

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 | 16 |
| 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 The Project 8 Phased Development Plan | 30 |
| 47 | 3.3 Phase II: First Tritium Beta Decay Spectrum and Neutrino Mass Measurement with CRES | 34 |
| 48 | 3.3.1 The Phase II CRES Apparatus | 34 |
| 49 | 3.3.2 CRES Track and Event Reconstruction | 38 |
| 50 | 3.3.3 Results from Phase II | 42 |
| 51 | 3.4 Phase III R&D: Antenna Array CRES | 46 |

| | | | |
|----|-------|--|----|
| 53 | 3.4.1 | The Basic Approach | 47 |
| 54 | 3.4.2 | The FSCD: Free-space CRES Demonstrator | 48 |
| 55 | 3.5 | Pilot-scale Experiments | 53 |
| 56 | 3.5.1 | Choice of Frequency | 53 |
| 57 | 3.5.2 | Pilot-scale Experiment Concepts | 56 |
| 58 | 3.6 | Phase IV | 58 |

| | | | |
|----|--|--|-----------|
| 59 | Chapter 4 | | |
| 60 | Signal Reconstruction Techniques for Antenna Array CRES and the | | |
| 61 | FSCD | | 60 |
| 62 | 4.1 | Introduction | 60 |
| 63 | 4.2 | FSCD Simulations | 61 |
| 64 | 4.2.1 | Kassiopeia | 62 |
| 65 | 4.2.2 | Locust | 66 |
| 66 | 4.2.3 | CRESana | 70 |
| 67 | 4.3 | Signal Detection and Reconstruction Techniques for Antenna Array CRES | 71 |
| 68 | 4.3.1 | Digital Beamforming | 75 |
| 69 | 4.3.2 | Matched Filtering | 85 |
| 70 | 4.3.3 | Machine Learning | 97 |
| 71 | 4.4 | Analysis of Signal Detection Algorithms for the Antenna Array Demonstrator | 102 |
| 72 | 4.4.1 | Introduction | 103 |
| 73 | 4.4.2 | Signal Detection with Antenna Array CRES | 105 |
| 74 | 4.4.2.1 | Antenna Array and DAQ System | 105 |
| 75 | 4.4.2.2 | Real-time Signal Detection | 108 |
| 76 | 4.4.3 | Signal Detection Algorithms | 112 |
| 77 | 4.4.3.1 | Power Threshold | 112 |
| 78 | 4.4.3.2 | Matched Filtering | 114 |
| 79 | 4.4.3.3 | Machine Learning | 118 |
| 80 | 4.4.4 | Methods | 120 |
| 81 | 4.4.4.1 | Data Generation | 120 |
| 82 | 4.4.4.2 | Template Number and Match Estimation | 121 |
| 83 | 4.4.4.3 | CNN Training and Data Augmentation | 122 |
| 84 | 4.4.5 | Results and Discussion | 123 |
| 85 | 4.4.5.1 | Trigger Classification Performance | 123 |
| 86 | 4.4.5.2 | Computational Cost and Hardware Requirements | 125 |
| 87 | 4.4.6 | Conclusion | 127 |

| | | | |
|----|---|---|------------|
| 88 | Chapter 5 | | |
| 89 | Antenna and Antenna Measurement System Development for the | | |
| 90 | Project 8 Experiment | | 129 |
| 91 | 5.1 | Introduction | 129 |
| 92 | 5.2 | Antenna Measurements for CRES experiments | 130 |
| 93 | 5.2.1 | Antenna Parameters | 130 |
| 94 | 5.2.1.1 | Radiation Patterns | 130 |

| | | | |
|-----|---------|--|-----|
| 95 | 5.2.1.2 | Directivity and Gain | 131 |
| 96 | 5.2.1.3 | Far-field and Near-field | 132 |
| 97 | 5.2.1.4 | Polarization | 133 |
| 98 | 5.2.1.5 | Antenna Factor and Effective Aperture | 134 |
| 99 | 5.2.2 | Antenna Measurement Fundamentals | 136 |
| 100 | 5.2.2.1 | Friis Transmission Equation | 136 |
| 101 | 5.2.2.2 | S-Parameters and Network Analyzers | 137 |
| 102 | 5.2.2.3 | Antenna Array Commissioning and Calibration Measurements | 138 |
| 103 | 5.2.3 | The Penn State Antenna Measurement System | 140 |
| 104 | 5.3 | Development of a Synthetic Cyclotron Antenna (SYNCA) for Antenna Array Calibration | 143 |
| 105 | 5.3.1 | Introduction | 143 |
| 106 | 5.3.2 | Cyclotron Radiation Phenomenology | 146 |
| 107 | 5.3.3 | SYNCA Simulations and Design | 152 |
| 108 | 5.3.4 | Characterization of the SYNCA | 157 |
| 109 | 5.3.5 | Beamforming Measurements with the SYNCA | 160 |
| 110 | 5.3.6 | Conclusions | 164 |
| 111 | 5.4 | FSCD Antenna Array Measurements with the SYNCA | 165 |
| 112 | 5.4.1 | Introduction | 165 |
| 113 | 5.4.2 | Measurement Setups | 166 |
| 114 | 5.4.2.1 | FSCD Array Setup | 166 |
| 115 | 5.4.2.2 | Synthetic Array Setup | 168 |
| 116 | 5.4.3 | Simulations, Analysis, and Results | 169 |
| 117 | 5.4.3.1 | Simulations | 169 |
| 118 | 5.4.3.2 | Phase Analysis | 170 |
| 119 | 5.4.3.3 | Magnitude Analysis | 175 |
| 120 | 5.4.3.4 | Beamforming Characterization | 177 |
| 121 | 5.4.4 | Conclusions | 180 |

| | | | |
|-----|--|--|-----|
| 124 | Chapter 6 | | |
| 125 | Development of Resonant Cavities for Large Volume CRES Measurements | 182 | |
| 126 | 6.1 | Introduction | 182 |
| 127 | 6.2 | Cylindrical Resonant Cavities | 183 |
| 128 | 6.2.1 | General Field Solutions | 183 |
| 129 | 6.2.2 | TE and TM Modes | 184 |
| 130 | 6.2.3 | Resonant Frequencies of a Cylindrical Cavity | 186 |
| 131 | 6.2.4 | Cavity Q-factors | 188 |
| 132 | 6.3 | The Cavity Approach to CRES | 192 |
| 133 | 6.3.1 | A Sketch of a Molecular Tritium Cavity CRES Experiment | 192 |
| 134 | 6.3.2 | Magnetic Field, Cavity Geometry, and Resonant Modes | 194 |
| 135 | 6.3.3 | Trade-offs Between the Antenna and Cavity Approaches | 197 |
| 136 | 6.4 | Single-mode Resonant Cavity Design and Simulations | 199 |

| | | | |
|-----|--|---|------------|
| 138 | 6.4.1 | Open Cylindrical Cavities with Coaxial Terminations | 199 |
| 139 | 6.4.2 | Mode Filtering | 202 |
| 140 | 6.4.3 | Simulations of Open, Mode-filtered Cavities | 204 |
| 141 | 6.5 | Single-mode Resonant Cavity Measurements | 207 |
| 142 | 6.5.1 | Cavities and Setup | 207 |
| 143 | 6.5.2 | Results and Discussion | 210 |
| 144 | Chapter 7 | | |
| 145 | Conclusion and Future Prospects | | 213 |
| 146 | Bibliography | | 215 |

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.

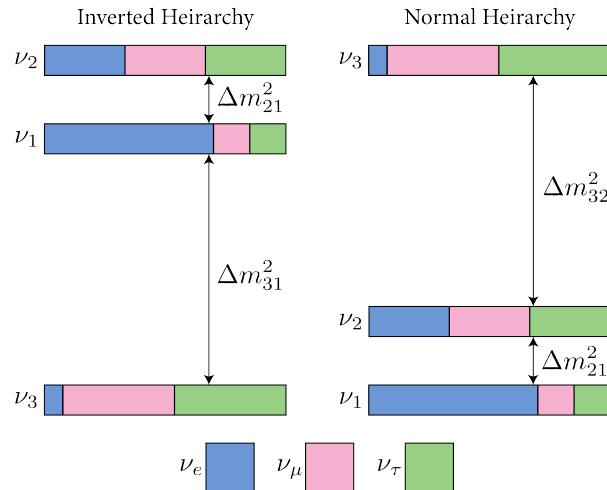


Figure 2.1: A diagram of two different neutrino mass ordering scenarios. In the inverted hierarchy (inverted mass ordering) the lightest neutrino mass is m_3 , whereas, in the normal hierarchy (normal mass ordering) m_1 is the lightest neutrino. What cannot be measured by neutrino oscillations is the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

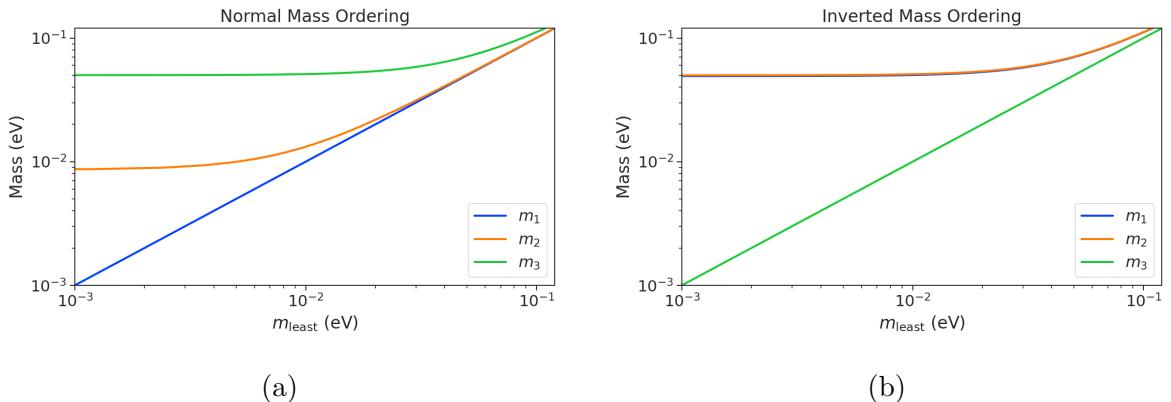


Figure 2.2: The masses of the neutrinos as a function of the lightest neutrino mass in both the normal (a) and inverted (b) mass ordering regimes.

2.4 Neutrino Masses in the Standard Model

In this section, I briefly summarize the current theoretical understanding of neutrino masses in the standard model [1–3]. Neutrinos are spin 1/2 particles, which are described using the Dirac equation.

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

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

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

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

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

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

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

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

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

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

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

⁴¹⁴ the neutrinos?

⁴¹⁵ An alternative approach is to allow the neutrinos to have a Majorana mass, which is
⁴¹⁶ possible because neutrinos are electrically neutral particles. The Majorana mass terms
⁴¹⁷ for the neutrino have the form

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

⁴¹⁸ where M_{Rij} and M_{Lij} are right-handed and left-handed Majorana mass matrices. A
⁴¹⁹ consequence of neutrinos being Majorana particles is lepton number violation, which
⁴²⁰ predicts the occurrence of neutrino-less double beta-decay at a rate proportional to the
⁴²¹ neutrino mass.

⁴²² In the most general case neutrinos have both Dirac and Majorana mass terms, which
⁴²³ allows one to generate neutrino masses with Yukawa couplings similar to the rest of
⁴²⁴ the standard model. Considering just one generation of neutrinos for illustration, the
⁴²⁵ combined Lagrangian can be written as

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

⁴²⁶ or equivalently,

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

⁴²⁷ An example mass generation mechanism with this approach is the Type-I see-saw
⁴²⁸ mechanism [4], in which we take $m_L = 0$ and $m_R \gg m_D$. By diagonalizing Equation 2.14
⁴²⁹ one obtains the mass eigenvalues that represent the physical masses of the neutrinos.
⁴³⁰ The light neutrino mass eigenstate, which represents the observed neutrino mass, has a
⁴³¹ mass given by

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

⁴³² and the heavy neutrino mass eigenstate, which represents the unobserved sterile neutrino,
⁴³³ has a mass

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

⁴³⁴ For m_D similar to the other quark or lepton masses, one obtains physical neutrino masses
⁴³⁵ consistent with observations from sterile neutrino masses of $m_R \approx O(10^{15})$ GeV. This
⁴³⁶ mass scale is well beyond the capabilities of modern particle accelerators.

2.5 Neutrino Absolute Mass Scale

⁴³⁷ The neutrino absolute mass scale or simply "neutrino mass" cannot be probed with
⁴³⁹ neutrino oscillations, since oscillation probabilities are determined by the squared mass
⁴⁴⁰ differences between neutrino mass eigenstates, therefore, alternative techniques are needed
⁴⁴¹ to perform an effective measurement of the neutrino mass.

442 2.5.1 Limits from Cosmology

⁴⁴³ The Λ CDM model summarizes our current cosmological understanding of our universe [5].
⁴⁴⁴ Λ CDM predicts that the universe originated from a single expansion event colloquially
⁴⁴⁵ called the "Big Bang". During the Big Bang, the universe originated as a hot spacetime
⁴⁴⁶ singularity, which abruptly experienced rapid expansion in a process known as inflation.
⁴⁴⁷ After expansion the inflationary field eventually decayed into a population of quarks,
⁴⁴⁸ gluons, leptons, and photons, which were kept in thermal equilibrium by the high-
⁴⁴⁹ temperatures of the early universe.

⁴⁵⁰ As the universe continued to expand its density and temperature decreased until
⁴⁵¹ the formation of neutral atoms, primarily hydrogen, was possible. At which point the
⁴⁵² population of photons produced during the Big Bang thermally decoupled. A direct
⁴⁵³ prediction of the Λ CDM model is that this population of photons should still be present,
⁴⁵⁴ but with a significantly reduced temperature due to the expansion of the universe. This
⁴⁵⁵ is consistent with the observation of the CMB (cosmic microwave background), which is
⁴⁵⁶ a population of microwave radiation with a blackbody temperature of 2.7 K. The CMB
⁴⁵⁷ is extremely uniform in all directions with slight anisotropies that can be analyzed to
⁴⁵⁸ study the evolution of the early universe. A series of experiments have measured the
⁴⁵⁹ CMB with increasing levels of precision, which has lead to a significant increase in our
⁴⁶⁰ current understanding of cosmology.

⁴⁶¹ In addition to the CMB, inflation predicts the existence of a $C\nu B$ (cosmic neutrino
⁴⁶² background) [6], which are the remnant neutrinos produced during the Big Bang. Since
⁴⁶³ neutrinos only interact via the weak force, they decouple from the hot Big Bang plasma
⁴⁶⁴ at an earlier time than the CMB radiation. The temperature at which the $C\nu B$ decouples
⁴⁶⁵ depends on the neutrino rest mass. Neutrinos play a unique role in the Λ CDM model,
⁴⁶⁶ due to the fact that neutrinos act as radiation early in the universe but as matter in the
⁴⁶⁷ late universe. This leads to specific signatures that impact the expected anisotropies
⁴⁶⁸ of the CMB as well as the distribution of matter in the universe [7]. By combining
⁴⁶⁹ measurements of the CMB with measurements of the large-scale structure (LSS) of the

universe one can constrain the neutrino mass scale by fitting these datasets with the Λ CDM model. This analysis results in some of the most stringent constraints on the neutrino mass. Recent analyses [5] have been able to constrain the neutrino mass scale to

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

where m_i are the neutrino mass eigenstates.

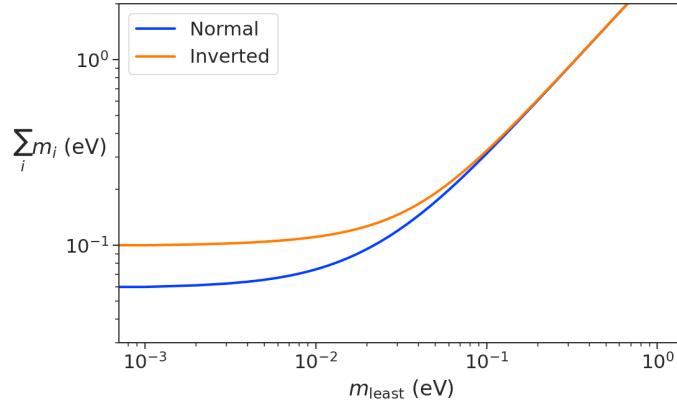


Figure 2.3: The neutrino mass observable measured by cosmology as a function of the lightest neutrino mass eigenstate.

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

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

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

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

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

Upcoming experiments [8] are planned to refine measurements of the CMB, LSS, and other cosmological observables. With this additional data it is possible that in the near future cosmological measurements will be able to positively constrain the neutrino absolute mass scale. However, the strength of these limits strictly depend on the accuracy of the Λ CDM model, which highlights the need for direct experimental measurements of

485 the neutrino mass to confirm the predictions of cosmology and to fix the neutrino mass
 486 parameter in future cosmological analyses.

487 2.5.2 Limits from Neutrinoless Double Beta-decay Searches

488 If neutrinos are Majorana fermions, then the neutrino is equivalent to its own antiparticle
 489 and lepton conservation is not an exact law of nature [9]. Limits on the rate of neutrinoless
 490 double beta-decay ($0\nu\beta\beta$), are some of the most powerful current tests of lepton number
 491 conservation [5]. If $0\nu\beta\beta$ were observed it would direct evidence that neutrinos are
 492 Majorana fermions, and provide a method for measuring the neutrino mass scale.

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

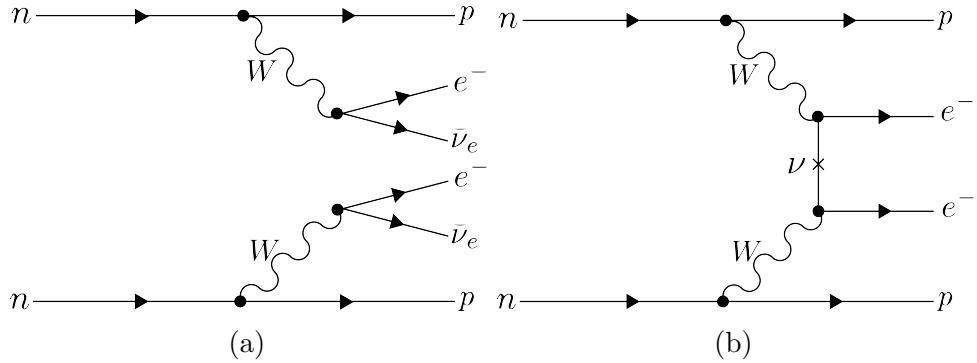


Figure 2.4: Feynman diagrams for double beta-decay (a) and $0\nu\beta\beta$ (b).

495
 496 producing only two electrons, which violates lepton number by two.

497 Assuming that the exchange of two Majorana neutrinos is the dominant channel for
 498 $0\nu\beta\beta$, then a measurement of the $0\nu\beta\beta$ half-life for a particular isotope can be used to
 499 set the neutrino absolute mass scale [10]. The half-life is written in terms of the effective
 500 neutrino mass for $0\nu\beta\beta$ ($m_{\beta\beta}$) using the equation

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

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

503 by the PMNS mixing matrix parameters,

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

504 The information provided from $0\nu\beta\beta$ on the neutrino mass scale can be visualized
 505 by expressing the value of $m_{\beta\beta}$ in terms of m_{least} and two relative Majorana phases [11].
 506 The allowed regions for $m_{\beta\beta}$ as a function of m_{least} are shown in Figure 2.5 as the regions
 507 bounded by the black curves overlayed with the discovery probabilities of future $0\nu\beta\beta$
 decay experiments based on current neutrino data.

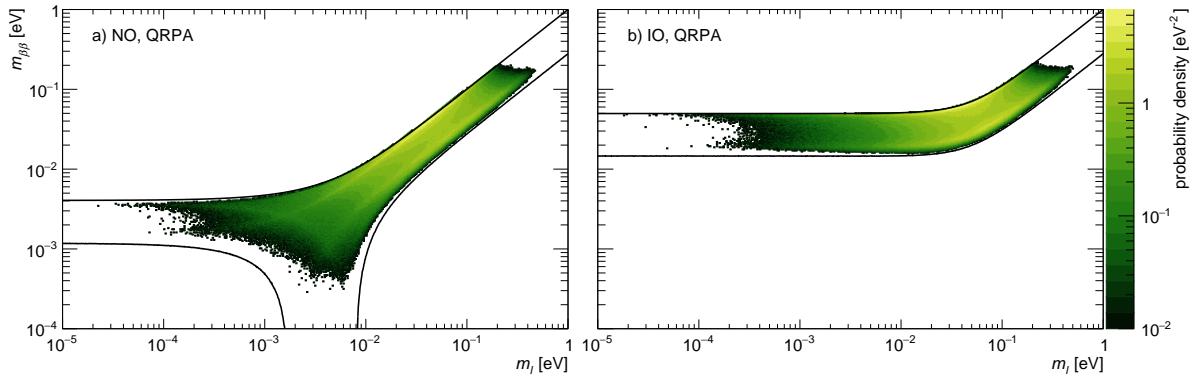


Figure 2.5: The discovery probabilities for the future generation of $0\nu\beta\beta$ experiments as a function of $m_{\beta\beta}$ and m_{least} . Figure from [11].

508
 509 Because of the possibility of cancellation due to the unknown Majorana phases
 510 included in the sum specified by Equation 2.21, the information gained is necessarily
 511 imperfect. Additionally, theoretical uncertainties in the calculation of the nuclear matrix
 512 elements complicates the calculation of $m_{\beta\beta}$ from a measurement of $0\nu\beta\beta$ half-life. Similar
 513 to cosmology, there is a high degree of complementarity between direct measurements
 514 of the neutrino mass and $0\nu\beta\beta$. In particular, a measurement of m_{least} to less than
 515 than 0.1 eV sensitivity provides significant information for $0\nu\beta\beta$ searches based of the
 516 discovery probabilities of Figure 2.5.

517 2.5.3 Limits from Beta-decay

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

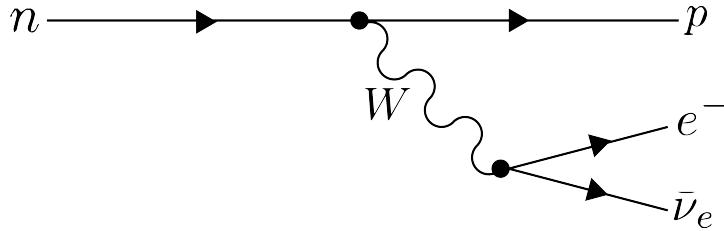


Figure 2.6: A Feynman diagram of beta decay

522 neutrino ("neutrino" for brevity) to conserve charge and lepton number [cite das and
 523 ferbel]. Therefore, by applying the principles of energy and momentum conservation, a
 524 measurement of the kinematics of the final state can be used to constrain the neutrino
 525 mass [12].

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

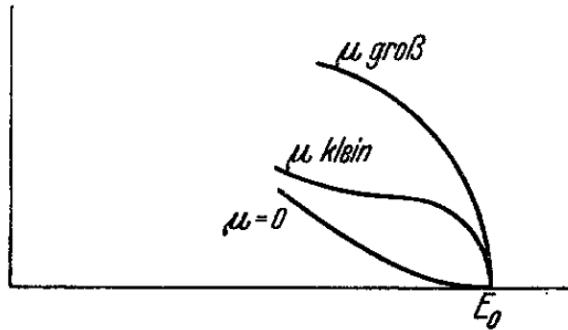


Figure 2.7: A figure from Fermi's 1934 paper on a theory of beta-decay depicting the kinetic energy spectrum of the emitted electron. The effect of the neutrino mass, written as μ , is to distort the shape of the spectrum near the endpoint from the zero-mass spectrum.

527
 528 the neutrino mass from beta-decay depend only on the final state measurement capabilities
 529 and the principles of energy and momentum conservation, neutrino mass measurements
 530 with beta-decay are sometimes called direct measurements. A direct measurement like
 531 beta-decay contrasts with other neutrino mass measurements approaches that are model-
 532 dependent such as cosmology and $0\nu\beta\beta$, which provide complementary ways to study
 533 the physics of massive neutrinos.

534 The isotope of choice for direct neutrino mass measurements with beta-decay has
 535 been tritium (3H_2) for many decades, because it conveniently fulfills many experimental
 536 requirements. Of upmost importance is a decay with a low Q-value, which is the available

537 kinetic energy based on the mass difference between the initial and final states. The
 538 effect of a massive neutrino on the shape of the spectrum is magnified for low Q-values
 539 and tritium decays have an unusually low Q-value of 18.6 keV.

540 Additionally, tritium beta-decay is a super-allowed decay, which results in a relatively
 541 short half-life of 12.3 years. Therefore, it is relatively easy to obtain a high-activity
 542 using a small source mass. High-activity is desireable because of the low-activity near
 543 the tritium spectrum endpoint. For tritium beta-decays only a factor of 3×10^{-13} of
 544 the decays occur in the last 1 eV of the spectrum. Isotopes with Q-values lower than
 545 tritium are known [12], but this is outweighed by exceedingly long half-lives leading to
 546 unobtainable source masses.

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

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

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

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

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

560 where m_β is the electron-weighted neutrino mass or simply "neutrino mass" for brevity.
 561 m_β corresponds to a particular weighted sum of the neutrino masses, which is distinct
 562 from effective neutrino masses such as $m_{\beta\beta}$ [12]. Assuming unitarity, the neutrino mass
 563 can be expressed in terms of the PMNS matrix elements, squared mass differences, and

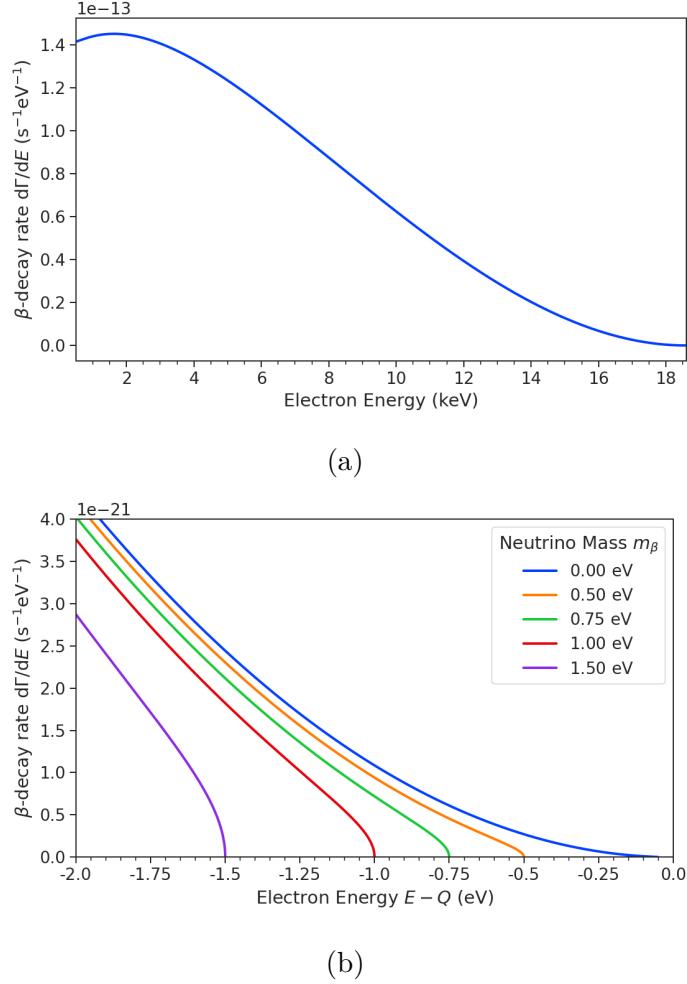


Figure 2.8: The tritium beta-decay spectrum. The affect of a massive neutrino on the spectrum is to change it's shape near the endpoint by an amount proportional to the size of the neutrino mass. This suggests that a sufficiently high-statistic and high-resolution measurement of the spectrum endpoint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

⁵⁶⁴ the lightest neutrino mass eigenstate. For the normal mass ordering the equation is

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

⁵⁶⁵ and for the inverted ordering the equation is

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

⁵⁶⁶ Therefore, a measurement of the neutrino mass in combination with neutrino mixing
⁵⁶⁷ parameters is effectively a measurement of m_{least} .

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

573 **Chapter 3** |

574 **Direct Measurement of the Neutrino Mass**

575 **with Project 8**

576 **3.1 Introduction**

577 A promising technique for direct measurements of the neutrino mass beyond the projected
578 limit of the ongoing KATRIN experiment [13] is tritium beta-decay spectroscopy with an
579 atomic tritium source [14]. Atomic tritium, combined with a large-volume, high-resolution
580 energy measurement technique, is capable of measuring the neutrino mass with sensitivity
581 below the 50 meV limit allowed by neutrino oscillations.

582 Cyclotron Radiation Emission Spectroscopy or CRES is a high-resolution energy
583 measurement technique compatible with atomic tritium production and storage that can
584 enable the next-generation of neutrino mass direct measurement experiments [15]. The
585 Project 8 collaboration is currently engaged in a program of research and development
586 (R&D) aimed at developing the technology necessary for a 40 meV sensitivity measurement
587 of the neutrino mass using CRES and atomic tritium [16].

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

592 In Section 3.3 I give a brief overview of Phase II of the Project 8 experiment [17, 18],
593 which completed early in 2023. Although the bulk of the work presented in this thesis is
594 relevant to designs of future Project 8 experiments, a description of the work in Phase II
595 provides useful context for the rest of the work.

596 In Section 3.4 I introduce a CRES measurement concept based on antenna arrays [19],
597 which could be the basis for the ultimate Project 8 neutrino mass experiment. A
598 significant portion of the R&D efforts of Project 8 in Phase III were directed towards

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

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

606 **3.2 Cyclotron Radiation Emission Spectroscopy and Project** 607 **8**

608 **3.2.1 Cyclotron Radiation Emission Spectroscopy — CRES**

609 Time and frequency are two of the most precisely measured quantities in physics. It is
610 often advantageous to convert measurements of other physical quantities like mass or
611 length into frequency measurements due to the digital nature of frequency measurements
612 that make them immune to many sources of noise. Atomic clocks, which operate by
613 measuring the frequencies of various atomic transitions, have been used to measure
614 time with astounding relative uncertainties of 10^{-18} seconds [20]. The extreme precision
615 possible with frequency measurements is often summarized using the a quote from the
616 Physicist Arthur Schawlow who said advise his students to "Never measure anything but
617 frequency!" [21].

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

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

631 emission of cyclotron radiation that can be detected using an array of antennas or a
 632 different RF sensor such as a resonant cavity. The frequency of the radiation gives the
 633 electron's kinetic energy, which is used to build the beta-decay spectrum and measure
 634 the neutrino mass. The name for this measurement technique is Cyclotron Radiation
 635 Emission Spectroscopy or CRES [15].

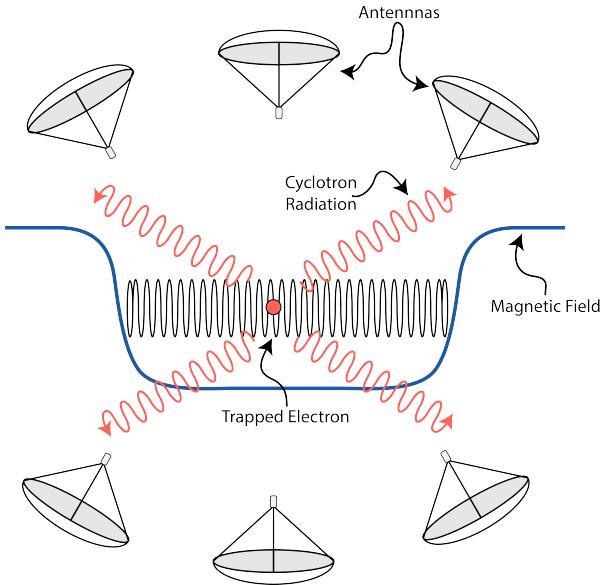


Figure 3.1: A cartoon illustration of the CRES technique. An electron is contained in a magnetic trap so that its cyclotron radiation can be detected by an array of antennas. Detecting the cyclotron radiation allows us to measure its cyclotron frequency and determine its kinetic energy.

636 For non-relativistic particles the cyclotron frequency is simply a function of the
 637 charge-to-mass ratio of the particle, however, from the relativistic form of the cyclotron
 638 frequency

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

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

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

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

644 from which we can obtain the relationship between fractional differences in energy and
645 frequency,

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

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

648 The minimum observation time required to achieve this resolution can be estimated
649 using the uncertainty principle as formulated by Gabor [22]. Electrons from tritium
650 beta-decay experience random collisions with the background gas particles, which limits
651 the uninterrupted radiation lifetime. The time between collision events, referred to
652 as track length in the context of CRES measurements, is an exponentially distributed
653 variable. Differences in the track lengths of a population of mono-energetic electrons leads
654 to uncertainty or broadening in the distribution of measured frequencies proportional to
655 the mean track length, τ_λ . The resulting frequency distribution has a Lorentzian profile,
656 whose width is given by the Gabor limit,

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

657 The cyclotron frequency for a 18.6-keV electron in a 1 T field is approximately
658 27 GHz, from which one can estimate the minimum observation time for 2 ppm frequency
659 resolution at approximately 3 μ sec. The Gabor limit is not the true lower bound on the
660 frequency resolution for a CRES signal, since it is based on the details of the Fourier
661 representation of a time-series with a fixed length. If one takes the approach of fitting
662 the CRES signal in the time-domain, then one finds that the limit on frequency precision
663 is given by the Cramér-Rao lower bound (CRLB) [23], which depends on the track length
664 and SNR. The CRLB allows for better precision on the cyclotron frequency, however,
665 the Gabor limit provides an intuitive limit with the correct order of magnitude.

666 Ensuring that an electron remains under observation long enough so that it's frequency
667 can be precisely measured requires a magnetic trap. A magnetic trap is a local minimum
668 in a background magnetic field generated an appropriate configuration of electromagnetic
669 coils. Since magnetic fields can do no work, there is no danger of the magnetic trap
670 affecting the kinetic energy electron after it is emitted from the beta-decay. One common
671 approach to creating a magnetic trap is the "bathtub" trap configuration, which in it's
672 simplest form consists of two high magnetic field pinch coils aligned on a central axis
673 that are well separated (see Figure 3.2). This configuration produces a trap with a flat
674 uniform bottom and relatively steep walls, which is ideal for CRES measurements.

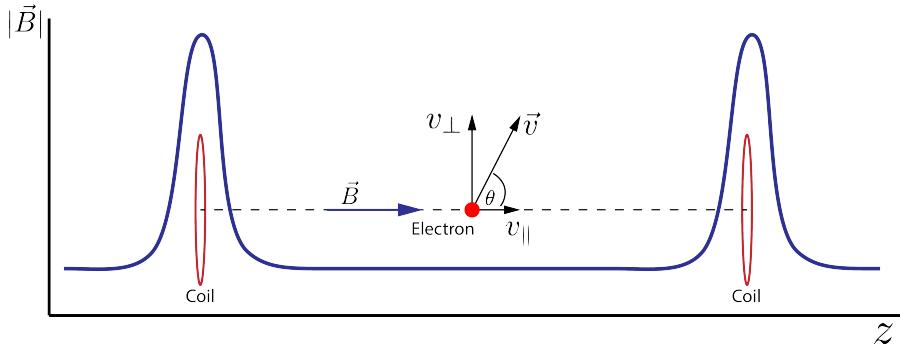


Figure 3.2: An illustration of an electron in a bathtub magnetic trap generated by two well-separated coils.

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

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

680 The axial motion of the electron leads to variation in the cyclotron frequency due to
 681 the changing value of the magnetic fields. This leads to frequency modulation that
 682 generate sidebands in the cyclotron radiation spectrum. Resolving these sideband
 683 frequency components is necessary for a complete reconstruction of the CRES signal in
 684 the experiment.

685 Electrons trapped in a cylindrically symmetric trap have three primary components of
 686 motion (see Figure 3.3). The dominant component, typically with the highest frequency,
 687 is the electron's cyclotron orbit, which encodes information on the electron's kinetic
 688 energy. Axial motion from the electron's pitch angle leads to frequency modulation but
 689 also a shift in the average magnetic field experienced by an electron. This leads to a
 690 correlation between the kinetic energy of the electron and the pitch angle depending on
 691 the particular shape of the magnetic trap, which can negatively impact energy resolution.
 692 To reduce this correlation one must engineer the trap to have a flat bottom with very
 693 steep wall both of which are more easily achieved with a small aspect ratio bathtub trap.
 694 Radial gradients in the trap leads to a third component of motion called grad-B drift [24].

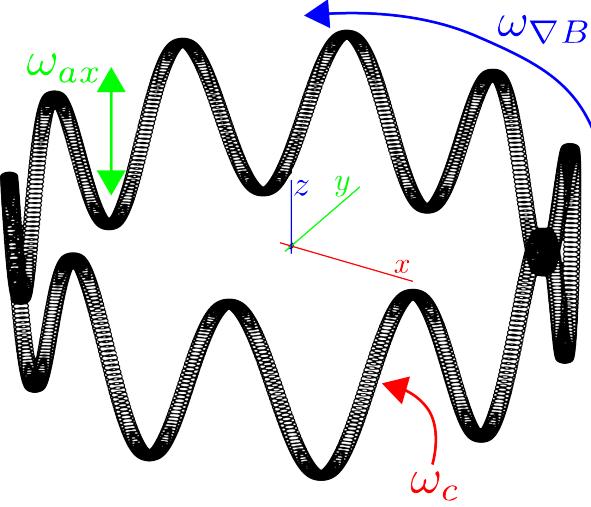


Figure 3.3: A plot of the main components of an electron's trajectory in a cylindrically symmetric trap.

695 The equation for the drift velocity is

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

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

698 The total power of the radiation emitted by an electron in a free-space environment
 699 is given by the Larmor equation [25]

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

700 where ω_c is the cyclotron frequency multiplied by 2π and θ_p is the pitch angle to distinguish
 701 it from the spherical angle coordinate. A single electron with a 90° pitch angle and
 702 18.6 keV of kinetic energy in a 1 T magnetic field emits a total radiation power of 1.2 fW,
 703 which is quite small compared with typical RF systems, furthermore, one is typically
 704 only able to receive a fraction of this total power with an antenna or other detection
 705 system. Therefore, RF systems in CRES experiments must be operated at cryogenic
 706 temperatures to limit the noise power such that adequate SNR can be achieved for signal
 707 detection and reconstruction. Alternatively, longer tracks enable detection of weaker
 708 signals due to the increase in the total signal energy available for the detection algorithm.

709 3.2.2 The Project 8 Collaboration

710 The Project 8 collaboration¹ is a group of institutions in the United States and Germany
711 aiming to measure the neutrino mass by developing a novel spectrometer technology
712 based on CRES. In the ultimate Project 8 experiment the CRES technique will be used
713 to measure the beta-decay spectrum using a large source of atomic tritium sufficient to
714 achieve the required statistics in the last $O(10)$ eV of the decay spectrum. Project 8 is
715 targeting a neutrino mass sensitivity below 50 meV [26], which exhausts the range of
716 possible neutrino masses under the inverted hierarchy and is a factor of four less than
717 sensitivity projections for the ongoing KATRIN experiment.

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

721 Atomic Tritium

722 Previous measurements of the tritium beta-decay spectrum for neutrino mass measure-
723 ments have relied on sources of molecular tritium for their measurements [13, 27, 28] due
724 to the technical challenges associated with the production and storage of atomic tritium.

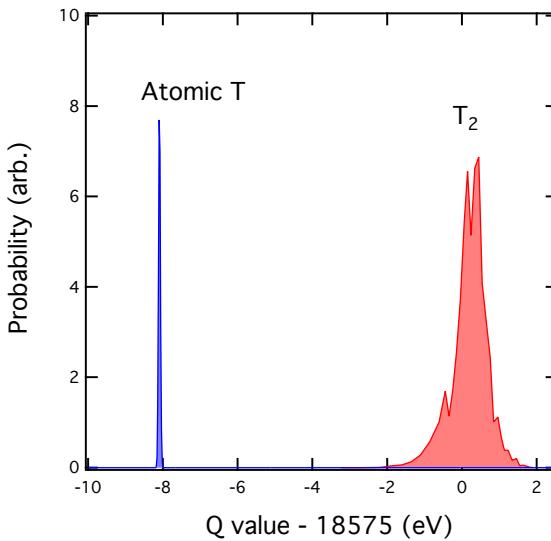


Figure 3.4: A plot of the final state distributions of atomic and molecular tritium. The final state distribution provides the primary contribution to the width of the molecular spectrum whereas thermal doppler broadening is responsible for the width of the atomic spectrum.

¹<https://www.project8.org/>

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

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

737 The significant experimental complexity caused by atomic tritium makes a molecular
738 source the obvious choice from practical considerations. However, the drawback of
739 molecular tritium for neutrino mass measurement is the irreducible broadening in the
740 electron's kinetic energy due to the final state spectrum of molecular tritium (see Figure
741 3.4). The broadening of the final state spectra has a RMS amplitude of 436 meV [29, 30]
742 caused by variation in the final vibrational state of the daughter molecule. For atomic
743 tritium the primary sources of broadening in the final state spectrum are magnetic
744 hyperfine splittings (magnitude of $O(10^{-5})$ eV) and thermal Doppler broadening caused
745 by the motion of the trapped atom. For atomic tritium at a temperature of 1 mK thermal
746 broadening is the dominant contribution, providing about 1 meV RMS of broadening to
747 the electron's kinetic energy.

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

753 **CRES for Neutrino Mass Measurement**

754 Several features of the CRES technique make it an attractive choice for a next generation
755 neutrino mass measurement experiment. For example, with a CRES experiment the
756 volume of the source gas can be the same as the volume of the CRES spectrometer.
757 This is due to the fact that CRES is a remote-sensing technique that can observe the
758 energy of the electron without altering its trajectory or directly interacting with the

759 electron. Given that tritium gas is transparent to cyclotron radiation the kinetic energies
760 of electrons can be measured with an appropriate sensing technology, such as a cavity or
761 antenna array, located directly outside the atom trapping volume.

762 The current state-of-the-art tritium beta-decay spectroscopy experiment, KATRIN,
763 utilizes the magnetic adiabatic collimation with an electrostatic filter (MAC-E filter)
764 technique to measure the beta-decay spectrum of molecular tritium. In this approach,
765 a source of molecular tritium is located outside the spectrometer. When a beta-decay
766 occurs the electron must exit the tritium source and travel through the MAC-E filter
767 before it can be detected on the other side of the filter using a charge sensor. The
768 measurement statistics of the MAC-E filter are limited by the transverse areas of the
769 tritium source and the filter due to the need to travel through the experiment without
770 scattering. This scaling is less favorable than the volumetric scaling of CRES due to the
771 ability to co-locate source and detector.

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

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

787 Neutrino Mass Sensitivity Goals

788 Project 8's ultimate goal is to combine CRES with atomic tritium to measure the neutrino
789 mass with 40 meV sensitivity at the 90% confidence level (see Figure 3.5). This sensitivity
790 is sufficient to fully exhaust the range of allowable neutrino masses under the inverted
791 neutrino mass ordering regime and is approximately an order of magnitude less than the
792 projected final sensitivity of the KATRIN experiment. Excluding the full neutrino mass

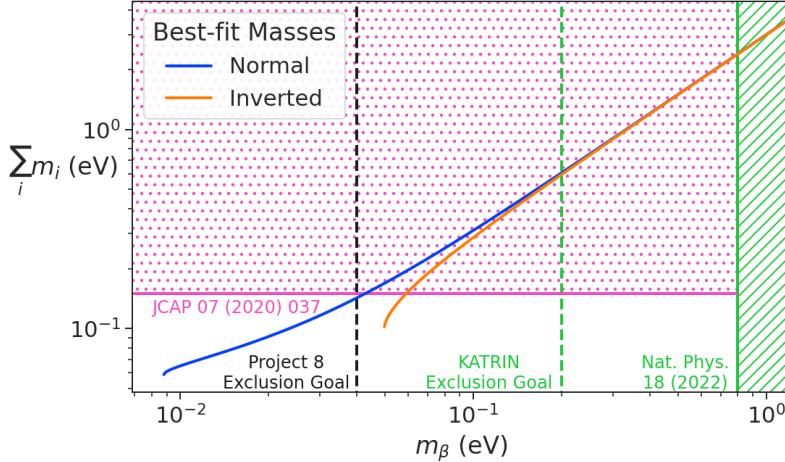


Figure 3.5: Neutrino mass exclusion plot including limits from cosmological measurements and the KATRIN experiment. Allowed ranges for neutrino masses under the normal and inverted hierarchies are shown as the blue and orange lines respectively. The black dashed line shows Project 8’s goal neutrino mass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

793 parameter space would require a sensitivity an order of magnitude lower than what is
 794 proposed by Project 8, which would require an experiment whose size and complexity
 795 are currently well beyond proposals for the next-generation of neutrino mass direct
 796 measurement experiments.

797 3.2.3 The Project 8 Phased Development Plan

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

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

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

808 In Phase I, Project 8 performed a proof-of-principle measurement of the ^{83m}Kr

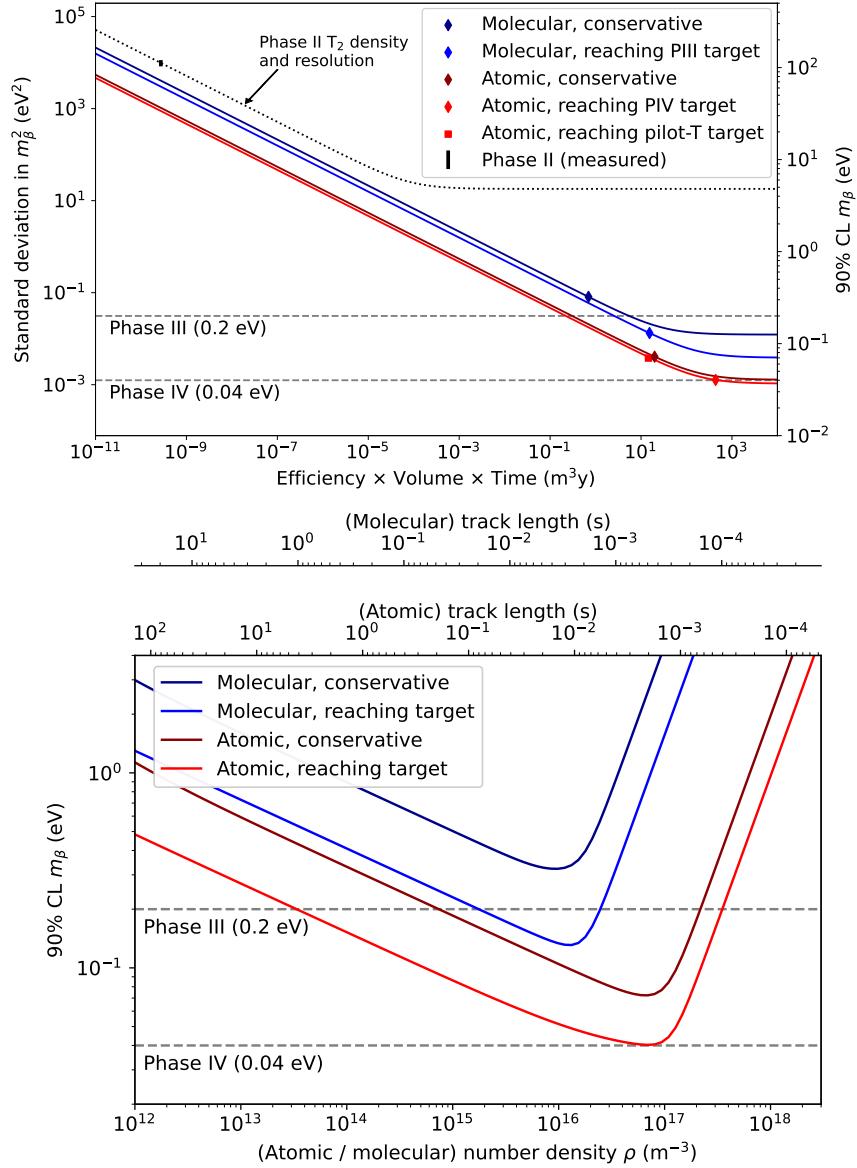


Figure 3.6: Sensitivity calculations for a cavity based CRES experiment that demonstrate the neutrino mass measurement goals of the Project 8 collaboration throughout the phased development plan. The blue curves indicate molecular tritium sources and the red curves indicate atomic tritium sources. In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

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

Success in Phase I was achieved with the 2014 publication of the measured ^{83m}Kr conversion spectrum [31], which contains a mono-energetic 17.8-keV as well as several other conversion lines at higher energies. Publication of this result marked the official end of Phase I and the start of Phase II in which Project 8 shifted its focus to the demonstration of the first tritium beta-decay spectrum using CRES. Phase II successfully concluded in 2023 with the submission of the papers demonstrating the first tritium beta-decay spectrum endpoint and neutrino mass measurement using CRES. For more information on Phase II please see Section 3.3.

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

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

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

The need for new CRES detection techniques is driven by the drastic increase in scale

843 from Phase II to the Phase IV and the pilot-scale experiments. The physical volume
844 used for CRES in Phase II was on the order of a few cubic-centimeters, and achieving
845 Project 8’s sensitivity target of 40 meV requires an experiment volume on the multi-cubic
846 meter scale. Therefore, the waveguide gas cell CRES detection technique used in Phase
847 II is not a feasible option for the future of Project 8 due to it’s inability to scale to the
848 required size.

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

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

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

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

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

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

In Phase II Project 8 demonstrated the first ever measurement of the tritium beta-decay spectrum endpoint using the CRES technique, which lead to the first neutrino mass measurement by the Project 8 collaboration. This milestone was made possible by many improvements in the CRES technique and in the understanding of CRES systematics, which takes an important first step towards larger scale measurements of the tritium beta-decay spectrum with CRES. In this section, I briefly describe some important elements of the Phase II experiment, with the goal of contextualizing the research and development efforts for Phases III and IV of Project 8. For more complete descriptions of the work that lead to Project 8’s Phase II results please refer to the relevant publications by the collaboration [17, 18].

3.3.1 The Phase II CRES Apparatus

Magnet and Cryogenics

The magnetic field for the the Phase II experiment is provided by a nuclear magnetic resonance (NMR) spectroscopy magnet with a central bore diameter of 52 mm (see Figure 3.7). The magnet produces a background magnetic field with an average value of 0.959 T and a 10 ppm variation across the bore diameter achieved using several shim coils built into the magnet. Using an external NMR field probe the variation of the magnetic field along the vertical axis of the magnet bore was measured to obtain an accurate model of the magnetic field so that the CRES cell could be positioned for optimal magnetic field uniformity.

An external solenoid magnet was installed inside the magnet bore to provide the ability to shift the magnitude of the background magnetic field by values on the order of a few mT. The solenoid has inside diameter of 46 mm and a length of 350 mm, which terminates in a vacuum flange that allows it to be inserted into the NMR magnet bore from the bottom. By shifting the value of the magnetic field by a few mT, the cyclotron frequencies of electrons produced by the 17.8 keV ^{83m}Kr internal-conversion line [32] can be shifted over a range of frequencies on the order of 100 MHz. This allows one to study the frequency dependent behavior of multiple CRES systematics such as detection efficiency that directly affect the measured shape of the tritium spectrum.

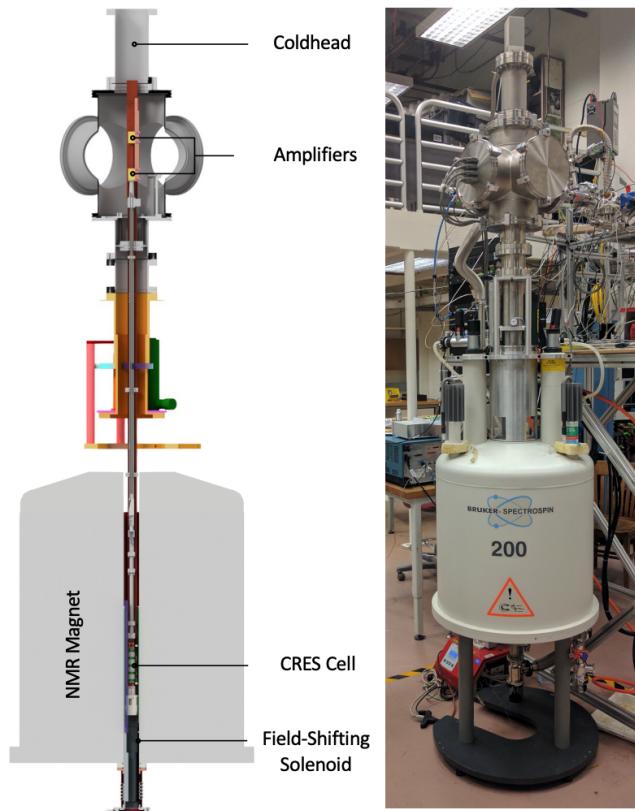


Figure 3.7: The Phase II CRES apparatus used to perform the first measurement of the tritium beta-decay spectrum using CRES.

906 The inside of the magnet bore diameter was pumped down to a vacuum of less than
 907 10 μtorr using a turbomolecular pump, which allows for cryogenic cooling of the CRES
 908 cell and RF system. Cooling power was supplied to the Phase II apparatus using a
 909 cryopump with its coldhead mounted above the primary magnet and CRES cell. This
 910 arrangement allowed for sufficient cooling power to be delivered to the amplifiers to cool
 911 them to a temperature of ≈ 40 K, while keeping the amplifiers far enough from the
 912 magnet so as not to be damaged by the large field strength. Thermal contact between
 913 the coldhead, amplifiers, RF system, and CRES cell is achieved using a copper bar that
 914 runs the full length of the apparatus. To prevent freeze-out of ^{83m}Kr on the walls of the
 915 CRES cell a separate heater was installed to keep the CRES cell near a temperature of
 916 85 K during the operation of the experiment.

917 **CRES Cell**

918 Located in the most uniform region of the magnetic field is the CRES cell, which is the
919 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

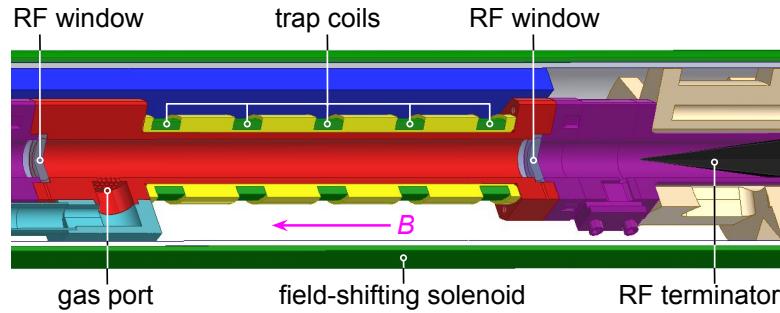


Figure 3.8: Diagram of the CRES cell portion of the Phase II apparatus.

920
921 from a segment of cylindrical waveguide designed to operate at K-band frequencies
922 near 26 GHz. The diameter of the waveguide determines which resonant modes of the
923 waveguide will couple to the electron and transmit its radiation to the amplifiers. For
924 Phase II a waveguide diameter of 1 cm was selected, which allows electrons to couple to
925 the TE₁₁ and TM₀₁ cylindrical waveguide modes. To reduce complexity in modeling and
926 analyzing the CRES data, it is ideal to select a diameter that prevents electrons from
927 coupling to higher-order waveguide modes beyond the fundamental TE and TM modes.

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

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

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

952 RF System

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

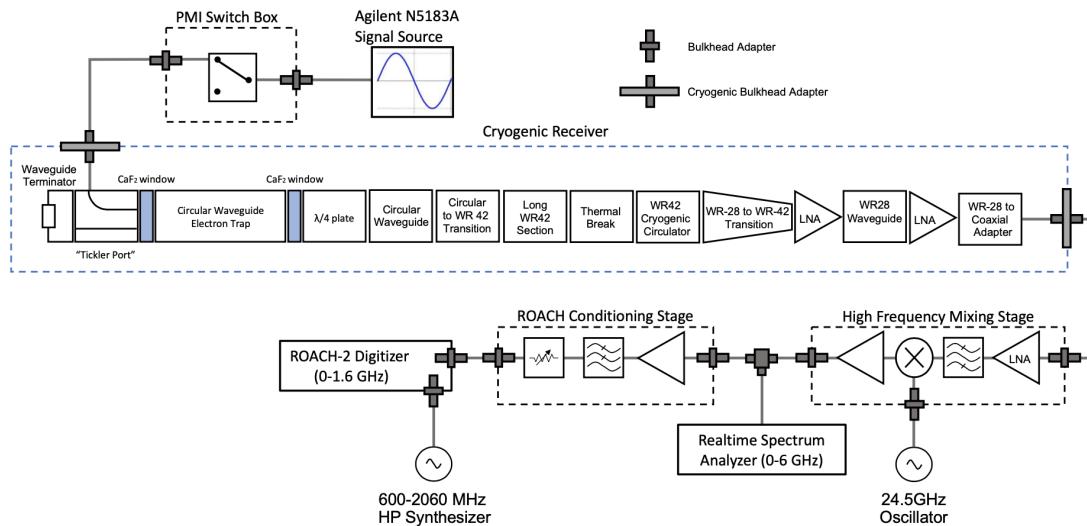


Figure 3.9: RF system diagram for the Phase II apparatus.

956 Below the CRES cell, at the bottom of the Phase II apparatus, is a tickler port and
 957 waveguide terminator. The tickler port is used to inject signals into the CRES cell and
 958 RF system for testing and calibration purposes. The waveguide terminator is designed to
 959 absorb cyclotron radiation emitted by electrons that transmits out of the bottom of the
 960 CRES cell. This lowers the total power received from electrons in the CRES cell, since all
 961 the energy radiated downwards is absorbed into the terminator. Earlier iterations of the
 962

963 Phase II apparatus used an RF short in this location that reflected this power up towards
964 the amplifiers, however, interference between the upward traveling and reflected radiation
965 led to a disappearance in the signal carrier that made reconstruction impossible.

966 Radiation traveling upward passes through the CaF_2 window passes through a $\lambda/4$ plate,
967 which transforms the circularly polarized cyclotron radiation into linear polarization.
968 The linearly polarized fields next travel through a segment of circular waveguide that
969 transitions into a long segment of WR-42 waveguide that carries the fields out of the
970 high magnetic field region. A thermal break segment is included, which consists of a
971 segment of gold-plated stainless steel WR-42 waveguide, to help thermally isolate the
972 relatively warm CRES cell from the colder amplifiers. The radiation then passes through
973 a cryogenic circular, which prevents signals reflected from the amplifiers from interfering
974 with the CRES cell before a WR-42 to WR-28 transition connects the waveguide to the
975 first of the cryogenic amplifiers. The radiation passes through two cryogenic amplifiers
976 before being coupled to a coaxial termination at the top of the Phase II apparatus.

977 The coaxial cable transfers the cyclotron radiation signals to a high-frequency mixing
978 stage that performs an analog frequency down-conversion using a 24.5 GHz LO. Two forms
979 of digitization can be used at this stage to readout the CRES data. One is a real-time
980 spectrum analyzer that digitizes the CRES signal data in time-domain and computes the
981 frequency spectrum in real-time, which allows for direct visualization of CRES signal
982 spectrograms as the experiment is running. The real-time spectrum analyzer is most
983 useful for taking small amount of streamed data for debugging and analysis of the system.
984 The other method, which was used to collect the majority of the CRES data in Phase II,
985 is a ROACH-2 FPGA and digitizer system. The ROACH system consists of a fast ADC
986 that samples the CRES signal data at 3.2 GSps. Internal digital down-conversion stages
987 implemented in the FPGA perform a mixing operation that reduces the bandwidth of the
988 CRES signals to 100 MHz. The FPGA implements a 8192 sample FFT and packetizes
989 time and frequency domain records in parallel. The packetized data is then transferred
990 from the ROACH to be analyzed by the data-processing pipeline.

991 **3.3.2 CRES Track and Event Reconstruction**

992 **Time-Frequency Spectrogram**

993 The online data-processing is intended to identify interesting data that could contain
994 CRES signals using a software real-time triggering algorithm. Interesting segments of
995 data identified by this algorithm are collected into files that are transferred to a server for

996 offline processing and analysis. The data files contain a continuous series of time-domain
997 samples, broken into a set of records, which are 4096 samples long. The time-series is
998 made up of 8-bit IQ samples acquired at 100 MHz.

999 Each time-series record is accompanied by an associated frequency spectrum consisting
1000 of 4096 frequency bins approximately 24.4 kHz wide, which is represented as a power
1001 spectral density. The individual frequency spectra can be organized temporally to create
1002 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

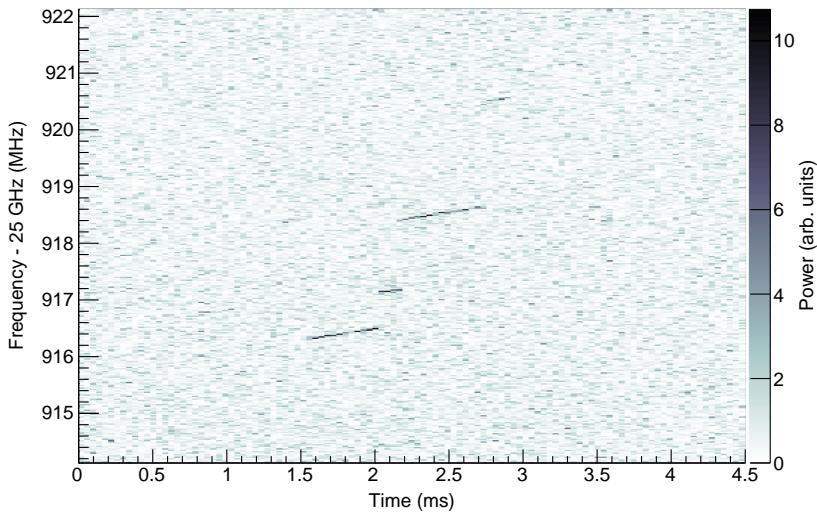


Figure 3.10: The time-frequency spectrogram of a tritium CRES event in the Phase II apparatus.

1003
1004 spectrogram is represented as a two-dimensional image where the color of each pixel is
1005 proportional to the power spectral density. Each vertical slice of pixels in the image
1006 represents a frequency spectrum, therefore, each horizontal bin represents the data
1007 obtained over a duration of $4096 \times 0.01 \text{ MHz}^{-1} = 40.96 \mu\text{sec}$.

1008 CRES Event Data Features

1009 Phenomenologically, a CRES signal appears as a sinusoidal signal whose frequency slow
1010 increases ("chirps") over time. Axial motion of the electron in the trap leads to the
1011 formation of frequency sidebands that surround the more powerful carrier frequency, due
1012 to Doppler modulation of the electron's frequency as it bounces between the walls of the
1013 magnetic trap. The critical piece of information that must be extracted from the track
1014 and event reconstruction procedure is the carrier frequency, since it is this frequency

1015 that gives the cyclotron frequency and thus the kinetic energy. While axial motion from
1016 non- 90° pitch angles does change the average magnetic field experienced by an electron
1017 and, therefore, changes the cyclotron frequency. Because of low-SNR sidebands were
1018 unable to be observed in Phase II, so a correction for the effect of the pitch angle on the
1019 cyclotron frequency was not possible.

1020 In the time-frequency spectrogram representation the chirping carrier frequency
1021 appears as a linear track of high-power frequency bins (see Figure 3.10). The vertical
1022 slope of the tracks is caused by the emission of energy from the electron in the form of
1023 cyclotron radiation, therefore, the size of the slope parameter is directly proportional
1024 to the Larmour power. The continuous track is periodically interrupted by random
1025 jumps to higher frequency and lower energy caused by random inelastic collisions with
1026 background gas molecules. The length of a track is an exponentially distributed variable
1027 whose mean value is inversely proportional to the gas density. The size of the frequency
1028 discontinuities is directly proportional to the energies of the rotational and vibrational
1029 states of background gas molecules.

1030 A CRES event refers to the collection of tracks produced by a trapped electron until
1031 it inevitably scatters into a pitch angle that can no longer be trapped. The goal of track
1032 and event reconstruction is to first identify the set of tracks present in a time-frequency
1033 spectrogram that represents a segment of data acquired in the Phase II apparatus. These
1034 tracks must then be clustered into events from which we can determine the first track
1035 produced by the electron and thus estimate it's starting cyclotron frequency and kinetic
1036 energy.

1037 **Track Reconstruction**

1038 The first step in CRES event reconstruction is the identification of tracks in the time-
1039 frequency spectrogram, which is essentially an image processing task. Track finding
1040 starts by normalizing the power spectral density based on the average noise power.
1041 Next a power threshold is applied to the normalized spectrogram where only bins that
1042 have a signal-to-noise ratio greater than five are selected to build tracks. In this case
1043 signal-to-noise ratio is defined as the ratio between the normalized, unitless power of a
1044 bin divided by the average normalized power across the full frequency spectrum.

1045 The sparse spectrogram produced by this power cut consists only of a sparse collection
1046 of high-power frequency bins that could be part of a CRES signal track (see Figure
1047 3.11). In this form is it much easier to identify tracks "by eye", however, for the Phase II
1048 analysis Project 8 developed its own custom-made track finding algorithm, called the

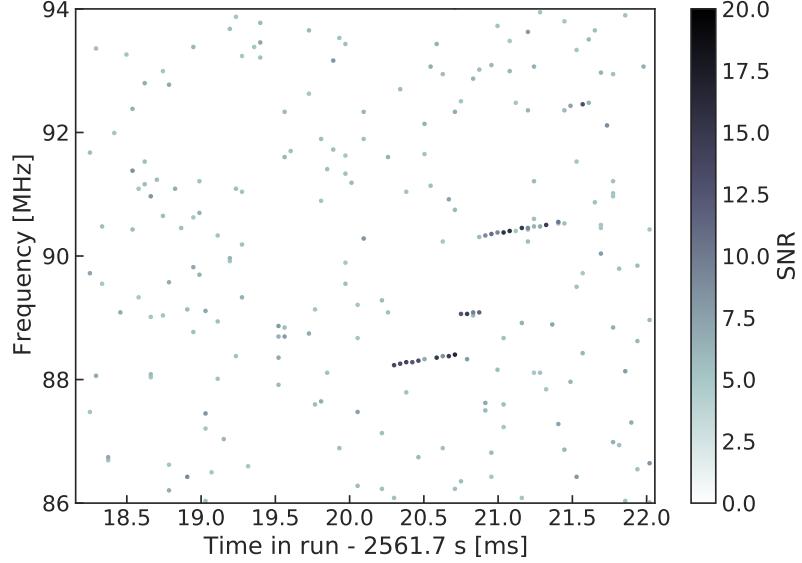


Figure 3.11: The sparse spectrogram obtained by placing a power cut on the raw spectrogram shown in Figure 3.10.

1049 sequential track finder (STF).

1050 The STF algorithm processes the sparse spectrogram in sequential fashion, processing
1051 each time-slice one-by-one until the end of the spectrogram is reached. Tracks are found
1052 by searching for points in the sparse spectrogram that appear to fall on a straight line.
1053 Multiple configurable parameters are built into the STF algorithm that allow the user to
1054 tune the criteria for adding a point to an existing track or creating a new track. These
1055 include parameters such as maximum time and frequency differences between subsequent
1056 points in a track as well as minimum SNR values for the start and endpoints of the track.
1057 Additionally, tracks are required to have a minimum length and slope to be considered
1058 potential CRES tracks rather than random noise fluctuations.

1059 The resulting output of the STF is a collection of track objects that consist of the track
1060 point objects and their properties. The final step is to calculate track-level properties
1061 and apply cuts to reject false tracks found by the STF. This involves the fitting of a
1062 line to the collection of track points as well as the total and average power of the track
1063 obtained by computing the sum and mean of the points powers. The starting frequency
1064 of the track is determined by calculating the time coordinate that intersects with the
1065 linear fit. A cut is performed to remove all tracks that do not have a specified average
1066 power over their duration, which helps to remove the majority of noise fluctuations that
1067 have passed all previous cuts up to this point.

1068 **Event Reconstruction**

1069 After track reconstruction comes event reconstruction where the identified tracks are
1070 grouped into events that correspond to the trajectory of a single electron in the trap. This
1071 procedure attempts to match tracks head to tail by checking if the start and end times
1072 of a pair of tracks falls within a certain tolerance. This tolerance is a configurable
1073 parameter that can be tuned to an optimal value using Monte Carlo simulations of events
1074 in the Phase II apparatus.

1075 After the event building procedure has completed there remains a small likelihood
1076 that false tracks have made it through to the event reconstruction stage. Typically, cuts
1077 at the track level are able to remove 95% of the false tracks identified by the STF, which
1078 leads to a significant number of false tracks at the event building stage. However, the
1079 additional event-level information makes it possible to reject events that contain these
1080 false tracks with a high degree of confidence.

1081 Two event level features are associated with events caused by real electrons — the
1082 duration of the first track as well as the number of tracks in the event. Real electrons
1083 tend to have event structures with longer first tracks and a higher number of total tracks.
1084 Based on the values of these two criteria, a minimum threshold on the average power in
1085 the first track was configured to reject false events. The average power in the first track
1086 was chosen due to the critical nature of the starting frequency of the first track in an
1087 event to the krypton and tritium spectrum analyses.

1088 **3.3.3 Results from Phase II**

1089 The main result from Phase II was the measurement of the tritium beta-decay spectrum
1090 using CRES, which lead to the first neutrino mass limit with CRES. However, Phase
1091 II also included a significant ^{83m}Kr measurement campaign to understand important
1092 systematics relevant to the tritium spectrum measurement, but also to understanding the
1093 fundamentals of the CRES technique itself. This required high-resolution measurements
1094 of the ^{83m}Kr internal-conversion spectrum [32], which is an interesting science result in
1095 its own right.

1096 The results from Phase II represents a significant effort from the entire Project 8
1097 collaboration over several years. Because the focus of my contributions to Project 8 is
1098 directed towards the research and development efforts for the Phase III experiments, the
1099 goal in this section is not to provide a detailed description of the analyses that lead to
1100 the Phase II results. Rather, I will provide brief descriptions of a few plots representative

1101 of the main results from Phase II.

1102 Measurements with Krypton

1103 Measurements with krypton were a key calibration tool for Phase II of the experiment and
1104 will continue to be useful in Phase III. In the context of Project 8 krypton measurements
1105 refers to CRES measurements of the internal-conversion spectrum of the metastable state
1106 of krypton-83, ^{83m}Kr , produced by electron capture decays of ^{83}Rb . A supply of ^{83}Rb
1107 was built into the Phase II apparatus gas system that supplied the CRES cell with ^{83m}Kr
1108 via emanation.

1109 The ^{83m}Kr internal-conversion spectrum consists of several lines based on the orbital
1110 of the electron ejected during the decay. The conversion lines useful to Project 8 are
1111 those that emit electrons with kinetic energies that fall inside the detectable frequency
1112 bandwidth of the Phase II apparatus. These are the K; L2 and L3; M2 and M3; and N2
1113 and N3 lines with kinetic energies of 17.8 keV, \approx 30.4 keV, \approx 31.9 keV, and \approx 32.1 keV,
1114 respectively. The different energies of the lines allow a onw to test the linearity of the
1115 relationship between kinetic energy and frequency across the range of frequencies covered
1116 by the continuous tritium spectrum.

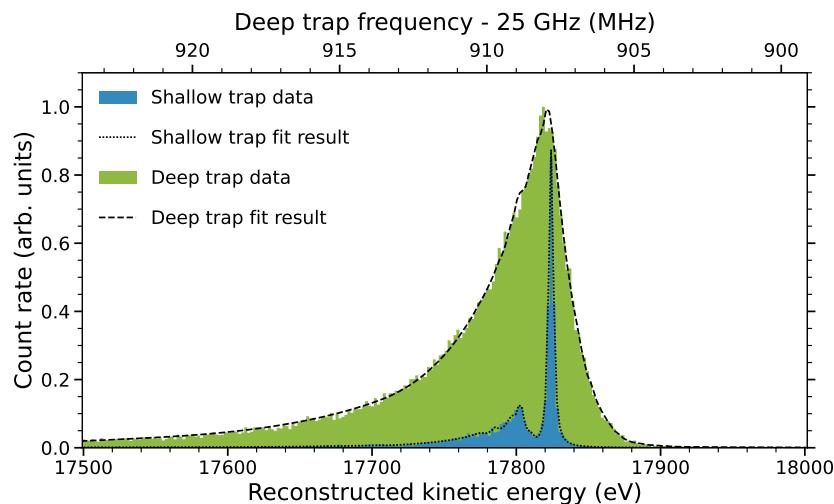


Figure 3.12: Fits to the measured 17.8-keV ^{83m}Kr conversion line using the deep and shallow trap configurations.

1117 Numerous detector related effects relevant to the tritium analysis can be characterized
1118 by measuring the shape of the krypton spectrum. Specific examples include variations
1119 in the magnetic field as a function of the radial position of the electron, variation in
1120 the magnetic field caused by the trap shape, variation in the average magnetic field for

1121 electrons with different pitch angles, and the effect of missing tracks due to scattering.
1122 These spectrum shape measurements focused on the 17.8-keV krypton line and utilized
1123 different trap geometries based on the particular goal of the dataset (see Figure 3.12).

1124 Krypton measurements with a shallow trap allow for high energy resolution, since
1125 variation in frequency due to pitch angle differences is sharply reduced in the shallow
1126 trap configuration. With this trap the main 17.8-keV peak of the conversion spectrum is
1127 clearly visible along with additional satellite peaks at lower energy, which correspond to
1128 the shakeup/shakeoff spectrum of the decay. The high accuracy of the fit demonstrates a
1129 high degree of understanding of the CRES systematics.

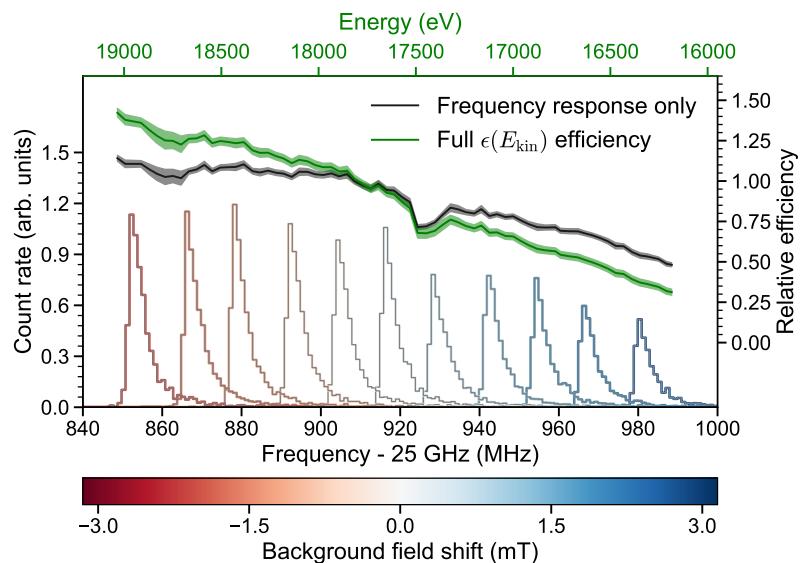


Figure 3.13: Measurements of the 17.8-keV ^{83m}Kr line using the deep trap configuration for different values of the magnetic field from the field shifting solenoid.

1130 The broadening of the krypton spectrum seen for the deeper track is due to the large
1131 range of electron pitch angles that can be trapped. Furthermore, with a deeper trap
1132 there is a larger parameter space of electron that could be produced with pitch angles
1133 that are trappable but not visible in the time-frequency spectrogram. These electrons
1134 live in the trap and can scatter multiple times before randomly scattering to a visible
1135 pitch angle. This leads to one or more missing tracks earlier in the event, which leads to
1136 a misreconstruction of the true starting frequency. By measuring the krypton spectrum
1137 shape in the same trap used to detect tritium events, the effect this has on the spectrum
1138 shape can be characterized to mitigate its impact on the tritium measurements.

1139 Changes in the Krypton spectrum shape as a function of CRES frequency were

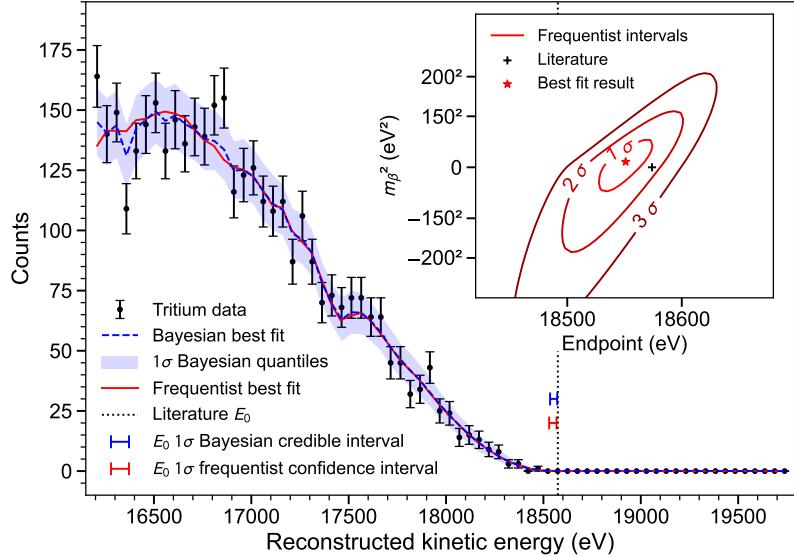


Figure 3.14: The measured tritium spectrum from Phase II with Bayesian and frequentist fits.

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

1148 Tritium Spectrum and Neutrino Mass Results

1149 The tritium measurement campaign resulted in the collection of 82 days of detector
 1150 live time during which 3770 total tritium events were detected. The track and event
 1151 reconstruction analysis extracted the starting frequencies of these tritium events, which
 1152 were used to build a frequency spectrum of tritium beta-decays. The resulting frequency
 1153 spectrum was then converted to an energy spectrum using the information gleaned from
 1154 the krypton measurement campaign to obtain the tritium beta-decay spectrum (see
 1155 Figure 3.14).

1156 CRES is inherently a very low background technique with the dominant source of
 1157 noise being random RF fluctuations. Monte Carlo simulations backed validated using

1158 measurements of the RF noise background were used to set track and event cuts to
1159 guarantee that zero false events would occur over the duration of the experiment with
1160 90% confidence. Notably, the measured spectrum has zero events beyond the tritium
1161 spectrum endpoint, which allows us to constrain the background rate in the Phase II
1162 apparatus to less than 3×10^{-10} counts/ev/s. Achieving a low background is critical for
1163 future neutrino mass experiments that seek to measure the neutrino mass with less than
1164 100 meV sensitivity.

1165 Bayesian and frequentist based fits to the measured tritium spectrum, incorporating
1166 information gained about CRES systematics from the krypton measurements, were
1167 performed to extract upper limits on the tritium beta-decay spectrum endpoint as well as
1168 the neutrino mass. The estimated spectrum endpoints are 18553^{+18}_{-19} eV for the Bayesian
1169 analysis and 18548^{+19}_{-19} eV for the frequentist analysis. The quoted uncertainties are
1170 $1-\sigma$, and both results are within $2-\sigma$ of the literature endpoint value of 15574 eV. The
1171 estimated neutrino mass for both results is consistent with $m_\beta^2 = 0$. The 90% confidence
1172 upper limits for the Bayesian analysis is $m_\beta < 155$ eV/c² and $m_\beta < 152$ eV/c for the
1173 frequentist analysis.

1174 Though the neutrino mass results from Phase II are not competitive with KATRIN
1175 the experiment was a promising first step towards the development of more precise
1176 neutrino mass measurements using CRES. The low-background and high-resolution
1177 achievable with krypton measurements are promising features of the technique that were
1178 demonstrated with the Phase II apparatus. As new technologies are developed to enable
1179 CRES measurements in larger volume, many of the lessons learned from Phase II will
1180 continue to influence the operation and design of future experiments.

1181 **3.4 Phase III R&D: Antenna Array CRES**

1182 The goal of Phase III in the Project 8 experimental program is to develop the technologies
1183 and expertise required to build an experiment that uses CRES to measure the neutrino
1184 mass with a target sensitivity of 40 meV. One of the key technologies is a method for
1185 performing high resolution CRES measurements in a large volume, which allows one to
1186 observe a sufficient quantity of tritium to measure the low-activity endpoint region of
1187 the tritium spectrum.

1188 **3.4.1 The Basic Approach**

1189 One possible approach, suggested in the original CRES publication [15], is to use many
1190 antennas to surround a volume of tritium gas in a magnetic field (see Figure 3.15). When
1191 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

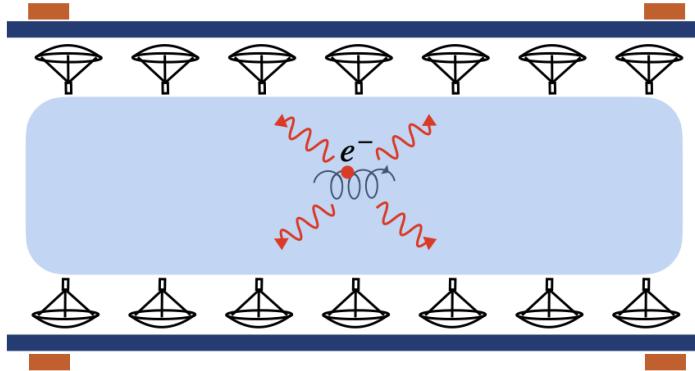


Figure 3.15: A cartoon illustration of the basics of the antenna array CRES technique.

1192
1193 fraction of the electron's signal power, which is less than 1 fW for a 18.6 keV kinetic
1194 energy electron in a 1 T magnetic field. Scaling to large volumes with the antenna
1195 array approach is accomplished by increasing the number of antennas in the array, which
1196 increases the volume under observation proportionally, so that a sufficient population of
1197 tritium atoms can be observed to measure the tritium spectrum endpoint shape.

1198 Several features of the antenna array approach make it an attractive candidate technol-
1199 ogy for a large volume experiment. One example is the accurate position reconstruction
1200 made possible by the multichannel nature of the array. Using techniques like digital
1201 beamforming it is possible to estimate the radial and azimuthal positions of the electron
1202 in the magnetic trap with a precision significantly less than the size of the cyclotron
1203 wavelength. This capability allows one to perform event-by-event estimations of the
1204 magnetic field experienced by an electron, which is crucial to achieving high energy
1205 resolution with the CRES technique.

1206 The easy availability of position information with the antennas array approach
1207 is potentially a unique advantage that provides significant flexibility in the magnetic
1208 field uniformity requirements compared to other proposed approaches to large volume
1209 CRES (see Chapter 6). Spatial discrimination using digital beamforming leads to pileup
1210 reduction, which helps to reduce the potential of background events caused by missing
1211 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 complexity of the antenna array CRES technique requires the construction of a small scale demonstration experiment to develop an understanding of technique itself and relevant systematics. Without a demonstrator experiment it is not possible to sufficiently retire the technical risks associated with the full-scale experiment. Therefore, Phase III of the Project 8 experimental program is primarily focused on the development and operation of demonstrator experiments to inform the design of the Phase IV experiment.

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

Magnetic Field

The background magnetic field for the FSCD is provided by a hospital-grade MRI magnet (see Figure 3.16). The magnet produces a magnetic field of approximately 0.958 T, which

corresponds to a tritium spectrum endpoint frequency of approximately 25.86 GHz. The



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

1244
1245 magnet is installed in the Project 8 laboratory located at the University of Washington,
1246 Seattle, and is shimmed to produce a uniform magnetic field with variations on the
1247 ppm-level. Measurements of the magnetic field non-uniformities are performed using a
1248 NMR probe and rotational gantry to capture measurements of the magnetic field around
1249 an elliptical surface in the center of the MRI magnet. During the operation of the FSCD
1250 an array of Hall or NMR magnetometers would be used to periodically measure the
1251 magnetic field to monitor its time stability.

1252 Inside the field of the MRI magnet additional electromagnets would be installed that
1253 provide the capability to shift the value of the background magnetic field and produce
1254 a magnetic trap. Shifting the background magnetic field by a few μ T lets one control
1255 the cyclotron frequencies of electrons with a fixed kinetic energy, which is key to an
1256 effective calibration of the FSCD. The preferred calibration method for the FSCD is
1257 a mono-energetic electron gun that can inject electrons into the magnetic trap with a
1258 known kinetic energy. In combination with the field shifting magnet, one can vary the
1259 cyclotron frequencies of the electrons to measure the response of the antenna array as a
1260 function of the radiation frequency and electron position. This procedure characterizes
1261 the response of the antenna array and provides further information on magnetic field
1262 uniformity, which important to achieving good energy resolution.

1263 The design of the magnetic trap is absolutely critical to the success of a CRES
1264 experiment. The ideal shape is the perfect magnetic box, which has a flat bottom and

1265 step function walls. Any variation in the average magnetic field experienced by an
1266 electron leads to changes in the cyclotron frequency that can make determining the true
1267 starting kinetic energy more difficult. This includes changes in the magnetic field caused
1268 by the walls of the magnetic trap as well as radial magnetic field variations.

1269 The ideal box trap is completely uniform and has infinitely steep walls that cause
1270 no change in the electron's cyclotron frequency as it is reflected from the trap wall,
1271 however, such a trap cannot be made from any combination of magnetic coils since it
1272 violates Maxwell's equations. One of the goals of magnetic trap design is to identify the
1273 configuration of coils that produces a trap that approximates the perfect box trap as
1274 closely as possible.

1275 **Antenna Array**

1276 The canonical antenna array design for CRES is a uniform cylindrical array of antennas
1277 that surrounds the magnetic trap volume. Since the FSCD is a demonstrator experiment,
1278 the antenna array design is the simplest form of the uniform cylindrical array, which is a
single circular ring of antennas with a diameter of 20 cm (see Figure 3.17). Along this

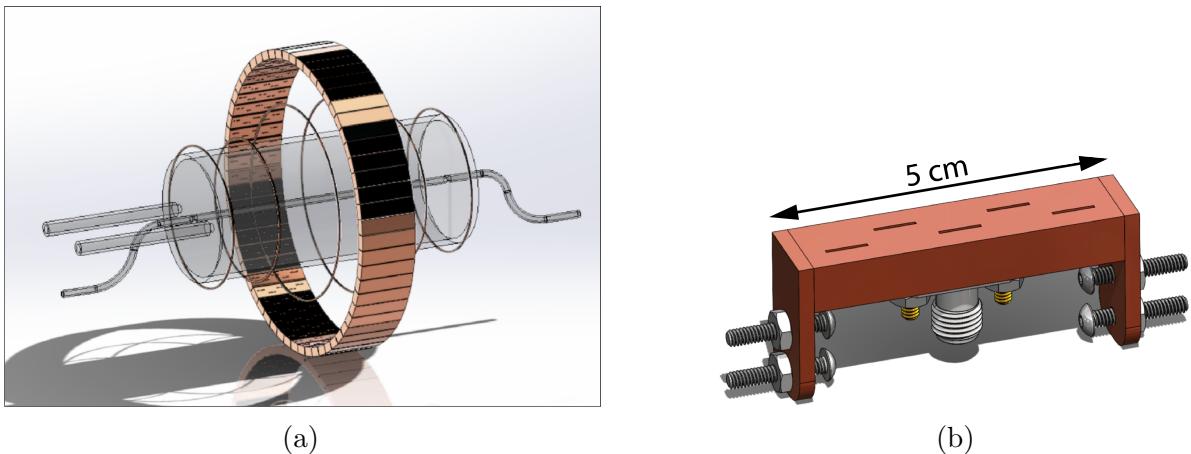


Figure 3.17: (a) A model of the FSCD antenna array, magnetic trap, and tritium containment vessel design.(b) A more detailed model of a prototype design for the 5-slot waveguide antenna design.

1279
1280 circle are sixty slotted waveguide antennas that fully populate the available space around
1281 the array circumference. In order to maximize the power collected from each electron
1282 it is optimal to cover as large a fraction of the solid angle around the magnetic trap as
1283 possible.

1284 The distance between antennas around the circumference of the array is proportional

1285 to the wavelength of the cyclotron radiation. Therefore, maximizing the solid angle
1286 coverage of the array, while minimizing channel count to keep the hardware and data
1287 acquisition costs manageable, biases one towards smaller array diameters. Antenna
1288 near-field effects limit the minimum diameter of the array for a given antenna design
1289 since the radiation from electrons that are too close to the array cannot be detected
1290 due to destructive interference caused by path-length differences from the electron to
1291 different points on the antenna surface.

1292 Slotted waveguide antennas are used in the FSCD antenna array due to their high
1293 efficiency and low loss, which comes from the lack of dielectric materials in the antenna
1294 structure. Coupling to the waveguide can be performed with a coaxial cable connected
1295 at the center or on either end of the waveguide. One of the drawbacks of waveguide
1296 antennas is the large amount of space required to fit them inside the limited MRI magnet
1297 volume. Alternative antenna designs, constructed from microstrip printed circuit boards
1298 require significantly less space at the cost of slightly higher energy loss in the antenna
1299 structure.

1300 The FSCD antenna design is a 5 cm long segment of WR-34 waveguide with 5 vertical
1301 slots cut into the side. The distance between slots along the length of the waveguide is
1302 a half wavelength for optimal power combination between the individual antenna slots.
1303 Each slot is offset from the center of the antenna face a small distance in order to most
1304 effectively couple the slot to waveguide modes inside the antenna.

1305 The passive power combination achieved by placing 5 slots in a single waveguide is a
1306 compromise intended to reduce the cost and complexity of the antenna array system.
1307 Each additional channel in the array requires it's own cryogenic amplifier and also increase
1308 the required computer power to process the raw data collected by digitizing each channel.
1309 Passive summation, achieved by combining antennas into arrays axially, reduces the array
1310 channel count at the cost of losses from imperfect passive combination. Imperfect passive
1311 combination is caused by effects such as re-radiation of energy from and destructive
1312 interference between slots in the waveguide antenna.

1313 Interference and re-radiation eventually limit the achievable the axial extent of passive
1314 power combination. The 5-slot designed developed for the FSCD is optimized to minimize
1315 the impact of these losses while achieving the maximum amount of axial coverage with a
1316 single ring of antennas. Scaling beyond the volume covered by a single ring of antennas is
1317 achieved by stacking additional rings of antennas together to cover a larger trap volume
1318 for a higher statistics measurement of the tritium spectrum endpoint region. A likely
1319 scenario for the FSCD experiment involves a staged experiment approach, where first

1320 a series of measurements is performed using only a single ring of antennas followed by
1321 experiments that add additional rings to the FSCD. The goal would be to first understand
1322 the principles of antenna array CRES using the simplest possible experiment, before
1323 attempting to scale the technique by expanding the antenna array size.

1324 **Tritium Source**

1325 While the primary purpose of the FSCD is as a technology demonstrator, it is unlikely
1326 for the collaboration to gain the required confidence in the antenna array CRES tech-
1327 nique to perform neutrino mass measurements at the 40 meV sensitivity level without
1328 an intermediate scale measurement of the neutrino mass using antenna array CRES.
1329 Therefore, the FSCD has an additional scientific goal of measuring the neutrino mass
1330 with a rough sensitivity goal of a few eV. This level of precision is achievable using a
1331 source of molecular tritium with a volume of approximately 1 L at a density comparable
1332 to potential Phase IV scenarios.

1333 Unlike previous CRES experiments, where the tritium source could be co-located
1334 with the receiving antenna inside a waveguide transmission line, the tritium source
1335 in the FSCD is thermally isolated from the antenna array to avoid freeze-out of the
1336 tritium molecules. The tiny radiation power emitted by electrons requires a system noise
1337 temperature of ≈ 10 K or less, in order to detect events at a high enough efficiency to
1338 reach the neutrino mass sensitivity goals of the experiment. Achieving a system noise of
1339 10 K requires that the antenna array and amplifiers operate at cryogenic, liquid helium
1340 temperatures of ≈ 4 K, which significantly lowers the vapor pressure of molecular tritium.
1341 By keeping the molecular tritium isolated in an RF-transparent vessel the tritium gas can
1342 be kept at a relatively warmer temperature in the range of 30 K to avoid the accumulation
1343 of tritium on the experiment surfaces.

1344 **Data Acquisition and Reconstruction**

1345 A fundamental change in the data acquisition system for the FSCD is the shift from
1346 single to multi-channel reconstruction. This transition results in a significant increase in
1347 the data-generation rate, which is linearly related to the number of independent channels
1348 in the array. The larger data volume coincides with an increased demand for computer
1349 processing power based on the need for more precise signal reconstruction algorithms
1350 driven by the FSCD and Phase IV sensitivity goals. Therefore, the data acquisition
1351 system for the FSCD is likely to represent a significantly larger fraction of the experiment
1352 cost and complexity than previous CRES experiments.

1353 Each antenna in the array is connected to a cryogenic amplifier and down-converted
1354 from the 26 GHz CRES frequency using an IQ-mixer to reduce the size of the analysis
1355 window in which the tritium spectrum is measured. Using an LO with a frequency of
1356 approximately 25.80 GHz the antenna array signals can be digitized at a rate of 200 MHz,
1357 which is sufficient bandwidth to resolve the complete sideband spectrum produced by
1358 axial oscillations of electrons in the FSCD magnetic trap.

1359 Direct storage of the raw FSCD antenna array data is undesirable, since the estimated
1360 amount of raw data generated is $O(1)$ exabyte per year. The management and storage
1361 of such a large dataset is infeasible for a demonstrator experiment on the scale of the
1362 FSCD and would represent a large fraction of the budget for a Phase IV scale antenna
1363 array based CRES experiment. Therefore, a sub-goal of the FSCD experiment is the
1364 development of real-time reconstruction methods that could reduce the raw data volume
1365 by detecting and reconstructing CRES events in real-time. The ultimate goal would be
1366 a complete real-time reconstruction pipeline that takes raw voltages samples from the
1367 antenna array and returns estimates for the starting kinetic energies of CRES events in
1368 the data.

1369 The feasibility of a real-time reconstruction pipeline rests on the development of
1370 computationally efficient algorithms that can be implemented without the need for
1371 enormous computing resources. One challenge with the antenna array approach is that
1372 the small radiation power of a single electron is distributed between each channel in
1373 the array, such that reconstruction using only the information in a single channel is not
1374 possible. Therefore, the simply performing the initial step in reconstruction — signal
1375 detection — requires orders of magnitude more computational power than previous CRES
1376 experiments. This operation will then be followed by other, potentially more expensive,
1377 reconstruction steps that are required in order to determine the kinetic energy of the
1378 electron.

1379 **3.5 Pilot-scale Experiments**

1380 **3.5.1 Choice of Frequency**

1381 The optimal CRES frequency for Project 8 is that which can reach our target sensitivity
1382 of 40 meV, while minimizing the cost and complexity of the overall experiment. The
1383 magnitude of the background magnetic field determines the cyclotron frequency, which
1384 affects the entirety of the CRES detection system design, specifying the operating

frequency of the CRES experiments is one of the first steps towards developing a full design.

Scaling Laws

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

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

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

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

¹⁴¹⁹ field correction system will need to be devised, which includes constant monitoring of
¹⁴²⁰ external fields.

¹⁴²¹ Atomic Tritium Considerations

¹⁴²² The pilot-scale experiments will be the first Project 8 experiments to combine CRES with
¹⁴²³ atomic tritium, therefore, the optimal frequency should take into account the affect of the background magnetic field on the atom trap. The primary influence of the background

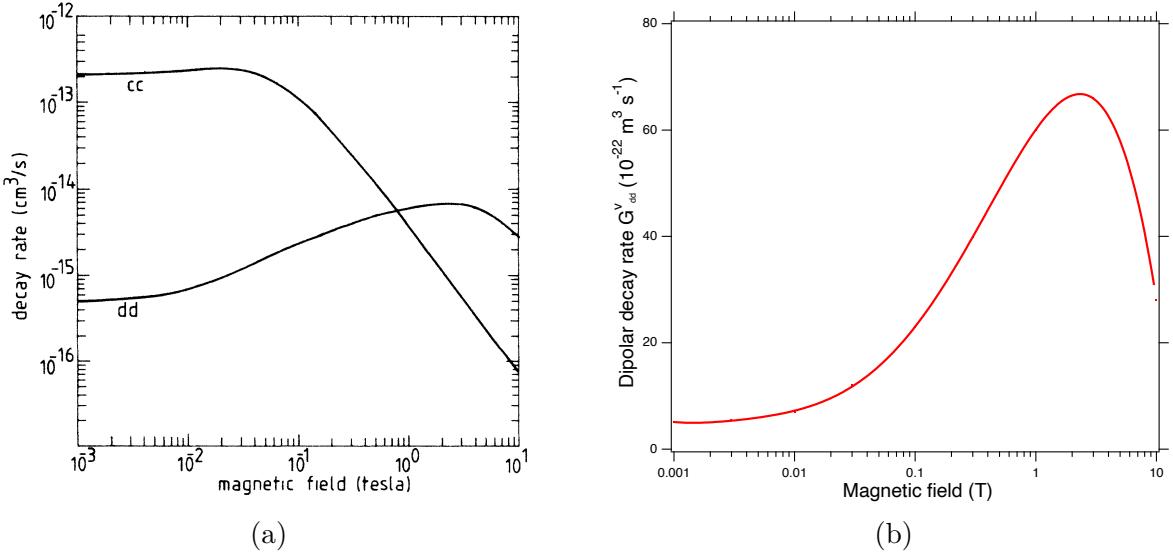


Figure 3.18: (a) A plot of the decay rate for the two-body dipolar spin exchange interaction for cc and dd state. (b) A plot of the decay rate of the dipolar spin exchange interaction for d+d states as a function of magnetic field magnitude. Lowering the magnetic field is key for reducing the losses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

¹⁴²⁴

¹⁴²⁵ field magnitude is through the rate of dipolar spin-flips caused by a spin exchange
¹⁴²⁶ interaction between trapped atoms [33].

¹⁴²⁷ Atomic tritium is a simple quantum system with a hyperfine structure given by the
¹⁴²⁸ addition of the nuclear and atomic spins. The addition of two spins leads to a hyperfine
¹⁴²⁹ structure with four states in the (m_s, m_I) basis [34]. The states with atomic spins directed
¹⁴³⁰ anti-parallel to the magnetic field have $m_s = -1/2$ and are labeled as the a and b states.
¹⁴³¹ The a and b states are colloquially known as high-field seeking states, since their energy is
¹⁴³² minimized when in regions of higher magnetic field. This leads to losses in the magnetic
¹⁴³³ trap as these atoms are drawn to higher fields away from the trap center. Alternatively,
¹⁴³⁴ the c and d states, with atomic spin $m_s = +1/2$, minimize their energy in low magnetic
¹⁴³⁵ fields because of the parallel alignment between spin and the magnetic field. Therefore,

1436 these low-field seeking states tend to stay trapped significantly longer than the high-field
1437 seeking states.

1438 It would be advantageous to prepare tritium atoms in purely c and d states before
1439 trapping, however, even in this case losses still occur due to dipolar interactions between
1440 pairs of c and d states leading to flipped atomic spins and subsequent losses from high-field
1441 seeking atoms. The rate of these interactions depends on the magnitude of the background
1442 magnetic field and is maximal for dd interactions around 1 T (see Figure 3.18). The rate
1443 of losses from these interactions at 1 T requires atomic tritium production at a rate two
1444 orders of magnitude larger than at 0.1 T, thus, requirements on the whole atomic tritium
1445 system are significantly relaxed at lower magnetic fields, which provides an additional
1446 argument for transitioning to lower frequencies with the pilot-scale experiments.

1447 3.5.2 Pilot-scale Experiment Concepts

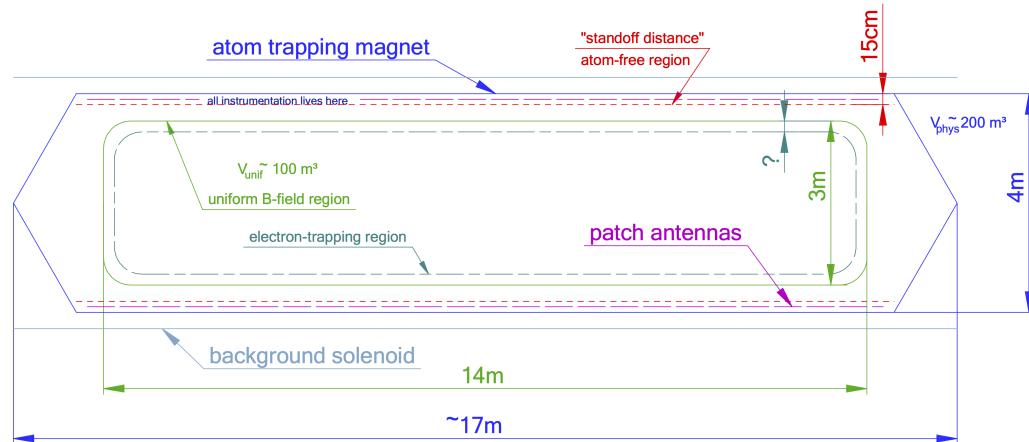


Figure 3.19: A conceptual sketch of a large-volume antenna array based CRES experiment to measure the neutrino mass.

1448 While the pilot-scale experiments are still in the early stages, enough is known to
1449 sketch the general features of these experiments at the conceptual level.

1450 Pilot-scale Antenna Array CRES Experiment Concept

1451 A conceptual design for an antenna-based CRES experiment is shown in Figure 3.19.
1452 A large solenoid magnet provides a uniform background magnetic field less than 0.1 T
1453 in magnitude. Inside this region is the atom trapping magnet that generates a high
1454 magnetic field at the walls, which decays exponentially towards the central region. Known

1455 magnet designs that produce suitable atom trapping fields include Ioffe-Prichard traps,
1456 which use conducting coils, as well as a Halbach array made from permanent magnets.
1457 Either magnet choice produces a region of high magnetic fields, which excludes atoms
1458 and allows for the placement of antennas inside the experiment.

1459 Inside this region an array of microstrip patch antennas is inserted to collect the
1460 cyclotron radiation without providing a surface for atomic tritium recombination. Due
1461 to the lower frequency of cyclotron radiation antennas of a larger size can be used,
1462 which lowers the total number of antennas required to observe the experiment volume.
1463 Because of this scaling, the lower frequency experiment uses a similar number of antennas
1464 compared to a much smaller demonstrator experiment with a 1 T magnetic field.

1465 The atomic tritium beamline that supplies fresh tritium atoms to the experiment is
1466 not shown in the figure. The general configuration would matches the one shown for the
1467 pilot-scale cavity experiment (see Figure 3.20).

1468 Pilot-scale Cavity CREs Experiment Concept

1469 The pilot-scale cavity experiment includes both an atomic tritium system and cavity
1470 CREs system. The atomic system consists of a thermal atom cracker located at the
1471 start of an evaporatively cooled atomic beamline. The atomic tritium system provides a
1472 supply of tritium atoms to the trap with temperatures on the order of a few mK. Atoms
1473 at this temperature can be trapped magneto-gravitationally, which is the reason for the
1474 vertical orientation of the cavity. At these low magnetic fields the trapping requirements
1475 for electrons and atoms differ enough such that it is advantageous to decouple the the
1476 trapping potentials to avoid radioactive heating of the tritium atoms from excess trapped
1477 electrons. Electron trapping is provided by a set of magnetic pinch coils at the top and
1478 bottom of the cavity and a multi-pole Ioffe or Halbach magnet serves to contain the
1479 atoms.

1480 The cavity design for the pilot-scale experiment consists of a large cylindrical cavity
1481 with a TE011 resonance of 325 MHz. Such a cavity is truly enormous, with a diameter
1482 of approximately 1.2 m and a height of 11 m. When an electron is produced inside
1483 the cavity with a cyclotron frequency that matches the TE011 resonant frequency it's
1484 cyclotron orbit couples the electron to the TE011, which drives a resonance in the cavity.
1485 These resonant fields can be read-out using an appropriate cavity coupling mechanism
1486 located at the center of the cavity. For more information on the cavity approach to
1487 CREs see Chapter 6.

1488 The bottom of the cavity has a cone termination to match the contour of the atom

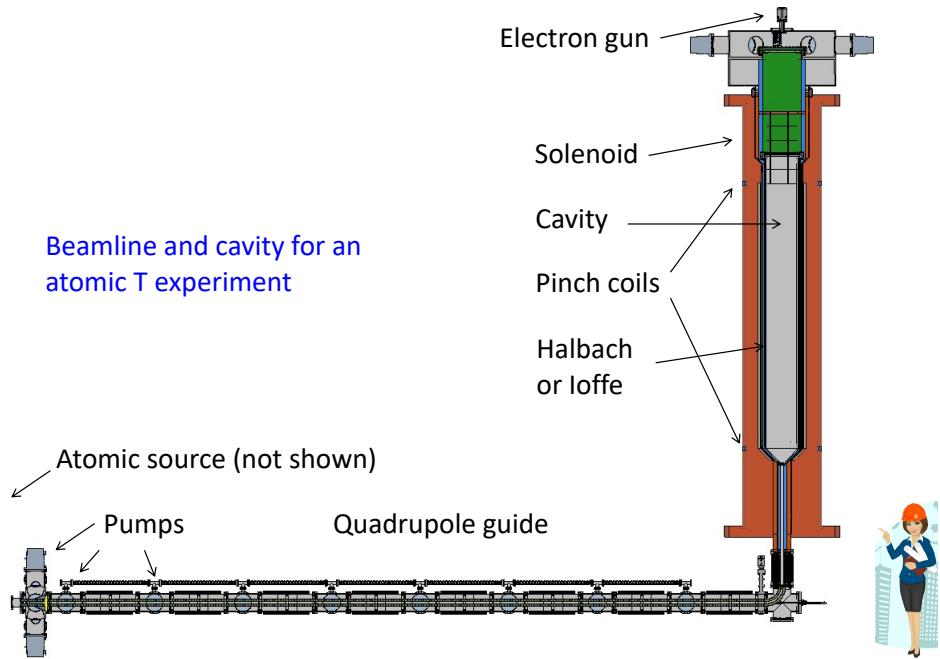


Figure 3.20: A conceptual sketch of a pilot-scale cavity CRES experiment with an atomic tritium beamline.

¹⁴⁸⁹ trapping magnet. This shape still allows for TE011 resonances with high internal Qs,
¹⁴⁹⁰ which are required for good SNR in the cavity experiment. A small opening in the bottom
¹⁴⁹¹ of the cone serves as an entry point for the tritium atoms. To allow for calibration of
¹⁴⁹² the magnetic field inhomogeneities with an electron gun, the top of the cavity is left
¹⁴⁹³ nearly completely open. Normally, this would drastically lower the Q-factor of the TE011
¹⁴⁹⁴ mode, but a specially configured coaxial partition is inserted at the top. This termination
¹⁴⁹⁵ scheme is designed to act as a perfect short for the TE011 mode since the circular shape
¹⁴⁹⁶ of the partition matches the electric field boundary conditions for the TE011 mode.
¹⁴⁹⁷ Simulations with HFSS have confirmed that this design results in a high quality TE011
¹⁴⁹⁸ resonance despite the nearly completely open end.

¹⁴⁹⁹ 3.6 Phase IV

¹⁵⁰⁰ The baseline CRES technology being pursued by the Project 8 collaboration are resonant
¹⁵⁰¹ cavities, which, due to their geometric properties, simple CRES signal structure, and low
¹⁵⁰² channel count, appear to be the better option for Phase IV. The current knowledge of the
¹⁵⁰³ antenna array CRES approach reveals no technical obstacles that would preclude it as a
¹⁵⁰⁴ baseline technology for Phase IV though it would most certainly be significantly more

1505 expensive. Therefore, antenna arrays represent a fallback approach if resonant cavities
1506 prove infeasible.

1507 The sensitivity of the pilot-scale atomic tritium experiment is estimated to be on
1508 the order of 0.1 eV, which means that increasing the sensitivity to reach the Phase IV
1509 goal will require an even larger experiment. Because of the direct coupling between the
1510 RF characteristics of a cavity and its geometry, the baseline plan is to build multiple
1511 copies of the pilot-scale experiment (see Figure 3.21) to obtain the required amount of
1512 volume rather than increase the size of the cavity beyond the pilot-scale. The built-in
1513 redundancy of this approach is useful in the sense that the experiment has no single
1514 point of failure, additionally, building several copies of the a pilot-scale experiment will
1515 minimize new engineering and design effort.

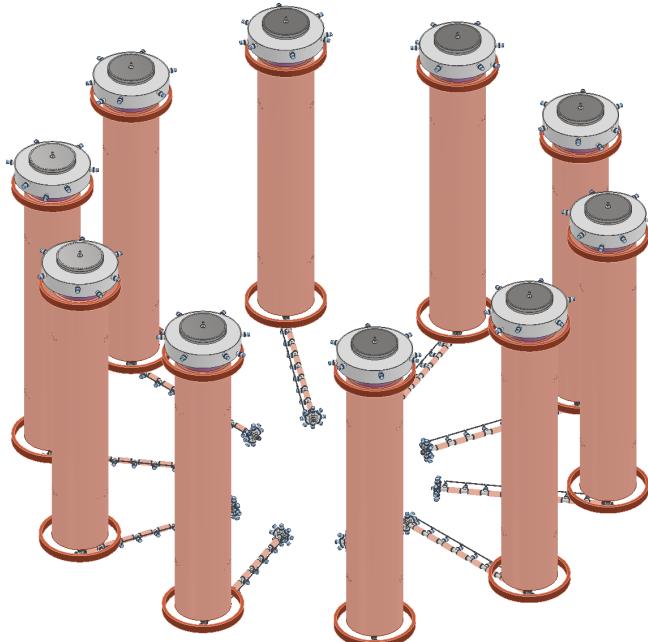


Figure 3.21: An illustration of a possible arrangement of ten pilot-scale cavity experiments for Phase IV. The experiments are arranged in a circle with an approximate diameter of 50 meters. Each atomic beamline connected to the bottom of each cavity is approximately 10 m in length. The cavities themselves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

1516 **Chapter 4**

1517 **Signal Reconstruction Techniques for An-**

1518 **tenna Array CRES and the FSCD**

1519 **4.1 Introduction**

1520 The transition from a waveguide CRES experiment to an antenna array CRES exper-
1521 iment introduces new challenges related to data acquisition, signal detection, and signal
1522 reconstruction caused by the multi-channel nature of the data. The development of signal
1523 reconstruction algorithms is crucial to the design of antenna array based experiments like
1524 the FSCD, because these algorithms directly influence the detection efficiency and energy
1525 resolution of the CRES experiment. In this Chapter I summarize my contributions to
1526 the development and analysis of signal reconstruction and detection algorithms for the
1527 FSCD experiment.

1528 In Section 4.2 I discuss the primary tool for this work, which is the Locust simulations
1529 package developed by the Project 8 experiment. Locust is used to simulate CRES events
1530 in the detector. Locust uses Kassiopeia to calculate particle trajectory solutions for
1531 electrons in the magnetic trap. The trajectories are then used to calculate the response
1532 of the antenna array to the cyclotron radiation produced by the electron, which results
1533 in signals that can be used to analyze the performance of different signal reconstruction
1534 algorithms. More recently, Project 8 has developed CREsana, which is a new simulations
1535 package that takes a more analytical approach to CRES signal simulations for antenna
1536 arrays. Although CREsana signals were not used for the signal reconstruction algorithm
1537 development detailed here, we introduce the software as it plays a role in the antenna
1538 array measurements presented in Section 5.4.

1539 In Section 4.3 I discuss the signal reconstruction and detection approaches analyzed
1540 for the FSCD experiment. In general there are two steps to signal reconstruction —
1541 detection and parameter estimation. With signal detection one is primarily concerned

1542 only with distinguishing between data that contains a signal versus data that contains only
1543 noise, whereas, with parameter estimation one is interested in extracting the kinematic
1544 parameters of the electron encoded in the cyclotron radiation signal shape. Due to
1545 the low signal power of electrons near the spectrum endpoint in the FSCD experiment,
1546 signal detection is a non-trivial problem. This is magnified by the need to maximize the
1547 detection efficiency of the experiment in order to achieve the neutrino mass sensitivity
1548 goals. My contributions to signal reconstruction analysis for the FSCD are focused on
1549 this signal detection component of reconstruction.

1550 After the discussion of various signal detection approaches, in Section 4.4 I present a
1551 more detailed analysis of the detection performance of three algorithms, which could be
1552 used to signal detection in the FSCD. This section was originally prepared for publication
1553 in JINST as a separate paper. The algorithms include a digital beamforming algorithm,
1554 a matched filter algorithm, and a neural network algorithm, which I analyze in terms of
1555 classification accuracy and estimated computational cost.

1556 **4.2 FSCD Simulations**

1557 Antenna array CRES and the FSCD requires a combination of different capabilities
1558 not often found in a single simulation tool. First of all, accurate calculations of the
1559 magneto-static fields produced by current-carrying coils are required in order to accurately
1560 model the magnetic trap and background magnets. The resulting magnetic fields must
1561 then be used to calculate the exact relativistic trajectory of electrons, which is required
1562 in order to calculate the electro-magnetic (EM) fields produced by the acceleration of
1563 the electron. Finally, the simulation has to model the interaction of the antenna and
1564 RF receiver chain with these EM-fields in order to produce the simulated voltage signals
1565 produced by the antenna array during the CRES event. At the time when Project 8 was
1566 developing this simulation capability, no single available simulation tool was known to
1567 adequately perform this suite of calculations, which prompted the development of custom
1568 simulation framework to simulate the FSCD. This simulation framework includes custom
1569 simulation tools developed by Project 8 as well as other open-source and proprietary
1570 software developed by third-parties.

1571 4.2.1 Kassiopeia

1572 Kassiopeia¹ is a particle tracking and static EM-field solver developed by the KATRIN
1573 collaboration for simulations of their spectrometer based on magnetic adiabatic collimation
1574 with an electrostatic filter [35]. Due to the measurement technique employed by the
1575 KATRIN collaboration, Kassiopeia is not designed to solve for the EM-fields produced by
1576 electrons in magnetic fields. However, it does provide efficient solvers for static electric
1577 and magnetic fields and charged particle trajectory solvers. Because of this, Project 8
1578 has incorporated parts of Kassiopeia into its own simulation framework.

1579 Magnetostatic Field Solutions

1580 The solutions to the electric and magnetic fields generated by a static configuration of
1581 charges and currents is given by Maxwell's equations in the limit where the time-dependent
1582 terms go to zero. In their static form Maxwell's equations [24] are

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

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

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

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

1583 where we can see that the electric and magnetic fields are now completely decoupled
1584 from each other. The solution for the magnetic field in this boundary value problem is
1585 given by the Biot-Savart law

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

1586 which Kassiopeia uses a variety of numeric integration techniques to solve for a user
1587 defined current distribution.

1588 Kassiopeia Simulation of the FSCD Magnetic Trap

1589 The trap developed for the FSCD experiment utilizes six current carrying coils, which
1590 surround a cylindrical tritium containment vessel (see Figure 4.1). Some critical aspects
1591 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 |

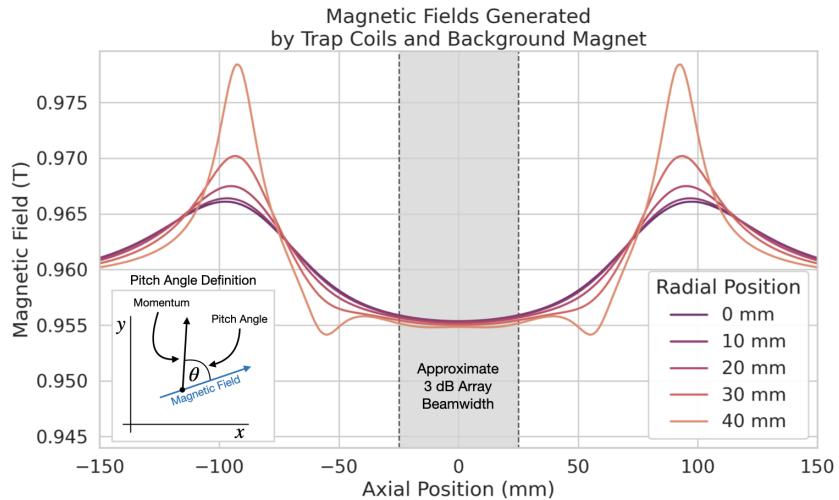
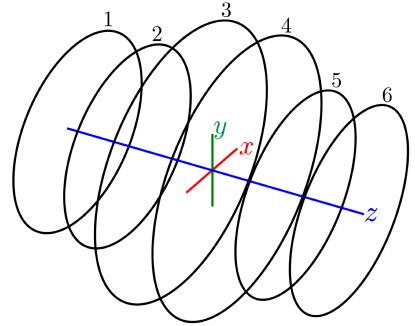


Figure 4.1: The geometry and parameters of the coils used to simulate the FSCD magnetic trap in Kassiopeia. Some axial profiles of the magnetic trap at different radial positions are show to demonstrate the shape of the magnetic field and trap depth as a function of position. Calculation of the magnetic 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° [36]. 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

1636 equations

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

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

1637 where \mathbf{r} is the position of the electron, \mathbf{p} is the electron's momentum, e is the charge of
1638 the electron, m is the electron's mass, and γ is the relativistic Lorentz term. To account
1639 for kinetic energy losses from radiation Kassiopeia includes an additional term in the
1640 momentum differential equation, which calculates the change in the electron's momentum
1641 induced by synchrotron radiation. Kassiopeia solves this pair of differential equations
1642 using numerical integration, however, the exact trajectory can be computationally
1643 intensive to solve. If the adiabatic approximation can be applied, then Kassiopeia can
1644 make use of a simpler set of equations that can be more readily solved numerically.

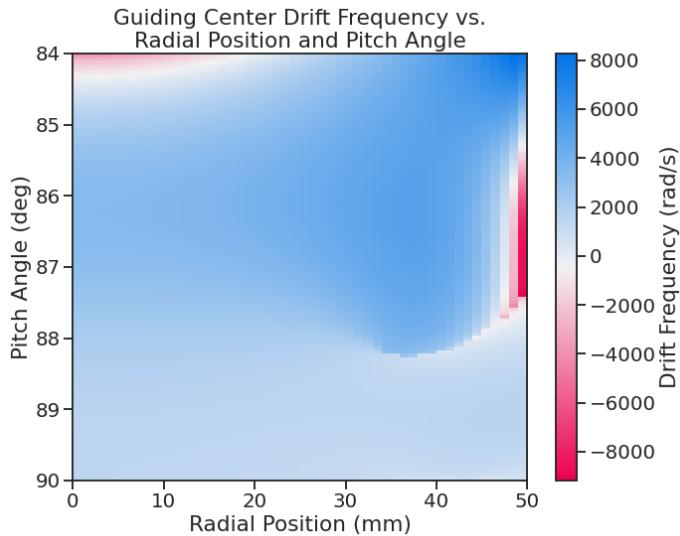


Figure 4.2: A map of the average ∇B -drift frequency for electrons trapped in the prototype FSCD trap shown in Figure 4.1. Negative drift frequencies indicate electrons that are drifting opposite to the standard direction, which means that they are close to escaping the magnetic trap.

1645 Even though Kassiopeia is not directly capable of simulating the cyclotron radiation,
1646 it is still an invaluable CRES simulation tool, due to the accurate trajectory solutions
1647 for electrons in magnetic traps. With Kassiopeia it is possible to test the efficiency of a
1648 particular trap design and analyze features of the electron trajectories that are important
1649 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 [37] 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 [38, 39], 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 [40],

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

1707 where V is the voltage and $|\mathbf{E}|$ is the magnitude of the incident electric field. To obtain the
 1708 antenna factor for the antennas developed for the FSCD Project 8 employs Ansys HFSS.
 1709 HFSS is a commercially available finite element method electromagnetic solver widely
 1710 used throughout the antenna engineering industry [41]. HFSS is capable of calculating
 1711 the antenna factor and gain patterns for complex antenna designs and outputting the
 1712 resulting quantities in the form of a text file that can be used as an input to the Locust
 1713 simulation.

1714 The antenna factor defines the steady-state response of the antenna to electromagnetic
 1715 plane waves and is a function of the frequency of the radiation. Therefore, in order to
 1716 apply the transfer function for the calculation of the antenna voltage response in the
 1717 time domain, Locust models the antenna as a linear time-invariant system [42]. In this
 1718 formalism the response of the system to the driving force is given by

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

1719 where $y[n]$ is the discretely sampled response, x is the driving force stimulus, and h is
 1720 the finite impulse response (FIR) filter. When applied to the FSCD array, this formalism
 1721 calculates the voltage time-series produced in each antenna by convolving the electric
 1722 field time-series with the antenna FIR filter, which is obtained by performing a inverse
 1723 Fourier transform on the transfer function from HFSS.

1724 Radio-frequency Receiver and Signal Processing

1725 After obtaining the voltage time-series by computing the electron trajectory and antenna
 1726 response, Locust simulates the signal processing associated with the radio-frequency
 1727 receiver chain. The standard receiver chain used in Locust simulations of the FSCD
 1728 attempts to mimic the operations that would actually occur in hardware (see Figure 4.3).

1729 Frequency down-conversion is used in the FSCD to reduce the digitization bandwidth
 1730 required to read-out CRES data. According to the Nyquist sampling theorem [43], the
 1731 minimal sampling rate that guarantees no information loss for a signal with a bandwidth
 1732 Δf is given by

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

1733 The total bandwidth of CRES signal frequencies from tritium beta-decay ranges from 0
 1734 to 26 GHz in a 0.95 T magnetic field, therefore, direct digitization of CRES signals from
 1735 the FSCD would require sampling frequencies greater than 50 GHz, which is infeasible for

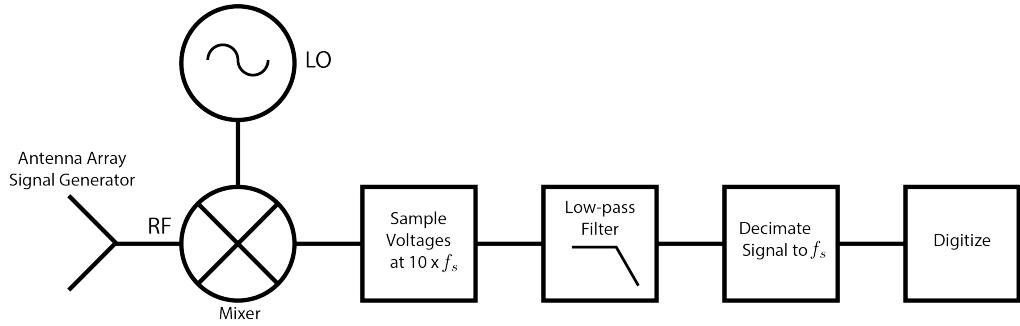


Figure 4.3: The receiver chain used by Locust when simulating CRES events in the FSCD.

1736 a real experiment. However, for the purposes of neutrino mass measurement we are only
 1737 interested in measuring the shape of the spectrum in the last 100 eV, which corresponds
 1738 to a frequency bandwidth of 5 MHz. Down-conversion is a technique for reducing the
 1739 base frequencies of signals in a bandwidth given by $[f_{\text{LO}}, f_{\text{LO}} + \Delta f]$ to the bandwidth
 1740 $[0, \Delta f]$, by performing the following multiplication

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

1741 In down-conversion the signal ($x(t)$) is multiplied by a sinusoidal signal with frequency
 1742 f_{LO} to reduce the absolute frequencies of the signals in the bandwidth. In the FSCD this
 1743 allows us to detect events in the last 100 eV of the tritium spectrum while sampling the
 1744 data far below 50 GHz. The standard bandwidth used in the FSCD is 200 MHz, which
 1745 allows for higher frequency resolution than the minimum sampling frequency for 100 eV
 1746 of energy bandwidth.

1747 Trying to directly simulate down-conversion with a frequency multiplication in Locust
 1748 would require the sampling of the electric fields at each antenna in the FSCD array with
 1749 a period of ≈ 20 ps, which is extremely slow computationally. To avoid this Locust
 1750 performs the down-conversion by intentionally under-sampling the electric fields with
 1751 a frequency of 2 GHz. Sampling below the Nyquist limit causes the higher frequency
 1752 components of the CRES signal to alias, however, Locust can remove these aliased
 1753 frequency peaks using a combination of low-pass filtering and decimation to recreate
 1754 frequency down-conversion. After filtering and decimation, Locust simulates digitization
 1755 by an 8-bit digitizer at a sampling frequency of 200 MHz to recreate the conditions of
 1756 the FSCD. The voltage offset and the digitizer range must be configured by the user
 1757 based on the characteristics of the simulation.

1758 **Data**

1759 The output of Locust simulations for the FSCD primarily consists of two data files. The
1760 first is the electron trajectory information calculated by Kassiopiea, which is output in
1761 the form of a `.root` file [44]. This file contains important kinematic information about
1762 the electron such as it's position and pitch angle as a function of time. The other file is
1763 produced by Locust and it contains the digitized signals acquired from each antenna in
1764 the FSCD array. The Locust output files conform to the Monarch specification developed
1765 by Project 8, which is based on the commonly used HDF5 file format, and matches the
1766 format of the files produced by the Project 8 data acquisition software. This makes it
1767 possible to use the same data analysis code to analyze both simulated and real data.

1768 **4.2.3 CRESana**

1769 Locust is the primary simulation tool used by Project 8 in the development and simulation
1770 of the FSCD. However, simulations of CRES events in larger antenna arrays (≥ 100
1771 antennas) using Locust can take several hours to complete, which is prohibitively long
1772 when one is performing a sensitivity analysis for a large scale antenna experiment. One
1773 of the reasons for Locust's slow operation is that the electric fields from the electron
1774 must be solved numerically for each time-step for each of the antennas in the array.
1775 These numerical solutions allow Locust to accurately simulate the electric fields from
1776 arbitrarily complicated electron trajectories at the cost of more computations and slower
1777 simulations. Therefore, an additional simulation tool that sacrifices some accuracy for
1778 computational efficiency would be extremely useful simulations and sensitivity analyses
1779 of larger antenna array experiments.

1780 To fill this need, Project has developed a new simulations package called CRESana³,
1781 specifically designed to perform analytical simulations of antenna array based CRES
1782 experiments. CRESana is not as flexible as Locust, but it provides a significant increase
1783 in simulation speed. It does this by using well-justified analytical approximations of the
1784 electrons motion in the magnetic field and the resulting electric fields from the electron's
1785 acceleration. The electric fields and signals generated by CRESana are consistent with
1786 theoretical calculations of the electron's radiation, and are test for accuracy using
1787 well-known test-case simulations and consistency checks.

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

4.3 Signal Detection and Reconstruction Techniques for Antenna Array CRES

1790 Antenna Array CRES Signal Reconstruction

1791 A robust set of FSCD simulation tools are vital to the development of the analysis
1792 algorithms necessary for antenna array CRES to succeed. In order to perform CRES
1793 measurements using an antenna array, one must develop an algorithm that uses the
1794 multi-channel time-series obtained by digitizing the array to estimate the starting kinetic
1795 energies of electrons produced in the magnetic trap. This procedure consists of a multi-
1796 stage process of detecting a CRES signal then estimating the parameters of the electron
1797 that produced and is often referred to as simply CRES signal reconstruction.

1798 Compared with the signal reconstruction approaches of the Phase I and II CRES
1799 experiments, antenna array CRES requires a significantly different approach to signal
1800 reconstruction. In Phase I and II, CRES was performed using a waveguide gas cell that
1801 could be directly connected to a waveguide transmission line. The transmission line
1802 efficiently transmits the cyclotron radiation along its length to an antenna at either end
1803 of the waveguide. However, with an antenna array the electron is essentially radiating
1804 into free-space, therefore, the cyclotron radiation power collected by the array is directly
1805 proportional to the solid angle surrounding the electron that is covered with antennas.
1806 Because it is not practical to fully surround the magnetic trap with antennas, some of the
1807 cyclotron radiation power that would have been collected by the waveguide escapes into
1808 free-space. Furthermore, the power that is collected by the antenna array is split between
1809 every channel in the antenna array, which significantly lowers the signal-to-noise ratio
1810 (SNR) of CRES signals in a single antenna channel compared to a waveguide apparatus.
1811 Therefore, a suite of completely new signal reconstruction techniques are needed in order
1812 to perform CRES in the FSCD.

1813 Changes to the approach to CRES signal reconstruction are also motivated by the
1814 more ambitious scientific goals of the FSCD experiment. A measurement of the tritium
1815 beta-decay spectrum that is sensitive to neutrino masses as small as 40 meV requires that
1816 we measure the kinetic energies of individual electrons with a total energy broadening
1817 of 115 meV [45]. This resolution includes all sources of uncertainty in the electron's
1818 kinetic energy such as magnetic field inhomogeneities. This level of energy resolution is
1819 compatible only with an event-by-event signal reconstruction approach where the kinetic
1820 energies, pitch angles, and other parameters of the CRES events are estimated before

1821 constructing the beta-decay spectrum.

1822 The event-by-event approach is distinct from the analysis done for the Phase I and
1823 Phase II experiments where only the starting cyclotron frequency of the event was
1824 estimated by analyzing the tracks formed by the carrier frequency in the time-frequency
1825 spectrogram. These frequencies were then combined into a frequency spectrogram, which
1826 was converted to the beta-decay energy spectrum using an ensemble approach that
1827 averaged over all other event parameters. The ensemble approach to signal reconstruction
1828 results in poor energy resolution because other kinematic parameters such as pitch angle
1829 change the cyclotron carrier frequency due to changes in the average magnetic field
1830 experience by the electron, and it is therefore incompatible with the future goals of the
1831 Project 8 collaboration.

1832 Components of Reconstruction: Signal Detection and Parameter Estimation

1833 CRES signal reconstruction can be viewed as a two-step procedure consisting of signal
1834 detection followed by parameter estimation. In the former, one is concerned with
1835 identifying CRES signals in the data regardless of the signal parameters, whereas, in the
1836 latter one operates under the assumption that a signal is present and then estimates its
1837 parameters.

1838 More formally, signal detection is essentially a binary hypothesis test between the
1839 signal and noise data classes and parameter estimation describes a procedure of fitting a
1840 model to the observed data. While both of these processes are required for a complete
1841 reconstruction (see Figure 4.4), the focus of my work and this chapter is on the signal
1842 detection aspect of antenna array CRES signal reconstruction.

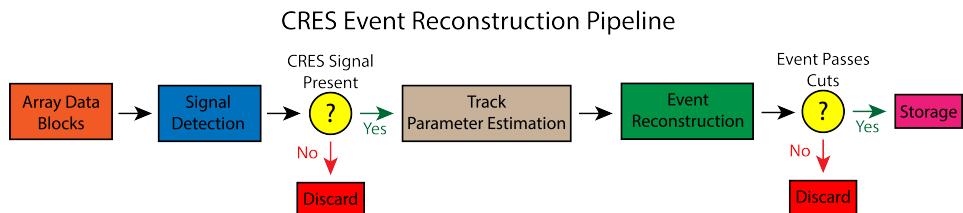


Figure 4.4: A high-level diagram depicting the process of CRES event reconstruction. The first step consists of identifying the presence of a signal in the data. This step is necessary to avoid the danger of performing a reconstruction of a false event, which would constitute a background contribution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

1843 **Detection Theory**

1844 The problem of signal detection can be posed as a statistical hypothesis test [46]. For
1845 CRES signals, which are essentially vectors with added white Gaussian noise (WGN),
1846 one needs to choose between two hypotheses

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

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

1847 where \mathbf{y} is the CRES data vector, $\boldsymbol{\nu}$ is a sample of WGN, and \mathbf{x} represents the CRES
1848 signal. The hypothesis that the data contains only noise is labeled \mathcal{H}_0 and the hypothesis
1849 that the data contains a signal is labeled \mathcal{H}_1 .

1850 For illustrative purposes one can examine the case where one the first sample of
1851 data is used to distinguish between \mathcal{H}_0 and \mathcal{H}_1 . The value of the first data sample is
1852 distributed according to two gaussian distributions corresponding to \mathcal{H}_0 and \mathcal{H}_1 (see
1853 Figure 4.5). By setting a decision threshold on the value of this sample, one can choose
1854 the correct hypothesis with a probability given by the areas underneath the probability
1855 distribution curves. A true positive corresponds to correctly identifying that the data
1856 contains signal, whereas, a true negative means that one has correctly identified the data
1857 as noise. The rate at which the detector performs a true positive classification is given
1858 by the green region underneath $p(\mathbf{y}[0]; \mathcal{H}_0)$, and the rate at which the detector performs
1859 a true negative classification is given by the orange region underneath $p(\mathbf{y}[0]; \mathcal{H}_1)$. Two
1860 types of misclassifications are possible. Either we declare noise data as signal, which is
1861 call a false positive, or we declare signal data as noise, which is a false negative. Note
1862 that it is only possible to trade off these two types of errors by tuning the detection
1863 threshold. One cannot simultaneously reduce the rate of false positives without also
1864 increasing the rate of false negatives.

1865 The approach taken with CRES signals is to fix the rate of false positives by setting
1866 a minimum value for a detection threshold. The rate of false positives that is acceptable
1867 at the detection stage depends upon the rate of background events compatible with the
1868 sensitivity goals of the experiment. The ultimate goal of a neutrino mass measurement
1869 with 40 meV sensitivity in general has strict requirements on the number of background
1870 events, which requires a relatively high detection threshold to achieve. Consequently,
1871 the ideal signal detection algorithm is the one that achieves the maximum rate of true
1872 positives for a fixed rate of false positives, so that the detection efficiency of the experiment
1873 is maximized and potential sources of background are kept to a minimum.

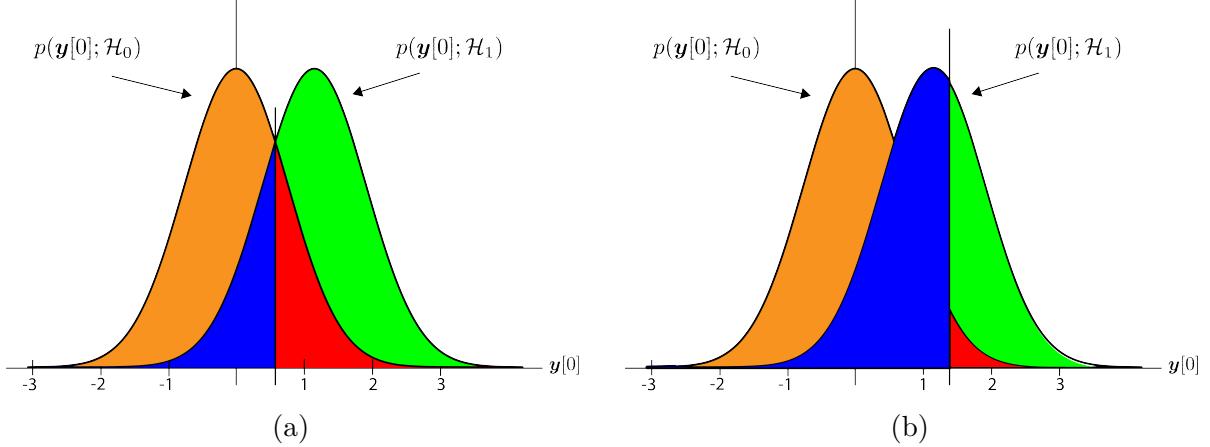


Figure 4.5: An illustration of two PDFs associated with a binary hypothesis test. The decision threshold is represented by the vertical line that partitions both distributions. The orange and red areas correspond to the true negative and false positive probabilities and the blue and green areas correspond 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 [47], 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

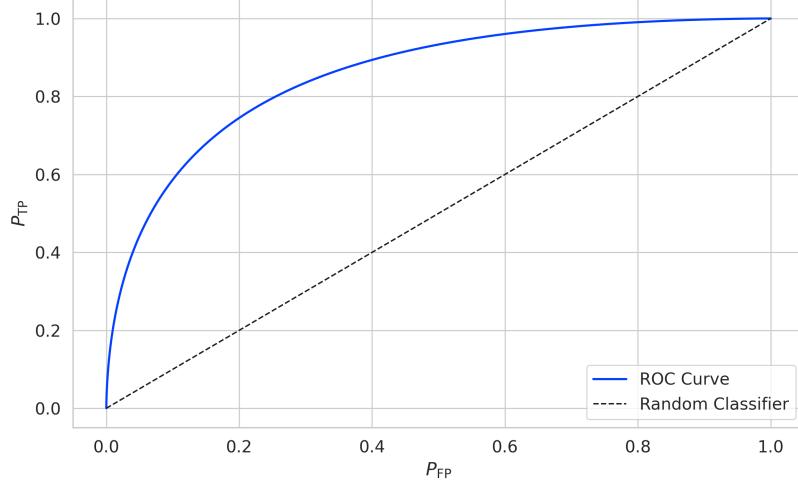


Figure 4.6: An example ROC curve formed by computing the P_{FP} and the P_{TP} for a given likelihood ratio test. As the decision threshold is increased P_{FP} decreases at the expense of a lower P_{TP} . The black dashed line indicates the lower bound ROC curve obtained by randomly deciding between \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 [40]. 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

1902 emitted by the array will constructively interfere at the point of interest (see Figure
 1903 4.7). As a consequence of the principle of reciprocity [48], when the array is operating in
 1904 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-

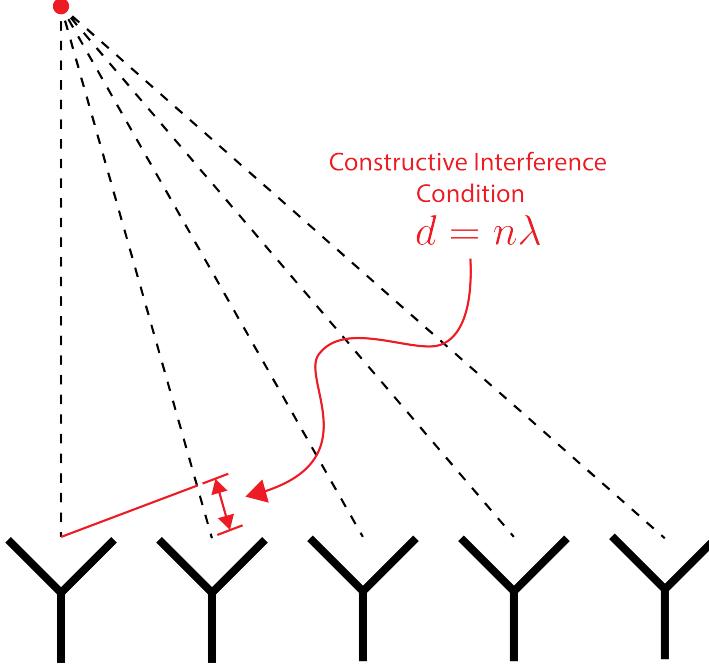


Figure 4.7: An illustration of the constructive interference condition which is the operating principle of digital beamforming using a uniform linear array as an example.

1905
 1906 length difference to the beamforming point between different antennas in the array. The
 1907 relationship between the phase delay and the path-length difference is given by the
 1908 familiar equation

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

1909 where ϕ is the phase delay, d is the path-length difference, and λ is the wavelength of
 1910 the radiation. In practice, one chooses the values of d by specifying the beamforming
 1911 positions of interest and then calculates the beamforming phases using Equation 4.20,
 1912 which is guaranteed to follow the constructive interference condition shown in Figure 4.7.

1913 Beamforming can be neatly expressed mathematically using the vector equation

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

1914 where $\mathbf{x}[n]$ is the array snapshot vector, $\Phi[n]$ is a vector of beamforming shifts, and
 1915 $y[n]$ is the resulting summed signal. The beamforming shifts consist of a set of complex

1916 numbers that contain the beamforming phase shift and an amplitude weighting factor,

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

1917 where the set of magnitudes $A_i[n]$ are amplitude weighting factors and $\phi_i[n]$ are the phase
1918 shifts from the path-length differences. The index i is used to denote the antenna channel
1919 number. The amplitude weighting factor is the relative magnitude of the signal received
1920 by a particular antenna to the other antennas in the array, such that the antennas that
1921 receive signals with higher amplitude, due to being closer to the source, have more
1922 weight in the beamforming summation. The input and outputs signals beamforming
1923 are naturally expected to be functions of time as indicated by the index $[n]$, however, it
1924 is also possible to use time dependent beamforming phases that shift the beamforming
1925 position of the array over time.

1926 Digital beamforming is the type of beamforming algorithm of interest to Project 8 for
1927 CRES. Specifically, digital beamforming means that the beamforming phases are applied
1928 to the array signals in software rather than employing fixed beamforming phase shifts in
1929 the receiver chain hardware. The advantage of digital beamforming is that for a given
1930 series of array snapshots one can specify a large number of beamforming positions and
1931 effectively search for electrons by performing the beamforming summation associated
1932 with each point and applying a signal detection algorithm to identify the presence of a
1933 CRES signal.

1934 One of the most attractive features of digital beamforming is the spatial filtering
1935 effect, which is a direct consequence of the constructive interference condition used to
1936 define the beamforming phases. Spatial filtering allows for signals from multiple electrons
1937 at different positions in the trap to be effectively separated, because the constructive
1938 interference condition will force the signals from electrons at positions different from the
1939 beamforming position to cancel. This helps to reduce signal pile-up that could become
1940 an issue for large scale CRES experiments using a dense tritium source.

1941 The digital beamforming positions can be specified with arbitrary densities limited
1942 only by the available computational resources. This provides a very straight-forward way
1943 to estimate the position of the electron in the trap by using a dense grid of beamforming
1944 positions and maximizing the output power of the beamforming summation over this
1945 grid. This natural approach to position reconstruction is attractive due the requirements
1946 of an event-by-event signal reconstruction, which needs an accurate estimation of the
1947 exact magnetic field experienced by the electron in order to correctly estimate it's kinetic

1948 energy. Combined with an accurate map of the magnetic field inhomogeneities of the
 1949 trap obtained from calibrations, beamforming allows one to apply this magnetic field
 1950 correction with a spatial resolution that is a fraction of the cyclotron wavelength.

1951 **Laboratory Beamforming Demonstrations**

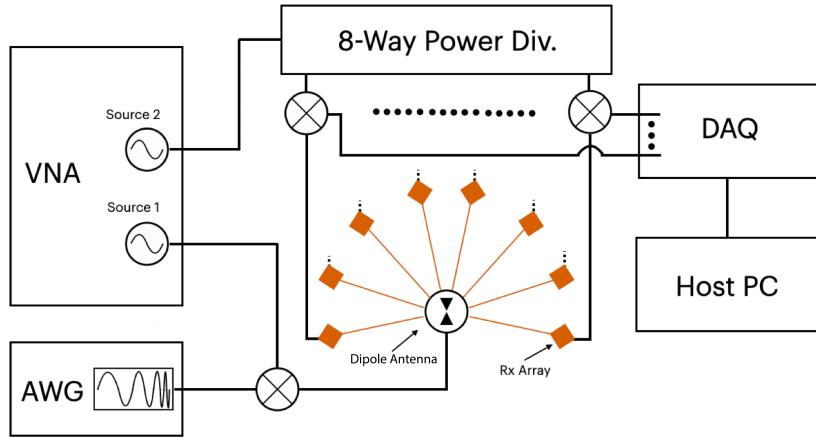


Figure 4.8: System level diagram of the laboratory setup used for beamforming demonstrations at Penn State. For more information on this system see Chapter 5. Signals near 26 GHz are fed to a dipole antenna using an arbitrary waveform generator (AWG) and vector network analyzer (VNA), which drive a mixer.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

1952 As part of the development of antenna array CRES for the FSCD, an antenna
 1953 measurement setup was constructed at Penn State to serve as a testbed for antenna
 1954 prototypes and to perform laboratory validations of array simulations. This system
 1955 is discussed in more detail in Chapter 5. Early versions of the antenna measurement
 1956 system (see Figure 4.8 and Figure 4.9) were used to perform beamforming reconstruction
 1957 studies of a simple probe antenna to better understand the principles of beamforming
 1958 and confirm the estimated beamforming performance of Locust.

1959 Signals from an arbitrary waveform generator were up-converted to 26 GHz using a
 1960 mixer and a high-frequency source from a vector network analyzer and fed to the dipole
 1961 antenna through a balun. The radiation from the dipole antenna was received by an
 1962 array of horn antennas. The signals from the horn antennas were then down-converted
 1963 to baseband using a collection of mixers and an 8-way power divider. The signals were
 1964 then digitized and saved to a host computer for analysis.

1965 The data collected using the dipole and horn antenna array is reconstructed using the

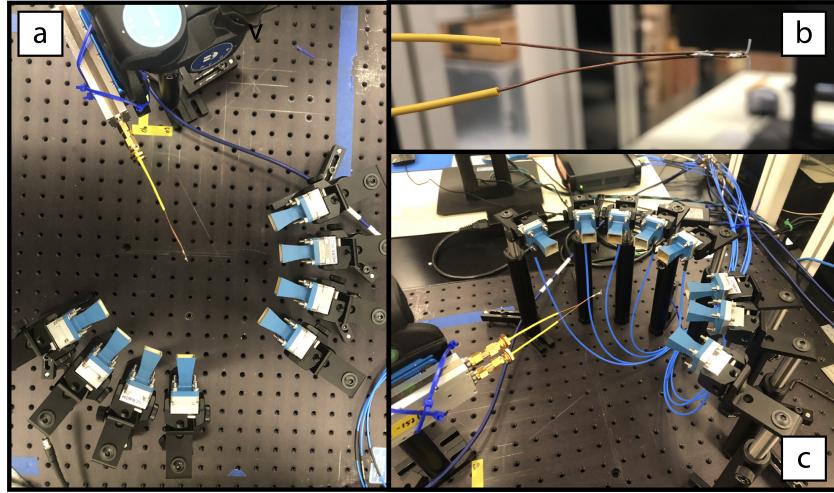


Figure 4.9: Photographs of the beamforming demonstration setup. In (a) I show a top-down view of the dipole antenna and the array of eight horn antennas. Manual repositioning of the horn antennas allows one to synthesize a full-circular antenna array. The dipole antenna is mounted on a camera tripod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

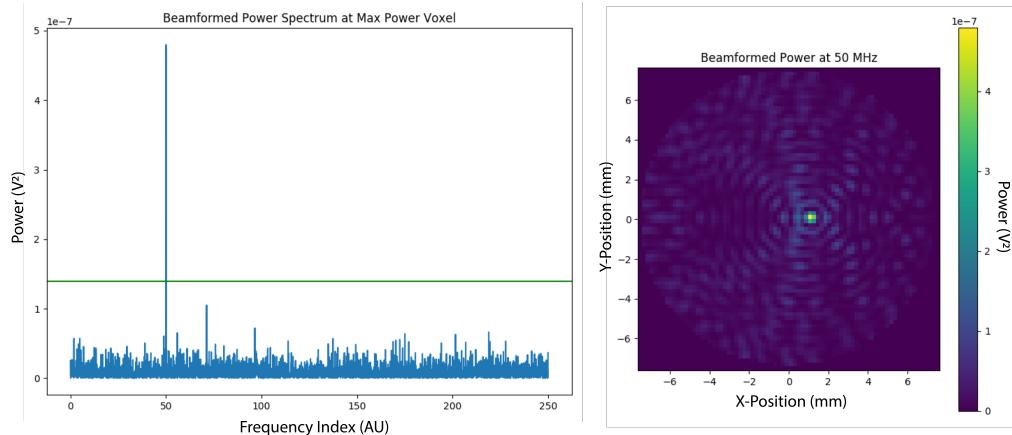


Figure 4.10: An example of digital beamforming reconstruction of a dipole antenna using a synthetic array of horn antennas. The beamforming image on the right is constructed by computing the time-averaged power of the summed signals for a two-dimensional grid of beamforming positions. In the image one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

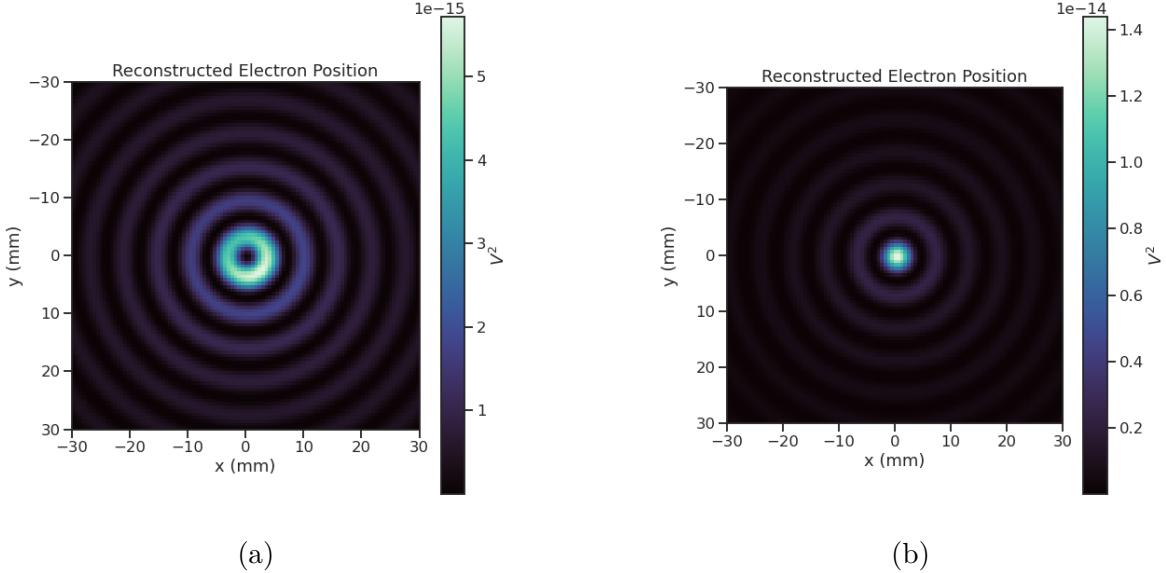


Figure 4.11: Beamforming images visualizing the reconstruction of an electron without (a) and with (b) the cyclotron phase correction. The images were generated using data from Locust simulations. The cyclotron phase refers to a phase offset equal to the relative azimuthal position of an antenna in the 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.

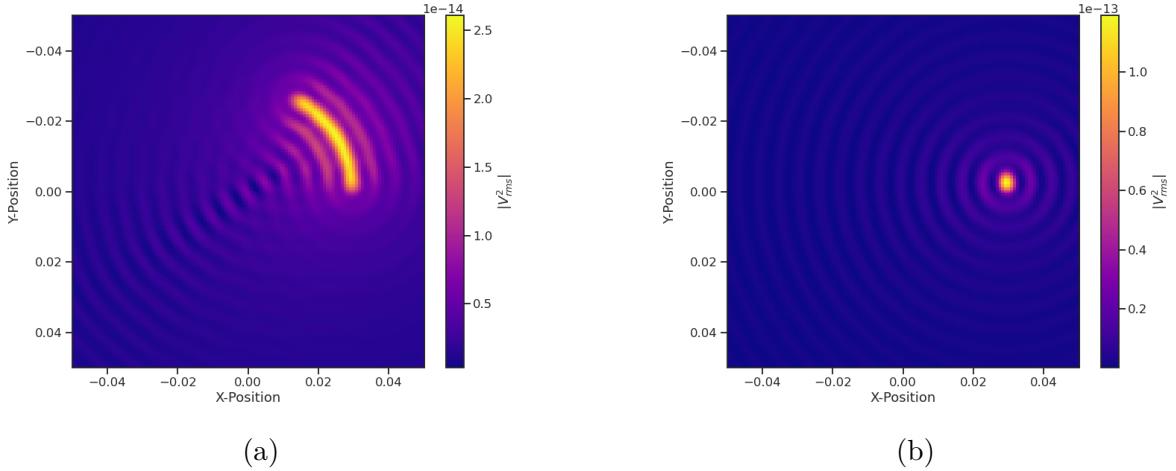
If the beamforming phases consist only of the spatial phase component from Equation 4.20, then the resulting beamforming image contains a relatively high-power ring-shaped region that is centered on the position of the electron (see Figure 4.11a). The origin of this shape is an additional phase offset 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

2016 antenna channel. Using the updated beamforming phases in the summation changes the
 2017 ring feature into a Bessel function peak whose maximum corresponds to the position of
 2018 the electron. Including this cyclotron phase correction significantly improves the signal
 2019 detection and reconstruction capabilities of beamforming by more than doubling the
 2020 summed signal power and shrinking the beamforming maxima feature size.

2021 The beamforming image examples in Figure 4.11 were produced using an electron
 2022 located on the central axis of the magnetic trap, which do not experience ∇B -drift.
 2023 However, for electrons produced at non-zero radial position the beamforming phases
 2024 must be made time-dependent in order to track the position of the electron's guiding
 2025 center over time. Without this correction the ∇B -drift causes the electron to move
 2026 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



2027 Figure 4.12: Beamforming images visualizing the reconstruction of an electron located
 2028 off the central axis of the FSCD trap. In (a) we performing beamforming without the
 2029 ∇B -drift correction, and in (b) we include the ∇B -drift correction.

2027 reduces the power of the beamforming maxima and increases the size of the beamforming
 2028 features, simultaneously harming detection efficiency and position reconstruction.

2029 The ∇B -drift correction simply adds a circular time-dependence to the beamforming
 2031 positions as a function of time,

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

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

2032 where $\omega_{\nabla B}$ is the drift frequency and $t[n]$ is the time vector. In the ideal case the ∇B -drift

2033 frequencies from Figure 4.2 for the correct pitch angle and radial position would be used,
2034 however, it is not possible to know the electron’s pitch angle a priori. In principle, one
2035 could perform multiple beamforming summations for a given beamforming position using
2036 different drift frequencies and choose the one that maximizes the summed power, but
2037 this approach leads to a huge computational burden that would be impractical for a
2038 real FSCD experiment. A compromise is to use an average value of $\omega_{\nabla B}$ obtained by
2039 averaging over the drift frequencies for electrons of different pitch angle at a particular
2040 radius. This approach keeps the computational cost of time-dependent beamforming to a
2041 minimum while still providing a significant increase in the detection efficiency of digital
2042 beamforming.

2043 **Signal Detection with Beamforming and a Power Threshold**

2044 Up to this point we have neglected any specific discussion of how digital beamforming is
2045 used for signal detection and reconstruction. This is because, strictly speaking, digital
2046 beamforming consists only of the phased summation of the array signals and cannot
2047 be used alone for signal detection. The example beamforming images shown in Figure
2048 4.11 and Figure 4.12 were produced using simulated data that contained no noise, which
2049 significantly degrades the utility of analyzing the beamforming images for signal detection
2050 and reconstruction.

2051 Digital beamforming as a detection algorithm is understood to mean digital beam-
2052 forming plus a detection threshold placed on the amplitude of the frequency spectrum
2053 obtained by applying a fast Fourier transform (FFT) to the summed time-series (see
2054 Figure 4.13). This approach is most similar to the time-frequency spectrogram analysis
2055 employed in previous CRES experiments, however, in principle any signal detection
2056 algorithm could be used after the beamforming procedure. In Section 4.4 I analyze the
2057 signal detection performance of the power threshold approach in detail.

2058 From the example frequency spectra in Figure 4.13 it is clear that without a re-
2059 construction technique that coherently combines the signals from the full antenna our
2060 ability to detect CRES signals will be drastically reduced. Because the CRES signals are
2061 in-phase at the correct beamforming position the summed power increases as a function
2062 of N^2 compared to a single antenna channel, where N is the number of antennas. It
2063 is true that the noise power is also increased by beamforming, but, because the noise
2064 is incoherent, its power only increases linearly. Consequently, the signal-to-noise ratio
2065 (SNR) of the CRES signal increases linearly with the number of antennas, which greatly
2066 improves detection efficiency compared to using only the information in a single antenna.

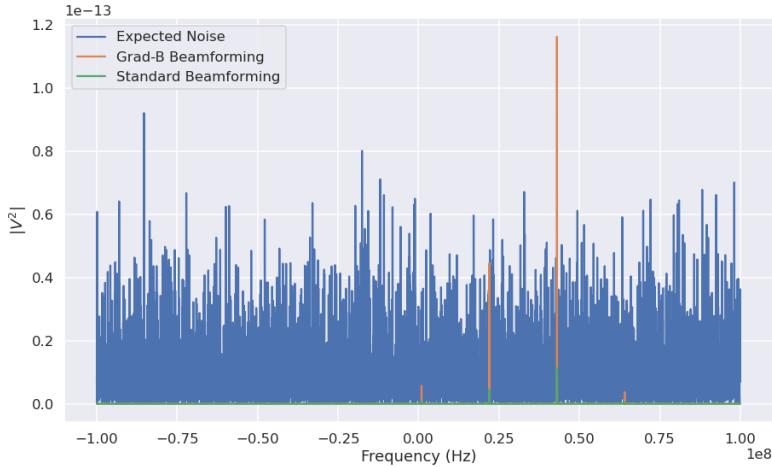


Figure 4.13: A plot of a typical frequency spectrum obtained by applying a Fourier transform to the time-series obtained from beamforming. The frequency spectra are plotted without noise on top of an example of a typical noise spectrum to visualize a realistic signal-to-noise ratio. In the example we 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.

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

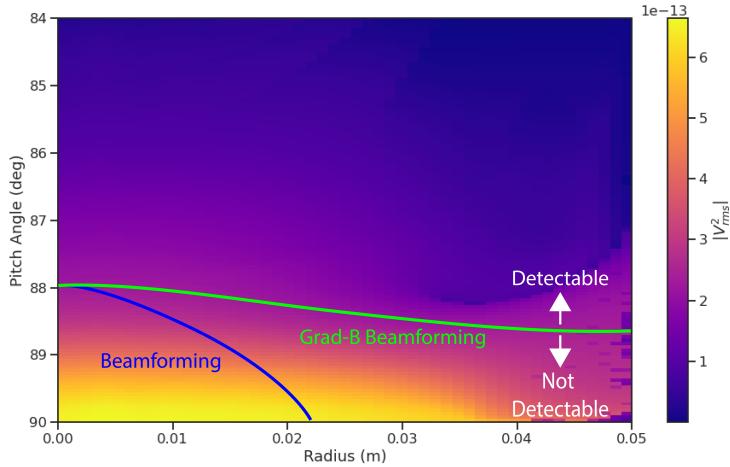


Figure 4.14: A plot of the total signal power received by the FSCD array from trapped electrons with different radial positions and pitch angles generated using Locust simulations. The lines on the plot indicate a 10 dB detection threshold above the mean value of the noise in the frequency spectrum. 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.

4.3.2 Matched Filtering

Introduction to Matched Filtering

The problem of CRES signal detection is the problem of detecting a signal buried in WGN, which has been examined at great depth in the signal processing literature [46]. For a fully known signal in WGN the optimal detector is the matched filter, which means that it achieves the highest true positive rate for a fixed rate of false positives. The matched filter test statistic is calculated by taking the inner product of the data with the matched filter template

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

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

2094 data is composed of the known signal with additive WGN. The matched filter template
 2095 is obtained by rescaling the known signal in the following way

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

2096 where τ is the variance of the WGN and $x[n]$ is the known signal. Strictly speaking,
 2097 Equation 4.27 is only true for noise with a diagonal covariance matrix, however, in the
 2098 context of the FSCD we are justified in assuming this to be true. Defining the matched
 2099 filter templates in this way guarantees that the expectation value of \mathcal{T} is equal to one
 2100 when the data contains only noise, which is the standard matched filter normalization in
 2101 the signal processing literature.

2102 Although matched filters are canonically formulated in terms of a perfectly known
 2103 signal, it is still possible to apply the matched filter technique given imperfect information
 2104 about the signal provided that the signal is deterministic. From our discussion of CRES
 2105 simulation tools for the FSCD (see Section 4.2) we know that the shape of CRES signals
 2106 are completely determined by the initial parameters of the electron. The random collisions
 2107 with background gas molecules which cause the formation of signal tracks are the only
 2108 stochastic component of the CRES event after the initial beta-decay, therefore, it is
 2109 possible to develop a matched filter for the detection of CRES signal tracks which are fully
 2110 determined by the parameters of the electron after the initial beta-decay or subsequent
 2111 collision events.

2112 The matched filter test statistic for CRES signals is a modified version of Equation
 2113 4.26

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

2114 where the matched filter inner product has been replaced with a convolution operation
 2115 and a maximization over the template and convolution delay (m). Replacing the inner
 2116 product with a convolution accounts for the fact that the start time of the CRES signal is
 2117 now an unknown parameter, in addition, we now perform a maximization of the matched
 2118 filter convolution over a number of different templates. Because the shape of the signal is
 2119 unknown we are forced to guess a number of different signal shapes to create a template
 2120 bank with which we can identify unknown signals by performing an exhaustive search.

2121 The template bank approach to matched filtering, while quite powerful, can quickly
 2122 become computationally intractable. This is especially true in the case of the FSCD
 2123 because of the large amount of raw data produced by the array that must be analyzed.

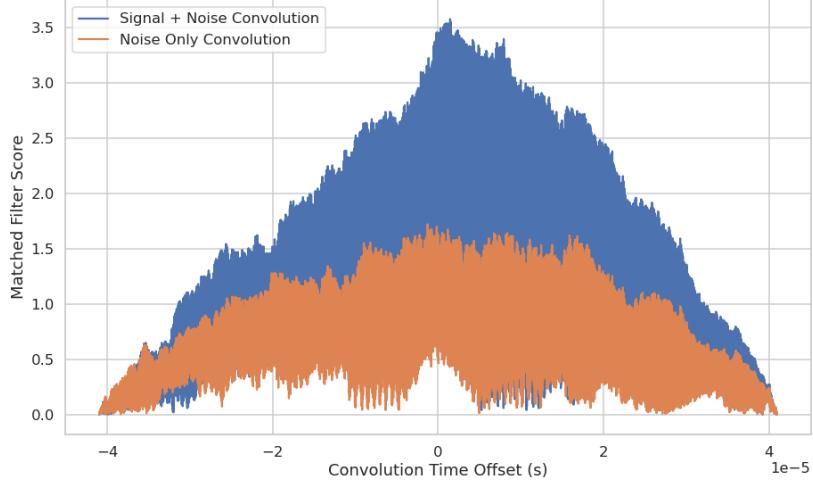


Figure 4.15: Example of a convolution of a CRES signal template with a segment of noisy data. A simulated CRES signal was simulated using Locust and normalized to create a matched filter template. When this template is convolved with noisy data the contains the matching signal the convolution output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

2124 Specifically, the time-domain convolution specified by Equation 4.28 is particularly
 2125 computationally intensive and is a major barrier towards the implementation of a
 2126 matched filter for signal detection in an experiment like the FSCD. This can be avoided
 2127 by using the convolution theorem to replace the time-domain convolution with an inner
 2128 product in the frequency domain.

2129 The convolution theorem states that

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

2130 where \mathbf{f} and \mathbf{g} are discretely sampled time-series, \mathbf{F} and \mathbf{G} are the respective discrete
 2131 Fourier transforms, and \mathcal{F}^{-1} is the inverse discrete Fourier transform operator. The
 2132 convolution theorem allows us to perform the matched filter convolution by first com-
 2133 puting the Fourier transform of the template and data, then performing a point-wise
 2134 multiplication of the two frequency series, and finally performing the inverse Fourier
 2135 transform to obtain the convolution output. Because discrete Fourier transforms can be
 2136 performed extremely efficiently, the convolution theorem is almost always used in lieu of
 2137 directly computing the convolution.

2138 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.

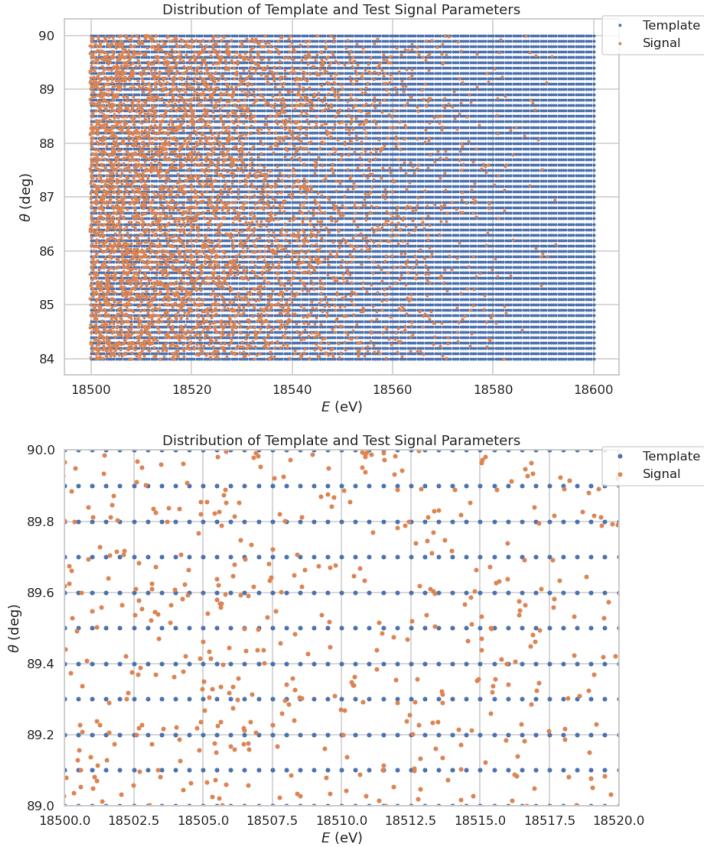


Figure 4.16: An example two-dimensional parameter distribution of a matched filter template bank and random test signals. θ refers to the pitch angle of the electron and E is the kinetic energy. The template bank forms a regular grid of in pitch angle and energy, whereas, the test signals are uniformly 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 $> 90\%$.

2173 This introduces uncertainty into the signal parameter estimation that manifests as an
 2174 energy broadening. Dense grids of matched filter templates allows one to quantify this
 2175 broadening by analyzing the parameter space of templates with matched filter scores
 2176 close to the ground truth. This approach is analogous to maximum likelihood estimation
 2177 and is one key component of a complete sensitivity analysis for an antenna array CRES
 2178 experiment.

2179 A key parameter for describing the performance of a matched filter template bank at
 2180 signal detection is match, which we define as the average ratio of the highest matched

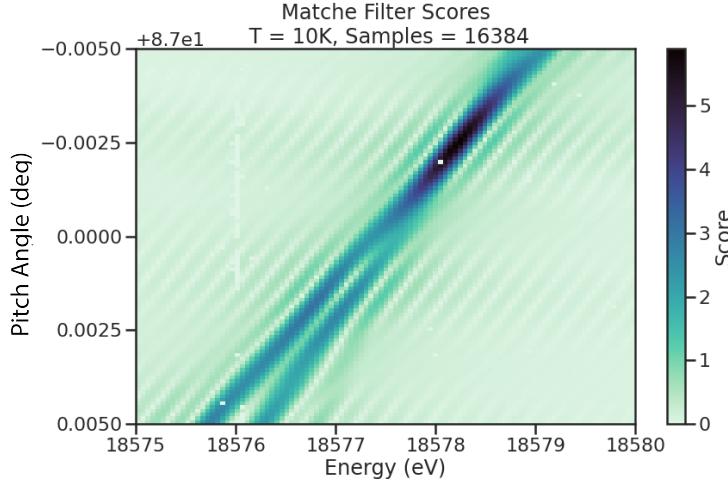


Figure 4.17: The matched filter scores of a dense grid of templates in pitch angle energy space. Dense template grids allow one to estimate the kinetic energy of the electron by identifying the best matching template. The uncertainty on this value is proportional to the space of templates that also match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

2181 filter score for a random signal to the matched filter score for a perfectly matching
2182 template. In equation form this is

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

2183 where $\mathcal{T}_{\text{best}}$ is the matched filter score of the best fitting template in the bank and $\mathcal{T}_{\text{ideal}}$ is
2184 the hypothetical matched filter score one would measure if the signal perfectly matched
2185 the template. Generally, one desires an average match as close to one as possible, however,
2186 the average match value is an exponential function of the number of templates in the
2187 template bank (see Figure 4.18). This behavior is observed for dense matched filter grids
2188 like the one in Figure 4.17. A dense grid was used to calculate the average value of match
2189 for different template bank sizes shown in Figure 4.18.

2190 The exponential relationship between match and template bank size is also evident
2191 for template banks that cover a wide range of parameters, such as the template bank
2192 visualized in Figure 4.16. Since no prior knowledge of the signal parameters is available,
2193 one has no choice but to use a template bank that covers a large range of parameters for
2194 signal detection. Achieving a high average match in this scenario can easily overwhelm
2195 the available computational resources, so in practice only a limited number of templates

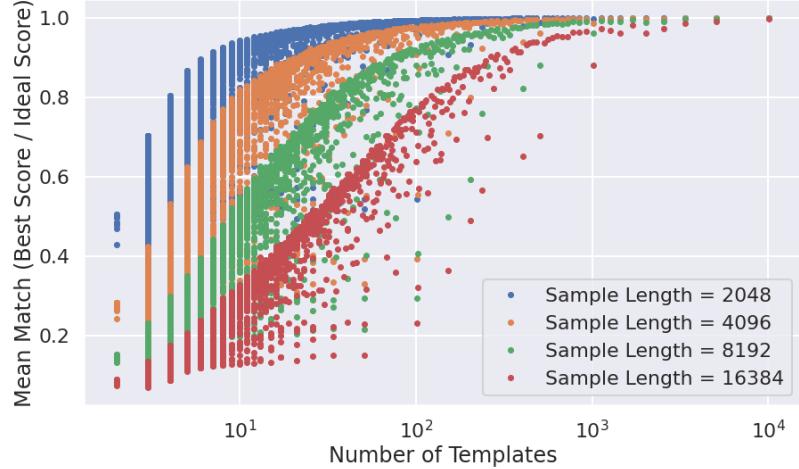


Figure 4.18: The mean match of the dense template grid shown in Figure 4.17 for different numbers of templates. Grids of different sizes were obtained by decimating a dense grid of templates and the average match for each grid was computed using the same set of randomly distributed test signals. Plotting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

²¹⁹⁶ could be used at the detection stage. Therefore, accurately modeling the effects of match
²¹⁹⁷ is key to correct sensitivity calculations.

²¹⁹⁸ The effect of match on the detection efficiency of the matched filter template bank can
²¹⁹⁹ be summarized using the ROC curve (see Figure 4.19). A single ROC curve is obtained
²²⁰⁰ by averaging over the PDFs that describe the detection probabilities of each individual
²²⁰¹ template. The matched filter score for a template follows a Rician distribution with a
²²⁰² mean value equal to the matched filter score multiplied by the match ratio between the
²²⁰³ template and signal. Therefore, the distribution that describes the average matched filter
²²⁰⁴ score when there is a signal in the data is obtained by averaging over the distributions
²²⁰⁵ for every template, whose expectation values are multiplied by the average match ratio.

²²⁰⁶ The distribution of the matched filter score when there is no signal in the data follows
²²⁰⁷ a Rayleigh distribution. Therefore, a trials penalty, which is the statistical penalty one
²²⁰⁸ pays for randomly checking many templates in order to avoid a random match between
²²⁰⁹ noise and a template, is included by computing the joint distribution of N_{template} Rayleigh
²²¹⁰ distributions, where N_{template} is the size of the template bank. For more information on
²²¹¹ the calculation of matched filter template bank ROC curves please refer to Section 4.4.

²²¹² An alternative way to visualize the detection performance for each algorithm is to

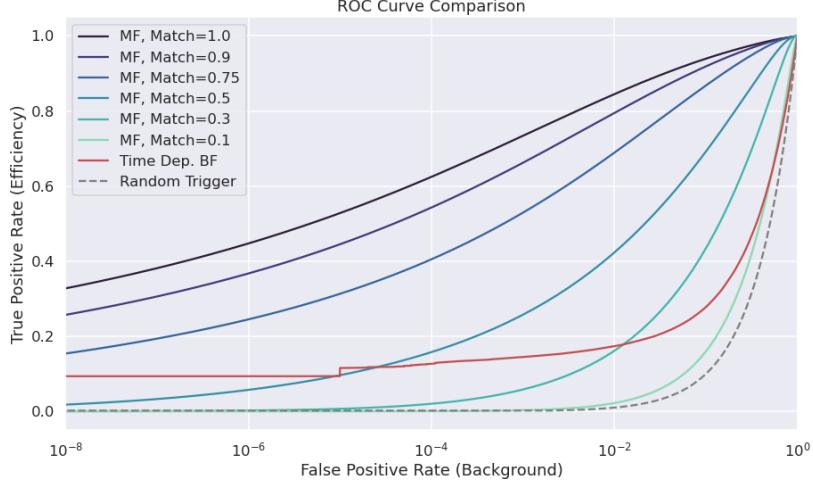


Figure 4.19: Matched filter template bank ROC curves as a function of mean match. One can see that for low match a matched filter is on average worse than the more straight forward beamforming detection approach.

2213 specify a minimum acceptable false positive rate at the trigger level. This is equivalent
 2214 to specifying a minimum threshold on the value of the matched filter score or the size of
 2215 a frequency peak for a beamforming power threshold trigger. One can then draw regions
 2216 of detectable signals as a function of the electron's pitch angle and radial position (see
 2217 Figure 4.20). A kinetic energy shift is equivalent to an overall frequency shift of the
 2218 signal and should have no effect on the detection probability assuming sufficient density
 2219 of matched filter templates in the energy dimension. A electron is declared "detectable"
 2220 for the regions in Figure 4.20 if the signal has at least 50% probability of falling above the
 2221 decision threshold of the respective classifier. One can see that the parameter space of
 2222 detectable signals is greatly expanded beyond the beamforming power threshold trigger
 2223 with a matched filter (MF) or deep neural network (DNN) (see Section 4.3.3). Plots such
 2224 as Figure 4.20 are useful for visualization, but, since the handling of detection likelihood
 2225 is not sufficiently rigorous, the detection probability boundaries are not well-suited to
 2226 sensitivity estimates.

2227 Optimized Matched Filtering Implementation for the FSCD

2228 The biggest practical obstacle to the implementation of a matched filter template bank
 2229 detection approach is oftentimes the computational cost associated with exhaustively
 2230 calculating the matched filter scores of the template bank, and the FSCD is no exception
 2231 in this regard. At a basic level computing a matched filter score requires the convolution

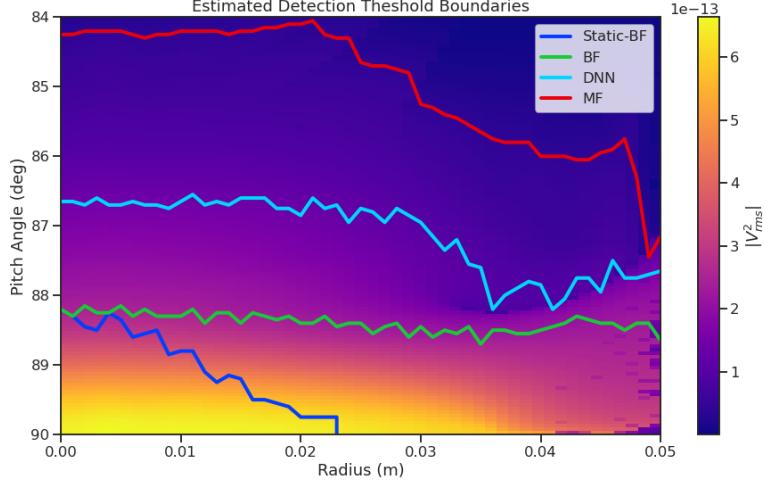


Figure 4.20: Boundaries of detectable electrons in pitch angle kinetic energy space for a series of different signal detection algorithms. A detectable signal is defined as a signal that is above a consistent decision with at least 50% probability. This non-rigorous treatment of detection probability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

2246 reduced by exploiting the convolution theorem and FFT algorithms. If we restrict
 2247 ourselves to signals and templates that contain equal numbers of samples then the
 2248 convolution can be calculated by Fourier transforming both vectors, performing the
 2249 point-wise multiplication, and then performing the inverse Fourier transform to obtain
 2250 the convolution result. The FFT algorithm is able to compute the Fourier transform
 2251 utilizing only $O(N \log N)$ operations compared to $O(N^2)$ for a naive Fourier transform
 2252 implementation. This optimization results in a computational cost per-template of

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

2253 A typical signal vector in the FSCD contains $O(10^4)$ samples in which case the FFT
 2254 reduces the computational cost of the matched filter by a factor of $O(10^3)$. This large
 2255 reduction in computational cost implies that a direct implementation of a matched filter
 2256 is completely infeasible in the FSCD due to resource constraints.

2257 Rather than relying solely on the matched filter it is tempting to consider using
 2258 digital beamforming as an initial step in the signal reconstruction for the purposes of
 2259 data reduction. The primary motivation is to reduce the dimensionality of the data by
 2260 a factor of N_{ch} by combining the array outputs coherently into a single channel. One
 2261 can view the beamforming operation as a partial matched filter, in the sense that the
 2262 matched filter convolution contains the beamforming phased summation along with a
 2263 prediction of the signal shape. By separating beamforming from the signal shape one
 2264 hopes to reduce the overall computational cost by effectively shrinking the number of
 2265 templates and reducing the number of operations required to check each one.

2266 The nature of this optimization requires that we account for the number of templates
 2267 used for pure matched filtering versus the hybrid approach. To first order, the total
 2268 number of templates at the trigger stage is a product of the number of guesses for each
 2269 of the electron's parameters

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

2270 where N_E is the number of kinetic energies, N_θ is the number of pitch angles, N_r is the
 2271 number of starting radial positions, and N_φ is the number of starting azimuthal positions.
 2272 The starting axial position and cyclotron motion phase are not necessary to include in
 2273 the template bank since these parameters manifest themselves as the starting phase of
 2274 the signal, which is effectively marginalized when using a FFT to compute the matched
 2275 filter convolution. Therefore, the total number of operations required by a matched filter

2276 to detect a signal in a segment of array data is on the order of

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

2277 With the hybrid approach we attempt to remove the spatial parameters from the
2278 template bank by using beamforming to combine the array signals into a single channel.
2279 Beamforming explicitly assumes a starting position, which allows us to only use matched
2280 filter templates that span the two-dimensional space of kinetic energy and pitch angle.
2281 The total computational cost of the hybrid method is directly proportional to the number
2282 of beamforming positions. For the time-dependent beamforming defined in Section 4.3.1,
2283 the number of beamforming positions is given by

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

2284 where N_r and N_φ are the same spatial parameters encountered in the pure matched
2285 filter template bank and $N_{\omega_{\nabla B}}$ is the number of ∇B -drift frequency assumptions. If a
2286 unique drift frequency is used for each pitch angle then the hybrid approach is effectively
2287 equivalent to a pure matched filter in the number of operations. The key efficiency gain
2288 of the hybrid approach is to exploit the relatively small differences in $\omega_{\nabla B}$ for electrons
2289 of different pitch angles by using only a small number of average drift frequencies.

2290 The total number of operations for the hybrid approach can be expressed as a sum of
2291 the operations required by the beamforming and matched filtering steps,

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

2292 The first product in the sum is the number of operations required by beamforming,
2293 which is simply the number of beamforming points times the computational cost of the
2294 beamforming matrix multiplication, and the second product is the computational cost
2295 of matched filtering the summed signal generated by each beamforming position. To
2296 compare this to pure matched filtering we take the ratio of Equations 4.34 and 4.36 to
2297 obtain

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

2298 This expression can be simplified by observing that $O(N_E N_\theta) \times O(\log N_s) \gg O(N_{ch})$,

2299 which means that the ratio of computational cost for the two methods can be reduced to

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

2300 If we limit ourselves to a number of estimated drift frequencies of $O(1)$ then we see that
2301 the estimated computational cost reduction of the hybrid approach is of $O(N_{\text{ch}})$. This is
2302 quite a large reduction considering that the FSCD antenna array contains sixty antennas
2303 in the baseline design.

2304 The main drawback of the hybrid approach is that the limited number of allowed
2305 drift frequency guesses can lead to detection efficiency loss due to phase mismatch. The
2306 degree of phase error from an incorrect drift frequency is proportional to the length of
2307 the array data vector used by the signal detection algorithm. For signals with lengths
2308 equal to the baseline FSCD Fourier analysis window of 8192 samples, typical phase errors
2309 from using an average versus the exact ∇B -drift frequency are on the order of a few
2310 percent in terms of the signal energy. This has a relatively small impact on the overall
2311 detection efficiency, however, future experiments with antenna array CRES will want to
2312 balance optimizations such as these during the design phase to keep experiment costs to
2313 a minimum while still achieving scientific goals.

2314 Kinetic Energy and Pitch Angle Degeneracy

2315 More accurate modeling of a matched filter requires that we consider the effects of
2316 mismatched signals and template, since this more accurately reflects the real-world usage
2317 of a matched filter where many incorrect templates are convolved with the data until the
2318 matching template is found. One way to study this is to use the grid of simulated signals
2319 to compute the matched filter scores between mismatched signals and templates and
2320 evaluate the matched filter scores under this scenario. What one finds when performing
2321 this analysis is that templates for kinetic energies and pitch angles that do not match
2322 the underlying signal can have matched filter scores that are indistinguishable from the
2323 matched filter score of the correct template (see Figure 4.21 and Figure 4.21).

2324 This degeneracy in matched filter score is the result of correlations between the kinetic
2325 energy of the electron and the pitch angle caused by changes in the average magnetic field
2326 experienced by an electron for different pitch angles. While in principle helpful for the
2327 purposes of signal detection these correlations are unacceptable since they greatly reduce
2328 the energy resolution of the experiment by causing electrons with specific kinetic energy
2329 to templates across a wide range of energies. It is important to emphasize that this

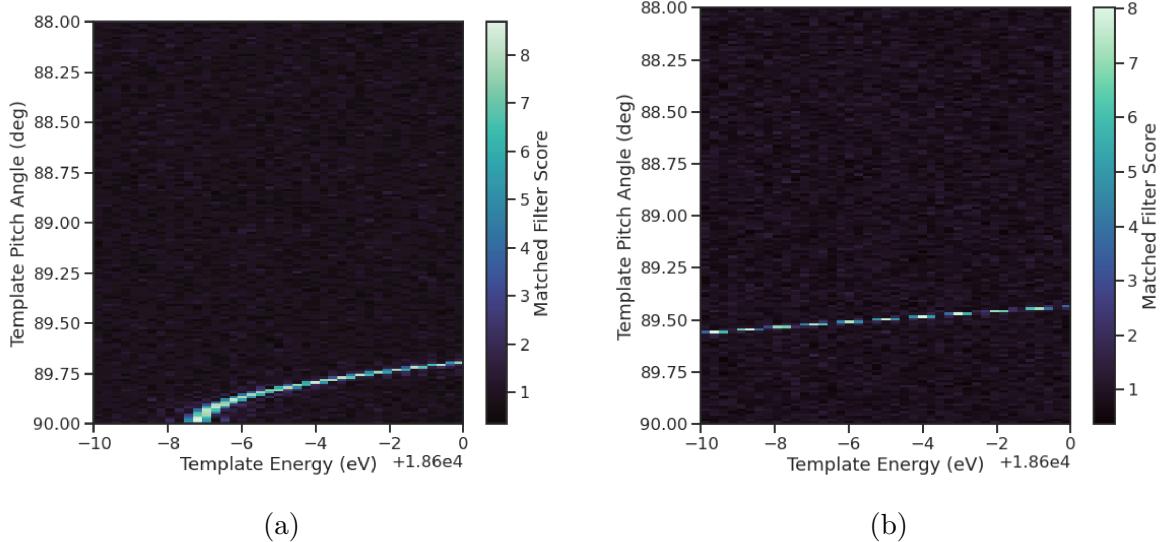


Figure 4.21: Two example illustrations of the correlation between kinetic energy and pitch angle imparted by the shape of the FSCD magnetic trap. The correlations manifest themselves as degeneracies in the matched filter score where multiple matched filter templates have the same matched filter for a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

2330 degeneracy cannot be fixed by implementing a different signal reconstruction algorithm.
 2331 As revealed by the matched filter scores the shapes of the signals for different parameters
 2332 are identical. Resolving this degeneracy between pitch angle and energy requires the
 2333 design of a new magnetic trap with steeper walls so that the average magnetic field
 2334 experienced by an electron is less dependent on pitch angle.

2335 **4.3.3 Machine Learning**

2336 Machine learning is a vast and rapidly developing field of research [49]. In this Section
 2337 we shall attempt to provided a brief introduction to some of the concepts and techniques
 2338 of machine learning that were applied to CRES signal detection rather than attempt a
 2339 comprehensive overview.

2340 **Introduction to Machine Learning**

2341 Digitization of the FSCD antenna array generates large amounts of data that must be
 2342 rapidly processed to enable real-time signal detection and reconstruction. While digital
 2343 beamforming combined with a power threshold is relatively computationally inexpensive,

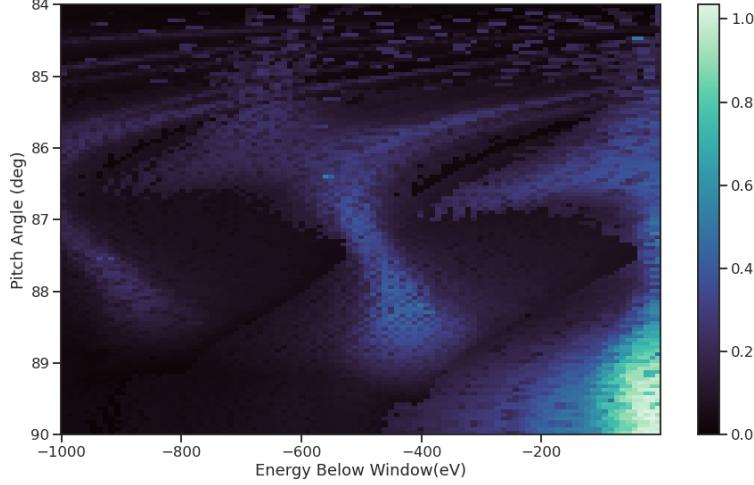


Figure 4.22: A visualization of the correlation between energy and pitch angle in the FSCD magnetic trap. The image is formed by computing the match of the best template from a grid consisting of pitch angles from 84 to 90 degrees in steps of 0.05 degrees, kinetic energies from 17574 to 18574 eV, located 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 [50] and recent developments have shown success in the detection of gravitational wave signals using machine learning techniques [51, 52] 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 [49]. 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

2359 in the case of regression problems.

2360 In supervised learning the model is trained to map from the data to the correct label
2361 by evaluating the output of the model using a training dataset consisting of a set of
2362 paired data and labels. To evaluate the difference between the model output and the
2363 correct label a loss function is used to quantify the error between the model prediction
2364 and the ground truth. For example, a common loss function in regression problems is the
2365 squared error loss function, which quantifies error using the squared difference between
2366 the model output and label.

2367 Using the outputs of the loss function the next step in supervised learning is to
2368 compute the gradient of error with respect to the model parameters in a process called
2369 backpropagation. Using the model parameter gradients the last step in the supervised
2370 learning process is to perform an update of the parameter values in order to minimize
2371 the error in the model predictions across the whole dataset. This loop is performed many
2372 times while randomly shuffling the dataset until the error converges to a minimum value
2373 at which point the training procedure has finished. It is standard practice to monitor
2374 the training procedure by evaluating the performance of the model using a separate
2375 validation dataset that matches the statistical distribution of the training data and to
2376 check the performance of the model after training using yet another dataset called the
2377 test dataset. These practices help to guard against overtraining which is a concern for
2378 models with many parameters.

2379 **Convolutional Neural Networks**

2380 A popular class of machine learning models are neural networks. A neural network is
2381 essentially a function composed of a series of linear operations called layers which take a
2382 piece of data typically represented as a matrix, multiplies the elements of the data by a
2383 weight, and then sums these products to produce an output matrix. Neural networks
2384 composed of purely linear operations are unable to model complex non-linear behavior,
2385 therefore, non-linear activation functions are applied to the outputs of each of the layers
2386 to increase the ability of the neural network to model complex relationships between the
2387 data.

2388 Neural networks are typically composed of at least three layers, but with the present
2389 capabilities of computer hardware they more often contain many more than this. The
2390 first layer in a neural network is called the input layer, because it takes the data objects
2391 as input, and the last layer in a neural network is known as the output layer. The
2392 output layer is trained by machine learning to map the data to a desired output using

2393 the supervised learning procedure described in Section 4.3.3. In between the input and
2394 the output layer are typically several hidden layers that receive inputs from and transmit
2395 outputs to other layers in the neural network model. The term deep neural network
2396 (DNN) refers to those neural networks that have at least one hidden layer, which have
2397 proven to be extremely powerful tools for pattern recognition and function approximation.

2398 An important type of DNN are convolutional neural networks (CNN) that typically
2399 contain several layers which perform a convolution of the input with a set of filters. These
2400 convolution operations are typically accompanied by layers that attempt to down-sample
2401 the data along with the standard neural network activation functions. A standard CNN
2402 is composed of several convolutional layers at the beginning of the network and ends
2403 with a series of fully-connected neural network layers at the output. Intuitively, one
2404 can imagine that the convolutional layers are extracting features from the data that
2405 fully-connected layers use to perform the classification or regression task.

2406 **Deep Filtering for Signal Detection in the FSCD**

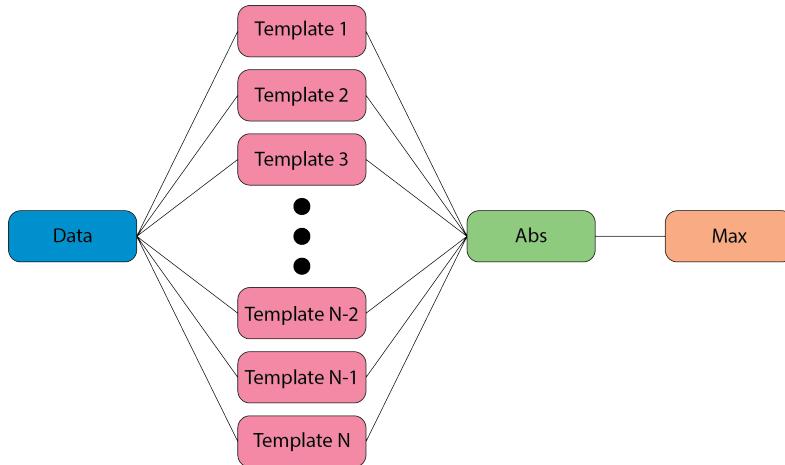


Figure 4.23: A representation of a matched filter template bank as a convolutional neural network. The network has a single layer composed of the templates, which act as convolutional filters. The activation of the neural network is an absolute value followed by a max operator.

2407 CNNs have been extremely influential in the field of computer vision, particularly tasks
2408 such as image segmentation and classification, but have also been applied in numerous
2409 experimental physics contexts. Given the particular challenge posed by signal detection
2410 and reconstruction in the FSCD we are interested in exploring the potential of machine
2411 learning as an effective algorithm for real-time signal detection, since this application

2412 requires both high efficiency and fast evaluation.

2413 In the machine learning paradigm signal detection is equivalent to a binary classifi-
2414 cation problem between the signal and noise data classes, and my investigation focuses
2415 specifically on the application of CNNs to signal detection in the FSCD, which is moti-
2416 vated by relatively recent demonstrations of CNNs achieving classification accuracies for
2417 gravitational wave time-series signals comparable to a matched filter template bank. In
2418 this framework it is possible to interpret the matched filter as a type of CNN composed
2419 of a single convolutional layer with the templates making up the layer filters (see Figure
2420 4.23). Since this neural network has no hidden layers, it is not a DNN like we have
2421 been discussing so far, but we can attempt to construct a proper CNN that attempts to
2422 reproduce the classification performance of the matched filter network.

2423 The name deep filtering refers to this scheme of replacing a matched filter template
2424 bank with a DNN. The reason why one might want to do this is that it may be possible to
2425 exploit redundancies and correlations between templates that may allow one to perform
2426 signal detection with similar accuracy but with fewer computations, which is important
2427 for real-time detection scenarios like the FSCD experiment. In Section 4.4 we perform a
2428 detailed comparison of the signal detection performance of a CNN to beamforming and a
2429 matched filter template bank.

2430 Deep filtering is conceptually a simple technique. Similar to a matched filter template
2431 bank a large number of simulated CRES signals are generated and used to train a model
2432 to distinguish between signal and noise data (see Figure 4.24). In order to reduce the
2433 dimensionality of the input FSCD data a digital beamforming summation is applied
2434 to the raw time-series data generated by Locust to compress the 60-channel data to a
2435 single time-series. CRES signal have a sparse frequency representation and experiments
2436 training CNN's on time-series and frequency series data found that models trained on
2437 frequency spectrum data performed significantly better, therefore, an FFT is applied to
2438 the summed time-series before being normalized and fed to the classification model.

2439 The data used to train the model consists of an equal proportion of signal and noise
2440 frequency spectra. Unique samples of WGN are generated and added to the signals during
2441 training time to avoid have to pre-generate and store large samples of noise data. The
2442 binary cross-entropy loss function combined with the ADAM optimizer proved effective
2443 at training the models to classify CRES data. A simple hyperparameter optimization
2444 was performed by manually tuning model, loss function, and optimizer parameters. The
2445 model and training loops was implemented in python using the PyTorch deep learning
2446 framework. Standard machine learning best practices were followed when training the

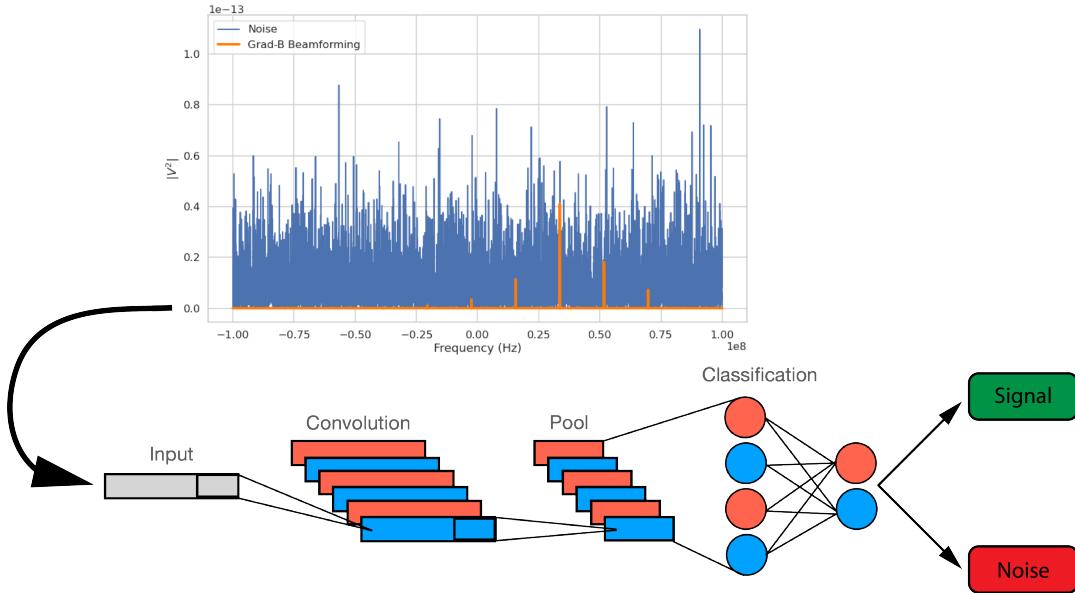


Figure 4.24: A graphical depiction of CRES signal detection using a CNN. A noisy segment of data is converted to a frequency series using digital beamforming and a FFT. The complex-valued frequency series is input into a trained CNN model that classifies the data as signal or noise using a decision threshold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

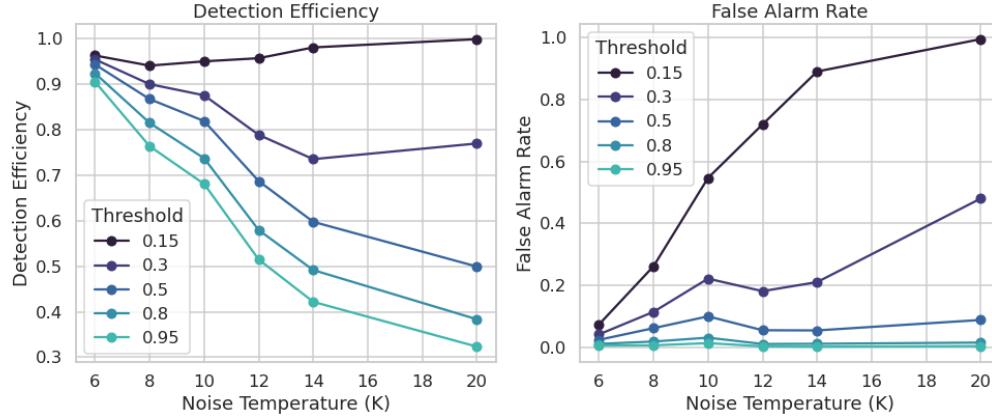


Figure 4.25: The detection efficiency and false alarm rate (false positive rate) as a function of the decision threshold for different values of the noise temperature. The model is trained to output a value close to one for data that contains a signal and outputs a value near zero when the data contains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

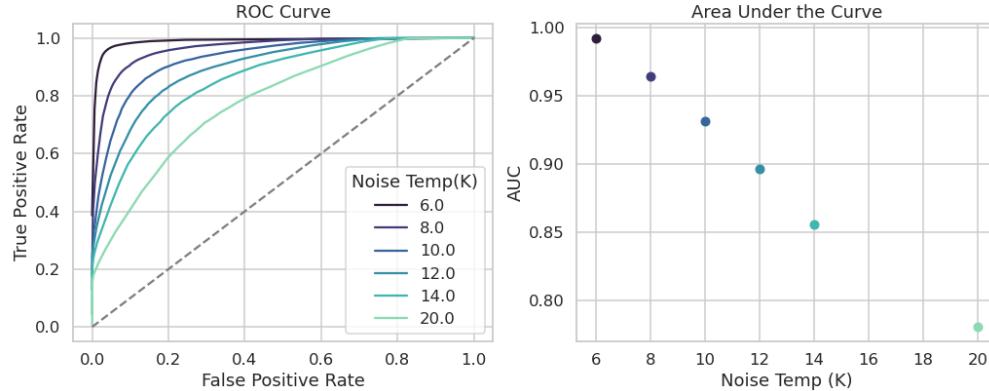


Figure 4.26: ROC curves for a CNN model classifying CRES signals. One can see that the area under the curve, which is a figure of merit that describes the performance of the classifier, is roughly linearly dependent with the noise temperature.

2462 population of simulated CRES signals generated with Locust. The focus of the paper is
 2463 on the performance of the signal detection algorithms for pitch angles below 88.5° where
 2464 the beamforming power threshold begins to fail.

2465 4.4.1 Introduction

2466 Cyclotron Radiation Emission Spectroscopy (CRES) is a technique for measuring the
 2467 kinetic energies of charged particles by observing the frequency of the cyclotron radiation

2468 that is emitted as they travel through a magnetic field [15]. The Project 8 Collaboration
2469 is developing the CRES technique as a next-generation approach to tritium beta-decay
2470 endpoint spectroscopy for neutrino mass measurement. Recently, Project 8 has used
2471 CRES to perform the first ever tritium beta-decay energy spectrum and neutrino mass
2472 measurement [17, 18].

2473 Previous CRES measurements have utilized relatively small volumes of gas that are
2474 directly integrated with a waveguide transmission line, which transmits the cyclotron
2475 radiation emitted by the trapped electrons to a cryogenic amplifier. While this technology
2476 has had demonstrable success, it is not a feasible option for scaling up to significantly
2477 larger measurement volumes. In particular, the goal of the Project 8 Collaboration
2478 is to use CRES combined with atomic tritium to measure the neutrino mass with a
2479 40 meV sensitivity. Achieving this sensitivity goal will require a multi-cubic-meter scale
2480 measurement volume in order to obtain the required event statistics in the tritium
2481 beta-spectrum endpoint region; hence, there is a need for new techniques to enable large
2482 volume CRES measurements for future experiments.

2483 One approach is to surround a large volume with an array of antennas that together
2484 collect the cyclotron radiation emitted by trapped electrons [16, 53]. A promising
2485 array design is an inward-facing uniform cylindrical array that surrounds the tritium
2486 containment volume. Increasing the size of the antenna array, by adding additional
2487 rings of antennas along vertical axis, allows one to grow the experimental volume until a
2488 sufficient amount of tritium gas can be observed by the array. A challenging aspect of
2489 this approach is that the total radiated power emitted by an electron near the tritium
2490 spectrum endpoint is on the order of 1 fW or less, which is then distributed between
2491 all the antennas in the array. Consequently, detecting the presence of a CRES signal
2492 and determining the electron’s kinetic energy requires reconstructing the entire antenna
2493 array output over the course of the CRES event, posing a significant data acquisition
2494 and signal reconstruction challenge.

2495 Project 8 has developed a triggering system to enable real-time identification of CRES
2496 events using an antenna array [54]. Previous measurements with the CRES technique
2497 have utilized a threshold on the frequency spectrum formed from a segment of CRES
2498 time-series data. This algorithm relies on the detection of a frequency peak above the
2499 thermal noise background, which limits the kinematic parameter space of detectable
2500 electrons. Due to the limitations of this power threshold, Project 8 has been investigating
2501 alternative signal identification approaches, including both matched filtering and machine
2502 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 [16, 53], 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 [19]. 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

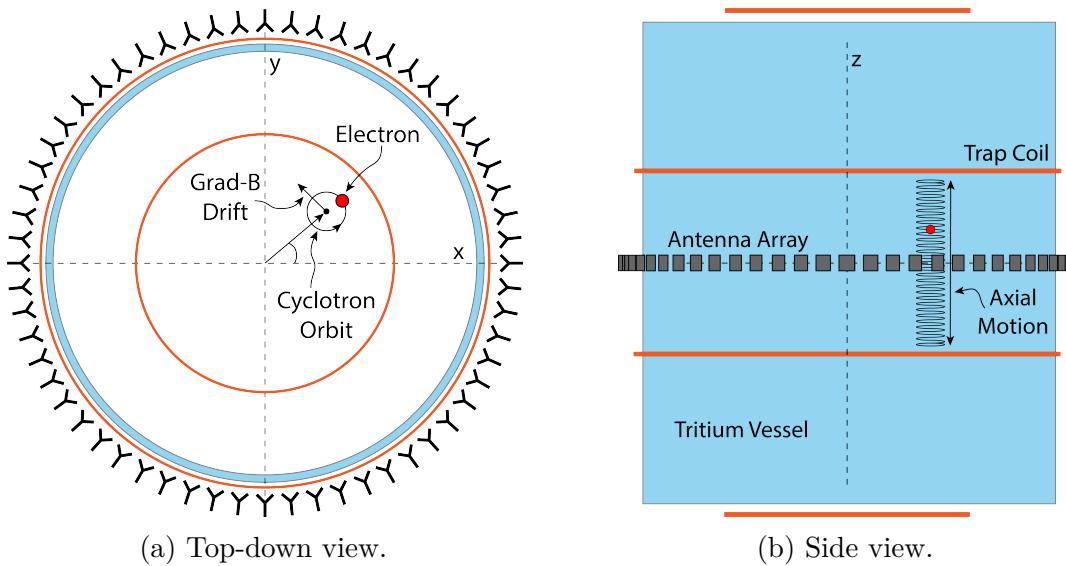


Figure 4.27: An illustration of the conceptual design for an antenna array CRES tritium beta-decay spectrum measurement. The antenna array geometry consists of a 20 cm interior diameter with 60 independent antenna channels arranged evenly around the circumference. The nominal antenna design is sensitive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

determined by the size of the background magnetic field. The FSCD design assumes a background magnetic field value of approximately 0.96 T, which results in cyclotron frequencies for electrons with kinetic energies near the tritium beta-spectrum endpoint from 25 to 26 GHz. The component with the next highest frequency is the axial oscillation experienced by electrons with pitch angles of less than 90° [36]. The flat region of the FSCD magnetic trap extends approximately 3 cm above and below the antenna array plane causing electrons to move back and forth as they are reflected from the trap walls. Typical oscillation frequencies are on the order of 10's of MHz, which results in an oscillation period that is $O(10^3)$ smaller than the duration of a typical CREES event. Therefore, when reconstructing CREES events we treat the electron as occupying only an average axial position at the center of the magnetic trap, since we are not able to resolve the axial position as a function of time. The component of motion with the smallest frequency is ∇B -drift caused by radial field gradients in the trap, producing an orbit of the electron around the central axis of the trap with a frequency on the order of a few

2550 kHz, dependent on the pitch angle and the radial position of the electron.

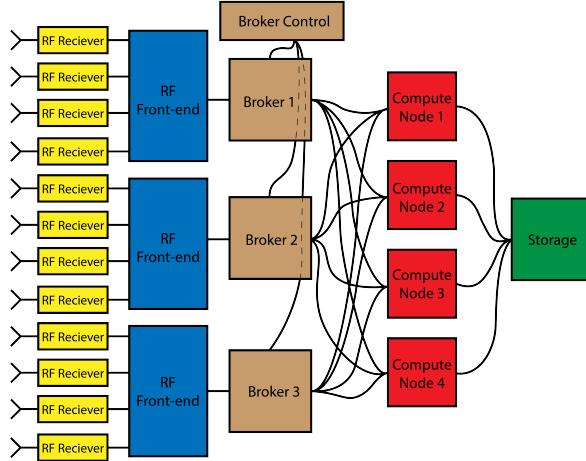


Figure 4.28: A high-level diagram of the DAQ system archctecture envisioned for the FSCD.

2551 The data acquisition (DAQ) system digitizes the signals from the antenna array and
2552 combines thee data streams into a time-ordered matrix of array snapshots that can be
2553 used by the reconstruction algorithms. The FSCD DAQ system design [54] is divided into
2554 three layers 4.28. The first layer is the RF front-end, which includes the antenna array,
2555 the RF receiver boards, and the digitization electronics. The receiver board contains an
2556 amplifier, RF mixer, and bandpass filter to enable down-conversion, and is followed by
2557 the digitization electronics that sample the CRES signals at 200 MHz. In order to achieve
2558 an adequate signal-to-noise ratio to detect CRES events, the DAQ system for the antenna
2559 array demonstrator must have a total system noise temperature of ≈ 10 K, which we
2560 can achieve by using low-noise amplifiers and operating at cryogenic temperatures. After
2561 digitization, the array data must be reorganized from individual data streams sorted
2562 by channel into array snapshots sorted by time. In order to solve this data transfer
2563 and networking problem the second layer of the DAQ system consists of a set of broker
2564 computer nodes that reorganize the array data into time-ordered chunks. This approach
2565 allows us accommodate different data transfer requirements by scaling the number of
2566 broker nodes in this layer accordingly. Next, the broker layer distributes these chunks
2567 of array data to the final layer of the DAQ system, which consists of a set of identical
2568 reconstruction nodes that perform the calculations required for CRES reconstruction.
2569 Similar to the broker layer, the number of reconstruction nodes can be increased or
2570 decreased depending on the amount of computer power required for real-time CRES
2571 reconstruction.

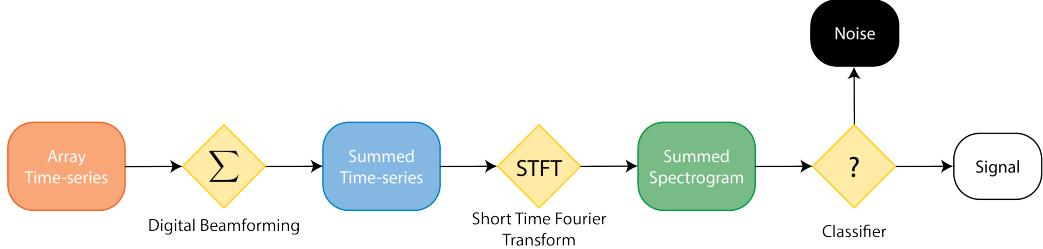


Figure 4.29: A block diagram illustration of the real-time triggering algorithm proposed for antenna array CRES reconstruction.

2572 The design of the FSCD DAQ system is intended to enable a significant portion of
 2573 the CRES event reconstruction to occur in real-time. The motivation for this comes from
 2574 the fact that the FSCD antenna array generates approximately 1 exabyte of raw data
 2575 per year of operation. Therefore, in order to reduce the data-storage requirements, it is
 2576 ideal to perform at least some of the CRES event reconstruction in real-time so that it
 2577 is possible to save a reduced form of the data for offline analysis. The first step of the
 2578 real-time reconstruction would be a real-time signal detection algorithm, which is the
 2579 focus of this paper. Our approach consists of three main operations performed on the
 2580 time-series data blocks including digital beamforming, a short time Fourier transform
 2581 (STFT), and a binary classification algorithm to distinguish between signal and noise
 2582 data (see Figure 4.29).

2583 4.4.2.2 Real-time Signal Detection

2584 The first step in the real-time detection algorithm is digital beamforming, which is a
 2585 phased summation of the signals received by individual antennas in the array (see Figure
 2586 5.21). The phase shifts correspond to the path length differences between a spatial
 2587 position and each individual antenna such that, when there is an electron located at
 2588 the beamforming position, all the signals received by the array constructively interfere.
 2589 Since we do not know ahead of time where an electron will be produced in the detector,
 2590 we define a grid of beamforming positions that cover the entire region where electrons
 2591 can be trapped and perform a phased summation for each of these points for every
 2592 time-step in the array data block. As we saw in Section 4.4.2.1, the axial oscillation
 2593 of the electrons prevents us from resolving its position along the Z-axis of the trap,
 2594 therefore our beamforming grid need only cover the possible positions of the electron in
 2595 the two-dimensional plane defined by the antenna array.

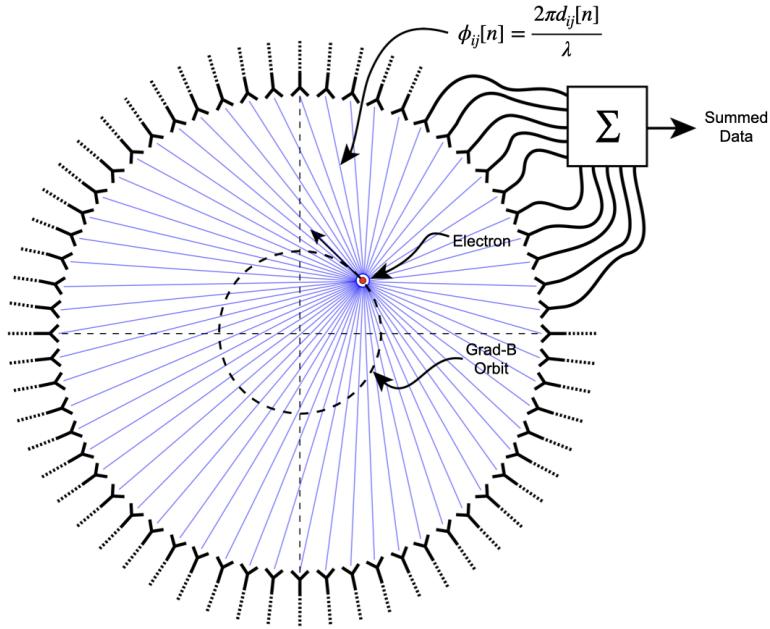


Figure 4.30: An illustration of the digital beamforming procedure. The blue lines indicate the various distances from the beamforming position to the antenna. In the situation depicted the actual position of the electron matches the beamforming position, so we should expect constructive interference 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.

2596

The equation defining digital beamforming can be expressed as

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

2597 where $\mathbf{x}[n]$ is array snapshot vector at the sampled time n , $\Phi[n]$ is the matrix of
 2598 beamforming phase shifts, and $\mathbf{y}[n]$ is summed output vector that contains the voltages
 2599 for each of the summed channels that correspond to a particular beamforming position.
 2600 The elements of the beamforming phase shift matrix can be expressed as a weighted
 2601 complex exponential

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

2602 where the indices i and j label the beamforming and antenna positions respectively. The
 2603 weight A_{ij} accounts for the relative power increase for antennas that are closer to the
 2604 position of the electron, and ϕ_{ij} is the total beamforming phase shift for the j -th antenna
 2605 at the i -th beamforming position.

2606 The beamforming phase shift is a sum of two terms

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

2607 where the first term is the phase shift originating from the path length difference ($d_{ij}[n]$)
2608 between the beamforming and antenna positions, which are represented by the vectors
2609 (r_j, θ_j) and ($r_i, \theta_i[n]$), and the second term is the angular separation ($\theta_{ij}[n]$) of the two
2610 positions. The angular separation enters into the beamforming phase due to an effect
2611 caused by the circular orbit of the electron that produces radiation whose phase is linearly
2612 dependent on the relative azimuthal position of the antenna [55,56]. The time-dependence
2613 of the beamforming phases is intended to correct for the effects of ∇B -drift, which cause
2614 the guiding centers of electrons to orbit the center of the magnetic trap. By including a
2615 linear time-dependence in the azimuthal beamforming position,

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

2616 where $\omega_{\nabla B}$ is the azimuthal grad-B drift frequency, $t[n]$ is the time vector and, $\theta_{i,0}$ is the
2617 starting azimuthal position, we can configure the beamforming phases to effectively track
2618 the XY-position of the guiding center over the event duration. Predicting accurate values
2619 of $\omega_{\nabla B}$ for a specific trap and set of kinematic parameters will be done by simulations,
2620 which are performed using the Kassiopeia software package [35] by Project 8.

2621 After digital beamforming, we apply a short-time Fourier transform (STFT) to the
2622 summed time-series to obtain the frequency spectrum representation of the signals (see
2623 Figure 4.31). From the detection perspective, the frequency representation of the CRES
2624 data is advantageous compared to the time domain, because the frequency spectra of
2625 CRES signals are well-approximated by a frequency and amplitude modulated sinusoidal
2626 whose carrier frequency increases as a linear chirp. The modulation is caused by the axial
2627 oscillation of the electron in the magnetic trap and produce frequency spectra that are
2628 well-described by a small number of frequency components. The linear chirp is caused
2629 by the energy loss due to cyclotron radiation, which results in a relatively slow increase
2630 in the frequency components of the CRES signal over time. During the standard Fourier
2631 analysis window for the FSCD of 40.96 μ sec, we expect a typical CRES signal to increase
2632 in frequency by approximately 15 kHz, which is smaller than the frequency bin width
2633 given the 200 MHz sample rate. Therefore when considering a single frequency spectrum
2634 it is justifiable to neglect the effects of the linear frequency chirp.

2635 In the cases where the electron's pitch angle is $\gtrsim 88.5^\circ$, the majority of the signal

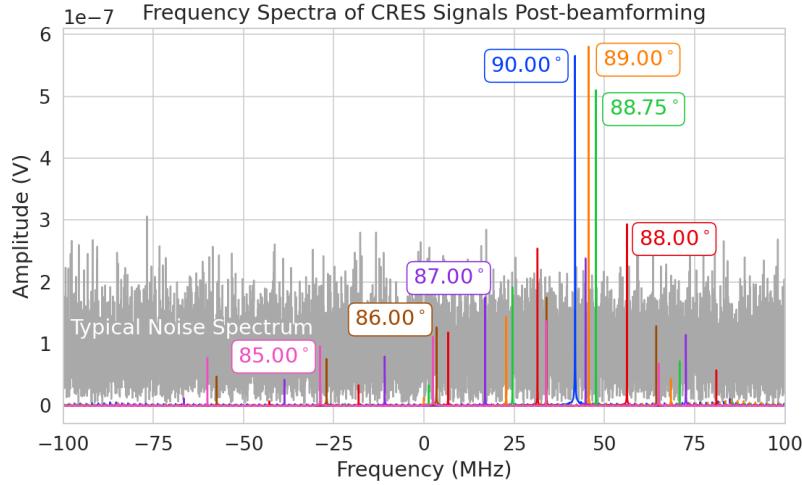


Figure 4.31: Frequency spectra of simulated CRES signals post-beamforming. The signal of a 90° electron consists of a single frequency component that is easy to detect with a power threshold on the frequency spectrum. This power threshold is still effective for signals with relatively large pitch angles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

2653 the FSCD requires that we pose the signal detection problem in a consistent manner.
2654 The approach we take is to perform a binary hypothesis test on the frequency spectra
2655 generated by the STFT. Mathematically, this is expressed as,

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

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

2656 Where under hypothesis \mathcal{H}_0 , the vector representing the frequency spectrum ($y[n]$) is
2657 composed of pure white Gaussian noise (WGN) represented by $\nu[n]$, and under hypothesis
2658 \mathcal{H}_1 the frequency spectrum is composed of a CRES signal ($x[n]$) with added WGN. The
2659 dominant source of noise in a FSCD-like experiment is expected to be thermal Nyquist-
2660 Johnson noise, which is well approximated by a WGN distribution. In order to decide
2661 between these two hypotheses we follow the standard Neyman-Pearson approach by
2662 performing a log-likelihood ratio test between the probability distributions of the signal
2663 classifier output under \mathcal{H}_1 and \mathcal{H}_0 [46]. The output of the log-likelihood ratio test is
2664 called the test statistic, which is used to assign the data as belonging to the noise (\mathcal{H}_0)
2665 or signal (\mathcal{H}_1) classes by setting a decision threshold on the value of the test statistic.

2666 In practice, we select the decision threshold by finding the value of the test statistic
2667 that guarantees an acceptable rate of false positives and then attempt to maximize
2668 the signal detection probability under that fixed false positive rate. Because the signal
2669 classifier will be used to evaluate the spectra of $O(10^2)$ beamforming positions every
2670 $40.96 \mu\text{sec}$, we will require the signal classifiers to operate with decision thresholds that
2671 provide false positive rates significantly smaller than 1%. This reduces the burden placed
2672 on later stages of the CRES reconstruction chain to reject these false positives and
2673 decreases the overall likelihood of reconstructing a false event. Below, we calculate the
2674 probability distributions that allow us characterize how different detection algorithms
2675 will perform for CRES signals in an FSCD experiment.

2676 **4.4.3 Signal Detection Algorithms**

2677 **4.4.3.1 Power Threshold**

2678 The power threshold detection algorithm uses the maximum amplitude of the frequency
2679 spectra as the detection test statistic. To model the performance of this approach,
2680 consider first the case where the signal is pure WGN. For a single bin in the frequency
2681 spectrum, the probability that the amplitude falls below a specific threshold value is

2682 given by the Rayleigh cumulative distribution function (CDF),

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

2683 where the complex amplitude of the frequency bin is x , and τ is the WGN variance.
 2684 Because the noise samples for each frequency bin are independent and identically dis-
 2685 tributed (IID), the probability that every bin in the frequency spectrum falls below the
 2686 threshold is the joint CDF formed by the product of each individual frequency bin CDF,

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

2687 The PDF for the power threshold classifier can then be obtained by differentiating the
 2688 CDF.

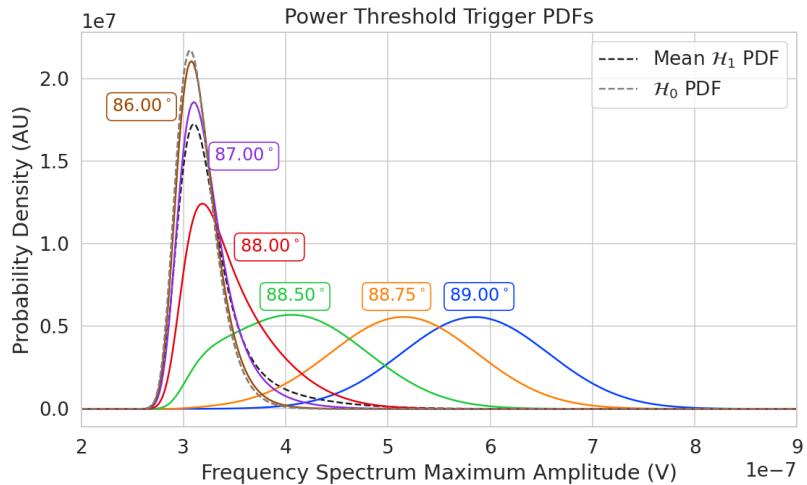


Figure 4.32: PDFs of the power threshold test statistic for CREs signals with various pitch angles as well as the PDF for the noise-only signal case. The average PDF computed for pitch angles ranging from 85.5 to 88.5° is also shown. As the pitch angle is decreased the signal PDF converges towards the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

2689 The probability distribution for the power threshold classifier under \mathcal{H}_1 is formed in
 2690 a similar way, but the frequency bins that contain signal must be treated separately. For
 2691 a frequency bin that contains both signal and noise we can describe the probability that
 2692 the amplitude of the bin will fall below our threshold using the Rician CDF,

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

2693 where the parameter $|\nu|$ defines the noise-free amplitude of the signal and Q_1 is the
 2694 Marcum Q-function. This time the CDF that describes the probability that the entire
 2695 spectrum falls below the decision threshold is the product of both signal and noise CDFs,

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

2696 The first half of Equation 4.48 is the contribution from the bins in the frequency spectrum
 2697 that contain only noise, and the second half is the product of the Rician CDFs for the
 2698 frequency bins that contain signal peaks with a noise-free amplitude of $|\nu_k|$. In Figure
 2699 4.32 we show plots of example PDFs under \mathcal{H}_1 and \mathcal{H}_0 .

2700 4.4.3.2 Matched Filtering

2701 The shape of a CRES signal is completely determined by the initial conditions of the
 2702 electron as it is emitted from beta-decay, which implies that it is possible to apply
 2703 matched filtering as a signal detection algorithm. With a matched filter one uses the
 2704 shape of the known signal, which is called a template, to filter the incoming data by
 2705 computing the convolution between the signal and the data [46]. For cases where the
 2706 signal is buried in WGN, the matched filter is the optimal detector in that it achieves
 2707 the maximum probability of a true detection for a fixed false positive rate. Since CRES
 2708 signals have an unknown shape but are deterministic, we can apply a matched filter by
 2709 using simulations to generate a large number of signal templates called a template bank,
 2710 which spans the parameter space of possible signals. Then at detection time, we use the
 2711 template bank to identify signals by performing the matched filter convolution for each
 2712 template in an exhaustive search.

2713 As we saw from the frequency spectra in Figure 4.31, CRES signals are highly periodic
 2714 in nature. In such cases, it is advantageous to utilize the convolution theorem to replace
 2715 the matched filter convolution with an inner product in the frequency-domain. With the
 2716 convolution theorem, the matched filter test statistic that describes the detection of a
 2717 signal buried in WGN using a matched filter template bank is given by

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

2718 where $h^\dagger[n]$ is the complex conjugate of the signal template. For the case when our
 2719 template bank consists of only a single template it is possible to derive an exact analytical

2720 form for the PDF describing the matched filter test statistic. First, we derive PDF under
 2721 the signal hypothesis, where the equation describing the matched filter test statistic, also
 2722 known as the matched filter score, becomes

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

2723 Each noisy frequency bin represented by $y[n]$ is the sum between value of the signal
 2724 at that bin and complex WGN, which means that $y[n]$ is itself Gaussian distributed.
 2725 Therefore, the value of the inner product between the template and the data is also a
 2726 complex Gaussian variable; and, since the matched filter score is the magnitude of this
 2727 inner product, it must follow a Rician distribution.

2728 We can derive the equation for the Rician PDF by expressing the matched filter
 2729 template \mathbf{h} in terms of the corresponding simulated signal, which we write as \mathbf{x}_h to
 2730 distinguish from the signal in the data. Using the standard normalization and assuming
 2731 uncorrelated WGN, the matched filter templates can be written as

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

2732 where τ is the noise variance. Inserting this into Equation 4.49 and expressing the data
 2733 as a sum between a signal and a WGN vector yields,

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

2734 Next, we transform the expression by isolating the randomly distributed components
 2735 giving

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

2736 The first term of 4.53 can be simplified by using the Cauchy-Schawrz inequality to express
 2737 the magnitude of the inner product in terms of the magnitudes of the signal and template
 2738 as well as an orthogonality constant which we call "match" (Γ). Using this we obtain,

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

2739 The second term is a sum of Gaussian distributed variables, which we should expect also
2740 follows a Gaussian distribution. Each of the samples $\nu[n]$ is described by

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

2741 where $\mathcal{N}(0, \tau)$ is a complex Gaussian distribution with zero mean and variance τ . There-
2742 fore,

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

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

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

2743 We see that \mathcal{T} is magnitude of a complex variable with mean $|\mathbf{h}||\mathbf{x}|\Gamma$ and variance one. In
2744 order to simply the expression a bit further, we define the quantity $\mathcal{T}_{\text{ideal}} = |\mathbf{h}||\mathbf{x}|\Gamma$, which
2745 we call the ideal matched filter score, because it represents the value of the matched
2746 filter inner product that we would expect if no noise was present in the signal. We can
2747 write the matched filter test statistic as the magnitude of a two-dimensional vector in
2748 the complex plane

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

2749 where n_r and n_i are the real and imaginary components of the noise each with variance
2750 $1/2$, which is modeled by a Rician distribution with shape factor $\mathcal{T}_{\text{ideal}}$. Therefore, the
2751 probability distribution of the matched filter test statistic is given by,

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

2752 where I_0 is the zeroth-order modified Bessel function.

2753 The shape of the matched filter score distribution is controlled by the parameter
2754 $\mathcal{T}_{\text{ideal}}$, which is effectively the value of the matched filter score if the data contained no
2755 noise. Without noise, the data vector reduces to the signal, \mathbf{x} , in which case Equation
2756 4.50 becomes the magnitude of an inner product between two vectors. We can write
2757 the magnitude of an inner product in terms of the lengths of the individual vectors and
2758 a constant that describes the degree of orthogonality between them. Applying this to

2759 Equation 4.50, we obtain

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

2760 where Γ describes the orthogonality between \mathbf{h} and \mathbf{x} . From the point of view of matched
2761 filtering, we can interpret Γ as describing how well the template matches the underlying
2762 signal in the data.

2763 The matched filter score PDF under the noise hypothesis can be readily obtained
2764 from Equation 4.60 by setting the value of $\mathcal{T}_{\text{ideal}}$ to zero, since the data contains no signal
2765 in the noise case. Doing this, we obtain the Rayleigh distribution that describes the
2766 matched filter score under \mathcal{H}_0 ,

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

2767 Equations 4.60 and 4.62 describe the behavior of the matched filter test statistic
2768 under \mathcal{H}_0 and \mathcal{H}_1 for a single template. However, defining a PDF that describes the
2769 matched filter test statistic in the case of multiple templates is in general a mathematically
2770 intractable problem, since there is no guarantee of orthogonality between matched filter
2771 templates. This leads to correlations between the matched filter scores of different
2772 templates because only one sample of noise is used to compute the matched filter scores
2773 of the template bank. In order to proceed, we need to make the simplifying assumption
2774 that we can treat the matched filter scores as IID variables, which allows to ignore
2775 correlations between templates. The overall effect of this will be an underestimate of the
2776 performance of the matched filter, since we are under counting the number of templates
2777 that could contribute a detectable score.

2778 For \mathcal{H}_0 we model the probability that the matched filter score falls below our threshold
2779 using the CDF obtained by integrating Equation 4.62. Because we are assuming that
2780 the matched filter scores using different templates are independent, the probability that
2781 the matched filter score for all templates falls below a threshold value is the joint CDF
2782 formed by multiplying the CDF for each template. Under \mathcal{H}_0 this is

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

2783 where x is the matched filter score threshold and N_t is the number of templates. We
2784 should expect that the distribution describing the matched filter template bank maximum
2785 score depends on N_t , because with more templates there is a greater chance of a random
2786 match between the template and data.

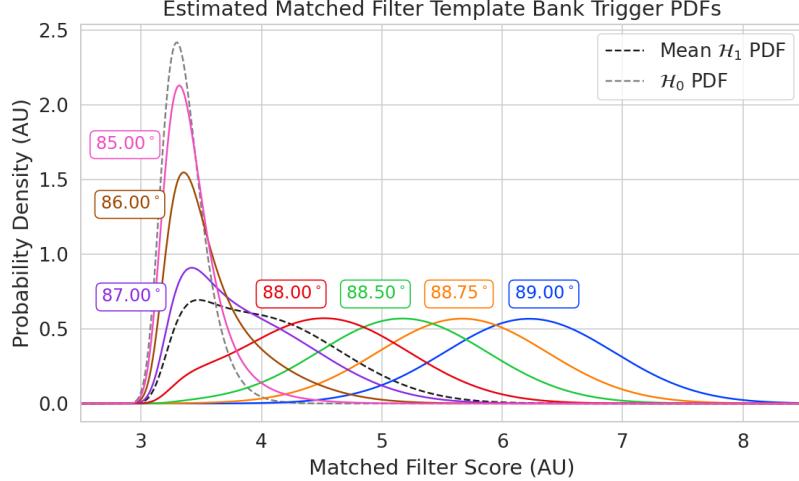


Figure 4.33: Plots of the estimated PDFs for the matched filter template bank test statistic for CRES signals with various pitch angles as well as the estimated PDF for the noise only signal case. We assume an estimated number of templates of 10^5 and perfect match between signal and template i.e. $\Gamma_{\text{best}}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

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

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

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

where \mathbf{y} is the noisy data vector and $G(\mathbf{y}; \boldsymbol{\Omega})$ is the machine learning model parameterized by the weights $\boldsymbol{\Omega}$. By using supervised learning on a labeled set of training signals, we can modify the function parameters to learn the mapping from the data to the likelihood 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 |

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

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

4.4.4 Methods

4.4.4.1 Data Generation

To test the triggering performance of the classifiers, simulated CRES signals were generated using the Locust simulations package [37, 55] developed by the Project 8 collaboration. Locust uses the separately developed Kassiopeia package to calculate the magnetic fields produced by a user defined set of current carrying coils along with any specified background magnetic fields, resulting in a magnetic trap. Next, Kassiopeia calculates the trajectory of an electron in this magnetic field starting from a set of user specified initial conditions. The Locust software then uses the electron trajectories from Kassiopeia to calculate the resulting electromagnetic fields using the Liénard-Wiechert equations, and determine the voltages generated in the antenna array with the antenna transfer function. Locust then simulates the down-conversion, filtering, and digitization steps resulting in the simulated CRES signals for an electron.

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

2852 50000 signals randomly parameterized in energy and pitch angle. This dataset provides
2853 the training and test data for the machine learning approach, and acts as a representative
2854 sample of signals to evaluate the performance of the matched filter template bank.

2855 Each signal was simulated for a duration of $40.96 \mu\text{s}$, which is equivalent to 8192
2856 samples at the FSCD digitization rate, and begins at time $t = 0 \text{ s}$ for all simulations.
2857 This duration represents a single frequency spectrum generated by the STFT. The output
2858 of the Locust simulation is a matrix of array snapshots with size given by the number of
2859 channels times the event length ($N_{\text{ch}} \times N_{\text{sample}}$), which we pre-process using the digital
2860 beamforming summation and STFT described in Section 4.4.2.2. The ∇B -drift correction
2861 uses the exact value of $\omega_{\nabla B}$, obtained from the Kassiopeia simulation of that electron.
2862 In practice, an average value for $\omega_{\nabla B}$ could be used, because there is limited variation in
2863 drift frequency across this parameter space.

2864 4.4.4.2 Template Number and Match Estimation

2865 The estimated PDF for the matched filter template bank depends on the score of the
2866 best matching template or equivalently the match of the best template (Γ_{best}) as well
2867 as the number of templates. One expects that with a higher number of templates the
2868 average value of Γ_{best} will increase, however, there is a point of diminishing returns at
2869 which more templates will not significantly increase match, but will still increase the
2870 likelihood of false positives. Therefore, it is desirable to use the minimum number of
2871 templates that provide an acceptable mean value of Γ_{best} .

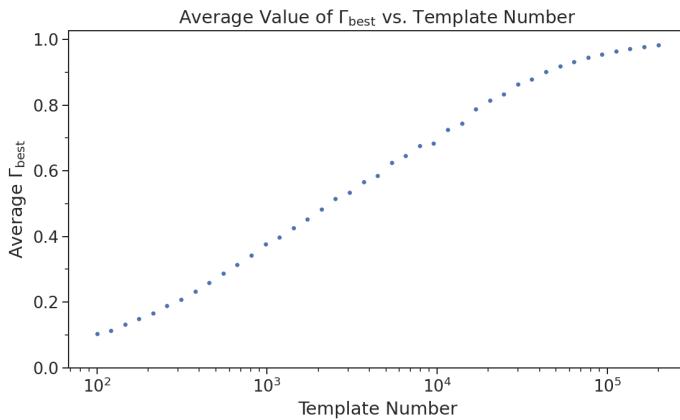


Figure 4.34: The mean match of the matched filter template bank to a test set of randomly parameterized signals as a function of the number or density of templates. The parameter space includes pitch angles from $85.5 - 88.5^\circ$ and energies from 18575 – 18580 eV.

2871

2872 To quantify the relationship between match and template number, we calculated
2873 the mean match of the random dataset to a selection of templates obtained from the
2874 regularly spaced dataset. The results are shown in Figure 4.34, where we find that the
2875 average value of Γ_{best} is an exponential function of the number of templates. From this
2876 plot we select the desired value of mean match at which we would like to evaluate the
2877 matched filter PDF and can infer the required number of templates.

2878 **4.4.4.3 CNN Training and Data Augmentation**

2879 To prepare the data for training the model, we split the random dataset in half to create
2880 distinct training and test datasets. Additionally, a randomly selected 20% of the training
2881 data is isolated for use as a validation set during the training loop. The size of the
2882 training, validation, and test datasets are then tripled by appending two additional copies
2883 of the data to increase the sample size of the dataset after data augmentation. The
2884 data is loaded with no noise, which is added to each data batch during the training
2885 phase by generating a new noise sample from a complex WGN distribution. In order to
2886 ensure an even split between signal and noise data we append to the noise-free signals an
2887 equal number of empty signals composed of all zeros. Therefore, as the data is randomly
2888 shuffled during training, on average an equal number of empty signals will be included
2889 with the training signals. After adding the sample of WGN to the data batch, the empty
2890 signals represent the noise-only data that the model must distinguish from signal data.

2891 As the training signals are loaded we apply a unique random phase shift as the
2892 first form of data augmentation. Since the data is generated using the same initial
2893 axial position and cyclotron orbit phase, the randomization is an attempt to prevent
2894 overtraining on these features. During each training epoch the data is randomly shuffled
2895 and split into batches of 2500 signals. Each batch of signals is then circularly shifted
2896 by a random number of frequency bins to simulate a kinetic energy shift from -75 to
2897 20 eV to simulate a training dataset with a larger energy range. Next, a sample of
2898 complex WGN, consistent with the expected 10 K Nyquist-Johnson noise expected for
2899 the FSCD, is generated and added to the signal, which prevents overtraining on noise
2900 features. As a final step, the data is renormalized by the standard deviation of the noise
2901 so that the range of values in the data is close to $[-1, 1]$, which helps ensure well-behaved
2902 back-propagation.

2903 The Binary Cross-entropy loss function is used to compute the loss for each batch of
2904 data and the model weights are updated using the ADAM optimizer with a learning rate
2905 of 5×10^{-3} . After each training epoch, the loss and classification accuracy of the validation

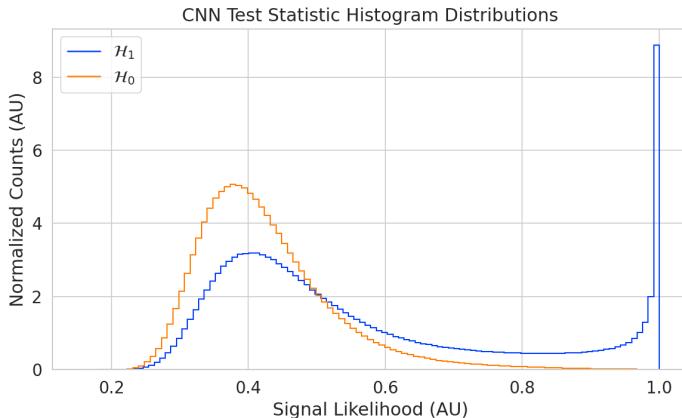


Figure 4.35: Histograms of the trained CNN model output from the test dataset. The blue histogram shows the model outputs for signal data. The oddly shaped peak near the end is the result of the softmax function mapping the long tail of the raw output distribution to the range [0, 1].

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 [57].

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

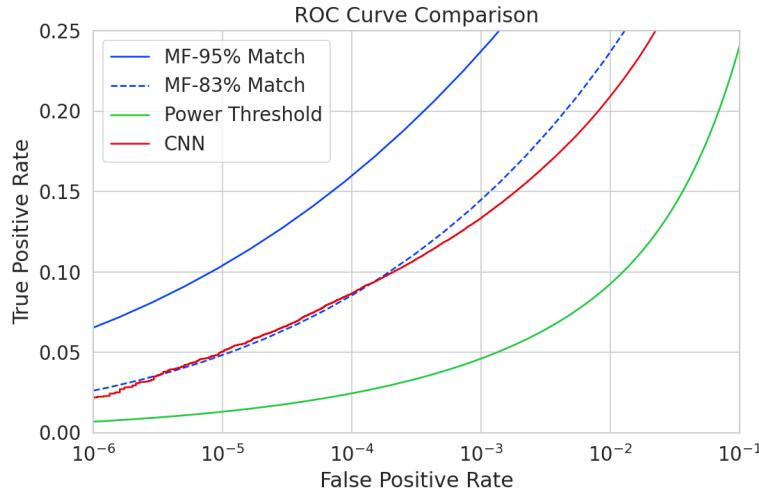


Figure 4.36: ROC curves describing the detection efficiency or true positive rates for the three signal classification algorithms examined in this paper.

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

The true positive rate of a signal classifier is equivalent to its detection efficiency, and we see that for the population of signals with pitch angles $< 88.5^\circ$ the power threshold has a consistently lower detection efficiency than the CNN and the matched filter. This result could have been predicted from the visualization of signal spectra in Figure 4.31, where we see that there is no way to distinguish between a noise peak and a signal peak with high confidence at small pitch angles. The CNN offers a significant and consistent increase in detection efficiency over the power threshold approach, with the relative improvement in detection efficiency increasing as the false positive rate decreases. If we compare the CNN to the matched filter, we see that the performance of the tested network is roughly equivalent to a matched filter detector with an average match of about 83%, which uses approximately 20000 matched filter templates. The overall best 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 [58], dependent on the required floating-point precision of the computation. Therefore, the

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

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

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

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

3015 cost considerations. The low estimated computational cost of the CNN is directly related
3016 to the small network size.

3017 Additional experiments with larger CNNs, generated by increasing the depth and
3018 width of the neural network, and we observed that these changes provided minimal
3019 ($\lesssim 1\%$) improvement in the classification accuracy of the model. A potential reason
3020 for this could be the sparse nature of the signals in the frequency domain and the low
3021 SNR which makes for a challenging dataset to learn from. Future work could investigate
3022 modifications to the neural network architecture such as sparse convolutions, which may
3023 improve the classification accuracy of the model or further reduce the computational
3024 costs of this approach. Alternatively, more complicated CNN architectures such as a
3025 ResNet [59] or VGG model [60] may provide improved classification performance over a
3026 basic CNN. An additional promising area of investigation are recurrent neural networks,
3027 which may be able to exploit the time-ordered features of the STFT for more accurate
3028 signal detection if the electron signals last for multiple Fourier transform windows.

3029 Our estimate of the computational cost of the matched filter is somewhat naive if
3030 we notice that the majority of the values that make up a CRES frequency spectra are
3031 zero (see Figure 4.31). Therefore, the majority of operations in the matched filter inner
3032 product are unnecessary, and we could instead evaluate the matched filter inner product
3033 using only the $\lesssim 10$ frequency peaks that make up CRES signal. This optimization
3034 reduces the number of operations required to check each template by a factor of $O(100)$
3035 to $O(1000)$, which brings the estimated computational cost of the matched filter in
3036 line with the CNN. Although this level of sparsity results in a multiplication with very
3037 low arithmetic complexity, the resulting sparse matched filter algorithm is still likely
3038 to be constrained by memory access speed rather than compute speed. Ultimately, the
3039 comparison of the relative computational and hardware costs between the matched filter
3040 and CNN will depend on the efficiency of the software implementation and hardware
3041 support for neural network and sparse matrix calculations.

3042 **4.4.6 Conclusion**

3043 Increasing the detection efficiency and overall event rate of the CRES technique represents
3044 a key developmental path towards new scientific results and broader applications of the
3045 CRES technique. It is what motivates both the antenna array detection approach and
3046 the development of real-time signal reconstruction algorithms. We have demonstrated
3047 that significant gains in the detection efficiency of the CRES technique are achievable
3048 by utilizing triggering algorithms that account for the specific shape of CRES signals in

3049 the detector. These algorithms emphasize the need for accurate and fast methods for
3050 CRES simulation, since they directly contribute to the success of matched filter methods
3051 by providing a way to generate expected signal templates and also serve as a source of
3052 training data for machine learning approaches.

3053 The improvements in detection efficiency offered by these alternative approaches to
3054 triggering are crucial to the success of efforts to develop scalable technologies for CRES
3055 measurement, since they provide a significant increase in the detectable parameter space
3056 of CRES events, which allows for a better utilization of the larger detection volume.
3057 While we have focused on the real-time detection of CRES signals from antenna arrays,
3058 these same signal classifiers could be used in CRES experiments utilizing a different
3059 detector technologies, since the same principles of signal detection will apply. For example,
3060 previous CRES measurements by the Project 8 collaboration that utilized a waveguide
3061 gas cell, could have improved their detection efficiency by employing a matched filter
3062 or neural network classifier to identify trapped electrons with pitch angles that are too
3063 small to be detected by the power threshold approach. Furthermore, alternative CRES
3064 detector technologies such as resonant cavities [16] could also see similar improvements
3065 in detection efficiency, which is of crucial importance to future efforts by the Project 8
3066 collaboration to utilize CRES to measure the neutrino mass.

Chapter 5

Antenna and Antenna Measurement System Development for the Project 8 Experiment

5.1 Introduction

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

In this chapter we summarize the development of an antenna measurement system at Penn State to implement and test the techniques of antenna array CRES on the bench-top, in order to support the efforts of the Project 8 collaboration. In Section 5.2 we provide an introduction to some fundamental parameters and concepts related to antenna measurements as well as an overview of the Penn State antenna measurement system hardware. In Section 5.3 we include the manuscript of a paper [56] which details the design and characterization of a specialized antenna developed to mimic the electric fields emitted by an electron in a CRES experiment. This antenna, called the Synthetic Cyclotron Antenna (SYNCA), is intended as a calibration tool for antenna arrays developed for CRES measurements. Lastly, in Section 5.4 we summarize a set of prototype FSCD antenna array measurements with the SYNCA [19], which we use to validate the simulated performance of the antenna array and estimate systematic errors associated with the antenna array.

3092 5.2 Antenna Measurements for CRES experiments

3093 5.2.1 Antenna Parameters

3094 Antenna characterization measurements are intended to validate simulations of the
3095 antenna array performance, which ultimately informs the neutrino mass sensitivity of
3096 the experiment. In this section, I shall summarize a few fundamental concepts relating
3097 to antennas and antenna measurement, before introducing how Project 8 uses antenna
3098 measurement for the development of antenna array CRES.

3099 5.2.1.1 Radiation Patterns

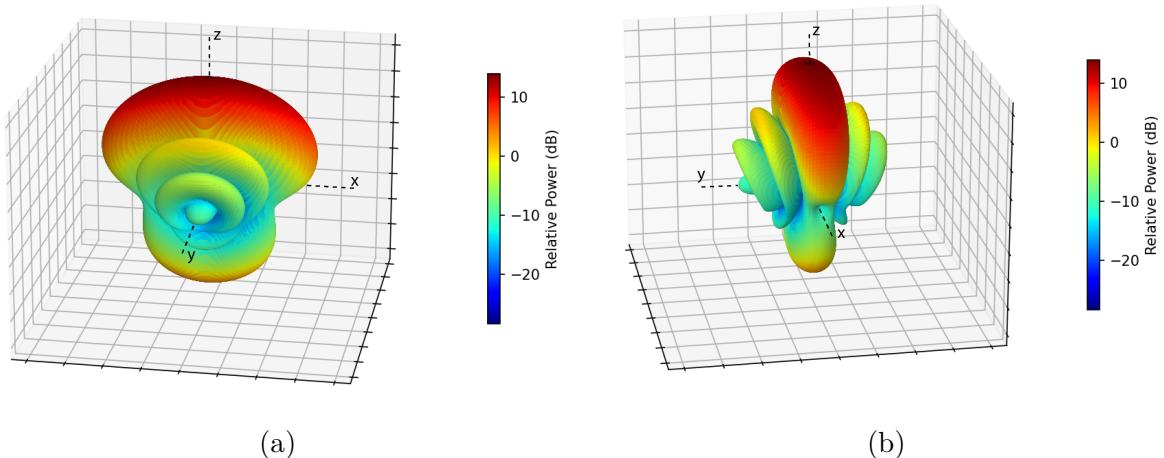


Figure 5.1: An example radiation pattern generated using HFSS simulations. The color and radial distance of the surface from the origin indicate the relative magnitude of radiation power emitted by the antenna in that direction. The primary goal of most antenna measurements is typically to measure the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

3100 Antennas are conductive structures designed to carry alternating electric currents
3101 in order to transmit energy in the form of electro-magnetic (EM) waves [40]. Perhaps
3102 the most fundamental way to characterize an antenna, is to map out the radiated power
3103 density as a function of position, which is called the radiation pattern (see Figure 5.1).
3104 We find the radiation power density by calculating the time-averaged Poynting vector for
3105 all positions surrounding the antenna, which in equation form is

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

3106 where $\mathbf{E}(x, y, z, t)$ and $\mathbf{H}(x, y, z, t)$ are the time-dependent electric and magnetic fields
 3107 produced by the antenna [24]. The radiation power density has units of W/m^2 and is
 3108 more typically called the energy flux density in physics applications, since it is a measure
 3109 of the amount of energy passing through a unit area over time.

3110 Because the radiation power density is a measure of power per unit area, its value
 3111 in a particular direction will depend on the distance from the antenna at which we are
 3112 measuring. This is undesirable for practical applications A related quantity, which is
 3113 distance independent, is the energy flux per unit solid angle or radiation intensity, which
 3114 is computed directly from the radition power density by multiplying by the squared
 3115 distance from the antenna. Specifically,

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

3116 where r is the distance from the antenna to the field measurement point. The radiation
 3117 intensity is typically defined in regions where the Poynting vector consists only of a radial
 3118 component where it is safe to treat as a scalar quantity.

3119 5.2.1.2 Directivity and Gain

3120 Since the radiation intensity is a measure of average power per unit solid angle, it is
 3121 independent of distance and more useful as feature for antenna measurement. However,
 3122 most antenna measurements are performed in terms of the directly related directivity
 3123 and gain quantities. Directivity is defined as the ratio between the radiation intensity at
 3124 particular point on the radiation pattern to the average radiation intensity computed
 3125 over all solid angles [40]. The equation that relates the radiation intensity to directivity
 3126 is

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

3127 where U_0 is the average radiation intensity over all solid angles, which simply the total
 3128 radiated power (P_{rad}) divided by 4π . Closely related to directivity is concept of gain,
 3129 which accounts for energy losses that occur inside then antenna when attempting to
 3130 transmit or receive a signal. The antenna gain is given by

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

3131 where P_{in} is the total power delivered to the antenna. Gain can be thought of as the ratio
 3132 of the antenna's radiation intensity to that of a hypothetical isotropic, lossless radiator.

3133 The maximum values of gain and directivity exhibited by the main lobe of the antenna
 3134 pattern as well as the ratio between the gain of the main lobe and any side-lobes are
 3135 important figures of merit used to evaluate antenna designs.

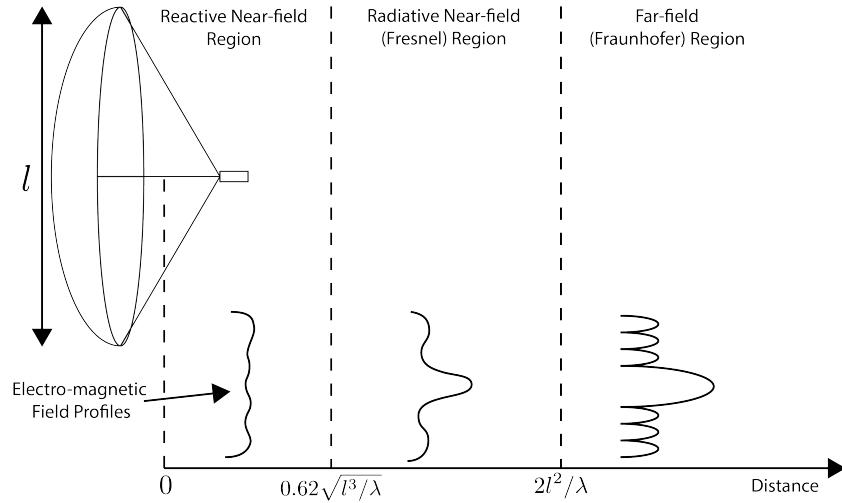


Figure 5.2: An illustration of the three field regions important for the analysis of an antenna system. Very close to the antenna the electric fields are primarily reactive so there is no radiation. If a receiving antenna were placed in this region most of the energy would be reflected back to the transmitter.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

3136 5.2.1.3 Far-field and Near-field

3137 Radiation patterns are only well-defined in regions where the shape of the radiation
 3138 pattern is independent of distance. The region where this approximation is valid is called
 3139 the "far-field", and in this region we can treat the EM fields from the antenna as spherical
 3140 plane waves. A rule of thumb for antennas is that the far-field approximation is valid
 3141 when the condition

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

3142 is met. In this expression, R is the distance from the antenna, l is the largest characteristic
 3143 dimension of the antenna, and λ is the wavelength of the radiation (see Figure 5.2).

3144 The region very close to the antenna is called the reactive near-field, because in this
3145 region the reactive component of the EM field is dominant. Unlike radiative electric
3146 fields, the reactive electric and magnetic fields are out of phase from each other by
3147 90°, since they are the result of electrostatic and magnetostatic effects coming from the
3148 self-capacitance and self-inductance of the antenna. The reactive fields are unable to
3149 transfer energy a significant distance from the antenna and are thus completely negligible
3150 for most antenna applications. The limit of the reactive near-field for an electrically-large
3151 antenna is typically taken to be

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

3152 The unique application of antennas by Project 8 is somewhat limited by reactive near-
3153 field effects in the form of a maximum radial position for electrons inside the uniform
3154 cylindrical antenna array. If electrons are too close to the edge of the array than reactive
3155 near-field effects leads to a large reduction in the received power and consequently
3156 detection efficiency. This leads to a significant volume inside of the antenna array that
3157 is unsuitable for CRES lowering the volumetric efficiency of the antenna array CRES
3158 technique relative to a cavity experiment.

3159 In between the reactive near-field and the far-field is the radiative near-field region.
3160 In this region the fields are primarily radiative, however we are still too close to the
3161 antenna for the spherical plane wave approximation to apply. Therefore, interference
3162 effects between EM waves emitted from different points on the antenna occur causing the
3163 shape of the radiation pattern to change as a function of distance from the antenna. If we
3164 evaluate the far-field distance limit for the FSCD one finds an estimated far-field distance
3165 of 43 cm, which is a factor of four larger than the radius of the antenna array designed for
3166 the experiment. Consequently, we expect near-field effects to influence the performance
3167 of the antenna array highlighting the importance of calibration and characterization
3168 measurements.

3169 **5.2.1.4 Polarization**

3170 The polarization of an EM wave defines the spatial orientation of the electric field
3171 oscillations in the plane perpendicular to the direction of the propagation, and is defined
3172 in terms of orthogonal polarization components. In our application, one analyzes the
3173 properties of radiation propagating along the radial (\hat{r}) direction away from the antenna,
3174 which implies that the electric fields can be described as a linear combination of orthogonal

3175 polarization components

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

3176 in Cartesian coordinates, or

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

3177 in spherical coordinates.

3178 In general, one defines partial radiation patterns, directivities, and gains so that the
3179 performance of the antenna for the desired polarization can be analyzed. The radiation
3180 pattern defined in terms of partial patterns is

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

3181 where U_ϕ and U_θ are the radiation intensities in a particular direction for the respective
3182 polarization components. Similarly, a quantity such as gain can be written in terms of
3183 partial gains,

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

3184 If we view an electron performing a circular orbit in the XY-plane from the side, that
3185 is, along the X or Y axes, then we would observe the electron to be performing a linear
3186 oscillation perpendicular to the viewing axis. From this intuitive picture, we can predict
3187 that the primary polarization of electric fields from CRES events to be linearly polarized
3188 in the $\hat{\phi}$ direction when viewed with an antenna positioned in the XY-plane.

3189 5.2.1.5 Antenna Factor and Effective Aperture

3190 A useful way to characterize the performance of an antenna is to measure the electric
3191 field magnitude required to produce a signal with an amplitude of one volt in the antenna
3192 terminals. This ratio between the magnitude of the incoming electric field and the
3193 magnitude of the signal produced by the antenna is called the antenna factor, which is
3194 written as

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

3195 where A_F is the antenna factor, E_{in} is the incoming electric field, and V_{ant} is the magnitude
3196 of the voltage produced by the antenna.

3197 The antenna factor can be expressed in terms of the antenna's gain through a related
3198 quantity called the effective aperture. The effective aperture defines for a given incident
3199 radiation power density (W/m^2) the power that is received by the antenna. Therefore,

3200 the effective aperture gives the equivalent area of the antenna,

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

3201 where the received power is P_r and the total incoming power is P_{in} .

3202 If we express the incident power in terms of the magnitude of the Poynting vector,
3203 then

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

3204 where η_0 is the impedance of free-space, which relates the magnitudes of the electric and
3205 magnetic fields in a vacuum, and is defined by

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

3206 The total received power by the antenna can therefore be expressed as

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

3207 To relate this to the antenna factor recall that we can relate the voltage produced by
3208 the antenna to the received power with

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

3209 where Z is the system impedance. Setting Equations 5.15 and 5.16 equal to each other,
3210 we obtain the following expression for antenna factor in terms of gain

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

3211 The second expression in Equation 5.17 is obtained by evaluating the constant terms
3212 assuming a system impedance of 50Ω .

3213 We have gone through the effort of expressing the antenna factor in terms of gain
3214 to highlight that the majority of antenna parameters that we care to measure for a
3215 CRES experiment can be obtained from the radiation or gain pattern of the antenna.
3216 The antenna factor is a particularly important parameter for CRES measurements
3217 due to its relevance to antenna array simulations with the Locust software [37, 55].
3218 Specifically, Locust simulates the trajectory of an electron in a magnetic trap by running

3219 the Kassiopeia software package [35] and then uses the Liénard-Wiechert equations [38,39]
3220 to calculate the electric fields that are incident on the antenna.

3221 To compute the response of the antenna to the electric field, Locust relies upon linear
3222 time-invariant system theory [42], which computes the response of the antenna (i.e. the
3223 voltage time series generated by the antenna) using a convolution between the electric field
3224 time-series and the antenna impulse response. This approach is necessary for correctly
3225 modeling the antenna response to the electric field due to the broadband and non-
3226 stationary nature of the electric fields from CRES events. Since antenna measurements
3227 take place under steady-state conditions, parameters such as the radiation pattern, gain,
3228 and antenna factor are defined in the frequency domain. However, by performing an
3229 inverse Fourier transform on the antenna factor we can obtain the antenna impulse
3230 response, which allows us to simulate CRES events in the antenna array demonstrator
3231 experiment.

3232 **5.2.2 Antenna Measurement Fundamentals**

3233 **5.2.2.1 Friis Transmission Equation**

3234 The antenna factor, sometimes called the antenna transfer function, is used to model
3235 how the antenna will respond to electric fields emitted from a CRES event. Therefore,
3236 being able to measure the antenna transfer function of the antenna array is a key step
3237 in the commissioning and calibration phases of an antenna array CRES experiment. A
3238 common approach to antenna characterization is to perform a two antenna transmit-
3239 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).

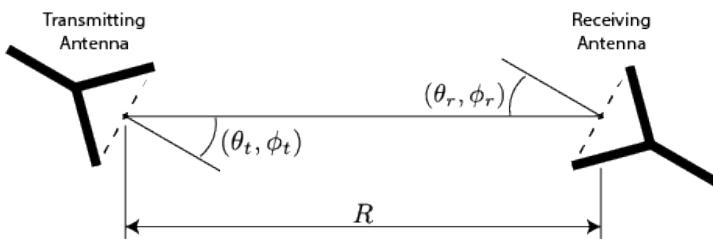


Figure 5.3: An illustration of the Friis measurement technique commonly used for antenna characterization measurements.

3240
3241 To analyze this two antenna setup we seek to calculate the amount of power from
3242 the transmitting antenna that we will detect with the receiving antenna. Using our
3243 understanding of antenna gain, we can calculate the power density transmitted by an

3244 antenna in a direction (θ_t, ϕ_t) at frequency f and distance R , which is given by

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

3245 Here, P_t is the total power delivered to the transmitting antenna and $G_t(\theta_t, \phi_t, f)$ is
3246 the value of the transmitting antenna gain. The power density is the power per unit
3247 area, so to calculate the total power delivered to the receiving antenna we multiply the
3248 transmitted power density by the effective area of the receiving antenna,

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

3249 where $G_r(\theta_r, \phi_r, f)$ is the gain of the receiving antenna. Equation 5.19 is called the
3250 Friis transmission equation [61, 62], which is of fundamental importance for antenna
3251 measurements, since it allows one to measure the gain of an unknown antenna by
3252 measuring the power received from an antenna with a known gain pattern. Alternatively,
3253 if no antenna with a known gain pattern is available, two identical antennas with unknown
3254 gain patterns can be used.

3255 5.2.2.2 S-Parameters and Network Analyzers

3256 Instead of directly measuring the power received by the antenna under test, it is more
3257 common to measure the ratio of the received power to the transmitted power,

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

3258 This power ratio can be easily measured using a vector network analyzer (VNA), which
3259 automates a significant fraction of the measurement process. Network analyzers are
3260 used to measure the scattering or S-parameters of a multi-port RF device [63], which
3261 describes how waves are scattered between the device ports. The antenna measurements
3262 we have been considering can be modeled as a two-port microwave device that we can
3263 characterize by measuring how incident voltage waves are transmitted or reflected (see
3264 Figure 5.4). We can write the scattered waves (V_1^- and V_2^-) in terms of the incident (V_1^+
3265 and V_2^+) waves using the scattering matrix

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

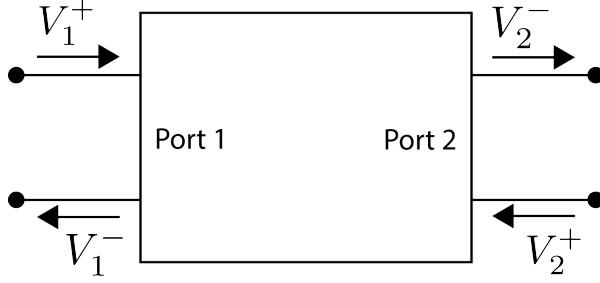


Figure 5.4: Illustration of a two-port S-parameter measurement setup. S-parameters characterize how incoming waves of voltage or power scatter off of the RF device under test. This allows you to measure important properties of the device. In particular, we can use this framework to model a two antenna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

3266 where the elements of the matrix are the device S-parameters. It is assumed that,
 3267 when exciting the device from a particular port, that all other ports in the network are
 3268 terminated at the system impedance. This ensures that the incident waves from other
 3269 ports in the network are zero. Therefore, the S-parameters are the ratios between the
 3270 scattered and incident waves,

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

3271 Alternatively, S-parameters can be defined as the ratio of the scattered and incident
 3272 power, which is proportional to the ratio of the squared voltage waves. Returning to
 3273 our antenna measurement setup, we see that measuring the ratio of the received to the
 3274 transmitted power is equivalent to measuring the ratio of power being scattered from port
 3275 1 to port 2 in a RF network. Therefore, measuring an antenna's gain can be accomplished
 3276 quite easily, by using a VNA to perform a two port S_{21} measurement.

3277 5.2.2.3 Antenna Array Commissioning and Calibration Measurements

3278 Up to this point we have been discussing calibration and commissioning measurements
 3279 as they apply to a single antenna. While these measurements play an important role
 3280 in validating the radiation patterns of the individual array elements, the ultimate goal
 3281 is to use a phased array of these antennas. Therefore, we must also consider antenna
 3282 measurement techniques that apply to the whole array system.

3283 By measuring the gain of each individual array element we can predict the features of
 3284 the signals received during a CRES event using the antenna factor (see Section 5.2.1.5).
 3285 However, unpredictable changes to the antenna performance can be introduced by the

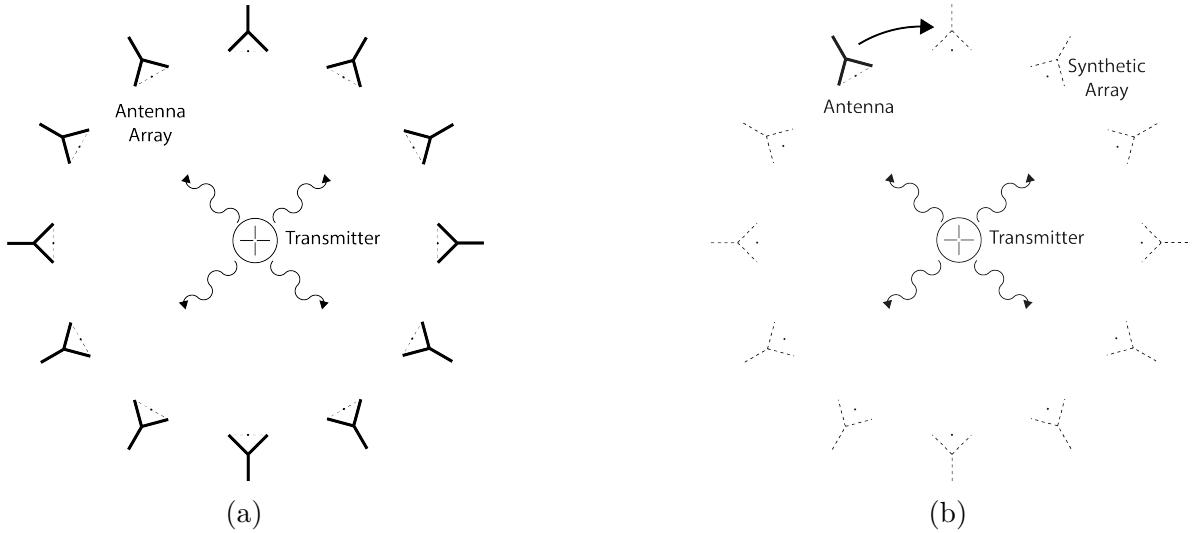


Figure 5.5: Two measurement approaches to characterize an antenna array for CRES measurements. The full-array approach (a) requires a complete antenna array with all the associated hardware. The synthetic array approach (b) utilizes a single antenna and a set of rotation/translation stages to reposition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

3299 5.2.3 The Penn State Antenna Measurement System

3300 The development of antenna array based CRES requires the capability to test and calibrate
3301 different antenna array designs to validate the performance of the as-built antenna array
3302 before and during the experiment. With these aims in mind we developed an antenna
3303 measurement system at Penn State specifically designed to mimic the characteristics of
3304 the antenna experiment designed for demonstration of the antenna array CRES technique
3305 by the Project 8 collaboration.

3306 The Penn State antenna measurement system utilizes a two antenna measurement
3307 configuration with a stationary reference antenna and a test antenna mounted on a set
3308 of motorized translation and rotation stages (see Figure 5.6). The antenna measurement
3309 system can be operated in two distinct modes, one focused on the characterization of
3310 the radiation patterns of prototype antennas and the other focused on the validation
3311 of data-acquisition (DAQ) and CRES signal reconstruction techniques to bridge the
3312 gap between real measurements and simulation. In both measurement configurations it
3313 is critical to isolate the antennas from the environment so that multi-path reflections
3314 do not negatively influence the measurement results. For this reason we surround the
3315 measurement volume with microwave absorber foam (AEMI AEC-1.5) [64] specifically
3316 designed to attenuate microwave radiation near the 26 GHz measurement range of the
3317 system.

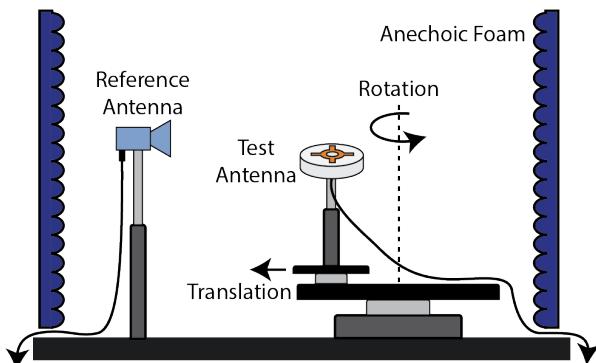


Figure 5.6: Illustration of the antenna measurement system developed for the Project 8 Collaboration. The reference and test antennas can be connected to different data acquisition configurations depending on the measurement goals. The reference antenna is typically a standard horn antenna and the test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

3318 In the first measurement configuration the reference antenna is typically a well-

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

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

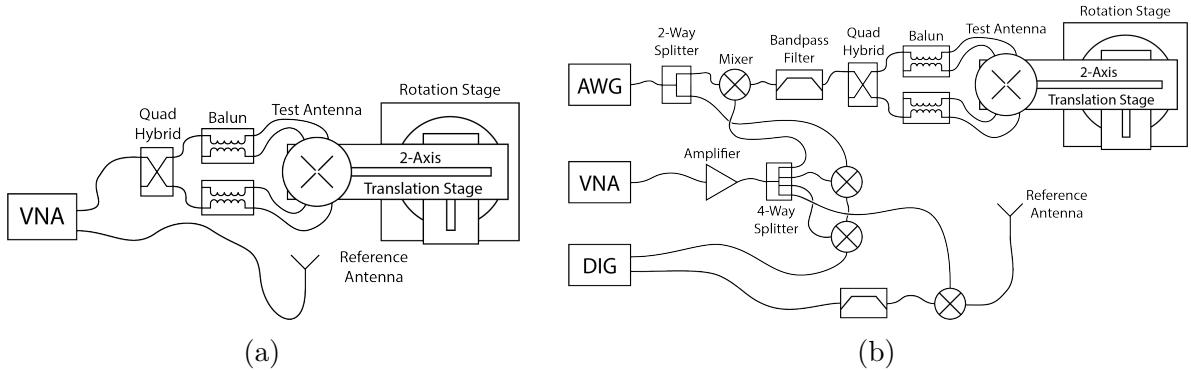


Figure 5.7: Diagrams of two measurement system configurations. Configuration (a) utilizes a VNA and is more suited to antenna characterization. Configuration (b) utilizes an AWG and VNA as a signal generation system and digitizer to collect measurement data, which is more suited to simulating CRES measurements.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

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

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

The characterization configuration utilizes a network analyzer (Keysight N5222A) [65, 66] with two independent sources and four measurement ports as the primary measurement tool. A standard reference antenna is connected to one measurement port, and the test antenna is connected to a separate port. The typical reference antenna used for these studies is a Pasternack PF9851 horn antenna [67]. In the measurement shown, the test antenna represents a SYNCA antenna, which requires a transmission chain consisting of quadrature hybrid coupler [68, 69] (Marki QH-0226) connected to two baluns [70] (Marki BAL-0026) to generate feed signals with the appropriate phases. The VNA measures the radiation pattern by performing a transmission S-parameter measurement, which can be used with the knowledge of the reference antenna's radiation pattern to determine the radiation pattern of the test antenna (see Section 5.2.1).

The second configuration is more complicated and incorporates more hardware components in order to more closely mimic the DAQ system envisioned for the FSCD experiment. The basic approach is to produce CRES-like radiation and use an antenna combined with a realistic RF receiver chain to acquire the signals. On the transmit side, an arbitrary waveform generator [71] (AWG, RIGOL DG5252) is used to generate a waveform that mimics a CRES signal at a baseband frequency up to 250 MHz. This frequency is then up-converted to the CRES signal frequency band of 25.8 to 26.0 GHz using a mixer [72] (Marki MM1-0832L) and a bandpass filter (K&L Microwave 3C62-25900/T200-K/K) to reject unwanted mixing components outside out of the 200 MHz CRES signal band. The local oscillator signal for mixing is provided by one of the VNA sources configured to run in a continuous wave setting. On the receive side, a prototype antenna is used to detect the radiation emitted by the test antenna, which is down-converted and filtered using the same mixer and bandpass filter as the transmission chain. Lastly, data acquisition is performed using a 14-bit ADC sampling at 500 MSa/s [73] (CAEN DT530) to digitize the down-converted signals.

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

3375 the starting phase of the transmitted signal. In order to monitor the random phase of the
3376 LO, a 2-way power splitter (MiniCircuits Z99SC-62-S+) is used to split the signal from
3377 the AWG between the transmission path and a LO monitoring path. The LO monitoring
3378 path consists of an up-conversion and down conversion using two mixers connected by a
3379 coaxial cable, and monitors the relative phase of the LO using a channel on the digitizer
3380 to sample this path. A phase shift in the LO will lead to a proportional phase shift in
3381 the mixed signal, which is measured and removed from the received signals.

3382 The test antenna is mounted on a set of motorized stages, which are identical for
3383 both measurement configurations. A rotational stage (ThorLabs PRMTZ8) is used as
3384 the base layer with additional translation stages mounted on top of this. The rotational
3385 stage is ideal for measuring a complete azimuthal scan of the test antenna's radiation
3386 pattern as well as for moving a SYNCA antenna in circular motion to recreate the
3387 symmetry of the FSCD antenna array. On top of the rotational stage we mount two
3388 linear translation stages (ThorLabs MTS50-Z8 and MTS25-Z8) in a cross-wise manner
3389 so that the test antenna can be moved along two perpendicular axes. Using the linear
3390 stages in combination with the rotational stage allows one to fine-tune the positioning of
3391 the test antenna so that it can be perfectly aligned with the central axis of the array.
3392 A LabView script was developed to automate the measurement of a full 360° radiation
3393 pattern and control the measurement electronics. Data from these acquisitions is stored
3394 on university provided cloud storage.

3395 **5.3 Development of a Synthetic Cyclotron Antenna (SYNCA) 3396 for Antenna Array Calibration**

3397 This section is the manuscript of the publication [56] detailing the development of a
3398 Synthetic Cyclotron Antenna (SYNCA) for antenna array characterization measurements
3399 by the Project 8 collaboration.

3400 **5.3.1 Introduction**

3401 Neutrinos are the most abundant standard model fermions in our universe, but due to
3402 weak interaction cross-sections with other particles, neutrinos are particularly difficult
3403 to study. Consequently, many fundamental properties of neutrinos are still unknown
3404 including the absolute scale of the neutrino mass [5]. Direct, kinematic measurements of
3405 the neutrino mass are particularly valuable due to their model independent nature [12].

3406 To date the most sensitive direct neutrino mass measurements have been performed by
 3407 the KATRIN collaboration [74], which measures the molecular tritium β -decay spectrum
 3408 to infer the neutrino mass. Current data from neutrino oscillation measurements [5]
 3409 allow for neutrino masses significantly smaller than the design sensitivity of the KATRIN
 3410 experiment; therefore, there is a need for new technologies for performing direct neutrino
 3411 mass measurements to probe lower neutrino masses.

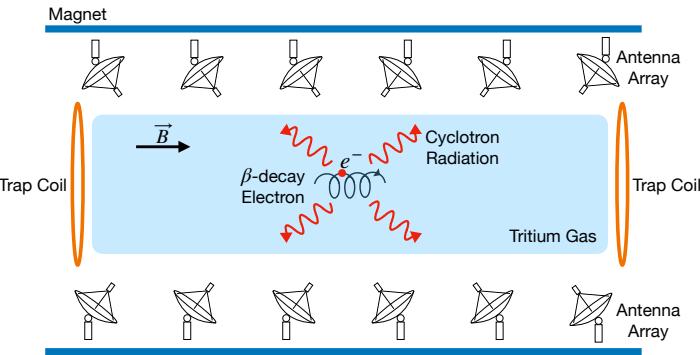


Figure 5.8: A sketch of an antenna array large-volume CRES experiment. Electrons from β -decays are confined in a magnetic field using a set of trap coils. The cyclotron radiation produced by the motion of the trapped electrons can be detected by a surrounding antenna array to determine the electron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

3412 The Project 8 collaboration is developing new methods for neutrino mass measurement
 3413 based on Cyclotron Radiation Emission Spectroscopy (CRES) [31, 75–77], with the goal
 3414 of measuring the absolute scale of the neutrino mass with a $40 \text{ meV}/c^2$ sensitivity [?, 12].
 3415 This sensitivity goal will require the development of two separate technical capabilities.
 3416 First is the development of an atomic tritium source, which avoids significant spectral
 3417 broadening due to molecular final states [30]. Second is the technology for performing
 3418 CRES in a multi-cubic-meter experimental volume with high combined detection and
 3419 reconstruction efficiency, which is required in order to obtain sufficient event statistics
 3420 near the tritium spectrum endpoint.

3421 One approach for a large-volume CRES experiment is to use an array of antennas,
 3422 which surrounds a volume of tritium gas, to detect the cyclotron radiation produced
 3423 by the β -decay electrons when they are trapped in a background magnetic field using a
 3424 set of magnetic trapping coils (see Figure 5.8). Project 8 has developed a conceptual
 3425 experiment design to study the feasibility of this approach. The design consists of a
 3426 single circular array of antennas with a radius of 10 cm and 60 independent channels
 3427 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 [55, 78], 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.

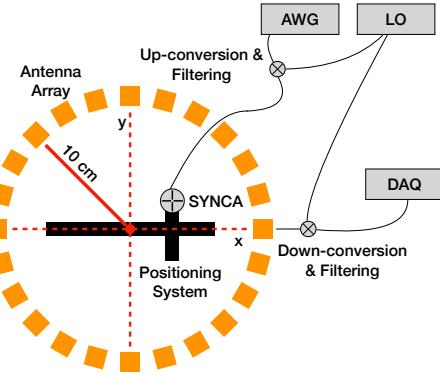


Figure 5.9: A schematic of the antenna array test stand. The circular antenna array has a radius of 10 cm with 60 independent channels (limited number shown for clarity). The test stand includes an arbitrary waveform generator (AWG), local oscillator (LO), and data acquisition (DAQ) hardware. Finally,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

3449 the fields generated by the SYNCA match simulation, and finally in Section 5.3.5 we
 3450 demonstrate an application of the SYNCA to test phased array reconstruction techniques
 3451 on the bench-top.

3452 5.3.2 Cyclotron Radiation Phenomenology

3453 To understand the cyclotron radiation phenomenology that the SYNCA should mimic,
 3454 we consider a charged particle moving at relativistic speed in the presence of an external
 3455 magnetic field (see Figure 5.10). In the special case we shall examine, the entirety of
 3456 the electron's momentum is directed perpendicular to the magnetic field; therefore, the
 3457 trajectory of the electron is confined to the cyclotron orbit plane. Because the momentum
 3458 vector is oriented perpendicular to the magnetic field, electrons with these trajectories
 3459 are said to have pitch angles of 90°.

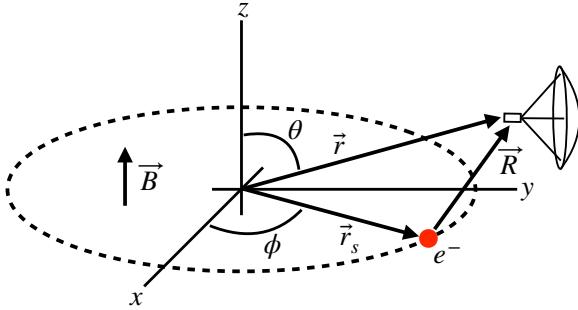


Figure 5.10: An electron (red dot) performing cyclotron motion in the x-y plane. The resulting cyclotron radiation is observed by an antenna located at the field point of interest.

3460 The cyclotron radiation fields generated by this circular trajectory are those which
 3461 we aim to reproduce with the SYNCA. We can describe the electromagnetic (EM) fields
 3462 using the Liénard-Wiechert equations [24, 55], which in non-covariant form express the
 3463 electric field as

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

3464 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
 3465 electron to the field measurement point, $\vec{\beta} = \dot{\vec{r}}_s/c$ is the velocity of the particle divided
 3466 by the speed of light, and γ is the relativistic Lorentz factor. The equation is meant to
 3467 be evaluated at the retarded time as indicated by $t_r = t - |\vec{R}|/c$, which accounts for the

³⁴⁶⁸ time delay due to the finite speed of light between the point where the field was emitted
³⁴⁶⁹ and the point where the field is detected.

³⁴⁷⁰ We would like to simplify Equation 5.23 it at all possible. As a first step we analyze
³⁴⁷¹ the relative magnitudes of the electric field polarization components. Consider an electron
³⁴⁷² following a circular cyclotron orbit in a uniform magnetic field whose guiding center
³⁴⁷³ is positioned at the origin of the coordinate system. The equation of motion can be
³⁴⁷⁴ expressed as

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

³⁴⁷⁵ For single antenna located along the y-axis at position $\vec{r} = r_a \hat{y}$ we are interested in the
³⁴⁷⁶ incident electric fields from the electron. The electric field is given by Equation 5.23,
³⁴⁷⁷ which we evaluate in the regime where $r_a \gg r_c$. This limit can be justified by comparing
³⁴⁷⁸ the radius of the cyclotron orbit for an electron with the tritium beta-spectrum endpoint
³⁴⁷⁹ energy of 18.6 keV in a 1 T magnetic field to the typical ($r_a \simeq 100$ mm) radial position
³⁴⁸⁰ of the receiving antenna. We find that the cyclotron orbit has a radius of 0.46 mm which
³⁴⁸¹ is approximately a factor of 200 smaller than the typical antenna radial position. In this
³⁴⁸² regime we can make the approximation $\vec{R} \simeq r_a \hat{y}$ and the expression for the electric field
³⁴⁸³ at the antenna's position becomes

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

³⁴⁸⁴ Since the receiving antenna is part of a circular array of antennas, it is useful to rewrite
³⁴⁸⁵ Equation 5.25 in terms of the azimuthal ($\hat{\phi}$) and radial (\hat{r}) polarizations. Making use of
³⁴⁸⁶ 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

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

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

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

³⁴⁸⁷ For the purposes of designing a synthetic cyclotron radiation antenna we are interested
³⁴⁸⁸ in the dominant electric field polarization emitted by the electron. The antenna is being
³⁴⁸⁹ designed to mimic the cyclotron radiation produced by electrons with kinetic energies of
³⁴⁹⁰ approximately 18.6 keV in a 1 T magnetic field [30]. Since the relativistic beta factor for

3491 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
 3492 justified. Inserting these expressions into the equations for the electric field components
 3493 above simplifies the comparison of the magnitudes of the two components. Additionally,
 3494 we compare the time-averaged magnitudes to evaluate the root mean squared electric
 3495 field ratio. The time-averaged ratio of the radial and azimuthally polarized electric fields
 3496 with the above simplifications is given by

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

3497 where we have made use of the fact that for these magnetic fields and kinetic energies
 3498 the cyclotron radius is much smaller than the radius of the antenna array.

3499 From Equation 5.29 we see that the time-averaged azimuthal polarization is larger than
 3500 the radial polarization by about a factor of 8, which makes it the dominant contribution
 3501 to the electric fields at the position of the antenna. We must also consider the directivity
 3502 of the receiving antenna which can have a gain that is disproportionately large for a
 3503 specific polarization component. Because the E_ϕ component is dominant, the receiving
 3504 antenna array is designed with an azimuthal polarization, which negates the voltages
 3505 induced in the antenna from the radially polarized fields. Therefore, we conclude that
 3506 for the purpose of designing the SYNCA antenna it is acceptable to approximate the
 3507 electric fields from Equation 5.23 as purely azimuthally or ϕ -polarized. The simplified
 3508 expression for the electric field received by an antenna becomes

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

3509 where the radius of the cyclotron orbit is called r_c , the cyclotron frequency is called ω_c ,
 3510 and the radial position of the receiving antenna is called r_a . Equation 5.30 has been
 3511 evaluated in the non-relativistic limit where $\gamma \simeq 1$, which is justified by the fact that
 3512 $|\vec{\beta}| \simeq \frac{e}{4}$ for an electron with an 18.6 keV kinetic energy in a 1 T magnetic field.

3513 This rather complicated expression can be simplified using Fourier analysis. Assuming
 3514 a background magnetic field of 1 T and a kinetic energy of 18.6 keV we calculate
 3515 numerically the electric field using Equation 5.30 and apply a discrete Fourier Transform
 3516 to visualize the frequency spectrum (see Figure 5.11).

3517 We observe that the azimuthally polarized electric field is periodic with a base cyclotron
 3518 frequency of 25.898 GHz corresponding to the highest power frequency component in

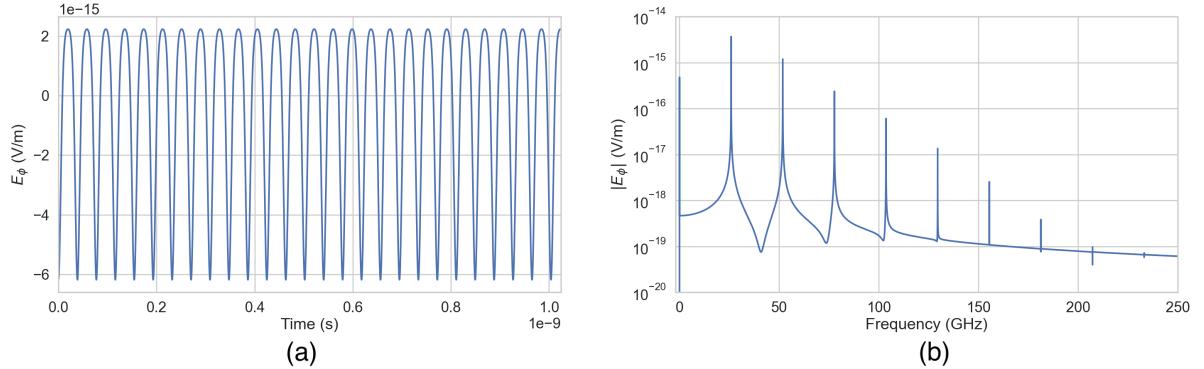


Figure 5.11: A plot of the numeric solution to Equation 5.31. The time-domain representation of the signal (a) is composed of a zero frequency term and a series of harmonics separated by the main cyclotron frequency as shown in the plot of the frequency spectrum (b). We can see that the relative amplitude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

3519 Figure 5.11. The frequency spectrum reveals that the signal is composed of a constant
 3520 term with zero frequency and a series of harmonics separated by 25.898 GHz. Therefore,
 3521 we can represent the azimuthal electric fields from the electron as a linear combination
 3522 of pure sinusoids with frequencies given by $\omega_k = k\omega_c$ ($k \in 0, 1, 2, \dots$) and amplitudes
 3523 extracted from the Fourier representation. Using this representation we can transform
 3524 the equation for the azimuthally polarized electric fields in Equation 5.30 into

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

3525 where we have truncated the sum over harmonics at the 7th order for completeness. The
 3526 amplitudes A_k are dimensionless complex numbers, which encode the relative powers of
 3527 the harmonics as well as the starting overall phase of the cyclotron radiation. Because
 3528 magnitude of the relative amplitudes exponentially decreases for higher harmonics, it is
 3529 usually sufficient to consider only the terms up to $k = 4$ where the relative amplitude
 3530 of the harmonics has decreased from the main carrier by a factor of approximately 100.
 3531 However, for completeness we include harmonics up to 7th order in Equation 5.31. The
 3532 range of frequencies to which the receiving antenna array in the antenna test stand is
 3533 sensitive is defined by the antenna's transfer function. The receptive bandwidth for
 3534 the antennas used in the test stand is a range of frequencies with a bandwidth on the
 3535 order of a few GHz centered around the main cyclotron carrier frequency of 25.898 GHz.
 3536 Therefore, the higher order harmonics as well as the zero frequency term can be ignored

3537 when considering only the signals that will be received by the antenna array.

3538 Considering only the 1st order harmonic term from Equation 5.31, which represents
3539 the portion of the electric field that will be detected by the array, and evaluating this at
3540 the retarded time we obtain the following for the ϕ -polarized electric fields

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

3541 where the arbitrary phase Δ is defined by $A_k = |A_k|e^{i\Delta}$. We are interested in the
3542 characteristics of the amplitude of the electric field as a function of the radial distance
3543 component ($|\vec{R}|$) of the retarded time. In particular, the maximum of E_ϕ occurs when
3544 the argument of the cosine function is equal $n\pi$ where $n \in \{0, \pm 2, \pm 4, \dots\}$; however, the
3545 solutions where n is negative can be discarded since they represent unphysical negative
3546 overall phases. Applying this condition to Equation 5.32 gives a condition on the radial
3547 position of the maximum of E_ϕ

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

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

3548 which is a function of time in the frame of the moving electron (t). Equation 5.33 can
3549 be further simplified by noticing that the azimuthal position of the electron ($\phi_e(t)$) as a
3550 function of time is defined by $\phi_e(t) = \omega_c t - \Delta$ which reduces Equation 5.33 to

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

3551 Equation 5.34 represents an archimedean spiral which is formed when plotting the
3552 amplitude of E_ϕ in the x-y plane. The solution where $n = 0$ represents the leading edge
3553 of the radiation spiral which propagates outward from the electron at the speed of light.
3554 The additional solutions for $n > 0$ represent the persistent spiral at radii inside the
3555 leading edge of the radiated fields that have not yet been detected by the receiver at the
3556 current time. In Figure 5.12a we show the expected spiral pattern for the maxima of the
3557 cyclotron radiation.

3558 In particular, we note that for the circular array geometry of the test stand, depicted
3559 as the series of circles in Figure 5.12a, each antenna receives a linearly polarized wave
3560 with a phase offset that corresponds to the azimuthal angle for that antenna element.
3561 Therefore, as we show in Figure 5.12b, when the relative phase of the received signal is
3562 plotted as a function of the receiving antenna's azimuthal position the result is also an

3563 Archimedean spiral.

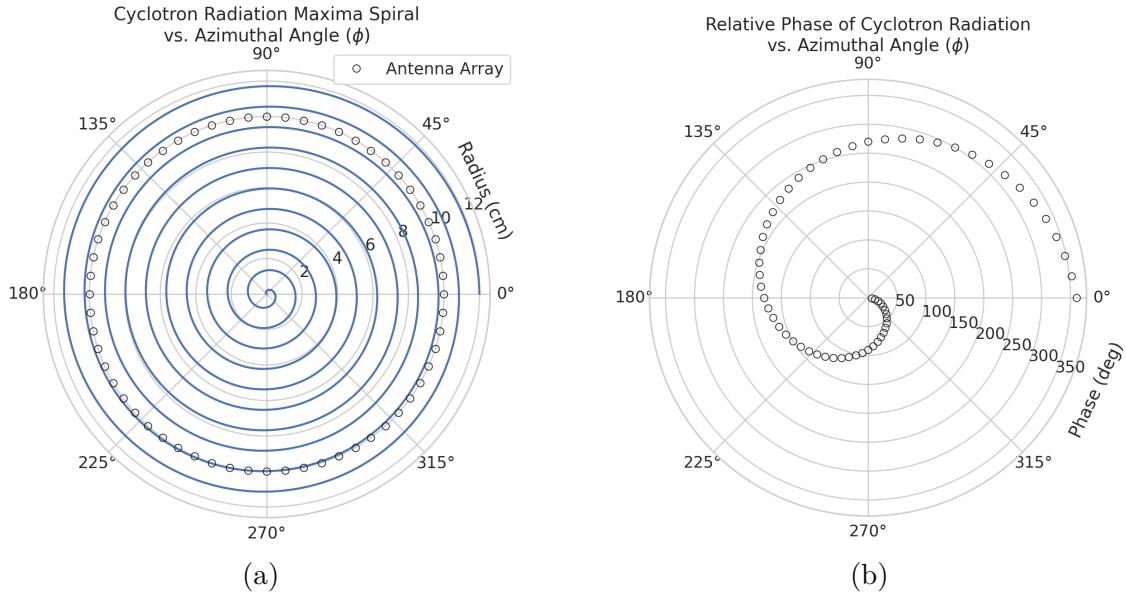


Figure 5.12: The amplitude maxima of the cyclotron radiation form an Archimedean spiral as the radiation propagates outward from the cyclotron orbit center (a). A circular antenna array located at a fixed radius from the orbit center will receive electric fields with equal magnitude in each of its channels, 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).

3564 Based on these analytical calculations we can characterize the magnitude, polarization,
3565 and phase of the signals received by the antenna array using three criteria. These criteria
3566 are the basis of comparison for the radiation produced by the SYNCA and cyclotron
3567 radiation emitted by electrons and will be used to evaluate the performance of antenna
3568 designs. The criteria are:

- 3569 1. Electric fields that are ϕ -polarized near $\theta = 90^\circ$
- 3570 2. Uniform time-averaged electric field magnitudes around the circumference of a
3571 circle centered on the antenna
- 3572 3. Electric fields whose phase is equal to the azimuthal angle at the point of measure-
3573 ment plus a constant

3574 The Locust simulation package [78] can be used to directly simulate the EM fields
3575 generated by electrons performing cyclotron motion to validate the analytical calculations.
3576 Locust simulates the EM fields by first calculating the trajectory of the electrons in

3577 the magnetic trap using the Kassiopeia software package [79]. The trajectory can then
 3578 be used to solve for the EM fields using the Liénard-Wiechert equations directly with
 3579 no approximations. The resulting electric field solutions drive a receiving antenna by
 3580 convolving the time-domain fields with the finite-impulse response filter of the antenna
 3581 or they can be examined directly to study the field characteristics that the SYNCA must
 3582 reproduce. In the next section we compare the radiation field patterns for electrons
 3583 simulated with Locust to patterns from a SYNCA antenna design.

3584 5.3.3 SYNCA Simulations and Design

3585 One potential SYNCA design is the crossed-dipole antenna [80]. A crossed-dipole antenna
 3586 consists of two dipole antennas, one of which is rotated 90° with respect to the other,
 3587 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

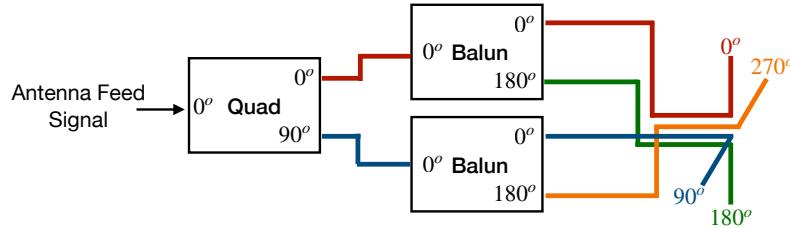


Figure 5.13: An idealized crossed-dipole antenna consists of two electric dipole antennas oriented perpendicular to each other and is fed with four signals with a quadrature phase relationship. An example antenna feed circuit is shown which is composed of a chained combination of a quadrature hybrid-coupler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

3588
 3589 out of phase from each of the neighboring arms by 90°, which mirrors the spatial phase
 3590 relationship of cyclotron radiation fields.

3591 A potential drawback of this design is that standard crossed-dipole antennas do not
 3592 radiate uniform electric fields near the $\theta = \pi/2$ plane. Typical crossed-dipole antennas
 3593 use dipole arm lengths equal to $\lambda/4$ or larger [80], where λ is the wavelength at the
 3594 desired operating frequency. Such large arm lengths cause the electric field magnitude
 3595 to vary significantly around the circumference of the antenna. However, making the
 3596 antenna electrically small by shrinking the arm length can improve the antenna pattern
 3597 uniformity.

3598 In general, the criterion for an electrically small antenna is that the largest dimension
3599 of the antenna (D) obey $D \lesssim \lambda/10$ [40]. In our application, we are attempting to mimic
3600 the cyclotron radiation emitted by electrons produced from tritium β -decay with energies
3601 near the spectrum endpoint. For a background magnetic field of 1 T, the corresponding
3602 cyclotron frequency of tritium endpoint electrons is approximately 26 GHz. Therefore, the
3603 electrically small condition would require that the largest dimension of the crossed-dipole
3604 antenna be smaller than 1.2 mm.

3605 A crossed-dipole antenna with an overall size of 1.2 mm is challenging to fabricate due
3606 to the small dimensions of the dipole arms that, in practice, are fragile and unsuitable
3607 for use as a calibration probe. To mitigate some of the challenges with the fabrication
3608 of such a small antenna, a variant crossed-dipole antenna design using printed circuit
3609 board (PCB) technology (see Figure 5.14) was developed in partnership with an antenna
prototyping company, Field Theory Consulting ¹.

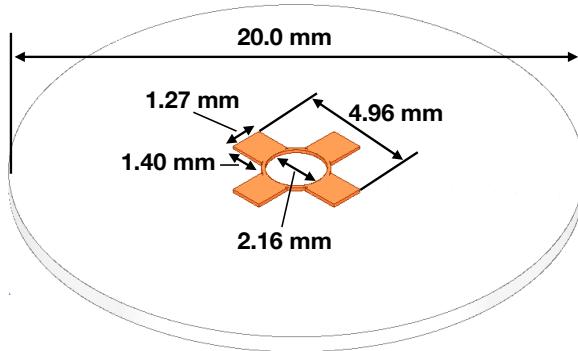


Figure 5.14: A model of the PCB crossed-dipole antenna with dimensions. The design has an inside diameter of 2.16 mm for the central circular trace, which is 0.13 mm wide. The dipole arms each have a width of 1.27 mm and protrude beyond the circular trace by 1.40 mm, which gives an overall width of 4.96 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.

3610
3611 The PCB crossed-dipole design uses four rectangular pads to represent the dipole arms,
3612 which are connected by a thin circular trace. The circular trace both adds mechanical
3613 stability to the antenna and improves the azimuthal uniformity of the electric fields
3614 compared to a more standard crossed-dipole geometry. Furthermore, the circular trace
3615 allows for a greater separation between dipole arms than standard crossed-dipoles, which

¹<https://fieldtheoryinc.com/>

3616 is required to accommodate the coaxial connections to each pad. The pads each contain
 3617 a through-hole solder joint to connect coaxial transmission lines using hand soldering.
 3618 The antenna PCB has no ground plane on the bottom layer as this was observed in
 3619 simulation to significantly distort the radiation pattern in the plane of the PCB. The
 3620 only ground planes present in the model are the outer conductors of the four coaxial
 3621 transmission lines which feed the antenna. These are left unterminated on the bottom of
 3622 the PCB dielectric material.

3623 The antenna design development utilized a combination of Locust electron simula-
 3624 tions and antenna simulations using ANSYS HFSS [41], a commercial finite-element
 3625 electromagnetic simulation software. Two antenna designs were simulated: an idealized
 3626 electrically small crossed-dipole antenna with an arm length of 0.40 mm and an arm
 3627 separation of 0.05 mm, as well as a PCB crossed-dipole antenna with the dimensions
 3628 shown in Figure 5.14. Plotting the magnitude of the electric fields generated by the
 3629 antennas across a 10 cm square located in the same plane as the respective antennas
 3630 reveals the expected cyclotron spiral pattern (see Figure 5.15) which closely matches
 3631 the prediction for simulated electrons. The spiral pattern demonstrates that the electric
 3632 fields have the appropriate phases to mimic cyclotron radiation, which fulfills SYNCA
 criterion 3 identified in Section 5.3.2.

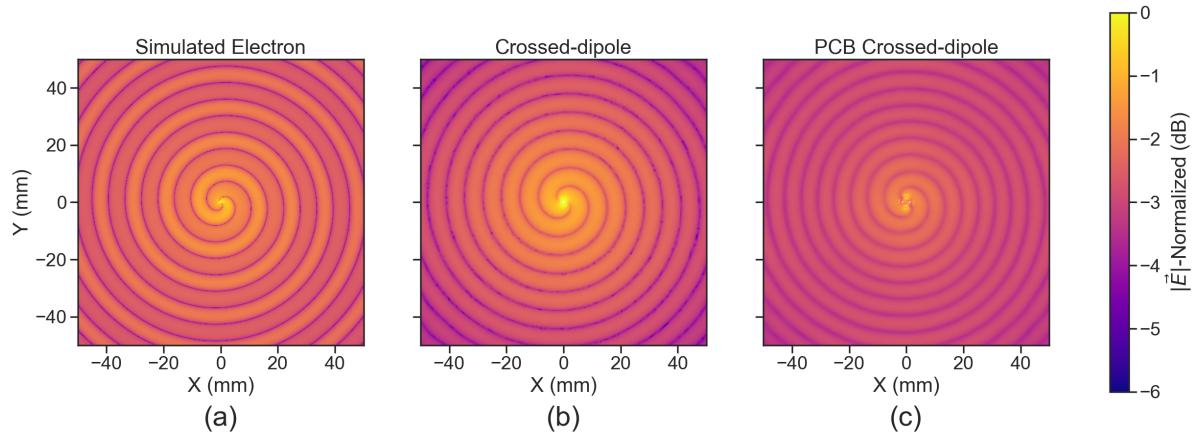


Figure 5.15: A comparison of the electric field magnitudes, normalized by the maximum value of the electric field in each simulation, plotted on a 10 cm square to visualize the Archimedean spirals formed by the electron (a), the crossed-dipole antenna (b), and a PCB crossed-dipole antenna (c). The matching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

3633

3634 As we can see from Figure 5.16, the crossed-dipole antenna, which uses an idealized

3635 geometry, exhibits good agreement with simulation. The antenna has a maximum
 3636 deviation from a simulated electron of approximately 0.5 dB in the total electric field, 1
 3637 dB for the ϕ -polarized electric field and 1 dB for the θ -polarized electric field.

3638 In comparison, the pattern of the PCB crossed-dipole antenna, because the simulation
 3639 incorporates the geometry of the coax transmission lines, exhibits some distortion from
 3640 the idealized cross-dipole simulations. The vertically oriented ground planes of the coax
 3641 lines introduce more θ -polarized electric fields than are observed for simulated electrons
 3642 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

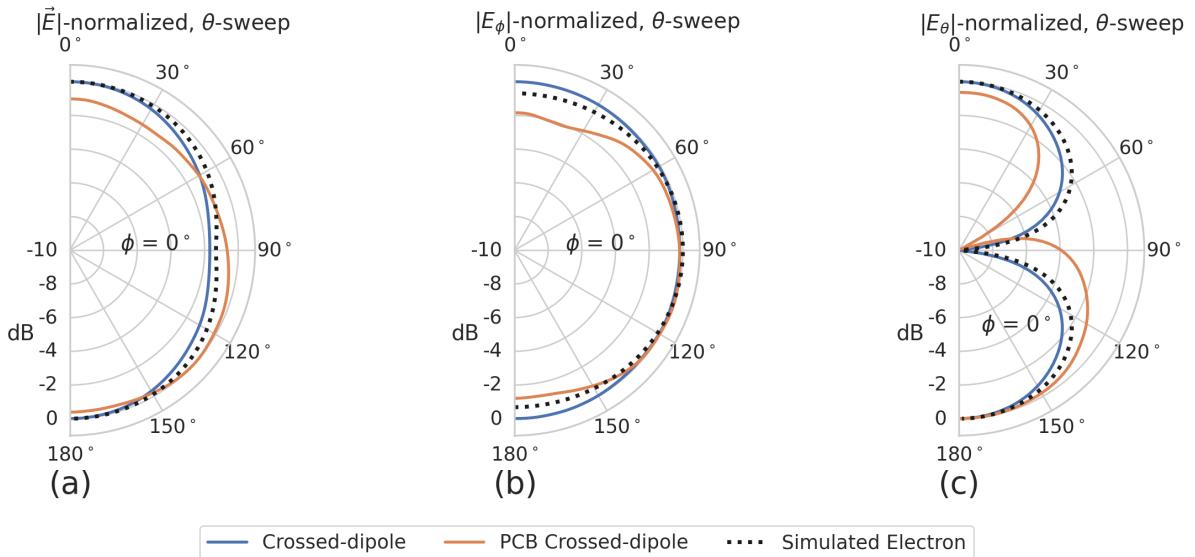


Figure 5.16: A comparison of the normalized electric field magnitudes for the ideal crossed-dipole, PCB crossed-dipole, and a simulated electron as a function of the polar angle (θ). (a) Shows the total electric field, (b) shows the ϕ -polarized electric field component, and (c) shows the θ -polarized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

3643
 3644 antenna should not greatly impact the performance of the antenna because the receiving
 3645 antenna array is primarily ϕ -polarized. Therefore deviations in the θ -polarized fields
 3646 will be suppressed due to the polarization mismatch. More importantly, the ϕ -polarized
 3647 electric field pattern generated by the PCB crossed-dipole closely matches simulated
 3648 electrons across the polar angle range of $50^\circ < \theta < 150^\circ$. In this region the PCB crossed-
 3649 dipole differs by less than 0.5 dB from simulated electrons. This range greatly exceeds
 3650 the beamwidth of the receiving antenna array which is designed to be most sensitive
 3651 to fields produced near $\theta = 90^\circ$. Therefore, we conclude that the PCB crossed-dipole

3652 antenna generates a ϕ -polarized radiation pattern that fulfills SYNCA criterion 1 from
3653 Section 5.3.2.

3654 The final SYNCA criterion is related to the uniformity of the electric fields when
3655 measured azimuthally around the antenna. As we saw for real electrons in Section 5.3.2
3656 it is expected that the magnitude of the electric field be completely uniform as a function
3657 of the azimuthal angle due to the symmetry of the cyclotron orbit. In Figure 5.17 we plot
3658 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

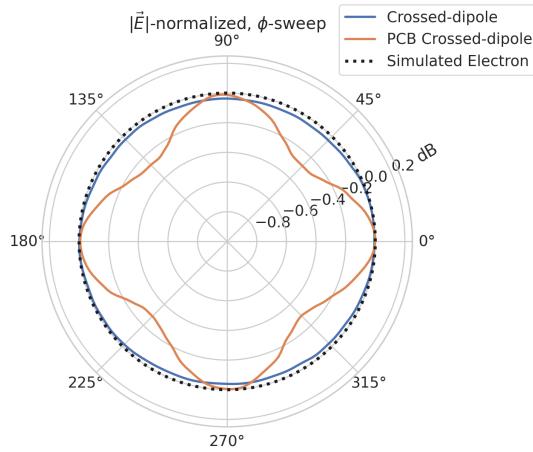


Figure 5.17: A comparison of the normalized electric field magnitudes for the crossed-dipole, PCB crossed-dipole, and a simulated electron as a function of the azimuthal angle (ϕ) evaluated at $\theta = 90^\circ$. This image was generated using Locust simulations for the electron and ANSYS HFSS for both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

3659
3660 perfect uniformity around the azimuthal angle, whereas the PCB crossed-dipole has a
3661 small periodic deviation with a maximum difference of 0.3 dB caused by the coaxial
3662 transmission lines below the PCB. Such a small deviation from uniformity is acceptable
3663 since it is smaller than the expected variation in uniformity caused by imperfections in
3664 the antenna fabrication process, which modifies the antenna shape in an uncontrolled
3665 manner by introducing solder blobs with a typical size of a few tenths of a millimeter on
3666 the dipole arms (see Figure 5.18). Additionally, the SYNCA will be separately calibrated
3667 to account for azimuthal differences in the electric field magnitude. Therefore we see
3668 from the simulated performance of the PCB crossed-dipole antenna that this antenna
3669 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.

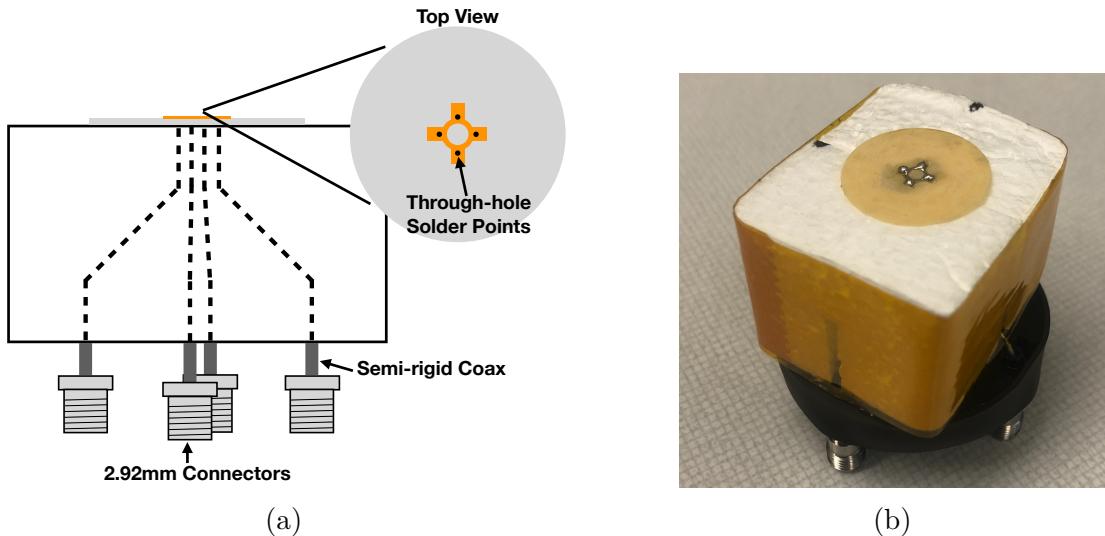


Figure 5.18: (a) A cartoon schematic which highlights the routing of the semi-rigid coax transmission lines. (b) A photograph of a SYNCA constructed using the modified crossed-dipole PCB antenna design. Visible in the photograph of the SYNCA are four blobs of solder which are an artifact of the SYNCA's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

3685 at one port through a hybrid-coupler whose outputs are connected to two baluns to
 3686 generate the signals with the appropriate phases to feed the SYNCA (see Figure 5.13).
 3687 The other port of the VNA is connected to a single reference horn antenna that serves
 3688 as a field probe. To position the SYNCA, a combination of translation and rotation
 3689 stages are used to characterize the antenna's fields across the entire radiation pattern
 3690 circumference. This measurement scheme is equivalent to measuring the fields generated
 3691 by the SYNCA using a full circular array of probe antennas.

3692 The antenna measurement space is surrounded by RF anti-reflective foam to isolate
 3693 the measurements from the lab environment (see Figure 5.19b) and remaining reflections
 3694 are removed using the VNA's time-gating feature. The SYNCA is affixed to the stages
 3695 by a custom RF transparent mount made of polystyrene foam. The coaxial cables deliver
 3696 the antenna feed signals generated by the VNA to the SYNCA while still allowing
 3697 unrestricted rotation. The horn antenna probe is nominally positioned in the plane
 3698 formed by the antenna PCB ($\theta = 90^\circ$ or $z = 0$ mm) at a distance of 10 cm from the
 3699 SYNCA, to match the expected position of the antenna array relative to the SYNCA in
 3700 the antenna array test stand. The horn antenna can be manually raised or lowered to
 3701 different relative vertical positions to characterize the radiation pattern at different polar
 3702 angles.

3703 Several 360° scans were performed with probe vertical offsets of -10.0 mm, -5.0 mm,
 3704 0.0 mm, 5.0 mm, and 10.0 mm relative to the antenna PCB plane. These probe offsets

3705 cover a 2 cm wide vertical region centered on the SYNCA PCB, approximately equal to
 3706 ± 6 degrees of polar angle. The measurements show that the SYNCA is generating fields
 3707 with nearly isotropic magnitude across the probed region. The standard deviation of the
 3708 electric field magnitude measured around the antenna circumference is approximately
 3709 2.9 dB for a typical rotational scan. The presence of a significant pattern null is noted
 3710 near 45° (see Figure 5.20), which we attribute to small imperfections in the antenna
 3711 PCB that could be introduced from the hand soldered terminations connecting the coax
 3712 cables to the antenna. There is no significant difference in the radiation pattern when
 3713 measured across the 2 cm vertical range. The measured relative phases closely follow
 3714 the expectation for an electron, being linear with the measurement rotation angle and
 3715 forming the expected spiral pattern. Other than the small phase imperfections there is
 3716 a slight sinusoidal bias to the phase data, which we determined is the result of a small
 3717 ($\lesssim 1$ mm) offset of the antenna's phase center from the rotation axis of the automated
 3718 stages.

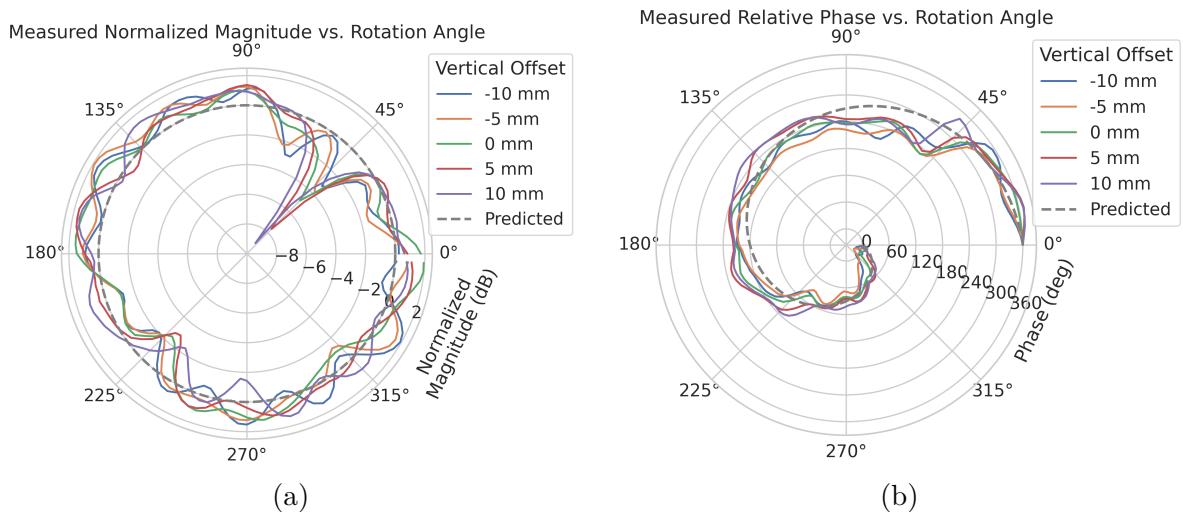


Figure 5.20: Linear interpolations of the measured electric field magnitude (a) and phase (b). The data was acquired using a VNA at 120 points spaced by 3 degrees from 0 to 357 degrees of azimuthal angle. The different color lines indicate the vertical offset of the horn antenna relative to the SYNCA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

3719 The characterization measurements confirm the simulated performance of the SYNCA.
 3720 As expected the fields generated by the antenna are nearly isotropic in magnitude, ϕ -
 3721 polarized, and are linearly out of phase around the circumference of the antenna as

3722 predicted for cyclotron radiation in Section 5.3.2. Small imperfections in the magnitude
 3723 and phase of the antenna are expected, particularly at the antenna's high operating
 3724 frequency of 26 GHz where small geometric changes can have significant impacts on
 3725 electrical properties. However, calibration through careful characterization measurements
 3726 can be used to remove the majority of these pattern imperfections, including the relatively
 3727 large pattern null near 45°, which will allow for the usage of the SYNCA as a test source
 3728 for free-space CRES experiments utilizing antenna arrays. In the next section we use the
 3729 VNA measurements obtained here as a calibration for signal reconstruction using digital
 3730 beamforming.

3731 **5.3.5 Beamforming Measurements with the SYNCA**

3732 Digital beamforming is a standard technique for signal reconstruction using a phased
 3733 array [81]. The SYNCA, since it exhibits the same cyclotron phases as a trapped electron,
 3734 can be used to perform simulated CRES digital beamforming reconstruction experiments
 3735 on the bench-top without the need for the magnet, cryogenics, and vacuum systems
 3736 required by a full CRES experiment. The fields received by the individual elements
 3737 of the antenna array will have phases dependent on the spatial position of the source
 3738 relative to the antennas. Therefore, a simple summation of the received signals will fail
 3739 to reconstruct the signal due to destructive interference between the individual channels
 3740 in the array. However, applying a phase shift associated with the source's spatial position

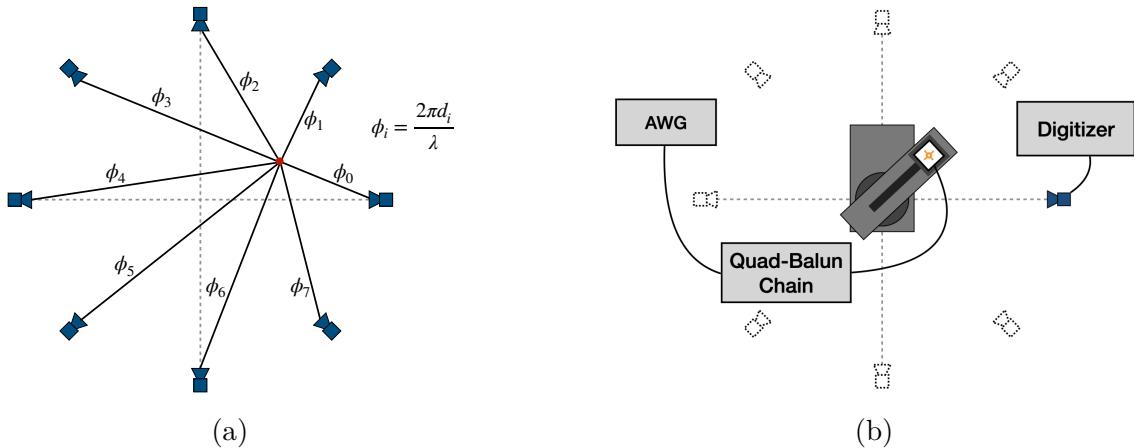


Figure 5.21: (a) A depiction of the relative phase differences for signals received by a circular antenna array from an isotropic source. The phases correspond to a unique spatial position. (b) A schematic of the setup used to perform digital beamforming.

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

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

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

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

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

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

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

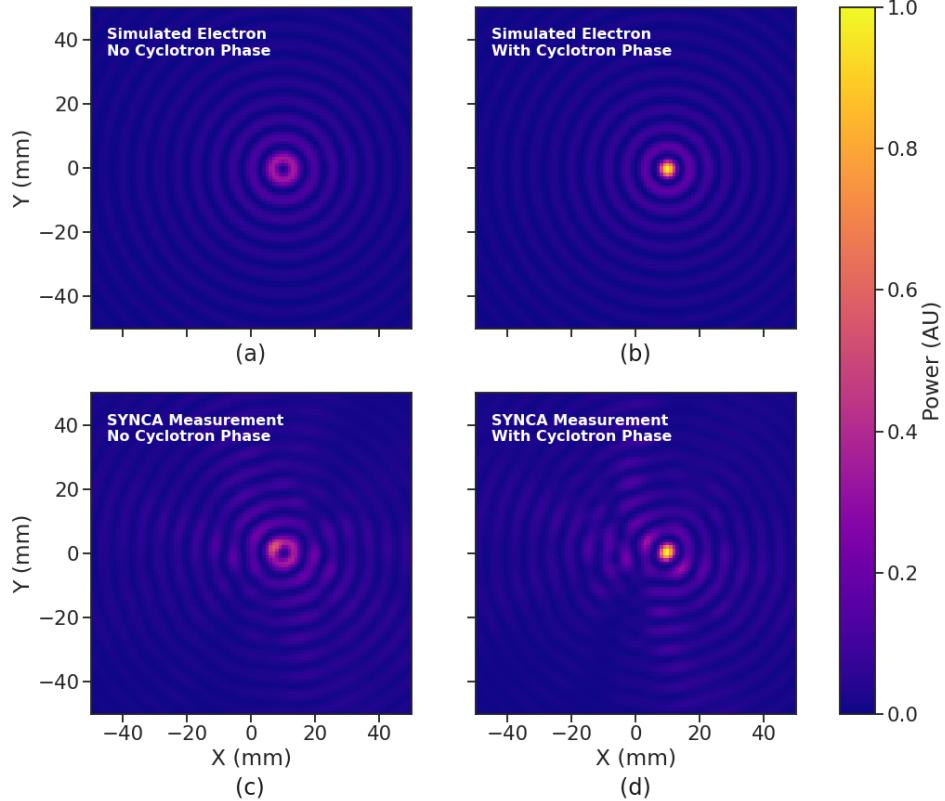


Figure 5.22: Digital beamforming maps generated using a simulated 60 channel array and electron simulated using the Locust package. (a) and (b) show the beamforming maps for simulated electrons without the cyclotron spiral phases and with the cyclotron spiral phases respectively. (c) and (d) show the 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

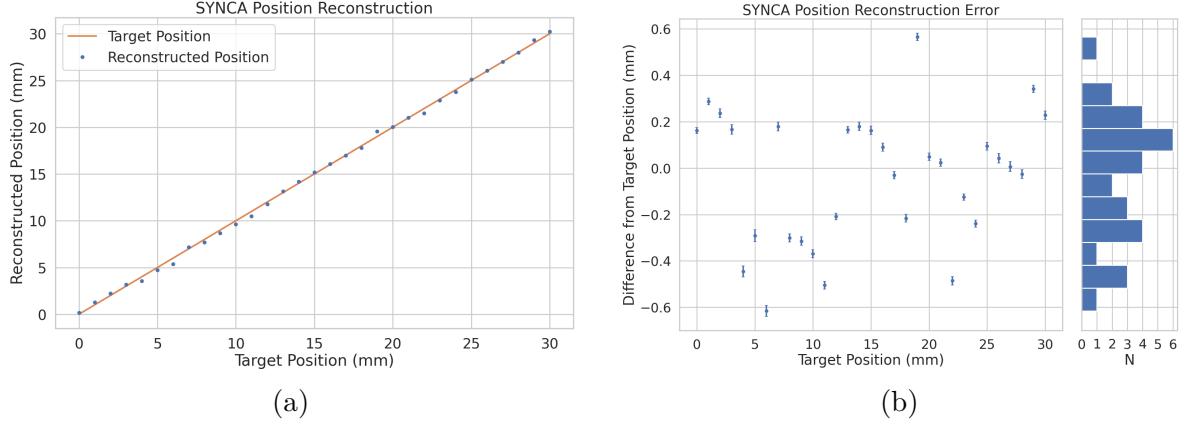


Figure 5.23: A plot of the SYNCA’s reconstructed position using the synthesized horn-antenna array and digital beamforming. (a) Shows the reconstructed position of the SYNCA compared with the target position indicated by the positioning system readout. (b) Shows the reconstruction error, which is the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

3798 **5.3.6 Conclusions**

3799 In this paper we have introduced the SYNCA, which is a novel antenna design that
3800 emits radiation that mimics the unique properties of the cyclotron radiation generated by
3801 charged particles moving in a magnetic field. The characterization measurements of the
3802 SYNCA validated the simulated performance of the PCB crossed-dipole antenna design.
3803 Additionally, the SYNCA was used to estimate the position reconstruction capabilities
3804 of a synthesized array of horn antennas and experimentally reproduced the simulated
3805 digital beamforming reconstruction of electrons.

3806 While the SYNCA performs well, there exist discrepancies in the phase and magnitude
3807 of the radiation pattern compared to the simulated SYNCA design that are related to
3808 the small geometric differences in the soldered connections. Future design iterations that
3809 replace the soldered connections with a fully surface mount design could improve the
3810 radiation pattern at the cost of some complexity and expense. Furthermore, improving
3811 the design of the antenna PCB and mounting system would allow the antenna to be
3812 inserted into a cryogenic and vacuum environment where in-situ antenna measurement
3813 calibrations could be performed.

3814 The discrepancies in the radiation pattern and phases exhibited by the as-built
3815 SYNCA should not greatly impact its performance as a calibration probe. Both magni-
3816 tude and phase variations can be accounted by applying the SYNCA characterization
3817 measurements as a calibration to the data collected by the antenna array test stand. The
3818 separate calibration of the SYNCA radiation does not impact the primary goals for the
3819 antenna array test stand which are array calibration and signal reconstruction algorithm
3820 performance characterization, because it can be performed with standard reference horn
3821 antennas with well understood characteristics.

3822 The SYNCA antenna technology advances the CRES technique by providing a
3823 mechanism to characterize free-space antenna arrays for CRES measurements without
3824 the need for a magnet and cryogenics system, which would be required for calibration
3825 using electron sources. Both the Project 8 collaboration as well as future collaborations
3826 which are developing antenna array based CRES experiments can make use of SYNCA
3827 antennas as an important component of their calibration and commissioning phases.

3828 **5.4 FSCD Antenna Array Measurements with the SYNCA**

3829 **5.4.1 Introduction**

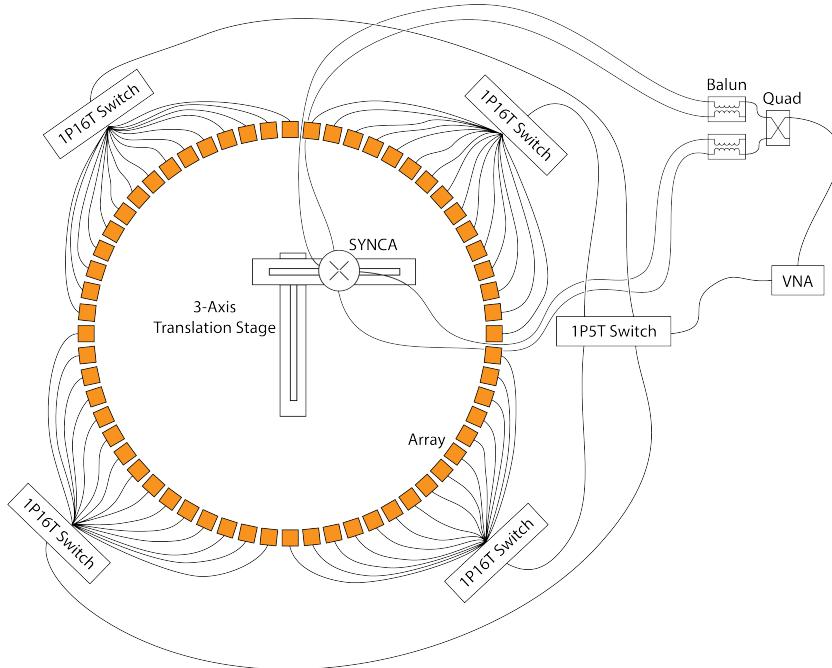


Figure 5.24: A diagram of the array measurement system used to test the prototype FSCD antenna array. A VNA is used as the primary measurement tool, which is connected to the array through a series of switches. The other port of the VNA connects to the SYNCA through the quad-balun chain used to provide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

3830 Using the SYNCA we can perform full-array measurements of prototype versions
3831 of the FSCD antenna array to test its performance with a realistic cyclotron radiation
3832 source (see Figure 5.24). The goal is to check how the measured power received by
3833 the array compares to FSCD simulations as a function of the radial and axial position
3834 of the SYNCA. These measurements are intended to validate the antenna research
3835 and development by Project 8, which has been driven primarily by simulations with
3836 Locust [37] and CREsana (see Section 4.2.3), and identify any discrepancies with these
3837 simulations tools. This knowledge will provide confidence in the simulations necessary
3838 for the analysis of the sensitivity of larger antenna array based CRES experiment designs
3839 to the neutrino mass.

3840 As shown in Section 5.3, the SYNCA does have some radiation pattern imperfections
3841 that complicate the comparison between measurement and simulation data. One way to
3842 disentangle some of the effects of these imperfections is to perform an additional set of
3843 measurements using a synthetic antenna array setup along with the SYNCA antenna.
3844 Since the synthetic array setup uses only a single array antenna, the data should be
3845 free of errors associated with individual antenna differences and multi-path interference,
3846 which are two error sources being tested with the full-array setup. By comparing the
3847 synthetic array data to the FSCD array data and to simulation data one can evaluate the
3848 significance of these effects relative to the errors introduced by SYNCA imperfections.

3849 **5.4.2 Measurement Setups**

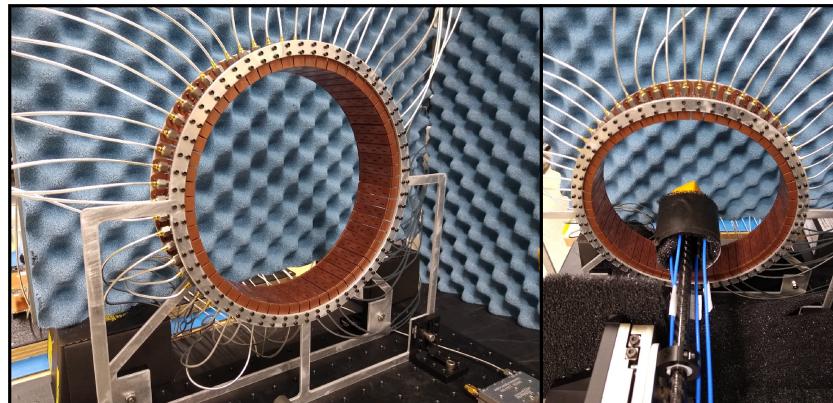
3850 **5.4.2.1 FSCD Array Setup**

3851 The antenna design that composes the array is the 5-slot waveguide antenna developed
3852 for the FSCD experiment (see Figure 5.25a). The antenna is 5 cm long and is constructed
3853 out of WR-34 waveguide with a 2.92 mm coax connector located at the center of the
3854 antenna. Copper flanges located on both ends of the antenna are used to mount the
3855 antenna in the array support structure. The antennas are supported by two circular steel
3856 brackets that can be bolted to both ends of the waveguide to construct the circular array
3857 (see Figure 5.25b). The antenna array consists of sixty identical waveguide antennas
3858 with a radius of 10 cm. The array is mounted perpendicular to an optical breadboard
3859 surface using a pair of the steel brackets, which provide sufficient space for the coaxial
3860 cable connections and allows for easy positioning of the SYNCA antenna. The SYNCA is
3861 mounted on the end of a carbon fiber rod attached to a set of manual translation stages,
3862 which are used to move the SYNCA antenna to different positions inside the array (see
3863 Figure 5.25c). The stages allow for independent motion in three different axes and can
3864 position the SYNCA at radial distances up to 5 cm from the center.

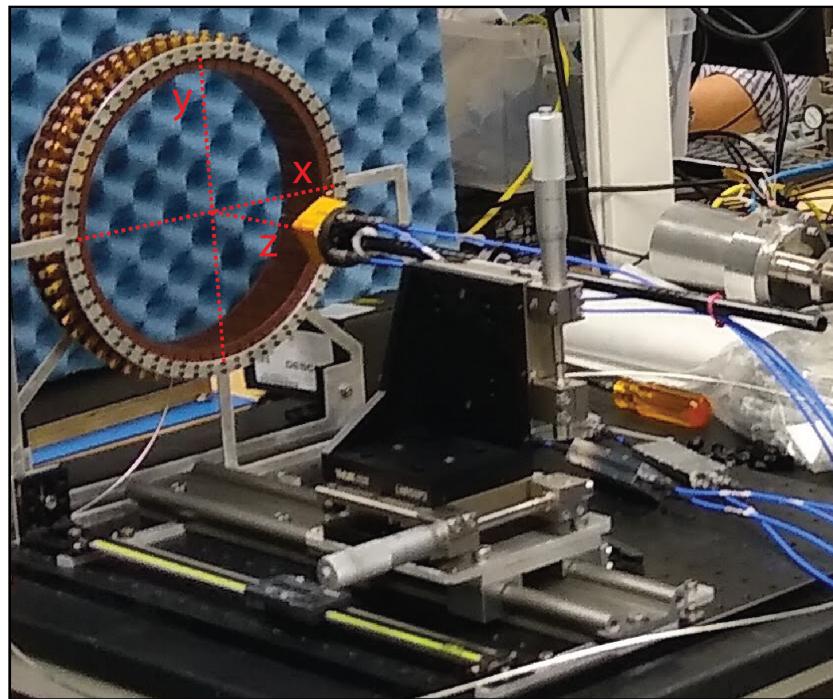
3865 Data acquisition is accomplished using a two-port VNA in combination with a series
3866 of microwave switches that allow the VNA to connect to each channel in the array . The
3867 first port of the VNA is connected to the quad-balun chain used to feed the SYNCA (see
3868 Section 5.3), and the second port of the VNA connects to a 1P5T microwave switch. The
3869 1P5T switch is connected to four separate 1P16T switch boards that connect directly
3870 to the array. The data acquisition is controlled by a python script running on a lab
3871 computer, which is connected to the VNA and an Arduino board programmed to control
3872 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).

3873 antennas in the array to the VNA. The VNA is configured to load a specific calibration
3874 file for each antenna channel and performs the measurements of all available S-parameters.
3875 The separate calibration files is an attempt to remove phase and magnitude errors caused
3876 by different propagation through the RF switches. Array measurements were performed
3877 for the set of SYNCA positions consisting of radial (x-axis) positions from 0 to 50 mm in
3878 5 mm steps and axial (z-axis) positions from 0 to 50 mm in 5 mm steps resulting in 121
3879 array measurements. At each SYNCA position we measured the two-port S-parameter
3880 matrix using a linear frequency sweep from 25.1 to 26.1 GHz with 101 discrete frequencies.

3881 **5.4.2.2 Synthetic Array Setup**

3882 A photograph of the setup used to perform the synthetic array measurements is shown
3883 in Figure 5.26. One important difference between this setup and the FSCD array setup
3884 is that the synthetic array measurements were performed with a waveform generator and
3885 digitizer instead of a VNA. The electronics configuration is identical to the diagram in
3886 Figure 5.7b. Despite the differences, one is still able to compare the measured phases of
3887 the synthetic array and the relative magnitude of the power, since the digitized signal
3888 power is directly proportional to S21.

3889 The arbitrary waveform generator in the setup is configured to produce a 64 MHz
3890 sine wave signal that is up-converted to 25.864 GHz using a mixer and the VNA source.
3891 This signal is passed through a bandpass filter and fed to the SYNCA quad-balun chain.
3892 A single FSCD antenna is positioned 10 cm from the SYNCA and aligned vertically so
3893 that the center of the 5-slot waveguide is in the plane of the SYNCA PCB (see Figure
3894 5.26). This position corresponds to $z = 0$ in Figure 5.25c. The SYNCA is rotated
3895 in three degree steps to synthesize an antenna array with 120 channels. This channel
3896 count is more than could physically fit in a 10 cm radius array, but there is no cost to
3897 over-sampling. Additionally, over-sampling allows for a check of the smoothness of the
3898 antenna array radiation pattern. The signals from the FSCD antenna are down-converted
3899 using the second mixer connected to the VNA source before being digitized at 250 MHz
3900 and saved to disk. Several synthetic array measurement scans were performed by using
3901 the linear translation stage to change the radial position of the SYNCA. In total eight
3902 scans were taken from 0 to 35 mm using a radial position step size of 5 mm.

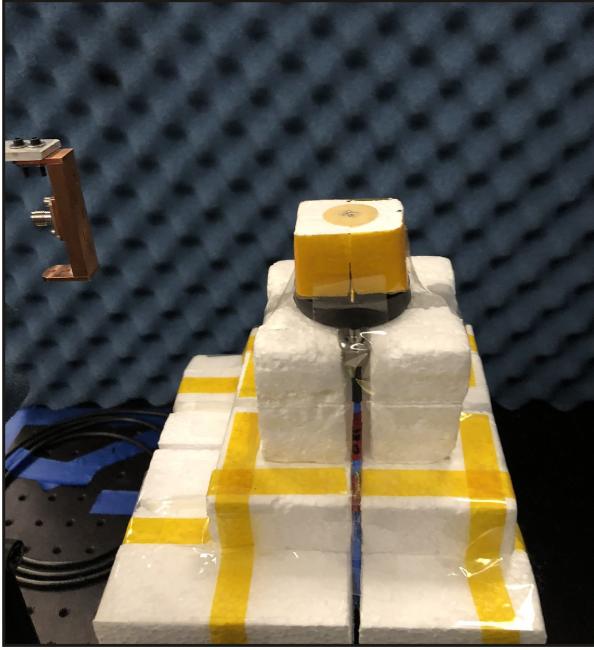


Figure 5.26: A photo of the FSCD antenna and the SYNCA in the synthetic array measurement setup at Penn State.

3903 **5.4.3 Simulations, Analysis, and Results**

3904 The Locust and CRESana simulation packages utilize the antenna transfer functions
3905 to calculate the power that would be received by each antenna from a CRES electron.
3906 The equivalent quantity in the measurement setup is the S21 matrix element, which
3907 indicates the ratio of the power received by an antenna in the array to the amount of
3908 power delivered to the SYNCA. Therefore, the analysis focuses on comparing the relative
3909 magnitudes and phase of the S21 parameters measured by the VNA as a function of
3910 the array channel and the SYNCA position. Additionally, we apply a beamforming
3911 reconstruction to the S21 data to evaluate how the summed power and beamforming
3912 images change as a function of the position of the SYNCA.

3913 **5.4.3.1 Simulations**

3914 Simulations for the FSCD array measurements were performed using CRESana, which
3915 performs analytical calculations of the EM-fields produced by an electron at the position
3916 of the antennas. At each sampled time CRESana computes the electric field vector at the
3917 antenna positions, which is projected onto the antenna polarization axis to obtain the
3918 co-polar electric field. The magnitude of the co-polar electric field is then multiplied by

3919 a flat antenna transfer function to calculate the corresponding voltage signal. CRESana
 3920 simulations exploit the flat transfer functions of the FSCD antennas, which allows the
 3921 electric field to be multiplied by the antenna transfer function rather than performing
 3922 the full FIR calculation. These calculations produce a voltage time-series for each of the
 3923 antennas in the array that can be compared to the laboratory measurements.

3924 CRESana was configured to simulate a 90° electron in a constant background magnetic
 3925 field of ≈ 0.958 T with a kinetic energy of 18.6 keV. These parameters were chosen
 3926 in order to mimic a CRES event near the tritium beta-decay spectrum endpoint in
 3927 the FSCD experiment. The constant background magnetic field guarantees that the
 3928 guiding center of the electron is stationary across the duration of the simulation which is
 3929 consistent with the SYNCA in the laboratory measurements. Simulations were performed
 3930 with the electron's guiding center at radial positions from 0 to 45 mm in steps of 1 mm
 3931 and axial positions from 0 to 30 mm in steps of 1 mm. The simulations generated time
 3932 series consisting of 8192 samples at 200 MHz for the sixty channel FSCD antenna array
 3933 geometry.

3934 5.4.3.2 Phase Analysis

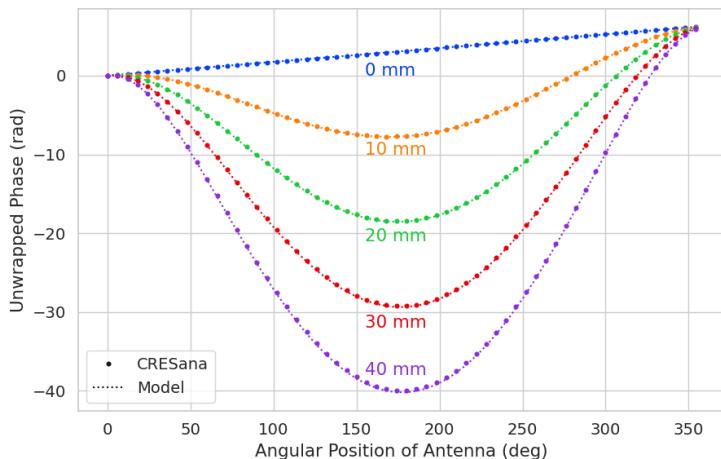


Figure 5.27: The unwrapped phases of signals received by the FSCD antenna array from an electron with a 90° pitch angle located in the plane of the antenna array. The data points indicated the phases extracted from simulation and the dashed lines show the model predictions.

3935 Correct modeling of the signal phases is fundamental to reconstruction for both
 3936 beamforming and matched filter approaches. The beamforming reconstruction relies on

3937 a signal phase model developed from Locust simulations, which allows one to predict the
3938 relative signal phases for a specific magnetic trap and electron position. The equation
3939 for the model is

$$\phi_{ij}(t) = \frac{2\pi d_{ij}(t)}{\lambda} + \theta_{ij}(t), \quad (5.37)$$

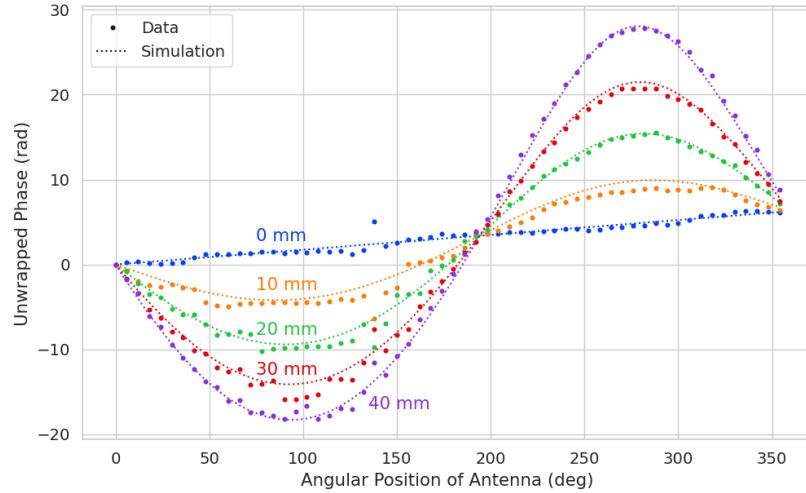
3940 where $d_{ij}(t)$ is distance between the assumed electron position and the antenna position,
3941 and $\theta_{ij}(t)$ is the angular separation between the electron and antenna positions. For
3942 details on the components of the phase model see Section 5.3.2. In Figure 5.27 we
3943 compare the phases predicted by Equation 5.37 to phases extracted from CREsana
3944 simulations of an electron located in the plane of the antenna array at a series of radial
3945 positions. One observes excellent agreement between the model and simulation.

3946 The measured signal phases from the FSCD array and synthetic array are shown
3947 in Figures 5.28a and 5.28b compared to the signal phase model. The axial position of
3948 the SYNCA in both plots is $z = 0$ mm, such that the plane of the PCB is aligned with
3949 the center of the FSCD antenna. The data shown in Figure 5.28a corresponds to the
3950 S-parameters measured at 25.80 GHz which is the frequency closest to the one used in
3951 the synthetic array setup. The different slope and sinusoidal phases exhibited by Figure
3952 5.28a and 5.28b reflects differences in the coordinate system for each setup. In general,
3953 we see that the phase model predicts the large scale features of the phases quite well,
3954 but there are some small scale deviations or errors from the phase model that do not
3955 appear to be present in simulation.

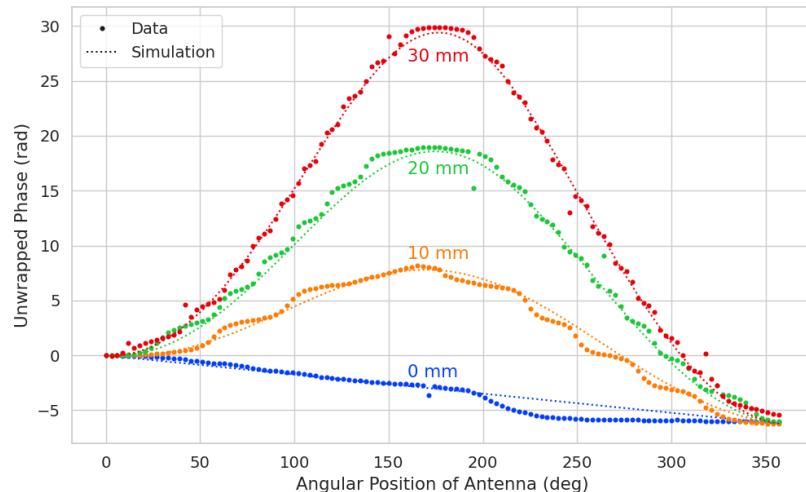
3956 A comparison of the phase errors, which are the difference between measurement and
3957 model is shown in Figure 5.29. The FSCD array data is referred to as the JUGAAD
3958 data in the plot legend, which is an alternative name for the FSCD array setup.

3959 The phase error at $R = 0$ in Figure 5.29 forms a smooth curve, with the exception of
3960 an outlier data point caused by a bug in the data acquisition script. One can attribute
3961 the observed phase error at this position to imperfections in the antenna pattern of the
3962 SYNCA. As the SYNCA is moved away from $R = 0$ mm one observes that the phase
3963 error exhibits oscillations whose frequency increases as a function of the radial position
3964 of the SYNCA. These oscillations have the appearance of a diffraction pattern, which
3965 is particularly clear for the radii ≥ 15 mm, due to the bilateral symmetry of the phase
3966 error peaks around 180° .

3967 One can observe a higher average variance in the phase errors measured for the FSCD
3968 array compared to the synthetic array. This is best seen by comparing the curves at
3969 $R \leq 15$ mm where the smooth synthetic array curves are distinct from the relatively
3970 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.

3971 differences in the radiation patterns of the antennas that make up the array as well as
 3972 differences in the transmission lines through the switch network that introduce additional
 3973 phase errors into the measurement. Since the synthetic array measurements use only
 3974 a single antenna, these extra error terms are not present, which explains the relatively
 3975 smoother phase error curves. Despite the extra phase errors in the FSCD array, it is still
 3976 possible to observe a similar phase error oscillation effect as the SYNCA is moved away

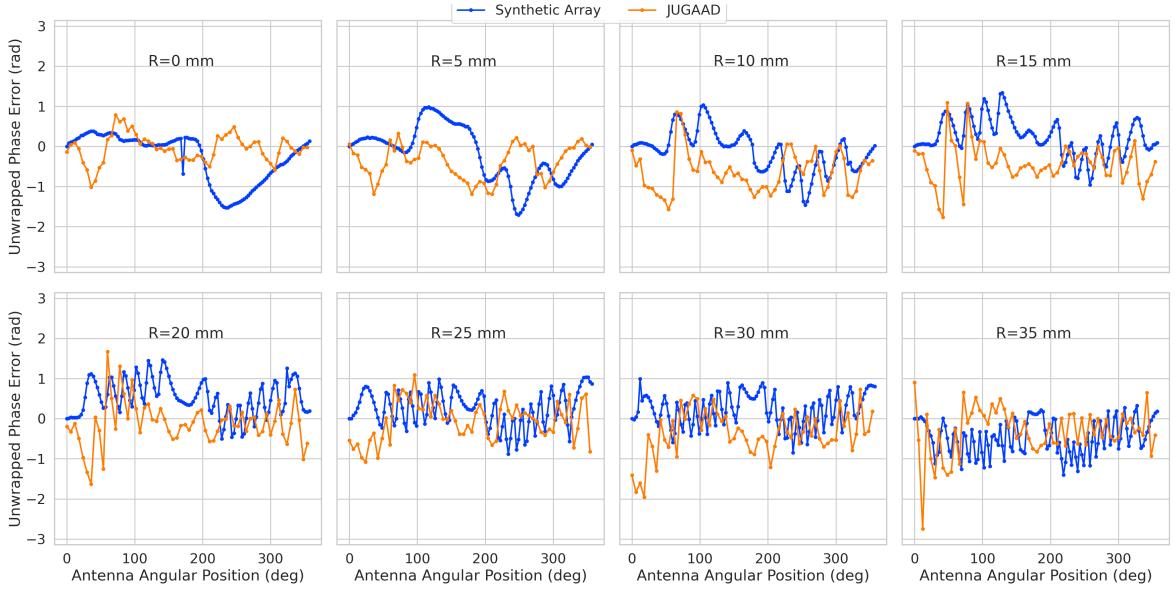
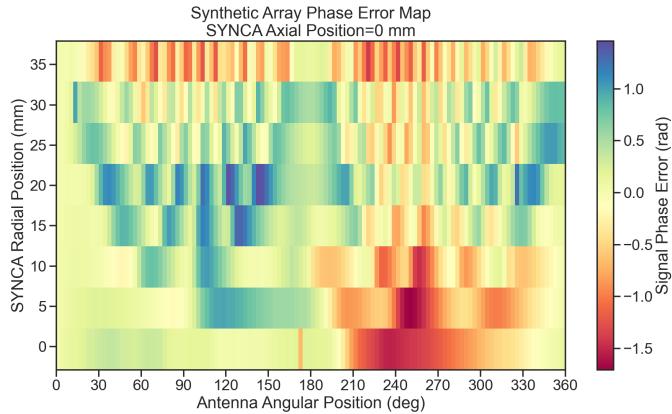


Figure 5.29: The phase errors between the measurement and model for the synthetic array (blue) and the FSCD array (orange) for a series of radial positions. The label JUGAAD refers to an alternative name for the FSCD array setup. As the SYNCA is translated off-axis phase errors with progressively higher oscillation frequency enter into the measurements.

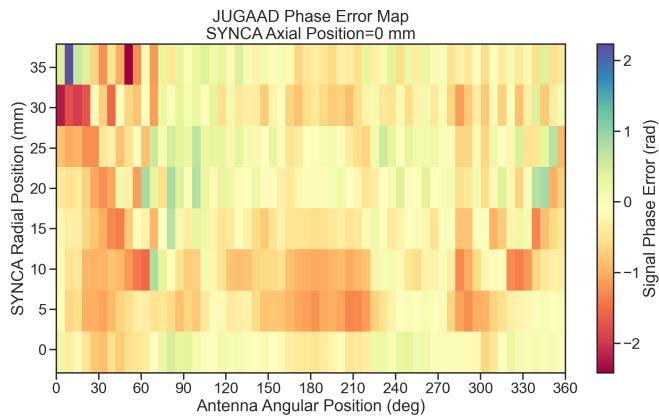
3977 from $R = 0$ mm.

3978 The diffraction pattern exhibited by the phase error oscillations is more easily observed
 3979 by plotting the phase errors in a two-dimensional map, which is done in Figures 5.30a and
 3980 5.30b. For the synthetic array ones observes a relatively clear diffraction pattern
 3981 that emerges as the SYNCA is moved radially. The bilateral symmetry of the diffraction
 3982 patterns is due to the bilateral symmetry of the circular synthetic array around the
 3983 translation axis of the SYNCA. A similar pattern is also visible in the FSCD array data,
 3984 although, it is obscured by the additional phase error that results from the multi-channel
 3985 array.

3986 The physical origin of the phase error diffraction pattern is attributed to interference
 3987 effects arising from path-length differences between the individual slots in the FSCD
 3988 antenna and the SYNCA transmitter. Since we are operating in the radiative near-field of
 3989 the FSCD antenna, the path length differences between the slots introduces a significant
 3990 change in the summation of the signals that occurs inside the waveguide, which causes
 3991 the radiation pattern of the antenna to change as a function of distance. Therefore, when
 3992 the SYNCA is positioned off-axis the different path-lengths from the SYNCA to each
 3993 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.

This near-field effect is not present in simulations, because in order to simplify the calculations we assume that the far-field approximation can be applied to the FSCD antennas. This means that the radiation pattern and antenna transfer functions are independent of the distance between the transmitter and the receiving antenna. In principle, we can account for these near-field effects with a more detailed simulation of the FSCD antennas either in CRESana or Locust, which would result in an additional term in the beamforming phase model. However, this would increase the computational intensity of the simulation software. In the next section we briefly discuss the impact of

4002 these near-field effects on the measured magnitude of the signals.

4003 5.4.3.3 Magnitude Analysis

4004 Exactly as for the signal phase, one can use simulations to construct a model that
4005 describes the magnitude of the signals received by each channel in the antenna array.
4006 By examining the results of simulations or by analyzing the Liénard-Wiechert equation
4007 one can show that radiation pattern from a 90° pitch angle electron in a magnetic field
4008 is omni-directional. Therefore the relative magnitudes of the signals received by each
4009 channel will be determined by the free-space power loss, which is proportional to the
4010 inverse distance between the assumed electron position and the antenna.

4011 A consequence of this is that the signals produced in the array for electrons off the
4012 central axis will have larger amplitudes for the antennas closer to the electron compared
4013 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.

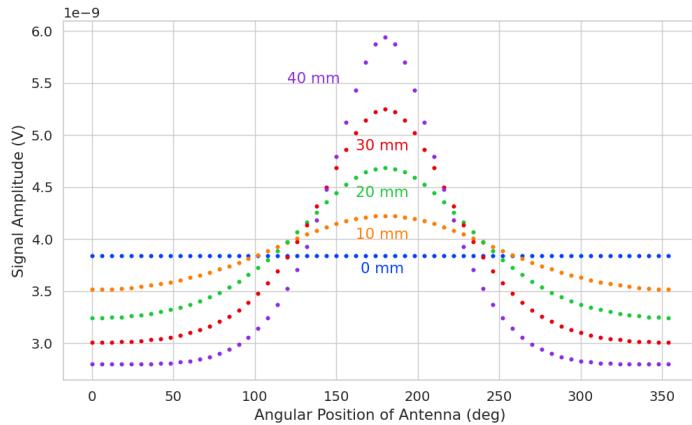


Figure 5.31: The amplitude of the signals from CREsana for the FSCD array from a 90° electron. As the electron is moved from $R = 0$ the signals begin to have unequal amplitudes depending on the distance from the electron to the antenna.

4014
4015 One expects to see a similar trend in the signal magnitudes in both the FSCD and
4016 synthetic arrays. The normalized signal magnitudes extracted from the full and synthetic
4017 array setups for a series of radial SYNCA positions are shown in Figure 5.32. The data
4018 corresponds to a SYNCA axial position of $z = 0$ mm and at a frequency 25.86 GHz. One
4019 complication is that the radiation pattern of the SYNCA is not perfectly omni-directional,
4020 which causes the measured magnitudes at $R = 0$ mm to diverge from the perfectly flat
4021 behavior exhibited by electrons.

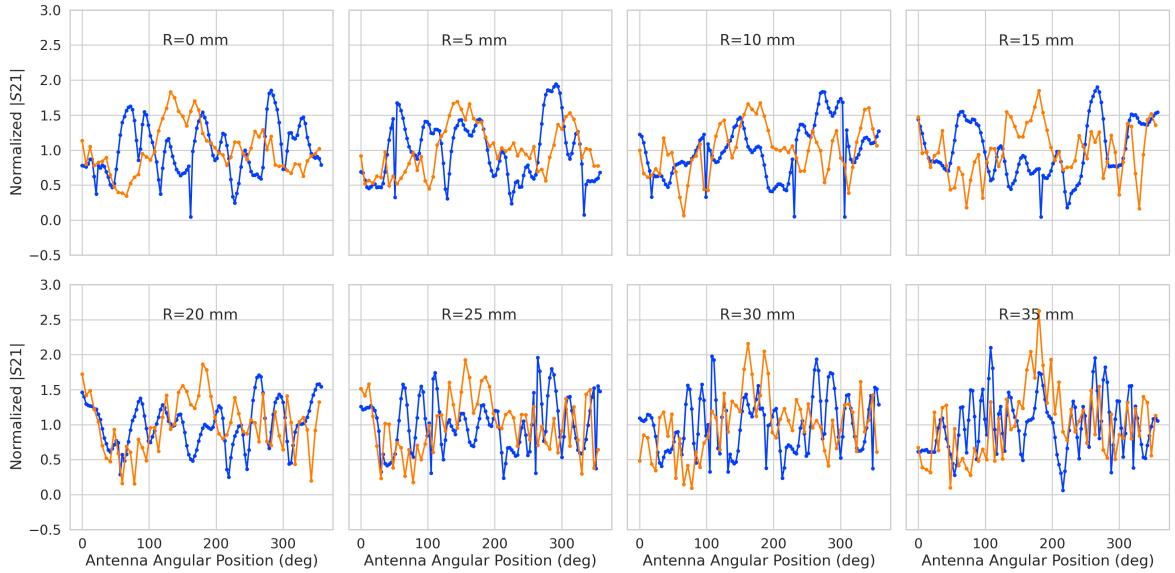
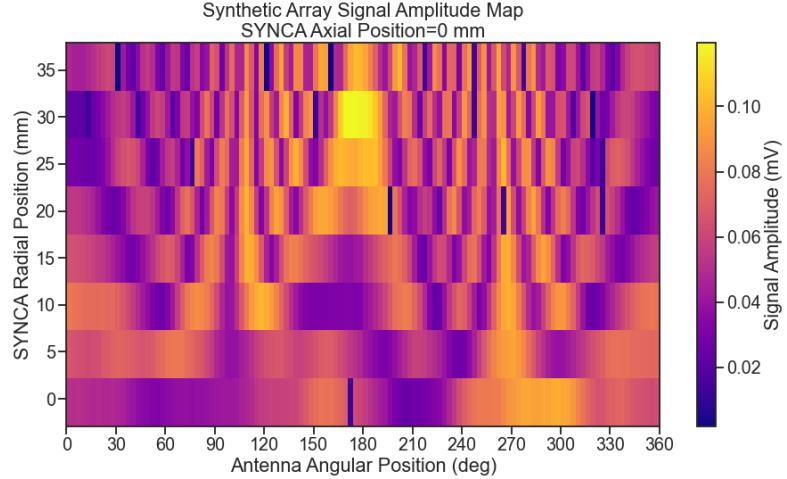


Figure 5.32: The normalized magnitudes of the S21 parameters measured in the FSCD (orange) and synthetic array (blue) setups. The dominant observed behavior as a function of radius is the increase in the number of magnitude peaks, which was noted in the phase error curves. There does not appear to be a strong change in the relative amplitude of a group of antennas as predicted by CREsana.

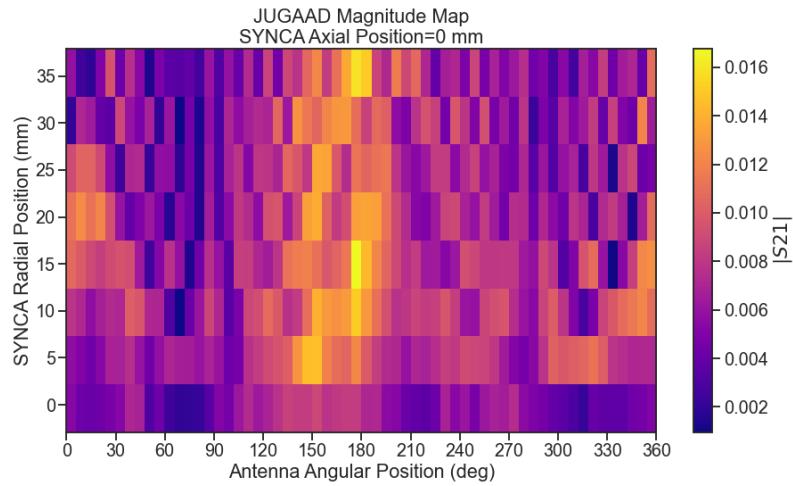
4022 As the SYNCA is moved off-axis one observes a similar increase in the number of
 4023 magnitude peaks in the synthetic array data that one would expect from a diffraction
 4024 pattern, although this trend is not as stark compared to the phase data. Noticeably,
 4025 there does not appear to be a set of channels with disproportionately larger amplitude at
 4026 large R , which would be expected based on the trends from CREsana.

4027 Comparing the magnitudes of the synthetic array to the FSCD array in Figure 5.32
 4028 we see that there is a similar amount of variability in the magnitudes at $R = 0$ mm,
 4029 although there is potentially more small scale error in the magnitude curve caused by
 4030 channel differences in the FSCD array. We observe a similar trend in the number of
 4031 magnitude error peaks in the FSCD array data to the synthetic array data, which mirrors
 4032 the diffraction effect observed in the phase data. The diffraction effect can be visualized
 4033 more clearly by plotting a similar two-dimensional map of the magnitudes (see Figure
 4034 5.33).

4035 The fact that one observes a similar diffraction pattern in the signal magnitudes
 4036 as a function the SYNCA position reinforces the conclusions from the phase analysis
 4037 that near-field effects are having a significant impact on the radiation pattern of the
 4038 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

4039 radiation pattern of the FSCD antenna as a function of distance. If left uncorrected these
 4040 errors reduce detection efficiency by causing power loss in the beamforming or matched
 4041 filter reconstruction due to phase mismatch. We explore the impact of these phase and
 4042 magnitude errors on beamforming in the next section.

4043 **5.4.3.4 Beamforming Characterization**

4044 Errors in the signal magnitudes and phases lead to errors in signal reconstruction. For
 4045 example, a matched filter reconstruction requires accurate knowledge of the signals in

4046 each channel to achieve optimal performance. Uncorrected errors leads to mismatches
 4047 between the template and signal, which reduces detection efficiency and introduces
 4048 uncertainty in the parameter estimation. In this section, we analyze the beamformed
 4049 signal amplitude as a function of the position of the SYNCA to quantify the impact of
 4050 the phase and magnitude errors on signal reconstruction. Because of the imperfections
 4051 in the SYNCA source, it is inappropriate to directly compare the beamformed signal
 4052 amplitude of the FSCD array or synthetic array. Such a comparison would not allow
 4053 one to disentangle losses that occur because of the antenna array from those that occur
 4054 because of the source. Therefore, we focus on comparing the beamforming of the FSCD
 4055 array to the synthetic array.

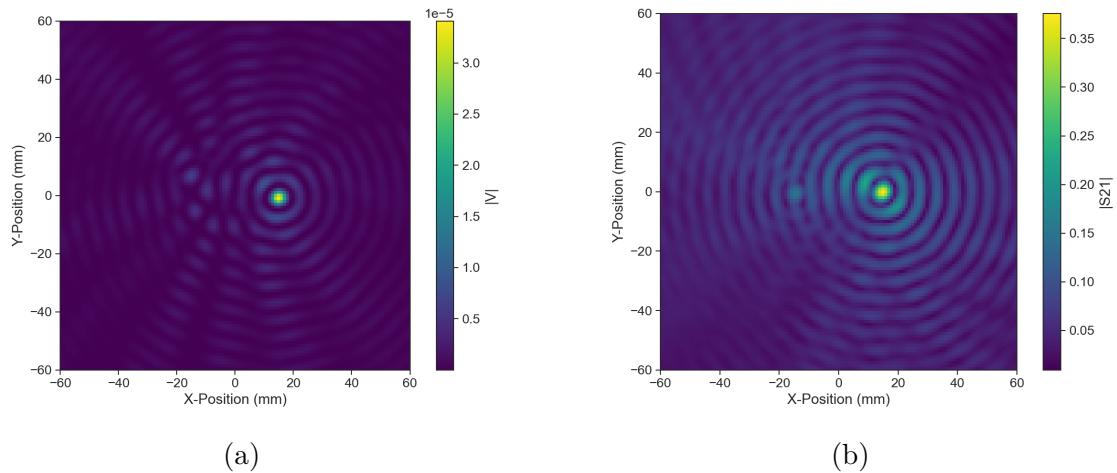


Figure 5.34: Beamforming images from the synthetic array (a) and FSCD array (b) setups with the SYNCA positioned 15 mm off the central axis. In both images we see a clear maxima that corresponds to the true SYNCA position. However, in the FSCD array there is an additional faint peak located at the opposite 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.

4056 The first method of comparison is to analyze the images generated by applying the
 4057 beamforming reconstruction specified in Section 4.3.1 to the FSCD and synthetic array
 4058 data (see Figure 5.34). The beamforming grid consisting of a square 121×121 grid
 4059 spanning a range of -60-mm to 60 mm in the x and y dimensions. The beamforming
 4060 images formed from the synthetic array produces a three-dimensional matrix where each
 4061 grid position contains a summed time series. A single beamforming image is formed from
 4062 this data matrix by taking the mean over the time dimension. In the case of the FSCD
 4063 array, the VNA generates frequency domain data such that each grid position contains a
 4064 summed frequency series produced by the VNA sweep. For this data a single image is

4065 formed by averaging in the frequency domain.

4066 There is a clear difference between the synthetic and FSCD array beamforming images,
 4067 which is the additional faint beamforming maxima located directly opposite the maxima
 4068 corresponding to the SYNCA position. The images in Figure 5.34 were generated with
 4069 data collected at a SYNCA radial position of 15 mm, which agrees well with the observed
 4070 beamforming maximum in both images. We observe that the faint beamforming peak is
 4071 located directly opposite of the true beamforming maximum similar to a mirror image.
 4072 Therefore, the origin of this additional feature appears to be reflections between the two
 4073 sides of the circular antenna array that are not present for the synthetic array since only
 4074 a single physical antenna is used.

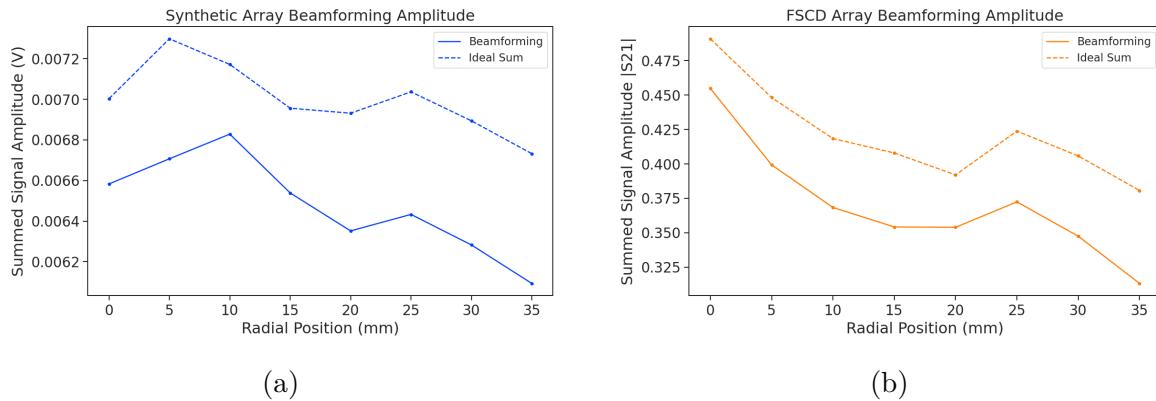


Figure 5.35: A comparison of the maximum signal amplitude obtained by beamforming to the signal amplitude obtained with an ideal summation as a function of the radial position of the SYNCA. The amplitudes for the synthetic array are shown in (a) and the FSCD array are shown in (b). In both setups we 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.

4075 From the beamforming images we extract the maximum amplitude, which we plot
 4076 as a function of the radial position of the SYNCA (see Figure 5.35). The phase errors
 4077 we observed in the FSCD and synthetic arrays leads to power loss at the beamforming
 4078 stage due to phase mismatches between the signals at different channels. This power
 4079 loss can be quantified by comparing the signal amplitude obtained from beamforming to
 4080 the amplitude which would be obtained from an ideal summation. We perform the ideal
 4081 summation by phase shifting each array channel to the same phase and then summing.
 4082 The comparison between the beamforming and ideal sums is shown in Figure 5.35, where
 4083 we observe that both the synthetic and FSCD arrays experience power losses from the
 4084 beamforming summation.

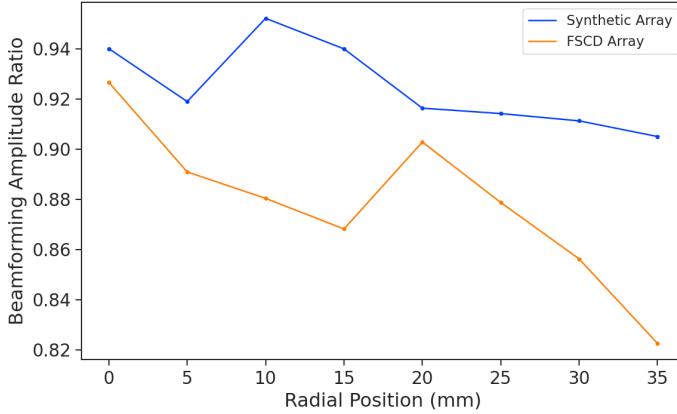


Figure 5.36: The ratio of the beamforming signal amplitude to the ideal signal amplitude for the FSCD and synthetic arrays. We see that the FSCD array has a larger power loss from phase error compare to the synthetic array which indicates that calibration errors associated with the multiple channels as well as reflections are impacting the signal reconstruction.

4085 The beamforming power loss can be quantified using the ratio of the beamforming to
 4086 ideal signal amplitudes. Computing this ratio as a function of SYNCA radial position
 4087 radius for the FSCD and synthetic arrays we find that the FSCD array has a uniformly
 4088 smaller beamforming amplitude ratio, which means that the FSCD array has a larger
 4089 beamforming power loss (see Figure 5.36). The primary contributions to the beamforming
 4090 power loss in the synthetic array are phase errors from the SYNCA and phase errors
 4091 from the FSCD antenna near-field. Both of these phase errors contribute to beamforming
 4092 losses in the FSCD array, but there are clearly additional phase errors in the FSCD array
 4093 measurements contributing to the smaller ratio. Two potential error sources include phase
 4094 differences in the different antenna channels that could not be corrected by calibration as
 4095 well as reflections between antennas in the array. The total effect of these additional phase
 4096 errors is to reduce the beamforming amplitude ratio by about 5% from the beamforming
 4097 ratio of the synthetic array. Therefore, we estimate that if no effort is made to correct
 4098 these phase errors in an FSCD-like experiment, then we expect approximately a 10%
 4099 total signal amplitude loss from a beamforming signal reconstruction.

4100 5.4.4 Conclusions

4101 The estimated power loss of a beamforming reconstruction obtained from this analysis
 4102 provides valuable inputs to sensitivity calculations of a FSCD-like antenna array exper-
 4103 iment to measure the neutrino mass, since it helps to bound systematic uncertainties

4104 from the antenna array and reconstruction pipeline. This power loss lowers the estimated
4105 detection efficiency of the experiment since some of the signal power is lost due to
4106 improper combining between channels and also increases the uncertainty in the electron's
4107 kinetic energy by contributing to errors in the estimation of the electron's cyclotron
4108 frequency.

4109 If these reconstruction losses prove unacceptable there are steps that can be taken
4110 to mitigate their effects. Some examples include the development of a more accurate
4111 antenna simulation approach that can reproduce the observed near-field interference
4112 patterns of the FSCD antennas and the implementation of a calibration approach that
4113 allows for the relative phase delays of the array to be measured without changing or
4114 disconnecting the antenna array configuration.

4115 **Chapter 6 |**

4116 **Development of Resonant Cavities for Large**

4117 **Volume CRES Measurements**

4118 **6.1 Introduction**

4119 The cavity approach was originally an alternative CRES measurement technology under
4120 consideration by the Project 8 collaboration for the Phase IV experiment. After pursuing
4121 an antenna array based CRES demonstrator design for several years the increasing costs
4122 and complexity of the antenna arrays led to a reconsideration of the baseline technology
4123 for the ultimate CRES experiment planned by Project 8. Currently, a cavity based CRES
4124 experiment is the preferred technology choice for future experiments by the Project 8
4125 collaboration including the Phase IV experiment.

4126 In this chapter I provide a brief summary of resonant cavities and sketch out the key
4127 features of a cavity based CRES experiment. In Section 6.2 I provide a brief introduction
4128 to cylindrical resonant cavities and the solutions for the electromagnetic fields in the
4129 cavity volume.

4130 In Section 6.3 I describe the main components of a cavity based CRES experiment,
4131 including the background and trap magnets, cavity geometry and design, and cavity
4132 coupling considerations. I also discuss some relevant trade-offs between an antenna array
4133 and cavity CRES experiment, and highlight some reasons for the transition of Project 8
4134 to the development of a cavity based experiment.

4135 Finally, in Sections 6.4 and 6.5, I present the design and development of an open
4136 mode-filtered cavity that could be used in a cavity based CRES experiment with atomic
4137 tritium. The results of the cavity simulations are confirmed by laboratory measurements
4138 of a proof-of-principle prototype that demonstrates key features of the design.

4139 6.2 Cylindrical Resonant Cavities

4140 Resonant cavities are sealed conductive containers, which allows us to describe the
4141 electromagnetic (EM) fields contained in the cavity volume as a superposition of resonant
4142 modes [63]. The field shapes of the resonant modes are determined by Maxwell's equations
4143 and the boundary conditions enforced by the cavity geometry. Of interest to Project 8
4144 for CRES measurements are cylindrical cavities due to their ease of construction and
4145 integration with atom and electron trapping magnets.

4146 6.2.1 General Field Solutions

4147 Consider a long segment of conducting material with a cylindrical cross-section (see
4148 Figure 6.1). A geometry such as this can be used as a waveguide transmission line to
4149 transfer EM energy from point to point, or, if conducting shorts are inserted on both
4150 ends of the cylinder, the waveguide becomes a resonant cavity.

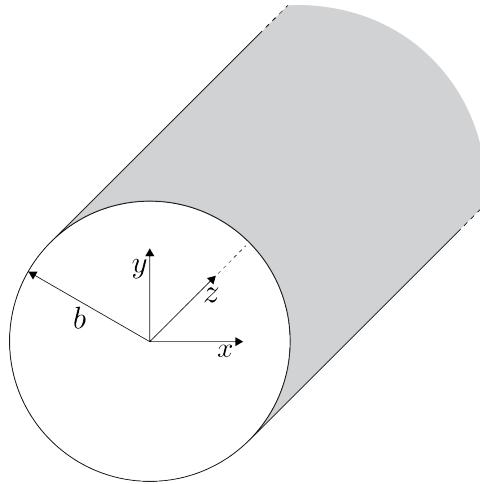


Figure 6.1: Geometry of a cylindrical waveguide with radius b .

4151 The fields allowed inside a cylindrical cavity are determined by the boundary conditions
4152 of the cylindrical geometry. The general approach to solving the fields begins by assuming
4153 solutions to Maxwell's equations of the form

$$\mathbf{E}(x, y, z) = (\mathbf{e}(x, y) + \hat{z}e_z(x, y))e^{-i\beta z}, \quad (6.1)$$

$$\mathbf{H}(x, y, z) = (\mathbf{h}(x, y) + \hat{z}h_z(x, y))e^{-i\beta z}. \quad (6.2)$$

4154 The solutions assume a harmonic time dependence of the form $e^{i\omega t}$ and propagation

4155 along the positive z-axis. The functions $\mathbf{e}(x, y)$ and $\mathbf{h}(x, y)$ represent the transverse
 4156 (\hat{x}, \hat{y}) components of the electric and magnetic fields respectively, and $e_z(x, y)$, $h_z(x, y)$
 4157 represent the longitudinal components. The version of Maxwell's equations in the case
 4158 where there are no source terms can be written as a pair of coupled differential equations,

$$\nabla \times \mathbf{E} = -i\omega\mu\mathbf{H}, \quad (6.3)$$

$$\nabla \times \mathbf{H} = i\omega\epsilon\mathbf{E}, \quad (6.4)$$

4159 where ϵ and μ are the permittivity and permeability of the material inside the waveguide
 4160 or cavity. Using the field solutions from Equations 6.1 and 6.2 one can solve for the
 4161 transverse components of the fields in terms of the longitudinal fields. Because we
 4162 are interested in cylindrical cavities it is advantageous to write the field solutions in
 4163 cylindrical coordinates. After performing this transformation the set of four equations
 4164 for the transverse field components are,

$$H_\rho = \frac{i}{k_c^2} \left(\frac{\omega\epsilon}{\rho} \frac{\partial E_z}{\partial\phi} - \beta \frac{\partial H_z}{\partial\rho} \right), \quad (6.5)$$

$$H_\phi = \frac{-i}{k_c^2} \left(\omega\epsilon \frac{\partial E_z}{\partial\rho} + \frac{\beta}{\rho} \frac{\partial H_z}{\partial\phi} \right), \quad (6.6)$$

$$E_\rho = \frac{-i}{k_c^2} \left(\beta \frac{\partial E_z}{\partial\rho} + \frac{\omega\mu}{\rho} \frac{\partial H_z}{\partial\phi} \right), \quad (6.7)$$

$$E_\phi = \frac{i}{k_c^2} \left(-\beta \frac{\partial E_z}{\partial\phi} + \omega\mu \frac{\partial H_z}{\partial\rho} \right), \quad (6.8)$$

4165 where k_c is the cutoff wavenumber defined by $k_c^2 = k^2 - \beta^2$ with $k = \omega\sqrt{\mu\epsilon}$ being the
 4166 wavenumber of the EM radiation.

4167 This set of equations can be used to solve for a variety of different modes that can be
 4168 obtained by setting conditions on E_z and H_z . For cylindrical cavities two types of modes
 4169 are allowed, which correspond to solutions where $E_z = 0$ and $H_z = 0$ respectively.

4170 6.2.2 TE and TM Modes

4171 The TE family of modes corresponds to the case where $E_z = 0$. This implies that H_z is
 4172 a solution to the Helmholtz wave equation

$$(\nabla^2 + k^2)H_z = 0. \quad (6.9)$$

⁴¹⁷³ 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
⁴¹⁷⁴ the standard technique of separation of variables. Rather than reproduce the derivation
⁴¹⁷⁵ here we shall simply quote the solutions for the transverse fields [63], which are

$$H_\rho = \frac{-i\beta}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.10)$$

$$H_\phi = \frac{-i\beta n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.11)$$

$$E_\rho = \frac{-i\omega\mu n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.12)$$

$$E_\phi = \frac{i\omega\mu}{k_{c_{nm}}} (A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}. \quad (6.13)$$

⁴¹⁷⁶ One can observe that the solutions have a periodic dependence on ϕ , and radial profiles
⁴¹⁷⁷ given by the Bessel functions of the first kind. The integer indices n and m arise from
⁴¹⁷⁸ continuity conditions on the EM fields in the azimuthal and radial directions. For the
⁴¹⁷⁹ TE modes $n \geq 0$ and $m \geq 1$. $k_{c_{nm}}$ is the cutoff wavenumber for the TE_{nm} mode given by

$$k_{c_{nm}} = \frac{p'_{nm}}{b}, \quad (6.14)$$

⁴¹⁸⁰ where b is the radius of the cavity or waveguide and p'_{nm} is the m -th root of the derivative
⁴¹⁸¹ 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 |

⁴¹⁸² The TM mode family corresponds to the case where $H_z = 0$, and $(\nabla^2 + k^2)E_z = 0$.
⁴¹⁸³ Again, we assume solutions of the form $E_z(\rho, \phi, z) = e_z(\rho, \phi)e^{-i\beta z}$, for which the general
⁴¹⁸⁴ form of the solutions is the same as for the TE modes. However, the different boundary
⁴¹⁸⁵ conditions for the TM modes results in particular solutions with a different from, which
⁴¹⁸⁶ we shall quote here without derivation. The transverse fields of the TM modes are given
⁴¹⁸⁷ by

$$H_\rho = \frac{-i\omega\epsilon n}{k_{c_{nm}}^2 \rho} (A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.15)$$

$$H_\phi = \frac{-i\omega\epsilon}{k_{c_{nm}}}(A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z} \quad (6.16)$$

$$E_\rho = \frac{-i\beta}{k_{c_{nm}}}(A \sin n\phi + B \cos n\phi) J'_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.17)$$

$$E_\phi = \frac{-i\beta n}{k_{c_{nm}}^2 \rho}(A \cos n\phi - B \sin n\phi) J_n(k_{c_{nm}}\rho) e^{-i\beta_{nm}z}, \quad (6.18)$$

⁴¹⁸⁸ which one may notice are the same solutions as the TE modes with H and E flipped.
⁴¹⁸⁹ The cutoff wavenumber for the TM modes is given by, $k_{c_{nm}} = p_{nm}/b$, where the values of
⁴¹⁹⁰ 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 |

⁴¹⁹¹ 6.2.3 Resonant Frequencies of a Cylindrical Cavity

⁴¹⁹² A cylindrical cavity is constructed by taking a section of cylindrical waveguide and
⁴¹⁹³ shorting both ends with conductive material. This means that the electric fields inside
⁴¹⁹⁴ 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).

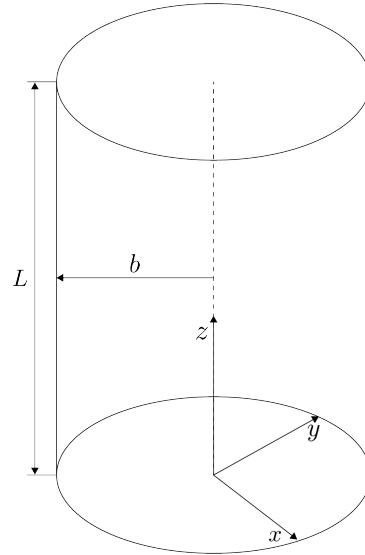


Figure 6.2: The geometry of a cylindrical cavity with length L and radius b .

4195

4196 The transverse electric field solutions for a cylindrical waveguide are of the form

$$\mathbf{E}(\rho, \phi, z) = \mathbf{e}(\rho, \phi) (A_+ e^{-i\beta_{nm}z} + A_- e^{i\beta_{nm}z}), \quad (6.19)$$

4197 where A_+ and A_- are arbitrary amplitudes of forward and backward propagating waves.

4198 In order to enforce that \mathbf{E} is zero at both ends of the cavity we require that

$$\beta_{nm}L = 2\pi\ell, \quad (6.20)$$

4199 where $\ell = 0, 1, 2, 3, \dots$. Using this constraint on the propagation constant we can solve

4200 for the resonant frequencies of the TE_{nml} and the TM_{nml} modes in a cylindrical cavity.

4201 For the TE modes the resonant frequencies are

$$f_{nml} = \frac{c}{2\pi\sqrt{\mu_r\epsilon_r}} \sqrt{\left(\frac{p'_{nm}}{b}\right)^2 + \left(\frac{\ell\pi}{L}\right)^2}, \quad (6.21)$$

4202 and the frequencies of the TM modes are

$$f_{nml} = \frac{c}{2\pi\sqrt{\mu_r\epsilon_r}} \sqrt{\left(\frac{p_{nm}}{b}\right)^2 + \left(\frac{\ell\pi}{L}\right)^2}. \quad (6.22)$$

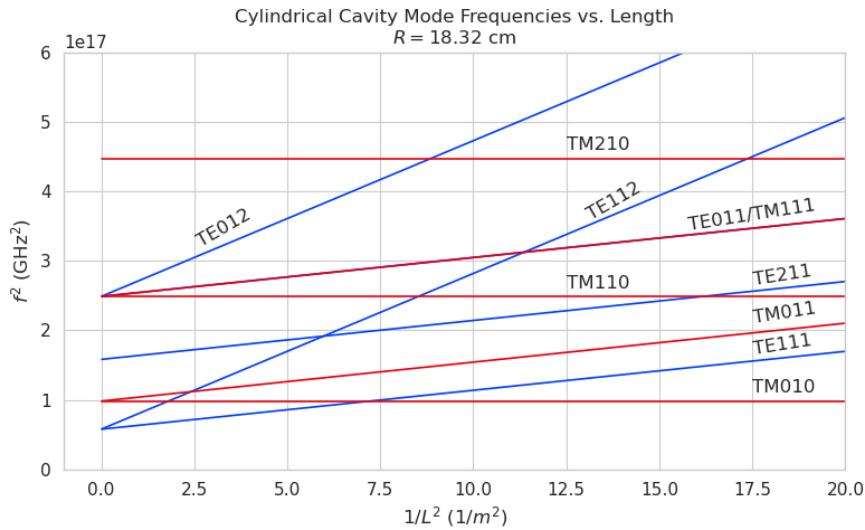


Figure 6.3: Relation of mode frequency to cavity length for a cylindrical cavity with a radius of 18.32 cm.

4203 6.2.4 Cavity Q-factors

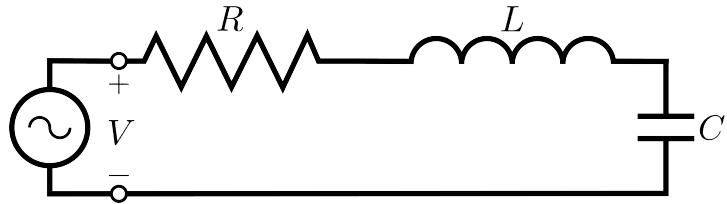


Figure 6.4: A series RLC circuit.

4204 The resonant behavior of cylindrical cavities can be modeled as a series RLC circuit
 4205 (see figure 6.4). The input impedance of the circuit can be obtained by applying
 4206 Kirchhoff's laws to calculate the impedance of the equivalent circuit. For a series RLC
 4207 circuit the input impedance is

$$Z_{\text{in}} = \left(\frac{1}{R} + \frac{1}{i\omega L} + i\omega C \right). \quad (6.23)$$

4208 The resistance in the circuit represents all sources of loss in the cavity, which is primarily
 4209 caused by the finite conductivity of the cavity walls. The inductor and capacitor represent
 4210 the energy stored in the cavity in the form of electric and magnetic fields. If the circuit
 4211 is being driven by an external power source we can write the input power in terms of the
 4212 circuit input impedance and the source voltage

$$P_{\text{in}} = \frac{1}{2} Z_{\text{in}} |I|^2 = \frac{1}{2} |I|^2 \left(\frac{1}{R} + \frac{1}{i\omega L} + i\omega C \right). \quad (6.24)$$

4213 The resistor introduces a loss into the system with a power given by

$$P_{\text{loss}} = \frac{1}{2} |I|^2 R, \quad (6.25)$$

4214 and the capacitor and inductor store energies given by

$$W_e = \frac{1}{4} \frac{|I|^2}{\omega^2 C}, \quad (6.26)$$

$$W_m = \frac{1}{4} |I|^2 L, \quad (6.27)$$

4215 respectively. Using these expressions we can write the input power and input impedance

4216 expressions in terms of the lost power and stored energy

$$P_{\text{in}} = P_{\text{loss}} + 2i\omega(W_m - W_e), \quad (6.28)$$

$$Z_{\text{in}} = \frac{P_{\text{loss}} + 2i\omega(W_m - W_e)}{\frac{1}{2}|I|^2}. \quad (6.29)$$

4217 The condition for resonance in the RLC circuit is that the stored magnetic energy
4218 is equal to the stored electric energy ($W_e = W_m$). When this occurs $Z_{\text{in}} = R$, which is a
4219 purely real impedance, and $P_{\text{in}} = P_{\text{loss}}$. The resonant frequency of the circuit can be
4220 determined from the condition $W_e = W_m$ from which one finds that

$$\omega_0 = \frac{1}{\sqrt{LC}}. \quad (6.30)$$

4221 An important performance parameter for any resonant system is the Q-factor, which
4222 quantifies the quality of the resonator as the ratio of the stored energy multiplied by the
4223 resonant frequency to the average energy lost per second. For the series RLC circuit, the
4224 Q-factor is given by the expression

$$Q_0 = \omega \frac{W_e + W_m}{P_{\text{loss}}} = \frac{1}{\omega_0 RC}, \quad (6.31)$$

4225 from which one observes that as the resistance of the RLC circuit is decreased the quality
4226 factor of the resonator increases. From the perspective of cylindrical cavities this implies
4227 that as one decreases the resistance of the cavity walls it is expected that the Q-factor of
4228 the cavity should increase, which is indeed the case. In certain applications where a high
4229 Q is desireable it is possible to manufacture a cavity out of superconducting materials in
4230 order to minimize the power losses of the system.

4231 The Q-factor of the resonator also determines with bandwidth (BW) of the system.
4232 A cavity with a high Q-factor will resonant with a smaller range of frequencies than a
4233 cavity with a low Q-factor. To see this we can examine the behavior of the RLC circuit
4234 when driven by frequencies near the resonance. For a frequency $\omega = \omega_0 + \Delta\omega$, where
4235 $\Delta\omega = \omega - \omega_0 \ll \omega_0$, we can write the input impedance as

$$Z_{\text{in}} = R + i\omega L \left(\frac{\omega^2 - \omega_0^2}{\omega^2} \right), \quad (6.32)$$

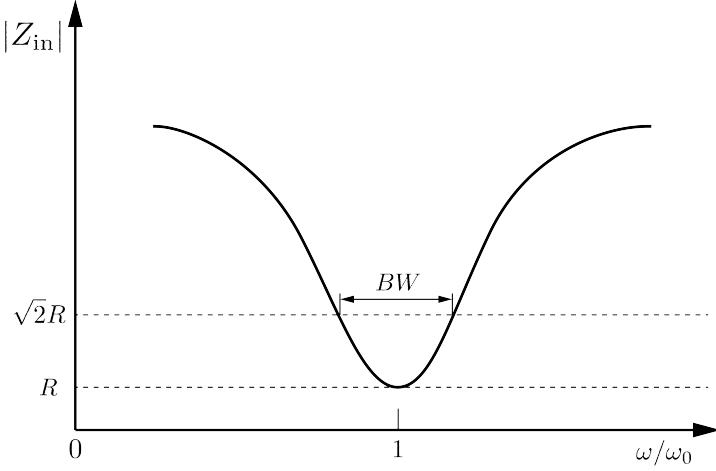


Figure 6.5: Illustration of the behavior of the input impedance of the series RLC circuit as a function of the driving frequency. The BW is proportion to the width of the resonance, which is inversely proportional to Q.

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

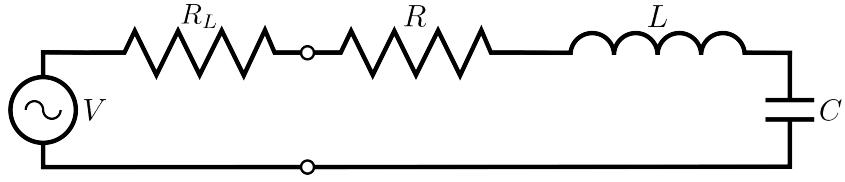


Figure 6.6: A series RLC circuit coupled to an external circuit with input impedance R_L .

external circuit to drive a cavity resonance or to measure the energy of a resonant mode. Coupling a cavity to an external circuit changes the Q by loading the equivalent cavity RLC circuit (see Figure 6.6). The Q-factor of the cavity when it is loaded by an external circuit is called the loaded Q, which is the quantity that one actually measures when exciting a resonance in the cavity. Using the series RLC circuit model one can see that the load resistor in Figure 6.6 will add in series with the resistor in the circuit for a total equivalent resistance of $R + R_L$. Therefore, the loaded Q is given by

$$Q_L = \frac{1}{\omega_0(R + R_L)C}, \quad (6.37)$$

from which one observes that the loaded Q is always less than the intrinsic Q of the cavity.

The amount of coupling that is desireable depends on the specific application of the resonator. If one wants a resonator that is particular frequency selective than it makes sense to limit the amount of coupling to the cavity to maintain a small BW, alternatively, if a larger BW is need one can increase the cavity coupling by tuning the input impedance of the external circuit. The critical point, where maximum power is transferred between the cavity and the external circuit, occurs when the input impedance of the cavity matches the input impedance of the external transmission line. For the series RLC circuit on resonance, this matching condition corresponds to

$$Z_0 = Z_{in} = R, \quad (6.38)$$

where Z_0 is the impedance of the transmission line. The loaded Q at this critical point is, therefore,

$$Q_L = \frac{1}{2\omega_0 Z_0 C} = \frac{Q_0}{2}. \quad (6.39)$$

One can described the degree of coupling between the cavity and an external circuit by

⁴²⁶⁷ defining a coupling factor, g , such that,

$$g = \frac{Q_0}{Q_L} - 1. \quad (6.40)$$

⁴²⁶⁸ When $g = 1$ then $Q_L = Q_0/2$, and the cavity is said to be critically coupled as we
⁴²⁶⁹ described. If $Q_L < Q_0/2$, then the cavity is undercoupled to the transmission line,
⁴²⁷⁰ corresponding to $g < 1$. Alternatively, if $Q_L > Q_0/2$, then $g > 1$, and the cavity is
⁴²⁷¹ overcoupled to the transmission line. Various specialized circuits can be used to tune the
⁴²⁷² input impedance of the external circuit as seen by the cavity to achieve a wide range of
⁴²⁷³ different coupling factors based on the desired application of the cavity.

⁴²⁷⁴ 6.3 The Cavity Approach to CRES

⁴²⁷⁵ 6.3.1 A Sketch of a Molecular Tritium Cavity CRES Experiment

⁴²⁷⁶ Resonant cavities can be used to perform CRES measurements, and they represent the
⁴²⁷⁷ current preferred technology by the Project 8 collaboration. The basic approach to a
⁴²⁷⁸ neutrino mass measurement using a resonant cavity and molecular tritium beta-decay
source is illustrated by Figure 6.7.

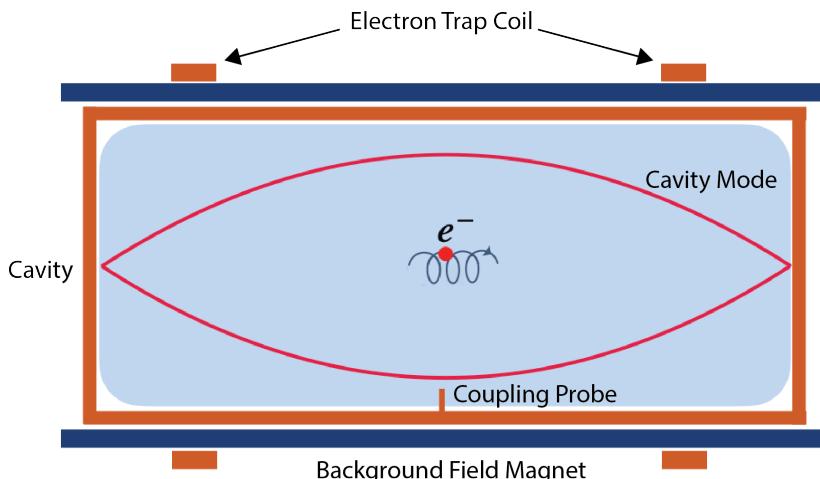


Figure 6.7: A cartoon depiction of a cavity CRES experiment. A metallic cavity filled with tritium gas is inserted into a uniform background magnetic field to perform CRES measurements. Electrons from beta-decays inside the cavity can be trapped and used to excite a resonant mode(s). By coupling to the cavity mode with a suitable probe one can measure the cyclotron frequency of the electron and perform CRES.

⁴²⁷⁹

4280 At the core of the experiment is a large resonant cavity filled with tritium gas. The
4281 filled cavity is then placed in a uniform magnetic field provided by a primary magnet
4282 that provides the background magnetic field. The value of the background magnetic field
4283 sets the range of cyclotron frequencies for electrons emitted near the tritium spectrum
4284 endpoint. When a beta-decay electron is produced in the cavity it is trapped using a set
4285 of magnetic pinch coils that keep electrons inside the cavity volume.

4286 Electrons trapped inside the cavity do not radiate in the same way as electrons
4287 in free-space. Effectively, the same boundary conditions that were used to derive the
4288 resonant modes of a cylindrical cavity in Section 6.2 apply to the radiation of the electron
4289 as well. The coupling of an electron performing cyclotron motion in a cavity has been
4290 studied in detail for measurements of the electron’s magnetic moment [82–84] If an
4291 electron is emitted with a kinetic energy that corresponds to a cyclotron frequency that
4292 matches a resonant frequency of the cavity, then energy radiated by the electron excites
4293 a corresponding resonance in the cavity. The strength of the electron’s coupling to the
4294 cavity is given to first order by the dot product between the electrons trajectory and
4295 the electric field vector of the resonant mode. Additional effects, such as the Purcell
4296 enhancement [85], alter the emitted power from the free-space Larmor equation [25]. If an
4297 electron is moving with a cyclotron frequency that is far from any resonant modes in the
4298 cavity, then radiation from the electron is suppressed. One can interpret this somewhat
4299 surprising effect as the metallic walls of the cavity reflecting the radiated energy back to
4300 the electron.

4301 Detecting an electron in the cavity is accomplished by coupling the cavity to an
4302 external transmission line that leads to an amplifier and RF receiver chain [86]. The
4303 coupling of the cavity resonance to the amplifier occurs through a coupling probe or
4304 aperture designed to read-out the excitation of the mode(s) excited by the electron. For
4305 CRES measurements, the placement of a wire antenna coupling probe inside the cavity
4306 volume leads to unacceptable losses of tritium atoms due to recombination to molecular
4307 tritium on the antenna surface, therefore, apertures are the preferred coupling method
4308 for cavity CRES experiments.

4309 One of the attractive features of the CRES technique for neutrino mass measurement
4310 is the gain in statistics that comes from the differential nature of the tritium spectrum
4311 measurement. Initially, this seems incompatible with cavities, due to the narrow reso-
4312 nances of cavity modes giving relatively small bandwidth. However, by intentionally
4313 over-coupling to a single cavity mode one can achieve bandwidths of a few 10’s of MHz
4314 (see Section 6.2), which is sufficient for a measurement of the tritium spectrum endpoint

4315 region.

4316 **6.3.2 Magnetic Field, Cavity Geometry, and Resonant Modes**

4317 **Magnetic Field and Volume Scaling**

4318 For a CRES experiment, cylindrical cavities are a natural choice since they match
4319 the geometry of standard solenoid magnets, which are needed in order to produce the
4320 background magnetic field for CRES measurements. Furthermore, the cylindrical shape is
4321 compatible with a Halbach array, which is the leading choice of atom trapping technology
4322 for future atomic tritium experiments by the Project 8 collaboration. Cylindrical
4323 cavities also benefit from well-established machining practices that are able to achieve
4324 high geometric precision at large lengths scales. More exotic cavity designs are under-
4325 consideration and there are on-going efforts to investigate the potential advantages these
4326 may have over the standard cylindrical geometry.

4327 As we saw in Section 6.2, the physical dimensions of the cavity are directly coupled
4328 to the resonant frequencies of the cavity. This dependency links the size of the cavity to
4329 the magnitude of the background magnetic field, because the magnetic field determines
4330 the cyclotron frequencies of trapped electrons. Specifically, as the size of the cavity is
4331 increased to accommodate larger volumes of tritium gas, the frequencies of the resonant
4332 modes decrease proportionally. This requires that the magnetic field also decrease in
4333 order to maintain coupling between electrons and the desired cavity mode.

4334 The required cavity size is ultimately determined by the required statistics in the
4335 tritium spectrum endpoint region. Because the gas density must be kept below a certain
4336 level to ensure that electrons have sufficient time to radiate before scattering, larger
4337 volumes become the only way to achieve higher event statistics. To achieve the sensitivity
4338 goals of Phase III and IV cavity volumes on the order of several cubic-meters are required,
4339 which pushes one towards frequencies in the range of 100's of MHz.

4340 **Single-mode Cavity CRES**

4341 It is tempting to consider maintaining a high magnetic field, while still increasing the size
4342 of the cavity, in order to increase the radiated power from trapped electrons for better
4343 SNR. However, if one were to maintain the same magnetic field while increasing the
4344 size of the cavity, the electrons would begin to couple to higher order modes with more
4345 complicated transverse geometries. The danger with this approach is that a complicated
4346 mode structure could introduce systematic errors into the CRES signals. Example

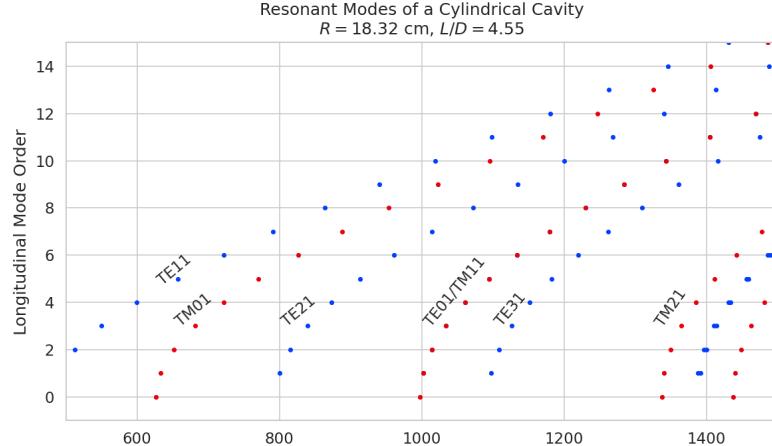
4347 systematics include unpredicted mode hybridization or changes in the mode shapes from
4348 imperfections in the cavity construction, which would prevent reconstruction of the
4349 electron's starting kinetic energies with adequate resolution. For this reason, it is ideal
4350 to operate with magnetic fields that give cyclotron frequencies near the fundamental
4351 frequency of the cavity, where the mode structure is relatively simple (see Figure 6.8).
4352 In this frequency region it is possible to perform CRES by coupling to only a single
4353 resonant mode, however, it is currently an open question if a single mode measurement
4354 will provide enough information about an individual electron's position to reconstruct
4355 the full event. Regardless, developing a solid understanding of the CRES phenomenology
4356 when an electron is coupling to a single mode will be a necessary step towards a future
4357 multi-mode cavity experiment.

4358 Considerations for Resonant Mode Selection

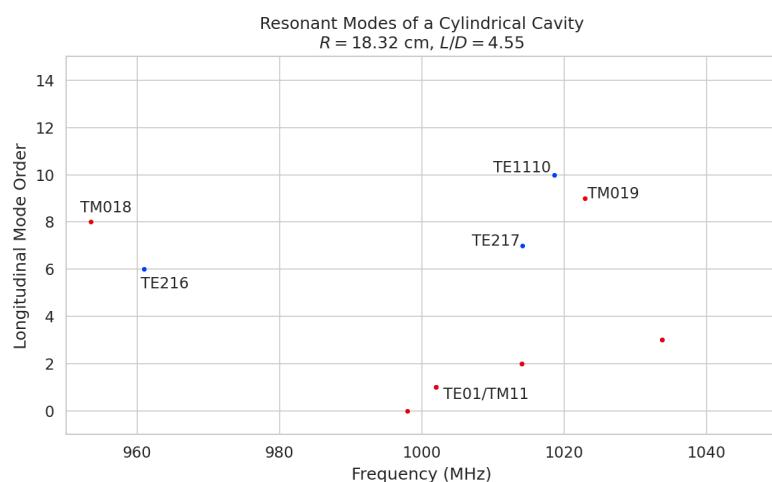
4359 A single-mode cavity experiment begs the question, which resonant mode is best for
4360 CRES measurements? There is an immediate bias towards low order TE_{nm} and TM_{nm}
4361 modes due to the multi-mode considerations discussed above. Additionally, there is a
4362 preference towards modes with longitudinal index $\ell = 1$ with a single antinode along the
4363 vertical axis of the cylindrical cavity. The reason for this is that there is a phase change
4364 in the electric fields between antinodes that leads to modulation effects that destroy the
4365 carrier frequency signal information.

4366 A second consideration for mode selection is the volumetric efficiency of the mode.
4367 Volumetric efficiency can be thought of as an integral over the volume of the cavity
4368 weighted by the relative amplitude of the mode. From the perspective of simply maximiz-
4369 ing the volume useable for CRES measurements this integral would be as close to unity
4370 as possible. However, there is a requirement to reconstruct the position of the electrons
4371 inside the cavity volume so that the local magnetic fields can be used to convert the
4372 measured cyclotron frequency to a kinetic energy. With a single mode this necessarily
4373 requires a variable transverse mode amplitude, which lowers the volumetric efficiency, so
4374 that position of the electron in the cavity can be estimated from the average amplitude
4375 of the CRES signal. Longitudinal indices of $\ell = 1$ have an advantage in volumetric
4376 efficiency over higher order ℓ modes, since there are only two longitudinal nodes, one at
4377 each end of the cavity. Therefore, the average coupling strength of trapped electrons as
4378 they oscillate axially is higher for $\ell = 1$ modes.

4379 The longitudinal variation in the mode strength is ultimately critical for achieving the
4380 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.

4381 the average magnetic fields experienced by electrons with different pitch angles requires
4382 that information on the axial motion of the electron be encoded into the CRES signal.
4383 The longitudinal variation in the mode amplitude leads to amplitude modulation of the
4384 CRES signal with a frequency proportional to the electron's pitch angle.

4385 An additional factor for mode selection is the intrinsic or unloaded Q of the mode. In
4386 terms of SNR it is advantageous to use a mode with a very high Q_0 , which is then highly
4387 overcoupled to achieve the necessary bandwidth to cover the tritium endpoint spectrum.
4388 This scheme leads to a decoupling of the physical cavity temperature from the effective
4389 noise temperature after the amplifier, which allows us to achieve adequate SNR without

4390 the requirement of cooling the entire cavity to single Kelvin temperatures.

4391 An example of a resonant mode that exhibits these traits is the TE₀₁₁ mode. At present
4392 the TE₀₁₁ mode is the preferred resonance for a single-mode cavity CRES experiment
4393 by the Project 8 collaboration. TE₀₁₁ is a low order mode located in a region relatively
4394 far from other cavity modes. Furthermore, the separation of the TE₀₁₁ mode can be
4395 improved by various mode-filtering techniques discussed in Section 6.4.2 below. TE₀₁₁
4396 consists of a single longitudinal antinode that can provide pitch angle information in the
4397 form of amplitude modulation, and has an electric field with a radial profile given by the
4398 J'_0 Bessel function allowing for radial position estimation. Lastly, the TE₀₁₁ mode has a
4399 relatively high intrinsic Q compared to nearby modes, which helps with SNR. Unloaded
4400 Q's greater than 80000 are achievable for a 1 GHz TE₀₁₁ resonance using a copper walled
4401 cavity.

4402 **6.3.3 Trade-offs Between the Antenna and Cavity Approaches**

4403 The choice between cavities and antennas for large-scale CRES measurements is not
4404 without trade-offs. Both the antenna array and cavity approaches are relatively immature
4405 techniques, at present there are no known obstacles that would prevent either approach
4406 from being used for a large scale neutrino mass experiment. The preference for cavities
4407 is largely driven by important practical considerations that could make a cavity based
4408 experiment significantly cheaper than an antenna experiment of similar size and scope.
4409 However, the switch to cavities also introduces new challenges less relevant to the
4410 antenna array, which must be solved in order for Project 8 to achieve its neutrino mass
4411 measurement goals.

4412 One of the major relative drawbacks of the antenna array approach is the size and
4413 complexity of the data-acquisition system. A large-scale antenna array experiment
4414 requires $O(100)$ antennas independently digitized at rates of $O(10)$ to $O(100)$ MHz. Since
4415 there is insufficient information in a single antenna channel to detect or reconstruct the
4416 CRES signal, the entire array output must be processed during the signal reconstruction.
4417 Because data storage becomes an issue with these data volumes, there is a real-time
4418 signal reconstruction requirement that allows one to detect CRES signals buried in the
4419 thermal noise. As we discuss in Section 4.4, the computational cost of these real-time
4420 detection algorithms are potentially quite large for even a small scale antenna array
4421 experiment. However, the operating principle of a cavity experiment allows the CRES
4422 signal to be detected using only a single read-out channel digitized at rates of $O(10)$ MHz,
4423 which reduces the cost of the data acquisition system by many orders of magnitude.

4424 From an engineering perspective, the simple geometry and thin-walls of a cylindrical
4425 cavity are simpler to interface with the cryogenic and magnetic subsystems needed for a
4426 CRES experiment. Whereas, the antenna array requires careful design and engineering
4427 to accommodate the antenna array and receiver electronics in proximity to the trapping
4428 magnets. Additionally, due to near-field interference effects, the antenna array is unable
4429 to reconstruct CRES events within the reactive near-field distance of the antennas.
4430 Because atom trapping requirements require magnetic fields which correspond to cyclotron
4431 frequencies for endpoint electrons less than 1 GHz, the required stand-off distance leads to
4432 a significant loss in useable experiment volume, necessitating larger and more expensive
4433 magnets.

4434 Another advantage to the cavity approach is the relatively compact sideband structure,
4435 which is a result of the low modulation index for cavity CRES signals. The axial motion
4436 in an antenna array experiment leads to frequency modulation and sidebands. The shape
4437 of the sideband structure is determined by the modulation index, $h = \frac{\Delta f}{f_a}$, where Δf
4438 is the size of the frequency deviation and f_a is the axial frequency. The large electron
4439 traps required for a cubic-meter-scale experiment leads to high modulation indices, which
4440 causes the signal spectrum to be made up of numerous low power sidebands that make
4441 reconstruction and detection challenging. This behavior was observed in simulations
4442 of the FSCD in which carrier power decreased with pitch angle due to the increase in
4443 modulation index (see Figure 4.31). For cavities, however, the modulation index remains
4444 near $h = 1$ even for very long magnetic traps due to the high phase velocity in cavities
4445 relative to the axial velocity of the electron. This results in an almost ideal spectrum
4446 shape that has a strong carrier frequency with a few sidebands whose relative amplitudes
4447 encode pitch angle information.

4448 A downside of the cavity approach is the apparent difficulty of estimating the position
4449 of the electron using only the coupling of the electron to a single mode. The amplitude of
4450 the TE₀₁₁ mode is completely independent of the azimuthal coordinate, therefore, position
4451 reconstruction using the TE₀₁₁ mode is only able to estimate the radial position of the
4452 electron. This position degeneracy may lead to magnetic field uniformity requirements
4453 that are too challenging to meet due to mechanical uncertainties in cavity and magnet
4454 construction, as well as uncertainties caused by nuisance external magnetic fields such
4455 as the Earth's field and magnetic fields from building materials. A multi-mode cavity
4456 experiment may provide a way to extract more precise information on the position of
4457 the electron by analyzing the coupling of the electron to several modes that overlap in
4458 different ways.

4459 **6.4 Single-mode Resonant Cavity Design and Simulations**

4460 The single-mode cylindrical cavities envisioned for the Phase III and IV experiments must
4461 be carefully engineered in order to measure the neutrino mass with the desired sensitivity.
4462 In this section I summarize some simulation studies performed to analyze early design
4463 concepts for a single-mode cavity. The primary tool for these investigations was Ansys
4464 HFSS, which was also used for the development of the SYNCA antenna described in
4465 Section 5.3.

4466 **6.4.1 Open Cylindrical Cavities with Coaxial Terminations**

4467 **Design Concept**

4468 A basic cavity design question relevant to Project 8's ultimate goal of an atomic tritium
4469 CRES experiment is how to build a cavity that can be efficiently filled with atomic
4470 tritium. To keep the rate of atom loss from recombination on surfaces it is ideal if the
4471 ends of the cylindrical cavity are as open as possible so that tritium atoms can flow
4472 inside unimpeded. Additionally, one of the primary calibration techniques planned for
4473 future CRES experiments involves CRES measurements using electrons injected from
4474 an electron gun source, which also requires an opening at the cavity end. Cylindrical
4475 cavities with open ends can be manufactured, however, the intrinsic Q-factors of these
4476 cavities are orders of magnitude less than their sealed counterparts, which reduces the
4477 signal-to-noise ratio when that cavity is used for CRES measurement.

4478 Cylindrical cavities with mostly open ends that also exhibit Q values for the $TE_{01\ell}$
4479 modes similar to sealed cavities can be built by using coaxial endcaps to terminate the
4480 cavity. Cavities of this type have been manufactured for specialized applications related
4481 to the measurements of the dielectric constants of liquefied gasses (see Figure 6.9) [87, 88].
4482 This cavity design leaves the ends of the cavity wide open, but retains high Q-values for
4483 the $TE_{01\ell}$ modes due to the coaxial endcap, which are designed to perfectly reflect the
4484 electric fields of $TE_{01\ell}$ modes. Coupling to the $TE_{01\ell}$ mode is achieved via an aperture
4485 located at the center of the cavity wall.

4486 A cavity similar to Figure 6.9 is a candidate design for the future CRES experiments
4487 by Project 8, since it appears to elegantly solve many practical issues that arise when
4488 combining cavity CRES and atomic tritium. The coaxial endcaps leave significant regions
4489 of the cavity ends completely open, which allows for the entrance of atomic tritium as
4490 well as the pumping away of molecular tritium that has recombined on the cavity walls.

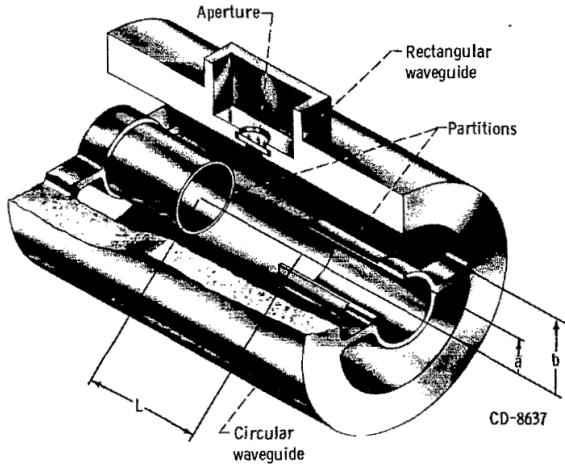


Figure 6.9: An image of an open cavity with coaxial terminations used for dielectric constant measurements. Figure from [88].

4491 These open ends are achieved while preserving the high Q-values of the $TE_{01\ell}$ modes,
 4492 which is important for extracting as much signal power from the electron as possible. In
 4493 subsequent sections we shall analyze this cavity design in more detail, primarily by using
 4494 HFSS simulations to analyze the resonant mode structure of this cavity geometry.

4495 Coaxial Terminator Constraints

4496 The reason that coaxial endcaps can be used to achieve high Q-values for the $TE_{01\ell}$
 4497 modes is that the electric fields for these modes are purely azimuthally polarized (see
 4498 Equations 6.12 and 6.13). Therefore, the boundary conditions that require the electric
 4499 field to go to zero at the cavity ends can be supplied using a coaxial partition of the
 4500 correct radius (see Figure 6.10). Because the cylindrical shape enforced by the partition
 4501 does not match the boundary conditions of other cavity modes, these terminations also
 4502 significantly suppress the Q-factors of non- $TE_{01\ell}$ modes, which is potentially beneficial
 4503 for a single-mode cavity CRES experiment.

4504 The correct radius of the cylindrical partition is derived by setting up the boundary
 4505 value problem in Figure 6.10, and analyzing the reflection and transmission coefficients
 4506 for waves incident on the coaxial terminators. The basic problem is to identify the radius
 4507 a where the reflection coefficient for the $TE_{01\ell}$ modes becomes equal to 1. One can show
 4508 that if the coaxial partitions are made sufficiently long relative to the wavelength of the
 4509 TE_{01} modes than perfect reflection can be achieved. This derivation is quite lengthy
 4510 and complex and is presented in full in [87]. Here, we shall simply explain the resulting

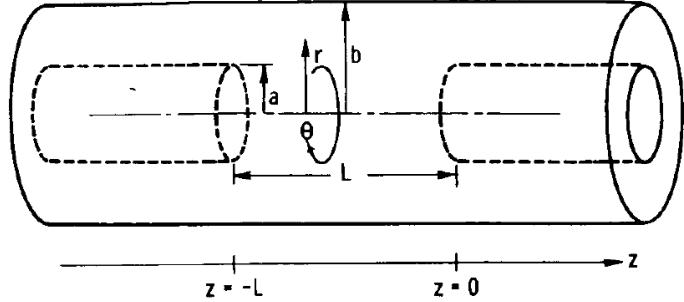


Figure 6.10: The simplified geometry of an open cavity with coaxial terminations. Figure from [87].

4511 conditions on the partition radius for perfect reflection.

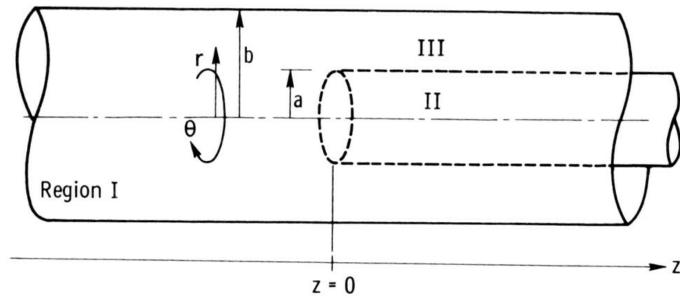


Figure 6.11: Electric field regions for the open cavity boundary value problem. Figure from [87].

4512 The open cavity boundary value problem is solved by expressing the forms of the
 4513 electric fields in the different regions of the cavity and requiring that the electric fields are
 4514 continuous. There are effectively three distinct regions in the open cavity corresponding
 4515 to the central cavity volume, the inner coaxial volume, and the outer coaxial volume (see
 4516 Figure 6.11).

4517 In Region I, the boundary conditions are those of a cylindrical waveguide, and we
 4518 require that E_ϕ for the TE_{0m} modes go to zero at the cavity wall ($r = b$). This requires
 4519 that $J'_{0m}(k_{c0m} b) = 0$. We aim to solve for the radius a in the specific situation where the
 4520 TE_{01} mode can propagate but all other TE_{0m} modes are below the cutoff frequency for
 4521 the circular waveguide. This is equivalent to requiring

$$3.832 < k_{c0m} b < 7.016, \quad (6.41)$$

4522 where the numbers 3.832 and 7.016 correspond to the first and second zeros of the Bessel

4523 function (see Table 6.1).

4524 In Region II the boundary conditions are those of a cylindrical waveguide, but with
4525 a smaller radius. The condition that $E_\phi = 0$ at the cylindrical partition radius is that
4526 $J'_{0m}(k_{c0m}a) = 0$. To ensure perfect reflection, we want all modes in Region 1 of the cavity
4527 to be below the cutoff frequency of the circular waveguide formed by the inner volume of
4528 the coaxial terminator. Therefore, we consider the solutions where

$$k_{c0m}a < 3.832. \quad (6.42)$$

4529 Finally, in Region III the boundary condition are those of a coaxial waveguide. We
4530 need to guarantee that $E_\phi = 0$ at both $r = b$ and $r = a$, which involves finding the
4531 eigenvalues of the following equation

$$J'_0(k_{c0m}a)Y'_0(k_{c0m}b) - J'_0(k_{c0m}b)Y'_0(k_{c0m}a) = 0, \quad (6.43)$$

4532 where Y'_0 the zeroth-order derivatives of the Bessel function of the second kind. The
4533 solutions to this equation depend on the value of the ratio b/a . The approximate solution
4534 is given by

$$\delta_n a \simeq \frac{n\pi}{b/a - 1}, \quad (6.44)$$

4535 where δ_n are eigenvalues of Equation 6.43. Similar to Region II, we are interested in
4536 solutions for which the TE₀₁ modes of Region I are below the cutoff of Region III.
4537 Therefore, we require that

$$k_{c0m} < \delta_1. \quad (6.45)$$

4538 In general, one has some freedom in specifying the value of b/a . A value typically used
4539 in practice is $b/a = 2.082$, which corresponds to positioning the radius of the cylindrical
4540 partition at the maxima of the TE₀₁ electrical fields.

4541 Using the constraints from the three field regions one can develop a coaxial terminator
4542 that acts as a virtual perfectly conducting surface for the TE₀₁ modes. The only required
4543 inputs are the desired frequency of the TE₀₁₁ mode and a choice for the value of b/a .

4544 6.4.2 Mode Filtering

4545 The general case of an electron coupling to a resonant cavity is complicated. This is
4546 because cavities contain an infinite number of resonant modes, which for higher order
4547 modes, have couplings to the electron with a complex spatial dependence. The danger is

4548 that improper modeling of the electron's coupling to the cavity can lead to systematic
4549 errors in the CRES measurements that prevent a high-resolution measurement of the
4550 electron's kinetic energy. This in part drives the preference for a single-mode cavity
4551 experiment that uses only the electron's coupling to the TE₀₁₁ mode to perform CRES,
4552 assuming that sufficient information on the electron's position can be obtained with a
4553 single mode.

4554 The TE₀₁₁ mode is in a region where there are relatively few other modes to which
4555 the electron could couple(see Figure 6.8). However, one can see that the frequency of
4556 the TE₀₁₁ is perfectly degenerate with the TM₁₁₁ mode, which means that electrons will
4557 inevitably couple to both modes if they have the correct cyclotron frequency.

4558 The magnitude of the impact of the electron coupling to both TE₀₁₁ and TM₁₁₁ is
4559 currently unknown. To first order an electron coupling to more both modes will lose more
4560 energy overtime, which can be measured by observing the frequency chirp rate of the
4561 signal. This effect may be small enough to be negligible or simple enough to model that
4562 the cavity can be treated as an effective single-mode cavity. Alternatively, the one could
4563 consider devising a coupling scheme that is sensitive to both the TE₀₁₁ and the TM₁₁₁
4564 modes. By measuring the coupling of the electron to both modes more information on
4565 the position of the electron could be obtained, which could improve the position and
4566 energy resolution of the CRES measurements.

4567 A different approach is the mode filtering approach, which seeks to obtain a single
4568 TE₀₁₁ mode cavity using perturbations to the cavity walls that selectively impede the
4569 TM modes, while leaving the TE modes mostly unperturbed. The type of perturbations
4570 required can be determined by visualizing the surface currents induced in the cavity
4571 walls by each type of mode (see Figure 6.12). By definition, all TM have electric fields
4572 directed along the vertical axis of the cylindrical cavity, which means that perturbations
4573 that impede currents in this direction will modify TM resonances. On the other hand,
4574 the TE₀₁ modes induce azimuthal currents in the cavity walls, therefore, it is possible to
4575 break the degeneracy between TE₀₁ and TM₁₁ using a cavity perturbation that impedes
4576 axial currents, but does not affect the flow of azimuthal currents.

4577 Figure 6.12 shows two cavity design concepts that achieve this selective current
4578 perturbation. The resistive approach inserts a series of thin dielectric rings into the walls
4579 of the cavity that introduces a resistive and capacitive impedance to the longitudinal
4580 currents, while leaving azimuthal current paths intact. Cavities of this type with high
4581 TE₀₁ Q's have also been constructed by tightly wrapping a thin, dielectric coated wire
4582 around a mold to form the cavity wall. An alternative method is to introduce an inductive

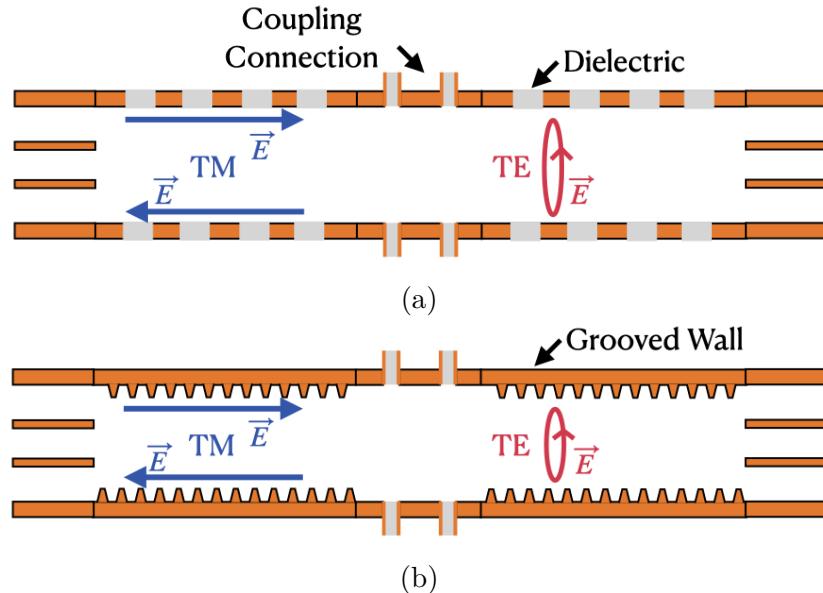


Figure 6.12: Two mode filtering concepts to break the degeneracy of TE_{01} and TM_{11} modes. The resistive approach uses dielectric materials to impede currents that travel vertically along the cavity while leaving azimuthal currents unperturbed. An alternative approach is to impede the currents using grooves cut into the cavity wall, which achieve the same effect with an inductive impedance.

4583 impedance by cutting grooves or a thread pattern on the inside wall of the cavity. For
 4584 reasons of manufacturability and compatibility with tritium the grooved cavity approach
 4585 is the preferred method for mode-filtered cavity construction by Project 8.

4586 6.4.3 Simulations of Open, Mode-filtered Cavities

4587 A candidate design for a single TE_{011} mode CRES experiment is a cavity that utilizes
 4588 the coaxial terminations combined with a mode-filtering wall. The first step towards
 4589 validating that a cavity that combines these two design features will operate as expected
 4590 is a thorough simulation effort for which finite element method (FEM) simulation software
 4591 is invaluable. The primary tool for electromagnetic FEM calculations inside Project 8 is
 4592 Ansys HFSS, which has a robust and well-established eigenmode solver that can identify
 4593 the resonant frequencies and associated Q-factors for given structure.

4594 Four variations of a cavity design with a ~ 1 GHz TE_{011} resonance were implemented
 4595 in HFSS (see Figure 6.13). The four designs include a standard cylindrical cavity, an
 4596 open cavity with smooth walls, an open cavity with resistive walls, and an open cavity
 4597 with grooved walls. The relevant design parameters are summarized in Table 6.3. All

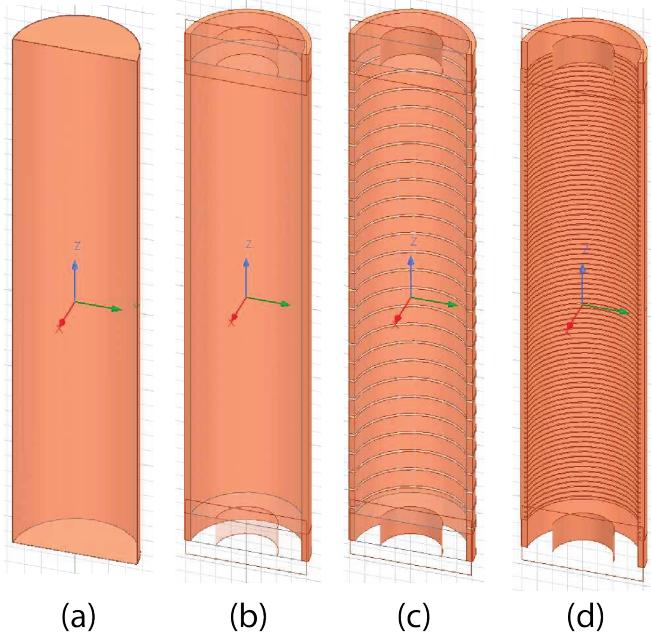


Figure 6.13: Four cavity design variations. (a) is a standard sealed cylindrical cavity, (b) is an open cavity with smooth walls, (c) is an open cavity with resistive walls, and (d) is an open cavity with grooved walls. The main cavity and coaxial terminator parameter are identical for all four cavities.

4598 cavities were simulated using copper walls and filled with a vacuum dielectric. The
 4599 identities of the resonant modes found by HFSS were validated by visual inspection of
 4600 the electric and magnetic field patterns and by comparison to analytical calculations of
 4601 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 |

4602 The results of the HFSS simulations validate our predictions of the resonant behavior
 4603 of an open, mode-filtered cavity developed in the preceding sections (see Figure 6.14) One

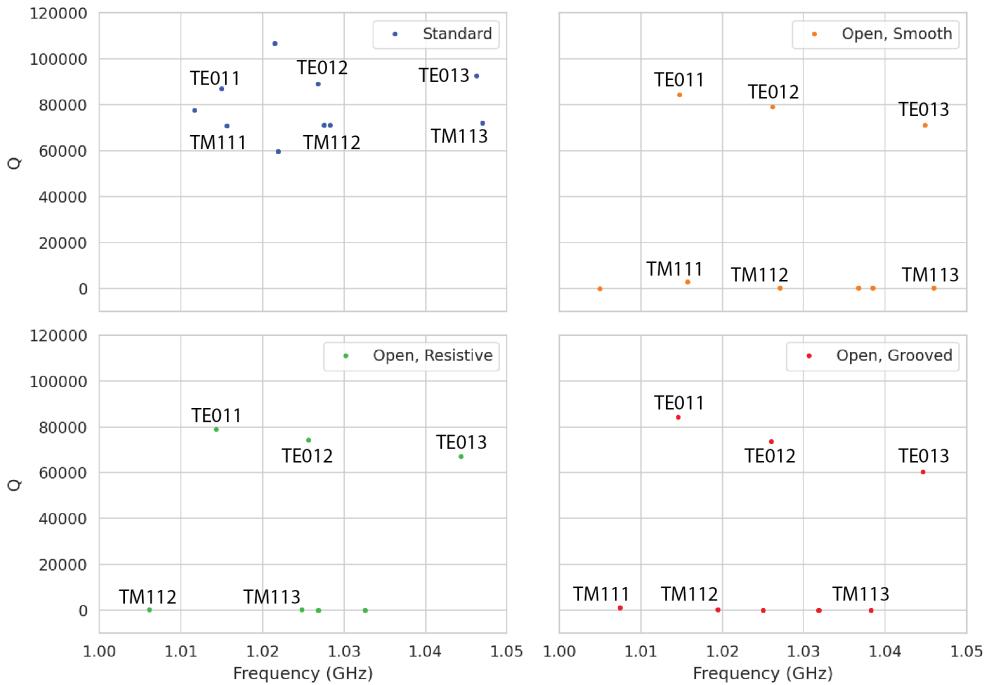


Figure 6.14: The frequencies and Q-factors of the resonant modes identified by HFSS for the cavity variations shown in Figure 6.13. The fully-sealed cavity with smooth walls has several high-Q modes near the TE₀₁₁ resonance. Introducing the open-termination preserves the Q-factors of the TE_{01 ℓ} modes 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 TE₀₁₁ mode. By properly tuning the geometry of the grooves or the resistive spacers several MHz of frequency separation can be achieved.

can see that for a standard cavity the TE₀₁ and the TM₁₁ are degenerate in frequency with relatively high Q-factors. The open-ended cavity preserves the high Q-factors of the TE₀₁ modes, while the other modes, since their boundary conditions do not match the coaxial geometry, have their Q-factors suppressed. One can see that the effect of the resistive and inductive mode-filtering schemes is to effectively shift the resonant frequencies of the TM₁₁ modes below those of the associated TE₀₁ modes, which breaks the degeneracy. Optimization of the dielectric spacer or groove parameters can ensure that the TE₀₁₁ mode is isolated from other modes by $O(10)$ MHz, which provides sufficient bandwidth for a measurement of the tritium spectrum endpoint.

Further optimization of the cavity design requires a more detailed cavity simulation that includes the cavity coupling mechanism as well as other geometry modifications required for integration into the magnetic and tritium gas subsystems. Perhaps more

4616 important is the development of the capability to simulate the interaction of electrons
4617 with the cavity so that simulated CRES signals can be generated using cavities designed
4618 for CRES measurements. Simulated CRES signals can then be used to estimate the
4619 neutrino mass sensitivity of the experiment, which allows for the optimization of the cavity
4620 design towards the configuration that provides the best measurement of the neutrino
4621 mass.

4622 **6.5 Single-mode Resonant Cavity Measurements**

4623 Measurement test stands play an important role in the research and development process
4624 that cannot be replaced by simulations. For example, constructing a prototype CRES
4625 cavity forces one to consider important practical issues such as manufacturability and
4626 machine tolerances that may require modifications to the design. Furthermore, by
4627 comparing laboratory measurements of a real cavity to simulations, one can quantify
4628 the impact of imperfections and real-life measurement systematics, which allows for
4629 more accurate sensitivity estimates of the experiment. Lastly, the development of these
4630 prototypes helps to build the necessary experience and expertise within the collaboration
4631 required for more complicated experiments to succeed.

4632 In this spirit a prototype cavity was constructed to demonstrate the open, mode-
4633 filtered cavity concept explored in the previous sections. The primary goal of the
4634 measurements was to validate that an open, mode-filtered cavity suppressed the TM_{11}
4635 modes as predicted by HFSS simulations.

4636 **6.5.1 Cavities and Setup**

4637 Two rudimentary, cavities were constructed using segments of copper pipe available from
4638 McMaster-Carr (see Figure 6.15). The design consists of copper pipes of two diameters.
4639 The larger diameter pipe forms the main cavity wall and the smaller diameter pipe is
4640 used to create a coaxial termination. The diameter of the outer pipe was chosen to
4641 produce a TE_{011} resonance of approximately 6 GHz, while the diameter of the smaller
4642 pipe was selected based on the open termination criteria introduced in Section 6.4.1. The
4643 approximate diameters and lengths of the copper pipe are summarized in Table 6.4.

4644 Coupling to the cavity was achieved using a hand-formable segment of coaxial cable
4645 stripped at one end to form a loop antenna. This was inserted into a small hole located
4646 at the center of the main cavity wall. The coaxial terminators were supported inside the

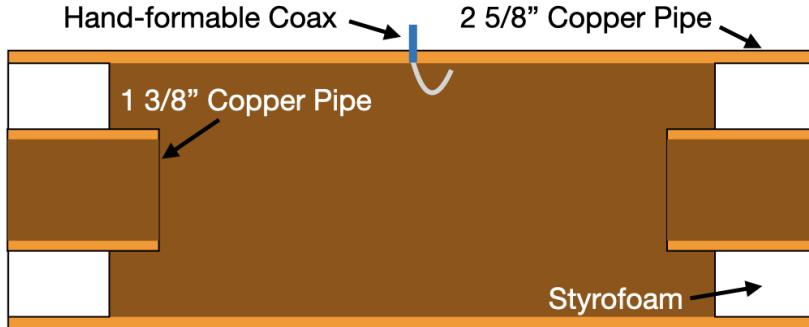


Figure 6.15: A cartoon depicting the design of the open-ended cavity prototype designed to operate at approximately 6 GHz. The main cavity wall was composed of a single copper pipe. A mode-filtered version of this cavity was constructed by

⁴⁶⁴⁷ main cavity by carving a spacer from polystyrene foam (styrofoam) so that they could
⁴⁶⁴⁸ be easily inserted into the cavity and repositioned. The dielectric constant of styrofoam
⁴⁶⁴⁹ is quite close to air at microwave frequencies so this is expected to have minimal impact
⁴⁶⁵⁰ 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 |

⁴⁶⁵¹ The actual length of the cavity is given by the distance between the inner edges of the
⁴⁶⁵² coaxial terminations. The length of the outer section of pipe that forms the main wall of
⁴⁶⁵³ the cavity is approximately 16" in length which leads to a cavity length of $\approx 13"$ when
⁴⁶⁵⁴ both terminators are inserted in the cavity. Because the terminators were not rigidly
⁴⁶⁵⁵ mounted this distance is only approximate, however, the uncertain length of the cavity
⁴⁶⁵⁶ will not prevent us from validating the open cavity design.

⁴⁶⁵⁷ Along with the smooth-walled open cavity a resistively mode-filtered cavity was
⁴⁶⁵⁸ constructed by creating dielectric spacers out of segments of clear PVC pipe (see Figure
⁴⁶⁵⁹ 6.16). The spacers were machined such that the conductive segments of the cavity would

4660 be separated by 0.75" when the cavity was fully assembled. Due to variations in the
 4661 lengths of the copper segments that make up the cavity wall the distance between spacers
 4662 has significant variation with average value of about 1.7". Eight total spacers were used
 4663 to build the cavity, which when assembled was approximately 16" in total length similar
 to the non-filtered cavity.

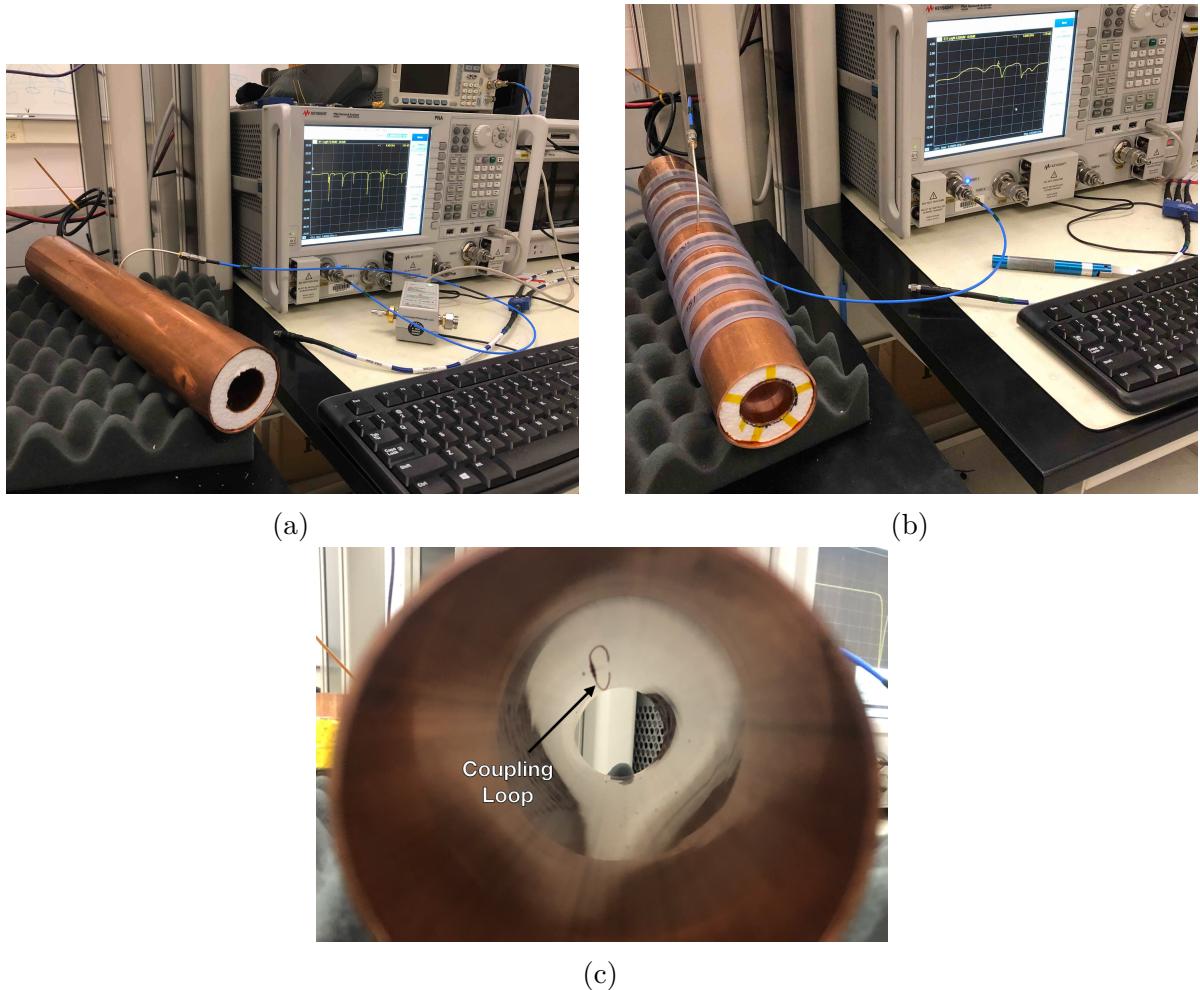


Figure 6.16: Images depicting the measurement of the filtered and non-filtered open cavities using the VNA. The coupling loop in the figure is shown in the TE orientation.

4664 Measurements of both cavities were performed using a VNA connected to the cavity
 4665 coupling probe (see Figure 6.16). By measuring the return loss over a range of frequencies
 4666 one can measure the frequencies and relative Q-factors of the resonant modes in the
 4667 cavity. Due to the opposite polarity of the electric fields for the TE and TM modes,
 4668 the loop coupling probe must be rotated 90° to change the polarity of the loop antenna.
 4669 When the antenna is oriented such that the loop opening faces the ends of the cavity, it

4671 couples primarily to the TE modes which have magnetic fields directed along the long
 4672 axis of the cavity (see Figure 6.16). If the coupling loop is turned by 90° from where
 4673 it is shown in the image then it will couple to the TM modes which have azimuthally
 4674 directed magnetic fields. In this way both the TE and TM resonances can be measured
 4675 independently.

4676 **6.5.2 Results and Discussion**

4677 The primary analysis method for the prototype cavities involved simply visualizing the
 4678 return loss measured by the VNA and comparing between the filtered and non-filtered
 4679 cavities. Since the resonances measured by the VNA are not labeled, there is some
 4680 uncertainty about the true identities of the modes measured by the VNA. To help with
 4681 this we performed a simulation of the simplest possible cavity that could be created from
 4682 the prototype components, which is a fully open cavity created by simply removing the
 4683 coaxial inserts from the non-filtered cavity configuration. The fully open cavity with the
 4684 as-built dimensions was simulated in HFSS to get estimates on the positions of the TE_{011}
 4685 and TM_{111} modes (see Figure 6.17).

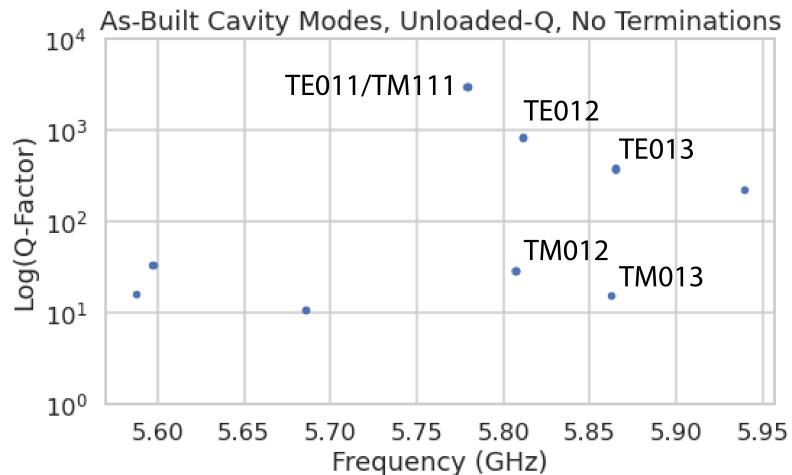


Figure 6.17: HFSS simulation results for a the as-built cavity with the coaxial terminators removed. The $\text{TE}_{011}/\text{TM}_{111}$ frequency is approximately 5.78 GHz.

4686 Simulation of the fully open cavity shows that the $\text{TE}_{011}/\text{TM}_{111}$ modes have a
 4687 frequency of approximately 5.78 GHz in the fully open cavity. If the frequency of this
 4688 mode is compared to the measurments of the fitered and non-filtered cavities with the
 4689 terminators removed we can easily identify the TE_{011} mode at approximately 5.75 GHz

4690 (see Figure 6.18).

4691 For the non-filtered cavity one sees that the TE_{011} mode is degenerate in frequency
 4692 with what appears to be a doublet of TM modes located at the TM_{111} frequency position.
 4693 This doublet is actually the TM_{111} mode, which has two polarizations with opposite
 4694 polarizations. Because the pipe used to construct the cavity is not perfectly round, the
 4695 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

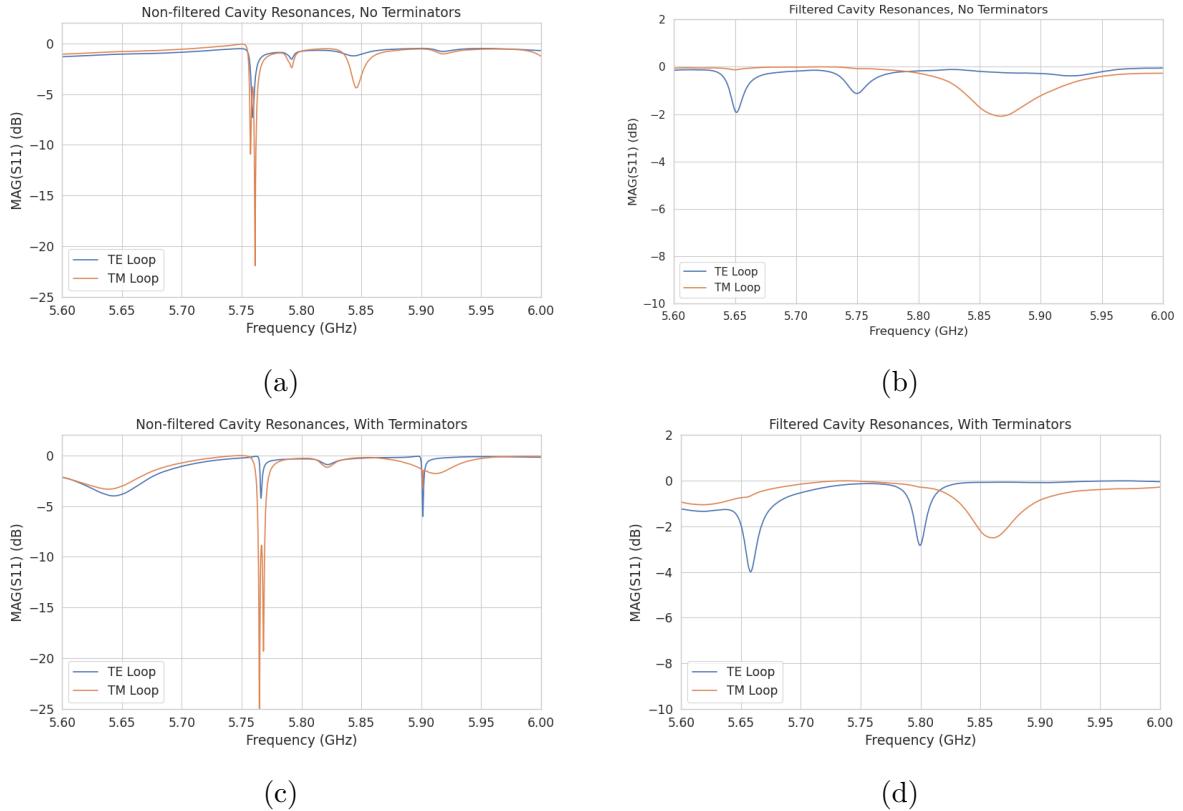


Figure 6.18: Measurements of the filtered and non-filtered prototype cavities acquired with the VNA.

4696
 4697 resonance at 5.75 GHz that appears to be the TE_{011} , however, there is no apparent TM_{111}
 4698 doublet at the same frequency. This is what one would expect if the mode-filtering was
 4699 effective at suppressing the TM modes. There is a notable difference in the Q of the
 4700 TE_{011} resonance for the non-filtered and filtered cavities indicated by the relative widths
 4701 of the resonances. This is likely caused by the large width of the dielectric spacers that
 4702 are partially impeding the TE modes. When the terminators are inserted into the cavity
 4703 one sees that Q-factors of the modes improves as expected, by noticing the narrowing of
 4704 the peaks compared to the no terminator plots.

⁴⁷⁰⁵ In conclusion, one see from these cavity measurements that, in principle, mode-
⁴⁷⁰⁶ filtering can be used to separate the TE₀₁₁ resonance from the degenerate TM₁₁₁ mode in
⁴⁷⁰⁷ combination with the an open cavity design. The ideal next step would be to construct a
⁴⁷⁰⁸ open, mode-filtered cavity that could be used to perform CRES measurements. In order
⁴⁷⁰⁹ to study the coupling of an electron to the isolated TE₀₁₁ mode.

4710 **Chapter 7 |**

4711 **Conclusion and Future Prospects**

4712 In this dissertation we have discussed research and development efforts towards the
4713 development of a scalable CRES measurement technology that can be used to build a
4714 CRES experiment at cubic-meter scales with sensitivity to neutrino masses of 40 meV.
4715 The primary contributions of my dissertation are the development and analysis of signal
4716 reconstruction algorithms for an antenna array based CRES experiment [89], which leads
4717 to estimates of the neutrino mass sensitivity; the development of a synthetic cyclotron
4718 radiation antenna (SYNCA) [56], which allowed for laboratory validation of antenna
4719 array CRES simulation models [19]; and the development of an open-ended cavity design
4720 compatible with atomic tritium for a cavity based CRES experiment. A measurable
4721 impact of this work is the transition of the Project 8 collaboration’s experimental plan
4722 from an antenna array based approach to a cavity based approach, where my work played
4723 a key role in demonstrating the significantly higher cost and complexity of the antenna
4724 array experiment.

4725 The transition from antenna arrays to cavities requires a new set of demonstrator
4726 experiments to make incremental progress towards a 40 meV measurement of the neutrino
4727 mass. At the time of writing, the near-term plan of Project 8 is to design and construct a
4728 small-scale cavity CRES experiment utilizing the 1 T magnet installed in the UW-Seattle.
4729 This cavity is designed to have a TE011 resonance with a frequency of about 26 GHz with
4730 a length-to-diameter ratio that mimics the larger cavities intended for the pilot-scale and
4731 Phase IV experiments. The goal of this experiment is to demonstrate cavity CRES as
4732 well as validate models of CRES systematics using electrons from ^{83m}Kr and an electron
4733 gun. Though the primary goal is demonstration, near-term physics measurements are
4734 available in the form of high-resolution measurements of the ^{83m}Kr conversion spectrum
4735 of interest to the KATRIN collaboration.

4736 Furthermore, Project 8 is currently constructing a low-frequency CRES setup located
4737 at Yale University to better understand the principles of cavity based CRES at lower

4738 magnetic fields. The Low, UHF Cavity Krypton Experiment at Yale (LUCKEY) is
4739 a 1.5 GHz cavity CRES experiment the will use conversion electrons from ^{83m}Kr to
4740 perform CRES measurements at the lowest frequencies ever attempted with the technique.
4741 LUCKEY will validate frequency scaling models developed by Project 8 and will pave
4742 the way for the future Low-Frequency Apparatus (LFA), which will be a larger, 1 GHz
4743 cavity CRES experiment that includes a molecular tritium source. The target for the
4744 LFA is a measurement of the neutrino mass with a sensitivity of approximately 0.2 eV,
4745 which will build towards the atomic pilot-scale CRES experiment.

4746 In parallel to the development of cavity CRES is the development of the atomic
4747 tritium source. Recent demonstrations of the production of atomic hydrogen are excellent
4748 steps towards the atomic tritium production needed for the pilot-scale experiment. One
4749 area of future study includes the development of a more detailed unstanding of the
4750 efficiency of atomic hydrogen production. Near-term plans include the development of a
4751 magnetic, evaporatively cooled beamline, as well as the prototyping of a Halbach array
4752 atoms trap. Nearly all of the components of the atomic tritium system will require
4753 demonstration before the complete system can be built. The long-term goal of the
4754 atomic tritium work is to construct a full atomic tritium prototype that demonstrates
4755 the production, cooling, trapping, and recycling of tritium at the rates needed for the
4756 pilot-scale experiment.

4757 More broadly, the long-term goal of the Project 8 collaboration is to fully develop
4758 both the atomic tritium and cavity CRES technologies so that both can be combined in
4759 a pilot-scale CRES experiment. It is envisioned that this process will take approximately
4760 10 years for both atomic tritium and cavity CRES. After these developments comes
4761 the pilot-scale experiment which will be the first CRES experiment that simultaneously
4762 demonstrates all the required technologies for Phase IV. Scaling to Phase IV with cavity
4763 CRES will require the construction of multiple copies (approximately 10) of the pilot-scale
4764 experiment to obtain sufficient statistics for 40 meV sensitivity.

4765 Development of the CRES experimental technique by Project 8 has led to new
4766 experiments utilizing the CRES technique for basic physics research, such as the ^6He -
4767 CRES collaboration [90], and has also found applications as a new approach to x-ray
4768 spectroscopy [91]. Recently, a new experimental effort called CRESDA has begun in
4769 the UK to develop new quantum technologies applied to CRES measurements for the
4770 neutrino mass [92]. This flourishing of new experimental efforts based on the CRES
4771 technique is likely to continue as Project 8 continues to develop the technique towards
4772 its neutrino mass measurement goal.

Bibliography

- 4773 [1] ZUBER, K. (2020) *Neutrino Physics*, CRC Press.
- 4775 [2] XING, Z. Z. and S. ZHOU (2011) *Neutrinos in Particle Physics, Astronomy and*
4776 *Cosmology*, Springer.
- 4777 [3] MOHAPATRA, R. N. ET AL. (2007) “Theory of neutrinos: A White paper,” *Rept.*
4778 *Prog. Phys.*, **70**, pp. 1757–1867, [hep-ph/0510213](#).
- 4779 [4] DE GOUVÉA, A. (2016) “Neutrino Mass Models,” *Ann. Rev. Nucl. Part. Sci.*, **66**,
4780 pp. 197–217.
- 4781 [5] WORKMAN, R. L. and OTHERS (2022) “Review of Particle Physics,” *PTEP*, **2022**,
4782 p. 083C01.
- 4783 [6] WONG, Y. Y. Y. (2011) “Neutrino mass in cosmology: status and prospects,” *Ann.*
4784 *Rev. Nucl. Part. Sci.*, **61**, pp. 69–98, [1111.1436](#).
- 4785 [7] ADE, P. A. R. ET AL. (2016) “Planck 2015 results. XIII. Cosmological parameters,”
4786 *Astron. Astrophys.*, **594**, p. A13, [1502.01589](#).
- 4787 [8] CARLSTROM, J., K. ABAZAJIAN, G. ADDISON, P. ADSHEAD, Z. AHMED, S. W.
4788 ALLEN, D. ALONSO, M. ALVAREZ, A. ANDERSON, K. S. ARNOLD, C. BACCI-
4789 GALUPI, K. BAILEY, D. BARKATS, D. BARRON, P. S. BARRY, J. G. BARTLETT,
4790 R. BASU THAKUR, N. BATTAGLIA, E. BAXTER, R. BEAN, C. BEBEK, A. N.
4791 BENDER, B. A. BENSON, E. BERGER, S. BHIMANI, C. A. BISCHOFF, L. BLEEM,
4792 S. BOCQUET, K. BODDY, M. BONATO, J. R. BOND, J. BORRILL, F. R. BOUCHET,
4793 M. L. BROWN, S. BRYAN, B. BURKHART, V. BUZA, K. BYRUM, E. CAL-
4794 ABRESE, V. CALAFUT, R. CALDWELL, J. E. CARLSTROM, J. CARRON, T. CECIL,
4795 A. CHALLINOR, C. L. CHANG, Y. CHINONE, H.-M. S. CHO, A. COORAY,
4796 T. M. CRAWFORD, A. CRITES, A. CUKIERMAN, F.-Y. CYR-RACINE, T. DE
4797 HAAN, G. DE ZOTTI, J. DELABROUILLE, M. DEMARTEAU, M. DEVLIN, E. DI
4798 VALENTINO, M. DOBBS, S. DUFF, A. DUIVENVOORDEN, C. DVORKIN, W. ED-
4799 WARDS, J. EIMER, J. ERRARD, T. ESSINGER-HILEMAN, G. FABBIAN, C. FENG,
4800 S. FERRARO, J. P. FILIPPINI, R. FLAUGER, B. FLAUGHER, A. A. FRAISSE,
4801 A. FROLOV, N. GALITZKI, S. GALLI, K. GANGA, M. GERBINO, M. GILCHRIESE,
4802 V. GLUSCEVIC, D. GREEN, D. GRIN, E. GROHS, R. GUALTIERI, V. GUARINO,

4803 J. E. GUDMUNDSSON, S. HABIB, G. HALLER, M. HALPERN, N. W. HALVERSON,
4804 S. HANANY, K. HARRINGTON, M. HASEGAWA, M. HASSELFIELD, M. HAZUMI,
4805 K. HEITMANN, S. HENDERSON, J. W. HENNING, J. C. HILL, R. HLOŽEK,
4806 G. HOLDER, W. HOLZAPFEL, J. HUBMAYR, K. M. HUFFENBERGER, M. HUFFER,
4807 H. HUI, K. IRWIN, B. R. JOHNSON, D. JOHNSTONE, W. C. JONES, K. KARKARE,
4808 N. KATAYAMA, J. KERBY, S. KERNOVSKY, R. KESKITALO, T. KISNER, L. KNOX,
4809 A. KOSOWSKY, J. KOVAC, E. D. KOVETZ, S. KUHLMANN, C.-L. KUO, N. KURITA,
4810 A. KUSAKA, A. LAHTEENMAKI, C. R. LAWRENCE, A. T. LEE, A. LEWIS, D. LI,
4811 E. LINDER, M. LOVERDE, A. LOWITZ, M. S. MADHAVACHERIL, A. MANTZ,
4812 F. MATSUDA, P. MAUSKOPF, J. McMAHON, P. D. MEERBURG, J. MELIN,
4813 J. MEYERS, M. MILLEA, J. MOHR, L. MONCELSI, T. MROCKOWSKI, S. MUKHER-
4814 JEE, M. MUNCHMEYER, D. NAGAI, J. NAGY, T. NAMIKAWA, F. NATI, T. NATOLI,
4815 M. NEGRELLO, L. NEWBURGH, M. D. NIEMACK, H. NISHINO, M. NORDBY,
4816 V. NOVOSAD, P. O'CONNOR, G. OBIED, S. PADIN, S. PANDEY, B. PARTRIDGE,
4817 E. PIERPAOLI, L. POGOSIAN, C. PRYKE, G. PUGLISI, B. RACINE, S. RAGHU-
4818 NATHAN, A. RAHIN, S. RAJAGOPALAN, M. RAVERI, M. REICHANADTER, C. L.
4819 REICHARDT, M. REMAZEILLES, G. ROCHA, N. A. ROE, A. ROY, J. RUHL,
4820 M. SALATINO, B. SALIWANCHIK, E. SCHAAN, A. SCHILLACI, M. M. SCHMIT-
4821 TFULL, D. SCOTT, N. SEHGAL, S. SHANDERA, C. SHEEHY, B. D. SHERWIN,
4822 E. SHIROKOFF, S. M. SIMON, A. SLOSAR, R. SOMERVILLE, S. T. STAGGS,
4823 A. STARK, R. STOMPOR, K. T. STORY, C. STOUGHTON, A. SUZUKI, O. TAJIMA,
4824 G. P. TEPLY, K. THOMPSON, P. TIMBIE, M. TOMASI, J. I. TREU, M. TRISTRAM,
4825 G. TUCKER, C. UMLITA, A. VAN ENGELEN, J. D. VIEIRA, A. G. VIEREGG,
4826 M. VOGELSBERGER, G. WANG, S. WATSON, M. WHITE, N. WHITEHORN, E. J.
4827 WOLLACK, W. L. K. WU, Z. XU, S. YASINI, J. YECK, K. W. YOON, E. YOUNG,
4828 and A. ZONCA (2019) "CMB-S4," in *Bulletin of the American Astronomical Society*,
4829 vol. 51, p. 209, 1908.01062.

- 4830 [9] ELLIOTT, S. R. and P. VOGEL (2002) "Double beta decay," *Ann. Rev. Nucl. Part. Sci.*, **52**, pp. 115–151, [hep-ph/0202264](#).
- 4831 [10] DOLINSKI, M. J., A. W. P. POON, and W. RODEJOHANN (2019) "Neutrinoless Double-Beta Decay: Status and Prospects," *Ann. Rev. Nucl. Part. Sci.*, **69**, pp. 219–251, [1902.04097](#).
- 4832 [11] AGOSTINI, M., G. BENATO, and J. DETWILER (2017) "Discovery probability of
4833 next-generation neutrinoless double- β decay experiments," *Phys. Rev. D*, **96**(5), p.
4834 053001, [1705.02996](#).
- 4835 [12] FORMAGGIO, J. A., A. L. C. DE GOUVÊA, and R. G. H. ROBERTSON (2021)
4836 "Direct measurements of neutrino mass," *Physics Reports*, **914**, pp. 1–54, direct
4837 measurements of neutrino mass.
4838 URL [https://www.sciencedirect.com/science/article/pii/
4839 S0370157321000636](https://www.sciencedirect.com/science/article/pii/S0370157321000636)

- 4843 [13] AKER, M. ET AL. (2022) “New Constraint on the Local Relic Neutrino Background
 4844 Overdensity with the First KATRIN Data Runs,” *Phys. Rev. Lett.*, **129**(1), p. 011806,
 4845 2202.04587.
- 4846 [14] FORMAGGIO, J. A., A. L. C. DE GOUVÊA, and R. G. H. ROBERTSON (2021)
 4847 “Direct Measurements of Neutrino Mass,” *Phys. Rept.*, **914**, pp. 1–54, 2102.00594.
- 4848 [15] MONREAL, B. and J. A. FORMAGGIO (2009) “Relativistic cyclotron radiation
 4849 detection of tritium decay electrons as a new technique for measuring the neutrino
 4850 mass,” *Phys. Rev. D*, **80**, p. 051301.
 4851 URL <https://link.aps.org/doi/10.1103/PhysRevD.80.051301>
- 4852 [16] ESFAHANI, A. A., S. BÖSER, N. BUZINSKY, M. C. CARMONA-BENITEZ,
 4853 C. CLAESSENS, L. DE VIVEIROS, P. J. DOE, S. ENOMOTO, M. FERTL, J. A.
 4854 FORMAGGIO, J. K. GAISON, M. GRANDO, K. M. HEEGER, X. HUYAN, A. M.
 4855 JONES, K. KAZKAZ, M. LI, A. LINDMAN, C. MATTHÉ, R. MOHIUDDIN, B. MON-
 4856 REAL, R. MUELLER, J. A. NIKKEL, E. NOVITSKI, N. S. OBLATH, J. I. PEÑA,
 4857 W. PETTUS, R. REIMANN, R. G. H. ROBERTSON, G. RYBKA, L. SALDAÑA,
 4858 M. SCHRAM, P. L. SLOCUM, J. STACHURSKA, Y. H. SUN, P. T. SURUKUCHI,
 4859 J. R. TEDESCHI, A. B. TELLES, F. THOMAS, M. THOMAS, L. A. THORNE,
 4860 T. THÜMMLER, W. V. D. PONTSEELE, B. A. VANDEVENDER, T. E. WEISS,
 4861 T. WENDLER, and A. ZIEGLER (2022) “The Project 8 Neutrino Mass Experiment,”
 4862 2203.07349.
- 4863 [17] ESFAHANI, A. A., S. BÖSER, N. BUZINSKY, M. C. CARMONA-BENITEZ,
 4864 C. CLAESSENS, L. DE VIVEIROS, P. J. DOE, M. FERTL, J. A. FORMAGGIO, J. K.
 4865 GAISON, L. GLADSTONE, M. GUIQUE, J. HARTSE, K. M. HEEGER, X. HUYAN,
 4866 A. M. JONES, K. KAZKAZ, B. H. LAROQUE, M. LI, A. LINDMAN, E. MACHADO,
 4867 A. MARSTELLER, C. MATTHÉ, R. MOHIUDDIN, B. MONREAL, R. MUELLER, J. A.
 4868 NIKKEL, E. NOVITSKI, N. S. OBLATH, J. I. PEÑA, W. PETTUS, R. REIMANN,
 4869 R. G. H. ROBERTSON, D. R. D. JESÚS, G. RYBKA, L. SALDAÑA, M. SCHRAM,
 4870 P. L. SLOCUM, J. STACHURSKA, Y. H. SUN, P. T. SURUKUCHI, J. R. TEDESCHI,
 4871 A. B. TELLES, F. THOMAS, M. THOMAS, L. A. THORNE, T. THÜMMLER,
 4872 L. TVRZNIKOVA, W. V. D. PONTSEELE, B. A. VANDEVENDER, J. WEINTROUB,
 4873 T. E. WEISS, T. WENDLER, A. YOUNG, E. ZAYAS, and A. ZIEGLER (2023)
 4874 “Cyclotron Radiation Emission Spectroscopy of Electrons from Tritium Beta Decay
 4875 and ^{83m}Kr Internal Conversion,” 2303.12055.
- 4876 [18] ESFAHANI, A. A., S. BÖSER, N. BUZINSKY, M. C. CARMONA-BENITEZ,
 4877 C. CLAESSENS, L. DE VIVEIROS, P. J. DOE, M. FERTL, J. A. FORMAGGIO, J. K.
 4878 GAISON, L. GLADSTONE, M. GRANDO, M. GUIQUE, J. HARTSE, K. M. HEEGER,
 4879 X. HUYAN, J. JOHNSTON, A. M. JONES, K. KAZKAZ, B. H. LAROQUE, M. LI,
 4880 A. LINDMAN, E. MACHADO, A. MARSTELLER, C. MATTHÉ, R. MOHIUDDIN,
 4881 B. MONREAL, R. MUELLER, J. A. NIKKEL, E. NOVITSKI, N. S. OBLATH, J. I.
 4882 PEÑA, W. PETTUS, R. REIMANN, R. G. H. ROBERTSON, D. R. D. JESÚS, G. RY-
 4883 BKA, L. SALDAÑA, M. SCHRAM, P. L. SLOCUM, J. STACHURSKA, Y. H. SUN,

- 4884 P. T. SURUKUCHI, J. R. TEDESCHI, A. B. TELLES, F. THOMAS, M. THOMAS,
 4885 L. A. THORNE, T. THÜMMLER, L. TVRZNIKOVA, W. V. D. PONTSEELE, B. A.
 4886 VANDEVENDER, J. WEINTROUB, T. E. WEISS, T. WENDLER, A. YOUNG, E. ZA-
 4887 YAS, and A. ZIEGLER (2023) “Tritium Beta Spectrum and Neutrino Mass Limit
 4888 from Cyclotron Radiation Emission Spectroscopy,” 2212.05048.
- 4889 [19] “Antenna Arrays for Physics Measurements with Large-scale CRES Detectors,” *In
 4890 preparation*.
- 4891 [20] NICHOLSON, T. ET AL. (2015) “Systematic evaluation of an atomic clock at 2 x
 4892 10-18 total uncertainty,” *Nature Communications*, **6**(6896).
- 4893 [21] HÄNSCH, T. W. (2006) “Nobel Lecture: Passion for precision,” *Rev. Mod. Phys.*,
 4894 **78**, pp. 1297–1309.
 4895 URL <https://link.aps.org/doi/10.1103/RevModPhys.78.1297>
- 4896 [22] GABOR, D. (1946) “Theory of communication. Part 1: The analysis of information,”
 4897 *Journal of the Institution of Electrical Engineers-part III: radio and communication
 4898 engineering*, **93**(26), pp. 429–441.
- 4899 [23] ESFAHANI ASHTARI, A. ET AL. (2022) “Viterbi decoding of CRES signals in Project
 4900 8,” *New J. Phys.*, **24**(5), p. 053013, 2112.05265.
- 4901 [24] JACKSON, J. D. (1999) *Classical electrodynamics*, 3rd ed., Wiley, New York, NY.
 4902 URL <http://cdsweb.cern.ch/record/490457>
- 4903 [25] F.R.S., J. L. D. (1897) “LXIII. On the theory of the magnetic influence on
 4904 spectra; and on the radiation from moving ions,” *The London, Edinburgh, and
 4905 Dublin Philosophical Magazine and Journal of Science*, **44**(271), pp. 503–512, <https://doi.org/10.1080/14786449708621095>.
 4906 URL <https://doi.org/10.1080/14786449708621095>
- 4908 [26] ESFAHANI, A. A. ET AL. (2022) “The Project 8 Neutrino Mass Experiment,” in
 4909 *2022 Snowmass Summer Study*, 2203.07349.
- 4910 [27] KRAUS, C. ET AL. (2005) “Final results from phase II of the Mainz neutrino mass
 4911 search in tritium beta decay,” *Eur. Phys. J. C*, **40**, pp. 447–468, [hep-ex/0412056](https://arxiv.org/abs/hep-ex/0412056).
- 4912 [28] ASEEV, V. N. ET AL. (2011) “An upper limit on electron antineutrino mass from
 4913 Troitsk experiment,” *Phys. Rev. D*, **84**, p. 112003, 1108.5034.
- 4914 [29] SAENZ, A., S. JONSELL, and P. FROELICH (2000) “Improved Molecular Final-State
 4915 Distribution of HeT+ for the β -Decay Process of T2,” *Phys. Rev. Lett.*, **84**(2), p.
 4916 242.
- 4917 [30] BODINE, L. I., D. S. PARNO, and R. G. H. ROBERTSON (2015) “Assessment of
 4918 molecular effects on neutrino mass measurements from tritium β decay,” *Phys. Rev.
 4919 C*, **91**(3), p. 035505, 1502.03497.

- 4920 [31] ASNER, D. M. ET AL. (2015) “Single electron detection and spectroscopy via
 4921 relativistic cyclotron radiation,” *Phys. Rev. Lett.*, **114**(16), p. 162501, 1408.5362.
- 4922 [32] VÉNOS, D., J. SENTKERESTIOVÁ, O. DRAGOUN, M. SLEZÁK, M. RYŠAVÝ, and
 4923 A. ŠPALEK (2018) “Properties of 83mKr conversion electrons and their use in the
 4924 KATRIN experiment,” *JINST*, **13**(02), p. T02012.
- 4925 [33] LAGENDIJK, A., I. F. SILVERA, and B. J. VERHAAR (1986) “Spin exchange and
 4926 dipolar relaxation rates in atomic hydrogen: Lifetimes in magnetic traps,” *Phys.
 4927 Rev. B*, **33**, pp. 626–628.
 4928 URL <https://link.aps.org/doi/10.1103/PhysRevB.33.626>
- 4929 [34] SILVERA, I. and J. WALRAVEN (1986) “Spin-polarized atomic hydrogen,” *Prog.
 4930 Low Temp. Phys.*, **X**, pp. 139–370.
- 4931 [35] FURSE, D. ET AL. (2017) “Kassiopeia: a modern, extensible C++ particle tracking
 4932 package,” *New Journal of Physics*, **19**(5), p. 053012.
 4933 URL <https://doi.org/10.1088/1367-2630/aa6950>
- 4934 [36] ESFAHANI, A. A., V. BANSAL, S. BÖSER, N. BUZINSKY, R. CERVANTES,
 4935 C. CLAESSENS, L. DE VIVEIROS, P. J. DOE, M. FERTL, J. A. FORMAGGIO,
 4936 L. GLADSTONE, M. GUIGUE, K. M. HEEGER, J. JOHNSTON, A. M. JONES,
 4937 K. KAZKAZ, B. H. LAROQUE, M. LEBER, A. LINDMAN, E. MACHADO, B. MON-
 4938 REAL, E. C. MORRISON, J. A. NIKKEL, E. NOVITSKI, N. S. OBLATH, W. PETTUS,
 4939 R. G. H. ROBERTSON, G. RYBKA, L. SALDAÑA, V. SIBILLE, M. SCHRAM, P. L.
 4940 SLOCUM, Y.-H. SUN, J. R. TEDESCHI, T. THÜMMLER, B. A. VANDEVENDER,
 4941 M. WACHTENDONK, M. WALTER, T. E. WEISS, T. WENDLER, and E. ZAYAS
 4942 (2019) “Electron radiated power in cyclotron radiation emission spectroscopy experi-
 4943 ments,” *Phys. Rev. C*, **99**, p. 055501.
 4944 URL <https://link.aps.org/doi/10.1103/PhysRevC.99.055501>
- 4945 [37] ASHTARI ESFAHANI, A. ET AL. (2019) “Locust: C++ software for simulation of
 4946 RF detection,” *New J. Phys.*, **21**, p. 113051, 1907.11124.
- 4947 [38] WIECHERT, E. (1901) “Elektrodynamische Elementargesetze,” .
 4948 URL <https://doi.org/10.1002/andp.19013090403>
- 4949 [39] LIÉARD, A. (1898) “Champ électrique et Magnétique,” *Léclairage électrique*, **16**(27-
 4950 29).
- 4951 [40] BALANIS, C. (2015) *Antenna Theory: Analysis and Design*, Wiley.
 4952 URL <https://books.google.com/books?id=PTFcCwAAQBAJ>
- 4953 [41] <https://www.ansys.com/products/electronics/ansys-hfss>.
- 4954 [42] https://en.wikipedia.org/wiki/Linear_time-invariant_system.

- 4955 [43] NYQUIST, H. (1928) “Certain Topics in Telegraph Transmission Theory,” *Transactions of the American Institute of Electrical Engineers*, **47**(2), pp. 617–644.
- 4956
- 4957 [44] BRUN, R. and F. RADEMAKERS (1997) “ROOT: An object oriented data analysis framework,” *Nucl. Instrum. Meth. A*, **389**, pp. 81–86.
- 4958
- 4959 [45] ASHTARI ESFAHANI, A. ET AL. (2021) “Bayesian analysis of a future β decay experiment’s sensitivity to neutrino mass scale and ordering,” *Phys. Rev. C*, **103**, p. 065501.
- 4960
- 4961
- 4962 URL <https://link.aps.org/doi/10.1103/PhysRevC.103.065501>
- 4963 [46] KAY, S. (1998) *Fundamentals of Statistical Signal Processing: Detection Theory, Volume II*, Pearson.
- 4964
- 4965 [47] NEYMAN, J. and E. PEARSON (1933) “On the problem of the the most efficient tests of statistical hypotheses,” *Phil. Trans. R. Soc. Lond. A*, **231**.
- 4966
- 4967 [48] STUMPF, M. (2018) *Electromagnetic Reciprocity in Antenna Theory*, Wiley.
- 4968 [49] BISHOP, C. (2016) *Pattern Recognition and Machine Learning*, Springer.
- 4969 [50] PLEHN, T., A. BUTTER, B. DILLON, and C. KRAUSE (2022) “Modern Machine Learning for LHC Physicists,” 2211.01421.
- 4970
- 4971 [51] GEORGE, D. and E. A. HUERTA (2018) “Deep Learning for Real-time Gravitational Wave Detection and Parameter Estimation: Results with Advanced LIGO Data,” *Phys. Lett. B*, **778**, pp. 64–70, 1711.03121.
- 4972
- 4973
- 4974 [52] GABBARD, H., C. MESSENGER, I. S. HENG, F. TONOLINI, and R. MURRAY-SMITH (2022) “Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy,” *Nature Phys.*, **18**(1), pp. 112–117, 1909.06296.
- 4975
- 4976
- 4977 [53] REIMANN, R. (2022) “Project 8: R&D for a next-generation neutrino mass experiment,” *PoS, PANIC2021*, p. 283.
- 4978
- 4979 [54] LAROQUE, B. (2020) “Zero-deadtime processing in beta spectroscopy for measurement of the non-zero neutrino mass,” *EPJ Web Conf.*, **245**, p. 01014.
- 4980
- 4981 [55] BUZINSKY, N. (2021) *Statistical Signal Processing and Detector Optimization in Project 8*, Ph.D. thesis, Massachusetts Institute of Technology.
- 4982
- 4983 [56] ESFAHANI, A. A., S. BÖSER, N. BUZINSKY, M. CARMONA-BENITEZ, C. CLAESSENS, L. DE VIVEIROS, M. FERTL, J. FORMAGGIO, L. GLADSTONE, M. GRANDO, J. HARTSE, K. HEEGER, X. HUYAN, A. JONES, K. KAZKAZ, M. LI, A. LINDMAN, C. MATTHÉ, R. MOHIUDDIN, B. MONREAL, R. MUELLER, J. NIKKEL, E. NOVITSKI, N. OBLATH, J. PEÑA, W. PETTUS, R. REIMANN,
- 4984
- 4985
- 4986
- 4987

- 4988 R. ROBERTSON, L. SALDAÑA, P. SLOCUM, J. STACHURSKA, Y.-H. SUN, P. SU-
 4989 RUKUCHI, A. TELLES, F. THOMAS, M. THOMAS, L. THORNE, T. THÜMM-
 4990 LER, L. TVRZNIKOVA, W. V. D. PONTSEELE, B. VANDEVENDER, T. WEISS,
 4991 T. WENDLER, E. ZAYAS, A. ZIEGLER, and P. . COLLABORATION (2023) “SYNCA:
 4992 A Synthetic Cyclotron Antenna for the Project 8 Collaboration,” *Journal of Instru-*
 4993 *mentation*, **18**(01), p. P01034.
 4994 URL <https://dx.doi.org/10.1088/1748-0221/18/01/P01034>
- 4995 [57] <https://www.nvidia.com/en-us/data-center/v100/>.
- 4996 [58] <https://www.nvidia.com/en-us/data-center/h100/>.
- 4997 [59] HE, K., X. ZHANG, S. REN, and J. SUN (2016) “Deep Residual Learning for
 4998 Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern
 4999 Recognition (CVPR)*, pp. 770–778.
- 5000 [60] SIMONYAN, K. and A. ZISSERMAN (2015) “Very Deep Convolutional Networks
 5001 for Large-Scale Image Recognition,” in *3rd International Conference on Learning
 5002 Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference
 5003 Track Proceedings* (Y. Bengio and Y. LeCun, eds.).
 5004 URL <http://arxiv.org/abs/1409.1556>
- 5005 [61] FRIIS, H. (1946) “A Note on a Simple Transmission Formula,” *Proceedings of the
 5006 IRE*, **34**(5), pp. 254–256.
- 5007 [62] https://en.wikipedia.org/wiki/Friis_transmission_equation.
- 5008 [63] POZAR, D. M. (2005) *Microwave engineering*; 3rd ed., Wiley, Hoboken, NJ.
 5009 URL <https://cds.cern.ch/record/882338>
- 5010 [64] [https://www.mvg-world.com/en/products/absorbers/standard-absorbers/
 5011 convoluted-absorbers-aec-series](https://www.mvg-world.com/en/products/absorbers/standard-absorbers/convoluted-absorbers-aec-series).
- 5012 [65] [https://en.wikipedia.org/wiki/Network_analyzer_\(electrical\)](https://en.wikipedia.org/wiki/Network_analyzer_(electrical)).
- 5013 [66] [https://www.keysight.com/us/en/product/N5222A/
 5014 pna-microwave-network-analyzer-265-ghz.html](https://www.keysight.com/us/en/product/N5222A/pna-microwave-network-analyzer-265-ghz.html).
- 5015 [67] <https://www.pasternack.com/standard-gain-horn-antennas-category.aspx>.
- 5017 [68] <https://www.markimicrowave.com/home/>.
- 5018 [69] https://en.wikipedia.org/wiki/Power_dividers_and_directional_couplers.
- 5020 [70] <https://en.wikipedia.org/wiki/Balun>.
- 5021 [71] <https://www.rigolna.com/products/waveform-generators/dg5000/>.

- 5022 [72] https://en.wikipedia.org/wiki/Frequency_mixer.
- 5023 [73] <https://www.caen.it/>.
- 5024 [74] AKER, M. ET AL. (2022) “Direct neutrino-mass measurement with sub-electronvolt
5025 sensitivity,” *Nature Physics*, **18**(2), pp. 160–166.
5026 URL <https://doi.org/10.1038/s41567-021-01463-1>
- 5027 [75] MONREAL, B. and J. A. FORMAGGIO (2009) “Relativistic Cyclotron Radiation
5028 Detection of Tritium Decay Electrons as a New Technique for Measuring the Neutrino
5029 Mass,” *Phys. Rev. D*, **80**, p. 051301, 0904.2860.
- 5030 [76] ASHTARI ESFAHANI, A. ET AL. (2017) “Determining the neutrino mass with
5031 cyclotron radiation emission spectroscopy—Project 8,” *J. Phys. G*, **44**(5), p. 054004,
5032 1703.02037.
- 5033 [77] ESFAHANI, A. A. ET AL. (2022) “The Project 8 Neutrino Mass Experiment,” in
5034 *2022 Snowmass Summer Study*, 2203.07349.
- 5035 [78] ASHTARI ESFAHANI, A. ET AL. (2019) “Locust: C++ software for simulation of
5036 RF detection,” *New Journal of Physics*, **21**(11), p. 113051.
5037 URL <https://doi.org/10.1088/1367-2630/ab550d>
- 5038 [79] FURSE, D. ET AL. (2017) “Kassiopeia: a modern, extensible C++ particle tracking
5039 package,” *New Journal of Physics*, **19**(5), p. 053012.
5040 URL <https://doi.org/10.1088/1367-2630/aa6950>
- 5041 [80] BALANIS, C. (2011) *Modern Antenna Handbook*, Wiley.
5042 URL <https://books.google.com/books?id=UYpV8L8GNcWc>
- 5043 [81] WIRTH, W. (2001) *Radar Techniques Using Array Antennas*, Institution of Engi-
5044 neering and Technology.
5045 URL <https://books.google.com/books?id=ALht42gkzLsC>
- 5046 [82] BROWN, L. S., G. GABRIELSE, K. HELMERSON, and J. TAN (1985) “Cyclotron
5047 motion in a microwave cavity: Lifetime and frequency shifts,” *Phys. Rev. A*, **32**, pp.
5048 3204–3218.
5049 URL <https://link.aps.org/doi/10.1103/PhysRevA.32.3204>
- 5050 [83] HANNEKE, D., S. FOGWELL HOOGERHEIDE, and G. GABRIELSE (2011) “Cavity
5051 control of a single-electron quantum cyclotron: Measuring the electron magnetic
5052 moment,” *Phys. Rev. A*, **83**, p. 052122.
5053 URL <https://link.aps.org/doi/10.1103/PhysRevA.83.052122>
- 5054 [84] HANNEKE, D. A. (2007) *Cavity control in a single-electron quantum cyclotron: an
5055 improved measurement of the electron magnetic moment*, Ph.D. thesis, Harvard U.

- 5056 [85] PURCELL, E. (1946) “Spontaneous Emission Probabilities at Radio Frequencies,”
5057 *Phys. Rev.*, **69**, pp. 674–674.
5058 URL <https://link.aps.org/doi/10.1103/PhysRev.69.674>
- 5059 [86] MOSKOWITZ, B. E. and J. ROGERS (1988) “ANALYSIS OF A MICROWAVE
5060 CAVITY DETECTOR COUPLED TO A NOISY AMPLIFIER,” *Nucl. Instrum.*
5061 *Meth. A*, **264**, pp. 445–452.
- 5062 [87] WENGER, N. (1967) “Resonant Frequency of Open-Ended Cylindrical Cavity,”
5063 *IEEE Transactions on Microwave Theory and Techniques*, **15**(6), pp. 334–340.
- 5064 [88] WENGER, N. C. and J. SMETANA (1972) “Hydrogen Density Measurements Using
5065 an Open-Ended Microwave Cavity,” *IEEE Transactions on Instrumentation and*
5066 *Measurement*, **21**(2), pp. 105–114.
- 5067 [89] “Real-time Signal Detection for Cyclotron Radiation Emission Spectroscopy Mea-
5068 surements using Antenna Arrays,” *In preparation*.
- 5069 [90] BYRON, W. ET AL. (2022) “First observation of cyclotron radiation from MeV-scale
5070 e^{pm} following nuclear beta decay,” [2209.02870](#).
- 5071 [91] KAZKAZ, K. and N. WOOLLETT (2021) “Using Cyclotron Radiation Emission
5072 for Ultra-high Resolution X-Ray Spectroscopy,” *New J. Phys.*, **23**(3), p. 033043,
5073 [1911.05869](#).
- 5074 [92] CANNING, J. A. L., F. F. DEPPISCH, and W. PEI (2023) “Sensitivity of future
5075 tritium decay experiments to New Physics,” *JHEP*, **03**, p. 144, [2212.06106](#).