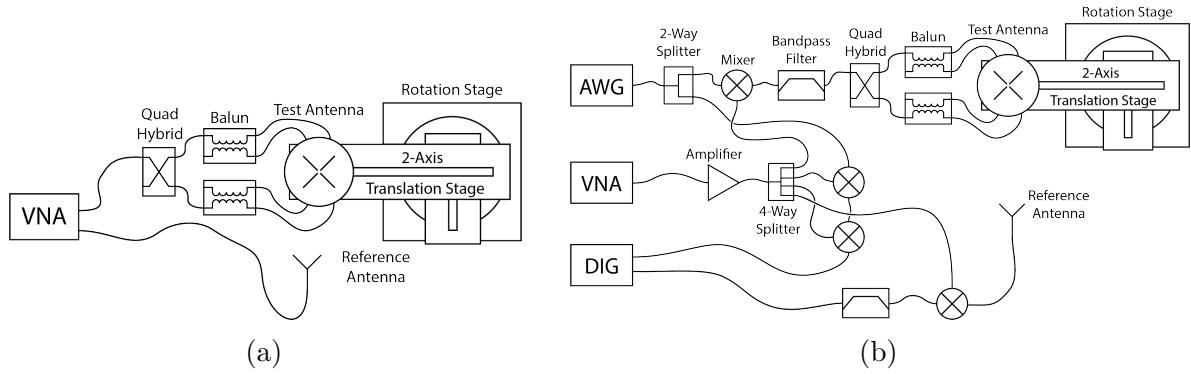


**Figure 5.6.** Illustration of the antenna measurement system developed for the Project 8 Collaboration. The reference and test antennas can be connected to different data acquisition configurations depending on the measurement goals. The reference antenna is typically a standard horn antenna and the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

In the first measurement configuration the reference antenna is typically 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 we wish to characterize. Mounting the test antenna on motorized rotation and translation stages allows us to automate the procedure significantly speeding up the radiation pattern measurement process.

In the second measurement configuration one is interested in recreating the conditions of an antenna array CRES experiment as it concerns the antenna array and DAQ system. In this case, the reference antenna is a prototype FSCD antenna, which will be used to construct the antenna array in the FSCD experiment, and the test antenna is a specially designed synthetic cyclotron antenna (SYNCA) as picture in Figure 5.6. The SYNCA is designed such that the radiation pattern mimics that of a CRES electron so that the signals received by the 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.

In Figure 5.7 we show two high-level system diagrams of the Penn State antenna measurement system that depict the important system components and the connections between them. The two configurations of the measurement system utilize different hardware. For characterization and radiation pattern measurements, one prefers the configuration shown in Figure 5.7a. In this case a vector network analyzer (VNA) is used as both the transmission source and data acquisition system and it is relatively easy to calibrate over a wide range of frequencies. Whereas, if one is more interested

3972 in recreating what would take place in the FSCD experiment then the configuration  
3973 shown in Figure 5.7b is preferable, since this system effectively mimics the receiver chain  
3974 envisioned for the FSCD experiment.

3975 The characterization configuration utilizes a network analyzer (Keysight N5222A)  
3976 with two independent sources and four measurement ports as the primary measurement  
3977 tool. A standard reference antenna is connected to one measurement port, and the test  
3978 antenna is connected to a separate port. The typical reference antenna used for these  
3979 studies is a Pasternack PF9851 horn antenna . In the measurement shown, the test  
3980 antenna represents a SYNCA antenna, which requires a transmission chain consisting of  
3981 quadrature hybrid coupler (Marki QH-0226) connected to two baluns (Marki BAL-0026)  
3982 to generate feed signals with the appropriate phases. The VNA measures the radiation  
3983 pattern by performing a transmission S-parameter measurement, which can be used with  
3984 the knowledge of the reference antenna's radiation pattern to determine the radiation  
3985 pattern of the test antenna (see Section 5.2.1).

3986 The second configuration is more complicated and incorporates more hardware  
3987 components in order to more closely mimic the DAQ system envisioned for the FSCD  
3988 experiment. The basic approach is to produce CRES-like radiation and use an antenna  
3989 combined with a realistic RF receiver chain to acquire the signals. On the transmit side,  
3990 an arbitrary waveform generator (AWG, RIGOL DG5252) is used to generate a waveform  
3991 that mimics a CRES signal at a baseband frequency up to 250 MHz. This frequency is  
3992 then up-converted to the CRES signal frequency band of 25.8 to 26.0 GHz using a mixer  
3993 (Marki MM1-0832L) and a bandpass filter (K&L Microwave 3C62-25900/T200-K/K) to  
3994 reject unwanted mixing components outside out of the 200 MHz CRES signal band. The  
3995 local oscillator signal for mixing is provided by one of the VNA sources configured to run  
3996 in a continuous wave setting. On the receive side, a prototype antenna is used to detect  
3997 the radiation emitted by the test antenna, which is down-converted and filtered using  
3998 the same mixer and bandpass filter as the transmission chain. Lastly, data acquisition is  
3999 performed using a 14-bit ADC sampling at 500 MSa/s (CAEN DT530) to digitize the  
4000 down-converted signals.

4001 In order to distribute the LO to all mixers a 4-way power splitter (MiniCircuits  
4002 ZC4PD-18263-S+) along with an amplifier (Marki APM-6848) is used to drive the four  
4003 mixers used in the measurement system. A limitation of using the VNA as an LO source  
4004 is that there is no control of the LO phase when a measurement is triggered by the  
4005 control script, which leads to a random phase offset between acquisitions. This makes it  
4006 impossible to perform synthetic array measurements, which require strict control over

4007 the starting phase of the transmitted signal. In order to monitor the random phase of the  
4008 LO, a 2-way power splitter (MiniCircuits Z99SC-62-S+) is used to split the signal from  
4009 the AWG between the transmission path and a LO monitoring path. The LO monitoring  
4010 path consists of an up-conversion and down conversion using two mixers connected by a  
4011 coaxial cable, and monitors the relative phase of the LO using a channel on the digitizer  
4012 to sample this path. A phase shift in the LO will lead to a proportional phase shift in  
4013 the mixed signal, which is measured and removed from the received signals.

4014 The test antenna is mounted on a set of motorized stages, which are identical for  
4015 both measurement configurations. A rotational stage (ThorLabs PRMTZ8) is used as  
4016 the base layer with additional translation stages mounted on top of this. The rotational  
4017 stage is ideal for measuring a complete azimuthal scan of the test antenna's radiation  
4018 pattern as well as for moving a SYNCA antenna in circular motion to recreate the  
4019 symmetry of the FSCD antenna array. On top of the rotational stage we mount two  
4020 linear translation stages (ThorLabs MTS50-Z8 and MTS25-Z8) in a cross-wise manner  
4021 so that the test antenna can be moved along two perpendicular axes. Using the linear  
4022 stages in combination with the rotational stage allows one to fine-tune the positioning of  
4023 the test antenna so that it can be perfectly aligned with the central axis of the array.  
4024 A LabView script was developed to automate the measurement of a full 360° radiation  
4025 pattern and control the measurement electronics. Data from these acquisitions is stored  
4026 on university provided cloud storage.

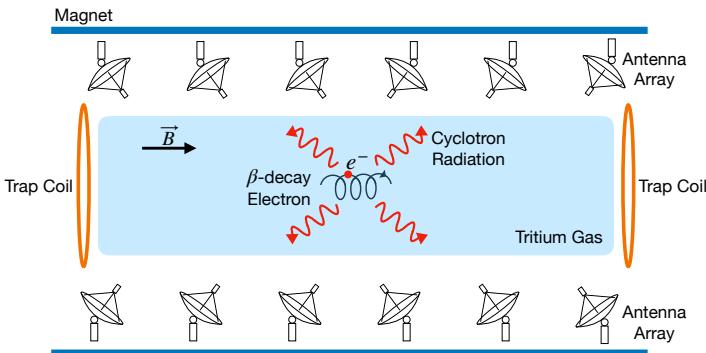
## 4027 **5.3 Development of a Synthetic Cyclotron Antenna (SYNCA)** 4028 **for Antenna Array Calibration**

4029 This section is the manuscript of the publication [81] detailing the development of a  
4030 Synthetic Cyclotron Antenna (SYNCA) for antenna array characterization measurements  
4031 by the Project 8 collaboration.

### 4032 **5.3.1 Introduction**

4033 Neutrinos are the most abundant standard model fermions in our universe, but due to  
4034 weak interaction cross-sections with other particles, neutrinos are particularly difficult  
4035 to study. Consequently, many fundamental properties of neutrinos are still unknown  
4036 including the absolute scale of the neutrino mass [31]. Direct, kinematic measurements of  
4037 the neutrino mass are particularly valuable due to their model independent nature [37].

4038 To date the most sensitive direct neutrino mass measurements have been performed by  
 4039 the KATRIN collaboration [88], which measures the molecular tritium  $\beta$ -decay spectrum  
 4040 to infer the neutrino mass. Current data from neutrino oscillation measurements [31]  
 4041 allow for neutrino masses significantly smaller than the design sensitivity of the KATRIN  
 4042 experiment; therefore, there is a need for new technologies for performing direct neutrino  
 4043 mass measurements to probe lower neutrino masses.

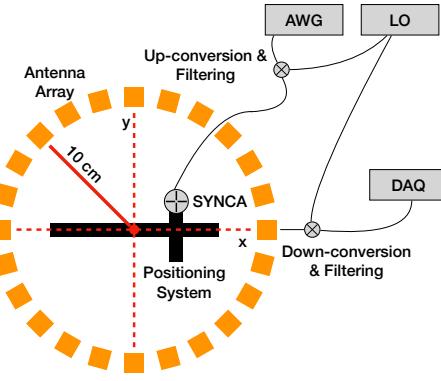


**Figure 5.8.** A sketch of an antenna array large-volume CRES experiment. Electrons from  $\beta$ -decays are confined in a magnetic field using a set of trap coils. The cyclotron radiation produced by the motion of the trapped electrons can be detected by a surrounding antenna array to determine the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

4044 The Project 8 collaboration is developing new methods for neutrino mass measurement  
 4045 based on Cyclotron Radiation Emission Spectroscopy (CRES) [56, 89–91], with the goal  
 4046 of measuring the absolute scale of the neutrino mass with a 40 meV/c<sup>2</sup> sensitivity [?, 37].  
 4047 This sensitivity goal will require the development of two separate technical capabilities.  
 4048 First is the development of an atomic tritium source, which avoids significant spectral  
 4049 broadening due to molecular final states [55]. Second is the technology for performing  
 4050 CRES in a multi-cubic-meter experimental volume with high combined detection and  
 4051 reconstruction efficiency, which is required in order to obtain sufficient event statistics  
 4052 near the tritium spectrum endpoint.

4053 One approach for a large-volume CRES experiment is to use an array of antennas,  
 4054 which surrounds a volume of tritium gas, to detect the cyclotron radiation produced  
 4055 by the  $\beta$ -decay electrons when they are trapped in a background magnetic field using a  
 4056 set of magnetic trapping coils (see Figure 5.8). Project 8 has developed a conceptual  
 4057 experiment design to study the feasibility of this approach. The design consists of a  
 4058 single circular array of antennas with a radius of 10 cm and 60 independent channels  
 4059 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 [80, 92], 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.

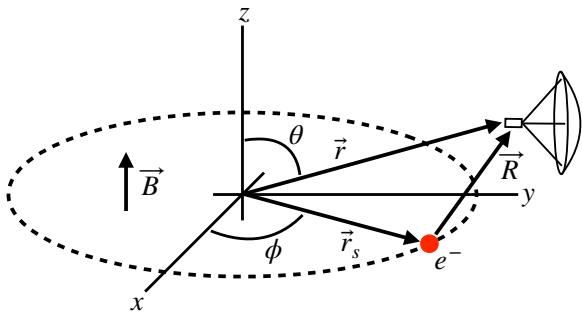
We call this probe antenna the Synthetic Cyclotron Antenna or SYNCA. The SYNCA is a novel antenna design that mimics the cyclotron radiation generated by individual charged particles trapped in a magnetic field, which will be used in the antenna test stand to perform characterization measurements, simulation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

In Section 5.3.2 we provide a description of the cyclotron radiation field characteristics that we recreate with the SYNCA. In Section 5.3.3 we give an overview of the simulations performed to develop an antenna design that mimics the characteristics of cyclotron radiation. In Section 5.3.4 we outline characterization measurements to validate that the fields generated by the SYNCA match simulation, and finally in Section 5.3.5 we

4082 demonstrate an application of the SYNCA to test phased array reconstruction techniques  
 4083 on the bench-top.

### 4084 5.3.2 Cyclotron Radiation Phenomenology

4085 To understand the cyclotron radiation phenomenology that the SYNCA should mimic,  
 4086 we consider a charged particle moving at relativistic speed in the presence of an external  
 4087 magnetic field (see Figure 5.10). In the special case we shall examine, the entirety of  
 4088 the electron's momentum is directed perpendicular to the magnetic field; therefore, the  
 4089 trajectory of the electron is confined to the cyclotron orbit plane. Because the momentum  
 4090 vector is oriented perpendicular to the magnetic field, electrons with these trajectories  
 4091 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.

4092 The cyclotron radiation fields generated by this circular trajectory are those which  
 4093 we aim to reproduce with the SYNCA. We can describe the electromagnetic (EM) fields  
 4094 using the Liénard-Wiechert equations [49, 80], which in non-covariant form express the  
 4095 electric field as

$$\vec{E} = e \left[ \frac{\hat{n} - \vec{\beta}}{\gamma^2 (1 - \vec{\beta} \cdot \hat{n})^3 |\vec{R}|^2} \right]_{t_r} + \frac{e}{c} \left[ \frac{\hat{n} \times [(\hat{n} - \vec{\beta}) \times \dot{\vec{\beta}}]}{(1 - \vec{\beta} \cdot \hat{n})^3 |\vec{R}|} \right]_{t_r}, \quad (5.23)$$

4096 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  
 4097 electron to the field measurement point,  $\vec{\beta} = \dot{\vec{r}}_s/c$  is the velocity of the particle divided  
 4098 by the speed of light, and  $\gamma$  is the relativistic Lorentz factor. The equation is meant to  
 4099 be evaluated at the retarded time as indicated by  $t_r = t - |\vec{R}|/c$ , which accounts for the  
 4100 time delay due to the finite speed of light between the point where the field was emitted  
 4101 and the point where the field is detected.

We would like to simplify Equation 5.23 it at all possible. As a first step we analyze the relative magnitudes of the electric field polarization components. Consider an electron following a circular cyclotron orbit in a uniform magnetic field whose guiding center is positioned at the origin of the coordinate system. The equation of motion can be expressed as

$$\vec{r}_s = (r_c \cos \omega_c t_r) \hat{x} + (r_c \sin \omega_c t_r) \hat{y}. \quad (5.24)$$

For single antenna located along the y-axis at position  $\vec{r} = r_a \hat{y}$  we are interested in the incident electric fields from the electron. The electric field is given by Equation 5.23, which we evaluate in the regime where  $r_a \gg r_c$ . This limit can be justified by comparing the radius of the cyclotron orbit for an electron with the tritium beta-spectrum endpoint energy of 18.6 keV in a 1 T magnetic field to the typical ( $r_a \simeq 100$  mm) radial position of the receiving antenna. We find that the cyclotron orbit has a radius of 0.46 mm which is approximately a factor of 200 smaller than the typical antenna radial position. In this regime we can make the approximation  $\vec{R} \simeq r_a \hat{y}$  and the expression for the electric field at the antenna's position becomes

$$\vec{E} = \frac{e}{\gamma^2 r_a^2} \frac{\hat{x}(\frac{r_c \omega_c}{c} \sin \omega_c t_r) + \hat{y}(1 - \frac{r_c \omega_c}{c} \cos \omega_c t_r)}{(1 - \frac{r_c \omega_c}{c} \cos \omega_c t_r)^3} - \frac{e}{cr_a} \frac{\hat{x}(\frac{r_c^2 \omega_c^3}{c^2} - \frac{r_c \omega_c^2}{c} \cos \omega_c t_r)}{(1 - \frac{r_c \omega_c}{c} \cos \omega_c t_r)^3}. \quad (5.25)$$

Since the receiving antenna is part of a circular array of antennas, it is useful to rewrite Equation 5.25 in terms of the azimuthal ( $\hat{\phi}$ ) and radial ( $\hat{r}$ ) polarizations. Making use of 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

$$\vec{E} = \hat{\phi} E_\phi + \hat{r} E_r \quad (5.26)$$

$$E_\phi = \frac{e}{(1 - \frac{r_c \omega_c}{c} \cos \omega_c t_r)^3} \left[ -\frac{\frac{r_c \omega_c}{c} \sin \omega_c t_r}{\gamma^2 r_a^2} + \frac{\omega_c \left( \frac{r_c^2 \omega_c^2}{c^2} - \frac{r_c \omega_c}{c} \cos \omega_c t_r \right)}{cr_a} \right] \quad (5.27)$$

$$E_r = \frac{e(1 - \frac{r_c \omega_c}{c} \sin \omega_c t_r)}{\gamma^2 r_a^2 (1 - \frac{r_c \omega_c}{c} \cos \omega_c t_r)^3}. \quad (5.28)$$

For the purposes of designing a synthetic cyclotron radiation antenna we are interested in the dominant electric field polarization emitted by the electron. The antenna is being designed to mimic the cyclotron radiation produced by electrons with kinetic energies of approximately 18.6 keV in a 1 T magnetic field [55]. Since the relativistic beta factor for 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 justified. Inserting these expressions into the equations for the electric field components

4125 above simplifies the comparison of the magnitudes of the two components. Additionally,  
 4126 we compare the time-averaged magnitudes to evaluate the root mean squared electric  
 4127 field ratio. The time-averaged ratio of the radial and azimuthally polarized electric fields  
 4128 with the above simplifications is given by

$$\frac{\langle |E_r| \rangle}{\langle |E_\phi| \rangle} = \frac{8 - \sqrt{2}}{\left| 1 - \frac{r_a}{r_c} \frac{1-2\sqrt{2}}{8} \right|} \simeq \frac{r_c}{r_a} \frac{8(8 - \sqrt{2})}{2\sqrt{2} - 1} = 0.13, \quad (5.29)$$

4129 where we have made use of the fact that for these magnetic fields and kinetic energies  
 4130 the cyclotron radius is much smaller than the radius of the antenna array.

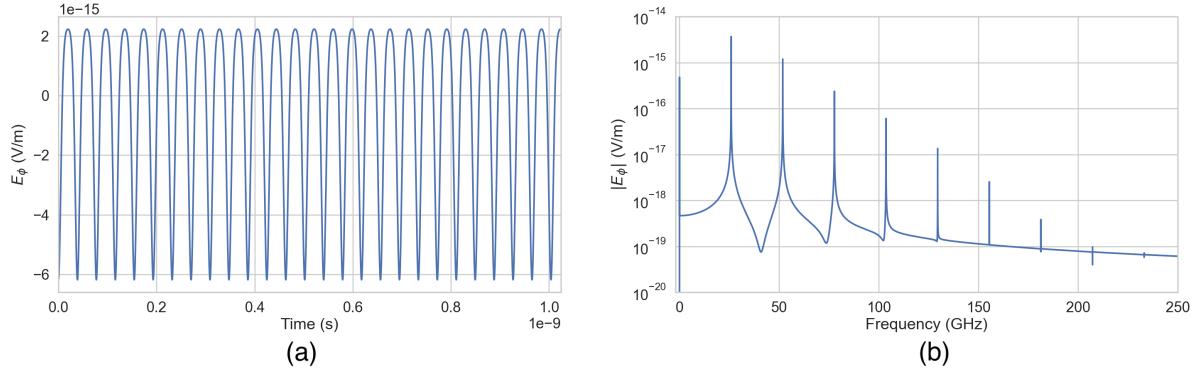
4131 From Equation 5.29 we see that the time-averaged azimuthal polarization is larger than  
 4132 the radial polarization by about a factor of 8, which makes it the dominant contribution  
 4133 to the electric fields at the position of the antenna. We must also consider the directivity  
 4134 of the receiving antenna which can have a gain that is disproportionately large for a  
 4135 specific polarization component. Because the  $E_\phi$  component is dominant, the receiving  
 4136 antenna array is designed with an azimuthal polarization, which negates the voltages  
 4137 induced in the antenna from the radially polarized fields. Therefore, we conclude that  
 4138 for the purpose of designing the SYNCA antenna it is acceptable to approximate the  
 4139 electric fields from Equation 5.23 as purely azimuthally or  $\phi$ -polarized. The simplified  
 4140 expression for the electric field received by an antenna becomes

$$\vec{E} = E_\phi \hat{\phi} = \frac{e \frac{r_c \omega_c}{c}}{4r_a r_c} \left[ \frac{\frac{r_c \omega_c}{c} - \cos \omega_c t - \frac{4r_c}{r_a} \sin \omega_c t}{(1 - \frac{r_c \omega_c}{c} \cos \omega_c t)^3} \right]_{t_r} \hat{\phi}, \quad (5.30)$$

4141 where the radius of the cyclotron orbit is called  $r_c$ , the cyclotron frequency is called  $\omega_c$ ,  
 4142 and the radial position of the receiving antenna is called  $r_a$ . Equation 5.30 has been  
 4143 evaluated in the non-relativistic limit where  $\gamma \simeq 1$ , which is justified by the fact that  
 4144  $|\vec{\beta}| \simeq \frac{c}{4}$  for an electron with an 18.6 keV kinetic energy in a 1 T magnetic field.

4145 This rather complicated expression can be simplified using Fourier analysis. Assuming  
 4146 a background magnetic field of 1 T and a kinetic energy of 18.6 keV we calculate  
 4147 numerically the electric field using Equation 5.30 and apply a discrete Fourier Transform  
 4148 to visualize the frequency spectrum (see Figure 5.11).

4149 We observe that the azimuthally polarized electric field is periodic with a base cyclotron  
 4150 frequency of 25.898 GHz corresponding to the highest power frequency component in  
 4151 Figure 5.11. The frequency spectrum reveals that the signal is composed of a constant  
 4152 term with zero frequency and a series of harmonics separated by 25.898 GHz. Therefore,



**Figure 5.11.** A plot of the numeric solution to Equation 5.31. 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.

we can represent the azimuthal electric fields from the electron as a linear combination of pure sinusoids with frequencies given by  $\omega_k = k\omega_c$  ( $k \in 0, 1, 2, \dots$ ) and amplitudes extracted from the Fourier representation. Using this representation we can transform the equation for the azimuthally polarized electric fields in Equation 5.30 into

$$E_\phi = \frac{e^{\frac{r_c \omega_c}{c}}}{4r_a r_c} \sum_{k=0}^7 A_k e^{i\omega_k t_r}, \quad (5.31)$$

where we have truncated the sum over harmonics at the 7th order for completeness. The amplitudes  $A_k$  are dimensionless complex numbers, which encode the relative powers of the harmonics as well as the starting overall phase of the cyclotron radiation. Because magnitude of the relative amplitudes exponentially decreases for higher harmonics, it is usually sufficient to consider only the terms up to  $k = 4$  where the relative amplitude of the harmonics has decreased from the main carrier by a factor of approximately 100. However, for completeness we include harmonics up to 7th order in Equation 5.31. The range of frequencies to which the receiving antenna array in the antenna test stand is sensitive is defined by the antenna's transfer function. The receptive bandwidth for the antennas used in the test stand is a range of frequencies with a bandwidth on the order of a few GHz centered around the main cyclotron carrier frequency of 25.898 GHz. Therefore, the higher order harmonics as well as the zero frequency term can be ignored when considering only the signals that will be received by the antenna array.

Considering only the 1st order harmonic term from Equation 5.31, which represents the portion of the electric field that will be detected by the array, and evaluating this at

<sup>4172</sup> the retarded time we obtain the following for the  $\phi$ -polarized electric fields

$$E_\phi \propto \cos \left( \omega_c \left( t - |\vec{R}|/c \right) - \Delta \right), \quad (5.32)$$

<sup>4173</sup> where the arbitrary phase  $\Delta$  is defined by  $A_k = |A_k|e^{i\Delta}$ . We are interested in the  
<sup>4174</sup> characteristics of the amplitude of the electric field as a function of the radial distance  
<sup>4175</sup> component ( $|\vec{R}|$ ) of the retarded time. In particular, the maximum of  $E_\phi$  occurs when  
<sup>4176</sup> the argument of the cosine function is equal  $n\pi$  where  $n \in \{0, \pm 2, \pm 4, \dots\}$ ; however, the  
<sup>4177</sup> solutions where  $n$  is negative can be discarded since they represent unphysical negative  
<sup>4178</sup> overall phases. Applying this condition to Equation 5.32 gives a condition on the radial  
<sup>4179</sup> position of the maximum of  $E_\phi$

$$\omega_c(t - |\vec{R}|/c) - \Delta = n\pi, \quad (5.33a)$$

$$|\vec{R}| = \frac{c}{\omega_c} ((\omega_c t - \Delta) - n\pi), \quad (5.33b)$$

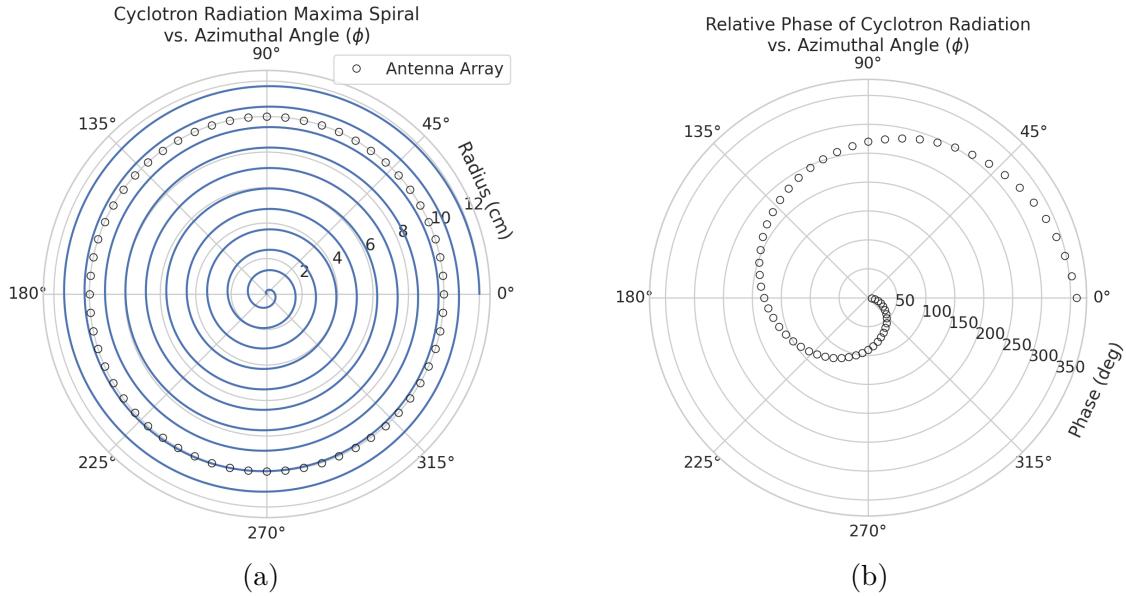
<sup>4180</sup> which is a function of time in the frame of the moving electron ( $t$ ). Equation 5.33 can  
<sup>4181</sup> be further simplified by noticing that the azimuthal position of the electron ( $\phi_e(t)$ ) as a  
<sup>4182</sup> function of time is defined by  $\phi_e(t) = \omega_c t - \Delta$  which reduces Equation 5.33 to

$$|\vec{R}| = \frac{c}{\omega_c} (\phi_e(t) - n\pi). \quad (5.34)$$

<sup>4183</sup> Equation 5.34 represents an archimedian spiral which is formed when plotting the  
<sup>4184</sup> amplitude of  $E_\phi$  in the x-y plane. The solution where  $n = 0$  represents the leading edge  
<sup>4185</sup> of the radiation spiral which propagates outward from the electron at the speed of light.  
<sup>4186</sup> The additional solutions for  $n > 0$  represent the persistent spiral at radii inside the  
<sup>4187</sup> leading edge of the radiated fields that have not yet been detected by the receiver at the  
<sup>4188</sup> current time. In Figure 5.12a we show the expected spiral pattern for the maxima of the  
<sup>4189</sup> cyclotron radiation.

<sup>4190</sup> In particular, we note that for the circular array geometry of the test stand, depicted  
<sup>4191</sup> as the series of circles in Figure 5.12a, each antenna receives a linearly polarized wave  
<sup>4192</sup> with a phase offset that corresponds to the azimuthal angle for that antenna element.  
<sup>4193</sup> Therefore, as we show in Figure 5.12b, when the relative phase of the received signal is  
<sup>4194</sup> plotted as a function of the receiving antenna's azimuthal position the result is also an  
<sup>4195</sup> Archimedean spiral.

<sup>4196</sup> Based on these analytical calculations we can characterize the magnitude, polarization,  
<sup>4197</sup> and phase of the signals received by the antenna array using three criteria. These criteria



**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).

4198 are the basis of comparison for the radiation produced by the SYNCA and cyclotron  
 4199 radiation emitted by electrons and will be used to evaluate the performance of antenna  
 4200 designs. The criteria are:

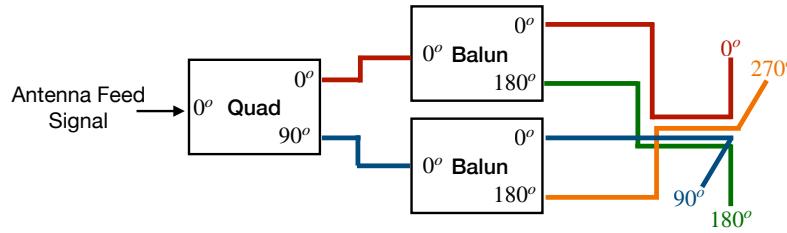
- 4201 1. Electric fields that are  $\phi$ -polarized near  $\theta = 90^\circ$
- 4202 2. Uniform time-averaged electric field magnitudes around the circumference of a  
 4203 circle centered on the antenna
- 4204 3. Electric fields whose phase is equal to the azimuthal angle at the point of measure-  
 4205 ment plus a constant

4206 The Locust simulation package [92] can be used to directly simulate the EM fields  
 4207 generated by electrons performing cyclotron motion to validate the analytical calculations.  
 4208 Locust simulates the EM fields by first calculating the trajectory of the electrons in  
 4209 the magnetic trap using the Kassiopeia software package [93]. The trajectory can then  
 4210 be used to solve for the EM fields using the Liénard-Wiechert equations directly with  
 4211 no approximations. The resulting electric field solutions drive a receiving antenna by

4212 convolving the time-domain fields with the finite-impulse response filter of the antenna  
 4213 or they can be examined directly to study the field characteristics that the SYNCA must  
 4214 reproduce. In the next section we compare the radiation field patterns for electrons  
 4215 simulated with Locust to patterns from a SYNCA antenna design.

### 4216 **5.3.3 SYNCA Simulations and Design**

4217 One potential SYNCA design is the crossed-dipole antenna [94]. A crossed-dipole antenna  
 4218 consists of two dipole antennas, one of which is rotated 90° with respect to the other,  
 4219 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.

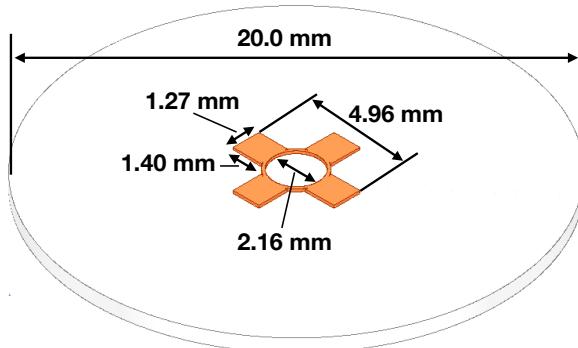
4220  
 4221 out of phase from each of the neighboring arms by 90°, which mirrors the spatial phase  
 4222 relationship of cyclotron radiation fields.

4223 A potential drawback of this design is that standard crossed-dipole antennas do not  
 4224 radiate uniform electric fields near the  $\theta = \pi/2$  plane. Typical crossed-dipole antennas  
 4225 use dipole arm lengths equal to  $\lambda/4$  or larger [94], where  $\lambda$  is the wavelength at the  
 4226 desired operating frequency. Such large arm lengths cause the electric field magnitude  
 4227 to vary significantly around the circumference of the antenna. However, making the  
 4228 antenna electrically small by shrinking the arm length can improve the antenna pattern  
 4229 uniformity.

4230 In general, the criterion for an electrically small antenna is that the largest dimension  
 4231 of the antenna ( $D$ ) obey  $D \lesssim \lambda/10$  [65]. In our application, we are attempting to mimic  
 4232 the cyclotron radiation emitted by electrons produced from tritium  $\beta$ -decay with energies  
 4233 near the spectrum endpoint. For a background magnetic field of 1 T, the corresponding

4234 cyclotron frequency of tritium endpoint electrons is approximately 26 GHz. Therefore, the  
4235 electrically small condition would require that the largest dimension of the crossed-dipole  
4236 antenna be smaller than 1.2 mm.

4237 A crossed-dipole antenna with an overall size of 1.2 mm is challenging to fabricate due  
4238 to the small dimensions of the dipole arms that, in practice, are fragile and unsuitable  
4239 for use as a calibration probe. To mitigate some of the challenges with the fabrication  
4240 of such a small antenna, a variant crossed-dipole antenna design using printed circuit  
4241 board (PCB) technology (see Figure 5.14) was developed in partnership with an antenna  
prototyping company, Field Theory Consulting <sup>1</sup>.



**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.

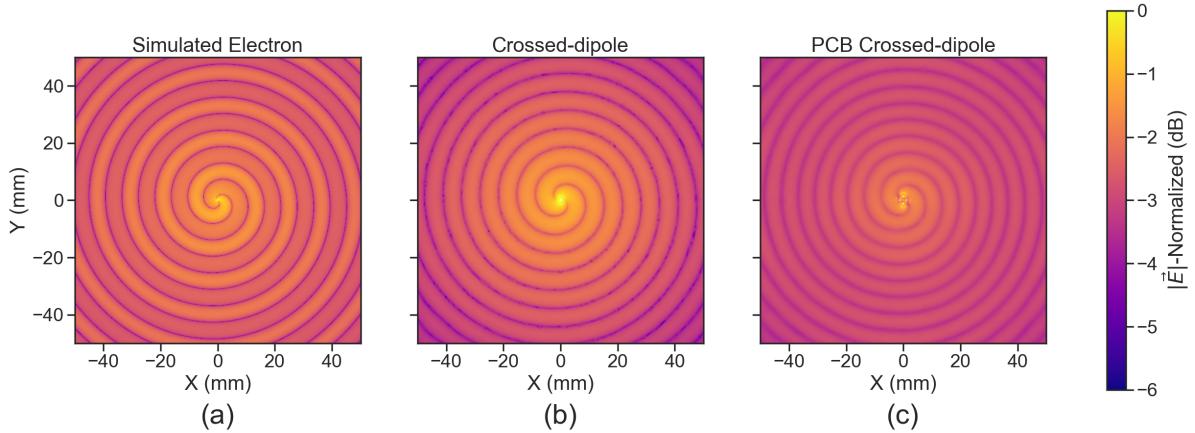
4242  
4243 The PCB crossed-dipole design uses four rectangular pads to represent the dipole arms,  
4244 which are connected by a thin circular trace. The circular trace both adds mechanical  
4245 stability to the antenna and improves the azimuthal uniformity of the electric fields  
4246 compared to a more standard crossed-dipole geometry. Furthermore, the circular trace  
4247 allows for a greater separation between dipole arms than standard crossed-dipoles, which  
4248 is required to accommodate the coaxial connections to each pad. The pads each contain  
4249 a through-hole solder joint to connect coaxial transmission lines using hand soldering.  
4250 The antenna PCB has no ground plane on the bottom layer as this was observed in  
4251 simulation to significantly distort the radiation pattern in the plane of the PCB. The  
4252 only ground planes present in the model are the outer conductors of the four coaxial

---

<sup>1</sup><https://fieldtheoryinc.com/>

4253 transmission lines which feed the antenna. These are left unterminated on the bottom of  
4254 the PCB dielectric material.

4255 The antenna design development utilized a combination of Locust electron simula-  
4256 tions and antenna simulations using ANSYS HFSS [66], a commercial finite-element  
4257 electromagnetic simulation software. Two antenna designs were simulated: an idealized  
4258 electrically small crossed-dipole antenna with an arm length of 0.40 mm and an arm  
4259 separation of 0.05 mm, as well as a PCB crossed-dipole antenna with the dimensions  
4260 shown in Figure 5.14. Plotting the magnitude of the electric fields generated by the  
4261 antennas across a 10 cm square located in the same plane as the respective antennas  
4262 reveals the expected cyclotron spiral pattern (see Figure 5.15) which closely matches  
4263 the prediction for simulated electrons. The spiral pattern demonstrates that the electric  
4264 fields have the appropriate phases to mimic cyclotron radiation, which fulfills SYNCA  
criterion 3 identified in Section 5.3.2.

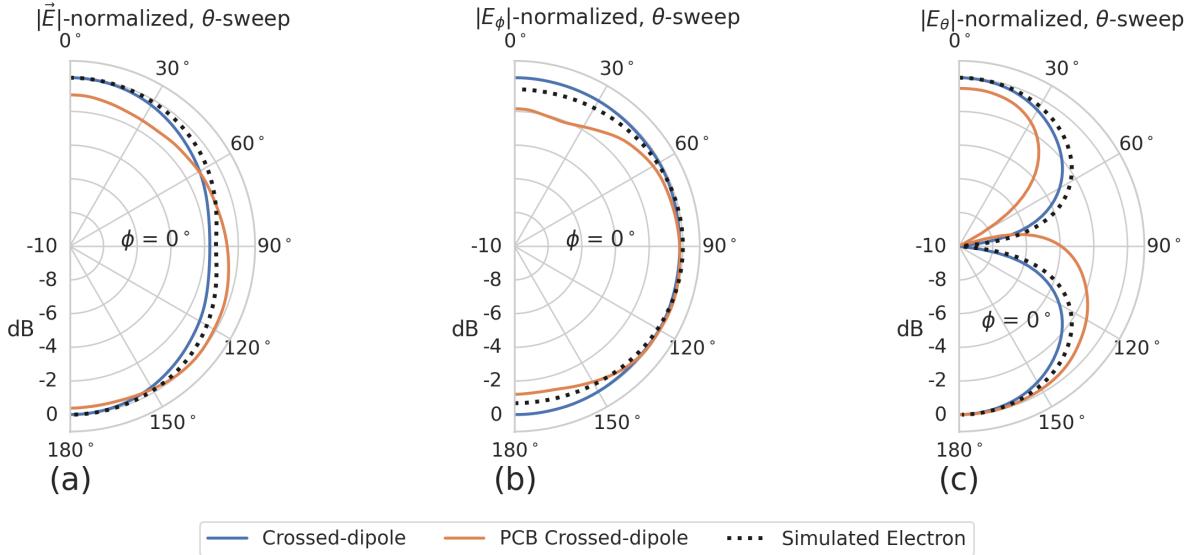


**Figure 5.15.** A comparison of the electric field magnitudes, normalized by the maximum value of the electric field in each simulation, plotted on a 10 cm square to visualize the Archimedean spirals formed by the electron (a), the crossed-dipole antenna (b), and a PCB crossed-dipole antenna (c). The 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.

4265  
4266 As we can see from Figure 5.16, the crossed-dipole antenna, which uses an idealized  
4267 geometry, exhibits good agreement with simulation. The antenna has a maximum  
4268 deviation from a simulated electron of approximately 0.5 dB in the total electric field, 1  
4269 dB for the  $\phi$ -polarized electric field and 1 dB for the  $\theta$ -polarized electric field.

4270 In comparison, the pattern of the PCB crossed-dipole antenna, because the simulation  
4271 incorporates the geometry of the coax transmission lines, exhibits some distortion from  
4272 the idealized cross-dipole simulations. The vertically oriented ground planes of the coax

4273 lines introduce more  $\theta$ -polarized electric fields than are observed for simulated electrons  
 4274 near  $\theta = 90^\circ$ . The significant  $\theta$ -polarized field minimum is still present but shifted  
 to approximately  $\theta = 65^\circ$ . The  $\theta$ -polarized field deviations of the PCB crossed-dipole

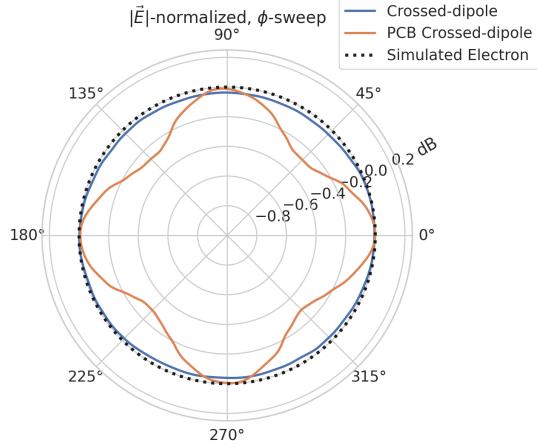


**Figure 5.16.** A comparison of the normalized electric field magnitudes for the ideal crossed-dipole, PCB crossed-dipole, and a simulated electron as a function of the polar angle ( $\theta$ ). (a) Shows the total electric field, (b) shows the  $\phi$ -polarized electric field component, and (c) shows the  $\theta$ -polarized electric field component. These images were generated using Locust simulations for the electron and ANSYS HFSS for both antennas.

4275  
 4276 antenna should not greatly impact the performance of the antenna because the receiving  
 4277 antenna array is primarily  $\phi$ -polarized. Therefore deviations in the  $\theta$ -polarized fields  
 4278 will be suppressed due to the polarization mismatch. More importantly, the  $\phi$ -polarized  
 4279 electric field pattern generated by the PCB crossed-dipole closely matches simulated  
 4280 electrons across the polar angle range of  $50^\circ < \theta < 150^\circ$ . In this region the PCB crossed-  
 4281 dipole differs by less than 0.5 dB from simulated electrons. This range greatly exceeds  
 4282 the beamwidth of the receiving antenna array which is designed to be most sensitive  
 4283 to fields produced near  $\theta = 90^\circ$ . Therefore, we conclude that the PCB crossed-dipole  
 4284 antenna generates a  $\phi$ -polarized radiation pattern that fulfills SYNCA criterion 1 from  
 4285 Section 5.3.2.

4286 The final SYNCA criterion is related to the uniformity of the electric fields when  
 4287 measured azimuthally around the antenna. As we saw for real electrons in Section 5.3.2  
 4288 it is expected that the magnitude of the electric field be completely uniform as a function  
 4289 of the azimuthal angle due to the symmetry of the cyclotron orbit. In Figure 5.17 we plot  
 4290 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 ( $\phi$ ) evaluated at  $\theta = 90^\circ$ . This image was generated using Locust simulations for the electron and ANSYS HFSS for both antennas.

4291

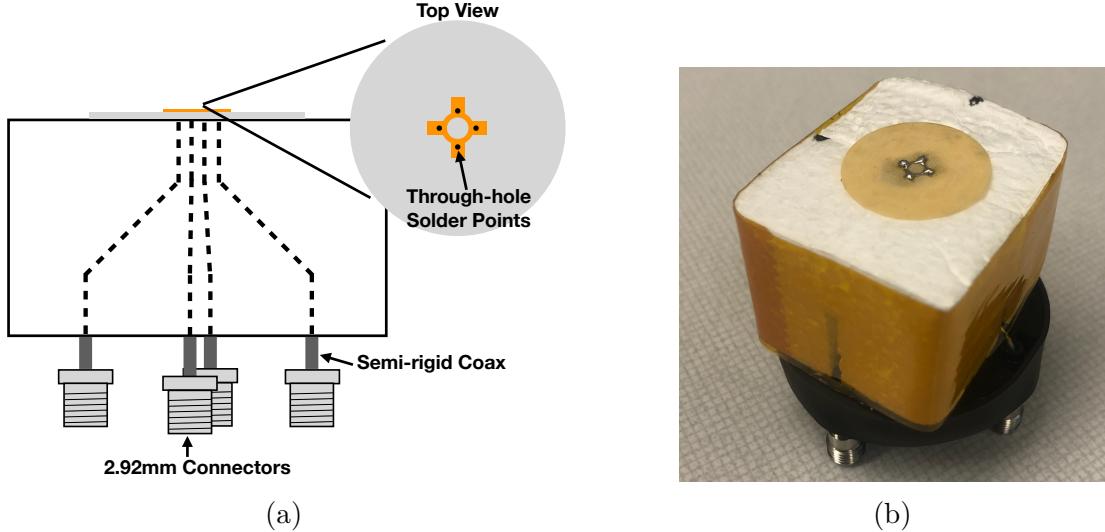
4292 perfect uniformity around the azimuthal angle, whereas the PCB crossed-dipole has a  
 4293 small periodic deviation with a maximum difference of 0.3 dB caused by the coaxial  
 4294 transmission lines below the PCB. Such a small deviation from uniformity is acceptable  
 4295 since it is smaller than the expected variation in uniformity caused by imperfections in  
 4296 the antenna fabrication process, which modifies the antenna shape in an uncontrolled  
 4297 manner by introducing solder blobs with a typical size of a few tenths of a millimeter on  
 4298 the dipole arms (see Figure 5.18). Additionally, the SYNCA will be separately calibrated  
 4299 to account for azimuthal differences in the electric field magnitude. Therefore we see  
 4300 from the simulated performance of the PCB crossed-dipole antenna that this antenna  
 4301 design meets all three of the SYNCA criteria.

4302

### 5.3.4 Characterization of the SYNCA

4303 Two SYNCAs were manufactured using the PCB crossed-dipole design (see Figure 5.18).  
 4304 The antenna PCB (Matrix Circuit Board Materials, MEGTRON 6) is connected to  
 4305 four 2.92 mm coaxial connectors (Fairview Microwave, SC5843) using semi-rigid coax  
 4306 (Fairview Microwave, FMBC002), which also physically support the antenna PCB. The  
 4307 antenna PCB consists only of two layers which correspond to the copper antenna trace  
 4308 and the PCB dielectric. Each coax line is connected to the associated dipole arm using  
 4309 through-hole soldering and phase matched to ensure that the electrical length of each  
 4310 of the transmission lines is identical at the operating frequency. The antenna PCB is

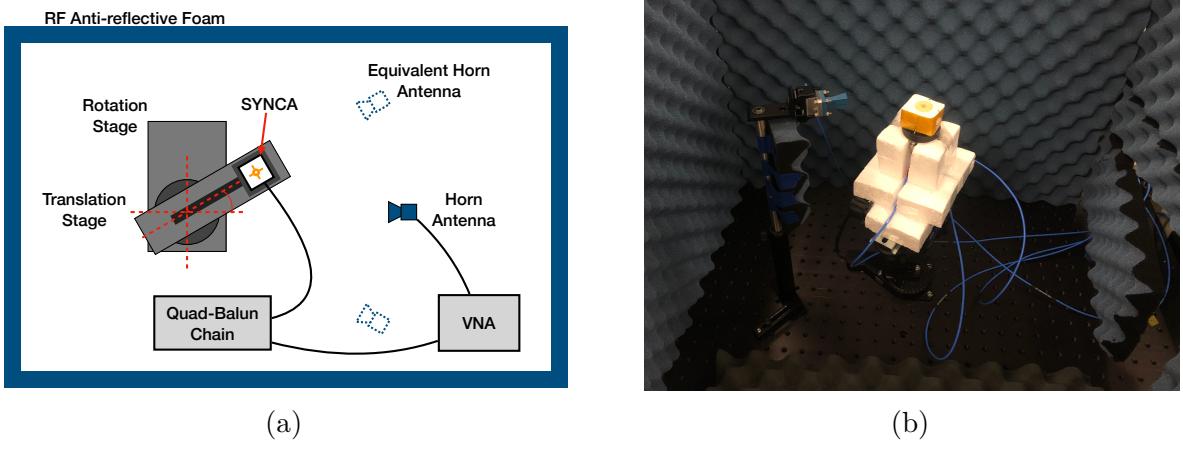
4311 further reinforced using custom cut polystyrene foam blocks, which have an electrical  
 4312 permittivity nearly identical to air. A custom 3D printed mount is included at the base  
 4313 of the antenna to support the coax connectors and to provide a sturdy mounting base.



**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.

4314 Characterization measurements were performed using a Vector Network Analyzer  
 4315 (VNA) to measure the electric field magnitude and phase radiated by the SYNCA to  
 4316 verify the radiation pattern (see Figure 5.19). The VNA is connected to the SYNCA  
 4317 at one port through a hybrid-coupler whose outputs are connected to two baluns to  
 4318 generate the signals with the appropriate phases to feed the SYNCA (see Figure 5.13).  
 4319 The other port of the VNA is connected to a single reference horn antenna that serves  
 4320 as a field probe. To position the SYNCA, a combination of translation and rotation  
 4321 stages are used to characterize the antenna's fields across the entire radiation pattern  
 4322 circumference. This measurement scheme is equivalent to measuring the fields generated  
 4323 by the SYNCA using a full circular array of probe antennas.

4324 The antenna measurement space is surrounded by RF anti-reflective foam to isolate  
 4325 the measurements from the lab environment (see Figure 5.19b) and remaining reflections  
 4326 are removed using the VNA's time-gating feature. The SYNCA is affixed to the stages  
 4327 by a custom RF transparent mount made of polystyrene foam. The coaxial cables deliver

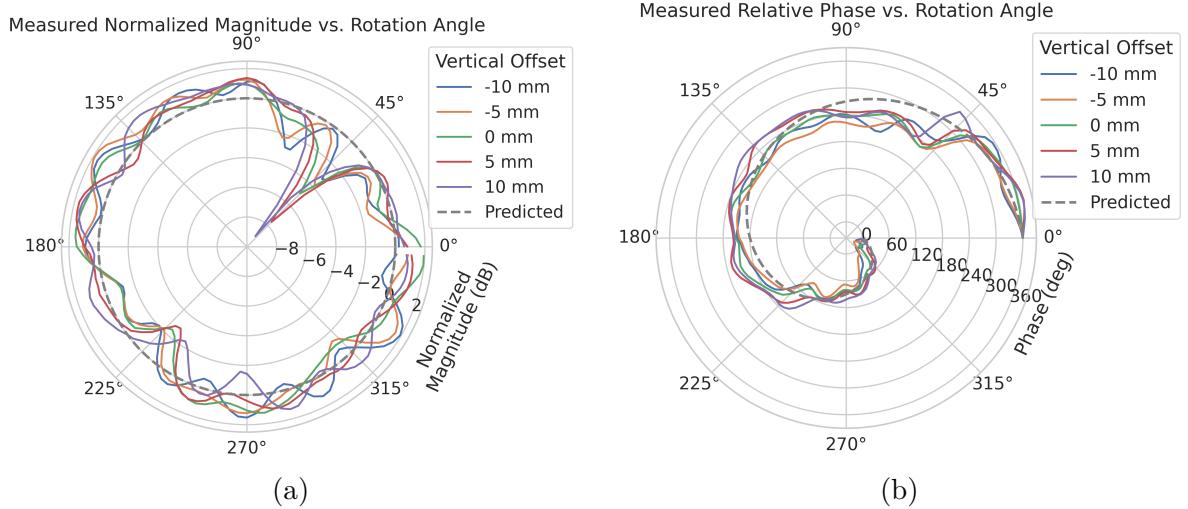


**Figure 5.19.** A schematic of the VNA characterization measurements (a). This setup allows for antenna gain and phase measurements across a full  $360^\circ$  of azimuthal angles using a motorized rotation stage and control of the radial position of the SYNCA using a translation stage. A photo of the setup in the lab is shown in (b).

4328 the antenna feed signals generated by the VNA to the SYNCA while still allowing  
 4329 unrestricted rotation. The horn antenna probe is nominally positioned in the plane  
 4330 formed by the antenna PCB ( $\theta = 90^\circ$  or  $z = 0$  mm) at a distance of 10 cm from the  
 4331 SYNCA, to match the expected position of the antenna array relative to the SYNCA in  
 4332 the antenna array test stand. The horn antenna can be manually raised or lowered to  
 4333 different relative vertical positions to characterize the radiation pattern at different polar  
 4334 angles.

4335 Several  $360^\circ$  scans were performed with probe vertical offsets of -10.0 mm, -5.0 mm,  
 4336 0.0 mm, 5.0 mm, and 10.0 mm relative to the antenna PCB plane. These probe offsets  
 4337 cover a 2 cm wide vertical region centered on the SYNCA PCB, approximately equal to  
 4338  $\pm 6$  degrees of polar angle. The measurements show that the SYNCA is generating fields  
 4339 with nearly isotropic magnitude across the probed region. The standard deviation of the  
 4340 electric field magnitude measured around the antenna circumference is approximately  
 4341 2.9 dB for a typical rotational scan. The presence of a significant pattern null is noted  
 4342 near  $45^\circ$  (see Figure 5.20), which we attribute to small imperfections in the antenna  
 4343 PCB that could be introduced from the hand soldered terminations connecting the coax  
 4344 cables to the antenna. There is no significant difference in the radiation pattern when  
 4345 measured across the 2 cm vertical range. The measured relative phases closely follow  
 4346 the expectation for an electron, being linear with the measurement rotation angle and  
 4347 forming the expected spiral pattern. Other than the small phase imperfections there is

4348 a slight sinusoidal bias to the phase data, which we determined is the result of a small  
 4349 ( $\lesssim$  1 mm) offset of the antenna's phase center from the rotation axis of the automated  
 4350 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.

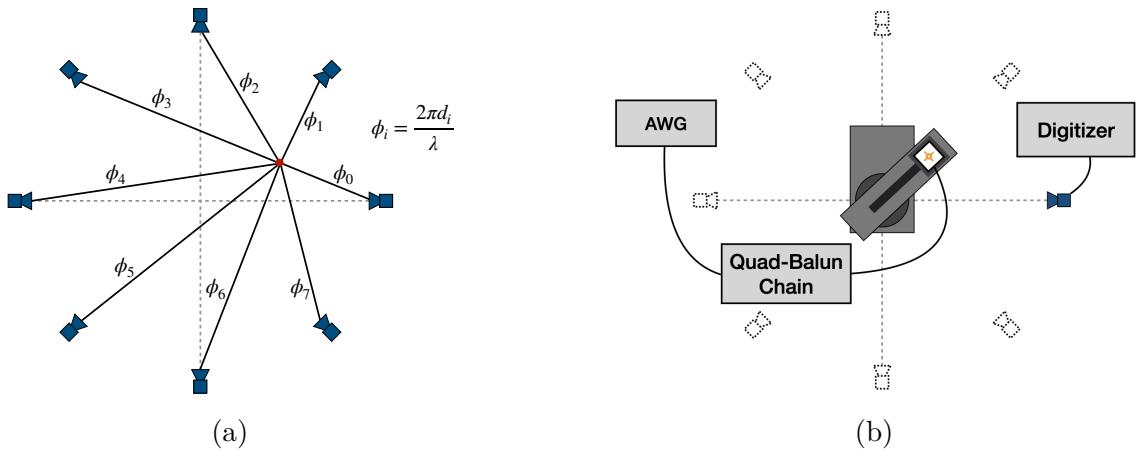
4351 The characterization measurements confirm the simulated performance of the SYNCA.  
 4352 As expected the fields generated by the antenna are nearly isotropic in magnitude,  $\phi$ -  
 4353 polarized, and are linearly out of phase around the circumference of the antenna as  
 4354 predicted for cyclotron radiation in Section 5.3.2. Small imperfections in the magnitude  
 4355 and phase of the antenna are expected, particularly at the antenna's high operating  
 4356 frequency of 26 GHz where small geometric changes can have significant impacts on  
 4357 electrical properties. However, calibration through careful characterization measurements  
 4358 can be used to remove the majority of these pattern imperfections, including the relatively  
 4359 large pattern null near 45°, which will allow for the usage of the SYNCA as a test source  
 4360 for free-space CRES experiments utilizing antenna arrays. In the next section we use the  
 4361 VNA measurements obtained here as a calibration for signal reconstruction using digital  
 4362 beamforming.

### 4363 5.3.5 Beamforming Measurements with the SYNCA

4364 Digital beamforming is a standard technique for signal reconstruction using a phased  
 4365 array [95]. The SYNCA, since it exhibits the same cyclotron phases as a trapped electron,  
 4366 can be used to perform simulated CRES digital beamforming reconstruction experiments  
 4367 on the bench-top without the need for the magnet, cryogenics, and vacuum systems  
 4368 required by a full CRES experiment. The fields received by the individual elements  
 4369 of the antenna array will have phases dependent on the spatial position of the source  
 4370 relative to the antennas. Therefore, a simple summation of the received signals will fail  
 4371 to reconstruct the signal due to destructive interference between the individual channels  
 4372 in the array. However, applying a phase shift associated with the source's spatial position  
 4373 removes phase differences and results in a constructive summation of the channel signals  
 4374 (see Figure 5.21). We can summarize the digital beamforming operation succinctly using  
 4375 the following equation

$$y[t_n] = \sum_{m=0}^{N-1} x_m[t_n] A_m e^{i\phi_m}, \quad (5.35)$$

4376 where  $y[t_n]$  represents the summed array signal at time  $t_n$ ,  $x_m[t_n]$  is the signal received  
 4377 by channel  $m$  at time  $t_n$ ,  $\phi_m$  is the phase shift applied to the signal received at channel  
 4378  $m$ , and  $A_m$  is an amplitude weighting factor that accounts for the different signal power  
 4379 received by individual channels. By changing the digital beamforming phases, the point  
 4380 of constructive interference can be scanned across the sensitive region of the array to

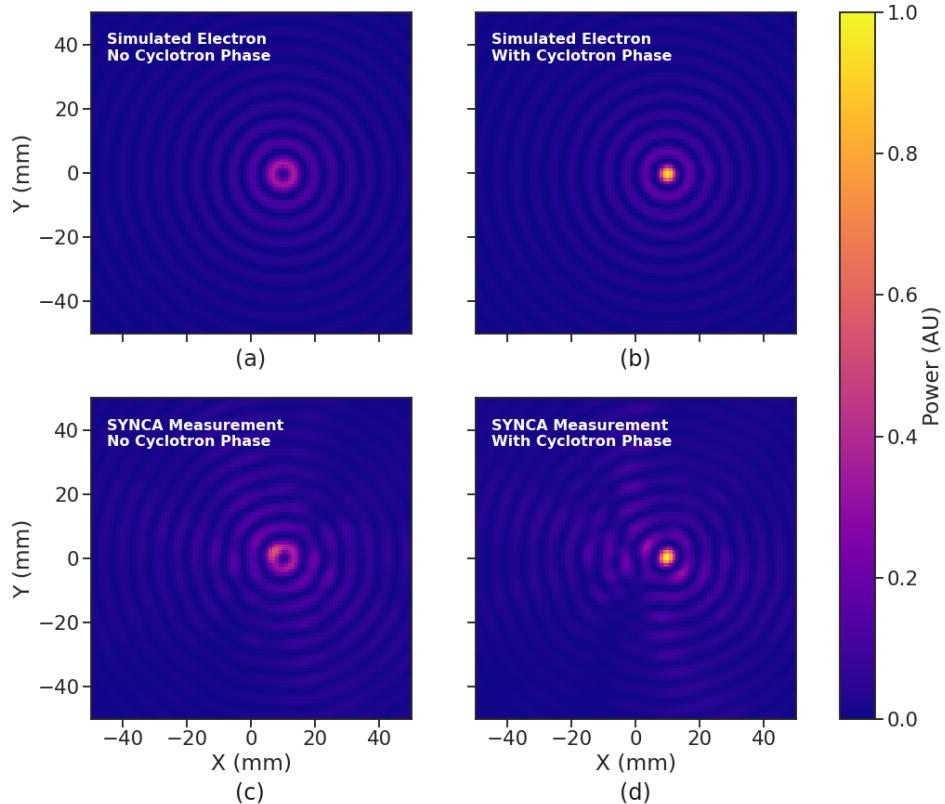


**Figure 5.21.** (a) A depiction of the relative phase differences for signals received by a circular antenna array from an isotropic source. The phases correspond to a unique spatial position. (b) A schematic of the setup used to perform digital beamforming.

4381 search for the location of a radiating source, which is identified as the point of maximum  
 4382 summed signal power above a specified threshold. The digital beamforming phases consist  
 4383 of two components,

$$\phi_m = 2\pi d_m/\lambda + \theta_m, \quad (5.36)$$

4384 where  $d_m$  is the distance from the  $m$ -th array element to the source, and  $\theta_m$  is the  
 4385 relative angle between the source position and the  $m$ -th antenna. The first component is  
 4386 the standard digital beamforming phase that corresponds to the spatial position of the  
 4387 source, and the second component is the cyclotron phase that corresponds to the relative  
 4388 azimuthal phase offset.



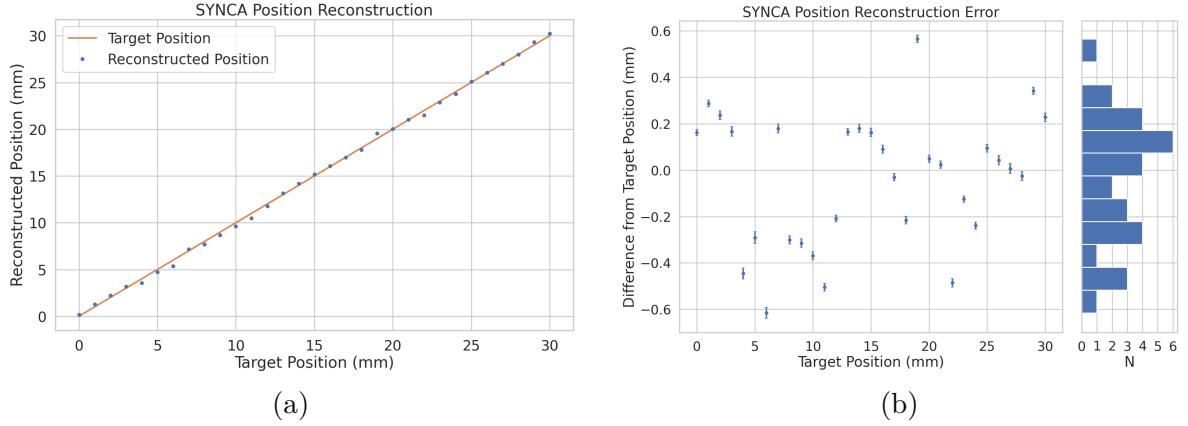
**Figure 5.22.** Digital beamforming maps generated using a simulated 60 channel array and electron simulated using the Locust package. (a) and (b) show the beamforming maps for simulated electrons without the cyclotron spiral phases and with the cyclotron spiral phases respectively. (c) and (d) show the beamforming maps produced from SYNCA measurements. We observe good agreement between simulated electrons and the SYNCA measurements.

4389 With a small modification to the hardware used to characterize the SYNCA (see  
 4390 Figure 5.19), we can perform a digital beamforming reconstruction of a synthetic CRES

4391 event. By replacing the VNA with an arbitrary waveform generator (AWG), the SYNCA  
4392 can be used to generate cyclotron radiation with an arbitrary signal structure, which  
4393 can then be detected by digitizing the signals received by the horn antenna. Rotational  
4394 symmetry allows us to use the rotational stage of the positioning system to rotate the  
4395 SYNCA to recreate the signals that would have been received by a complete circular  
4396 array of antennas.

4397 Using this setup, signals from a 60 channel circular array of equally spaced horn  
4398 antennas were generated with the SYNCA positioned 10 mm off the central array axis,  
4399 reconstructed using digital beamforming, and compared to Locust simulation (see Figure  
4400 5.22). When the cyclotron spiral phases are not used, which is equivalent to setting  $\theta_m$   
4401 in Equation 5.36 to zero, the SYNCA's position is reconstructed as a relatively faint ring  
4402 as predicted by simulation. However, when the appropriate cyclotron phases are used  
4403 during the beamforming procedure, both the simulated electron and the SYNCA appear  
4404 as a single peak of high relative power corresponding to the source position. Therefore,  
4405 we observe good agreement between the simulated and SYNCA reconstructions. While it  
4406 may seem that for the case with no cyclotron phase corrections the ring reconstructs the  
4407 position of the electron as effectively as beamforming with the cyclotron phase corrections,  
4408 it is important to note that the simulations and measurements were generated without a  
4409 realistic level of thermal noise. The larger maxima region and lower signal power, which  
4410 occurs without the cyclotron phase corrections, significantly reduce the probability of  
4411 detecting an electron in a realistic noise background.

4412 To bound the beamforming capabilities of the synthetic array of horn antennas, we  
4413 performed a series of beamforming reconstructions where the SYNCA was progressively  
4414 moved off the central axis of the array (see Figure 5.23). To extract an estimate of the  
4415 position of the SYNCA using the digital beamforming image we apply a 2-dimensional  
4416 (2D) Gaussian fit to the image data and extract the estimated centroid value. We find  
4417 that the synthetic horn antenna array reconstructs the position of the SYNCA with a  
4418  $1\sigma$ -error of 0.3 mm with no apparent trend across the 30 mm measurement range. This  
4419 reconstruction error is an order of magnitude larger than mean fit position uncertainty  
4420 of 0.02 mm indicating that systematic effects related to the SYNCA positioning system  
4421 could be contributing additional uncertainty to the measurements. Note that the current  
4422 mean reconstruction error of 0.3 mm is a factor of 20 smaller than the full width at half  
4423 maximum of the digital beamforming peak (6 mm), which could be interpreted as a naive  
4424 estimate of the position reconstruction performance of this technique. Because these  
4425 measurements are intended as a proof-of-principle demonstration, we do not investigate



**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.

4426 potential sources of systematic errors further; however, we expect that a similar and  
 4427 more thorough investigation will be performed using the Project 8 antenna array test  
 4428 stand, where typical reconstruction errors can be used to estimate the energy resolution  
 4429 limits of antenna array designs.

### 4430 5.3.6 Conclusions

4431 In this paper we have introduced the SYNCA, which is a novel antenna design that  
 4432 emits radiation that mimics the unique properties of the cyclotron radiation generated by  
 4433 charged particles moving in a magnetic field. The characterization measurements of the  
 4434 SYNCA validated the simulated performance of the PCB crossed-dipole antenna design.  
 4435 Additionally, the SYNCA was used to estimate the position reconstruction capabilities  
 4436 of a synthesized array of horn antennas and experimentally reproduced the simulated  
 4437 digital beamforming reconstruction of electrons.

4438 While the SYNCA performs well, there exist discrepancies in the phase and magnitude  
 4439 of the radiation pattern compared to the simulated SYNCA design that are related to  
 4440 the small geometric differences in the soldered connections. Future design iterations that  
 4441 replace the soldered connections with a fully surface mount design could improve the

4442 radiation pattern at the cost of some complexity and expense. Furthermore, improving  
4443 the design of the antenna PCB and mounting system would allow the antenna to be  
4444 inserted into a cryogenic and vacuum environment where in-situ antenna measurement  
4445 calibrations could be performed.

4446 The discrepancies in the radiation pattern and phases exhibited by the as-built  
4447 SYNCA should not greatly impact its performance as a calibration probe. Both magni-  
4448 tude and phase variations can be accounted by applying the SYNCA characterization  
4449 measurements as a calibration to the data collected by the antenna array test stand. The  
4450 separate calibration of the SYNCA radiation does not impact the primary goals for the  
4451 antenna array test stand which are array calibration and signal reconstruction algorithm  
4452 performance characterization, because it can be performed with standard reference horn  
4453 antennas with well understood characteristics.

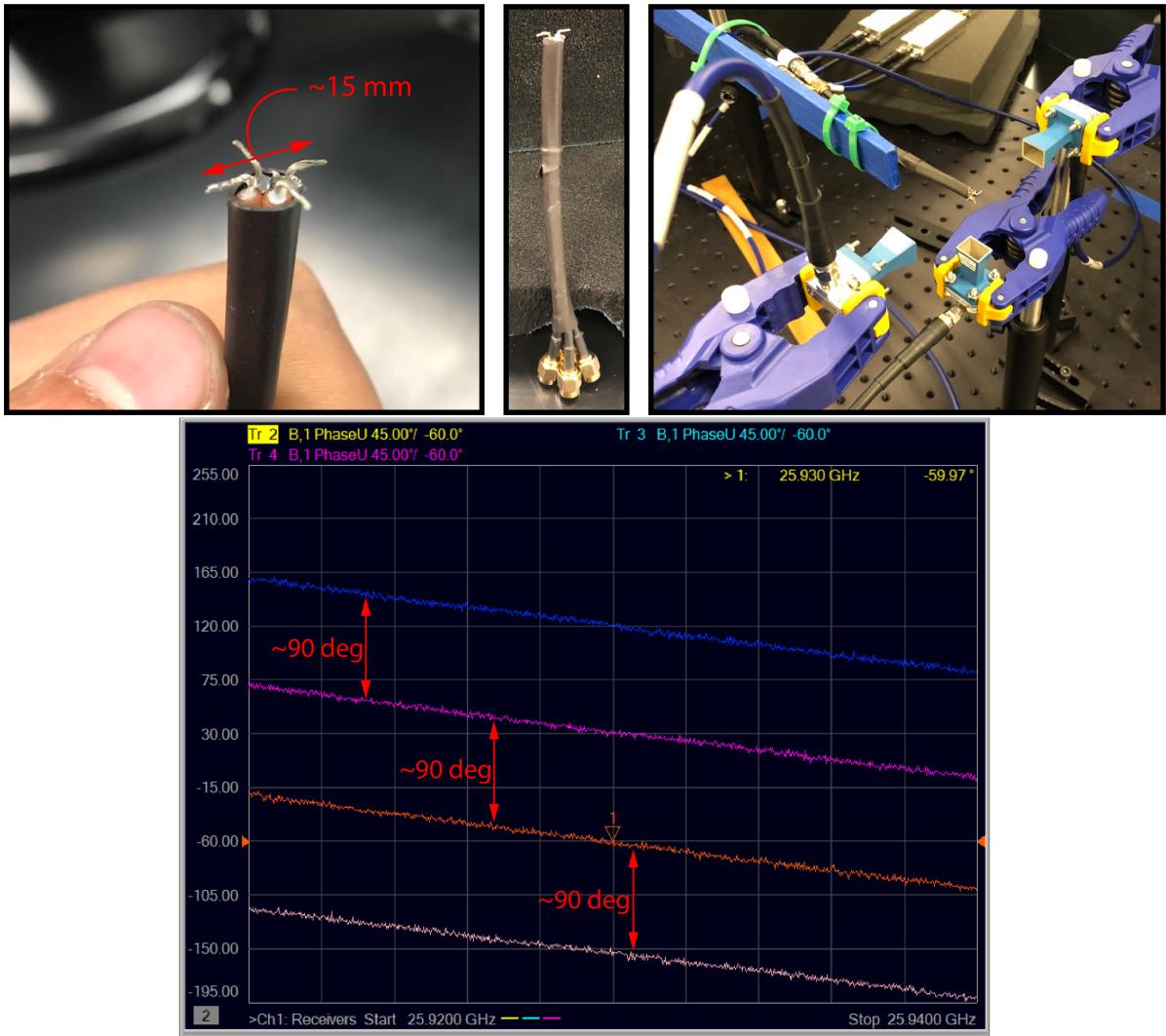
4454 The SYNCA antenna technology advances the CRES technique by providing a  
4455 mechanism to characterize free-space antenna arrays for CRES measurements without  
4456 the need for a magnet and cryogenics system, which would be required for calibration  
4457 using electron sources. Both the Project 8 collaboration as well as future collaborations  
4458 which are developing antenna array based CRES experiments can make use of SYNCA  
4459 antennas as an important component of their calibration and commissioning phases.

## 4460 5.4 SYNCA Development Discussion

4461 A crossed-dipole antenna (see Figure 5.24) was identified early on as a candidate SYNCA  
4462 design. The crossed-dipole is a circularly polarized antenna, consequently, the electric  
4463 fields measured in the plane of the dipole antenna exhibit the same relative phase offsets  
4464 as a  $90^\circ$  electron in a magnetic trap. This is explained in greater detail in Section 5.3.  
4465 These phase offsets were measured with the first rudimentary crossed-dipole prototype  
4466 manufactured from coaxial cables with the insulation and shield stripped away.

4467 Because the SYNCA is ultimately a calibration tool, it is desireable that the antenna  
4468 have a well-characterized and robust antenna pattern. Therefore, manufacturing a  
4469 SYNCA using the stripped wire method shown in Figure 5.24 is infeasible. Studies of  
4470 crossed-dipole antennas manufactured out of printed circuit boards were performed using  
4471 HFSS to identify an antenna design that imitated an electron, while being more robust  
4472 and simpler to manufacture (see Figure 5.25).

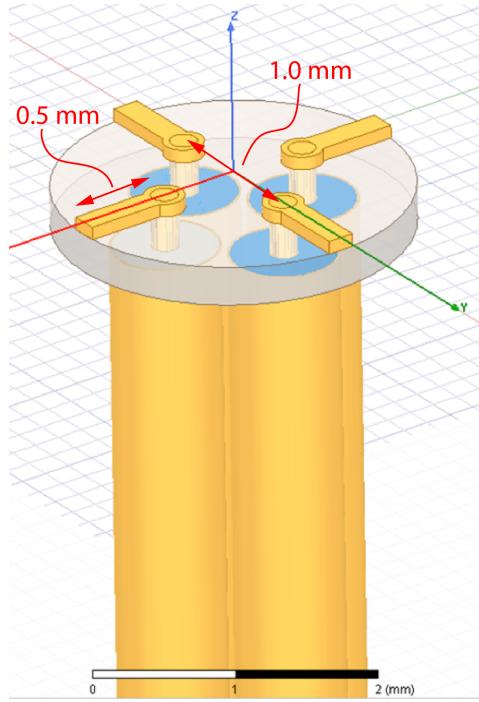
4473 Identifying a design that was robust, manufacturable, and most importantly matched  
4474 the electric fields of a trapped electron proved to be a non-trivial task. The primary



**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.

factor driving the difficulty was the high operating frequency of the antenna (26 GHz) combined with the requirement that the antenna be electrically-small. An antenna that is electrically-small at 26 GHz has a largest dimension on the order of 1 mm, which poses significant manufacturability challenges given the limited available budget for SYCNA fabrication.

One of the key limitations with the small size requirements is the diameter of the



**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.

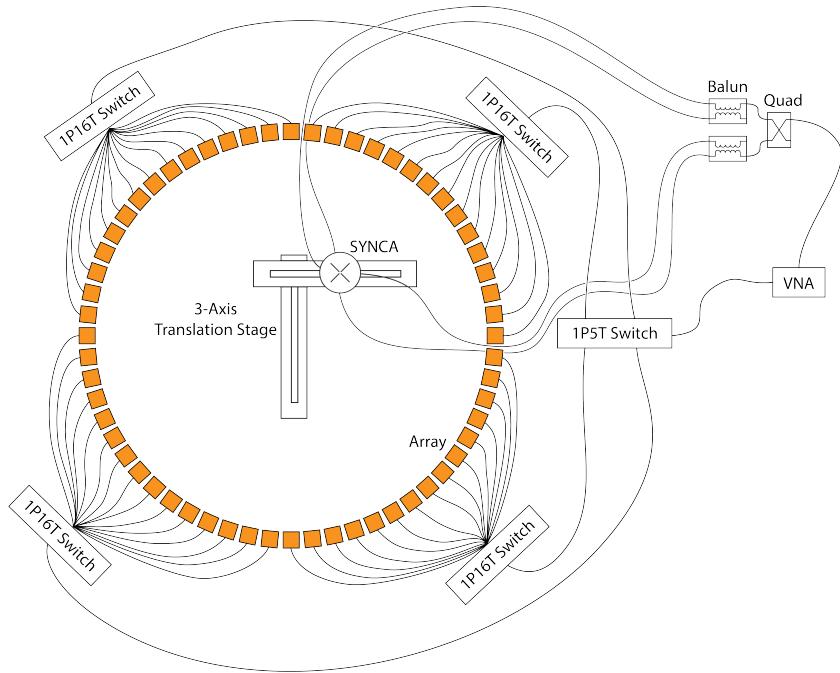
4481 coaxial cables needed to feed the crossed-dipole antenna. The smallest commonly available  
 4482 rigid coaxial cables available on the market have diameters of approximately 0.5 mm,  
 4483 which limited the spacing between dipole arms to a minimum of about 1 mm. The  
 4484 crossed-dipole antenna performs better as a SYNCA if the dipole arm separation is  
 4485 significantly less than the operating wavelength. Therefore, the high operating frequency  
 4486 ultimately limited how well the SYNCA could mimic an electron. If the desired cyclotron  
 4487 frequency was lowered by an order of magnitude to approximately 3 GHz a significantly  
 4488 higher quality SYNCA could be manufactured at lower cost.

4489 The decision to use coaxial transmission lines terminated on the antenna PCB with a  
 4490 hand-soldered connection was driven primarily to limit the costs of SYNCA development  
 4491 and contributed to the observable variations in the SYNCA's gain and phase patterns.  
 4492 A second iteration of the SYNCA design that minimized hand-soldering by using surface-  
 4493 mount components could significantly reduce variations in the antenna pattern. The  
 4494 major drawback in the development of a surface-mount SYNCA is the cost, and given the

4495 transition to a cavity based design for Phase IV, such a design was never investigated.

## 4496 5.5 FSCD Antenna Array Measurements with the SYNCA

### 4497 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.

4498 Using the SYNCA we can perform full-array measurements of prototype versions  
4499 of the FSCD antenna array to test its performance with a realistic cyclotron radiation  
4500 source (see Figure 5.26). The goal is to check how the measured power received by  
4501 the array compares to FSCD simulations as a function of the radial and axial position  
4502 of the SYNCA. These measurements are intended to validate the antenna research  
4503 and development by Project 8, which has been driven primarily by simulations with  
4504 Locust [62] and CRESana (see Section 4.2.3), and identify any discrepancies with these  
4505 simulations tools. This knowledge will provide confidence in the simulations necessary  
4506 for the analysis of the sensitivity of larger antenna array based CRES experiment designs

4507 to the neutrino mass.

4508 As shown in Section 5.3, the SYNCA does have some radiation pattern imperfections  
4509 that complicate the comparison between measurement and simulation data. One way to  
4510 disentangle some of the effects of these imperfections is to perform an additional set of  
4511 measurements using a synthetic antenna array setup along with the SYNCA antenna.  
4512 Since the synthetic array setup uses only a single array antenna, the data should be  
4513 free of errors associated with individual antenna differences and multi-path interference,  
4514 which are two error sources being tested with the full-array setup. By comparing the  
4515 synthetic array data to the FSCD array data and to simulation data one can evaluate the  
4516 significance of these effects relative to the errors introduced by SYNCA imperfections.

## 4517 **5.5.2 Measurement Setups**

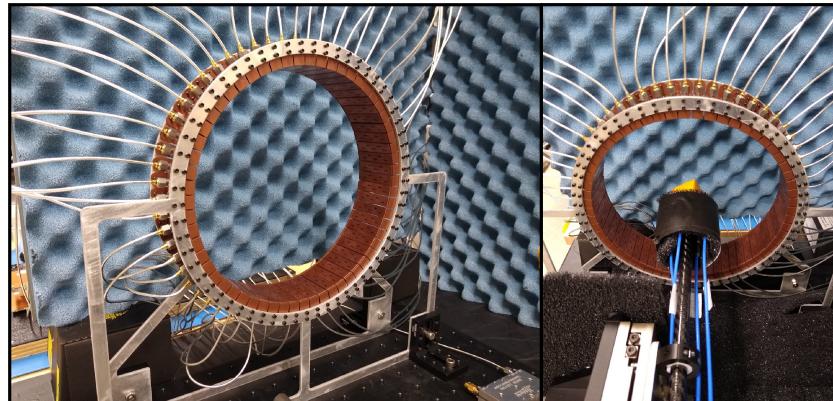
### 4518 **5.5.2.1 FSCD Array Setup**

4519 The antenna design that composes the array is the 5-slot waveguide antenna developed  
4520 for the FSCD experiment (see Figure 5.27a). The antenna is 5 cm long and is constructed  
4521 out of WR-34 waveguide with a 2.92 mm coax connector located at the center of the  
4522 antenna. Copper flanges located on both ends of the antenna are used to mount the  
4523 antenna in the array support structure. The antennas are supported by two circular steel  
4524 brackets that can be bolted to both ends of the waveguide to construct the circular array  
4525 (see Figure 5.27b). The antenna array consists of sixty identical waveguide antennas  
4526 with a radius of 10 cm. The array is mounted perpendicular to an optical breadboard  
4527 surface using a pair of the steel brackets, which provide sufficient space for the coaxial  
4528 cable connections and allows for easy positioning of the SYNCA antenna. The SYNCA is  
4529 mounted on the end of a carbon fiber rod attached to a set of manual translation stages,  
4530 which are used to move the SYNCA antenna to different positions inside the array (see  
4531 Figure 5.27c). The stages allow for independent motion in three different axes and can  
4532 position the SYNCA at radial distances up to 5 cm from the center.

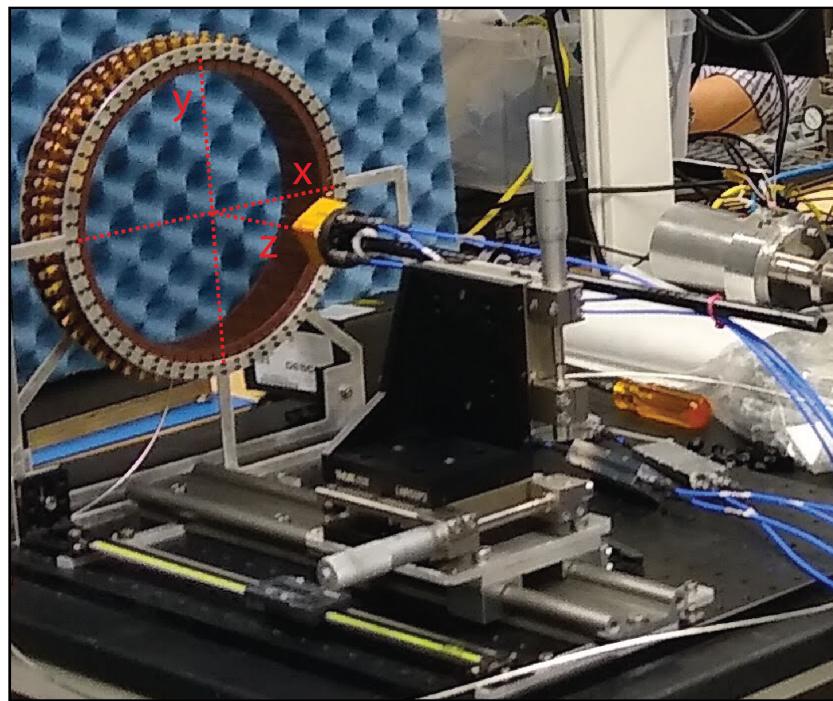
4533 Data acquisition is accomplished using a two-port VNA in combination with a series  
4534 of microwave switches that allow the VNA to connect to each channel in the array . The  
4535 first port of the VNA is connected to the quad-balun chain used to feed the SYNCA (see  
4536 Section 5.3), and the second port of the VNA connects to a 1P5T microwave switch. The  
4537 1P5T switch is connected to four separate 1P16T switch boards that connect directly  
4538 to the array. The data acquisition is controlled by a python script running on a lab  
4539 computer, which is connected to the VNA and an Arduino board programmed to control



(a)



(b)



(c)

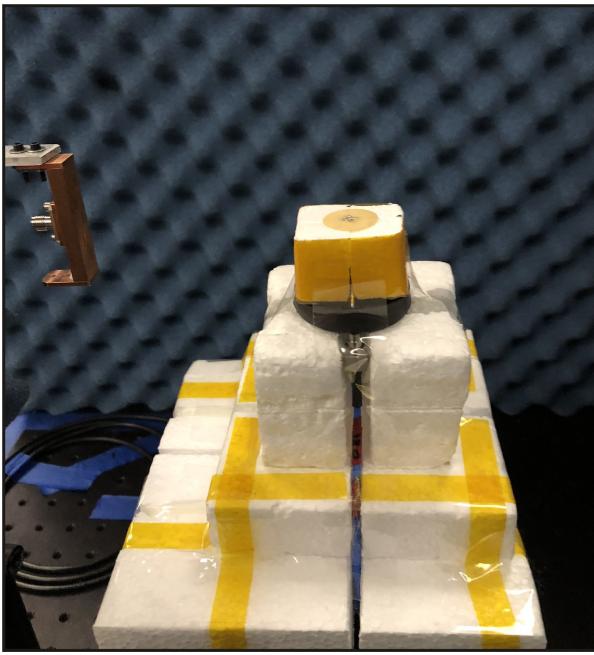
**Figure 5.27.** Photos of the prototype FSCD antenna (a), the FSCD array and SYNCA (b), and the translation stages and coordinate system used to position the SYNCA (c).

4540 the microwave switches. The script uses the switches to iteratively connect each of the  
4541 antennas in the array to the VNA. The VNA is configured to load a specific calibration  
4542 file for each antenna channel and performs the measurements of all available S-parameters.  
4543 The separate calibration files is an attempt to remove phase and magnitude errors caused  
4544 by different propagation through the RF switches. Array measurements were performed  
4545 for the set of SYNCA positions consisting of radial (x-axis) positions from 0 to 50 mm in  
4546 5 mm steps and axial (z-axis) positions from 0 to 50 mm in 5 mm steps resulting in 121  
4547 array measurements. At each SYNCA position we measured the two-port S-parameter  
4548 matrix using a linear frequency sweep from 25.1 to 26.1 GHz with 101 discrete frequencies.

4549 **5.5.2.2 Synthetic Array Setup**

4550 A photograph of the setup used to perform the synthetic array measurements is shown  
4551 in Figure 5.28. One important difference between this setup and the FSCD array setup  
4552 is that the synthetic array measurements were performed with a waveform generator and  
4553 digitizer instead of a VNA. The electronics configuration is identical to the diagram in  
4554 Figure 5.7b. Despite the differences, one is still able to compare the measured phases of  
4555 the synthetic array and the relative magnitude of the power, since the digitized signal  
4556 power is directly proportional to S21.

4557 The arbitrary waveform generator in the setup is configured to produce a 64 MHz  
4558 sine wave signal that is up-converted to 25.864 GHz using a mixer and the VNA source.  
4559 This signal is passed through a bandpass filter and fed to the SYNCA quad-balun chain.  
4560 A single FSCD antenna is positioned 10 cm from the SYNCA and aligned vertically so  
4561 that the center of the 5-slot waveguide is in the plane of the SYNCA PCB (see Figure  
4562 5.28). This position corresponds to  $z = 0$  in Figure 5.27c. The SYNCA is rotated  
4563 in three degree steps to synthesize an antenna array with 120 channels. This channel  
4564 count is more than could physically fit in a 10 cm radius array, but there is no cost to  
4565 over-sampling. Additionally, over-sampling allows for a check of the smoothness of the  
4566 antenna array radiation pattern. The signals from the FSCD antenna are down-converted  
4567 using the second mixer connected to the VNA source before being digitized at 250 MHz  
4568 and saved to disk. Several synthetic array measurement scans were performed by using  
4569 the linear translation stage to change the radial position of the SYNCA. In total eight  
4570 scans were 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.

### 4571    5.5.3 Simulations, Analysis, and Results

4572    The Locust and CRESana simulation packages utilize the antenna transfer functions  
4573    to calculate the power that would be received by each antenna from a CRES electron.  
4574    The equivalent quantity in the measurement setup is the S21 matrix element, which  
4575    indicates the ratio of the power received by an antenna in the array to the amount of  
4576    power delivered to the SYNCA. Therefore, the analysis focuses on comparing the relative  
4577    magnitudes and phase of the S21 parameters measured by the VNA as a function of  
4578    the array channel and the SYNCA position. Additionally, we apply a beamforming  
4579    reconstruction to the S21 data to evaluate how the summed power and beamforming  
4580    images change as a function of the position of the SYNCA.

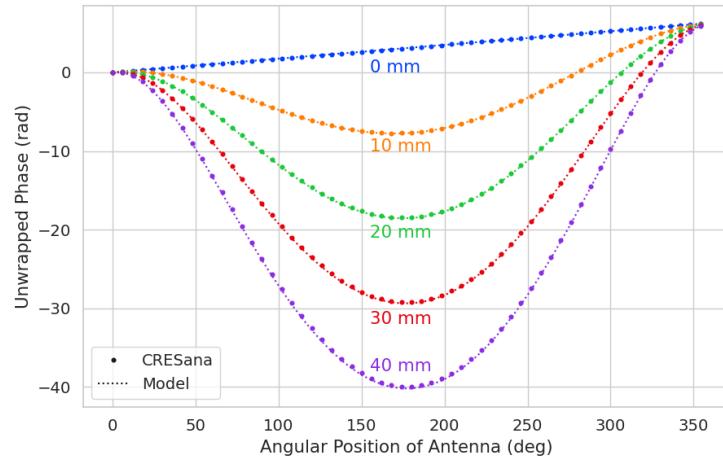
#### 4581    5.5.3.1 Simulations

4582    Simulations for the FSCD array measurements were performed using CRESana, which  
4583    performs analytical calculations of the EM-fields produced by an electron at the position  
4584    of the antennas. At each sampled time CRESana computes the electric field vector at the  
4585    antenna positions, which is projected onto the antenna polarization axis to obtain the  
4586    co-polar electric field. The magnitude of the co-polar electric field is then multiplied by  
4587    a flat antenna transfer function to calculate the corresponding voltage signal. CRESana

4588 simulations exploit the flat transfer functions of the FSCD antennas, which allows the  
 4589 electric field to be multiplied by the antenna transfer function rather than performing  
 4590 the full FIR calculation. These calculations produce a voltage time-series for each of the  
 4591 antennas in the array that can be compared to the laboratory measurements.

4592 CRESana was configured to simulate a  $90^\circ$  electron in a constant background magnetic  
 4593 field of  $\approx 0.958$  T with a kinetic energy of 18.6 keV. These parameters were chosen  
 4594 in order to mimic a CRES event near the tritium beta-decay spectrum endpoint in  
 4595 the FSCD experiment. The constant background magnetic field guarantees that the  
 4596 guiding center of the electron is stationary across the duration of the simulation which is  
 4597 consistent with the SYNCA in the laboratory measurements. Simulations were performed  
 4598 with the electron's guiding center at radial positions from 0 to 45 mm in steps of 1 mm  
 4599 and axial positions from 0 to 30 mm in steps of 1 mm. The simulations generated time  
 4600 series consisting of 8192 samples at 200 MHz for the sixty channel FSCD antenna array  
 4601 geometry.

#### 4602 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  $90^\circ$  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.

4603 Correct modeling of the signal phases is fundamental to reconstruction for both  
 4604 beamforming and matched filter approaches. The beamforming reconstruction relies on  
 4605 a signal phase model developed from Locust simulations, which allows one to predict the  
 4606 relative signal phases for a specific magnetic trap and electron position. The equation

4607 for the model is

$$\phi_{ij}(t) = \frac{2\pi d_{ij}(t)}{\lambda} + \theta_{ij}(t), \quad (5.37)$$

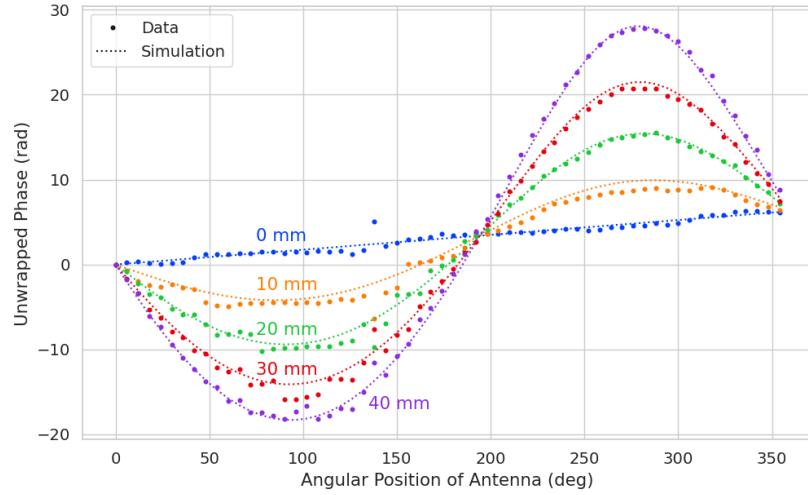
4608 where  $d_{ij}(t)$  is distance between the assumed electron position and the antenna position,  
4609 and  $\theta_{ij}(t)$  is the angular separation between the electron and antenna positions. For  
4610 details on the components of the phase model see Section 5.3.2. In Figure 5.29 we  
4611 compare the phases predicted by Equation 5.37 to phases extracted from CREsana  
4612 simulations of an electron located in the plane of the antenna array at a series of radial  
4613 positions. One observes excellent agreement between the model and simulation.

4614 The measured signal phases from the FSCD array and synthetic array are shown  
4615 in Figures 5.30a and 5.30b compared to the signal phase model. The axial position of  
4616 the SYNCA in both plots is  $z = 0$  mm, such that the plane of the PCB is aligned with  
4617 the center of the FSCD antenna. The data shown in Figure 5.30a corresponds to the  
4618 S-parameters measured at 25.80 GHz which is the frequency closest to the one used in  
4619 the synthetic array setup. The different slope and sinusoidal phases exhibited by Figure  
4620 5.30a and 5.30b reflects differences in the coordinate system for each setup. In general,  
4621 we see that the phase model predicts the large scale features of the phases quite well,  
4622 but there are some small scale deviations or errors from the phase model that do not  
4623 appear to be present in simulation.

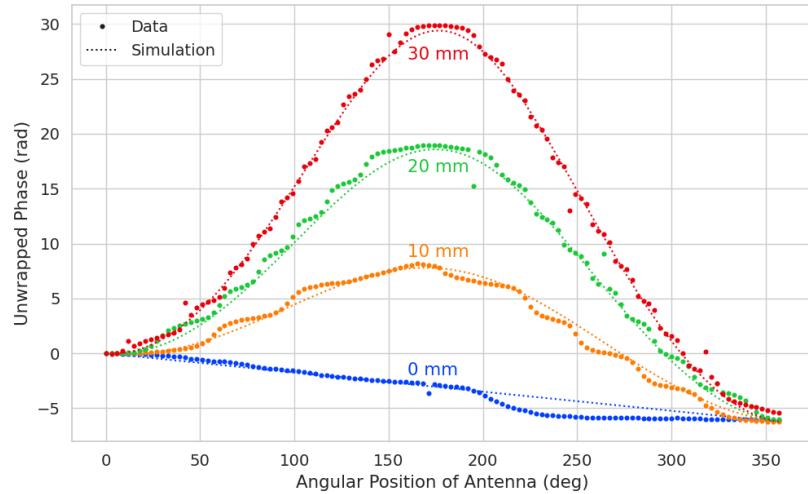
4624 A comparison of the phase errors, which are the difference between measurement and  
4625 model is shown in Figure 5.31. The FSCD array data is referred to as the JUGAAD  
4626 data in the plot legend, which is an alternative name for the FSCD array setup.

4627 The phase error at  $R = 0$  in Figure 5.31 forms a smooth curve, with the exception of  
4628 an outlier data point caused by a bug in the data acquisition script. One can attribute  
4629 the observed phase error at this position to imperfections in the antenna pattern of the  
4630 SYNCA. As the SYNCA is moved away from  $R = 0$  mm one observes that the phase  
4631 error exhibits oscillations whose frequency increases as a function of the radial position  
4632 of the SYNCA. These oscillations have the appearance of a diffraction pattern, which  
4633 is particularly clear for the radii  $\geq 15$  mm, due to the bilateral symmetry of the phase  
4634 error peaks around  $180^\circ$ .

4635 One can observe a higher average variance in the phase errors measured for the FSCD  
4636 array compared to the synthetic array. This is best seen by comparing the curves at  
4637  $R \leq 15$  mm where the smooth synthetic array curves are distinct from the relatively  
4638 noisy FSCD array errors. The extra noise in the FSCD array is most likely caused by  
4639 differences in the radiation patterns of the antennas that make up the array as well as  
4640 differences in the transmission lines through the switch network that introduce additional



(a)

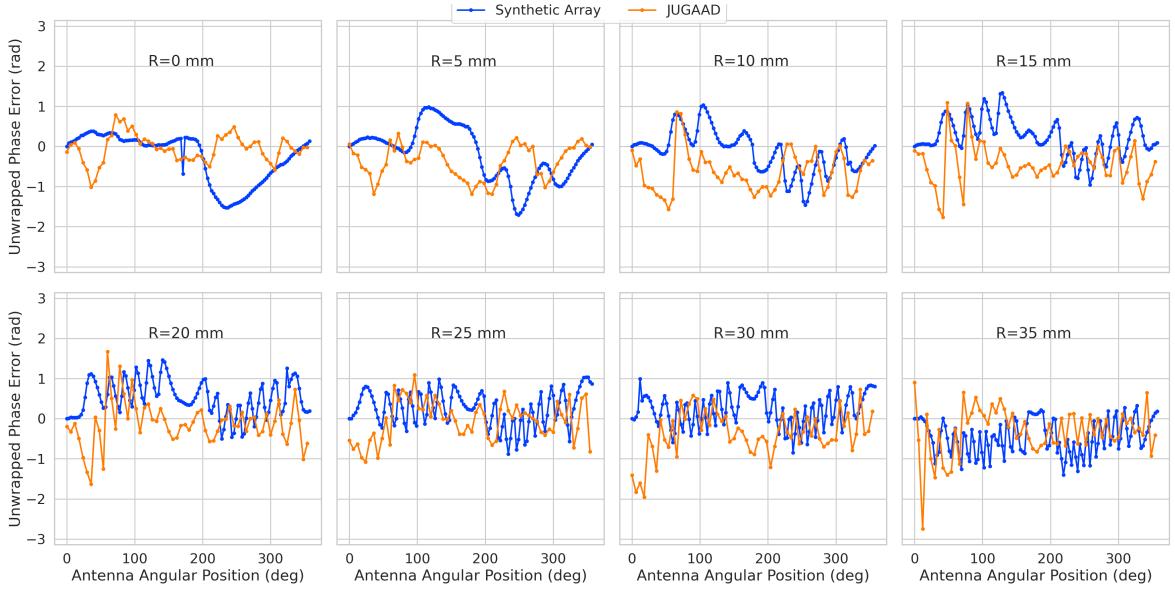


(b)

**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

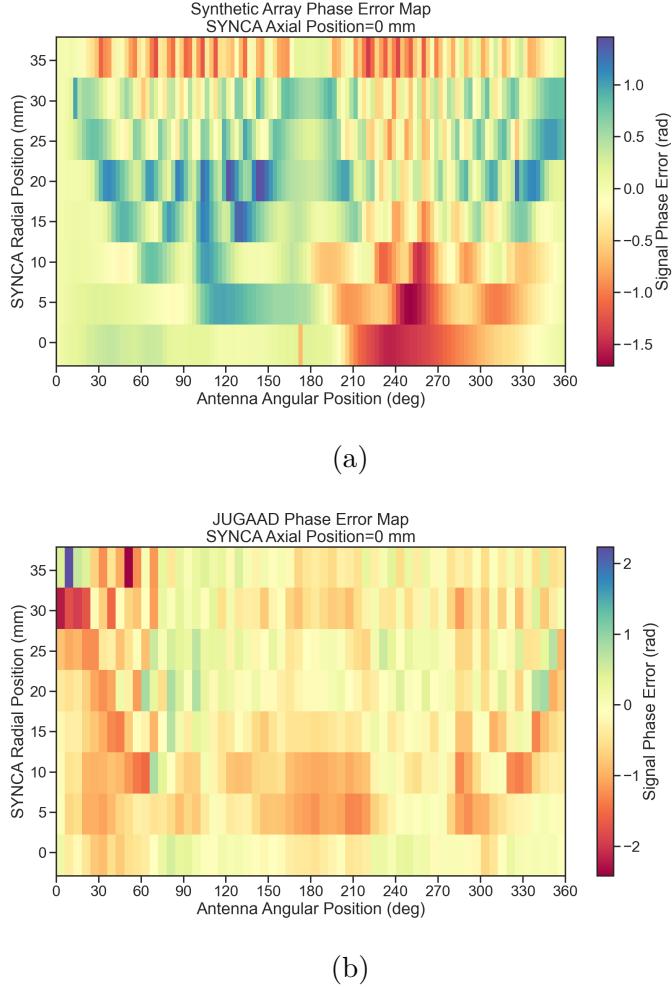


**Figure 5.31.** The phase errors between the measurement and model for the synthetic array (blue) and the FSCD array (orange) for a series of radial positions. The label JUGAAD refers to an alternative name for the FSCD array setup. As the SYNCA is translated off-axis phase errors with progressively higher oscillation frequency enter into the measurements.

4647 by plotting the phase errors in a two-dimensional map, which is done in Figures 5.32a and  
 4648 5.32b. For the synthetic array data ones observes a relatively clear diffraction pattern  
 4649 that emerges as the SYNCA is moved radially. The bilateral symmetry of the diffraction  
 4650 patterns is due to the bilateral symmetry of the circular synthetic array around the  
 4651 translation axis of the SYNCA. A similar pattern is also visible in the FSCD array data,  
 4652 although, it is obscured by the additional phase error that results from the multi-channel  
 4653 array.

4654 The physical origin of the phase error diffraction pattern is attributed to interference  
 4655 effects arising from path-length differences between the individual slots in the FSCD  
 4656 antenna and the SYNCA transmitter. Since we are operating in the radiative near-field of  
 4657 the FSCD antenna, the path length differences between the slots introduces a significant  
 4658 change in the summation of the signals that occurs inside the waveguide, which causes  
 4659 the radiation pattern of the antenna to change as a function of distance. Therefore, when  
 4660 the SYNCA is positioned off-axis the different path-lengths from the SYNCA to each  
 4661 antenna results in different radiation patterns leading to the observed diffraction pattern.

4662 This near-field effect is not present in simulations, because in order to simplify the  
 4663 calculations we assume that the far-field approximation can be applied to the FSCD  
 4664 antennas. This means that the radiation pattern and antenna transfer functions are



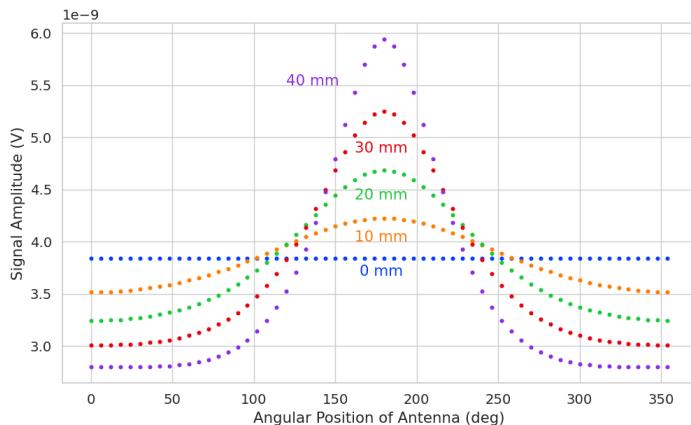
**Figure 5.32.** Two dimensional plots of the phase errors for the synthetic array (a) and the FSCD (JUGAAD) array (b). In both plots we observe evidence of a similar diffraction pattern with bilateral symmetry, but the FSCD array measurements have an additional phase error contribution from the different antennas and paths through the switch network.

independent of the distance between the transmitter and the receiving antenna. In principle, we can account for these near-field effects 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 we briefly discuss the impact of these near-field effects on the measured magnitude of the signals.

### 4671 5.5.3.3 Magnitude Analysis

4672 Exactly as for the signal phase, one can use simulations to construct a model that  
4673 describes the magnitude of the signals received by each channel in the antenna array.  
4674 By examining the results of simulations or by analyzing the Liénard-Wiechert equation  
4675 one can show that radiation pattern from a  $90^\circ$  pitch angle electron in a magnetic field  
4676 is omni-directional. Therefore the relative magnitudes of the signals received by each  
4677 channel will be determined by the free-space power loss, which is proportional to the  
4678 inverse distance between the assumed electron position and the antenna.

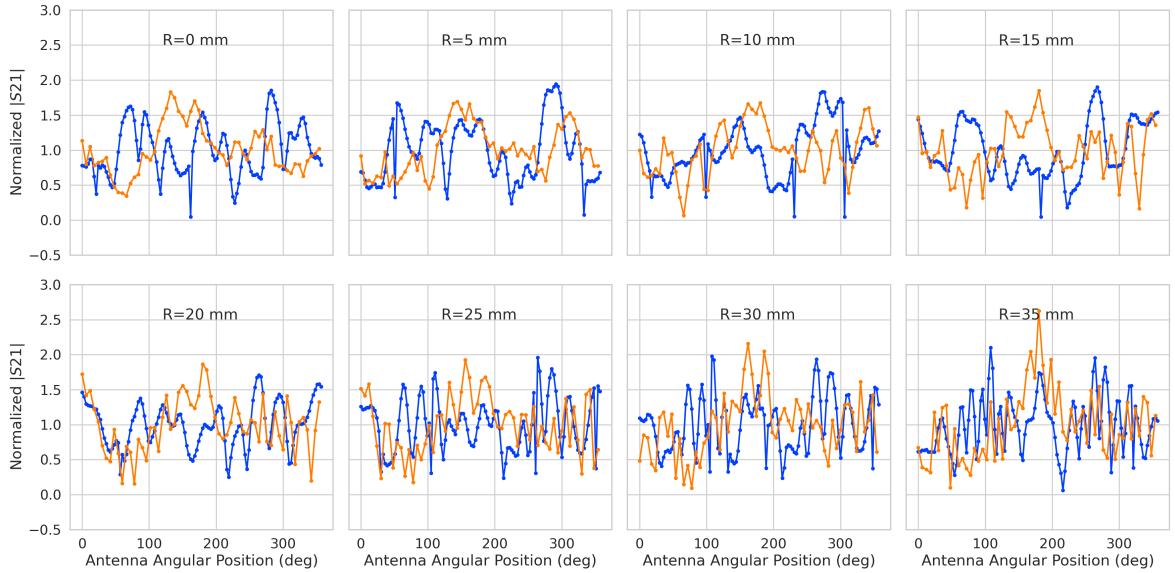
4679 A consequence of this is that the signals produced in the array for electrons off the  
4680 central axis will have larger amplitudes for the antennas closer to the electron compared  
4681 to those which are further away. The amplitudes of the signals received by the array  
from an electron located 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  $90^\circ$  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.

4682  
4683 One expects to see a similar trend in the signal magnitudes in both the FSCD and  
4684 synthetic arrays. The normalized signal magnitudes extracted from the full and synthetic  
4685 array setups for a series of radial SYNCA positions are shown in Figure 5.34. The data  
4686 corresponds to a SYNCA axial position of  $z = 0$  mm and at a frequency 25.86 GHz. One  
4687 complication is that the radiation pattern of the SYNCA is not perfectly omni-directional,  
4688 which causes the measured magnitudes at  $R = 0$  mm to diverge from the perfectly flat  
4689 behavior exhibited by electrons.

4690 As the SYNCA is moved off-axis one observes a similar increase in the number of  
4691 magnitude peaks in the synthetic array data that one would expect from a diffraction

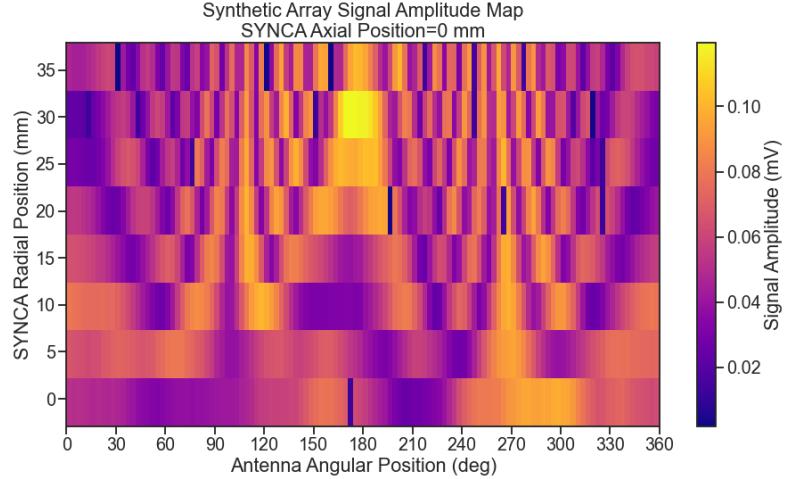


**Figure 5.34.** The normalized magnitudes of the S21 parameters measured in the FSCD (orange) and synthetic array (blue) setups. The dominant observed behavior as a function of radius is the increase in the number of magnitude peaks, which was noted in the phase error curves. There does not appear to be a strong change in the relative amplitude of a group of antennas as predicted by CRESana.

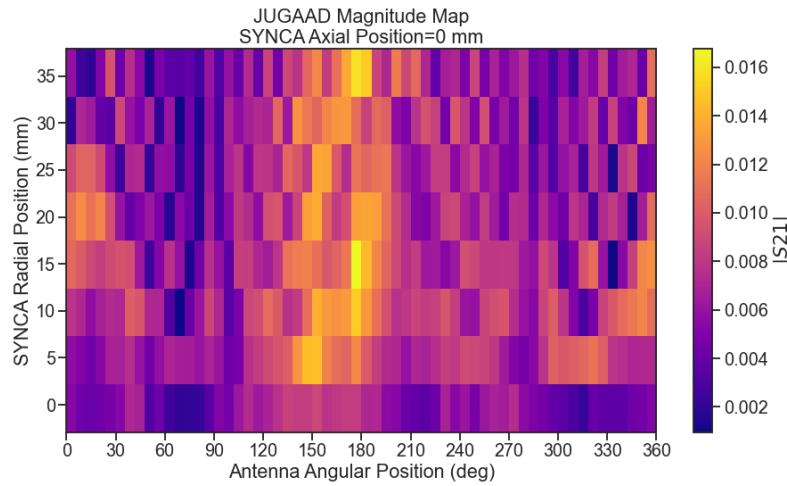
pattern, although this trend is not as stark compared to the phase data. Noticeably, there does not appear to be a set of channels with disproportionately larger amplitude at large  $R$ , which would be expected based on the trends from CRESana.

Comparing the magnitudes of the synthetic array to the FSCD array in Figure 5.34 we see that there is 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. We observe a similar trend 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



(a)



(b) The two-dimensional maps showing the diffractive pattern exhibited by the FSCD and synthetic array signal magnitudes.

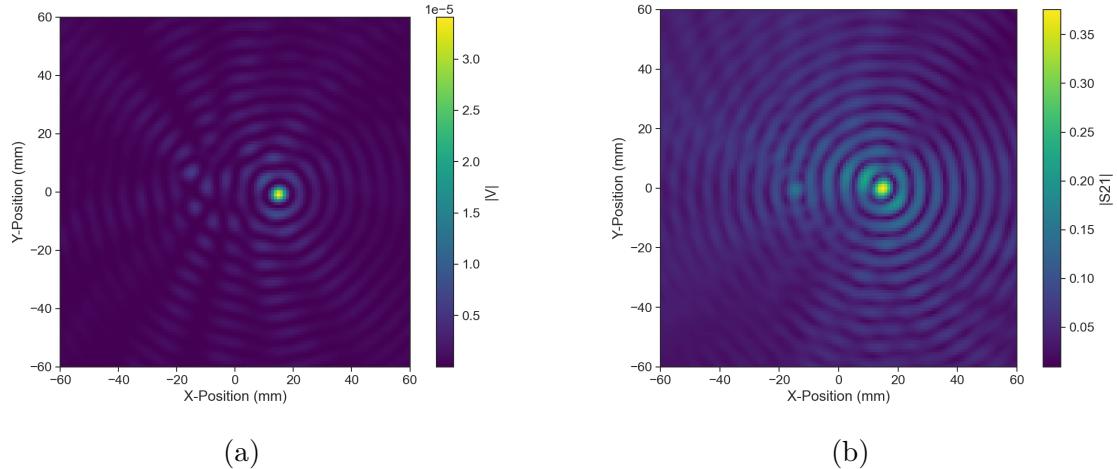
**Figure 5.35.**

<sup>4709</sup> filter reconstruction due to phase mismatch. We explore the impact of these phase and  
<sup>4710</sup> magnitude errors on beamforming in the next section.

#### <sup>4711</sup> 5.5.3.4 Beamforming Characterization

<sup>4712</sup> Errors in the signal magnitudes and phases lead to errors in signal reconstruction. For  
<sup>4713</sup> example, a matched filter reconstruction requires accurate knowledge of the signals in  
<sup>4714</sup> each channel to achieve optimal performance. Uncorrected errors leads to mismatches  
<sup>4715</sup> between the template and signal, which reduces detection efficiency and introduces

uncertainty in the parameter estimation. In this section, we analyze the beamformed signal amplitude as a function of the position of the SYNCA to quantify the impact of the phase and magnitude errors on signal reconstruction. Because of the imperfections in the SYNCA source, it is inappropriate to directly compare the beamformed signal amplitude of the FSCD array or synthetic array. Such a comparison would not allow one to disentangle losses that occur because of the antenna array from those that occur because of the source. Therefore, we focus on comparing the beamforming of the FSCD array to the synthetic array.

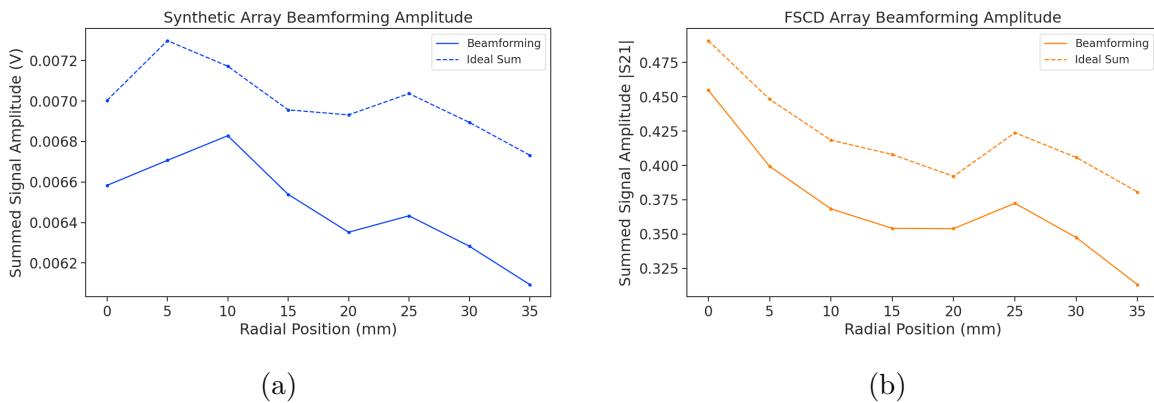


**Figure 5.36.** Beamforming images from the synthetic array (a) and FSCD array (b) setups with the SYNCA positioned 15 mm off the central axis. In both images we see a clear maxima that corresponds to the true SYNCA position. However, in the FSCD array there is an additional faint peak located at the 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.

The first method of comparison is to analyze the images generated by applying the beamforming reconstruction specified in Section 4.3.1 to the FSCD and synthetic array data (see Figure 5.36). The beamforming grid consisting of a square  $121 \times 121$  grid spanning a range of -60-mm to 60 mm in the x and y dimensions. The beamforming images formed from the synthetic array produces a three-dimensional matrix where each grid position contains a summed time series. A single beamforming image is formed from this data matrix by taking the mean over the time dimension. In the case of the FSCD array, the VNA generates frequency domain data such that each grid position contains a summed frequency series produced by the VNA sweep. For this data a single image is formed by averaging in the frequency domain.

There is a clear difference between the synthetic and FSCD array beamforming images,

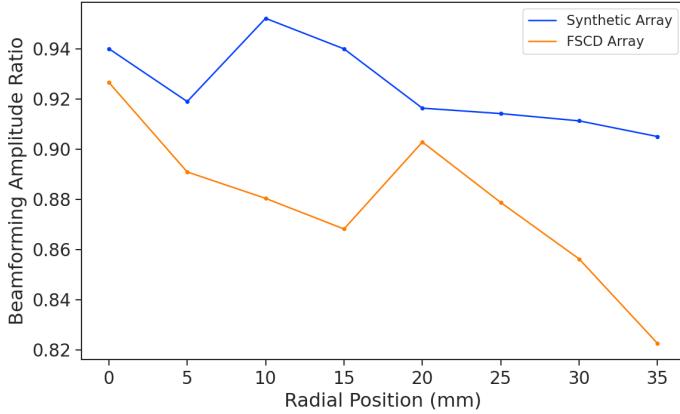
4735 which is the additional faint beamforming maxima located directly opposite the maxima  
 4736 corresponding to the SYNCA position. The images in Figure 5.36 were generated with  
 4737 data collected at a SYNCA radial position of 15 mm, which agrees well with the observed  
 4738 beamforming maximum in both images. We observe that the faint beamforming peak is  
 4739 located directly opposite of the true beamforming maximum similar to a mirror image.  
 4740 Therefore, the origin of this additional feature appears to be reflections between the two  
 4741 sides of the circular antenna array that are not present for the synthetic array since only  
 4742 a single 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 we observe that the signal amplitudes obtained from beamforming are smaller than the signal amplitude that could be attained with the ideal summation without phase mismatch.

4743 From the beamforming images we extract the maximum amplitude, which we plot  
 4744 as a function of the radial position of the SYNCA (see Figure 5.37). The phase errors  
 4745 we observed in the FSCD and synthetic arrays leads to power loss at the beamforming  
 4746 stage due to phase mismatches between the signals at different channels. This power  
 4747 loss can be quantified by comparing the signal amplitude obtained from beamforming to  
 4748 the amplitude which would be obtained from an ideal summation. We perform the ideal  
 4749 summation by phase shifting each array channel to the same phase and then summing.  
 4750 The comparison between the beamforming and ideal sums is shown in Figure 5.37, where  
 4751 we observe that both the synthetic and FSCD arrays experience power losses from the  
 4752 beamforming summation.

4753 The beamforming power loss can be quantified using the ratio of the beamforming to  
 4754 ideal signal amplitudes. Computing this ratio as a function of SYNCA radial position



**Figure 5.38.** The ratio of the beamforming signal amplitude to the ideal signal amplitude for the FSCD and synthetic arrays. We see that the FSCD array has a larger power loss from phase error compare to the synthetic array which indicates that calibration errors associated with the multiple channels as well as reflections are impacting the signal reconstruction.

radius for the FSCD and synthetic arrays we find that the FSCD array has a uniformly smaller beamforming amplitude ratio, which means that the FSCD array has a larger 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, we estimate that if no effort is made to correct these phase errors in an FSCD-like experiment, then we 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

<sup>4775</sup> kinetic energy by contributing to errors in the estimation of the electron's cyclotron  
<sup>4776</sup> frequency.

<sup>4777</sup> If these reconstruction losses prove unacceptable there are steps that can be taken  
<sup>4778</sup> to mitigate their effects. Some examples include the development of a more accurate  
<sup>4779</sup> antenna simulation approach that can reproduce the observed near-field interference  
<sup>4780</sup> patterns of the FSCD antennas and the implementation of a calibration approach that  
<sup>4781</sup> allows for the relative phase delays of the array to be measured without changing or  
<sup>4782</sup> 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.

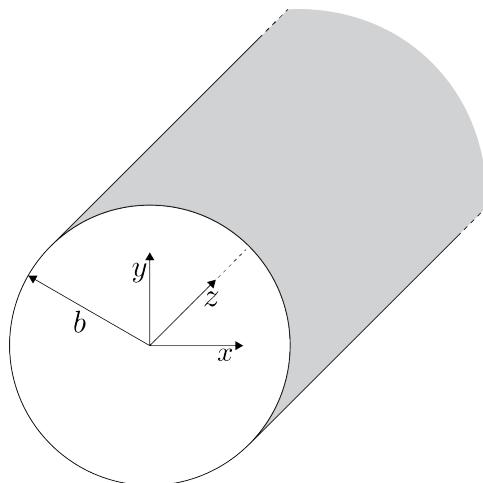
Finally, in Sections 6.4 and 6.5, I present the design and development of an open mode-filtered cavity that could be used in a cavity based CRES experiment with atomic tritium. The results of the cavity simulations are confirmed by laboratory measurements of a proof-of-principle prototype that demonstrates key features of the design.

## 6.2 Cylindrical Resonant Cavities

Resonant cavities are sealed conductive containers, which allows us to describe the electromagnetic (EM) fields contained in the cavity volume as a superposition of resonant modes [87]. The field shapes of the resonant modes are determined by Maxwell's equations and the boundary conditions enforced by the cavity geometry. Of interest to Project 8 for CRES measurements are cylindrical cavities due to their ease of construction and integration with atom and electron trapping magnets.

### 6.2.1 General Field Solutions

Consider a long segment of conducting material with a cylindrical cross-section (see Figure 6.1). A geometry such as this can be used as a waveguide transmission line to transfer EM energy from point to point, or, if conducting shorts are inserted on both ends of the cylinder, the waveguide becomes a resonant cavity.



**Figure 6.1.** Geometry of a cylindrical waveguide with radius  $b$ .

The fields allowed inside a cylindrical cavity are determined by the boundary conditions of the cylindrical geometry. The general approach to solving the fields begins by assuming solutions to Maxwell's equations of the form

$$\mathbf{E}(x, y, z) = (\mathbf{e}(x, y) + \hat{z}e_z(x, y))e^{-i\beta z}, \quad (6.1)$$

$$\mathbf{H}(x, y, z) = (\mathbf{h}(x, y) + \hat{z}h_z(x, y))e^{-i\beta z}. \quad (6.2)$$

The solutions assume a harmonic time dependence of the form  $e^{i\omega t}$  and propagation

4823 along the positive z-axis. The functions  $\mathbf{e}(x, y)$  and  $\mathbf{h}(x, y)$  represent the transverse  
 4824 ( $\hat{x}, \hat{y}$ ) components of the electric and magnetic fields respectively, and  $e_z(x, y)$ ,  $h_z(x, y)$   
 4825 represent the longitudinal components. The version of Maxwell's equations in the case  
 4826 where there are no source terms can be written as a pair of coupled differential equations,

$$\nabla \times \mathbf{E} = -i\omega\mu\mathbf{H}, \quad (6.3)$$

$$\nabla \times \mathbf{H} = i\omega\epsilon\mathbf{E}, \quad (6.4)$$

4827 where  $\epsilon$  and  $\mu$  are the permittivity and permeability of the material inside the waveguide  
 4828 or cavity. Using the field solutions from Equations 6.1 and 6.2 one can solve for the  
 4829 transverse components of the fields in terms of the longitudinal fields. Because we  
 4830 are interested in cylindrical cavities it is advantageous to write the field solutions in  
 4831 cylindrical coordinates. After performing this transformation the set of four equations  
 4832 for the transverse field components are,

$$H_\rho = \frac{i}{k_c^2} \left( \frac{\omega\epsilon}{\rho} \frac{\partial E_z}{\partial\phi} - \beta \frac{\partial H_z}{\partial\rho} \right), \quad (6.5)$$

$$H_\phi = \frac{-i}{k_c^2} \left( \omega\epsilon \frac{\partial E_z}{\partial\rho} + \frac{\beta}{\rho} \frac{\partial H_z}{\partial\phi} \right), \quad (6.6)$$

$$E_\rho = \frac{-i}{k_c^2} \left( \beta \frac{\partial E_z}{\partial\rho} + \frac{\omega\mu}{\rho} \frac{\partial H_z}{\partial\phi} \right), \quad (6.7)$$

$$E_\phi = \frac{i}{k_c^2} \left( -\beta \frac{\partial E_z}{\partial\phi} + \omega\mu \frac{\partial H_z}{\partial\rho} \right), \quad (6.8)$$

4833 where  $k_c$  is the cutoff wavenumber defined by  $k_c^2 = k^2 - \beta^2$  with  $k = \omega\sqrt{\mu\epsilon}$  being the  
 4834 wavenumber of the EM radiation.

4835 This set of equations can be used to solve for a variety of different modes that can be  
 4836 obtained by setting conditions on  $E_z$  and  $H_z$ . For cylindrical cavities two types of modes  
 4837 are allowed, which correspond to solutions where  $E_z = 0$  and  $H_z = 0$  respectively.

### 4838 6.2.2 TE and TM Modes

4839 The TE family of modes corresponds to the case where  $E_z = 0$ . This implies that  $H_z$  is  
 4840 a solution to the Helmholtz wave equation

$$(\nabla^2 + k^2)H_z = 0. \quad (6.9)$$

4841 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  
 4842 the standard technique of separation of variables. Rather than reproduce the derivation  
 4843 here we shall simply quote the solutions for the transverse fields [87], which are

$$H_\rho = \frac{-i\beta}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.10)$$

$$H_\phi = \frac{-i\beta n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.11)$$

$$E_\rho = \frac{-i\omega\mu n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.12)$$

$$E_\phi = \frac{i\omega\mu}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}. \quad (6.13)$$

4844 One can observe that the solutions have a periodic dependence on  $\phi$ , and radial profiles  
 4845 given by the Bessel functions of the first kind. The integer indices  $n$  and  $m$  arise from  
 4846 continuity conditions on the EM fields in the azimuthal and radial directions. For the  
 4847 TE modes  $n \geq 0$  and  $m \geq 1$ .  $k_{c_{nm}}$  is the cutoff wavenumber for the  $\text{TE}_{nm}$  mode given by

$$k_{c_{nm}} = \frac{p'_{nm}}{b}, \quad (6.14)$$

4848 where  $b$  is the radius of the cavity or waveguide and  $p'_{nm}$  is the  $m$ -th root of the derivative  
 4849 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

4850 The TM mode family corresponds to the case where  $H_z = 0$ , and  $(\nabla^2 + k^2)E_z = 0$ .  
 4851 Again, we assume solutions of the form  $E_z(\rho, \phi, z) = e_z(\rho, \phi)e^{-i\beta z}$ , for which the general  
 4852 form of the solutions is the same as for the TE modes. However, the different boundary  
 4853 conditions for the TM modes results in particular solutions with a different form, which  
 4854 we shall quote here without derivation. The transverse fields of the TM modes are given  
 4855 by

$$H_\rho = \frac{-i\omega\epsilon n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.15)$$

$$H_\phi = \frac{-i\omega\epsilon}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z} \quad (6.16)$$

$$E_\rho = \frac{-i\beta}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.17)$$

$$E_\phi = \frac{-i\beta n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.18)$$

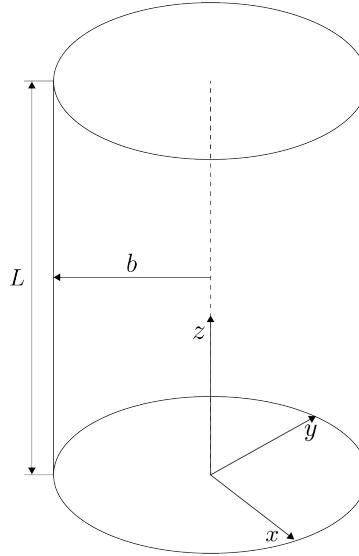
which one may notice are the same solutions as the TE modes with  $H$  and  $E$  flipped.  
 The cutoff wavenumber for the TM modes is given by,  $k_{c_{nm}} = p_{nm}/b$ , where the values of  $p_{nm}$  correspond to the  $m$ -th zero of the  $n$ -th order Bessel function (see Table 6.2).

**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

### 6.2.3 Resonant Frequencies of a Cylindrical Cavity

A cylindrical cavity is constructed by taking a section of cylindrical waveguide and shorting both ends with conductive material. This means that the electric fields inside a cylindrical cavity are exactly those we 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  $b$ .

4864

The transverse electric field solutions for a cylindrical waveguide are of the form

$$\mathbf{E}(\rho, \phi, z) = \mathbf{e}(\rho, \phi) (A_+ e^{-i\beta_{nm}z} + A_- e^{i\beta_{nm}z}), \quad (6.19)$$

4865 where  $A_+$  and  $A_-$  are arbitrary amplitudes of forward and backward propagating waves.4866 In order to enforce that  $\mathbf{E}$  is zero at both ends of the cavity we require that

$$\beta_{nm}L = 2\pi\ell, \quad (6.20)$$

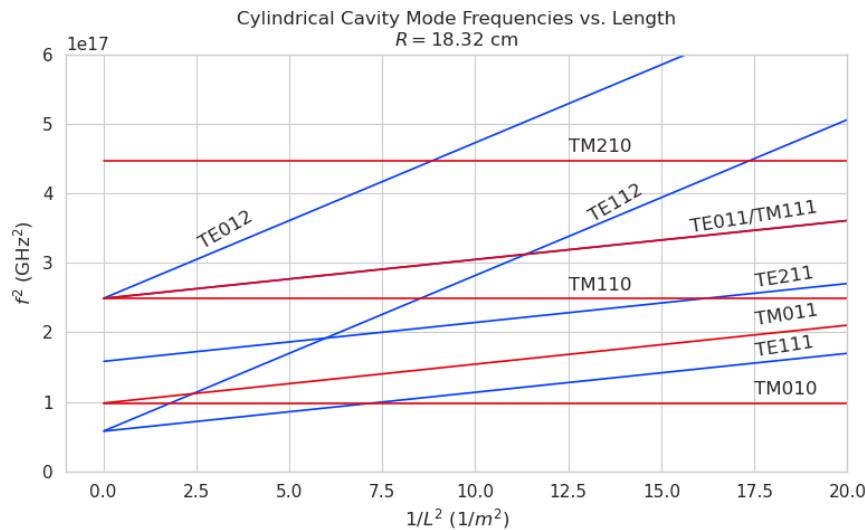
4867 where  $\ell = 0, 1, 2, 3, \dots$ . Using this constraint on the propagation constant we can solve4868 for the resonant frequencies of the  $\text{TE}_{nml}$  and the  $\text{TM}_{nml}$  modes in a cylindrical cavity.

4869 For the TE modes the resonant frequencies are

$$f_{nml} = \frac{c}{2\pi\sqrt{\mu_r\epsilon_r}} \sqrt{\left(\frac{p'_{nm}}{b}\right)^2 + \left(\frac{\ell\pi}{L}\right)^2}, \quad (6.21)$$

4870 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.

## 4871 6.2.4 Cavity Q-factors

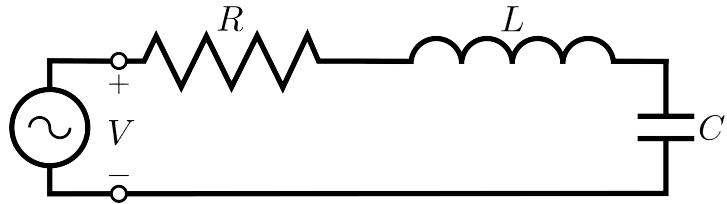


Figure 6.4. A series RLC circuit.

4872 The resonant behavior of cylindrical cavities can be modeled as a series RLC circuit  
 4873 (see figure 6.4). The input impedance of the circuit can be obtained by applying  
 4874 Kirchhoff's laws to calculate the impedance of the equivalent circuit. For a series RLC  
 4875 circuit the input impedance is

$$Z_{\text{in}} = \left( \frac{1}{R} + \frac{1}{i\omega L} + i\omega C \right). \quad (6.23)$$

4876 The resistance in the circuit represents all sources of loss in the cavity, which is primarily  
 4877 caused by the finite conductivity of the cavity walls. The inductor and capacitor represent  
 4878 the energy stored in the cavity in the form of electric and magnetic fields. If the circuit  
 4879 is being driven by an external power source we can write the input power in terms of the  
 4880 circuit input impedance and the source voltage

$$P_{\text{in}} = \frac{1}{2} Z_{\text{in}} |I|^2 = \frac{1}{2} |I|^2 \left( \frac{1}{R} + \frac{1}{i\omega L} + i\omega C \right). \quad (6.24)$$

4881 The resistor introduces a loss into the system with a power given by

$$P_{\text{loss}} = \frac{1}{2} |I|^2 R, \quad (6.25)$$

4882 and the capacitor and inductor store energies given by

$$W_e = \frac{1}{4} \frac{|I|^2}{\omega^2 C}, \quad (6.26)$$

$$W_m = \frac{1}{4} |I|^2 L, \quad (6.27)$$

4883 respectively. Using these expressions we can write the input power and input impedance

4884 expressions in terms of the lost power and stored energy

$$P_{\text{in}} = P_{\text{loss}} + 2i\omega(W_m - W_e), \quad (6.28)$$

$$Z_{\text{in}} = \frac{P_{\text{loss}} + 2i\omega(W_m - W_e)}{\frac{1}{2}|I|^2}. \quad (6.29)$$

4885 The condition for resonance in the RLC circuit is that the stored magnetic energy  
 4886 is equal to the stored electric energy ( $W_e = W_m$ ). When this occurs  $Z_{\text{in}} = R$ , which is a  
 4887 purely real impedance, and  $P_{\text{in}} = P_{\text{loss}}$ . The resonant frequency of the circuit can be  
 4888 determined from the condition  $W_e = W_m$  from which one finds that

$$\omega_0 = \frac{1}{\sqrt{LC}}. \quad (6.30)$$

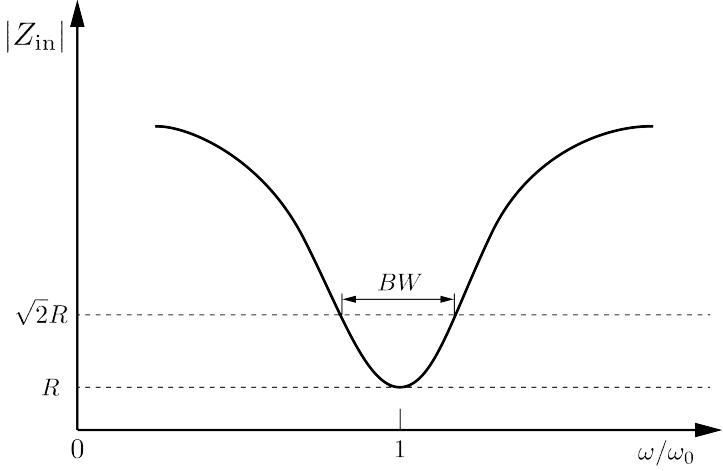
4889 An important performance parameter for any resonant system is the Q-factor, which  
 4890 quantifies the quality of the resonator as the ratio of the stored energy multiplied by the  
 4891 resonant frequency to the average energy lost per second. For the series RLC circuit, the  
 4892 Q-factor is given by the expression

$$Q_0 = \omega \frac{W_e + W_m}{P_{\text{loss}}} = \frac{1}{\omega_0 RC}, \quad (6.31)$$

4893 from which one observes that as the resistance of the RLC circuit is decreased the quality  
 4894 factor of the resonator increases. From the perspective of cylindrical cavities this implies  
 4895 that as one decreases the resistance of the cavity walls it is expected that the Q-factor of  
 4896 the cavity should increase, which is indeed the case. In certain applications where a high  
 4897 Q is desireable it is possible to manufacture a cavity out of superconducting materials in  
 4898 order to minimize the power losses of the system.

4899 The Q-factor of the resonator also determines with bandwidth (BW) of the system.  
 4900 A cavity with a high Q-factor will resonant with a smaller range of frequencies than a  
 4901 cavity with a low Q-factor. To see this we can examine the behavior of the RLC circuit  
 4902 when driven by frequencies near the resonance. For a frequency  $\omega = \omega_0 + \Delta\omega$ , where  
 4903  $\Delta\omega = \omega - \omega_0 \ll \omega_0$ , we can write the input impedance 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.

and by expanding  $(\omega^2 - \omega_0^2)/\omega^2$  to first order in  $\Delta\omega$ , we obtain

$$Z_{\text{in}} \approx R + i \frac{2RQ_0\Delta\omega}{\omega_0}. \quad (6.33)$$

Therefore, the magnitude of the input impedance near the resonance is given by

$$|Z_{\text{in}}| = R \sqrt{1 + 4Q_0^2 \frac{\Delta\omega^2}{\omega^2}}, \quad (6.34)$$

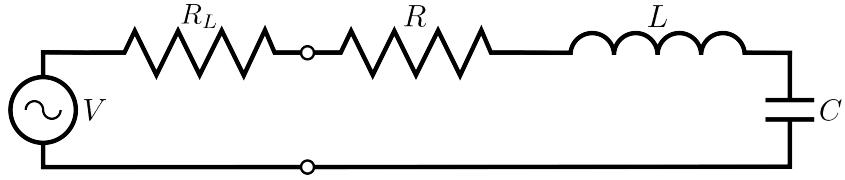
from which we observe that for the series RLC circuit the input impedance is minimized at the resonant frequency, which corresponds to the maximum input power (see Figure 6.5). The half-power BW is the range of frequencies over which the input power drops to half the input power on resonance. This occurs when  $|Z_{\text{in}}| = \sqrt{2}R$ , which corresponds to  $\Delta\omega/\omega = \text{BW}/2$ . Using Equation 6.34 one can find that

$$2R^2 = R^2(1 + Q_0^2\text{BW}^2), \quad (6.35)$$

which implies

$$\text{BW} = \frac{1}{Q_0} \quad (6.36)$$

It is important to emphasize that the Q-factor defined here,  $Q_0$ , is technically the unloaded Q. It reflects the quality of the cavity or resonant circuit without the influence 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  $R_L$ .

4915 external circuit to drive a cavity resonance or to measure the energy of a resonant mode.  
 4916 Coupling a cavity to an external circuit changes the Q by loading the equivalent cavity  
 4917 RLC circuit (see Figure 6.6). The Q-factor of the cavity when it is loaded by an external  
 4918 circuit is called the loaded Q, which is the quantity that one actually measures when  
 4919 exciting a resonance in the cavity. Using the series RLC circuit model one can see that  
 4920 the load resistor in Figure 6.6 will add in series with the resistor in the circuit for a total  
 4921 equivalent resistance of  $R + R_L$ . Therefore, the loaded Q is given by

$$Q_L = \frac{1}{\omega_0(R + R_L)C}, \quad (6.37)$$

4922 from which one observes that the loaded Q is always less than the intrinsic Q of the  
 4923 cavity.

4924 The amount of coupling that is desireable depends on the specific application of  
 4925 the resonator. If one wants a resonator that is particular frequency selective than it  
 4926 makes sense to limit the amount of coupling to the cavity to maintain a small BW,  
 4927 alternatively, if a larger BW is need one can increase the cavity coupling by tuning the  
 4928 input impedance of the external circuit. The critical point, where maximum power is  
 4929 transferred between the cavity and the external circuit, occurs when the input impedance  
 4930 of the cavity matches the input impedance of the external transmission line. For the  
 4931 series RLC circuit on resonance, this matching condition corresponds to

$$Z_0 = Z_{in} = R, \quad (6.38)$$

4932 where  $Z_0$  is the impedance of the transmission line. The loaded Q at this critical point  
 4933 is, therefore,

$$Q_L = \frac{1}{2\omega_0 Z_0 C} = \frac{Q_0}{2}. \quad (6.39)$$

4934 One can described the degree of coupling between the cavity and an external circuit by

4935 defining a coupling factor,  $g$ , such that,

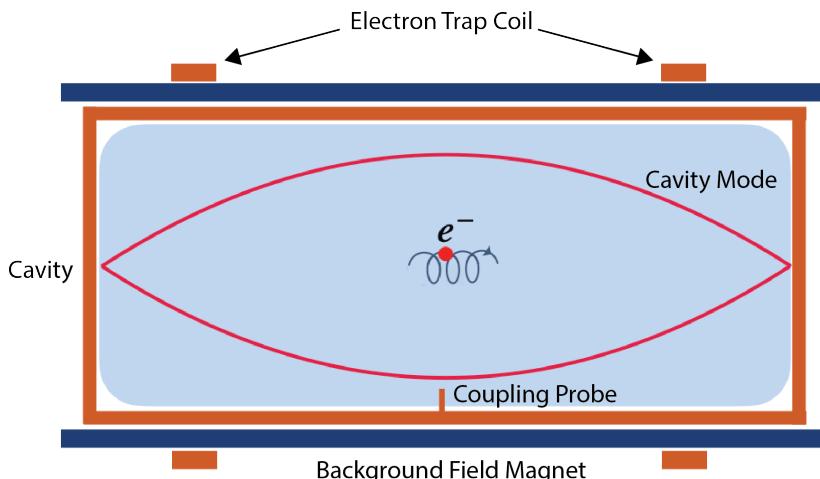
$$g = \frac{Q_0}{Q_L} - 1. \quad (6.40)$$

4936 When  $g = 1$  then  $Q_L = Q_0/2$ , and the cavity is said to be critically coupled as we  
4937 described. If  $Q_L < Q_0/2$ , then the cavity is undercoupled to the transmission line,  
4938 corresponding to  $g < 1$ . Alternatively, if  $Q_L > Q_0/2$ , then  $g > 1$ , and the cavity is  
4939 overcoupled to the transmission line. Various specialized circuits can be used to tune the  
4940 input impedance of the external circuit as seen by the cavity to achieve a wide range of  
4941 different coupling factors based on the desired application of the cavity.

## 4942 6.3 The Cavity Approach to CRES

### 4943 6.3.1 A Sketch of a Molecular Tritium Cavity CRES Experiment

4944 Resonant cavities can be used to perform CRES measurements, and they represent the  
4945 current preferred technology by the Project 8 collaboration. The basic approach to a  
4946 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.

4947

4948 At the core of the experiment is a large resonant cavity filled with tritium gas. The  
4949 filled cavity is then placed in a uniform magnetic field provided by a primary magnet  
4950 that provides the background magnetic field. The value of the background magnetic field  
4951 sets the range of cyclotron frequencies for electrons emitted near the tritium spectrum  
4952 endpoint. When a beta-decay electron is produced in the cavity it is trapped using a set  
4953 of magnetic pinch coils that keep electrons inside the cavity volume.

4954 Electrons trapped inside the cavity do not radiate in the same way as electrons  
4955 in free-space. Effectively, the same boundary conditions that were used to derive the  
4956 resonant modes of a cylindrical cavity in Section 6.2 apply to the radiation of the electron  
4957 as well. The coupling of an electron performing cyclotron motion in a cavity has been  
4958 studied in detail for measurements of the electron’s magnetic moment [96–98] If an  
4959 electron is emitted with a kinetic energy that corresponds to a cyclotron frequency that  
4960 matches a resonant frequency of the cavity, then energy radiated by the electron excites  
4961 a corresponding resonance in the cavity. The strength of the electron’s coupling to the  
4962 cavity is given to first order by the dot product between the electrons trajectory and  
4963 the electric field vector of the resonant mode. Additional effects, such as the Purcell  
4964 enhancement [99], alter the emitted power from the free-space Larmor equation [50]. If an  
4965 electron is moving with a cyclotron frequency that is far from any resonant modes in the  
4966 cavity, then radiation from the electron is suppressed. One can interpret this somewhat  
4967 surprising effect as the metallic walls of the cavity reflecting the radiated energy back to  
4968 the electron.

4969 Detecting an electron in the cavity is accomplished by coupling the cavity to an  
4970 external transmission line that leads to an amplifier and RF receiver chain [100]. The  
4971 coupling of the cavity resonance to the amplifier occurs through a coupling probe or  
4972 aperture designed to read-out the excitation of the mode(s) excited by the electron. For  
4973 CRES measurements, the placement of a wire antenna coupling probe inside the cavity  
4974 volume leads to unacceptable losses of tritium atoms due to recombination to molecular  
4975 tritium on the antenna surface, therefore, apertures are the preferred coupling method  
4976 for cavity CRES experiments.

4977 One of the attractive features of the CRES technique for neutrino mass measurement  
4978 is the gain in statistics that comes from the differential nature of the tritium spectrum  
4979 measurement. Initially, this seems incompatible with cavities, due to the narrow reso-  
4980 nances of cavity modes giving relatively small bandwidth. However, by intentionally  
4981 over-coupling to a single cavity mode one can achieve bandwidths of a few 10’s of MHz  
4982 (see Section 6.2), which is sufficient for a measurement of the tritium spectrum endpoint

4983 region.

### 4984 **6.3.2 Magnetic Field, Cavity Geometry, and Resonant Modes**

#### 4985 **Magnetic Field and Volume Scaling**

4986 For a CRES experiment, cylindrical cavities are a natural choice since they match  
4987 the geometry of standard solenoid magnets, which are needed in order to produce the  
4988 background magnetic field for CRES measurements. Furthermore, the cylindrical shape is  
4989 compatible with a Halbach array, which is the leading choice of atom trapping technology  
4990 for future atomic tritium experiments by the Project 8 collaboration. Cylindrical  
4991 cavities also benefit from well-established machining practices that are able to achieve  
4992 high geometric precision at large lengths scales. More exotic cavity designs are under-  
4993 consideration and there are on-going efforts to investigate the potential advantages these  
4994 may have over the standard cylindrical geometry.

4995 As we saw in Section 6.2, the physical dimensions of the cavity are directly coupled  
4996 to the resonant frequencies of the cavity. This dependency links the size of the cavity to  
4997 the magnitude of the background magnetic field, because the magnetic field determines  
4998 the cyclotron frequencies of trapped electrons. Specifically, as the size of the cavity is  
4999 increased to accommodate larger volumes of tritium gas, the frequencies of the resonant  
5000 modes decrease proportionally. This requires that the magnetic field also decrease in  
5001 order to maintain coupling between electrons and the desired cavity mode.

5002 The required cavity size is ultimately determined by the required statistics in the  
5003 tritium spectrum endpoint region. Because the gas density must be kept below a certain  
5004 level to ensure that electrons have sufficient time to radiate before scattering, larger  
5005 volumes become the only way to achieve higher event statistics. To achieve the sensitivity  
5006 goals of Phase III and IV cavity volumes on the order of several cubic-meters are required,  
5007 which pushes one towards frequencies in the range of 100's of MHz.

#### 5008 **Single-mode Cavity CRES**

5009 It is tempting to consider maintaining a high magnetic field, while still increasing the size  
5010 of the cavity, in order to increase the radiated power from trapped electrons for better  
5011 SNR. However, if one were to maintain the same magnetic field while increasing the  
5012 size of the cavity, the electrons would begin to couple to higher order modes with more  
5013 complicated transverse geometries. The danger with this approach is that a complicated  
5014 mode structure could introduce systematic errors into the CRES signals. Example

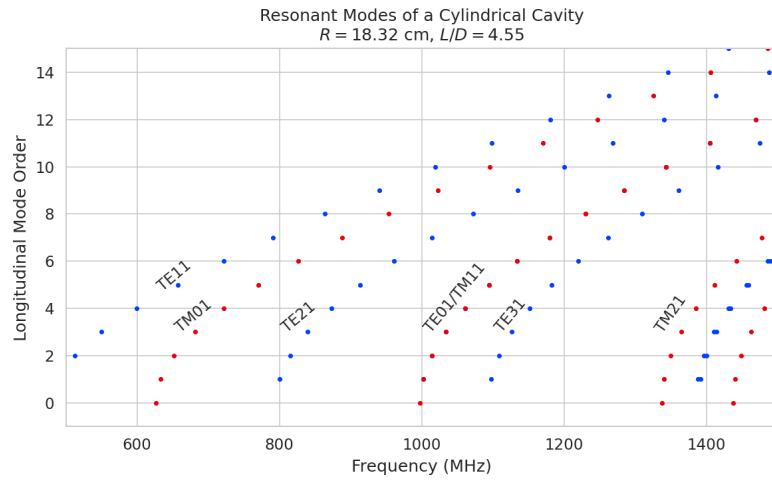
5015 systematics include unpredicted mode hybridization or changes in the mode shapes from  
5016 imperfections in the cavity construction, which would prevent reconstruction of the  
5017 electron's starting kinetic energies with adequate resolution. For this reason, it is ideal  
5018 to operate with magnetic fields that give cyclotron frequencies near the fundamental  
5019 frequency of the cavity, where the mode structure is relatively simple (see Figure 6.8).  
5020 In this frequency region it is possible to perform CRES by coupling to only a single  
5021 resonant mode, however, it is currently an open question if a single mode measurement  
5022 will provide enough information about an individual electron's position to reconstruct  
5023 the full event. Regardless, developing a solid understanding of the CRES phenomenology  
5024 when an electron is coupling to a single mode will be a necessary step towards a future  
5025 multi-mode cavity experiment.

## 5026 Considerations for Resonant Mode Selection

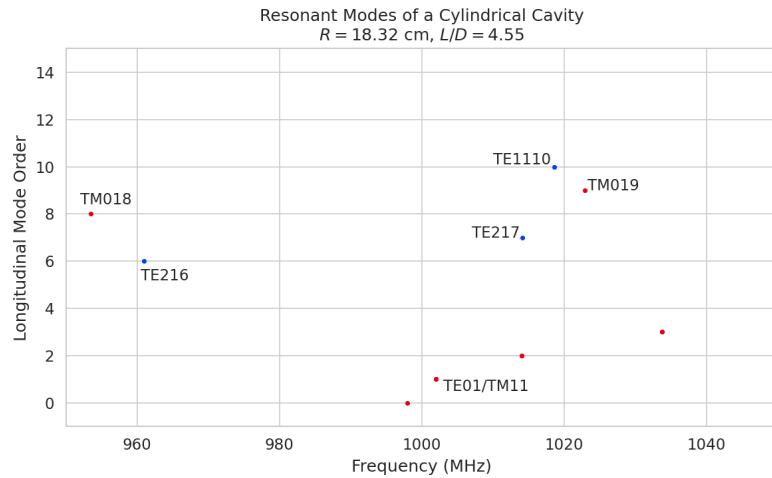
5027 A single-mode cavity experiment begs the question, which resonant mode is best for  
5028 CRES measurements? There is an immediate bias towards low order  $TE_{nm}$  and  $TM_{nm}$   
5029 modes due to the multi-mode considerations discussed above. Additionally, there is a  
5030 preference towards modes with longitudinal index  $\ell = 1$  with a single antinode along the  
5031 vertical axis of the cylindrical cavity. The reason for this is that there is a phase change  
5032 in the electric fields between antinodes that leads to modulation effects that destroy the  
5033 carrier frequency signal information.

5034 A second consideration for mode selection is the volumetric efficiency of the mode.  
5035 Volumetric efficiency can be thought of as an integral over the volume of the cavity  
5036 weighted by the relative amplitude of the mode. From the perspective of simply maximiz-  
5037 ing the volume useable for CRES measurements this integral would be as close to unity  
5038 as possible. However, there is a requirement to reconstruct the position of the electrons  
5039 inside the cavity volume so that the local magnetic fields can be used to convert the  
5040 measured cyclotron frequency to a kinetic energy. With a single mode this necessarily  
5041 requires a variable transverse mode amplitude, which lowers the volumetric efficiency, so  
5042 that position of the electron in the cavity can be estimated from the average amplitude  
5043 of the CRES signal. Longitudinal indices of  $\ell = 1$  have an advantage in volumetric  
5044 efficiency over higher order  $\ell$  modes, since there are only two longitudinal nodes, one at  
5045 each end of the cavity. Therefore, the average coupling strength of trapped electrons as  
5046 they oscillate axially is higher for  $\ell = 1$  modes.

5047 The longitudinal variation in the mode strength is ultimately critical for achieving the  
5048 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.

5049 the average magnetic fields experienced by electrons with different pitch angles requires  
 5050 that information on the axial motion of the electron be encoded into the CRE signal.  
 5051 The longitudinal variation in the mode amplitude leads to amplitude modulation of the  
 5052 CRE signal with a frequency proportional to the electron's pitch angle.

5053 An additional factor for mode selection is the intrinsic or unloaded  $Q$  of the mode. In  
 5054 terms of SNR it is advantageous to use a mode with a very high  $Q_0$ , which is then highly  
 5055 overcoupled to achieve the necessary bandwidth to cover the tritium endpoint spectrum.  
 5056 This scheme leads to a decoupling of the physical cavity temperature from the effective  
 5057 noise temperature after the amplifier, which allows us to achieve adequate SNR without

5058 the requirement of cooling the entire cavity to single Kelvin temperatures.

5059 An example of a resonant mode that exhibits these traits is the TE<sub>011</sub> mode. At present  
5060 the TE<sub>011</sub> mode is the preferred resonance for a single-mode cavity CRES experiment  
5061 by the Project 8 collaboration. TE<sub>011</sub> is a low order mode located in a region relatively  
5062 far from other cavity modes. Furthermore, the separation of the TE<sub>011</sub> mode can be  
5063 improved by various mode-filtering techniques discussed in Section 6.4.2 below. TE<sub>011</sub>  
5064 consists of a single longitudinal antinode that can provide pitch angle information in the  
5065 form of amplitude modulation, and has an electric field with a radial profile given by the  
5066  $J'_0$  Bessel function allowing for radial position estimation. Lastly, the TE<sub>011</sub> mode has a  
5067 relatively high intrinsic Q compared to nearby modes, which helps with SNR. Unloaded  
5068 Q's greater than 80000 are achievable for a 1 GHz TE<sub>011</sub> resonance using a copper walled  
5069 cavity.

### 5070 **6.3.3 Trade-offs Between the Antenna and Cavity Approaches**

5071 The choice between cavities and antennas for large-scale CRES measurements is not  
5072 without trade-offs. Both the antenna array and cavity approaches are relatively immature  
5073 techniques, at present there are no known obstacles that would prevent either approach  
5074 from being used for a large scale neutrino mass experiment. The preference for cavities  
5075 is largely driven by important practical considerations that could make a cavity based  
5076 experiment significantly cheaper than an antenna experiment of similar size and scope.  
5077 However, the switch to cavities also introduces new challenges less relevant to the  
5078 antenna array, which must be solved in order for Project 8 to achieve its neutrino mass  
5079 measurement goals.

5080 One of the major relative drawbacks of the antenna array approach is the size and  
5081 complexity of the data-acquisition system. A large-scale antenna array experiment  
5082 requires  $O(100)$  antennas independently digitized at rates of  $O(10)$  to  $O(100)$  MHz. Since  
5083 there is insufficient information in a single antenna channel to detect or reconstruct the  
5084 CRES signal, the entire array output must be processed during the signal reconstruction.  
5085 Because data storage becomes an issue with these data volumes, there is a real-time  
5086 signal reconstruction requirement that allows one to detect CRES signals buried in the  
5087 thermal noise. As we discuss in Section 4.4, the computational cost of these real-time  
5088 detection algorithms are potentially quite large for even a small scale antenna array  
5089 experiment. However, the operating principle of a cavity experiment allows the CRES  
5090 signal to be detected using only a single read-out channel digitized at rates of  $O(10)$  MHz,  
5091 which reduces the cost of the data acquisition system by many orders of magnitude.

5092 From an engineering perspective, the simple geometry and thin-walls of a cylindrical  
5093 cavity are simpler to interface with the cryogenic and magnetic subsystems needed for a  
5094 CRES experiment. Whereas, the antenna array requires careful design and engineering  
5095 to accommodate the antenna array and receiver electronics in proximity to the trapping  
5096 magnets. Additionally, due to near-field interference effects, the antenna array is unable  
5097 to reconstruct CRES events within the reactive near-field distance of the antennas.  
5098 Because atom trapping requirements require magnetic fields which correspond to cyclotron  
5099 frequencies for endpoint electrons less than 1 GHz, the required stand-off distance leads to  
5100 a significant loss in useable experiment volume, necessitating larger and more expensive  
5101 magnets.

5102 Another advantage to the cavity approach is the relatively compact sideband structure,  
5103 which is a result of the low modulation index for cavity CRES signals. The axial motion  
5104 in an antenna array experiment leads to frequency modulation and sidebands. The shape  
5105 of the sideband structure is determined by the modulation index,  $h = \frac{\Delta f}{f_a}$ , where  $\Delta f$   
5106 is the size of the frequency deviation and  $f_a$  is the axial frequency. The large electron  
5107 traps required for a cubic-meter-scale experiment leads to high modulation indices, which  
5108 causes the signal spectrum to be made up of numerous low power sidebands that make  
5109 reconstruction and detection challenging. This behavior was observed in simulations  
5110 of the FSCD in which carrier power decreased with pitch angle due to the increase in  
5111 modulation index (see Figure 4.31). For cavities, however, the modulation index remains  
5112 near  $h = 1$  even for very long magnetic traps due to the high phase velocity in cavities  
5113 relative to the axial velocity of the electron. This results in an almost ideal spectrum  
5114 shape that has a strong carrier frequency with a few sidebands whose relative amplitudes  
5115 encode pitch angle information.

5116 A downside of the cavity approach is the apparent difficulty of estimating the position  
5117 of the electron using only the coupling of the electron to a single mode. The amplitude of  
5118 the TE<sub>011</sub> mode is completely independent of the azimuthal coordinate, therefore, position  
5119 reconstruction using the TE<sub>011</sub> mode is only able to estimate the radial position of the  
5120 electron. This position degeneracy may lead to magnetic field uniformity requirements  
5121 that are too challenging to meet due to mechanical uncertainties in cavity and magnet  
5122 construction, as well as uncertainties caused by nuisance external magnetic fields such  
5123 as the Earth's field and magnetic fields from building materials. A multi-mode cavity  
5124 experiment may provide a way to extract more precise information on the position of  
5125 the electron by analyzing the coupling of the electron to several modes that overlap in  
5126 different ways.

## 5127 **6.4 Single-mode Resonant Cavity Design and Simulations**

5128 The single-mode cylindrical cavities envisioned for the Phase III and IV experiments must  
5129 be carefully engineered in order to measure the neutrino mass with the desired sensitivity.  
5130 In this section I summarize some simulation studies performed to analyze early design  
5131 concepts for a single-mode cavity. The primary tool for these investigations was Ansys  
5132 HFSS, which was also used for the development of the SYNCA antenna described in  
5133 Section 5.3.

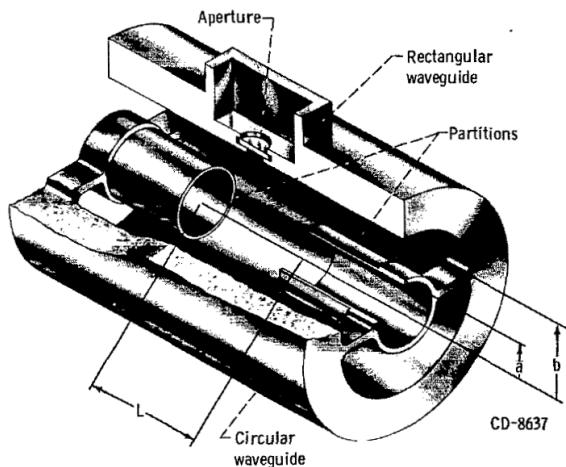
### 5134 **6.4.1 Open Cylindrical Cavities with Coaxial Terminations**

#### 5135 **Design Concept**

5136 A basic cavity design question relevant to Project 8's ultimate goal of an atomic tritium  
5137 CRES experiment is how to build a cavity that can be efficiently filled with atomic  
5138 tritium. To keep the rate of atom loss from recombination on surfaces it is ideal if the  
5139 ends of the cylindrical cavity are as open as possible so that tritium atoms can flow  
5140 inside unimpeded. Additionally, one of the primary calibration techniques planned for  
5141 future CRES experiments involves CRES measurements using electrons injected from  
5142 an electron gun source, which also requires an opening at the cavity end. Cylindrical  
5143 cavities with open ends can be manufactured, however, the intrinsic Q-factors of these  
5144 cavities are orders of magnitude less than their sealed counterparts, which reduces the  
5145 signal-to-noise ratio when that cavity is used for CRES measurement.

5146 Cylindrical cavities with mostly open ends that also exhibit Q values for the  $TE_{01\ell}$   
5147 modes similar to sealed cavities can be built by using coaxial endcaps to terminate the  
5148 cavity. Cavities of this type have been manufactured for specialized applications related  
5149 to the measurements of the dielectric constants of liquefied gasses (see Figure 6.9) [2, 3].  
5150 This cavity design leaves the ends of the cavity wide open, but retains high Q-values for  
5151 the  $TE_{01\ell}$  modes due to the coaxial endcap, which are designed to perfectly reflect the  
5152 electric fields of  $TE_{01\ell}$  modes. Coupling to the  $TE_{01\ell}$  mode is achieved via an aperture  
5153 located at the center of the cavity wall.

5154 A cavity similar to Figure 6.9 is a candidate design for the future CRES experiments  
5155 by Project 8, since it appears to elegantly solve many practical issues that arise when  
5156 combining cavity CRES and atomic tritium. The coaxial endcaps leave significant regions  
5157 of the cavity ends completely open, which allows for the entrance of atomic tritium as  
5158 well as the pumping away of molecular tritium that has recombined on the cavity walls.



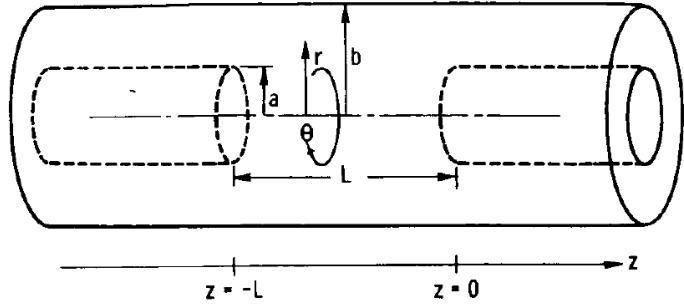
**Figure 6.9.** An image of an open cavity with coaxial terminations used for dielectric constant measurements. Figure from [2].

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 we shall analyze this cavity design in more detail, primarily by using HFSS simulations to analyze the resonant mode structure of this cavity geometry.

### Coaxial Terminator Constraints

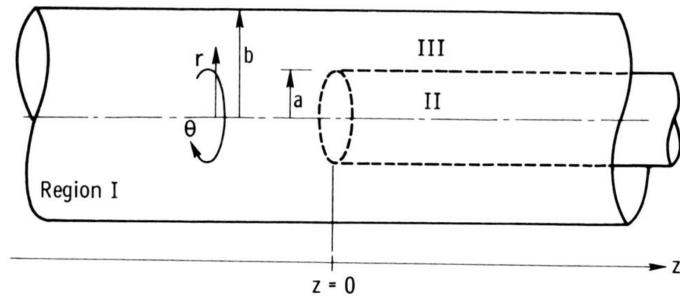
The reason that coaxial endcaps can be used to achieve high Q-values for the  $\text{TE}_{01\ell}$  modes is that the electric fields for these modes are purely azimuthally polarized (see Equations 6.12 and 6.13). Therefore, the boundary conditions that require the electric field to go to zero at the 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  $\text{TE}_{01}$  modes than perfect reflection can be achieved. This derivation is quite lengthy and complex and is presented in full in [3]. Here, we shall simply explain the resulting



**Figure 6.10.** The simplified geometry of an open cavity with coaxial terminations. Figure from [3].

5179 conditions on the partition radius for perfect reflection.



**Figure 6.11.** Electric field regions for the open cavity boundary value problem. Figure from [3].

5180 The open cavity boundary value problem is solved by expressing the forms of the  
 5181 electric fields in the different regions of the cavity and requiring that the electric fields are  
 5182 continuous. There are effectively three distinct regions in the open cavity corresponding  
 5183 to the central cavity volume, the inner coaxial volume, and the outer coaxial volume (see  
 5184 Figure 6.11).

5185 In Region I, the boundary conditions are those of a cylindrical waveguide, and we  
 5186 require that  $E_\phi$  for the  $TE_{0m}$  modes go to zero at the cavity wall ( $r = b$ ). This requires  
 5187 that  $J'_{0m}(k_{c0m} b) = 0$ . We aim to solve for the radius  $a$  in the specific situation where the  
 5188  $TE_{01}$  mode can propagate but all other  $TE_{0m}$  modes are below the cutoff frequency for  
 5189 the circular waveguide. This is equivalent to requiring

$$3.832 < k_{c0m} b < 7.016, \quad (6.41)$$

5190 where the numbers 3.832 and 7.016 correspond to the first and second zeros of the Bessel  
 5191 function (see Table 6.1).

5192 In Region II the boundary conditions are those of a cylindrical waveguide, but with  
 5193 a smaller radius. The condition that  $E_\phi = 0$  at the cylindrical partition radius is that  
 5194  $J'_{0m}(k_{c0m}a) = 0$ . To ensure perfect reflection, we want all modes in Region 1 of the cavity  
 5195 to be below the cutoff frequency of the circular waveguide formed by the inner volume of  
 5196 the coaxial terminator. Therefore, we consider the solutions where

$$k_{c0m}a < 3.832. \quad (6.42)$$

5197 Finally, in Region III the boundary condition are those of a coaxial waveguide. We  
 5198 need to guarantee that  $E_\phi = 0$  at both  $r = b$  and  $r = a$ , which involves finding the  
 5199 eigenvalues of the following equation

$$J'_0(k_{c0m}a)Y'_0(k_{c0m}b) - J'_0(k_{c0m}b)Y'_0(k_{c0m}a) = 0, \quad (6.43)$$

5200 where  $Y'_0$  the zeroth-order derivatives of the Bessel function of the second kind. The  
 5201 solutions to this equation depend on the value of the ratio  $b/a$ . The approximate solution  
 5202 is given by

$$\delta_n a \simeq \frac{n\pi}{b/a - 1}, \quad (6.44)$$

5203 where  $\delta_n$  are eigenvalues of Equation 6.43. Similar to Region II, we are interested in  
 5204 solutions for which the TE<sub>01</sub> modes of Region I are below the cutoff of Region III.  
 5205 Therefore, we require that

$$k_{c0m} < \delta_1. \quad (6.45)$$

5206 In general, one has some freedom in specifying the value of  $b/a$ . A value typically used  
 5207 in practice is  $b/a = 2.082$ , which corresponds to positioning the radius of the cylindrical  
 5208 partition at the maxima of the TE<sub>01</sub> electrical fields.

5209 Using the constraints from the three field regions one can develop a coaxial terminator  
 5210 that acts as a virtual perfectly conducting surface for the TE<sub>01</sub> modes. The only required  
 5211 inputs are the desired frequency of the TE<sub>011</sub> mode and a choice for the value of  $b/a$ .

### 5212 6.4.2 Mode Filtering

5213 The general case of an electron coupling to a resonant cavity is complicated. This is  
 5214 because cavities contain an infinite number of resonant modes, which for higher order  
 5215 modes, have couplings to the electron with a complex spatial dependence. The danger is  
 5216 that improper modeling of the electron's coupling to the cavity can lead to systematic

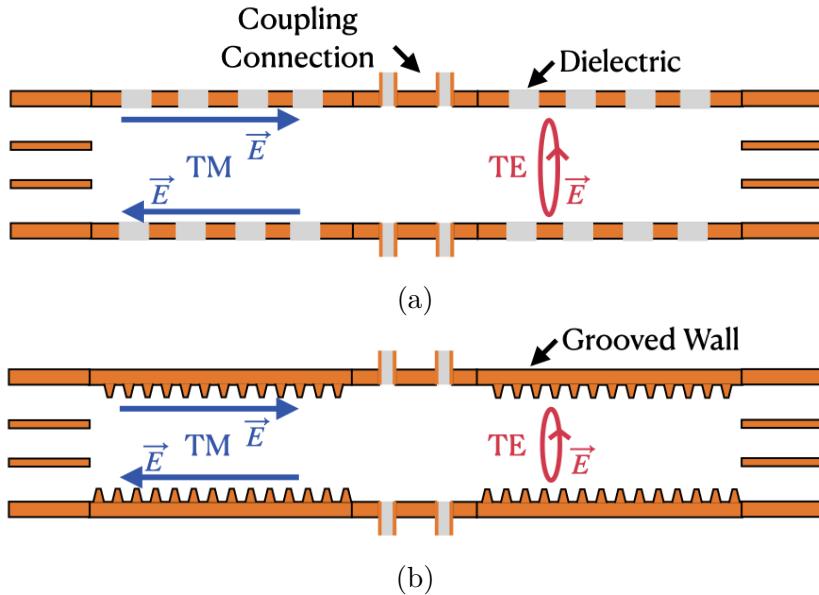
5217 errors in the CRES measurements that prevent a high-resolution measurement of the  
5218 electron's kinetic energy. This in part drives the preference for a single-mode cavity  
5219 experiment that uses only the electron's coupling to the  $TE_{011}$  mode to perform CRES,  
5220 assuming that sufficient information on the electron's position can be obtained with a  
5221 single mode.

5222 The  $TE_{011}$  mode is in a region where there are relatively few other modes to which  
5223 the electron could couple(see Figure 6.8). However, one can see that the frequency of  
5224 the  $TE_{011}$  is perfectly degenerate with the  $TM_{111}$  mode, which means that electrons will  
5225 inevitably couple to both modes if they have the correct cyclotron frequency.

5226 The magnitude of the impact of the electron coupling to both  $TE_{011}$  and  $TM_{111}$  is  
5227 currently unknown. To first order an electron coupling to more both modes will lose more  
5228 energy overtime, which can be measured by observing the frequency chirp rate of the  
5229 signal. This effect may be small enough to be negligible or simple enough to model that  
5230 the cavity can be treated as an effective single-mode cavity. Alternatively, the one could  
5231 consider devising a coupling scheme that is sensitive to both the  $TE_{011}$  and the  $TM_{111}$   
5232 modes. By measuring the coupling of the electron to both modes more information on  
5233 the position of the electron could be obtained, which could improve the position and  
5234 energy resolution of the CRES measurements.

5235 A different approach is the mode filtering approach, which seeks to obtain a single  
5236  $TE_{011}$  mode cavity using perturbations to the cavity walls that selectively impede the  
5237 TM modes, while leaving the TE modes mostly unperturbed. The type of perturbations  
5238 required can be determined by visualizing the surface currents induced in the cavity  
5239 walls by each type of mode (see Figure 6.12). By definition, all TM have electric fields  
5240 directed along the vertical axis of the cylindrical cavity, which means that perturbations  
5241 that impede currents in this direction will modify TM resonances. On the other hand,  
5242 the  $TE_{01}$  modes induce azimuthal currents in the cavity walls, therefore, it is possible to  
5243 break the degeneracy between  $TE_{01}$  and  $TM_{11}$  using a cavity perturbation that impedes  
5244 axial currents, but does not affect the flow of azimuthal currents.

5245 Figure 6.12 shows two cavity design concepts that achieve this selective current  
5246 perturbation. The resistive approach inserts a series of thin dielectric rings into the walls  
5247 of the cavity that introduces a resistive and capacitive impedance to the longitudinal  
5248 currents, while leaving azimuthal current paths intact. Cavities of this type with high  
5249  $TE_{01}$  Q's have also been constructed by tightly wrapping a thin, dielectric coated wire  
5250 around a mold to form the cavity wall. An alternative method is to introduce an inductive  
5251 impedance by cutting grooves or a thread pattern on the inside wall of the cavity. For



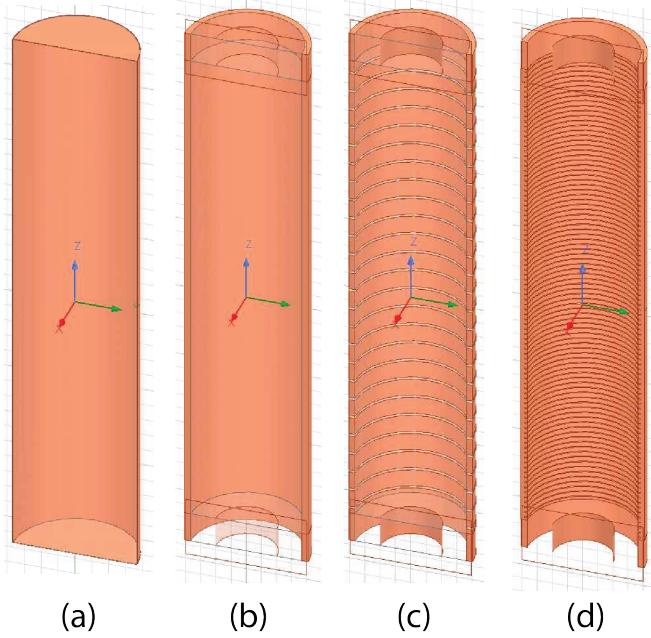
**Figure 6.12.** Two mode filtering concepts to break the degeneracy of  $TE_{01}$  and  $TM_{11}$  modes. The resistive approach uses dielectric materials to impede currents that travel vertically along the cavity while leaving azimuthal currents unperturbed. An alternative approach is to impede the currents using grooves cut into the cavity wall, which achieve the same effect with an inductive impedance.

5252 reasons of manufacturability and compatibility with tritium the grooved cavity approach  
 5253 is the preferred method for mode-filtered cavity construction by Project 8.

### 5254 **6.4.3 Simulations of Open, Mode-filtered Cavities**

5255 A candidate design for a single  $TE_{011}$  mode CRES experiment is a cavity that utilizes  
 5256 the coaxial terminations combined with a mode-filtering wall. The first step towards  
 5257 validating that a cavity that combines these two design features will operate as expected  
 5258 is a thorough simulation effort for which finite element method (FEM) simulation software  
 5259 is invaluable. The primary tool for electromagnetic FEM calculations inside Project 8 is  
 5260 Ansys HFSS, which has a robust and well-established eigenmode solver that can identify  
 5261 the resonant frequencies and associated Q-factors for given structure.

5262 Four variations of a cavity design with a  $\sim 1$  GHz  $TE_{011}$  resonance were implemented  
 5263 in HFSS (see Figure 6.13). The four designs include a standard cylindrical cavity, an  
 5264 open cavity with smooth walls, an open cavity with resistive walls, and an open cavity  
 5265 with grooved walls. The relevant design parameters are summarized in Table 6.3. All  
 5266 cavities were simulated using copper walls and filled with a vacuum dielectric. The  
 5267 identities of the resonant modes found by HFSS were validated by visual inspection of



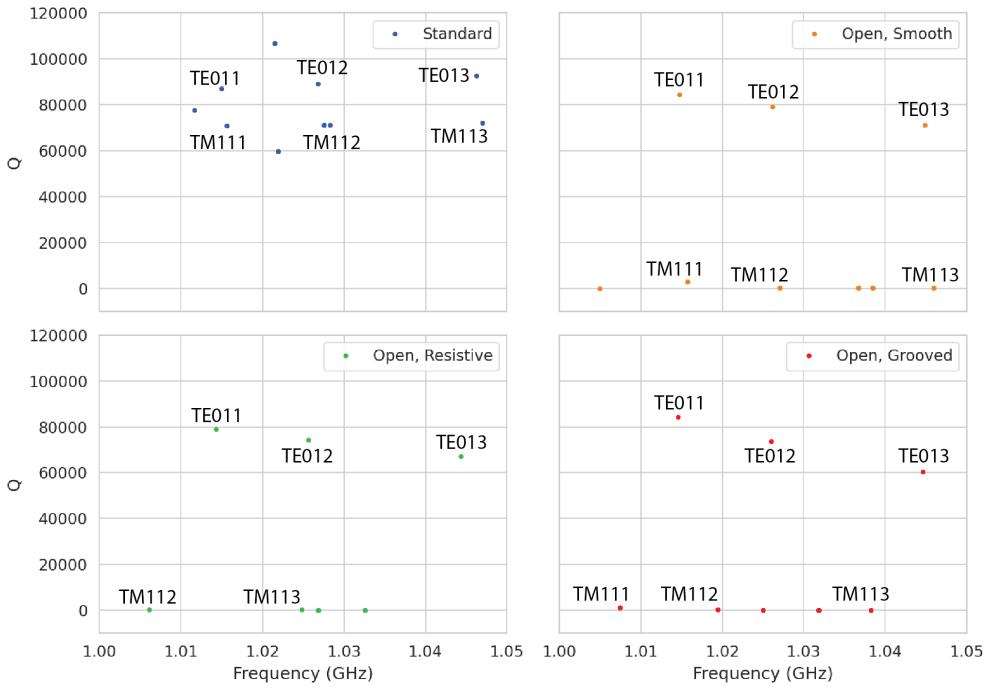
**Figure 6.13.** Four cavity design variations. (a) is a standard sealed cylindrical cavity, (b) is an open cavity with smooth walls, (c) is an open cavity with resistive walls, and (d) is an open cavity with grooved walls. The main cavity and coaxial terminator parameter are identical for all four cavities.

5268 the electric and magnetic field patterns and by comparison to analytical calculations of  
5269 the mode frequencies.

**Table 6.3.** A table of cavity design parameters used for HFSS simulations.

Name	Qty.	Unit	Description
$D_{\text{cav}}$	326.4	mm	Cavity diameter
$L_{\text{cav}}$	1668.0	mm	Cavity length
$D_{\text{term}}$	200.2	mm	Inner diameter of coaxial terminator
$L_{\text{term}}$	100.0	mm	Terminator length
$l_{\text{die}}$	8.3	mm	Dielectric spacer thickness
$\Delta l_{\text{die}}$	66.7	mm	Distance between dielectric spacers
$l_{\text{groove}}$	3.0	mm	Groove height
$d_{\text{groove}}$	9.0	mm	Groove depth
$\Delta l_{\text{groove}}$	18.3	mm	Distance between grooves

5270 The results of the HFSS simulations validate our predictions of the resonant behavior  
5271 of an open, mode-filtered cavity developed in the preceding sections (see Figure 6.14) One  
5272 can see that for a standard cavity the  $\text{TE}_{01}$  and the  $\text{TM}_{11}$  are degenerate in frequency  
5273 with relatively high Q-factors. The open-ended cavity preserves the high Q-factors of  
5274 the  $\text{TE}_{01}$  modes, while the other modes, since their boundary conditions do not match



**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 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 neutrino mass sensitivity of the experiment, which allows for the optimization of the cavity

5288 design towards the configuration that provides the best measurement of the neutrino  
5289 mass.

## 5290 **6.5 Single-mode Resonant Cavity Measurements**

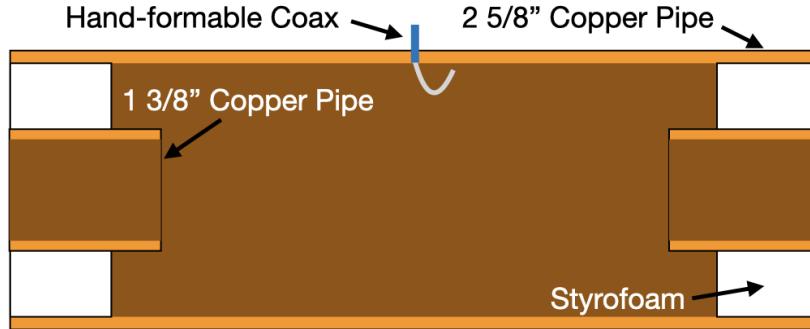
5291 Measurement test stands play an important role in the research and development process  
5292 that cannot be replaced by simulations. For example, constructing a prototype CRES  
5293 cavity forces one to consider important practical issues such as manufacturability and  
5294 machine tolerances that may require modifications to the design. Furthermore, by  
5295 comparing laboratory measurements of a real cavity to simulations, one can quantify  
5296 the impact of imperfections and real-life measurement systematics, which allows for  
5297 more accurate sensitivity estimates of the experiment. Lastly, the development of these  
5298 prototypes helps to build the necessary experience and expertise within the collaboration  
5299 required for more complicated experiments to succeed.

5300 In this spirit a prototype cavity was constructed to demonstrate the open, mode-  
5301 filtered cavity concept explored in the previous sections. The primary goal of the  
5302 measurements was to validate that an open, mode-filtered cavity suppressed the  $\text{TM}_{11}$   
5303 modes as predicted by HFSS simulations.

### 5304 **6.5.1 Cavities and Setup**

5305 Two rudimentary, cavities were constructed using segments of copper pipe available from  
5306 McMaster-Carr (see Figure 6.15). The design consists of copper pipes of two diameters.  
5307 The larger diameter pipe forms the main cavity wall and the smaller diameter pipe is  
5308 used to create a coaxial termination. The diameter of the outer pipe was chosen to  
5309 produce a  $\text{TE}_{011}$  resonance of approximately 6 GHz, while the diameter of the smaller  
5310 pipe was selected based on the open termination criteria introduced in Section 6.4.1. The  
5311 approximate diameters and lengths of the copper pipe are summarized in Table 6.4.

5312 Coupling to the cavity was achieved using a hand-formable segment of coaxial cable  
5313 stripped at one end to form a loop antenna. This was inserted into a small hole located  
5314 at the center of the main cavity wall. The coaxial terminators were supported inside the  
5315 main cavity by carving a spacer from polystyrene foam (styrofoam) so that they could  
5316 be easily inserted into the cavity and repositioned. The dielectric constant of styrofoam  
5317 is quite close to air at microwave frequencies so this is expected to have minimal impact  
5318 on the resonant properties of the cavity.



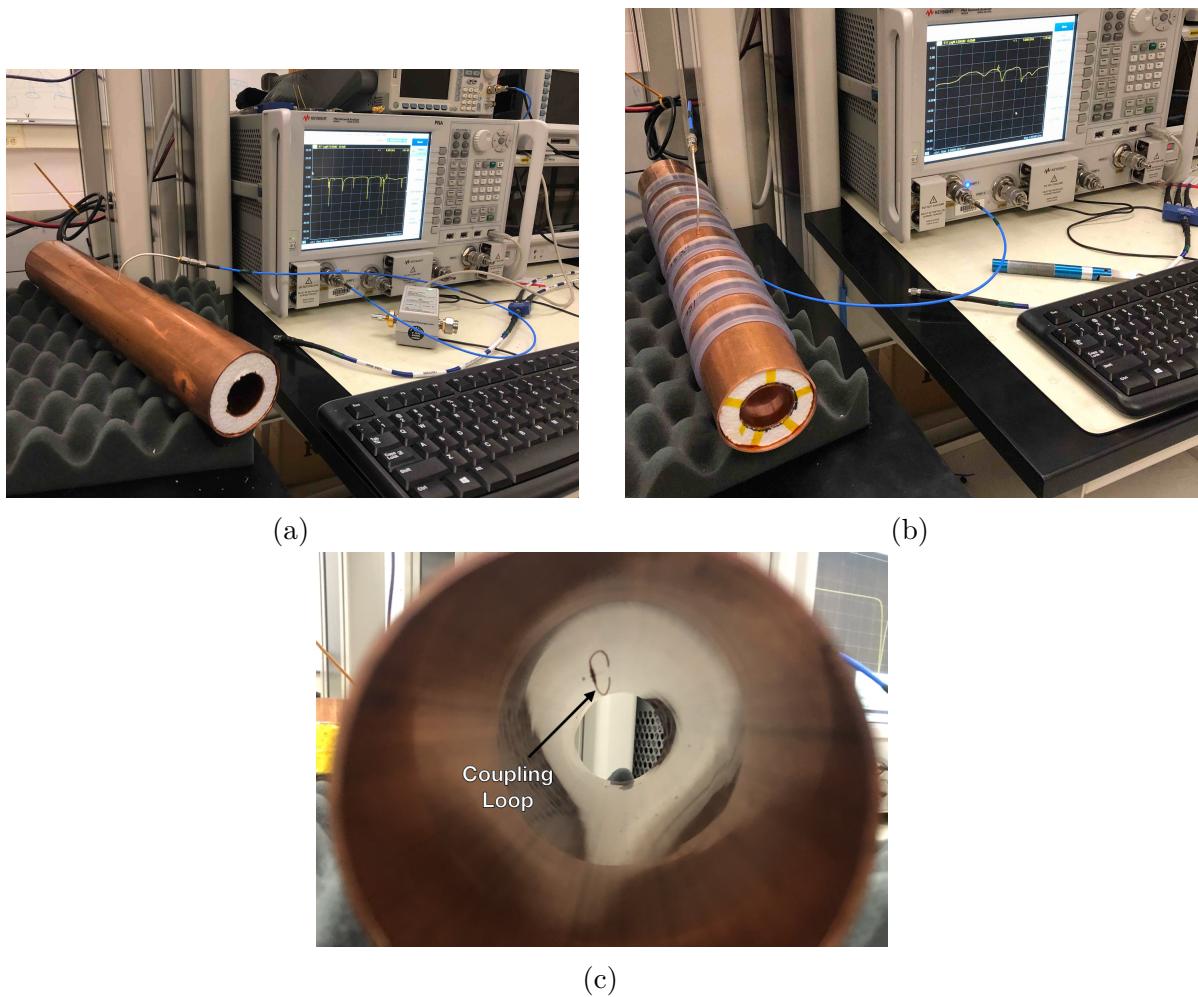
**Figure 6.15.** A cartoon depicting the design of the open-ended cavity prototype designed to operate at approximately 6 GHz. The main cavity wall was composed of a single copper pipe. A mode-filtered version of this cavity was constructed by

**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
$D_{\text{cav}}$	2.625	in	Cavity diameter
$L_{\text{cav}}$	$\approx 13$	in	Cavity length
$D_{\text{term}}$	1.375	in	Inner diameter of coaxial terminator
$L_{\text{term}}$	1.575	in	Terminator length
$l_{\text{die}}$	0.75	in	Dielectric spacer thickness
$\Delta l_{\text{die}}$	$\approx 1.50$ to $1.85$	in	Distance between dielectric spacers

5319 The actual length of the cavity is given by the distance between the inner edges of the  
 5320 coaxial terminations. The length of the outer section of pipe that forms the main wall of  
 5321 the cavity is approximately 16" in length which leads to a cavity length of  $\approx 13"$  when  
 5322 both terminators are inserted in the cavity. Because the terminators were not rigidly  
 5323 mounted this distance is only approximate, however, the uncertain length of the cavity  
 5324 will not prevent us from validating the open cavity design.

5325 Along with the smooth-walled open cavity a resistively mode-filtered cavity was  
 5326 constructed by creating dielectric spacers out of segments of clear PVC pipe (see Figure  
 5327 6.16). The spacers were machined such that the conductive segments of the cavity would  
 5328 be separated by 0.75" when the cavity was fully assembled. Due to variations in the  
 5329 lengths of the copper segments that make up the cavity wall the distance between spacers  
 5330 has significant variation with average value of about 1.7". Eight total spacers were used  
 5331 to build the cavity, which when assembled was approximately 16" in total length similar  
 5332 to the non-filtered cavity.

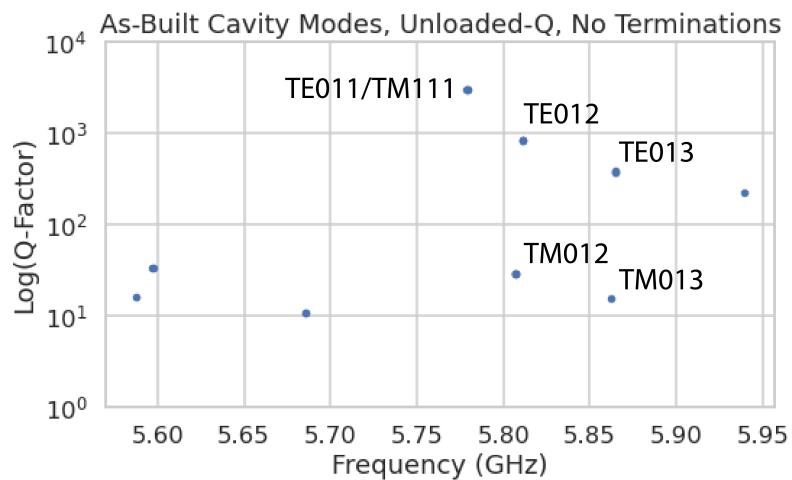


**Figure 6.16.** Images depicting the measurement of the filtered and non-filtered open cavities using the VNA. The coupling loop in the figure is shown in the TE orientation.

5333 Measurements of both cavities were performed using a VNA connected to the cavity  
 5334 coupling probe (see Figure 6.16). By measuring the return loss over a range of frequencies  
 5335 one can measure the frequencies and relative Q-factors of the resonant modes in the  
 5336 cavity. Due to the opposite polarity of the electric fields for the TE and TM modes,  
 5337 the loop coupling probe must be rotated 90° to change the polarity of the loop antenna.  
 5338 When the antenna is oriented such that the loop opening faces the ends of the cavity, it  
 5339 couples primarily to the TE modes which have magnetic fields directed along the long  
 5340 axis of the cavity (see Figure 6.16). If the coupling loop is turned by 90° from where  
 5341 it is shown in the image then it will couple to the TM modes which have azimuthally  
 5342 directed magnetic fields. In this way both the TE and TM resonances can be measured  
 5343 independently.

## 5344 6.5.2 Results and Discussion

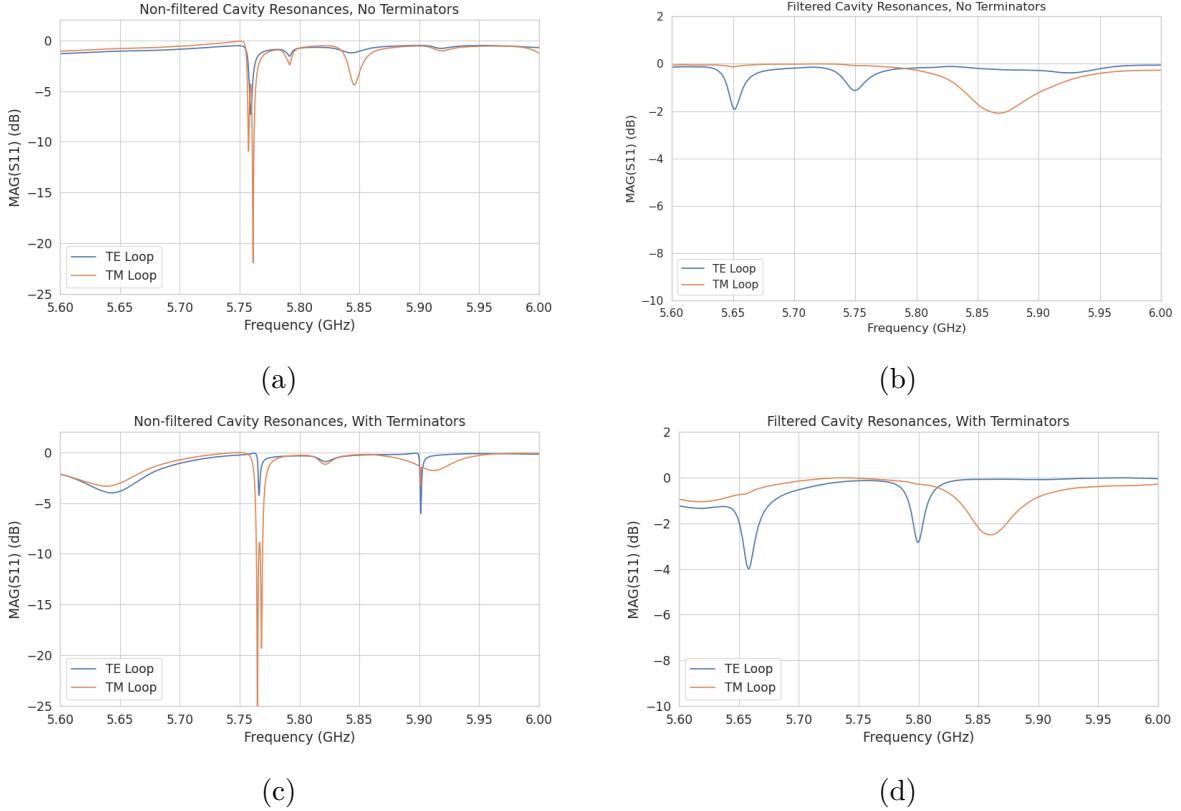
5345 The primary analysis for the prototype cavities involved a simple visualization of the  
5346 return loss as measured by the VNA and a comparison between the filtered and non-  
5347 filtered variations. Since the resonances measured by the VNA are not labeled, there is  
5348 an uncertainty about the true identities of the modes measured by the VNA. To resolve  
5349 this I performed a simulation of the simplest possible cavity that could be created from  
5350 the prototype components, which is a fully open cavity created by removing the coaxial  
5351 inserts. The fully-open cavity with the as-built dimensions was simulated in HFSS to get  
5352 estimates on the positions of the TE<sub>011</sub> and TM<sub>111</sub> modes (see Figure 6.17).



**Figure 6.17.** HFSS simulation results for a the as-built cavity with the coaxial terminators removed. The TE<sub>011</sub>/TM<sub>111</sub> frequency is approximately 5.78 GHz.

5353 Simulation of the fully open cavity shows that the TE<sub>011</sub>/TM<sub>111</sub> modes have a  
5354 frequency of approximately 5.78 GHz in the fully open cavity. If the frequency of this  
5355 mode is compared to the measurements of the filtered and non-filtered cavities with the  
5356 terminators removed one can easily identify the TE<sub>011</sub> mode at approximately 5.75 GHz  
5357 (see Figure 6.18).

5358 Both variations of the non-filtered cavities one sees that the TE<sub>011</sub> mode is degenerate  
5359 in frequency with what appears to be a doublet of TM modes located at the TM<sub>111</sub>  
5360 frequency position. This doublet is actually the TM<sub>111</sub> mode, which has two polarizations  
5361 with opposite polarizations. Because the pipe used to construct the cavity is not perfectly  
5362 round, the frequency degeneracy between the two polarizations is broken resulting in the  
5363 doublet peaks.



**Figure 6.18.** Measurements of the filtered and non-filtered prototype cavities acquired with the VNA.

5364      The S-parameter plot for the filtered cavity without terminators has an isolated TE  
 5365      resonance at 5.65 GHz, associated with the  $TE_{011}$  mode. The frequency of this mode  
 5366      is lower than the non-filtered cavity due to a difference in the overall lengths of the  
 5367      cavities. An obvious difference between the filtered and non-filtered cavities is that  
 5368      there is no  $TM_{111}$  doublet at the  $TE_{011}$  frequency. This is what one would expect if  
 5369      the mode-filtering was suppressing the TM modes. There appears to be a noticeable  
 5370      difference in the Q of the  $TE_{011}$  resonance between non-filtered and filtered variations as  
 5371      indicated by the increased resonance depth for the filtered cavity. Overall, the Q-factors  
 5372      of the filtered cavity appear significantly smaller than the non-filtered cavity due to the  
 5373      increase in resonance width. This is likely caused by the relatively large widths of the  
 5374      dielectric spacers, which are partially impeding the TE modes.

5375      One can see from these cavity measurements that, in principle, resistive mode-filtering  
 5376      can be used to separate the  $TE_{011}$  resonance from the degenerate  $TM_{111}$  modes in  
 5377      combination with the open cavity endcaps. This finding agrees with the expectations  
 5378      from HFSS, which should provide confidence that the eigenmode solver is correctly

5379 modeling the behavior of the cavity. Although I did not perform a similar study using  
5380 a cavity with grooved walls it is expected that the resonant mode structure would be  
5381 similar to the cavity studied here.

5382 While this prototype cavity is a good first step, several deficiencies prevent this setup  
5383 from providing more than qualitative information to the design of cavities for CRES. This  
5384 includes the rudimentary approach to cavity coupling using a stripped coax antenna and  
5385 the inability to map the field density in the cavity volume. Improvements in these areas  
5386 are required so that measurements from a real cavity can provide useful information to  
5387 cavity CRES simulations that will ultimately inform neutrino mass sensitivity estimates.

5388 Future work with prototype cavities must include an improved cavity coupling scheme,  
5389 which is robust and compatible with atomic tritium. Since the cavity will ultimately  
5390 be filled with atomic tritium, a coupling antenna cannot be used due to the losses of  
5391 atomic tritium caused by recombination on the antenna surfaces. Possible non-invasive  
5392 coupling schemes include aperture coupling, where the cavity is coupled to an external  
5393 waveguide structure through an aperture, or a split-ring coupling approach, where the  
5394 center segment of the cylindrical cavity wall is replaced an isolated conductive ring with  
5395 a small vertical slit. The aperture coupling approach is a standard coupling scheme [87]  
5396 used in a wide range of applications, but at low frequencies the size of the external  
5397 waveguide conflicts with design of the atom trapping magnet and cryogenics system.  
5398 The split-ring approach could potentially be coupled to a small coaxial transmission line  
5399 which is more compatible with the rest of the experiment design. A challenge is achieving  
5400 adequate coupling through impedance tuning, which is a focus of current research.

5401 The robustness of the coupling mechanism is relevant due to the difficulty in modeling  
5402 its effect on the cavity modes. Small changes in geometry can have a large influence on  
5403 the coupling and hence the performance of the cavity, therefore, correctly modeling the  
5404 cavity coupling is critical for accurate CRES simulations. Coupling schemes that rely  
5405 on connections to coaxial lines are potentially at a disadvantage in this regard due to  
5406 the affect of soldering imperfections or unintended bends in the coax on the coupling.  
5407 Future work will identify a coupling scheme for the cavity compatible with the neutrino  
5408 mass goals of Project 8.

5409 Imperfections in the geometry of a real cavity will necessarily distort the resonant  
5410 modes away from simulation predictions. This will change the coupling of an electron  
5411 to the cavity and thus change the expected signal structure. Ultimately, this effect will  
5412 limit the achievable energy resolution of the experiment unless the differences between  
5413 simulation and a real cavity can be sufficiently characterized and calibrated. One possible

5414 approach to this is to utilize a "bead puller" system [101] to strategically perturb the  
5415 cavity by moving a conductive bead through the cavity volume. The small perturbation  
5416 caused by the bead affects the phase of the cavity resonances proportional to the total  
5417 magnitude of the electric field at that position, so by moving the bead through the  
5418 cavity volume the total electric field can be mapped and compared to simulation. This  
5419 information can provide bounds on the relative perturbations to the cavity mode structure  
5420 from real-life imperfections compared to the idealized cavity in HFSS.

5421 **Chapter 7 |**

5422 **Conclusion and Future Prospects**

5423 In this dissertation we have discussed research and development efforts towards the  
5424 development of a scalable CRES measurement technology that can be used to build a  
5425 CRES experiment at cubic-meter scales with sensitivity to neutrino masses of 40 meV.  
5426 The primary contributions of my dissertation are the development and analysis of signal  
5427 reconstruction algorithms for an antenna array based CRES experiment [102], which leads  
5428 to estimates of the neutrino mass sensitivity; the development of a synthetic cyclotron  
5429 radiation antenna (SYNCA) [81], which allowed for laboratory validation of antenna  
5430 array CRES simulation models [44]; and the development of an open-ended cavity design  
5431 compatible with atomic tritium for a cavity based CRES experiment. A measurable  
5432 impact of this work is the transition of the Project 8 collaboration's experimental plan  
5433 from an antenna array based approach to a cavity based approach, where my work played  
5434 a key role in demonstrating the significantly higher cost and complexity of the antenna  
5435 array experiment.

5436 The transition from antenna arrays to cavities requires a new set of demonstrator  
5437 experiments to make incremental progress towards a 40 meV measurement of the neutrino  
5438 mass. At the time of writing, the near-term plan of Project 8 is to design and construct a  
5439 small-scale cavity CRES experiment utilizing the 1 T magnet installed in the UW-Seattle.  
5440 This cavity is designed to have a TE011 resonance with a frequency of about 26 GHz with  
5441 a length-to-diameter ratio that mimics the larger cavities intended for the pilot-scale and  
5442 Phase IV experiments. The goal of this experiment is to demonstrate cavity CRES as  
5443 well as validate models of CRES systematics using electrons from  $^{83m}\text{Kr}$  and an electron  
5444 gun. Though the primary goal is demonstration, near-term physics measurements are  
5445 available in the form of high-resolution measurements of the  $^{83m}\text{Kr}$  conversion spectrum  
5446 of interest to the KATRIN collaboration.

5447 Furthermore, Project 8 is currently constructing a low-frequency CRES setup located  
5448 at Yale University to better understand the principles of cavity based CRES at lower

5449 magnetic fields. The Low, UHF Cavity Krypton Experiment at Yale (LUCKEY) is  
5450 a 1.5 GHz cavity CRES experiment the will use conversion electrons from  $^{83m}\text{Kr}$  to  
5451 perform CRES measurements at the lowest frequencies ever attempted with the technique.  
5452 LUCKEY will validate frequency scaling models developed by Project 8 and will pave  
5453 the way for the future Low-Frequency Apparatus (LFA), which will be a larger, 1 GHz  
5454 cavity CRES experiment that includes a molecular tritium source. The target for the  
5455 LFA is a measurement of the neutrino mass with a sensitivity of approximately 0.2 eV,  
5456 which will build towards the atomic pilot-scale CRES experiment.

5457 In parallel to the development of cavity CRES is the development of the atomic  
5458 tritium source. Recent demonstrations of the production of atomic hydrogen are excellent  
5459 steps towards the atomic tritium production needed for the pilot-scale experiment. One  
5460 area of future study includes the development of a more detailed understanding of the  
5461 efficiency of atomic hydrogen production. Near-term plans include the development of a  
5462 magnetic, evaporatively cooled beamline, as well as the prototyping of a Halbach array  
5463 atoms trap. Nearly all of the components of the atomic tritium system will require  
5464 demonstration before the complete system can be built. The long-term goal of the  
5465 atomic tritium work is to construct a full atomic tritium prototype that demonstrates  
5466 the production, cooling, trapping, and recycling of tritium at the rates needed for the  
5467 pilot-scale experiment.

5468 More broadly, the long-term goal of the Project 8 collaboration is to fully develop  
5469 both the atomic tritium and cavity CRES technologies so that both can be combined in  
5470 a pilot-scale CRES experiment. It is envisioned that this process will take approximately  
5471 10 years for both atomic tritium and cavity CRES. After these developments comes  
5472 the pilot-scale experiment which will be the first CRES experiment that simultaneously  
5473 demonstrates all the required technologies for Phase IV. Scaling to Phase IV with cavity  
5474 CRES will require the construction of multiple copies (approximately 10) of the pilot-scale  
5475 experiment to obtain sufficient statistics for 40 meV sensitivity.

5476 Development of the CRES experimental technique by Project 8 has led to new  
5477 experiments utilizing the CRES technique for basic physics research, such as the  $^6\text{He}$ -  
5478 CRES collaboration [103], and has also found applications as a new approach to x-ray  
5479 spectroscopy [104]. Recently, a new experimental effort called CRESDA has begun in  
5480 the UK to develop new quantum technologies applied to CRES measurements for the  
5481 neutrino mass [105]. This flourishing of new experimental efforts based on the CRES  
5482 technique is likely to continue as Project 8 continues to develop the technique towards  
5483 its neutrino mass measurement goal.

5484

# Bibliography

- 5485 [1] AGOSTINI, M., G. BENATO, and J. DETWILER (2017) “Discovery probability of  
5486 next-generation neutrinoless double-  $\beta$  decay experiments,” *Phys. Rev. D*, **96**(5), p.  
5487 053001, 1705.02996.
- 5488 [2] WENGER, N. C. and J. SMETANA (1972) “Hydrogen Density Measurements Using  
5489 an Open-Ended Microwave Cavity,” *IEEE Transactions on Instrumentation and  
5490 Measurement*, **21**(2), pp. 105–114.
- 5491 [3] WENGER, N. (1967) “Resonant Frequency of Open-Ended Cylindrical Cavity,”  
5492 *IEEE Transactions on Microwave Theory and Techniques*, **15**(6), pp. 334–340.
- 5493 [4] DAS, A. and T. FERBEL (2003) *Introduction to Nuclear and Particle Physics*,  
5494 World Scientific.
- 5495 [5] CURIE, P. (1898) “Sur une nouvelle substance fortement radioactive, contenue  
5496 dans la pechblende,” *Comptes rendus de l’Academie des Sciences, Paris*, **127**, pp.  
5497 1215–1217.
- 5498 [6] RUTHERFORD, E. (1899) “Uranium radiation and the electrical conduction pro-  
5499 duced by it,” *Philos. Mag.*, **47**, p. 109.  
5500 URL <https://cds.cern.ch/record/261896>
- 5501 [7] GERWARD, L. (1999) “Paul Villard and his Discovery of Gamma Rays,” *Physics  
5502 in Perspective*, **1**, pp. 367–383.
- 5503 [8] BECQUEREL, H. (1900) “Deviation du rayonnement du radium dans un champ  
5504 électrique,” *C.R.*, **130**, pp. 809–815.
- 5505 [9] D.Sc., E. R. M. and F. S. B.A. (1902) “XLI. The cause and nature of ra-  
5506 dioactivity.—Part I,” *The London, Edinburgh, and Dublin Philosophical Mag-  
5507 azine and Journal of Science*, **4**(21), pp. 370–396, [https://doi.org/10.1080/  
5508 14786440209462856](https://doi.org/10.1080/14786440209462856).  
5509 URL <https://doi.org/10.1080/14786440209462856>
- 5510 [10] CHADWICK, J. (1914) *Intensitätsverteilung im magnetischen Spektrum der  $\beta$ -  
5511 Strahlen von Radium B+C*, Druck von Friedr. Vieweg und Sohn.  
5512 URL <https://books.google.com/books?id=8s8UmQEACAAJ>

- 5513 [11] BROWN, L. M. (1978) “The idea of the neutrino,” *Physics Today*, **31**(9).
- 5514 [12] PAULI, W. (1978) “Dear radioactive ladies and gentlemen,” *Phys. Today*, **31N9**,  
5515 p. 27.
- 5516 [13] AMALDI, E. (1984) “From the discovery of the neutron to the discovery of nuclear  
5517 fission,” *Physics Reports*, **111**(1), pp. 1–331.  
5518 URL <https://www.sciencedirect.com/science/article/pii/037015738490214X>
- 5520 [14] CHADWICK, J. (1932) “The existence of a neutron,” *Proceedings of the Royal  
5521 Society of London. Series A, Containing Papers of a Mathematical and Physical  
5522 Character*, **136**(830), pp. 692–708, <https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1932.0112>.  
5524 URL <https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1932.0112>
- 5526 [15] WILSON, F. L. (1968) “Fermi’s Theory of Beta Decay,” *American Journal of  
5527 Physics*, **36**(12), pp. 1150–1160, [https://pubs.aip.org/aapt/ajp/article-pdf/36/12/1150/11891052/1150\\_1\\_online.pdf](https://pubs.aip.org/aapt/ajp/article-pdf/36/12/1150/11891052/1150_1_online.pdf).  
5529 URL <https://doi.org/10.1119/1.1974382>
- 5530 [16] CRANE, H. R. and J. HALPERN (1938) “New Experimental Evidence for the  
5531 Existence of a Neutrino,” *Phys. Rev.*, **53**, pp. 789–794.  
5532 URL <https://link.aps.org/doi/10.1103/PhysRev.53.789>
- 5533 [17] COWAN, C. L., F. REINES, F. B. HARRISON, H. W. KRUSE, and A. D.  
5534 MC GUIRE (1956) “Detection of the Free Neutrino: a Confirmation,” *Science*,  
5535 **124**(3212), pp. 103–104, <https://www.science.org/doi/pdf/10.1126/science.124.3212.103>.  
5537 URL <https://www.science.org/doi/abs/10.1126/science.124.3212.103>
- 5538 [18] ANICIN, I. V. (2005) “The Neutrino - Its Past, Present and Future,” *arXiv e-prints*,  
5539 [physics/0503172](https://arxiv.org/abs/physics/0503172), [physics/0503172](https://arxiv.org/abs/physics/0503172).
- 5540 [19] DAVIS, R. (2003) “Nobel Lecture: A half-century with solar neutrinos,” *Rev. Mod.*  
5541 *Phys.*, **75**, pp. 985–994.
- 5542 [20] McDONALD, A. B. ET AL. (2002) “Direct evidence for neutrino flavor transfor-  
5543 mation from neutral-current interactions in SNO,” *AIP Conf. Proc.*, **646**(1), pp.  
5544 43–58.
- 5545 [21] FUKUDA, Y. ET AL. (1998) “Evidence for oscillation of atmospheric neutrinos,”  
5546 *Phys. Rev. Lett.*, **81**, pp. 1562–1567, [hep-ex/9807003](https://arxiv.org/abs/hep-ex/9807003).
- 5547 [22] GIGANTI, C., S. LAVIGNAC, and M. ZITO (2018) “Neutrino oscillations: The rise  
5548 of the PMNS paradigm,” *Prog. Part. Nucl. Phys.*, **98**, pp. 1–54, [1710.00715](https://arxiv.org/abs/1710.00715).

- 5549 [23] AN, F. ET AL. (2016) “Neutrino Physics with JUNO,” *J. Phys. G*, **43**(3), p. 030401,  
 5550 1507.05613.
- 5551 [24] BIAN, J. ET AL. (2022) “Hyper-Kamiokande Experiment: A Snowmass White  
 5552 Paper,” in *Snowmass 2021*, 2203.02029.
- 5553 [25] ABI, B. ET AL. (2020) “Long-baseline neutrino oscillation physics potential of the  
 5554 DUNE experiment,” *Eur. Phys. J. C*, **80**(10), p. 978, 2006.16043.
- 5555 [26] WORKMAN, R. L. ET AL. (2022) “Review of Particle Physics,” *PTEP*, **2022**, p.  
 5556 083C01.
- 5557 [27] ZUBER, K. (2020) *Neutrino Physics*, CRC Press.
- 5558 [28] XING, Z. Z. and S. ZHOU (2011) *Neutrinos in Particle Physics, Astronomy and*  
 5559 *Cosmology*, Springer.
- 5560 [29] MOHAPATRA, R. N. ET AL. (2007) “Theory of neutrinos: A White paper,” *Rept.*  
 5561 *Prog. Phys.*, **70**, pp. 1757–1867, hep-ph/0510213.
- 5562 [30] DE GOUVÊA, A. (2016) “Neutrino Mass Models,” *Ann. Rev. Nucl. Part. Sci.*, **66**,  
 5563 pp. 197–217.
- 5564 [31] WORKMAN, R. L. and OTHERS (2022) “Review of Particle Physics,” *PTEP*, **2022**,  
 5565 p. 083C01.
- 5566 [32] WONG, Y. Y. Y. (2011) “Neutrino mass in cosmology: status and prospects,”  
 5567 *Ann. Rev. Nucl. Part. Sci.*, **61**, pp. 69–98, 1111.1436.
- 5568 [33] ADE, P. A. R. ET AL. (2016) “Planck 2015 results. XIII. Cosmological parameters,”  
 5569 *Astron. Astrophys.*, **594**, p. A13, 1502.01589.
- 5570 [34] CARLSTROM, J. ET AL. (2019) “CMB-S4,” in *Bulletin of the American Astronomical Society*, vol. 51, p. 209, 1908.01062.
- 5572 [35] ELLIOTT, S. R. and P. VOGEL (2002) “Double beta decay,” *Ann. Rev. Nucl. Part.*  
 5573 *Sci.*, **52**, pp. 115–151, hep-ph/0202264.
- 5574 [36] DOLINSKI, M. J., A. W. P. POON, and W. RODEJOHANN (2019) “Neutrinoless  
 5575 Double-Beta Decay: Status and Prospects,” *Ann. Rev. Nucl. Part. Sci.*, **69**, pp.  
 5576 219–251, 1902.04097.
- 5577 [37] FORMAGGIO, J. A., A. L. C. DE GOUVÊA, and R. G. H. ROBERTSON (2021)  
 5578 “Direct measurements of neutrino mass,” *Physics Reports*, **914**, pp. 1–54, direct  
 5579 measurements of neutrino mass.  
 5580 URL [https://www.sciencedirect.com/science/article/pii/  
 5581 S0370157321000636](https://www.sciencedirect.com/science/article/pii/S0370157321000636)

- 5582 [38] AKER, M. ET AL. (2022) “New Constraint on the Local Relic Neutrino Background  
 5583 Overdensity with the First KATRIN Data Runs,” *Phys. Rev. Lett.*, **129**(1), p.  
 5584 011806, 2202.04587.
- 5585 [39] FORMAGGIO, J. A., A. L. C. DE GOUVÊA, and R. G. H. ROBERTSON (2021)  
 5586 “Direct Measurements of Neutrino Mass,” *Phys. Rept.*, **914**, pp. 1–54, 2102.00594.
- 5587 [40] MONREAL, B. and J. A. FORMAGGIO (2009) “Relativistic cyclotron radiation  
 5588 detection of tritium decay electrons as a new technique for measuring the neutrino  
 5589 mass,” *Phys. Rev. D*, **80**, p. 051301.  
 5590 URL <https://link.aps.org/doi/10.1103/PhysRevD.80.051301>
- 5591 [41] ESFAHANI, A. A. ET AL. (2022) “The Project 8 Neutrino Mass Experiment,”  
 5592 2203.07349.
- 5593 [42] ——— (2023) “Cyclotron Radiation Emission Spectroscopy of Electrons from  
 5594 Tritium Beta Decay and  $^{83m}\text{Kr}$  Internal Conversion,” 2303.12055.
- 5595 [43] ——— (2023) “Tritium Beta Spectrum and Neutrino Mass Limit from Cyclotron  
 5596 Radiation Emission Spectroscopy,” 2212.05048.
- 5597 [44] “Antenna Arrays for Physics Measurements with Large-scale CRES Detectors,” *In  
 5598 preparation*.
- 5599 [45] NICHOLSON, T. ET AL. (2015) “Systematic evaluation of an atomic clock at 2 x  
 5600 10-18 total uncertainty,” *Nature Communications*, **6**(6896).
- 5601 [46] HÄNSCH, T. W. (2006) “Nobel Lecture: Passion for precision,” *Rev. Mod. Phys.*,  
 5602 **78**, pp. 1297–1309.  
 5603 URL <https://link.aps.org/doi/10.1103/RevModPhys.78.1297>
- 5604 [47] GABOR, D. (1946) “Theory of communication. Part 1: The analysis of information,”  
 5605 *Journal of the Institution of Electrical Engineers-part III: radio and communication  
 5606 engineering*, **93**(26), pp. 429–441.
- 5607 [48] ESFAHANI ASHTARI, A. ET AL. (2022) “Viterbi decoding of CRES signals in  
 5608 Project 8,” *New J. Phys.*, **24**(5), p. 053013, 2112.05265.
- 5609 [49] JACKSON, J. D. (1999) *Classical electrodynamics*, 3rd ed., Wiley, New York, NY.  
 5610 URL <http://cdsweb.cern.ch/record/490457>
- 5611 [50] F.R.S., J. L. D. (1897) “LXIII. On the theory of the magnetic influence on  
 5612 spectra; and on the radiation from moving ions,” *The London, Edinburgh, and  
 5613 Dublin Philosophical Magazine and Journal of Science*, **44**(271), pp. 503–512,  
 5614 <https://doi.org/10.1080/14786449708621095>.  
 5615 URL <https://doi.org/10.1080/14786449708621095>

- 5616 [51] ESFAHANI, A. A. ET AL. (2022) “The Project 8 Neutrino Mass Experiment,” in  
 5617 *2022 Snowmass Summer Study*, 2203.07349.
- 5618 [52] KRAUS, C. ET AL. (2005) “Final results from phase II of the Mainz neutrino mass  
 5619 search in tritium beta decay,” *Eur. Phys. J. C*, **40**, pp. 447–468, [hep-ex/0412056](#).
- 5620 [53] ASEEV, V. N. ET AL. (2011) “An upper limit on electron antineutrino mass from  
 5621 Troitsk experiment,” *Phys. Rev. D*, **84**, p. 112003, [1108.5034](#).
- 5622 [54] SAENZ, A., S. JONSELL, and P. FROELICH (2000) “Improved Molecular Final-  
 5623 State Distribution of HeT+ for the  $\beta$ -Decay Process of T2,” *Phys. Rev. Lett.*, **84**(2),  
 5624 p. 242.
- 5625 [55] BODINE, L. I., D. S. PARNO, and R. G. H. ROBERTSON (2015) “Assessment  
 5626 of molecular effects on neutrino mass measurements from tritium  $\beta$  decay,” *Phys.  
 5627 Rev. C*, **91**(3), p. 035505, [1502.03497](#).
- 5628 [56] ASNER, D. M. ET AL. (2015) “Single electron detection and spectroscopy via  
 5629 relativistic cyclotron radiation,” *Phys. Rev. Lett.*, **114**(16), p. 162501, [1408.5362](#).
- 5630 [57] VÉNOS, D., J. SENTKERESTIOVÁ, O. DRAGOUN, M. SLEZÁK, M. RYŠAVÝ, and  
 5631 A. ŠPALEK (2018) “Properties of 83mKr conversion electrons and their use in the  
 5632 KATRIN experiment,” *JINST*, **13**(02), p. T02012.
- 5633 [58] LAGENDIJK, A., I. F. SILVERA, and B. J. VERHAAR (1986) “Spin exchange and  
 5634 dipolar relaxation rates in atomic hydrogen: Lifetimes in magnetic traps,” *Phys.  
 5635 Rev. B*, **33**, pp. 626–628.  
 5636 URL <https://link.aps.org/doi/10.1103/PhysRevB.33.626>
- 5637 [59] SILVERA, I. and J. WALRAVEN (1986) “Spin-polarized atomic hydrogen,” *Prog.  
 5638 Low Temp. Phys.*, **X**, pp. 139–370.
- 5639 [60] FURSE, D. ET AL. (2017) “Kassiopeia: a modern, extensible C++ particle tracking  
 5640 package,” *New Journal of Physics*, **19**(5), p. 053012.  
 5641 URL <https://doi.org/10.1088/1367-2630/aa6950>
- 5642 [61] ESFAHANI, A. A. ET AL. (2019) “Electron radiated power in cyclotron radiation  
 5643 emission spectroscopy experiments,” *Phys. Rev. C*, **99**, p. 055501.  
 5644 URL <https://link.aps.org/doi/10.1103/PhysRevC.99.055501>
- 5645 [62] ASHTARI ESFAHANI, A. ET AL. (2019) “Locust: C++ software for simulation of  
 5646 RF detection,” *New J. Phys.*, **21**, p. 113051, [1907.11124](#).
- 5647 [63] WIECHERT, E. (1901) “Elektrodynamische Elementargesetze,” .  
 5648 URL <https://doi.org/10.1002/andp.19013090403>
- 5649 [64] LIÉARD, A. (1898) “Champ électrique et Magnétique,” *Léclairage électrique*, **16**(27–  
 5650 29).

- 5651 [65] BALANIS, C. (2015) *Antenna Theory: Analysis and Design*, Wiley.  
 5652 URL <https://books.google.com/books?id=PTFcCwAAQBAJ>
- 5653 [66] <https://www.ansys.com/products/electronics/ansys-hfss>.
- 5654 [67] [https://en.wikipedia.org/wiki/Linear\\_time-invariant\\_system](https://en.wikipedia.org/wiki/Linear_time-invariant_system).
- 5655 [68] NYQUIST, H. (1928) “Certain Topics in Telegraph Transmission Theory,” *Transactions of the American Institute of Electrical Engineers*, **47**(2), pp. 617–644.
- 5656
- 5657 [69] BRUN, R. and F. RADEMAKERS (1997) “ROOT: An object oriented data analysis  
 5658 framework,” *Nucl. Instrum. Meth. A*, **389**, pp. 81–86.
- 5659 [70] ASHTARI ESFAHANI, A. ET AL. (2021) “Bayesian analysis of a future  $\beta$  decay  
 5660 experiment’s sensitivity to neutrino mass scale and ordering,” *Phys. Rev. C*, **103**,  
 5661 p. 065501.  
 5662 URL <https://link.aps.org/doi/10.1103/PhysRevC.103.065501>
- 5663 [71] KAY, S. (1998) *Fundamentals of Statistical Signal Processing: Detection Theory,*  
 5664 *Volume II*, Pearson.
- 5665 [72] NEYMAN, J. and E. PEARSON (1933) “On the problem of the the most efficient  
 5666 tests of statistical hypotheses,” *Phil. Trans. R. Soc. Lond. A*, **231**.
- 5667 [73] STUMPF, M. (2018) *Electromagnetic Reciprocity in Antenna Theory*, Wiley.
- 5668 [74] BISHOP, C. (2016) *Pattern Recognition and Machine Learning*, Springer.
- 5669 [75] PLEHN, T., A. BUTTER, B. DILLON, and C. KRAUSE (2022) “Modern Machine  
 5670 Learning for LHC Physicists,” 2211.01421.
- 5671 [76] GEORGE, D. and E. A. HUERTA (2018) “Deep Learning for Real-time Gravitational  
 5672 Wave Detection and Parameter Estimation: Results with Advanced LIGO Data,”  
 5673 *Phys. Lett. B*, **778**, pp. 64–70, 1711.03121.
- 5674 [77] GABBARD, H., C. MESSENGER, I. S. HENG, F. TONOLINI, and R. MURRAY-  
 5675 SMITH (2022) “Bayesian parameter estimation using conditional variational au-  
 5676 toencoders for gravitational-wave astronomy,” *Nature Phys.*, **18**(1), pp. 112–117,  
 5677 1909.06296.
- 5678 [78] REIMANN, R. (2022) “Project 8: R&D for a next-generation neutrino mass experi-  
 5679 ment,” *PoS, PANIC2021*, p. 283.
- 5680 [79] LAROQUE, B. (2020) “Zero-deadtime processing in beta spectroscopy for measure-  
 5681 ment of the non-zero neutrino mass,” *EPJ Web Conf.*, **245**, p. 01014.
- 5682 [80] BUZINSKY, N. (2021) *Statistical Signal Processing and Detector Optimization in*  
 5683 *Project 8*, Ph.D. thesis, Massachusetts Institute of Technology.

- 5684 [81] ESFAHANI, A. A. ET AL. (2023) “SYNCA: A Synthetic Cyclotron Antenna for  
5685 the Project 8 Collaboration,” *Journal of Instrumentation*, **18**(01), p. P01034.  
5686 URL <https://dx.doi.org/10.1088/1748-0221/18/01/P01034>
- 5687 [82] <https://www.nvidia.com/en-us/data-center/v100/>.
- 5688 [83] <https://www.nvidia.com/en-us/data-center/h100/>.
- 5689 [84] HE, K., X. ZHANG, S. REN, and J. SUN (2016) “Deep Residual Learning for  
5690 Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern  
5691 Recognition (CVPR)*, pp. 770–778.
- 5692 [85] SIMONYAN, K. and A. ZISSERMAN (2015) “Very Deep Convolutional Networks  
5693 for Large-Scale Image Recognition,” in *3rd International Conference on Learning  
5694 Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference  
5695 Track Proceedings* (Y. Bengio and Y. LeCun, eds.).  
5696 URL <http://arxiv.org/abs/1409.1556>
- 5697 [86] FRIIS, H. (1946) “A Note on a Simple Transmission Formula,” *Proceedings of the  
5698 IRE*, **34**(5), pp. 254–256.
- 5699 [87] POZAR, D. M. (2005) *Microwave engineering*; 3rd ed., Wiley, Hoboken, NJ.  
5700 URL <https://cds.cern.ch/record/882338>
- 5701 [88] AKER, M. ET AL. (2022) “Direct neutrino-mass measurement with sub-electronvolt  
5702 sensitivity,” *Nature Physics*, **18**(2), pp. 160–166.  
5703 URL <https://doi.org/10.1038/s41567-021-01463-1>
- 5704 [89] MONREAL, B. and J. A. FORMAGGIO (2009) “Relativistic Cyclotron Radiation  
5705 Detection of Tritium Decay Electrons as a New Technique for Measuring the  
5706 Neutrino Mass,” *Phys. Rev. D*, **80**, p. 051301, 0904.2860.
- 5707 [90] ASHTARI ESFAHANI, A. ET AL. (2017) “Determining the neutrino mass with  
5708 cyclotron radiation emission spectroscopy—Project 8,” *J. Phys. G*, **44**(5), p. 054004,  
5709 1703.02037.
- 5710 [91] ESFAHANI, A. A. ET AL. (2022) “The Project 8 Neutrino Mass Experiment,” in  
5711 *2022 Snowmass Summer Study*, 2203.07349.
- 5712 [92] ASHTARI ESFAHANI, A. ET AL. (2019) “Locust: C++ software for simulation of  
5713 RF detection,” *New Journal of Physics*, **21**(11), p. 113051.  
5714 URL <https://doi.org/10.1088/1367-2630/ab550d>
- 5715 [93] FURSE, D. ET AL. (2017) “Kassiopeia: a modern, extensible C++ particle tracking  
5716 package,” *New Journal of Physics*, **19**(5), p. 053012.  
5717 URL <https://doi.org/10.1088/1367-2630/aa6950>

- 5718 [94] BALANIS, C. (2011) *Modern Antenna Handbook*, Wiley.  
 5719 URL <https://books.google.com/books?id=UYpV8L8GNCwC>
- 5720 [95] WIRTH, W. (2001) *Radar Techniques Using Array Antennas*, Institution of Engi-  
 5721 neering and Technology.  
 5722 URL <https://books.google.com/books?id=ALht42gkzLsC>
- 5723 [96] BROWN, L. S., G. GABRIELSE, K. HELMERSON, and J. TAN (1985) “Cyclotron  
 5724 motion in a microwave cavity: Lifetime and frequency shifts,” *Phys. Rev. A*, **32**,  
 5725 pp. 3204–3218.  
 5726 URL <https://link.aps.org/doi/10.1103/PhysRevA.32.3204>
- 5727 [97] HANNEKE, D., S. FOGWELL HOOGERHEIDE, and G. GABRIELSE (2011) “Cavity  
 5728 control of a single-electron quantum cyclotron: Measuring the electron magnetic  
 5729 moment,” *Phys. Rev. A*, **83**, p. 052122.  
 5730 URL <https://link.aps.org/doi/10.1103/PhysRevA.83.052122>
- 5731 [98] HANNEKE, D. A. (2007) *Cavity control in a single-electron quantum cyclotron: an*  
 5732 *improved measurement of the electron magnetic moment*, Ph.D. thesis, Harvard U.
- 5733 [99] PURCELL, E. (1946) “Spontaneous Emission Probabilities at Radio Frequencies,”  
 5734 *Phys. Rev.*, **69**, pp. 674–674.  
 5735 URL <https://link.aps.org/doi/10.1103/PhysRev.69.674>
- 5736 [100] MOSKOWITZ, B. E. and J. ROGERS (1988) “ANALYSIS OF A MICROWAVE  
 5737 CAVITY DETECTOR COUPLED TO A NOISY AMPLIFIER,” *Nucl. Instrum.*  
 5738 *Meth. A*, **264**, pp. 445–452.
- 5739 [101] RAPIDIS, N. M. (2018) “Application of the Bead Perturbation Technique to a  
 5740 Study of a Tunable 5 GHz Annular Cavity,” *Springer Proc. Phys.*, **211**, pp. 45–51,  
 5741 1708.04276.
- 5742 [102] “Real-time Signal Detection for Cyclotron Radiation Emission Spectroscopy Mea-  
 5743 surements using Antenna Arrays,” *In preparation*.
- 5744 [103] BYRON, W. ET AL. (2022) “First observation of cyclotron radiation from MeV-scale  
 5745  $e^{pm}$  following nuclear beta decay,” 2209.02870.
- 5746 [104] KAZKAZ, K. and N. WOOLLETT (2021) “Using Cyclotron Radiation Emission  
 5747 for Ultra-high Resolution X-Ray Spectroscopy,” *New J. Phys.*, **23**(3), p. 033043,  
 5748 1911.05869.
- 5749 [105] CANNING, J. A. L., F. F. DEPPISCH, and W. PEI (2023) “Sensitivity of future  
 5750 tritium decay experiments to New Physics,” *JHEP*, **03**, p. 144, 2212.06106.

5751

## Vita

5752

**Andrew Douglas Ziegler**

5753

### **Education**

5754

- Doctor of Philosophy, Physics, The Pennsylvania State University, 2023
- Bachelor of Science, Physics, The University of Minnesota, 2017

5755

### **Selected Publications**

5756

### **Selected Presentations**

5757