

The Pennsylvania State University
The Graduate School
Eberly College of Science

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

A Thesis in
The Physics Department
by
Andrew Douglas Ziegler

10 © 2023 Andrew Douglas Ziegler

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

14 Doctor of Philosophy

15 December 2023

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

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

Carmen Carmona
Assistant Professor of Physics

¹⁷ Doug Cowan
Professor of Physics and Astrophysics

Xingjie Ni
Associate Professor of Electrical Engineering

¹⁸ Stephanie Wissel
Associate Professor of Physics

¹⁹ *Signatures are on file in the Graduate School.

²⁰ **Abstract**

²¹ Some shit goes here.

Table of Contents

23	List of Figures	viii
24	List of Tables	viii
25	Acknowledgments	viii
26	Chapter 1	
27	Introduction	1
28	1.1 Summary	1
29	1.2 Outline	4
30	Chapter 2	
31	Neutrinos and Neutrino Masses	6
32	2.1 Introduction	6
33	2.2 Neutrinos and Beta-decay	6
34	2.3 Neutrino Oscillations	7
35	2.4 Neutrino Masses in the Standard Model	10
36	2.5 Neutrino Absolute Mass Scale	13
37	2.5.1 Limits from Cosmology	13
38	2.5.2 Limits from Neutrinoless Double Beta-decay Searches	15
39	2.5.3 Limits from Beta-decay	17
40	Chapter 3	
41	Direct Measurement of the Neutrino Mass with Project 8	21
42	3.1 Introduction	21
43	3.2 Cyclotron Radiation Emission Spectroscopy and Project 8	22
44	3.2.1 Cyclotron Radiation Emission Spectroscopy — CRES	22
45	3.2.2 The Project 8 Collaboration	27
46	3.2.3 Project 8 Phased Development Plan	30
47	3.3 Phase II: First Tritium Beta Decay Spectrum and Neutrino Mass Measurement with CRES	34
49	3.3.1 The Phase II CRES Apparatus	34
50	3.3.2 CRES Track and Event Reconstruction	39
51	3.3.3 Results from Phase II	43
52	3.4 Phase III R&D: Antenna Array CRES	47

53	3.4.1	The Basic Approach	47
54	3.4.2	The FSCD: Free-space CRES Demonstrator	49
55	3.5	Pilot-scale Experiments	54
56	3.5.1	Choice of Frequency	54
57	3.5.2	Pilot-scale Experiment Concepts	57
58	3.6	Phase IV	59

59	Chapter 4		
60	Signal Reconstruction Techniques for Antenna Array CRES and the		
61	FSCD		61
62	4.1	Introduction	61
63	4.2	FSCD Simulations	62
64	4.2.1	Kassiopeia	63
65	4.2.2	Locust	67
66	4.2.3	CRESana	71
67	4.3	Signal Detection and Reconstruction Techniques for Antenna Array CRES	72
68	4.3.1	Digital Beamforming	76
69	4.3.2	Matched Filtering	86
70	4.3.3	Machine Learning	98
71	4.4	Analysis of Signal Detection Algorithms for the Antenna Array Demonstrator	103
72	4.4.1	Introduction	104
73	4.4.2	Signal Detection with Antenna Array CRES	106
74	4.4.2.1	Antenna Array and DAQ System	106
75	4.4.2.2	Real-time Signal Detection	109
76	4.4.3	Signal Detection Algorithms	113
77	4.4.3.1	Power Threshold	113
78	4.4.3.2	Matched Filtering	115
79	4.4.3.3	Machine Learning	119
80	4.4.4	Methods	121
81	4.4.4.1	Data Generation	121
82	4.4.4.2	Template Number and Match Estimation	122
83	4.4.4.3	CNN Training and Data Augmentation	123
84	4.4.5	Results and Discussion	124
85	4.4.5.1	Trigger Classification Performance	124
86	4.4.5.2	Computational Cost and Hardware Requirements	126
87	4.4.6	Conclusion	128

88	Chapter 5		
89	Antenna and Antenna Measurement System Development for the		
90	Project 8 Experiment		130
91	5.1	Introduction	130
92	5.2	Antenna Measurements for CRES experiments	131
93	5.2.1	Antenna Parameters	131
94	5.2.1.1	Radiation Patterns	131

95	5.2.1.2	Directivity and Gain	132
96	5.2.1.3	Far-field and Near-field	133
97	5.2.1.4	Polarization	134
98	5.2.1.5	Antenna Factor and Effective Aperture	135
99	5.2.2	Antenna Measurement Fundamentals	137
100	5.2.2.1	Friis Transmission Equation	137
101	5.2.2.2	S-Parameters and Network Analyzers	138
102	5.2.2.3	Antenna Array Commissioning and Calibration Measurements	139
103	5.2.3	The Penn State Antenna Measurement System	141
104	5.3	Development of a Synthetic Cyclotron Antenna (SYNCA) for Antenna Array Calibration	144
105	5.3.1	Introduction	144
106	5.3.2	Cyclotron Radiation Phenomenology	147
107	5.3.3	SYNCA Simulations and Design	153
108	5.3.4	Characterization of the SYNCA	158
109	5.3.5	Beamforming Measurements with the SYNCA	161
110	5.3.6	Conclusions	165
111	5.4	FSCD Antenna Array Measurements with the SYNCA	166
112	5.4.1	Introduction	166
113	5.4.2	Measurement Setups	167
114	5.4.2.1	FSCD Array Setup	167
115	5.4.2.2	Synthetic Array Setup	169
116	5.4.3	Simulations, Analysis, and Results	170
117	5.4.3.1	Simulations	170
118	5.4.3.2	Phase Analysis	171
119	5.4.3.3	Magnitude Analysis	176
120	5.4.3.4	Beamforming Characterization	178
121	5.4.4	Conclusions	181

124	Chapter 6		
125	Development of Resonant Cavities for Large Volume CRES Measurements	183	
126	6.1	Introduction	183
127	6.2	Cylindrical Resonant Cavities	184
128	6.2.1	General Field Solutions	184
129	6.2.2	TE and TM Modes	185
130	6.2.3	Resonant Frequencies of a Cylindrical Cavity	187
131	6.2.4	Cavity Q-factors	189
132	6.3	The Cavity Approach to CRES	193
133	6.3.1	A Sketch of a Molecular Tritium Cavity CRES Experiment	193
134	6.3.2	Magnetic Field, Cavity Geometry, and Resonant Modes	195
135	6.3.3	Trade-offs Between the Antenna and Cavity Approaches	198
136	6.4	Single-mode Resonant Cavity Design and Simulations	200

138	6.4.1	Open Cylindrical Cavities with Coaxial Terminations	200
139	6.4.2	Mode Filtering	203
140	6.4.3	Simulations of Open, Mode-filtered Cavities	205
141	6.5	Single-mode Resonant Cavity Measurements	208
142	6.5.1	Cavities and Setup	208
143	6.5.2	Results and Discussion	211
144	Chapter 7		
145	Conclusion and Future Prospects		214
146	Bibliography		216

¹⁴⁷ **Acknowledgments**

¹⁴⁸ Shout out to all the haters.

¹⁴⁹ **Dedication**

¹⁵⁰ Something heartfelt.

¹⁵¹ **Chapter 1** |
¹⁵² **Introduction**

¹⁵³ **1.1 Summary**

¹⁵⁴ Neutrinos are one of the fundamental particles that comprise the standard model of
¹⁵⁵ particle physics and account for a significant fraction of the matter in the universe.
¹⁵⁶ Neutrinos are the most abundant fermions in the universe, but due to their weak
¹⁵⁷ interactions neutrinos seldom interact with other particles. Regardless, neutrinos play a
¹⁵⁸ unique role in the evolution of the early-universe, therefore, a detailed understanding
¹⁵⁹ of the properties of the neutrino is important to understanding the cosmology of the
¹⁶⁰ universe as well as understanding the universe at the fundamental particle physics scale.

¹⁶¹ Unlike other fermions it was unclear that neutrinos had nonzero mass until neutrino
¹⁶² flavor oscillations were definitively observed in the late 90's and early 00's. Flavor
¹⁶³ oscillations require that neutrinos experience time so that when acted upon by the
¹⁶⁴ time-evolution operator the initial neutrino state can evolve to a new flavor state. This
¹⁶⁵ implies that the neutrino flavor states are really a superposition of at least three separate
¹⁶⁶ neutrino states with well-defined masses. Measurements of neutrino oscillations that have
¹⁶⁷ taken place over the past couple of decades have measured the differences between
¹⁶⁸ neutrino mass eigenstates with increasing precision. However, oscillation measurements
¹⁶⁹ cannot tell us the mass scale of the neutrinos, which is required in order to measure the
¹⁷⁰ absolute neutrino masses.

¹⁷¹ The neutrino mass scale remains an unknown quantity in the standard model of
¹⁷² particle physics. The value of the neutrino mass influences the evolution of the early
¹⁷³ universe and is likely relevant to the energy-scale of new physics responsible for the factor
¹⁷⁴ of 10^{-6} difference between the neutrino and electron masses. A model-independent way
¹⁷⁵ to measure the neutrino mass is to measure the tritium beta-decay spectrum near its
¹⁷⁶ endpoint. Energy conservation requires that the neutrino mass carry away some kinetic
¹⁷⁷ energy from the beta-decay electron in the form of its mass, which causes a distortion in

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

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

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

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

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

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

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

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

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

- 233 • the development and analysis of signal reconstruction algorithms for antenna array
234 CRES, which provided key inputs to sensitivity analyses of antenna array CRES
235 experiments,
- 236 • the development of a specialized antenna designed to synthesize fake CRES radia-
237 tion, which enabled bench-top testing and validation of the antenna array CRES
238 technique,
- 239 • the development of an open-cavity design for CRES measurement whose mode
240 structure can be tuned using perturbations that modify the impedance of the cavity
241 walls. The development of this cavity concept was one of many developments that
242 eventually lead to the adoption of cavities as the CRES technology of choice for
243 the future of Project 8.

²⁴⁴ 1.2 Outline

²⁴⁵ The outline of this dissertation is as follows. In Chapter 2 I provide an introduction to
²⁴⁶ the basic physics of neutrinos and beta-decay, which provides context for a discussion of
²⁴⁷ various methods to measure the neutrino absolute mass scale.

²⁴⁸ Chapter 3 is an overview of the CRES technique and the Project 8 collaboration.
²⁴⁹ I highlight the Project 8 Phase II experiment, which was the first measurement of
²⁵⁰ the tritium beta-decay spectrum with CRES, and I discuss the planned research and
²⁵¹ development for an antenna array CRES experiment in Phase III of the Project 8
²⁵² collaboration’s experiment plan. I end Chapter 3 with a discussion of the pilot-scale and
²⁵³ Phase IV experiments, that will combine a scalable CRES measurement technology with
²⁵⁴ atomic tritium and measure the neutrino mass with 40 meV sensitivity.

²⁵⁵ Chapter 4 discusses the first of the contributions mentioned above, which is the
²⁵⁶ development of signal reconstruction techniques for antenna array CRES and an antenna
²⁵⁷ array demonstrator experiment called the FSCD. I discuss the important tools that Project
²⁵⁸ 8 uses to simulate antenna array CRES before introducing three signal reconstruction
²⁵⁹ algorithms that can be used to detect CRES signals using the array. I end Chapter 4
²⁶⁰ with a paper that summarizes a detailed analysis and comparison of the signal detection
²⁶¹ performance of each algorithm.

²⁶² Chapter 5 describes my contributions to the development of antennas and an antenna
²⁶³ measurement system for Project 8, which is the second major contribution of this
²⁶⁴ dissertation. I begin with a general overview of basic principle of antennas and antenna
²⁶⁵ measurements, before including a paper that describes the development of unique antenna
²⁶⁶ designed to mimic the cyclotron radiation emitted by electrons in free-space when trapped
²⁶⁷ in a magnetic field. I call this antenna the synthetic cyclotron radiation antenna (SYNCA)
²⁶⁸ and its main purpose is to serve a fake electron for laboratory validation measurements
²⁶⁹ of Project 8’s antenna array CRES simulations. Chapter 5 ends with an overview
²⁷⁰ of laboratory measurements of a prototype antenna array that were compared with
²⁷¹ simulations to provide upper bounds on reconstruction errors caused by imperfections in
²⁷² real-life measurements.

²⁷³ Chapter 6 discusses the cavity approach to CRES, which was adopted as the preferred
²⁷⁴ CRES technology for Phase IV late into my dissertation work. The chapter stars by
²⁷⁵ discussing resonant cavities in general before introducing the operating principles of the
²⁷⁶ cavity approach to CRES. I end the chapter by discussing a study of and open-cavity
²⁷⁷ design that could be used for CRES measurements and integrated with atomic tritium

²⁷⁸ and an electron gun calibration source for the pilot-scale and Phase IV experiments.

²⁷⁹ Finally, in Chapter 7 I conclude by briefly discussing the future directions of the
²⁸⁰ Project 8 collaboration as we continue towards a direct measurement of the neutrino
²⁸¹ mass.

²⁸² **Chapter 2 |**

²⁸³ **Neutrinos and Neutrino Masses**

²⁸⁴ **2.1 Introduction**

²⁸⁵ In this chapter I provide a cursory overview of background information relevant to
²⁸⁶ neutrinos and neutrino mass measurements.

²⁸⁷ In Section 2.2 I provide some background information on the history of neutrinos and
²⁸⁸ beta-decay. In Section 2.3 I describe the discover of neutrino oscillations, which proved
²⁸⁹ unambiguously that neutrinos have non-zero masses. In Section 2.4 I discuss the current
²⁹⁰ state of the theoretical understanding of neutrino masses in the standard model. Lastly,
²⁹¹ in Section 2.5 I discuss methods for measuring the absolute scale of the neutrino mass.

²⁹² **2.2 Neutrinos and Beta-decay**

²⁹³ Late in the 19th century the phenomena of radioactivity was first observed in experiments
²⁹⁴ performed by Henri Becquerel with uranium, and further studied using thorium and
²⁹⁵ radium by Marie and Pierre Curie. Early work in radioactivity classified different forms
²⁹⁶ of radiation based on it's ability to penetrate different materials. Rutherford was the first
²⁹⁷ to separate radioactive emissions into two types, alpha and beta radiation. Alpha rays
²⁹⁸ can be easily stopped by a piece of paper or thin foil of metal, whereas beta radiation
²⁹⁹ could penetrate metals several millimeters thick. Later a third form of radiation was
³⁰⁰ identified by Villard, which was still more penetrating, and was eventually termed gamma
³⁰¹ radiation by Rutherford.

³⁰² When these forms of radioactivity were first discovered it was unclear what physically
³⁰³ constituted an alpha, beta, or gamma particle. Experiments with radioactivity in
³⁰⁴ magnetic fields was eventually able to identify the charge composition of different forms
³⁰⁵ of radiation. In particular, experiments by Becquerel identified that beta radiation had

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

308 Further studies of beta radiation lead to the discovery that radioactivity resulted
309 in the transmutation of elements caused by the decay of a heavier nucleus to a lighter
310 species. One feature of beta radiation, which we now properly call beta-decay, that
311 was different from alpha-decays and gamma radiation is that the electrons produced by
312 beta-decay have a continuous spectrum of kinetic energies, whereas, alpha and gamma
313 particles are emitted with discrete energies. This feature of beta-decay was first observed
314 by Chadwick in 1914, and was extremely puzzling at the time since the continuous
315 spectrum apparently violates energy conservation.

316 Famously, in 1930 Pauli proposed the existence of a new neutral particle, which he
317 termed the "neutron", that was also produced during beta-decay in order to resolve the
318 missing energy problem posed by the beta-decay spectrum. Because this particle carried
319 no charge, it was hypothesized at the time that it had simply not been observed in any
320 experiments up to that time. This "neutron", which was initially estimated to have a
321 mass no larger than that of an electron, was eventually renamed the "neutrino" by Fermi
322 after the discovery of the neutron by Chadwick in 1932. Later, in 1933, Fermi developed
323 a quantum mechanical theory for beta-decay in which both an electron and neutrino are
324 produced by the decay of a neutron to a proton inside the radioactive nucleus.

325 Little more than a speculation when first introduced, indirect evidence for the existence
326 of neutrinos was obtained in 1938 by the simultaneous observation of the electron and
327 recoiling nucleus in cloud chambers by Crane and Halpern. However, it wasn't until the
328 Cowan-Reines experiment in 1956 that direct evidence for the existence of neutrinos
329 was observed by detecting the inverse beta-decays caused by neutrinos from a nuclear
330 reactor interacting with protons contained in water molecules. The difficulty in detecting
331 neutrinos is caused by their weak interactions with other particles. Further, experiments
332 revealed that different types of neutrinos existed based on the nature of the leptons
333 produced in neutrino charged-current interactions, but the existence of a neutrino mass
334 remained an open question that would take more than 40 year to resolve.

335 **2.3 Neutrino Oscillations**

336 The first hint of neutrino flavor transitions or neutrino oscillations was indicated by
337 the solar neutrino problem, which referred to discrepancies between the predicted flux
338 of ν_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. 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. 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.

The origin of neutrino oscillations is that the weak eigenstates are distinct from the mass eigenstates. 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.

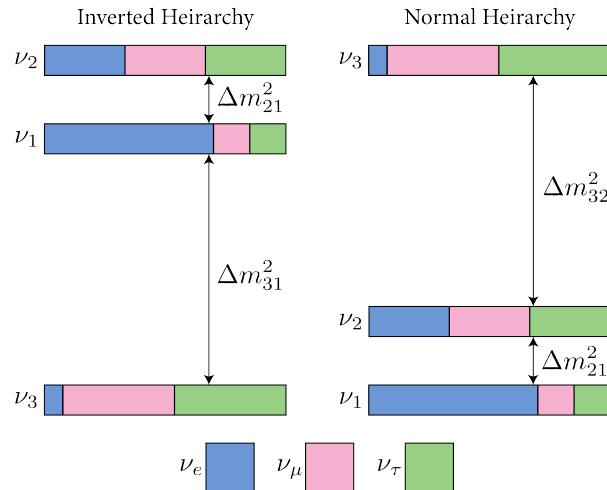


Figure 2.1: A diagram of two different neutrino mass ordering scenarios. In the inverted hierarchy (inverted mass ordering) the lightest neutrino mass is m_3 , whereas, in the normal hierarchy (normal mass ordering) m_1 is the lightest neutrino. What cannot be measured by neutrino oscillations is the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

356 One standard parameterization of the PMNS matrix is

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

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

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

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

365 A giant experimental effort over the past couple of decades has greatly contained the
366 majority of parameters in the PMNS matrix, many to relative uncertainties of only a
367 few percent. However, some parameters still remain relatively unconstrained, which is
368 the origin of the current uncertainty in the ordering of the neutrino masses (see Figure
369 2.1). The neutrino masses can be organized by their relative mass. The current neutrino
370 oscillation data can confirm that $m_2 > m_1$, however, the sign of Δm_{32}^2 is still unknown.
371 This leads to two scenarios where neutrino masses follow the ordering $m_3 > m_2 > m_1$,
372 which is called the normal mass ordering (NMO), or alternatively neutrino masses may
373 be ordered $m_2 > m_1 > m_3$, which is called the inverted mass ordering (IMO). Next-
374 generation neutrino oscillation experiments such as JUNO, Hyper-Kamiokande, and
375 DUNE are poised to resolve this ambiguity in the coming years.

376 Neutrino oscillation probabilities are only sensitive to the neutrino masses via the
377 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 with 1σ -uncertainties for the squared mass differences are

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

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

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_{least}) take on a range of values one can visualize the relative masses of the neutrinos as a function of m_{least} (see Figure 2.2).

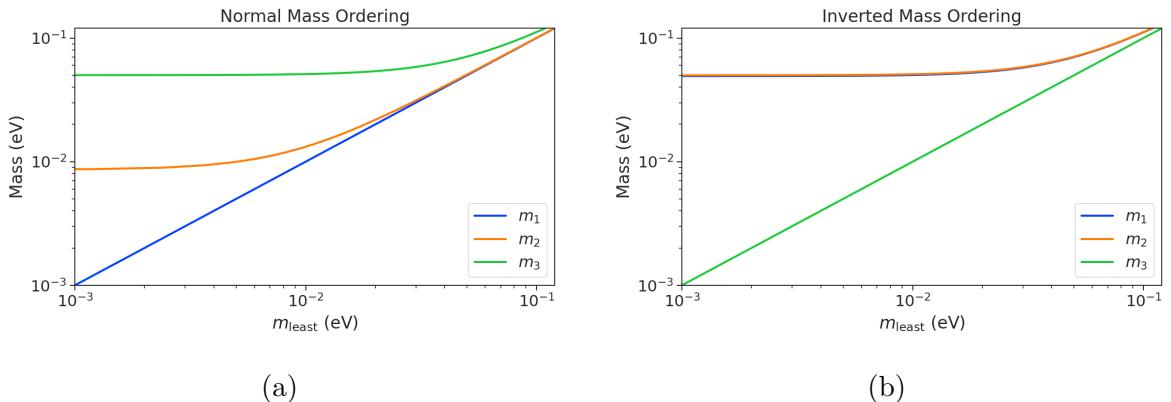


Figure 2.2: The masses of the neutrinos as a function of the lightest neutrino mass in both the normal (a) and inverted (b) mass ordering regimes.

2.4 Neutrino Masses in the Standard Model

Neutrinos are spin 1/2 particles and in modern quantum field theory spin-1/2 particles, or fermions, are described using the Dirac equation.

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

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

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

392 where Y_{ij}^ℓ is an element of the 3×3 Yukawa coupling matrix for leptons, L_{Li} is the
 393 left-handed lepton doublet for generation i , ϕ is the Higgs doublet, and E_{Rj} is the
 394 right-handed lepton field for generation j . In the standard model neutrinos are only
 395 represented as left-handed neutrinos and right-handed antineutrinos, consistent with
 396 experimental observations. Since there are no right-handed neutrino singlet fields and
 397 no Yukawa interaction terms for neutrinos are strictly massless, and non-zero neutrino
 398 masses is evidence for physics beyond the standard model. For the charged leptons, the
 399 Yukawa interaction leads to masses of the form

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

400 where v is the Higgs vacuum expectation value.

401 The observation of massive neutrinos motivates the extension of the standard model
 402 to explain the origin of neutrino masses, which can be approached in different way, but
 403 all methods add additional degrees of freedom to the standard model. One approach
 404 is to introduce to the standard model a right-handed neutrino field that allows one to
 405 introduce Yukawa terms of the form

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

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

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

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

414 An alternative approach is to allow the neutrinos to have a Majorana mass, which is

⁴¹⁵ possible because neutrinos are electrically neutral particles. The Majorana mass terms
⁴¹⁶ for the neutrino have the form

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

⁴¹⁷ where M_{Rij} and M_{Lij} are right-handed and left-handed Majorana mass matrices. A
⁴¹⁸ consequence of neutrinos being Majorana particles is lepton number violation, which
⁴¹⁹ predicts the occurrence of neutrino-less double beta-decay at a rate proportional to the
⁴²⁰ neutrino mass.

⁴²¹ In the most general case neutrinos have both Dirac and Majorana mass terms, which
⁴²² allows one to generate neutrino masses with Yukawa couplings similar to the rest of
⁴²³ the standard model. Considering just one generation of neutrinos for illustration, the
⁴²⁴ combined Lagrangian can be written as

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

⁴²⁵ or equivalently,

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

⁴²⁶ An example mass generation mechanism with this approach is the Type-I see-saw
⁴²⁷ mechanism, in which we take $m_L = 0$ and $m_R \gg m_D$. By diagonalizing Equation 2.14
⁴²⁸ one obtains the mass eigenvalues that represent the physical masses of the neutrinos.
⁴²⁹ The light neutrino mass eigenstate, which represents the observed neutrino mass, has a
⁴³⁰ mass given by

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

⁴³¹ and the heavy neutrino mass eigenstate, which represents the unobserved sterile neutrino,
⁴³² has a mass

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

⁴³³ For m_D similar to the other quark or lepton masses, one obtains physical neutrino masses
⁴³⁴ consistent with observations from sterile neutrino masses of $m_R \approx O(10^{15})$ GeV. This
⁴³⁵ mass scale is well beyond the capabilities of modern particle accelerators.

⁴³⁶ 2.5 Neutrino Absolute Mass Scale

⁴³⁷ The neutrino absolute mass scale or simply "neutrino mass" cannot be probed with
⁴³⁸ neutrino oscillations, since oscillation probabilities are determined by the squared mass
⁴³⁹ differences between neutrino mass eigenstates, therefore, alternative techniques are needed
⁴⁴⁰ to perform an effective measurement of the neutrino mass.

⁴⁴¹ 2.5.1 Limits from Cosmology

⁴⁴² In the Λ CDM model, which summarizes our current cosmological understanding of our
⁴⁴³ universe, the mass-energy content of the universe is composed of approximately 27%
⁴⁴⁴ dark matter and only 5% normal matter including neutrinos. From this observation, a
⁴⁴⁵ rough limit on the neutrino mass can be obtained from the condition that neutrinos are
⁴⁴⁶ not responsible for the entirety of the matter content of the universe. Using only this
⁴⁴⁷ condition one can constrain the neutrino mass to be ...

⁴⁴⁸ A prediction of the Λ CDM model is that the universe originated from a single
⁴⁴⁹ expansion event colloquially called the "Big Bang". In the Big Bang scenario, our
⁴⁵⁰ universe originated as a hot spacetime singularity, which abruptly experience rapid
⁴⁵¹ expansion in a process called inflation. After the inflationary epoch the universe entered
⁴⁵² the reheating phase where the potential energy responsible for inflation decays into
⁴⁵³ standard model particles such as electrons, quarks, and gluons. The universe continued to
⁴⁵⁴ expand in size resulting in a decrease in energy density and lower temperature. Eventually
⁴⁵⁵ the temperature of the universe decreased enough to allow the formation of protons,
⁴⁵⁶ neutrons, and other baryons from quarks and gluons produced from the decays of the
⁴⁵⁷ inflationary fields.

⁴⁵⁸ Also produced during the Big Bang are electrons, neutrinos and other leptons as
⁴⁵⁹ well as a population of photons. These particles are kept in thermal equilibrium with
⁴⁶⁰ the rest of the quark-gluon plasma through interactions that take place at the high
⁴⁶¹ temperatures and densities of the early universe. However, as the universe continues
⁴⁶² to expand it's density and temperatures decreases leading to the eventual decoupling
⁴⁶³ of photons and leptons from the quarks and gluons. A prediction of inflation is that
⁴⁶⁴ this population of photons produced during the Big Bang should still be present, but
⁴⁶⁵ with a significantly reduced temperature due to the expansion of the universe. This is
⁴⁶⁶ consistent with the observation of the CMB (cosmic microwave background), which is a
⁴⁶⁷ population of microwave radiation with a blackbody temperature of 2.7 K. The CMB
⁴⁶⁸ 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), 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 somewhat unique role in the Λ CDM model due to the fact that neutrinos act as radiation early in the universe but as matter in the late universe. This leads to unique signatures that impact anisotropies of the CMB as well as the distribution of matter in the universe. By combining measurements of the CMB with measurements of the large-scale structure (LSS) of the universe one can constrain the neutrino mass scale by fitting these datasets with the Λ CDM model. This analysis results in some of the most stringent constraints on the neutrino mass. A recent analysis was able to constrain the neutrino mass scale to

$$\Sigma_{m_\nu} \equiv \sum_i m_i < 0.12 \text{ eV}, \quad (2.17)$$

where m_i are the neutrino mass eigenstates.

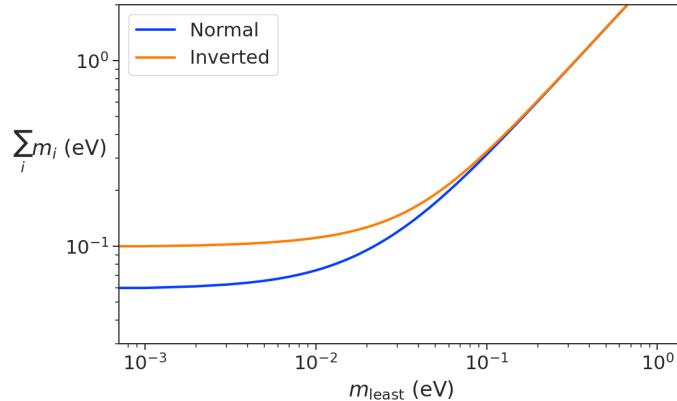


Figure 2.3: The neutrino mass observable measured by cosmology as a function of the lightest neutrino mass eigenstate.

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

⁴⁸⁷ the form

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

⁴⁸⁸ where it is clear that a measurement of Σ_{m_ν} effectively sets the neutrino mass scale
⁴⁸⁹ through m_{least} . The analogous formula for the inverted mass ordering is

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

⁴⁹⁰ In figure 2.3 we plot the observable Σ_{m_ν} as a function of m_{least} .

⁴⁹¹ Upcoming experiments are planned to refine measurements of the CMB, LSS, and
⁴⁹² other cosmological observables. With this additional data it is possible that in the
⁴⁹³ near future cosmological measurements will be able to positively constrain the neutrino
⁴⁹⁴ absolute mass scale. However, the strength of these limits strictly depend on the accuracy
⁴⁹⁵ of the Λ CDM model, which highlights the need for direct experimental measurements of
⁴⁹⁶ the neutrino mass to confirm the predictions of cosmology and to fix the neutrino mass
⁴⁹⁷ parameter in future cosmological analyses.

⁴⁹⁸ 2.5.2 Limits from Neutrinoless Double Beta-decay Searches

⁴⁹⁹ If neutrinos are Majorana fermions then the neutrino is equivalent to its own antiparticle
⁵⁰⁰ and lepton conservation is not an exact law of nature. Searches for lepton number
⁵⁰¹ violation, specifically the neutrinoless double beta-decay ($0\nu\beta\beta$) process, are some of the
⁵⁰² most powerful tests of lepton number conservation, which depend on the neutrinos being
⁵⁰³ Majorana fermions. In double beta-decay two neutrons contained in the decay species
⁵⁰⁴ nucleus spontaneously decay into two protons resulting in the production of two electrons
and two neutrinos (see Figure 2.4). However, for $0\nu\beta\beta$ the two neutrinos self-annihilate

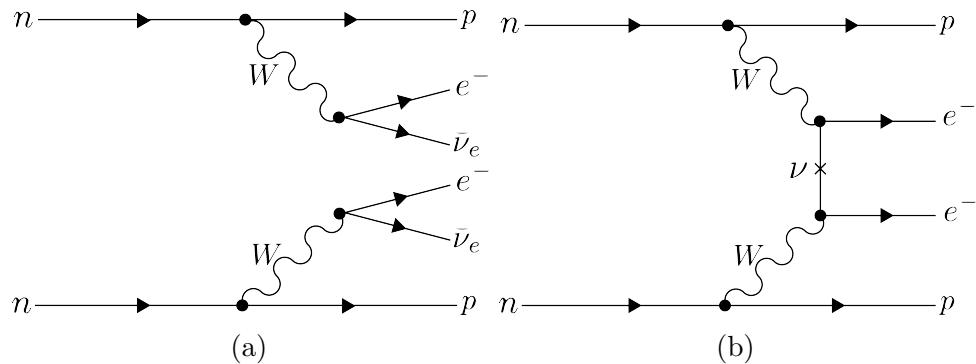


Figure 2.4: Feynman diagrams for double beta-decay (a) and $0\nu\beta\beta$ (b).

⁵⁰⁵

506 during the decay resulting only in the production of two electrons and a violation of
 507 lepton number by two.

508 Assuming that the exchange of two Majorana neutrinos is the dominant channel for
 509 $0\nu\beta\beta$, then a measurement of the $0\nu\beta\beta$ half-life for a particular isotope can be used to
 510 set the neutrino absolute mass scale. The half-life is written in terms of the effective
 511 neutrino mass for $0\nu\beta\beta$ ($m_{\beta\beta}$) using the equation

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

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

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

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

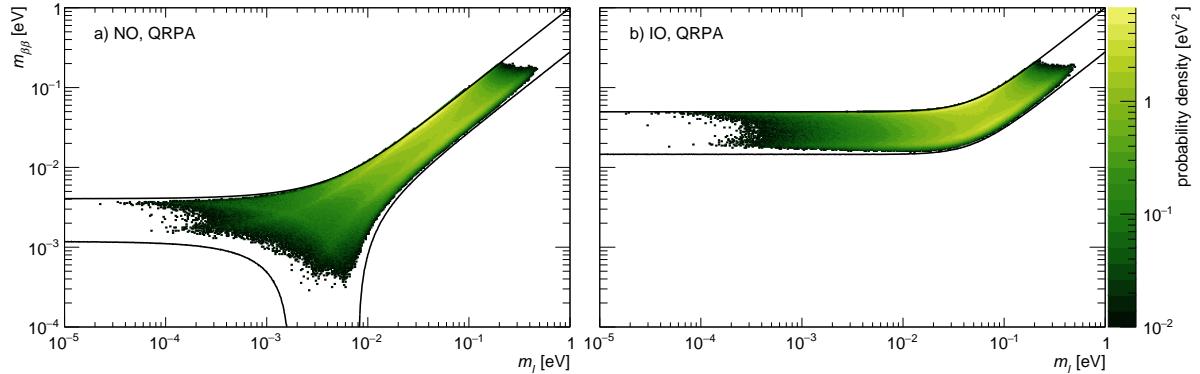


Figure 2.5: The discovery probabilities for the future generation of $0\nu\beta\beta$ experiments as
 a function of $m_{\beta\beta}$ and m_{least} .

519
 520 of cancellation due to the unknown Majorana phases included in the sum specified by
 521 Equation 2.21, the information gained is necessarily imperfect. Additionally, theoretical
 522 uncertainty in the calculation of the nuclear matrix elements complicates the calculation
 523 of $m_{\beta\beta}$ from a measurement of $0\nu\beta\beta$ half-life. Similar to cosmology there is a high degree

524 of complementarity between direct measurements of the neutrino mass and $0\nu\beta\beta$. In
 525 particular, a measurement of m_{least} to less than than 0.1 eV sensitivity provides significant
 526 information for $0\nu\beta\beta$ searches based of the discovery probabilities of Figure 2.5.

527 2.5.3 Limits from Beta-decay

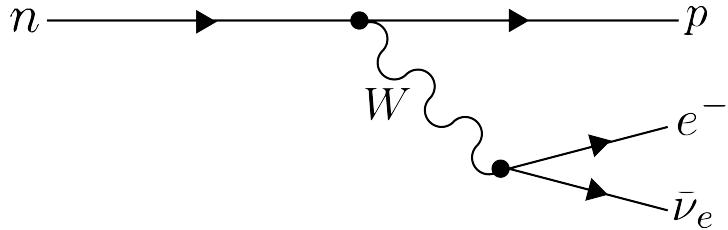


Figure 2.6: A Feynman diagram of beta decay

528 Certain processes involving neutrinos, in particular beta-decay (see Figure 2.6), have
 529 initial states with well-defined total energies and final states where the kinetic energies
 530 of each of the particles can be measured with high accuracy and precision. Beta-decay
 531 involves the decay of an unstable isotope where a neutron spontaneously converts to
 532 a proton and emits and electron and anti-neutrino ("neutrino" for brevity) to conserve
 533 charge and lepton number. Therefore, by applying the principles of energy and momentum
 534 conservation a measurement of the kinematics of the final state can be used to constrain
 535 the neutrino mass as proposed by Fermi in his 1934 description of nuclear beta-decay
 (see Figure 2.7). Because the constraint on the neutrino mass from beta-decay depends

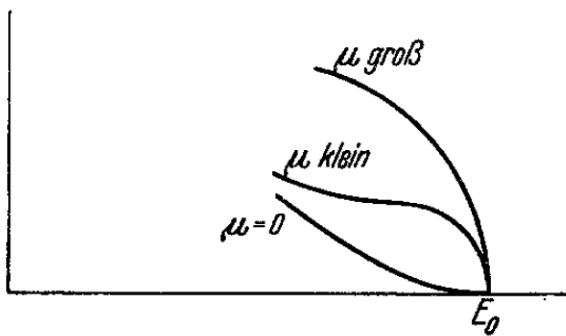


Figure 2.7: A figure from Fermi's 1934 paper on a theory of beta-decay depicting the kinetic energy spectrum of the emitted electron. The effect of the neutrino mass, written as μ , is to distort the shape of the spectrum near the endpoint from the zero-mass spectrum.

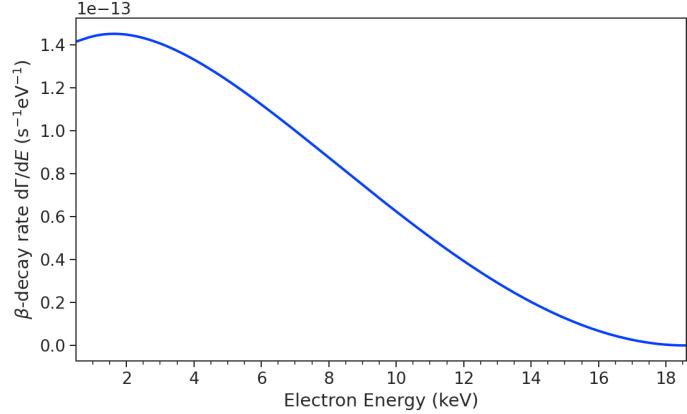
only on the final state measurement capabilities, such measurements of the neutrino mass are often referred to as model-independent or direct in contrast to constraints on the neutrino mass from cosmology and $0\nu\beta\beta$.

The beta-decay isotope of choice for direct neutrino mass measurements with beta-decay has been tritium (3H_2) for many decades, because it conveniently fulfills many experimental requirements. Of upmost importance is a decay with a low Q-value, which is the available kinetic energy based on the mass difference between the initial and final states. The effect of a massive neutrino on the shape of the spectrum is magnified for low Q-values and tritium decays have an unusually low Q-value of 18.6 keV. Additionally, tritium beta-decay is a super-allowed decay, which means that it has a relatively short half-life of 12.3 years making it easy to obtain a high-activity source with a relatively small source mass. High-activity is desireable due to the low-activity near the tritium spectrum endpoint. For tritium beta-decays only a factor of 3×10^{-13} of the decays occur in the last 1 eV of the spectrum. Isotopes with Q-values lower than tritium are known, but this is outweighed by exceedingly long half-lives leading to unobtainable source masses.

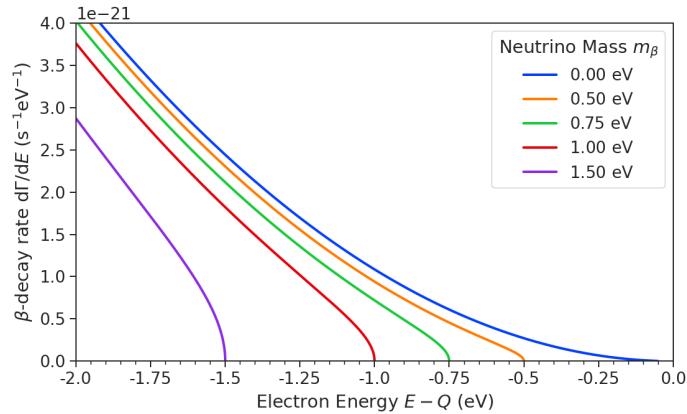
Neutrino mass measurements using beta-decay measure the effect of the neutrino's mass on shape of the electron's kinetic energy spectrum near the endpoint. The kinetic energy spectrum (see Figure ??) is given by

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

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



(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.

565

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

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

566 where m_{beta} is the effective mass of the neutrino in beta-decay or simply neutrino mass
 567 for brevity. By assuming unitarity, the neutrino mass can be expressed in terms of
 568 the PMNS matrix elements, squared mass differences, and the lightest neutrino mass

⁵⁶⁹ eigenstate. For the normal mass ordering the equation is

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

⁵⁷⁰ and for the inverted ordering the equation is

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

⁵⁷¹ Therefore, a measurement of the neutrino mass in combination with neutrino mixing
⁵⁷² parameters is effectively a measurement of m_{least} .

⁵⁷³ Since the neutrino mass is small (< 1 eV), it's effect on the spectrum is limited to
⁵⁷⁴ the endpoint region. The affect of a non-zero neutrino mass on the endpoint spectrum is
⁵⁷⁵ plotted for the reader in Figure 2.8. Resolving the small changes in the spectrum shape
⁵⁷⁶ requires an experimental technique with high statistics, excellent energy resolution, and
⁵⁷⁷ low background activity.

⁵⁷⁸ The KATRIN collaboration, utilizing a large MAC-E (magnetic adiabatic collimation
⁵⁷⁹ with electrostactic) filter spectrometer recently obtained the best direct measurement of
⁵⁸⁰ the neutrino mass, with a 90% confidence upper limit of 0.8 eV. With more statistics the
⁵⁸¹ KATRIN collaboration estimates an ultimate sensitivity to neutrino masses of 0.2 eV.

582 **Chapter 3** |

583 **Direct Measurement of the Neutrino Mass**

584 **with Project 8**

585 **3.1 Introduction**

586 A promising technique for direct measurements of the neutrino mass beyond the projected
587 limit of the ongoing KATRIN experiment is tritium beta-decay spectroscopy with an
588 atomic tritium source [1]. Atomic tritium, combined with a large-volume, high-resolution
589 energy measurement technique, is capable of measuring the neutrino mass with sensitivity
590 below the 50 meV limit allowed by neutrino oscillations.

591 Cyclotron Radiation Emission Spectroscopy or CRES is a high-resolution energy
592 measurement technique compatible with atomic tritium production and storage that can
593 enable the next-generation of neutrino mass direct measurement experiments [2]. The
594 Project 8 collaboration is currently engaged in a program of research and development
595 (R&D) aimed at developing the technology necessary for a 40 meV sensitivity measurement
596 of the neutrino mass using CRES and atomic tritium [3].

597 In Section 3.2 we provide an introduction to the basics of the CRES technique
598 as well as the goals of the Project 8 experiment. Additionally, we sketch out the
599 phased experiment development plan being implemented by Project 8 to build towards a
600 next-generation neutrino mass experiment.

601 In Section 3.3 we give a brief overview of Phase II of the Project 8 experiment [4, 5],
602 which completed early in 2023. Although the bulk of the work presented in this thesis is
603 relevant to designs of future Project 8 experiments, a description of the work in Phase II
604 provides useful context for the rest of the work.

605 In Section 3.4 we introduce a CRES measurement concept based on antenna arrays [6],
606 which could be the basis for the ultimate Project 8 neutrino mass experiment. A significant
607 portion of the R&D efforts of Project 8 in Phase III were directed towards simulating

608 and modeling this experimental concept in order to understand the achievable sensitivity
609 to the neutrino mass.

610 Lastly, in Section 3.5 we introduce conceptual designs of pilot-scale experiments that
611 combine atomic CRES with a large-volume CRES detection technique. This includes a
612 design concept for an antenna array based experiment, but also a design for a resonant
613 cavity based experiment. Resonant cavities are discussed in more depth in Chapter 6
614 and have become the preferred choice for future CRES experiments in Project 8 over
615 antenna arrays.

616 **3.2 Cyclotron Radiation Emission Spectroscopy and Project** 617 **8**

618 **3.2.1 Cyclotron Radiation Emission Spectroscopy — CRES**

619 Of the standard physical quantities the one that can be measured with the highest
620 precision is time and the inversely related quantity frequency. In fact it is often advan-
621 tageous to convert measurements of other physical quantities like mass or length into
622 frequency measurements due to the digital nature of frequency measurements that make
623 them immune to many sources of noise. Atomic clocks, which operate by measuring the
624 frequencies of various atomic transitions, have been used to measure time with astounding
625 relative uncertainties of 10^{-18} seconds. The extreme precision possible with frequency
626 measurements is often summarized using the a quote from the Physicist Arthur Schawlow
627 who said advise his students to "Never measure anything but frequency!".

628 Neutrino mass measurements using tritium beta-decay require us to measure a
629 perturbation of the 18600 eV tritium endpoint to precisions as low as 0.1 eV, therefore, a
630 spectroscopic technique with extremely high resolution is required for this measurement.
631 Part of the reason that frequency measurements are capable of such high resolutions is
632 that they are essentially counting measurements, which average the number of oscillations
633 of a physical system over time. By observing a rapidly oscillating system over a sufficient
634 length of time one can obtain essentially arbitrary precision on a frequency limited only
635 by the time available for measurement and the SNR of the system.

636 In order to perform frequency-based high-resolution spectroscopy of the tritium beta-
637 decay spectrum one needs to translate the kinetic energy of the electron into a frequency.
638 The simplest way to accomplish this is to place a gaseous supply of tritium into a magnetic
639 field. When one of the atoms decays the resulting electron will immediately begin to

640 orbit around a magnetic field line at the cyclotron frequency which is proportional to
 641 its kinetic energy (see Figure 3.1). The acceleration caused by the orbit leads to the
 642 emission of cyclotron radiation that can be detected using an array of antennas or a
 643 different RF sensor such as a resonant cavity. The frequency of the radiation gives the
 644 electron's kinetic energy, which is used to build the beta-decay spectrum and measure
 645 the neutrino mass. The name for this measurement technique is Cyclotron Radiation
 646 Emission Spectroscopy or CRES.

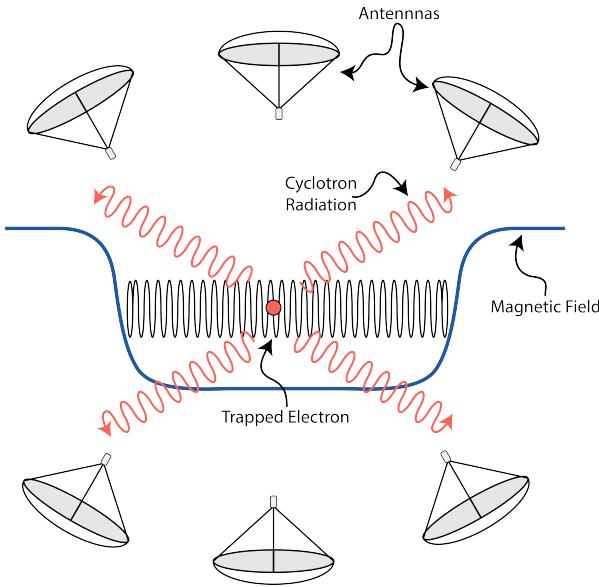


Figure 3.1: A cartoon illustration of the CRES technique. An electron is contained in a magnetic trap so that its cyclotron radiation can be detected by an array of antennas. Detecting the cyclotron radiation allows us to measure its cyclotron frequency and determine its kinetic energy.

647 For non-relativistic particles the cyclotron frequency is only a function of the charge-
 648 to-mass ratio of the particle, however, from the relativistic form of the cyclotron frequency

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

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

652 The required frequency resolution needed for neutrino mass measurement can be

653 obtained by differentiating Equation 3.1,

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

654 from which we can obtain the relationship between fractional differences in energy and
655 frequency,

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

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

658 The minimum observation time required to achieve this resolution can be estimated
659 using the uncertainty principle as formulated by Gabor. Electron's from tritium beta-
660 decay experience random collisions with the background gas particles, which limits the
661 uninterrupted radiation lifetime. The time between collision events, referred to as track
662 length in the context of CRES measurements, is an exponentially distributed variable.
663 Differences in the track lengths of a population of mono-energetic electrons leads to
664 uncertainty or broadening in the distribution of measured frequencies proportional to
665 the mean track length, τ_λ . The resulting frequency distribution has a Lorentzian profile,
666 whose width is given by the Gabor limit,

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

667 The cyclotron frequency for a 18.6-keV electron in a 1 T field is approximately
668 27 GHz, from which one can estimate the minimum observation time for 2 ppm frequency
669 resolution at approximately 3 μ sec. The Gabor limit is not the true lower bound on the
670 frequency resolution for a CRES signal, since it is based on the details of the Fourier
671 representation of a time-series with a fixed length. If one takes the approach of fitting the
672 CRES signal in the time-domain, then one finds that the limit on frequency precision is
673 given by the Cramér-Rao lower bound (CRLB), which depends on both the track length
674 as well as the SNR. In general, the CRLB allows for better precision on the cyclotron
675 frequency, however, the Gabor limit provides an illustrative limit with the correct order
676 of magnitude.

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

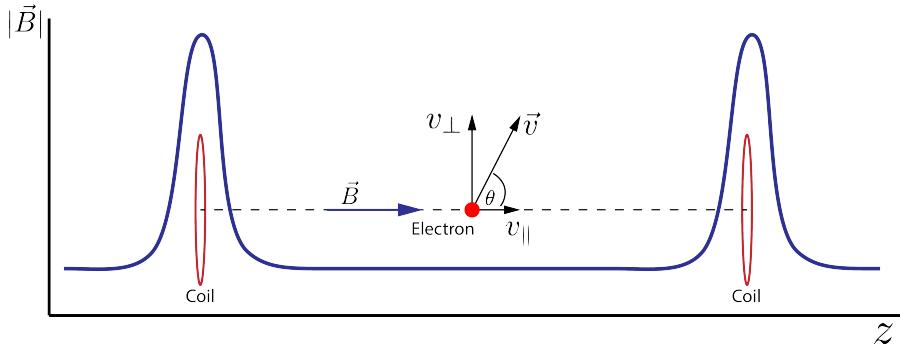


Figure 3.2: An illustration of an electron in a bathtub magnetic trap generated by two well-separated coils.

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 its 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 generates 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.

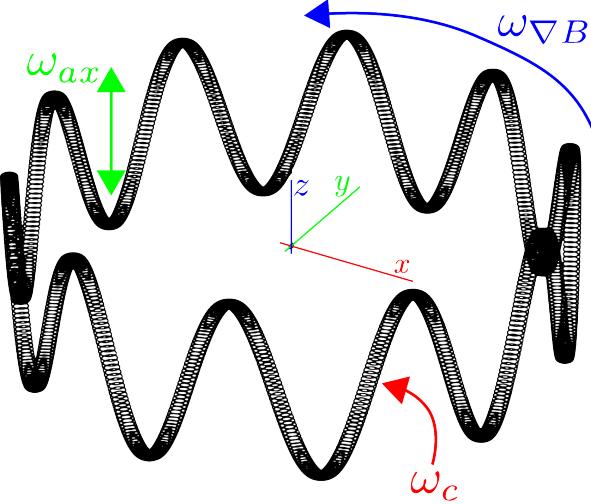


Figure 3.3: A plot of the main components of an electron's trajectory in a cylindrically symmetric trap.

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 oftentimes leads to a third component of motion called grad-B drift. The equation for the drift velocity is

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

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

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

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

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

719 signals due to the increase in the total signal energy available for the detection algorithm.

720 **3.2.2 The Project 8 Collaboration**

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

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

732 **Atomic Tritium**

733 Previous measurements of the tritium beta-decay spectrum for neutrino mass measure-
734 ments have all relied on a sources of molecular tritium for their measurements due to the
735 numerous practical and technical challenges associated with the production and storage
736 of hydrogen isotopes.

737 To produce atomic hydrogen one must supply sufficient energy to the tritium molecule
738 to break the molecular bond between. Common approaches to this include the use of hot
739 coaxial filament atom crackers as well as plasma atom sources. Both approaches heat the
740 tritium atoms to temperatures > 2500 K, which must then be cooled to temperatures
741 on the order of a few mK so that the tritium atoms can be trapped. Cooling the atoms
742 requires the construction of a large tritium infrastructure and cooling system that can
743 supply a source of cold atoms to the trap.

744 Once cold tritium atoms are produced they cannot make contact with any surfaces
745 to avoid recombination of the atoms to molecules. Therefore, a magnetic trap is required
746 to store the atoms for a sufficient length of time that they have a chance to decay before
747 escaping the trap. Trapping the atoms at this scale requires the construction of a large
748 and complex magnet system that must be cooled to cryogenic temperatures to avoid
749 heating of the atoms, which leads to their escape from the trap.

750 The significant experimental complexity caused by atomic tritium makes a molecular

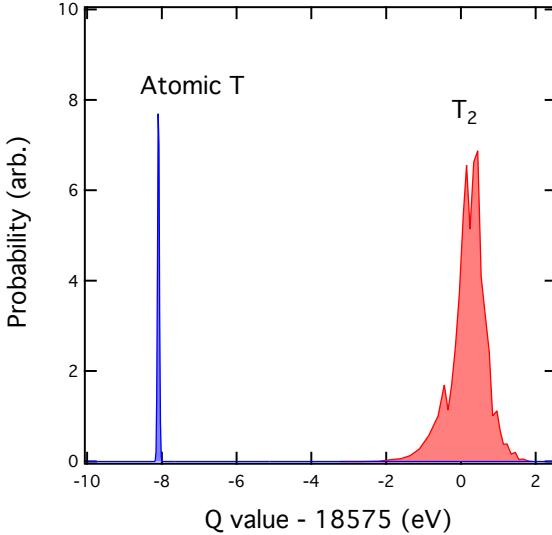


Figure 3.4: A plot of the final state distributions of atomic and molecular tritium. The final state distribution provides the primary contribution to the width of the molecular spectrum whereas thermal doppler broadening is responsible for the width of the atomic spectrum.

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

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

766 CRES for Neutrino Mass Measurement

767 Several promising features of the CRES technique make it a particularly attractive choice
768 for a next generation neutrino mass measurement experiment. For example, with a CRES

769 experiment the volume of the source gas can be the same as the volume of the CRES
770 spectrometer. This is due to the fact that CRES is a remote-sensing technique that can
771 observe the energy of the electron without altering its trajectory or directly interacting
772 with the electron. Given that tritium gas is transparent to cyclotron radiation the kinetic
773 energies of electrons can be measured with an appropriate sensing technology, such as a
774 cavity or antenna array, located directly outside the atom trapping volume.

775 The current state-of-the-art tritium beta-decay spectroscopy experiment, KATRIN,
776 utilizes the magnetic adiabatic collimation with an electrostatic filter (MAC-E filter)
777 technique to measure the beta-decay spectrum of molecular tritium. In this approach, a
778 source of molecular tritium is located outside of the spectrometer. When a beta-decay
779 occurs the electron must exit the tritium source and travel through the MAC-E filter
780 before it can be detected on the other side of the filter using a charge sensor. With this
781 approach the measurement statistics are limited by the transverse areas of the tritium
782 source and MAC-E filter due to the need to travel through the detector without scattering.
783 This scaling is less favorable than the volumetric scaling that one has with CRES due to
784 the ability to co-locate source and detector.

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

794 A final aspect of the CRES technique that is attractive for a next-generation experi-
795 ment is the relative immunity to backgrounds. Since CRES operates via non-destructive
796 measurements of the electron's cyclotron frequency potential sources of background elec-
797 trons are effectively filtered out by limiting the frequency bandwidth of the measurement.
798 The fiducial volume of the experiment is free from any surfaces that could introduce
799 stray electrons and electrons from sources outside the fiducial volume can be prevented
800 from entering the experiment.

801 **Neutrino Mass Sensitivity Goals**

802 Project 8's ultimate goal is to combine CRES with atomic tritium to measure the neutrino mass with 40 meV sensitivity at the 90% confidence level (see Figure 3.5). This sensitivity

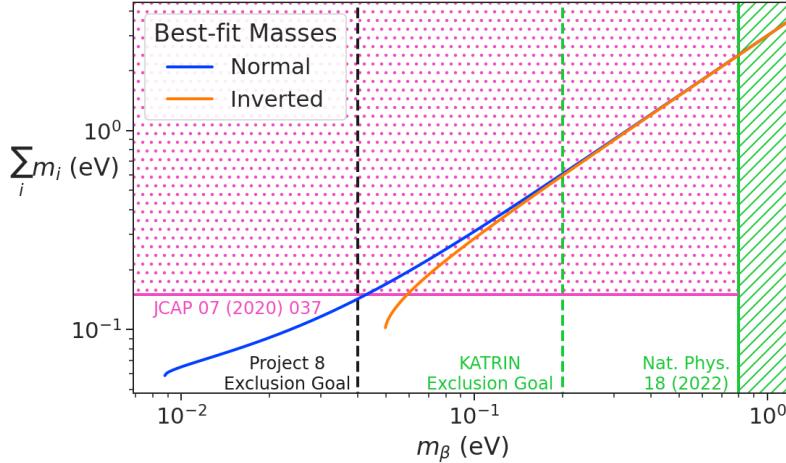


Figure 3.5: Neutrino mass exclusion plot including limits from cosmological measurements and the KATRIN experiment. Allowed ranges for neutrino masses under the normal and inverted hierarchies are shown as the blue and orange lines respectively. The black dashed line shows Project 8's goal neutrino mass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

803
804 is sufficient to fully exhaust the range of allowable neutrino masses under the inverted
805 neutrino mass ordering regime and is approximately an order of magnitude less than the
806 projected final sensitivity of the KATRIN experiment. Excluding the full neutrino mass
807 parameter space would require a sensitivity an order of magnitude lower than what is
808 proposed by Project 8, which would require an experiment whose size and complexity
809 are currently well beyond proposals for the next-generation of neutrino mass direct
810 measurement experiments.

811 **3.2.3 Project 8 Phased Development Plan**

812 Reaching 40 meV sensitivity will require the simultaneous development and eventually
813 combination of two novel technologies. The first is the technology required to supply a
814 source of atomic tritium of the appropriate size, density, purity, and temperature along
815 so that the atoms can be trapped and their beta-decays measured in the spectrometer.
816 The second is a CRES measurement technology that is both compatible with the tritium
817 atom trap and is capable of reconstructing CRES events with sufficient energy resolution

818 to achieve the required sensitivity.

819 These technologies require a significant up-front research and development (R&D)
820 investment to build-out the required capabilities for a 40 meV CRES experiment. There-
821 fore, Project 8 is following a phased experiment plan in which incremental progress can
822 be made towards the ultimate goal of a 40 meV neutrino mass measurement with CRES.

823 **Phase I and II: Proof of Principle and First Tritium Measurements**

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

827 In Phase I, Project 8 performed a proof-of-principle measurement of the ^{83m}Kr
828 spectrum using CRES, which marked the first ever energy spectrum measurement with
829 CRES. The experiment included all of the main components expected for the full-scale
830 version of the experiment. An electron source consisting of a gas of ^{83m}Kr was supplied
831 to a waveguide gas cell constructed out of a segment of WR-42 waveguide and sealed
832 with Kapton windows at the top and bottom. A magnetic trapping region was created
833 in the waveguide cell using a single electromagnetic coil wrapped around the waveguide
834 which provided a trapping volume on the order of a few cubic-millimeters. Detection of
835 the cyclotron radiation was performed by connecting the waveguide cell to an additional
836 segment of waveguide that transmitted the radiation to a cryogenic amplifier.

837 Success in Phase I was achieved with the 2014 publication of the measured ^{83m}Kr
838 conversion spectrum, which contains a mono-energetic 17.8-keV as well as several other
839 conversion lines at higher energies. Publication of this result marked the official end of
840 Phase I and the start of Phase II in which Project 8 shifted its focus to the demonstration
841 of the first tritium beta-decay spectrum using CRES. Phase II successfully concluded
842 in 2023 with the submission of the papers demonstrating the first tritium beta-decay
843 spectrum endpoint and neutrino mass measurement using CRES. For more information
844 on Phase II please see Section 3.3.

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

846 With the completion of Phase II Project 8 has shifted into a phase focused on the
847 construction of an experiment that demonstrates all the technologies of the final experi-
848 ment in Phase IV. The goal for this pilot-scale experiment is to successfully retire all
849 technological and engineering risks associated with the Phase IV experiment, while being

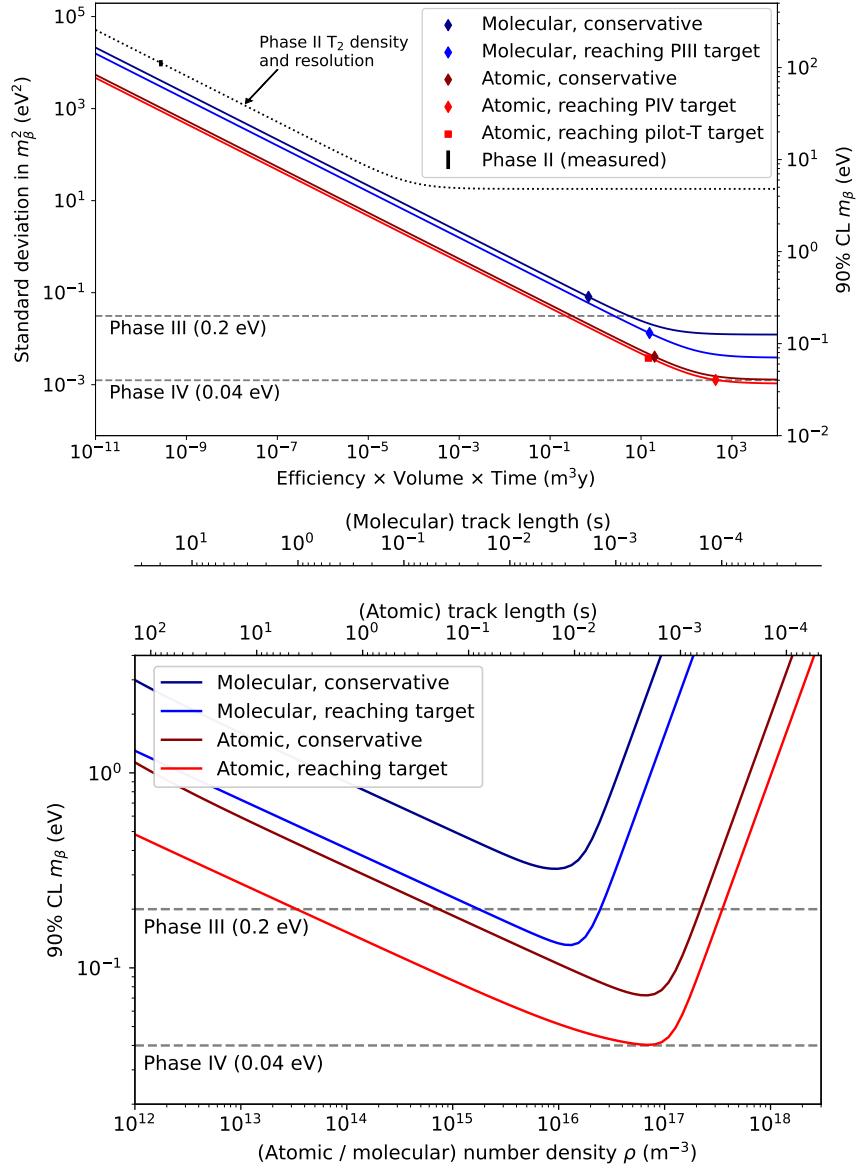


Figure 3.6: Sensitivity calculations for a cavity based CRES experiment that demonstrate the neutrino mass measurement goals of the Project 8 collaboration throughout the phased development plan. The blue tinged curves indicate molecular tritium sources and the red tinged curves indicate atomic tritium 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.

850 a scientifically interesting experiment in it's own right that has sensitivity to neutrino
851 masses on par with KATRIN's final projected sensitivity.

852 Phase III R&D is divided into two equally important efforts — atomic tritium and
853 CRES detection techniques. Atomic tritium development in Phase III includes the
854 development of all aspects of the tritium system required for the pilot-scale experiment.
855 This includes the production of tritium atoms, atomic cooling and recirculation systems,
856 purity and isotope concentration monitoring, and trapping. Currently, Project 8 is
857 operating small scale demonstrator systems developing atom crackers to show that atom
858 production at the estimated rates needed for Phase IV is achievable. Future efforts
859 will continue the current developments on atom production and expand to include
860 demonstrations of atomic cooling with an evaporative beam line as well as atom trapping
861 using Halbach magnet arrays.

862 The need for new CRES detection techniques is driven by the drastic increase in scale
863 from Phase II to the Phase IV and the pilot-scale experiments. The physical volume
864 used for CRES in Phase II was on the order of a few cubic-centimeters, and achieving
865 Project 8's sensitivity target of 40 meV requires an experiment volume on the multi-cubic
866 meter scale. Therefore, the waveguide gas cell CRES detection technique used in Phase
867 II is not a feasible option for the future of Project 8 due to it's inability to scale to the
868 required size.

869 Two alternative CRES detection techniques have been proposed for the pilot-scale
870 experiment — antenna arrays and resonant cavities (see Section 3.4 and Chapter 6
871 respectively). Both approaches have relative advantages and disadvantages, however,
872 the improved understanding of the antenna array and cavity approaches to CRES in
873 the recent years has led to cavities being the preferred technology for the pilot-scale
874 experiment due to the estimated reduced cost and complexity of this approach. Since a
875 large degree of the work presented in this thesis is focused on the development of the
876 antenna array CRES technique as well as the design of demonstrator experiments, we
877 described the proposed R&D plan for antenna array CRES in Phase III in Section 3.4.

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

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

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

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

895 **3.3 Phase II: First Tritium Beta Decay Spectrum and**
896 **Neutrino Mass Measurement with CRES**

897 In Phase II Project 8 demonstrate the first ever measurement of the tritium beta-decay
898 spectrum endpoint using the CRES technique, which lead to the first neutrino mass
899 measurement by the Project 8 collaboration. This milestone was made possible by
900 many improvements in the CRES technique and more developed understanding of CRES
901 systematics, which takes an important first step towards larger scale measurements of
902 the tritium beta-decay spectrum with CRES. In this section, I shall briefly describe some
903 the important elements of the Phase II experiment, with the goal of contextualizing the
904 research and development efforts for Phases III and IV of Project 8. For more complete
905 descriptions of the work that lead to Project 8's Phase II results please refer to the many
906 Phase II papers produced by the collaboration.

907 **3.3.1 The Phase II CRES Apparatus**

908 **Magnet and Cryogenics**

909 The magnetic field for the the Phase II experiment is provided by a nuclear magnetic
910 resonance (NMR) spectroscopy magnet with a central bore diameter of 52 mm (see Figure
911 3.7). The magnet produces a background magnetic field with an average value of 0.959 T
912 and a 10 ppm variation across the bore diameter achieved using several shim coils built
913 into the magnet. Using an external NMR field probe the variation of the magnetic field

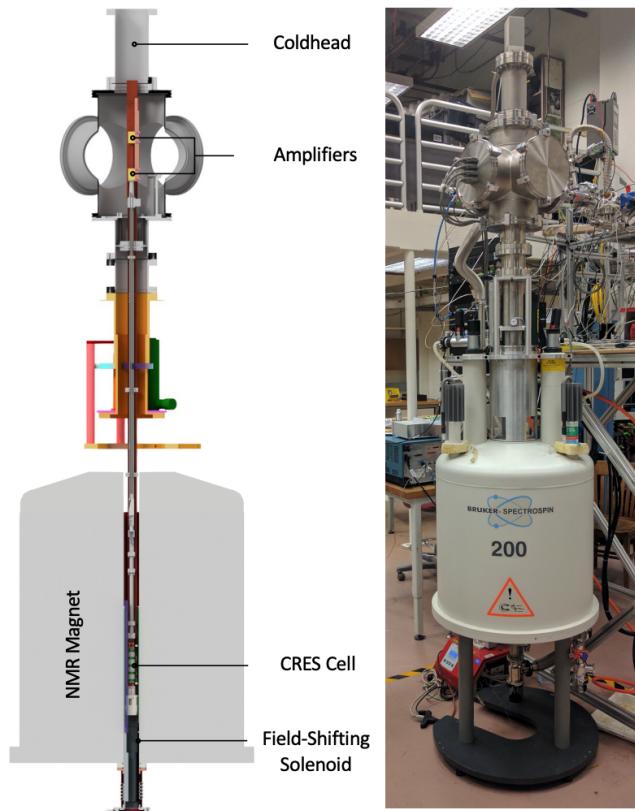


Figure 3.7: The Phase II CRES apparatus used to perform the first measurement of the tritium beta-decay spectrum using CRES.

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

917 An external solenoid magnet was installed inside the magnet bore to provide the
 918 ability to shift the magnitude of the background magnetic field by values on the order of
 919 a few mT. The solenoid has inside diameter of 46 mm and a length of 350 mm, which
 920 terminates in a vacuum flange that allows it to be inserted into the NMR magnet bore
 921 from the bottom. By shifting the value of the magnetic field by a few mT, the cyclotron
 922 frequencies of electrons produced by the 17.8 keV ^{83m}Kr internal-conversion line can be
 923 shifted over a range of frequencies on the order of 100 MHz. This allows one to study the
 924 frequency dependent behavior of multiple CRES systematics such as detection efficiency
 925 that directly affect the measured shape of the tritium spectrum.

926 The inside of the magnet bore diameter was pumped down to a vacuum of less than
 927 10 μtorr using a turbomolecular pump, which allows for cryogenic cooling of the CRES

cell and RF system. Cooling power was supplied to the Phase II apparatus using a cryopump with its coldhead mounted above the primary magnet and CRES cell. This arrangement allowed for sufficient cooling power to be delivered to the amplifiers to cool them to a temperature of ≈ 40 K, while keeping the amplifiers far enough from the magnet so as not to be damaged by the large field strength. Thermal contact between the coldhead, amplifiers, RF system, and CRES cell is achieved using a copper bar that runs the full length of the apparatus. To prevent freeze-out of ^{83m}Kr on the walls of the CRES cell a separate heater was installed to keep the CRES cell near a temperature of 85 K during the operation of the experiment.

937 CRES Cell

938 Located in the most uniform region of the magnetic field is the CRES cell, which is the
 939 region of the apparatus where radioactive decays of ^{83m}Kr and T_2 emit electrons that can
 be trapped and measured using CRES (see Figure 3.8). The CRES cell is manufactured

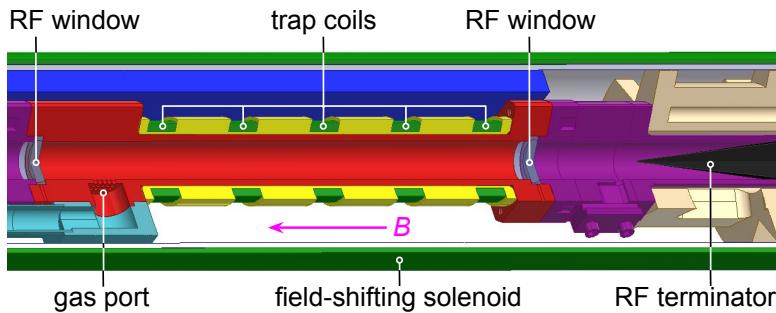


Figure 3.8: Diagram of the CRES cell portion of the Phase II apparatus.

940
 941 from a segment of cylindrical waveguide designed to operate at K-band frequencies
 942 near 26 GHz. The diameter of the waveguide determines which resonant modes of the
 943 waveguide will couple to the electron and transmit its radiation to the amplifiers. For
 944 Phase II a waveguide diameter of 1 cm was selected, which allows electrons to couple to
 945 the TE₁₁ and TM₀₁ cylindrical waveguide modes. To reduce complexity in modeling and
 946 analyzing the CRES data, it is ideal to select a diameter that prevents electrons from
 947 coupling to higher-order waveguide modes beyond the fundamental TE and TM modes.

948 Around the exterior of the cylindrical waveguide are several magnetic coils used
 949 to produce magnetic traps inside the CRES cell volume. Without a magnetic trap
 950 electrons produced from decays inside the CRES cell quickly impact the cell wall, which
 951 prevents a measurement of their cyclotron frequency using CRES. Each coil along the
 952 length of the waveguide produces a separate trap that is approximately harmonic in

shape. By independently controlling the currents provided to each coil the traps could be configured to have equal values of the magnetic field at the trap bottom despite a variable background magnetic field from the NMR magnet.

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

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

RF System

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

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

Radiation traveling upward passes through the CaF_2 window passes through a $\lambda/4$

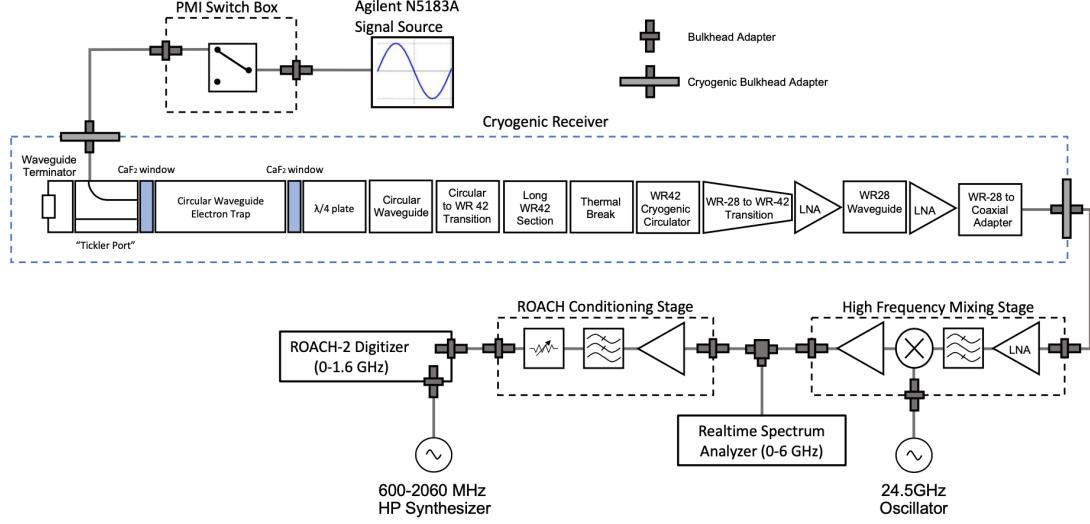


Figure 3.9: RF system diagram for the Phase II apparatus.

plate, which transforms the circularly polarized cyclotron radiation into linear polarization. The linearly polarized fields next travel through a segment of circular waveguide that transitions into a long segment of WR-42 waveguide that carries the fields out of the high magnetic field region. A thermal break segment is included, which consists of a a segment of gold-plated stainless steel WR-42 waveguide, to help thermally isolate the relatively warm CRES cell from the colder amplifiers. The radiation then passes through a cryogenic circular, which prevents signals reflected from the amplifiers from interfering with the CRES cell before a WR-42 to WR-28 transition connects the waveguide to the first of the cryogenic amplifiers. The radiation passes through two cryogenic amplifiers before being coupled to a coaxial termination at the top of the Phase II apparatus.

The coaxial cable transfers the cyclotron radiation signals to a high-frequency mixing stage that performs an analog frequency down-conversion using a 24.5 GHz LO. Two forms of digitization can be used at this stage to readout the CRES data. One is a real-time spectrum analyzer that digitizes the CRES signal data in time-domain and computes the frequency spectrum in real-time, which allows for direct visualization of CRES signal spectrograms as the experiment is running. The real-time spectrum analyzer is most useful for taking small amount of streamed data for debugging and analysis of the system. The other method, which was used to collect the majority of the CRES data in Phase II, is a ROACH-2 FPGA and digitizer system. The ROACH system consists of a fast ADC that samples the CRES signal data at 3.2 GSps. Internal digital down-conversion stages implemented in the FPGA perform a mixing operation that reduces the bandwidth of the

1008 CRES signals to 100 MHz. The FPGA implements a 8192 sample FFT and packetizes
1009 time and frequency domain records in parallel. The packetized data is then transferred
1010 from the ROACH to be analyzed by the data-processing pipeline.

1011 **3.3.2 CRES Track and Event Reconstruction**

1012 **Time-Frequency Spectrogram**

1013 The online data-processing is intended to identify interesting data that could contain
1014 CRES signals using a software real-time triggering algorithm. Interesting segments of
1015 data identified by this algorithm are collected into files that are transferred to a server for
1016 offline processing and analysis. The data files contain a continuous series of time-domain
1017 samples, broken into a set of records, which are 4096 samples long. The time-series is
1018 made up of 8-bit IQ samples acquired at 100 MHz.

1019 Each time-series record is accompanied by an associated frequency spectrum consisting
1020 of 4096 frequency bins approximately 24.4 kHz wide, which is represented as a power
1021 spectral density. The individual frequency spectra can be organized temporally to create
1022 a time-frequency spectrogram that represents the evolution of the cyclotron frequency
spectrum over the course of the CRES event (see Figure 3.10). The time-frequency

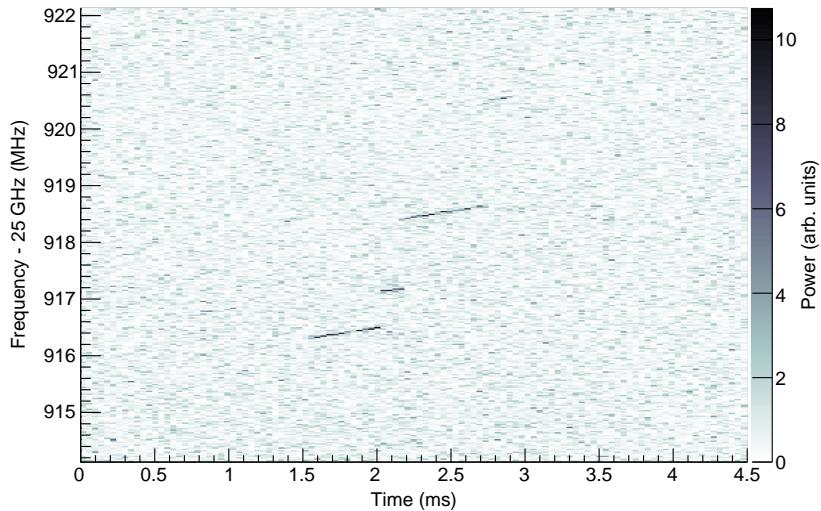


Figure 3.10: The time-frequency spectrogram of a tritium CRES event in the Phase II apparatus.

1023
1024 spectrogram is represented as a two-dimensional image where the color of each pixel is
1025 proportional to the power spectral density. Each vertical slice of pixels in the image

represents a frequency spectrum, therefore, each horizontal bin represents the data obtained over a duration of $4096 \times 0.01 \text{ MHz}^{-1} = 40.96 \mu\text{sec}$.

CRES Event Data Features

Phenomenologically, a CRES signal appears as a sinusoidal signal whose frequency slow increases ("chirps") over time. Axial motion of the electron in the trap leads to the formation of frequency sidebands that surround the more powerful carrier frequency, due to doppler modulation of the electron's frequency as it bounces between the walls of the magnetic trap. The critical piece of information that must be extracted from the track and event reconstruction procedure is the carrier frequency, since it is this frequency that gives the cyclotron frequency and thus the kinetic energy. While axial motion from non- 90° pitch angles does change the average magnetic field experienced by an electron and, therefore, changes the cyclotron frequency. We were not able to resolve sidebands in Phase II, so a correction for the effect of the pitch angle on the cyclotron frequency was not possible.

In the time-frequency spectrogram representation the chirping carrier frequency appears as a linear track of high-power frequency bins (see Figure 3.10). The vertical slope of the tracks is caused by the emission of energy from the electron in the form of cyclotron radiation, therefore, the size of the slope parameter is directly proportional to the Larmor power. The continuous track is periodically interrupted by random jumps to higher frequency and lower energy caused by random inelastic collisions with background gas molecules. The length of a track is an exponentially distributed variable whose mean value is inversely proportional to the gas density. The size of the frequency discontinuities is directly proportional to the energies of the rotational and vibrational states of background gas species such as CO_2 .

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

1057 **Track Reconstruction**

1058 The first step in this process is the identification of tracks in the time-frequency spectrogram, which is essentially an image processing feature identification task. The first step
1059 in the track finding procedure is to normalize the power spectral density based on the
1060 average noise power to obtain the time-frequency spectrogram in the form of normalized,
1061 unitless power. Next a power threshold is applied to the normalized spectrogram
1062 where only bins that have a signal-to-noise ratio greater than five are selected to build
1063 tracks. In this case signal-to-noise ratio is defined as the ratio between the normalized,
1064 unitless power of a bin divided by the average normalized power across the full frequency
1065 spectrum.

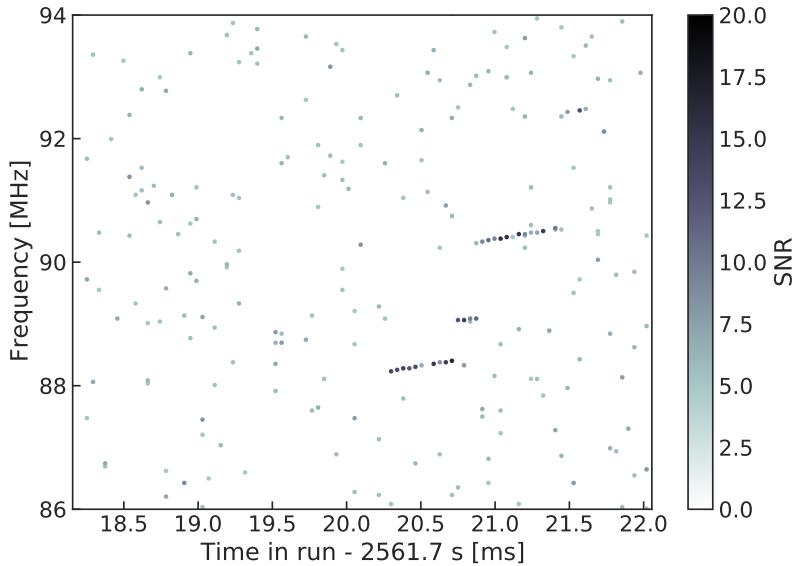


Figure 3.11: The sparse spectrogram obtained by placing a power cut on the raw spectrogram shown in Figure 3.10.

1067 The spectrogram produced by this power cut, termed the sparse spectrogram, consists
1068 only of a sparse collection of high-power frequency bins that could be part of a CRES
1069 signal track (see Figure 3.11). In this form is it much easier to identify tracks "by eye",
1070 however, for the Phase II analysis Project 8 developed it's own custom-made track finding
1071 algorithm, called the sequential track finder (STF).

1072 The STF algorithm processes the sparse spectrogram in sequential fashion, processing
1073 each time-slice one-by-one until the end of the spectrogram is reached. Tracks are found
1074 by searching for points in the sparse spectrogram that appear to fall on a straight line.
1075 Multiple configurable parameters are built into the STF algorithm that allow the user to

1076 tune the criteria for adding a point to an existing track or creating a brand new track.
1077 These include parameters such as maximum time and frequency differences between
1078 subsequent points in a track as well as minimum SNR values for the start and endpoints
1079 of the track. Additionally, tracks are required to have a minimum length and slope to be
1080 considered potential CRES tracks rather than random noise fluctuations.

1081 The resulting output of the STF is a collection of track objects that consist of all of the
1082 points that make up the track and their properties. The final step in track reconstruction
1083 is to calculate the track properties and apply final cuts to reject the majority of false
1084 tracks found by the STF. This involves the fitting of a line to the collection of track
1085 points as well as the total and average power of the track obtained by computing the
1086 sum and mean of the points powers. The starting frequency of the track is determined by
1087 calculating the time coordinate that intersects with the linear fit. A cut is performed
1088 to remove all tracks that do not have a specified average power over their duration, which
1089 helps to remove the majority of noise fluctuations that have passed all previous cuts up
1090 to this point.

1091 Event Reconstruction

1092 The final step is event reconstruction where the identified tracks are grouped into events
1093 that contain all tracks likely caused by the same electron. This procedure simply attempts
1094 to match tracks head to tail by checking if the start and end times of a pair of tracks
1095 falls withing a certain tolerance. This tolerance is an additional configurable parameter
1096 that can be tuned to an optimal value using monte carlo simulations of events in the
1097 Phase II apparatus.

1098 After the event building procedure has completed there is still a small likelihood that
1099 false tracks have made it through to this stage in the reconstruction. Typically, cuts at
1100 the track level are able to remove 95% of the false tracks identified by the STF, which
1101 leads to a significant number of false tracks at the event building stage. However, the
1102 additional event-level information makes it possible to reject events that contain these
1103 false tracks with a high degree of confidence.

1104 Two event level features are associated with events caused by real electrons — the
1105 duration of the first track as well as the number of tracks in the event. Real electrons
1106 tend to have event structures with longer first tracks and a higher number of total tracks.
1107 Based on the values of these two criteria, a minimum threshold on the average power in
1108 the first track was configured to reject false events. The average power in the first track
1109 was chosen due to the critical nature of the starting frequency of the first track in an

1110 event to the krypton and tritium spectrum analyses.

1111 3.3.3 Results from Phase II

1112 The primary result from Phase II is the first-ever measurement of the tritium beta-decay
1113 spectrum using CRES, which lead to the first neutrino mass limit using the CRES
1114 technique. However, Phase II also included a significant ^{83m}Kr measurement campaign
1115 to understand important systematics relevant to the tritium spectrum measurement, but
1116 also to understanding the fundamentals of the CRES technique itself. This required
1117 high-resolution measurements of the ^{83m}Kr internal-conversion spectrum, which is an
1118 interesting science result in its own right.

1119 The results from Phase II represents a significant effort from the entire Project 8
1120 collaboration over several years. Because the focus of my contributions to Project 8 is
1121 directed towards the research and development efforts for the Phase III experiments, the
1122 goal in this section is not to provide a detailed description of the the analyses that lead to
1123 the Phase II results. Rather, I will provide brief descriptions of a few plots representative
1124 of the main results from Phase II and direct the interested reader to the relevant Phase
1125 II papers.

1126 Measurements with Krypton

1127 Measurements with krypton were a key calibration tool for Phase II of the experiment
1128 and will most likely continue to be useful in future Phases of Project 8. In the context of
1129 Project 8 krypton measurements refers to CRES measurements of the internal-conversion
1130 spectrum of the metastable state of krypton-83, ^{83m}Kr , produced by electron capture
1131 decays of ^{83}Rb . A supply of ^{83}Rb was built into the Phase II apparatus gas system that
1132 supplied the CRES cell with ^{83m}Kr via emanation.

1133 The ^{83m}Kr internal-conversion spectrum consists of several lines based on the orbital
1134 of the electron ejected during the decay. The conversion lines useful to Project 8 are
1135 those that emit electrons with kinetic energies that fall inside the detectable frequency
1136 bandwidth of the Phase II apparatus. These are the K; L2 and L3; M2 and M3; and N2
1137 and N3 lines with kinetic energies of 17.8 keV, \approx 30.4 keV, \approx 31.9 keV, and \approx 32.1 keV,
1138 respectively. The different energies of the lines allow us to test the linearity of the
1139 relationship between kinetic energy and frequency across the range of frequencies covered
1140 by the continuous tritium spectrum.

1141 By measuring the shape of the krypton spectrum we can characterize the effects of

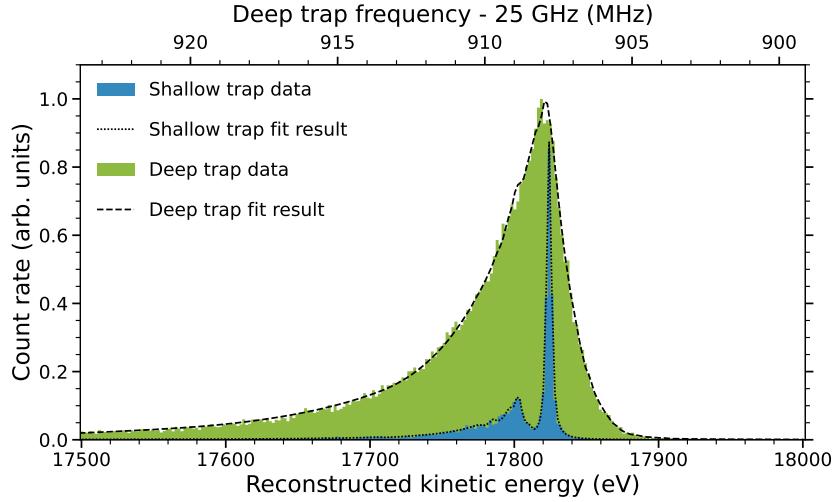


Figure 3.12: Fits to the measured 17.8-keV ^{83m}Kr conversion line using the deep and shallow trap configurations.

numerous detector related effects relevant to the tritium analysis. Specific examples include the variation in the magnetic field as a function of the radial position of the electron, variation in the magnetic field caused by the trap shape, variation in the average magnetic field for electron of different pitch angles, the effect of missing tracks due to scattering, among others. These spectrum shape measurements focused on the 17.8-keV krypton line and utilized different trap geometries based on the particular goal of the dataset (see Figure 3.12).

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

The broadening of the krypton spectrum seen for the deeper track is due to the higher range of electron pitch angles that can be trapped. Furthermore, with a deeper trap there is a larger parameter space of electron that could be produced with pitch angles that are trappable but not visible in the time-frequency spectrogram. These electrons live in the trap and can scatter multiple times before randomly scattering to a pitch angle that is now visible. This causes us to miss one to several of the electron's tracks earlier in the event, which leads us to mis-reconstruct the true starting frequency of an event. By measuring the krypton spectrum shape in the same deep trap used to detect tritium

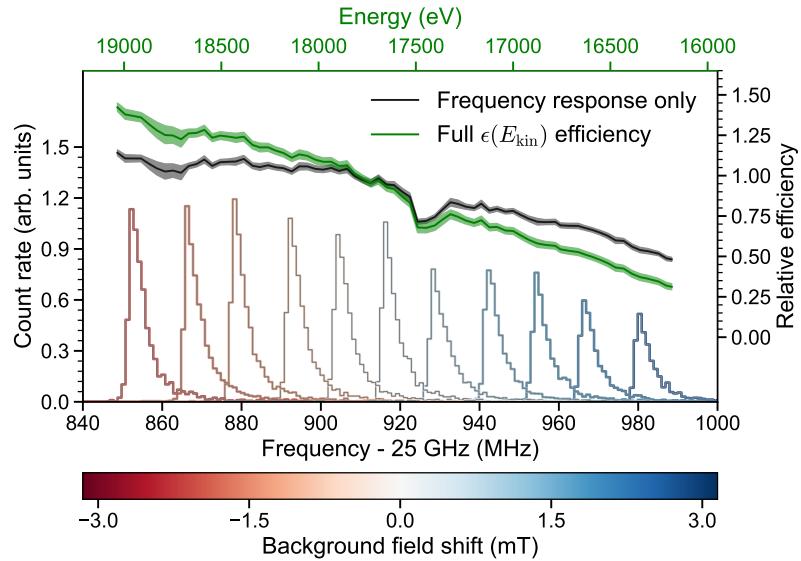


Figure 3.13: Measurements of the 17.8-keV ^{83m}Kr line using the deep trap configuration for different values of the magnetic field from the field shifting solenoid.

events we can characterize the affect that this has on the spectrum shape to mitigate it's impact on the tritium measurements.

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

1174 Tritium Spectrum and Neutrino Mass Results

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

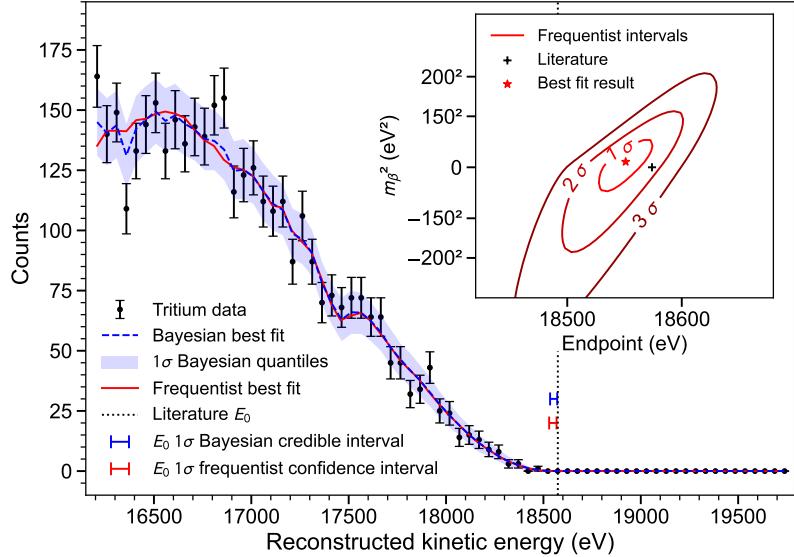


Figure 3.14: The measured tritium spectrum from Phase II with Bayesian and frequentist fits.

1181 Figure 3.14).

1182 CRES is inherently a very low background technique with the dominant source of noise
 1183 being random RF fluctuations. Monte carlo simulations backed up by measurements
 1184 of the RF noise background were used to set track and event characteristic cuts to
 1185 guarantee that zero false events would occur over the duration of the experiment with
 1186 90% confidence. Notably, the measured spectrum has zero events beyond the tritium
 1187 spectrum endpoint, which allows us to constrain the background rate in the Phase II
 1188 apparatus to less than 3×10^{-10} counts/ev/s. Achieving a low background is critical for
 1189 future neutrino mass experiments that seek to measure the neutrino mass with less than
 1190 100 meV sensitivity.

1191 Bayesian and frequentist based fits to the measured tritium spectrum, incorporating
 1192 information gained about CRES systematics from the krypton measurements, were
 1193 performed to extract upper limits on the tritium beta-decay spectrum endpoint as well as
 1194 the neutrino mass. The estimated spectrum endpoints are 18553^{+18}_{-19} eV for the Bayesian
 1195 analysis and 18548^{+19}_{-19} eV for the frequentist analysis. The quoted uncertainties are
 1196 1- σ , and both results are within 2- σ of the literature endpoint value of 15574 eV. The
 1197 estimated neutrino mass for both results is consistent with $m_\beta^2 = 0$. The 90% confidence
 1198 upper limits for the Bayesian analysis is $m_\beta < 155$ eV/c² and $m_\beta < 152$ eV/c for the
 1199 frequentist analysis.

Though the neutrino mass results from Phase II are not competitive with KATRIN it is a promising first step towards the development of more precise neutrino mass measurements using CRES. The low background and demonstrated high resolution with krypton measurements are promising features of the technique that were able to be demonstrated with the Phase II apparatus. As new technologies are developed to enable CRES measurements in larger volume, many of the lessons learned from Phase II will continue to influence the operation and design of the detectors.

3.4 Phase III R&D: Antenna Array CRES

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

3.4.1 The Basic Approach

One possible approach, suggested in the original CRES publication, is to use many antennas to surround a volume of tritium gas in a magnetic field (see Figure 3.15). When a decay occurs the electron will begin to emit cyclotron radiation that can be collected by the array and used to perform CRES. Each antenna in the array collects only a small

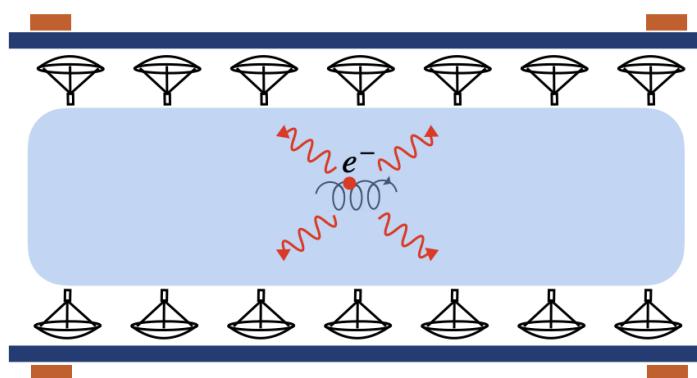


Figure 3.15: A cartoon illustration of the basics of the antenna array CRES technique.

fraction of the electron's signal power, which is less than 1 fW for a 18.6 keV kinetic energy electron in a 1 T magnetic field. Scaling to large volumes with the antenna

array approach is accomplished by increasing the number of antennas in the array, which increases the volume under observation proportionally, so that a sufficient population of tritium atoms can be observed to measure the tritium spectrum endpoint shape.

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

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

Another beneficial quality of the antenna array approach is that the volume of the experiment can be scaled independent of frequency by simply adding more antennas to the array (see Figure 3.19). Resonant cavities, the proposed alternative large volume CRES technology, are ideally operated in magnetic fields that cause electrons to move with cyclotron frequencies near the fundamental cavity resonance, to avoid complex coupling of the electron to many cavity modes simultaneously. This leads to a coupling between the cavity volume and the magnetic field magnitude, which forces one to lower the magnetic field in order to increase the experiment scale. Whereas, for antenna arrays, in principle there is no physical limitation on the size of the antenna array that can be used at a particular magnetic field. However, the nature of scaling an antenna array based experiment leads to rapidly increasing cost and complexity due to the large number of antennas, amplifiers, and data streams that require substantial computer processing power to effectively analyze.

3.4.2 The FSCD: Free-space CRES Demonstrator

The complex collection of new experimental techniques and methods that come together in the antenna array CRES technique require the construction of a small scale demonstration experiment designed to develop an understanding of the principles of antenna array CRES measurements and the relevant systematics. Without operating such an experiment it is not possible to develop a design for a large scale CRES experiment with sufficient confidence that the experiment is capable of measuring the shape of the tritium spectrum endpoint to the degree of accuracy required for 40 meV sensitivity to the neutrino mass. Therefore, Phase III of the Project 8 experimental program is primarily focused on the development and operation of demonstrator experiments to inform the design of the final Phase IV experiment.

Specifically for antenna array CRES, the associated demonstrator experiment in Phase III is called the Free-space CRES Demonstrator or FSCD. The goals of the FSCD include not only the development of antenna array CRES itself, but is also a capable neutrino mass measurement experiment in it's own right, with a target neutrino mass sensitivity of a few eV using a molecular tritium source.

Magnetic Field

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

Inside the main magnetic field of the MRI magnet are additional magnets that provide the capability to shift the value of the background magnetic field as well as the magnets that produce the magnetic trap. Shifting the background value of the magnetic field on a scale of $O(\mu T)$ allows one to control the cyclotron frequencies of electrons with a fixed kinetic energy, which is key to effectively calibrating the FSCD. The preferred calibration method for the FSCD is a mono-energetic electron gun that can inject electrons into



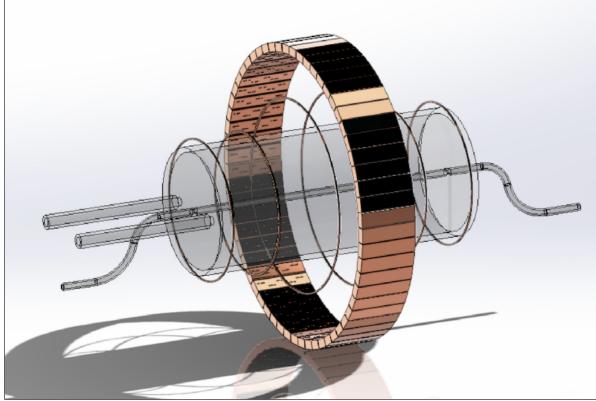
Figure 3.16: An image of the MRI magnet installed in the Project 8 laboratory at the University of Washington, Seattle.

1288 the magnetic trap with a known kinetic energy. In combination with the field shifting
 1289 magnet one can vary the cyclotron frequencies of the electrons to measure the response
 1290 of the antenna array as a function of the radiation frequency and electron position. This
 1291 procedure not only characterizes the response of the antenna array but also provides
 1292 further information on magnetic field uniformity, which important to achieving optimal
 1293 energy resolution.

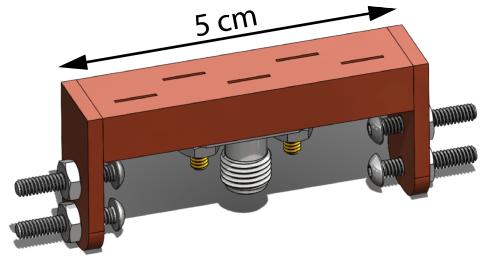
1294 Several additional magnetic coils will need to included inside the MRI magnet to
 1295 produce the magnetic trap. The ideal trap shape for CRES is the perfect magnetic box,
 1296 which has a flat bottom and step function walls. Any variation in the average magnetic
 1297 field experienced by an electron leads to changes in the cyclotron frequency that can
 1298 make determining the true starting kinetic energy more difficult. This includes changes
 1299 in the magnetic field caused by the walls of the magnetic trap as well as radial magnetic
 1300 field variations. The perfect box trap is completely uniform and has infinitely steep walls
 1301 that cause no change in the electron's cyclotron frequency as it is reflected from the
 1302 trap wall, however, such a trap cannot be made from any combination of magnetic coils
 1303 since it violates Maxwell's equations. The goal of magnetic trap design is to identify the
 1304 configuration of coils that produces a trap that approximates the perfect box trap as
 1305 closely as possible.

1306 **Antenna Array**

1307 The canonical antenna array design for a CRES experiment is a uniform cylindrical array
1308 of antennas that surrounds the magnetic trap volume. Since the FSCD is a demonstrator
1309 experiment, the antenna array design is the simplest form of the uniform cylindrical
1310 array, which is a single circular ring of antennas with a diameter of 20 cm (see Figure
3.17). Along this circle are sixty slotted waveguide antennas that fully populate the



(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.

1311
1312 available space around the array circumference. In order to maximize the power collected
1313 from each electron it is optimal to cover as large a fraction of the solid angle around the
1314 magnetic trap as possible.

1315 The distance between antennas around the circumference of the array is proportional
1316 to the wavelength of the cyclotron radiation. Therefore, maximizing the solid angle
1317 coverage of the array, while minimizing channel count to keep the hardware and data
1318 acquisition costs manageable, biases one towards smaller array diameters. Antenna
1319 near-field effects limit the minimum diameter of the array for a given antenna design
1320 since the radiation from electrons that are too close to the array cannot be detected
1321 due to destructive interference caused by path-length differences from the electron to
1322 different points on the antenna surface.

1323 Slotted waveguide antennas are used in the FSCD antenna array due to their high
1324 efficiency and low loss, which comes from the lack of dielectric materials in the antenna
1325 structure. Coupling to the waveguide can be performed with a coaxial cable connected
1326 at the center or on either end of the waveguide. One of the drawbacks of waveguide

1327 antennas is the large amount of space required to fit them inside the limited MRI magnet
1328 volume. Alternative antenna designs, constructed from microstrip printed circuit boards
1329 require significantly less space at the cost of slightly higher energy loss in the antenna
1330 structure.

1331 The FSCD antenna design is a 5 cm long segment of WR-34 waveguide with 5 vertical
1332 slots cut into the side. The distance between slots along the length of the waveguide is
1333 a half wavelength for optimal power combination between the individual antenna slots.
1334 Each slot is offset from the center of the antenna face a small distance in order to most
1335 effectively couple the slot to waveguide modes inside the antenna.

1336 The passive power combination achieved by placing 5 slots in a single waveguide is a
1337 compromise intended to reduce the cost and complexity of the antenna array system.
1338 Each additional channel in the array requires it's own cryogenic amplifier and also increase
1339 the required computer power to process the raw data collected by digitizing each channel.
1340 Passive summation, achieved by combining antennas into arrays axially, reduces the array
1341 channel count at the cost of losses from imperfect passive combination. Imperfect passive
1342 combination is caused by effects such as re-radiation of energy from and destructive
1343 interference between slots in the waveguide antenna.

1344 Interference and re-radiation eventually limit the achievable the axial extent of passive
1345 power combination. The 5-slot designed developed for the FSCD is optimized to minimize
1346 the impact of these losses while achieving the maximum amount of axial coverage with a
1347 single ring of antennas. Scaling beyond the volume covered by a single ring of antennas is
1348 achieved by stacking additional rings of antennas together to cover a larger trap volume
1349 for a higher statistics measurement of the tritium spectrum endpoint region. A likely
1350 scenario for the FSCD experiment involves a staged experiment approach, where first
1351 a series of measurements is performed using only a single ring of antennas followed by
1352 experiments that add additional rings to the FSCD. The goal would be to first understand
1353 the principles of antenna array CRES using the simplest possible experiment, before
1354 attempting to scale the technique by expanding the antenna array size.

1355 **Tritium Source**

1356 While the primary purpose of the FSCD is as a technology demonstrator, it is unlikely
1357 for the collaboration to gain the required confidence in the antenna array CRES tech-
1358 nique to perform neutrino mass measurements at the 40 meV sensitivity level without
1359 an intermediate scale measurement of the neutrino mass using antenna array CRES.
1360 Therefore, the FSCD has an additional scientific goal of measuring the neutrino mass

1361 with a rough sensitivity goal of a few eV. This level of precision is achievable using a
1362 source of molecular tritium with a volume of approximately 1 L at a density comparable
1363 to potential Phase IV scenarios.

1364 Unlike previous CRES experiments, where the tritium source could be co-located
1365 with the receiving antenna inside a waveguide transmission line, the tritium source
1366 in the FSCD is thermally isolated from the antenna array to avoid freeze-out of the
1367 tritium molecules. The tiny radiation power emitted by electrons requires a system noise
1368 temperature of ≈ 10 K or less, in order to detect events at a high enough efficiency to
1369 reach the neutrino mass sensitivity goals of the experiment. Achieving a system noise of
1370 10 K requires that the antenna array and amplifiers operate at cryogenic, liquid helium
1371 temperatures of ≈ 4 K, which significantly lowers the vapor pressure of molecular tritium.
1372 By keeping the molecular tritium isolated in an RF-transparent vessel the tritium gas can
1373 be kept at a relatively warmer temperature in the range of 30 K to avoid the accumulation
1374 of tritium on the experiment surfaces.

1375 Data Acquisition and Reconstruction

1376 A fundamental change in the data acquisition system for the FSCD is the shift from
1377 single to multi-channel reconstruction. This transition results in a significant increase in
1378 the data-generation rate, which is linearly related to the number of independent channels
1379 in the array. The larger data volume coincides with an increased demand for computer
1380 processing power based on the need for more precise signal reconstruction algorithms
1381 driven by the FSCD and Phase IV sensitivity goals. Therefore, the data acquisition
1382 system for the FSCD is likely to represent a significantly larger fraction of the experiment
1383 cost and complexity than previous CRES experiments.

1384 Each antenna in the array is connected to a cryogenic amplifier and down-converted
1385 from the 26 GHz CRES frequency using an IQ-mixer to reduce the size of the analysis
1386 window in which the tritium spectrum is measured. Using an LO with a frequency of
1387 approximately 25.80 GHz the antenna array signals can be digitized at a rate of 200 MHz,
1388 which is sufficient bandwidth to resolve the complete sideband spectrum produced by
1389 axial oscillations of electrons in the FSCD magnetic trap.

1390 Direct storage of the raw FSCD antenna array data is undesirable, since the estimated
1391 amount of raw data generated is $O(1)$ exabyte per year. The management and storage
1392 of such a large dataset is infeasible for a demonstrator experiment on the scale of the
1393 FSCD and would represent a large fraction of the budget for a Phase IV scale antenna
1394 array based CRES experiment. Therefore, a sub-goal of the FSCD experiment is the

1395 development of real-time reconstruction methods that could reduce the raw data volume
1396 by detecting and reconstructing CRES events in real-time. The ultimate goal would be
1397 a complete real-time reconstruction pipeline that takes raw voltages samples from the
1398 antenna array and returns estimates for the starting kinetic energies of CRES events in
1399 the data.

1400 The feasibility of a real-time reconstruction pipeline rests on the development of
1401 computationally efficient algorithms that can be implemented without the need for
1402 enormous computing resources. One challenge with the antenna array approach is that
1403 the small radiation power of a single electron is distributed between each channel in
1404 the array, such that reconstruction using only the information in a single channel is not
1405 possible. Therefore, the simply performing the initial step in reconstruction — signal
1406 detection — requires orders of magnitude more computational power than previous CRES
1407 experiments. This operation will then be followed by other, potentially more expensive,
1408 reconstruction steps that are required in order to determine the kinetic energy of the
1409 electron.

1410 **3.5 Pilot-scale Experiments**

1411 **3.5.1 Choice of Frequency**

1412 The optimal CRES frequency for Project 8 is that which can reach our target sensitivity
1413 of 40 meV, while minimizing the cost and complexity of the overall experiment. Since the
1414 size of the background magnetic field determines the cyclotron frequency, which affects
1415 the entirety of the CRES detection system design, specifying the operating frequency of
1416 the CRES experiments is one of the first steps towards developing a full design.

1417 **Scaling Laws**

1418 In Phases I and II the background magnetic field was provided by an NMR magnet with
1419 a 0.959 T magnetic field. This magnetic field was selected primarily for convenience,
1420 however, the cyclotron frequencies for electrons near the tritium endpoint in a 0.959 T
1421 field ranges from 25 to 26 GHz, which is within the standard RF Ka-band. Therefore,
1422 microwave electronics specialized for these frequencies are easily obtainable for relatively
1423 low cost. Frequency choice for the upcoming large-scale experiments must be selected
1424 in a more rigorous manner than in the earlier phases due to the increasing scale and
1425 complexity of the systems and the 40 meV neutrino mass science goal.

1426 Naturally, for a larger volume experiment there is a bias towards lower frequencies, due
1427 to the direct relationship between wavelength and the physical size of the compatible RF
1428 components like antennas and cavities. With a longer wavelength a larger volume can be
1429 surrounded by an array with fewer antennas, which reduces hardware and data-processing
1430 costs. On the other hand, for a cavity experiment, the volume of the experiment is
1431 directly proportional to the wavelength since this sets the physical dimensions of the
1432 cavity. Furthermore, it is easier to engineer a magnet that provides a uniform magnetic
1433 field across several cubic-meters of space at a lower magnetic fields, which provides
1434 advantages in terms of cost-reduction as well as more uniform magnetic fields for CRES.

1435 A concern with lower magnetic fields and frequencies is the scaling of the Larmor
1436 power equation, which is proportional to the square of the frequency. Naively, one would
1437 predict that the SNR would decrease with lower fields, however, two additional scaling
1438 laws that affect the noise power also come into play. Noise power is directly proportional
1439 to the required bandwidth, which decreases linearly with the magnetic field. Furthermore,
1440 at lower frequencies it is possible to purchase amplifiers with lower noise temperatures
1441 until approximately 300 MHz at which point this relationship tends to flatten. Therefore,
1442 it is expected that the SNR remains approximately constant as the frequency decreases.

1443 The SNR directly impacts the overall efficiency of the experiment through its affects
1444 on CRES signal detection probabilities as well as energy resolution. Thus, the expectation
1445 that SNR remains the same at lower frequencies clearly biases large-scale experiments
1446 in this direction. One drawback of lower magnetic fields is the increased influence of
1447 external magnetic fields on the experiment. This includes magnetic fields from the
1448 building materials as well as variations in the earth's magnetic field. To deal with these
1449 affects a suitable magnetic field correction system will need to be devised, which includes
1450 constant monitoring of external fields.

1451 **Atomic Tritium Considerations**

1452 The pilot-scale experiments will be the first Project 8 experiments to combine CRES
1453 with atomic tritium, therefore, the optimal frequency should take into account the affect
1454 of the background magnetic field size on atom trapping. The primary influence of the
1455 background field magnitude is through the rate of dipolar spin-flips caused by a spin
1456 exchange interaction between trapped atoms.

1457 Atomic tritium is a simple quantum system with a hyperfine structure given by the
1458 addition of the nuclear and atomic spins. The addition of two spins leads to a hyperfine
1459 structure with four states in the (m_s, m_I) basis. The states with atomic spins directed

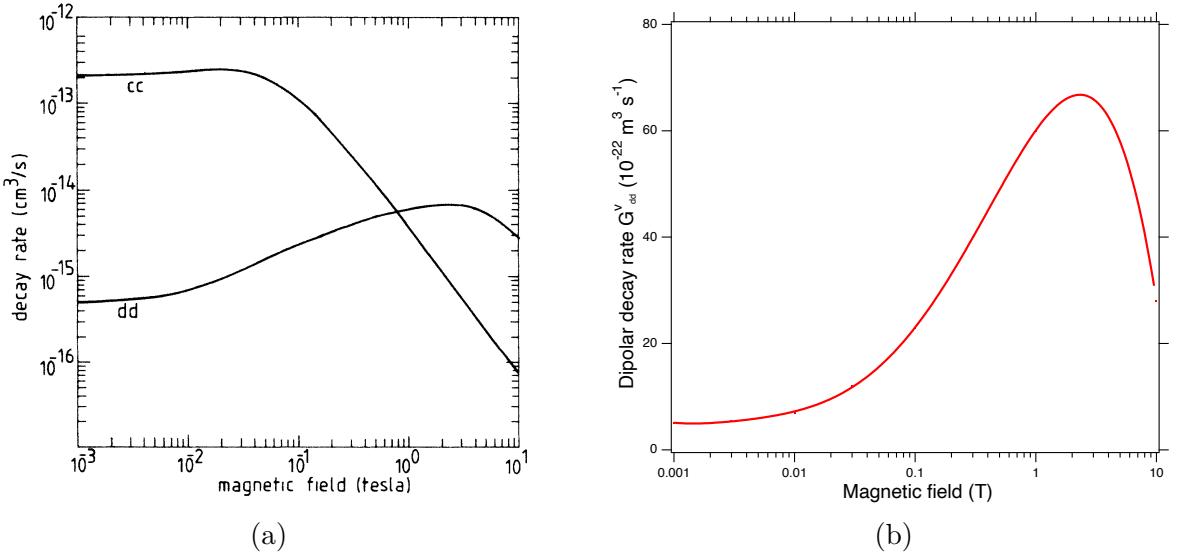


Figure 3.18: (a) A plot of the decay rate for the two-body dipolar spin exchange interaction for $c+c$ and $d+d$ state. (b) A plot of the decay rate of the dipolar spin exchange interaction for $d+d$ states as a function of magnetic field magnitude. Lowering the magnetic field is key for reducing the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

1460 anti-parallel to the magnetic field have $m_s = -1/2$ and are labeled as the a and b states.
 1461 The a and b states are colloquially known as high-field seeking states, since their energy is
 1462 minimized when in regions of higher magnetic field. This leads to losses in the magnetic
 1463 trap as these atoms are drawn to higher fields away from the trap center. Alternatively,
 1464 the c and d states, with atomic spin $m_s = +1/2$, minimize their energy in low magnetic
 1465 fields because of the parallel alignment between spin and the magnetic field. Therefore,
 1466 these low-field seeking states tend to stay trapped significantly longer than the high-field
 1467 seeking states.

1468 Project 8 would do well to prepare the tritium atoms in purely c and d states before
 1469 trapping, however, even in this case losses still occur due to dipolar interactions between
 1470 pairs of c and d states leading to a flipped atomic spins and subsequent losses due
 1471 to high-field seeking atoms. The rate of these interactions depends on the magnitude
 1472 of the background magnetic field and is maximal for dd interactions around 1 T (see
 1473 Figure 3.18). The rate of losses from these interactions at 1 T requires atomic tritium
 1474 production at a rate two orders of magnitude larger than at 0.1 T, thus, requirements
 1475 on the whole atomic tritium system are significantly relaxed at lower magnetic fields,
 1476 which provides an additional argument for transitioning to lower frequencies with the
 1477 pilot-scale experiments.

1478 3.5.2 Pilot-scale Experiment Concepts

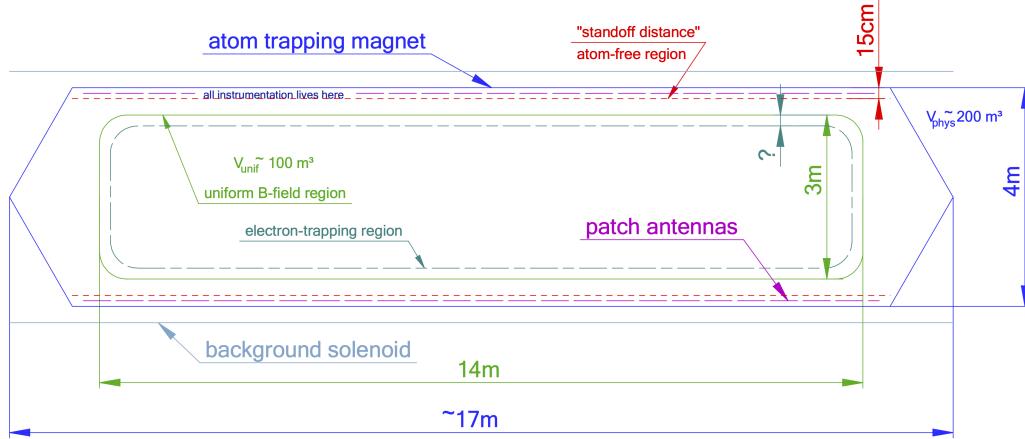


Figure 3.19: A conceptual sketch of a large-volume antenna array based CRES experiment to measure the neutrino mass.

1479 While the pilot-scale experiments are still in the early stages, enough is known to
 1480 sketch the general features of these experiments at the cartoon level.

1481 Pilot-scale Antenna Array CRES Experiment Concept

1482 A conceptual design for an antenna-based CRES experiment is shown in Figure 3.19.
 1483 A large solenoid magnet provides a uniform background magnetic field less than 0.1 T
 1484 in magnitude. Inside this region is the atom trapping magnet that generates a high
 1485 magnetic field at the walls, which decays exponentially towards the central region. Known
 1486 magnet designs that produce suitable atom trapping fields include Ioffe-Prichard traps,
 1487 which use conducting coils, as well as a Halbach array made from permanent magnets.
 1488 Either magnet choice produces a region of high magnetic fields, which excludes atoms
 1489 and allows for the placement of antennas inside the experiment.

1490 Inside this region an array of microstrip patch antennas is inserted to collect the
 1491 cyclotron radiation without providing a surface for atomic tritium recombination. Due
 1492 to the lower frequency of cyclotron radiation antennas of a larger size can be used,
 1493 which lowers the total number of antennas required to observe the experiment volume.
 1494 Because of this scaling, the lower frequency experiment uses a similar number of antennas
 1495 compared to a much smaller demonstrator experiment with a 1 T magnetic field.

1496 The atomic tritium beamline that supplies fresh tritium atoms to the experiment is
 1497 not shown in the figure. The general configuration would matches the one shown for the

¹⁴⁹⁸ pilot-scale cavity experiment (see Figure 3.20).

¹⁴⁹⁹ Pilot-scale Cavity CRES Experiment Concept

¹⁵⁰⁰ The pilot-scale cavity experiment includes both an atomic tritium system and cavity
¹⁵⁰¹ CRES system. The atomic system consists of a thermal atom cracker located at the
¹⁵⁰² start of an evaporatively cooled atomic beamline. The atomic tritium system provides a
¹⁵⁰³ supply of tritium atoms to the trap with temperatures on the order of a few mK. Atoms
¹⁵⁰⁴ at this temperature can be trapped magneto-gravitationally, which is the reason for the
¹⁵⁰⁵ vertical orientation of the cavity. At these low magnetic fields the trapping requirements
¹⁵⁰⁶ for electrons and atoms differ enough such that it is advantageous to decouple the the
¹⁵⁰⁷ trapping potentials to avoid radioactive heating of the tritium atoms from excess trapped
¹⁵⁰⁸ electrons. Electron trapping is provided by a set of magnetic pinch coils at the top and
¹⁵⁰⁹ bottom of the cavity and a multi-pole Ioffe or Halbach magnet serves to contain the
¹⁵¹⁰ atoms.

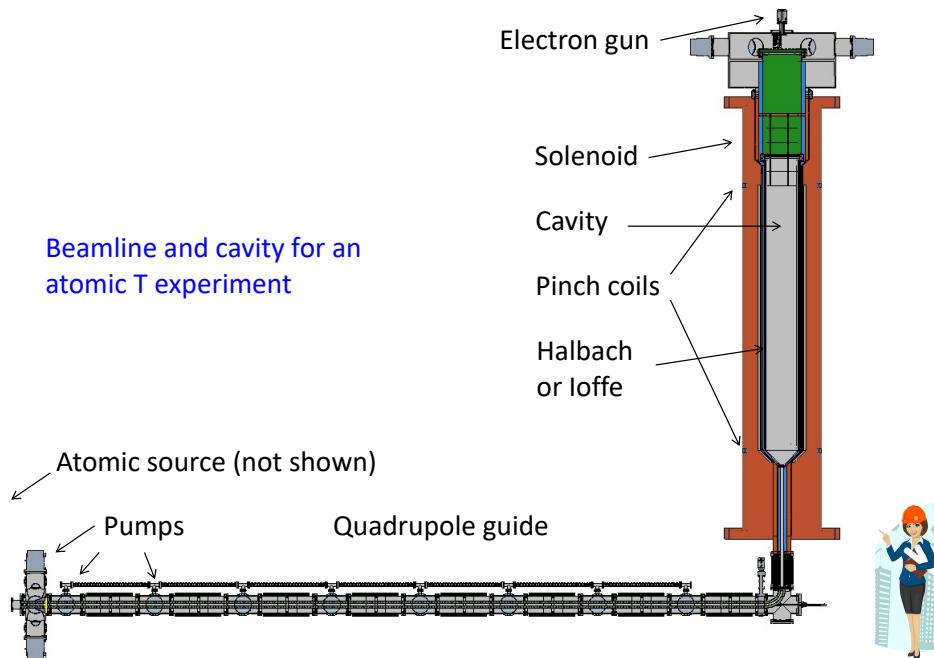


Figure 3.20: A conceptual sketch of a pilot-scale cavity CRES experiment with an atomic tritium beamline.

¹⁵¹¹ The cavity design for the pilot-scale experiment consists of a large cylindrical cavity
¹⁵¹² with a TE011 resonance of 325 MHz. Such a cavity is truly enormous, with a diameter
¹⁵¹³ of approximately 1.2 m and a height of 10 m. When an electron is produced inside
¹⁵¹⁴ the cavity with a cyclotron frequency that matches the TE011 resonant frequency it's

1515 cyclotron orbit couples the electron to the TE011, which drives a resonance in the cavity.
1516 These resonant fields can be read-out using an appropriate cavity coupling mechanism
1517 located at the center of the cavity. For more information on the cavity approach to
1518 CRES see Chapter 6.

1519 The bottom of the cavity has a cone termination to match the contour of the atom
1520 trapping magnet. This shape still allows for TE011 resonances with high internal Qs,
1521 which are required for good SNR in the cavity experiment. A small opening in the bottom
1522 of the cone serves as an entry point for the tritium atoms. To allow for calibration of
1523 the magnetic field inhomogeneities with an electron gun, the top of the cavity is left
1524 nearly completely open. Normally, this would drastically lower the Q-factor of the TE011
1525 mode, but a specially configured coaxial partition is inserted at the top. This termination
1526 scheme is designed to act as a perfect short for the TE011 mode since the circular shape
1527 of the partition matches the electric field boundary conditions for the TE011 mode.
1528 Simulations with HFSS have confirmed that this design results in a high quality TE011
1529 resonance despite the nearly completely open end.

1530 3.6 Phase IV

1531 The baseline CRES technology being pursued by the Project 8 collaboration are resonant
1532 cavities, which, due to their geometric properties, simple CRES signal structure, and low
1533 channel count, appear to be the better option for Phase IV. The current knowledge of the
1534 antenna array CRES approach reveals no technical obstacles that would preclude it as a
1535 baseline technology for Phase IV though it would most certainly be significantly more
1536 expensive. Therefore, antenna arrays represent a fallback approach if resonant cavities
1537 prove infeasible.

1538 The sensitivity of the pilot-scale atomic tritium experiment is estimated to be on
1539 the order of 0.1 eV, which means that increasing the sensitivity to reach the Phase IV
1540 goal will require a larger volume experiment. Because of the direct coupling between the
1541 RF characteristics of a cavity and its geometry, the baseline plan is to build multiple
1542 copies of the pilot-scale experiment (see Figure 3.21) to obtain the required amount of
1543 volume rather than increase the size of the cavity beyond the pilot-scale. The built-in
1544 redundancy of this approach is attractive in that the experiment has no single point of
1545 failure, additionally, building several copies of the a pilot-scale experiment will require
1546 minimal new engineering and design.

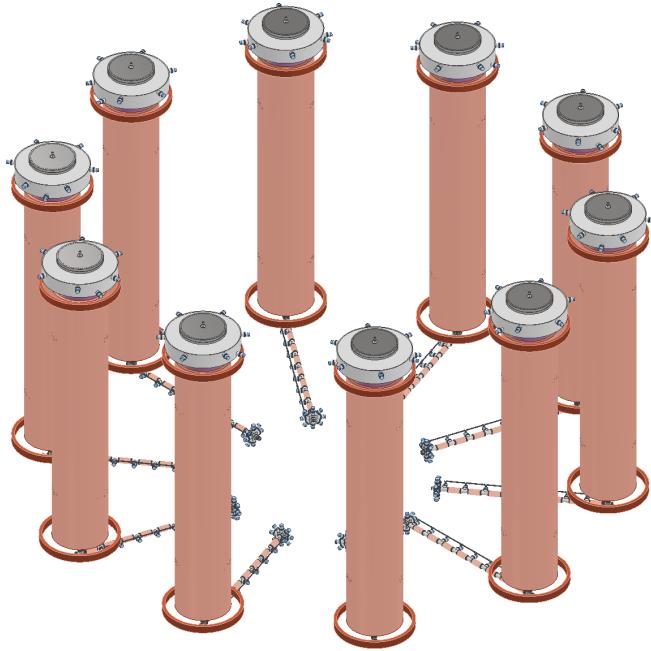


Figure 3.21: An illustration of a possible arrangement of ten pilot-scale cavity experiments for Phase IV. The experiments are arranged in a circle with an approximate diameter of 50 meters. Each atomic beamline connected to the bottom of each cavity is approximately 10 m in length. The cavities themselves 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 5.4.

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

1573 only with distinguishing between data that contains a signal versus data that contains only
1574 noise, whereas, with parameter estimation one is interested in extracting the kinematic
1575 parameters of the electron encoded in the cyclotron radiation signal shape. Due to
1576 the low signal power of electrons near the spectrum endpoint in the FSCD experiment,
1577 signal detection is a non-trivial problem. This is magnified by the need to maximize the
1578 detection efficiency of the experiment in order to achieve the neutrino mass sensitivity
1579 goals. My contributions to signal reconstruction analysis for the FSCD are focused on
1580 this signal detection component of reconstruction.

1581 After the discussion of various signal detection approaches, in Section 4.4 I present a
1582 more detailed analysis of the detection performance of three algorithms, which could be
1583 used to signal detection in the FSCD. This section was originally prepared for publication
1584 in JINST as a separate paper. The algorithms include a digital beamforming algorithm,
1585 a matched filter algorithm, and a neural network algorithm, which I analyze in terms of
1586 classification accuracy and estimated computational cost.

1587 **4.2 FSCD Simulations**

1588 Antenna array CRES and the FSCD requires a combination of different capabilities
1589 not often found in a single simulation tool. First of all, accurate calculations of the
1590 magneto-static fields produced by current-carrying coils are required in order to accurately
1591 model the magnetic trap and background magnets. The resulting magnetic fields must
1592 then be used to calculate the exact relativistic trajectory of electrons, which is required
1593 in order to calculate the electro-magnetic (EM) fields produced by the acceleration of
1594 the electron. Finally, the simulation has to model the interaction of the antenna and
1595 RF receiver chain with these EM-fields in order to produce the simulated voltage signals
1596 produced by the antenna array during the CRES event. At the time when Project 8 was
1597 developing this simulation capability, no single available simulation tool was known to
1598 adequately perform this suite of calculations, which prompted the development of custom
1599 simulation framework to simulate the FSCD. This simulation framework includes custom
1600 simulation tools developed by Project 8 as well as other open-source and proprietary
1601 software developed by third-parties.

₁₆₀₂ **4.2.1 Kassiopeia**

₁₆₀₃ Kassiopeia¹ is a particle tracking and static EM-field solver developed by the KATRIN
₁₆₀₄ collaboration for simulations of their spectrometer based on magnetic adiabatic collimation
₁₆₀₅ with an electrostatic filter [7]. Due to the measurement technique employed by the
₁₆₀₆ KATRIN collaboration, Kassiopeia is not designed to solve for the EM-fields produced by
₁₆₀₇ electrons in magnetic fields. However, it does provide efficient solvers for static electric
₁₆₀₈ and magnetic fields and charged particle trajectory solvers. Because of this, Project 8
₁₆₀₉ has incorporated parts of Kassiopeia into its own simulation framework.

₁₆₁₀ **Magnetostatic Field Solutions**

₁₆₁₁ The solutions to the electric and magnetic fields generated by a static configuration of
₁₆₁₂ charges and currents is given by Maxwell's equations in the limit where the time-dependent
₁₆₁₃ terms go to zero. In their static form Maxwell's equations [8] are

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

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

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

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

₁₆₁₄ where we can see that the electric and magnetic fields are now completely decoupled
₁₆₁₅ from each other. The solution for the magnetic field in this boundary value problem is
₁₆₁₆ given by the Biot-Savart law

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

₁₆₁₇ which Kassiopeia uses a variety of numeric integration techniques to solve for a user
₁₆₁₈ defined current distribution.

₁₆₁₉ **Kassiopeia Simulation of the FSCD Magnetic Trap**

₁₆₂₀ The trap developed for the FSCD experiment utilizes six current carrying coils, which
₁₆₂₁ surround a cylindrical tritium containment vessel (see Figure 4.1). Some critical aspects
₁₆₂₂ of the trap design include the total trapping volume, the maximum trap depth, the

¹<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

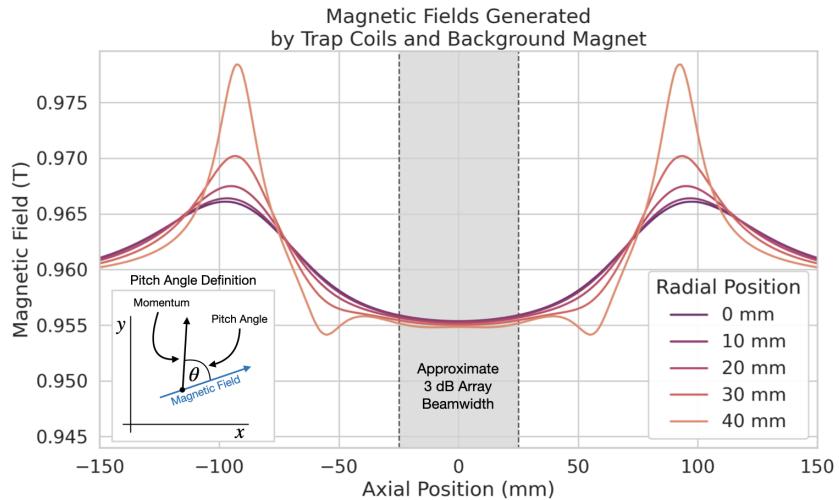
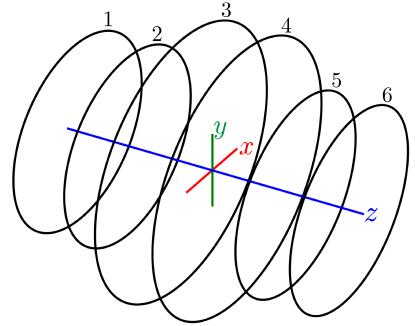


Figure 4.1: The geometry and parameters of the coils used to simulate the FSCD magnetic trap in Kassiopeia. Some axial profiles of the magnetic trap at different radial positions are show to demonstrate the shape of the magnetic field and trap depth as a function of position. Calculation of the magnetic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

The depth of the FSCD trap is approximately 10 mT when measured along the central axis, which is sufficient to trap electrons with pitch angles as small as 84° . The

trap depth 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° [9]. 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

1667 equations

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

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

1668 where \mathbf{r} is the position of the electron, \mathbf{p} is the electron's momentum, e is the charge of
1669 the electron, m is the electron's mass, and γ is the relativistic Lorentz term. To account
1670 for kinetic energy losses from radiation Kassiopeia includes an additional term in the
1671 momentum differential equation, which calculates the change in the electron's momentum
1672 induced by synchrotron radiation. Kassiopeia solves this pair of differential equations
1673 using numerical integration, however, the exact trajectory can be computationally
1674 intensive to solve. If the adiabatic approximation can be applied, then Kassiopeia can
1675 make use of a simpler set of equations that can be more readily solved numerically.

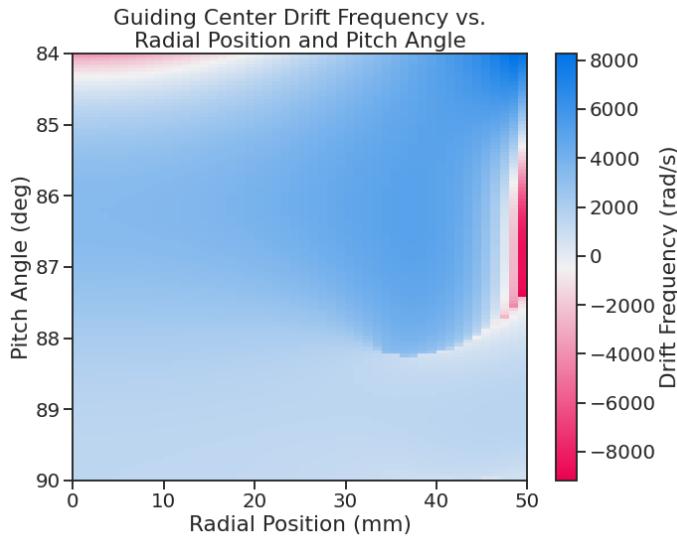


Figure 4.2: A map of the average ∇B -drift frequency for electrons trapped in the prototype FSCD trap shown in Figure 4.1. Negative drift frequencies indicate electrons that are drifting opposite to the standard direction, which means that they are close to escaping the magnetic trap.

1676 Even though Kassiopeia is not directly capable of simulating the cyclotron radiation,
1677 it is still an invaluable CRES simulation tool, due to the accurate trajectory solutions
1678 for electrons in magnetic traps. With Kassiopeia it is possible to test the efficiency of a
1679 particular trap design and analyze features of the electron trajectories that are important
1680 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 ∇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 Locust² software package [10] 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 [11, 12], which are obtained by
 differentiating the Liénard-Wiechert potentials. In their full form the Liénard-Wiechert
 field equations are

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

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

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

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

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

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

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

Antenna Response Modeling

With the electric field it is possible, in principle, to calculate the resulting voltages produced in the antenna. However, direct simulation of the antenna itself is computationally expensive since it would require the modeling of complex interactions of the electron's electric fields with charge carriers in the conductive elements of the antenna. Direct simulation of the antenna in Locust can be avoided by modeling the antenna response using the antenna factor, or antenna transfer function, approach. The antenna factor defines the voltage produced in the antenna terminal for an incident electric field [13],

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

1738 where V is the voltage and $|\mathbf{E}|$ is the magnitude of the incident electric field. To obtain the
 1739 antenna factor for the antennas developed for the FSCD Project 8 employs Ansys HFSS.
 1740 HFSS is a commercially available finite element method electromagnetic solver widely
 1741 used throughout the antenna engineering industry [14]. HFSS is capable of calculating
 1742 the antenna factor and gain patterns for complex antenna designs and outputting the
 1743 resulting quantities in the form of a text file that can be used as an input to the Locust
 1744 simulation.

1745 The antenna factor defines the steady-state response of the antenna to electromagnetic
 1746 plane waves and is a function of the frequency of the radiation. Therefore, in order to
 1747 apply the transfer function for the calculation of the antenna voltage response in the
 1748 time domain, Locust models the antenna as a linear time-invariant system [15]. In this
 1749 formalism the response of the system to the driving force is given by

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

1750 where $y[n]$ is the discretely sampled response, x is the driving force stimulus, and h is
 1751 the finite impulse response (FIR) filter. When applied to the FSCD array, this formalism
 1752 calculates the voltage time-series produced in each antenna by convolving the electric
 1753 field time-series with the antenna FIR filter, which is obtained by performing a inverse
 1754 Fourier transform on the transfer function from HFSS.

1755 Radio-frequency Receiver and Signal Processing

1756 After obtaining the voltage time-series by computing the electron trajectory and antenna
 1757 response, Locust simulates the signal processing associated with the radio-frequency
 1758 receiver chain. The standard receiver chain used in Locust simulations of the FSCD
 1759 attempts to mimic the operations that would actually occur in hardware (see Figure 4.3).

1760 Frequency down-conversion is used in the FSCD to reduce the digitization bandwidth
 1761 required to read-out CRES data. According to the Nyquist sampling theorem [16], the
 1762 minimal sampling rate that guarantees no information loss for a signal with a bandwidth
 1763 Δf is given by

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

1764 The total bandwidth of CRES signal frequencies from tritium beta-decay ranges from 0
 1765 to 26 GHz in a 0.95 T magnetic field, therefore, direct digitization of CRES signals from
 1766 the FSCD would require sampling frequencies greater than 50 GHz, which is infeasible for

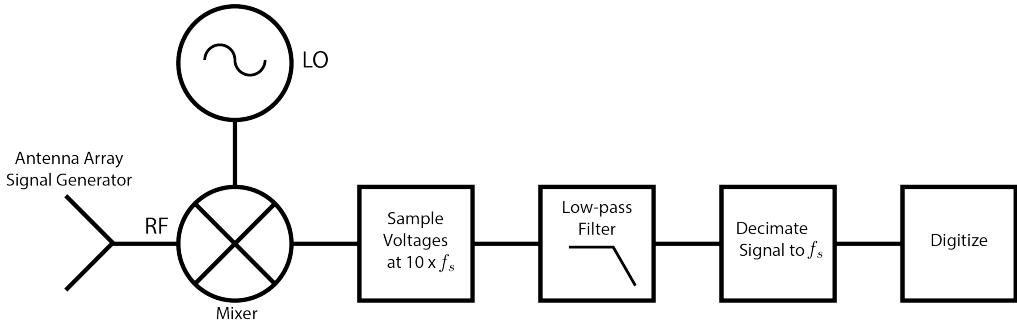


Figure 4.3: The receiver chain used by Locust when simulating CRES events in the FSCD.

1767 a real experiment. However, for the purposes of neutrino mass measurement we are only
 1768 interested in measuring the shape of the spectrum in the last 100 eV, which corresponds
 1769 to a frequency bandwidth of 5 MHz. Down-conversion is a technique for reducing the
 1770 base frequencies of signals in a bandwidth given by $[f_{\text{LO}}, f_{\text{LO}} + \Delta f]$ to the bandwidth
 1771 $[0, \Delta f]$, by performing the following multiplication

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

1772 In down-conversion the signal ($x(t)$) is multiplied by a sinusoidal signal with frequency
 1773 f_{LO} to reduce the absolute frequencies of the signals in the bandwidth. In the FSCD this
 1774 allows us to detect events in the last 100 eV of the tritium spectrum while sampling the
 1775 data far below 50 GHz. The standard bandwidth used in the FSCD is 200 MHz, which
 1776 allows for higher frequency resolution than the minimum sampling frequency for 100 eV
 1777 of energy bandwidth.

1778 Trying to directly simulate down-conversion with a frequency multiplication in Locust
 1779 would require the sampling of the electric fields at each antenna in the FSCD array with
 1780 a period of ≈ 20 ps, which is extremely slow computationally. To avoid this Locust
 1781 performs the down-conversion by intentionally under-sampling the electric fields with
 1782 a frequency of 2 GHz. Sampling below the Nyquist limit causes the higher frequency
 1783 components of the CRES signal to alias, however, Locust can remove these aliased
 1784 frequency peaks using a combination of low-pass filtering and decimation to recreate
 1785 frequency down-conversion. After filtering and decimation, Locust simulates digitization
 1786 by an 8-bit digitizer at a sampling frequency of 200 MHz to recreate the conditions of
 1787 the FSCD. The voltage offset and the digitizer range must be configured by the user
 1788 based on the characteristics of the simulation.

1789 **Data**

1790 The output of Locust simulations for the FSCD primarily consists of two data files. The
1791 first is the electron trajectory information calculated by Kassiopiea, which is output in
1792 the form of a `.root` file [17]. This file contains important kinematic information about
1793 the electron such as it's position and pitch angle as a function of time. The other file is
1794 produced by Locust and it contains the digitized signals acquired from each antenna in
1795 the FSCD array. The Locust output files conform to the Monarch specification developed
1796 by Project 8, which is based on the commonly used HDF5 file format, and matches the
1797 format of the files produced by the Project 8 data acquisition software. This makes it
1798 possible to use the same data analysis code to analyze both simulated and real data.

1799 **4.2.3 CRESana**

1800 Locust is the primary simulation tool used by Project 8 in the development and simulation
1801 of the FSCD. However, simulations of CRES events in larger antenna arrays (≥ 100
1802 antennas) using Locust can take several hours to complete, which is prohibitively long
1803 when one is performing a sensitivity analysis for a large scale antenna experiment. One
1804 of the reasons for Locust's slow operation is that the electric fields from the electron
1805 must be solved numerically for each time-step for each of the antennas in the array.
1806 These numerical solutions allow Locust to accurately simulate the electric fields from
1807 arbitrarily complicated electron trajectories at the cost of more computations and slower
1808 simulations. Therefore, an additional simulation tool that sacrifices some accuracy for
1809 computational efficiency would be extremely useful simulations and sensitivity analyses
1810 of larger antenna array experiments.

1811 To fill this need, Project has developed a new simulations package called CRESana³,
1812 specifically designed to perform analytical simulations of antenna array based CRES
1813 experiments. CRESana is not as flexible as Locust, but it provides a significant increase
1814 in simulation speed. It does this by using well-justified analytical approximations of the
1815 electrons motion in the magnetic field and the resulting electric fields from the electron's
1816 acceleration. The electric fields and signals generated by CRESana are consistent with
1817 theoretical calculations of the electron's radiation, and are test for accuracy using
1818 well-known test-case simulations and consistency checks.

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

1819 4.3 Signal Detection and Reconstruction Techniques for 1820 Antenna Array CRES

1821 Antenna Array CRES Signal Reconstruction

1822 A robust set of FSCD simulation tools are vital to the development of the analysis
1823 algorithms necessary for antenna array CRES to succeed. In order to perform CRES
1824 measurements using an antenna array, one must develop an algorithm that uses the
1825 multi-channel time-series obtained by digitizing the array to estimate the starting kinetic
1826 energies of electrons produced in the magnetic trap. This procedure consists of a multi-
1827 stage process of detecting a CRES signal then estimating the parameters of the electron
1828 that produced and is often referred to as simply CRES signal reconstruction.

1829 Compared with the signal reconstruction approaches of the Phase I and II CRES
1830 experiments, antenna array CRES requires a significantly different approach to signal
1831 reconstruction. In Phase I and II, CRES was performed using a waveguide gas cell that
1832 could be directly connected to a waveguide transmission line. The transmission line
1833 efficiently transmits the cyclotron radiation along its length to an antenna at either end
1834 of the waveguide. However, with an antenna array the electron is essentially radiating
1835 into free-space, therefore, the cyclotron radiation power collected by the array is directly
1836 proportional to the solid angle surrounding the electron that is covered with antennas.
1837 Because it is not practical to fully surround the magnetic trap with antennas, some of the
1838 cyclotron radiation power that would have been collected by the waveguide escapes into
1839 free-space. Furthermore, the power that is collected by the antenna array is split between
1840 every channel in the antenna array, which significantly lowers the signal-to-noise ratio
1841 (SNR) of CRES signals in a single antenna channel compared to a waveguide apparatus.
1842 Therefore, a suite of completely new signal reconstruction techniques are needed in order
1843 to perform CRES in the FSCD.

1844 Changes to the approach to CRES signal reconstruction are also motivated by the
1845 more ambitious scientific goals of the FSCD experiment. A measurement of the tritium
1846 beta-decay spectrum that is sensitive to neutrino masses as small as 40 meV requires that
1847 we measure the kinetic energies of individual electrons with a total energy broadening
1848 of 115 meV [18]. This resolution includes all sources of uncertainty in the electron's
1849 kinetic energy such as magnetic field inhomogeneities. This level of energy resolution is
1850 compatible only with an event-by-event signal reconstruction approach where the kinetic
1851 energies, pitch angles, and other parameters of the CRES events are estimated before

1852 constructing the beta-decay spectrum.

1853 The event-by-event approach is distinct from the analysis done for the Phase I and
1854 Phase II experiments where only the starting cyclotron frequency of the event was
1855 estimated by analyzing the tracks formed by the carrier frequency in the time-frequency
1856 spectrogram. These frequencies were then combined into a frequency spectrogram, which
1857 was converted to the beta-decay energy spectrum using an ensemble approach that
1858 averaged over all other event parameters. The ensemble approach to signal reconstruction
1859 results in poor energy resolution because other kinematic parameters such as pitch angle
1860 change the cyclotron carrier frequency due to changes in the average magnetic field
1861 experience by the electron, and it is therefore incompatible with the future goals of the
1862 Project 8 collaboration.

1863 Components of Reconstruction: Signal Detection and Parameter Estimation

1864 CRES signal reconstruction can be viewed as a two-step procedure consisting of signal
1865 detection followed by parameter estimation. In the former, one is concerned with
1866 identifying CRES signals in the data regardless of the signal parameters, whereas, in the
1867 latter one operates under the assumption that a signal is present and then estimates its
1868 parameters.

1869 More formally, signal detection is essentially a binary hypothesis test between the
1870 signal and noise data classes and parameter estimation describes a procedure of fitting a
1871 model to the observed data. While both of these processes are required for a complete
1872 reconstruction (see Figure 4.4), the focus of my work and this chapter is on the signal
1873 detection aspect of antenna array CRES signal reconstruction.

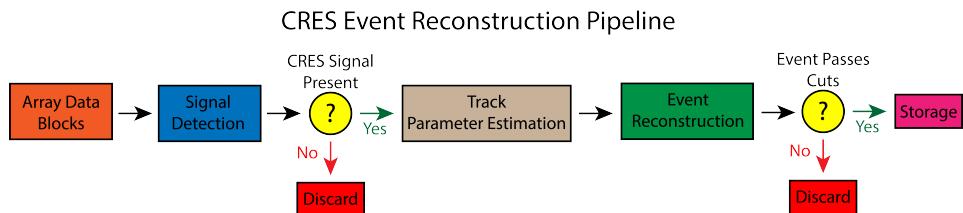


Figure 4.4: A high-level diagram depicting the process of CRES event reconstruction. The first step consists of identifying the presence of a signal in the data. This step is necessary to avoid the danger of performing a reconstruction of a false event, which would constitute a background contribution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

1874 **Detection Theory**

1875 The problem of signal detection can be posed as a statistical hypothesis test [19]. For
1876 CRES signals, which are essentially vectors with added white Gaussian noise (WGN),
1877 one needs to choose between two hypotheses

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

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

1878 where \mathbf{y} is the CRES data vector, $\boldsymbol{\nu}$ is a sample of WGN, and \mathbf{x} represents the CRES
1879 signal. The hypothesis that the data contains only noise is labeled \mathcal{H}_0 and the hypothesis
1880 that the data contains a signal is labeled \mathcal{H}_1 .

1881 For illustrative purposes one can examine the case where one the first sample of
1882 data is used to distinguish between \mathcal{H}_0 and \mathcal{H}_1 . The value of the first data sample is
1883 distributed according to two gaussian distributions corresponding to \mathcal{H}_0 and \mathcal{H}_1 (see
1884 Figure 4.5). By setting a decision threshold on the value of this sample, one can choose
1885 the correct hypothesis with a probability given by the areas underneath the probability
1886 distribution curves. A true positive corresponds to correctly identifying that the data
1887 contains signal, whereas, a true negative means that one has correctly identified the data
1888 as noise. The rate at which the detector performs a true positive classification is given
1889 by the green region underneath $p(\mathbf{y}[0]; \mathcal{H}_0)$, and the rate at which the detector performs
1890 a true negative classification is given by the orange region underneath $p(\mathbf{y}[0]; \mathcal{H}_1)$. Two
1891 types of misclassifications are possible. Either we declare noise data as signal, which is
1892 call a false positive, or we declare signal data as noise, which is a false negative. Note
1893 that it is only possible to trade off these two types of errors by tuning the detection
1894 threshold. One cannot simultaneously reduce the rate of false positives without also
1895 increasing the rate of false negatives.

1896 The approach taken with CRES signals is to fix the rate of false positives by setting
1897 a minimum value for a detection threshold. The rate of false positives that is acceptable
1898 at the detection stage depends upon the rate of background events compatible with the
1899 sensitivity goals of the experiment. The ultimate goal of a neutrino mass measurement
1900 with 40 meV sensitivity in general has strict requirements on the number of background
1901 events, which requires a relatively high detection threshold to achieve. Consequently,
1902 the ideal signal detection algorithm is the one that achieves the maximum rate of true
1903 positives for a fixed rate of false positives, so that the detection efficiency of the experiment
1904 is maximized and potential sources of background are kept to a minimum.

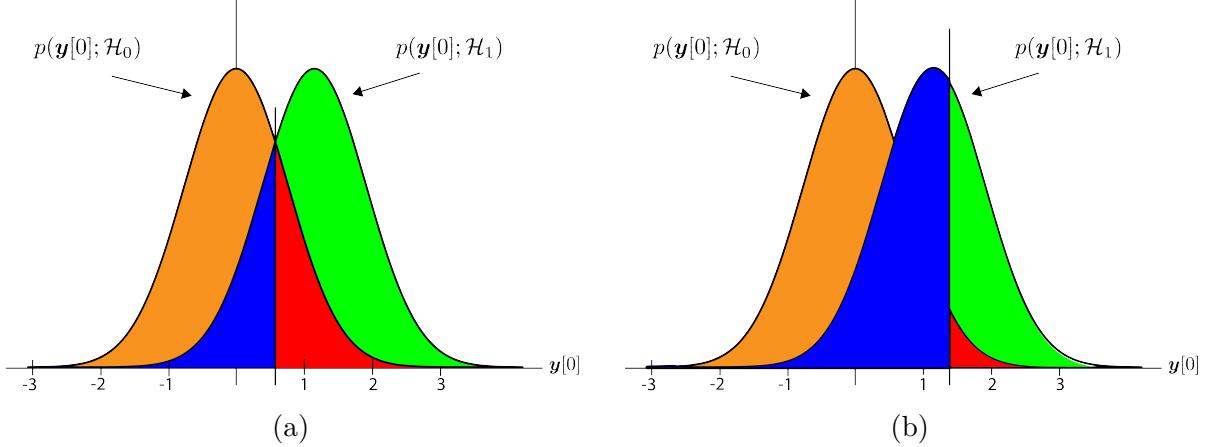


Figure 4.5: An illustration of two PDFs associated with a binary hypothesis test. The decision threshold is represented by the vertical line that partitions both distributions. The orange and red areas correspond to the true negative and false positive probabilities and the blue and green areas correspond 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 [20], 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 γ 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 α 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_{TP} and P_{FP} as

appropriate for the given situation. It is common to summarize the relationship between P_{TP} and P_{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

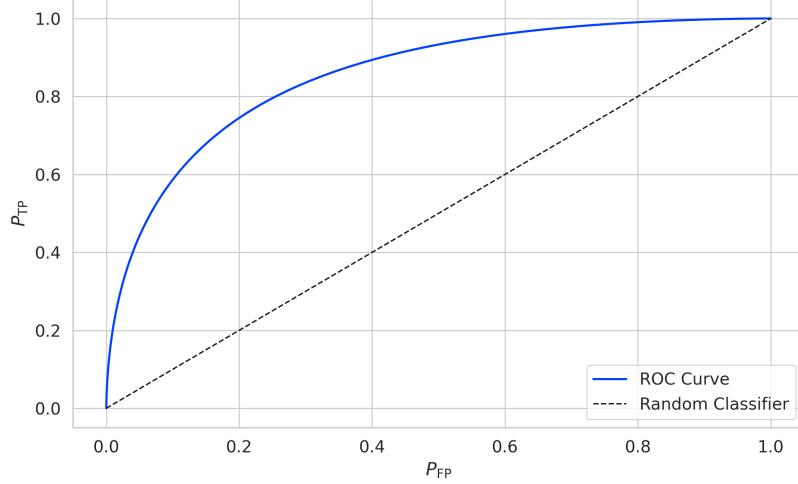


Figure 4.6: An example ROC curve formed by computing the P_{FP} and the P_{TP} for a given likelihood ratio test. As the decision threshold is increased P_{FP} decreases at the expense of a lower P_{TP} . The black dashed line indicates the lower bound ROC curve obtained by randomly deciding between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

the performance of different signal detection algorithms. In general, a classifier with a higher the P_{TP} as a function of P_{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 [13]. 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

1933 emitted by the array will constructively interfere at the point of interest (see Figure
 1934 4.7). As a consequence of the principle of reciprocity [21], when the array is operating in
 1935 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-

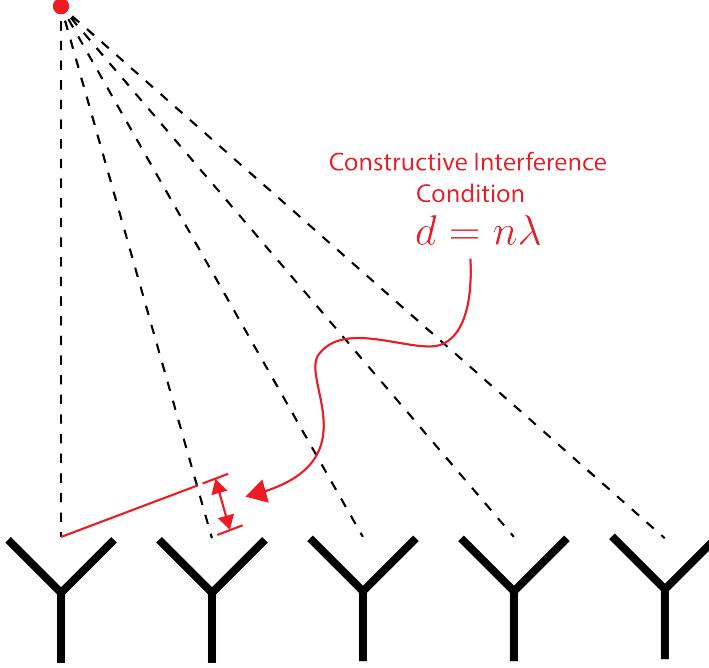


Figure 4.7: An illustration of the constructive interference condition which is the operating principle of digital beamforming using a uniform linear array as an example.

1936
 1937 length difference to the beamforming point between different antennas in the array. The
 1938 relationship between the phase delay and the path-length difference is given by the
 1939 familiar equation

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

1940 where ϕ is the phase delay, d is the path-length difference, and λ is the wavelength of
 1941 the radiation. In practice, one chooses the values of d by specifying the beamforming
 1942 positions of interest and then calculates the beamforming phases using Equation 4.20,
 1943 which is guaranteed to follow the constructive interference condition shown in Figure 4.7.

1944 Beamforming can be neatly expressed mathematically using the vector equation

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

1945 where $\mathbf{x}[n]$ is the array snapshot vector, $\Phi[n]$ is a vector of beamforming shifts, and
 1946 $y[n]$ is the resulting summed signal. The beamforming shifts consist of a set of complex

1947 numbers that contain the beamforming phase shift and an amplitude weighting factor,

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

1948 where the set of magnitudes $A_i[n]$ are amplitude weighting factors and $\phi_i[n]$ are the phase
1949 shifts from the path-length differences. The index i is used to denote the antenna channel
1950 number. The amplitude weighting factor is the relative magnitude of the signal received
1951 by a particular antenna to the other antennas in the array, such that the antennas that
1952 receive signals with higher amplitude, due to being closer to the source, have more
1953 weight in the beamforming summation. The input and outputs signals beamforming
1954 are naturally expected to be functions of time as indicated by the index $[n]$, however, it
1955 is also possible to use time dependent beamforming phases that shift the beamforming
1956 position of the array over time.

1957 Digital beamforming is the type of beamforming algorithm of interest to Project 8 for
1958 CRES. Specifically, digital beamforming means that the beamforming phases are applied
1959 to the array signals in software rather than employing fixed beamforming phase shifts in
1960 the receiver chain hardware. The advantage of digital beamforming is that for a given
1961 series of array snapshots one can specify a large number of beamforming positions and
1962 effectively search for electrons by performing the beamforming summation associated
1963 with each point and applying a signal detection algorithm to identify the presence of a
1964 CRES signal.

1965 One of the most attractive features of digital beamforming is the spatial filtering
1966 effect, which is a direct consequence of the constructive interference condition used to
1967 define the beamforming phases. Spatial filtering allows for signals from multiple electrons
1968 at different positions in the trap to be effectively separated, because the constructive
1969 interference condition will force the signals from electrons at positions different from the
1970 beamforming position to cancel. This helps to reduce signal pile-up that could become
1971 an issue for large scale CRES experiments using a dense tritium source.

1972 The digital beamforming positions can be specified with arbitrary densities limited
1973 only by the available computational resources. This provides a very straight-forward way
1974 to estimate the position of the electron in the trap by using a dense grid of beamforming
1975 positions and maximizing the output power of the beamforming summation over this
1976 grid. This natural approach to position reconstruction is attractive due the requirements
1977 of an event-by-event signal reconstruction, which needs an accurate estimation of the
1978 exact magnetic field experienced by the electron in order to correctly estimate it's kinetic

1979 energy. Combined with an accurate map of the magnetic field inhomogeneities of the
1980 trap obtained from calibrations, beamforming allows one to apply this magnetic field
1981 correction with a spatial resolution that is a fraction of the cyclotron wavelength.

1982 **Laboratory Beamforming Demonstrations**

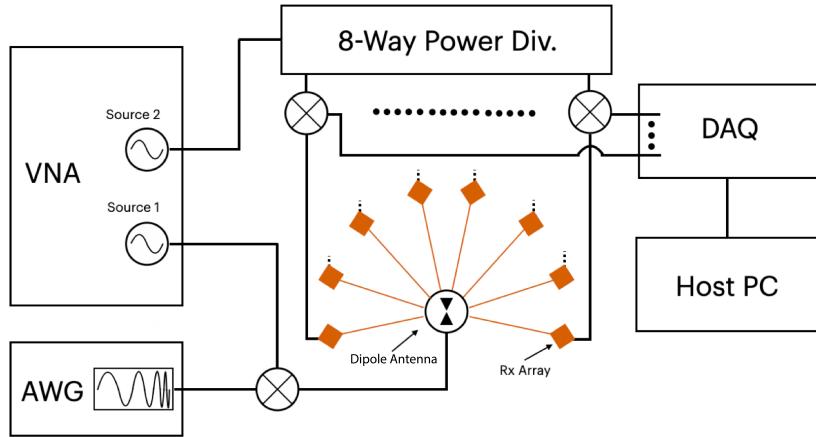


Figure 4.8: System level diagram of the laboratory setup used for beamforming demonstrations at Penn State. For more information on this system see Chapter 5. Signals near 26 GHz are fed to a dipole antenna using an arbitrary waveform generator (AWG) and vector network analyzer (VNA), which drive a mixer.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

1983 As part of the development of antenna array CRES for the FSCD, an antenna
1984 measurement setup was constructed at Penn State to serve as a testbed for antenna
1985 prototypes and to perform laboratory validations of array simulations. This system
1986 is discussed in more detail in Chapter 5. Early versions of the antenna measurement
1987 system (see Figure 4.8 and Figure 4.9) were used to perform beamforming reconstruction
1988 studies of a simple probe antenna to better understand the principles of beamforming
1989 and confirm the estimated beamforming performance of Locust.

1990 Signals from an arbitrary waveform generator were up-converted to 26 GHz using a
1991 mixer and a high-frequency source from a vector network analyzer and fed to the dipole
1992 antenna through a balun. The radiation from the dipole antenna was received by an
1993 array of horn antennas. The signals from the horn antennas were then down-converted
1994 to baseband using a collection of mixers and an 8-way power divider. The signals were
1995 then digitized and saved to a host computer for analysis.

1996 The data collected using the dipole and horn antenna array is reconstructed using the

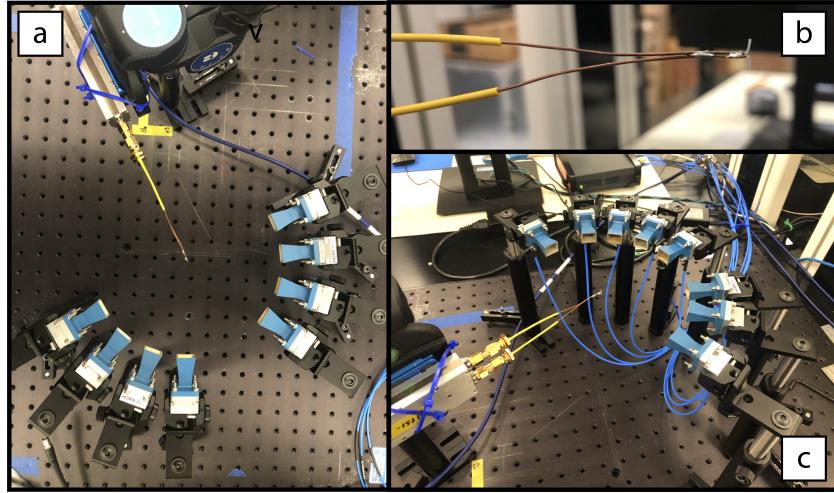


Figure 4.9: Photographs of the beamforming demonstration setup. In (a) I show a top-down view of the dipole antenna and the array of eight horn antennas. Manual repositioning of the horn antennas allows one to synthesize a full-circular antenna array. The dipole antenna is mounted on a camera tripod 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.

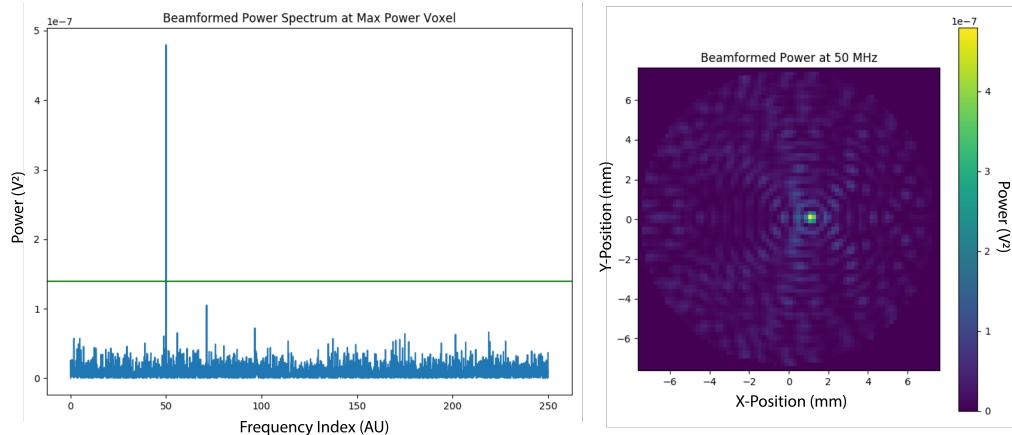


Figure 4.10: An example of digital beamforming reconstruction of a dipole antenna using a synthetic array of horn antennas. The beamforming image on the right is constructed by computing the time-averaged power of the summed signals for a two-dimensional grid of beamforming positions. In the image one 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.

beamforming reconstruction approach specified in Section 4.3.1. A two-dimensional grid 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 beings 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

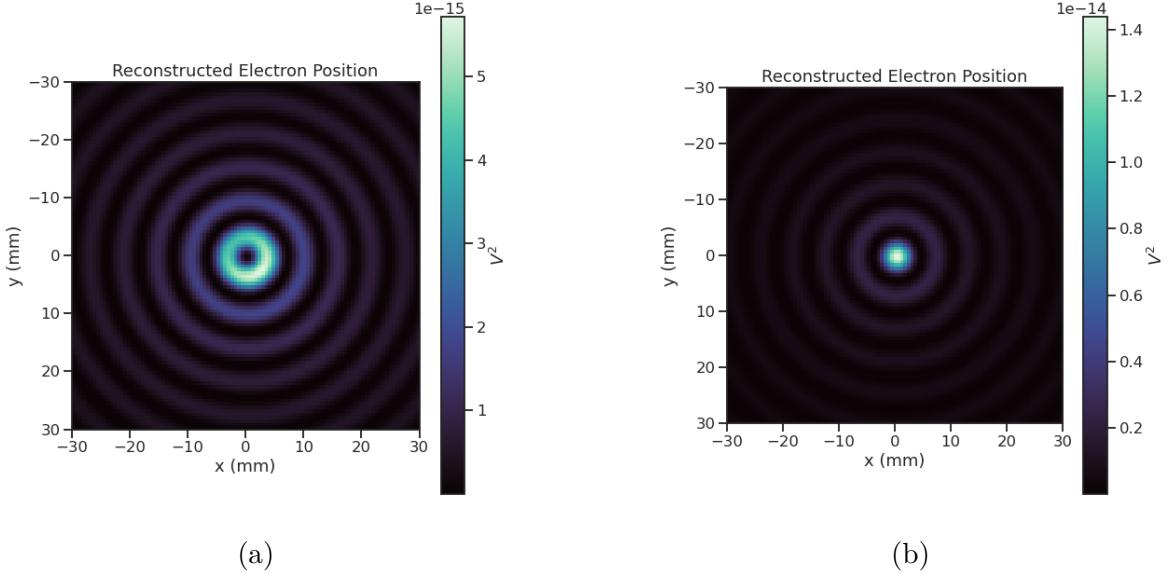


Figure 4.11: Beamforming images visualizing the reconstruction of an electron without (a) and with (b) the cyclotron phase correction. The images were generated using data from Locust simulations. The cyclotron phase refers to a phase offset equal to the relative azimuthal position of an antenna in the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

the third dimension contains the summed time-series. The image is formed by taking the time-averaged power (see Figure 4.11). Beamforming images are purely for the purposes of visualization and are not particularly useful for signal detection or reconstruction.

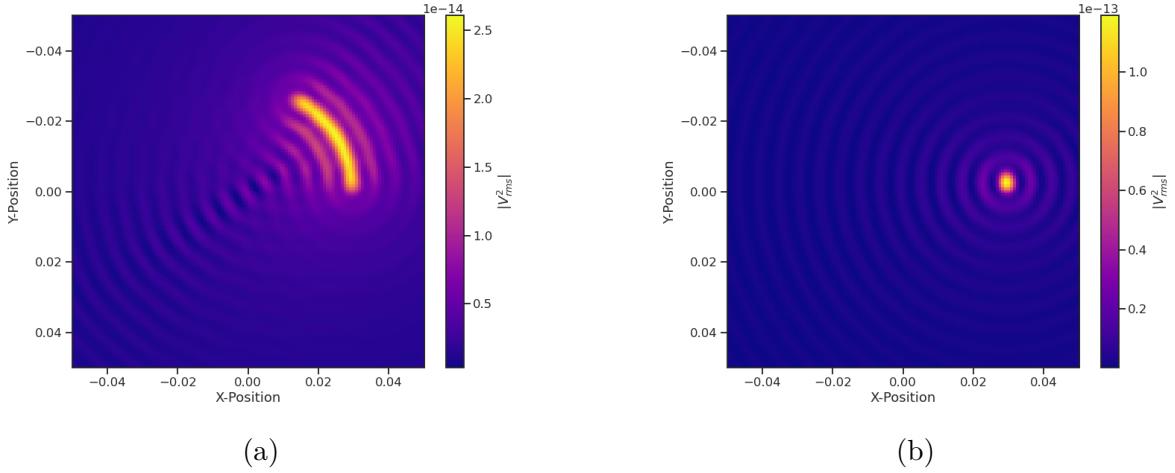
If the beamforming phases consist only of the spatial phase component from Equation 4.20, then the resulting beamforming image contains a relatively high-power ring-shaped region that is centered on the position of the electron (see Figure 4.11a). The origin of this shape is an additional phase offset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

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

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

2047 antenna channel. Using the updated beamforming phases in the summation changes the
 2048 ring feature into a Bessel function peak whose maximum corresponds to the position of
 2049 the electron. Including this cyclotron phase correction significantly improves the signal
 2050 detection and reconstruction capabilities of beamforming by more than doubling the
 2051 summed signal power and shrinking the beamforming maxima feature size.

2052 The beamforming image examples in Figure 4.11 were produced using an electron
 2053 located on the central axis of the magnetic trap, which do not experience ∇B -drift.
 2054 However, for electrons produced at non-zero radial position the beamforming phases
 2055 must be made time-dependent in order to track the position of the electron's guiding
 2056 center over time. Without this correction the ∇B -drift causes the electron to move
 2057 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



2058 Figure 4.12: Beamforming images visualizing the reconstruction of an electron located
 2059 off the central axis of the FSCD trap. In (a) we performing beamforming without the
 2060 ∇B -drift correction, and in (b) we include the ∇B -drift correction.

2061 reduces the power of the beamforming maxima and increases the size of the beamforming
 2062 features, simultaneously harming detection efficiency and position reconstruction.

2063 The ∇B -drift correction simply adds a circular time-dependence to the beamforming
 2064 positions as a function of time,

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

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

2063 where $\omega_{\nabla B}$ is the drift frequency and $t[n]$ is the time vector. In the ideal case the ∇B -drift

2064 frequencies from Figure 4.2 for the correct pitch angle and radial position would be used,
2065 however, it is not possible to know the electron’s pitch angle a priori. In principle, one
2066 could perform multiple beamforming summations for a given beamforming position using
2067 different drift frequencies and choose the one that maximizes the summed power, but
2068 this approach leads to a huge computational burden that would be impractical for a
2069 real FSCD experiment. A compromise is to use an average value of $\omega_{\nabla B}$ obtained by
2070 averaging over the drift frequencies for electrons of different pitch angle at a particular
2071 radius. This approach keeps the computational cost of time-dependent beamforming to a
2072 minimum while still providing a significant increase in the detection efficiency of digital
2073 beamforming.

2074 **Signal Detection with Beamforming and a Power Threshold**

2075 Up to this point we have neglected any specific discussion of how digital beamforming is
2076 used for signal detection and reconstruction. This is because, strictly speaking, digital
2077 beamforming consists only of the phased summation of the array signals and cannot
2078 be used alone for signal detection. The example beamforming images shown in Figure
2079 4.11 and Figure 4.12 were produced using simulated data that contained no noise, which
2080 significantly degrades the utility of analyzing the beamforming images for signal detection
2081 and reconstruction.

2082 Digital beamforming as a detection algorithm is understood to mean digital beam-
2083 forming plus a detection threshold placed on the amplitude of the frequency spectrum
2084 obtained by applying a fast Fourier transform (FFT) to the summed time-series (see
2085 Figure 4.13). This approach is most similar to the time-frequency spectrogram analysis
2086 employed in previous CRES experiments, however, in principle any signal detection
2087 algorithm could be used after the beamforming procedure. In Section 4.4 I analyze the
2088 signal detection performance of the power threshold approach in detail.

2089 From the example frequency spectra in Figure 4.13 it is clear that without a re-
2090 construction technique that coherently combines the signals from the full antenna our
2091 ability to detect CRES signals will be drastically reduced. Because the CRES signals are
2092 in-phase at the correct beamforming position the summed power increases as a function
2093 of N^2 compared to a single antenna channel, where N is the number of antennas. It
2094 is true that the noise power is also increased by beamforming, but, because the noise
2095 is incoherent, its power only increases linearly. Consequently, the signal-to-noise ratio
2096 (SNR) of the CRES signal increases linearly with the number of antennas, which greatly
2097 improves detection efficiency compared to using only the information in a single antenna.

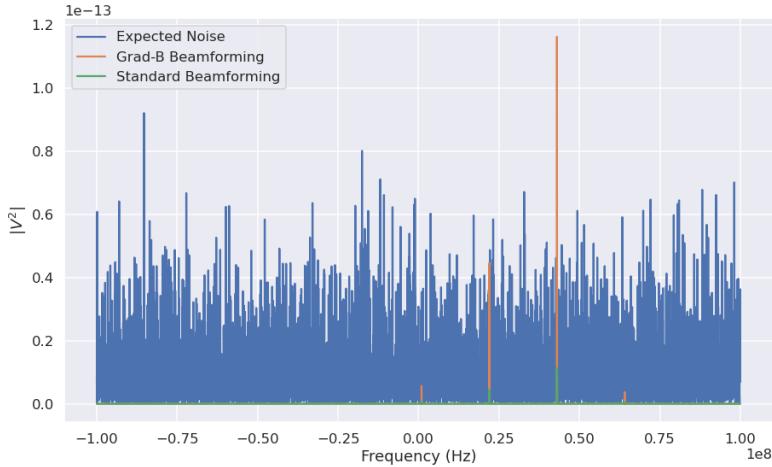


Figure 4.13: A plot of a typical frequency spectrum obtained by applying a Fourier transform to the time-series obtained from beamforming. The frequency spectra are plotted without noise on top of an example of a typical noise spectrum to visualize a realistic signal-to-noise ratio. In the example we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

2098 The power threshold detection algorithm searches for high-power frequency bins that
 2099 should correspond to a frequency component of the CRES signal. In order to prevent
 2100 random noise fluctuations from being mistaken as CRES signals the power threshold
 2101 must be set high enough so that it is unlikely that random noise could be responsible. A
 2102 consequence of this is that many electrons that can be trapped will go undetected because
 2103 the modulation caused by axial oscillations leads to the cyclotron carrier power to falling
 2104 below the decision threshold. The time-dependent beamforming used to correct for the
 2105 ∇B -drift increases the volume of the magnetic trap where electrons can be detected,
 2106 but it is ineffective at increasing the range of detectable pitch angles (see Figure 4.14).
 2107 Fundamentally, this is because the power threshold only uses a fraction of the signal
 2108 power to detect electrons and ignores the power present in the frequency sidebands. In
 2109 the subsequent sections I examine two other signal detection algorithms that seek to
 2110 improve the detection efficiency of the FSCD by utilizing the more of the signal shape to
 2111 compute the detection test statistics.

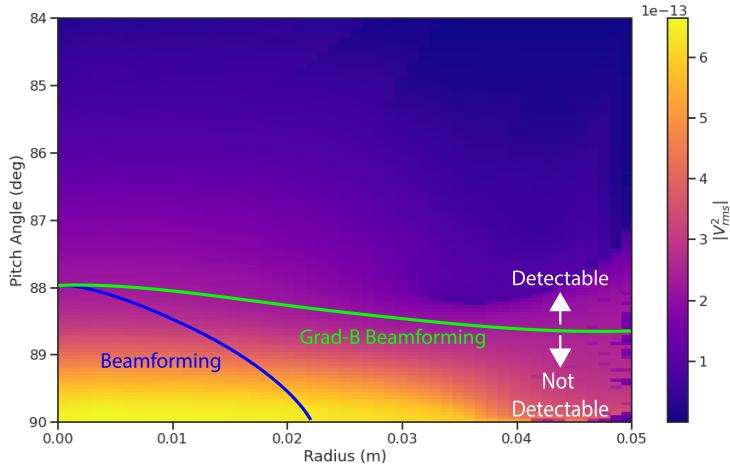


Figure 4.14: A plot of the total signal power received by the FSCD array from trapped electrons with different radial positions and pitch angles generated using Locust simulations. The lines on the plot indicate a 10 dB detection threshold above the mean value of the noise in the frequency spectrum.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

2112 4.3.2 Matched Filtering

2113 Introduction to Matched Filtering

2114 The problem of CRES signal detection is the problem of detecting a signal buried in
 2115 WGN, which has been examined at great depth in the signal processing literature [19].
 2116 For a fully known signal in WGN the optimal detector is the matched filter, which means
 2117 that it achieves the highest true positive rate for a fixed rate of false positives. The
 2118 matched filter test statistic is calculated by taking the inner product of the data with
 2119 the matched filter template

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

2120 where $h[n]$ is the matched filter template and $y[n]$ is the data. The matched filter test
 2121 statistic defines a binary hypothesis test in which the data vector is assumed to be an
 2122 instance of two possible data classes. By setting a decision threshold on the value of \mathcal{T} ,
 2123 one can classify a given data vector as belonging to two distinct hypotheses. Under the
 2124 first hypothesis the data is composed of pure WGN, and under the second hypothesis the

2125 data is composed of the known signal with additive WGN. The matched filter template
 2126 is obtained by rescaling the known signal in the following way

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

2127 where τ is the variance of the WGN and $x[n]$ is the known signal. Strictly speaking,
 2128 Equation 4.27 is only true for noise with a diagonal covariance matrix, however, in the
 2129 context of the FSCD we are justified in assuming this to be true. Defining the matched
 2130 filter templates in this way guarantees that the expectation value of \mathcal{T} is equal to one
 2131 when the data contains only noise, which is the standard matched filter normalization in
 2132 the signal processing literature.

2133 Although matched filters are canonically formulated in terms of a perfectly known
 2134 signal, it is still possible to apply the matched filter technique given imperfect information
 2135 about the signal provided that the signal is deterministic. From our discussion of CRES
 2136 simulation tools for the FSCD (see Section 4.2) we know that the shape of CRES signals
 2137 are completely determined by the initial parameters of the electron. The random collisions
 2138 with background gas molecules which cause the formation of signal tracks are the only
 2139 stochastic component of the CRES event after the initial beta-decay, therefore, it is
 2140 possible to develop a matched filter for the detection of CRES signal tracks which are fully
 2141 determined by the parameters of the electron after the initial beta-decay or subsequent
 2142 collision events.

2143 The matched filter test statistic for CRES signals is a modified version of Equation
 2144 4.26

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

2145 where the matched filter inner product has been replaced with a convolution operation
 2146 and a maximization over the template and convolution delay (m). Replacing the inner
 2147 product with a convolution accounts for the fact that the start time of the CRES signal is
 2148 now an unknown parameter, in addition, we now perform a maximization of the matched
 2149 filter convolution over a number of different templates. Because the shape of the signal is
 2150 unknown we are forced to guess a number of different signal shapes to create a template
 2151 bank with which we can identify unknown signals by performing an exhaustive search.

2152 The template bank approach to matched filtering, while quite powerful, can quickly
 2153 become computationally intractable. This is especially true in the case of the FSCD
 2154 because of the large amount of raw data produced by the array that must be analyzed.

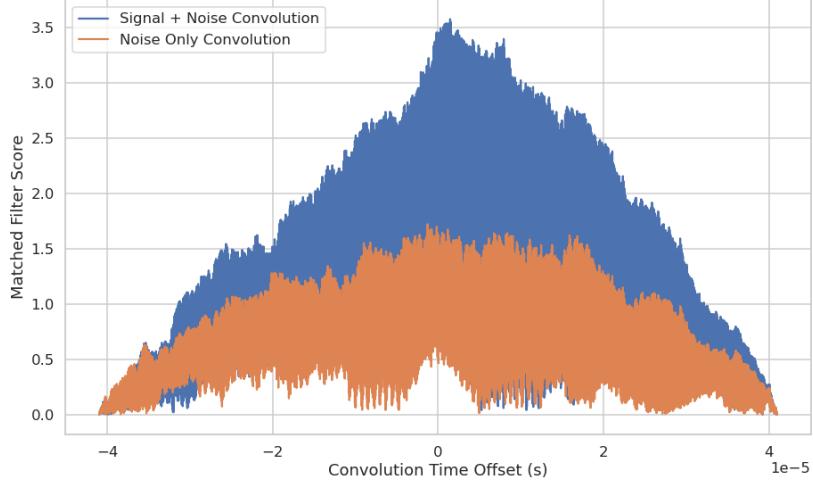


Figure 4.15: Example of a convolution of a CRES signal template with a segment of noisy data. A simulated CRES signal was simulated using Locust and normalized to create a matched filter template. When this template is convolved with noisy data the contains the matching signal the convolution output 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.

2155 Specifically, the time-domain convolution specified by Equation 4.28 is particularly
 2156 computationally intensive and is a major barrier towards the implementation of a
 2157 matched filter for signal detection in an experiment like the FSCD. This can be avoided
 2158 by using the convolution theorem to replace the time-domain convolution with an inner
 2159 product in the frequency domain.

2160 The convolution theorem states that

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

2161 where \mathbf{f} and \mathbf{g} are discretely sampled time-series, \mathbf{F} and \mathbf{G} are the respective discrete
 2162 Fourier transforms, and \mathcal{F}^{-1} is the inverse discrete Fourier transform operator. The
 2163 convolution theorem allows us to perform the matched filter convolution by first com-
 2164 puting the Fourier transform of the template and data, then performing a point-wise
 2165 multiplication of the two frequency series, and finally performing the inverse Fourier
 2166 transform to obtain the convolution output. Because discrete Fourier transforms can be
 2167 performed extremely efficiently, the convolution theorem is almost always used in lieu of
 2168 directly computing the convolution.

2169 One thing to note here is that the convolution theorem for discrete sequences shown

here, is technically valid only for circular convolutions, which is not directly specified in Equation 4.28. However, because typical CRES track lengths are much longer than the Fourier analysis window and also that the frequency chirp rates are small compared to the time-slice duration, it is relatively safe to use circular convolutions to evaluate matched filter scores for CRES signals, which allows us to apply the convolution theorem to compute matched filter scores using the frequency representation of the data and matched filter template.

Matched Filter Analysis of the FSCD

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

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

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

The matched filter approach can also be used to estimate the achievable energy resolution of the experiment by using a dense grid of templates generated with parameters close to the unknown signal (see figure 4.17). Because matched filter templates with similar parameters have signal shapes that are also similar, templates with incorrect parameters can have nearly identical matched filter scores as the correct template. Since only one sample of noise is included in a sample of real data, one cannot guarantee that the best matching template corresponds to the ground truth parameters of the signal.

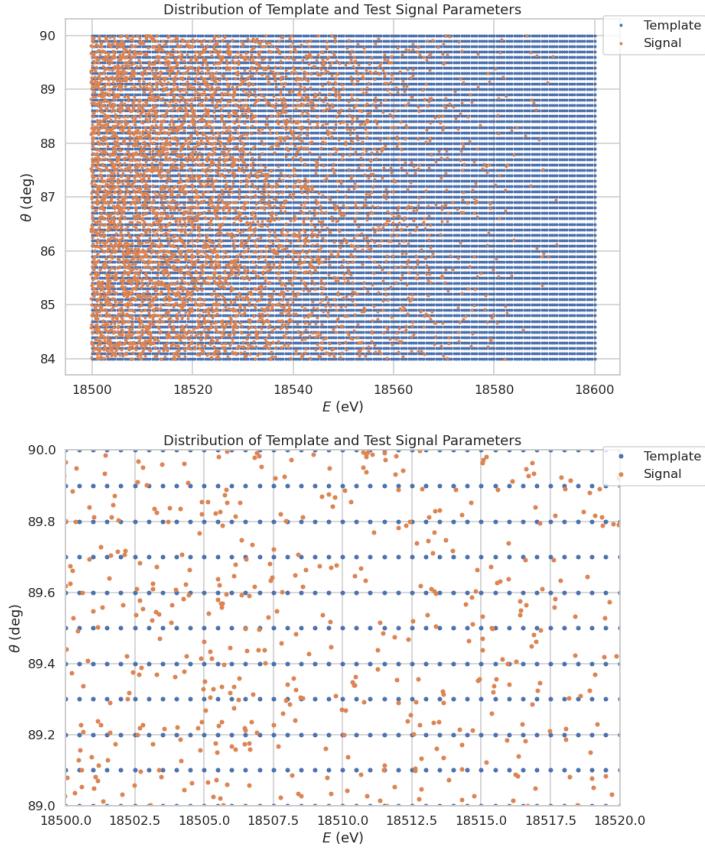


Figure 4.16: An example two-dimensional parameter distribution of a matched filter template bank and random test signals. θ refers to the pitch angle of the electron and E is the kinetic energy. The template bank forms a regular grid of in pitch angle and energy, whereas, the test signals are uniformly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

This introduces uncertainty into the signal parameter estimation that manifests as an energy broadening. Dense grids of matched filter templates allows one to quantify this 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

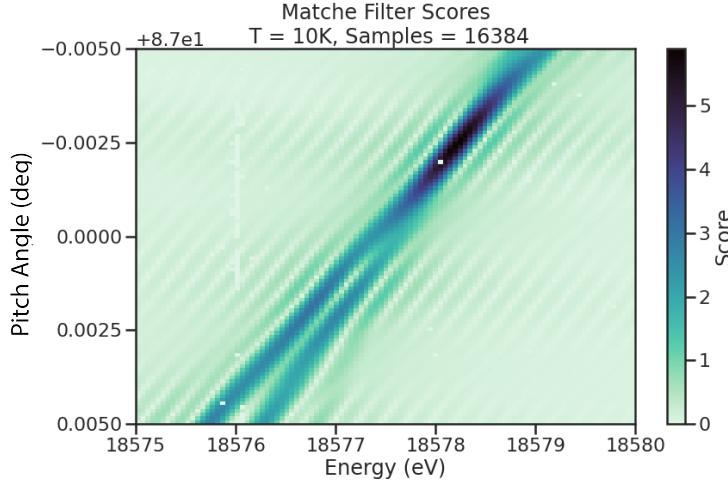


Figure 4.17: The matched filter scores of a dense grid of templates in pitch angle energy space. Dense template grids allow one to estimate the kinetic energy of the electron by identifying the best matching template. The uncertainty on this value is proportional to the space of templates that also match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

2212 filter score for a random signal to the matched filter score for a perfectly matching
2213 template. In equation form this is

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

2214 where $\mathcal{T}_{\text{best}}$ is the matched filter score of the best fitting template in the bank and $\mathcal{T}_{\text{ideal}}$ is
2215 the hypothetical matched filter score one would measure if the signal perfectly matched
2216 the template. Generally, one desires an average match as close to one as possible, however,
2217 the average match value is an exponential function of the number of templates in the
2218 template bank (see Figure 4.18). This behavior is observed for dense matched filter grids
2219 like the one in Figure 4.17. A dense grid was used to calculate the average value of match
2220 for different template bank sizes shown in Figure 4.18.

2221 The exponential relationship between match and template bank size is also evident
2222 for template banks that cover a wide range of parameters, such as the template bank
2223 visualized in Figure 4.16. Since no prior knowledge of the signal parameters is available,
2224 one has no choice but to use a template bank that covers a large range of parameters for
2225 signal detection. Achieving a high average match in this scenario can easily overwhelm
2226 the available computational resources, so in practice only a limited number of templates

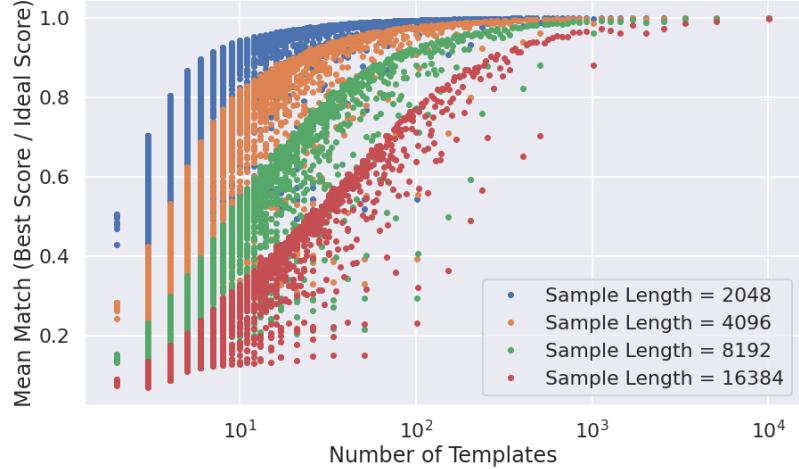


Figure 4.18: The mean match of the dense template grid shown in Figure 4.17 for different numbers of templates. Grids of different sizes were obtained by decimating a dense grid of templates and the average match for each grid was computed using the same set of randomly distributed test signals. Plotting 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.

2227 could be used at the detection stage. Therefore, accurately modeling the effects of match
 2228 is key to correct sensitivity calculations.

2229 The effect of match on the detection efficiency of the matched filter template bank can
 2230 be summarized using the ROC curve (see Figure 4.19). A single ROC curve is obtained
 2231 by averaging over the PDFs that describe the detection probabilities of each individual
 2232 template. The matched filter score for a template follows a Rician distribution with a
 2233 mean value equal to the matched filter score multiplied by the match ratio between the
 2234 template and signal. Therefore, the distribution that describes the average matched filter
 2235 score when there is a signal in the data is obtained by averaging over the distributions
 2236 for every template, whose expectation values are multiplied by the average match ratio.

2237 The distribution of the matched filter score when there is no signal in the data follows
 2238 a Rayleigh distribution. Therefore, a trials penalty, which is the statistical penalty one
 2239 pays for randomly checking many templates in order to avoid a random match between
 2240 noise and a template, is included by computing the joint distribution of N_{template} Rayleigh
 2241 distributions, where N_{template} is the size of the template bank. For more information on
 2242 the calculation of matched filter template bank ROC curves please refer to Section 4.4.

2243 An alternative way to visualize the detection performance for each algorithm is to

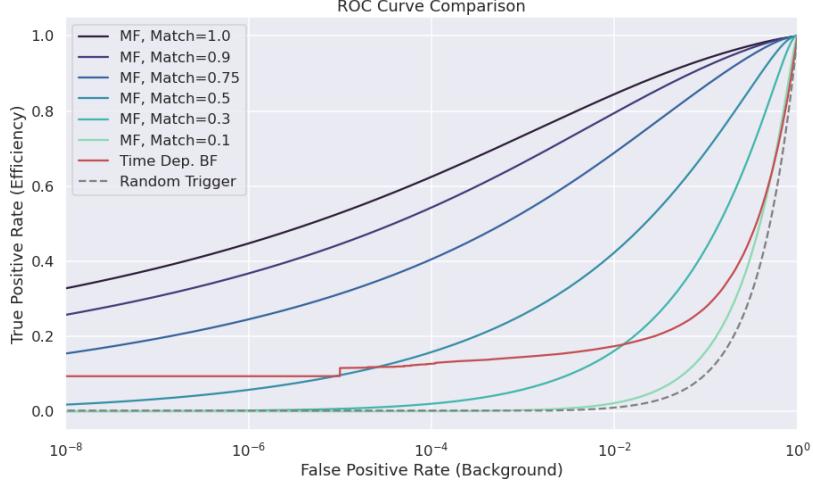


Figure 4.19: Matched filter template bank ROC curves as a function of mean match. One can see that for low match a matched filter is on average worse than the more straight forward beamforming detection approach.

2244 specify a minimum acceptable false positive rate at the trigger level. This is equivalent
 2245 to specifying a minimum threshold on the value of the matched filter score or the size of
 2246 a frequency peak for a beamforming power threshold trigger. One can then draw regions
 2247 of detectable signals as a function of the electron's pitch angle and radial position (see
 2248 Figure 4.20). A kinetic energy shift is equivalent to an overall frequency shift of the
 2249 signal and should have no effect on the detection probability assuming sufficient density
 2250 of matched filter templates in the energy dimension. A electron is declared "detectable"
 2251 for the regions in Figure 4.20 if the signal has at least 50% probability of falling above the
 2252 decision threshold of the respective classifier. One can see that the parameter space of
 2253 detectable signals is greatly expanded beyond the beamforming power threshold trigger
 2254 with a matched filter (MF) or deep neural network (DNN) (see Section 4.3.3). Plots such
 2255 as Figure 4.20 are useful for visualization, but, since the handling of detection likelihood
 2256 is not sufficiently rigorous, the detection probability boundaries are not particularly well-suited to sensitivity estimates.
 2257

2258 Optimized Matched Filtering Implementation for the FSCD

2259 The biggest practical obstacle to the implementation of a matched filter template bank
 2260 detection approach is oftentimes the computational cost associated with exhaustively
 2261 calculating the matched filter scores of the template bank, and the FSCD is no exception
 2262 in this regard. At a basic level computing a matched filter score requires the convolution

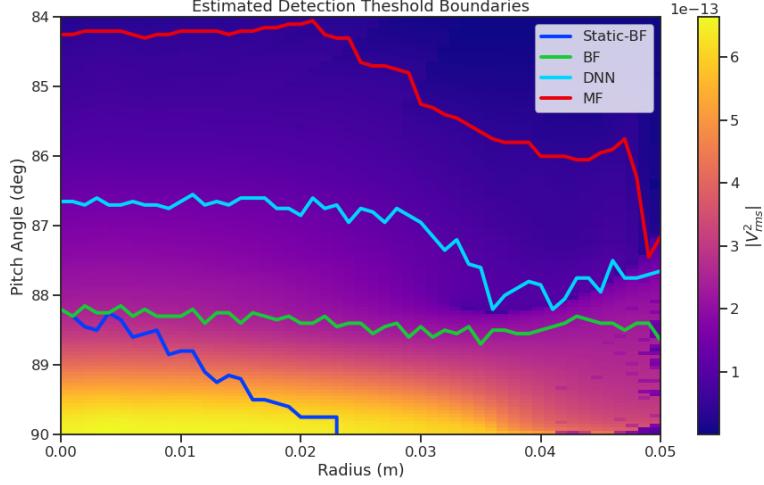


Figure 4.20: Boundaries of detectable electrons in pitch angle kinetic energy space for a series of different signal detection algorithms. A detectable signal is defined as a signal that is above a consistent decision with at least 50% probability. This non-rigorous treatment of detection probability 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 ∇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_{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

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

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

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

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

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

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

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

2307 to detect a signal in a segment of array data is on the order of

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

2308 With the hybrid approach we attempt to remove the spatial parameters from the
2309 template bank by using beamforming to combine the array signals into a single channel.
2310 Beamforming explicitly assumes a starting position, which allows us to only use matched
2311 filter templates that span the two-dimensional space of kinetic energy and pitch angle.
2312 The total computational cost of the hybrid method is directly proportional to the number
2313 of beamforming positions. For the time-dependent beamforming defined in Section 4.3.1,
2314 the number of beamforming positions is given by

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

2315 where N_r and N_φ are the same spatial parameters encountered in the pure matched
2316 filter template bank and $N_{\omega_{\nabla B}}$ is the number of ∇B -drift frequency assumptions. If a
2317 unique drift frequency is used for each pitch angle then the hybrid approach is effectively
2318 equivalent to a pure matched filter in the number of operations. The key efficiency gain
2319 of the hybrid approach is to exploit the relatively small differences in $\omega_{\nabla B}$ for electrons
2320 of different pitch angles by using only a small number of average drift frequencies.

2321 The total number of operations for the hybrid approach can be expressed as a sum of
2322 the operations required by the beamforming and matched filtering steps,

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

2323 The first product in the sum is the number of operations required by beamforming,
2324 which is simply the number of beamforming points times the computational cost of the
2325 beamforming matrix multiplication, and the second product is the computational cost
2326 of matched filtering the summed signal generated by each beamforming position. To
2327 compare this to pure matched filtering we take the ratio of Equations 4.34 and 4.36 to
2328 obtain

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

2329 This expression can be simplified by observing that $O(N_E N_\theta) \times O(\log N_s) \gg O(N_{ch})$,

2330 which means that the ratio of computational cost for the two methods can be reduced to

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

2331 If we limit ourselves to a number of estimated drift frequencies of $O(1)$ then we see that
2332 the estimated computational cost reduction of the hybrid approach is of $O(N_{\text{ch}})$. This is
2333 quite a large reduction considering that the FSCD antenna array contains sixty antennas
2334 in the baseline design.

2335 The main drawback of the hybrid approach is that the limited number of allowed
2336 drift frequency guesses can lead to detection efficiency loss due to phase mismatch. The
2337 degree of phase error from an incorrect drift frequency is proportional to the length of
2338 the array data vector used by the signal detection algorithm. For signals with lengths
2339 equal to the baseline FSCD Fourier analysis window of 8192 samples, typical phase errors
2340 from using an average versus the exact ∇B -drift frequency are on the order of a few
2341 percent in terms of the signal energy. This has a relatively small impact on the overall
2342 detection efficiency, however, future experiments with antenna array CRES will want to
2343 balance optimizations such as these during the design phase to keep experiment costs to
2344 a minimum while still achieving scientific goals.

2345 Kinetic Energy and Pitch Angle Degeneracy

2346 More accurate modeling of a matched filter requires that we consider the effects of
2347 mismatched signals and template, since this more accurately reflects the real-world usage
2348 of a matched filter where many incorrect templates are convolved with the data until the
2349 matching template is found. One way to study this is to use the grid of simulated signals
2350 to compute the matched filter scores between mismatched signals and templates and
2351 evaluate the matched filter scores under this scenario. What one finds when performing
2352 this analysis is that templates for kinetic energies and pitch angles that do not match
2353 the underlying signal can have matched filter scores that are indistinguishable from the
2354 matched filter score of the correct template (see Figure 4.21 and Figure 4.21).

2355 This degeneracy in matched filter score is the result of correlations between the kinetic
2356 energy of the electron and the pitch angle caused by changes in the average magnetic field
2357 experienced by an electron for different pitch angles. While in principle helpful for the
2358 purposes of signal detection these correlations are unacceptable since they greatly reduce
2359 the energy resolution of the experiment by causing electrons with specific kinetic energy
2360 to templates across a wide range of energies. It is important to emphasize that this

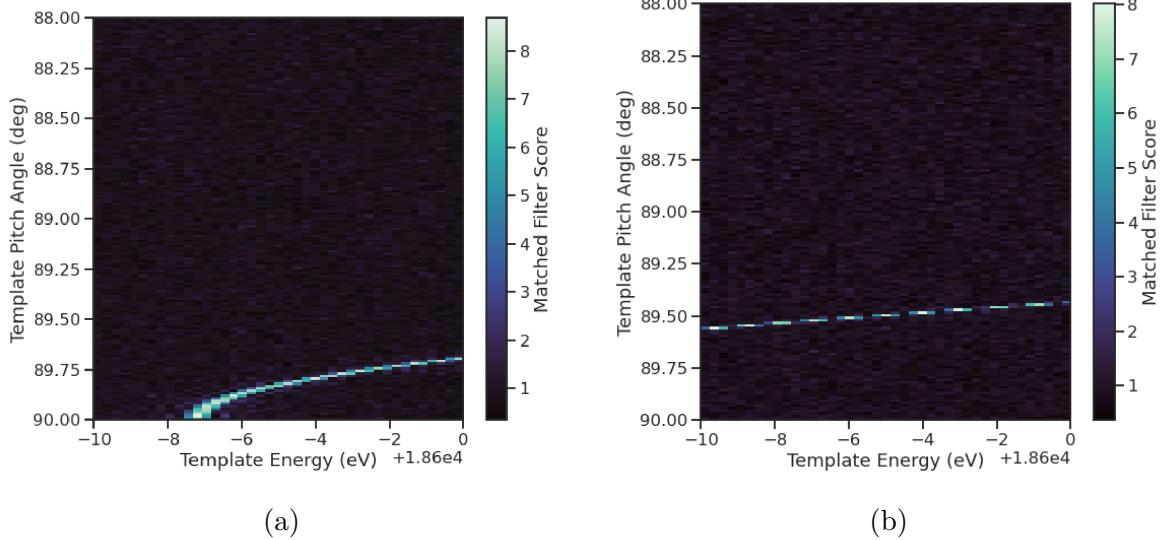


Figure 4.21: Two example illustrations of the correlation between kinetic energy and pitch angle imparted by the shape of the FSCD magnetic trap. The correlations manifest themselves as degeneracies in the matched filter score where multiple matched filter templates have the same matched filter for a 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.

2361 degeneracy cannot be fixed by implementing a different signal reconstruction algorithm.
 2362 As revealed by the matched filter scores the shapes of the signals for different parameters
 2363 are identical. Resolving this degeneracy between pitch angle and energy requires the
 2364 design of a new magnetic trap with steeper walls so that the average magnetic field
 2365 experienced by an electron is less dependent on pitch angle.

2366 4.3.3 Machine Learning

2367 Machine learning is a vast and rapidly developing field of research [22]. In this Section
 2368 we shall attempt to provided a brief introduction to some of the concepts and techniques
 2369 of machine learning that were applied to CRES signal detection rather than attempt a
 2370 comprehensive overview.

2371 Introduction to Machine Learning

2372 Digitization of the FSCD antenna array generates large amounts of data that must be
 2373 rapidly processed to enable real-time signal detection and reconstruction. While digital
 2374 beamforming combined with a power threshold is relatively computationally inexpensive,

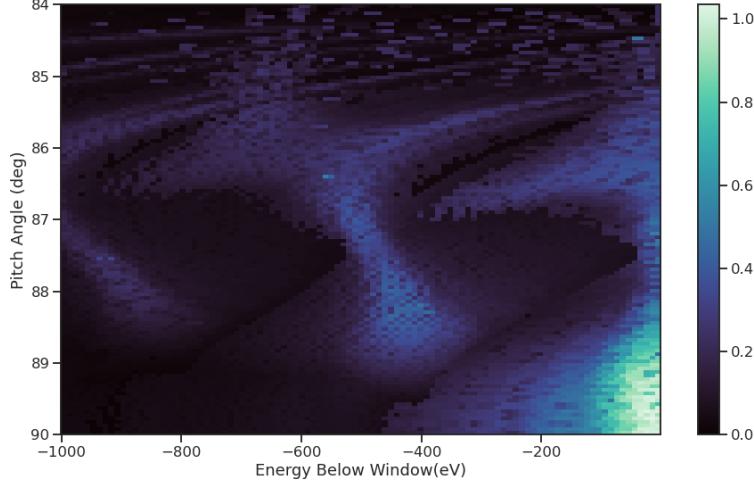


Figure 4.22: A visualization of the correlation between energy and pitch angle in the FSCD magnetic trap. The image is formed by computing the match of the best template from a grid consisting of pitch angles from 84 to 90 degrees in steps of 0.05 degrees, kinetic energies from 17574 to 18574 eV, located 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.

it is relatively ineffective at detecting CRES signal with small pitch angles, since it relies on a visible frequency peak above the noise. On the other hand, a matched filter is able to detect signals with a significantly larger range of parameters, however, the exhaustive search of matched filter templates can be computationally expensive. Machine learning based triggering algorithms have been used successfully in many different high-energy physics experiments [23] and recent developments have shown success in the detection of gravitational wave signals using machine learning techniques [24, 25] in place of the more traditional matched filtering method. This motivates the exploration of machine learning as a potential CRES signal detection algorithm.

There are several different approaches to machine learning, but the one most important to our discussion here is the supervised learning approach. In supervised machine learning one uses a differentiable model or function that is designed to map the input data to the appropriate label [22]. The data is represented as a multidimensional matrix of floating point values such as an image or a time-series, and the label is generally a class name such as signal or noise for classification problems or a continuous value like kinetic energy

2390 in the case of regression problems.

2391 In supervised learning the model is trained to map from the data to the correct label
2392 by evaluating the output of the model using a training dataset consisting of a set of
2393 paired data and labels. To evaluate the difference between the model output and the
2394 correct label a loss function is used to quantify the error between the model prediction
2395 and the ground truth. For example, a common loss function in regression problems is the
2396 squared error loss function, which quantifies error using the squared difference between
2397 the model output and label.

2398 Using the outputs of the loss function the next step in supervised learning is to
2399 compute the gradient of error with respect to the model parameters in a process called
2400 backpropagation. Using the model parameter gradients the last step in the supervised
2401 learning process is to perform an update of the parameter values in order to minimize
2402 the error in the model predictions across the whole dataset. This loop is performed many
2403 times while randomly shuffling the dataset until the error converges to a minimum value
2404 at which point the training procedure has finished. It is standard practice to monitor
2405 the training procedure by evaluating the performance of the model using a separate
2406 validation dataset that matches the statistical distribution of the training data and to
2407 check the performance of the model after training using yet another dataset called the
2408 test dataset. These practices help to guard against overtraining which is a concern for
2409 models with many parameters.

2410 **Convolutional Neural Networks**

2411 A popular class of machine learning models are neural networks. A neural network is
2412 essentially a function composed of a series of linear operations called layers which take a
2413 piece of data typically represented as a matrix, multiplies the elements of the data by a
2414 weight, and then sums these products to produce an output matrix. Neural networks
2415 composed of purely linear operations are unable to model complex non-linear behavior,
2416 therefore, non-linear activation functions are applied to the outputs of each of the layers
2417 to increase the ability of the neural network to model complex relationships between the
2418 data.

2419 Neural networks are typically composed of at least three layers, but with the present
2420 capabilities of computer hardware they more often contain many more than this. The
2421 first layer in a neural network is called the input layer, because it takes the data objects
2422 as input, and the last layer in a neural network is known as the output layer. The
2423 output layer is trained by machine learning to map the data to a desired output using

2424 the supervised learning procedure described in Section 4.3.3. In between the input and
2425 the output layer are typically several hidden layers that receive inputs from and transmit
2426 outputs to other layers in the neural network model. The term deep neural network
2427 (DNN) refers to those neural networks that have at least one hidden layer, which have
2428 proven to be extremely powerful tools for pattern recognition and function approximation.

2429 An important type of DNN are convolutional neural networks (CNN) that typically
2430 contain several layers which perform a convolution of the input with a set of filters. These
2431 convolution operations are typically accompanied by layers that attempt to down-sample
2432 the data along with the standard neural network activation functions. A standard CNN
2433 is composed of several convolutional layers at the beginning of the network and ends
2434 with a series of fully-connected neural network layers at the output. Intuitively, one
2435 can imagine that the convolutional layers are extracting features from the data that
2436 fully-connected layers use to perform the classification or regression task.

2437 **Deep Filtering for Signal Detection in the FSCD**

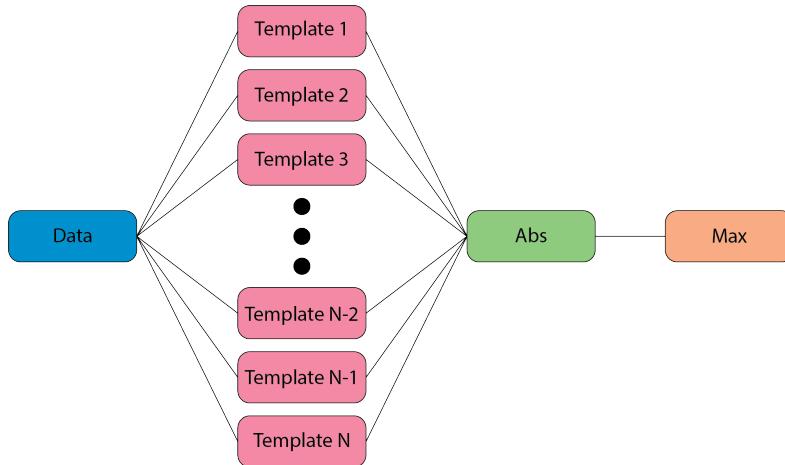


Figure 4.23: A representation of a matched filter template bank as a convolutional neural network. The network has a single layer composed of the templates, which act as convolutional filters. The activation of the neural network is an absolute value followed by a max operator.

2438 CNNs have been extremely influential in the field of computer vision, particularly tasks
2439 such as image segmentation and classification, but have also been applied in numerous
2440 experimental physics contexts. Given the particular challenge posed by signal detection
2441 and reconstruction in the FSCD we are interested in exploring the potential of machine
2442 learning as an effective algorithm for real-time signal detection, since this application

2443 requires both high efficiency and fast evaluation.

2444 In the machine learning paradigm signal detection is equivalent to a binary classifi-
2445 cation problem between the signal and noise data classes, and my investigation focuses
2446 specifically on the application of CNNs to signal detection in the FSCD, which is moti-
2447 vated by relatively recent demonstrations of CNNs achieving classification accuracies for
2448 gravitational wave time-series signals comparable to a matched filter template bank. In
2449 this framework it is possible to interpret the matched filter as a type of CNN composed
2450 of a single convolutional layer with the templates making up the layer filters (see Figure
2451 4.23). Since this neural network has no hidden layers, it is not a DNN like we have
2452 been discussing so far, but we can attempt to construct a proper CNN that attempts to
2453 reproduce the classification performance of the matched filter network.

2454 The name deep filtering refers to this scheme of replacing a matched filter template
2455 bank with a DNN. The reason why one might want to do this is that it may be possible to
2456 exploit redundancies and correlations between templates that may allow one to perform
2457 signal detection with similar accuracy but with fewer computations, which is important
2458 for real-time detection scenarios like the FSCD experiment. In Section 4.4 we perform a
2459 detailed comparison of the signal detection performance of a CNN to beamforming and a
2460 matched filter template bank.

2461 Deep filtering is conceptually a simple technique. Similar to a matched filter template
2462 bank a large number of simulated CRES signals are generated and used to train a model
2463 to distinguish between signal and noise data (see Figure 4.24). In order to reduce the
2464 dimensionality of the input FSCD data a digital beamforming summation is applied
2465 to the raw time-series data generated by Locust to compress the 60-channel data to a
2466 single time-series. CRES signal have a sparse frequency representation and experiments
2467 training CNN's on time-series and frequency series data found that models trained on
2468 frequency spectrum data performed significantly better, therefore, an FFT is applied to
2469 the summed time-series before being normalized and fed to the classification model.

2470 The data used to train the model consists of an equal proportion of signal and noise
2471 frequency spectra. Unique samples of WGN are generated and added to the signals during
2472 training time to avoid have to pre-generate and store large samples of noise data. The
2473 binary cross-entropy loss function combined with the ADAM optimizer proved effective
2474 at training the models to classify CRES data. A simple hyperparameter optimization
2475 was performed by manually tuning model, loss function, and optimizer parameters. The
2476 model and training loops was implemented in python using the PyTorch deep learning
2477 framework. Standard machine learning best practices were followed when training the

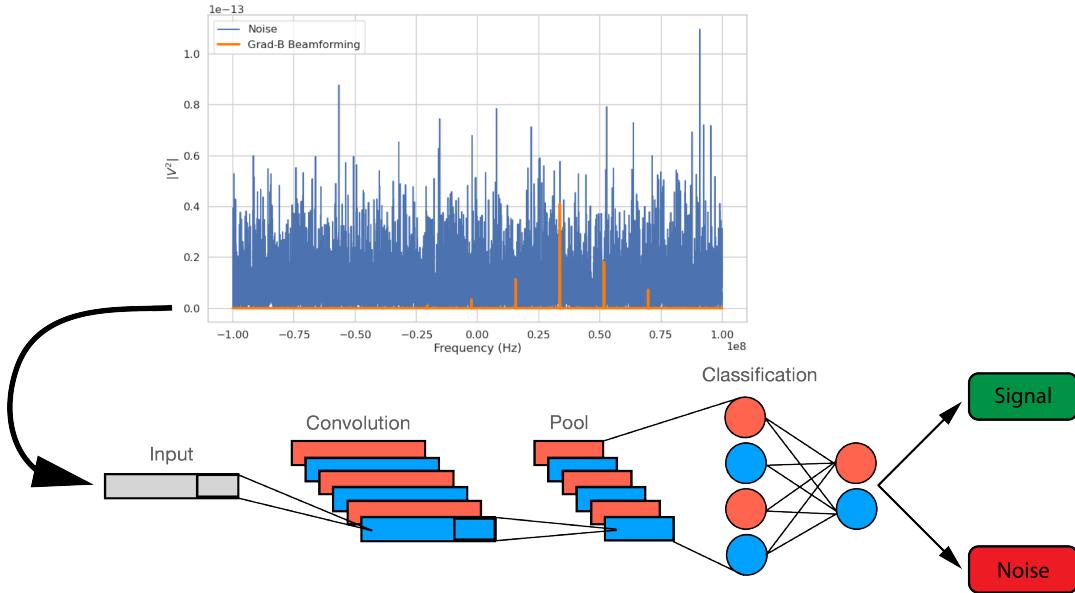


Figure 4.24: A graphical depiction of CRES signal detection using a CNN. A noisy segment of data is converted to a frequency series using digital beamforming and a FFT. The complex-valued frequency series is input into a trained CNN model that classifies the data as signal or noise using a decision threshold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

models, such as overtraining monitoring using a validation dataset. Models were trained until the training loss and accuracy converged and then evaluated using a separate test data set.

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

4.4 Analysis of Signal Detection Algorithms for the Antenna Array Demonstrator

This section contains an early version of the manuscript for the triggering paper prepared for publication in JINST. In it I present a relatively detailed analysis of the signal detection performance of the three signal detection approaches discussed so far using a

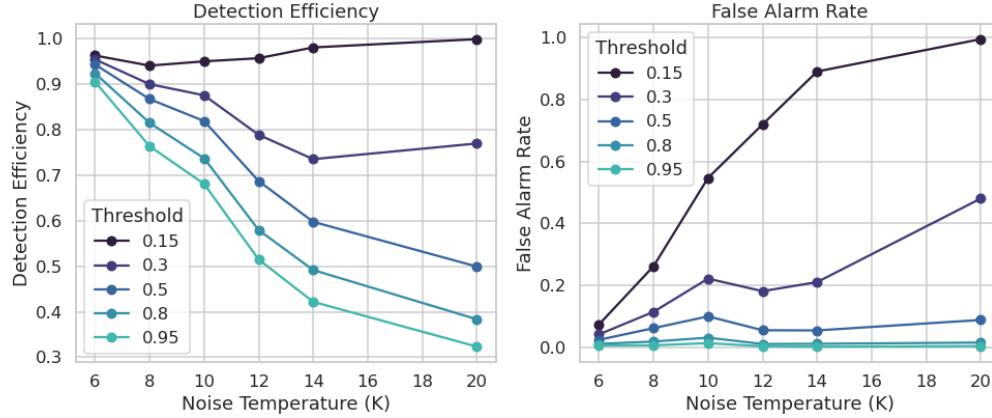


Figure 4.25: The detection efficiency and false alarm rate (false positive rate) as a function of the decision threshold for different values of the noise temperature. The model is trained to output a value close to one for data that contains a signal and outputs a value near zero when the data contains 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.

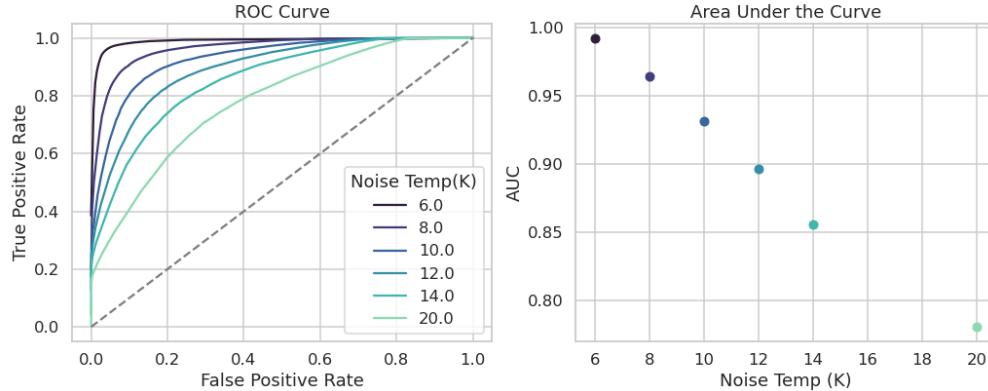


Figure 4.26: ROC curves for a CNN model classifying CRES signals. One can see that the area under the curve, which is a figure of merit that describes the performance of the classifier, is roughly linearly dependent with the noise temperature.

2493 population of simulated CRES signals generated with Locust. The focus of the paper is
 2494 on the performance of the signal detection algorithms for pitch angles below 88.5° where
 2495 the beamforming power threshold begins to fail.

2496 4.4.1 Introduction

2497 Cyclotron Radiation Emission Spectroscopy (CRES) is a technique for measuring the
 2498 kinetic energies of charged particles by observing the frequency of the cyclotron radiation

2499 that is emitted as they travel through a magnetic field [2]. The Project 8 Collaboration
2500 is developing the CRES technique as a next-generation approach to tritium beta-decay
2501 endpoint spectroscopy for neutrino mass measurement. Recently, Project 8 has used
2502 CRES to perform the first ever tritium beta-decay energy spectrum and neutrino mass
2503 measurement [4, 5].

2504 Previous CRES measurements have utilized relatively small volumes of gas that are
2505 directly integrated with a waveguide transmission line, which transmits the cyclotron
2506 radiation emitted by the trapped electrons to a cryogenic amplifier. While this technology
2507 has had demonstrable success, it is not a feasible option for scaling up to significantly
2508 larger measurement volumes. In particular, the goal of the Project 8 Collaboration
2509 is to use CRES combined with atomic tritium to measure the neutrino mass with a
2510 40 meV sensitivity. Achieving this sensitivity goal will require a multi-cubic-meter scale
2511 measurement volume in order to obtain the required event statistics in the tritium
2512 beta-spectrum endpoint region; hence, there is a need for new techniques to enable large
2513 volume CRES measurements for future experiments.

2514 One approach is to surround a large volume with an array of antennas that together
2515 collect the cyclotron radiation emitted by trapped electrons [3, 26]. A promising array de-
2516 sign is an inward-facing uniform cylindrical array that surrounds the tritium containment
2517 volume. Increasing the size of the antenna array, by adding additional rings of antennas
2518 along vertical axis, allows one to grow the experimental volume until a sufficient amount
2519 of tritium gas can be observed by the array. A challenging aspect of this approach is
2520 that the total radiated power emitted by an electron near the tritium spectrum endpoint
2521 is on the order of 1 fW or less, which is then distributed between all the antennas in
2522 the array. Consequently, detecting the presence of a CRES signal and determining the
2523 electron's kinetic energy requires reconstructing the entire antenna array output over the
2524 course of the CRES event, posing a significant data acquisition and signal reconstruction
2525 challenge.

2526 Project 8 has developed a triggering system to enable real-time identification of CRES
2527 events using an antenna array [27]. Previous measurements with the CRES technique
2528 have utilized a threshold on the frequency spectrum formed from a segment of CRES
2529 time-series data. This algorithm relies on the detection of a frequency peak above the
2530 thermal noise background, which limits the kinematic parameter space of detectable
2531 electrons. Due to the limitations of this power threshold, Project 8 has been investigating
2532 alternative signal identification approaches, including both matched filtering and machine
2533 learning based classifiers, to improve the detection efficiency of the experiment. In

order to evaluate the relative gains in detection efficiency that come from utilizing these alternative algorithms, we develop analytical models for the power threshold and matched filter signal classifier performance applicable to an antenna array based CRES detector. In addition, we implement and test a basic convolutional neural network (CNN) as a first step towards the development of neural-network based classifiers for CRES measurements. These results allow us to compare the estimated detection efficiencies of each of these methods, which we weigh against the associated computational costs for real-time applications.

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

4.4.2 Signal Detection with Antenna Array CRES

4.4.2.1 Antenna Array and DAQ System

In order to explore the potential of antenna array CRES for neutrino mass measurement, the Project 8 Collaboration has developed a conceptual design for a prototype antenna array CRES experiment [3, 26], called the Free-space CRES Demonstrator or FSCD, which could be used as a demonstration of the antenna array measurement technique (see Figure 4.27). The FSCD utilizes a single ring of antennas, which is the simplest form of a uniform cylindrical array configuration, to surround a radio-frequency (RF) transparent tritium gas vessel. A prototype version of this antenna array has been built and tested by the Project 8 collaboration to validate simulations of the array radiation pattern and beamforming algorithms [6]. In the FSCD the antenna array is positioned at the center of the magnetic trap formed by a set of electro-magnetic coils that are designed to produce a magnetic trap with flat central region and steep walls both radially and axially.

When a beta-decay electron is trapped its motion consists of three primary components. The component with the highest frequency is the cyclotron orbit whose frequency is

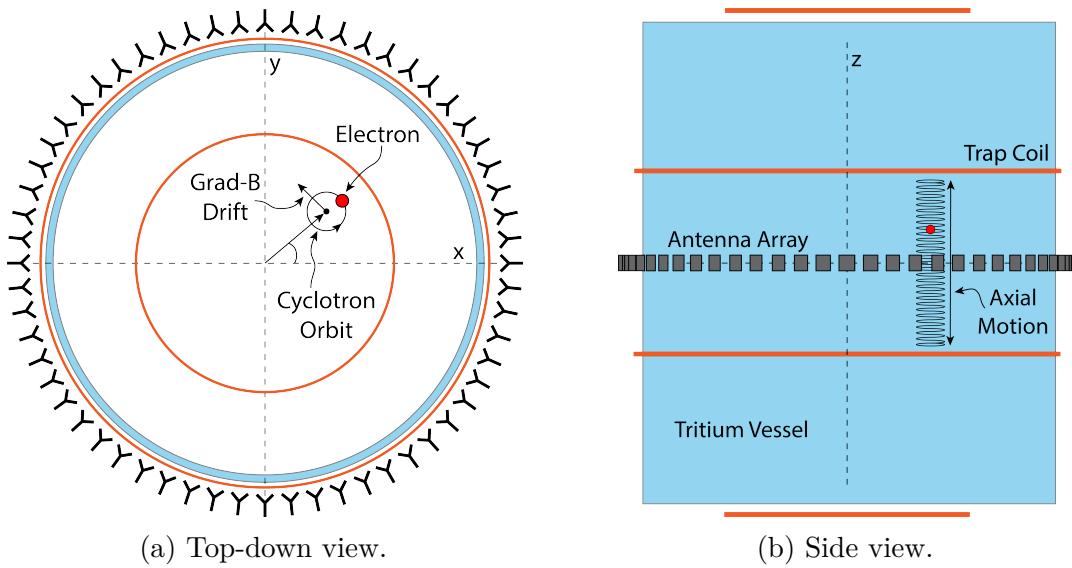


Figure 4.27: An illustration of the conceptual design for an antenna array CRES tritium beta-decay spectrum measurement. The antenna array geometry consists of a 20 cm interior diameter with 60 independent antenna channels arranged evenly around the circumference. The nominal antenna design is sensitive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

2567 determined by the size of the background magnetic field. The FSCD design assumes
 2568 a background magnetic field value of approximately 0.96 T, which results in cyclotron
 2569 frequencies for electrons with kinetic energies near the tritium beta-spectrum endpoint
 2570 from 25 to 26 GHz. The component with the next highest frequency is the axial oscillation
 2571 experienced by electrons with pitch angles of less than 90° [9]. The flat region of the
 2572 FSCD magnetic trap extends approximately 3 cm above and below the antenna array
 2573 plane causing electrons to move back and forth as they are reflected from the trap walls.
 2574 Typical oscillation frequencies are on the order of 10's of MHz, which results in an
 2575 oscillation period that is $O(10^3)$ smaller than the duration of a typical CRES event.
 2576 Therefore, when reconstructing CRES events we treat the electron as occupying only an
 2577 average axial position at the center of the magnetic trap, since we are not able to resolve
 2578 the axial position as a function of time. The component of motion with the smallest
 2579 frequency is ∇B -drift caused by radial field gradients in the trap, producing an orbit of
 2580 the electron around the central axis of the trap with a frequency on the order of a few

2581 kHz, dependent on the pitch angle and the radial position of the electron.

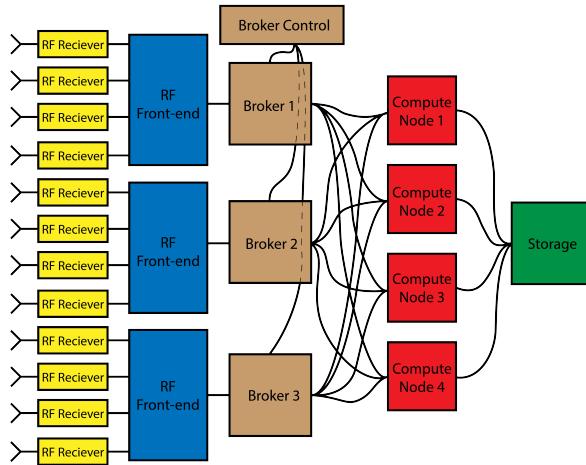


Figure 4.28: A high-level diagram of the DAQ system architecture envisioned for the FSCD.

The data acquisition (DAQ) system digitizes the signals from the antenna array and combines three data streams into a time-ordered matrix of array snapshots that can be used by the reconstruction algorithms. The FSCD DAQ system design [27] is divided into three layers 4.28. The first layer is the RF front-end, which includes the antenna array, the RF receiver boards, and the digitization electronics. The receiver board contains an amplifier, RF mixer, and bandpass filter to enable down-conversion, and is followed by the digitization electronics that sample the CRES signals at 200 MHz. In order to achieve an adequate signal-to-noise ratio to detect CRES events, the DAQ system for the antenna array demonstrator must have a total system noise temperature of ≈ 10 K, which we can achieve by using low-noise amplifiers and operating at cryogenic temperatures. After digitization, the array data must be reorganized from individual data streams sorted by channel into array snapshots sorted by time. In order to solve this data transfer and networking problem the second layer of the DAQ system consists of a set of broker computer nodes that reorganize the array data into time-ordered chunks. This approach allows us to accommodate different data transfer requirements by scaling the number of broker nodes in this layer accordingly. Next, the broker layer distributes these chunks of array data to the final layer of the DAQ system, which consists of a set of identical reconstruction nodes that perform the calculations required for CRES reconstruction. Similar to the broker layer, the number of reconstruction nodes can be increased or decreased depending on the amount of computer power required for real-time CRES reconstruction.

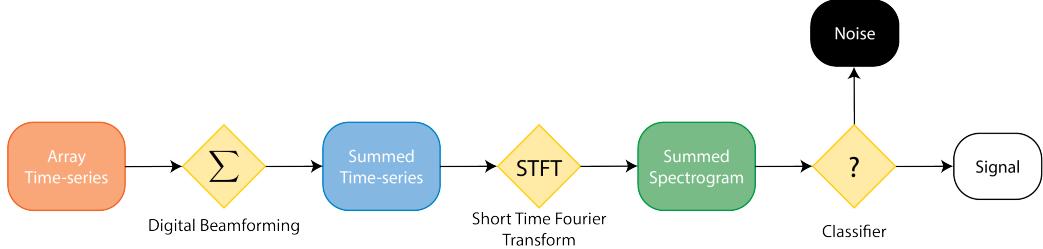


Figure 4.29: A block diagram illustration of the real-time triggering algorithm proposed for antenna array CRES reconstruction.

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

4.4.2.2 Real-time Signal Detection

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

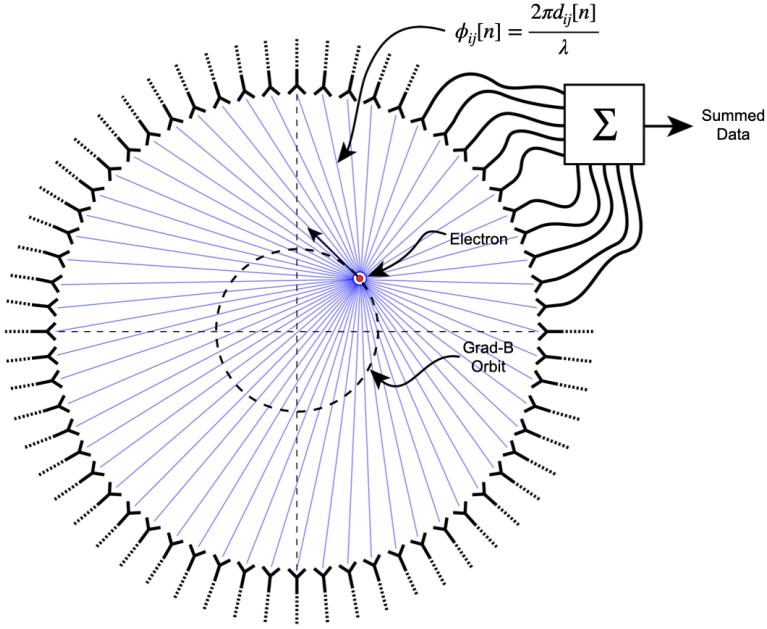


Figure 4.30: An illustration of the digital beamforming procedure. The blue lines indicate the various distances from the beamforming position to the antenna. In the situation depicted the actual position of the electron matches the beamforming position, so we should expect constructive interference when the phase shifted signals are summed. To prevent the electron's ∇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.

2627 The equation defining digital beamforming can be expressed as

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

2628 where $\mathbf{x}[n]$ is array snapshot vector at the sampled time n , $\Phi[n]$ is the matrix of
2629 beamforming phase shifts, and $\mathbf{y}[n]$ is summed output vector that contains the voltages
2630 for each of the summed channels that correspond to a particular beamforming position.
2631 The elements of the beamforming phase shift matrix can be expressed as a weighted
2632 complex exponential

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

2633 where the indices i and j label the beamforming and antenna positions respectively. The
2634 weight A_{ij} accounts for the relative power increase for antennas that are closer to the
2635 position of the electron, and ϕ_{ij} is the total beamforming phase shift for the j -th antenna
2636 at the i -th beamforming position.

2637 The beamforming phase shift is a sum of two terms

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

2638 where the first term is the phase shift originating from the path length difference ($d_{ij}[n]$)
2639 between the beamforming and antenna positions, which are represented by the vectors
2640 (r_j, θ_j) and ($r_i, \theta_i[n]$), and the second term is the angular separation ($\theta_{ij}[n]$) of the two
2641 positions. The angular separation enters into the beamforming phase due to an effect
2642 caused by the circular orbit of the electron that produces radiation whose phase is linearly
2643 dependent on the relative azimuthal position of the antenna [28,29]. The time-dependence
2644 of the beamforming phases is intended to correct for the effects of ∇B -drift, which cause
2645 the guiding centers of electrons to orbit the center of the magnetic trap. By including a
2646 linear time-dependence in the azimuthal beamforming position,

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

2647 where $\omega_{\nabla B}$ is the azimuthal grad-B drift frequency, $t[n]$ is the time vector and, $\theta_{i,0}$ is the
2648 starting azimuthal position, we can configure the beamforming phases to effectively track
2649 the XY-position of the guiding center over the event duration. Predicting accurate values
2650 of $\omega_{\nabla B}$ for a specific trap and set of kinematic parameters will be done by simulations,
2651 which are performed using the Kassiopeia software package [7] by Project 8.

2652 After digital beamforming, we apply a short-time Fourier transform (STFT) to the
2653 summed time-series to obtain the frequency spectrum representation of the signals (see
2654 Figure 4.31). From the detection perspective, the frequency representation of the CRES
2655 data is advantageous compared to the time domain, because the frequency spectra of
2656 CRES signals are well-approximated by a frequency and amplitude modulated sinusoidal
2657 whose carrier frequency increases as a linear chirp. The modulation is caused by the axial
2658 oscillation of the electron in the magnetic trap and produce frequency spectra that are
2659 well-described by a small number of frequency components. The linear chirp is caused
2660 by the energy loss due to cyclotron radiation, which results in a relatively slow increase
2661 in the frequency components of the CRES signal over time. During the standard Fourier
2662 analysis window for the FSCD of 40.96 μ sec, we expect a typical CRES signal to increase
2663 in frequency by approximately 15 kHz, which is smaller than the frequency bin width
2664 given the 200 MHz sample rate. Therefore when considering a single frequency spectrum
2665 it is justifiable to neglect the effects of the linear frequency chirp.

2666 In the cases where the electron's pitch angle is $\gtrsim 88.5^\circ$, the majority of the signal

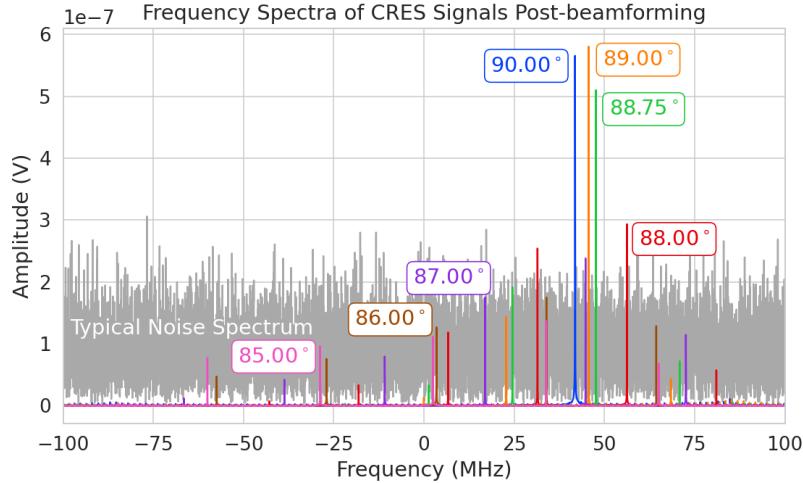


Figure 4.31: Frequency spectra of simulated CRES signals post-beamforming. The signal of a 90° electron consists of a single frequency component that is easy to detect with a power threshold on the frequency spectrum. This power threshold is still effective for signals with relatively large pitch angles 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.

power is contained in a single frequency component, with the remaining signal power contained in a small number of sidebands proportional to the electron's axial modulation (see Figure 4.31). In these cases detection is relatively straight-forward by implementing a power threshold on the STFT, since the amplitude of the main signal peak is distinct from the thermal noise spectrum. However, as the pitch angle of the electron is decreased below 88.5° , the modulation index of the signal increases causing the maximum amplitude of the frequency spectrum to be comparable to typical noise fluctuations. At this point, the power threshold trigger is no longer able to distinguish between signal and noise leading to a reduction in detection efficiency. The neutrino mass sensitivity of the FSCD is directly linked to the overall detection efficiency. And, because the distribution of electron pitch angles is effectively uniformly distributed across the range of pitch angles that can be trapped, the overall detection efficiency is directly influenced by the range of pitch angles that have detectable signals. Therefore, utilizing a signal detection algorithm that can more effectively identify signals with pitch angles less than 88.5° will improve both detection efficiency and ultimately the neutrino mass sensitivity of the FSCD and other CRES experiments.

Modeling the detection performance of alternative signal detection algorithms for

2684 the FSCD requires that we pose the signal detection problem in a consistent manner.
2685 The approach we take is to perform a binary hypothesis test on the frequency spectra
2686 generated by the STFT. Mathematically, this is expressed as,

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

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

2687 Where under hypothesis \mathcal{H}_0 , the vector representing the frequency spectrum ($y[n]$) is
2688 composed of pure white Gaussian noise (WGN) represented by $\nu[n]$, and under hypothesis
2689 \mathcal{H}_1 the frequency spectrum is composed of a CRES signal ($x[n]$) with added WGN. The
2690 dominant source of noise in a FSCD-like experiment is expected to be thermal Nyquist-
2691 Johnson noise, which is well approximated by a WGN distribution. In order to decide
2692 between these two hypotheses we follow the standard Neyman-Pearson approach by
2693 performing a log-likelihood ratio test between the probability distributions of the signal
2694 classifier output under \mathcal{H}_1 and \mathcal{H}_0 [19]. The output of the log-likelihood ratio test is
2695 called the test statistic, which is used to assign the data as belonging to the noise (\mathcal{H}_0)
2696 or signal (\mathcal{H}_1) classes by setting a decision threshold on the value of the test statistic.

2697 In practice, we select the decision threshold by finding the value of the test statistic
2698 that guarantees an acceptable rate of false positives and then attempt to maximize
2699 the signal detection probability under that fixed false positive rate. Because the signal
2700 classifier will be used to evaluate the spectra of $O(10^2)$ beamforming positions every
2701 $40.96 \mu\text{sec}$, we will require the signal classifiers to operate with decision thresholds that
2702 provide false positive rates significantly smaller than 1%. This reduces the burden placed
2703 on later stages of the CRES reconstruction chain to reject these false positives and
2704 decreases the overall likelihood of reconstructing a false event. Below, we calculate the
2705 probability distributions that allow us characterize how different detection algorithms
2706 will perform for CRES signals in an FSCD experiment.

2707 **4.4.3 Signal Detection Algorithms**

2708 **4.4.3.1 Power Threshold**

2709 The power threshold detection algorithm uses the maximum amplitude of the frequency
2710 spectra as the detection test statistic. To model the performance of this approach,
2711 consider first the case where the signal is pure WGN. For a single bin in the frequency
2712 spectrum, the probability that the amplitude falls below a specific threshold value is

2713 given by the Rayleigh cumulative distribution function (CDF),

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

2714 where the complex amplitude of the frequency bin is x , and τ is the WGN variance.
 2715 Because the noise samples for each frequency bin are independent and identically dis-
 2716 tributed (IID), the probability that every bin in the frequency spectrum falls below the
 2717 threshold is the joint CDF formed by the product of each individual frequency bin CDF,

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

2718 The PDF for the power threshold classifier can then be obtained by differentiating the
 2719 CDF.

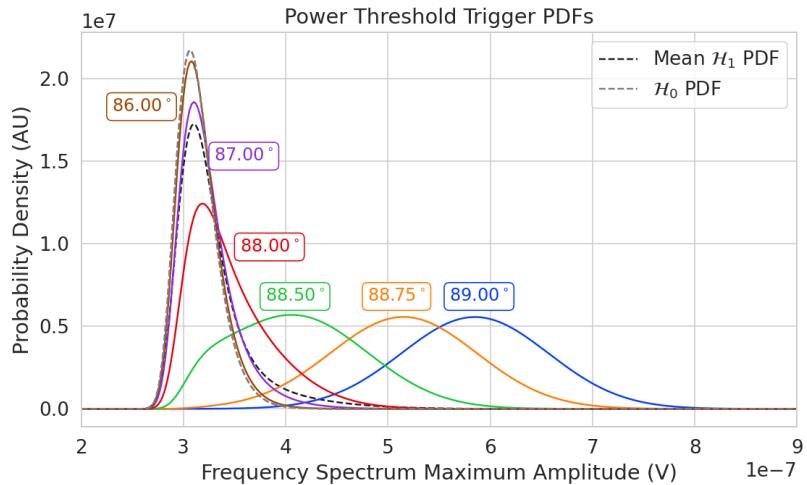


Figure 4.32: PDFs of the power threshold test statistic for CREs signals with various pitch angles as well as the PDF for the noise-only signal case. The average PDF computed for pitch angles ranging from 85.5 to 88.5° is also shown. As the pitch angle is decreased the signal PDF converges towards the noise PDF which indicates that the power threshold trigger is unable to distinguish between small pitch angle signals and noise.

2720 The probability distribution for the power threshold classifier under \mathcal{H}_1 is formed in
 2721 a similar way, but the frequency bins that contain signal must be treated separately. For
 2722 a frequency bin that contains both signal and noise we can describe the probability that
 2723 the amplitude of the bin will fall below our threshold using the Rician CDF,

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

2724 where the parameter $|\nu|$ defines the noise-free amplitude of the signal and Q_1 is the
 2725 Marcum Q-function. This time the CDF that describes the probability that the entire
 2726 spectrum falls below the decision threshold is the product of both signal and noise CDFs,

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

2727 The first half of Equation 4.48 is the contribution from the bins in the frequency spectrum
 2728 that contain only noise, and the second half is the product of the Rician CDFs for the
 2729 frequency bins that contain signal peaks with a noise-free amplitude of $|\nu_k|$. In Figure
 2730 4.32 we show plots of example PDFs under \mathcal{H}_1 and \mathcal{H}_0 .

2731 4.4.3.2 Matched Filtering

2732 The shape of a CRES signal is completely determined by the initial conditions of the
 2733 electron as it is emitted from beta-decay, which implies that it is possible to apply
 2734 matched filtering as a signal detection algorithm. With a matched filter one uses the
 2735 shape of the known signal, which is called a template, to filter the incoming data by
 2736 computing the convolution between the signal and the data [19]. For cases where the
 2737 signal is buried in WGN, the matched filter is the optimal detector in that it achieves
 2738 the maximum probability of a true detection for a fixed false positive rate. Since CRES
 2739 signals have an unknown shape but are deterministic, we can apply a matched filter by
 2740 using simulations to generate a large number of signal templates called a template bank,
 2741 which spans the parameter space of possible signals. Then at detection time, we use the
 2742 template bank to identify signals by performing the matched filter convolution for each
 2743 template in an exhaustive search.

2744 As we saw from the frequency spectra in Figure 4.31, CRES signals are highly periodic
 2745 in nature. In such cases, it is advantageous to utilize the convolution theorem to replace
 2746 the matched filter convolution with an inner product in the frequency-domain. With the
 2747 convolution theorem, the matched filter test statistic that describes the detection of a
 2748 signal buried in WGN using a matched filter template bank is given by

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

2749 where $h^\dagger[n]$ is the complex conjugate of the signal template. For the case when our
 2750 template bank consists of only a single template it is possible to derive an exact analytical

2751 form for the PDF describing the matched filter test statistic. First, we derive PDF under
 2752 the signal hypothesis, where the equation describing the matched filter test statistic, also
 2753 known as the matched filter score, becomes

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

2754 Each noisy frequency bin represented by $y[n]$ is the sum between value of the signal
 2755 at that bin and complex WGN, which means that $y[n]$ is itself Gaussian distributed.
 2756 Therefore, the value of the inner product between the template and the data is also a
 2757 complex Gaussian variable; and, since the matched filter score is the magnitude of this
 2758 inner product, it must follow a Rician distribution.

2759 We can derive the equation for the Rician PDF by expressing the matched filter
 2760 template \mathbf{h} in terms of the corresponding simulated signal, which we write as \mathbf{x}_h to
 2761 distinguish from the signal in the data. Using the standard normalization and assuming
 2762 uncorrelated WGN, the matched filter templates can be written as

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

2763 where τ is the noise variance. Inserting this into Equation 4.49 and expressing the data
 2764 as a sum between a signal and a WGN vector yields,

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

2765 Next, we transform the expression by isolating the randomly distributed components
 2766 giving

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

2767 The first term of 4.53 can be simplified by using the Cauchy-Schawrz inequality to express
 2768 the magnitude of the inner product in terms of the magnitudes of the signal and template
 2769 as well as an orthogonality constant which we call "match" (Γ). Using this we obtain,

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

2770 The second term is a sum of Gaussian distributed variables, which we should expect also
 2771 follows a Gaussian distribution. Each of the samples $\nu[n]$ is described by

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

2772 where $\mathcal{N}(0, \tau)$ is a complex Gaussian distribution with zero mean and variance τ . There-
 2773 fore,

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

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

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

2774 We see that \mathcal{T} is magnitude of a complex variable with mean $|\mathbf{h}||\mathbf{x}|\Gamma$ and variance one. In
 2775 order to simply the expression a bit further, we define the quantity $\mathcal{T}_{\text{ideal}} = |\mathbf{h}||\mathbf{x}|\Gamma$, which
 2776 we call the ideal matched filter score, because it represents the value of the matched
 2777 filter inner product that we would expect if no noise was present in the signal. We can
 2778 write the matched filter test statistic as the magnitude of a two-dimensional vector in
 2779 the complex plane

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

2780 where n_r and n_i are the real and imaginary components of the noise each with variance
 2781 $1/2$, which is modeled by a Rician distribution with shape factor $\mathcal{T}_{\text{ideal}}$. Therefore, the
 2782 probability distribution of the matched filter test statistic is given by,

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

2783 where I_0 is the zeroth-order modified Bessel function.

2784 The shape of the matched filter score distribution is controlled by the parameter
 2785 $\mathcal{T}_{\text{ideal}}$, which is effectively the value of the matched filter score if the data contained no
 2786 noise. Without noise, the data vector reduces to the signal, \mathbf{x} , in which case Equation
 2787 4.50 becomes the magnitude of an inner product between two vectors. We can write
 2788 the magnitude of an inner product in terms of the lengths of the individual vectors and
 2789 a constant that describes the degree of orthogonality between them. Applying this to

2790 Equation 4.50, we obtain

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

2791 where Γ describes the orthogonality between \mathbf{h} and \mathbf{x} . From the point of view of matched
2792 filtering, we can interpret Γ as describing how well the template matches the underlying
2793 signal in the data.

2794 The matched filter score PDF under the noise hypothesis can be readily obtained
2795 from Equation 4.60 by setting the value of $\mathcal{T}_{\text{ideal}}$ to zero, since the data contains no signal
2796 in the noise case. Doing this, we obtain the Rayleigh distribution that describes the
2797 matched filter score under \mathcal{H}_0 ,

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

2798 Equations 4.60 and 4.62 describe the behavior of the matched filter test statistic
2799 under \mathcal{H}_0 and \mathcal{H}_1 for a single template. However, defining a PDF that describes the
2800 matched filter test statistic in the case of multiple templates is in general a mathematically
2801 intractable problem, since there is no guarantee of orthogonality between matched filter
2802 templates. This leads to correlations between the matched filter scores of different
2803 templates because only one sample of noise is used to compute the matched filter scores
2804 of the template bank. In order to proceed, we need to make the simplifying assumption
2805 that we can treat the matched filter scores as IID variables, which allows to ignore
2806 correlations between templates. The overall effect of this will be an underestimate of the
2807 performance of the matched filter, since we are under counting the number of templates
2808 that could contribute a detectable score.

2809 For \mathcal{H}_0 we model the probability that the matched filter score falls below our threshold
2810 using the CDF obtained by integrating Equation 4.62. Because we are assuming that
2811 the matched filter scores using different templates are independent, the probability that
2812 the matched filter score for all templates falls below a threshold value is the joint CDF
2813 formed by multiplying the CDF for each template. Under \mathcal{H}_0 this is

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

2814 where x is the matched filter score threshold and N_t is the number of templates. We
2815 should expect that the distribution describing the matched filter template bank maximum
2816 score depends on N_t , because with more templates there is a greater chance of a random
2817 match between the template and data.

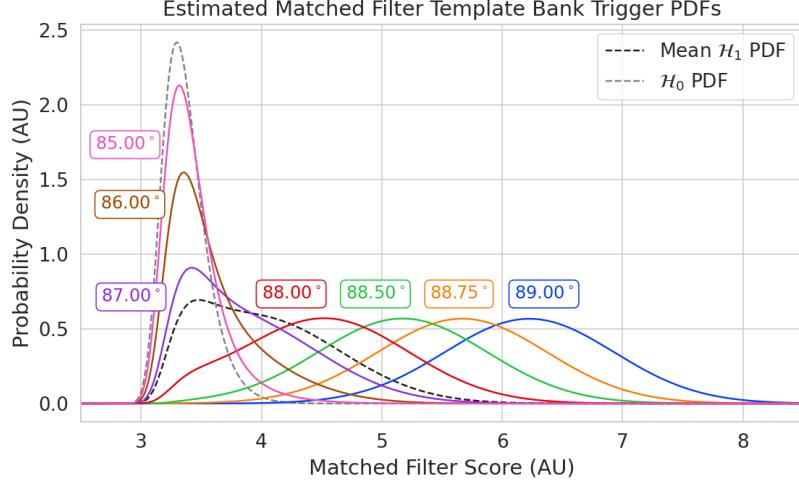


Figure 4.33: Plots of the estimated PDFs for the matched filter template bank test statistic for CRES signals with various pitch angles as well as the estimated PDF for the noise only signal case. We assume an estimated number of templates of 10^5 and perfect match between signal and template i.e. $\Gamma_{\text{best}} = 1$. 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.

For \mathcal{H}_1 , we start by denoting the CDF of the best matching template as $F_{\text{best}}(x; \mathcal{T}_{\text{best}})$, and treat the matched filter scores for all other templates as negligible ($\mathcal{T}_{\text{ideal}} \approx 0$). Then we form the joint CDF by combining the distributions for all templates used during detection. Since we are exhaustively checking the matched filter scores, the number of templates checked will be a randomly distributed variable that ranges from zero to the total number of available templates. If we assume that signals are uniformly distributed across the parameter space spanned by the template bank then on average we check $(N_t - 1)/2 \approx N_t/2$ templates for each inference. Therefore, the estimated CDF under \mathcal{H}_1 is

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

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

4.4.3.3 Machine Learning

In this paper we focus on Convolutional Neural Networks (CNN) as an example of a machine learning based signal classifier. CNNs are constructed using a series of

2832 convolutional layers, each composed of a set of filters that are convolved with the input
 2833 data. The individual convolutional filters can be viewed as matched filter templates that
 2834 are learned from a set of simulated data rather than being directly generated. This opens
 2835 the possibility of finding a more efficient representation of the matched filter templates
 2836 during the training process that can potentially reduce computational cost at inference
 2837 time while still offering good classification performance.

2838 The machine learning approach is distinct from both the power threshold and matched
 2839 filtering in that we do not attempt to manually engineer a test statistic that is computed
 2840 from the data for classification. Instead, we attempt calculate the test statistic by
 2841 constructing a differentiable function that maps the complex frequency series generated
 2842 by the STFT to a binary classification as either signal or noise. The test statistic for the
 2843 machine learning classifier can be expressed as

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

2844 where \mathbf{y} is the noisy data vector and $G(\mathbf{y}; \boldsymbol{\Omega})$ is the machine learning model parameterized
 2845 by the weights $\boldsymbol{\Omega}$. By using supervised learning on a labeled set of training signals, we
 2846 can modify the function parameters to learn the mapping from the data to the likelihood
 2847 of \mathbf{y} belonging to either \mathcal{H}_1 or \mathcal{H}_0 .

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

2848 The CNN architecture used for this work is summarized by Table 4.1. No strategic
 2849 hyper-parameter optimization approach was implemented beyond the manual testing

2850 of different CNN architecture variations, so this particular model is best viewed as a
2851 proof-of-concept rather than a rigorously optimized design. Numerous model variations
2852 were tested, some with significantly more layers and convolutions filters per layer, as
2853 well as others that were even smaller than the architecture in Table 4.1. Ultimately, the
2854 model architecture choice was driven by the motivation to find the minimal model whose
2855 classification performance was still comparable to the larger CNN’s tested, because of
2856 the importance of minimizing computational cost in real-time applications. It is possible
2857 that more sophisticated machine learning models could improve upon the classification
2858 results achieved here, but we leave this investigation for future work.

2859 **4.4.4 Methods**

2860 **4.4.4.1 Data Generation**

2861 To test the triggering performance of the classifiers, simulated CRES signals were
2862 generated using the Locust simulations package [10, 28] developed by the Project 8
2863 collaboration. Locust uses the separately developed Kassiopeia package to calculate the
2864 magnetic fields produced by a user defined set of current carrying coils along with any
2865 specified background magnetic fields, resulting in a magnetic trap. Next, Kassiopeia
2866 calculates the trajectory of an electron in this magnetic field starting from a set of user
2867 specified initial conditions. The Locust software then uses the electron trajectories from
2868 Kassiopeia to calculate the resulting electromagnetic fields using the Liénard-Wiechert
2869 equations, and determine the voltages generated in the antenna array with the antenna
2870 transfer function. Locust then simulates the down-conversion, filtering, and digitization
2871 steps resulting in the simulated CRES signals for an electron.

2872 The shape of the received CRES signal is determined by the initial kinematic param-
2873 eters, including the starting position of the electron, the starting kinetic energy of the
2874 electron, and the pitch angle. For the studies performed here we constrain ourselves to a
2875 single initial electron position located at $(x, y, z) = (5, 0, 0)$ mm, and using this starting
2876 position we generate two datasets by varying the initial kinetic energy and the starting
2877 pitch angle. The first dataset consists of a two-dimensional square grid of kinetic energy
2878 and pitch angle spanning an energy range from 18575-18580 eV with a spacing of 0.1 eV,
2879 and pitch angles from 85.5-88.5° with a spacing of 0.001°, resulting in 153051 signals with
2880 a unique energy-pitch angle combination. This dataset is intended to represent a matched
2881 filter template bank. The second dataset was generated by randomly sampling uniform
2882 probability distributions covering the same parameter space to produce approximately

2883 50000 signals randomly parameterized in energy and pitch angle. This dataset provides
2884 the training and test data for the machine learning approach, and acts as a representative
2885 sample of signals to evaluate the performance of the matched filter template bank.

2886 Each signal was simulated for a duration of $40.96 \mu\text{s}$, which is equivalent to 8192
2887 samples at the FSCD digitization rate, and begins at time $t = 0 \text{ s}$ for all simulations.
2888 This duration represents a single frequency spectrum generated by the STFT. The output
2889 of the Locust simulation is a matrix of array snapshots with size given by the number of
2890 channels times the event length ($N_{\text{ch}} \times N_{\text{sample}}$), which we pre-process using the digital
2891 beamforming summation and STFT described in Section 4.4.2.2. The ∇B -drift correction
2892 uses the exact value of $\omega_{\nabla B}$, obtained from the Kassiopeia simulation of that electron.
2893 In practice, an average value for $\omega_{\nabla B}$ could be used, because there is limited variation in
2894 drift frequency across this parameter space.

2895 4.4.4.2 Template Number and Match Estimation

2896 The estimated PDF for the matched filter template bank depends on the score of the
2897 best matching template or equivalently the match of the best template (Γ_{best}) as well
2898 as the number of templates. One expects that with a higher number of templates the
2899 average value of Γ_{best} will increase, however, there is a point of diminishing returns at
2900 which more templates will not significantly increase match, but will still increase the
2901 likelihood of false positives. Therefore, it is desirable to use the minimum number of
2902 templates that provide an acceptable mean value of Γ_{best} .

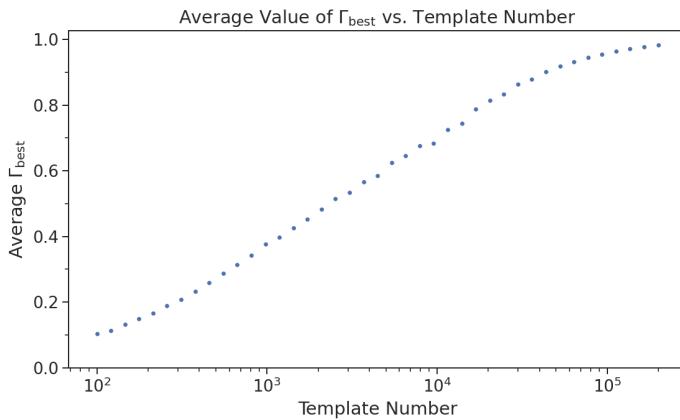


Figure 4.34: The mean match of the matched filter template bank to a test set of randomly parameterized signals as a function of the number or density of templates. The parameter space includes pitch angles from $85.5 - 88.5^\circ$ and energies from 18575 – 18580 eV.

2902

2903 To quantify the relationship between match and template number, we calculated
2904 the mean match of the random dataset to a selection of templates obtained from the
2905 regularly spaced dataset. The results are shown in Figure 4.34, where we find that the
2906 average value of Γ_{best} is an exponential function of the number of templates. From this
2907 plot we select the desired value of mean match at which we would like to evaluate the
2908 matched filter PDF and can infer the required number of templates.

2909 **4.4.4.3 CNN Training and Data Augmentation**

2910 To prepare the data for training the model, we split the random dataset in half to create
2911 distinct training and test datasets. Additionally, a randomly selected 20% of the training
2912 data is isolated for use as a validation set during the training loop. The size of the
2913 training, validation, and test datasets are then tripled by appending two additional copies
2914 of the data to increase the sample size of the dataset after data augmentation. The
2915 data is loaded with no noise, which is added to each data batch during the training
2916 phase by generating a new noise sample from a complex WGN distribution. In order to
2917 ensure an even split between signal and noise data we append to the noise-free signals an
2918 equal number of empty signals composed of all zeros. Therefore, as the data is randomly
2919 shuffled during training, on average an equal number of empty signals will be included
2920 with the training signals. After adding the sample of WGN to the data batch, the empty
2921 signals represent the noise-only data that the model must distinguish from signal data.

2922 As the training signals are loaded we apply a unique random phase shift as the
2923 first form of data augmentation. Since the data is generated using the same initial
2924 axial position and cyclotron orbit phase, the randomization is an attempt to prevent
2925 overtraining on these features. During each training epoch the data is randomly shuffled
2926 and split into batches of 2500 signals. Each batch of signals is then circularly shifted
2927 by a random number of frequency bins to simulate a kinetic energy shift from -75 to
2928 20 eV to simulate a training dataset with a larger energy range. Next, a sample of
2929 complex WGN, consistent with the expected 10 K Nyquist-Johnson noise expected for
2930 the FSCD, is generated and added to the signal, which prevents overtraining on noise
2931 features. As a final step, the data is renormalized by the standard deviation of the noise
2932 so that the range of values in the data is close to $[-1, 1]$, which helps ensure well-behaved
2933 back-propagation.

2934 The Binary Cross-entropy loss function is used to compute the loss for each batch of
2935 data and the model weights are updated using the ADAM optimizer with a learning rate
2936 of 5×10^{-3} . After each training epoch, the loss and classification accuracy of the validation

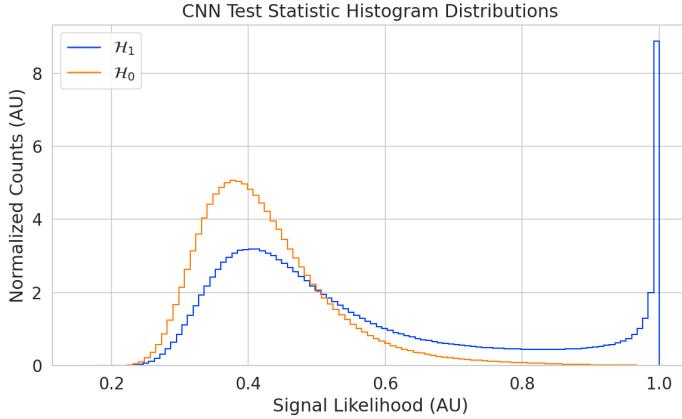


Figure 4.35: Histograms of the trained CNN model output from the test dataset. The blue histogram shows the model outputs for signal data. The oddly shaped peak near the end is the result of the softmax function mapping the long tail of the raw output distribution to the range [0, 1].

dataset are computed to monitor for overtraining. It was noticed that the relatively high noise power and the fact that a new sample of noise was used for each batch together provided a strong form of regularization, since no evidence of over-training was observed even after several thousand epochs. Typically, the loss and classification accuracy of the model converged after a few hundred training epochs, but the training loop was extended to 3000 epochs to attempt to achieve the best possible performance. The training procedure generally took about 24 hrs using a single NVIDIA V100 GPU [30].

After training the model, we use it to classify the test dataset and generate 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.

4.4.5 Results and Discussion

4.4.5.1 Trigger Classification Performance

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

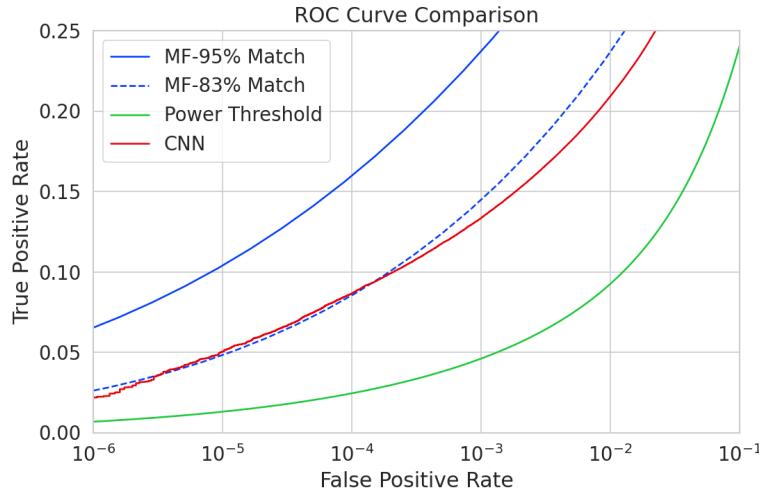


Figure 4.36: ROC curves describing the detection efficiency or true positive rates for the three signal classification algorithms examined in this paper.

2956 over the full signal parameter space in order to get a measure of the overall detection
 2957 efficiency achieved by each algorithm. For the power threshold and matched filter
 2958 algorithms, we obtain the mean ROC curve by taking the average over all signals in the
 2959 regularly spaced dataset. In the case of the matched filter, we examine two cases using
 2960 different numbers of templates, which have different values of mean match. The ROC
 2961 curve describing the CNN is obtained by forming a histogram of the network outputs
 2962 for each class of signal and from this computing the estimated CDFs and ROC curve.
 2963 In Figure 4.36, we show the ROC curves obtained for each of the detection algorithms,
 2964 visualized in terms of true positive rate and false positive rate.

2965 The true positive rate of a signal classifier is equivalent to its detection efficiency, and
 2966 we see that for the population of signals with pitch angles $< 88.5^\circ$ the power threshold
 2967 has a consistently lower detection efficiency than the CNN and the matched filter. This
 2968 result could have been predicted from the visualization of signal spectra in Figure 4.31,
 2969 where we see that there is no way to distinguish between a noise peak and a signal peak
 2970 with high confidence at small pitch angles. The CNN offers a significant and consistent
 2971 increase in detection efficiency over the power threshold approach, with the relative
 2972 improvement in detection efficiency increasing as the false positive rate decreases. If
 2973 we compare the CNN to the matched filter, we see that the performance of the tested
 2974 network is roughly equivalent to a matched filter detector with an average match of
 2975 about 83%, which uses approximately 20000 matched filter templates. The overall best
 2976 detection efficiency is achieved by the matched filter classifier if a large enough template

bank is used. We show in the plot the ROC curve for a matched filter template bank with 95% average match, which is achieved with approximately 100000 templates. Since the matched filter is known to be statistically optimal for detecting a known signal in WGN, it is somewhat expected that this algorithm has the highest detection efficiency.

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

4.4.5.2 Computational Cost and Hardware Requirements

In the process of investigating triggering approaches for an antenna array CRES experiment, we have uncovered a strong tension between detection efficiency and computational resources. To relate the computational cost estimates to actual costs, we compare the theoretical amount of computer hardware required to implement the signal classifiers for real-time detection in an FSCD experiment. To do this we shall utilize order of magnitude estimates of the theoretical peak performance values for currently available Graphics Processing Units (GPUs) as a metric. This approach will underestimate the amount of required hardware, since it is unlikely that any CRES detection algorithm could reach the theoretical peak performance of the hardware.

Of the three detection algorithms tested, the power threshold classifier is the least expensive. It requires that we check whether the amplitude of each frequency bin in the STFT is below or above our decision threshold. The STFT combined with digital beamforming produces $N_{\text{bin}}N_b$ frequency bins that must be checked every N_{bin}/f_s seconds. This requires approximately $O(10^{10})$ FLOPS to check in real-time. Current generations of GPUs have peak theoretical performances in the range of $O(10^{13}) - O(10^{14})$ FLOPS [31], dependent on the required floating-point precision of the computation. Therefore, the

3011 entire computational needs of a real-time triggering system using a power threshold
3012 classifier, including digital beamforming and generation of the STFT, could be met by a
3013 single high-end GPU or a small number of less powerful GPUs. Since triggering is only
3014 one step of the full real-time signal reconstruction approach, limiting the computational
3015 cost of this stage is ideal. However, we have seen that the power threshold classifier does
3016 not provided sufficient detection efficiency across the entire range of possible signals,
3017 which is the primary motivation for exploring more complicated triggering solutions.

3018 As discussed, the computational cost of the matched filter approach requires counting
3019 the number of templates that must be checked for each frequency spectra produced by the
3020 STFT. Computing the matched filter scores requires $O(N_b N_t N_{\text{bin}})$ operations, since for
3021 each of the N_b beamforming positions we must multiply N_t templates with a data vector
3022 that has length N_{bin} . The time within which we must perform this calculation is equal
3023 to N_{bin}/f_s to keep up with the data generation rate. To cover the 5 eV kinetic energy
3024 range spanned by the template bank, we saw that 10^4 to 10^5 templates are required in
3025 order to match or exceed the detection efficiency of the CNN. If the number of templates
3026 scales linearly with then kinetic energy range of interest as expected, then we would
3027 require 10^5 to 10^6 matched filter templates with this more realistic range of energies.
3028 Considering this, the estimated computational cost of the matched filter is between
3029 $O(10^{15})$ to $O(10^{16})$ FLOPS, which is $O(10^2)$ to $O(10^3)$ high-end GPUs.

3030 Lastly, we have the CNN classifier. To estimate the computational cost we simply
3031 sum the number of convolutions and matrix multiplications specified by the network
3032 architecture shown in Table 4.1. Each convolutional layer consists of $N_{\text{in}} N_{\text{out}} N_{\text{kernel}} L_{\text{input}}$
3033 floating-point operations, where N_{in} is the number of input channels, N_{out} is the number
3034 of output channels, N_{kernel} is the size of the convolutional kernel, and L_{input} is the length
3035 of the input vector, and the fully connected layers each contribute $N_{\text{in}} N_{\text{out}}$ operations.
3036 Summing all the neural network layers we estimate that the CNN would require $O(10^6)$
3037 floating point operations for each frequency spectra; therefore, the total computation
3038 cost of the CNN trigger is this cost times the number of beamforming positions per the
3039 data acquisition time, which is $O(10^{13})$ FLOPS or $O(10^0)$ GPUs.

3040 Compared with the matched filter approach the CNN requires $O(100)$ to $O(1000)$
3041 fewer GPUs to implement, dependent on the exact number of templates used in the
3042 template bank. The 100 eV kinetic energy range is motivated by the application of these
3043 detection algorithms to an FSCD-like neutrino mass measurement experiment. However,
3044 if a significantly larger range of kinetic energies is required, a CNN may be the preferred
3045 detection approach despite the lower average detection efficiency due to computational

3046 cost considerations. The low estimated computational cost of the CNN is directly related
3047 to the small network size.

3048 Additional experiments with larger CNNs, generated by increasing the depth and
3049 width of the neural network, and we observed that these changes provided minimal
3050 ($\lesssim 1\%$) improvement in the classification accuracy of the model. A potential reason
3051 for this could be the sparse nature of the signals in the frequency domain and the low
3052 SNR which makes for a challenging dataset to learn from. Future work could investigate
3053 modifications to the neural network architecture such as sparse convolutions, which may
3054 improve the classification accuracy of the model or further reduce the computational
3055 costs of this approach. Alternatively, more complicated CNN architectures such as a
3056 ResNet [32] or VGG model [33] may provide improved classification performance over a
3057 basic CNN. An additional promising area of investigation are recurrent neural networks,
3058 which may be able to exploit the time-ordered features of the STFT for more accurate
3059 signal detection if the electron signals last for multiple Fourier transform windows.

3060 Our estimate of the computational cost of the matched filter is somewhat naive if
3061 we notice that the majority of the values that make up a CRES frequency spectra are
3062 zero (see Figure 4.31). Therefore, the majority of operations in the matched filter inner
3063 product are unnecessary, and we could instead evaluate the matched filter inner product
3064 using only the $\lesssim 10$ frequency peaks that make up CRES signal. This optimization
3065 reduces the number of operations required to check each template by a factor of $O(100)$
3066 to $O(1000)$, which brings the estimated computational cost of the matched filter in
3067 line with the CNN. Although this level of sparsity results in a multiplication with very
3068 low arithmetic complexity, the resulting sparse matched filter algorithm is still likely
3069 to be constrained by memory access speed rather than compute speed. Ultimately, the
3070 comparison of the relative computational and hardware costs between the matched filter
3071 and CNN will depend on the efficiency of the software implementation and hardware
3072 support for neural network and sparse matrix calculations.

3073 **4.4.6 Conclusion**

3074 Increasing the detection efficiency and overall event rate of the CRES technique represents
3075 a key developmental path towards new scientific results and broader applications of the
3076 CRES technique. It is what motivates both the antenna array detection approach and
3077 the development of real-time signal reconstruction algorithms. We have demonstrated
3078 that significant gains in the detection efficiency of the CRES technique are achievable
3079 by utilizing triggering algorithms that account for the specific shape of CRES signals in

3080 the detector. These algorithms emphasize the need for accurate and fast methods for
3081 CRES simulation, since they directly contribute to the success of matched filter methods
3082 by providing a way to generate expected signal templates and also serve as a source of
3083 training data for machine learning approaches.

3084 The improvements in detection efficiency offered by these alternative approaches to
3085 triggering are crucial to the success of efforts to develop scalable technologies for CRES
3086 measurement, since they provide a significant increase in the detectable parameter space
3087 of CRES events, which allows for a better utilization of the larger detection volume.
3088 While we have focused on the real-time detection of CRES signals from antenna arrays,
3089 these same signal classifiers could be used in CRES experiments utilizing a different
3090 detector technologies, since the same principles of signal detection will apply. For example,
3091 previous CRES measurements by the Project 8 collaboration that utilized a waveguide
3092 gas cell, could have improved their detection efficiency by employing a matched filter
3093 or neural network classifier to identify trapped electrons with pitch angles that are too
3094 small to be detected by the power threshold approach. Furthermore, alternative CRES
3095 detector technologies such as resonant cavities [3] could also see similar improvements
3096 in detection efficiency, which is of crucial importance to future efforts by the Project 8
3097 collaboration to utilize CRES to measure the neutrino mass.

3098 **Chapter 5** |

3099 **Antenna and Antenna Measurement Sys-**

3100 **tem Development for the Project 8 Ex-**

3101 **periment**

3102 **5.1 Introduction**

3103 The FSCD and antenna array CRES represent an innovative approach to beta-decay
3104 spectroscopy. While much can be learned from simulations about the systematics of
3105 CRES with antenna arrays, laboratory measurements and demonstrations provide critical
3106 inputs to sensitivity and simulation models as well as provide a means for calibration and
3107 commissioning of the experiment. Therefore, a robust program of antenna and antenna
3108 measurement hardware development is important to the success of the FSCD and the
3109 development of antenna array CRES more broadly.

3110 In this chapter we summarize the development of an antenna measurement system
3111 at Penn State to implement and test the techniques of antenna array CRES on the
3112 bench-top, in order to support the efforts of the Project 8 collaboration. In Section 5.2
3113 we provide an introduction to some fundamental parameters and concepts related to
3114 antenna measurements as well as an overview of the Penn State antenna measurement
3115 system hardware. In Section 5.3 we include the manuscript of a paper [29] which
3116 details the design and characterization of a specialized antenna developed to mimic the
3117 electric fields emitted by an electron in a CRES experiment. This antenna, called the
3118 Synthetic Cyclotron Antenna (SYNCA), is intended as a calibration tool for antenna
3119 arrays developed for CRES measurements. Lastly, in Section 5.4 we summarize a set
3120 of prototype FSCD antenna array measurements with the SYNCA [6], which we use to
3121 validate the simulated performance of the antenna array and estimate systematic errors
3122 associated with the antenna array.

3123 5.2 Antenna Measurements for CRES experiments

3124 5.2.1 Antenna Parameters

3125 Antenna characterization measurements are intended to validate simulations of the
3126 antenna array performance, which ultimately informs the neutrino mass sensitivity of
3127 the experiment. In this section, I shall summarize a few fundamental concepts relating
3128 to antennas and antenna measurement, before introducing how Project 8 uses antenna
3129 measurement for the development of antenna array CRES.

3130 5.2.1.1 Radiation Patterns

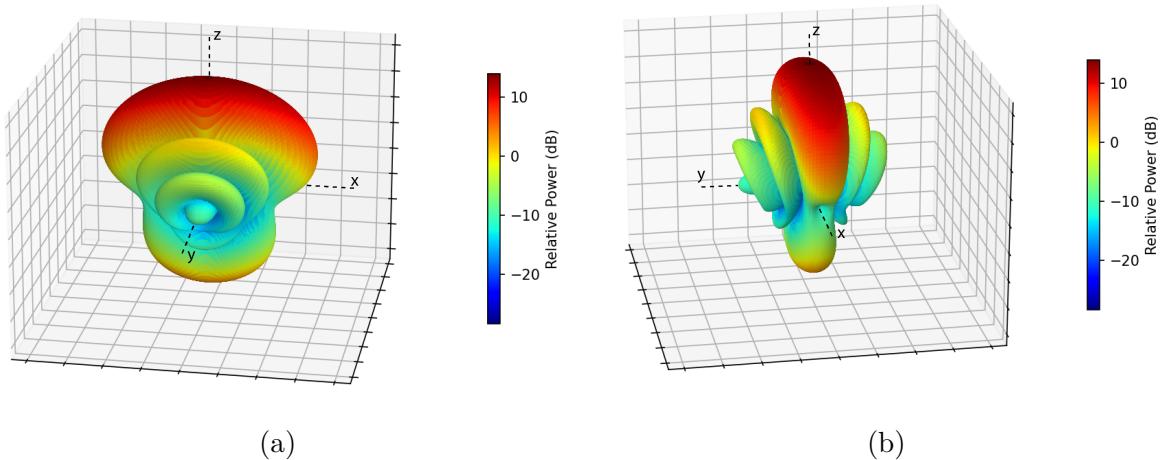


Figure 5.1: An example radiation pattern generated using HFSS simulations. The color and radial distance of the surface from the origin indicate the relative magnitude of radiation power emitted by the antenna in that direction. The primary goal of most antenna measurements is typically to measure the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

3131 Antennas are conductive structures designed to carry alternating electric currents
3132 in order to transmit energy in the form of electro-magnetic (EM) waves [13]. Perhaps
3133 the most fundamental way to characterize an antenna, is to map out the radiated power
3134 density as a function of position, which is called the radiation pattern (see Figure 5.1).
3135 We find the radiation power density by calculating the time-averaged Poynting vector for
3136 all positions surrounding the antenna, which in equation form is

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

3137 where $\mathbf{E}(x, y, z, t)$ and $\mathbf{H}(x, y, z, t)$ are the time-dependent electric and magnetic fields
 3138 produced by the antenna [8]. The radiation power density has units of W/m^2 and is
 3139 more typically called the energy flux density in physics applications, since it is a measure
 3140 of the amount of energy passing through a unit area over time.

3141 Because the radiation power density is a measure of power per unit area, its value
 3142 in a particular direction will depend on the distance from the antenna at which we are
 3143 measuring. This is undesirable for practical applications A related quantity, which is
 3144 distance independent, is the energy flux per unit solid angle or radiation intensity, which
 3145 is computed directly from the radition power density by multiplying by the squared
 3146 distance from the antenna. Specifically,

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

3147 where r is the distance from the antenna to the field measurement point. The radiation
 3148 intensity is typically defined in regions where the Poynting vector consists only of a radial
 3149 component where it is safe to treat as a scalar quantity.

3150 5.2.1.2 Directivity and Gain

3151 Since the radiation intensity is a measure of average power per unit solid angle, it is
 3152 independent of distance and more useful as feature for antenna measurement. However,
 3153 most antenna measurements are performed in terms of the directly related directivity
 3154 and gain quantities. Directivity is defined as the ratio between the radiation intensity at
 3155 particular point on the radiation pattern to the average radiation intensity computed
 3156 over all solid angles [13]. The equation that relates the radiation intensity to directivity
 3157 is

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

3158 where U_0 is the average radiation intensity over all solid angles, which simply the total
 3159 radiated power (P_{rad}) divided by 4π . Closely related to directivity is concept of gain,
 3160 which accounts for energy losses that occur inside then antenna when attempting to
 3161 transmit or receive a signal. The antenna gain is given by

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

3162 where P_{in} is the total power delivered to the antenna. Gain can be thought of as the ratio
 3163 of the antenna's radiation intensity to that of a hypothetical isotropic, lossless radiator.

³¹⁶⁴ The maximum values of gain and directivity exhibited by the main lobe of the antenna
³¹⁶⁵ pattern as well as the ratio between the gain of the main lobe and any side-lobes are
³¹⁶⁶ important figures of merit used to evaluate antenna designs.

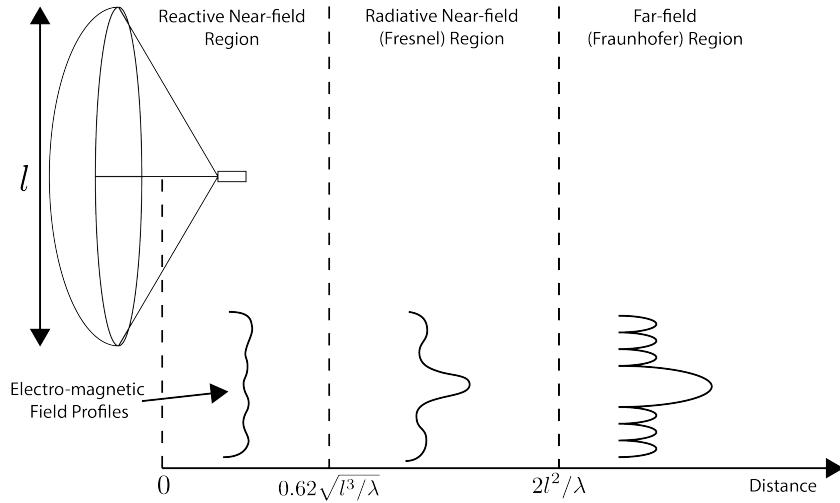


Figure 5.2: An illustration of the three field regions important for the analysis of an antenna system. Very close to the antenna the electric fields are primarily reactive so there is no radiation. If a receiving antenna were placed in this region most of the energy would be reflected back to the transmitter. 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.

³¹⁶⁷ 5.2.1.3 Far-field and Near-field

³¹⁶⁸ Radiation patterns are only well-defined in regions where the shape of the radiation
³¹⁶⁹ pattern is independent of distance. The region where this approximation is valid is called
³¹⁷⁰ the "far-field", and in this region we can treat the EM fields from the antenna as spherical
³¹⁷¹ plane waves. A rule of thumb for antennas is that the far-field approximation is valid
³¹⁷² when the condition

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

³¹⁷³ is met. In this expression, R is the distance from the antenna, l is the largest characteristic
³¹⁷⁴ dimension of the antenna, and λ is the wavelength of the radiation (see Figure 5.2).

3175 The region very close to the antenna is called the reactive near-field, because in this
3176 region the reactive component of the EM field is dominant. Unlike radiative electric
3177 fields, the reactive electric and magnetic fields are out of phase from each other by
3178 90°, since they are the result of electrostatic and magnetostatic effects coming from the
3179 self-capacitance and self-inductance of the antenna. The reactive fields are unable to
3180 transfer energy a significant distance from the antenna and are thus completely negligible
3181 for most antenna applications. The limit of the reactive near-field for an electrically-large
3182 antenna is typically taken to be

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

3183 The unique application of antennas by Project 8 is somewhat limited by reactive near-
3184 field effects in the form of a maximum radial position for electrons inside the uniform
3185 cylindrical antenna array. If electrons are too close to the edge of the array than reactive
3186 near-field effects leads to a large reduction in the received power and consequently
3187 detection efficiency. This leads to a significant volume inside of the antenna array that
3188 is unsuitable for CRES lowering the volumetric efficiency of the antenna array CRES
3189 technique relative to a cavity experiment.

3190 In between the reactive near-field and the far-field is the radiative near-field region.
3191 In this region the fields are primarily radiative, however we are still too close to the
3192 antenna for the spherical plane wave approximation to apply. Therefore, interference
3193 effects between EM waves emitted from different points on the antenna occur causing the
3194 shape of the radiation pattern to change as a function of distance from the antenna. If we
3195 evaluate the far-field distance limit for the FSCD one finds an estimated far-field distance
3196 of 43 cm, which is a factor of four larger than the radius of the antenna array designed for
3197 the experiment. Consequently, we expect near-field effects to influence the performance
3198 of the antenna array highlighting the importance of calibration and characterization
3199 measurements.

3200 **5.2.1.4 Polarization**

3201 The polarization of an EM wave defines the spatial orientation of the electric field
3202 oscillations in the plane perpendicular to the direction of the propagation, and is defined
3203 in terms of orthogonal polarization components. In our application, one analyzes the
3204 properties of radiation propagating along the radial (\hat{r}) direction away from the antenna,
3205 which implies that the electric fields can be described as a linear combination of orthogonal

3206 polarization components

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

3207 in Cartesian coordinates, or

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

3208 in spherical coordinates.

3209 In general, one defines partial radiation patterns, directivities, and gains so that the
3210 performance of the antenna for the desired polarization can be analyzed. The radiation
3211 pattern defined in terms of partial patterns is

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

3212 where U_ϕ and U_θ are the radiation intensities in a particular direction for the respective
3213 polarization components. Similarly, a quantity such as gain can be written in terms of
3214 partial gains,

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

3215 If we view an electron performing a circular orbit in the XY-plane from the side, that
3216 is, along the X or Y axes, then we would observe the electron to be performing a linear
3217 oscillation perpendicular to the viewing axis. From this intuitive picture, we can predict
3218 that the primary polarization of electric fields from CRES events to be linearly polarized
3219 in the $\hat{\phi}$ direction when viewed with an antenna positioned in the XY-plane.

3220 5.2.1.5 Antenna Factor and Effective Aperture

3221 A useful way to characterize the performance of an antenna is to measure the electric
3222 field magnitude required to produce a signal with an amplitude of one volt in the antenna
3223 terminals. This ratio between the magnitude of the incoming electric field and the
3224 magnitude of the signal produced by the antenna is called the antenna factor, which is
3225 written as

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

3226 where A_F is the antenna factor, E_{in} is the incoming electric field, and V_{ant} is the magnitude
3227 of the voltage produced by the antenna.

3228 The antenna factor can be expressed in terms of the antenna's gain through a related
3229 quantity called the effective aperture. The effective aperture defines for a given incident
3230 radiation power density (W/m^2) the power that is received by the antenna. Therefore,

3231 the effective aperture gives the equivalent area of the antenna,

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

3232 where the received power is P_r and the total incoming power is P_{in} .

3233 If we express the incident power in terms of the magnitude of the Poynting vector,
3234 then

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

3235 where η_0 is the impedance of free-space, which relates the magnitudes of the electric and
3236 magnetic fields in a vacuum, and is defined by

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

3237 The total received power by the antenna can therefore be expressed as

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

3238 To relate this to the antenna factor recall that we can relate the voltage produced by
3239 the antenna to the received power with

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

3240 where Z is the system impedance. Setting Equations 5.15 and 5.16 equal to each other,
3241 we obtain the following expression for antenna factor in terms of gain

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

3242 The second expression in Equation 5.17 is obtained by evaluating the constant terms
3243 assuming a system impedance of 50Ω .

3244 We have gone through the effort of expressing the antenna factor in terms of gain
3245 to highlight that the majority of antenna parameters that we care to measure for a
3246 CRES experiment can be obtained from the radiation or gain pattern of the antenna.
3247 The antenna factor is a particularly important parameter for CRES measurements
3248 due to its relevance to antenna array simulations with the Locust software [10, 28].
3249 Specifically, Locust simulates the trajectory of an electron in a magnetic trap by running

3250 the Kassiopeia software package [7] and then uses the Liénard-Wiechert equations [11, 12]
3251 to calculate the electric fields that are incident on the antenna.

3252 To compute the response of the antenna to the electric field, Locust relies upon linear
3253 time-invariant system theory [15], which computes the response of the antenna (i.e. the
3254 voltage time series generated by the antenna) using a convolution between the electric field
3255 time-series and the antenna impulse response. This approach is necessary for correctly
3256 modeling the antenna response to the electric field due to the broadband and non-
3257 stationary nature of the electric fields from CRES events. Since antenna measurements
3258 take place under steady-state conditions, parameters such as the radiation pattern, gain,
3259 and antenna factor are defined in the frequency domain. However, by performing an
3260 inverse Fourier transform on the antenna factor we can obtain the antenna impulse
3261 response, which allows us to simulate CRES events in the antenna array demonstrator
3262 experiment.

3263 **5.2.2 Antenna Measurement Fundamentals**

3264 **5.2.2.1 Friis Transmission Equation**

3265 The antenna factor, sometimes called the antenna transfer function, is used to model
3266 how the antenna will respond to electric fields emitted from a CRES event. Therefore,
3267 being able to measure the antenna transfer function of the antenna array is a key step
3268 in the commissioning and calibration phases of an antenna array CRES experiment. A
3269 common approach to antenna characterization is to perform a two antenna transmit-
3270 receive measurement where an antenna with a known gain is used to characterize the
unknown gain of the antenna under test (see Figure 5.3).

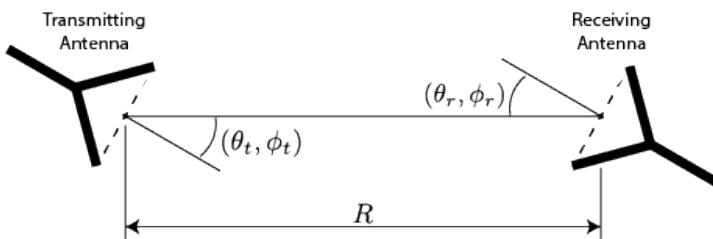


Figure 5.3: An illustration of the Friis measurement technique commonly used for antenna characterization measurements.

3271
3272 To analyze this two antenna setup we seek to calculate the amount of power from
3273 the transmitting antenna that we will detect with the receiving antenna. Using our
3274 understanding of antenna gain, we can calculate the power density transmitted by an

3275 antenna in a direction (θ_t, ϕ_t) at frequency f and distance R , which is given by

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

3276 Here, P_t is the total power delivered to the transmitting antenna and $G_t(\theta_t, \phi_t, f)$ is
 3277 the value of the transmitting antenna gain. The power density is the power per unit
 3278 area, so to calculate the total power delivered to the receiving antenna we multiply the
 3279 transmitted power density by the effective area of the receiving antenna,

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

3280 where $G_r(\theta_r, \phi_r, f)$ is the gain of the receiving antenna. Equation 5.19 is called the
 3281 Friis transmission equation [34, 35], which is of fundamental importance for antenna
 3282 measurements, since it allows one to measure the gain of an unknown antenna by
 3283 measuring the power received from an antenna with a known gain pattern. Alternatively,
 3284 if no antenna with a known gain pattern is available, two identical antennas with unknown
 3285 gain patterns can be used.

3286 5.2.2.2 S-Parameters and Network Analyzers

3287 Instead of directly measuring the power received by the antenna under test, it is more
 3288 common to measure the ratio of the received power to the transmitted power,

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

3289 This power ratio can be easily measured using a vector network analyzer (VNA), which
 3290 automates a significant fraction of the measurement process. Network analyzers are
 3291 used to measure the scattering or S-parameters of a multi-port RF device [36], which
 3292 describes how waves are scattered between the device ports. The antenna measurements
 3293 we have been considering can be modeled as a two-port microwave device that we can
 3294 characterize by measuring how incident voltage waves are transmitted or reflected (see
 3295 Figure 5.4). We can write the scattered waves (V_1^- and V_2^-) in terms of the incident (V_1^+
 3296 and V_2^+) waves using the scattering matrix

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

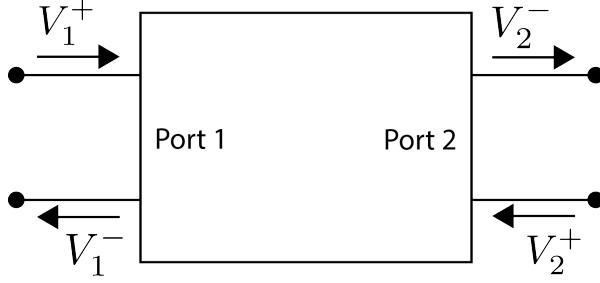


Figure 5.4: Illustration of a two-port S-parameter measurement setup. S-parameters characterize how incoming waves of voltage or power scatter off of the RF device under test. This allows you to measure important properties of the device. In particular, we can use this framework to model a two antenna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

3297 where the elements of the matrix are the device S-parameters. It is assumed that,
 3298 when exciting the device from a particular port, that all other ports in the network are
 3299 terminated at the system impedance. This ensures that the incident waves from other
 3300 ports in the network are zero. Therefore, the S-parameters are the ratios between the
 3301 scattered and incident waves,

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

3302 Alternatively, S-parameters can be defined as the ratio of the scattered and incident
 3303 power, which is proportional to the ratio of the squared voltage waves. Returning to
 3304 our antenna measurement setup, we see that measuring the ratio of the received to the
 3305 transmitted power is equivalent to measuring the ratio of power being scattered from port
 3306 1 to port 2 in a RF network. Therefore, measuring an antenna's gain can be accomplished
 3307 quite easily, by using a VNA to perform a two port S_{21} measurement.

3308 5.2.2.3 Antenna Array Commissioning and Calibration Measurements

3309 Up to this point we have been discussing calibration and commissioning measurements
 3310 as they apply to a single antenna. While these measurements play an important role
 3311 in validating the radiation patterns of the individual array elements, the ultimate goal
 3312 is to use a phased array of these antennas. Therefore, we must also consider antenna
 3313 measurement techniques that apply to the whole array system.

3314 By measuring the gain of each individual array element we can predict the features of
 3315 the signals received during a CRES event using the antenna factor (see Section 5.2.1.5).
 3316 However, unpredictable changes to the antenna performance can be introduced by the

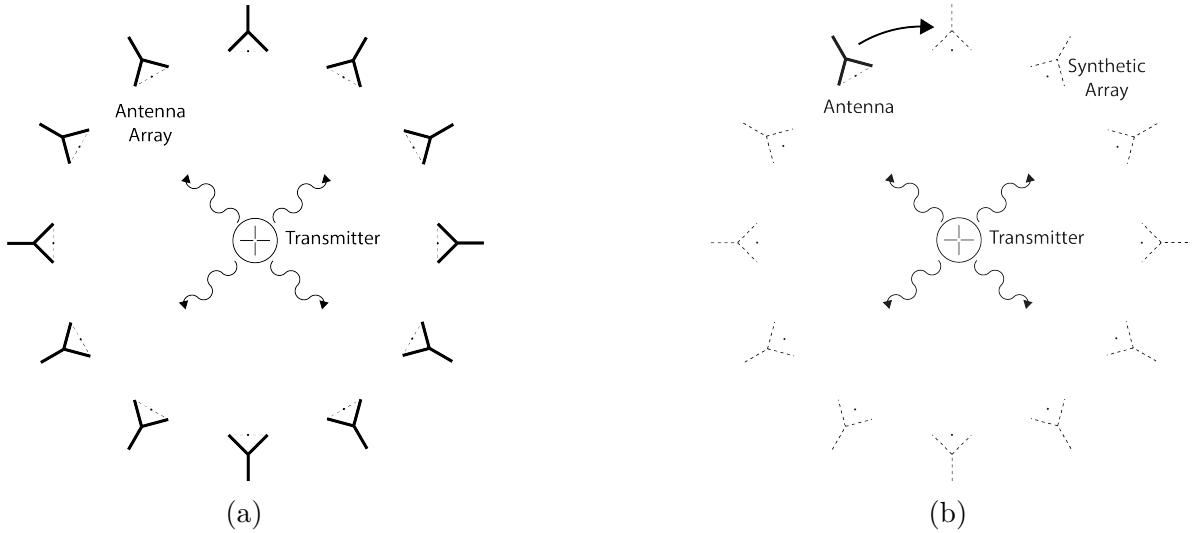


Figure 5.5: Two measurement approaches to characterizing an antenna array for CRES measurements. The full-array approach (a) requires a complete antenna array with all the associated hardware. The synthetic array approach (b) utilizes a single antenna and a set of rotation/translation stages to reposition 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.

incorporation of the antennas into the circular array geometry, therefore, we employ both individual antenna and full-array measurements in the commissioning of the FSCD to account for these effects.

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

5.2.3 The Penn State Antenna Measurement System

The development of antenna array based CRES requires the capability to test and calibrate different antenna array designs to validate the performance of the as-built antenna array before and during the experiment. With these aims in mind 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) [37] specifically designed to attenuate microwave radiation near the 26 GHz measurement range of the system.

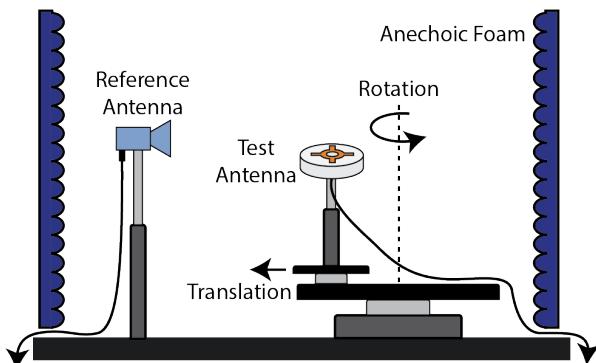


Figure 5.6: Illustration of the antenna measurement system developed for the Project 8 Collaboration. The reference and test antennas can be connected to different data acquisition configurations depending on the measurement goals. The reference antenna is typically a standard horn antenna and the test 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 form designed to mitigate RF reflections surrounds the setup.

In the first measurement configuration the reference antenna is typically a well-

3350 characterized horn antenna as pictured, since horn antennas have well-known and stable
 3351 radiation patterns making them ideal as standard references. For characterization
 3352 measurements, the test antenna represents the antenna-under-test whose pattern we wish
 3353 to characterize. Mounting the test antenna on motorized rotation and translation stages
 3354 allows us to automate the procedure significantly speeding up the radiation pattern
 3355 measurement process.

3356 In the second measurement configuration one is interested in recreating the conditions
 3357 of an antenna array CRES experiment as it concerns the antenna array and DAQ system.
 3358 In this case, the reference antenna is a prototype FSCD antenna, which will be used to
 3359 construct the antenna array in the FSCD experiment, and the test antenna is a specially
 3360 designed synthetic cyclotron antenna (SYNCA) as picture in Figure 5.6. The SYNCA is
 3361 designed such that the radiation pattern mimics that of a CRES electron so that the
 3362 signals received by the prototype CRES array antenna mimic what is expected for a real
 3363 CRES experiment.

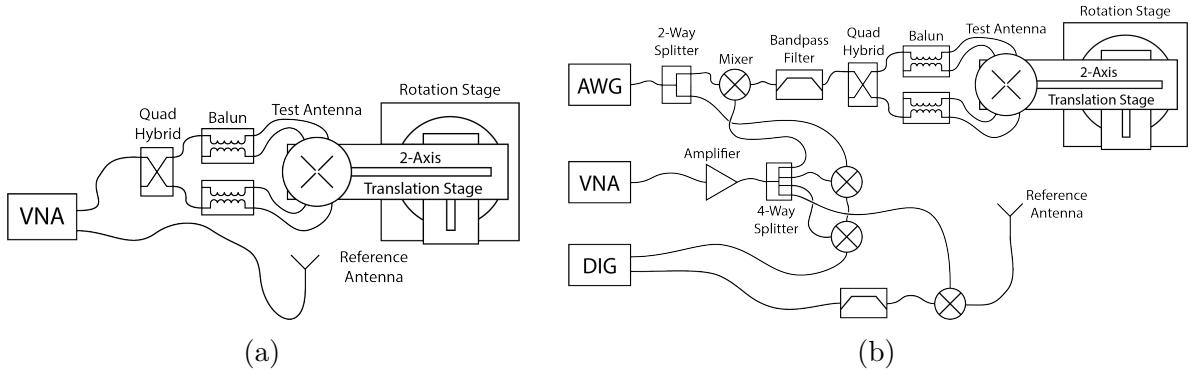


Figure 5.7: Diagrams of two measurement system configurations. Configuration (a) utilizes a VNA and is more suited to antenna characterization. Configuration (b) utilizes an AWG and VNA as a signal generation system and digitizer to collect measurement data, which is more suited to simulating CRES measurements. 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.

3364 In Figure 5.7 we show two high-level system diagrams of the Penn State antenna
 3365 measurement system that depict the important system components and the connections
 3366 between them. The two configurations of the measurement system utilize different
 3367 hardware. For characterization and radiation pattern measurements, one prefers the
 3368 configuration shown in Figure ???. In this case a vector network analyzer (VNA) is used
 3369 as both the transmission source and data acquisition system and it is relatively easy to
 3370 calibrate over a wide range of frequencies. Whereas, if one is more interested in recreating

3371 what would take place in the FSCD experiment then the configuration shown in Figure
3372 ?? is preferable, since this system effectively mimics the receiver chain envisioned for the
3373 FSCD experiment.

3374 The characterization configuration utilizes a network analyzer (Keysight N5222A)
3375 [38, 39] with two independent sources and four measurement ports as the primary
3376 measurement tool. A standard reference antenna is connected to one measurement port,
3377 and the test antenna is connected to a separate port. The typical reference antenna
3378 used for these studies is a Pasternack PF9851 horn antenna [40]. In the measurement
3379 shown, the test antenna represents a SYNCA antenna, which requires a transmission
3380 chain consisting of quadrature hybrid coupler [41, 42] (Marki QH-0226) connected to
3381 two baluns [43] (Marki BAL-0026) to generate feed signals with the appropriate phases.
3382 The VNA measures the radiation pattern by performing a transmission S-parameter
3383 measurement, which can be used with the knowledge of the reference antenna's radiation
3384 pattern to determine the radiation pattern of the test antenna (see Section 5.2.1).

3385 The second configuration is more complicated and incorporates more hardware
3386 components in order to more closely mimic the DAQ system envisioned for the FSCD
3387 experiment. The basic approach is to produce CRES-like radiation and use an antenna
3388 combined with a realistic RF receiver chain to acquire the signals. On the transmit side,
3389 an arbitrary waveform generator [44] (AWG, RIGOL DG5252) is used to generate a
3390 waveform that mimics a CRES signal at a baseband frequency up to 250 MHz. This
3391 frequency is then up-converted to the CRES signal frequency band of 25.8 to 26.0 GHz
3392 using a mixer [45] (Marki MM1-0832L) and a bandpass filter (K&L Microwave 3C62-
3393 25900/T200-K/K) to reject unwanted mixing components outside out of the 200 MHz
3394 CRES signal band. The local oscillator signal for mixing is provided by one of the VNA
3395 sources configured to run in a continuous wave setting. On the receive side, a prototype
3396 antenna is used to detect the radiation emitted by the test antenna, which is down-
3397 converted and filtered using the same mixer and bandpass filter as the transmission chain.
3398 Lastly, data acquisition is performed using a 14-bit ADC sampling at 500 MSa/s [46]
3399 (CAEN DT530) to digitize the down-converted signals.

3400 In order to distribute the LO to all mixers a 4-way power splitter (MiniCircuits
3401 ZC4PD-18263-S+) along with an amplifier (Marki APM-6848) is used to drive the four
3402 mixers used in the measurement system. A limitation of using the VNA as an LO source
3403 is that there is no control of the LO phase when a measurement is triggered by the
3404 control script, which leads to a random phase offset between acquisitions. This makes it
3405 impossible to perform synthetic array measurements, which require strict control over

3406 the starting phase of the transmitted signal. In order to monitor the random phase of the
3407 LO, a 2-way power splitter (MiniCircuits Z99SC-62-S+) is used to split the signal from
3408 the AWG between the transmission path and a LO monitoring path. The LO monitoring
3409 path consists of an up-conversion and down conversion using two mixers connected by a
3410 coaxial cable, and monitors the relative phase of the LO using a channel on the digitizer
3411 to sample this path. A phase shift in the LO will lead to a proportional phase shift in
3412 the mixed signal, which is measured and removed from the received signals.

3413 The test antenna is mounted on a set of motorized stages, which are identical for
3414 both measurement configurations. A rotational stage (ThorLabs PRMTZ8) is used as
3415 the base layer with additional translation stages mounted on top of this. The rotational
3416 stage is ideal for measuring a complete azimuthal scan of the test antenna's radiation
3417 pattern as well as for moving a SYNCA antenna in circular motion to recreate the
3418 symmetry of the FSCD antenna array. On top of the rotational stage we mount two
3419 linear translation stages (ThorLabs MTS50-Z8 and MTS25-Z8) in a cross-wise manner
3420 so that the test antenna can be moved along two perpendicular axes. Using the linear
3421 stages in combination with the rotational stage allows one to fine-tune the positioning of
3422 the test antenna so that it can be perfectly aligned with the central axis of the array.
3423 A LabView script was developed to automate the measurement of a full 360° radiation
3424 pattern and control the measurement electronics. Data from these acquisitions is stored
3425 on university provided cloud storage.

3426 **5.3 Development of a Synthetic Cyclotron Antenna (SYNCA) 3427 for Antenna Array Calibration**

3428 This section is the manuscript of the publication [29] detailing the development of a
3429 Synthetic Cyclotron Antenna (SYNCA) for antenna array characterization measurements
3430 by the Project 8 collaboration.

3431 **5.3.1 Introduction**

3432 Neutrinos are the most abundant standard model fermions in our universe, but due to
3433 weak interaction cross-sections with other particles, neutrinos are particularly difficult
3434 to study. Consequently, many fundamental properties of neutrinos are still unknown
3435 including the absolute scale of the neutrino mass [47]. Direct, kinematic measurements of
3436 the neutrino mass are particularly valuable due to their model independent nature [48].

3437 To date the most sensitive direct neutrino mass measurements have been performed by
 3438 the KATRIN collaboration [49], which measures the molecular tritium β -decay spectrum
 3439 to infer the neutrino mass. Current data from neutrino oscillation measurements [47]
 3440 allow for neutrino masses significantly smaller than the design sensitivity of the KATRIN
 3441 experiment; therefore, there is a need for new technologies for performing direct neutrino
 3442 mass measurements to probe lower neutrino masses.

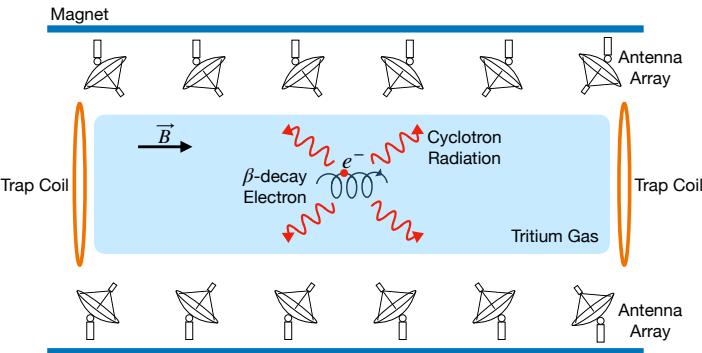


Figure 5.8: A sketch of an antenna array large-volume CRES experiment. Electrons from β -decays are confined in a magnetic field using a set of trap coils. The cyclotron radiation produced by the motion of the trapped electrons can be detected by a surrounding antenna array to determine the electron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

3443 The Project 8 collaboration is developing new methods for neutrino mass measurement
 3444 based on Cyclotron Radiation Emission Spectroscopy (CRES) [50–53], with the goal of
 3445 measuring the absolute scale of the neutrino mass with a 40 meV/c² sensitivity [?, 48].
 3446 This sensitivity goal will require the development of two separate technical capabilities.
 3447 First is the development of an atomic tritium source, which avoids significant spectral
 3448 broadening due to molecular final states [54]. Second is the technology for performing
 3449 CRES in a multi-cubic-meter experimental volume with high combined detection and
 3450 reconstruction efficiency, which is required in order to obtain sufficient event statistics
 3451 near the tritium spectrum endpoint.

3452 One approach for a large-volume CRES experiment is to use an array of antennas,
 3453 which surrounds a volume of tritium gas, to detect the cyclotron radiation produced
 3454 by the β -decay electrons when they are trapped in a background magnetic field using a
 3455 set of magnetic trapping coils (see Figure 5.8). Project 8 has developed a conceptual
 3456 experiment design to study the feasibility of this approach. The design consists of a
 3457 single circular array of antennas with a radius of 10 cm and 60 independent channels
 3458 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 [28, 55], 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.

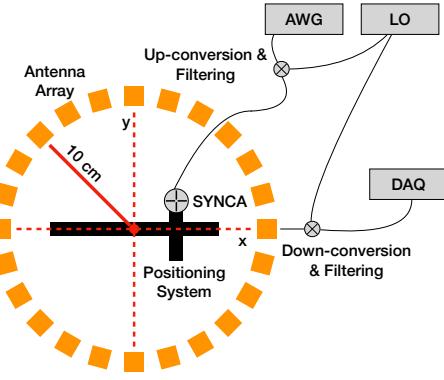


Figure 5.9: A schematic of the antenna array test stand. The circular antenna array has a radius of 10 cm with 60 independent channels (limited number shown for clarity). The test stand includes an arbitrary waveform generator (AWG), local oscillator (LO), and data acquisition (DAQ) hardware. Finally, 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 reconstruc-
 tion 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

3480 the fields generated by the SYNCA match simulation, and finally in Section 5.3.5 we
 3481 demonstrate an application of the SYNCA to test phased array reconstruction techniques
 3482 on the bench-top.

3483 5.3.2 Cyclotron Radiation Phenomenology

3484 To understand the cyclotron radiation phenomenology that the SYNCA should mimic,
 3485 we consider a charged particle moving at relativistic speed in the presence of an external
 3486 magnetic field (see Figure 5.10). In the special case we shall examine, the entirety of
 3487 the electron's momentum is directed perpendicular to the magnetic field; therefore, the
 3488 trajectory of the electron is confined to the cyclotron orbit plane. Because the momentum
 3489 vector is oriented perpendicular to the magnetic field, electrons with these trajectories
 3490 are said to have pitch angles of 90°.

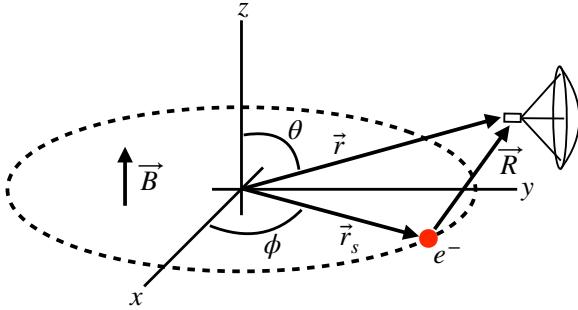


Figure 5.10: An electron (red dot) performing cyclotron motion in the x-y plane. The resulting cyclotron radiation is observed by an antenna located at the field point of interest.

3491 The cyclotron radiation fields generated by this circular trajectory are those which
 3492 we aim to reproduce with the SYNCA. We can describe the electromagnetic (EM) fields
 3493 using the Liénard-Wiechert equations [8, 28], which in non-covariant form express the
 3494 electric field as

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

3495 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
 3496 electron to the field measurement point, $\vec{\beta} = \dot{\vec{r}}_s/c$ is the velocity of the particle divided
 3497 by the speed of light, and γ is the relativistic Lorentz factor. The equation is meant to
 3498 be evaluated at the retarded time as indicated by $t_r = t - |\vec{R}|/c$, which accounts for the

3499 time delay due to the finite speed of light between the point where the field was emitted
3500 and the point where the field is detected.

3501 We would like to simplify Equation 5.23 it at all possible. As a first step we analyze
3502 the relative magnitudes of the electric field polarization components. Consider an electron
3503 following a circular cyclotron orbit in a uniform magnetic field whose guiding center
3504 is positioned at the origin of the coordinate system. The equation of motion can be
3505 expressed as

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

3506 For single antenna located along the y-axis at position $\vec{r} = r_a \hat{y}$ we are interested in the
3507 incident electric fields from the electron. The electric field is given by Equation 5.23,
3508 which we evaluate in the regime where $r_a \gg r_c$. This limit can be justified by comparing
3509 the radius of the cyclotron orbit for an electron with the tritium beta-spectrum endpoint
3510 energy of 18.6 keV in a 1 T magnetic field to the typical ($r_a \simeq 100$ mm) radial position
3511 of the receiving antenna. We find that the cyclotron orbit has a radius of 0.46 mm which
3512 is approximately a factor of 200 smaller than the typical antenna radial position. In this
3513 regime we can make the approximation $\vec{R} \simeq r_a \hat{y}$ and the expression for the electric field
3514 at the antenna's position becomes

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

3515 Since the receiving antenna is part of a circular array of antennas, it is useful to rewrite
3516 Equation 5.25 in terms of the azimuthal ($\hat{\phi}$) and radial (\hat{r}) polarizations. Making use of
3517 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

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

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

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

3518 For the purposes of designing a synthetic cyclotron radiation antenna we are interested
3519 in the dominant electric field polarization emitted by the electron. The antenna is being
3520 designed to mimic the cyclotron radiation produced by electrons with kinetic energies of
3521 approximately 18.6 keV in a 1 T magnetic field [54]. Since the relativistic beta factor for

3522 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
 3523 justified. Inserting these expressions into the equations for the electric field components
 3524 above simplifies the comparison of the magnitudes of the two components. Additionally,
 3525 we compare the time-averaged magnitudes to evaluate the root mean squared electric
 3526 field ratio. The time-averaged ratio of the radial and azimuthally polarized electric fields
 3527 with the above simplifications is given by

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

3528 where we have made use of the fact that for these magnetic fields and kinetic energies
 3529 the cyclotron radius is much smaller than the radius of the antenna array.

3530 From Equation 5.29 we see that the time-averaged azimuthal polarization is larger than
 3531 the radial polarization by about a factor of 8, which makes it the dominant contribution
 3532 to the electric fields at the position of the antenna. We must also consider the directivity
 3533 of the receiving antenna which can have a gain that is disproportionately large for a
 3534 specific polarization component. Because the E_ϕ component is dominant, the receiving
 3535 antenna array is designed with an azimuthal polarization, which negates the voltages
 3536 induced in the antenna from the radially polarized fields. Therefore, we conclude that
 3537 for the purpose of designing the SYNCA antenna it is acceptable to approximate the
 3538 electric fields from Equation 5.23 as purely azimuthally or ϕ -polarized. The simplified
 3539 expression for the electric field received by an antenna becomes

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

3540 where the radius of the cyclotron orbit is called r_c , the cyclotron frequency is called ω_c ,
 3541 and the radial position of the receiving antenna is called r_a . Equation 5.30 has been
 3542 evaluated in the non-relativistic limit where $\gamma \simeq 1$, which is justified by the fact that
 3543 $|\vec{\beta}| \simeq \frac{c}{4}$ for an electron with an 18.6 keV kinetic energy in a 1 T magnetic field.

3544 This rather complicated expression can be simplified using Fourier analysis. Assuming
 3545 a background magnetic field of 1 T and a kinetic energy of 18.6 keV we calculate
 3546 numerically the electric field using Equation 5.30 and apply a discrete Fourier Transform
 3547 to visualize the frequency spectrum (see Figure 5.11).

3548 We observe that the azimuthally polarized electric field is periodic with a base cyclotron
 3549 frequency of 25.898 GHz corresponding to the highest power frequency component in

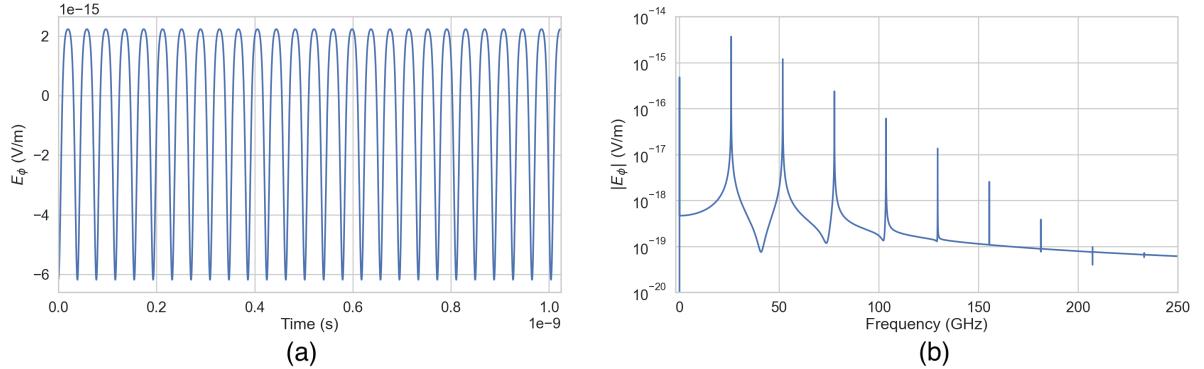


Figure 5.11: A plot of the numeric solution to Equation 5.31. The time-domain representation of the signal (a) is composed of a zero frequency term and a series of harmonics separated by the main cyclotron frequency as shown in the plot of the frequency spectrum (b). We can see that the relative amplitude of the harmonics beyond $k = 7$ are smaller than the main carrier by a factor of about 10^{-5} and are completely negligible.

3550 Figure 5.11. The frequency spectrum reveals that the signal is composed of a constant
 3551 term with zero frequency and a series of harmonics separated by 25.898 GHz. Therefore,
 3552 we can represent the azimuthal electric fields from the electron as a linear combination
 3553 of pure sinusoids with frequencies given by $\omega_k = k\omega_c$ ($k \in 0, 1, 2, \dots$) and amplitudes
 3554 extracted from the Fourier representation. Using this representation we can transform
 3555 the equation for the azimuthally polarized electric fields in Equation 5.30 into

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

3556 where we have truncated the sum over harmonics at the 7th order for completeness. The
 3557 amplitudes A_k are dimensionless complex numbers, which encode the relative powers of
 3558 the harmonics as well as the starting overall phase of the cyclotron radiation. Because
 3559 magnitude of the relative amplitudes exponentially decreases for higher harmonics, it is
 3560 usually sufficient to consider only the terms up to $k = 4$ where the relative amplitude
 3561 of the harmonics has decreased from the main carrier by a factor of approximately 100.
 3562 However, for completeness we include harmonics up to 7th order in Equation 5.31. The
 3563 range of frequencies to which the receiving antenna array in the antenna test stand is
 3564 sensitive is defined by the antenna's transfer function. The receptive bandwidth for
 3565 the antennas used in the test stand is a range of frequencies with a bandwidth on the
 3566 order of a few GHz centered around the main cyclotron carrier frequency of 25.898 GHz.
 3567 Therefore, the higher order harmonics as well as the zero frequency term can be ignored

3568 when considering only the signals that will be received by the antenna array.

3569 Considering only the 1st order harmonic term from Equation 5.31, which represents
3570 the portion of the electric field that will be detected by the array, and evaluating this at
3571 the retarded time we obtain the following for the ϕ -polarized electric fields

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

3572 where the arbitrary phase Δ is defined by $A_k = |A_k|e^{i\Delta}$. We are interested in the
3573 characteristics of the amplitude of the electric field as a function of the radial distance
3574 component ($|\vec{R}|$) of the retarded time. In particular, the maximum of E_ϕ occurs when
3575 the argument of the cosine function is equal $n\pi$ where $n \in \{0, \pm 2, \pm 4, \dots\}$; however, the
3576 solutions where n is negative can be discarded since they represent unphysical negative
3577 overall phases. Applying this condition to Equation 5.32 gives a condition on the radial
3578 position of the maximum of E_ϕ

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

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

3579 which is a function of time in the frame of the moving electron (t). Equation 5.33 can
3580 be further simplified by noticing that the azimuthal position of the electron ($\phi_e(t)$) as a
3581 function of time is defined by $\phi_e(t) = \omega_c t - \Delta$ which reduces Equation 5.33 to

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

3582 Equation 5.34 represents an archimedean spiral which is formed when plotting the
3583 amplitude of E_ϕ in the x-y plane. The solution where $n = 0$ represents the leading edge
3584 of the radiation spiral which propagates outward from the electron at the speed of light.
3585 The additional solutions for $n > 0$ represent the persistent spiral at radii inside the
3586 leading edge of the radiated fields that have not yet been detected by the receiver at the
3587 current time. In Figure 5.12a we show the expected spiral pattern for the maxima of the
3588 cyclotron radiation.

3589 In particular, we note that for the circular array geometry of the test stand, depicted
3590 as the series of circles in Figure 5.12a, each antenna receives a linearly polarized wave
3591 with a phase offset that corresponds to the azimuthal angle for that antenna element.
3592 Therefore, as we show in Figure 5.12b, when the relative phase of the received signal is
3593 plotted as a function of the receiving antenna's azimuthal position the result is also an

3594 Archimedean spiral.

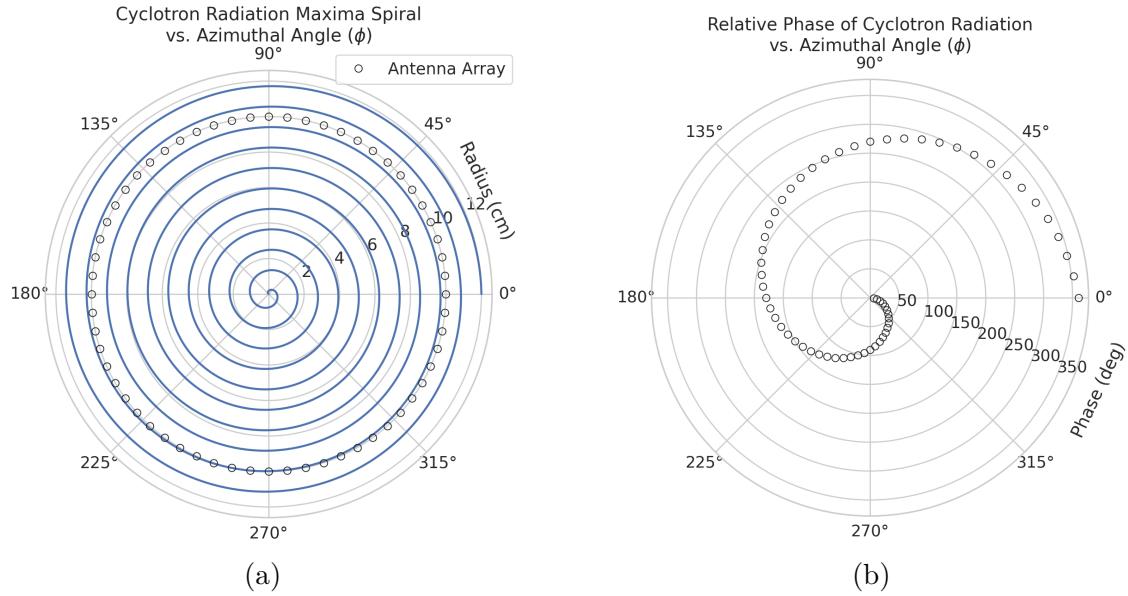


Figure 5.12: The amplitude maxima of the cyclotron radiation form an Archimedean spiral as the radiation propagates outward from the cyclotron orbit center (a). A circular antenna array located at a fixed radius from the orbit center will receive electric fields with equal magnitude in each of its channels, but the phase of the electric field incident on each array channel will be linearly out of phase from its neighbor antennas by an amount equal to the angular separation of the two channels (b).

3595 Based on these analytical calculations we can characterize the magnitude, polarization,
3596 and phase of the signals received by the antenna array using three criteria. These criteria
3597 are the basis of comparison for the radiation produced by the SYNCA and cyclotron
3598 radiation emitted by electrons and will be used to evaluate the performance of antenna
3599 designs. The criteria are:

- 3600 1. Electric fields that are ϕ -polarized near $\theta = 90^\circ$
3601 2. Uniform time-averaged electric field magnitudes around the circumference of a
3602 circle centered on the antenna
3603 3. Electric fields whose phase is equal to the azimuthal angle at the point of measure-
3604 ment plus a constant

3605 The Locust simulation package [55] can be used to directly simulate the EM fields
3606 generated by electrons performing cyclotron motion to validate the analytical calculations.
3607 Locust simulates the EM fields by first calculating the trajectory of the electrons in

3608 the magnetic trap using the Kassiopeia software package [56]. The trajectory can then
 3609 be used to solve for the EM fields using the Liénard-Wiechert equations directly with
 3610 no approximations. The resulting electric field solutions drive a receiving antenna by
 3611 convolving the time-domain fields with the finite-impulse response filter of the antenna
 3612 or they can be examined directly to study the field characteristics that the SYNCA must
 3613 reproduce. In the next section we compare the radiation field patterns for electrons
 3614 simulated with Locust to patterns from a SYNCA antenna design.

3615 5.3.3 SYNCA Simulations and Design

3616 One potential SYNCA design is the crossed-dipole antenna [57]. A crossed-dipole antenna
 3617 consists of two dipole antennas, one of which is rotated 90° with respect to the other,
 3618 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

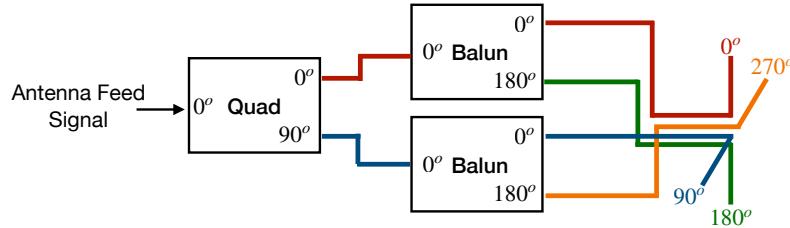


Figure 5.13: An idealized crossed-dipole antenna consists of two electric dipole antennas oriented perpendicular to each other and is fed with four signals with a quadrature phase relationship. An example antenna feed circuit is shown which is composed of a chained combination of a quadrature hybrid-coupler (Quad) and two baluns.

3619
 3620 out of phase from each of the neighboring arms by 90°, which mirrors the spatial phase
 3621 relationship of cyclotron radiation fields.

3622 A potential drawback of this design is that standard crossed-dipole antennas do not
 3623 radiate uniform electric fields near the $\theta = \pi/2$ plane. Typical crossed-dipole antennas
 3624 use dipole arm lengths equal to $\lambda/4$ or larger [57], where λ is the wavelength at the
 3625 desired operating frequency. Such large arm lengths cause the electric field magnitude
 3626 to vary significantly around the circumference of the antenna. However, making the
 3627 antenna electrically small by shrinking the arm length can improve the antenna pattern
 3628 uniformity.

3629 In general, the criterion for an electrically small antenna is that the largest dimension
3630 of the antenna (D) obey $D \lesssim \lambda/10$ [13]. In our application, we are attempting to mimic
3631 the cyclotron radiation emitted by electrons produced from tritium β -decay with energies
3632 near the spectrum endpoint. For a background magnetic field of 1 T, the corresponding
3633 cyclotron frequency of tritium endpoint electrons is approximately 26 GHz. Therefore, the
3634 electrically small condition would require that the largest dimension of the crossed-dipole
3635 antenna be smaller than 1.2 mm.

3636 A crossed-dipole antenna with an overall size of 1.2 mm is challenging to fabricate due
3637 to the small dimensions of the dipole arms that, in practice, are fragile and unsuitable
3638 for use as a calibration probe. To mitigate some of the challenges with the fabrication
3639 of such a small antenna, a variant crossed-dipole antenna design using printed circuit
3640 board (PCB) technology (see Figure 5.14) was developed in partnership with an antenna
prototyping company, Field Theory Consulting ¹.

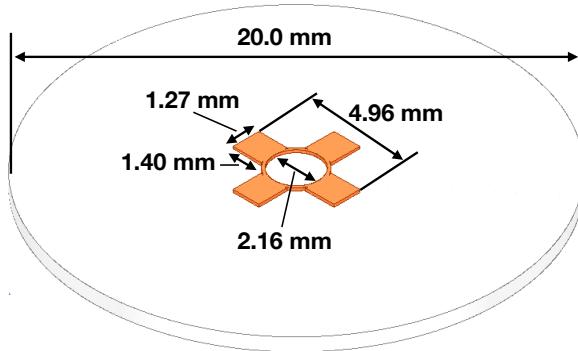


Figure 5.14: A model of the PCB crossed-dipole antenna with dimensions. The design has an inside diameter of 2.16 mm for the central circular trace, which is 0.13 mm wide. The dipole arms each have a width of 1.27 mm and protrude beyond the circular trace by 1.40 mm, which gives an overall width of 4.96 mm for the length of the antenna PCB trace from end-to-end. The overall size of the antenna is 20.0 mm the majority of which is the PCB dielectric material. This design was observed in simulation to maintain the field characteristics of the idealized crossed-dipole while being simpler to fabricate due to the increased size of the antenna.

3641
3642 The PCB crossed-dipole design uses four rectangular pads to represent the dipole arms,
3643 which are connected by a thin circular trace. The circular trace both adds mechanical
3644 stability to the antenna and improves the azimuthal uniformity of the electric fields
3645 compared to a more standard crossed-dipole geometry. Furthermore, the circular trace
3646 allows for a greater separation between dipole arms than standard crossed-dipoles, which

¹<https://fieldtheoryinc.com/>

3647 is required to accommodate the coaxial connections to each pad. The pads each contain
 3648 a through-hole solder joint to connect coaxial transmission lines using hand soldering.
 3649 The antenna PCB has no ground plane on the bottom layer as this was observed in
 3650 simulation to significantly distort the radiation pattern in the plane of the PCB. The
 3651 only ground planes present in the model are the outer conductors of the four coaxial
 3652 transmission lines which feed the antenna. These are left unterminated on the bottom of
 3653 the PCB dielectric material.

3654 The antenna design development utilized a combination of Locust electron simula-
 3655 tions and antenna simulations using ANSYS HFSS [14], a commercial finite-element
 3656 electromagnetic simulation software. Two antenna designs were simulated: an idealized
 3657 electrically small crossed-dipole antenna with an arm length of 0.40 mm and an arm
 3658 separation of 0.05 mm, as well as a PCB crossed-dipole antenna with the dimensions
 3659 shown in Figure 5.14. Plotting the magnitude of the electric fields generated by the
 3660 antennas across a 10 cm square located in the same plane as the respective antennas
 3661 reveals the expected cyclotron spiral pattern (see Figure 5.15) which closely matches
 3662 the prediction for simulated electrons. The spiral pattern demonstrates that the electric
 3663 fields have the appropriate phases to mimic cyclotron radiation, which fulfills SYNCA
 criterion 3 identified in Section 5.3.2.

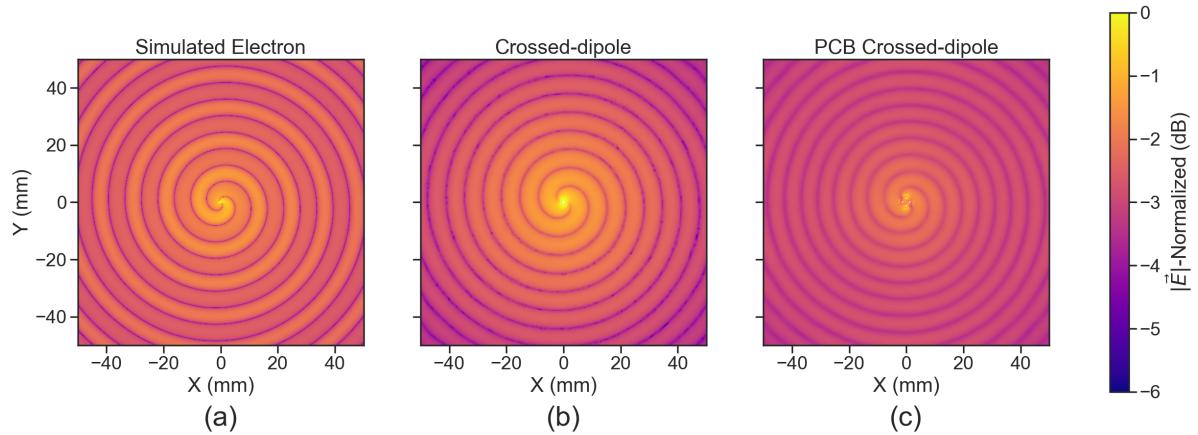


Figure 5.15: A comparison of the electric field magnitudes, normalized by the maximum value of the electric field in each simulation, plotted on a 10 cm square to visualize the Archimedean spirals formed by the electron (a), the crossed-dipole antenna (b), and a PCB crossed-dipole antenna (c). The matching 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.

3664

3665 As we can see from Figure 5.16, the crossed-dipole antenna, which uses an idealized

3666 geometry, exhibits good agreement with simulation. The antenna has a maximum
 3667 deviation from a simulated electron of approximately 0.5 dB in the total electric field, 1
 3668 dB for the ϕ -polarized electric field and 1 dB for the θ -polarized electric field.

3669 In comparison, the pattern of the PCB crossed-dipole antenna, because the simulation
 3670 incorporates the geometry of the coax transmission lines, exhibits some distortion from
 3671 the idealized cross-dipole simulations. The vertically oriented ground planes of the coax
 3672 lines introduce more θ -polarized electric fields than are observed for simulated electrons
 3673 near $\theta = 90^\circ$. The significant θ -polarized field minimum is still present but shifted
 to approximately $\theta = 65^\circ$. The θ -polarized field deviations of the PCB crossed-dipole

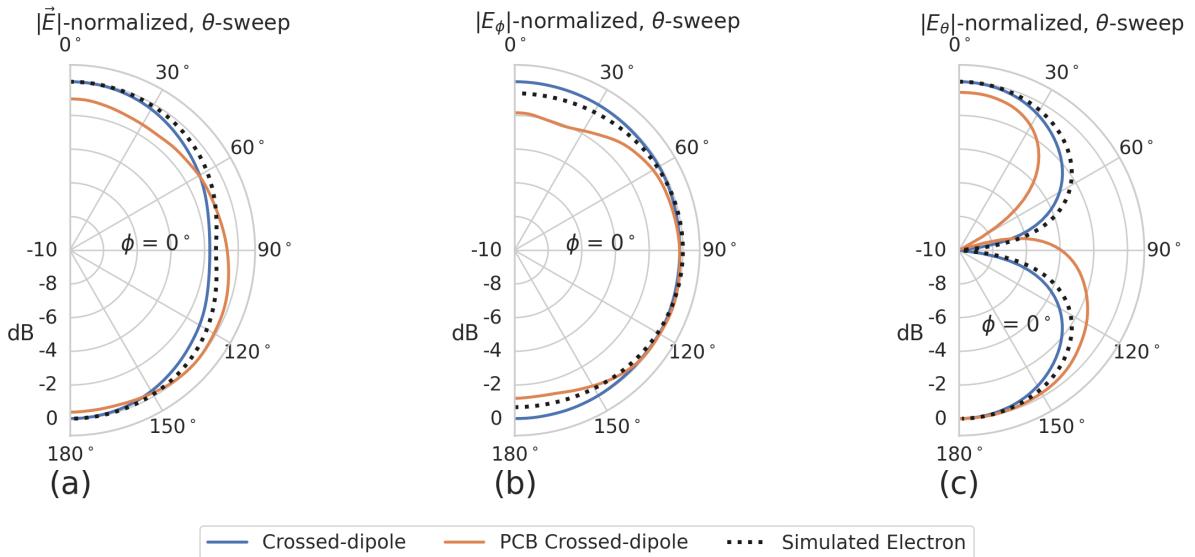


Figure 5.16: A comparison of the normalized electric field magnitudes for the ideal crossed-dipole, PCB crossed-dipole, and a simulated electron as a function of the polar angle (θ). (a) Shows the total electric field, (b) shows the ϕ -polarized electric field component, and (c) shows the θ -polarized electric field component. These images were generated using Locust simulations for the electron and ANSYS HFSS for both antennas.

3674
 3675 antenna should not greatly impact the performance of the antenna because the receiving
 3676 antenna array is primarily ϕ -polarized. Therefore deviations in the θ -polarized fields
 3677 will be suppressed due to the polarization mismatch. More importantly, the ϕ -polarized
 3678 electric field pattern generated by the PCB crossed-dipole closely matches simulated
 3679 electrons across the polar angle range of $50^\circ < \theta < 150^\circ$. In this region the PCB crossed-
 3680 dipole differs by less than 0.5 dB from simulated electrons. This range greatly exceeds
 3681 the beamwidth of the receiving antenna array which is designed to be most sensitive
 3682 to fields produced near $\theta = 90^\circ$. Therefore, we conclude that the PCB crossed-dipole

3683 antenna generates a ϕ -polarized radiation pattern that fulfills SYNCA criterion 1 from
3684 Section 5.3.2.

3685 The final SYNCA criterion is related to the uniformity of the electric fields when
3686 measured azimuthally around the antenna. As we saw for real electrons in Section 5.3.2
3687 it is expected that the magnitude of the electric field be completely uniform as a function
3688 of the azimuthal angle due to the symmetry of the cyclotron orbit. In Figure 5.17 we plot
3689 the total electric field as a function of azimuthal angle for an electron, the crossed-dipole
antenna, and the PCB crossed-dipole antenna. The crossed-dipole antenna exhibits

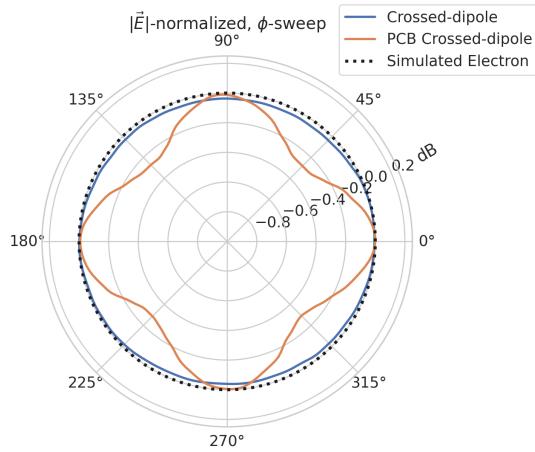


Figure 5.17: A comparison of the normalized electric field magnitudes for the crossed-dipole, PCB crossed-dipole, and a simulated electron as a function of the azimuthal angle (ϕ) evaluated at $\theta = 90^\circ$. This image was generated using Locust simulations for the electron and ANSYS HFSS for both antennas.

3690
3691 perfect uniformity around the azimuthal angle, whereas the PCB crossed-dipole has a
3692 small periodic deviation with a maximum difference of 0.3 dB caused by the coaxial
3693 transmission lines below the PCB. Such a small deviation from uniformity is acceptable
3694 since it is smaller than the expected variation in uniformity caused by imperfections in
3695 the antenna fabrication process, which modifies the antenna shape in an uncontrolled
3696 manner by introducing solder blobs with a typical size of a few tenths of a millimeter on
3697 the dipole arms (see Figure 5.18). Additionally, the SYNCA will be separately calibrated
3698 to account for azimuthal differences in the electric field magnitude. Therefore we see
3699 from the simulated performance of the PCB crossed-dipole antenna that this antenna
3700 design meets all three of the SYNCA criteria.

5.3.4 Characterization of the SYNCA

Two SYNCAs were manufactured using the PCB crossed-dipole design (see Figure 5.18). The antenna PCB (Matrix Circuit Board Materials, MEGTRON 6) is connected to four 2.92 mm coaxial connectors (Fairview Microwave, SC5843) using semi-rigid coax (Fairview Microwave, FMBC002), which also physically support the antenna PCB. The antenna PCB consists only of two layers which correspond to the copper antenna trace and the PCB dielectric. Each coax line is connected to the associated dipole arm using through-hole soldering and phase matched to ensure that the electrical length of each of the transmission lines is identical at the operating frequency. The antenna PCB is further reinforced using custom cut polystyrene foam blocks, which have an electrical permittivity nearly identical to air. A custom 3D printed mount is included at the base of the antenna to support the coax connectors and to provide a sturdy mounting base.

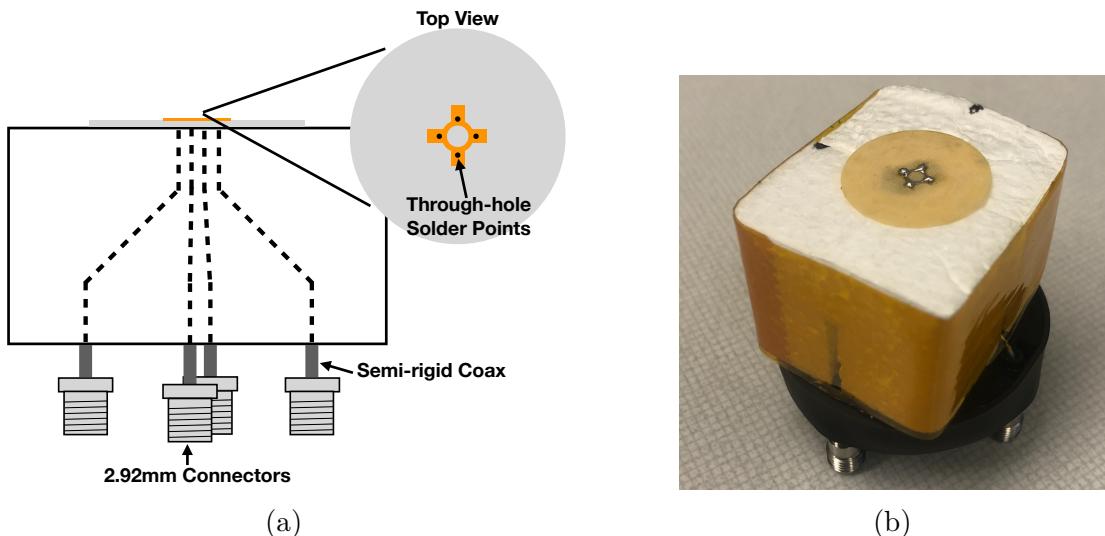


Figure 5.18: (a) A cartoon schematic which highlights the routing of the semi-rigid coax transmission lines. (b) A photograph of a SYNCA constructed using the modified crossed-dipole PCB antenna design. Visible in the photograph of the SYNCA are four blobs of solder which are an artifact of the SYNCA's 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.

Characterization measurements were performed using a Vector Network Analyzer (VNA) to measure the electric field magnitude and phase radiated by the SYNCA to verify the radiation pattern (see Figure 5.19). The VNA is connected to the SYNCA

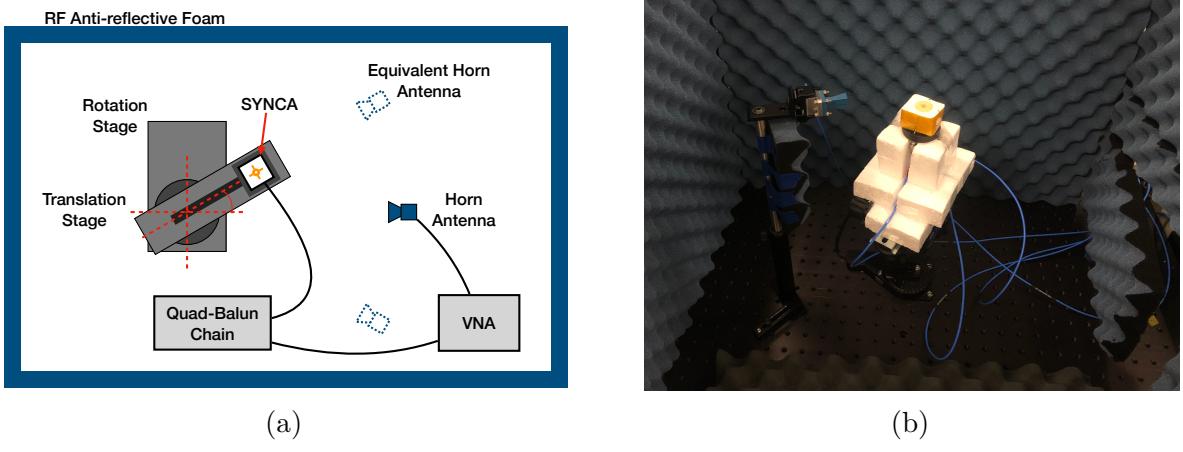


Figure 5.19: A schematic of the VNA characterization measurements (a). This setup allows for antenna gain and phase measurements across a full 360° of azimuthal angles using a motorized rotation stage and control of the radial position of the SYNCA using a translation stage. A photo of the setup in the lab is shown in (b).

3716 at one port through a hybrid-coupler whose outputs are connected to two baluns to
 3717 generate the signals with the appropriate phases to feed the SYNCA (see Figure 5.13).
 3718 The other port of the VNA is connected to a single reference horn antenna that serves
 3719 as a field probe. To position the SYNCA, a combination of translation and rotation
 3720 stages are used to characterize the antenna's fields across the entire radiation pattern
 3721 circumference. This measurement scheme is equivalent to measuring the fields generated
 3722 by the SYNCA using a full circular array of probe antennas.

3723 The antenna measurement space is surrounded by RF anti-reflective foam to isolate
 3724 the measurements from the lab environment (see Figure 5.19b) and remaining reflections
 3725 are removed using the VNA's time-gating feature. The SYNCA is affixed to the stages
 3726 by a custom RF transparent mount made of polystyrene foam. The coaxial cables deliver
 3727 the antenna feed signals generated by the VNA to the SYNCA while still allowing
 3728 unrestricted rotation. The horn antenna probe is nominally positioned in the plane
 3729 formed by the antenna PCB ($\theta = 90^\circ$ or $z = 0$ mm) at a distance of 10 cm from the
 3730 SYNCA, to match the expected position of the antenna array relative to the SYNCA in
 3731 the antenna array test stand. The horn antenna can be manually raised or lowered to
 3732 different relative vertical positions to characterize the radiation pattern at different polar
 3733 angles.

3734 Several 360° scans were performed with probe vertical offsets of -10.0 mm, -5.0 mm,
 3735 0.0 mm, 5.0 mm, and 10.0 mm relative to the antenna PCB plane. These probe offsets

3736 cover a 2 cm wide vertical region centered on the SYNCA PCB, approximately equal to
 3737 ± 6 degrees of polar angle. The measurements show that the SYNCA is generating fields
 3738 with nearly isotropic magnitude across the probed region. The standard deviation of the
 3739 electric field magnitude measured around the antenna circumference is approximately
 3740 2.9 dB for a typical rotational scan. The presence of a significant pattern null is noted
 3741 near 45° (see Figure 5.20), which we attribute to small imperfections in the antenna
 3742 PCB that could be introduced from the hand soldered terminations connecting the coax
 3743 cables to the antenna. There is no significant difference in the radiation pattern when
 3744 measured across the 2 cm vertical range. The measured relative phases closely follow
 3745 the expectation for an electron, being linear with the measurement rotation angle and
 3746 forming the expected spiral pattern. Other than the small phase imperfections there is
 3747 a slight sinusoidal bias to the phase data, which we determined is the result of a small
 3748 ($\lesssim 1$ mm) offset of the antenna's phase center from the rotation axis of the automated
 3749 stages.

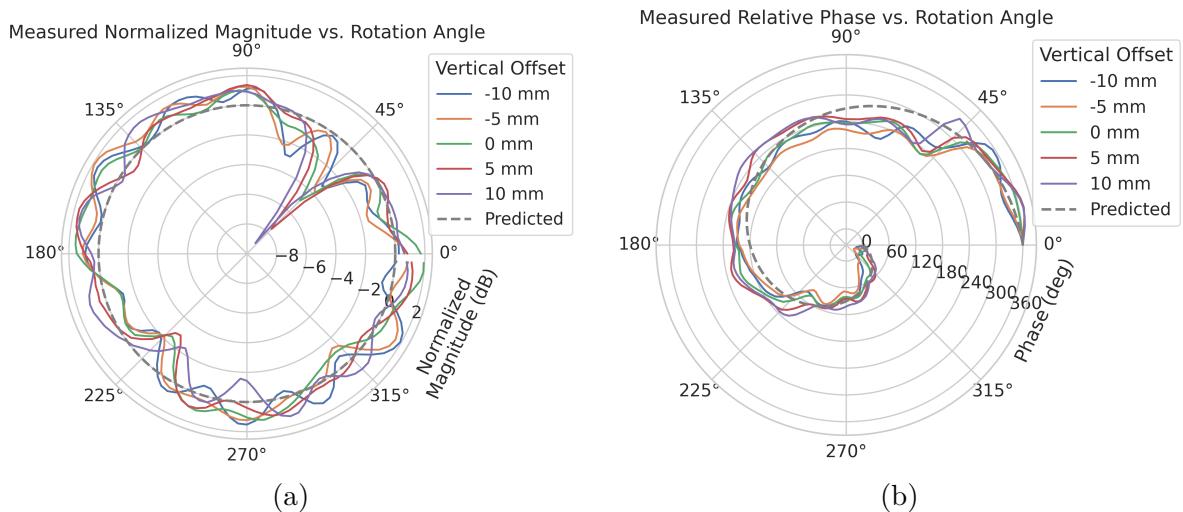


Figure 5.20: Linear interpolations of the measured electric field magnitude (a) and phase (b). The data was acquired using a VNA at 120 points spaced by 3 degrees from 0 to 357 degrees of azimuthal angle. The different color lines indicate the vertical offset of the horn antenna relative to the SYNCA 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.

3750 The characterization measurements confirm the simulated performance of the SYNCA.
 3751 As expected the fields generated by the antenna are nearly isotropic in magnitude, ϕ -
 3752 polarized, and are linearly out of phase around the circumference of the antenna as

3753 predicted for cyclotron radiation in Section 5.3.2. Small imperfections in the magnitude
 3754 and phase of the antenna are expected, particularly at the antenna's high operating
 3755 frequency of 26 GHz where small geometric changes can have significant impacts on
 3756 electrical properties. However, calibration through careful characterization measurements
 3757 can be used to remove the majority of these pattern imperfections, including the relatively
 3758 large pattern null near 45°, which will allow for the usage of the SYNCA as a test source
 3759 for free-space CRES experiments utilizing antenna arrays. In the next section we use the
 3760 VNA measurements obtained here as a calibration for signal reconstruction using digital
 3761 beamforming.

3762 **5.3.5 Beamforming Measurements with the SYNCA**

3763 Digital beamforming is a standard technique for signal reconstruction using a phased
 3764 array [58]. The SYNCA, since it exhibits the same cyclotron phases as a trapped electron,
 3765 can be used to perform simulated CRES digital beamforming reconstruction experiments
 3766 on the bench-top without the need for the magnet, cryogenics, and vacuum systems
 3767 required by a full CRES experiment. The fields received by the individual elements
 3768 of the antenna array will have phases dependent on the spatial position of the source
 3769 relative to the antennas. Therefore, a simple summation of the received signals will fail
 3770 to reconstruct the signal due to destructive interference between the individual channels
 3771 in the array. However, applying a phase shift associated with the source's spatial position

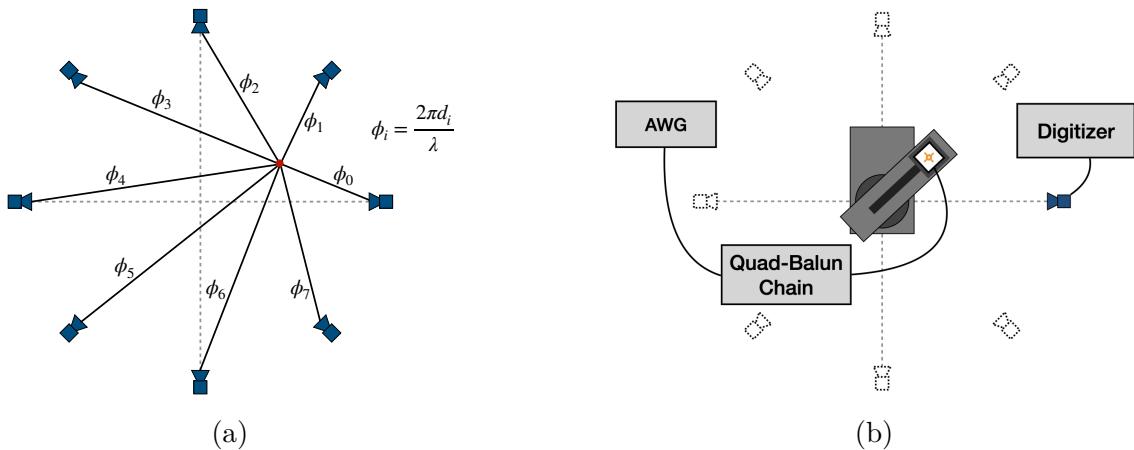


Figure 5.21: (a) A depiction of the relative phase differences for signals received by a circular antenna array from an isotropic source. The phases correspond to a unique spatial position. (b) A schematic of the setup used to perform digital beamforming.

3772 removes phase differences and results in a constructive summation of the channel signals
3773 (see Figure 5.21). We can summarize the digital beamforming operation succinctly using
3774 the following equation

$$y[t_n] = \sum_{m=0}^{N-1} x_m[t_n] A_m e^{i\phi_m}, \quad (5.35)$$

3775 where $y[t_n]$ represents the summed array signal at time t_n , $x_m[t_n]$ is the signal received
3776 by channel m at time t_n , ϕ_m is the phase shift applied to the signal received at channel
3777 m , and A_m is an amplitude weighting factor that accounts for the different signal power
3778 received by individual channels. By changing the digital beamforming phases, the point
3779 of constructive interference can be scanned across the sensitive region of the array to
3780 search for the location of a radiating source, which is identified as the point of maximum
3781 summed signal power above a specified threshold. The digital beamforming phases consist
3782 of two components,

$$\phi_m = 2\pi d_m / \lambda + \theta_m, \quad (5.36)$$

3783 where d_m is the distance from the m -th array element to the source, and θ_m is the
3784 relative angle between the source position and the m -th antenna. The first component is
3785 the standard digital beamforming phase that corresponds to the spatial position of the
3786 source, and the second component is the cyclotron phase that corresponds to the relative
3787 azimuthal phase offset.

3788 With a small modification to the hardware used to characterize the SYNCA (see
3789 Figure 5.19), we can perform a digital beamforming reconstruction of a synthetic CRES
3790 event. By replacing the VNA with an arbitrary waveform generator (AWG), the SYNCA
3791 can be used to generate cyclotron radiation with an arbitrary signal structure, which
3792 can then be detected by digitizing the signals received by the horn antenna. Rotational
3793 symmetry allows us to use the rotational stage of the positioning system to rotate the
3794 SYNCA to recreate the signals that would have been received by a complete circular
3795 array of antennas.

3796 Using this setup, signals from a 60 channel circular array of equally spaced horn
3797 antennas were generated with the SYNCA positioned 10 mm off the central array axis,
3798 reconstructed using digital beamforming, and compared to Locust simulation (see Figure
3799 5.22). When the cyclotron spiral phases are not used, which is equivalent to setting θ_m
3800 in Equation 5.36 to zero, the SYNCA's position is reconstructed as a relatively faint ring
3801 as predicted by simulation. However, when the appropriate cyclotron phases are used
3802 during the beamforming procedure, both the simulated electron and the SYNCA appear

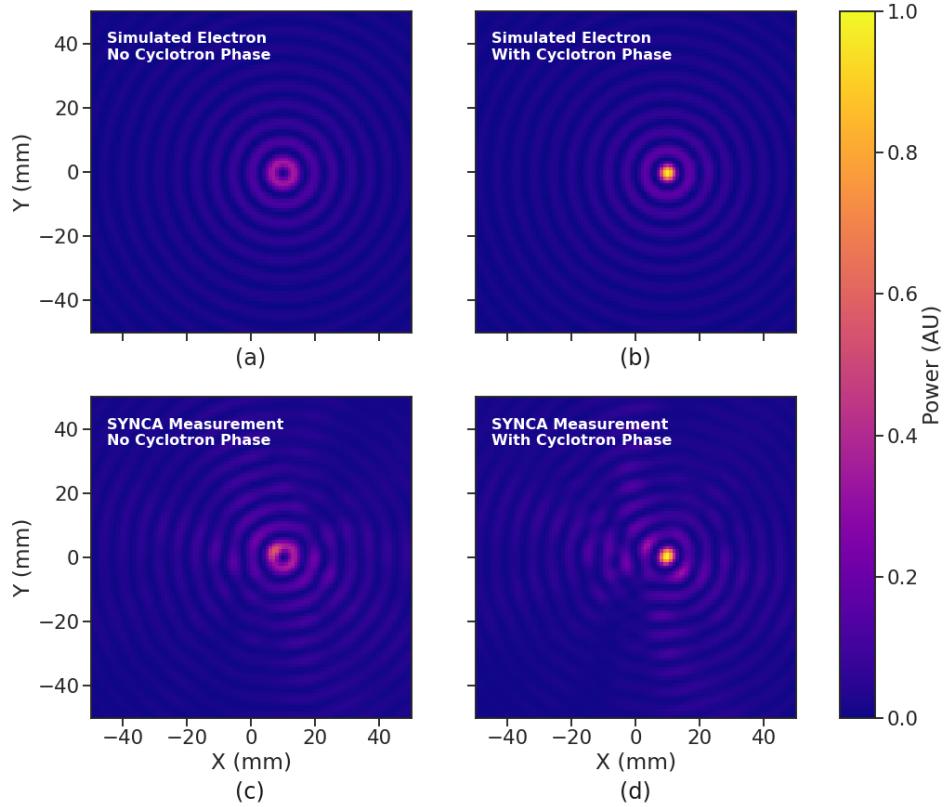


Figure 5.22: Digital beamforming maps generated using a simulated 60 channel array and electron simulated using the Locust package. (a) and (b) show the beamforming maps for simulated electrons without the cyclotron spiral phases and with the cyclotron spiral phases respectively. (c) and (d) show the beamforming maps produced from SYNCA measurements. We observe good agreement between simulated electrons and the SYNCA measurements.

as a single peak of high relative power corresponding to the source position. Therefore, we observe good agreement between the simulated and SYNCA reconstructions. While it may seem that for the case with no cyclotron phase corrections the ring reconstructs the position of the electron as effectively as beamforming with the cyclotron phase corrections, it is important to note that the simulations and measurements were generated without a realistic level of thermal noise. The larger maxima region and lower signal power, which occurs without the cyclotron phase corrections, significantly reduce the probability of detecting an electron in a realistic noise background.

To bound the beamforming capabilities of the synthetic array of horn antennas, we performed a series of beamforming reconstructions where the SYNCA was progressively moved off the central axis of the array (see Figure 5.23). To extract an estimate of the

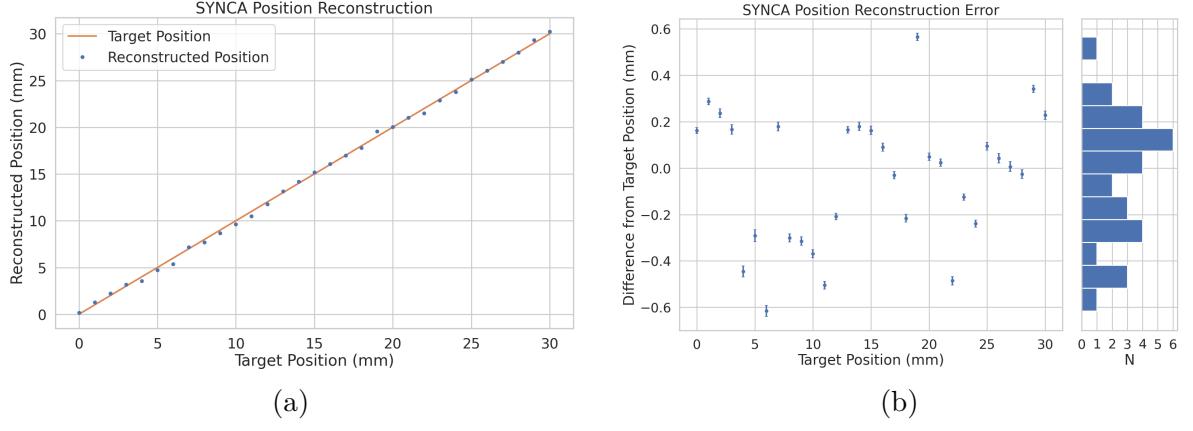


Figure 5.23: A plot of the SYNCA’s reconstructed position using the synthesized horn-antenna array and digital beamforming. (a) Shows the reconstructed position of the SYNCA compared with the target position indicated by the positioning system readout. (b) Shows the reconstruction error, which is the 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.

position of the SYNCA using the digital beamforming image we apply a 2-dimensional (2D) Gaussian fit to the image data and extract the estimated centroid value. We find that the synthetic horn antenna array reconstructs the position of the SYNCA with a 1σ -error of 0.3 mm with no apparent trend across the 30 mm measurement range. This reconstruction error is an order of magnitude larger than mean fit position uncertainty of 0.02 mm indicating that systematic effects related to the SYNCA positioning system could be contributing additional uncertainty to the measurements. Note that the current mean reconstruction error of 0.3 mm is a factor of 20 smaller than the full width at half maximum of the digital beamforming peak (6 mm), which could be interpreted as a naive estimate of the position reconstruction performance of this technique. Because these measurements are intended as a proof-of-principle demonstration, we do not investigate potential sources of systematic errors further; however, we expect that a similar and more thorough investigation will be performed using the Project 8 antenna array test stand, where typical reconstruction errors can be used to estimate the energy resolution limits of antenna array designs.

3829 **5.3.6 Conclusions**

3830 In this paper we have introduced the SYNCA, which is a novel antenna design that
3831 emits radiation that mimics the unique properties of the cyclotron radiation generated by
3832 charged particles moving in a magnetic field. The characterization measurements of the
3833 SYNCA validated the simulated performance of the PCB crossed-dipole antenna design.
3834 Additionally, the SYNCA was used to estimate the position reconstruction capabilities
3835 of a synthesized array of horn antennas and experimentally reproduced the simulated
3836 digital beamforming reconstruction of electrons.

3837 While the SYNCA performs well, there exist discrepancies in the phase and magnitude
3838 of the radiation pattern compared to the simulated SYNCA design that are related to
3839 the small geometric differences in the soldered connections. Future design iterations that
3840 replace the soldered connections with a fully surface mount design could improve the
3841 radiation pattern at the cost of some complexity and expense. Furthermore, improving
3842 the design of the antenna PCB and mounting system would allow the antenna to be
3843 inserted into a cryogenic and vacuum environment where in-situ antenna measurement
3844 calibrations could be performed.

3845 The discrepancies in the radiation pattern and phases exhibited by the as-built
3846 SYNCA should not greatly impact its performance as a calibration probe. Both magni-
3847 tude and phase variations can be accounted by applying the SYNCA characterization
3848 measurements as a calibration to the data collected by the antenna array test stand. The
3849 separate calibration of the SYNCA radiation does not impact the primary goals for the
3850 antenna array test stand which are array calibration and signal reconstruction algorithm
3851 performance characterization, because it can be performed with standard reference horn
3852 antennas with well understood characteristics.

3853 The SYNCA antenna technology advances the CRES technique by providing a
3854 mechanism to characterize free-space antenna arrays for CRES measurements without
3855 the need for a magnet and cryogenics system, which would be required for calibration
3856 using electron sources. Both the Project 8 collaboration as well as future collaborations
3857 which are developing antenna array based CRES experiments can make use of SYNCA
3858 antennas as an important component of their calibration and commissioning phases.

3859 5.4 FSCD Antenna Array Measurements with the SYNCA

3860 5.4.1 Introduction

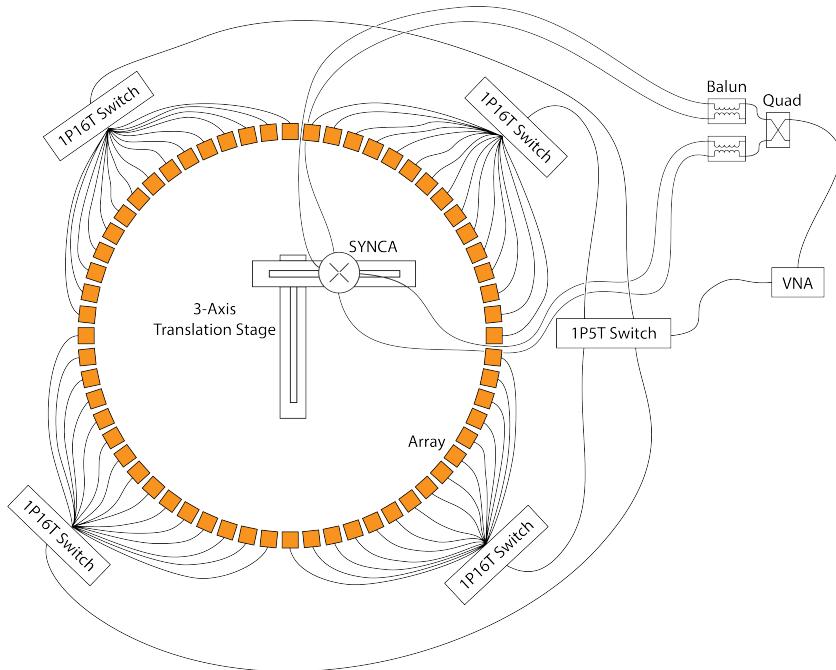


Figure 5.24: A diagram of the array measurement system used to test the prototype FSCD antenna array. A VNA is used as the primary measurement tool, which is connected to the array through a series of switches. The other port of the VNA connects to the SYNCA through the quad-balun chain used to provide 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.

3861 Using the SYNCA we can perform full-array measurements of prototype versions
3862 of the FSCD antenna array to test its performance with a realistic cyclotron radiation
3863 source (see Figure 5.24). The goal is to check how the measured power received by
3864 the array compares to FSCD simulations as a function of the radial and axial position
3865 of the SYNCA. These measurements are intended to validate the antenna research
3866 and development by Project 8, which has been driven primarily by simulations with
3867 Locust [10] and CREsana (see Section 4.2.3), and identify any discrepancies with these
3868 simulations tools. This knowledge will provide confidence in the simulations necessary
3869 for the analysis of the sensitivity of larger antenna array based CRES experiment designs
3870 to the neutrino mass.

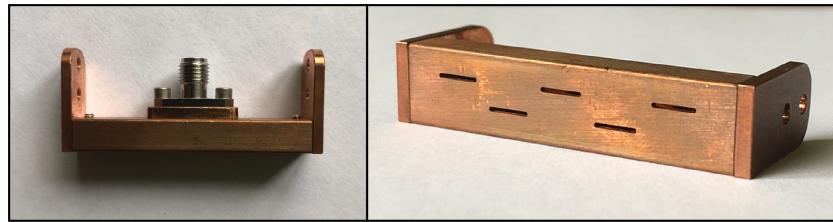
3871 As shown in Section 5.3, the SYNCA does have some radiation pattern imperfections
3872 that complicate the comparison between measurement and simulation data. One way to
3873 disentangle some of the effects of these imperfections is to perform an additional set of
3874 measurements using a synthetic antenna array setup along with the SYNCA antenna.
3875 Since the synthetic array setup uses only a single array antenna, the data should be
3876 free of errors associated with individual antenna differences and multi-path interference,
3877 which are two error sources being tested with the full-array setup. By comparing the
3878 synthetic array data to the FSCD array data and to simulation data one can evaluate the
3879 significance of these effects relative to the errors introduced by SYNCA imperfections.

3880 **5.4.2 Measurement Setups**

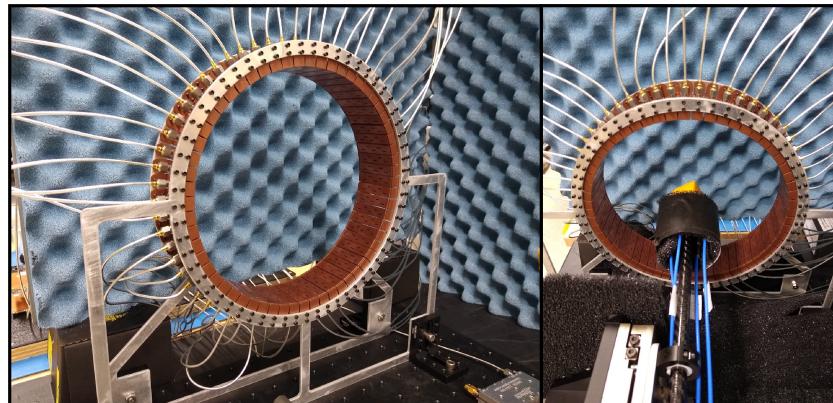
3881 **5.4.2.1 FSCD Array Setup**

3882 The antenna design that composes the array is the 5-slot waveguide antenna developed
3883 for the FSCD experiment (see Figure 5.25a). The antenna is 5 cm long and is constructed
3884 out of WR-34 waveguide with a 2.92 mm coax connector located at the center of the
3885 antenna. Copper flanges located on both ends of the antenna are used to mount the
3886 antenna in the array support structure. The antennas are supported by two circular steel
3887 brackets that can be bolted to both ends of the waveguide to construct the circular array
3888 (see Figure 5.25b). The antenna array consists of sixty identical waveguide antennas
3889 with a radius of 10 cm. The array is mounted perpendicular to an optical breadboard
3890 surface using a pair of the steel brackets, which provide sufficient space for the coaxial
3891 cable connections and allows for easy positioning of the SYNCA antenna. The SYNCA is
3892 mounted on the end of a carbon fiber rod attached to a set of manual translation stages,
3893 which are used to move the SYNCA antenna to different positions inside the array (see
3894 Figure 5.25c). The stages allow for independent motion in three different axes and can
3895 position the SYNCA at radial distances up to 5 cm from the center.

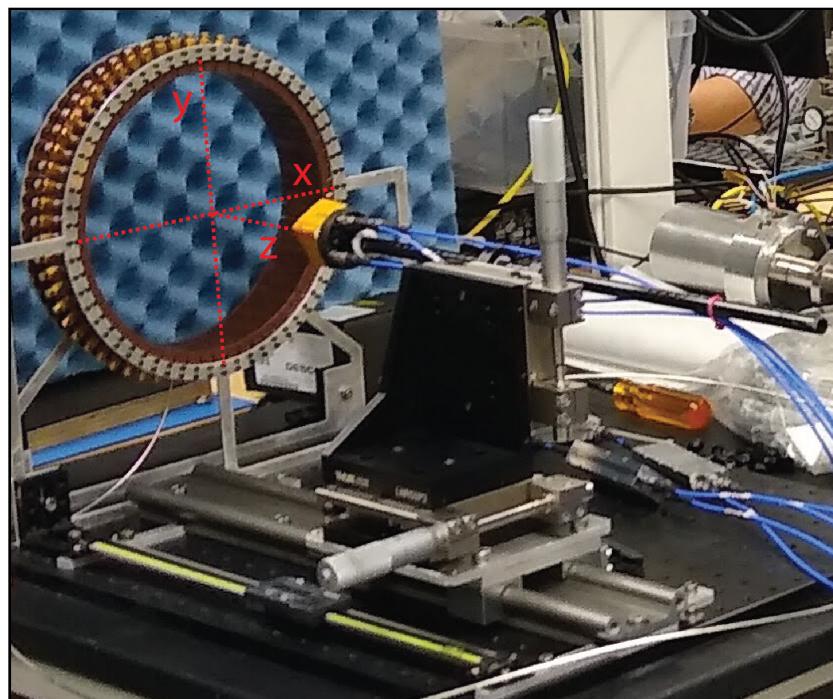
3896 Data acquisition is accomplished using a two-port VNA in combination with a series
3897 of microwave switches that allow the VNA to connect to each channel in the array . The
3898 first port of the VNA is connected to the quad-balun chain used to feed the SYNCA (see
3899 Section 5.3), and the second port of the VNA connects to a 1P5T microwave switch. The
3900 1P5T switch is connected to four separate 1P16T switch boards that connect directly
3901 to the array. The data acquisition is controlled by a python script running on a lab
3902 computer, which is connected to the VNA and an Arduino board programmed to control
3903 the microwave switches. The script uses the switches to iteratively connect each of the



(a)



(b)



(c)

Figure 5.25: 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).

3904 antennas in the array to the VNA. The VNA is configured to load a specific calibration
3905 file for each antenna channel and performs the measurements of all available S-parameters.
3906 The separate calibration files is an attempt to remove phase and magnitude errors caused
3907 by different propagation through the RF switches. Array measurements were performed
3908 for the set of SYNCA positions consisting of radial (x-axis) positions from 0 to 50 mm in
3909 5 mm steps and axial (z-axis) positions from 0 to 50 mm in 5 mm steps resulting in 121
3910 array measurements. At each SYNCA position we measured the two-port S-parameter
3911 matrix using a linear frequency sweep from 25.1 to 26.1 GHz with 101 discrete frequencies.

3912 5.4.2.2 Synthetic Array Setup

3913 A photograph of the setup used to perform the synthetic array measurements is shown
3914 in Figure 5.26. One important difference between this setup and the FSCD array setup
3915 is that the synthetic array measurements were performed with a waveform generator and
3916 digitizer instead of a VNA. The electronics configuration is identical to the diagram in
3917 Figure 5.7b. Despite the differences, one is still able to compare the measured phases of
3918 the synthetic array and the relative magnitude of the power, since the digitized signal
3919 power is directly proportional to S21.

3920 The arbitrary waveform generator in the setup is configured to produce a 64 MHz
3921 sine wave signal that is up-converted to 25.864 GHz using a mixer and the VNA source.
3922 This signal is passed through a bandpass filter and fed to the SYNCA quad-balun chain.
3923 A single FSCD antenna is positioned 10 cm from the SYNCA and aligned vertically so
3924 that the center of the 5-slot waveguide is in the plane of the SYNCA PCB (see Figure
3925 5.26). This position corresponds to $z = 0$ in Figure 5.25c. The SYNCA is rotated
3926 in three degree steps to synthesize an antenna array with 120 channels. This channel
3927 count is more than could physically fit in a 10 cm radius array, but there is no cost to
3928 over-sampling. Additionally, over-sampling allows for a check of the smoothness of the
3929 antenna array radiation pattern. The signals from the FSCD antenna are down-converted
3930 using the second mixer connected to the VNA source before being digitized at 250 MHz
3931 and saved to disk. Several synthetic array measurement scans were performed by using
3932 the linear translation stage to change the radial position of the SYNCA. In total eight
3933 scans were taken from 0 to 35 mm using a radial position step size of 5 mm.

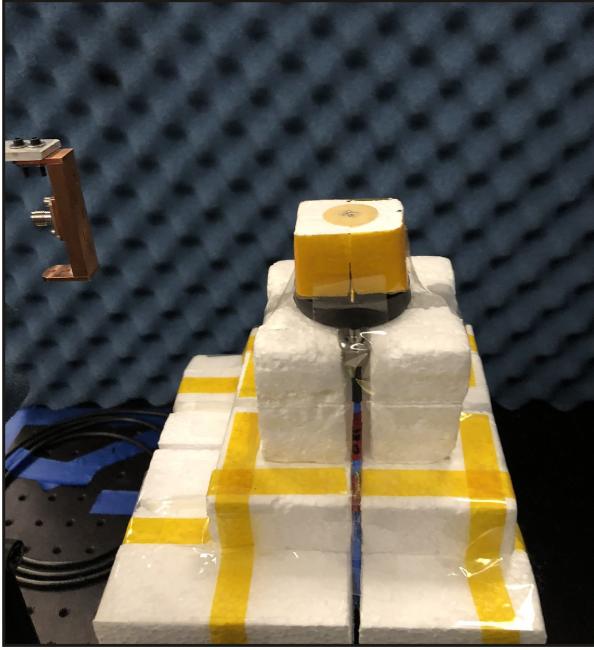


Figure 5.26: A photo of the FSCD antenna and the SYNCA in the synthetic array measurement setup at Penn State.

5.4.3 Simulations, Analysis, and Results

The Locust and CRESana simulation packages utilize the antenna transfer functions to calculate the power that would be received by each antenna from a CRES electron. The equivalent quantity in the measurement setup is the S21 matrix element, which indicates the ratio of the power received by an antenna in the array to the amount of power delivered to the SYNCA. Therefore, the analysis focuses on comparing the relative magnitudes and phase of the S21 parameters measured by the VNA as a function of the array channel and the SYNCA position. Additionally, we apply a beamforming reconstruction to the S21 data to evaluate how the summed power and beamforming images change as a function of the position of the SYNCA.

5.4.3.1 Simulations

Simulations for the FSCD array measurements were performed using CRESana, which performs analytical calculations of the EM-fields produced by an electron at the position of the antennas. At each sampled time CRESana computes the electric field vector at the antenna positions, which is projected onto the antenna polarization axis to obtain the co-polar electric field. The magnitude of the co-polar electric field is then multiplied by

3950 a flat antenna transfer function to calculate the corresponding voltage signal. CRESana
 3951 simulations exploit the flat transfer functions of the FSCD antennas, which allows the
 3952 electric field to be multiplied by the antenna transfer function rather than performing
 3953 the full FIR calculation. These calculations produce a voltage time-series for each of the
 3954 antennas in the array that can be compared to the laboratory measurements.

3955 CRESana was configured to simulate a 90° electron in a constant background magnetic
 3956 field of ≈ 0.958 T with a kinetic energy of 18.6 keV. These parameters were chosen
 3957 in order to mimic a CRES event near the tritium beta-decay spectrum endpoint in
 3958 the FSCD experiment. The constant background magnetic field guarantees that the
 3959 guiding center of the electron is stationary across the duration of the simulation which is
 3960 consistent with the SYNCA in the laboratory measurements. Simulations were performed
 3961 with the electron's guiding center at radial positions from 0 to 45 mm in steps of 1 mm
 3962 and axial positions from 0 to 30 mm in steps of 1 mm. The simulations generated time
 3963 series consisting of 8192 samples at 200 MHz for the sixty channel FSCD antenna array
 3964 geometry.

3965 5.4.3.2 Phase Analysis

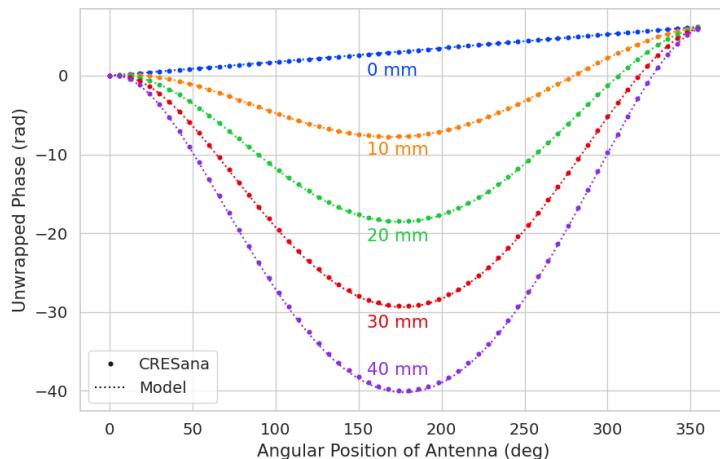


Figure 5.27: The unwrapped phases of signals received by the FSCD antenna array from an electron with a 90° pitch angle located in the plane of the antenna array. The data points indicated the phases extracted from simulation and the dashed lines show the model predictions.

3966 Correct modeling of the signal phases is fundamental to reconstruction for both
 3967 beamforming and matched filter approaches. The beamforming reconstruction relies on

3968 a signal phase model developed from Locust simulations, which allows one to predict the
3969 relative signal phases for a specific magnetic trap and electron position. The equation
3970 for the model is

$$\phi_{ij}(t) = \frac{2\pi d_{ij}(t)}{\lambda} + \theta_{ij}(t), \quad (5.37)$$

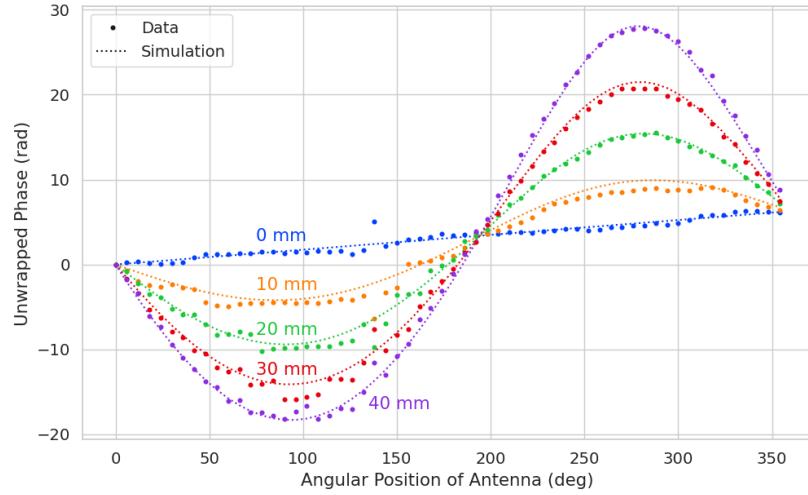
3971 where $d_{ij}(t)$ is distance between the assumed electron position and the antenna position,
3972 and $\theta_{ij}(t)$ is the angular separation between the electron and antenna positions. For
3973 details on the components of the phase model see Section 5.3.2. In Figure 5.27 we
3974 compare the phases predicted by Equation 5.37 to phases extracted from CREsana
3975 simulations of an electron located in the plane of the antenna array at a series of radial
3976 positions. One observes excellent agreement between the model and simulation.

3977 The measured signal phases from the FSCD array and synthetic array are shown
3978 in Figures 5.28a and 5.28b compared to the signal phase model. The axial position of
3979 the SYNCA in both plots is $z = 0$ mm, such that the plane of the PCB is aligned with
3980 the center of the FSCD antenna. The data shown in Figure 5.28a corresponds to the
3981 S-parameters measured at 25.80 GHz which is the frequency closest to the one used in
3982 the synthetic array setup. The different slope and sinusoidal phases exhibited by Figure
3983 5.28a and 5.28b reflects differences in the coordinate system for each setup. In general,
3984 we see that the phase model predicts the large scale features of the phases quite well,
3985 but there are some small scale deviations or errors from the phase model that do not
3986 appear to be present in simulation.

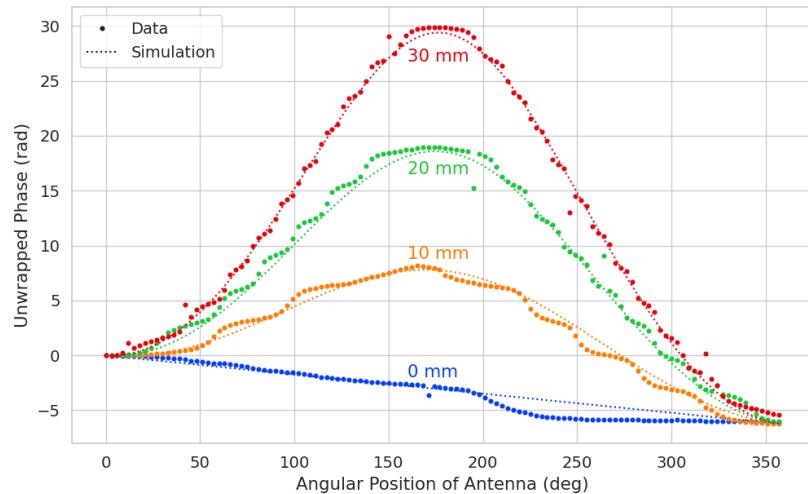
3987 A comparison of the phase errors, which are the difference between measurement and
3988 model is shown in Figure 5.29. The FSCD array data is referred to as the JUGAAD
3989 data in the plot legend, which is an alternative name for the FSCD array setup.

3990 The phase error at $R = 0$ in Figure 5.29 forms a smooth curve, with the exception of
3991 an outlier data point caused by a bug in the data acquisition script. One can attribute
3992 the observed phase error at this position to imperfections in the antenna pattern of the
3993 SYNCA. As the SYNCA is moved away from $R = 0$ mm one observes that the phase
3994 error exhibits oscillations whose frequency increases as a function of the radial position
3995 of the SYNCA. These oscillations have the appearance of a diffraction pattern, which
3996 is particularly clear for the radii ≥ 15 mm, due to the bilateral symmetry of the phase
3997 error peaks around 180° .

3998 One can observe a higher average variance in the phase errors measured for the FSCD
3999 array compared to the synthetic array. This is best seen by comparing the curves at
4000 $R \leq 15$ mm where the smooth synthetic array curves are distinct from the relatively
4001 noisy FSCD array errors. The extra noise in the FSCD array is most likely caused by



(a)



(b)

Figure 5.28: 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.

4002 differences in the radiation patterns of the antennas that make up the array as well as
 4003 differences in the transmission lines through the switch network that introduce additional
 4004 phase errors into the measurement. Since the synthetic array measurements use only
 4005 a single antenna, these extra error terms are not present, which explains the relatively
 4006 smoother phase error curves. Despite the extra phase errors in the FSCD array, it is still
 4007 possible to observe a similar phase error oscillation effect as the SYNCA is moved away

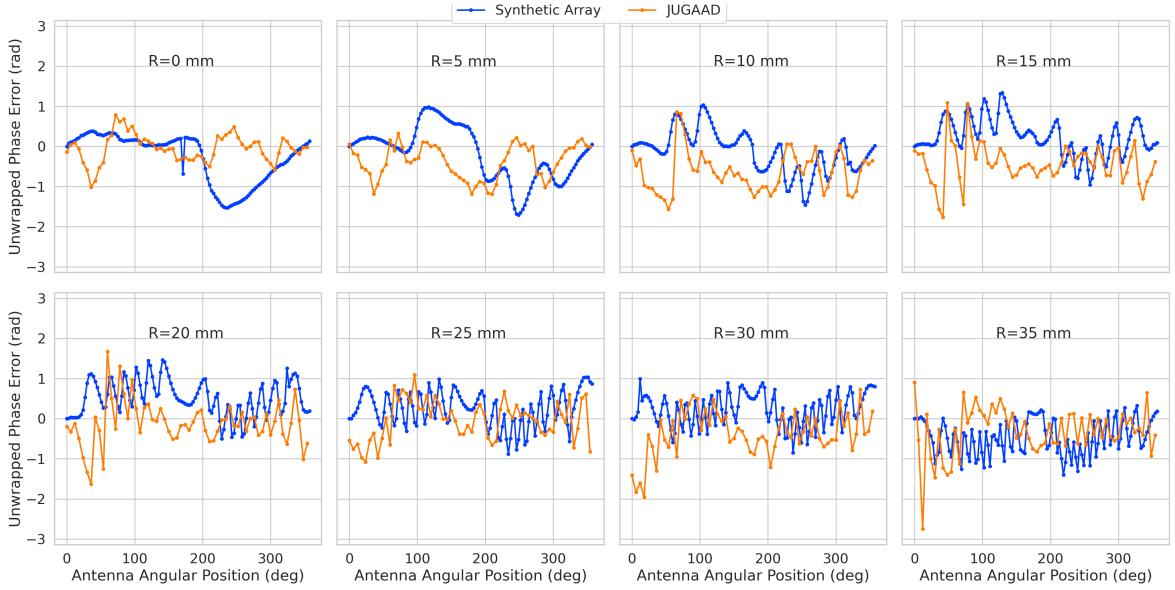
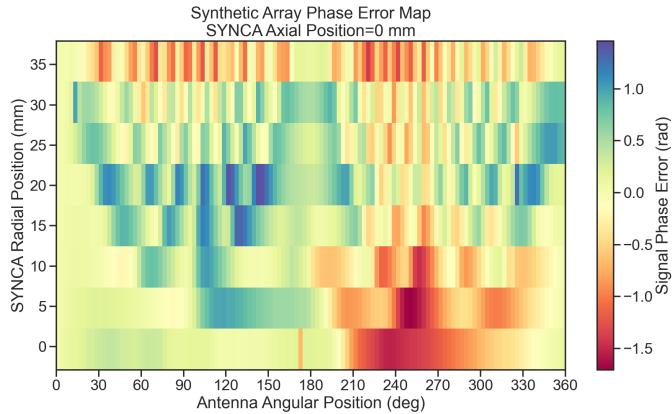


Figure 5.29: The phase errors between the measurement and model for the synthetic array (blue) and the FSCD array (orange) for a series of radial positions. The label JUGAAD refers to an alternative name for the FSCD array setup. As the SYNCA is translated off-axis phase errors with progressively higher oscillation frequency enter into the measurements.

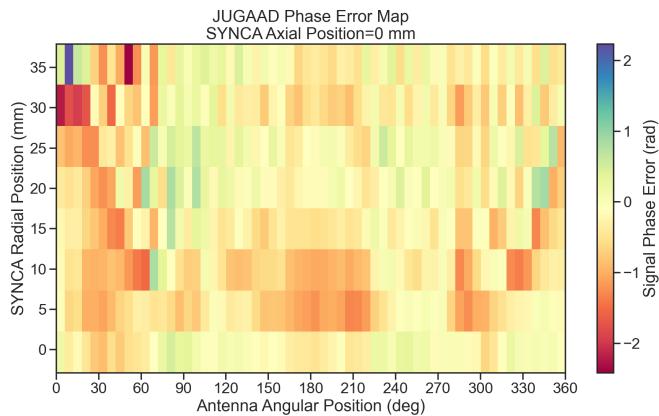
4008 from $R = 0$ mm.

4009 The diffraction pattern exhibited by the phase error oscillations is more easily observed
 4010 by plotting the phase errors in a two-dimensional map, which is done in Figures 5.30a and
 4011 5.30b. For the synthetic array ones observes a relatively clear diffraction pattern
 4012 that emerges as the SYNCA is moved radially. The bilateral symmetry of the diffraction
 4013 patterns is due to the bilateral symmetry of the circular synthetic array around the
 4014 translation axis of the SYNCA. A similar pattern is also visible in the FSCD array data,
 4015 although, it is obscured by the additional phase error that results from the multi-channel
 4016 array.

4017 The physical origin of the phase error diffraction pattern is attributed to interference
 4018 effects arising from path-length differences between the individual slots in the FSCD
 4019 antenna and the SYNCA transmitter. Since we are operating in the radiative near-field of
 4020 the FSCD antenna, the path length differences between the slots introduces a significant
 4021 change in the summation of the signals that occurs inside the waveguide, which causes
 4022 the radiation pattern of the antenna to change as a function of distance. Therefore, when
 4023 the SYNCA is positioned off-axis the different path-lengths from the SYNCA to each
 4024 antenna results in different radiation patterns leading to the observed diffraction pattern.



(a)



(b)

Figure 5.30: 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.

4025 This near-field effect is not present in simulations, because in order to simplify the
 4026 calculations we assume that the far-field approximation can be applied to the FSCD
 4027 antennas. This means that the radiation pattern and antenna transfer functions are
 4028 independent of the distance between the transmitter and the receiving antenna. In
 4029 principle, we can account for these near-field effects with a more detailed simulation of
 4030 the FSCD antennas either in CRESana or Locust, which would result in an additional
 4031 term in the beamforming phase model. However, this would increase the computational
 4032 intensity of the simulation software. In the next section we briefly discuss the impact of

4033 these near-field effects on the measured magnitude of the signals.

4034 **5.4.3.3 Magnitude Analysis**

4035 Exactly as for the signal phase, one can use simulations to construct a model that
4036 describes the magnitude of the signals received by each channel in the antenna array.
4037 By examining the results of simulations or by analyzing the Liénard-Wiechert equation
4038 one can show that radiation pattern from a 90° pitch angle electron in a magnetic field
4039 is omni-directional. Therefore the relative magnitudes of the signals received by each
4040 channel will be determined by the free-space power loss, which is proportional to the
4041 inverse distance between the assumed electron position and the antenna.

4042 A consequence of this is that the signals produced in the array for electrons off the
4043 central axis will have larger amplitudes for the antennas closer to the electron compared
4044 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 5.31.

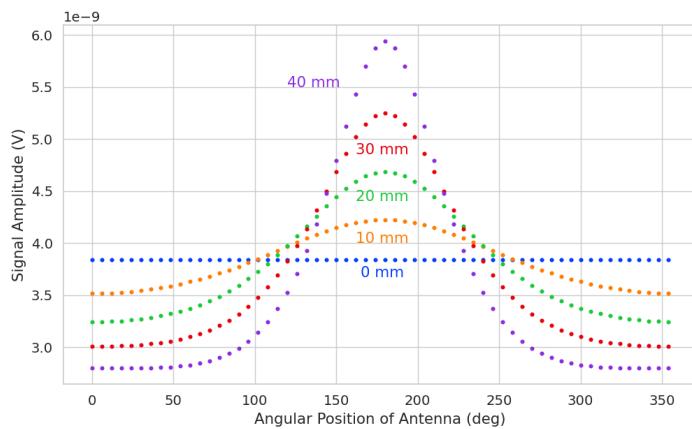


Figure 5.31: The amplitude of the signals from CREsana for the FSCD array from a 90° electron. As the electron is moved from $R = 0$ the signals begin to have unequal amplitudes depending on the distance from the electron to the antenna.

4045
4046 One expects to see a similar trend in the signal magnitudes in both the FSCD and
4047 synthetic arrays. The normalized signal magnitudes extracted from the full and synthetic
4048 array setups for a series of radial SYNCA positions are shown in Figure 5.32. The data
4049 corresponds to a SYNCA axial position of $z = 0$ mm and at a frequency 25.86 GHz. One
4050 complication is that the radiation pattern of the SYNCA is not perfectly omni-directional,
4051 which causes the measured magnitudes at $R = 0$ mm to diverge from the perfectly flat
4052 behavior exhibited by electrons.

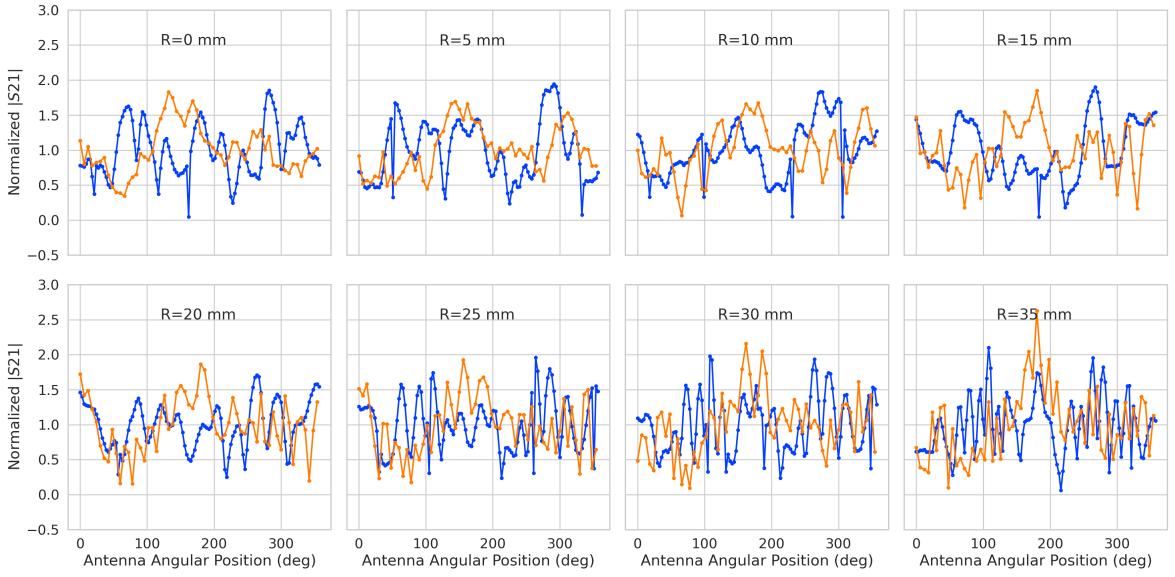
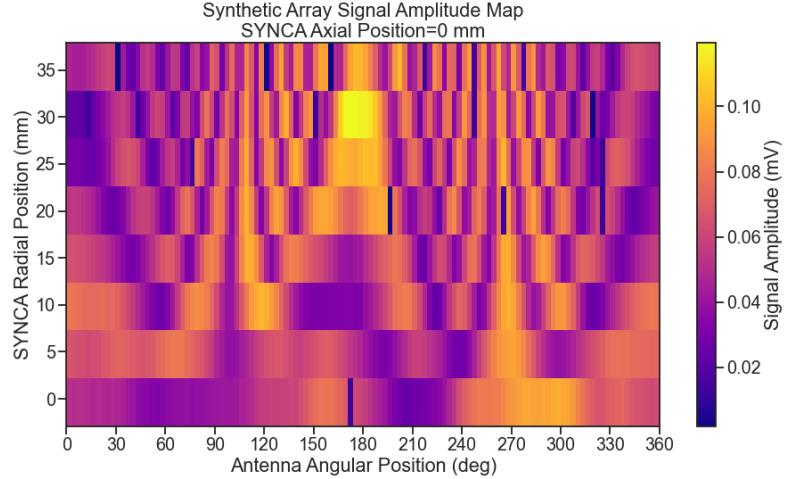


Figure 5.32: The normalized magnitudes of the S21 parameters measured in the FSCD (orange) and synthetic array (blue) setups. The dominant observed behavior as a function of radius is the increase in the number of magnitude peaks, which was noted in the phase error curves. There does not appear to be a strong change in the relative amplitude of a group of antennas as predicted by CREsana.

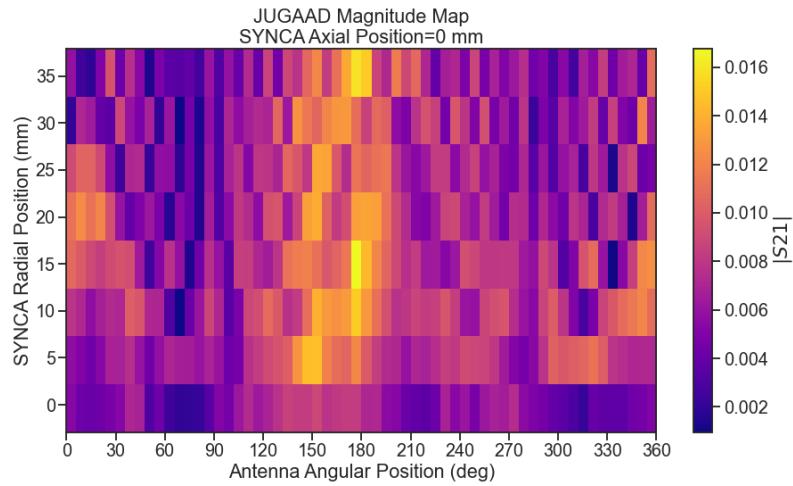
As the SYNCA is moved off-axis one observes a similar increase in the number of magnitude peaks in the synthetic array data that one would expect from a diffraction pattern, although this trend is not as stark compared to the phase data. Noticeably, there does not appear to be a set of channels with 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 5.32 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 5.33).

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



(a)



(b) The two-dimensional maps showing the diffractive pattern exhibited by the FSCD and synthetic array signal magnitudes.

Figure 5.33

radiation pattern of the FSCD antenna as a function of distance. If left uncorrected these errors reduce detection efficiency by causing power loss in the beamforming or matched filter reconstruction due to phase mismatch. We explore the impact of these phase and magnitude errors on beamforming in the next section.

5.4.3.4 Beamforming Characterization

Errors in the signal magnitudes and phases lead to errors in signal reconstruction. For example, a matched filter reconstruction requires accurate knowledge of the signals in

4077 each channel to achieve optimal performance. Uncorrected errors leads to mismatches
 4078 between the template and signal, which reduces detection efficiency and introduces
 4079 uncertainty in the parameter estimation. In this section, we analyze the beamformed
 4080 signal amplitude as a function of the position of the SYNCA to quantify the impact of
 4081 the phase and magnitude errors on signal reconstruction. Because of the imperfections
 4082 in the SYNCA source, it is inappropriate to directly compare the beamformed signal
 4083 amplitude of the FSCD array or synthetic array. Such a comparison would not allow
 4084 one to disentangle losses that occur because of the antenna array from those that occur
 4085 because of the source. Therefore, we focus on comparing the beamforming of the FSCD
 4086 array to the synthetic array.

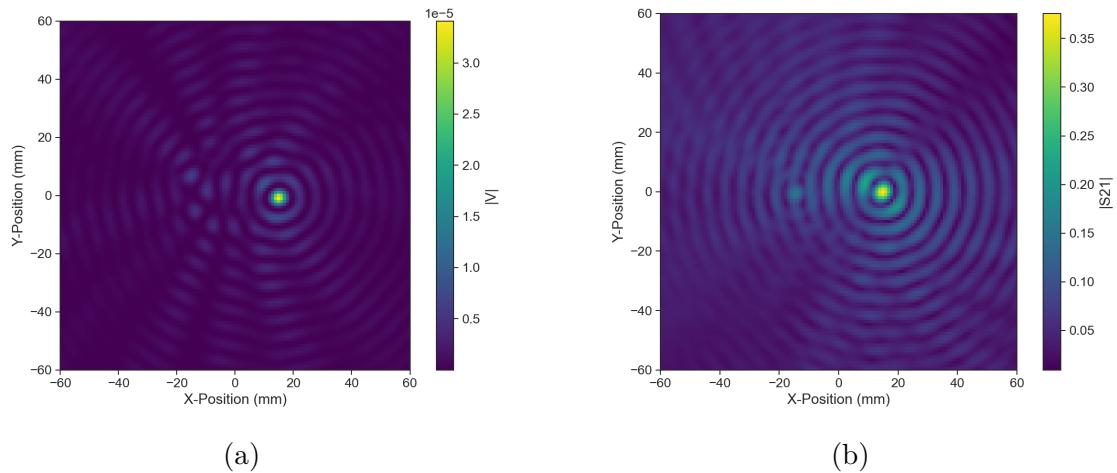


Figure 5.34: Beamforming images from the synthetic array (a) and FSCD array (b) setups with the SYNCA positioned 15 mm off the central axis. In both images we see a clear maxima that corresponds to the true SYNCA position. However, in the FSCD array there is an additional faint peak located at the opposite 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.

4087 The first method of comparison is to analyze the images generated by applying the
 4088 beamforming reconstruction specified in Section 4.3.1 to the FSCD and synthetic array
 4089 data (see Figure 5.34). The beamforming grid consisting of a square 121×121 grid
 4090 spanning a range of -60-mm to 60 mm in the x and y dimensions. The beamforming
 4091 images formed from the synthetic array produces a three-dimensional matrix where each
 4092 grid position contains a summed time series. A single beamforming image is formed from
 4093 this data matrix by taking the mean over the time dimension. In the case of the FSCD
 4094 array, the VNA generates frequency domain data such that each grid position contains a
 4095 summed frequency series produced by the VNA sweep. For this data a single image is

4096 formed by averaging in the frequency domain.

4097 There is a clear difference between the synthetic and FSCD array beamforming images,
 4098 which is the additional faint beamforming maxima located directly opposite the maxima
 4099 corresponding to the SYNCA position. The images in Figure 5.34 were generated with
 4100 data collected at a SYNCA radial position of 15 mm, which agrees well with the observed
 4101 beamforming maximum in both images. We observe that the faint beamforming peak is
 4102 located directly opposite of the true beamforming maximum similar to a mirror image.
 4103 Therefore, the origin of this additional feature appears to be reflections between the two
 4104 sides of the circular antenna array that are not present for the synthetic array since only
 4105 a single physical antenna is used.

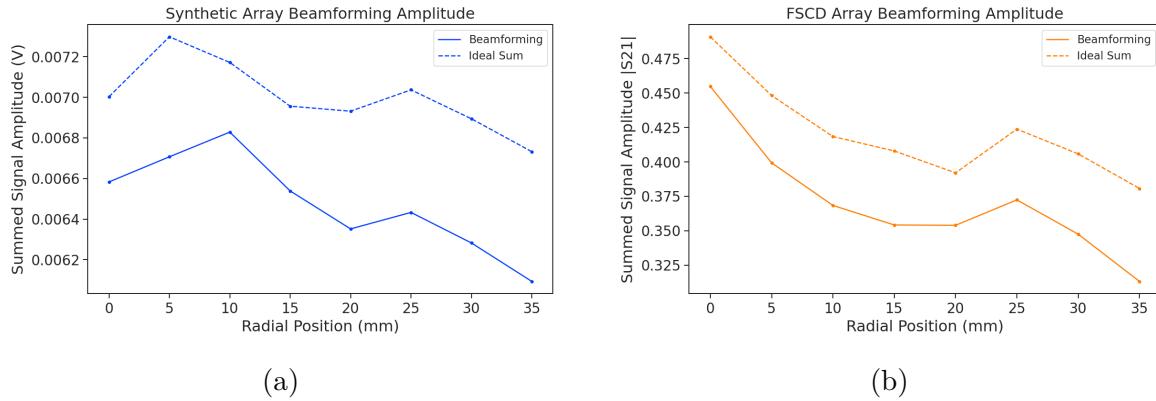


Figure 5.35: A comparison of the maximum signal amplitude obtained by beamforming to the signal amplitude obtained with an ideal summation as a function of the radial position of the SYNCA. The amplitudes for the synthetic array are shown in (a) and the FSCD array are shown in (b). In both setups we 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.

4106 From the beamforming images we extract the maximum amplitude, which we plot
 4107 as a function of the radial position of the SYNCA (see Figure 5.35). The phase errors
 4108 we observed in the FSCD and synthetic arrays leads to power loss at the beamforming
 4109 stage due to phase mismatches between the signals at different channels. This power
 4110 loss can be quantified by comparing the signal amplitude obtained from beamforming to
 4111 the amplitude which would be obtained from an ideal summation. We perform the ideal
 4112 summation by phase shifting each array channel to the same phase and then summing.
 4113 The comparison between the beamforming and ideal sums is shown in Figure 5.35, where
 4114 we observe that both the synthetic and FSCD arrays experience power losses from the
 4115 beamforming summation.

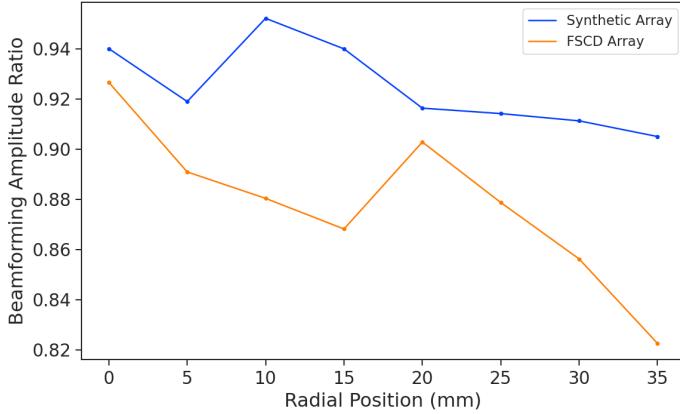


Figure 5.36: The ratio of the beamforming signal amplitude to the ideal signal amplitude for the FSCD and synthetic arrays. We see that the FSCD array has a larger power loss from phase error compare to the synthetic array which indicates that calibration errors associated with the multiple channels as well as reflections are impacting the signal reconstruction.

4116 The beamforming power loss can be quantified using the ratio of the beamforming to
 4117 ideal signal amplitudes. Computing this ratio as a function of SYNCA radial position
 4118 radius for the FSCD and synthetic arrays we find that the FSCD array has a uniformly
 4119 smaller beamforming amplitude ratio, which means that the FSCD array has a larger
 4120 beamforming power loss (see Figure 5.36). The primary contributions to the beamforming
 4121 power loss in the synthetic array are phase errors from the SYNCA and phase errors
 4122 from the FSCD antenna near-field. Both of these phase errors contribute to beamforming
 4123 losses in the FSCD array, but there are clearly additional phase errors in the FSCD array
 4124 measurements contributing to the smaller ratio. Two potential error sources include phase
 4125 differences in the different antenna channels that could not be corrected by calibration as
 4126 well as reflections between antennas in the array. The total effect of these additional phase
 4127 errors is to reduce the beamforming amplitude ratio by about 5% from the beamforming
 4128 ratio of the synthetic array. Therefore, we estimate that if no effort is made to correct
 4129 these phase errors in an FSCD-like experiment, then we expect approximately a 10%
 4130 total signal amplitude loss from a beamforming signal reconstruction.

4131 **5.4.4 Conclusions**

4132 The estimated power loss of a beamforming reconstruction obtained from this analysis
 4133 provides valuable inputs to sensitivity calculations of a FSCD-like antenna array exper-
 4134 iment to measure the neutrino mass, since it helps to bound systematic uncertainties

4135 from the antenna array and reconstruction pipeline. This power loss lowers the estimated
4136 detection efficiency of the experiment since some of the signal power is lost due to
4137 improper combining between channels and also increases the uncertainty in the electron's
4138 kinetic energy by contributing to errors in the estimation of the electron's cyclotron
4139 frequency.

4140 If these reconstruction losses prove unacceptable there are steps that can be taken
4141 to mitigate their effects. Some examples include the development of a more accurate
4142 antenna simulation approach that can reproduce the observed near-field interference
4143 patterns of the FSCD antennas and the implementation of a calibration approach that
4144 allows for the relative phase delays of the array to be measured without changing or
4145 disconnecting the antenna array configuration.

4146 **Chapter 6 |**

4147 **Development of Resonant Cavities for Large**

4148 **Volume CRES Measurements**

4149 **6.1 Introduction**

4150 The cavity approach was originally an alternative CRES measurement technology under
4151 consideration by the Project 8 collaboration for the Phase IV experiment. After pursuing
4152 an antenna array based CRES demonstrator design for several years the increasing costs
4153 and complexity of the antenna arrays led to a reconsideration of the baseline technology
4154 for the ultimate CRES experiment planned by Project 8. Currently, a cavity based CRES
4155 experiment is the preferred technology choice for future experiments by the Project 8
4156 collaboration including the Phase IV experiment.

4157 In this chapter I provide a brief summary of resonant cavities and sketch out the key
4158 features of a cavity based CRES experiment. In Section 6.2 I provide a brief introduction
4159 to cylindrical resonant cavities and the solutions for the electromagnetic fields in the
4160 cavity volume.

4161 In Section 6.3 I describe the main components of a cavity based CRES experiment,
4162 including the background and trap magnets, cavity geometry and design, and cavity
4163 coupling considerations. I also discuss some relevant trade-offs between an antenna array
4164 and cavity CRES experiment, and highlight some reasons for the transition of Project 8
4165 to the development of a cavity based experiment.

4166 Finally, in Sections 6.4 and 6.5, I present the design and development of an open
4167 mode-filtered cavity that could be used in a cavity based CRES experiment with atomic
4168 tritium. The results of the cavity simulations are confirmed by laboratory measurements
4169 of a proof-of-principle prototype that demonstrates key features of the design.

4170 6.2 Cylindrical Resonant Cavities

4171 Resonant cavities are sealed conductive containers, which allows us to describe the
4172 electromagnetic (EM) fields contained in the cavity volume as a superposition of resonant
4173 modes [36]. The field shapes of the resonant modes are determined by Maxwell's equations
4174 and the boundary conditions enforced by the cavity geometry. Of interest to Project 8
4175 for CRES measurements are cylindrical cavities due to their ease of construction and
4176 integration with atom and electron trapping magnets.

4177 6.2.1 General Field Solutions

4178 Consider a long segment of conducting material with a cylindrical cross-section (see
4179 Figure 6.1). A geometry such as this can be used as a waveguide transmission line to
4180 transfer EM energy from point to point, or, if conducting shorts are inserted on both
4181 ends of the cylinder, the waveguide becomes a resonant cavity.

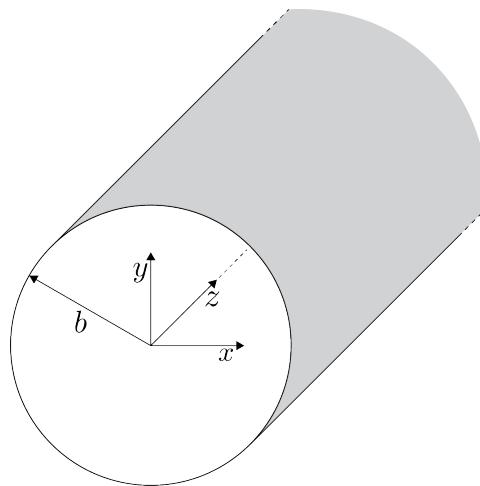


Figure 6.1: Geometry of a cylindrical waveguide with radius b .

4182 The fields allowed inside a cylindrical cavity are determined by the boundary conditions
4183 of the cylindrical geometry. The general approach to solving the fields begins by assuming
4184 solutions to Maxwell's equations of the form

$$\mathbf{E}(x, y, z) = (\mathbf{e}(x, y) + \hat{z}e_z(x, y))e^{-i\beta z}, \quad (6.1)$$

$$\mathbf{H}(x, y, z) = (\mathbf{h}(x, y) + \hat{z}h_z(x, y))e^{-i\beta z}. \quad (6.2)$$

4185 The solutions assume a harmonic time dependence of the form $e^{i\omega t}$ and propagation

4186 along the positive z-axis. The functions $\mathbf{e}(x, y)$ and $\mathbf{h}(x, y)$ represent the transverse
 4187 (\hat{x}, \hat{y}) components of the electric and magnetic fields respectively, and $e_z(x, y)$, $h_z(x, y)$
 4188 represent the longitudinal components. The version of Maxwell's equations in the case
 4189 where there are no source terms can be written as a pair of coupled differential equations,

$$\nabla \times \mathbf{E} = -i\omega\mu\mathbf{H}, \quad (6.3)$$

$$\nabla \times \mathbf{H} = i\omega\epsilon\mathbf{E}, \quad (6.4)$$

4190 where ϵ and μ are the permittivity and permeability of the material inside the waveguide
 4191 or cavity. Using the field solutions from Equations 6.1 and 6.2 one can solve for the
 4192 transverse components of the fields in terms of the longitudinal fields. Because we
 4193 are interested in cylindrical cavities it is advantageous to write the field solutions in
 4194 cylindrical coordinates. After performing this transformation the set of four equations
 4195 for the transverse field components are,

$$H_\rho = \frac{i}{k_c^2} \left(\frac{\omega\epsilon}{\rho} \frac{\partial E_z}{\partial\phi} - \beta \frac{\partial H_z}{\partial\rho} \right), \quad (6.5)$$

$$H_\phi = \frac{-i}{k_c^2} \left(\omega\epsilon \frac{\partial E_z}{\partial\rho} + \frac{\beta}{\rho} \frac{\partial H_z}{\partial\phi} \right), \quad (6.6)$$

$$E_\rho = \frac{-i}{k_c^2} \left(\beta \frac{\partial E_z}{\partial\rho} + \frac{\omega\mu}{\rho} \frac{\partial H_z}{\partial\phi} \right), \quad (6.7)$$

$$E_\phi = \frac{i}{k_c^2} \left(-\beta \frac{\partial E_z}{\partial\phi} + \omega\mu \frac{\partial H_z}{\partial\rho} \right), \quad (6.8)$$

4196 where k_c is the cutoff wavenumber defined by $k_c^2 = k^2 - \beta^2$ with $k = \omega\sqrt{\mu\epsilon}$ being the
 4197 wavenumber of the EM radiation.

4198 This set of equations can be used to solve for a variety of different modes that can be
 4199 obtained by setting conditions on E_z and H_z . For cylindrical cavities two types of modes
 4200 are allowed, which correspond to solutions where $E_z = 0$ and $H_z = 0$ respectively.

4201 6.2.2 TE and TM Modes

4202 The TE family of modes corresponds to the case where $E_z = 0$. This implies that H_z is
 4203 a solution to the Helmholtz wave equation

$$(\nabla^2 + k^2)H_z = 0. \quad (6.9)$$

4204 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
 4205 the standard technique of separation of variables. Rather than reproduce the derivation
 4206 here we shall simply quote the solutions for the transverse fields [36], which are

$$H_\rho = \frac{-i\beta}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.10)$$

$$H_\phi = \frac{-i\beta n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.11)$$

$$E_\rho = \frac{-i\omega\mu n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.12)$$

$$E_\phi = \frac{i\omega\mu}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}. \quad (6.13)$$

4207 One can observe that the solutions have a periodic dependence on ϕ , and radial profiles
 4208 given by the Bessel functions of the first kind. The integer indices n and m arise from
 4209 continuity conditions on the EM fields in the azimuthal and radial directions. For the
 4210 TE modes $n \geq 0$ and $m \geq 1$. $k_{c_{nm}}$ is the cutoff wavenumber for the TE_{nm} mode given by

$$k_{c_{nm}} = \frac{p'_{nm}}{b}, \quad (6.14)$$

4211 where b is the radius of the cavity or waveguide and p'_{nm} is the m -th root of the derivative
 4212 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

4213 The TM mode family corresponds to the case where $H_z = 0$, and $(\nabla^2 + k^2)E_z = 0$.
 4214 Again, we assume solutions of the form $E_z(\rho, \phi, z) = e_z(\rho, \phi)e^{-i\beta z}$, for which the general
 4215 form of the solutions is the same as for the TE modes. However, the different boundary
 4216 conditions for the TM modes results in particular solutions with a different form, which
 4217 we shall quote here without derivation. The transverse fields of the TM modes are given
 4218 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).

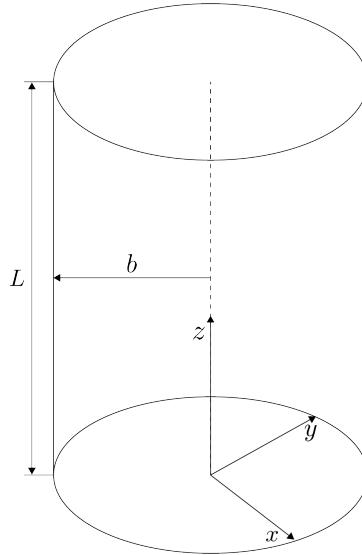


Figure 6.2: The geometry of a cylindrical cavity with length L and radius b .

4226

4227 The transverse electric field solutions for a cylindrical waveguide are of the form

$$\mathbf{E}(\rho, \phi, z) = \mathbf{e}(\rho, \phi) (A_+ e^{-i\beta_{nm}z} + A_- e^{i\beta_{nm}z}), \quad (6.19)$$

4228 where A_+ and A_- are arbitrary amplitudes of forward and backward propagating waves.

4229 In order to enforce that \mathbf{E} is zero at both ends of the cavity we require that

$$\beta_{nm}L = 2\pi\ell, \quad (6.20)$$

4230 where $\ell = 0, 1, 2, 3, \dots$. Using this constraint on the propagation constant we can solve

4231 for the resonant frequencies of the TE_{nml} and the TM_{nml} modes in a cylindrical cavity.

4232 For the TE modes the resonant frequencies are

$$f_{nml} = \frac{c}{2\pi\sqrt{\mu_r\epsilon_r}} \sqrt{\left(\frac{p'_{nm}}{b}\right)^2 + \left(\frac{\ell\pi}{L}\right)^2}, \quad (6.21)$$

4233 and the frequencies of the TM modes are

$$f_{nml} = \frac{c}{2\pi\sqrt{\mu_r\epsilon_r}} \sqrt{\left(\frac{p_{nm}}{b}\right)^2 + \left(\frac{\ell\pi}{L}\right)^2}. \quad (6.22)$$

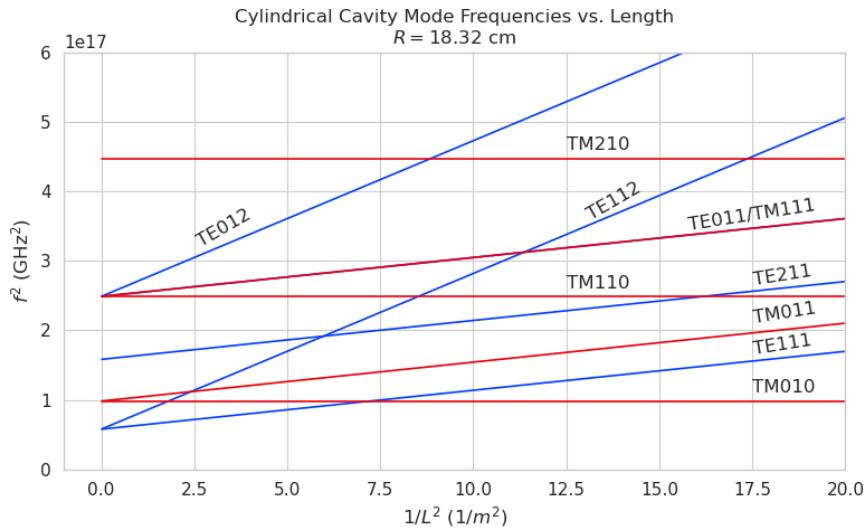


Figure 6.3: Relation of mode frequency to cavity length for a cylindrical cavity with a radius of 18.32 cm.

4234 6.2.4 Cavity Q-factors

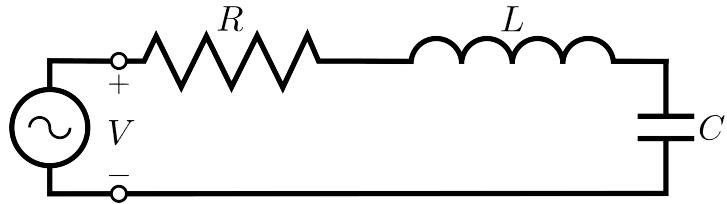


Figure 6.4: A series RLC circuit.

4235 The resonant behavior of cylindrical cavities can be modeled as a series RLC circuit
 4236 (see figure 6.4). The input impedance of the circuit can be obtained by applying
 4237 Kirchhoff's laws to calculate the impedance of the equivalent circuit. For a series RLC
 4238 circuit the input impedance is

$$Z_{\text{in}} = \left(\frac{1}{R} + \frac{1}{i\omega L} + i\omega C \right). \quad (6.23)$$

4239 The resistance in the circuit represents all sources of loss in the cavity, which is primarily
 4240 caused by the finite conductivity of the cavity walls. The inductor and capacitor represent
 4241 the energy stored in the cavity in the form of electric and magnetic fields. If the circuit
 4242 is being driven by an external power source we can write the input power in terms of the
 4243 circuit input impedance and the source voltage

$$P_{\text{in}} = \frac{1}{2} Z_{\text{in}} |I|^2 = \frac{1}{2} |I|^2 \left(\frac{1}{R} + \frac{1}{i\omega L} + i\omega C \right). \quad (6.24)$$

4244 The resistor introduces a loss into the system with a power given by

$$P_{\text{loss}} = \frac{1}{2} |I|^2 R, \quad (6.25)$$

4245 and the capacitor and inductor store energies given by

$$W_e = \frac{1}{4} \frac{|I|^2}{\omega^2 C}, \quad (6.26)$$

$$W_m = \frac{1}{4} |I|^2 L, \quad (6.27)$$

4246 respectively. Using these expressions we can write the input power and input impedance

4247 expressions in terms of the lost power and stored energy

$$P_{\text{in}} = P_{\text{loss}} + 2i\omega(W_m - W_e), \quad (6.28)$$

$$Z_{\text{in}} = \frac{P_{\text{loss}} + 2i\omega(W_m - W_e)}{\frac{1}{2}|I|^2}. \quad (6.29)$$

4248 The condition for resonance in the RLC circuit is that the stored magnetic energy
4249 is equal to the stored electric energy ($W_e = W_m$). When this occurs $Z_{\text{in}} = R$, which is a
4250 purely real impedance, and $P_{\text{in}} = P_{\text{loss}}$. The resonant frequency of the circuit can be
4251 determined from the condition $W_e = W_m$ from which one finds that

$$\omega_0 = \frac{1}{\sqrt{LC}}. \quad (6.30)$$

4252 An important performance parameter for any resonant system is the Q-factor, which
4253 quantifies the quality of the resonator as the ratio of the stored energy multiplied by the
4254 resonant frequency to the average energy lost per second. For the series RLC circuit, the
4255 Q-factor is given by the expression

$$Q_0 = \omega \frac{W_e + W_m}{P_{\text{loss}}} = \frac{1}{\omega_0 RC}, \quad (6.31)$$

4256 from which one observes that as the resistance of the RLC circuit is decreased the quality
4257 factor of the resonator increases. From the perspective of cylindrical cavities this implies
4258 that as one decreases the resistance of the cavity walls it is expected that the Q-factor of
4259 the cavity should increase, which is indeed the case. In certain applications where a high
4260 Q is desireable it is possible to manufacture a cavity out of superconducting materials in
4261 order to minimize the power losses of the system.

4262 The Q-factor of the resonator also determines with bandwidth (BW) of the system.
4263 A cavity with a high Q-factor will resonant with a smaller range of frequencies than a
4264 cavity with a low Q-factor. To see this we can examine the behavior of the RLC circuit
4265 when driven by frequencies near the resonance. For a frequency $\omega = \omega_0 + \Delta\omega$, where
4266 $\Delta\omega = \omega - \omega_0 \ll \omega_0$, we can write the input impedance as

$$Z_{\text{in}} = R + i\omega L \left(\frac{\omega^2 - \omega_0^2}{\omega^2} \right), \quad (6.32)$$

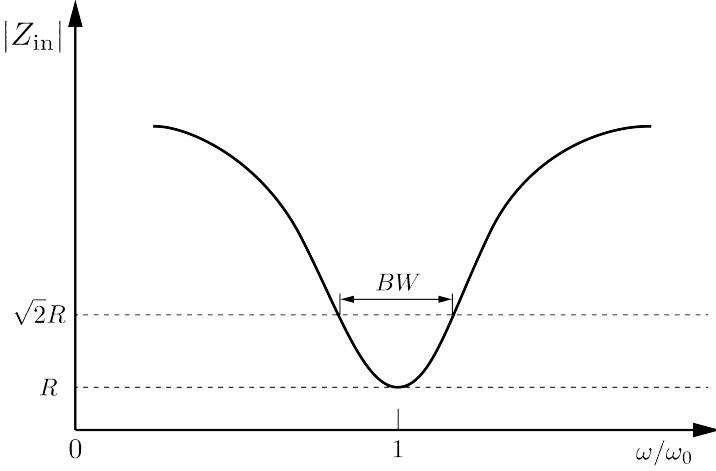


Figure 6.5: Illustration of the behavior of the input impedance of the series RLC circuit as a function of the driving frequency. The BW is proportion to the width of the resonance, which is inversely proportional to Q.

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

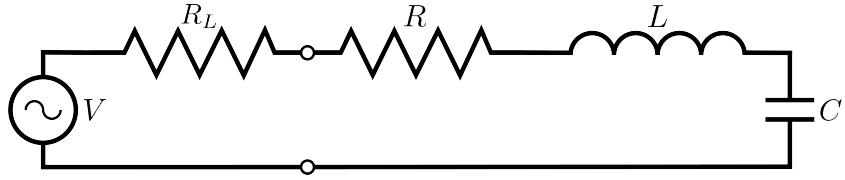


Figure 6.6: A series RLC circuit coupled to an external circuit with input impedance R_L .

4278 external circuit to drive a cavity resonance or to measure the energy of a resonant mode.
 4279 Coupling a cavity to an external circuit changes the Q by loading the equivalent cavity
 4280 RLC circuit (see Figure 6.6). The Q-factor of the cavity when it is loaded by an external
 4281 circuit is called the loaded Q, which is the quantity that one actually measures when
 4282 exciting a resonance in the cavity. Using the series RLC circuit model one can see that
 4283 the load resistor in Figure 6.6 will add in series with the resistor in the circuit for a total
 4284 equivalent resistance of $R + R_L$. Therefore, the loaded Q is given by

$$Q_L = \frac{1}{\omega_0(R + R_L)C}, \quad (6.37)$$

4285 from which one observes that the loaded Q is always less than the intrinsic Q of the
 4286 cavity.

4287 The amount of coupling that is desireable depends on the specific application of
 4288 the resonator. If one wants a resonator that is particular frequency selective than it
 4289 makes sense to limit the amount of coupling to the cavity to maintain a small BW,
 4290 alternatively, if a larger BW is need one can increase the cavity coupling by tuning the
 4291 input impedance of the external circuit. The critical point, where maximum power is
 4292 transferred between the cavity and the external circuit, occurs when the input impedance
 4293 of the cavity matches the input impedance of the external transmission line. For the
 4294 series RLC circuit on resonance, this matching condition corresponds to

$$Z_0 = Z_{\text{in}} = R, \quad (6.38)$$

4295 where Z_0 is the impedance of the transmission line. The loaded Q at this critical point
 4296 is, therefore,

$$Q_L = \frac{1}{2\omega_0 Z_0 C} = \frac{Q_0}{2}. \quad (6.39)$$

4297 One can described the degree of coupling between the cavity and an external circuit by

4298 defining a coupling factor, g , such that,

$$g = \frac{Q_0}{Q_L} - 1. \quad (6.40)$$

4299 When $g = 1$ then $Q_L = Q_0/2$, and the cavity is said to be critically coupled as we
4300 described. If $Q_L < Q_0/2$, then the cavity is undercoupled to the transmission line,
4301 corresponding to $g < 1$. Alternatively, if $Q_L > Q_0/2$, then $g > 1$, and the cavity is
4302 overcoupled to the transmission line. Various specialized circuits can be used to tune the
4303 input impedance of the external circuit as seen by the cavity to achieve a wide range of
4304 different coupling factors based on the desired application of the cavity.

4305 6.3 The Cavity Approach to CRES

4306 6.3.1 A Sketch of a Molecular Tritium Cavity CRES Experiment

4307 Resonant cavities can be used to perform CRES measurements, and they represent the
4308 current preferred technology by the Project 8 collaboration. The basic approach to a
4309 neutrino mass measurement using a resonant cavity and molecular tritium beta-decay
source is illustrated by Figure 6.7.

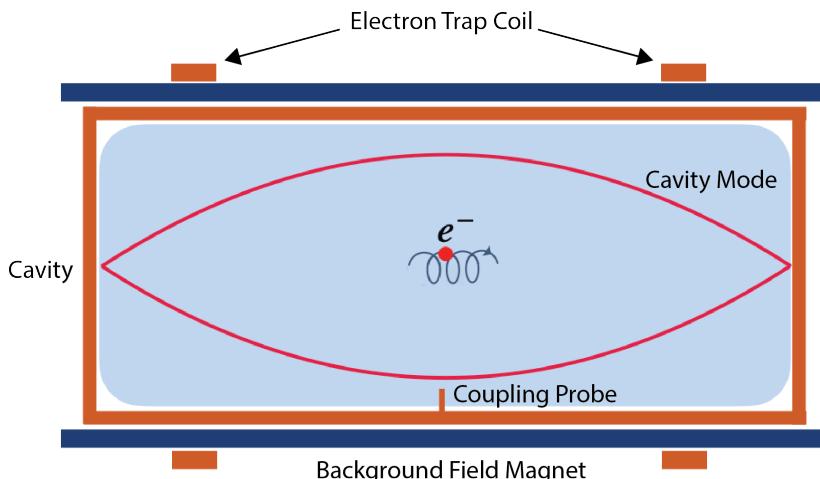


Figure 6.7: A cartoon depiction of a cavity CRES experiment. A metallic cavity filled with tritium gas is inserted into a uniform background magnetic field to perform CRES measurements. Electrons from beta-decays inside the cavity can be trapped and used to excite a resonant mode(s). By coupling to the cavity mode with a suitable probe one can measure the cyclotron frequency of the electron and perform CRES.

4310

4311 At the core of the experiment is a large resonant cavity filled with tritium gas. The
4312 filled cavity is then placed in a uniform magnetic field provided by a primary magnet
4313 that provides the background magnetic field. The value of the background magnetic field
4314 sets the range of cyclotron frequencies for electrons emitted near the tritium spectrum
4315 endpoint. When a beta-decay electron is produced in the cavity it is trapped using a set
4316 of magnetic pinch coils that keep electrons inside the cavity volume.

4317 Electrons trapped inside the cavity do not radiate in the same way as electrons
4318 in free-space. Effectively, the same boundary conditions that were used to derive the
4319 resonant modes of a cylindrical cavity in Section 6.2 apply to the radiation of the electron
4320 as well. The coupling of an electron performing cyclotron motion in a cavity has been
4321 studied in detail for measurements of the electron’s magnetic moment [59–61] If an
4322 electron is emitted with a kinetic energy that corresponds to a cyclotron frequency that
4323 matches a resonant frequency of the cavity, then energy radiated by the electron excites
4324 a corresponding resonance in the cavity. The strength of the electron’s coupling to the
4325 cavity is given to first order by the dot product between the electrons trajectory and
4326 the electric field vector of the resonant mode. Additional effects, such as the Purcell
4327 enhancement [62], alter the emitted power from the free-space Larmor equation [63]. If an
4328 electron is moving with a cyclotron frequency that is far from any resonant modes in the
4329 cavity, then radiation from the electron is suppressed. One can interpret this somewhat
4330 surprising effect as the metallic walls of the cavity reflecting the radiated energy back to
4331 the electron.

4332 Detecting an electron in the cavity is accomplished by coupling the cavity to an
4333 external transmission line that leads to an amplifier and RF receiver chain [64]. The
4334 coupling of the cavity resonance to the amplifier occurs through a coupling probe or
4335 aperture designed to read-out the excitation of the mode(s) excited by the electron. For
4336 CRES measurements, the placement of a wire antenna coupling probe inside the cavity
4337 volume leads to unacceptable losses of tritium atoms due to recombination to molecular
4338 tritium on the antenna surface, therefore, apertures are the preferred coupling method
4339 for cavity CRES experiments.

4340 One of the attractive features of the CRES technique for neutrino mass measurement
4341 is the gain in statistics that comes from the differential nature of the tritium spectrum
4342 measurement. Initially, this seems incompatible with cavities, due to the narrow reso-
4343 nances of cavity modes giving relatively small bandwidth. However, by intentionally
4344 over-coupling to a single cavity mode one can achieve bandwidths of a few 10’s of MHz
4345 (see Section 6.2), which is sufficient for a measurement of the tritium spectrum endpoint

4346 region.

4347 **6.3.2 Magnetic Field, Cavity Geometry, and Resonant Modes**

4348 **Magnetic Field and Volume Scaling**

4349 For a CRES experiment, cylindrical cavities are a natural choice since they match
4350 the geometry of standard solenoid magnets, which are needed in order to produce the
4351 background magnetic field for CRES measurements. Furthermore, the cylindrical shape is
4352 compatible with a Halbach array, which is the leading choice of atom trapping technology
4353 for future atomic tritium experiments by the Project 8 collaboration. Cylindrical
4354 cavities also benefit from well-established machining practices that are able to achieve
4355 high geometric precision at large lengths scales. More exotic cavity designs are under-
4356 consideration and there are on-going efforts to investigate the potential advantages these
4357 may have over the standard cylindrical geometry.

4358 As we saw in Section 6.2, the physical dimensions of the cavity are directly coupled
4359 to the resonant frequencies of the cavity. This dependency links the size of the cavity to
4360 the magnitude of the background magnetic field, because the magnetic field determines
4361 the cyclotron frequencies of trapped electrons. Specifically, as the size of the cavity is
4362 increased to accommodate larger volumes of tritium gas, the frequencies of the resonant
4363 modes decrease proportionally. This requires that the magnetic field also decrease in
4364 order to maintain coupling between electrons and the desired cavity mode.

4365 The required cavity size is ultimately determined by the required statistics in the
4366 tritium spectrum endpoint region. Because the gas density must be kept below a certain
4367 level to ensure that electrons have sufficient time to radiate before scattering, larger
4368 volumes become the only way to achieve higher event statistics. To achieve the sensitivity
4369 goals of Phase III and IV cavity volumes on the order of several cubic-meters are required,
4370 which pushes one towards frequencies in the range of 100's of MHz.

4371 **Single-mode Cavity CRES**

4372 It is tempting to consider maintaining a high magnetic field, while still increasing the size
4373 of the cavity, in order to increase the radiated power from trapped electrons for better
4374 SNR. However, if one were to maintain the same magnetic field while increasing the
4375 size of the cavity, the electrons would begin to couple to higher order modes with more
4376 complicated transverse geometries. The danger with this approach is that a complicated
4377 mode structure could introduce systematic errors into the CRES signals. Example

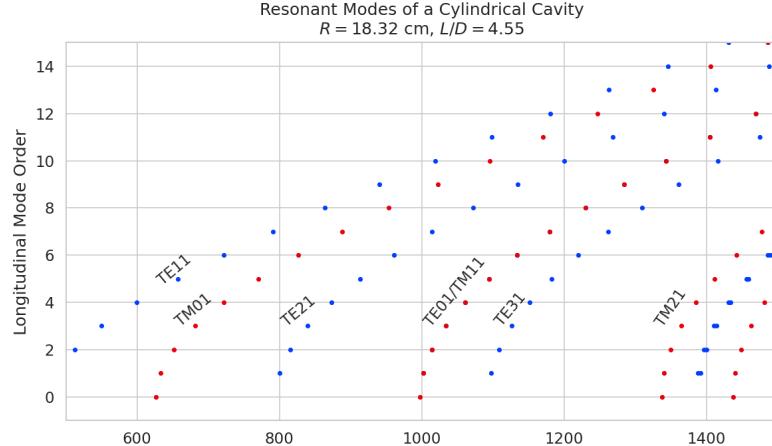
4378 systematics include unpredicted mode hybridization or changes in the mode shapes from
4379 imperfections in the cavity construction, which would prevent reconstruction of the
4380 electron's starting kinetic energies with adequate resolution. For this reason, it is ideal
4381 to operate with magnetic fields that give cyclotron frequencies near the fundamental
4382 frequency of the cavity, where the mode structure is relatively simple (see Figure 6.8).
4383 In this frequency region it is possible to perform CRES by coupling to only a single
4384 resonant mode, however, it is currently an open question if a single mode measurement
4385 will provide enough information about an individual electron's position to reconstruct
4386 the full event. Regardless, developing a solid understanding of the CRES phenomenology
4387 when an electron is coupling to a single mode will be a necessary step towards a future
4388 multi-mode cavity experiment.

4389 Considerations for Resonant Mode Selection

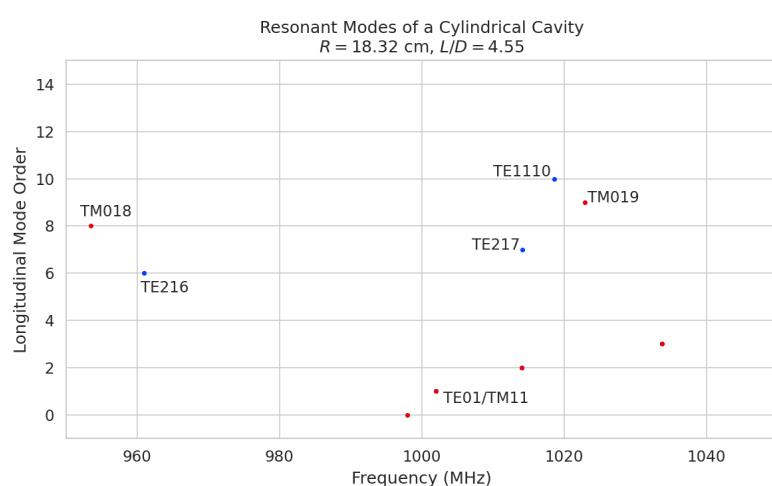
4390 A single-mode cavity experiment begs the question, which resonant mode is best for
4391 CRES measurements? There is an immediate bias towards low order TE_{nm} and TM_{nm}
4392 modes due to the multi-mode considerations discussed above. Additionally, there is a
4393 preference towards modes with longitudinal index $\ell = 1$ with a single antinode along the
4394 vertical axis of the cylindrical cavity. The reason for this is that there is a phase change
4395 in the electric fields between antinodes that leads to modulation effects that destroy the
4396 carrier frequency signal information.

4397 A second consideration for mode selection is the volumetric efficiency of the mode.
4398 Volumetric efficiency can be thought of as an integral over the volume of the cavity
4399 weighted by the relative amplitude of the mode. From the perspective of simply maximiz-
4400 ing the volume useable for CRES measurements this integral would be as close to unity
4401 as possible. However, there is a requirement to reconstruct the position of the electrons
4402 inside the cavity volume so that the local magnetic fields can be used to convert the
4403 measured cyclotron frequency to a kinetic energy. With a single mode this necessarily
4404 requires a variable transverse mode amplitude, which lowers the volumetric efficiency, so
4405 that position of the electron in the cavity can be estimated from the average amplitude
4406 of the CRES signal. Longitudinal indices of $\ell = 1$ have an advantage in volumetric
4407 efficiency over higher order ℓ modes, since there are only two longitudinal nodes, one at
4408 each end of the cavity. Therefore, the average coupling strength of trapped electrons as
4409 they oscillate axially is higher for $\ell = 1$ modes.

4410 The longitudinal variation in the mode strength is ultimately critical for achieving the
4411 energy resolution required for neutrino mass measurements. Correcting for the change in



(a)



(b)

Figure 6.8: Examples of the resonant mode frequencies of a cylindrical cavity. This cavity has a radius of 18.32 cm and a length to diameter ratio of 4.55.

the average magnetic fields experienced by electrons with different pitch angles requires that information on the axial motion of the electron be encoded into the CRES signal. The longitudinal variation in the mode amplitude leads to amplitude modulation of the CRES signal with a frequency proportional to the electron's pitch angle.

An additional factor for mode selection is the intrinsic or unloaded Q of the mode. In terms of SNR it is advantageous to use a mode with a very high Q_0 , which is then highly overcoupled to achieve the necessary bandwidth to cover the tritium endpoint spectrum. This scheme leads to a decoupling of the physical cavity temperature from the effective noise temperature after the amplifier, which allows us to achieve adequate SNR without

4421 the requirement of cooling the entire cavity to single Kelvin temperatures.

4422 An example of a resonant mode that exhibits these traits is the TE₀₁₁ mode. At present
4423 the TE₀₁₁ mode is the preferred resonance for a single-mode cavity CRES experiment
4424 by the Project 8 collaboration. TE₀₁₁ is a low order mode located in a region relatively
4425 far from other cavity modes. Furthermore, the separation of the TE₀₁₁ mode can be
4426 improved by various mode-filtering techniques discussed in Section 6.4.2 below. TE₀₁₁
4427 consists of a single longitudinal antinode that can provide pitch angle information in the
4428 form of amplitude modulation, and has an electric field with a radial profile given by the
4429 J'_0 Bessel function allowing for radial position estimation. Lastly, the TE₀₁₁ mode has a
4430 relatively high intrinsic Q compared to nearby modes, which helps with SNR. Unloaded
4431 Q's greater than 80000 are achievable for a 1 GHz TE₀₁₁ resonance using a copper walled
4432 cavity.

4433 **6.3.3 Trade-offs Between the Antenna and Cavity Approaches**

4434 The choice between cavities and antennas for large-scale CRES measurements is not
4435 without trade-offs. Both the antenna array and cavity approaches are relatively immature
4436 techniques, at present there are no known obstacles that would prevent either approach
4437 from being used for a large scale neutrino mass experiment. The preference for cavities
4438 is largely driven by important practical considerations that could make a cavity based
4439 experiment significantly cheaper than an antenna experiment of similar size and scope.
4440 However, the switch to cavities also introduces new challenges less relevant to the
4441 antenna array, which must be solved in order for Project 8 to achieve its neutrino mass
4442 measurement goals.

4443 One of the major relative drawbacks of the antenna array approach is the size and
4444 complexity of the data-acquisition system. A large-scale antenna array experiment
4445 requires $O(100)$ antennas independently digitized at rates of $O(10)$ to $O(100)$ MHz. Since
4446 there is insufficient information in a single antenna channel to detect or reconstruct the
4447 CRES signal, the entire array output must be processed during the signal reconstruction.
4448 Because data storage becomes an issue with these data volumes, there is a real-time
4449 signal reconstruction requirement that allows one to detect CRES signals buried in the
4450 thermal noise. As we discuss in Section 4.4, the computational cost of these real-time
4451 detection algorithms are potentially quite large for even a small scale antenna array
4452 experiment. However, the operating principle of a cavity experiment allows the CRES
4453 signal to be detected using only a single read-out channel digitized at rates of $O(10)$ MHz,
4454 which reduces the cost of the data acquisition system by many orders of magnitude.

4455 From an engineering perspective, the simple geometry and thin-walls of a cylindrical
4456 cavity are simpler to interface with the cryogenic and magnetic subsystems needed for a
4457 CRES experiment. Whereas, the antenna array requires careful design and engineering
4458 to accommodate the antenna array and receiver electronics in proximity to the trapping
4459 magnets. Additionally, due to near-field interference effects, the antenna array is unable
4460 to reconstruct CRES events within the reactive near-field distance of the antennas.
4461 Because atom trapping requirements require magnetic fields which correspond to cyclotron
4462 frequencies for endpoint electrons less than 1 GHz, the required stand-off distance leads to
4463 a significant loss in useable experiment volume, necessitating larger and more expensive
4464 magnets.

4465 Another advantage to the cavity approach is the relatively compact sideband structure,
4466 which is a result of the low modulation index for cavity CRES signals. The axial motion
4467 in an antenna array experiment leads to frequency modulation and sidebands. The shape
4468 of the sideband structure is determined by the modulation index, $h = \frac{\Delta f}{f_a}$, where Δf
4469 is the size of the frequency deviation and f_a is the axial frequency. The large electron
4470 traps required for a cubic-meter-scale experiment leads to high modulation indices, which
4471 causes the signal spectrum to be made up of numerous low power sidebands that make
4472 reconstruction and detection challenging. This behavior was observed in simulations
4473 of the FSCD in which carrier power decreased with pitch angle due to the increase in
4474 modulation index (see Figure 4.31). For cavities, however, the modulation index remains
4475 near $h = 1$ even for very long magnetic traps due to the high phase velocity in cavities
4476 relative to the axial velocity of the electron. This results in an almost ideal spectrum
4477 shape that has a strong carrier frequency with a few sidebands whose relative amplitudes
4478 encode pitch angle information.

4479 A downside of the cavity approach is the apparent difficulty of estimating the position
4480 of the electron using only the coupling of the electron to a single mode. The amplitude of
4481 the TE₀₁₁ mode is completely independent of the azimuthal coordinate, therefore, position
4482 reconstruction using the TE₀₁₁ mode is only able to estimate the radial position of the
4483 electron. This position degeneracy may lead to magnetic field uniformity requirements
4484 that are too challenging to meet due to mechanical uncertainties in cavity and magnet
4485 construction, as well as uncertainties caused by nuisance external magnetic fields such
4486 as the Earth's field and magnetic fields from building materials. A multi-mode cavity
4487 experiment may provide a way to extract more precise information on the position of
4488 the electron by analyzing the coupling of the electron to several modes that overlap in
4489 different ways.

6.4 Single-mode Resonant Cavity Design and Simulations

The single-mode cylindrical cavities envisioned for the Phase III and IV experiments must be carefully engineered in order to measure the neutrino mass with the desired sensitivity. In this section I summarize some simulation studies performed to analyze early design concepts for a single-mode cavity. The primary tool for these investigations was Ansys HFSS, which was also used for the development of the SYNCA antenna described in Section 5.3.

6.4.1 Open Cylindrical Cavities with Coaxial Terminations

Design Concept

A basic cavity design question relevant to Project 8's ultimate goal of an atomic tritium CRES experiment is how to build a cavity that can be efficiently filled with atomic tritium. To keep the rate of atom loss from recombination on surfaces it is ideal if the ends of the cylindrical cavity are as open as possible so that tritium atoms can flow inside unimpeded. Additionally, one of the primary calibration techniques planned for future CRES experiments involves CRES measurements using electrons injected from an electron gun source, which also requires an opening at the cavity end. Cylindrical cavities with open ends can be manufactured, however, the intrinsic Q-factors of these cavities are orders of magnitude less than their sealed counterparts, which reduces the signal-to-noise ratio when that cavity is used for CRES measurement.

Cylindrical cavities with mostly open ends that also exhibit Q values for the $TE_{01\ell}$ modes similar to sealed cavities can be built by using coaxial endcaps to terminate the cavity. Cavities of this type have been manufactured for specialized applications related to the measurements of the dielectric constants of liquefied gasses (see Figure 6.9) [65, 66]. This cavity design leaves the ends of the cavity wide open, but retains high Q-values for the $TE_{01\ell}$ modes due to the coaxial endcap, which are designed to perfectly reflect the electric fields of $TE_{01\ell}$ modes. Coupling to the $TE_{01\ell}$ mode is achieved via an aperture located at the center of the cavity wall.

A cavity similar to Figure 6.9 is a candidate design for the future CRES experiments by Project 8, since it appears to elegantly solve many practical issues that arise when combining cavity CRES and atomic tritium. The coaxial endcaps leave significant regions of the cavity ends completely open, which allows for the entrance of atomic tritium as well as the pumping away of molecular tritium that has recombined on the cavity walls.

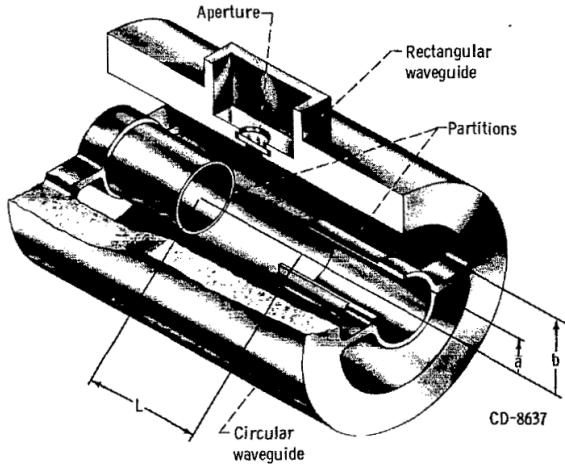


Figure 6.9: An image of an open cavity with coaxial terminations used for dielectric constant measurements. Figure from [66].

4522 These open ends are achieved while preserving the high Q-values of the $TE_{01\ell}$ modes,
 4523 which is important for extracting as much signal power from the electron as possible. In
 4524 subsequent sections we shall analyze this cavity design in more detail, primarily by using
 4525 HFSS simulations to analyze the resonant mode structure of this cavity geometry.

4526 Coaxial Terminator Constraints

4527 The reason that coaxial endcaps can be used to achieve high Q-values for the $TE_{01\ell}$
 4528 modes is that the electric fields for these modes are purely azimuthally polarized (see
 4529 Equations 6.12 and 6.13). Therefore, the boundary conditions that require the electric
 4530 field to go to zero at the cavity ends can be supplied using a coaxial partition of the
 4531 correct radius (see Figure 6.10). Because the cylindrical shape enforced by the partition
 4532 does not match the boundary conditions of other cavity modes, these terminations also
 4533 significantly suppress the Q-factors of non- $TE_{01\ell}$ modes, which is potentially beneficial
 4534 for a single-mode cavity CRES experiment.

4535 The correct radius of the cylindrical partition is derived by setting up the boundary
 4536 value problem in Figure 6.10, and analyzing the reflection and transmission coefficients
 4537 for waves incident on the coaxial terminators. The basic problem is to identify the radius
 4538 a where the reflection coefficient for the $TE_{01\ell}$ modes becomes equal to 1. One can show
 4539 that if the coaxial partitions are made sufficiently long relative to the wavelength of the
 4540 TE_{01} modes than perfect reflection can be achieved. This derivation is quite lengthy
 4541 and complex and is presented in full in [65]. Here, we shall simply explain the resulting

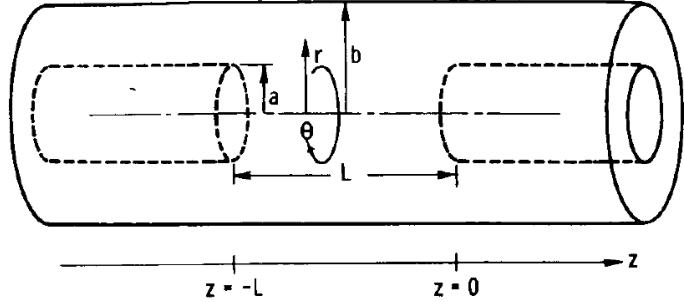


Figure 6.10: The simplified geometry of an open cavity with coaxial terminations. Figure from [65].

4542 conditions on the partition radius for perfect reflection.

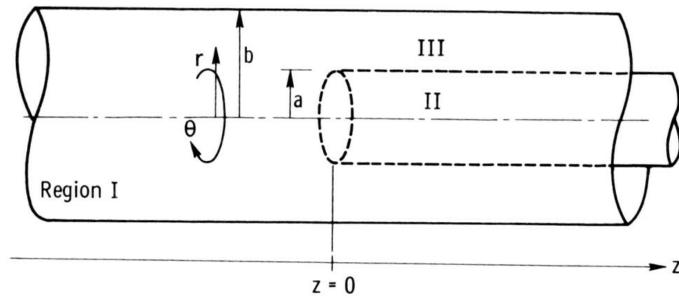


Figure 6.11: Electric field regions for the open cavity boundary value problem. Figure from [65].

4543 The open cavity boundary value problem is solved by expressing the forms of the
 4544 electric fields in the different regions of the cavity and requiring that the electric fields are
 4545 continuous. There are effectively three distinct regions in the open cavity corresponding
 4546 to the central cavity volume, the inner coaxial volume, and the outer coaxial volume (see
 4547 Figure 6.11).

4548 In Region I, the boundary conditions are those of a cylindrical waveguide, and we
 4549 require that E_ϕ for the TE_{0m} modes go to zero at the cavity wall ($r = b$). This requires
 4550 that $J'_{0m}(k_{c0m} b) = 0$. We aim to solve for the radius a in the specific situation where the
 4551 TE_{01} mode can propagate but all other TE_{0m} modes are below the cutoff frequency for
 4552 the circular waveguide. This is equivalent to requiring

$$3.832 < k_{c0m} b < 7.016, \quad (6.41)$$

4553 where the numbers 3.832 and 7.016 correspond to the first and second zeros of the Bessel

4554 function (see Table 6.1).

4555 In Region II the boundary conditions are those of a cylindrical waveguide, but with
4556 a smaller radius. The condition that $E_\phi = 0$ at the cylindrical partition radius is that
4557 $J'_{0m}(k_{c0m}a) = 0$. To ensure perfect reflection, we want all modes in Region 1 of the cavity
4558 to be below the cutoff frequency of the circular waveguide formed by the inner volume of
4559 the coaxial terminator. Therefore, we consider the solutions where

$$k_{c0m}a < 3.832. \quad (6.42)$$

4560 Finally, in Region III the boundary condition are those of a coaxial waveguide. We
4561 need to guarantee that $E_\phi = 0$ at both $r = b$ and $r = a$, which involves finding the
4562 eigenvalues of the following equation

$$J'_0(k_{c0m}a)Y'_0(k_{c0m}b) - J'_0(k_{c0m}b)Y'_0(k_{c0m}a) = 0, \quad (6.43)$$

4563 where Y'_0 the zeroth-order derivatives of the Bessel function of the second kind. The
4564 solutions to this equation depend on the value of the ratio b/a . The approximate solution
4565 is given by

$$\delta_n a \simeq \frac{n\pi}{b/a - 1}, \quad (6.44)$$

4566 where δ_n are eigenvalues of Equation 6.43. Similar to Region II, we are interested in
4567 solutions for which the TE₀₁ modes of Region I are below the cutoff of Region III.
4568 Therefore, we require that

$$k_{c0m} < \delta_1. \quad (6.45)$$

4569 In general, one has some freedom in specifying the value of b/a . A value typically used
4570 in practice is $b/a = 2.082$, which corresponds to positioning the radius of the cylindrical
4571 partition at the maxima of the TE₀₁ electrical fields.

4572 Using the constraints from the three field regions one can develop a coaxial terminator
4573 that acts as a virtual perfectly conducting surface for the TE₀₁ modes. The only required
4574 inputs are the desired frequency of the TE₀₁₁ mode and a choice for the value of b/a .

4575 **6.4.2 Mode Filtering**

4576 The general case of an electron coupling to a resonant cavity is complicated. This is
4577 because cavities contain an infinite number of resonant modes, which for higher order
4578 modes, have couplings to the electron with a complex spatial dependence. The danger is

4579 that improper modeling of the electron's coupling to the cavity can lead to systematic
4580 errors in the CRES measurements that prevent a high-resolution measurement of the
4581 electron's kinetic energy. This in part drives the preference for a single-mode cavity
4582 experiment that uses only the electron's coupling to the TE₀₁₁ mode to perform CRES,
4583 assuming that sufficient information on the electron's position can be obtained with a
4584 single mode.

4585 The TE₀₁₁ mode is in a region where there are relatively few other modes to which
4586 the electron could couple(see Figure 6.8). However, one can see that the frequency of
4587 the TE₀₁₁ is perfectly degenerate with the TM₁₁₁ mode, which means that electrons will
4588 inevitably couple to both modes if they have the correct cyclotron frequency.

4589 The magnitude of the impact of the electron coupling to both TE₀₁₁ and TM₁₁₁ is
4590 currently unknown. To first order an electron coupling to more both modes will lose more
4591 energy overtime, which can be measured by observing the frequency chirp rate of the
4592 signal. This effect may be small enough to be negligible or simple enough to model that
4593 the cavity can be treated as an effective single-mode cavity. Alternatively, the one could
4594 consider devising a coupling scheme that is sensitive to both the TE₀₁₁ and the TM₁₁₁
4595 modes. By measuring the coupling of the electron to both modes more information on
4596 the position of the electron could be obtained, which could improve the position and
4597 energy resolution of the CRES measurements.

4598 A different approach is the mode filtering approach, which seeks to obtain a single
4599 TE₀₁₁ mode cavity using perturbations to the cavity walls that selectively impede the
4600 TM modes, while leaving the TE modes mostly unperturbed. The type of perturbations
4601 required can be determined by visualizing the surface currents induced in the cavity
4602 walls by each type of mode (see Figure 6.12). By definition, all TM have electric fields
4603 directed along the vertical axis of the cylindrical cavity, which means that perturbations
4604 that impede currents in this direction will modify TM resonances. On the other hand,
4605 the TE₀₁ modes induce azimuthal currents in the cavity walls, therefore, it is possible to
4606 break the degeneracy between TE₀₁ and TM₁₁ using a cavity perturbation that impedes
4607 axial currents, but does not affect the flow of azimuthal currents.

4608 Figure 6.12 shows two cavity design concepts that achieve this selective current
4609 perturbation. The resistive approach inserts a series of thin dielectric rings into the walls
4610 of the cavity that introduces a resistive and capacitive impedance to the longitudinal
4611 currents, while leaving azimuthal current paths intact. Cavities of this type with high
4612 TE₀₁ Q's have also been constructed by tightly wrapping a thin, dielectric coated wire
4613 around a mold to form the cavity wall. An alternative method is to introduce an inductive

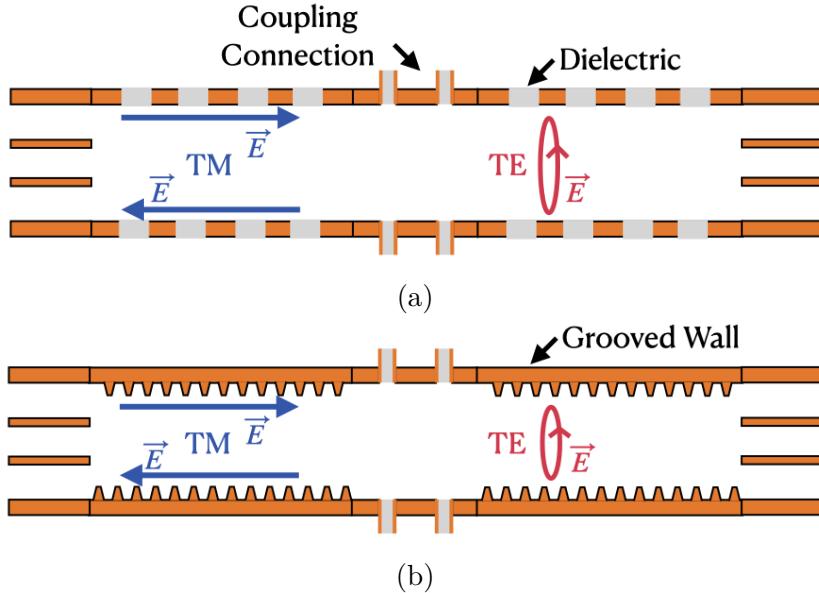


Figure 6.12: Two mode filtering concepts to break the degeneracy of TE_{01} and TM_{11} modes. The resistive approach uses dielectric materials to impede currents that travel vertically along the cavity while leaving azimuthal currents unperturbed. An alternative approach is to impede the currents using grooves cut into the cavity wall, which achieve the same effect with an inductive impedance.

⁴⁶¹⁴ impedance by cutting grooves or a thread pattern on the inside wall of the cavity. For
⁴⁶¹⁵ reasons of manufacturability and compatibility with tritium the grooved cavity approach
⁴⁶¹⁶ is the preferred method for mode-filtered cavity construction by Project 8.

⁴⁶¹⁷ 6.4.3 Simulations of Open, Mode-filtered Cavities

⁴⁶¹⁸ A candidate design for a single TE_{011} mode CRES experiment is a cavity that utilizes
⁴⁶¹⁹ the coaxial terminations combined with a mode-filtering wall. The first step towards
⁴⁶²⁰ validating that a cavity that combines these two design features will operate as expected
⁴⁶²¹ is a thorough simulation effort for which finite element method (FEM) simulation software
⁴⁶²² is invaluable. The primary tool for electromagnetic FEM calculations inside Project 8 is
⁴⁶²³ Ansys HFSS, which has a robust and well-established eigenmode solver that can identify
⁴⁶²⁴ the resonant frequencies and associated Q-factors for given structure.

⁴⁶²⁵ Four variations of a cavity design with a ~ 1 GHz TE_{011} resonance were implemented
⁴⁶²⁶ in HFSS (see Figure 6.13). The four designs include a standard cylindrical cavity, an
⁴⁶²⁷ open cavity with smooth walls, an open cavity with resistive walls, and an open cavity
⁴⁶²⁸ with grooved walls. The relevant design parameters are summarized in Table 6.3. All

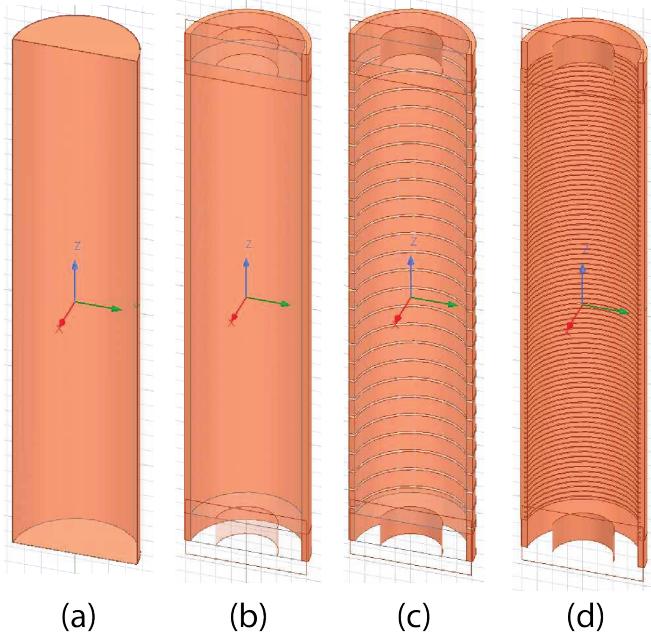


Figure 6.13: Four cavity design variations. (a) is a standard sealed cylindrical cavity, (b) is an open cavity with smooth walls, (c) is an open cavity with resistive walls, and (d) is an open cavity with grooved walls. The main cavity and coaxial terminator parameter are identical for all four cavities.

⁴⁶²⁹ cavities were simulated using copper walls and filled with a vacuum dielectric. The
⁴⁶³⁰ identities of the resonant modes found by HFSS were validated by visual inspection of
⁴⁶³¹ the electric and magnetic field patterns and by comparison to analytical calculations of
⁴⁶³² the mode frequencies.

Table 6.3: A table of cavity design parameters used for HFSS simulations.

Name	Qty.	Unit	Description
D_{cav}	326.4	mm	Cavity diameter
L_{cav}	1668.0	mm	Cavity length
D_{term}	200.2	mm	Inner diameter of coaxial terminator
L_{term}	100.0	mm	Terminator length
l_{die}	8.3	mm	Dielectric spacer thickness
Δl_{die}	66.7	mm	Distance between dielectric spacers
l_{groove}	3.0	mm	Groove height
d_{groove}	9.0	mm	Groove depth
Δl_{groove}	18.3	mm	Distance between grooves

⁴⁶³³ The results of the HFSS simulations validate our predictions of the resonant behavior
⁴⁶³⁴ of an open, mode-filtered cavity developed in the preceding sections (see Figure 6.14) One

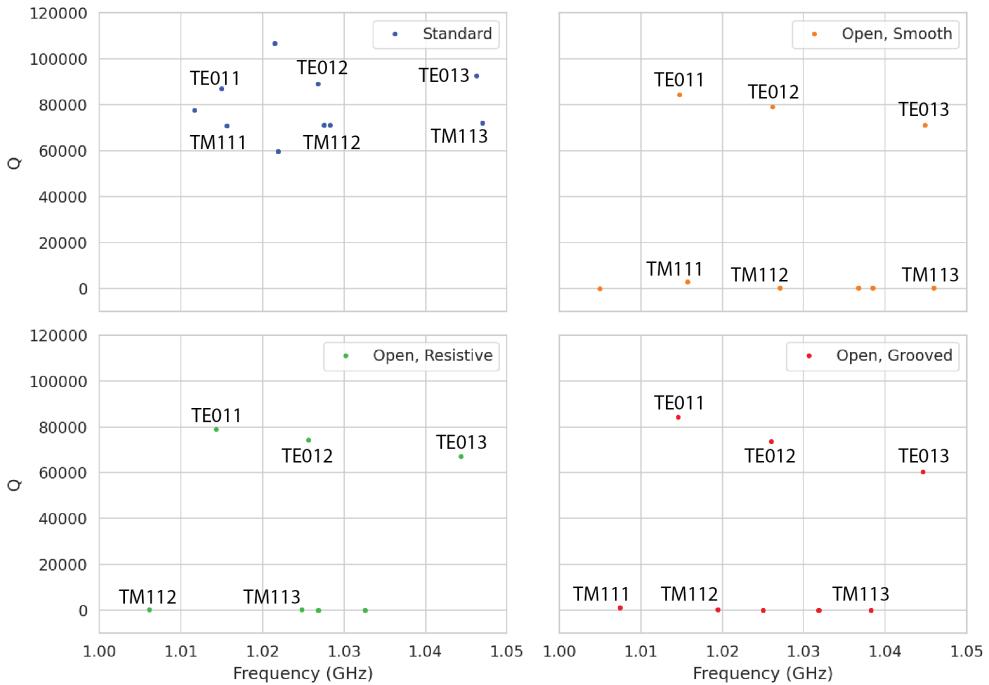


Figure 6.14: The frequencies and Q-factors of the resonant modes identified by HFSS for the cavity variations shown in Figure 6.13. The fully-sealed cavity with smooth walls has several high-Q modes near the TE_{01} resonance. Introducing the open-termination preserves the Q-factors of the $TE_{01\ell}$ 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.

4635 can see that for a standard cavity the TE_{01} and the TM_{11} are degenerate in frequency
 4636 with relatively high Q-factors. The open-ended cavity preserves the high Q-factors of
 4637 the TE_{01} modes, while the other modes, since their boundary conditions do not match
 4638 the coaxial geometry, have their Q-factors suppressed. One can see that the effect of
 4639 the resistive and inductive mode-filtering schemes is to effectively shift the resonant
 4640 frequencies of the TM_{11} modes below those of the associated TE_{01} modes, which breaks
 4641 the degeneracy. Optimization of the dielectric spacer or groove parameters can ensure
 4642 that the TE_{011} mode is isolated from other modes by $O(10)$ MHz, which provides sufficient
 4643 bandwidth for a measurement of the tritium spectrum endpoint.

4644 Further optimization of the cavity design requires a more detailed cavity simulation
 4645 that includes the cavity coupling mechanism as well as other geometry modifications
 4646 required for integration into the magnetic and tritium gas subsystems. Perhaps more

4647 important is the development of the capability to simulate the interaction of electrons
4648 with the cavity so that simulated CRES signals can be generated using cavities designed
4649 for CRES measurements. Simulated CRES signals can then be used to estimate the
4650 neutrino mass sensitivity of the experiment, which allows for the optimization of the cavity
4651 design towards the configuration that provides the best measurement of the neutrino
4652 mass.

4653 **6.5 Single-mode Resonant Cavity Measurements**

4654 Measurement test stands play an important role in the research and development process
4655 that cannot be replaced by simulations. For example, constructing a prototype CRES
4656 cavity forces one to consider important practical issues such as manufacturability and
4657 machine tolerances that may require modifications to the design. Furthermore, by
4658 comparing laboratory measurements of a real cavity to simulations, one can quantify
4659 the impact of imperfections and real-life measurement systematics, which allows for
4660 more accurate sensitivity estimates of the experiment. Lastly, the development of these
4661 prototypes helps to build the necessary experience and expertise within the collaboration
4662 required for more complicated experiments to succeed.

4663 In this spirit a prototype cavity was constructed to demonstrate the open, mode-
4664 filtered cavity concept explored in the previous sections. The primary goal of the
4665 measurements was to validate that an open, mode-filtered cavity suppressed the TM_{11}
4666 modes as predicted by HFSS simulations.

4667 **6.5.1 Cavities and Setup**

4668 Two rudimentary, cavities were constructed using segments of copper pipe available from
4669 McMaster-Carr (see Figure 6.15). The design consists of copper pipes of two diameters.
4670 The larger diameter pipe forms the main cavity wall and the smaller diameter pipe is
4671 used to create a coaxial termination. The diameter of the outer pipe was chosen to
4672 produce a TE_{011} resonance of approximately 6 GHz, while the diameter of the smaller
4673 pipe was selected based on the open termination criteria introduced in Section 6.4.1. The
4674 approximate diameters and lengths of the copper pipe are summarized in Table 6.4.

4675 Coupling to the cavity was achieved using a hand-formable segment of coaxial cable
4676 stripped at one end to form a loop antenna. This was inserted into a small hole located
4677 at the center of the main cavity wall. The coaxial terminators were supported inside the

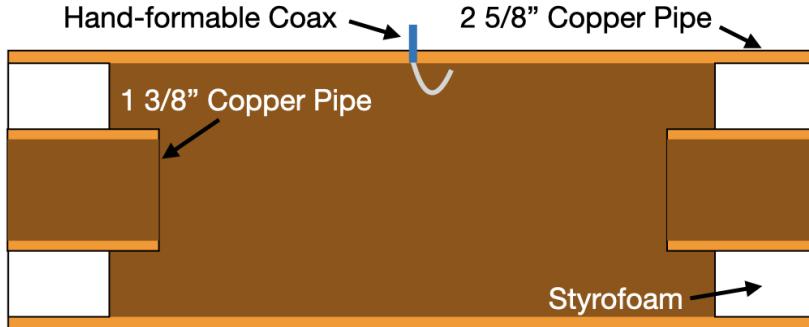


Figure 6.15: A cartoon depicting the design of the open-ended cavity prototype designed to operate at approximately 6 GHz. The main cavity wall was composed of a single copper pipe. A mode-filtered version of this cavity was constructed by

4678 main cavity by carving a spacer from polystyrene foam (styrofoam) so that they could
 4679 be easily inserted into the cavity and repositioned. The dielectric constant of styrofoam
 4680 is quite close to air at microwave frequencies so this is expected to have minimal impact
 4681 on the resonant properties of the cavity.

Table 6.4: A table of parameters describing the cavity prototypes. Certain values such as the cavity length and the distance between dielectric spacers are approximate due to variation in the machining of the copper. In particular, the filtered cavity was constructed from conducting copper segments that varied in size from 1.50" to 1.85".

Name	Qty.	Unit	Description
D_{cav}	2.625	in	Cavity diameter
L_{cav}	≈ 13	in	Cavity length
D_{term}	1.375	in	Inner diameter of coaxial terminator
L_{term}	1.575	in	Terminator length
l_{die}	0.75	in	Dielectric spacer thickness
Δl_{die}	≈ 1.50 to 1.85	in	Distance between dielectric spacers

4682 The actual length of the cavity is given by the distance between the inner edges of the
 4683 coaxial terminations. The length of the outer section of pipe that forms the main wall of
 4684 the cavity is approximately 16" in length which leads to a cavity length of $\approx 13"$ when
 4685 both terminators are inserted in the cavity. Because the terminators were not rigidly
 4686 mounted this distance is only approximate, however, the uncertain length of the cavity
 4687 will not prevent us from validating the open cavity design.

4688 Along with the smooth-walled open cavity a resistively mode-filtered cavity was
 4689 constructed by creating dielectric spacers out of segments of clear PVC pipe (see Figure
 4690 6.16). The spacers were machined such that the conductive segments of the cavity would

4691 be separated by 0.75" when the cavity was fully assembled. Due to variations in the
 4692 lengths of the copper segments that make up the cavity wall the distance between spacers
 4693 has significant variation with average value of about 1.7". Eight total spacers were used
 4694 to build the cavity, which when assembled was approximately 16" in total length similar
 to the non-filtered cavity.

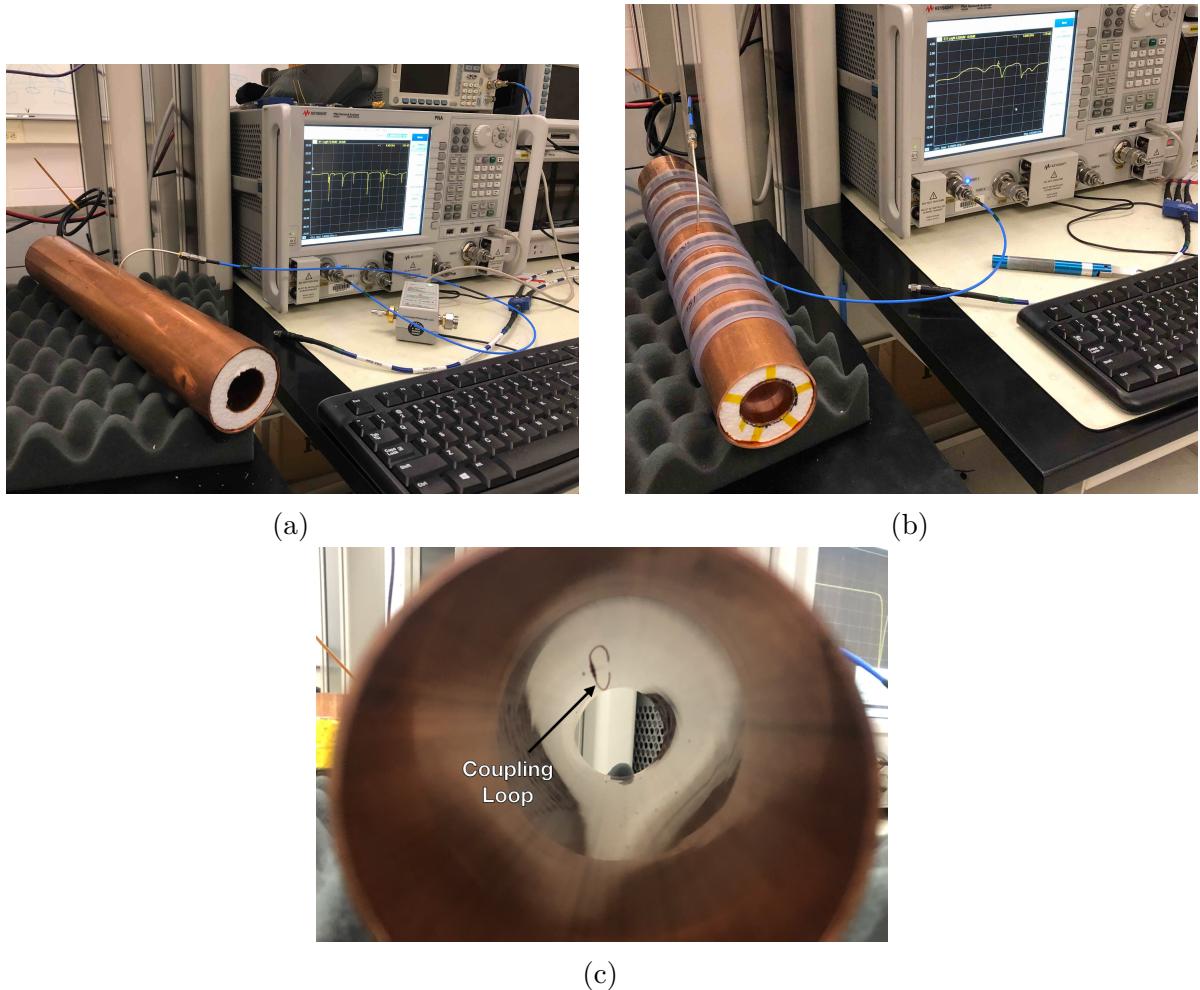


Figure 6.16: Images depicting the measurement of the filtered and non-filtered open cavities using the VNA. The coupling loop in the figure is shown in the TE orientation.

4695 Measurements of both cavities were performed using a VNA connected to the cavity
 4696 coupling probe (see Figure 6.16). By measuring the return loss over a range of frequencies
 4697 one can measure the frequencies and relative Q-factors of the resonant modes in the
 4698 cavity. Due to the opposite polarity of the electric fields for the TE and TM modes,
 4699 the loop coupling probe must be rotated 90° to change the polarity of the loop antenna.
 4700 When the antenna is oriented such that the loop opening faces the ends of the cavity, it

4702 couples primarily to the TE modes which have magnetic fields directed along the long
 4703 axis of the cavity (see Figure 6.16). If the coupling loop is turned by 90° from where
 4704 it is shown in the image then it will couple to the TM modes which have azimuthally
 4705 directed magnetic fields. In this way both the TE and TM resonances can be measured
 4706 independently.

4707 **6.5.2 Results and Discussion**

4708 The primary analysis method for the prototype cavities involved simply visualizing the
 4709 return loss measured by the VNA and comparing between the filtered and non-filtered
 4710 cavities. Since the resonances measured by the VNA are not labeled, there is some
 4711 uncertainty about the true identities of the modes measured by the VNA. To help with
 4712 this we performed a simulation of the simplest possible cavity that could be created from
 4713 the prototype components, which is a fully open cavity created by simply removing the
 4714 coaxial inserts from the non-filtered cavity configuration. The fully open cavity with the
 4715 as-built dimensions was simulated in HFSS to get estimates on the positions of the TE_{011}
 4716 and TM_{111} modes (see Figure 6.17).

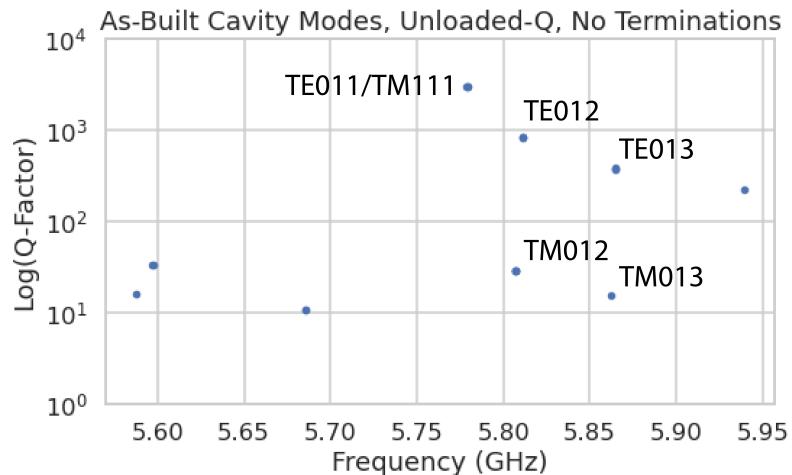


Figure 6.17: HFSS simulation results for a the as-built cavity with the coaxial terminators removed. The $\text{TE}_{011}/\text{TM}_{111}$ frequency is approximately 5.78 GHz.

4717 Simulation of the fully open cavity shows that the $\text{TE}_{011}/\text{TM}_{111}$ modes have a
 4718 frequency of approximately 5.78 GHz in the fully open cavity. If the frequency of this
 4719 mode is compared to the measurments of the fitered and non-filtered cavities with the
 4720 terminators removed we can easily identify the TE_{011} mode at approximately 5.75 GHz

4721 (see Figure 6.18).

4722 For the non-filtered cavity one sees that the TE_{011} mode is degenerate in frequency
 4723 with what appears to be a doublet of TM modes located at the TM_{111} frequency position.
 4724 This doublet is actually the TM_{111} mode, which has two polarizations with opposite
 4725 polarizations. Because the pipe used to construct the cavity is not perfectly round, the
 4726 frequency degeneracy between the two polarizations is broken producing the doublet
 peak. In the case of the filtered cavity with no terminators there is an isolated TE

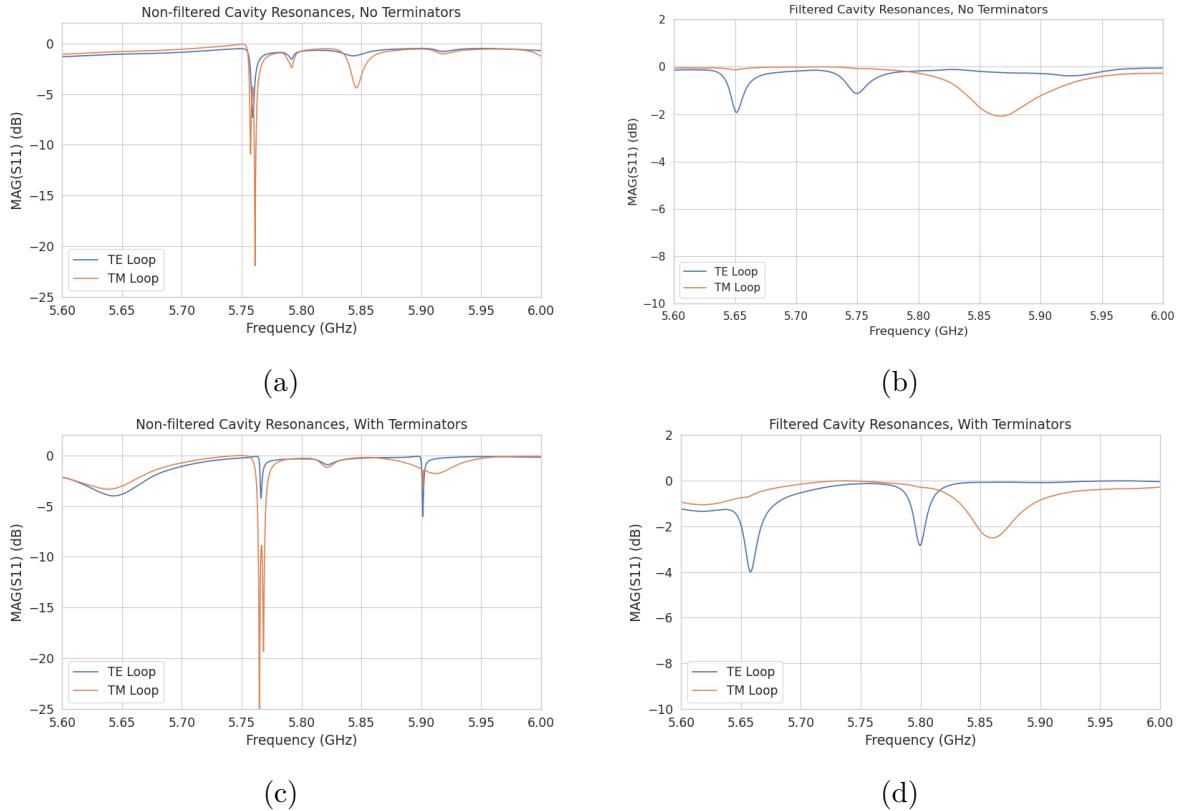


Figure 6.18: Measurements of the filtered and non-filtered prototype cavities acquired with the VNA.

4727
 4728 resonance at 5.75 GHz that appears to be the TE_{011} , however, there is no apparent TM_{111}
 4729 doublet at the same frequency. This is what one would expect if the mode-filtering was
 4730 effective at suppressing the TM modes. There is a notable difference in the Q of the
 4731 TE_{011} resonance for the non-filtered and filtered cavities indicated by the relative widths
 4732 of the resonances. This is likely caused by the large width of the dielectric spacers that
 4733 are partially impeding the TE modes. When the terminators are inserted into the cavity
 4734 one sees that Q-factors of the modes improves as expected, by noticing the narrowing of
 4735 the peaks compared to the no terminator plots.

⁴⁷³⁶ In conclusion, one see from these cavity measurements that, in principle, mode-
⁴⁷³⁷ filtering can be used to separate the TE₀₁₁ resonance from the degenerate TM₁₁₁ mode in
⁴⁷³⁸ combination with the an open cavity design. The ideal next step would be to construct a
⁴⁷³⁹ open, mode-filtered cavity that could be used to perform CRES measurements. In order
⁴⁷⁴⁰ to study the coupling of an electron to the isolated TE₀₁₁ mode.

4741 **Chapter 7 |**

4742 **Conclusion and Future Prospects**

4743 In this dissertation we have discussed research and development efforts towards the
4744 development of a scalable CRES measurement technology that can be used to build a
4745 CRES experiment at cubic-meter scales with sensitivity to neutrino masses of 40 meV.
4746 The primary contributions of my dissertation are the development and analysis of signal
4747 reconstruction algorithms for an antenna array based CRES experiment [67], which leads
4748 to estimates of the neutrino mass sensitivity; the development of a synthetic cyclotron
4749 radiation antenna (SYNCA) [29], which allowed for laboratory validation of antenna
4750 array CRES simulation models [6]; and the development of an open-ended cavity design
4751 compatible with atomic tritium for a cavity based CRES experiment. A measurable
4752 impact of this work is the transition of the Project 8 collaboration's experimental plan
4753 from an antenna array based approach to a cavity based approach, where my work played
4754 a key role in demonstrating the significantly higher cost and complexity of the antenna
4755 array experiment.

4756 The transition from antenna arrays to cavities requires a new set of demonstrator
4757 experiments to make incremental progress towards a 40 meV measurement of the neutrino
4758 mass. At the time of writing, the near-term plan of Project 8 is to design and construct a
4759 small-scale cavity CRES experiment utilizing the 1 T magnet installed in the UW-Seattle.
4760 This cavity is designed to have a TE011 resonance with a frequency of about 26 GHz with
4761 a length-to-diameter ratio that mimics the larger cavities intended for the pilot-scale and
4762 Phase IV experiments. The goal of this experiment is to demonstrate cavity CRES as
4763 well as validate models of CRES systematics using electrons from ^{83m}Kr and an electron
4764 gun. Though the primary goal is demonstration, near-term physics measurements are
4765 available in the form of high-resolution measurements of the ^{83m}Kr conversion spectrum
4766 of interest to the KATRIN collaboration.

4767 Furthermore, Project 8 is currently constructing a low-frequency CRES setup located
4768 at Yale University to better understand the principles of cavity based CRES at lower

4769 magnetic fields. The Low, UHF Cavity Krypton Experiment at Yale (LUCKEY) is
4770 a 1.5 GHz cavity CRES experiment the will use conversion electrons from ^{83m}Kr to
4771 perform CRES measurements at the lowest frequencies ever attempted with the technique.
4772 LUCKEY will validate frequency scaling models developed by Project 8 and will pave
4773 the way for the future Low-Frequency Apparatus (LFA), which will be a larger, 1 GHz
4774 cavity CRES experiment that includes a molecular tritium source. The target for the
4775 LFA is a measurement of the neutrino mass with a sensitivity of approximately 0.2 eV,
4776 which will build towards the atomic pilot-scale CRES experiment.

4777 In parallel to the development of cavity CRES is the development of the atomic
4778 tritium source. Recent demonstrations of the production of atomic hydrogen are excellent
4779 steps towards the atomic tritium production needed for the pilot-scale experiment. One
4780 area of future study includes the development of a more detailed unstanding of the
4781 efficiency of atomic hydrogen production. Near-term plans include the development of a
4782 magnetic, evaporatively cooled beamline, as well as the prototyping of a Halbach array
4783 atoms trap. Nearly all of the components of the atomic tritium system will require
4784 demonstration before the complete system can be built. The long-term goal of the
4785 atomic tritium work is to construct a full atomic tritium prototype that demonstrates
4786 the production, cooling, trapping, and recycling of tritium at the rates needed for the
4787 pilot-scale experiment.

4788 More broadly, the long-term goal of the Project 8 collaboration is to fully develop
4789 both the atomic tritium and cavity CRES technologies so that both can be combined in
4790 a pilot-scale CRES experiment. It is envisioned that this process will take approximately
4791 10 years for both atomic tritium and cavity CRES. After these developments comes
4792 the pilot-scale experiment which will be the first CRES experiment that simultaneously
4793 demonstrates all the required technologies for Phase IV. Scaling to Phase IV with cavity
4794 CRES will require the construction of multiple copies (approximately 10) of the pilot-scale
4795 experiment to obtain sufficient statistics for 40 meV sensitivity.

4796 Development of the CRES experimental technique by Project 8 has led to new
4797 experiments utilizing the CRES technique for basic physics research, such as the ^6He -
4798 CRES collaboration [68], and has also found applications as a new approach to x-ray
4799 spectroscopy [69]. Recently, a new experimental effort called CRESDA has begun in
4800 the UK to develop new quantum technologies applied to CRES measurements for the
4801 neutrino mass [70]. This flourishing of new experimental efforts based on the CRES
4802 technique is likely to continue as Project 8 continues to develop the technique towards
4803 its neutrino mass measurement goal.

Bibliography

- [1] FORMAGGIO, J. A., A. L. C. DE GOUVÊA, and R. G. H. ROBERTSON (2021) “Direct Measurements of Neutrino Mass,” *Phys. Rept.*, **914**, pp. 1–54, 2102.00594.
- [2] MONREAL, B. and J. A. FORMAGGIO (2009) “Relativistic cyclotron radiation detection of tritium decay electrons as a new technique for measuring the neutrino mass,” *Phys. Rev. D*, **80**, p. 051301.
URL <https://link.aps.org/doi/10.1103/PhysRevD.80.051301>
- [3] ESFAHANI, A. A., S. BÖSER, N. BUZINSKY, M. C. CARMONA-BENITEZ, C. CLAESSENS, L. DE VIVEIROS, P. J. DOE, S. ENOMOTO, M. FERTL, J. A. FORMAGGIO, J. K. GAISON, M. GRANDO, K. M. HEEGER, X. HUYAN, A. M. JONES, K. KAZKAZ, M. LI, A. LINDMAN, C. MATTHÉ, R. MOHIUDDIN, B. MONREAL, R. MUELLER, J. A. NIKKEL, E. NOVITSKI, N. S. OBLATH, J. I. PEÑA, W. PETTUS, R. REIMANN, R. G. H. ROBERTSON, G. RYBKA, L. SALDAÑA, M. SCHRAM, P. L. SLOCUM, J. STACHURSKA, Y. H. SUN, P. T. SURUKUCHI, J. R. TEDESCHI, A. B. TELLES, F. THOMAS, M. THOMAS, L. A. THORNE, T. THÜMMLER, W. V. D. PONTSEELE, B. A. VANDEVENDER, T. E. WEISS, T. WENDLER, and A. ZIEGLER (2022) “The Project 8 Neutrino Mass Experiment,” 2203.07349.
- [4] ESFAHANI, A. A., S. BÖSER, N. BUZINSKY, M. C. CARMONA-BENITEZ, C. CLAESSENS, L. DE VIVEIROS, P. J. DOE, M. FERTL, J. A. FORMAGGIO, J. K. GAISON, L. GLADSTONE, M. GUIQUE, J. HARTSE, K. M. HEEGER, X. HUYAN, A. M. JONES, K. KAZKAZ, B. H. LAROQUE, M. LI, A. LINDMAN, E. MACHADO, A. MARSTELLER, C. MATTHÉ, R. MOHIUDDIN, B. MONREAL, R. MUELLER, J. A. NIKKEL, E. NOVITSKI, N. S. OBLATH, J. I. PEÑA, W. PETTUS, R. REIMANN, R. G. H. ROBERTSON, D. R. D. JESÚS, G. RYBKA, L. SALDAÑA, M. SCHRAM, P. L. SLOCUM, J. STACHURSKA, Y. H. SUN, P. T. SURUKUCHI, J. R. TEDESCHI, A. B. TELLES, F. THOMAS, M. THOMAS, L. A. THORNE, T. THÜMMLER, L. TVRZNKOVA, W. V. D. PONTSEELE, B. A. VANDEVENDER, J. WEINTROUB, T. E. WEISS, T. WENDLER, A. YOUNG, E. ZAYAS, and A. ZIEGLER (2023) “Cyclotron Radiation Emission Spectroscopy of Electrons from Tritium Beta Decay and ^{83m}Kr Internal Conversion,” 2303.12055.
- [5] ESFAHANI, A. A., S. BÖSER, N. BUZINSKY, M. C. CARMONA-BENITEZ, C. CLAESSENS, L. DE VIVEIROS, P. J. DOE, M. FERTL, J. A. FORMAGGIO, J. K.

- GAISON, L. GLADSTONE, M. GRANDO, M. GUIGUE, J. HARTSE, K. M. HEEGER,
 X. HUYAN, J. JOHNSTON, A. M. JONES, K. KAZKAZ, B. H. LAROQUE, M. LI,
 A. LINDMAN, E. MACHADO, A. MARSTELLER, C. MATTHÉ, R. MOHIUDDIN,
 B. MONREAL, R. MUELLER, J. A. NIKKEL, E. NOVITSKI, N. S. OBLATH, J. I.
 PEÑA, W. PETTUS, R. REIMANN, R. G. H. ROBERTSON, D. R. D. JESÚS, G. RY-
 BKA, L. SALDAÑA, M. SCHRAM, P. L. SLOCUM, J. STACHURSKA, Y. H. SUN,
 P. T. SURUKUCHI, J. R. TEDESCHI, A. B. TELLES, F. THOMAS, M. THOMAS,
 L. A. THORNE, T. THÜMMLER, L. TVRZNIKOVA, W. V. D. PONTSEELE, B. A.
 VANDEVENDER, J. WEINTROUB, T. E. WEISS, T. WENDLER, A. YOUNG, E. ZA-
 YAS, and A. ZIEGLER (2023) “Tritium Beta Spectrum and Neutrino Mass Limit
 from Cyclotron Radiation Emission Spectroscopy,” 2212.05048.
- [6] “Antenna Arrays for Physics Measurements with Large-scale CRES Detectors,” *In preparation*.
- [7] FURSE, D. ET AL. (2017) “Kassiopeia: a modern, extensible C++ particle tracking package,” *New Journal of Physics*, **19**(5), p. 053012.
 URL <https://doi.org/10.1088/1367-2630/aa6950>
- [8] JACKSON, J. D. (1999) *Classical electrodynamics*, 3rd ed., Wiley, New York, NY.
 URL <http://cdsweb.cern.ch/record/490457>
- [9] ESFAHANI, A. A., V. BANSAL, S. BÖSER, N. BUZINSKY, R. CERVANTES,
 C. CLAESSENS, L. DE VIVEIROS, P. J. DOE, M. FERTL, J. A. FORMAGGIO,
 L. GLADSTONE, M. GUIGUE, K. M. HEEGER, J. JOHNSTON, A. M. JONES,
 K. KAZKAZ, B. H. LAROQUE, M. LEBER, A. LINDMAN, E. MACHADO, B. MON-
 REAL, E. C. MORRISON, J. A. NIKKEL, E. NOVITSKI, N. S. OBLATH, W. PETTUS,
 R. G. H. ROBERTSON, G. RYBKA, L. SALDAÑA, V. SIBILLE, M. SCHRAM, P. L.
 SLOCUM, Y.-H. SUN, J. R. TEDESCHI, T. THÜMMLER, B. A. VANDEVENDER,
 M. WACHTENDONK, M. WALTER, T. E. WEISS, T. WENDLER, and E. ZAYAS
 (2019) “Electron radiated power in cyclotron radiation emission spectroscopy experiments,” *Phys. Rev. C*, **99**, p. 055501.
 URL <https://link.aps.org/doi/10.1103/PhysRevC.99.055501>
- [10] ASHTARI ESFAHANI, A. ET AL. (2019) “Locust: C++ software for simulation of RF detection,” *New J. Phys.*, **21**, p. 113051, 1907.11124.
- [11] WIECHERT, E. (1901) “Elektrodynamische Elementargesetze,” .
 URL <https://doi.org/10.1002/andp.19013090403>
- [12] LIÉARD, A. (1898) “Champ électrique et Magnétique,” *Léclairage électrique*, **16**(27-29).
- [13] BALANIS, C. (2015) *Antenna Theory: Analysis and Design*, Wiley.
 URL <https://books.google.com/books?id=PTFcCwAAQBAJ>
- [14] <https://www.ansys.com/products/electronics/ansys-hfss>.

- 4875 [15] https://en.wikipedia.org/wiki/Linear_time-invariant_system.
- 4876 [16] NYQUIST, H. (1928) “Certain Topics in Telegraph Transmission Theory,” *Transactions of the American Institute of Electrical Engineers*, **47**(2), pp. 617–644.
- 4877
- 4878 [17] BRUN, R. and F. RADEMAKERS (1997) “ROOT: An object oriented data analysis
4879 framework,” *Nucl. Instrum. Meth. A*, **389**, pp. 81–86.
- 4880 [18] ASHTARI ESFAHANI, A. ET AL. (2021) “Bayesian analysis of a future β decay
4881 experiment’s sensitivity to neutrino mass scale and ordering,” *Phys. Rev. C*, **103**, p.
4882 065501.
4883 URL <https://link.aps.org/doi/10.1103/PhysRevC.103.065501>
- 4884 [19] KAY, S. (1998) *Fundamentals of Statistical Signal Processing: Detection Theory, Volume II*, Pearson.
- 4885
- 4886 [20] NEYMAN, J. and E. PEARSON (1933) “On the problem of the the most efficient
4887 tests of statistical hypotheses,” *Phil. Trans. R. Soc. Lond. A*, **231**.
- 4888 [21] STUMPF, M. (2018) *Electromagnetic Reciprocity in Antenna Theory*, Wiley.
- 4889 [22] BISHOP, C. (2016) *Pattern Recognition and Machine Learning*, Springer.
- 4890 [23] PLEHN, T., A. BUTTER, B. DILLON, and C. KRAUSE (2022) “Modern Machine
4891 Learning for LHC Physicists,” [2211.01421](https://arxiv.org/abs/2211.01421).
- 4892 [24] GEORGE, D. and E. A. HUERTA (2018) “Deep Learning for Real-time Gravitational
4893 Wave Detection and Parameter Estimation: Results with Advanced LIGO Data,”
4894 *Phys. Lett. B*, **778**, pp. 64–70, [1711.03121](https://arxiv.org/abs/1711.03121).
- 4895 [25] GABBARD, H., C. MESSENGER, I. S. HENG, F. TONOLINI, and R. MURRAY-SMITH
4896 (2022) “Bayesian parameter estimation using conditional variational autoencoders
4897 for gravitational-wave astronomy,” *Nature Phys.*, **18**(1), pp. 112–117, [1909.06296](https://arxiv.org/abs/1909.06296).
- 4898 [26] REIMANN, R. (2022) “Project 8: R&D for a next-generation neutrino mass experi-
4899 ment,” *PoS, PANIC2021*, p. 283.
- 4900 [27] LAROCHE, B. (2020) “Zero-deadtime processing in beta spectroscopy for measure-
4901 ment of the non-zero neutrino mass,” *EPJ Web Conf.*, **245**, p. 01014.
- 4902 [28] BUZINSKY, N. (2021) *Statistical Signal Processing and Detector Optimization in
4903 Project 8*, Ph.D. thesis, Massachusetts Institute of Technology.
- 4904 [29] ESFAHANI, A. A., S. BÖSER, N. BUZINSKY, M. CARMONA-BENITEZ,
4905 C. CLAESSENS, L. DE VIVEIROS, M. FERTL, J. FORMAGGIO, L. GLADSTONE,
4906 M. GRANDO, J. HARTSE, K. HEEGER, X. HUYAN, A. JONES, K. KAZKAZ,
4907 M. LI, A. LINDMAN, C. MATTHÉ, R. MOHIUDDIN, B. MONREAL, R. MUELLER,
4908 J. NIKKEL, E. NOVITSKI, N. OBLATH, J. PEÑA, W. PETTUS, R. REIMANN,

- 4909 R. ROBERTSON, L. SALDAÑA, P. SLOCUM, J. STACHURSKA, Y.-H. SUN, P. SU-
 4910 RUKUCHI, A. TELLES, F. THOMAS, M. THOMAS, L. THORNE, T. THÜMM-
 4911 LER, L. TVRZNIKOVA, W. V. D. PONTSEELE, B. VANDEVENDER, T. WEISS,
 4912 T. WENDLER, E. ZAYAS, A. ZIEGLER, and P. . COLLABORATION (2023) “SYNCA:
 4913 A Synthetic Cyclotron Antenna for the Project 8 Collaboration,” *Journal of Instru-*
 4914 *mentation*, **18**(01), p. P01034.
 4915 URL <https://dx.doi.org/10.1088/1748-0221/18/01/P01034>
- 4916 [30] <https://www.nvidia.com/en-us/data-center/v100/>.
- 4917 [31] <https://www.nvidia.com/en-us/data-center/h100/>.
- 4918 [32] HE, K., X. ZHANG, S. REN, and J. SUN (2016) “Deep Residual Learning for
 4919 Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern
 4920 Recognition (CVPR)*, pp. 770–778.
- 4921 [33] SIMONYAN, K. and A. ZISSERMAN (2015) “Very Deep Convolutional Networks
 4922 for Large-Scale Image Recognition,” in *3rd International Conference on Learning
 4923 Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference
 4924 Track Proceedings* (Y. Bengio and Y. LeCun, eds.).
 4925 URL <http://arxiv.org/abs/1409.1556>
- 4926 [34] FRIIS, H. (1946) “A Note on a Simple Transmission Formula,” *Proceedings of the
 4927 IRE*, **34**(5), pp. 254–256.
- 4928 [35] https://en.wikipedia.org/wiki/Friis_transmission_equation.
- 4929 [36] POZAR, D. M. (2005) *Microwave engineering*; 3rd ed., Wiley, Hoboken, NJ.
 4930 URL <https://cds.cern.ch/record/882338>
- 4931 [37] [https://www.mvg-world.com/en/products/absorbers/standard-absorbers/
 4932 convoluted-absorbers-aec-series](https://www.mvg-world.com/en/products/absorbers/standard-absorbers/convoluted-absorbers-aec-series).
- 4933 [38] [https://en.wikipedia.org/wiki/Network_analyzer_\(electrical\)](https://en.wikipedia.org/wiki/Network_analyzer_(electrical)).
- 4934 [39] [https://www.keysight.com/us/en/product/N5222A/
 4935 pna-microwave-network-analyzer-265-ghz.html](https://www.keysight.com/us/en/product/N5222A/pna-microwave-network-analyzer-265-ghz.html).
- 4936 [40] <https://www.pasternack.com/standard-gain-horn-antennas-category.aspx>.
- 4938 [41] <https://www.markimicrowave.com/home/>.
- 4939 [42] https://en.wikipedia.org/wiki/Power_dividers_and_directional_couplers.
- 4941 [43] <https://en.wikipedia.org/wiki/Balun>.
- 4942 [44] <https://www.rigolna.com/products/waveform-generators/dg5000/>.

- 4943 [45] https://en.wikipedia.org/wiki/Frequency_mixer.
- 4944 [46] <https://www.caen.it/>.
- 4945 [47] WORKMAN, R. L. and OTHERS (2022) “Review of Particle Physics,” *PTEP*, **2022**,
4946 p. 083C01.
- 4947 [48] FORMAGGIO, J. A., A. L. C. DE GOUVÊA, and R. G. H. ROBERTSON (2021)
4948 “Direct measurements of neutrino mass,” *Physics Reports*, **914**, pp. 1–54, direct
4949 measurements of neutrino mass.
4950 URL <https://www.sciencedirect.com/science/article/pii/S0370157321000636>
- 4952 [49] AKER, M. ET AL. (2022) “Direct neutrino-mass measurement with sub-electronvolt
4953 sensitivity,” *Nature Physics*, **18**(2), pp. 160–166.
4954 URL <https://doi.org/10.1038/s41567-021-01463-1>
- 4955 [50] MONREAL, B. and J. A. FORMAGGIO (2009) “Relativistic Cyclotron Radiation
4956 Detection of Tritium Decay Electrons as a New Technique for Measuring the Neutrino
4957 Mass,” *Phys. Rev. D*, **80**, p. 051301, 0904.2860.
- 4958 [51] ASNER, D. M. ET AL. (2015) “Single electron detection and spectroscopy via
4959 relativistic cyclotron radiation,” *Phys. Rev. Lett.*, **114**(16), p. 162501, 1408.5362.
- 4960 [52] ASHTARI ESFAHANI, A. ET AL. (2017) “Determining the neutrino mass with
4961 cyclotron radiation emission spectroscopy—Project 8,” *J. Phys. G*, **44**(5), p. 054004,
4962 1703.02037.
- 4963 [53] ESFAHANI, A. A. ET AL. (2022) “The Project 8 Neutrino Mass Experiment,” in
4964 *2022 Snowmass Summer Study*, 2203.07349.
- 4965 [54] BODINE, L. I., D. S. PARNO, and R. G. H. ROBERTSON (2015) “Assessment of
4966 molecular effects on neutrino mass measurements from tritium β decay,” *Phys. Rev.*
4967 *C*, **91**(3), p. 035505, 1502.03497.
- 4968 [55] ASHTARI ESFAHANI, A. ET AL. (2019) “Locust: C++ software for simulation of
4969 RF detection,” *New Journal of Physics*, **21**(11), p. 113051.
4970 URL <https://doi.org/10.1088/1367-2630/ab550d>
- 4971 [56] FURSE, D. ET AL. (2017) “Kassiopeia: a modern, extensible C++ particle tracking
4972 package,” *New Journal of Physics*, **19**(5), p. 053012.
4973 URL <https://doi.org/10.1088/1367-2630/aa6950>
- 4974 [57] BALANIS, C. (2011) *Modern Antenna Handbook*, Wiley.
4975 URL <https://books.google.com/books?id=UYpV8L8GNcwc>
- 4976 [58] WIRTH, W. (2001) *Radar Techniques Using Array Antennas*, Institution of Engi-
4977 neering and Technology.
4978 URL <https://books.google.com/books?id=ALht42gkzLsC>

- 4979 [59] BROWN, L. S., G. GABRIELSE, K. HELMERSON, and J. TAN (1985) “Cyclotron
 4980 motion in a microwave cavity: Lifetime and frequency shifts,” *Phys. Rev. A*, **32**, pp.
 4981 3204–3218.
 4982 URL <https://link.aps.org/doi/10.1103/PhysRevA.32.3204>
- 4983 [60] HANNEKE, D., S. FOGWELL HOOGERHEIDE, and G. GABRIELSE (2011) “Cavity
 4984 control of a single-electron quantum cyclotron: Measuring the electron magnetic
 4985 moment,” *Phys. Rev. A*, **83**, p. 052122.
 4986 URL <https://link.aps.org/doi/10.1103/PhysRevA.83.052122>
- 4987 [61] HANNEKE, D. A. (2007) *Cavity control in a single-electron quantum cyclotron: an
 4988 improved measurement of the electron magnetic moment*, Ph.D. thesis, Harvard U.
- 4989 [62] PURCELL, E. (1946) “Spontaneous Emission Probabilities at Radio Frequencies,”
 4990 *Phys. Rev.*, **69**, pp. 674–674.
 4991 URL <https://link.aps.org/doi/10.1103/PhysRev.69.674>
- 4992 [63] F.R.S., J. L. D. (1897) “LXIII. On the theory of the magnetic influence on
 4993 spectra; and on the radiation from moving ions,” *The London, Edinburgh, and
 4994 Dublin Philosophical Magazine and Journal of Science*, **44**(271), pp. 503–512, <https://doi.org/10.1080/14786449708621095>.
 4995 URL <https://doi.org/10.1080/14786449708621095>
- 4997 [64] MOSKOWITZ, B. E. and J. ROGERS (1988) “ANALYSIS OF A MICROWAVE
 4998 CAVITY DETECTOR COUPLED TO A NOISY AMPLIFIER,” *Nucl. Instrum.
 4999 Meth. A*, **264**, pp. 445–452.
- 5000 [65] WENGER, N. (1967) “Resonant Frequency of Open-Ended Cylindrical Cavity,”
 5001 *IEEE Transactions on Microwave Theory and Techniques*, **15**(6), pp. 334–340.
- 5002 [66] WENGER, N. C. and J. SMETANA (1972) “Hydrogen Density Measurements Using
 5003 an Open-Ended Microwave Cavity,” *IEEE Transactions on Instrumentation and
 5004 Measurement*, **21**(2), pp. 105–114.
- 5005 [67] “Real-time Signal Detection for Cyclotron Radiation Emission Spectroscopy Mea-
 5006 surements using Antenna Arrays,” *In preparation*.
- 5007 [68] BYRON, W. ET AL. (2022) “First observation of cyclotron radiation from MeV-scale
 5008 e^{pm} following nuclear beta decay,” [2209.02870](#).
- 5009 [69] KAZKAZ, K. and N. WOOLLETT (2021) “Using Cyclotron Radiation Emission
 5010 for Ultra-high Resolution X-Ray Spectroscopy,” *New J. Phys.*, **23**(3), p. 033043,
 5011 [1911.05869](#).
- 5012 [70] CANNING, J. A. L., F. F. DEPPISCH, and W. PEI (2023) “Sensitivity of future
 5013 tritium decay experiments to New Physics,” *JHEP*, **03**, p. 144, [2212.06106](#).