Title

Searching for Strongly-Interacting Dark Matter with the Heavy Photon Search Experiment

Author

Spellman, Alic Shen

Publication Date

2024

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

SEARCHING FOR STRONGLY-INTERACTING DARK MATTER WITH THE HEAVY PHOTON SEARCH EXPERIMENT

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PHYSICS

by

Alic Spellman

December 2024

The Dissertation of Alic Spellman
is approved:

Robert Johnson, Chair

Jason Nielsen

Stefania Gori

Peter F. Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Searching for Strongly-Interacting Dark Matter with the Heavy Photon Search Experiment

by

Alic Spellman

The Heavy Photon Search Experiment (HPS) is a fixed-target experiment at Jefferson Lab's Hall B, designed to explore a hidden sector (HS) of particles containing dark matter and a new force mediator known as the "heavy photon" (A'). The A' is a massive spin-1 gauge boson associated with a new $U(1)_D$ symmetry in the HS that kinetically mixes with the Standard Model () photon with a weak coupling strength parameterized by ϵ, with $\epsilon^2 \sim 10^{-2} - 10^{-10}$. HPS utilizes a high-intensity electron beam on a thin tungsten target to produce heavy photons in the MeV-GeV mass range via "dark bremsstrahlung," a process analogous to SM bremsstrahlung but suppressed by ϵ^2. The A' can decay resonantly to SM leptons, allowing HPS to conduct both mass resonance searches for prompt decays (large ϵ) and displaced vertex searches for long-lived particles (small ϵ). In addition to the minimal A' model, HPS probes more complex extensions such as the QCD-like strongly-interacting massive particles (SIMPs) HS containing "dark" pions (π_D) and vector mesons (V_D), with π_D as dark matter candidates. These particles introduce new thermal dark matter freeze-out scenarios and visible signals through long-lived V_D decays to SM leptons, which are accessible to HPS.

This analysis conducted a displaced vertex search for $V_D \rightarrow e^-e^+$ in the mass range 30 MeV to 124 MeV and ϵ between $10^{-6} < \epsilon < 10^{-2}$ using data from the 2016 Engineering Run ($10.753\,\mathrm{nb}^{-1}$) at $2.3\,\mathrm{GeV}$. Unlike the minimal A' search, SIMP signal kinematics required new approaches to signal normalization and SM background rejection. The strongest signal evidence was a local p-value of 0.01317 for $m_{V_D} = 119\,\mathrm{MeV}$, corresponding to a global significance of 0.9σ. Although no signal was found, this search excluded a region of the SIMP parameter space at $90\,\%$ confidence. This work demonstrates HPS's competitive capability to probe SIMP sectors within cosmologically significant parameters and introduces a new method for HPS displaced vertex searches using track vertical impact parameter cuts.

Acknowledgments

This dissertation would not have been possible without the mentorship, guidance, and support I received from my colleagues on the Heavy Photon Search experiment. Throughout my journey I struggled with self-doubts, insecurities, and feelings of resignation, but the unwavering encouragement from our team kept me on course, and I consider everyone that I worked with on HPS a friend. I am especially grateful to Robert Johnson for recruiting me into particle physics and giving me the opportunity to gain hands-on experience and find my path. Robert's guidance throughout this process has been critical, and his friendship and insight have meant a great deal to me. My deepest thanks go to my mentor, Cameron Bravo, for taking me under his wing and developing my skills in everything from hardware and software to constructing a robust analysis from the ground up. His guidance has shaped my experience above all else, and will be forever dear to me. I also thank Matthew Graham for his contributions to this analysis. Matt reassured me of my progress at every critical point and made the path forward feel tractable. I'm honored to thank Timothy Nelson for his leadership and technical insight, as well as his empathy. I will always cherish the lunches we had together, where we graduate students played the role of unruly children, Tim shaking his head with a smile at each shenanigan. Thank you to Pierfrancesco Butti (PF) for his mentorship in tracking, software, and analysis, as well as his incisive and frank perspective on various aspects of my life and career. On top of it all, I don't think I've laughed more working with anyone else. I want to thank Omar Moreno for his mentorship with hardware, especially during those stressful trips to Virginia, and Stepan Stepanyan for helping me survive those worker shifts establishing beam. Thank you to Vitaliy Fadeyev for introducing me to silicon detectors and guiding me through the hardware work that brought me into particle physics. And of course, thanks to the HPS analysis committee members who approved my unblinding of the data, Natalia Toro, Alessandra Filippi, and Lauren Tompkins.

I also want to thank my friends and peers at HPS, Sarah Geiser, Emrys Peets, Rory O'Dwyer, and Tom Eichlersmith. Sarah and Tom both made contributions to this analysis, and Sarah also put in significant efforts highlighting the myriad spelling and grammar errors in the draft of this dissertation. Emrys and I worked closely on hardware, installation, and operations, and we always had a great time together. I also had the opportunity to work with Rory, an endless source of joy, on data calibration and event reconstruction, where he carried our project through to completion and saved my efforts from having been in vain. I would also like to thank my friends and peers at UCSC, especially Nolan Smyth, David Raftrey, Nick Hamer, and Evan Frangipane. Not only did we suffer courses together, but the former three let me sleep in their garage when I couldn't find a place to live.

Lastly, I have to thank my family. To my mom, you have always believed in and supported me, and every risk I've taken in life has been motivated in part by the knowledge that if I fail, I can always go home. To my sister Melissa, I'm incredibly proud of everything you've achieved and of the wonderful mother that you are to my nephew, and will be to the child you carry now. To my nephew Jesse, you're the most adorable person to have ever lived, and you're growing up too quickly. I am excited to see you more often, and spoil you and your future sibling completely. To my dad, I can't describe how grateful I am that we were able to reconnect after so many years, and how much your support has helped me through this. The greatest outcome of my time here is without a doubt the relationship that we now share. To my brother Logan, I'm excited to see you grow into a confident and passionate young man, and I'm always waiting for our next adventure together. To my Great-Aunt Joyce, thank you for taking care of my mom and I in our many times of need; it's not clear to me how we would have survived without you. Finally, I want to thank my grandparents Jerry and Aileen, and Mark and Linda, for their unyielding fawning over me, and their endless words of wisdom. And to my step-grandmother Linda who passed during my time in undergraduate school, I remember you every day. I love you all.

Chapter 1

Introduction

The Standard Model (SM) of particle physics has successfully classified all known 17 fundamental particles and predicted the existence of more than a few of them, such as the W and Z bosons [5, 6], the top quark [7], charm quark [8], and the Higgs boson [9, 10]. The SM is summarized in Figure 1.1. There are 12 fundamental half-integer spin fermions, each with their own antiparticle partner, four fundamental force-carrying spin-1 gauge bosons responsible for the electromagnetic, strong nuclear, and weak forces, and the most recently discovered spin-0 Higgs boson that gives particles mass and makes possible the existence of complex structures, such as the person reading this. Despite the ability of the SM to precisely and accurately describe (most of) the complex interactions of these particles, there is a gaping hole in the model that fails to account for both the gravitational force and the existence of Dark Matter (DM).

The nature of DM is arguably one of the most interesting unanswered questions in fundamental physics. For thousands of years, humanity's exploration of the cosmos was restricted to those points of light on the night sky visible to the naked eye. Only in the last

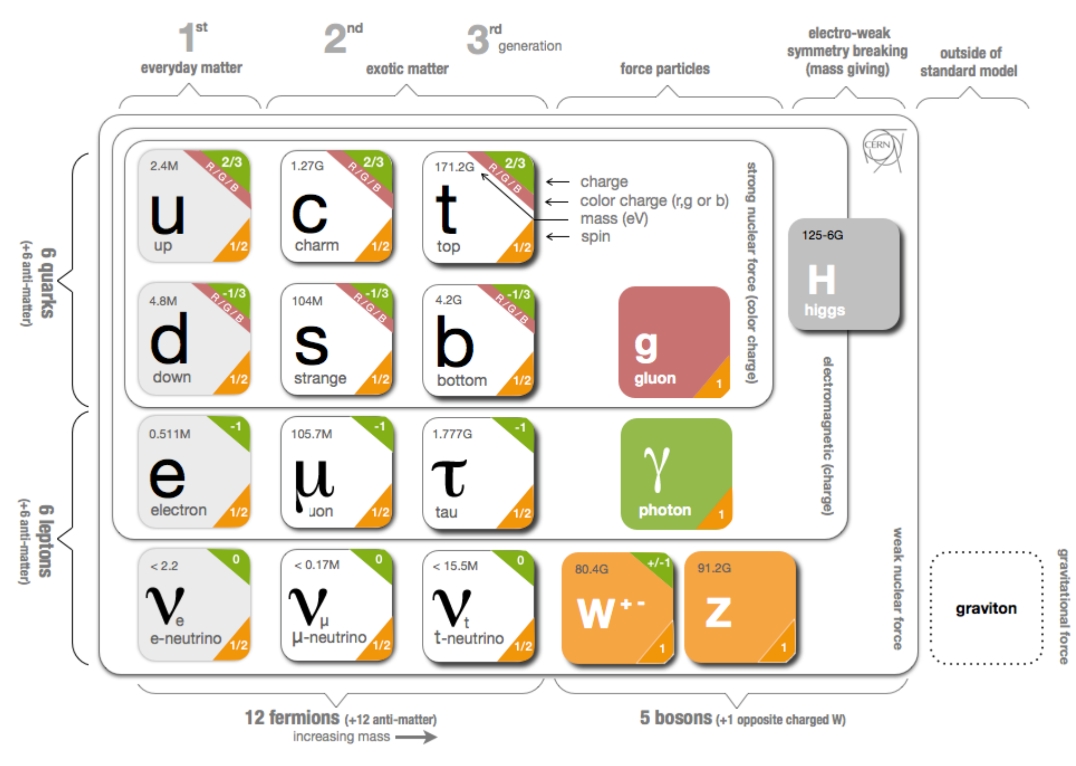


Figure 1.1: Diagram of the Standard Model of particle physics. There are six quarks, six leptons, four force-carrying gauge bosons, and the spin-0 Higgs boson [11]

500 years were observational tools developed to pierce deeper into the cosmos, revealing an astonishing breadth of complex luminous objects across the electromagnetic spectrum, from exploding stars and planetary nebulae to clusters of galaxies millions of light-years across. The visible universe is vast and rich, serving up more than enough mysteries to occupy humanity for the foreseeable future, so it is profoundly humbling to realize that the visible universe is only a fraction of the cosmos, dominated by some invisible form of matter that appears to provide the very structure upon which the stars and galaxies are created.

DM is non-luminous, and its existence is supported by an elaborate history of astrophysical observations of its gravitational effects on visible matter, as well as its consistency with the most well-supported cosmological models. This large body of evidence has led to the conclusion that DM constitutes approximately 27 % of the mass-energy of the universe, compared to only 5 % of "ordinary" SM matter [12, 13, 14, 15, 16, 17].

The DM density and structure formation of the universe are some of the most critical observables used to constrain its nature, and are well described by the Lambda Cold Dark Matter (ΛCDM) model [16]. According to ΛCDM, DM is thought to be composed of one or more new particles beyond the SM, stable on cosmological timescales, distributed evenly throughout the universe, not very strongly self-interacting, and only interacting appreciably with SM matter through gravity. Beyond the gravitational characterization of DM, however, its fundamental nature remains elusive, and the search has expanded into the realm of particle physics.

The SM does not offer any viable DM candidates, and thus physicists have invested increasing efforts in exploring new and exotic particles. In order to observe DM, there must be some coupling between SM and DM particles. Operating under the assumption that this coupling exists, particle physicists have been searching for DM using direct detection experiments [18, 19, 20] and indirect detection experiments using astrophysical signals [21, 22, 23], as well as by attempting to produce DM at high-energy accelerator facilities such as the Large Hadron Collider (LHC) and the Continuous Electron Beam Accelerator Facility (CEBAF) at the Thomas Jefferson National Lab (JLab) [24, 25, 26, 1]. The approach of these efforts is informed by a wide array of well-motivated models that determine DM's possible origins, mass and coupling scales, interaction mechanisms, signal signatures, etc.

One of the leading ΛCDM paradigms proposes that DM was created thermally in the early Universe, with DM and SM particles in thermal equilibrium mediated through some form of interaction. As the universe expanded and cooled, the DM density naturally decreased and eventually froze out, resulting in the observed relic abundance [27]. This model is consistent with the large-scale structure of the universe and current astrophysical data, and also directly relates the properties of DM to observable phenomena, making these models highly motivated. Thermal DM is defined in part by the interaction mechanisms through which the relic abundance is achieved. For instance, in a simple scenario where the number density is reduced by DM annihilation to SM particles, the correct relic abundance is achieved if DM has a mass of $\mathcal{O}$(GeV-TeV) and coupling strength similar to the weak scale. The coincidence that Weakly-Interacting Massive Particles (WIMPs) achieve the correct thermal relic abundance is known as the "WIMP Miracle" and has motivated an extensive search for these particles. In fact, the search has been so thorough that much of the viable WIMP parameter space has been excluded, and there is as of yet no compelling evidence for their existence. In light of this, interest has massively grown in searching for lighter (sub-GeV) thermal candidates.

The challenge with thermal “light DM” candidates, if one assumes the same type of WIMP-like interaction, is that they tend to overproduce DM due to the scaling of the interaction cross-section with mass [28]. However, the cross-section can be naturally boosted by introducing a new force mediator to achieve the correct relic abundance. The need for an additional mediator introduces the concept of a “dark” or “hidden” sector (HS) beyond the SM, which provides a versatile framework for accommodating a variety of new particles and interactions. One such benchmark HS model contains DM and a new force mediator in the form of a “heavy photon” or A', which is a massive spin-1 gauge boson associated with a new $U(1)_D$ gauge symmetry [29, 30]. The canonical A' model maintains thermal equilibrium with the SM through kinetic mixing with the SM photon, parameterized by the coupling strength ϵ, and achieves the correct thermal relic abundance for a wide range of viable masses and couplings [30]. The additional force mediator also addresses some of the discrepancies between the CDM-only simulations and observations. One notable discrepancy, referred to as the “too big to fail” problem, refers to the difference between the predicted and observed number and density of satellite galaxies. CDM simulations predict that there should be several dense DM subhalos around galaxies like the Milky Way that are massive enough to form visible satellite galaxies. However, the number of observed satellites is significantly smaller than what ΛCDM predicts [31], suggesting that the properties of DM may differ from the ΛCDM model. This discrepancy can be reconciled by self-interacting DM, such as through a new force mediator like the minimal A'. DM self-interactions can reduce the central density of the massive halos, making them no longer “too big” to fail at forming satellite galaxies.

Another widely recognized challenge for the ΛCDM model is called the “core vs. cup” problem. ΛCDM predicts that the DM density of galaxies should sharply increase towards the center, forming a “cusp”. However, many observations of the DM density profiles in dwarf galaxies and low-surface-brightness galaxies suggest that the DM density profile flattens towards the center, in direct contradiction with ΛCDM simulations [32, 33]. Self-interacting DM particles can transfer energy and flatten the density profile at the center of the halo from a cusp to a core and resolving the tension. Searching for A' hidden sectors is further motivated by the possibility that they can be terrestrially produced in high-energy experiments, due to the kinetic mixing between the A' and the SM photon. The ability to probe hidden sectors in high-energy experiments makes it much easier to explore the viable DM parameter space than through direct detection experiments alone. There are additionally more complex extensions of the minimal A' hidden sector that also conform with observations and produce distinct visible signatures. One particularly interesting model is the Strongly-Interacting Massive Particles (SIMPs) hidden sector, which as the name suggests, contains particles that interact strongly with each other through new forces [34, 35]. The additional degrees of freedom resulting from the more complex sector provide a richer framework to explain astrophysical observations such as the DM density, the distribution of DM in galaxies, structure formation, and self-interactions, etc.

This dissertation focuses on a QCD-like SIMP hidden sector characterized by an $SU(3)$ gauge symmetry in addition to the dark $U(1)$ gauge symmetry associated with the heavy photon. This sector contains “dark” versions of the SM vector mesons and pions, referred to as V_D and π_D, respectively, where the dark pions are stable DM. This

SIMPs model introduces two new DM freeze-out mechanisms: $3\pi_D \rightarrow 2\pi_D$ interactions and V_D decays to SM. It also provides an additional process to maintain thermal equilibrium through $V_D \leftrightarrow$ SM interactions. Including these new freeze-out and thermal equilibrium processes dramatically increases the range of mass and coupling strengths that are consistent with thermal targets [36]. The search for SIMPs in high-energy experiments is further motivated by the fact that the dark mesons could be produced in the lab through A' production. While the dark pions would remain invisible, the dark vector mesons are coupled to SM leptons through the A' and can visibly decay. Additionally, the dark vectors are naturally long-lived for most of the cosmologically favored parameter space, making them well-suited to displaced vertex search experiments like the Heavy Photon Search (HPS) Experiment. This dissertation motivates the SIMPs hidden sector as a cosmologically viable model of DM, and details the methods and results of the search for displaced dark vector meson decays using the HPS 2016 Engineering Run data. As a necessary complement to the results, the HPS experimental apparatus and event reconstruction are also described.

Chapter 2

Motivation for the SIMP Hidden Sector

This section references some of the strongest evidence for the existence of DM and motivates the search for a SIMP hidden sector of particles containing "dark" scalar and vector mesons, analogous to SM QCD, with the lightest stable dark scalar mesons being dark matter.

2.1 Evidence for Dark Matter (DM)

The evidence for DM is primarily gravitational in nature and stems from observations of galaxy rotation curves, structure formation, weak gravitational lensing, and measurements of the Cosmic Microwave Background (CMB) radiation.

2.1.1 Galactic Rotation Curves

The term dark matter, or "dunkle Materie," was first coined by Swiss astronomer Fritz Zwicky in a paper published in 1933 on the observations of nebulae outside of the Milky Way galaxy [12]. By measuring the distances and redshifts of the extra-galactic nebulae clusters, Zwicky calculated that the clusters have velocities proportional to their distances, given by

$$v = \frac{558\,\mathrm{km/s}}{1 \times 10^6\mathrm{parsec}}, \tag{2.1}$$

where the cluster speed is the average of the individual nebulae in each cluster, meaning the individual cluster member velocities can deviate from the mean. Zwicky noted an especially large variation in the Coma cluster velocities of between 1500 km/s to 2000 km/s, and by using the Virial theorem he found that the average density of the Coma cluster would need to be 400 times greater than suggested by visible matter in order to achieve the observed velocity dispersion, thus motivating the need for some massive quantity of invisible (or dark) matter.

The idea that invisible matter dominated the universe, however, was largely ignored for almost 40 years, until astrophysicists Vera Rubin and Kent Ford provided critical evidence for its existence in their 1970 publication on the observation of the rotational velocity of stars in the Andromeda galaxy (M31) [13]. According to Kepler's laws of planetary motion, the rotational velocity of the planets orbiting the sun decreases with distance, and the same behavior would be expected for stars orbiting the center of a galaxy, assuming the mass of the galaxy as centrally dominated as the distribution of visible matter suggests. However, when Rubin and Ford measured the rotational velocity of the stars in M31 as a function of their distance from the center, they observed that the velocity increases between around 2 kpc to 4 kpc, and remains nearly constant out to 24 kpc. They used the galactic rotational velocity curve to calculate the mass distribution and total mass of M31, and found the mass of the nucleus of the galaxy to be approximately $(6 \pm 1) \times 10^9$ solar masses, while total mass of the galaxy increases linearly out to around 4 kpc for a total mass of

$(1.85 \pm 0.1) \times 10^{11}$ solar masses. The total mass is substantially larger than that inferred from luminous matter alone, and the mass distribution indicates a significant amount of unseen matter extending well beyond the visible edges of the galaxy, now attributed to DM. Rubin and Ford went on to observe numerous spiral galaxies over a range of luminosities and sizes (4 kpc to 122 kpc), and found for every case that the rotational velocity curves not only remained flat with distance, but more often actually rose, further indicating that the mass of galaxies is not centrally dominated, but rather extends out significantly to far distances beyond the bulk of luminous matter. Rubin concluded that "...non-luminous matter exists beyond the optical galaxy" [14].

2.1.2 Bullet Cluster Merger

One could argue that the flat rotation curves and galaxy cluster velocity dispersions are the result of some modified theory of gravity that only applies at large cosmological scales, thus removing the need for DM. However, observations of the Bullet Cluster merger published in 2006 provided direct empirical evidence that DM exists, is non-baryonic, and indeed dominates the mass of the merging clusters [15]. The Bullet cluster consists of two primary galaxy centers, the western (right) subcluster and the eastern (left) main cluster, imaged in optical and x-ray in Figure 2.1, where the green contours indicate the gravitational potential measured by weak lensing. During the merger, the galaxies behave as collisionless particles, while the intracluster plasma experiences ram pressure and becomes spatially separated from the galaxies, as shown in the x-ray image. Since the plasma dominates the visible matter component of the clusters, one would expect the gravitational potential to coincide with the plasma if there were no DM. Instead, Clowe et al. observed

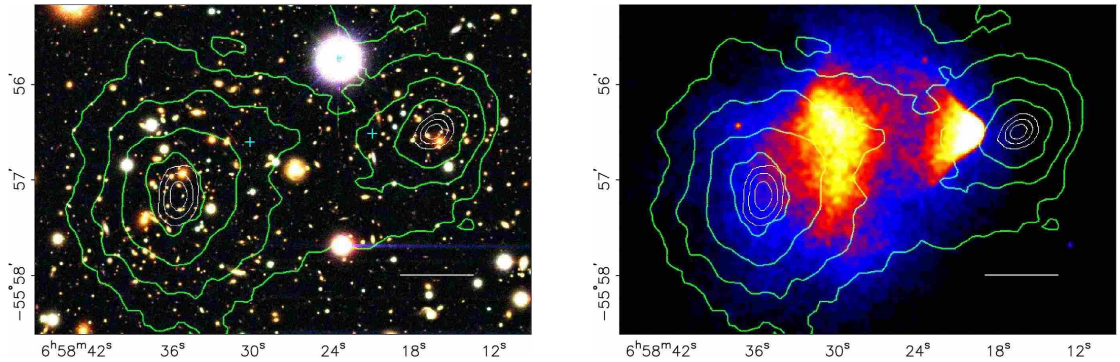


Figure 2.1: The left panel shows the optical color image of the bullet cluster merger. The right panel shows the Chandra X-ray emission image of the cluster. The green contours in both panels are weak lensing observations that represent mass-density [15].

that the plasma peak in each cluster lags behind the gravitational potential with a significance of 8σ, thus confirming the presence of a collisionless DM that dominates visible matter.

2.1.3 Cosmic Microwave Background (CMB) and Baryon Acoustic Oscillations (BAO)

The existence of DM beyond the SM is further evidenced by Planck CMB observations [16] combined with BAO measurements [17], which align well with the Lambda Cold Dark Matter (ΛCDM) model of Big Bang cosmology. In this model, Λ represents the cosmological constant associated with dark energy, which is responsible for the accelerating expansion of the universe observed through galaxy redshifts and supernova luminosity curves. The model posits that DM is non-baryonic, does not cool through radiation, interacts only gravitationally with itself and SM matter (and possibly marginally through a weak force), and is cold, meaning its velocity during the period of radiation-matter equality was much less than the speed of light. Shortly after the Big Bang, the universe was a

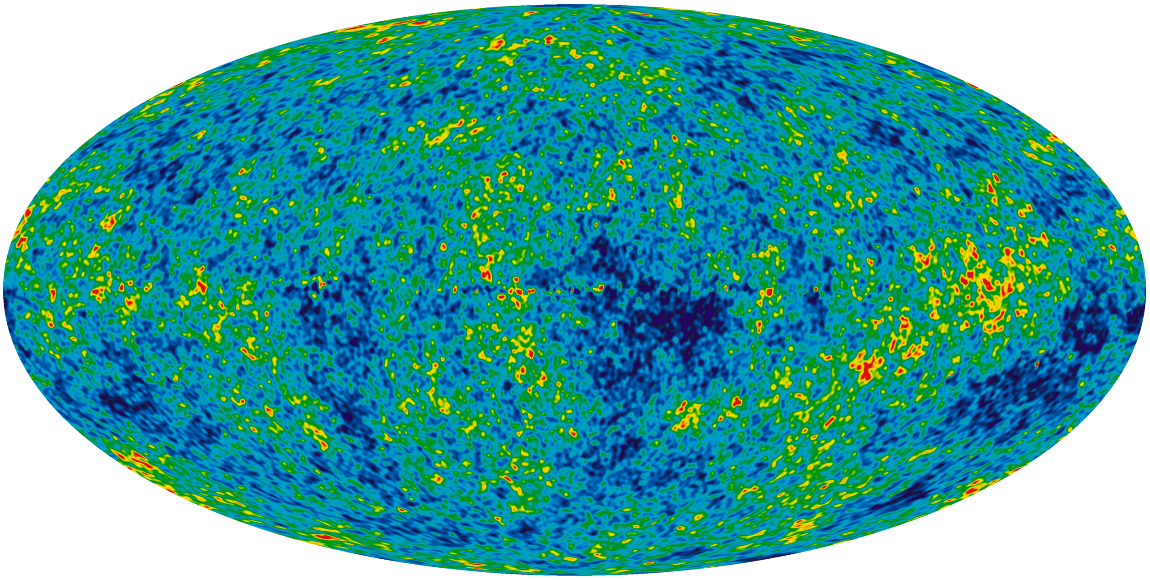


Figure 2.2: Heat map of the temperature fluctuations in the cosmic microwave background [37].

"hot primordial soup" composed of a plasma of electrons, protons, and photons. In this state, the photons constantly interacted with the plasma, rendering the universe opaque to light. As the universe expanded and cooled, electrons and protons combined to form neutral atoms (mostly hydrogen), at which point the universe became transparent to light and the photons propagated freely. These photons, redshifted due to the expansions of the universe, are observed today as the CMB radiation, shown in Figure 2.2.

Precision measurements of the CMB reveal faint anisotropies of 10 parts per million in the form of slightly colder and hotter regions, which are directly correlated to acoustic density waves in the photon-baryon plasma at recombination, referred to as Baryon Acoustic Oscillations (BAO). Gravity compresses the photon-baryon fluid, and this compression is resisted by radiation pressure, creating acoustic waves that result in hotter temperatures where there is compression and colder temperatures where there is expansion. These oscillations are frozen at recombination when the baryons release the photons, and the oscillations that freeze in at their extrema have enhanced temperature fluctuations over the characteristic sound horizon at recombination, where the sound horizon is the distance that sound waves in the fluid could have traveled in the time before recombination. These fluctuations appear as angular variations in temperature in the CMB data. The peaks in the CMB power spectrum, shown in Figure 2.3, correspond to the different oscillation modes, and the overall shape of the power spectrum directly depends on the relative amounts of DM and baryonic matter. For example, a higher baryon density results in larger compression relative to rarefaction, enhancing the odd-numbered acoustic peaks over even ones. Conversely, DM density affects compression and rarefaction symmetrically and influences the overall amplitude of the acoustic peaks.

The CMB power spectrum data is excellently described by the ΛCDM model, which measures the DM density of the early universe to be $\Omega_c h^2 = 0.120 \pm 0.001$, more than five times greater than the baryon density $\Omega_b h^2 = 0.0224 \pm 0.0001$. These results confirm that DM significantly dominates visible matter, and explicitly separates DM from baryonic matter, motivating the search for new particles beyond the standard model. Further evidence supporting ΛCDM comes from the observed imprint of CMB acoustic oscillations on structure formation in the universe. At the epoch of recombination, photons decoupled from baryons, relieving the pressure that supported the acoustic oscillations and causing the baryons to expand in a spherical shell around the initial DM perturbation site. The first resonance of this baryon shell is correlated with the sound horizon, roughly 150 Mpc in scale. As these DM perturbations grow, they combine with the baryon shells to seed the formation of large-scale structures, such as galaxies and clusters. The acoustic resonance feature at the sound horizon manifests as a peak in the galaxy correlation function and was

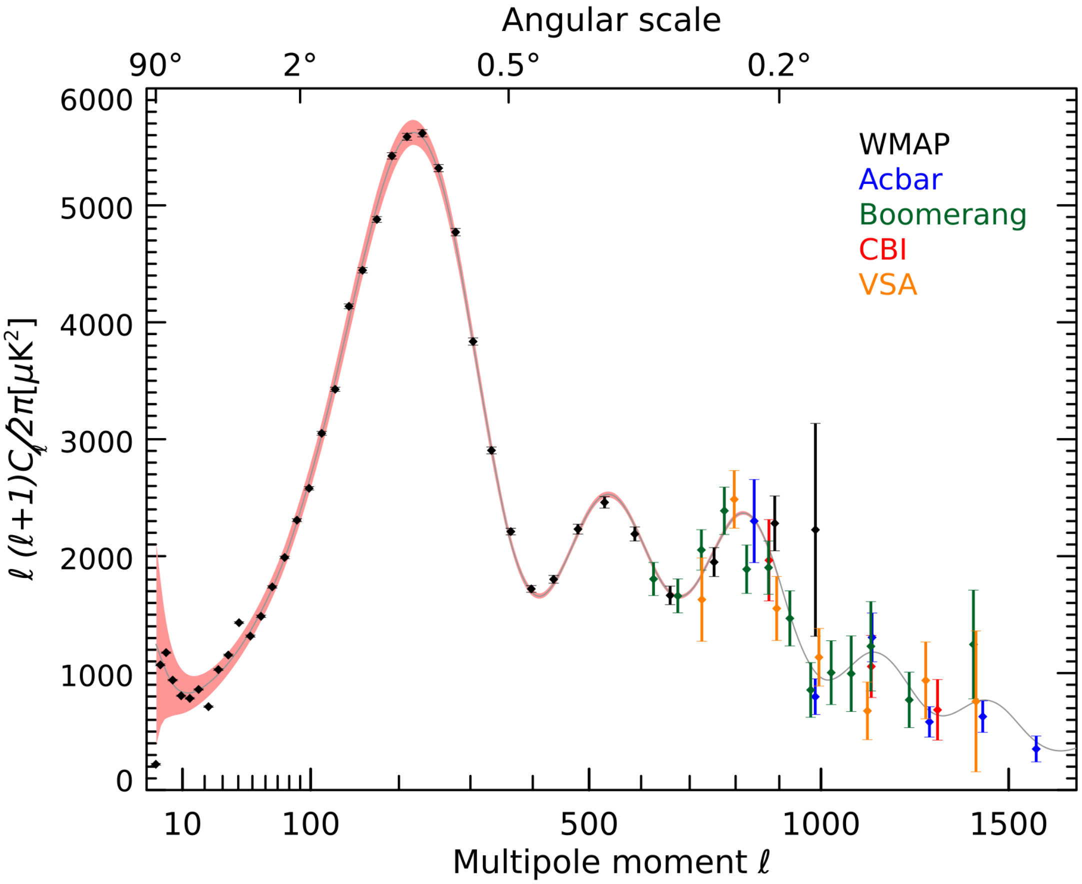


Figure 2.3: The power spectrum of the cosmic microwave background radiation temperature fluctuations. The power spectrum is fit with the Λ cold DM model (solid line) [38].

ultimately detected using 46,748 galaxies from the Sloan Digital Sky Survey (SDSS) [17]. The direct measurement of the sound horizon confirms the role of DM in galaxy formation and is consistent with the theoretical prediction that DM is critical to the overall large-scale structure of the universe.

2.2 Thermal Relic DM

Evidence from CMB measurements, and large-scale structure, and galaxy formation suggests that DM played a critical role in the evolution of the early universe. The fact that the DM density is comparable in magnitude to that of SM matter suggests that they may have a shared origin in the early universe. One prominent hypodissertation based on this assumption is that DM is a “thermal relic” of the early universe. In the thermal relic scenario, SM and DM particles were in thermal equilibrium in the hot early universe, annihilating into each other at similar rates and maintaining a constant number density. However, as the universe expanded and cooled, DM continued to annihilate to the SM, but the back-reaction rate decreased with temperature. This led to a decline in the overall DM number density. Eventually, the universe expanded to the point where the DM interaction rate could no longer keep up with the Hubble expansion rate, and the DM number density “froze out” to the thermal relic abundance, which aligns with the observed DM abundance today. The thermal relic framework guides the development of DM models by establishing a relationship between the relic abundance, $\Omega h^2 \approx 0.12$, and the thermally averaged DM annihilation cross-section, $\langle \sigma v \rangle$. In the standard thermal relic framework, the relic abundance and cross-section are inversely proportional, typically expressed as:

$$\Omega_{DM} h^2 \approx \frac{3 \times 10^{-27} \mathrm{cm}^3 \mathrm{s}^{-1}}{\langle \sigma v \rangle}, \tag{2.2}$$

where a larger $\langle \sigma v \rangle$ results in a lower relic abundance, and vice versa.

One of the most popular thermal relic DM candidates is Weakly-Interacting Massive Particles (WIMPs), which primarily undergo freeze-out through an interaction of the type $\mathrm{DM} + \overline{\mathrm{DM}} \rightarrow \mathrm{SM} + \overline{\mathrm{SM}}$. Generally, the WIMP thermally averaged annihilation cross-section is a function of its mass. For typical masses of the order of a few GeV to a few TeV, the correct thermal relic abundance is achieved when $\langle \sigma v \rangle \sim 3 \times 10^{-26} \text{cm}^3 \text{s}^{-1}$. This natural emergence of the electroweak-scale masses and interactions within this simple thermal relic paradigm is often referred to as the "WIMP miracle." This has led particle physicists to spend the last few decades developing and running experiments to search for these candidates. These searches have been highly successful in excluding significant portions of the WIMP parameter space for DM masses greater than a few GeV. As the viable mass parameter space continues to shrink, there is growing interest in exploring "light DM" candidates with sub-GeV masses. However, incorporating light DM into the traditional WIMP paradigm introduces some challenges. Since the thermally averaged annihilation cross-section is typically related to the WIMP mass, if the WIMP mass is less than $\sim 2\,\text{GeV}$, the predicted relic abundance becomes too large [28]. To address this issue, the WIMP paradigm can be extended to include light DM within the framework of "hidden/dark sectors," where new interactions and mediators can change the freeze-out dynamics and restore the correct relic abundance.

The hidden sector framework posits that DM and at least one new mediator exist within a hidden sector of particles. However, the new mediator must have some indirect coupling to the SM in order to regulate thermal equilibrium between DM and the SM in the early universe. The indirect coupling is also required to boost the DM thermally averaged annihilation cross-section and achieve the correct DM relic abundance [39]. One interesting class of hidden sector models that satisfies these conditions introduces a new massive $U(1)_D$ gauge boson, commonly referred to as the "heavy photon", or A' [29]. The A' kinetically mixes with the SM photon, allowing thermal equilibrium between the hidden and visible sectors in the early universe and providing a "portal" through which the hidden sector can be probed. The minimal (or canonical) A' HS model sufficiently produces the correct DM relic abundance over a wide range of A' masses and kinetic mixing strengths. However, this minimal A' HS framework can be also be extended to include a more complex internal structure. By introducing new particles and dynamics, the model can further expand the range of viable thermal equilibrium and freeze-out mechanisms, broadening the possible scenarios in the early universe.

A simple and well-motivated extension of the A' HS introduces a QCD-like $SU(3)_D$ symmetry, in addition to the A' $U(1)_D$ symmetry, resulting in a HS of Strongly-Interacting Massive Particles (SIMPs). The SIMP HS contains dark pions (π_D) and dark vector mesons (V_D), analogous to SM QCD, where the π_D are the lightest HS states and constitute DM. Dark pion self-interactions introduce a completely new freeze-out mechanism for thermal DM through $3\pi_D \rightarrow 2\pi_D$ annihilation, rather than the typical $2\pi_D \rightarrow 2\text{SM}$ process alone. Similar to the WIMP Miracle, the SIMPs model produces the correct observed DM relic abundance for strong-scale DM masses (MeV-GeV) and coupling, a scenario thus referred to as the "SIMPs Miracle" [34]. While the newly introduced $3 \rightarrow 2$ process alone can achieve the correct DM relic abundance [34], the dark vector mesons also play a significant role in SIMP cosmology by providing an additional freeze-out mechanism through a process called "semi-annihilation". In this process, the DM number density is reduced by $\pi_D\pi_D \rightarrow \pi_D V_D$, followed by $V_D \rightarrow \text{SM}$ through a virtual A'. This additional freeze-out mechanism often dominates over both the WIMP-like and $3\pi_D \rightarrow 2\pi_D$ processes for much of the model's parameter space. Including V_D semi-annihilation also dramatically expands the viable thermal DM mass range and the range of sector coupling ϵ to several

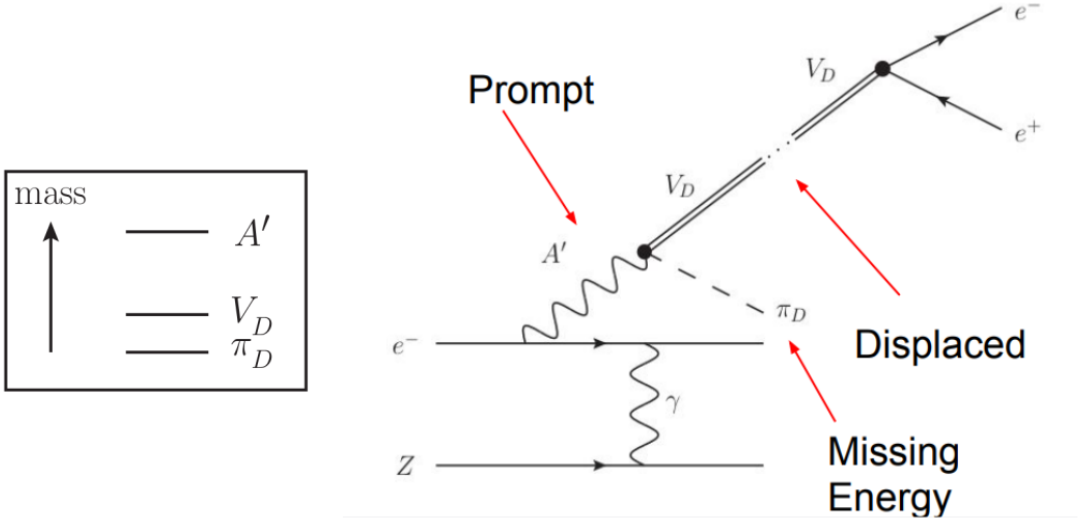


Figure 2.4: Diagram representing an electron beam incident on a fixed target with atomic number Z. An A' is produced at the target and promptly decays to a $V_D\pi_D$ pair. The π_D carries away some fraction of the A' energy. The V_D is long lived, and decays to e^+e^- with a displaced vertex. This process requires that the A' mass is greater than the sum of the V_D and π_D masses, and that the V_D mass is less than twice the π_D mass.

orders of magnitude, $\epsilon \sim 10^{-7} - 10^{-3}$ [36]. The potential for dark vector particles to decay into the SM motivates the search for SIMPs candidates in high-energy and high-intensity experiments. This dissertation focuses on conducting such a search using data from the Heavy Photon Search. HPS is capable of probing these complex HS models through the kinetic mixing of the SM photon with the A'. In a fixed target experiment, heavy photons could be produced at the target and subsequently generate long-lived dark vector mesons. The dark vector mesons could then visibly decay to the SM through a virtual A', as as illustrated in Figure 2.4. The distinctive signature of a displaced vertex associated with a visible resonant decay is particularly well-suited for the HPS.

2.3 A' Theory

The A' is a massive spin-1 $U(1)_D$ gauge boson that weakly couples to the SM photon and is a well-motivated and simple candidate for mediating interactions between the HS and the SM. The concept of an additional broken $U(1)$ Abelian gauge symmetry that mixes with SM hypercharge was first introduced in Ref. [29]. The model is insensitive to the mechanism that gives the A' mass, where it's generally assumed that the A' mass is generated by a Higgs or Stueckelberg mechanism [30]. In a theory with $U(1)_Y \times U(1)_D$ symmetry, $U(1)_D$ mixes with $U(1)_Y$ hypercharge through loops of massive fields that are charged under both symmetries. This kinetic mixing induces an effective gauge coupling between the SM photon (after electroweak symmetry breaking) and the heavy photon, with the interaction strength suppressed by the dimensionless kinetic mixing parameter ϵ. The Lagrangian that describes the mixing between the A' and the SM photon is

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + \frac{1}{4}F'^{\mu\nu}F'_{\mu\nu} + m_{A'}^2 A'^{\mu}A'_{\mu} + \epsilon F^{\mu\nu}F'_{\mu\nu}, \tag{2.3}$$

where $\mathcal{L}_{\text{SM}}$ is the Standard Model Lagrangian, $F_{\mu\nu}$ is the electromagnetic field strength tensor, and $F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu$ is the heavy photon field strength tensor.

The kinetic mixing term $\epsilon F^{\mu\nu}F'_{\mu\nu}$ in Equation (2.3) leads to mixing between the SM photon and the A' through the interactions of massive fields charged under both $U(1)_{\text{SM}}$ and $U(1)_D$, such as M_Φ and $M_{\Phi'}$ shown in Figure 2.5. Redefining the photon field $A^\mu \to A^\mu + \epsilon A'^\mu$, and eliminating all ϵ^2 terms, removes the kinetic mixing term and diagonalizes the gauge Lagrangian,

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}F'^{\mu\nu}F'_{\mu\nu} - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{2.4}$$

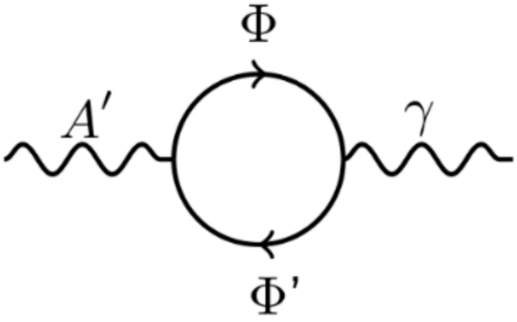


Figure 2.5: Diagram illustrating the kinetic mixing of the A' and photon through the one-loop interaction of two massive fields that are charged under both $U_{SM}(1)$ and $U(1)_D$.

This redefinition also changes the interaction term of the SM Lagrangian from $A^\mu J_\mu^{EM}$ to

$$A^\mu J_\mu^{\mathrm{EM}} \to (A^\mu + \epsilon A'^\mu) J_\mu^{\mathrm{EM}}, \tag{2.5}$$

which induces an effective coupling between the heavy photon and electrically charged SM particles, proportional to the kinetic mixing parameter ϵ.

Loop processes involving massive fields charged under both electric and dark charges lead to natural values of ϵ in the range of $\sim 10^{-2} - 10^{-8}$ [30]. This weak coupling between the heavy photon and the SM photon allows for thermal equilibrium between the HS and SM matter in the early universe. Consequently, the A' is an ideal candidate for the light mediator required by SIMP HS models. The coupling between the A' and the SM photon also makes A' production, and therefore exploration of SIMPs, possible with high-energy experiments.

2.3.1 Fixed Target A' Kinematics

High energy electrons incident on the fixed target can generate heavy photons through a process known as "dark bremsstrahlung." This process is analogous to standard photon bremsstrahlung but is suppressed by $\sim \epsilon^2 \frac{m_e^2}{m_{A'}^2}$. A' production is sharply peaked near the full beam energy, with a soft recoil electron, and the A' emission angle is approximately co-linear with the beam [30]. The minimal A' is long-lived with a decay length given by:

$$l_0 = \frac{0.8\,\mathrm{cm}}{N_{\mathrm{eff}}} \left(\frac{E_0}{10\,\mathrm{GeV}}\right) \left(\frac{10^{-4}}{\epsilon}\right)^2 \left(\frac{100\,\mathrm{MeV}}{m_{A'}}\right)^2, \tag{2.6}$$

However, in the SIMP model, the A' generally decays promptly to the hidden sector. This is because the coupling of the A' to the HS, typically characterized by $\alpha_D \sim 0.1$, is significantly stronger than the A' kinetic mixing with the SM, which is proportional to proportional to ϵ^2, and is typically $\epsilon < 10^{-2}$.

2.4 SIMP Theory

The SIMP HS is cosmologically motivated by a new 3DM $\rightarrow$ 2DM DM freeze-out mechanism. When the $3 \rightarrow 2$ process dominates the WIMP-like 2DM $\rightarrow$ 2SM process at freeze-out, and assuming the DM self-interaction strength is $\alpha_{\mathrm{eff}} \sim 1$ (strongly interacting), the SIMP framework achieves the correct DM relic abundance for a mass scale given by:

$$m_{DM} \approx \alpha_{\mathrm{eff}}(T_{eq}^2 M_{Pl}^{1/3}) \approx 100\,\mathrm{MeV}, \tag{2.7}$$

where T_{eq} is the matter-radiation equality temperature, and M_{Pl} is the reduced Planck mass [34].

As mentioned previously, thermal relic DM models require a measurable coupling between DM and the SM to maintain thermal equilibrium up until freeze-out. This coupling strength is parameterized by ϵ in anticipation of the role of the A'. The coupling must be small enough to ensure that the 2DM $\rightarrow$ 2SM annihilation process is inefficient at freeze-out, but sufficiently large to keep the hidden and visible sectors at a unified temperature. For SIMP models with strong-scale DM masses and self-interactions, these requirements can be met with a broad range of coupling strengths [34].

This dissertation focuses on a specific version of the SIMP HS characterized by an SU(3) gauge symmetry with $N_C = 3$ colors and $N_f = 3$ flavors of light HS quarks. The dynamics of this model are analogous to those of SM QCD. The global chiral symmetry $SU(N_f = 3)_L \times SU(N_f = 3)_R$ is spontaneously broken to $SU(N_f = 3)_V$, resulting in 5 massive pseudo-Goldstone bosons, π_D, which are the lightest states in the theory and are the DM particle candidates [35]. Additionally, this SIMP sector also necessarily includes heavier spin-1 dark vector mesons (V_D), which play a critical role in SIMP cosmology and its experimental signatures. The A' gauge boson serves as the new mediator between the SIMP HS and the SM. The SIMP HS is thus additionally charged under the A' $U(1)_D$ symmetry, indirectly coupling SIMPs to the SM through kinetic mixing with the A'. The SIMP 3DM $\rightarrow$ 2DM freeze-out mechanism is driven by the π_D self interactions, which are described by the 5-point interaction term, known as the Wess-Zumino-Witten (WZW) term [35]. The strength of these self-interactions is characterized by f_{π_D}, analogous to the pion decay constant in the SM.

The ratio of the dark pion mass to the dark pion decay constant, m_{π_D}/f_{π_D}, is a critical SIMP model parameter. If the $3\pi_D \rightarrow 2\pi_D$ process dominates freeze-out, the thermally averaged cross-section is proportional to $\frac{m_\pi^5}{f_\pi^{10}}$, and the SIMP model predicts the correct DM energy density for m_{π_D}/f_{π_D} on the order of a few, with m_{π_D} ranging from a few MeV to approximately 1 GeV. Observations from the Bullet Cluster merger and DM halo studies enforce a lower bound on m_{π_D} of a few hundred MeV, while the perturbativity bound requires $m_{\pi_D}/f_{\pi_D} \lesssim 4\pi$ [35].

2.4.1 Dark Vector Semi-Annihilation

The SIMP model described thus far constrains the dark matter mass to hundreds of MeV. However, the lower bound on m_{π_D} due to self-interactions is significantly reduced by incorporating the role of dark vector mesons in the DM freeze-out process. This effect was first demonstrated in Ref. [36], where it was shown that dark vector semi-annihilation ($\pi\pi \to \pi V$, followed by V → SM through a virtual A') can significantly broaden the range of DM masses and coupling strengths to the SM. The viable range of m_{π_D}/f_{π_D} found in the previous section implies that $m_{V_D} \sim m_{\pi_D}$, where the expected vector meson masses are

$$m_{V_D} \sim 4\pi f_{\pi_D}/\sqrt{N_{c=3}}. \tag{2.8}$$

At this mass scale, the dark vector mesons are expected to contribute significantly to freeze-out through semi-annihilation. In fact, this semi-annihilation process typically dominates freeze-out over $3\pi_D \to 2\pi_D$ for much of the parameter space, and dramatically expands the thermal DM mass range to $m_{\pi_D} \sim 0.01\,\text{GeV}$ to $1\,\text{GeV}$. Additionally, semi-annihilation contributes to maintaining kinetic equilibrium between the two sectors, expanding the range of the SM coupling ϵ by several orders of magnitude [36]. Finally, semi-annihilation provides a visible signature through the decay of dark vector mesons to the SM.

This more detailed view of the QCD-like SIMP model significantly broadens the viable parameter space for thermal DM and gives HPS sensitivity to the hidden sector visible decays. However, this sensitivity depends critically on enforcing certain mass hierarchies and parameter values that determine the dark vector meson production rates, visible decay rates, and lifetimes. The following section defines the key SIMP parameters, constraints, and rates, as they apply to the HPS displaced vertex search presented in this dissertation.

2.5 HPS SIMP Model

The SIMP model involves six key parameters, detailed in this section. Some of these parameters are constrained by quantum field theory and cosmological arguments. Additional constraints are enforced by HPS to ensure a visible signal. To simplify the analysis, the two well-motivated "benchmark" cases are established. In these cases, the HS particle mass ratios are fixed according to $m_{A'} : m_{V_D} : m_{\pi_D} = 3 : 1.8 : 1$. The $U(1)_D$ coupling is fixed to $\alpha_D = 0.01$, and the ratio m_{π_D}/f_{π_D} is evaluated at 3 and 4π [36].

2.5.1 Parameter Constraints

The key SIMP parameters are listed as follows:

1. $m_{A'}$: The mass of the A' gauge boson.
2. ϵ: The kinetic mixing strength between the SM photon and the A'. This parameter is critical for maintaining kinetic equilibrium between the two sectors, determining the efficiency of the semi-annihilation process in reducing the DM number density,

and influencing the V_D lifetime.

3. α_D: The hidden sector $U(1)_D$ gauge coupling constant, analogous to SM QED fine structure constant α.
4. m_{π_D}: The mass of the dark pions (π_D).
5. m_{V_D}: The mass of the dark vector mesons (V_D).
6. m_{π_D}/f_{π_D}: The ratio of the DM mass to the dark pion decay constant. This ratio impacts the dark meson interactions and the relative importance of the various freeze-out mechanisms. It also impacts the A' branching ratios to the HS.

The $U(1)_D$ gauge coupling α_D is constrained by perturbativity to be less than 1, and is typically fixed to $\alpha_D = 10^{-2}$. The kinetic mixing parameter ϵ has a natural range of $10^{-2} - 10^{-8}$. If $\epsilon \gtrsim 10^{-2}$., semi-annihilation becomes non-dominant, thus $\epsilon < 10^{-2}$ is required. For many well-motivated benchmark parameter choices, as $\epsilon \to 10^{-6}$, DM-SM elastic scattering and vector meson decays no longer maintain kinetic equilibrium during freeze-out, so it is required that $\epsilon > 10^{-6}$ [36]. The perturbativity of the meson theory constraints m_{π_D}/f_{π_D} to be less than 4π , since $m_{\pi_D}/f_{\pi_D} \approx g_D = 4\pi\alpha_D$, and α_D must be less than 1.

For semi-annihilation to dominate freeze-out, the A' mass must be greater than twice the π_D mass to suppress $\pi_D\pi_D \to A'\pi_D$. Additionally, with $m_{V_D} \sim m_{\pi_D}$, the A' mass must be greater than m_{V_D}. HPS requires that $m_{V_D} < 2m_{\pi_D}$, which excludes V$\to \pi_D\pi_D$ and forces the dark vector mesons to visibly decay to the SM. This requirement also restricts $m_{A'} < 2m_{V_D}$. The visibly decaying dark vector mesons are generated by fixed-target A'

production via $A' \to \pi_D V_D$, followed by $V_D \to e^+e^-$. The $A' \to \pi_D V_D$ process requires that $m_{A'} > m_{\pi_D} + m_{V_D}$. Additionally, in order for the V_D decay to be visible to HPS, it must decay to e^+e^-, thus requiring that $m_{V_D} < 2m_\mu$. Lastly, this analysis requires that $m_{A'} < 2m_\mu$. This is necessary due to the way that HPS scales the expected signal rate using data, and is described later in Section 5.1.4. The complete set of SIMP parameter constraints used in this analysis are summarized in Section 2.5.1.

$m_{A'} < 2m_\mu$ and $m_{A'} < 2m_{V_D}$
$m_{A'} > m_{V_D} + m_{\pi_D}$ and $m_{A'} > 2m_{\pi_D}$
$m_{V_D} < 2m_{\pi_D}$ and $m_{V_D} < 2m_\mu$
$\alpha_D < 1$
$10^{-6} < \epsilon < 10^{-2}$
$m_{\pi_D}/f_{\pi_D} < 4\pi$

Table 2.1: HPS SIMP Parameter Constraints

2.6 HPS Rates

The SIMP parameter values determine the dark vector production rates and lifetimes as functions of mass and the kinetic mixing strength ϵ. The rates presented in this section are calculated assuming the previously defined parameter constraints in Section 2.5.1.

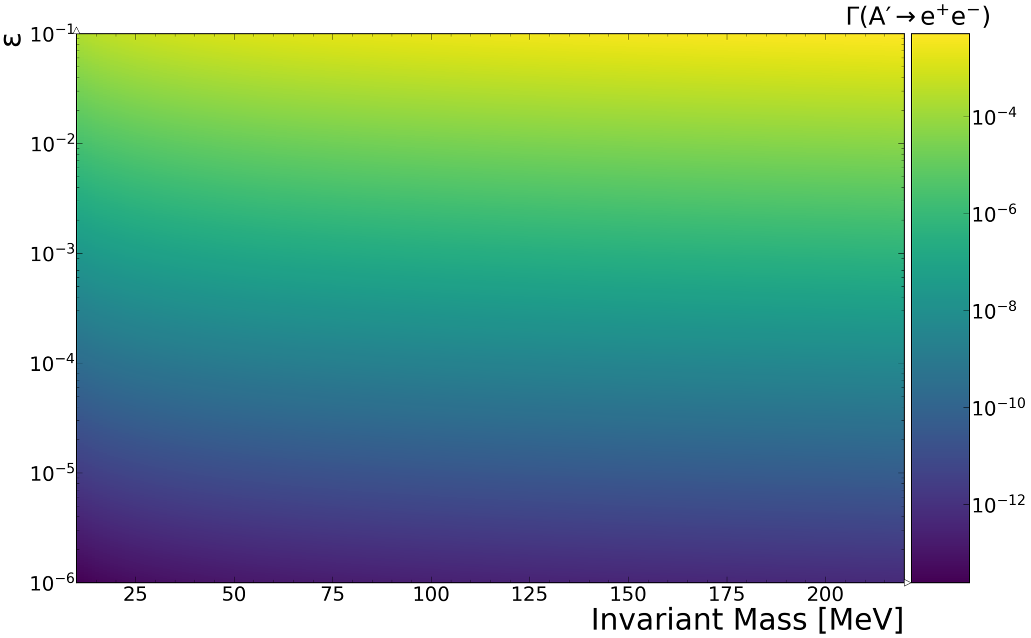


Figure 2.6: A' decay rate directly to e^- e^+ as a function of kinetic mixing ϵ. The decay rate is negligible across all of the relevant phase space in this analysis.

2.6.1 A' Decays

The HPS SIMP signal is characterized by visibly decaying dark vector mesons generated through A' decays into the HS. Therefore, the expected signal rate depends on the various A' branching ratios. A's can decay into both the HS, and to electrically charged SM particles through the its kinetic mixing with the photon. However, the kinematics of A' decays directly to the SM do not align with those of visibly decaying dark vector mesons, thus suppressing the expected signal production rate. Nonetheless, because $\epsilon \ll 1$, A' decays into the SM are highly suppressed, and the A' decays almost exclusively to HS mesons. Figure 2.6 shows $\Gamma(A' \to e^+e^-)$ over the relevant phase space, where this width is much less than $A' \to$ HS.

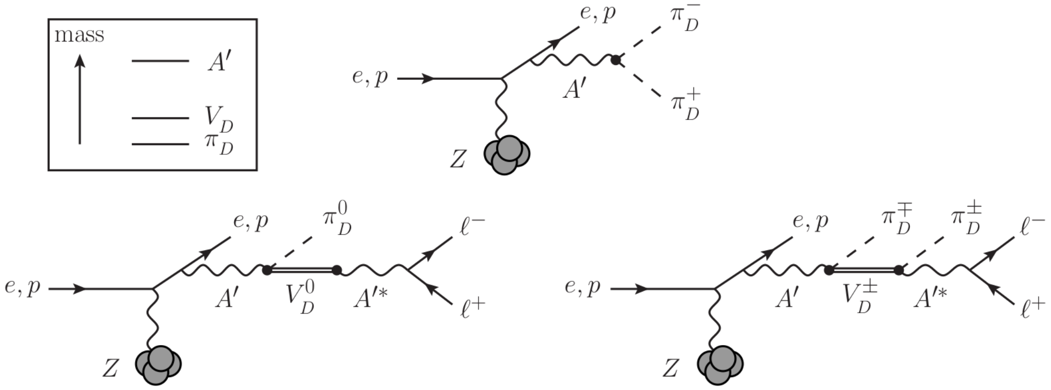


Figure 2.7: Production of hidden sector mesons for electron/proton fixed target experiments. The A' is produced through a "dark Bremsstrahlung" process and promptly decays to a pair of hidden sector mesons. The top diagram shows the invisible decay of A' to two hidden sector pions. The bottom left diagram shows the A' decay to a neutral dark vector and dark pion pair, followed by the 2-body vector decay to two leptons. The bottom right diagram shows the A' decay to a $U(1)_D$ charged vector/pion pair, followed by the 3-body decay of the charged vector meson to a dark pion and two leptons. Figure taken from Ref. [36]

The A' can decay into the HS through three additional processes:

1. $A' \to V_D^0 \pi_D^0$
2. $A' \to V_D^\pm \pi_D^\mp$
3. $A' \to \pi_D \pi_D$

These processes are shown in Figure 2.7, and the following rate calculations are defined in Ref. [36].

The decay $A' \to \pi\pi$ is invisible and therefore not detectable by the HPS experiment. The other two processes involve A' decays to dark vector mesons. V_D^0 refers to the neutral dark vector mesons, ρ_D and ϕ_D, which mix with the A' and decay into e^+e^-. $V_D^\pm$ denotes the remaining vector mesons that are charged under $U(1)_D$ and undergo three-body visible decay processes through an off-shell A'. The invisible A' decay rate to two dark pions is given by:

$$\Gamma(A' \to \pi\pi) = \frac{2\alpha_D}{3} m_{A'} \left(1 - \frac{4m_{\pi_D}^2}{m_{A'}^2}\right)^{3/2} \left(\frac{m_{V_D}^2}{m_{A'}^2 - m_{V_D}^2}\right)^2. \tag{2.9}$$

The A' decay to two dark vector mesons is forbidden by $m_{A'} < 2m_{V_D}$, however the rate is included here for completion, given by:

$$\Gamma(A' \to VV) = \frac{\alpha_D}{6} f(r) m_{A'}, \tag{2.10}$$

where $r = m_{V_D}/m_{A'}$ and

$$f(r) = \left(\frac{1 + 16r^2 - 68r^4 - 48r^6}{(1 - r^2)^2}\right) \sqrt{1 - 4r^2}. \tag{2.11}$$

The A' decay rate to V_D π_D is given by:

$$\Gamma\left(A' \to V_D \pi_D\right) = \frac{\alpha_D T_{V_D}}{192\pi^4} \left(\frac{m_{A'}}{m_{\pi_D}}\right)^2 \left(\frac{m_{V_D}}{m_{\pi_D}}\right)^2 \left(\frac{m_{\pi_D}}{f_{\pi_D}}\right)^4 m_{A'} \beta(x, y)^{3/2}, \tag{2.12}$$

where $x = m_{\pi_D}/m_{A'}, y = m_{V_D}/m_{A'}$, and

$$\beta(x, y) = (1 + y^2 - x^2 - 2y)(1 + y^2 - x^2 + 2y). \tag{2.13}$$

The factor T_{V_D} corresponds to the different types of final state vector mesons,

$$T_V = \begin{cases} 3/4 & \rho_D \\ 3/2 & \phi_D \\ 18 & \text{total.} \end{cases}, \tag{2.14}$$

Figure 2.8 illustrates how the A' branching ratios vary with the ratio m_{π_D}/f_{π_D}, where the remaining SIMP parameters have been fixed as described in Section 2.5.1. When

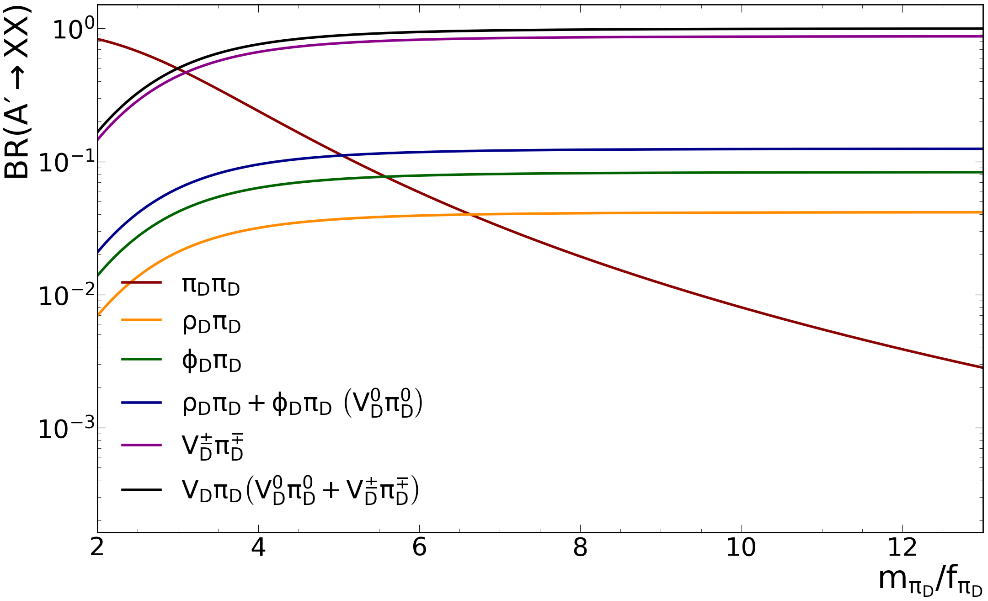


Figure 2.8: Branching ratios for A' to hidden sector mesons as a function of m_{π_D}/f_{π_D}, where the remaining SIMP parameters are fixed to: $m_{A'}/m_{\pi_D} = \frac{3}{1.8}, m_{V_D}/m_{\pi_D} = 1.8,$ and $\alpha_D = 0.01$. The green line represents the total branching ratio of the A' to all hidden sector vector/pion pairs (charged and neutral). The blue line represents the combined branching ratios of the two neutral dark vector mesons $\rho_D + \pi_D$ and $\phi_D + \pi_D$. The red line shows the branching ratio for A' to two dark pions, which is invisible. The relative rate of invisible to visible decays increases as m_{π_D}/f_{π_D} increases.

$m_{\pi_D}/f_{\pi_D} \lesssim 3$, the invisible decay process $A' \to \pi_D\pi_D$ dominates the branching fraction and suppresses the visible signal. However, as m_{π_D}/f_{π_D} approaches its perturbative limit of 4π, the visible branching fraction asymptotically approaches one, while the invisible decay rate decreases towards zero. Given this relationship, the SIMP analysis focuses on two benchmark cases, where $m_{\pi_D}/f_{\pi_D} = 3$ and $m_{\pi_D}/f_{\pi_D} = 4\pi$, corresponding to both extremes of the branching fraction spectrum.

The visible branching fractions are categorized into two main types: charged and neutral dark vector mesons. The maximum combined branching fraction for the neutral

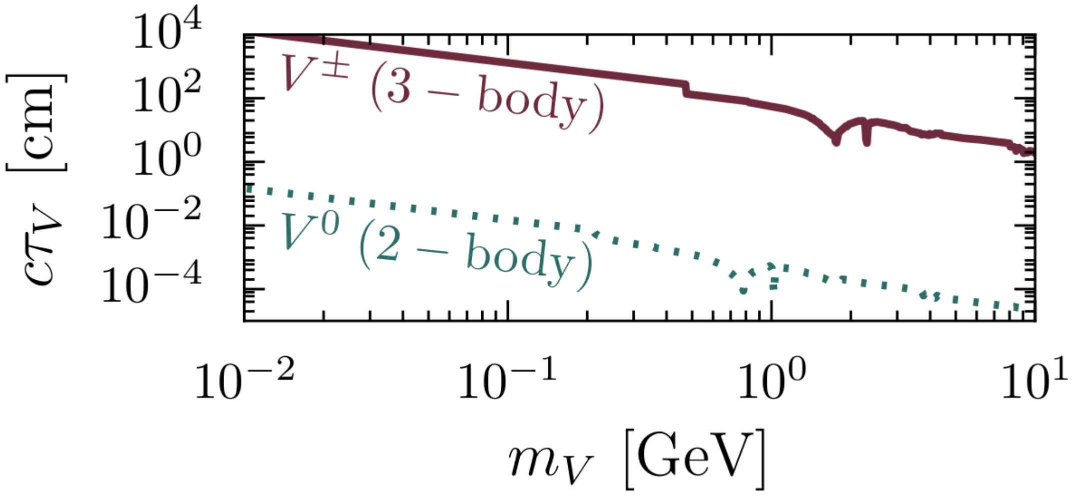


Figure 2.9: Proper lifetime of neutral (dotted) and charged (solid) dark vector mesons as a function of mass. The SIMP parameters are fixed to: $\alpha_D = 0.01, \epsilon = 10^{-3}$, and $m_{\pi_D}/f_{\pi_D} = 3$. The lifetime of the 3-body decay is many orders of magnitude larger than the 2-body case. Figure taken from Ref. [36].

dark vectors ($\rho_D + \phi_D$) is approximately 11 %, while for charged dark vectors, it is significantly higher, nearly 90 %. This may initially suggest that the charged dark vectors dominate the expected signal in this analysis. However, the charged vectors are ultimately excluded from this analysis due to their complicated kinematics and reduced acceptance within the HPS detector. Figure 2.9 shows the proper lifetime of the neutral (2-body) and charged (3-body) decaying dark vectors, again with the SIMP parameters fixed to those in Section 2.5.1, and $\epsilon = 10^{-3}$.

Detecting the charged dark vector mesons presents several challenges to the HPS experiment. The charged dark vector meson lifetimes are several orders of magnitude longer than those of their neutral counterparts. HPS is primarily designed to search for displaced decays within approximately 10 cm of the target, making it particularly sensitive to charged dark vectors for very small values of ϵ. However, the A' production rate scales with ϵ^2, resulting in a low charged dark vector production rate within the HPS detector acceptance downstream of the target. Additionally, the 3-body decay process leads to a significant transverse momentum kick from the additional π_D production, This tends to place the signal outside of the detector's acceptance in the transverse direction. Furthermore, the signal associated with charged dark vectors is non-resonant, further complicating the analysis. As a result of these complexities, this analysis excludes visible signals from charged dark vectors by restricting their masses to $m_{V_D^\pm} > 2m_{\pi_D}$, ensuring that they decay invisibly to two dark pions. Therefore, this analysis focuses exclusively on the 2-body decay of neutral dark vector mesons, ρ_D and ϕ_D.

The neutral dark vector decay rate to a pair of leptons is given by:

$$\Gamma(V \to \ell^+\ell^-) = \frac{16\pi\alpha_D\epsilon^2 f_{\pi_D}^2}{3m_{V_D}^2}\left(\frac{m_{V_D}^2}{m_{A'}^2 - m_{V_D}^2}\right)^2 \sqrt{1 - \frac{4m_\ell^2}{m_{V_D}^2}}\left(1 + \frac{2m_\ell^2}{m_{V_D}^2}\right) m_{V_D} \times \begin{cases} 2 & \rho_D \\ 1 & \phi_D \end{cases} . \tag{2.15}$$

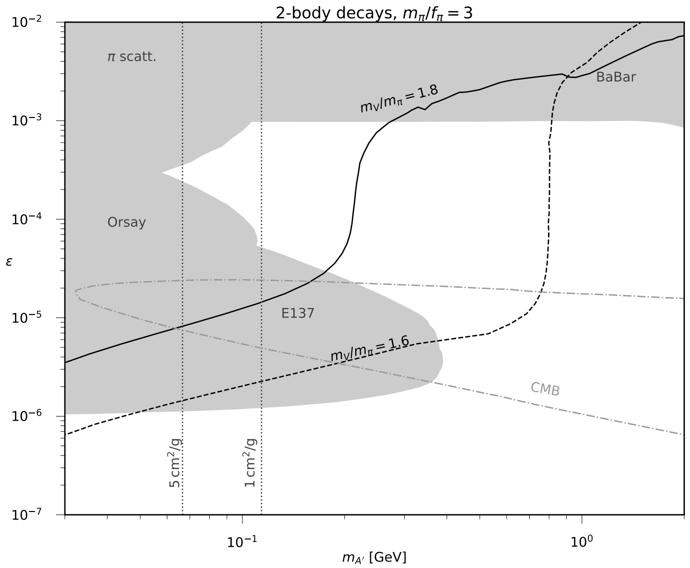
Since the neutral dark vectors have different A' branching fractions and lifetimes, the expected signal rate contribution from each particle is calculated separately and combined to get the total expected signal rate as a function of the A' mass and kinetic mixing strength ϵ. The expected signal calculation is derived later in Section 5.1.4.

2.7 SIMP Search Overview

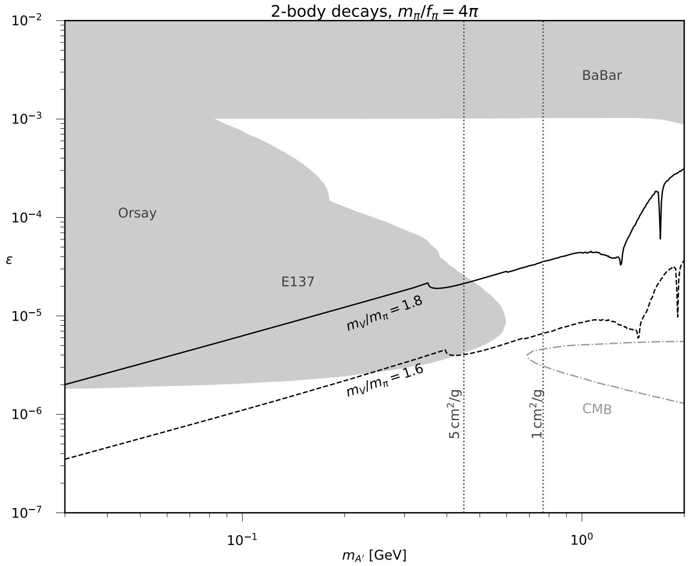
Figure 2.10 presents existing constraints on strongly interacting hidden sectors with parameters fixed as described in Section 2.5.1 for two benchmark values of m_{π_D}/f_{π_D}. The solid and dashed black lines represent the thermal target contours. Variations in the

V_D-π_D mass splitting for a given choice of m_{π_D}/f_{π_D} have a strong impact on the thermal contours, but not on the constraints. These constraints were calculated in Ref. [36] using a combination of bounds on the self-scattering cross section from astrophysical observations, CMB anisotropies, existing beam dump data, and collider searches for minimal A' decays, recasted to include A' decays to hidden sector pions and vector mesons.

The presented constraints specifically apply to a mass hierarchy where dark vector mesons decay to the SM exclusively through two-body processes ($V_D \rightarrow \ell^+\ell^-$), which is the scenario analyzed in this dissertation. The thermal targets, shown as solid and dashed black lines, vary with the V_D-π_D mass splitting and m_{π_D}/f_{π_D}. However, the exclusions themselves do not alter significantly with small variations in m_{V_D}/m_{π_D}. Large A' masses and large coupling strengths ϵ are largely excluded by dark photon missing momentum and energy searches conducted by the BaBar experiment [36]. The signal associated with A' production and invisible decay is consistent with both $A' \rightarrow \pi_D\pi_D$ and $A' \rightarrow V_D\pi_D$, where the V_D visibly decays outside of the detector. Conversely, the SIMP parameter space corresponding to sub-GeV masses and smaller coupling strengths (longer V_D lifetimes) is constrained by existing beam dump experiments. The SLAC E137 beam dump experiment, for example, was designed to search for long-lived particles that are neutral under the SM. E137 used a 20 GeV electron beam incident on a fixed target [40]. The long-lived particles were expected to travel through 179 m of dirt shielding before entering a 204 m long open decay region with an electromagnetic shower calorimeter located approximately 400 m downstream of the target. The experiment observed no signal events but was sensitive to long-lived V_Ds decay to e^+e^- via both two-body and three-body processes, thus excluding the regions with $\epsilon \sim 10^{-4} - 10^{-6}$, as shown in Figure 2.10. The constraints from beam



(a) 2 body decays with $m_\pi / f_\pi = 3$.



(b) 2 body decays with $m_\pi / f_\pi = 4\pi$.

Figure 2.10: Existing constraints on strongly interacting hidden sectors calculated in Ref. [36], assuming parameters defined in Section 2.5.1. (a) Shows $m_{\pi_D}/f_{\pi_D} = 3$, and (b) shows $m_{\pi_D}/f_{\pi_D} = 4\pi$. The two dashed vertical lines represent constraints on the self-scattering cross-section per DM mass from galaxy cluster observations. The dot-dashed gray contour is excluded by CMB measurements.

dump experiments are further extended to larger values of ϵ (shorter lifetimes) by shorter baseline experiments such as Orsay [41].

Another significant contribution to the excluded SIMP parameter space comes from searches for DM production via $A' \rightarrow \pi_D \pi_D$ followed by DM scattering. For example the high-luminosity proton fixed-target experiments at LSND and MiniBooNE were expected to generate a high intensity "DM beam" of particles boosted along the proton beam direction. This beam would be detectable through DM scattering off of electrons (LSND and MiniBoone) and nucleons (MinBoone) [42]. These constraints, combined with similar constraints from E137, are labeled as "$\pi -$ scatt." in Figure 2.10. It is evident that experiments well-suited to searching for heavy photon production are also ideal for probing the SIMP hidden sector, and much of the SIMP parameter space has already been explored through such experiments. However, a significant portion of the cosmologically motivated parameter space remains unexplored. This region lies primarily where $50\,\mathrm{MeV} < m_{A'} < 1\,\mathrm{GeV}$ with coupling strengths $10^{-5} < \epsilon < 10^{-3}$. This range in ϵ is particularly well-suited for precision vertexing fixed-target experiments searching for long-lived particles that visibly decay in the range of $\mathrm{mm} - \mathrm{cm}$, such as the Heavy Photon Search (HPS).

2.8 HPS Overview

HPS is a fixed target electron beam experiment hosted in Hall B at the Thomas Jefferson National Lab (JLab). It is designed to search for visibly decaying A's in the mass range 20 MeV to 220 MeV with a SM coupling strength $\epsilon^2 > 10^{-10}$. The experiment is unique in its ability to reconstruct displaced events with decay lengths of $\mathcal{O}(1\,\text{cm to } 10\,\text{cm})$ at millimeter-level precision, thanks to advanced silicon microstrip sensor technology that operates within approximately 500 µm of the high-energy electron beam. In 2016, HPS completed both a resonance and a displaced vertex search for the canonical A'. The resonance search successfully constrained the SM coupling to $\epsilon^2 \geq 10^{-5}$ for A' masses between 39 MeV to 179 MeV. However, the displaced vertex search did not achieve sufficient luminosity to set limits on the canonical A' cross-section [43].

The displaced vertex search was limited because the canonical A' production rate is suppressed by the kinetic mixing strength ϵ^2, which is directly related to the A' lifetime. Consequently, the production rate of A's that were long-lived enough to decay beyond the prompt QED background and be identifiable was too low. In contrast, for the SIMPs search, the additional degrees of freedom introduced by the self-interacting particles decouple the long-lived V_D lifetime from the production rate, significantly increasing the expected signal rate beyond the prompt QED background compared to the canonical A'. Despite the more complex SIMPs hidden sector involving a larger set of model parameters, which impacts HPS's sensitivity, HPS anticipates sensitivity in the 2016 data for one of the "benchmark" sets of SIMPs parameters used in the literature [36]. This expected sensitivity motivates this dissertation.

A significant portion of this dissertation is dedicated to developing SM background rejection techniques to minimize the falsely-displaced prompt QED backgrounds and enhance sensitivity to the signal across a broad range of masses and couplings (ϵ). It is important to emphasize that this dissertation uses data from the 2016 Engineering Run, conducted three year before the first official data run in 2019. HPS has since completed two additional data runs in 2019 and 2021. These runs feature significantly higher luminosities and an upgraded detector that improves the vertex resolution by approximately a factor of two. The 2019 and 2021 datasets are currently undergoing calibration, and are expected to significantly enhance HPS's ability to probe both the canonical A' and SIMPs hidden sectors.

Chapter 3

The Continuous Electron Beam Accelerator Facility and the HPS Detector

The HPS detector is a forward acceptance spectrometer consisting of a three magnet chicane, the Silicon Vertex Tracker, and the Electromagnetic Calorimeter, depicted in Figure 3.1. The SVT is an array of silicon micro-strip sensors housed in a stainless steel vacuum box roughly 1 m long. It is used to reconstruct the trajectory of charged particles as they traverse the approximately uniform magnetic field of the analyzing magnet. The Ecal is an array of Lead-Tungstate crystals attached to avalanche photodiodes. It used to reconstruct particle energy and trigger on events of interest with precision timing. HPS searches for highly-boosted e^+e^- decays with opening angles on the order of tens of mrad, with the SVT and Ecal divided into top and bottom volumes that closely straddle the through-going electron beam from above and below. The vertical proximity of the detector

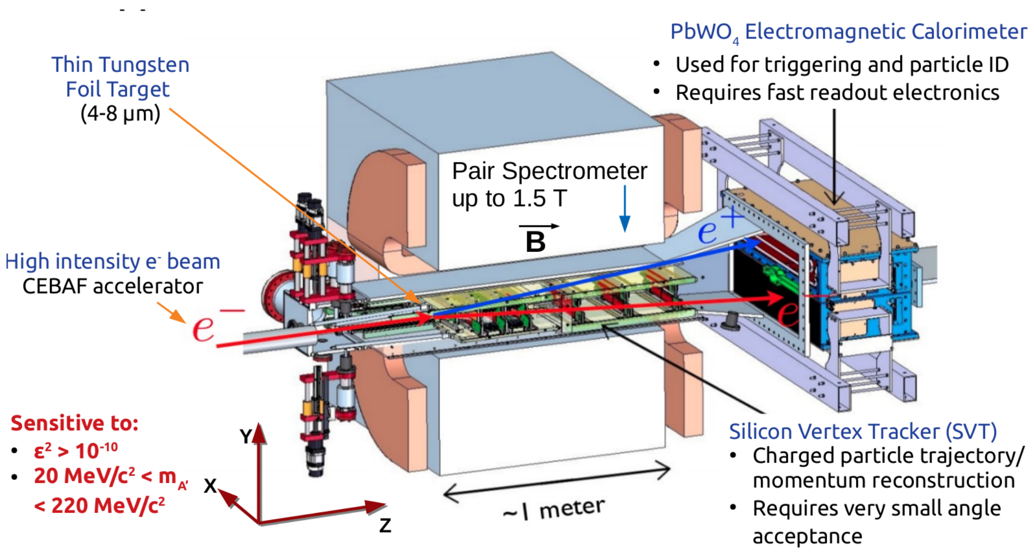


Figure 3.1: Schematic of the HPS detector.

to the beam is limited by the intense flux of multiple Coulomb scattered electrons and secondary radiatives from the target, creating a so-called "dead-zone" near the beam.

In order to achieve the necessary vertex resolution for identifying rare displaced events beyond the prompt QED background, the first SVT silicon tracking layer is placed as close to the target as possible. However, the closer the SVT is placed to the target, the larger the opening angle must be to avoid the dead-zone. Balancing angular acceptance, vertex resolution, and radiation exposure, the SVT is designed with a 15 mrad opening angle. In the 2015-2016 iterations of the SVT, the first silicon sensor layer is positioned approximately 10 cm downstream from the target, and as close as 0.5 mm vertically from the electron beam. The upgraded version of the SVT constructed in 2019 positions the first layer approximately 5 cm from the target with the same vertical distance, improving the vertex resolution by roughly a factor of two. To minimize radiation damage and decrease multiple scattering which degrades the vertex resolution, the tracking volume material budget is minimized. Radiation damage to the Ecal is managed by removing crystals from the region that has the largest beam background occupancy.

3.1 CEBAF

The Continuous Electron Beam Accelerator Facility (CEBAF) is a superconducting radio-frequency accelerator capable of simultaneously delivering an electron beam with energies up to 11 GeV to three experimental halls (A, B, and C), and up to 12 GeV to Hall D [44]. Figure 3.2 shows the layout of the CEBAF halls, beam-pipes, and cryomodules. The electron beam is produced using four independent lasers (one for each Hall) that are combined through a series of polarization filters into a collinear beam incident on the electron gun photo-cathode. The resulting electrons are passed into the injector, which provides an additional 108 MeV boost before the beam is injected into the accelerator. The accelerator consists of two linacs that each provide an 1100 MeV energy gain, and two arcs that steer the beam around the accelerator. Beam delivered to Halls A-C may circulate the accelerator for up to 5 passes (for a maximum energy of 11 GeV), where the number of passes is determined by the energy requirements of a given hall. The beam bunches are delivered to Halls A, B, and C independently by vertically separating the combined beam using a 500 MHz RF separator. A "5th pass separator" operating at 750 MHz separates the Hall D beam from the others, and allows an additional 0.5 pass through the accelerator, delivering a maximum beam energy of 12 GeV to Hall D exclusively [45][46].

Figure 3.3 shows a diagram of the HPS detector installation in the Hall B alcove, with the electron beam entering from the left of the diagram. It is critical that the elec-

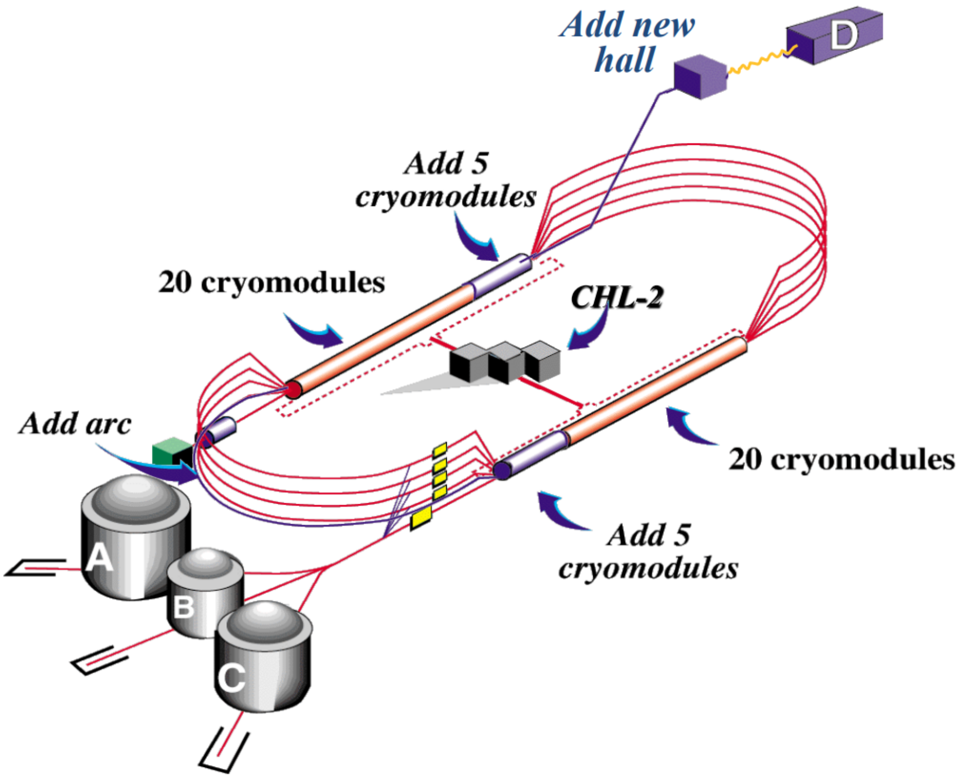


Figure 3.2: Schematic of the Continuous Electron Beam Accelerator Facility (CEBAF) linear accelerators and steering arcs. Beam is delivered simultaneously to four halls at JLab. This image shows the now completed inclusion of additional hall, Hall D, as well as new cryomodules. [44]

tron beam does not strike the HPS detector or any upstream materials that could scatter radiation into the detector. Therefore, establishing a safe beam position is paramount. During beam alignment, the SVT is retracted from the beamline and shielded by a series of collimators for protection. The beam position is monitored at various points along the beamline. The beam first enters Hall B through the Hall B collimator (12 mm or 20 mm), located roughly 40 m upstream from the detector. This collimator is designed to prevent direct beam exposure of the SVT in the event of a beam excursion upstream of Hall B. Beyond the collimator, the beam halo (the tails of the beam) is monitored using beam halo

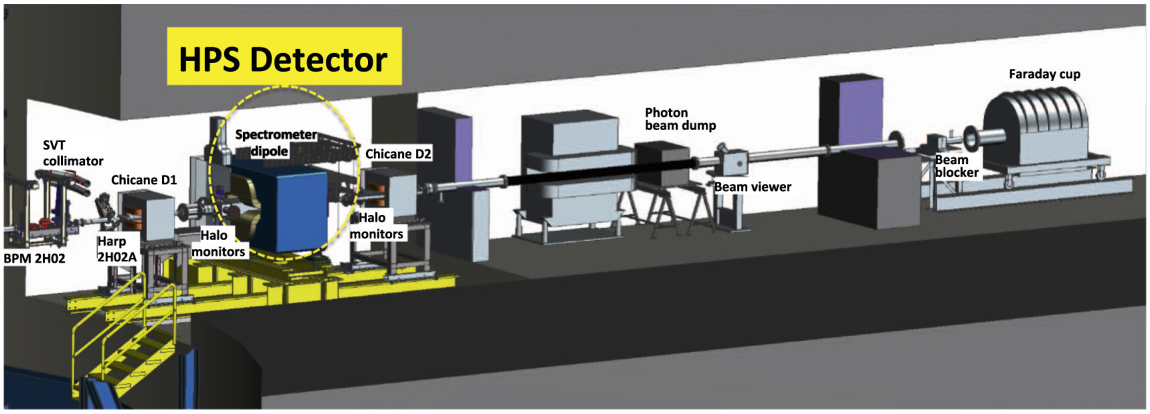


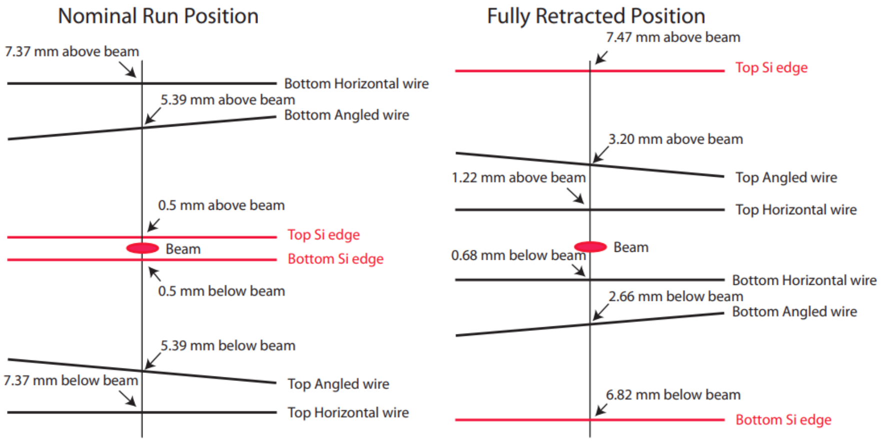
Figure 3.3: Diagram of HPS installation in the alcove at Hall B showing the full chicane. The electron beam enters from the left, and ends at the Faraday cup on the right, which measures the beam current and provides a visual marker for the beam position. Beam-position monitors (BPM) 2H02 and Harp 2H02A are used to measure the beam position along the beamline. The halo monitor counts are used to check for inadvertent beam scattering from things such as the beam scraping the beam-pipe.

counters that measure scattered electrons along the beamline. These counters are designed to trigger a beam shutoff if the count rate exceeds ∼ 100 Hz, indicating that the beam is "scraping" some material and need to be tuned accordingly.

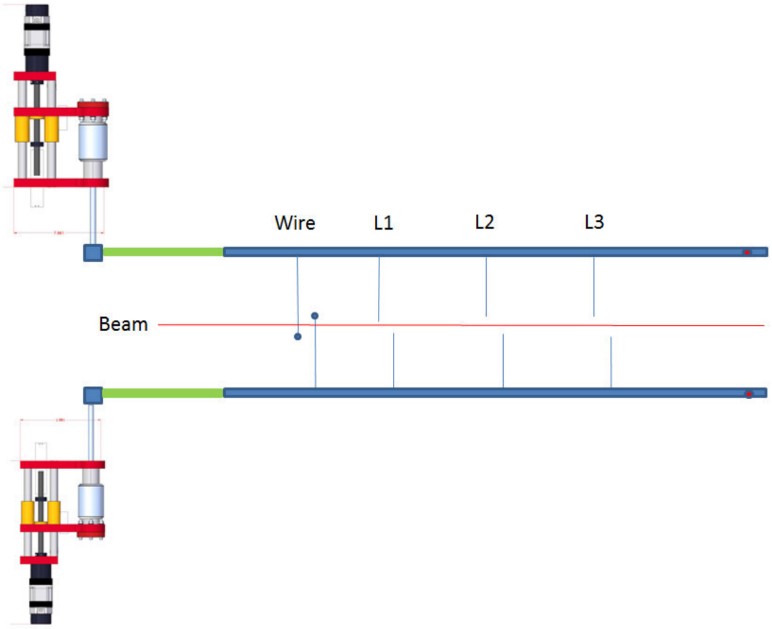
Beyond this collimator, the beam is passed through the SVT protection collimator, of which there are three different sizes: $8 \times 10\text{mm}^2$, $2.82 \times 10\text{mm}^2$, and $2.25 \times 10\text{mm}^2$. Once the beam is successfully passed through the largest SVT collimator, it is transported to the "Faraday Cup," which provides a visual aid to assess the beam position and shape. The Faraday Cup is also used to measure the beam current. The beam shape and position are then iteratively tested and changed as necessary to ensure that the beam quality is acceptable and centered within the SVT protection collimator.

After a high-quality beam is established (the beam is well-centered, visible on the Faraday Cup, has minimal tails, and is on the order of 100 µm in x and y), the beam is finely tuned relative to the SVT detector. The wide SVT protection collimator is replaced with one of the smaller versions used for data collection. The beam profile (position and transverse shape) is measured using two “SVT wire frames,” each attached to the top and bottom volumes of the SVT assembly. The wire frames house two tungsten wires: a 20 µm horizontal wire and a 30 µm wire tilted approximately 9 degrees relative to the horizontal.

The profile of the beam at the front of the SVT is measured by scanning the position of the entire SVT and moving the wires into the path of the electron beam. Beam electrons are scattered by the wires into a set of halo counters, where the count rate is proportional to the intensity of the beam. The count rates are used in combination with the known mapping of the SVT wires to measure the mean and Gaussian width of the beam, as demonstrated for the vertical direction in Figure 3.5. To safely operate the SVT, the beam must be centered vertically within 50 µm of the plane between the top and bottom volumes of the SVT. The beam profile must be less than 50 µm wide in the vertical direction and less than 150 µm wide in the horizontal direction. The tighter constraint in the vertical direction is due to the SVT’s operating position, 500 µm from the beam plane. A tightly constrained beamspot also aids in the analysis selection by requiring that reconstructed vertices project back to the beamspot on the target, as discussed in Section 5.2.2.3.



(a) Beam's eye view of the SVT wire frame.



(b) Side view of the SVT wire frame.

Figure 3.4: Diagrams of the SVT wire frame system used to measure the beam profile and establish quality physics beam before moving the SVT into operating (nominal) position. Figures taken from Ref. [47].

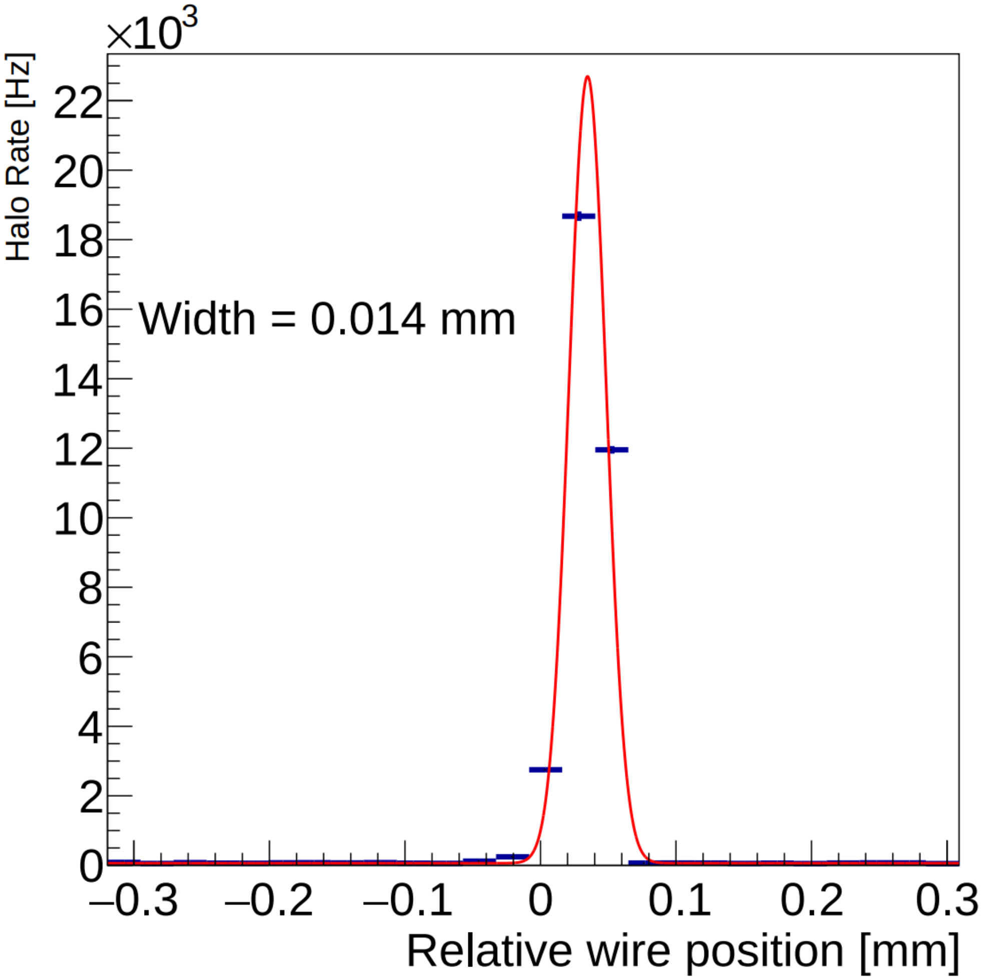


Figure 3.5: Example of a vertical beam profile measurement using the SVT wire scan. Figure taken from Ref. [1].

3.1.1 Datasets

HPS performed engineering runs in 2015 and 2016, and data runs in 2019 and 2021. The analysis presented in this thesis uses the 2016 engineering data taken with a 2.3 GeV electron beam incident on an 8 µm tungsten foil target at a beam current of 200 nA, for a total luminosity of $\sim 10\,753\,\mathrm{nb}^{-1}$. 81 of the 2016 data runs are considered "golden runs," meaning the runs used the production-level trigger, target, and beam current, the SVT was operating at its nominal position of 0.5 mm from the beam-axis, and the data acquisition (DAQ) was functioning properly. The total luminosity of the 2016 golden runs is $10\,703.81\,\mathrm{nb}^{-1}$. The 2015 and 2016 runs are summarized in Figure 3.6.

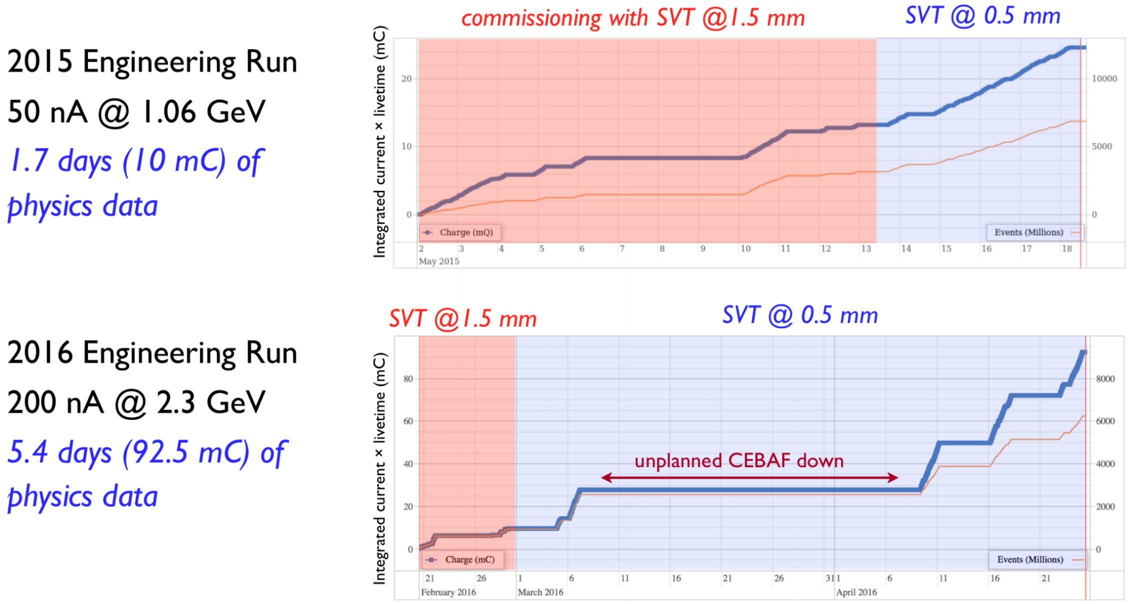


Figure 3.6: Plot of integrated beam current × DAQ livetime over the duration of the HPS run. (Top) 2015 Engineering Run. (Bottom) 2016 Engineering Run. Blue shading indicates where the SVT was operating at nominal position with layer 1 500 µm from the beam in the vertical direction.

Also presented in this thesis is work related to the detector and tracking algorithm upgrades implemented ahead of the 2019 run, which operated at a beam energy of 4.55 GeV, with a ~ 150 nA beam current incident on an 8 µm tungsten foil target for "low-luminosity" runs, and a 20 µm tungsten foil for "high-luminosity" runs. The 2015 and 2016 engineering runs are summarized in Figure 3.6, and the 2019 run is summarized in Figure 3.7.

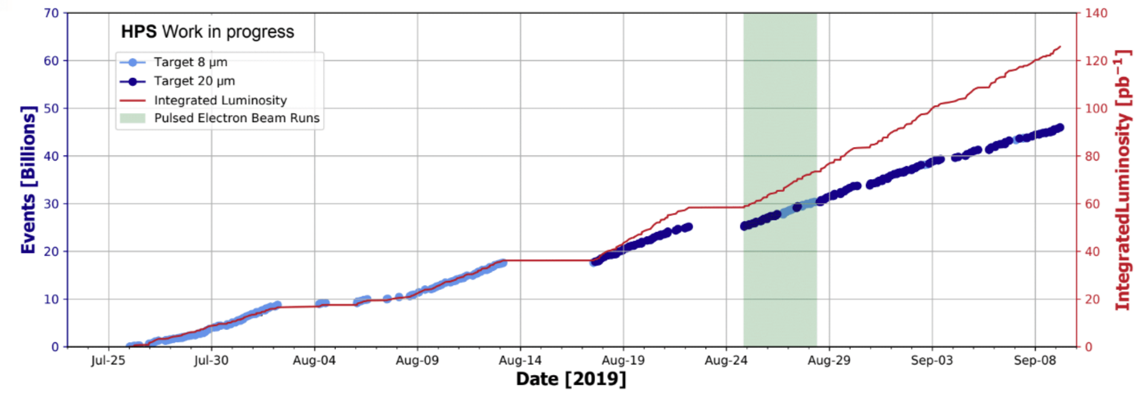


Figure 3.7: Plot showing the number of recorded events (left axis) and integrated luminosity (right) axis for the HPS 2019 Run. Two targets thicknesses were used in 2019, the "low luminosity" 8 µm thick target, and the "high luminosity" 20 µm thick target.

3.2 Silicon Vertex Tracker (SVT)

The SVT is responsible for measuring the trajectory and momentum of charged particles as they travel through the uniform magnetic field of the Pair Spectrometer. The reconstructed particle trajectories, referred to as “tracks,” are extrapolated to the Ecal where they are matched with energy deposits (“Ecal clusters”). The reconstructed track and matched Ecal cluster form reconstructed particles. Oppositely charged particle tracks are then fit with a common vertex to create reconstructed vertex candidates. Ecal clusters that are not matched to a track are tagged as photons.

Figure 3.8 shows a diagram of the SVT tracking layers inside of the vacuum box. The SVT is divided into two separate volumes (Top and Bottom) with a 15 mrad opening angle between them. Each volume in the 2015-2016 version of the SVT consists of six silicon-strip sensor tracking layers positioned at various locations downstream of the target along the z-axis of the HPS detector coordinate system, where the positive x direction is defined as beam-left and the positive y direction is defined as up. From this point onward, all coordinate references will be in the HPS detector coordinate system, unless otherwise stated.

Each SVT layer is composed of a pair of “half-modules,” the combination of either a single or a double silicon strip sensor. The sensors are mounted to a thin carbon fiber structure support with a hybrid readout circuit board, isolated from the carbon fiber support with a thin layer of Kapton insulation. Two half-modules are paired back-to-back with their strips angle relative to one another to form a complete module. The “axial” half-module is oriented with strips pointing along the x-direction, parallel to the beam plane,

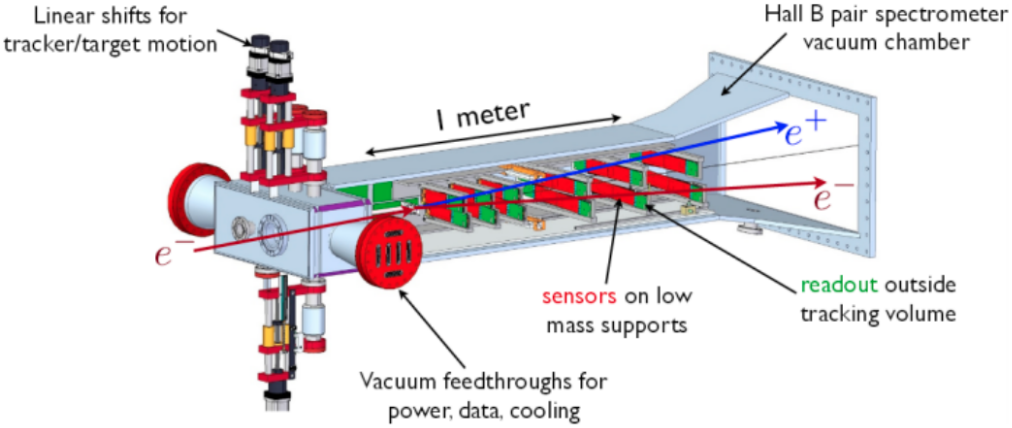


Figure 3.8: Diagram of the 2016 SVT, which consists of 6 silicon microstrip layers inside of a vacuum box, and submerged in a uniform magnetic field.

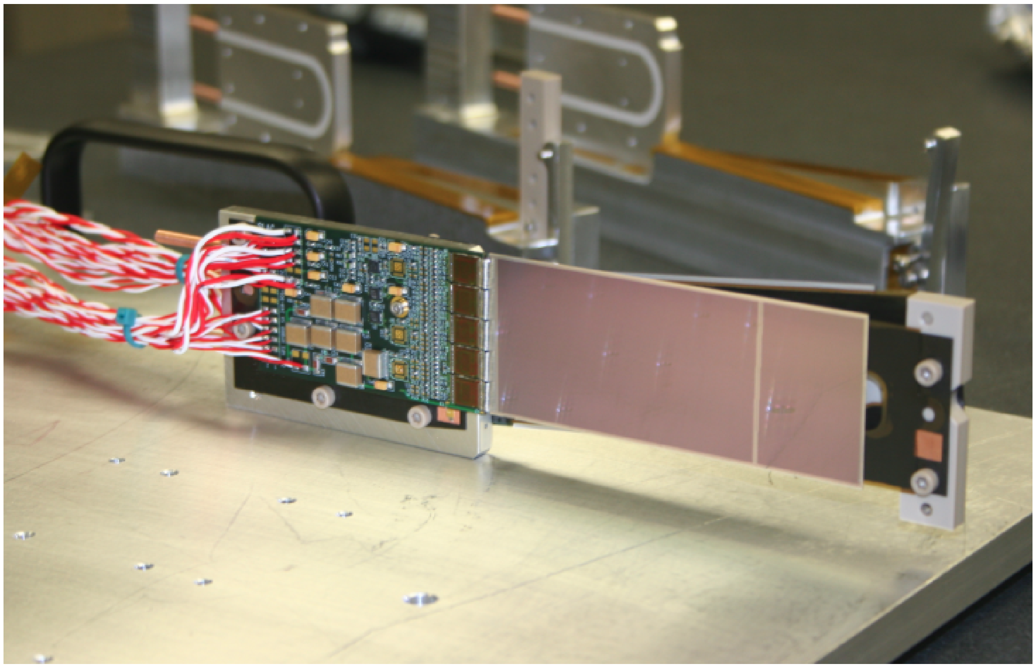
and provides position information in the vertical y-direction. The "stereo" half-module is rotated in the xy-plane by a small angle and, in combination with the axial sensor and the overall module position along z, provides three-dimensional position information. The stereo angle is designed to be 0.100 mrad for the first three layers, and 0.05 mrad for the last three layers. With the strips on adjacent layers overlapping at an angle, hits on adjacent strips can be used to provide two-dimensional position information at the layer position. The stereo angle also prevents the stereo sensors from extending into the high radiation dead-zone near the electron beam.

The modules are mounted on aluminum support modules, providing mechanical support and acting as a heat sink. The support modules use a spring pivot to apply constant tension on the half-modules and keep minimize thermal contraction effects and keep the half-modules flat at operating temperature (∼ 0 °C). The modules in the first three layers use single silicon strip sensors, while the last three layers use two silicon strip sensors joined end-to-end in the x-direction, increasing their acceptance to particles in the bending plane. Figure 3.9(a) shows an image of a single-sensor module. Figure 3.9(b) shows all six of the double-sensor modules laid out on a table in the clean-room. The modules are installed in the SVT using aluminum "U-channel" supports. Three modules each are mounted to a single U-channel, for a total of four U-channels corresponding to L1-3 and L4-6 in the top and bottom volumes. The previously described SVT wire frames are mounted to the front of the L1-3 U-channels, which are guided by motion levers that move the sensors and wire frames towards/away from the electron beam.

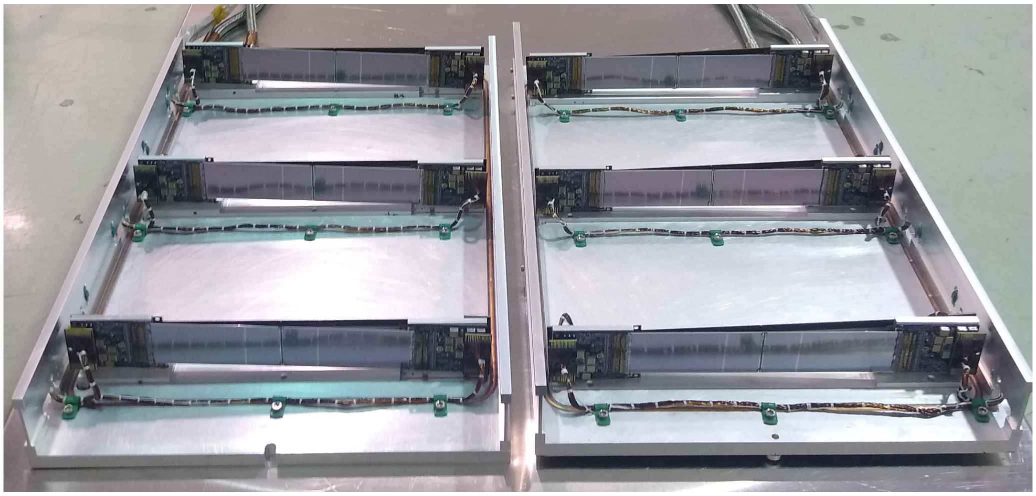
The first three SVT layers are spaced 100 mm apart, while layers 3-6 are spaced 200 mm apart. The first layer is placed as close to the target as possible to maximize the detector vertex resolution. However, the distance is limited by the maximum allowed occupancy of 1 % to 2 %, which is the fraction of strips with hits above threshold. The active region of the silicon sensors is 1 mm from the edge, and therefore must be placed at least 1.5 mm away from the beam plane. Using the small angle approximation for the 15 mrad opening angle, the minimum distance between the first SVT layer and the target is $\sim \frac{1.5\,\mathrm{mm}}{15\,\mathrm{mrad}} = 10\,\mathrm{cm}$. The SVT design features are summarized in Table 3.1.

Layer Number	1	2	3	4	5	6
Distance z from target (mm)	100	200	300	500	700	900
Dead Zone Distance y (mm)	±1.5	±3.0	±4.5	±7.5	±10.5	±13.5
Number of Sensors	4	4	4	8	8	8
Stereo Angle (mrad)	100	100	100	50	50	50
Bend Plane Resolution (μm)	≈ 60	≈ 60	≈ 60	≈ 120	≈ 120	≈ 120
Non-Bend Plane Resolution (μm)	≈ 6	≈ 6	≈ 6	≈ 6	≈ 6	≈ 6
Material Budget (%X_0)	0.7	0.7	0.7	0.7	0.7	0.7
Module Power Consumption (W)	6.9	6.9	6.9	13.8	13.8	13.8

Table 3.1: Summary of SVT Layer Specifications



(a) Two single-sensor half-modules mounted back to back.



(b) Six complete double sensor modules

Figure 3.9: Photographs of (a)two single-sensor half-modules (used in layers 1-3), and (b) all six complete double-sensor modules for layers SVT layers 4-6. The axial and stereo sensors are mounted back-to-back on their carbon fiber support structures, forming a complete module.

3.2.1 SVT Sensors and Readout

The silicon strip sensors employed by HPS are designed to minimize the material budget (therefore reducing multiple scattering), and to reliably operate in the high radiation environment near the electron beam, necessitating a high breakdown voltage and a low leakage current under prolonged irradiation. HPS uses single-sided p^+ on n-bulk silicon microstrip sensors that were originally developed for the DØ collaboration at Fermilab [48], designed to have a breakdown voltage greater than 350 V after irradiation. The sensor wafers are $\sim 320\,\mu\text{m}$ thick, with cut dimensions of $100\,\text{mm} \times 40.3\,\text{mm}$. They have an n-type bulk very lightly doped with phosphorus to achieve high resistivity and allow the bulk to be fully depleted by the bias voltage, and p-type implant strips 8 µm wide. There are 1277 implant strips with a 30 µm strip pitch, however, only every other strip is read out, resulting in 639 readout strips with a 60 µm readout pitch. The readout strips themselves are strips of aluminum that are AC coupled over the p-implants. The strips are biased using polysilicon resistors at both ends, and the p-implant strips are capacitively coupled to the aluminum readout strips. When a charged particle passes through the silicon sensor, it ionizes the bulk material and creates electron-hole pairs that are swept into the aluminum readout strips, generating a short current pulse. The strip current is read out by the APV25 readout chip, originally developed for use in the CMS tracker [49][50].

The APV25 chip schematic is shown in Figure 3.10. It has 128 input channels, each equipped with a pre-amplifier, a CR-RC shaper, and a 192 cell analog pipeline. The chip multiplexes all 128 channel signals onto a single differential output current. The channel pre-amplifier integrates the strip current pulse into a voltage step impulse with a long decay. The CR-RC shaper then transforms the voltage signal into a pulse with a peaking time and a fixed pulse shape that is independent of the hit charge collected. Since the pulse shape is independent of the signal magnitude, the pulse height is directly proportional to the amount of charge collected, and therefore to the energy of the charged particle.

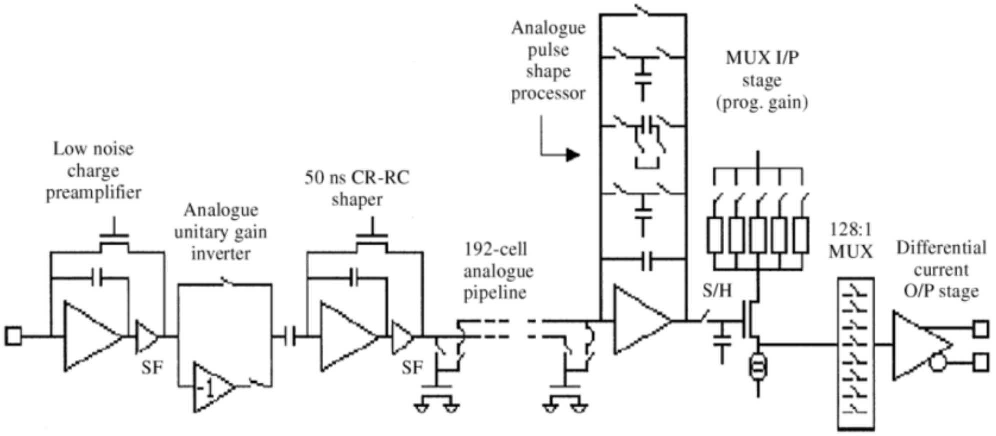


Figure 3.10: Block diagram of the APV25 readout chip [51]

It's critical for the hit pulse to return quickly to baseline in order to minimize the pile-up of successive hits, but it should not return so quickly that significant amplitude is lost. The "CR" differentiator sets the pulse duration using the calibrated decay time constant τ_{decay}. The "RC" integrator sets the rise time using the calibrated rise time constant τ_{rise}. The CR-RC circuit constants, along with the channel input capacitance load, define the hit response pulse shape for the channel, and are calibrated considering pile-up minimization, signal-to-noise ratio, charge collection efficiency, etc. The APV25 samples the shaper output of all 128 channels every clock cycle (24 ns at JLab), and stores the measurement in the 192-cell analog pipeline. HPS runs the APV25 in "multi-peak" mode [51], and reads six consecutive pipeline samples for each trigger, corresponding to a 144 ns hit readout window. The six ADC samples form a "raw SVT hit".

3.2.2 SVT DAQ

Each silicon sensor in the 2016 version of the SVT is mounted to a Hybrid readout circuit board that houses five APV25s, and the 639 sensor readout channels are directly wire-bonded to the APV25 readout channel pads (128 per chip). As mentioned previously, the first three layers consist of single-sensor half-modules, so each half-module requires a single Hybrid. The last three layer half-modules consist of two sensors joined end-to-end, and therefore require two Hybrids each. The Hybrid serves to filter the high voltage bias for the sensor, monitor the readout chip temperatures, and provides the interface for passing controls and data between the APV25s and the Front End Board (FEB), which is the next layer of hardware in the DAQ chain.

The FEBs also operate in vacuum, and are responsible for distributing controls, data, and both low-voltage power, and the high-voltage sensor bias, to the Hybrids. Each FEB can service 4 Hybrids, for a total of 10 FEBs required to operate the 36 Hybrids. Each FEB houses a pre-amplifier, a 14-bit ADC, and an Xilinx Artix-7 field programmable gate array (FPGA). The pre-amplifier converts the multiplexed APV25 differential current signals into a voltage compatible with the ADC, which samples the signal at 41.667 MHz, digitizes the signal, and passes it to the FPGA. The FPGA passes the clock and triggers downstream to the Hybrids, and monitors the Hybrids status and configuration. The FPGA also passes the ADC data upstream out of the vacuum through the vacuum penetrating data flange board (data flanges).

There is one data (vacuum) flange, with four flange boards, each capable of servicing three FEBs. The data and controls are passed between FEB and flange through mini SAS cables, where each cable has 4 data links (one for each Hybrid), and 2 control links. Each data flange converts the digitized signal from up to 3 FEBs (up to 12 Hybrids) to optical signal using QFSP fiber transceivers. This optical conversion is necessary to transfer the signal and controls out of the installation site to the Advanced Telecommunications Computing Architecture (ATCA) crate that houses the DAQ, located about 20 m away. The high and low-voltage power is also distributed from outside the vacuum to the FEBs through a single power flange with two flange boards through it, each connected to the FEBs through twisted pair cables. The Wiener MPOD power supplies are also located on the ATCA crate. The DAQ itself consists of two ATCA blades, each called a Cluster on Board (COB). The optical data and control signals interface with each COB through the COB's Rear Transmission Module (RTM) QSFP transceivers. Each COB has nine Data Processing Modules (DPM), and one Data Transport Module (DTM). Six of the DPMs run data reduction algorithms and build the events in data, while the fourth DPM processes upstream slow controls, such as commands to turn on FEBs, Hybrids, and individual APVs. The DTMs distribute the clock and triggers.

3.3 2019 and 2021 SVT Upgrade

HPS upgraded the SVT detector in 2019 by integrating an additional tracking layer closer to the target, utilizing state-of-the-art "slim-edge" sensor technology. Referred to as "L0," this new layer is positioned just $\sim 5\,\text{cm}$ downstream of the target. In addition, the original first layer was also replaced with the new slim-edge sensor modules. Results from a Monte Carlo study depicted in Figure 3.11 show that this upgrade improves the detector's vertex resolution by a factor of two [52].

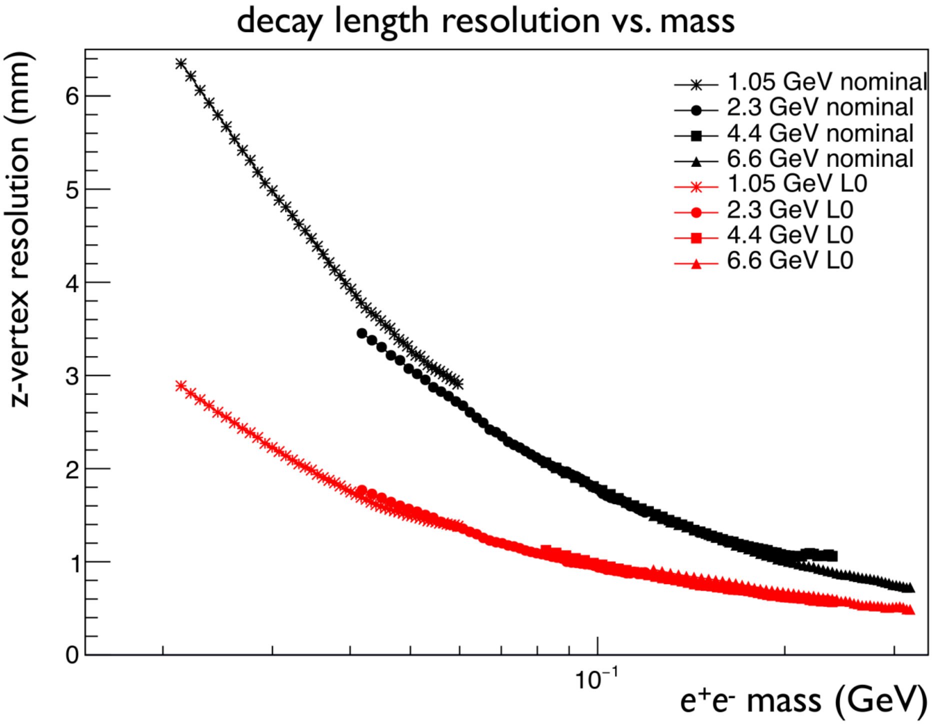


Figure 3.11: MC simulated z vertex resolution as a function of invariant mass. The red points show the expected improvement in vertex resolution after upgrading the SVT, compared to before in black.

3.3.1 Slim-Edge Sensor Technology

In order for L0 to operate with its active region at the designed 0.5 mm vertical distance from the beam, the inactive region of the sensor close to the beam had to be reduced from 1000 µm to 250 µm, otherwise the sensor material would be inserted into the beam path. This is illustrated in Figure 3.12. The thickness of the new sensors is reduced to 200 µm to limit scattering, and the strip length was shortened to manage increased occupancy. The occupancy is additionally reduced by dividing the new sensors into two isolated rows of strips.

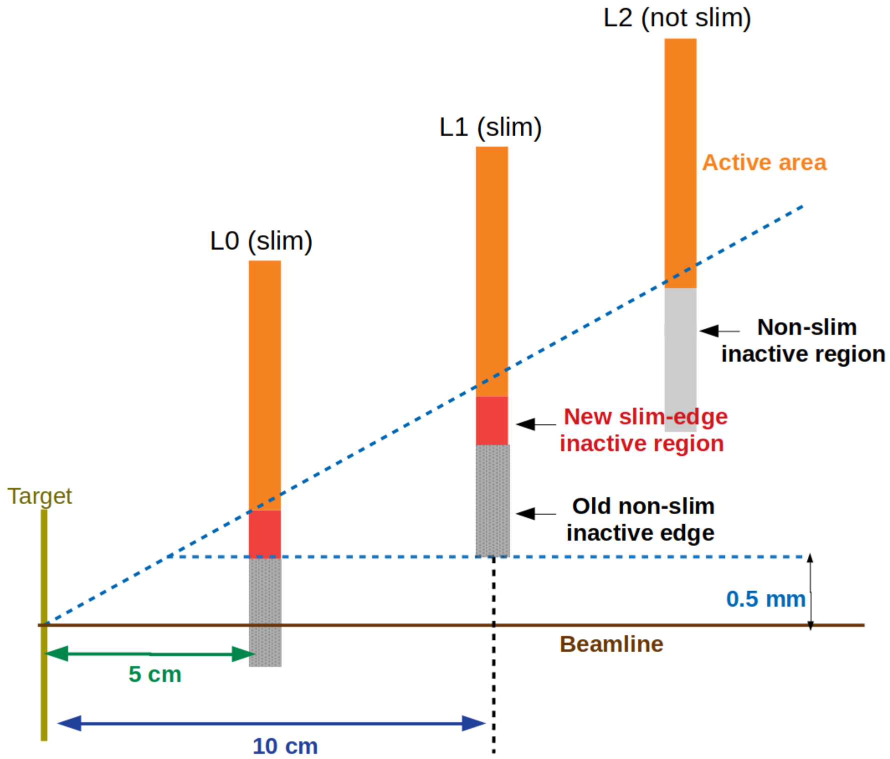


Figure 3.12: Schematic showing how the reduced inactive region of the new slim-edge sensors allows them to be placed closer to the target, given the designed 15 mrad SVT opening angle.

The slim-edge sensors were manufactured by Centro Nacional de Microelectronica (CNM) in Barcelona, Spain. The sensors are fabricated on 4 inch wafers, with six sensors per wafer, and are cut using a special "Scribe-Cleave-Passivate" method to achieve the designed 250 µm inactive material edge along the side of the sensor that is oriented towards the beam [53]. Figure 3.13 depicts an L0 hybrid with the cleaved edge of the sensor at the top of the image. The upgrade required a minimum of 8 slim-edge sensors to produce four new tracking layer modules, L0 and L1 in the Top and Bottom volumes. Ultimately, CNM produced 72 sensors, and 10 of them were selected to produce four modules and two spare half-modules. The selected sensors were chosen based on the criteria described in the following section.

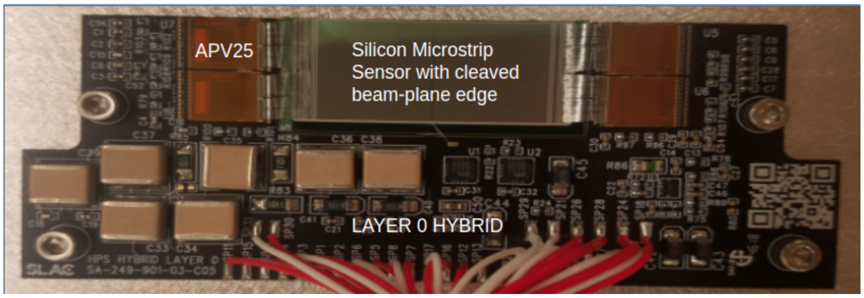


Figure 3.13: Photograph of a slim-edge sensor hybrid. The cleaved edge of the sensor is at the top, and is installed oriented towards the beam.

3.3.2 Sensor Quality Assurance

The sensors were imaged upon reception and examined for visible manufacturing defects and damage. Sensors with obvious physical damage, such as the one depicted in Figure 3.14, were pulled from the selection. The inactive region along the cleaved edge of the sensors was measured at various points along the edge to check for a consistent width less than 250 µm. Figure 3.15 shows an example of this measurement, where the imaged sensor has an inactive region 200 µm wide. It proved difficult to achieve the designed cleaved edge, and many of the sensors had an inconsistent edge orientation and width larger than desired. Due to timeline constraints, some sensors with edges larger than 250 µm were ultimately used.

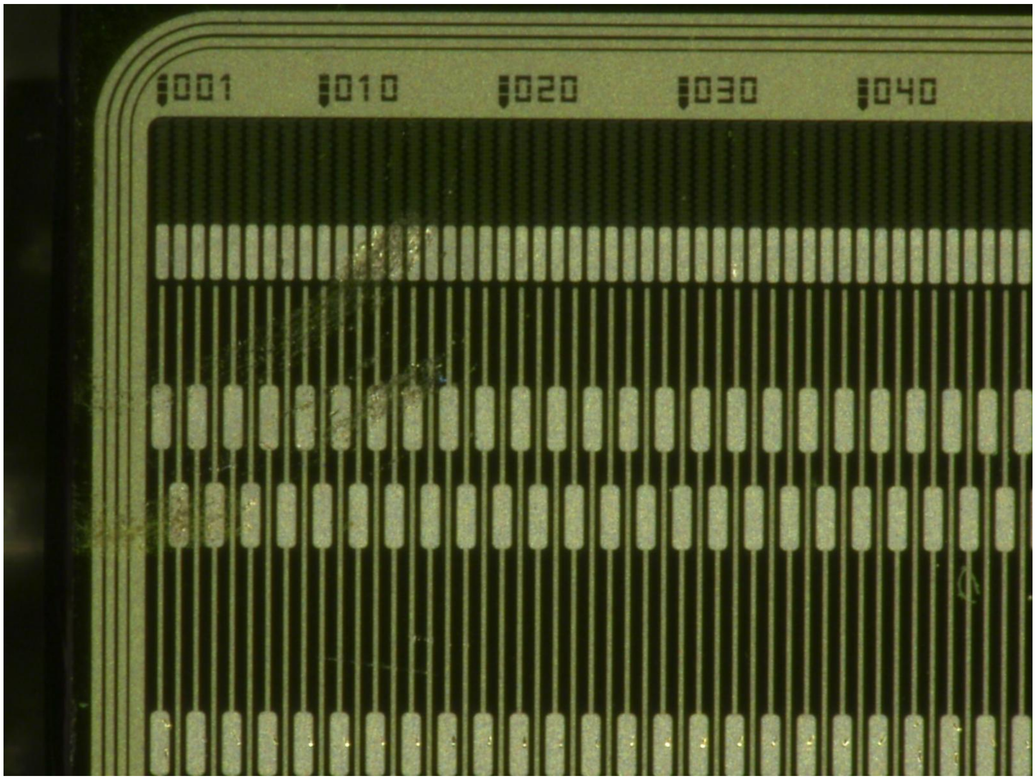
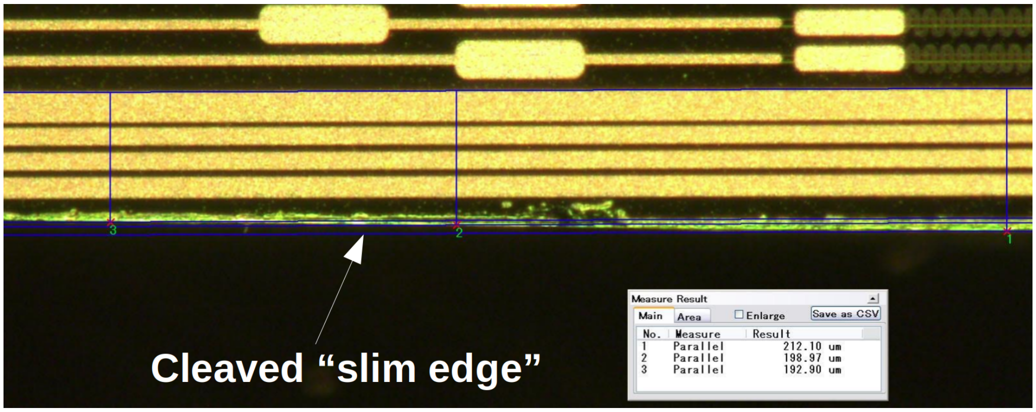


Figure 3.14: Microscopic photograph of the corner of a slim edge sensor candidate. This sensor is visibly damaged by scratches across multiple readout strips.

The sensor breakdown voltages were evaluated by reverse biasing the sensors with high-voltage and measuring their leakage current. The breakdown voltage is the voltage at which the leakage current increases rapidly, and is a critical selection criterion. These sensors are designed to operate in the extremely high occupancy region close to the target, and a high breakdown voltage is necessary to minimize leakage currents that may be induced by radiation damage. In addition, radiation damage over time will reduce the size



No.	Measure	Result
1	Parallel	212.10 um
2	Parallel	198.97 um
3	Parallel	192.90 um

Figure 3.15: Microscopic image of the cleaved edge of a slim-edge sensor candidate. The inactive material edge extends from the start of the guard ring out to the edge of the wafer, indicated by the horizontal blue lines. This sensor shows a consistent inactive region close to 250 µm.

of the sensor’s depletion region through a process called type-inversion, which effectively reduces the sensor’s sensitivity and charge collection efficiency. The size of the depletion region can be restored by increasing the reverse bias on the sensor, but this bias cannot exceed the breakdown voltage. Figure 3.16 shows the leakage currents as a function of bias voltage, with no radiation present, for a subset of the sensors. The breakdown voltage for sensor “W04-S3-d,” plotted in dark yellow, can be identified by the rapid increase in leakage current near 360 V.

The sensors were also evaluated based on the presence of shorts between the strip implant and the aluminum readout strip, called “pinholes.” Strips with pinholes do not provide signal, creating dead regions in the tracking detector, and are disconnected from the readout chips in the quality assurance stage of production to avoid interference with neighboring channels. The sensor production suffered from a relatively high rate of pinholes, such that their presence in the selected sensors was unavoidable. Instead, the sensors were evaluated primarily based on the location of the pinholes. The maximum track

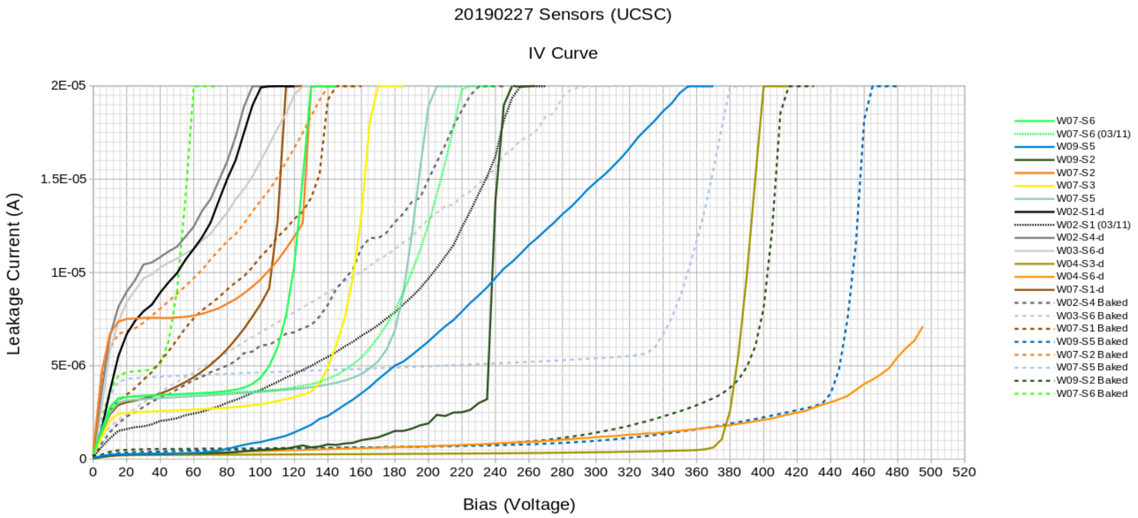


Figure 3.16: IV curves for a set of slim-edge sensor candidates. The sensor breakdown voltage is determined where the current rapidly increases.

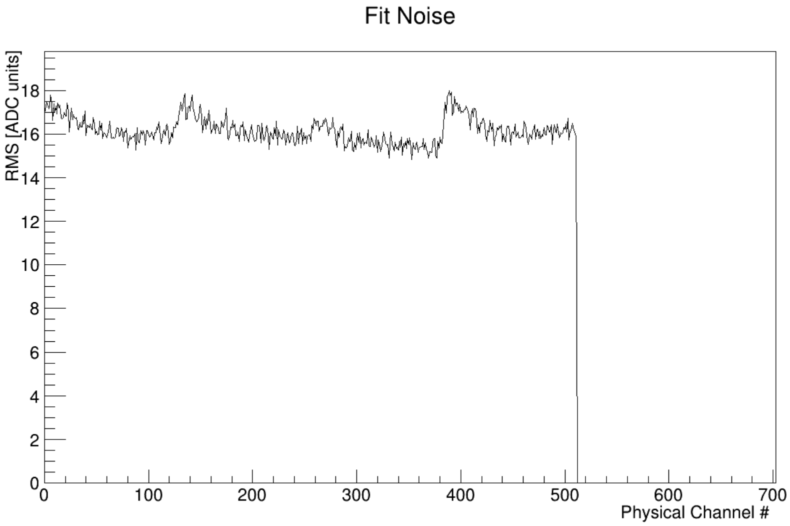
occupancy in the first layers, and the maximum signal acceptance, is expected to be closest to the edge of the sensors, so pinholes in these strips are particularly problematic. The locations of the pinholes on each sensor were mapped during the module production QA phase, after mounting the sensors to the hybrid readout electronic boards and wire-bonding the sensor readout strips to the APV25 readout chips.

3.3.3 Hybrid Quality Assurance

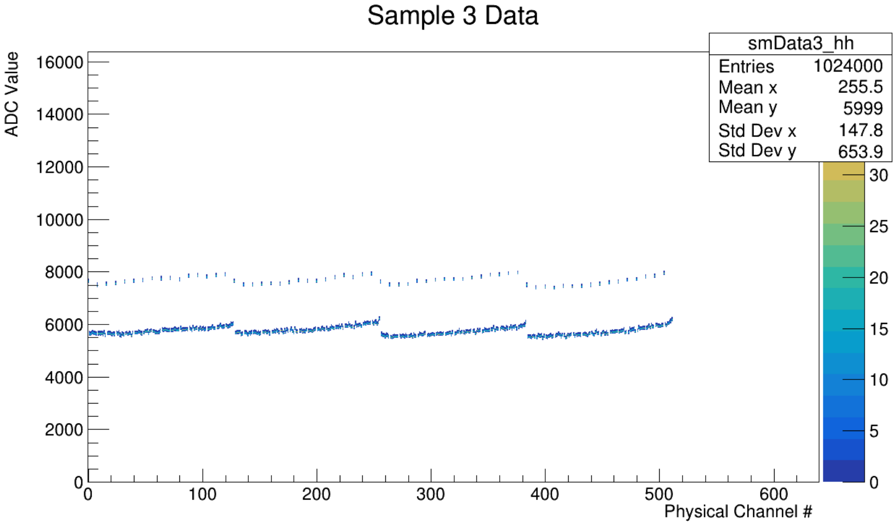
The module production process involves a series of QA tests at each stage of development. First, the bare hybrid readout electronic boards are tested without any readout chips or sensors. The hybrids in the first three layers transmit power, data, and communications through a bundle of twisted-pair cables ("pigtail") wired to a D-sub connector. Connecting a wire to the wrong pin is an easy mistake, so the connections are checked by simple voltage tests. The hybrid low voltage inputs are biased according to operational specifications, and the currents are measured to ensure that the correct amount of current is consumed. The high voltage inputs are tested up to 500 V, and the current is checked to ensure it stays less than 20 µA.

After passing the initial pigtail tests, the APV25 readout chips are mounted to the hybrids. The slim-edge hybrids utilize 4 APV25's, in contrast to the 5 APV25 chips on the standard hybrids. The baseline noise of the hybrid for all 512 APV25 channels is measured by reading out the channels over a few thousand readout windows. As a reminder, HPS reads out six consecutive ADC samples at 24 ns intervals upon receiving a trigger, and the first sample, "sample 0," generally represents the state of the channel before the trigger. The sample 0 ADC values are stored in a 2D histogram as a function of the APV25 channel number (0-512). The ADC distribution for each channel is isolated and fit with a Gaussian function, and the width of the fit represents the channel noise. The noise for one of the 2019 hybrids is shown in Figure 3.17(a). The noise distribution for each hybrid is visually inspected to ensure that there are no anomalous channels and that every APV25 channel is reading out correctly.

The response of each channel to signal is tested using the APV25's internal calibration generator. The generator injects a configurable amount of charge into the channel, which mimics the signal the channel would see if it was connected to a silicon strip that was ionized by charged particle. The charge injections are separated into eight calibration groups, where each group drives 16 distinct pre-amplifier channels, so the charge injection test must be repeated for each group in order to see the response of all 128 channels. Figure 3.17(b) shows the charge injection result for a single calibration group. The mean ADC value of every eighth channel is many standard deviations above the mean ADC value of channels that do not have charge injected. The internal calibration is described further in the context of pulse calibration in Section 4.2.2. If the baseline, noise, and charge injection scan results pass scrutiny, the hybrid is mounted with a silicon strip sensor to form a complete half-module.



(a) Layer 0 hybrid channel RMS noise.



(b) Layer 0 hybrid channel response.

Figure 3.17: (a) Layer 0 hybrid channel RMS noise. (b) Layer 0 hybrid channel response to APV25 calibrated internal charge injection for a single calibration group.

3.3.4 Half-Module Quality Assurance

The 22 best sensors were prioritized based on slim-edge width, breakdown voltage, and known manufacturing defects, and used to build 22 half-modules. However, the presence of pinholes in the sensors could only be determined after wire-bonding the sensors to the APV25 readout chips and measuring the channel noise, so the sensors needed to be re-prioritized after mounting, based on the location and number of pinholes. Channels with pinholes can be identified by their extremely low noise when the sensor is not biased. Figure 3.18 shows the channel noise for the "L0M3" half-module biased at 0 V, where five pinholes can be seen as low-noise spikes. The wire-bonds of these channels were plucked to prevent the shorted DC current from influencing neighboring channels, but this creates holes in the detector. Since essentially every sensor had multiple pinholes, they were prioritized based on the location of the pinholes. The track occupancy is highest near the edge of the sensor, so pinholes near the edge represent the largest loss in acceptance. Again looking at Figure 3.18, the edge of the sensor is represented by the first and last 50 physical channel numbers, so there is only one pinhole in the critical region of L0M03, making it overall a strong candidate to be used in the detector. A summary of the slim-edge sensors that were used in the SVT upgrade, and the two spare half-modules, is shown in Table 3.2.

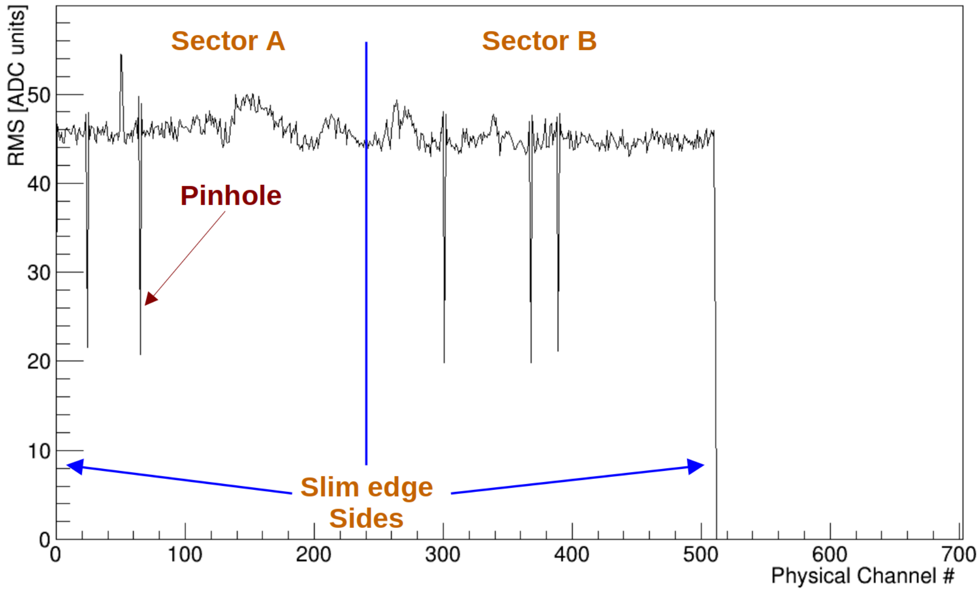


Figure 3.18: Example of layer 0 half-module RMS noise. The downward fluctuations in RMS represent pinholes where the wirebonds must be removed. The occupancy is concentrated towards the edge of the sensor, so pinholes there result in the greatest loss in acceptance, and must be minimzed.

Sensor ID	**Module Number**	**Tracker Location**	**Vfd [V]**	**Vbd [V]**	**Edge width [µm]**	**N(pinholes for first 50 strips)**	
						Strip row A	**Strip row B**
W11S2	22	L0TS	32	420	337 (358)	0	2
W10S1	19	L0TA	30	270	298 (338)	3	0
W11S1	16	L0BS	32	330	352 (398)	0	0
W11S3	17	L0BA	32	340	318	0	2
W11S6	18	L1TS	32	370	377 (425)	1	0
W09S2	7	L1TA	30	300	360 (399)	2	0
W10S3	15	L1BS	30	290	395 (442)	0	0
W10S4	3	L1BA	30	200	366	0	1
W4S3	9	Spare 1	25	370	349	1	2
W10S6	20	Spare 2	30	340	274	3	0

Table 3.2: 2019 Slim Sensor Data

3.4 Electromagnetic Calorimeter (Ecal)

The Electromagnetic Calorimeter (Ecal), shown in Figure 3.19, is an array of 442 lead tungstate (PbWO4) crystal readout modules that uses charged particle energy depositions to trigger on trident-like events with a very high background rate (up to 1 MHz per crystal) [3]. The Ecal is split into top and bottom halves above and below the through-going electron beam at a distance of 2 cm from the beam plane, which maintains the 15 mrad opening angle requirement, with the face of the Ecal located approximately 139.3 cm downstream of the target. Each Ecal half consists of 4 rows of 46 crystal modules, and an additional row of 37 modules closest to the beam, where the missing 9 crystals are removed to better accommodate the electron beam and the large Bremsstrahlung photon flux in that location.

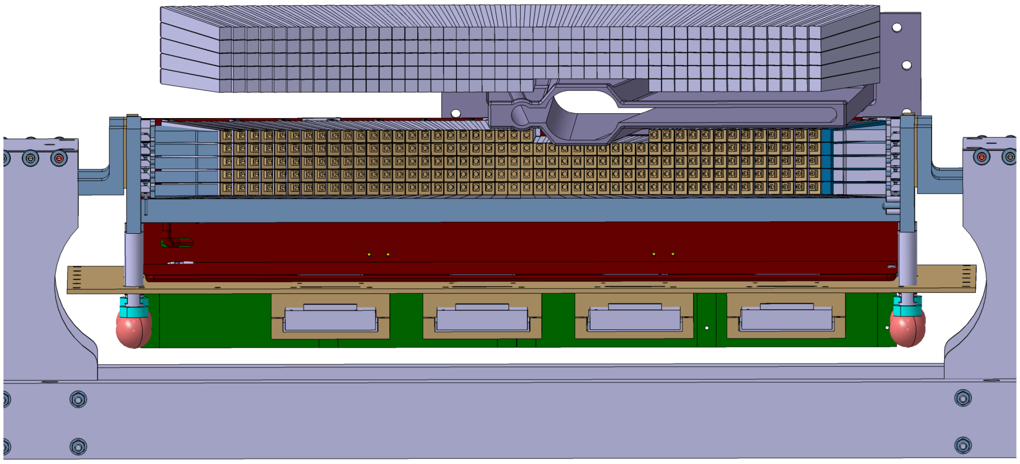


Figure 3.19: Ecal layout as seen in the beam-direction. Motherboards (green) and copper plates (red) for heat shielding are visible in the bottom half, with those elements removed from the top half for visual clarity [3].

Figure 3.20 shows a schematic of an Ecal module. The PbWO4 crystals are 160 mm long (18 radiation lengths), and tapered with a front (rear) face of 13.3×13.3 mm^2

(16×16 mm^2) The crystals are wrapped in VM2000 multilayer polymer mirror film to increase the scintillation light collection from energy depositions in the crystal. The light yield is amplified by a 16×16 mm^2 Hamamatsu avalanche photo-diode (APD) glued to the rear face, and the APD current is converted to a voltage pulse by a preamplifier. Each Ecal half is enclosed in a temperature controlled box operating at 17 °C to stabilize the crystal scintillation yield and APD gains. The preamplifier signal is routed to a flash analog-to-digital converter (FADC) board that digitizes up to 16 crystal signals at 250 MHz, and stores the samples in an 8 µs deep pipeline. When a trigger is received, a programmable range of the pipeline samples is read out, time-coincident with the trigger, and used to reconstruct Ecal clusters offline. Simultaneously, the FADC samples are also used to generate the trigger in the first place, which involves a faster online Ecal energy cluster reconstruction.

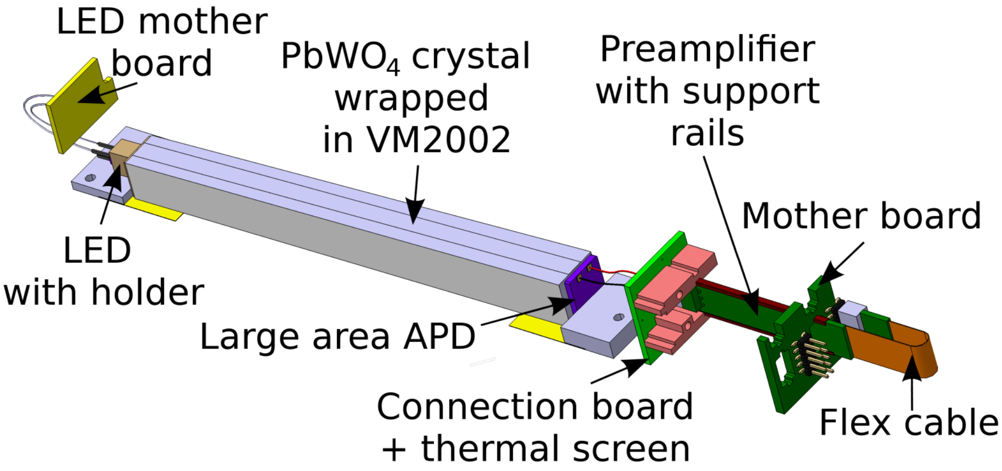


Figure 3.20: Schematic view of a single Ecal module [3].

3.5 Trigger

HPS uses a set of analysis triggers based on the energy, timing, and kinematics of online reconstructed Ecal energy clusters to reject the extremely high beam background in favor of trident-like events. As mentioned in the previous section, energy deposited in the Ecal crystals is read out by an FADC board every 4 ns. The FADC continuously checks for channels that cross some programmable ADC threshold, integrates all of the samples above threshold on a given channel (with a maximum of 128 samples), and sends the serialized summed values, as well as the threshold-crossing time and crystal position, to the Crate Trigger Processor (CTP) in 32 ns time frames. The summed ADC is converted into an energy using offline calibrated gains and forms a hit with energy, position, and time information. The hits are used by the General Trigger Processor (GTP) clustering algorithm to identify energy depositions consistent with the defined triggers [54].

The energy deposited by a particle in the Ecal is typically spread out across multiple crystals, so multiple hits are often combined into a single cluster. The clustering algorithm first identifies a seed-hit, which is a hit that has more energy than all of its neighboring hits in a 3×3 crystal window, across some programmable window of time. The algorithm then checks whether the seed-hit is above some minimum energy threshold, as the trident events of interest are expected to deposit significant energy. If the hit passes this cut, an Ecal energy cluster is formed as the energy sum of all hits in the 3×3 region centered on the seed-hit, across the full programmed time window. Seed-hits that occur on edge crystals of course use a smaller clustering window, as there can be no neighbors where there are no crystals. The clusters are sent to the Single Subsystem Processing (SSP) board to make trigger decisions. When the SSP identifies a positive trigger, the trigger is sent to the Trigger Supervisor board (TS), which checks whether DAQ can accept the trigger, and then sends the trigger to the Trigger Interface (TI) boards that initiate readout.

HPS used a set of five individual triggers for the 2016 run. The "pulser" trigger fires at a fixed 100 Hz. The "singles0" and "singles1" triggers fire on single cluster events that fall within some upper and lower energy thresholds, with a minimum number of hits on the cluster. The "pairs0" and "pairs1" triggers are the primary analysis triggers for the 2016 analysis, and trigger on pairs of clusters that are consistent with highly-boosted electron and positron decay products that have some opening angle between them, and bend oppositely in the uniform magnetic field permeating the SVT. The basic pairs trigger algorithm is illustrated in Figure 3.21. The pairs triggers first place cuts on the individual cluster energies and hit count, as in the singles triggers. There is a loose cut on the minimum energy sum of the two clusters, such that they are consistent with the expected highly-boosted decay products, but that allows for significant energy losses in the Ecal. The maximum energy sum must be less than the beam energy to reduce the trigger rate due to coincident scattered beam electrons. Additionally, a cut is placed on the energy difference between the two clusters. A geometric cut is placed on the "coplanarity" of the cluster pair, which requires that each cluster impinge the Ecal in opposite vertical and horizontal halves, as expected for the two-body decay of oppositely charged particles with any significant curvature. Lastly, there is a loose cut on the time-coincidence of the two clusters [55]. The pair1 trigger used in the 2016 analysis is summarized in Table 3.3.

Cut description	**Cut value**
Hits Per Cluster Min:	2
Cluster Time Coincidence:	±12 ns
Cluster Energy Min:	150 MeV
Cluster Energy Max:	1400 MeV
2-Cluster Energy-Sum Min:	600 MeV
2-Cluster Energy-Sum Max:	2000 MeV
2-Cluster Energy-Diff Max:	1100 MeV
Coplanarity Maximum:	40 deg
Energy-Dist Slope:	5.5 MeV/mm
Energy-Dist Minimum:	700 MeV

Table 3.3: Parameters of pair1 trigger. used in the 2016 analysis.

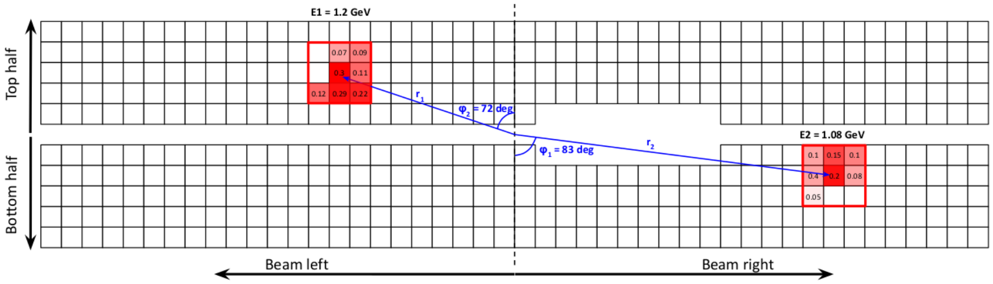


Figure 3.21: Example event triggered by the pairs trigger in HPS. The Ecal energy clusters are formed in 3x3 grids around a seed crystal with the largest energy. The pairs trigger requires two time-coincident clusters, with a minimum individual and combined energy threshold, to be in opposite quadrants of the Ecal [3].

3.6 Fixed Target Backgrounds

The primary backgrounds for the HPS experiment are prompt QED Bethe-Heitler and radiative tridents ($e^-Z \rightarrow e^-e^+e^-Z$), and Wide-Angle Bremsstrahlung (WABs) conversions ($e^-Z \rightarrow e^-\gamma Z$ followed by $Z\gamma \rightarrow e^-e^+$). The WABs photons can pair-produce inside the target, or convert downstream in the detector material. The Feynman diagrams for these processes are shown in Figure 3.22.

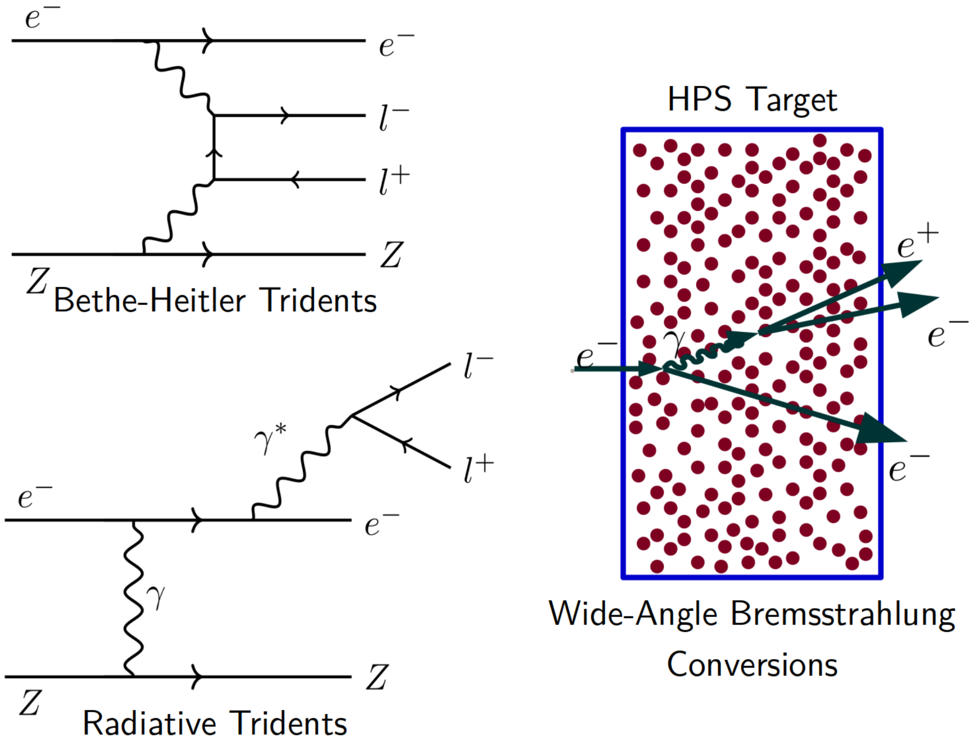


Figure 3.22: Feynman diagrams for Bethe-Heitler tridents (top left), radiative tridents (bottom left), and Wide-Angle Bremsstrahlung (WABs) conversions (right).

Another important background contribution is from elastically and quasi-elastically scattered beam electrons that can contaminate tracking. A background vertex can be incorrectly reconstructed using a scattered beam e^-, or the true tracks can be contaminated by beam e^- hits, and these reconstruction errors can lead to falsely displaced events. Consequently, the QED background rejection is also developed in consideration with the electron beam background, and all MC background samples are mixed with MC generated beam to simulate the data, as described in Section 4.4.

The radiative trident process is related to the A' production rate in Equation (5.6), and is used to scale the expected signal rate relative to the data. However, the Bethe-Heitler and radiative trident processes have identical final states, so the squared sum of the amplitudes are observed in data, and the processes are indistinguishable. Additionally, either trident process can be mistakenly reconstructed using the recoil electron rather than the daughter electron, which acts another background source. The combination of reconstructed radiative and Bethe-Heitlers tridents are collectively referred to as "tridents" in the data, and "tritrig" in MC (meaning triggered tridents). A MC sample of the truth-matched radiative tridents alone is used to estimate the "pure" radiative trident rate in the data, as discussed in Section 5.1.4.2 and Section 5.1.4.3.

WABs are called "wide-angle" because the photon and electron at the target are emitted at large angles from the beam axis, pushing them into the SVT acceptance. In addition to the prompt converted WABs in the target, WABs may convert anywhere in the detector material, such as in the first layer of silicon. In this case, perhaps only the e^+ is within the detector acceptance, however, it can be vertexed with the recoil e^- or a beam e^- and create a problematic displaced background. These kinds of events can be eliminated by enforcing certain hit on track requirements, as discussed in Section 5.2.1. Figure 3.23 depicts the relative contribution of each background component as a function of invariant mass using MC simulated events with MC beam included. The events shown do not reflect the final selection, but do have the "Preselection" cuts applied, which are detailed in Section 5.1.2.

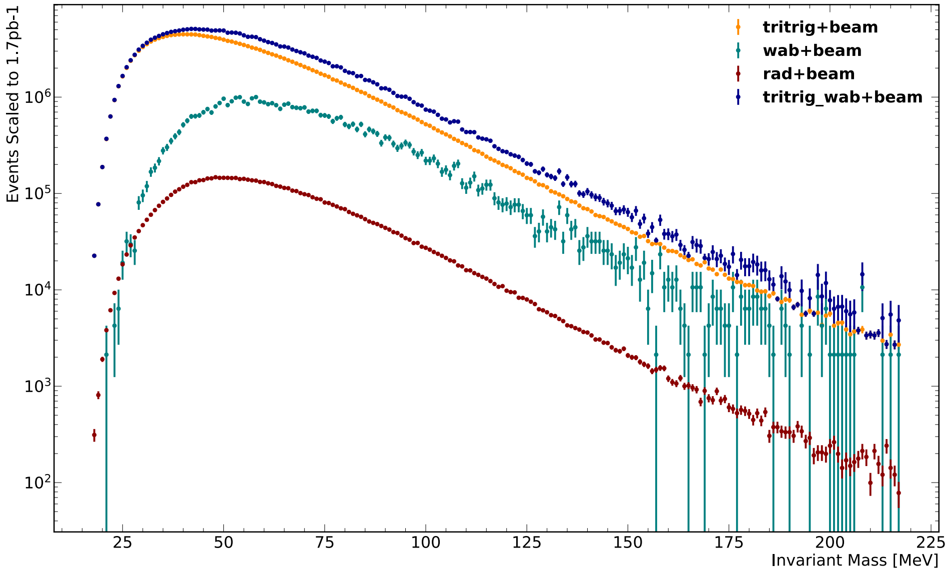


Figure 3.23: Histogram of MC-reconstructed vertex invariant mass for primary QED background processes. The histogram includes several components: "rad" represents the pure radiative trident component, while "tritrig" combines Bethe-Heitler and radiative tridents, referred to as tridents in the text. The "wab" component corresponds to the converted WABs contribution, and "tritrig+wab+beam" denotes the complete QED background including all components. Each component is scaled to $\sim 10\,\%$ of the 2016 luminosity($10.7\,\mathrm{pb}^{-1}$), and all events have had "Preselection" criteria applied (Section 5.1.2).

Figure 3.24 shows the summed momenta of the e^- and e^+ tracks used to reconstruct the vertex, called "Psum." This kinematic variable is used later in Section 5.1.3 to define the SIMPs signal region, and eliminates a large fraction of the tridents and WABs. Finally, Figure 3.25 presents the reconstructed vertex z position and highlights the challenges posed by these QED background processes in a displaced vertex search. The target is situated at $-4.3\,\mathrm{mm}$, yet prompt background events are being reconstructed significantly

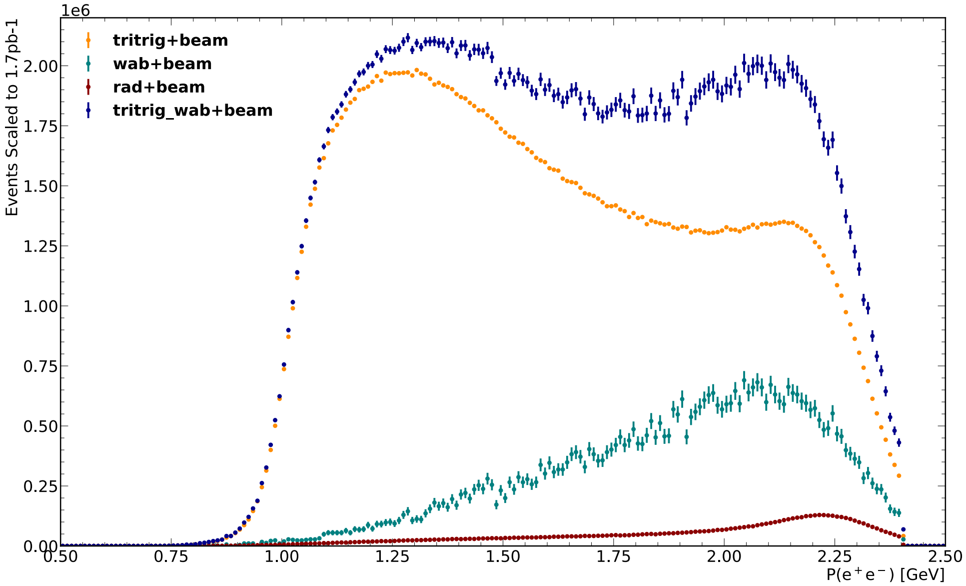


Figure 3.24: Histogram of MC-reconstructed vertex Psum ($|p_{e^-}| + |p_{e^+}|$) for primary QED background processes. The histogram includes several components: "rad" represents the pure radiative trident component, while "tritrig" combines Bethe-Heitler and radiative tridents, referred to as tridents in the text. The "wab" component corresponds to the converted WABs contribution, and "tritrig+wab+beam" denotes the complete QED background including all components. Each component is scaled to $\sim 10\,\%$ of the 2016 luminosity ($10.7\,\text{pb}^{-1}$), and all events have had "Preselection" criteria applied (Section 5.1.2).

beyond 124 mm downstream of the target. Given that the maximum SIMPs signal production rate is several orders of magnitude lower than the QED trident rate and is inversely related to the signal's lifetime, it is crucial to constrain the background effectively downstream of the target in order to identify a small displaced signal rate. A significant amount of effort in this thesis is dedicated to efficiently rejecting these falsely displaced background events, which are detailed in Section 5.1.

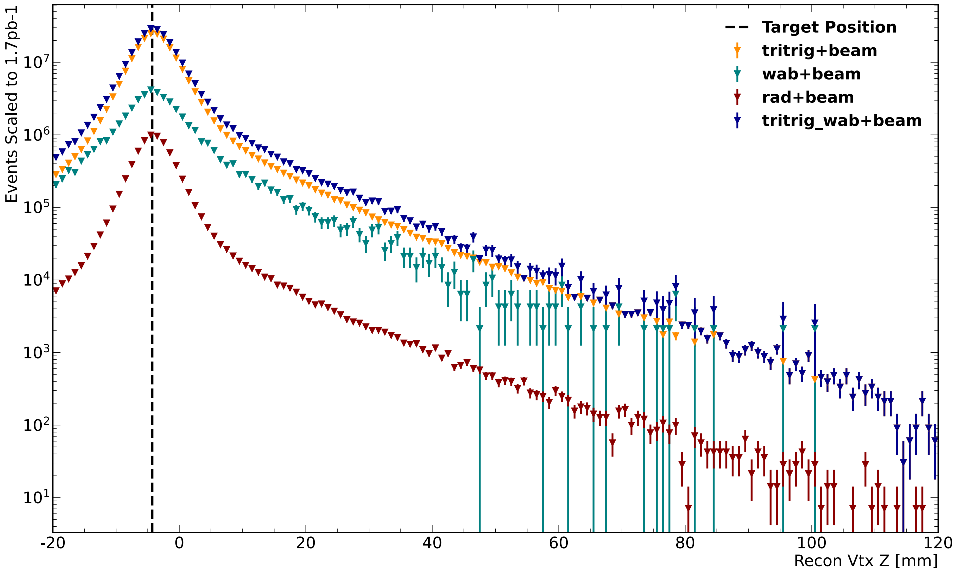


Figure 3.25: Histogram of MC-reconstructed vertex z position for primary QED background processes. The histogram includes several components: "rad" represents the pure radiative trident component, while "tritrig" combines Bethe-Heitler and radiative tridents, referred to as tridents in the text. The "wab" component corresponds to the converted WABs contribution, and "tritrig+wab+beam" denotes the complete QED background including all components. Each component is scaled to $\sim 10\,\%$ of the 2016 luminosity ($10.7\,\mathrm{pb}^{-1}$), and all events have had "Preselection" criteria applied (Section 5.1.2).

Chapter 4

Event Reconstruction

HPS is searching for displaced secondary vertices by triggering on the energy depositions of highly-boosted e^- e^+ pairs that are coincident in time and have some small opening angle between them. The charged particle Ecal energy clusters, SVT hits, trajectories through the SVT, and vertex candidates, are reconstructed offline using hps-java [56], which extends the lcsim software toolkit. The following sections describe each of the main stages of offline reconstruction.

4.1 Ecal Clusters

Charged particles that strike the Ecal deposit energy within its crystals, often across multiple crystals. The individual crystal hits are then characterized and clustered together to form Ecal clusters. Each crystal is sampled at 250 MHz, and 100 samples are read out upon triggering. The crystal pulse shape samples are fitted with the sum of a pedestal P and a 3-pole function to get the crystal energy in terms of ADC counts, given by:

$$\mathrm{ADC}(t) = P + \frac{A}{2\tau^2}(t - t_0)^2 e^{-(t-t_0)/\tau}, \tag{4.1}$$

where t_0 is the hit time, and τ is the pulse width. Time variations between each crystal are corrected using offline calibrations derived from the CEBAF accelerator's 499 MHz radio-frequency signal. Additionally, corrections are applied for the energy dependence of the time offsets, referred to as "time walk" [3].

4.1.1 Cluster MC Truth Matching

Ecal clusters in MC simulations can be "truth-matched" to the MC particle that generated the energy cluster, essential for various calibration and analysis studies. Ecal clusters are composed of multiple crystal hits, referred to as "readout hit" in the software. The particle associated with the seed hit of the cluster is presumed to be the primary generator of the cluster, so the seed readout hit is used to match the particle to the cluster. Each readout hit is linked to multiple "simulated calorimeter hits," where each simulated hit carries the truth information that relates it to the MC particles that generated it. The simulated hit that contributes the largest fraction of the readout hit's energy is selected, and if this hit is associated with multiple MC particles, it is matched to the MC particle with the highest energy contribution.

To ensure a reduced rate of false matches, several cuts are applied to the truth-matching process. Since the primary use of the cluster truth matching is to evaluate the performance of the track-cluster matching algorithm using prompt backgrounds, the MC particle matched to the cluster must originate at the target. Additionally, the sign of the MC particle's vertical momentum must correspond to whether the cluster is found in the top or bottom of the Ecal. Finally, the clusters are checked to see if they've been matched to duplicate MC particles, and if so, the MC particle position at the Ecal is used to select the best match. All remaining duplicates are removed.

4.2 SVT Track Reconstruction

Particle tracks are reconstructed offline using hits from the SVT silicon strip sensors. These reconstructed tracks are extrapolated to the face of the Ecal and matched with energy clusters to form final state e^- and e^+ objects, which are then used to identify vertex candidates. The following section describes the hit reconstruction, track finding and fitting, track-cluster matching, and the vertex fitting algorithms.

4.2.1 SVT Raw Hit Reconstruction

The APV25 readout chip samples the shaper output of each channel every 24 ns and stores the ADC values in a 192-cell deep analog pipeline. Upon receiving a trigger, six consecutive pipeline samples are read out, with the readout window centered on the trigger time. Channels with at least three samples above an offline calibrated ADC threshold are considered. These six consecutive pipeline measurements collectively form a "raw" hit. Raw hits are reconstructed into "strip hits" by fitting the samples with the APV25 channel pulse response shape, which is modeled by the four-pole transfer function:

$$F_{\mathrm{4pole}}(t) = \frac{\tau_1^2}{(\tau_1 - \tau_2)^3} \left(e^{-\frac{t}{\tau_1}} - e^{-\frac{t}{\tau_2}} \sum_{k=0}^{2} \frac{\left(\frac{\tau_1-\tau_2}{\tau_1\tau_2} t\right)^2}{k!} \right), \tag{4.2}$$

where τ_1 and τ_2 are shaping constants calibrated offline for each channel. The fitted strip hit returns the hit amplitude in ADC units and the time of the hit relative to the clock.

These times are corrected for run-dependent phase shifts, trigger time, and time of flight. Figure 4.1 illustrates the fitted strip hit pulse shape for a single channel.

Pile-up events occur when an out-of-time hit contributes to the pulse of an in-time hit. A pile-up algorithm checks whether the χ^2 probability of a single pulse fit is less than 0.5. If true, the pulse is fit again using a double pulse fit. If the χ^2 probability improves with the double pulse fit, the pile-up fit is kept. All strip hits must have a pulse fit χ^2 probability greater than 3.20×10^{-6}.

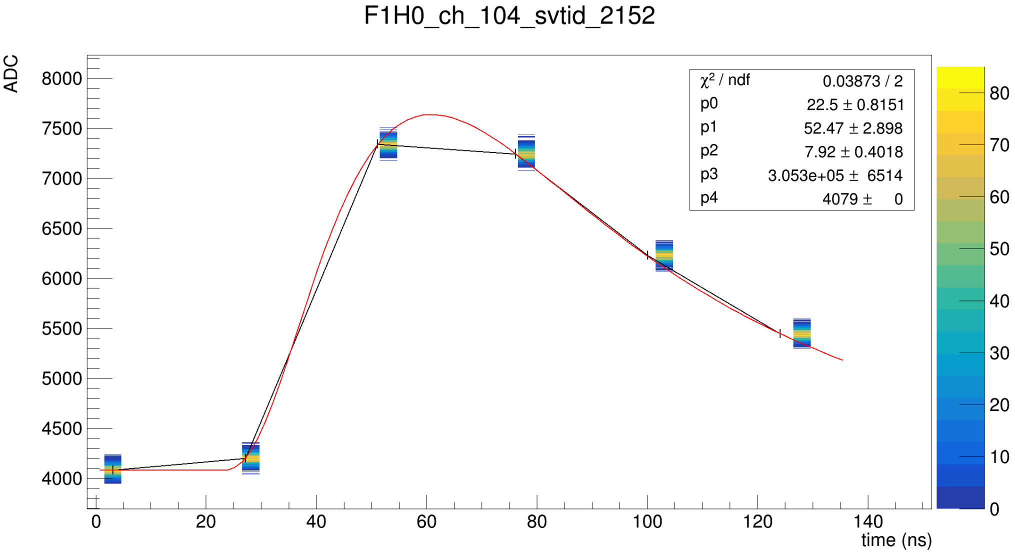


Figure 4.1: Example of an APV25 channel pulse. The ADC values are sampled every 24 ns and fit to the offline calibrated four-pole transfer function in Equation (4.2). This figure actually shows a 2D histogram of many pulses aligned perfectly in time.

4.2.2 CR-RC Calibrations

The SVT hit time resolution is critical for the analysis, as it affects the clustering of strip hits, associating hits along tracks, and matching tracks to clusters for reconstructed particles. Accurate reconstruction of hit times relies on high-quality fitting of the APV25 channel hit response pulse shape. Figure 4.2 illustrates the APV25 channel hit response curves for various signal magnitudes. The response shape is characterized by the CR-RC circuit time constants τ_1 and τ_2, which are used to fit the SVT hits according to Equation (4.2). These timing constants are calibrated offline using a dedicated "calibration run" after the detector is installed at JLab. The calibration run is performed without the electron beam, using the APV25's internal charge calibration generator. The generator injects a fixed, known quantity of charge into each channel, producing a pulse shape response that can be fitted using the equation above.

Charge is not injected into every channel simultaneously. Instead, it is sequentially injected into eight different calibration groups (0-7), with each comprising 16 channels. The calibration group number assigned to a channel is determined by:

$$\mathrm{Cal}_{\mathrm{Group}} = \mathrm{Channel}_{\mathrm{N}} \mod 8. \tag{4.3}$$

Figure 4.3 shows the APV25 internal calibration timing diagram. A calibration request signal ("11") is sent to the APV25 trigger line, which routes the request to the delay line ($\mathrm{req}_{\mathrm{in}}$), where the request output ($\mathrm{req}_{\mathrm{back}}$) is delayed by a configurable time t_{d}. After t_{d}, the calibration switch injects charge into every channel within the configured calibration group, as set using the "CRDV" control register on the APV25. The delay time t_{d} is adjusted by setting the time of charge injection, which is restricted to multiples of $\frac{1}{8}$

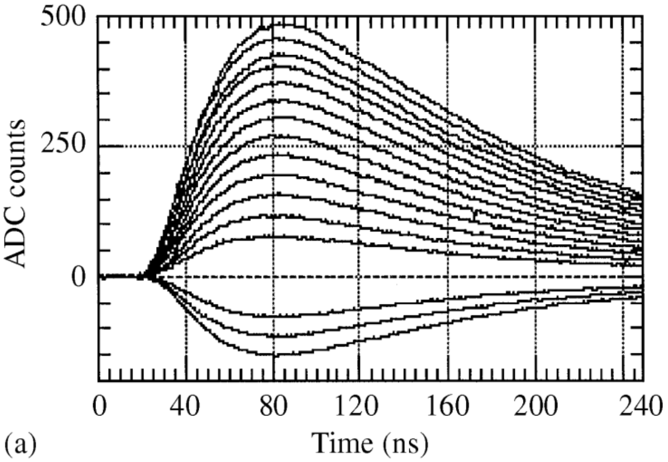


Figure 4.2: Response curve of the APV25 as a function of signal magnitude [4].

of the external clock cycle (24 ns at JLab) The delay time, t_d, is configured by setting the APV25 “CSEL” control register to an integer between 0 and 8. This delay is restricted to multiples of 1/8 of the external clock cycle (24 ns at JLab). For example, setting CSEL=2 results in a 6 ns delay between the request and the charge injection, while setting CSEL=8 results in a full 24 ns delay.

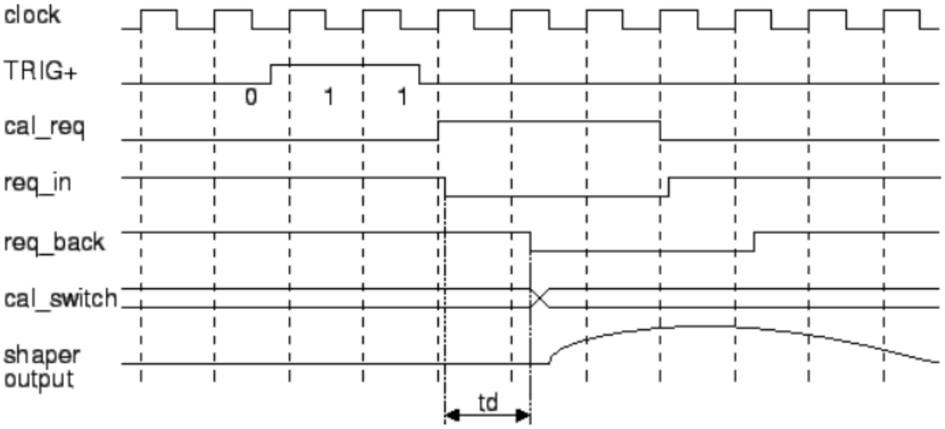


Figure 4.3: APV25 internal charge calibration generator timing diagram.

Recall from the previous section that HPS operates the APV25 in “multi-peak” mode, reading out only six ADC samples at 24 ns intervals. This sampling rate alone is insufficient to accurately calibrate the timing constants τ_1 and τ_2. However, the pulse shape sampling rate used in the calibration run can be improved from 24 ns to 3 ns by utilizing the charge injection time delay feature of the internal calibration generator. As illustrated in Figure 4.4, this process involves adjusting the charge injection time delay t_d, which shifts the sampling point along the fixed-shape and fixed-amplitude calibration pulse. The figure shows a series of pulses with different time delay settings, where the colored circles represent the shaper output samples aligned with the 24 ns clock cycle. By delaying the charge injection by t_d, the APV samples the shaper output at new time points along the pulse. Each CSEL setting therefore captures a different six-sample segment of the shaper output. All eight of the CSEL samples are combined to construct a pulse with 48 ADC samples at 3 ns intervals. These higher resolution calibration pulses are then fitted with Equation (4.2) to calibrate τ_1 and τ_2 for every SVT APV25 channel.

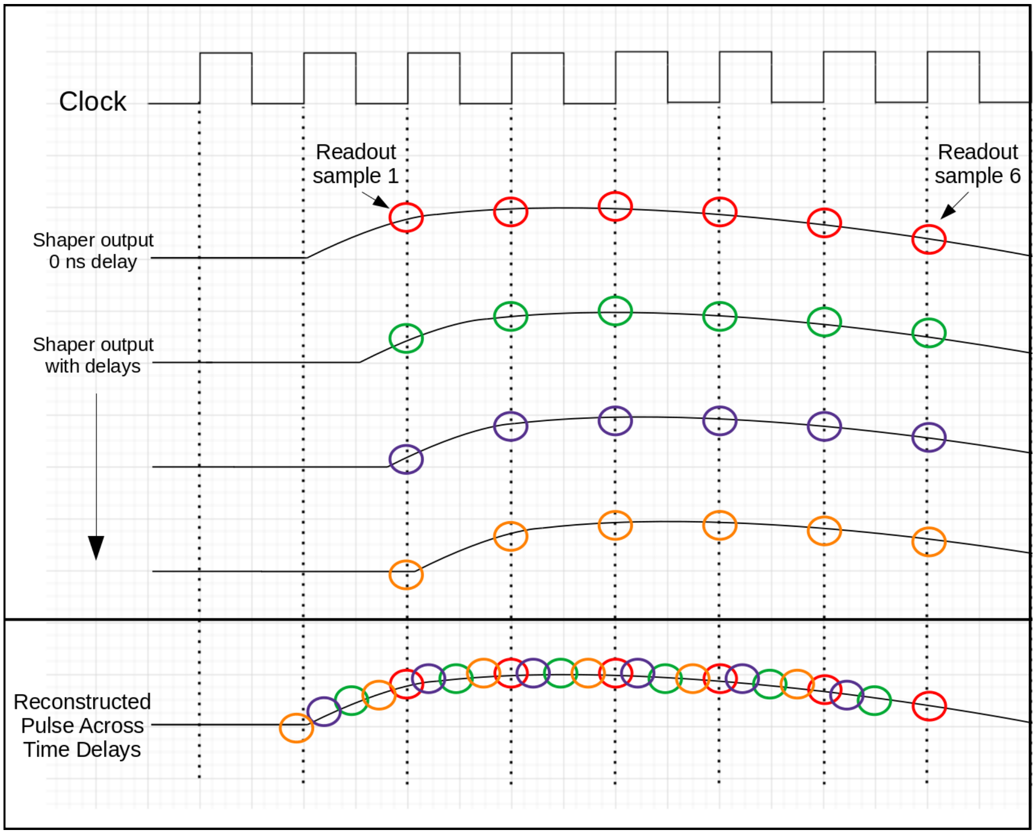


Figure 4.4: Illustration showing how the time delay setting in the APV25 internal charge calibration generator can be used to reconstruct calibration pulses with improved time resolution. The calibration pulses are used to calibrate τ_1 and τ_2 in Equation (4.2).

4.2.3 Hit Clustering

The charge deposited by a hit may span multiple readout strips, so strip hits are clustered together using the nearest neighbor RMS Clusterer algorithm [57]. A strip hit is considered to be a cluster seed candidate if its ADC amplitude is 4σ above the RMS noise of the channel, as determined from offline calibrations. Neighboring hits are included in the cluster if their amplitudes are 3σ above threshold, and if they occur within 24 ns of the seed. The strip cluster position is taken as the amplitude-weighted centroid of the strip hits on the individual sensor, $\sum x_i A_i / \sum A_i$, which improves the hit position resolution compared to using a single strip. The timing of the cluster is also improved by combining strips, with the cluster time given by:

$$t = \sum t_i A_i^2 / \sum A_i^2. \tag{4.4}$$

4.2.4 SVT Track Parameters

Particle tracks through the uniform magnetic field are described by a helix with five parameters: Ω, $d0$, $z0$, $\tan\lambda$, and ϕ These five track parameters are defined relative to the tracking coordinate system, where the x-axis points along the beam direction, the y-axis points beam-left, and the z-axis points vertically, roughly corresponding to the uniform B-field direction. The parameters are defined with respect to an arbitrary reference point, which is defaulted to the tracking coordinate system origin. However, the reference point can be configured in reconstruction to measure the parameters with respect to other locations, such as the beamspot position at the target. The track parameters, along with the covariance matrix, form a track state. Since the track trajectory changes at each detector layer due to multiple scattering, a reconstructed track is actually a collection of track states defined at each layer in the detector. Below is a description of each track parameter, illustrated in Figure 4.5.

- d0: The point of closest approach between the reference point and the projection of the helix onto the xy-plane.
- ϕ: The angle between the direction of the helix projected onto the xy-plane and the x-axis, evaluated at the point of closest approach. It represents how far to the left or right a track is pointing in the HPS detector frame.
- Ω: The inverse of the signed helix curvature at the point of closest approach, used to reconstruct the track momentum.
- λ: The angle between the xy-plane and the track direction in the sz-plane, where s is the direction of the track path.
- The tangent of this angle, $\tan\lambda$, represents the vertical slope of the track in the HPS detector frame.
- z0: The distance along the z-axis between the helix and the xy-plane at the point of closest approach. This parameter represents the vertical distance of the track in the HPS detector frame, known as the "vertical impact parameter." This is a critical parameter used to eliminate background in the analysis.

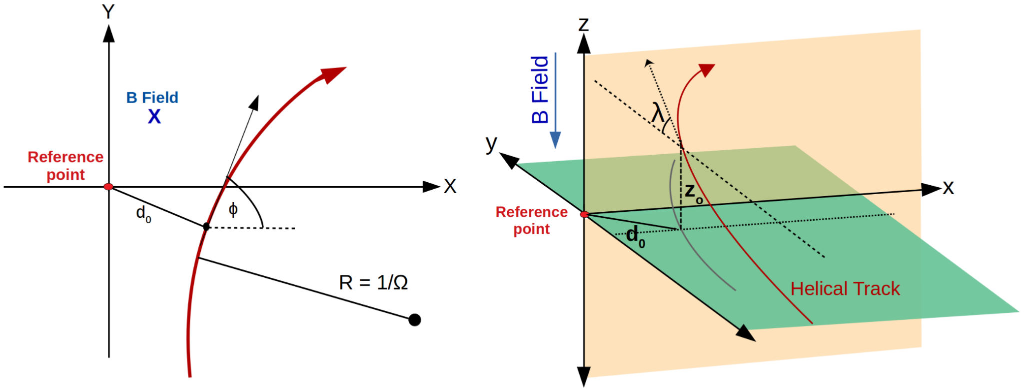


Figure 4.5: Illustration of the helix track parameters in the HPS tracking coordinate system. The track parameters are defined in the perigee parameterization.

4.2.5 Track Finding and Fitting

HPS originally utilized the SeedTracker [58] algorithm to form track candidates and then refined these candidates using the General Broken Lines (GBL) algorithm to account for changes in the trajectory due to detector interactions [59]. However, as discussed later in this section, the track finding and fitting was upgraded to use a Kalman Filter (KF) [60] algorithm, which improved the tracking performance by better accounting for the charged particle detector interactions during track finding. The benefits of this upgrade will be detailed later in the section.

4.2.5.1 SeedTracker and General Broken Lines Fitting

The implementation of SeedTracker and GBL in HPS requires the use of "3D" SVT strip clusters, which are the combination of axial and stereo sensor strip clusters. The matching criteria is not described here. SeedTracker constructs a seed helical track from a combination of consecutive 3D hits and extrapolates the helix to an additional tracking layer to look search for consistent hits. If a consistent hit is found, the helix is extended to the remaining layers. However, SeedTracker's helix extrapolation fails to consider the effects of energy loss in the detector material due to ionization and radiation, or the random changes in direction caused by multiple scattering at each layer. To address these issues, the SeedTracker track is refitted using the fast global track refit GBL algorithm. GBL approximates the particle trajectory as a series of track segments, each defined by a 5 parameter helix at the sensor planes. This approach accounts for multiple scattering deflections in the material by minimizing the hit residuals and scattering angles.

The implementation of this tracking strategy in HPS requires 3D hits, which results in reduced acceptance in regions where the axial and stereo sensors do not overlap, and also when there is a hit inefficiency in either sensor. While this efficiency loss was minimal for the 2016 run [57], it posed a significant issue for future runs with the upgraded SVT detector where the first layers are placed closer to the target and experience radiation damage leading to decreased hit efficiency over time. Since the vertex resolution is dominated by whether the tracks have a hit in the first layer, maximizing the hit efficiency in these layers is critical. As a result, HPS updated the tracking approach by using single strip clusters instead of 3D hits. Ultimately, the decision was made to transition to the KF based tracking, rather than modifying the existing tracking. In addition to recovering hits lost from requiring 3D hits, KF tracking incorporates multiple scattering and the non-uniform magnetic field into both track finding and fitting, which is expected to reduce the mis-identification of hits on track compared to the SeedTracker+GBL tracking. While this tracking upgrade was not crucial for the 2016 dataset, it was implemented as part of the transition to the new tracking system, and is used in this analysis.

4.2.5.2 Kalman Filter Tracking

The KF pattern recognition algorithm begins by seeding track candidates using a linear fit (line + parabola) of at least five silicon strip clusters (at exactly two axial and three stereo). The track parameters form the initial state vector $\vec{a}_{k-1}$, which is then propagated through the detector layers while accumulating hits. The KF process consists of three main stages:prediction, filtering, and smoothing [60].

The predicted state vector at the next silicon sensor is given by:

$$\vec{a}_k = \mathbf{F}_{k-1}(\vec{a}_{k-1}) + \vec{w}_{k-1}, \tag{4.5}$$

where $\mathbf{F}_{k-1}$ is the analytic track propagator, and $\vec{w}_{k-1}$ represents random multiple scattering effects, which primarily affects the track angles ϕ and λ. The covariance matrix $\mathbf{C}_k$ is updated according to:

$$\mathbf{C}_k = \mathbf{F}_{k-1}\mathbf{C}_{k-1}\mathbf{F}_{k-1}^{\mathrm{T}} + \mathbf{Q}_{k-1}, \tag{4.6}$$

where $\mathbf{Q}$ is the multiple scattering covariance matrix.

The predicted measurement at the silicon sensor is a linear function of the predicted state vector:

$$\vec{m}_k^{'} = \mathbf{H}_k\vec{a}_k, \tag{4.7}$$

where the covariance of the predicted measurement is:

$$\sigma_r^2 = \sigma^2 + \mathbf{H}_k\mathbf{C}_k^{-1}\mathbf{H}_k^{\mathrm{T}}. \tag{4.8}$$

The residual $r = \vec{m} - \vec{m}_k^{'}$ between the measured hit "$\vec{m}$" and the predicted hit is used by the pattern recognition algorithm to pick up hits based on the minimum residual. The residual covariance includes the measurement uncertainty σ. The filtering step updates the state vector based on the measurement using the Kalman gain matrix $\mathbf{K}_k$, which is defined by:

$$\mathbf{K}_k = \frac{\mathbf{C}_k \mathbf{H}_k^{\mathrm{T}}}{\sigma + \mathbf{H}_k \mathbf{C}_k^{-1} \mathbf{H}_k^{\mathrm{T}}}. \tag{4.9}$$

The filtered state vector is:

$$\vec{a}_f = \vec{a}_k + \mathbf{K}_k(\vec{m} - \mathbf{H}_k \vec{a}_k), \tag{4.10}$$

and the filtered covariance matrix is:

$$\mathbf{C}_f = (\mathbf{I} - \mathbf{k}_k \mathbf{H}_k)\mathbf{C}_k. \tag{4.11}$$

The filtered state vector and covariance are then used to predict the next measurement, continuing through the layers.

After completing the prediction and filter steps, the smoothing process runs in reverse, layer by layer, from the end of the track back to the beginning. The smoothed helix parameters at each layer provide the best estimate of the particle's trajectory. The full implementation of this tracking algorithm involves additional complexities, including coordinate transformations and the detailed modeling of track propagation, measurement functions, detector uncertainties, and multiple scattering. These comprehensive details can be found in Ref. [61].

4.2.5.3 MC Track Truth Matching and Tracking Performance Comparison

To compare the performance of the KF and SeedTracker+GBL tracking algorithms, a MC track truth matching algorithm was developed that accomodates both approaches. All MC particles in an event are checked to see if they are "trackable", meaning they left some minimum number of simulated hits in the SVT detector layers. The maximum possible number of hits is 12 in 2016, and 14 in 2019 and later years.

The tracks are decomposed into their constituent tracker hits. For SeedTracker+GBL tracks, this involves the 3D strip clusters (combinations of axial and stereo strip clusters), whereas for KF tracking, it involves the 2D single-sensor strip clusters. Each 2D/3D strip cluster is then further broken into raw hits, which are the 6 ADC samples from each strip in the cluster. The raw hits are sorted by SVT layer and sensor, and the raw hits for each layer+sensor combination are matched using the simulated tracker hits to the MC particle that generated the hit. Since each raw hit may be composed of multiple simulated tracker hits, it can be matched to multiple MC particles. The list of MC particles matched matched to the raw hits on a given sensor is filtered so that each MC particle is only counted once per sensor. Each track is then matched to the MC particle that left the most hits, allowing for duplicate matches if necessary. The KF tracking upgrade was originally intended to be used only for the 2019+ data sets, so the truth studies comparing Kalman Filter (KF) and General Broken Lines (GBL) tracking were done using 2019 data.

4.2.5.4 Track Probability Comparison

One metric for comparing the two tracking algorithms is the “track probability”, defined as:

$$\text{Track Probability} = \frac{N_{\text{truth hits}}}{N_{\text{hits on track}}}, \tag{4.12}$$

where $N_{\text{truth hits}}$ is the number of hits on track attributed to the matched MC particle; the number of hits due to the matched MC particle cannot exceed the number of hits on track, so the track probability cannot exceed one. A track with a probability of 1 represents a “perfect” reconstruction. For example, if a MC particle leaves 14 single-sensor hits in the SVT, but only 11 of these hits are used in the track reconstruction, the track probability would be approximately 0.78. Figure 4.6 compares the track probability for the two tracking algorithms, showing that the KF tracking results in roughly 25 % more perfect tracks and substantially fewer “low quality” tracks with track probability ≤ 0.6.

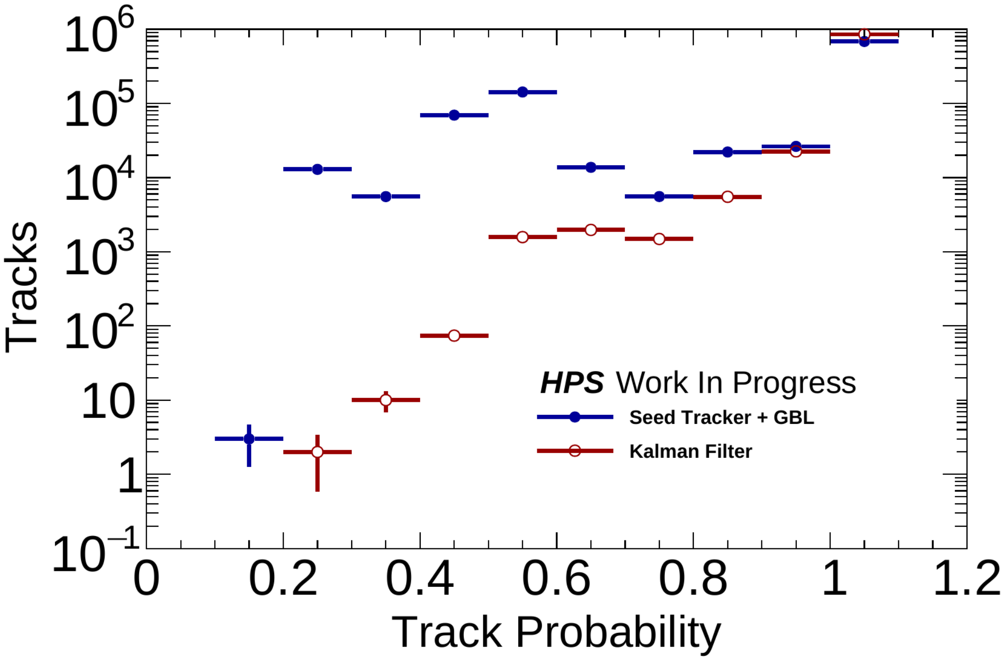


Figure 4.6: Track probability comparison for 2019 data. Track probability is defined as the ratio of the number of reconstructed tracks matched to a MC particle to the number of MC particles that could have been reconstructed. Results are shown for high-quality tracks with track probability ≥ 0.8.

4.2.5.5 Tracking Efficiency

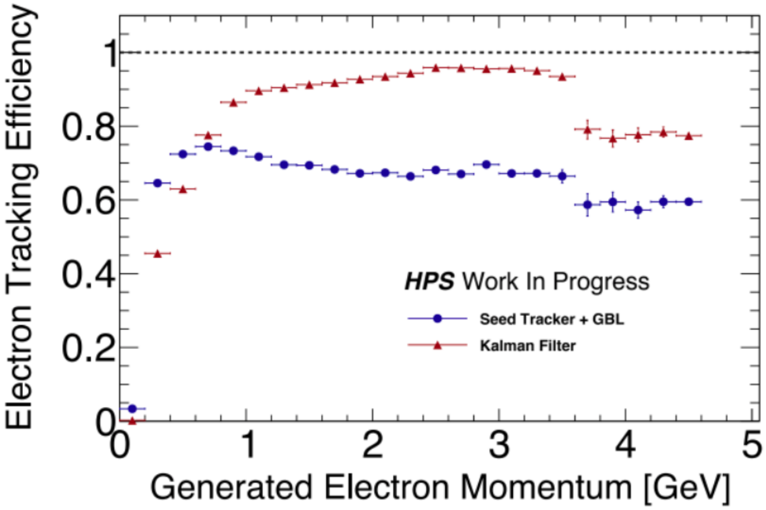
The “tracking efficiency” as a function of the generated momentum p_{truth} is defined as:

$$\epsilon(p_{truth}) = \frac{N_{\text{matched}}(p_{\text{truth}})}{N_{\text{trackable MCP}}(p_{\text{truth}})}, \tag{4.13}$$

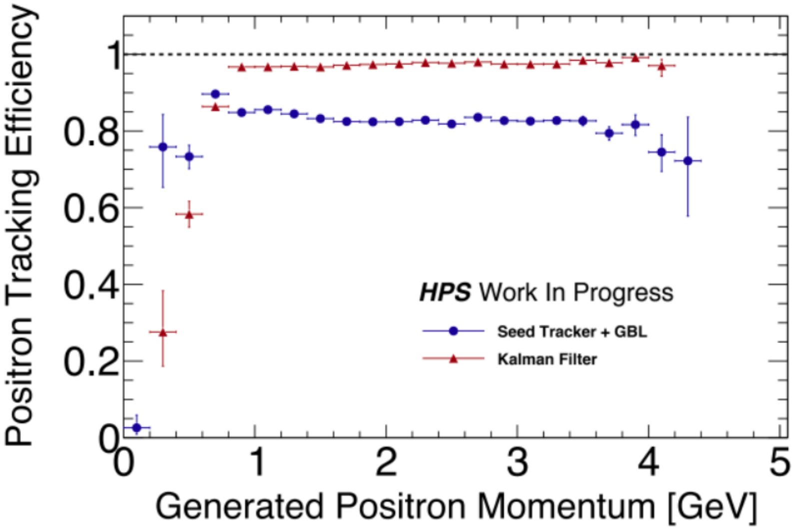
where N_{matched} is the number of reconstructed tracks in the momentum bin matched to a MC particle, and $N_{\text{trackable MCP}}$ is the number of MC particles that could potentially be reconstructed in the momentum bin. Figure 4.7 illustrates that KF tracking shows a significant improvement in "high quality" tracking efficiency.

4.2.5.6 Fake Reconstructed Tracks

The KF tracking algorithm also demonstrates a negligible rate of so-called "fake tracks," which are defined as tracks with a track probability less than 0.8. Figure 4.8 shows the fake track rate as a function of generated momentum p_{truth}. The KF tracking shows a negligible fake rate, in contrast to SeedTracker+GBL, which demonstrates a fake rate on the order of $30\,\%$. This difference highlights the improved reliability of the KF tracking in reducing the occurrence of fake tracks.

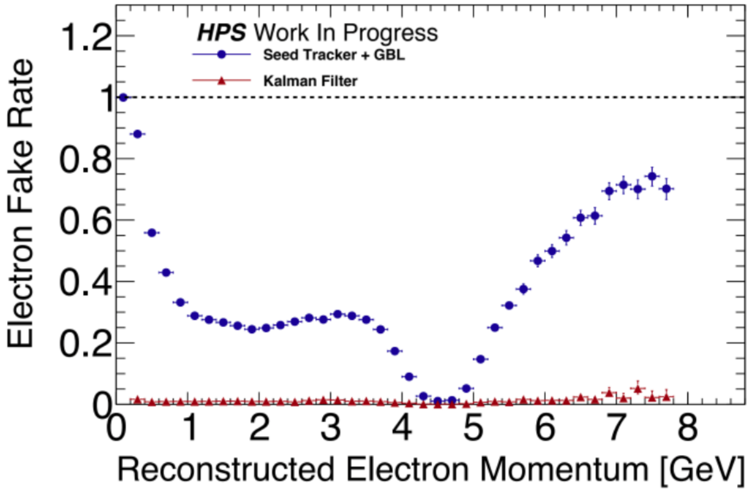


(a) Electrons.

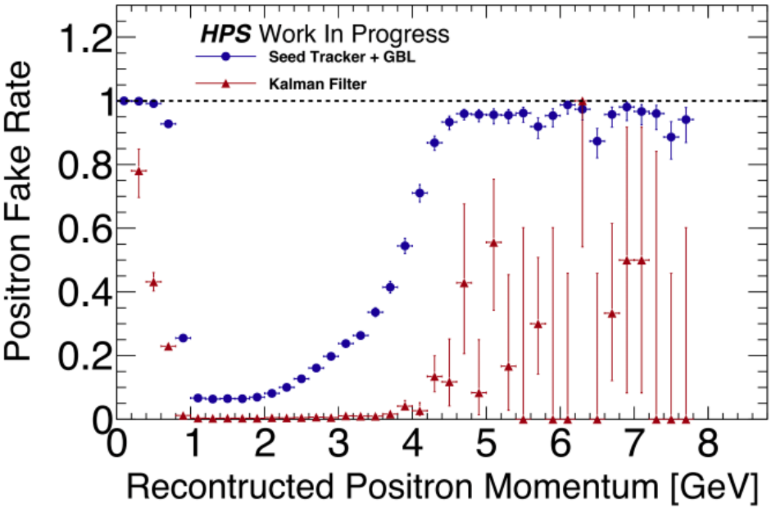


(b) Positrons.

Figure 4.7: Tracking efficiency for electrons and positrons in 2019.



(a) Electron fake rate (2019).

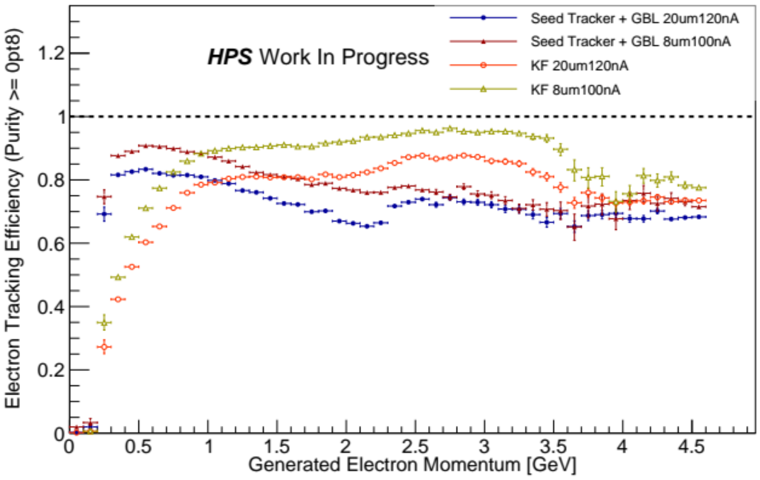


(b) Positron fake rate (2019).

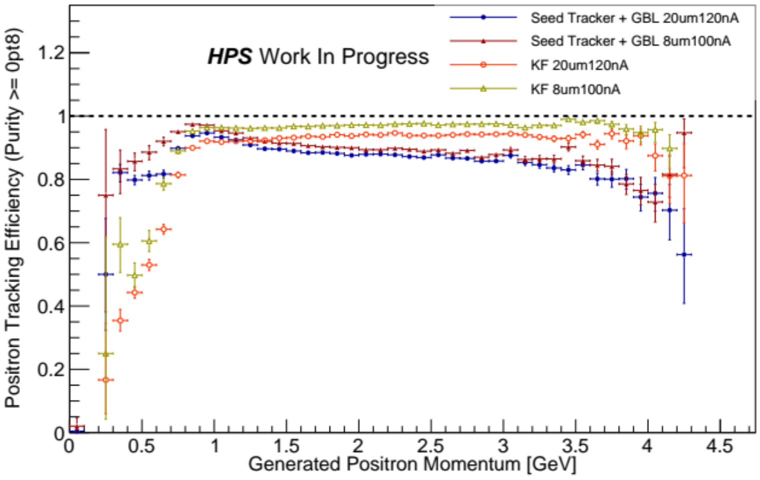
Figure 4.8: Fake rate for electrons and positrons in 2019. The fake rate is defined by track probability track probability < 0.8. KF tracking algorithm results in a negligible fake rate compared to SeedTracker+GBL, which shows a fake rate of approximately 30 %.

4.2.5.7 Tracking Performance as a Function of Sensor Occupancy

To evaluate the tracking performance under different sensor occupancy conditions, the same metrics above were used to compare two runs in 2019: a “high luminosity” run with a 20 µm thick target and a 120 nA beam current, and a “low luminosity” run with an 8 µm target and a 100 nA current. Figure 4.9 presents the tracking efficiency for high quality electron and positron tracks. In general, the electron tracking efficiency degrades with higher luminosity for both tracking algorithms. However, KF tracking demonstrates a superior electron tracking efficiency, particularly under high luminosity conditions. Figure 4.10 shows the tracking efficiency as a function of track $\tan\lambda$, which can be interpreted as the track angle in the vertical direction (y in HPS detector coordinates). The smallest values of $\tan\lambda$ with zero entries are extremely close to the beam and fall outside of the detector’s opening angle acceptance. As $\tan\lambda$ increases to cover the edge of the first tracking layer, the occupancy increases sharply before decreasing again. The primary region of interest includes the smallest values of $\tan\lambda$ within the acceptance. Here, the KF tracking shows improved electron tracking efficiency compared to SeedTracker+GBL, particularly in the highest occupancy regions. Lastly, Figure 4.11 shows the tracking fake rate as a function of momentum for different luminosities (corresponding to the combination of target thickness and beam current). The fake rate for SeedTracker+GBL tracks increases with luminosity, as expected due to increased pile-up and scattered beam electrons. However, the KF tracking fake rate remains low regardless of luminosity.

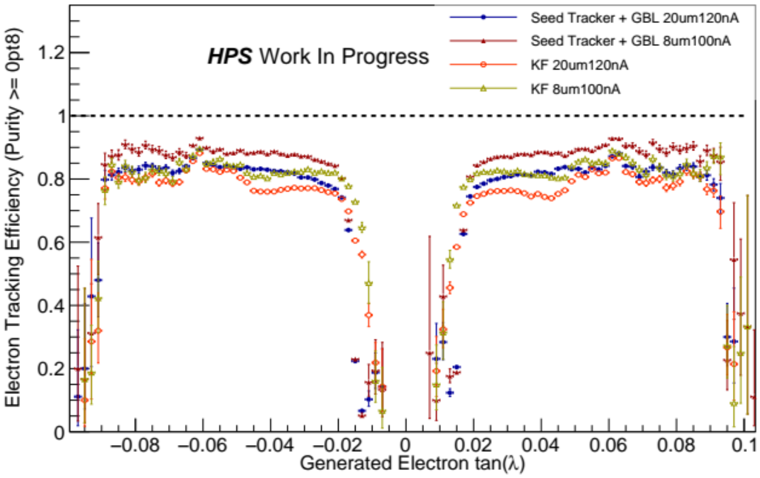


(a) Electron tracking efficiency (2019).

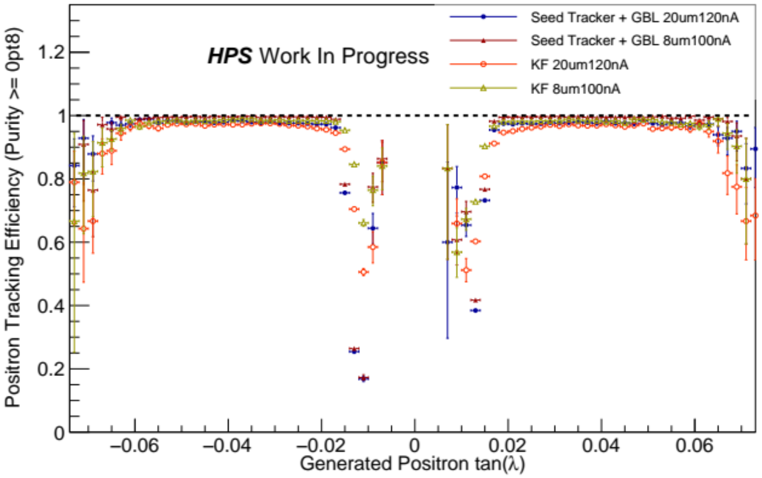


(b) Positron tracking efficiency (2019).

Figure 4.9: Luminosity dependent tracking efficiency for electrons and positrons in 2019. Both KF and GBL tracking are shown.

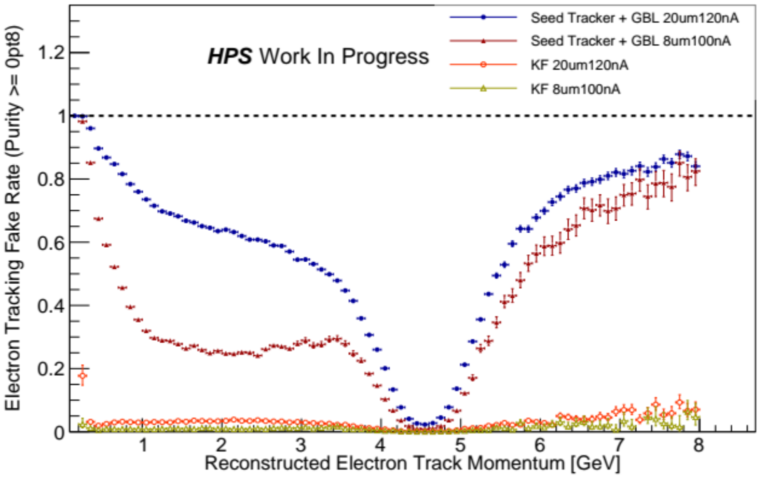


(a) Electron tracking efficiency vs. tanλ (2019).

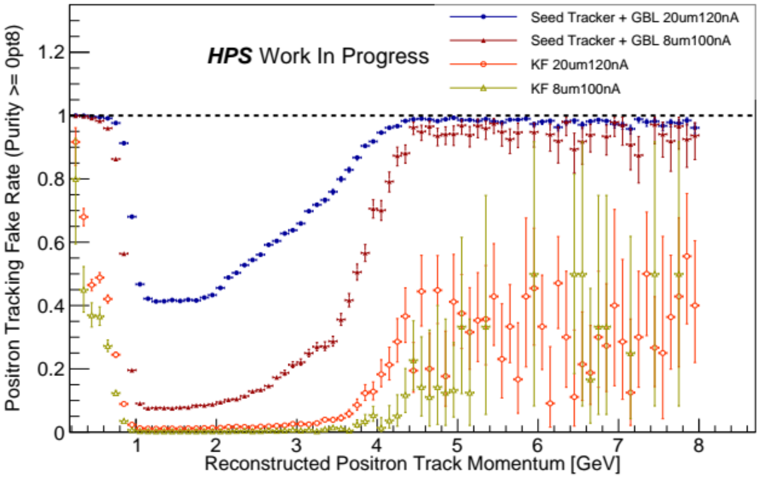


(b) Positron tracking efficiency vs. tanλ (2019).

Figure 4.10: Luminosity dependent tracking efficiency for electrons and positrons vs. tanλ in 2019. Both KF and GBL tracking are shown.



(a) Luminosity dependent electron fake rate vs. momentum (2019).



(b) Luminosity dependent positron fake rate vs. momentum (2019).

Figure 4.11: Luminosity dependent fake rate versus momentum for electrons and positrons in 2019. Both KF and GBL tracking are shown.

4.2.6 Track Cluster Matching

The 2016 analysis requires that both electron and positron tracks are matched with calorimeter clusters. To achieve this, it is necessary to extrapolate tracks to the face of the Ecal so that the track position information can be used to match the track with an energy cluster. While the magnetic field is fairly uniform through the tracking layer volume of the SVT, the field strength decreases rapidly between the last SVT layer and the face of the Ecal, where the field is approximately 0 T. The Kalman Filter tracking code accounts for this non-uniform field and propagates tracks from the last SVT hit location to the Ecal using Runge-Kutta integration [61].

The original track-cluster matching algorithm used with SeedTracker+GBL is not compatible with KF tracking, so a new track-cluster matching algorithm was developed. The new algorithm iterates through each track in the event and maps the position residual between the track extrapolated to the Ecal and every time-coincident Ecal cluster ($\mathcal{O}$(ns)) in the same volume (Top/Bottom of Ecal). After mapping all of the track-cluster position residuals, each track is matched to the nearest cluster to form track-cluster pair candidates. At this stage, some Ecal clusters may be matched to multiple tracks. Each pair candidate is checked to see if any other pair candidate shares the same cluster. If so, all pair candidates with that cluster are added to a list of duplicates. Only the track-cluster pair candidate with the smallest position residual in the list of duplicates is retained. The track-cluster position residuals map is modified to reflect this, and all other track-cluster position residuals corresponding to this cluster are removed. This cluster can no longer be matched to the duplicate tracks. Each track is then again matched to the nearest cluster in the updated track-cluster residuals map, and the duplicate removal process is repeated. This chain of matching pairs and removing duplicates is performed n times for n Ecal clusters in the event, ensuring that each cluster has the opportunity to be uniquely matched to a track. Clusters that are not matched to any tracks are considered photon objects. The position and time coincidence requirements for track-cluster matching are given in the Table 5.1.

4.3 Vertexing

The displaced vertex search utilizes the fast Billoir vertex fitting algorithm [62] to efficiently and accurately reconstruct both primary and displaced secondary vertices. This algorithm provides the vertex position and the track momenta at the vertex, which are then used to calculate the vertex mass and momentum. The vertex fitting uses the "perigee" helix parameterization, where tracks are parameterized based on their point of closest approach to the z-axis in the tracking coordinate system (along the B-field direction). The tracks are extrapolated from the perigee to an initial guessed vertex position. The extrapolated track parameters, as a function of the vertex position, are then are linearly expanded around this initial guess. The improved vertex position and track momenta at the vertex are found by minimizing the χ^2 function, which measures the deviation of the measured track parameters from the predicted parameters at the vertex location. This processes is iteratively repeated until the vertex position converges.

The local track parameterization at perigee becomes less accurate the further away the vertex is from the reference point, so highly displaced vertices may require an equally displaced reference point. However, using the tracking coordinate origin (close to the interaction point) as the reference is acceptable for secondary vertices within 15 mm to 20 mm from the reference [62]. All reconstructed e^- and e^+ objects are paired and fitted with the Billoir algorithm to form vertex candidates. The candidates are split into two categories: "Unconstrained" and "Target Constrained" vertices. Target Constrained vertices are consistent with originating from the beamspot at the target location in z, and are excluded from the displaced vertex search.

The Unconstrained vertices are further divided based on whether the two tracks are in the same detector hemisphere ("Unconstrained Vc collection"), or in opposite hemispheres ("Unconstrained V0 collection"). This analysis focuses on highly boosted e^- e^+ pairs originating from decays near the beam-axis, so only the Unconstrained V0 collection is used. Reconstruction requirements for the Unconstrained V0 collection used in this analysis are described in Table 5.1. Additional quality and analysis-specific selections are detailed in Section 5.1. In addition to opposite charge vertices, e^- pairs are also vertexed to form Moller candidates ($e^- + e^- \to e^- + e^-$), both target-constrained and unconstrained. These Moller objects are used to study and correct the mass resolution differences between MC and data, as detailed in in Section 4.5.

4.4 Monte Carlo

The data analysis requires realistic Monte Carlo (MC) simulations for various aspects of the experiment, including the primary backgrounds processes and signal, the complete detector geometry, and the detector readout response. MC background samples are necessary for comparing simulation and data, developing analysis techniques, and scaling the expected signal rate to the data as described in Section 5.1.4. Additionally, reconstructed beam electrons and Moller events provide standard candles for comparing the track momentum and invariant mass resolutions between MC and data. These comparisons are used to correct the discrepancies in Section 4.5.

This section introduces the MC software packages used to simulate events in HPS, and describes the MC production chain and samples used in this analysis. The MC chain generally includes event generation, target interactions, preparing events to pass through the simulated detector, passing events through the complete simulated detector, emulating the detector readout response, and then finally reconstructing the readout data as though it were real data. Electron beam events are generated and reconstructed independently, and mixed in with other event categories as needed.

4.4.1 Event Generation

The three signal-like background processes, Radiative Tridents (RAD), Tridents (RAD + Bethe-Heitler (BH), and Wide-Angle Bremsstrahlung (WABs) are generated using MadGraph4 [63]. The Feynman diagrams for these processes were previously shown in Section 3.6. The diagram for the SIMPs signal is found in Section 2.5.1.

Strongly-Interacting Massive Particles (SIMPs) signal events (prompt A' production followed by $A' \to V_D + \pi_D$) are also generated with MadGraph4. Long-lived V_Ds are initially simulated with a flat lifetime distribution $c\tau = 200\,\mathrm{mm}$ (extending out to the second tracking layer), with the appropriate decay function applied post reconstruction as described in Section 5.1.4. The generator level mass ratios of the A', V_D, and π_D are fixed according to the benchmark SIMPs parameters provided in Section 2.5. The MC SIMP samples were generated in the range of 24 MeV to 124 MeV with 2 MeV intervals.

Primary backgrounds, including elastically/quasi-elastically scattered beam electrons, Moller scattering, pair production, and bremsstrahlung, are collectively categorized as "beam events" and are generated using the EGS5 electron gamma shower generator software [64] rather than MadGraph. EGS5 also simulates the interaction of MadGraph final-state particles with the HPS target.

4.4.2 Stdhep Tools

Stdhep [65] tools are used to adjust and configure the EGS5 output events before they undergo simulation in the detector. Stdhep adds the parent particle truth information, such as PDG ID and Lorentz vector, as well as daughter particle information, to the EGS5 events. It also applies the beam rotation, diffusion, and target offsetting to accurately reflect events in data. Stdhep also creates realistic beam events by sampling and building Poisson-distributed beam bunches, with the Poisson mean representing the number of electrons per beam bunch.

4.4.3 HPS Detector Geometry Simulation

The Stdhep-configured final-state particles are passed through a complete simulation of the Heavy Photon Search (HPS) detector geometry using the GEANT4-based package SLIC (Simulator for the Linear Collider) [66]. SLIC simulates interactions with all of the sensitive detector materials, such as the tracking silicon strips and the Ecal crystals, generating simulated hits that include time and energy deposition information. These hits are later processed in the MC readout phase, generating a realistic detector hardware response to the particles. The simulation also accounts for interactions with structural support materials, such as the aluminum support modules and vacuum flanges.

4.4.4 Signal-Beam Merging

Signal events are initially generated without beam background. They are merged with the beam background after the SLIC detector simulation step and before the readout emulation step. The signal sample is filtered and spaced by 250 events, which roughly corresponds to the event window in the trigger system. Each signal event should have a corresponding beam bunch, so if a signal file has 10,000 events, and each event occurs every 250th event window, the signal is mixed with 250,000 beam bunches.

4.4.5 MC Readout and Reconstruction

The SLIC simulated tracker and Ecal hits are converted into raw hits as they would be read out by the HPS detector hardware under real operating conditions, using the readout simulation package in hps-java. This package simulates the physical response of the digitization and readout hardware for the SVT and Ecal and emulates the trigger system, producing realistic simulated raw data that mirrors real data. All of the MC samples used in this analysis were reconstructed using theRun 7984 conditions in the HPS database, which corresponds to a large statistics sample with ideal running conditions.

4.4.6 MC Sample Normalization

The trident and WAB MC simulations are normalized by the Integrated Cross Section (ICS) over the total number of generated events. The normalization calculation is detailed in [67]. The normalization parameters for the MC background samples are presented in Table 4.1.

Sample	μ of ICS	σ of ICS	# of good files	# of generated events per file
RAD	66.36 µb	0.6678 µb	10k	10k
Trident-Trig	1.416 mb	0.004 310 mb	10k	50k
WAB	0.1985 b	0.019 73 b	10k	100k

Table 4.1: Normalization parameters for the RAD, Trident-Trig, and WAB samples

4.4.7 MC Hit Killing

The SVT hit efficiency varies across the tracking layers, with the first layer showing significantly reduced efficiency due to occupancy effects. The hit efficiencies are not simulated in the MC samples, so the MC signal efficiency for events that leave hits in the two tracking layers is overestimated compared to data. This discrepancy is corrected by applying a hit-killing algorithm to the MC samples in order to mimic the observed hit efficiencies in data. The hit efficiencies for a given layer are measured in an unbiased manner by removing a hit on track from the layer of interest. The track is then refit without using the removed hit and extrapolated back to the layer where the hit was removed to see if a viable alternative hit exists within some proximity window. The hit killing algorithm uses the hit efficiencies observed in data as a function of sensor channel number, and probabilistic-ally removes MC hits on track in the first layer. These tracks are then re-reconstructed. The new tracks cannot use the removed hit, but it can pick up other nearby hits that pass the tracking requirements.

4.5 MC Mass Resolution

The signal search is performed over a range of invariant mass windows, and it's necessary to accurately estimate the signal mass resolution in each window. The natural width of V_D is negligible compared to the SVT detector resolution, which is on the order of a few MeV, so the observed mass resolution is that of the detector alone.

The mass resolution can be estimated by fitting the MC signal invariant mass distribution with a Gaussian function, using the standard deviation of the fit as the measure of the mass resolution. However, it's not necessarily true that the MC mass resolution is actually equivalent to the experiment, and in fact the 2016 bump-hunt analysis demonstrated that the MC mass resolution is too good compared to data [68]. The MC mass resolution was compared to data using the Moller process ($e^-e^- \rightarrow e^-e^-$), which acts as a standard candle with a fixed and known invariant mass depending on the beam energy. The invariant mass of the Moller final state is equal to the square root of the total four-momentum of the initial electrons, and is calculated in the center of mass frame by

$$M(e^-e^-) = \sqrt{S} = \sqrt{2m_{e^-} + 2E_{\text{beam}} m_{e^-}} \approx \sqrt{2E_{\text{beam}} m_{e^-}} = 48.498\,\text{MeV}, \tag{4.14}$$

where $E_{\text{beam}} = 2.3\,\text{GeV}$ in the 2016 run. The Moller mass has no width, so the reconstructed Moller mass resolution is equivalent to the detector mass resolution at the mass given by Equation (4.14). The Moller final state particles are identical in mass to the signal final state and have the same electromagnetic properties, so the Moller mass resolution is expected to be similar to the signal mass resolution at the Moller mass.

The 2016 bump-hunt analysis [68] used MC-generated Moller events with MC simulated beam to measure the detector acceptance for Moller final states. It was found that typically only one of the final state electrons is within the Ecal acceptance. However, both electrons are generally within the SVT acceptance, and the e^- tracks tend to point to four distinct regions on the face of the Ecal. These four regions define the fiducial area, which is used in combination with track time and track momentum sum cuts to select a relatively pure sample of Moller events. The cuts differ slightly between MC and data, as summarized in Table 4.2 and Table 4.3. The same cuts are used here to select Moller events; however, this analysis employs unconstrained vertices rather than constrained ones, and uses Kalman tracking instead of SeedTrack+GBL. The MC Moller samples used here are described in Section 4.4, and the data Moller samples are described in Section 3.1.1.

The unconstrained vertex Moller mass resolution in data and MC is calculated by fitting the invariant mass distributions with a Gaussian function, as shown in Figure 4.12. The fit results are summarized in Table 4.4. The fitted Moller mass peak means are consistent with Equation (4.14), but the MC mass resolution is approximately 26 % better

Data Sample	Δt_{min} [ns]	Δt_{max} [ns]	$\Delta\lvert t_{\text{Tracks}}\rvert$ [ns]	$P_{\text{sum,min}}$ [GeV]	$P_{\text{sum,max}}$ [GeV]
Data	-3.0	+2.5	N/A	2.1	2.45
MC	N/A	N/A	2.5	2.15	2.42

Table 4.2: Summary of the timing and momentum sum cuts used for the Møller selection.

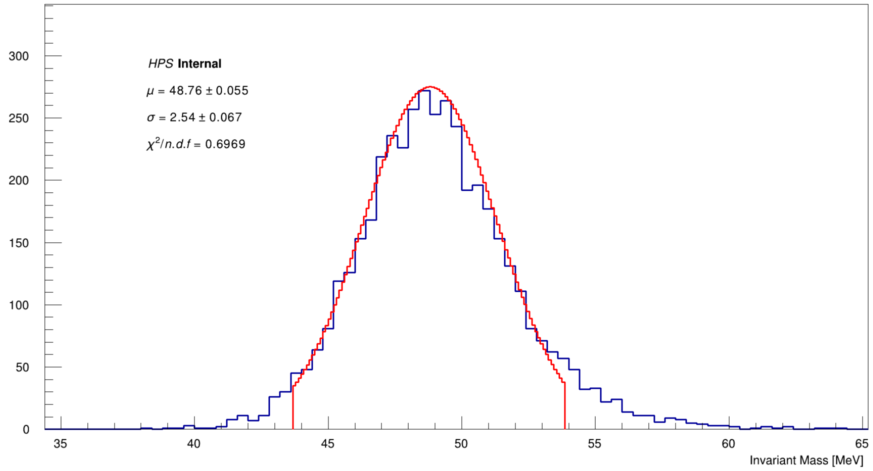
Data Sample	$\Delta x_{\text{top,min}}$ [mm]	$\Delta x_{\text{top,max}}$ [mm]	$\Delta x_{\text{bot,min}}$ [mm]	$\Delta x_{\text{bot,max}}$ [mm]
Data	-4.72	6.15	-7.51	2.98
MC	-4.89	4.82	-4.98	4.52

Table 4.3: Summary of the positional cuts used for the Møller selection. Note that cut values for MC and data are different.

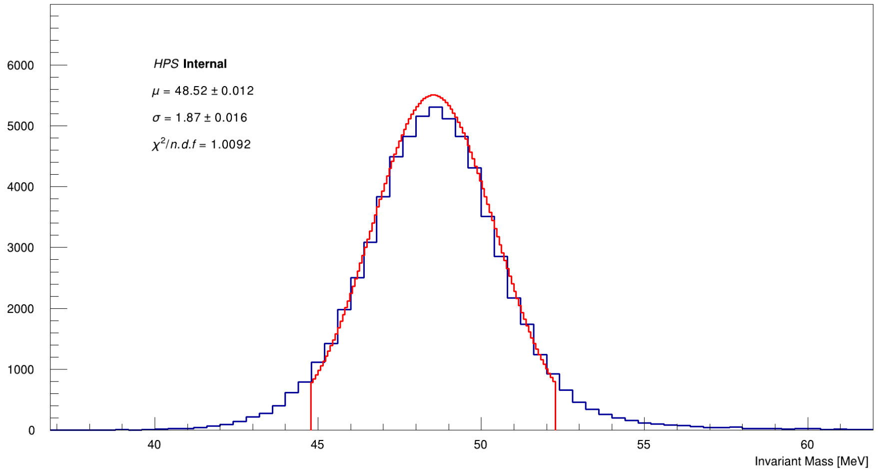
than that of the data. This discrepancy must be corrected to accurately estimate the signal mass resolution using the MC signal.

	μ [MeV]	σ [MeV]	σ_{err} [MeV]
Data	48.76	2.54	0.067
MC	48.52	1.87	0.016

Table 4.4: Comparison of mean (μ), standard deviation (σ), and standard deviation error (σ_{err}) values for the fitted invariant mass of Moller events in Data and MC in MeV.



(a) Data.



(b) MC without any corrections.

Figure 4.12: Fitted $e^{-}e^{-}$ reconstructed invariant mass spectrum using the Moller selection.

4.5.1 Momentum Resolution with FEEs

The mass resolution for a vertex that is reconstructed using the two daughter particles is directly related to the momentum resolution because the invariant mass of two particles in the ultra-relativistic limit is given by

$$M(ee) = 2\sqrt{P_1 P_2}\sin\left(\frac{\theta}{2}\right), \tag{4.15}$$

where θ is the angle between the two particles. Consequently, if the momentum resolution is better in MC compared to data, the invariant mass resolution will also be better, as observed. Additionally, if the MC momentum resolution can be adjusted to match that of the data, it is expected that the mass resolution will be corrected as well.

The momentum resolution is studied using the only events in data with precisely known momentum—elastically scattered beam electrons with momentum peaked at $E_{\text{beam}} = 2.3\,\text{GeV}$, known as Full Energy Electron (FEE)s.

MC FEEs are skimmed from the trident+beam MC samples described in Section 4.4, and the FEEs in data are taken from run 7800 in the 10% data sample described in Section 3.1.1. FEEs are initially selected using the "singles0" trigger defined in Section 3.5. Accidental contamination is reduced by placing a minimum and maximum cut on the x position of the track at the face of the Ecal, where the FEEs tend to point to the central region of the Ecal near the "electron-hole", and additional cuts are placed on the track timing and quality to clean the samples. The FEE selection is the same for both MC and data, and is summarized in Table 4.5

The momentum resolution depends slightly on the number of hits on track, and whether the track is reconstructed in the top or bottom volume of the SVT, so the FEE

Cut Description	**Requirement**
Track Quality	$\chi^2_{\text{Track}} < 6$
P_{Track} Min	$1.0\,\text{GeV}$
P_{Track} Max	$4.0\,\text{GeV}$
X_{Track} at Ecal Max	$X_{\text{Track}} < 50.0\,\text{mm}$
X_{Track} at Ecal Min	$X_{\text{Track}} > -100.0\,\text{mm}$
Track Time Max	$t_{\text{Track}} > -10.0\,\text{ns}$
Track Time Min	$t_{\text{Track}} < 10.0\,\text{ns}$

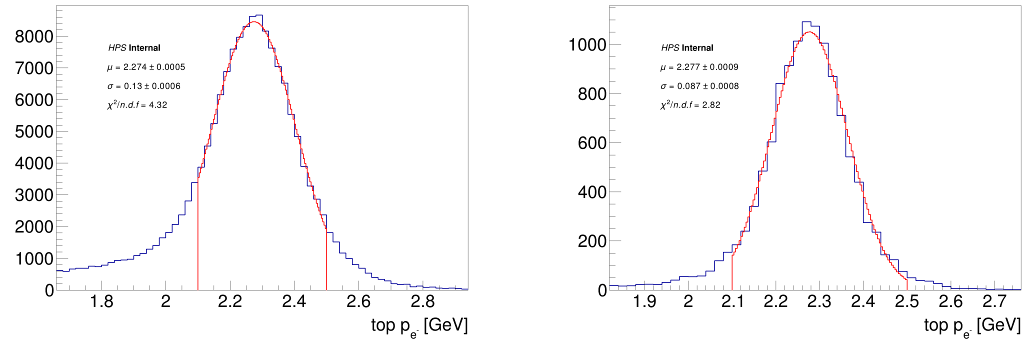
Table 4.5: Track selection used to select a clean sample of FEE candidates. The same cuts are used in both MC and data.

selection is split into six categories based on the volume, and whether a track has 12 hits, 11 hits, or 10 hits. The momentum resolution for each category of FEE in data and MC is determined by fitting the track momentum distribution with a Gaussian function. Figure 4.13 shows the fit results for tracks with 12 hits, Figure 4.14 presents the results for tracks with 11 hits, and Figure 4.15 displays the results for tracks with 10 hits. As expected, the MC FEE track momentum resolution is better than that of the data. The MC track momentum can be smeared to match the resolution observed in the data, which will also smear the mass resolution according to Equation (4.15).

The MC track momentum smearing factor is defined relative to the reconstructed momentum by

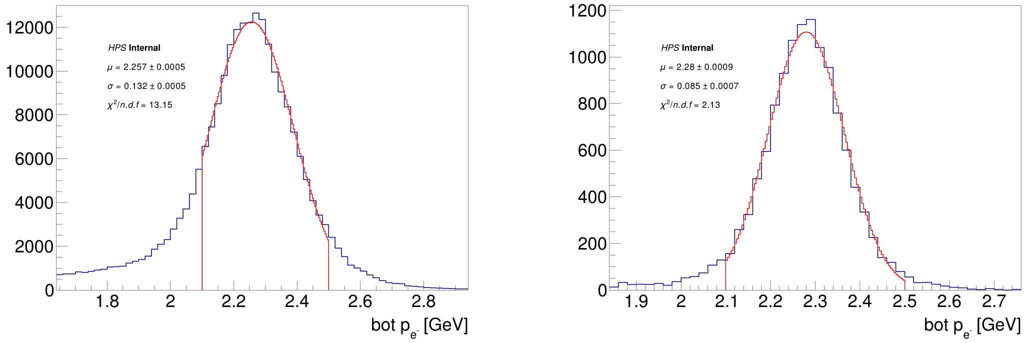
$$\Sigma_{\text{smear}} \equiv \frac{\sigma_{\text{smear}}}{P_{\text{mc}}} = \sqrt{\left(\frac{\sigma_{\text{data}}}{\mu_{\text{data}}}\right)^2 - \left(\frac{\sigma_{\text{mc}}}{\mu_{\text{mc}}}\right)^2}, \tag{4.16}$$

where σ_{data} and σ_{mc} are the fitted FEE mass resolutions in data and MC, and μ_{data} and μ_{mc}



(a) Data, Top volume.

(b) MC, Top volume.



(c) Data, Bottom volume.

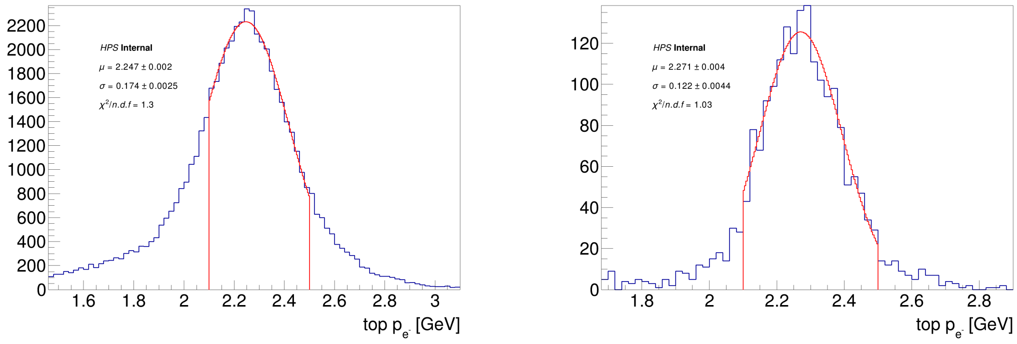
(d) MC, Bottom volume.

Figure 4.13: Fitted reconstructed e^- track momentum using the FEE selection for 12 hit tracks.

are the respective fitted momentum means. The FEE momentum fit results and smearing factors for each volume and hit category are summarized in Table 4.6. Tracks with less than 10 hits will use the “10hit” smearing coefficient. The smeared MC track momentum is therefore given by

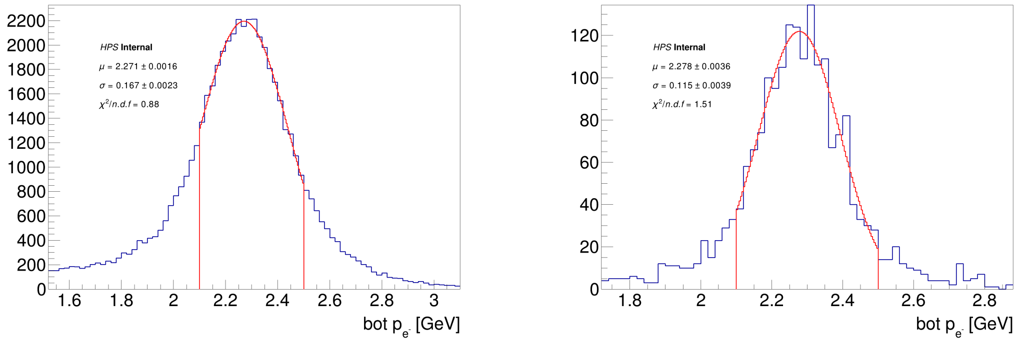
$$P_{\text{smear}} = P_{\text{reco}} + \mathcal{X}\Sigma^{\text{smear}} P_{\text{reco}}, \tag{4.17}$$

where $\mathcal{X} \sim \mathcal{N}(0,1)$, and P_{reco} is the reconstructed track momentum.



(a) Data, Top volume.

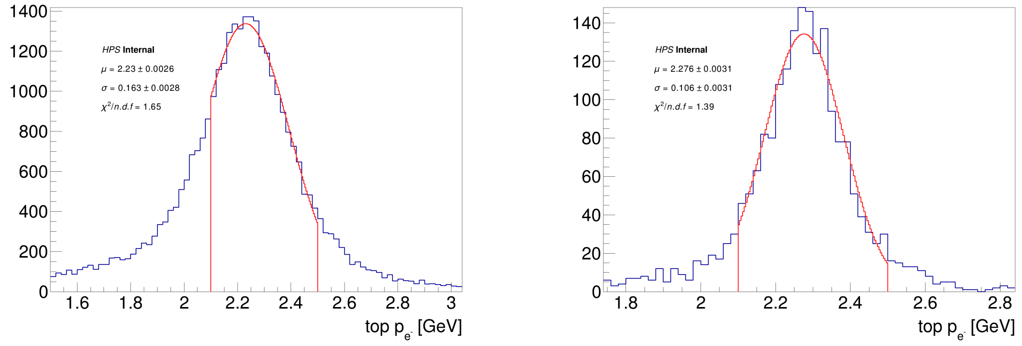
(b) MC, Top volume.



(c) Data, Bottom volume.

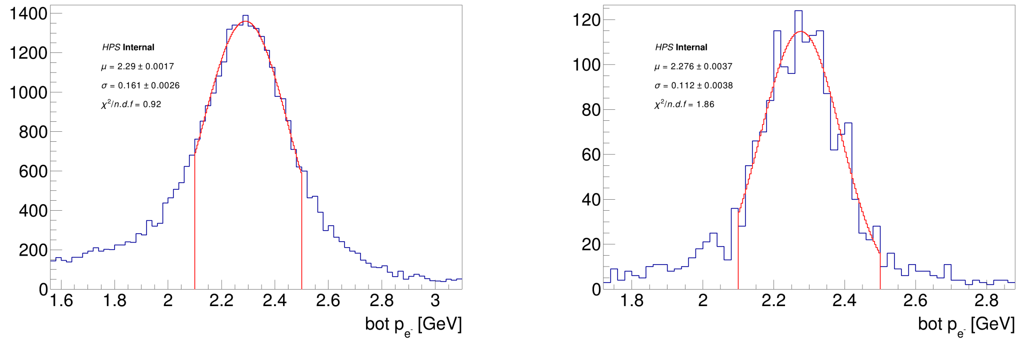
(d) MC, Bottom volume.

Figure 4.14: Fitted reconstructed e^- track momentum using the FEE selection for 11 hit tracks.



(a) Data, Top volume.

(b) MC, Top volume.



(c) Data, Bottom volume.

(d) MC, Bottom volume.

Figure 4.15: Fitted reconstructed e^- track momentum using the FEE selection for 10 hit tracks.

Variable/Category	**Data**	**mc**
$\mu_{\text{Top12hits}}$ [GeV]	2.274 ± 0.0005	2.277 ± 0.0009
$\mu_{\text{Bot12hits}}$ GeV	2.257 ± 0.0005	2.28 ± 0.0009
$\mu_{\text{Top11hits}}$ GeV	2.247 ± 0.002	2.271 ± 0.004
$\mu_{\text{Bot11hits}}$ GeV	2.271 ± 0.0016	2.78 ± 0.0036
$\mu_{\text{Top10hits}}$ GeV	2.23 ± 0.0026	2.276 ± 0.0037
$\mu_{\text{Bot10hits}}$ GeV	2.29 ± 0.0017	2.276 ± 0.0037
$\sigma_{\text{Top12hits}}$ GeV	0.13 ± 0.0006	0.087 ± 0.0006
$\sigma_{\text{Bot12hits}}$ GeV	0.132 ± 0.0005	0.085 ± 0.0007
$\sigma_{\text{Top11hits}}$ GeV	0.174 ± 0.0025	0.122 ± 0.0044
$\sigma_{\text{Bot11hits}}$ GeV	0.167 ± 0.0023	0.115 ± 0.0039
$\sigma_{\text{Top10hits}}$ GeV	0.163 ± 0.0028	0.106 ± 0.0031
$\sigma_{\text{Bot10hits}}$ GeV	0.161 ± 0.0026	0.112 ± 0.0038
$\Sigma^{\text{smear}}_{\text{Top12hits}}$ GeV	N/A	0.0427
$\Sigma^{\text{smear}}_{\text{Bot12hits}}$ GeV	N/A	0.0448
$\Sigma^{\text{smear}}_{\text{Top11hits}}$ GeV	N/A	0.0554
$\Sigma^{\text{smear}}_{\text{Bot11hits}}$ GeV	N/A	0.0535
$\Sigma^{\text{smear}}_{\text{Top10hits}}$ GeV	N/A	0.0561
$\Sigma^{\text{smear}}_{\text{Bot10hits}}$ GeV	N/A	0.0504

Table 4.6: Summary of mean (μ) and standard deviation (σ) values for Top and Bottom tracks, with their uncertainties and smearing factors,(Σ^{smear}) for different hit counts, given in GeV.

4.5.2 Corrected MC Mass Resolution

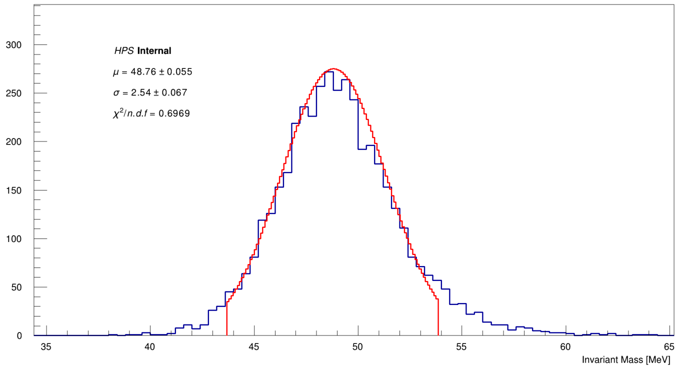
The MC mass resolution is corrected by solving Equation (4.15) using the smeared track momentum from Equation (4.17):

$$M(ee)^{\text{smear}} = 2\sqrt{P_{1,\text{smear}} P_{2,\text{smear}}} \sin\frac{\theta}{2} = \sqrt{\frac{P_{1,\text{smear}}}{P_{1,\text{rec}}} \frac{P_{2,\text{smear}}}{P_{2,\text{rec}}}} M(ee)^{\text{rec}}, \tag{4.18}$$

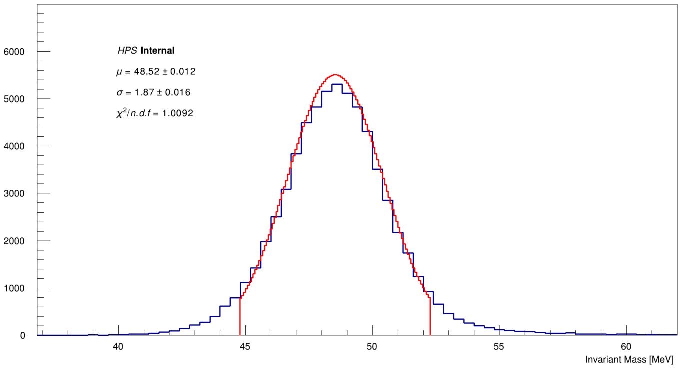
where $P_{1,\text{smear}}$ and $P_{2,\text{smear}}$ are the smeared track momenta given by Equation (4.17), $P_{1,\text{rec}}$ and $P_{2,\text{rec}}$ are the reconstructed track momenta, and $M(ee)^{\text{rec}}$ is the reconstructed vertex invariant mass before smearing. The corrected MC Moller mass resolution, using the track momentum smearing coefficients defined in Table 4.6, is shown in Table 4.7, with the fit presented in Figure 4.16. The smeared MC mass resolution now differs from data by approximately −8.6 %, significantly reduced from the initial −26 % difference, with the difference now within 0.22 MeV.

	μ [MeV]	σ [MeV]	σ_{err} [MeV]
Data	48.76	2.54	0.067
MC	48.52	1.87	0.016
$\text{mc}_{\text{smeared}}$	48.41	2.32	0.017

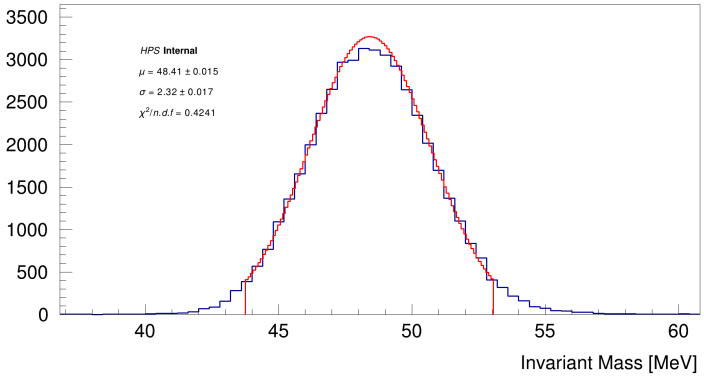
Table 4.7: Comparison of mean (μ), standard deviation (σ), and standard deviation error (σ_{err}) values for the Moller selected invariant mass distribution in Data, MC, and Smeared MC in MeV.



(a) Data.



(b) MC without corrections.



(c) MC with track momentum smearing.

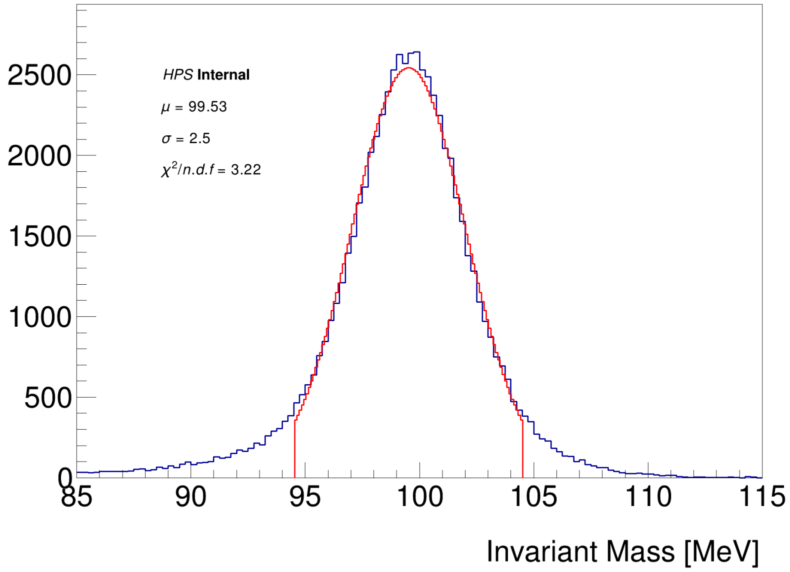
Figure 4.16: Fitted e^-e^- reconstructed invariant mass spectrum using the Moller selection. The corrected MC Moller mass resolution is now approximately equivalent to data after smearing the track momentum.

The MC signal mass resolution is corrected in the same manner used to correct the MC Moller mass resolution. The results presented here refer to signal events with the preselection, signal region, L1L1, and $n_{\mathrm{vtx}} = 1$ selections applied. The momentum resolution will differ for other layer hit categories, but those are outside of the scope of this analysis. The left plot in Figure 4.17 shows the fitted reconstructed invariant mass distribution for 100 MeV MC signal, and the right plot shows the corrected mass distribution; the corrected mass resolution nearly doubles from $\sigma_m = 2.5\,\mathrm{MeV}$ to $\sigma_m = 4.48\,\mathrm{MeV}$.

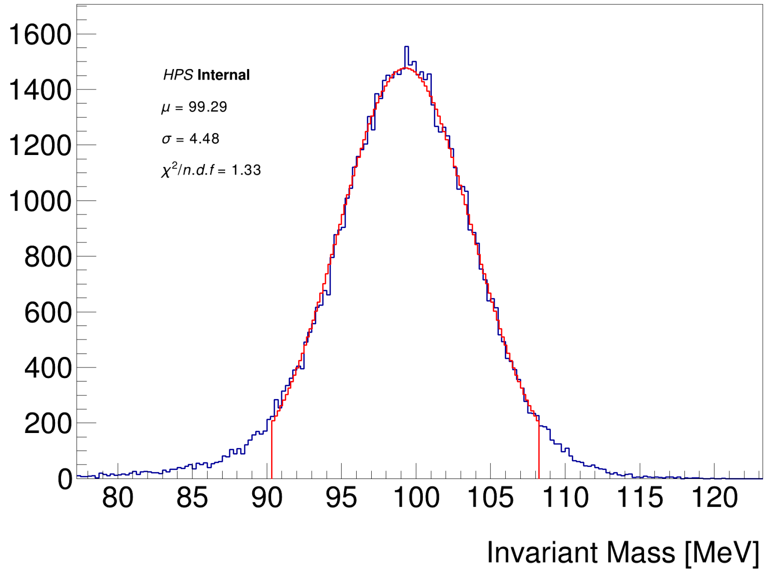
The corrected mass resolution is calculated for MC signal masses between 30 MeV to 124 MeV at 2 MeV intervals, and is then parameterized as a function of mass by fitting the results with a 2nd degree polynomial. The mass resolution polynomial fit is shown by the red fit line in Figure 4.18, and is given by

$$\sigma_m(m) = 0.75739851 + 0.031621002m + 5.2949672e^{-5}m^2, \tag{4.19}$$

where m is the mass in MeV.



(a) Unsmeared invariant mass distribution.



(b) Corrected MC signal mass distribution after applying momentum smearing.

Figure 4.17: Example of MC signal at 100 MeV.

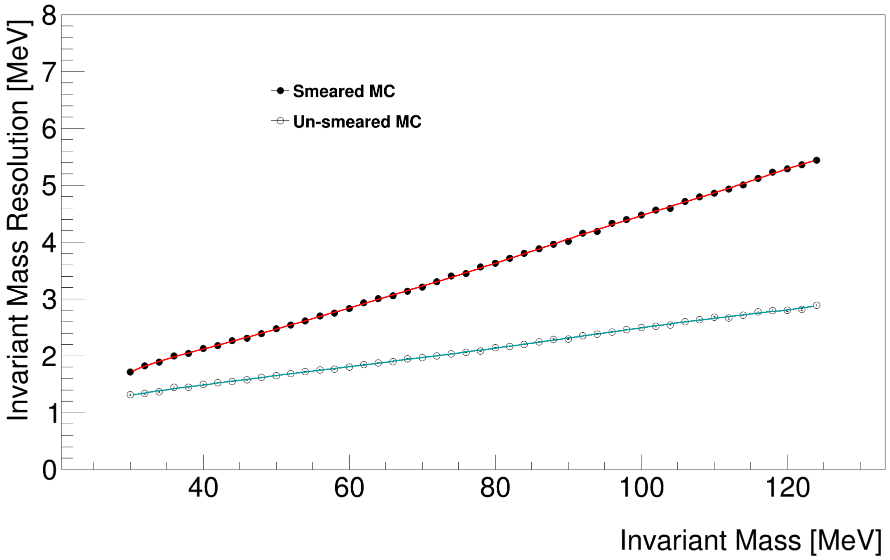


Figure 4.18: Invariant mass resolution parameterization. The solid black markers correspond to the corrected mass resolution, and are fit with the 2nd degree polynomial shown in red. The open circle markers show mass resolution before momentum smearing is applied.

Chapter 5

Event Selection

5.1 Event Selection

The fixed-target signal production rate is suppressed by the factor ϵ^2 compared to QED trident production, resulting in an expected signal rate that is roughly 4-12 orders of magnitude smaller than the background rate. The primary backgrounds include prompt QED tridents, WABs photon conversions at the target, and a dense population of multiple Coulomb and elastically scattered electrons that contaminate tracking. Due to the nature of these background processes, this analysis employs a displaced vertex search, which aims to resolve the low rate of signal events downstream from the prompt QED background. Despite the backgrounds being prompt in nature, reconstruction errors and large scattering events lead to a significant rate of falsely displaced background events downstream of the target location in the z-direction ($\mathrm{z}_{\mathrm{vtx}}$), as illustrated in Section 3.6.

In order to minimize the rate of falsely displaced background in the search, a series of analysis cuts are employed to eliminate poorly reconstructed background events, and events inconsistent with displaced signals. The selection process is divided into four categories: reconstruction, preselection, signal and control momentum regions, and the tight selection. Each of these selection stages applies increasingly strict criteria, detailed in the following sections. This event selection and analysis procedure was tuned using the MC samples described in Section 4.4 and 10 % of the 2016 data detailed in Section 3.1.1, corresponding to approximately $1096.27\,\mathrm{nb}^{-1}$. The MC samples used to finalize the preselection cuts did not have the hit-killing algorithm implemented (see Section 4.4.7). The MC samples also did not have the momentum smearing corrections derived in Section 4.5 applied. However, the preselection cuts are primarily data driven, and the impact of these differences between the MC and data are negligible at this stage. With this in mind, the tight selection does use the corrected MC samples with the hit-killing algorithm and momentum smearing corrections applied.

Some of the tight selection variables discussed in this section are correlated with the signal mass and lifetime, which is a function of ϵ. Given this correlation, the MC signal samples used to optimize the tight selection are reweighted based on the complete expected signal calculation, accounting for the particle lifetime. The expected signal calculation is derived in Section 5.1.4. All signal calculations presented in this section use the benchmark SIMP parameters detailed in Section 2.5.1, with $m_\pi/f_\pi = 4\pi$, corresponding to the maximum possible $\Gamma\left(A' \rightarrow V_D\pi_D\right)$.

5.1.1 Reconstruction-Level Cuts

The reconstruction-level cuts are intended to eliminate obviously mis-reconstructed objects and accidental associations between two uncorrelated particles. These cuts are kept intentionally loose to maintain large statistics samples for further analysis. Many of the selection criteria here are tightened further at a later stage.

For a given vertex candidate, both reconstructed tracks must be matched to an Ecal energy cluster, with the clusters in opposite volumes of the detector, consistent with an e^+e^- event. The track-cluster time difference must be less than $10\,\mathrm{ns}$ after applying a $56\,\mathrm{ns}$ ($43\,\mathrm{ns}$) timing correction in data (MC) [68]. The track-cluster transverse position differences at the Ecal must be less than $20\,\mathrm{mm}$ to eliminate accidental matches. The electron beam bunches provided by CEBAF are spaced at $2\,\mathrm{ns}$ intervals. Therefore, the two Ecal clusters belonging to a vertex candidate must occur within $2.5\,\mathrm{ns}$ of each other. The cluster timing is used instead of the track timing because the Ecal has significantly better time resolution. This cut is further tightened in the preselection stage. Since the analysis is concerned with trident events and not beam electrons, full-energy beam electrons are excluded by requiring that vertex electrons tracks have a momentum less than $2.15\,\mathrm{GeV}$, which is close to $E_{\mathrm{beam}} = 2.3\,\mathrm{GeV}$. Finally, candidates with vertex momentum greater than $2.8\,\mathrm{GeV}$ are removed to exclude obviously mis-reconstructed candidates that have momentum well above the beam energy.

The reconstruction level cuts are summarized in Table 5.1, with many of these cuts tightened in the subsequent preselection stage.

Cut Description	**Requirement**
Ecal clusters in opposite volumes	$e^{-}\text{Cluster}_y \times e^{+}\text{Cluster}_y$ <0
Track-Cluster Time Difference (Data)	$\lvert\text{Track}_t - \text{Cluster}_t - 56\,\text{ns}\rvert < 10\,\text{ns}$
Track-Cluster Time Difference (MC)	$\lvert\text{Track}_t - \text{Cluster}_t - 43\,\text{ns}\rvert < 10\,\text{ns}$
Track-Cluster X Position Difference	$\lvert x_{\text{TrackAtEcal}} - x_{\text{Cluster}}\rvert < 20.0\,\text{mm}$
Track-Cluster Y Position Difference	$\lvert\text{TrackAtEcal}_y - \text{Cluster}_y\rvert < 20.0\,\text{mm}$
Cluster Time Difference	$\Delta_t(\text{Cluster}_{e^-}, \text{Cluster}_{e^+}) < 2.5\,\text{ns}$
Beam electron cut	$p_{e^-} < 2.15\,\text{GeV}$
Vertex Momentum	$p_{\text{vtx}} < 2.8\,\text{GeV}$

Table 5.1: Reconstruction level requirements. Track-Cluster time difference in MC and data is corrected using offsets calibrated in [68]. The track positions are found by extrapolating the track from the last layer hit to the face of the Ecal.

5.1.2 Preselection

The preselection cuts are designed to provide a tighter selection on events than those required at the reconstruction stage while maintaining large-statistics samples of high-quality, signal-like vertex candidates for further study. The preselection cuts are detailed in this section and summarized in Table 5.2.

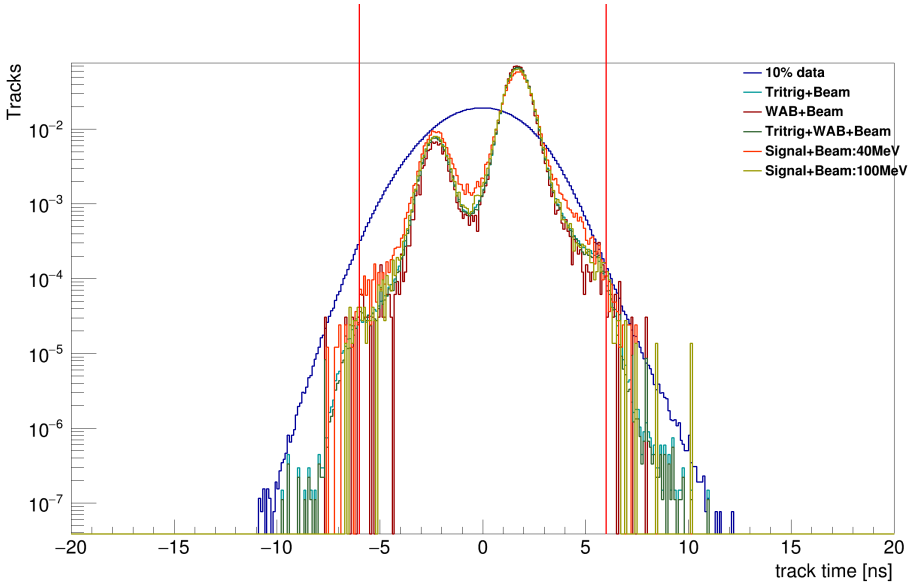
Cut Description	**Requirement**
Trigger	Pair1
Track Time	$\lvert \text{Track}_t \rvert < 6\,\text{ns}$
Cluster Time Difference	$\Delta_t(\text{Cluster}_{e^-}, \text{Cluster}_{e^+}) < 1.45\,\text{ns}$
Track-Cluster Time Difference	$\Delta_t(\text{Track,Cluster}) < 4.0\,\text{ns}$
Track Quality	Track $\chi^2/\text{n.d.f.} < 20.0$
Beam electron cut	$p_{e^-} < 1.75\,\text{GeV}$
Minimum Hits on Track	$\text{N}_{\text{2d hits on Track}} > 7.0$
Unconstrained Vertex Quality	$\text{vtx}_{\chi^2} < 20.0$
Vertex Momentum	$p_{e^- + e^+} < 2.4\,\text{GeV}$

Table 5.2: V_0 selection. The time offset for data is 56 ns and the time offset for MC is 43 ns.

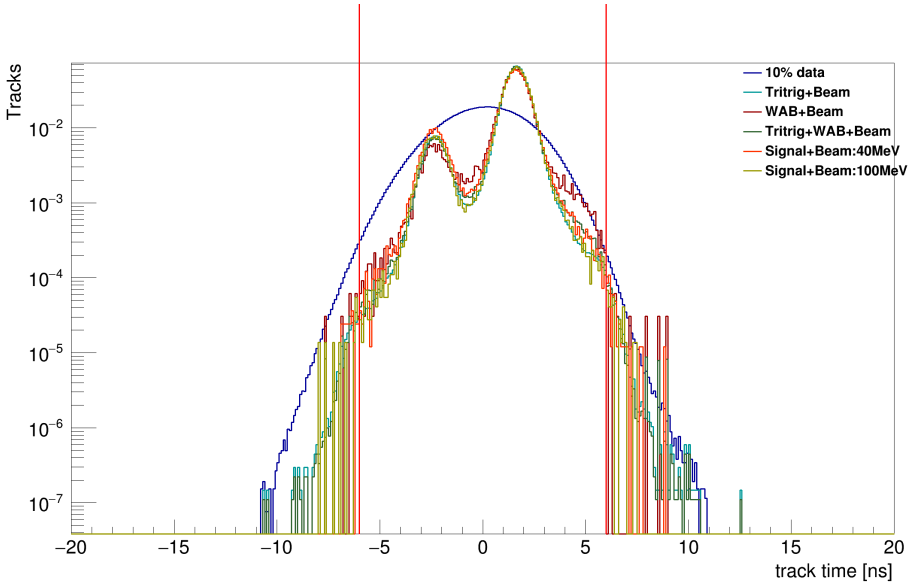
First, every vertex candidate in the 2016 analysis must satisfy the "Pair 1" trigger as described in Section 3.5. Tracks and clusters that are reconstructed out-of-time, meaning they likely do not correspond to the trigger, are reduced using track and cluster timing cuts. Tracks must occur within ±6 ns of the trigger time, as shown in Figure 5.1. Additionally, the time-coincidence between the two Ecal clusters for a vertex candidate is tightened to

1.45 ns, as shown in Figure 5.2. Lastly, the time difference between track and cluster is required to be less than ±4 ns, as shown in Figure 5.3.

Additional cuts are placed on the track quality to eliminate poorly reconstructed events than can lead to falsely displaced background. First, the track χ^2/dof is required to be less than 20. Figure 5.4 compares the track chi-squared distributions in MC and data, showing an overall discrepancy in the shapes. For this reason, the cut is kept fairly loose in order to minimize the systematic uncertainty. Electron tracks with a momentum greater than 1.75 GeV are eliminated to further reduce the rate of events reconstructed using elastically scattered beam electrons, which is a large component of the background. The impact of this cut is shown in Figure 5.5. The final track-level preselection requires that each track is reconstructed using at least 8 (out of 12 possible) SVT hits. The distribution of the number of hits on track is shown in Figure 5.6.



(a) Track time for electrons.



(b) Track time for positrons.

Figure 5.1: Reconstructed track time. Preselection positron and electron track time is shifted by $-1.5\,\mathrm{ns}$ in data, and $-2.2\,\mathrm{ns}$ in MC. A loose cut on $|\mathrm{Track}_{\mathrm{time}}| < 6\,\mathrm{ns}$ is applied.

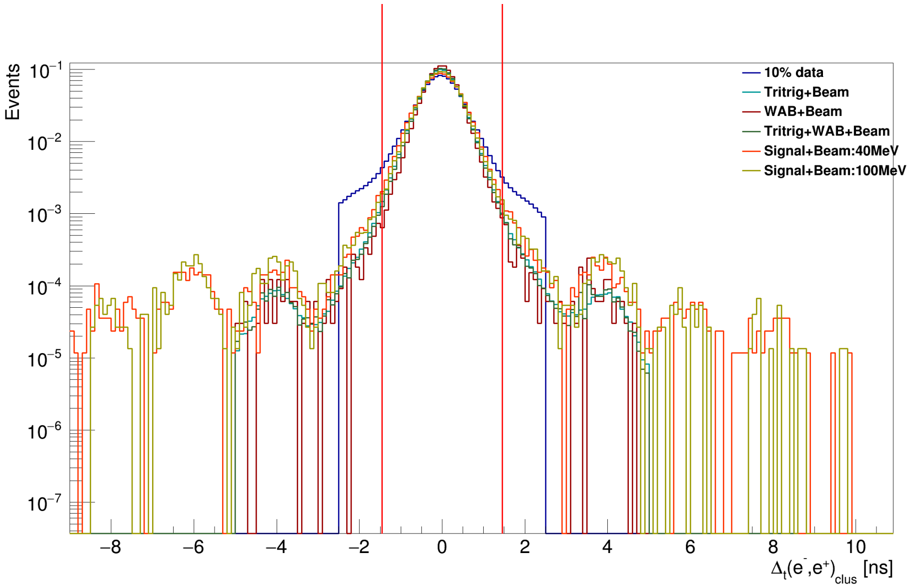
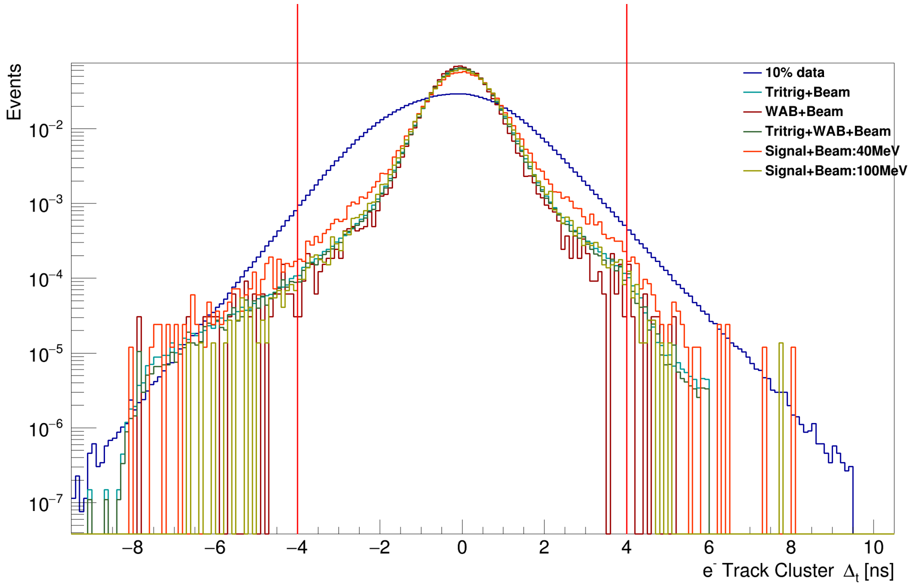
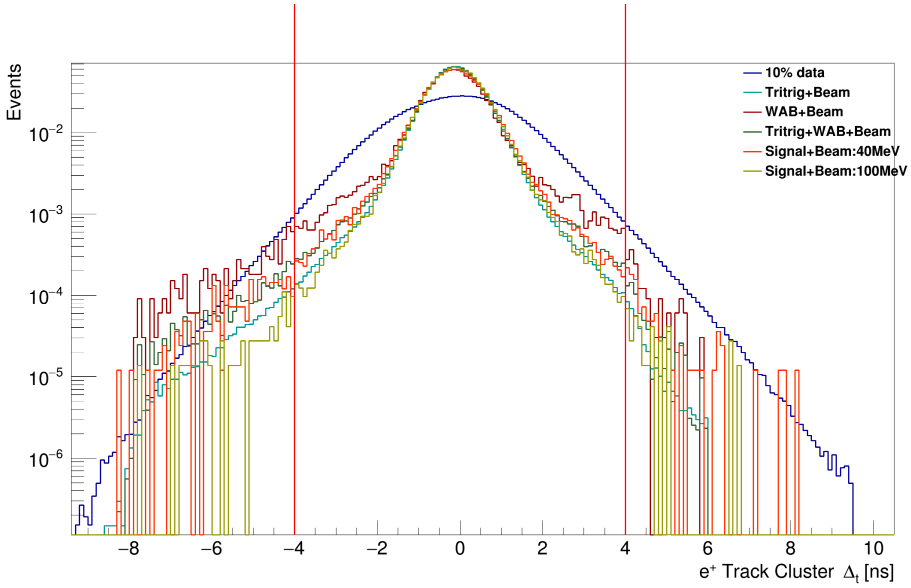


Figure 5.2: Preselection electron and Positron Cluster time difference. A cut is placed at $< 1.45\,\mathrm{ns}$ to avoid using accidental events from other beam bunches (which arrive at $2\,\mathrm{ns}$ intervals).

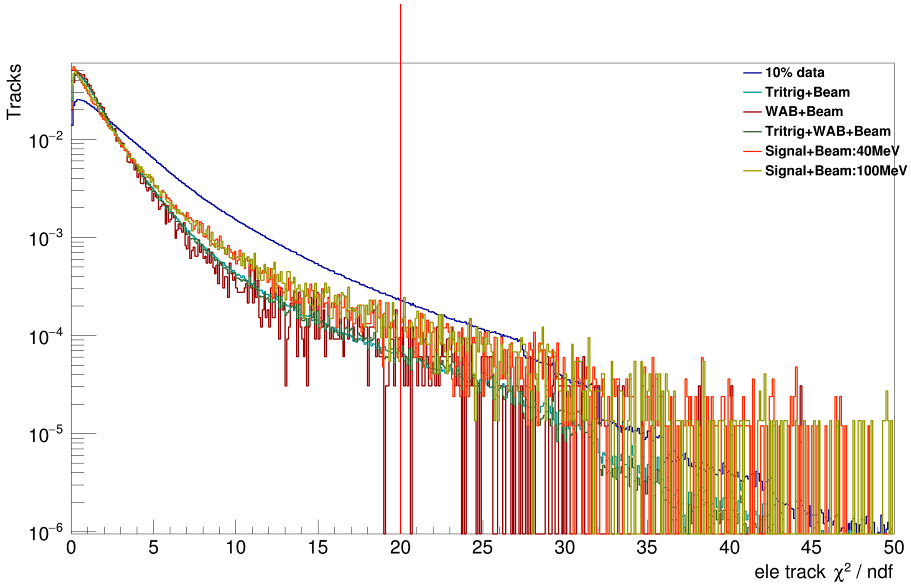


(a) Track-cluster time difference for electrons.

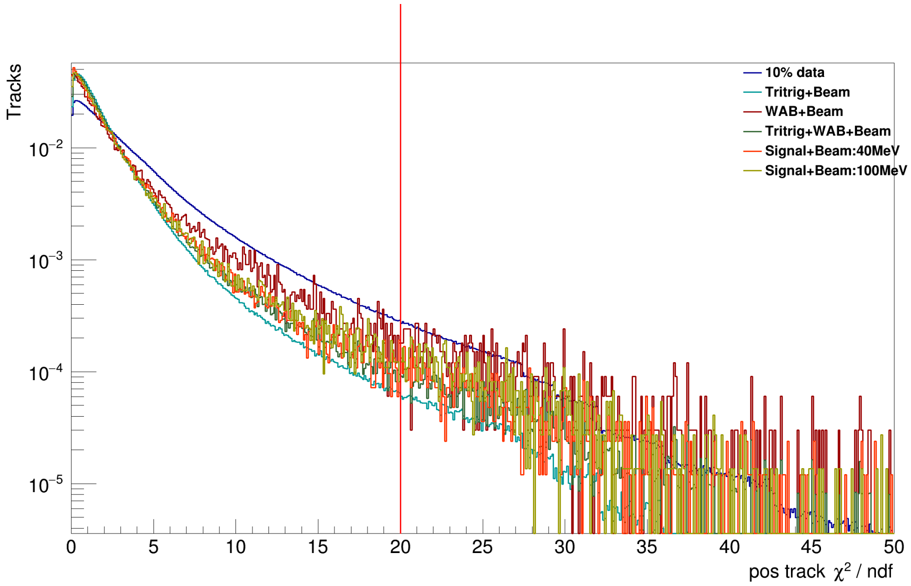


(b) Track-cluster time difference for positrons.

Figure 5.3: Preselection track-cluster time difference for electrons and positrons. Cluster times are shifted by 56 ns in Data, and 43 ns in MC. Track times are shifted by values in Figure 5.1.



(a) Track χ^2/ndf for electrons.



(b) Track χ^2/ndf for positrons.

Figure 5.4: Preselection reconstructed track χ^2/ndf (number of degrees of freedom) for electrons and positrons. The track χ^2 is worse in data than in MC, so the cut is kept very loose at less than 20.

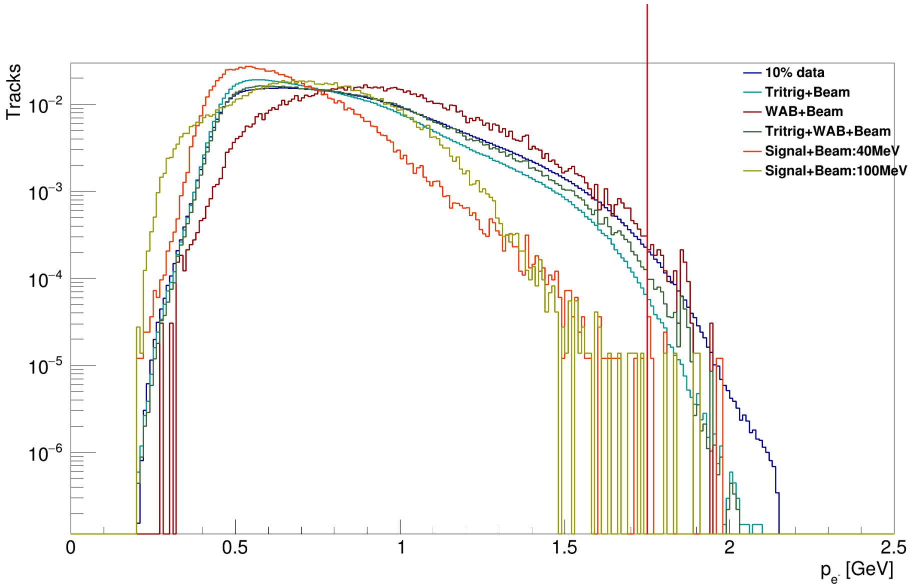
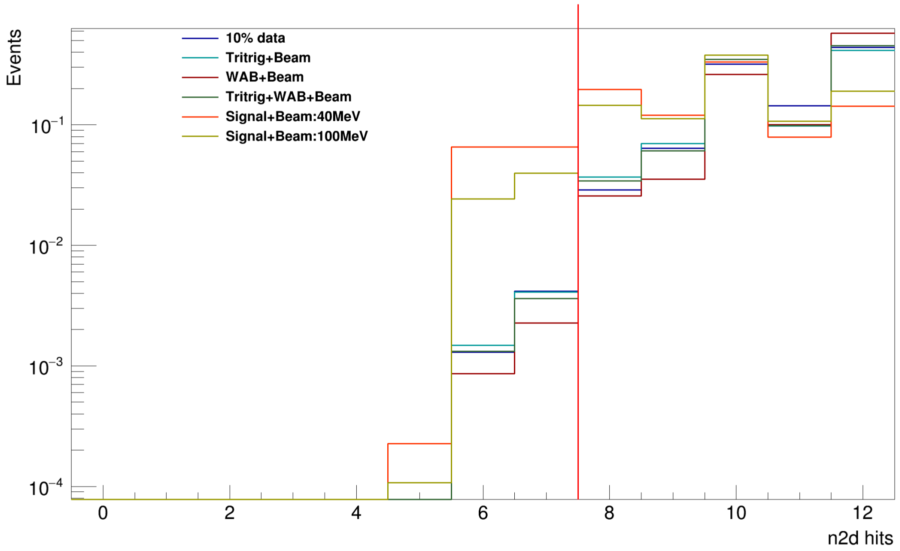
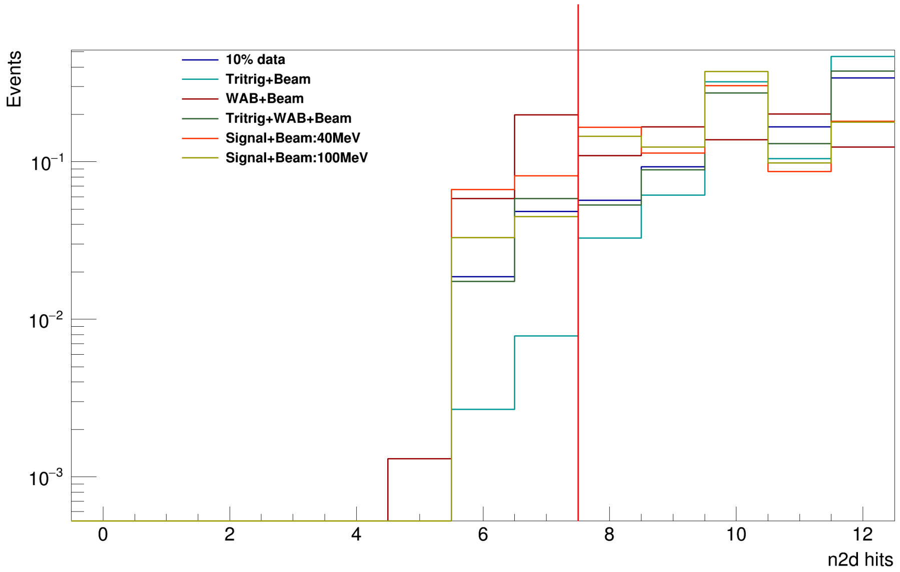


Figure 5.5: Preselection electron Track momentum. A maximum electron momentum cut of 1.75 GeV eliminates vertices that are reconstructed using elastically scattered electrons from the target.



(a) Number of 2D hits on track for electrons.



(b) Number of 2D hits on track for positrons.

Figure 5.6: Preselection number of 2D hits on track for electrons and positrons. Both tracks have a loose requirement of 8/12 2D hits on track.

The last of the preselection cuts are applied to the reconstructed vertex objects. First, a loose cut requires $\mathrm{Vtx}_{\chi^2} < 20$ to eliminate poorly fitted vertices. Figure 5.7 presents these distribution for MC and data, and demonstrates how this cut in particular is capable of eliminating a significant fraction of the reconstructed WABs conversions, according to the MC samples. The final vertex preselection cut removes obviously mis-reconstructed events where the sum of the track momenta (Psum), which is a proxy for the magnitude of the vertex momentum, is greater than 2.4 GeV. The impact of this cut is shown in Figure 5.8, reducing the high-momentum tails in both MC and data.

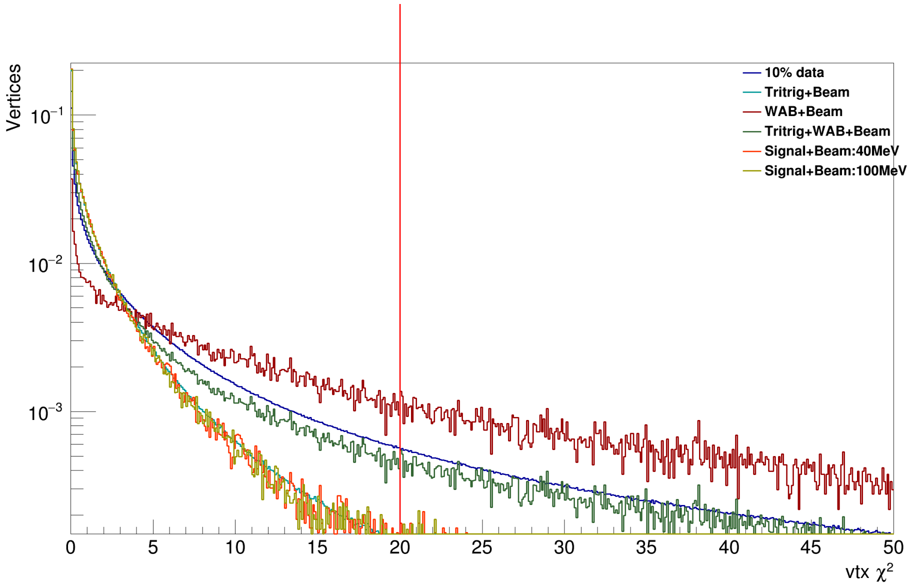


Figure 5.7: Preselection reconstructed Vertex Chi-square. A loose cut on unconstrained vertex fit chi-square is placed at 20 to eliminate poorly reconstructed vertices.

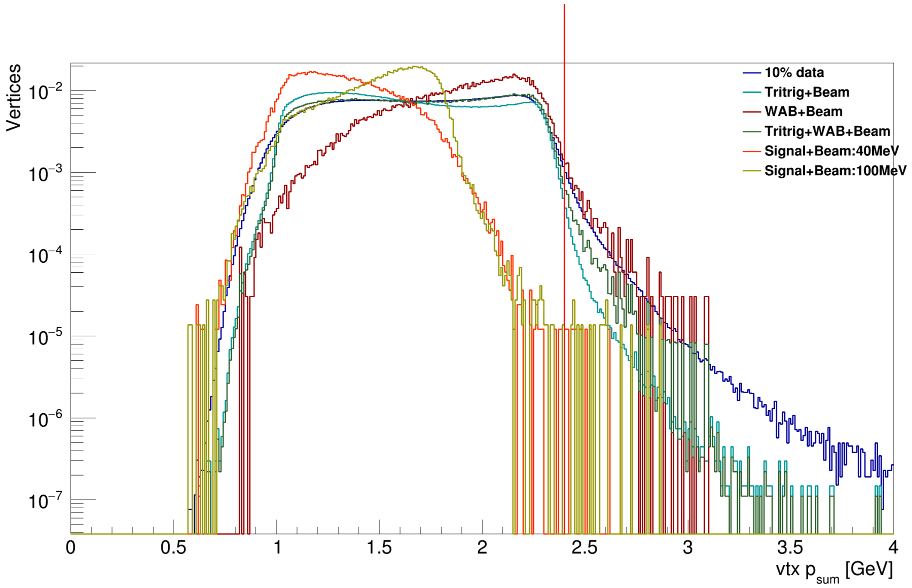


Figure 5.8: Preselection vertex track momenta sum. A maximum cut of 2.4 GeV eliminates obviously mis-reconstructed events with momentum above the full beam energy.

The relative efficiency of each individual cut is summarized in Table 5.3, and the invariant mass and reconstructed vertex z distributions for data and MC are shown in Appendix A.

Preselection N-1 Cutflow Efficiency

	Data Eff	Tritrig-Beam Eff	WAB-Beam Eff	Tritrig-WAB-Beam Eff	40 MeV Signal Eff	100 MeV Signal Eff
$\lvert e^{-}\mathrm{Track}_t\rvert < 6.0\,\mathrm{ns}$	1	1	1	1	1	1
$\lvert e^{+}\mathrm{Track}_t\rvert < 6.0\,\mathrm{ns}$	1	1	1	1	1	1
$\Delta_t(\mathrm{Cluster}_{e^-}, \mathrm{Cluster}_{e^+}) < 1.45\,\mathrm{ns}$	0.96	0.99	0.99	0.99	0.98	0.98
$e^{-}\Delta_t(\mathrm{Track,Cluster}) < 4.0\,\mathrm{ns}$	0.99	1	1	1	1	1
$e^{+}\Delta_t(\mathrm{Track,Cluster}) < 4.0\,\mathrm{ns}$	0.99	1	0.99	1	1	1
$e^{-}\mathrm{Track}\chi^2/\mathrm{n.d.f.} < 20.0$	0.99	1	1	1	0.99	0.99
$e^{+}\mathrm{Track}\chi^2/\mathrm{n.d.f.} < 20.0$	0.98	1	0.98	0.99	0.99	0.99
$p_{e^-} < 1.75\,\mathrm{GeV}$	1	1	1	1	1	1
$\mathrm{N}_{\mathrm{2d\ hits}} e^{-}_{\mathrm{Track}} > 7.0$	1	1	1	1	0.93	0.98
$\mathrm{N}_{\mathrm{2d\ hits}} e^{+}_{\mathrm{Track}} > 7.0$	0.98	1	0.94	0.98	0.93	0.97
$\mathrm{vtx}_{\chi^2} < 20.0$	0.83	0.97	0.65	0.86	0.97	0.97
$p_{e^- + e^+} < 2.4\,\mathrm{GeV}$	0.99	1	0.99	1	1	1

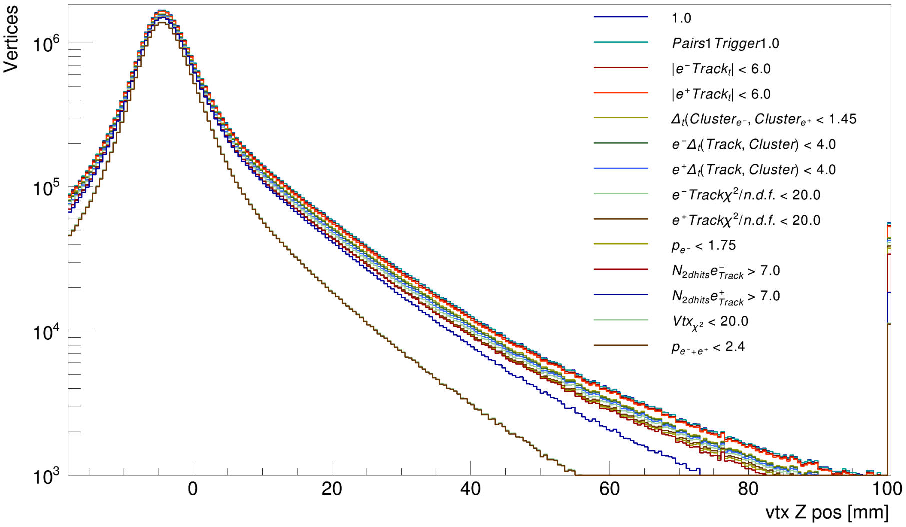
Table 5.3: “n-1” cut efficiency. The efficiency of the cut under consideration is calculated assuming that all other cuts applied correspond to an efficiency of 1.

The ordered preselection cutflow efficiency is summarized in Table 5.4, and the reconstructed vertex z preselection cutflow for data, MC background, and 100 MeV MC signal is shown in Figure 5.9. The total preselection signal efficiencies are relatively high, with 76 % for the 40 MeV MC signal and 83 % for the 100 MeV MC signal. In comparison, the data and MC background efficiencies are 68 % and 79 % respectively. The 11 % discrepancy between MC background and data background is largely attributed to the MC track χ^2 values being systematically lower ("too good") compared to data. This difference suggests that the MC simulation tracking efficiency is artificially high compared to data. Among the preselection cuts, the requirement of a minimum of 8 hits on track has the most significant impact on the signal efficiency. However, lowering the requirement could allow a higher rate of low-quality tracks and increase the occurrence of the falsely displaced vertices. It can also degrade the mass and vertex resolution. For these reasons, this requirement is maintained. The complete set of preselection cutflow plots for the invariant mass, vertex z, and Psum distributions can be found in Appendix A.

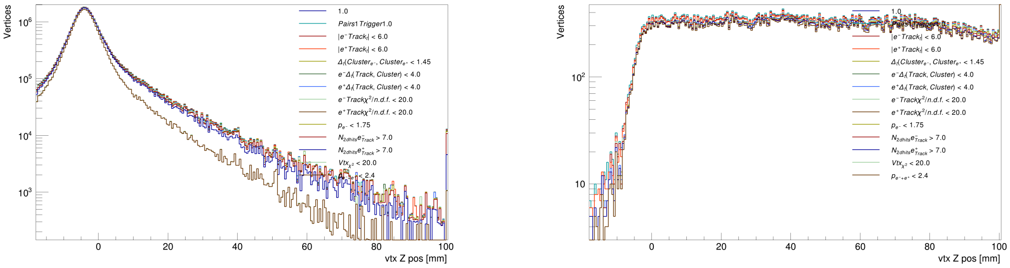
Preselection Cutflow Efficiency

	Data	Eff	Tritrig-Beam	Eff	WAB-Beam	Eff	Tritrig-WAB-Beam	Eff	40 MeV Signal	Eff	100 MeV Signal	Eff
1.0	3.8328e+07	1	7.31364e+06	1	59196	1	3.64637e+07	1	108200	1	87633	1
Pairs1 Trigger1.0	3.82279e+07	1	7.31364e+06	1	59196	1	3.64637e+07	1	108200	1	87633	1
$\|e^-\text{Track}_t\| < 6.0\,\text{ns}$	3.76094e+07	0.98	7.24228e+06	0.99	58656	0.99	3.61163e+07	0.99	105878	0.98	85862	0.98
$\|e^+\text{Track}_t\| < 6.0\,\text{ns}$	3.71464e+07	0.97	7.22328e+06	0.99	58079	0.98	3.59273e+07	0.99	104908	0.97	85004	0.97
$\Delta_t(\text{Cluster}_{e^-}, \text{Cluster}_{e^+}) < 1.45\,\text{ns}$	3.53385e+07	0.92	7.13085e+06	0.98	57439	0.97	3.54905e+07	0.97	101999	0.94	82205	0.94
$e^-\Delta_t(\text{Track,Cluster}) < 4.0\,\text{ns}$	3.49091e+07	0.91	7.10968e+06	0.97	57347	0.97	3.54027e+07	0.97	101442	0.94	82043	0.94
$e^+\Delta_t(\text{Track,Cluster}) < 4.0\,\text{ns}$	3.42158e+07	0.89	7.08966e+06	0.97	56420	0.95	3.51324e+07	0.96	100844	0.93	81853	0.93
$e^-\text{Track}\chi^2/\text{n.d.f.} < 20.0$	3.36416e+07	0.88	7.04596e+06	0.96	56102	0.95	3.49225e+07	0.96	99417	0.92	80639	0.92
$e^+\text{Track}\chi^2/\text{n.d.f.} < 20.0$	3.26225e+07	0.85	6.99543e+06	0.96	54423	0.92	3.43877e+07	0.94	98087	0.91	79315	0.91
$p_{e^-} < 1.75\,\text{GeV}$	3.24534e+07	0.85	6.9883e+06	0.96	54254	0.92	3.43273e+07	0.94	97986	0.91	79244	0.9
$\text{N}_{\text{2d hits}} e^-_{\text{Track}} > 7.0$	3.24111e+07	0.85	6.97724e+06	0.95	54212	0.92	3.42828e+07	0.94	91414	0.84	77312	0.88
$\text{N}_{\text{2d hits}} e^+_{\text{Track}} > 7.0$	3.17285e+07	0.83	6.95771e+06	0.95	51328	0.87	3.35782e+07	0.92	85459	0.79	74772	0.85
$\text{vtx}_{\chi^2} < 20.0$	2.62837e+07	0.69	6.75204e+06	0.92	32982	0.56	2.88368e+07	0.79	82675	0.76	72744	0.83
$p_{e^-+e^+} < 2.4\,\text{GeV}$	2.60451e+07	0.68	6.73881e+06	0.92	32579	0.55	2.87049e+07	0.79	82664	0.76	72728	0.83

Table 5.4: The preselection cutflow efficiency after each cut is applied in order.



(a) Unblinded ∼ 10 % data sample.



(b) Tritrig+WAB+Beam (background) MC scaled to ∼ 10 % luminosity.

(c) 100 MeV Signal MC, unscaled.

Figure 5.9: Vertex z position preselection cutflow for different samples: (a) Data, (b) Tritrig+WAB+Beam MC, and (c) 100 MeV Signal MC.

Figure 5.10 shows a comparison between the MC background and 10 % data invariant mass distributions with preselection applied. The MC backgrounds consisting of tridents and WABs appear consistent with the data for all masses, indicating that those processes indeed are the primary backgrounds in the data.

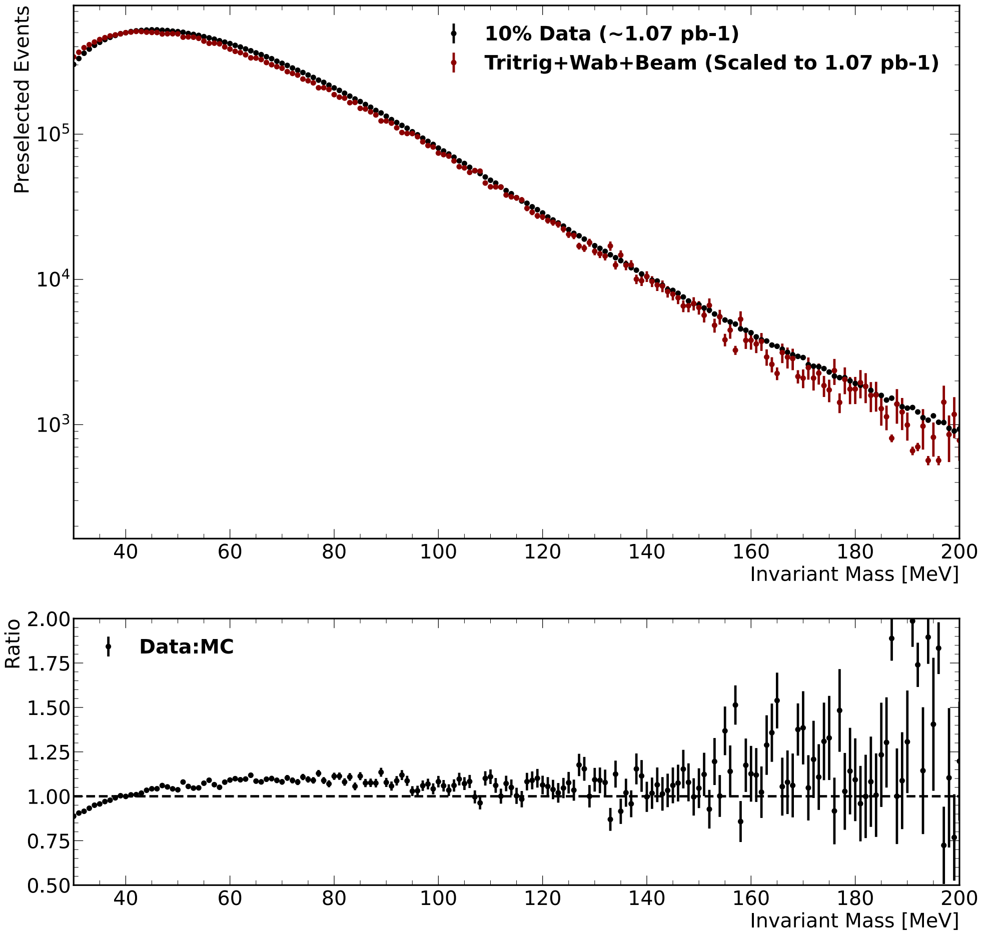


Figure 5.10: Preselection invariant mass. 10 % Data is shown in black. MC background is shown in red. The flat ratio of data to MC background indicates that the MC background processes accurately represent the selected processes in data.

5.1.3 Signal and Control Momentum Regions

As mentioned earlier, the invisible π_D produced in $A' \rightarrow V_D\pi_D$ carries away a significant fraction of the A' energy and leaves the V_D energy and momentum peaked at considerably less than $E_{\text{beam}} = 2.3\,\text{GeV}$. Figure 5.11 shows the vertex Psum distribution, defined as:

$$\text{Psum} = |P^{e^-}_{\text{Track}}| + |P^{e^+}_{\text{Track}}| \tag{5.1}$$

The MC background, shown in green, is relatively flat with a radiative peak at around 2.3 GeV. In contrast, the signal Psum is peaked at much less than 2.0 GeV for all masses. These kinematic differences between signal and background can be used to define a signal region (SR) at "low Psum" and a control region (CR) at "high Psum." The SR significantly reduces the background rate by eliminating the radiative peak while maintaining a high signal efficiency. In addition, the signal contamination is very low in CR. This is critical to the analysis because the normalization of the expected signal rate is ultimately calculated using the reconstructed background rate in the CR. The expected signal calculation is defined later in this section.

The signal and control region Psum boundaries were selected by studying the signal efficiency in both regions. Section 5.1.3 shows the MC signal efficiency (relative to preselection) in both regions over a range of masses and region boundaries. The minimum Psum of the SR is limited to Psum > 1.0 because the MC background is not simulated below this threshold. This is because the MC background was originally generated for the minimal A' search, which was only concerned with the high Psum region of the data, so a generator level cut was used to eliminate events with Psum < 1.0. Therefore, the SR

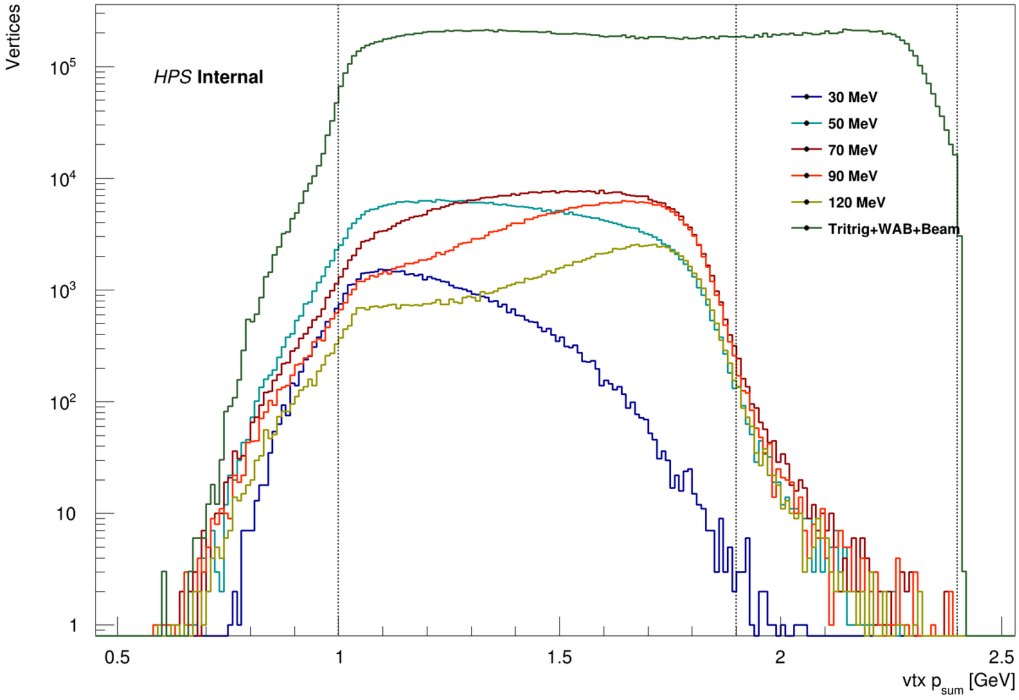


Figure 5.11: Preselection MC reconstructed vertex momentum, not scaled to luminosity. The green plot labeled Tritrig+WAB+Beam shows the MC background. All other plots show different MC signal masses. The MC signal vertex momentum for all masses is peaked at less than 2.0 GeV. The dashed vertical lines define the signal region in vertex momentum.

is restricted to the phase space where the MC background can be validated using data. The signal efficiency is low below Psum < 1.0, so the analysis result is not significantly affected. The maximum Psum of the SR is chosen to be 1.9 GeV, corresponding to over 90 % efficiency for most masses. The SR is therefore defined as:

$$SR = 1.0\,\mathrm{GeV} < \mathrm{Psum} < 1.9\,\mathrm{GeV}. \tag{5.2}$$

The CR is critical to the analysis because it is used to both calculate the expected signal rate, and to normalize the signal rate to the data. For these reasons, it is crucial that the signal efficiency in the CR is minimized. Figure 5.12(b) shows the signal efficiency leaking into the CR based on the minimum Psum value. Ideally the signal efficiency in the CR is zero. However, the expected signal rate is many orders of magnitude less than the reconstructed background rate, and a low signal contamination rate is acceptable. In addition, as described later, the rate of selected radiative tridents in the CR is used to calculate the expected signal rate, so maintaining a high radiative peak acceptance is crucial. Under these considerations, the CR is defined as:

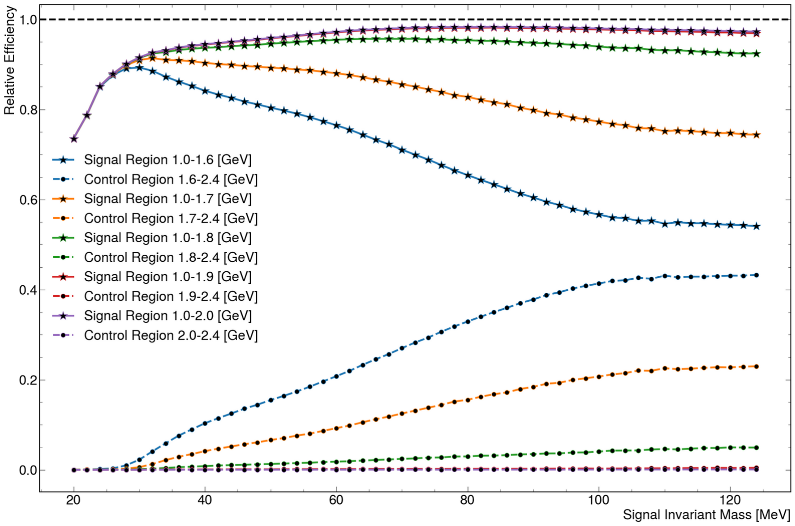
$$CR = 1.9\,\mathrm{GeV} < \mathrm{Psum} < 2.4\,\mathrm{GeV}, \tag{5.3}$$

where the upper limit on Psum was established in the preselection cuts.

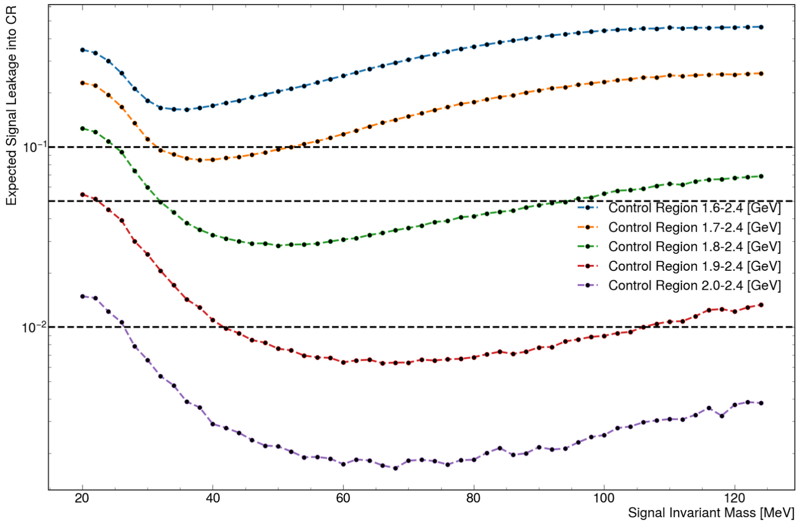
Discrepancies between the Psum distributions in data and MC background, particularly at lower values, are evident in Figure 5.13, even with the preselection and L1L1 requirements applied. This issue likely results from an overestimation of the trigger efficiency in the MC simulation, leading to an inflated signal efficiency at low Psum. To correct this, the ratio of the Psum distributions between the 10 % data sample and MC background is calculated and fitted with a second order polynomial:

$$R_{\mathrm{MC}}(\mathrm{Psum}) = 0.094272950 + 0.87334446 \times \mathrm{Psum} - 0.19641796 \times \mathrm{Psum}^2, \tag{5.4}$$

where Psum is in units of GeV. The MC signal events are then reweighted according to this fit, but only when $R_{\mathrm{MC}}(\mathrm{Psum}) < 1.0$, to align the signal efficiency in Psum with the data. Figure 5.14 shows the improved agreement between the reweighted MC background Psum distribution and data within the SR, demonstrating the effectiveness of this correction.



(a) MC signal efficiency in the signal and control regions as a function of the SR Psum upper edge and CR Psum lower edge.



(b) Fraction of MC signal events leaking into the control region as a function of invariant mass.

Figure 5.12: (a) Efficiency of MC signal in Psum signal and control region and (b) Contamination in the Psum control region based on the invariant mass.

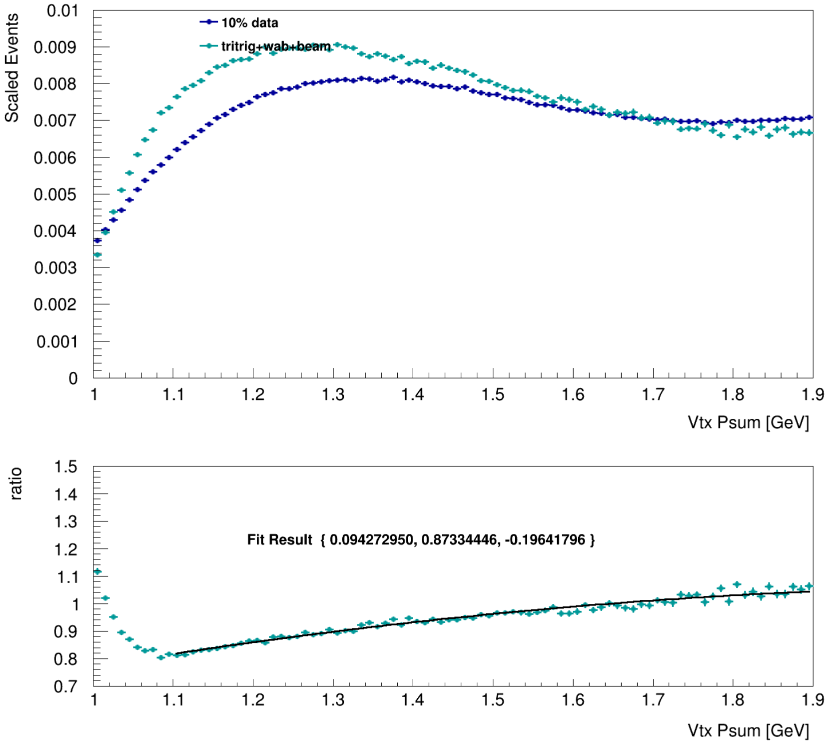


Figure 5.13: Top: 10 % data and MC background Psum distributions. Bottom: Fitted ratio of data to MC background.

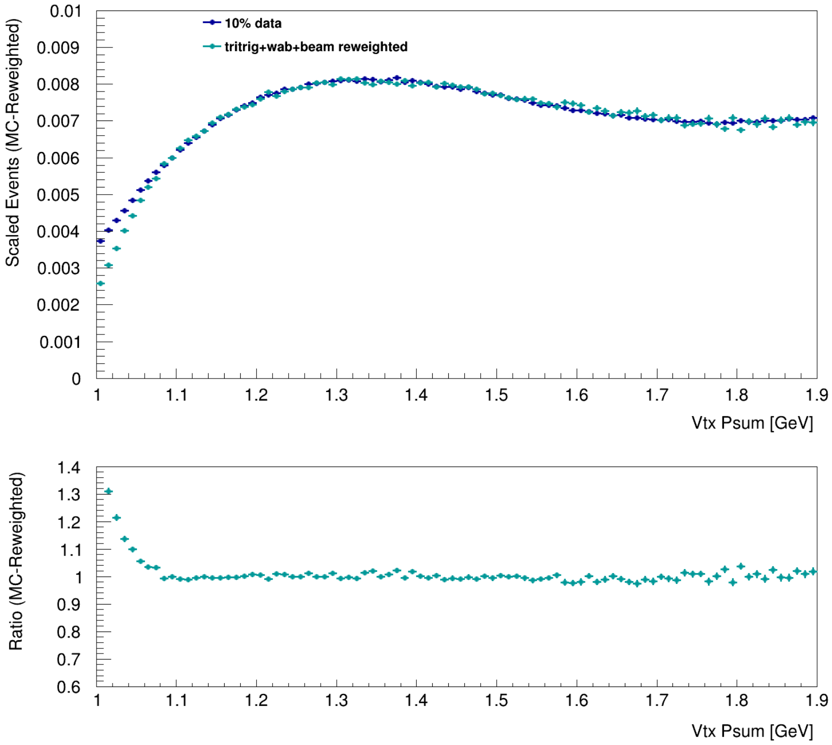


Figure 5.14: Comparison of data and MC background Psum distributions after re-weighting MC background events according to the fit shown in Figure 5.13 and defined in Equation (5.4).

5.1.4 Expected Signal Calculation

The expected signal rate, $N_{\mathrm{sig}}(m_{A'}, \epsilon)$, is calculated as a function of $m_{A'}$ and the kinetic mixing strength ϵ. All of the remaining SIMP parameters ($m_{A'} : m_{V_D}$, $m_{A'} : m_{\pi_D}$, α_D, $m_{\pi_D} : f_{\pi_D}$) are fixed according to Section 2.5.1.

The total A' production rate, $N_{A'}(m_{A'}, \epsilon)$, is first calculated using two terms derived from MC simulations: "radiative fraction" and "radiative acceptance." The A' can then visibly decay to the two different neutral dark vectors, ρ_D and ϕ_D. The expected signal contribution from each dark vector is calculated by applying the relevant branching fraction to the A' production rate, and then measuring the acceptance times efficiency using V_D MC simulations, where the acceptance times efficiency for each mass is a function of ϵ. The total expected signal rate is the summed expected signal rate for ρ_D and ϕ_D.

5.1.4.1 A' Production Rate

The A' production process is analogous to "dark bremsstrahlung" so the cross-section for producing heavy photons of mass $m_{A'}$ is related to the radiative trident production cross-section by:

$$\sigma_{A'} = \frac{3\pi m_{A'}\epsilon^2}{2N_{\mathrm{eff}=1}\alpha}\frac{\mathrm{d}\sigma_{\gamma^*}}{\mathrm{d}m_{l^+l^-}}\Big|_{m_{l^+l^-}=m_{A'}}, \tag{5.5}$$

where the differential cross-section is evaluated at the mass $m_{A'}$ [30]. By multiplying both sides of Equation (5.5) by the integrated luminosity, the A' production rate is proportional to the differential radiative trident rate:

$$N_{A'}(m_{A'}, \epsilon) = \frac{3\pi m_{A'}\epsilon^2}{2N_{\mathrm{eff}=1}\alpha}\frac{\mathrm{d}N_{\gamma^*}}{\mathrm{d}m_{A'}}. \tag{5.6}$$

The differential radiative trident rate in Equation (5.6) can be decomposed into three components:

$$\frac{\mathrm{d}N_{\gamma^*}}{\mathrm{d}m_{A'}} = \left(\frac{\mathrm{d}N_{\gamma^* CR}}{\mathrm{d}m_{A'}} \Big/ \frac{\mathrm{d}N_{CR}}{\mathrm{d}m_{\mathrm{reco}}}\right) \left(\frac{\mathrm{d}N_{\gamma^*}}{\mathrm{d}m_{A'}} \Big/ \frac{\mathrm{d}N_{\gamma^* CR}}{\mathrm{d}m_{A'}}\right) \frac{\mathrm{d}N_{CR}}{\mathrm{d}m_{\mathrm{reco}}}. \tag{5.7}$$

In this decomposition:

- The first term is the inverse of the "radiative acceptance" ($A_{\mathrm{rad}}(m_{A'})$), which is radiative trident acceptance times efficiency in the CR.

- The second term is the "radiative fraction" ($f_{\mathrm{rad}}(m_{A'})$) and represents the expected contribution of radiative tridents to the reconstructed and selected background in the CR.

- The last term, $\frac{\mathrm{d}N_{CR}}{\mathrm{d}m_{\mathrm{reco}}}$, is the reconstructed and selected background rate in the CR. This term scales the A' production rate according to the dataset, whether from MC simulations or data.

This approach allows the A' production rate to be calculated directly with respect to the data, minimizing the systematic uncertainties associated with using MC simulations. Although the radiative acceptance term introduces systematic uncertainties related to the detector acceptance and selection, these contributions largely cancel out when the MC signal acceptance times efficiency is considered.

5.1.4.2 Radiative Fraction

The radiative fraction is given by:

$$f_{\mathrm{rad}}(m_{A'}) = \frac{\mathrm{d}N_{\gamma^* CR}}{\mathrm{d}m_{A'}} / \frac{\mathrm{d}N_{CR}}{\mathrm{d}m_{\mathrm{reco}}}, \tag{5.8}$$

where:

- The numerator is the rate of reconstructed and selected radiative tridents as a function of truth mass, derived from pure radiative trident MC simulations. Each event is truth-matched to the radiative trident mother to exclude events reconstructed with the recoil electron.

- The denominator is the reconstructed and selected background rate as a function of reconstructed mass. The use of reconstructed mass instead of truth mass is necessary because it is not possible to determine whether an event is reconstructed with the truth daughter versus the recoil electron in background processes.

The radiative fraction is shown in Figure 5.15, and parameterized as a function of mass by:

$$\begin{aligned} f_{\text{rad}}(m) = 0.10541434 &- 0.0011737697m \\ &+ 7.4487930 \times 10^{-6}m^2 \\ &- 1.6766332 \times 10^{-8}m^3. \end{aligned} \tag{5.9}$$

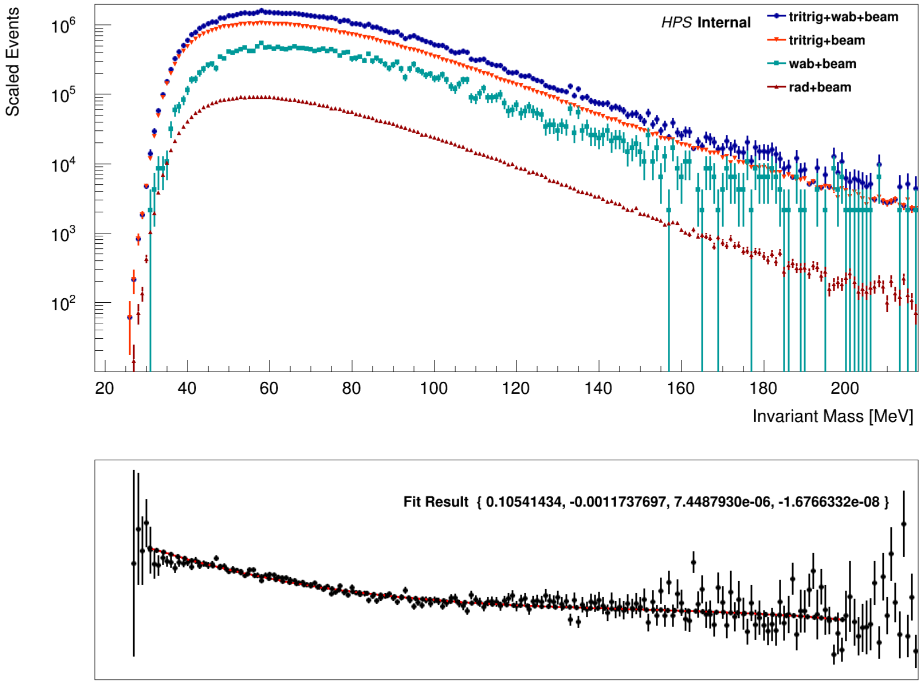


Figure 5.15: The Top plot shows all of the MC background components used to calculate the radiative fraction. The radiative component (rad) is a function of the truth mass for the γ^* event. The WABs and tritrig components, both mixed with MC simulated beam, are a function of reconstructed and selected unconstrained vertex invariant mass. The bottom plot shows the radiative fraction, which is fit by the 3rd-order polynomial shown in red. **fix bottom**

5.1.4.3 Radiative Acceptance

The radiative acceptance is given by:

$$A_{\text{rad}}(m_{A'}) = \frac{\mathrm{d}N_{\gamma^* CR}}{\mathrm{d}m_{A'}} \bigg/ \frac{\mathrm{d}N_{\gamma^*}}{\mathrm{d}m_{A'}}, \tag{5.10}$$

where both the numerator and denominator are evaluated as a function of truth mass $m_{A'}$. The calculation is shown in Figure 5.16, and parameterized in mass by:

$$\begin{aligned} \mathrm{A}_{\text{rad}} = &-0.48922505 + 0.073733061m - 0.0043873158m^2 + 0.00013455495m^3 \\ &- 2.3630535 \times 10^{-6}m^4 + 2.5402516 \times 10^{-8}m^5 \\ &- 1.7090900 \times 10^{-10}m^6 + 7.0355585 \times 10^{-13}m^7 \\ &- 1.6215982 \times 10^{-15}m^8 + 1.6032317 \times 10^{-18}m^9 \end{aligned} \tag{5.11}$$

Using Equation (5.8) and Equation (5.10), the differential radiative trident rate in Equation (5.7) simplifies to:

$$\frac{\mathrm{d}N_{\gamma^*}}{\mathrm{d}m_{A'}} = \frac{f_{\text{rad}}(m_{A'})}{A_{\text{rad}}(m_{A'})} \frac{\mathrm{d}N_{CR}}{\mathrm{d}m_t extreco}, \tag{5.12}$$

Thus, the expected A' production rate in Equation (5.6) is rewritten as:

$$N_{A'}(m_{A'}, \epsilon) = \frac{3\pi m_{A'}\epsilon^2}{2N_{\text{eff}=1}\alpha} \frac{f_{\text{rad}}(m_{A'})}{A_{\text{rad}}(m_{A'})} \frac{\mathrm{d}N_{CR}}{\mathrm{d}m_{\text{reco}}}. \tag{5.13}$$

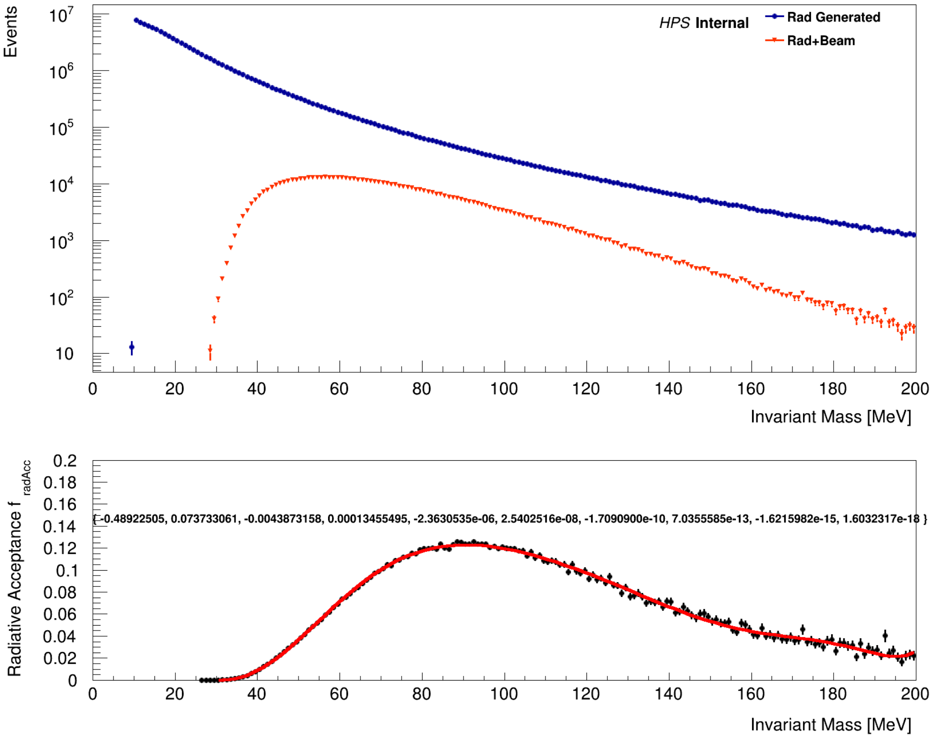


Figure 5.16: The Top plot shows all of the MC background components used to calculate the radiative acceptance. The generated radiative component (Rad Generated) and the reconstructed and selected radiative tridents (Rad+Beam) are both functions of the truth mass for the γ^* event. The bottom plot shows the ratio of Rad+Beam to Rad Generated, which is the radiative acceptance. The radiative acceptance is fit with a 9th-order polynomial shown in red.

5.1.4.4 V_D Acceptance and Efficiency

The number of expected signal events in the data, based on the A' production rate given by Equation (5.13), depends on the acceptance times efficiency of the visible decay $V_D \to e^+e^-$. The acceptance times efficiency is calculated as a function of $m_{A'}$ and ϵ using MC simulated signal over a range of m_{V_D}, where the choice of $m_{A'}$ fixes m_{V_D} according to the SIMP parameters fixed in Section 2.5.1.

The fraction of generated signal events that pass acceptance, reconstruction, and selection as a function of the true vertex z position is defined as

$$F(z) = \frac{R_{\mathrm{MC}}(\mathrm{Psum}) \left(\left. \frac{\mathrm{d}N_{V_D,\mathrm{selected}}}{\mathrm{d}z_{\mathrm{vtx}_{\mathrm{true}}}} \right|_{z_{\mathrm{vtx}}=z} \right)}{\left(\left. \frac{\mathrm{d}N_{V_D,\mathrm{generated}}}{\mathrm{d}z_{\mathrm{vtx}_{\mathrm{true}}}} \right|_{z_{\mathrm{vtx}}=z} \right)}, \tag{5.14}$$

where $R_{\mathrm{MC}}(\mathrm{Psum})$ is the Psum efficiency correction defined in Equation (5.4).

Since the MC signal is generated with a flat lifetime out to 20.0 cm, the signal events used in Equation (5.14) must be reweighted as a function of ϵ according to the expected lifetime of each dark vector meson, ρ_D and ϕ_D. The normalized probability distribution for displaced V_D decays in true vertex z is given by:

$$p(\epsilon, z) = \frac{\exp\left(\frac{z_{\mathrm{target}} - z}{\gamma c \tau}\right)}{\gamma c \tau}, \tag{5.15}$$

where the lifetime τ_{V_D} is a function of ϵ.

The function $F(z)$ in Equation (5.14) is re-weighted according to the probability distribution in Equation (5.15) to obtain the acceptance times efficiency as a function of truth z and ϵ, given by:

$$f_{V_D}(\epsilon, z) = \frac{\exp\left(\frac{z_{\mathrm{target}} - z}{\gamma c \tau_{V_D}}\right)}{\gamma c \tau_{V_D}} F(z). \tag{5.16}$$

The integral of Equation (5.16) gives the total acceptance times efficiency for a V_D with mass m_{V_D} and coupling strength ϵ. Thus, the total expected signal for mass $m_{A'}$ and coupling ϵ is:

$$N_{\text{sig}}(m_{A'}, \epsilon) = N_{A'} \int_{-\infty}^{\infty} \big(\text{BR}(\rho_D) f_{\rho_D}(\epsilon, z) + \text{BR}(\phi_D) f_{\phi_D}(\epsilon, z)\big) \mathrm{d}z, \tag{5.17}$$

where $N_{A'}$ is calculated in Equation (5.13), and the branching ratios are $\text{BR}(A' \rightarrow V_D \pi_D)$.

5.2 Tight Selection

The “tight” selection is the final selection stage. It splits the analysis into mutually exclusive categories based on the track hit content and applies a series of “high-z” cuts that are developed to eliminate the falsely displaced background in that category. The analysis depends on the track hit content due to the correlation between hits on track and the mass and vertex resolution. In addition, the nature of the falsely displaced background can depend on which layers the tracks leave hits on.

5.2.1 Layer 1 Hit Category

The first category, referred to as “L1L1,” consists of vertex candidates where both tracks leave hits in the first two tracking layers (L1 and L2). These events have the best vertex resolution, though the acceptance is restricted to decay lengths smaller than the position of L1, as illustrated in Figure 5.17. The requirement for a hit in L2 is crucial because it improves the tracking algorithm’s ability to accurately extrapolate the track backward toward L1 and correctly identify the L1 hit. The second category, “L1L2,” encompasses events where one track misses L1 due to inefficiencies in the hit detection or reduced acceptance for longer lifetimes. This category introduces more complicated backgrounds, such as an increased rate of unwanted WABs conversions originating in the L1 material. The vertex resolution is also degraded. These effects require a distinct approach to the analysis compared to L1L1. The third analysis category, ‘L2L2,” involves events where both tracks miss L1. This category also suffers from more complicated backgrounds compared to L1L1, in addition to a further reduced vertex resolution. However, this

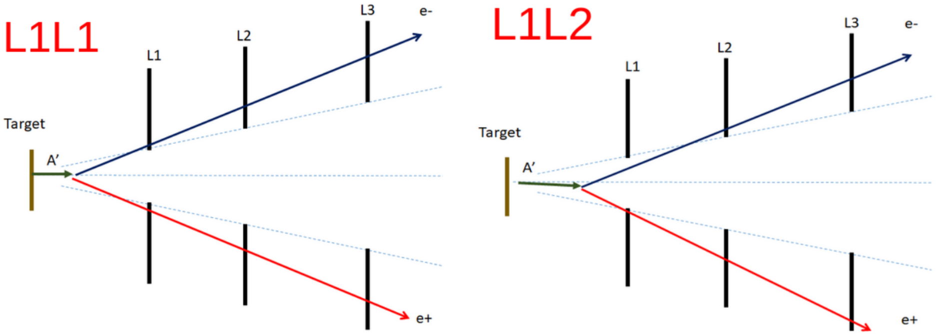


Figure 5.17: Left: short-lived signal is expected to leave hits in both L1 sensors. Right: Longer-lived signal tracks can miss one (or both) of the L1 sensors.

category does have acceptance to longer lifetimes where both tracks decay beyond the L1 position; this is particularly interesting in the context of SIMPs because the signal production rate is decoupled from the lifetime. While the signal production rate still decreases with ϵ, the rate for long-lived signal beyond L1 is potentially large enough to give HPS competitive sensitivity within the existing data sets.

This analysis focuses solely on the L1L1 category and does not address the L1L2 or L2L2 categories, which remain blinded for further studies. Therefore, the first tight selection criteria requires requires that both tracks leave hits in both the axial and stereo L1 and L2 sensors. Although the L1L1 category can be expanded to include events where only the axial sensor in L1 and/or L2 have a hit on track, the gain in acceptance is expected to be marginal was not pursued. After applying the L1L1 hit requirement, each event must additionally have exactly one vertex candidate to avoid duplicates, which are not anticipated for the rare signal events. This criterion results in less than a 1 % loss in efficiency for MC and data.

5.2.2 High-Z Cuts

The high-$z$cuts presented here are specifically designed for the L1L1 hit category but are likely applicable to other categories in a similar manner. These cuts are designed to eliminate falsely displaced background based on simple geometric motivations. The first cut, known as the target projected vertex significance cut ($\mathrm{N}\sigma_{\mathrm{V0_{proj}}}$), assumes that a truly displaced vertex should project back to the beamspot on the target. This cut ensures that only vertices that are consistent with the interaction point are considered. The second cut requires that both tracks for a displaced vertex have symmetric vertical impact parameters ($y0$) that increase linearly with the vertex displacement in $\mathrm{z_{vtx}}$, as illustrated in Figure 5.18. These two criteria alone are effective in reducing the falsely displaced background rate due to track reconstruction errors and large scattering events.

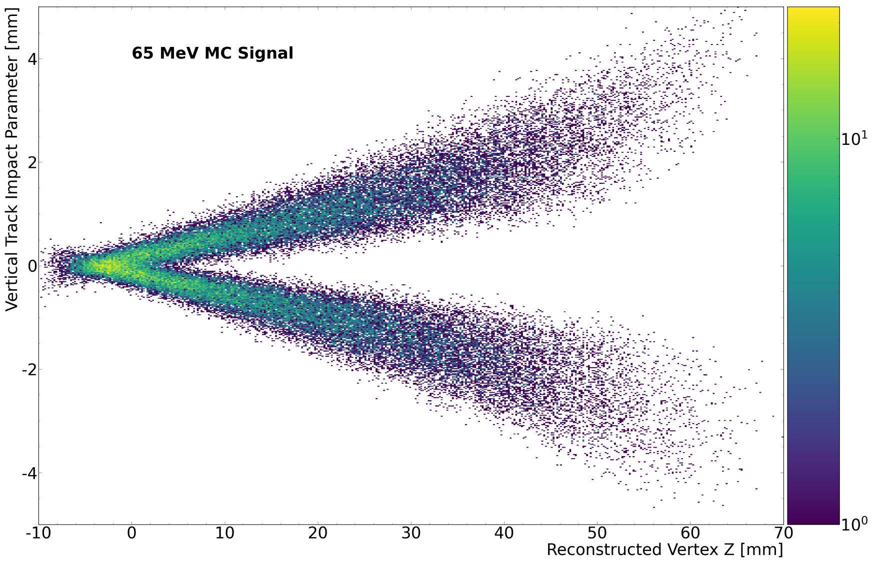


Figure 5.18: Example of (65 MeV) MC signal track vertical impact parameters versus reconstructed vertex z. Each vertex has an electron and positron track entry. Signal shows expected linear correlation between impact parameters and vertex z.

5.2.2.1 $y0$ Versus $\mathbf{z_{cut}}$ Approach

Originally, the analysis aimed to mitigate the falsely displaced background events using these high-$z$cuts while preserving the prompt background component. The signal region, as a function of invariant mass, was then defined by locating a position downstream from the target in $\mathrm{z_{vtx}}$ where the expected background rate was minimal compared to the signal rate. This position, referred to as the “$\mathrm{z_{cut}}$,” was determined by fitting the prompt background high-$z$tail with an exponential decay function, and identifying the $\mathrm{z_{vtx}}$ position where the background fit predicted less than a specified threshold number of events, given by:

$$\int$$

$\mathrm{z}_{\mathrm{cut}}^{\infty} A \exp(-z)\,\mathrm{d}z \leq N_{\mathrm{threshold}}$(5.18) This method is depicted in Figure 5.19, where the background threshold of 0.5 events was selected and the final signal region is defined to the right of the vertical dashed lined. This approach was inspired by a similar strategy used in the HPS 2016 displaced A' search [43]. However, during the optimization of the track vertical impact parameter cut, an improved method was discovered. Replacing the $\mathrm{z_{cut}}$ with a flat cut on $y0$ significantly enhanced the expected signal sensitivity across a range of lifetimes. The flat $y0$ cut requires that both tracks for a vertex candidate have large vertical impact parameters well away from the core of the prompt background. This requirement can effectively eliminate nearly all of the background events, prompt and falsely displaced, while maintaining a high level of signal efficiency nearest to the target.

An example comparing the $y0$ and $\mathrm{z_{cut}}$ approaches is shown in Figure 5.20, where the expected signal and background rates are shown with and without the $\mathrm{z_{cut}}$. The $\mathrm{z_{cut}}$

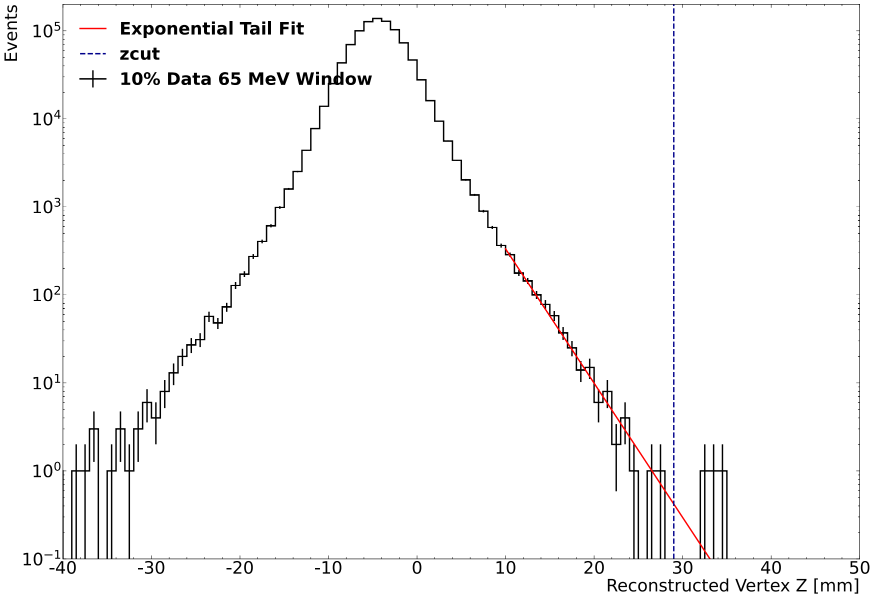


Figure 5.19: Example of the background exponential tail fit in the 65 MeV invariant mass window. The exponential fit is shown by the red line. The zcut location (dashed vertical line) is determined as the fit position beyond which less than 0.5 background events are expected. The signal region is defined as reconstructed vertex z greater than zcut.

approach predicts 4.73 signal events with 3 background events, while $y0$ approach predicts 10.45 signal events with 4 background events, demonstrating an improvement in the signal sensitivity. Although the $\mathrm{z}_{\mathrm{cut}}$ approach is not used in the final analysis, it is mentioned here because it played a role in the optimization process of the high-$z$cuts.

5.2.2.2 High-Z Cut Optimization

The high-$z$cut variables and values were analyzed as a function of invariant mass using the 10 % data sample (considered to be background-only) and the MC-generated signal described in Section 4.4. The binomial significance of the expected signal rate relative to the expected background rate in the invariant mass window, known as ZBi, was chosen

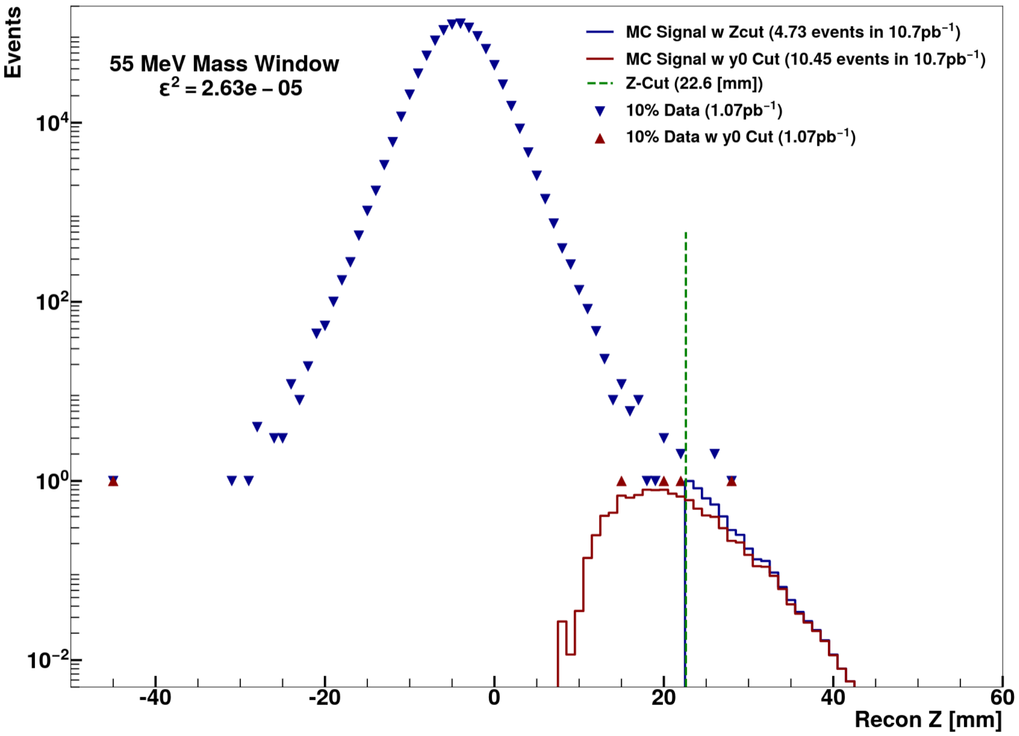


Figure 5.20: Reconstructed vertex z for tight selected events in a search window centered at 55 MeV, with the target located at -4.3 mm. The blue (red) triangles show the remaining events in 10 % data before (after) cutting on the vertical impact parameter $y0$. The expected signal rate for $\epsilon^2 = 2.63 \times 10^{-5}$ is shown for each selection by the color-coordinated solid lines. The 2016 A' analysis technique eliminates the background by placing a cut on recon z, indicated by the vertical dashed line, while the impact parameter selection used in this analysis eliminates the background without the need to cut on recon z.

as figure of merit with the goal of optimizing the cuts by maximizing ZBi. The binomial significance is related to the expected significance of the data if the signal is present, so maximizing this figure is expected to improve the signal sensitivity in the final analysis. The ZBi calculation, which is detailed in Appendix A, is only valid when the expected signal rate is greater than approximately 0.5 events, so the expected signal rate was scaled up during the optimization process when necessary. Additionally, the ZBi calculation requires the background rate to be relatively low, on the order of a hundred events, which is well above the acceptable background rate for this analysis. The optimization for each cut is discussed in the following sections.

5.2.2.3 Target Projected Vertex Significance Cut

The target projected vertex significance cut aims to eliminate poorly reconstructed events that are inconsistent with originating from the interaction point of the beam on the target. Real reconstructed vertices of interest should be consistent with originating from the beamspot on the target, as illustrated in Figure 5.21. This consistency is verified by projecting a vertex candidate back towards the target, upstream along the beam direction z, using the reconstructed vertex momentum. The target-projected vertex ($\mathrm{V0}_{\mathrm{proj}}$) has new coordinates x_{target} (beam left), y_{target} (beam up), and z_{target}. If the $\mathrm{V0}_{\mathrm{proj}}$ coordinates fall outside the beamspot mean position beyond a predefined significance threshold ($\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$), the candidate is rejected.

In order to measure $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ relative to the beamspot, the shape, size, and position of the beamspot on the target must be characterized on a run-by-run basis, as the beam conditions vary across runs. The beamspot is also characterized for the MC samples, though without run-specific variations.

5.2.2.4 Run Dependent beamspot Parameterization

The beamspot position and width in x and y, as well as the correlation between x and y, is estimated using approximately 1 % of the available data files for each run. Unconstrained vertices that pass preselection are projected back to the target using:

$$(x_{\mathrm{target}}, y_{\mathrm{target}}) = \left((x_{\mathrm{recon}} - z_{\mathrm{target}})\frac{p_x}{p_z},\ (y_{\mathrm{recon}} - z_{\mathrm{target}})\frac{p_y}{p_z} \right) \tag{5.19}$$

where $z_{\mathrm{target}} = -4.3\,\mathrm{mm}$, and $p_{x,y,z}$ are the reconstructed vertex momentum components. If MC momentum smearing is implemented, track momentum smearing factors are applied

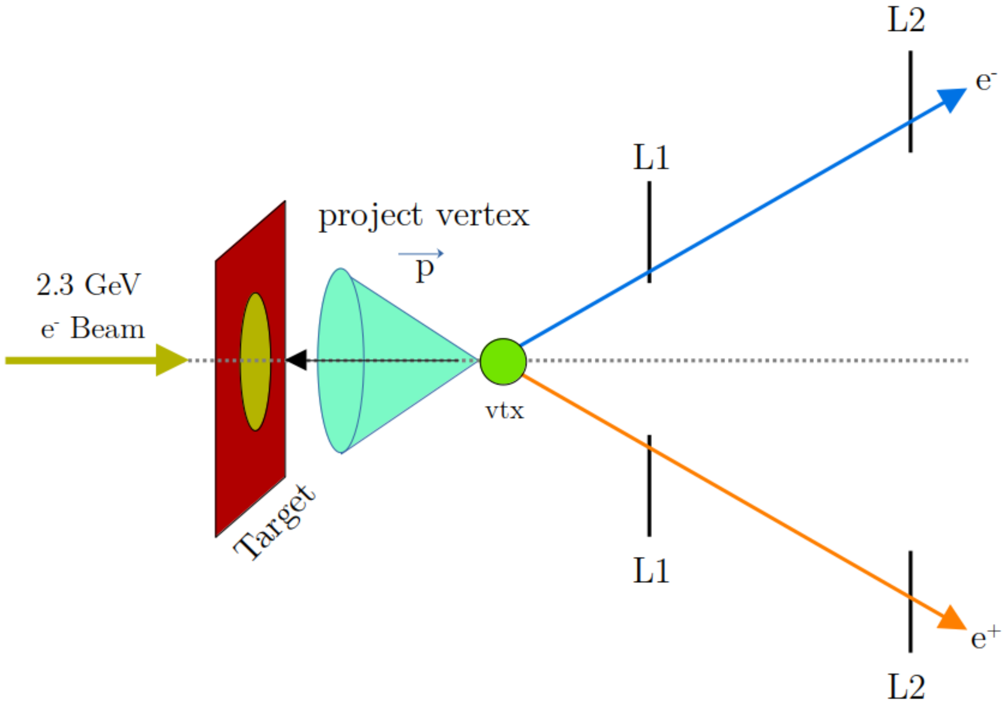


Figure 5.21: Illustration of vertex projection to target location in z. Vertex is projected back to the target using the vertex momentum. If the projected vertex falls outside of the beamspot, the event may be cut.

to approximate the vertex momentum components.

The beamspot is characterized by transverse Gaussian distributions in x and y, and their linear correlation, known as the beam rotation angle θ_{beam}. The beamspot is parameterized by the mean and width of the x and y distributions in the rotated coordinate system $(x_{\text{rot}}, y_{\text{rot}})$, defined along θ_{beam}. The V0$_{\text{proj}}$ coordinates are binned in a 2D histogram for each run, and the histogram is fit with a rotated 2D Gaussian function, given by:

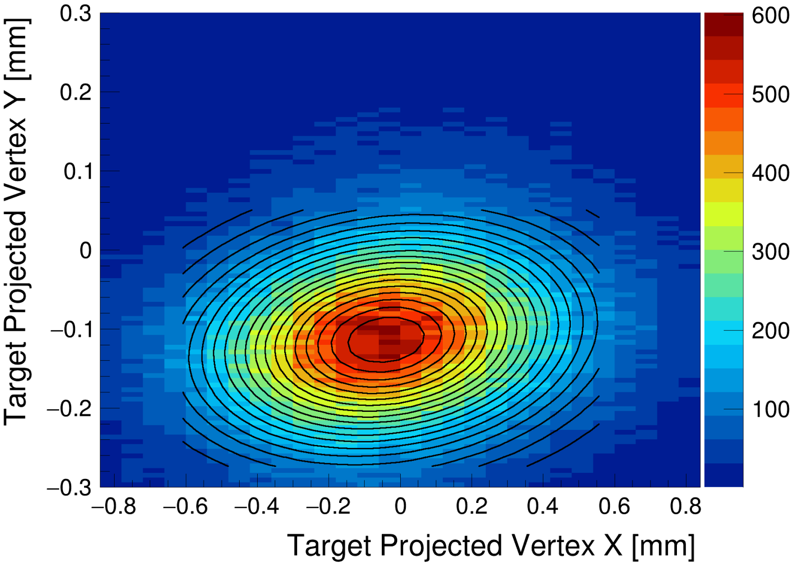
$$f(x,y) = \exp\left(-\frac{(x_{\text{rot}} - \mu_{x_{\text{rot}}})^2}{2\sigma^2_{x_{\text{rot}}}} - \frac{(y_{\text{rot}} - \mu_{y_{\text{rot}}})^2}{2\sigma^2_{y_{\text{rot}}}}\right), \tag{5.20}$$

where

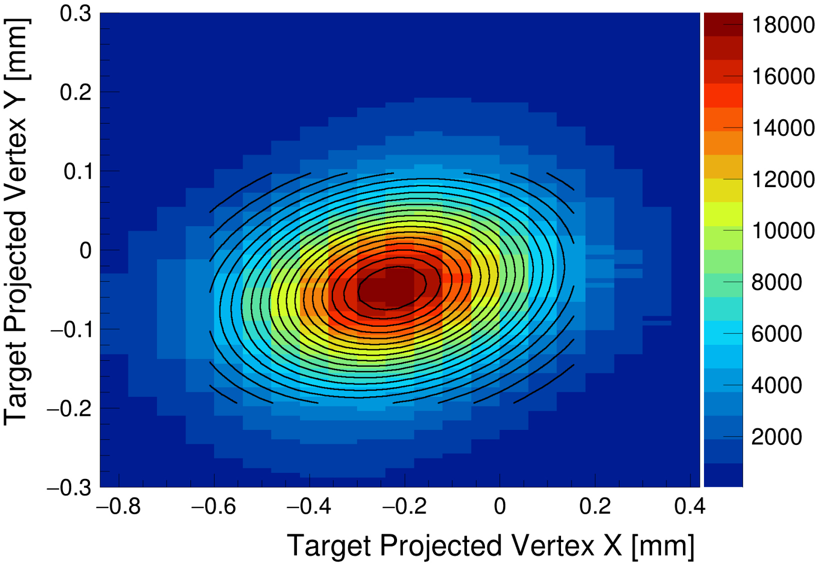
$$x_{\text{rot}} = x\cos(\theta_{\text{beam}}) - y\sin(\theta_{\text{beam}}), \quad y_{\text{rot}} = x\sin(\theta_{\text{beam}}) + y\cos(\theta_{\text{beam}}). \tag{5.21}$$

The fit range in rotated coordinates is defined as $1.5\sigma_{x,y}$ in the non-rotated coordinates, which is a reasonable approximation since the rotation angle is small. Figure 5.22 shows the V0_{proj} coordinates and an example of the rotated 2D Gaussian fit using data and MC background.

The fitted beamspot positions and widths are summarized as a function of run number in un-rotated coordinates in Figure 5.23, and rotated coordinates in Figure 5.24. The beam rotation angle is shown as a function of run number in Figure 5.25.

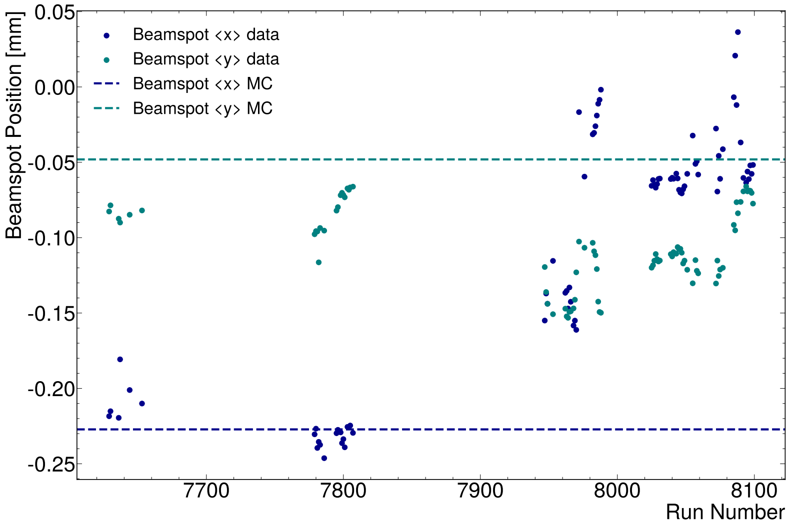


(a) Data Run 7800.

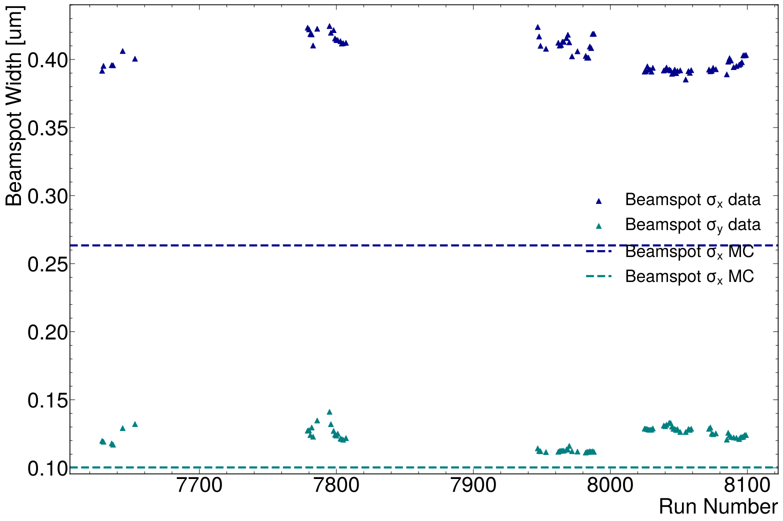


(b) MC background (Tritrig+Beam) Run 7984.

Figure 5.22: Example of unconstrained preselection vertices projected back to the target location in z (-4.3 mm). The x and y distributions are fit with a rotated 2D Gaussian function to estimate the beamspot position, width, and rotation angle.

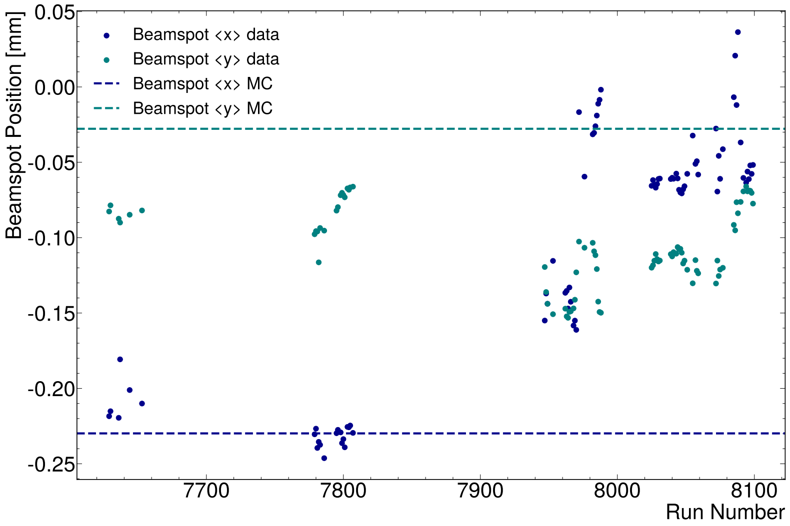


(a) Fitted x and y mean positions.

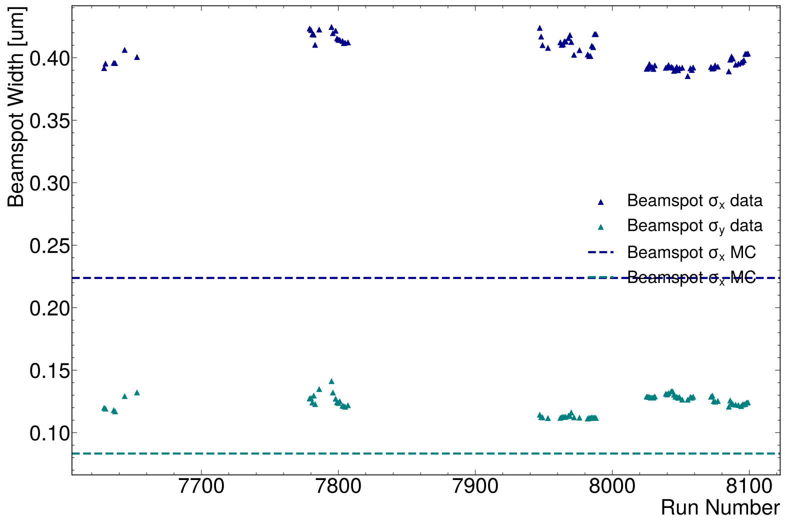


(b) Fitted x and y widths in the θ_{beam}.

Figure 5.23: Run-dependent fitted beamspot characteristics in the un-rotated coordinate system.



(a) Fitted x and y mean positions in the θ_{beam}.



(b) Fitted x and y widths in the θ_{beam}.

Figure 5.24: Run-dependent fitted beamspot characteristics in the rotated coordinate system (according to beam rotation angle).

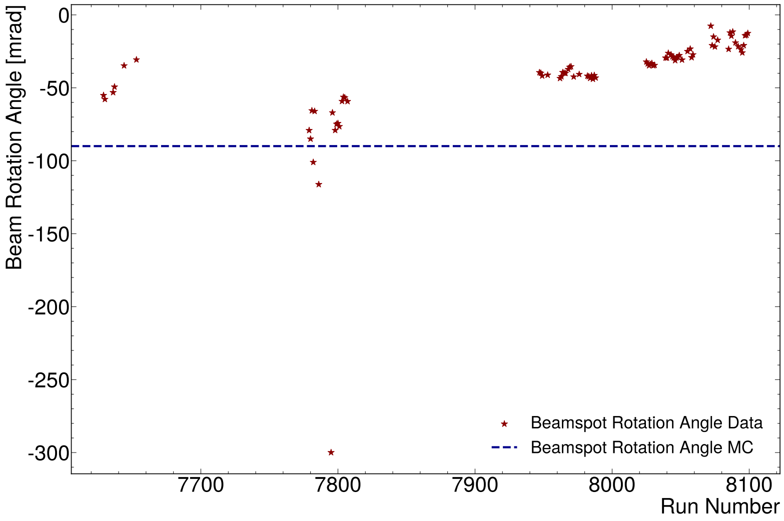


Figure 5.25: The run-dependent beam rotation angle θ_{beam} resulting from the rotated 2D Gaussian fit.

5.2.2.5 $\mathbf{V0_{proj}}$ Significance ($\mathbf{N\sigma_{V0_{proj}}}$)

Mis-reconstructed events, which result in falsely displaced vertices, often have momenta that are inconsistent with the expected displaced vertex position, and therefore tend to project back to a location far outside the known beamspot. The decision to cut an event is based on the distance of the projected vertex $\mathrm{V0_{proj}}$ from the fitted beamspot, measured in terms of standard deviations in both the transverse and longitudinal directions. The $\mathrm{V0_{proj}}$ transverse and longitudinal significance, $N_{\sigma x_{\mathrm{rot}}}$ and $N_{\sigma y_{\mathrm{rot}}}$, is calculated using

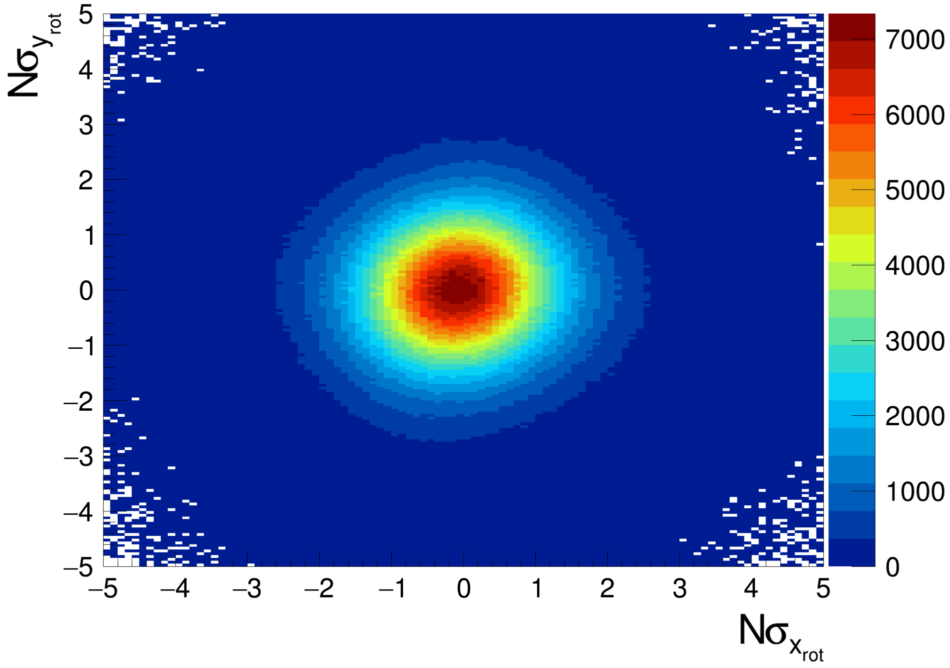
$$(N_{\sigma x_{\mathrm{rot}}}, N_{\sigma y_{\mathrm{rot}}}) = \left(\frac{x_{\mathrm{rot}} - \mu_{x_{\mathrm{rot}}}}{\sigma_{x_{\mathrm{rot}}}}, \frac{y_{\mathrm{rot}} - \mu_{y_{\mathrm{rot}}}}{\sigma_{y_{\mathrm{rot}}}}\right), \tag{5.22}$$

where x and y are in the rotated coordinate system defined by θ_{beam}, and $\mu_{x_{\mathrm{rot}},y_{\mathrm{rot}}}$ and $\sigma_{x_{\mathrm{rot}},y_{\mathrm{rot}}}$ are the run dependent fitted beamspot parameters in the rotated coordinate system. Figure 5.26 shows the $\mathrm{V0_{proj}}$ significance for 10 % data, MC background, and an MC signal sample.

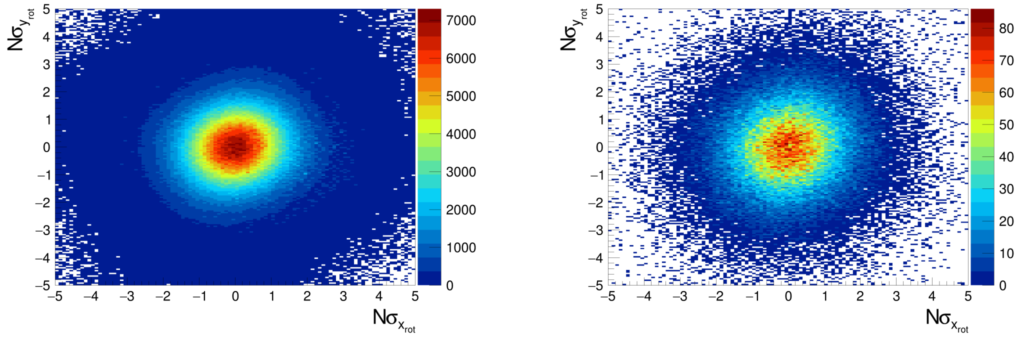
The $\mathrm{V0_{proj}}$ significance in both directions is combined to form a new elliptical cut variable, $\mathrm{N}\sigma_{\mathrm{V0_{proj}}}$, defined as:

$$\mathrm{N}\sigma_{\mathrm{V0_{proj}}} = \sqrt{N^2_{\sigma x_{\mathrm{rot}}} + N^2_{\sigma y_{\mathrm{rot}}}}. \tag{5.23}$$

The correlation between $\mathrm{N}\sigma_{\mathrm{V0_{proj}}}$ and falsely displaced background is demonstrated in Figure 5.27 using L1L1 MC background and signal. The distribution in data, representing background only, shows that falsely displaced events with large values of $\mathrm{z_{vtx}}$ are linearly correlated with $\mathrm{N}\sigma_{\mathrm{V0_{proj}}}$, while the displaced signal remains uncorrelated. A simple cut $\mathrm{N}\sigma_{\mathrm{V0_{proj}}}$, such as requiring $\mathrm{N}\sigma_{\mathrm{V0_{proj}}} < 2.0$, significantly reduces the falsely displaced while maintaining a high relative signal efficiency. Since $\mathrm{N}\sigma_{\mathrm{V0_{proj}}}$ and $\mathrm{z_{vtx}}$ are uncorrelated for



(a) Preselected target projected vertex significance $N_{\sigma x_{\mathrm{rot}}}$ versus $N_{\sigma y_{\mathrm{rot}}}$ for 10% Data.



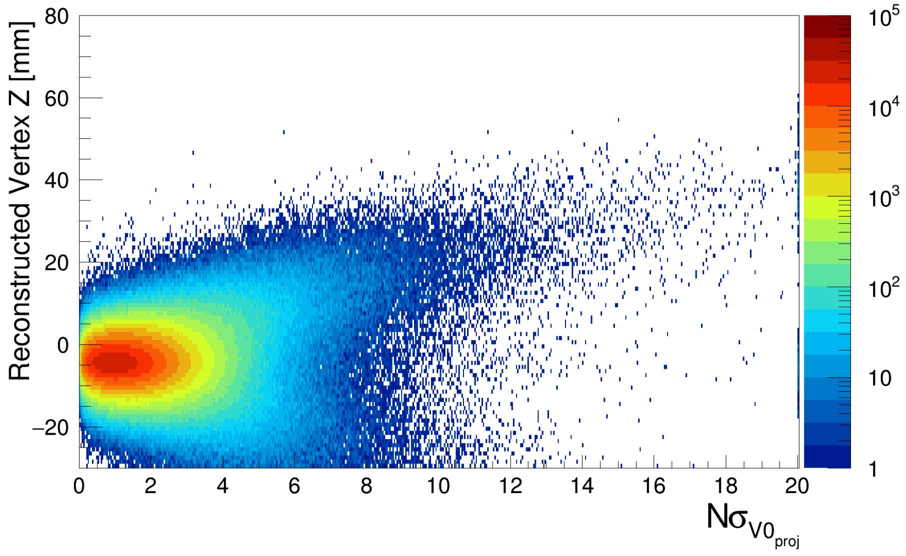
(b) Tritrig+Beam MC.

(c) 55 MeV MC Signal.

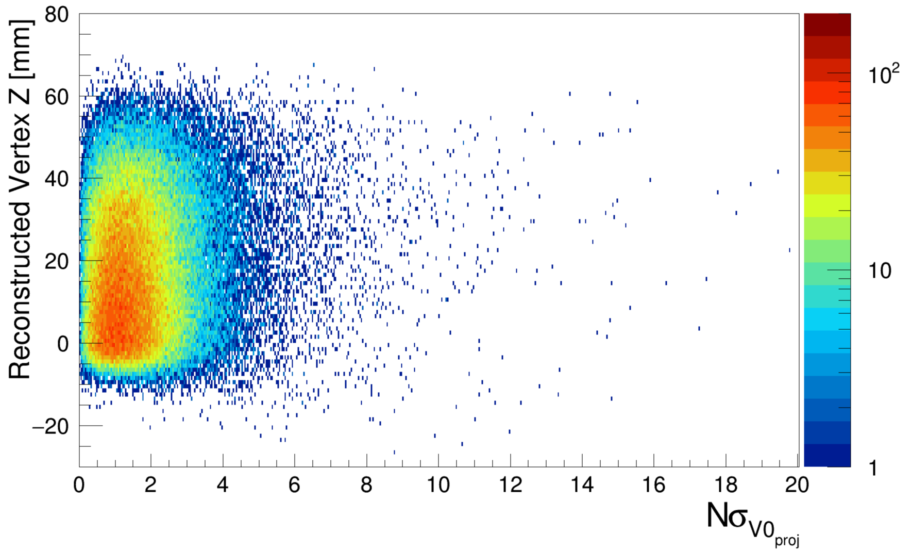
Figure 5.26: Preselection target projected vertex significance $N_{\sigma x_{\mathrm{rot}}}$ versus $N_{\sigma y_{\mathrm{rot}}}$ for different datasets.

the MC signal, the resulting efficiency is approximately constant across different lifetimes and does not need to be optimized as a function of ϵ.

The $\mathrm{N}\sigma_{\mathrm{V0_{proj}}}$ cut was optimized using the zcut-style analysis previously men-



(a) MC background.



(b) MC signal (55 MeV).

Figure 5.27: Histograms of z_{vtx} versus $N\sigma_{V0_{\mathrm{proj}}}$ for MC signal and MC background. The background shows a correlation between high z_{vtx} events and $N\sigma_{V0_{\mathrm{proj}}}$, while the displaced signal is uncorrelated.

tioned. The optimization process was conducted for each MC-generated signal mass point within a search window centered on the signal mass, with a width equal to four times the mass resolution. The expected signal was calculated using Equation (5.17), with ϵ^2 set to 3×10^{-6}, roughly corresponding to the maximum expected signal rate for all masses. For each search window, the $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ cut was initially set to the value that removes 0 % of the signal events, and then iteratively tightened. In each iteration, after applying the $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ cut, the high-$z$background tail was fitted using the exponential function in Section 5.2.2.1, and the expected background and signal rates were determined based on the $\mathrm{z}_{\mathrm{cut}}$ position. The optimal $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ cut values was selected by maximizing ZBi across all iterations and possible $\mathrm{z}_{\mathrm{cut}}$ values.

Figure 5.28 shows the optimal $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ cut values for each signal mass window. Although there is a slight linear relationship between the optimal $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ cut value and mass, the final $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ cut was set uniformly for all masses, requiring $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}} < 2.0$. Figure 5.29 shows the invariant mass distribution versus $\mathrm{z}_{\mathrm{vtx}}$ in 10 % data before and after the cut on $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$. The core of the background $\mathrm{z}_{\mathrm{vtx}}$ distribution before the $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ cut extends out as far as 44 mm downstream of the target, and is limited to around 20 mm after the cut. Additionally, the largest displacement events are reduced from nearly 64 mm downstream to around 20 mm.

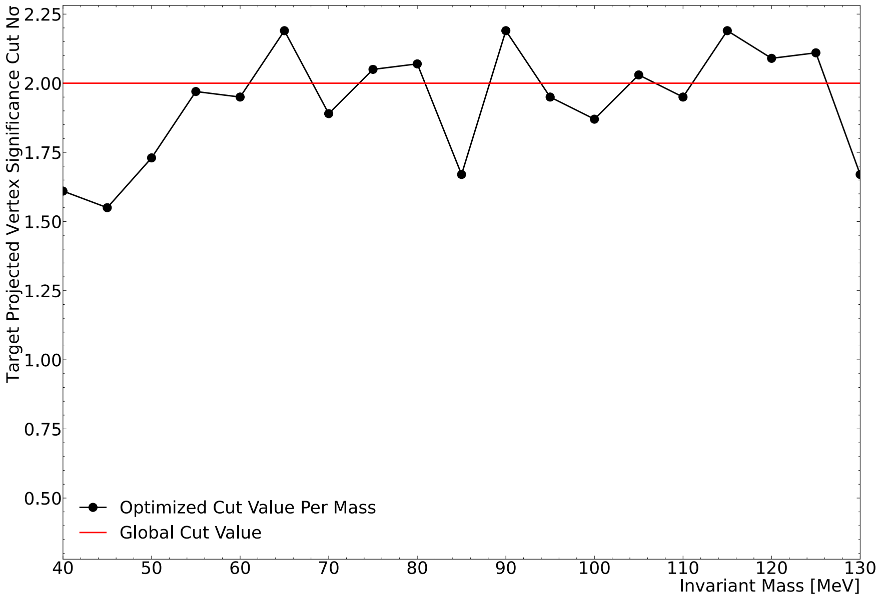
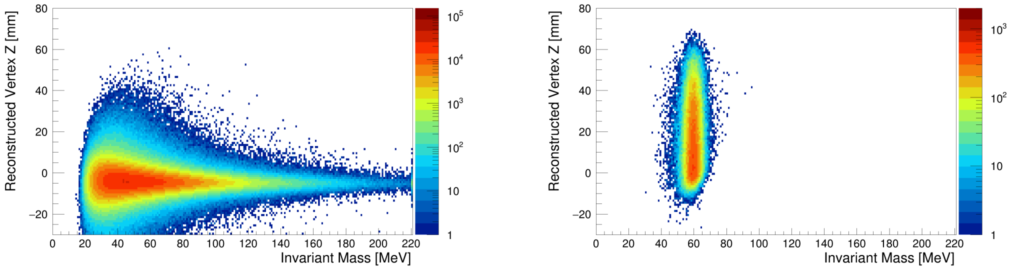
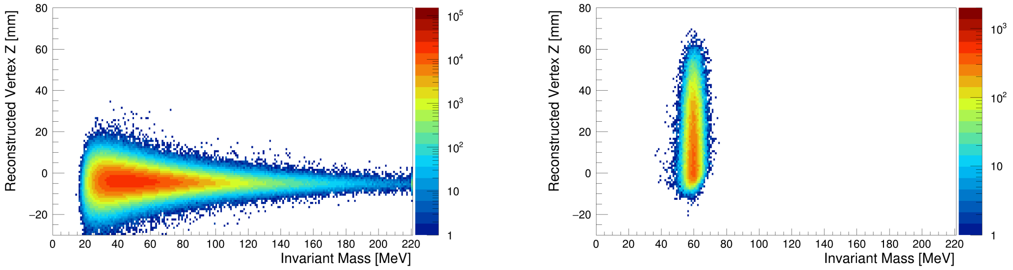


Figure 5.28: Plot showing $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ cut values optimized as a function of invariant mass. The cut values were optimized using the ZBi figure of merit with the zcut-style analysis. A flat cut requiring $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}} < 2.0$ was ultimately chosen.



(a) 10 % Data before N$\sigma_{\mathrm{V0_{proj}}}$ cut.

(b) 60 MeV MC signal before N$\sigma_{\mathrm{V0_{proj}}}$ cut.



(c) 10 % data after N$\sigma_{\mathrm{V0_{proj}}}$ cut.

(d) 60 MeV MC signal after N$\sigma_{\mathrm{V0_{proj}}}$ cut.

Figure 5.29: Mass versus reconstructed vertex z for 10 % data and 60 MeV MC signal, shown both before and after applying the vertex projection significance cut (N$\sigma_{\mathrm{V0_{proj}}} <$ 2.0).

5.2.2.6 Impact Parameter Cut

As discussed previously, truly displaced signals should exhibit tracks with symmetric vertical impact parameters ($y0$) that increase linearly with the z_{vtx} as shown in the top and bottom right diagrams of Figure 5.30. In contrast, correctly reconstructed prompt events should have $y0 \sim 0$ mm for both tracks, as illustrated in the top left of Figure 5.30. However, in cases of large scattering or track mis-reconstruction (e.g., using the wrong L1 hit), one track may exhibit a larger impact parameter than its partner, resulting in a falsely displaced vertex, depicted in the bottom left of Figure 5.30.

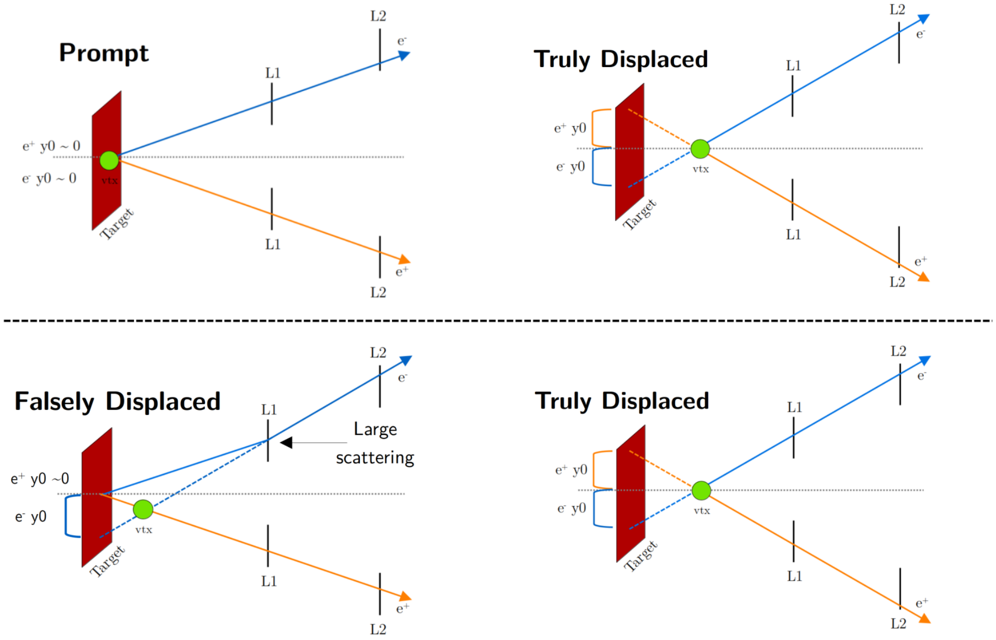


Figure 5.30: Illustrations of the track impact parameters $y0$ at the target for truly displaced events, prompt events, and falsely displaced events due to scattering or reconstruction errors.

Figure 5.31 shows $y0$ versus z_{vtx} for 45 MeV MC signal (top) and MC background in the same mass window. Each vertex has two entries, one for the electron track and the positron track. The MC signal demonstrates the expected linear correlation between $y0$

and $\mathrm{z_{vtx}}$ for both tracks. However, the data displays four distinct characteristics:

- A prompt core, uncorrelated with $\mathrm{z_{vtx}}$, where both tracks are correctly reconstructed with $y0$ centered on 0 mm, consistent with originating from the target.
- Two displaced $\mathrm{z_{vtx}}$ tails that are linearly correlated with $y0$, corresponding to mis-reconstructed/large-scattered tracks that are inconsistent with originating from the target.
- A single displaced $\mathrm{z_{vtx}}$ tail centered at 0 mm, corresponding to correctly reconstructed tracks that are partnered with a mis-reconstructed/large-scattered track with the same $\mathrm{z_{vtx}}$ position.

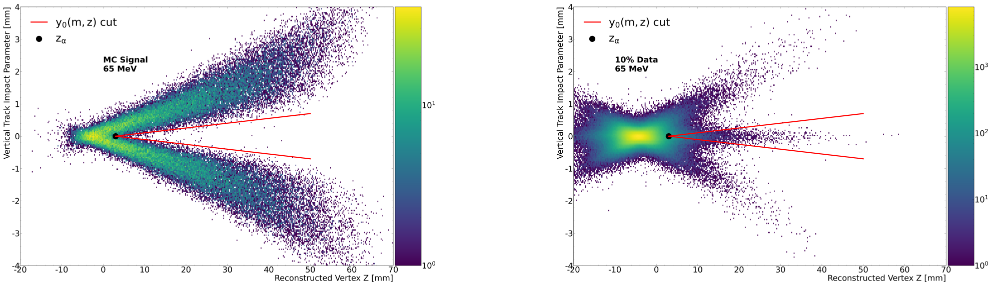
The asymmetry in the track vertical impact parameters for falsely displaced events can be exploited by cutting events where either track has a small track vertical impact parameter (near-zero), which is inconsistent with long-lived signal. An example of this cut is demonstrated by the red cut-lines in Figure 5.31.

The $y0$ cut can be parameterized as a function of signal mass by:

$$y0_{+}(m,z) > \alpha(m)\,($$

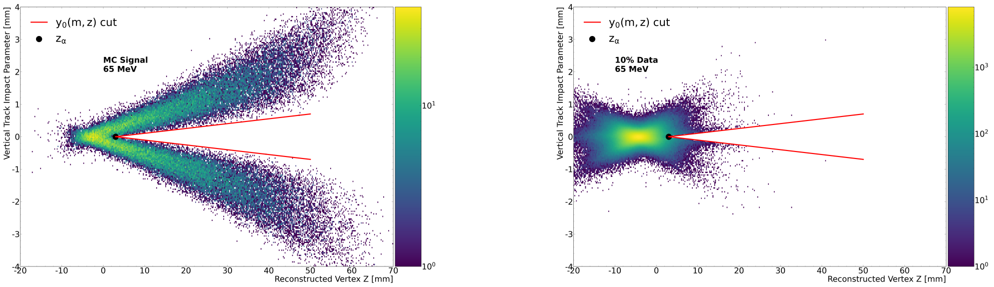
$\mathrm{z_{vtx}} - z_{\alpha}$

and $y0_{-}(m,z) < -\alpha(m)\,(\mathrm{z_{vtx}} - z_{\alpha},$ (5.24) where $\alpha(m)$ represents the slope of the cut as a function of mass. The displaced background rate can be further reduced by increasing the $y0$ cut slope $\alpha(m)$, and translating the cut in terms of z_{α}. However, optimizing these parameters is challenging due to the correlation between $y0$, $\mathrm{z_{vtx}}$, and the signal lifetime. As the lifetime increases, events at higher $\mathrm{z_{vtx}}$ contribute more to the expected signal. Therefore, this cut must be parameterized as function of both mass and ϵ.



(a) MC signal with mass 65 MeV before the cut.

(b) MC background before the cut.



(c) MC signal with mass 65 MeV after the cut.

(d) MC background after the cut.

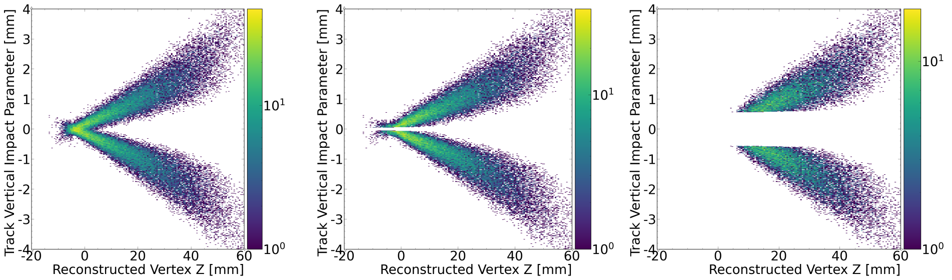
Figure 5.31: Example of the track vertical impact parameter versus reconstructed vertex z before and after applying the impact parameter cut defined in Section 5.2.2.6. Each events has two entries, one for each track. (a) and (c) show MC signal with mass 65 MeV, while (b) and (d) show the data for MC background. The top row represents the data before the cut, and the bottom row represents the data after the cut.

5.2.2.7 Track Vertical Impact Parameter Cut Optimization

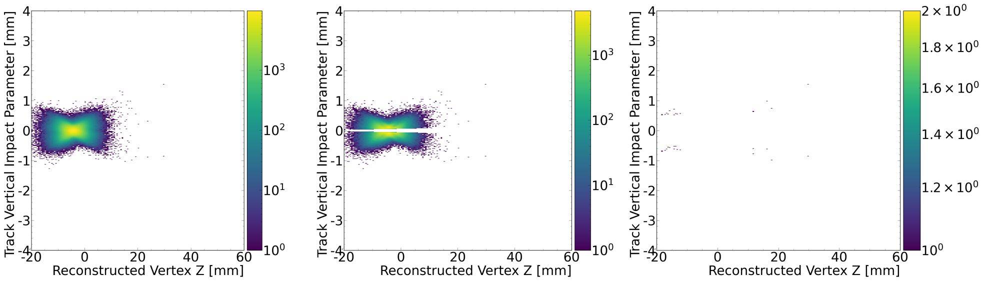
The $y0$ cut parameters α and z_α were optimized using the same ZBi procedure outlined for the $\mathrm{N}\sigma_{\mathrm{V0_{proj}}}$ optimization in Section 5.2.2.5. Since the motivation behind the $\mathrm{N}\sigma_{\mathrm{V0_{proj}}}$ cut is unrelated to $y0$, despite significantly reducing the rate of large $y0$ events, the cut $y0 < 2.0$ was applied prior to this optimization. Various values of the $y0$ cut slope α were tested across different MC-generated signal mass windows. The optimization process began by identifying the value of z_α the retains 100 % of the relative signal efficiency. The cut was then iteratively tightened by shifting z_α upstream towards the target in $\mathrm{z_{vtx}}$. During this process, it was found that very small values of α were capable of eliminating nearly all of the background events without needing a zcut, as long as z_α is shifted very far upstream in the negative $\mathrm{z_{vtx}}$ direction. This approach not only creates a near-zero background environment, but also maintains a relatively high efficiency for long-lived signal. The effectiveness of this cutting method is demonstrated in Figure 5.32.

One consequence of this method is that the remaining background rate quickly becomes too small to model by fitting. Therefore, instead of fitting the tail of the background with an exponential function as done in the zcut-based analysis, the expected rate background simply taken as the number of remaining events. Figure 5.33 shows the $y0$ parameter optimization results for a mass window centered on 55 MeV. The y-axis represents the $y0$ cut slope α and the x-axis represents the translation of z_α upstream with iteration number. The ZBi is apparently maximized for the smallest values of α.

The case where $\alpha = 0.0$ is a special scenario where a flat cut on $|y0|$ is applied and iteratively tightened in $y0$ rather thanz_α. The expected signal rate is relatively stable as



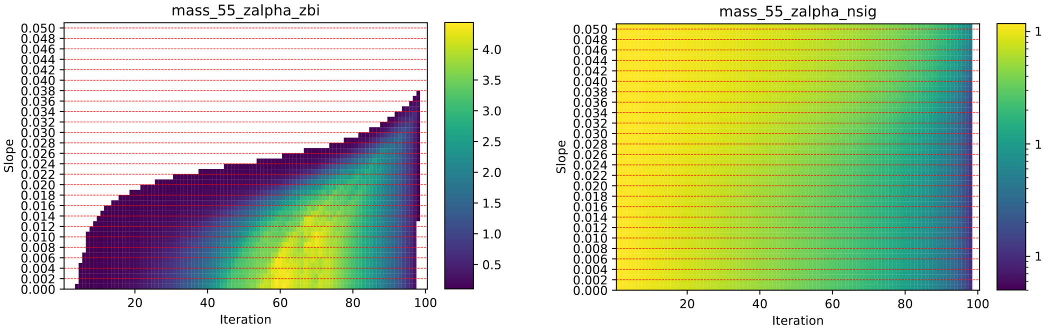
(a) 65 MeV MC signal events as z_α is pulled upstream (larger negative values).



(b) MC background rate in a 65 MeV mass window as z_α is pulled upstream (larger negative values).

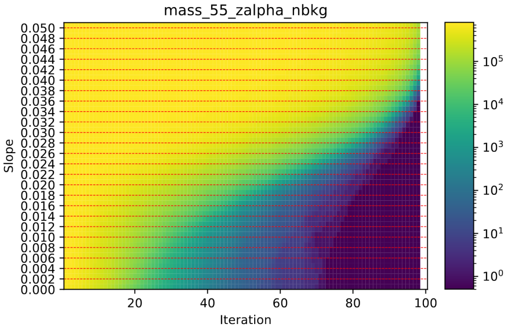
Figure 5.32: Example demonstrating the effectiveness of the track vertical impact parameter cut defined in Section 5.2.2.6 when using using small angles (α), and pulling z_α upstream in $\mathrm{z_{vtx}}$. (a) Signal efficiency remains high even as (b) the background is nearly eliminated in 65 MeV mass window.

the cut is tightened in z_α, while the number of background events decreases exponentially. The $y0$ optimization results are summarized across different masses in Figure 5.34, where each bin corresponds to the value of z_α that maximizes the ZBi. Once again, the smallest values of z_α result in the maximum ZBi with a significantly reduced background rate, independent of $\mathrm{z_{vtx}}$. Based on these findings the simplest parameterization of the $y0$ cut was chosen, using a flat cut on $|y0|$. As a result of this approach, the zcut is no longer required to define the final signal region. Instead, a simple zcut is applied to eliminate all



(a) ZBi vs. Slope for 55 MeV

(b) Nsig vs. Slope for 55 MeV



(c) Nbkg vs. Slope for 55 MeV

Figure 5.33: Optimization results for $y0$ in the 55 MeV mass window. The y-axis shows the slope α of the cut defined in Section 5.2.2.6. Each plot corresponds to different performance metrics: (a) ZBi, (b) Nsig, and (c) Nbkg.

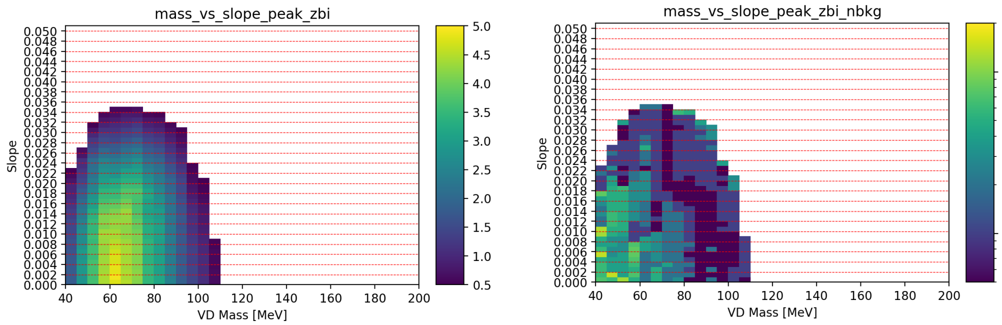
events at or behind the target location in z, defined by:

$\mathrm{z_{vtx}} > z_{\mathrm{target}}$, (5.25) where $z_{\mathrm{target}} = -4.3\,\mathrm{mm}$.

5.2.2.8 $y0^{\mathrm{min}}$ Cut Optimization

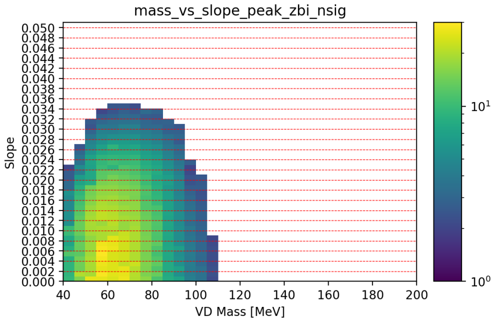
The flat cut on the track vertical impact parameter $y0$ is implemented by defining the new cut variable $y0^{\mathrm{min}}$ as:

$$y0^{\mathrm{min}} = \min(|y0_{e^-}|, |y0_{e^+}|). \tag{5.26}$$



(a) Slope (α) vs. ZBi for all masses

(b) Slope (α) vs. Nbkg for all masses



(c) Slope (α) vs. Nsig for all masses

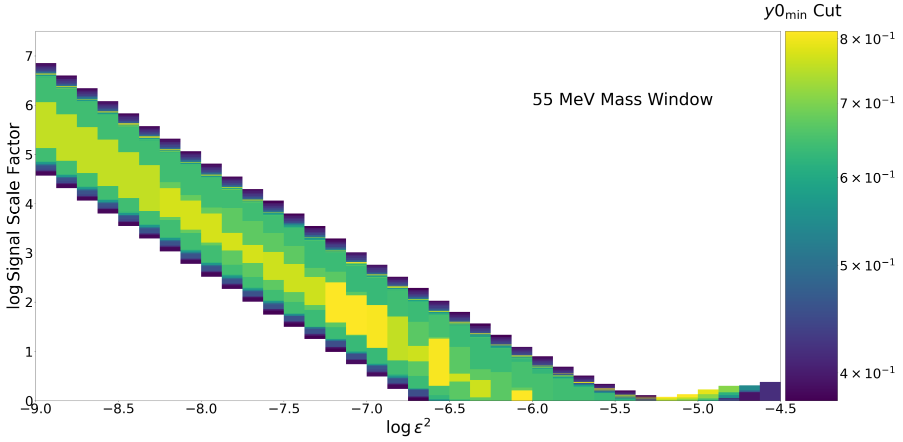
Figure 5.34: Optimization results for the $y0$ cut defined in Section 5.2.2.6 as a function of invariant mass and slope α. The iteration number on the x-axis corresponds to a particular value of z_α, where z_α is tightened in each iteration to whatever value cuts n% of the signal. The entry in each bin corresponds to the z_α parameter value that results in the maximum ZBi figure of merit. Each plot corresponds to different performance metrics: (a) ZBi, (b) Nsig, and (c) Nbkg.

Similar to the original $y0$ impact parameter cut, the $y0^{\mathrm{min}}$ cut value is optimized as a function of mass and ϵ. The expected signal rate, which is exponentially distributed in $\mathrm{z_{vtx}}$, results in a corresponding correlation between $y0^{\mathrm{min}}$ and $\mathrm{z_{vtx}}$. Consequently, the $y0^{\mathrm{min}}$ cut signal efficiency is a function of lifetime, parameterized by the coupling strength ϵ for a given mass. One could optimize the cut by simply using the expected signal rate in the 10 % data sample, but that is effectively zero for most of the signal parameter space. One could also scale up the expected signal rate to the full luminosity, but the background model still only represents approximately 10 % of the data, so the resulting ZBi doesn't necessarily represent expectations for the final result. These things considered, the optimization is performed over a range of signal scale factors to test how the optimized $y0^{\mathrm{min}}$ cut value varies with the expected signal rate.

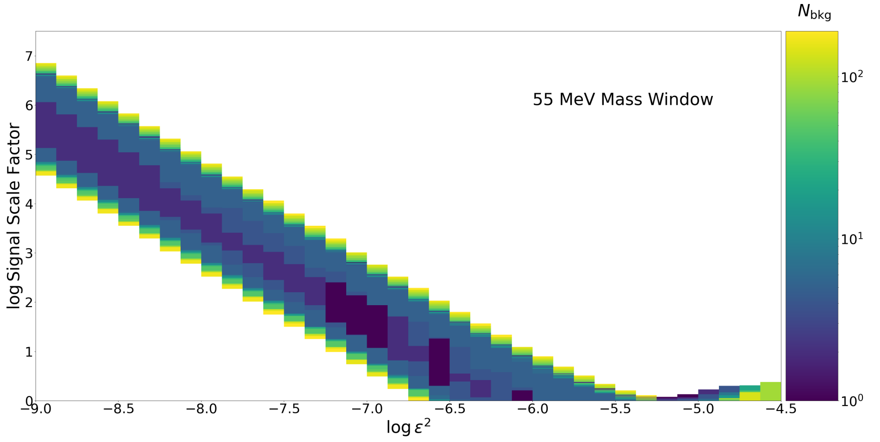
The $y0^{\mathrm{min}}$ cut values are optimized to maximize sensitivity to longer-lived signals (smaller values of ϵ), where the $y0^{\mathrm{min}}$ cut is most capable of distinguishing signal from background. Figure 5.35 presents the optimization results for 55 MeV MC signal as a function of ϵ on the x-axis, with the signal rate multiplied by the scale factor on the y-axis. The histogram density shows the value of the $y0^{\mathrm{min}}$ cut that maximizes the ZBi for a given ϵ and scale factor. Longer lifetimes, corresponding to smaller ϵ values, are represented on the left-hand side of the plot. For prompt-like signals on the right side of the plot, the $y0^{\mathrm{min}}$ cut tends to be too close to 0 mm, resulting in an unacceptably high background rate. However, the optimum $y0^{\mathrm{min}}$ cut value for smaller values of ϵ reaches a consistent maximum as a function of the signal scale factor on the y-axis, $y0^{\mathrm{min}} = 0.8\,\mathrm{mm}$ for the results shown in Figure 5.35. If the signal rate is too low, cutting any amount of the exponentially-distributed signal in $y0$ results in zero expected signal, so the $y0^{\mathrm{min}}$ cut remains loose, resulting in a significant background rate. As the signal rate is artificially scaled up, the expected signal rate for larger values of $y0$ increases, and the optimum $y0^{\mathrm{min}}$ cut value increases. The signal scale factor eventually becomes large enough that the exponential core of the signal $y0^{\mathrm{min}}$ distribution becomes significant relative to the background, and the optimum $y0^{\mathrm{min}}$ cut value loosens. The overall optimum $y0^{\mathrm{min}}$ cut value is chosen to be this turning point in $y0^{\mathrm{min}}$, resulting in a very low, but not necessarily zero, background rate.

Figure 5.36 shows the optimized $y0^{\min}$ cut values as a function of mass, parameterized by:

$$y0^{\min}(m) > \left(1.0762 - 7.44534 \times 10^{-3} m + 1.58746 \times 10^{-5} m^2\right) \text{ mm} \tag{5.27}$$



(a) $y0^{\min}$ cut value as a function of coupling strength ϵ and signal scale factor.



(b) Number of background events as a function of the $y0^{\min}$ cut value in (a).

Figure 5.35: Optimization results for $y0^{\min}$ using 10 % data and 55 MeV MC-generated signal. (a) The $y0^{\min}$ cut value (in mm) that maximizes the ZBi as a function of the coupling strength ϵ and the overall scale factor on the signal. (b) The number of background events as a function of the optimized $y0^{\min}$ cut in (a).

The performance of this cut is illustrated in Figure 5.37(a), which shows the

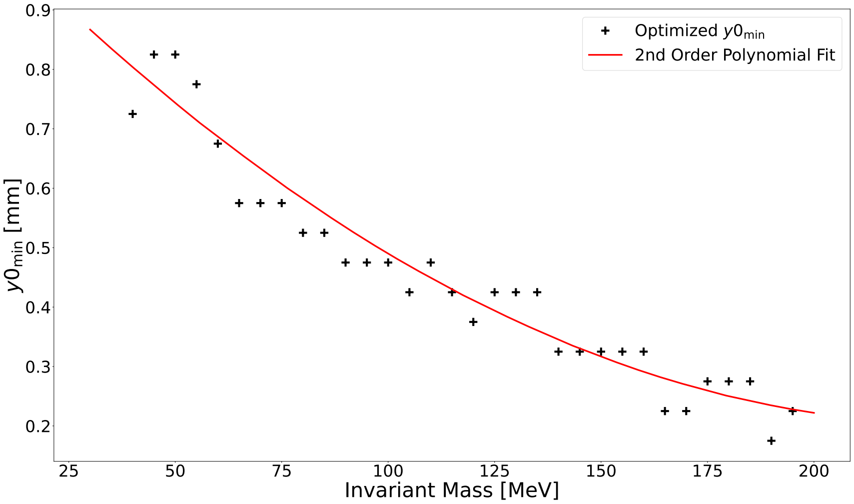
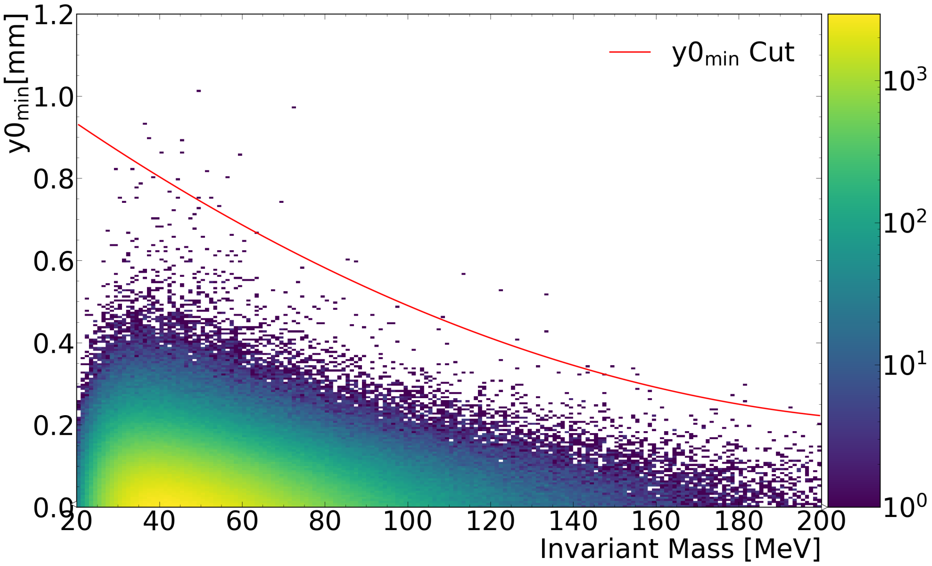


Figure 5.36: The solid black points represent the optimized $y0^{\mathrm{min}}$ cut value as a function of signal mass, parameterized by the second-order polynomial fit shown by the solid red line. The number of remaining tight selected events in 10 % data is depicted by the blue solid line, also as a function of invariant mass.

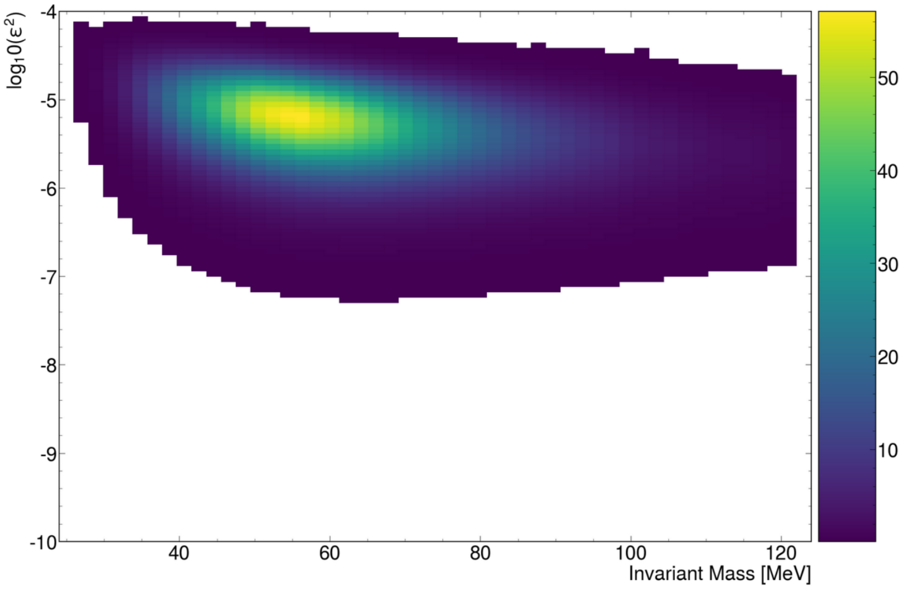
complete invariant mass distribution for 10 % data with all cuts applied, except for the optimized $y0^{\mathrm{min}}$ cut indicated in red. The total remaining background rate is roughly 30 events, while the maximum expected signal rate with all cuts applied, shown in Figure 5.37(b), is more than 50 events. The expected signal rate assumes a search window that is four times the mass resolution of the mass bin, corresponding to roughly 90 % of the signal. While the shape of the $y0^{\mathrm{min}}$ cut as a function of mass is optimized in this section, the cut is further tightened later in Section 6.2.0.1 by adding 0.1 mm to the fit function defined in Equation (5.27). This adjustment is made to account for statistical fluctuations in the full luminosity data, which aren't expected to alter the shape of the best cut. The fully optimized tight selection is summarized in Table 5.5, with the final $y0^{\mathrm{min}}$ cut parameterization given, rather than the version derived here.

Cut	**Condition**
Layer 1 Requirement	e^- and e^+ have L1 axial+stereo hit
Layer 2 Requirement	e^- and e^+ have L2 axial+stereo hit
One Vertex Per Event	$\mathrm{N_{vtx}} = 1$
Target Projected Vertex Significance Cut ($\mathrm{N}\sigma_{\mathrm{V0_{proj}}}$)	$\mathrm{N}\sigma_{\mathrm{V0_{proj}}} < 2.0$
Target Z Cut	$\mathrm{z_{vtx}} > -4.3\,[\mathrm{mm}]$
Minimum Track Vertical Impact Parameter Cut	$y0^{\mathrm{min}}(m) > 1.0762 - 7.44534 \times 10^{-3}m$
	$+1.58746 \times 10^{-5}m^2$ [mm]

Table 5.5: Tight Cuts Optimized Using 10% Data. The impact parameter cut is parameterized as a function of mass m in MeV.



(a) Minimum track vertical impact parameter cut versus mass



(b) Reach estimate for tight selection

Figure 5.37: Plots have all preselection and tight cuts applied, except for the minimum track vertical impact parameter cut. (a) Minimum track vertical impact parameter cut as a function of mass. (b) Reach estimate for the tight selection criteria with a $\pm 2\sigma_m$ search window.

Chapter 6

Displaced V_D Vertex Search

This analysis conducts a displaced vertex search for long-lived hidden sector vector mesons (V_D) that invisibly decay to e^+e^- between 1 cm to 10 cm downstream of a fixed target. The search is optimized to identify signals with an invariant mass in the range of 30 MeV to 124 MeV, and kinetic mixing strength $10^{-6} < \epsilon < 10^{-2}$. The signal is expected to appear as an excess number of displaced events distributed around a specific invariant mass. To search for signal in the data, the invariant mass spectrum between 30 MeV to 124 MeV is divided into overlapping search windows of some predetermined size given in terms of the mass resolution σ_m, with window centers spaced apart by 1 MeV. The signal rate is suppressed by many orders of magnitude relative to the QED backgrounds. Thus the analysis applies a series of selection stages, defined in Section 5.1, selecting only high-quality reconstructed events with signal-like kinematics. This is achieved primarily by using the target projected vertex significance ($\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$) and minimum track vertical impact parameter ($y0^{\mathrm{min}}$) cuts. The former eliminates events inconsistent with the beam interaction point, while the latter requires large track vertical impact parameters, thus eliminating nearly all of the prompt background.

Based on the selection studies using 10 % data, the background rate in the 100 % data SR is expected to be on the order of dozens of events distributed across the invariant mass spectrum, while the peak expected signal is ~ 50 events in a $4\sigma_m$ wide search window. However, the studies presented in this section demonstrated that tightening the $y0^{\mathrm{min}}$ cut relative to the ZBi optimized version further constrains the background without significantly impacting the expected signal sensitivity. In addition, since the search takes place in invariant mass windows, the background rate in each search window increases with window size based on the mass resolution. Thus, the search window size was slightly reduced from $4\sigma_m$ to further minimize background while balancing signal sensitivity. This section describes the methods used to estimate the expected background in each search window and calculate the significance of the data before presenting the 100 % data results. The OIM is also described, used to set upper limits on the SIMP signal rate. These methods were developed and validated before unblinding using the 10 % data sample and a larger statistics sample of 100 % data in the CR.

6.1 Estimating the Expected Background

Accurately estimating the background rate in each search window is crucial for signal detection. However, estimating the background is challenging due to the lack of a control data set that accurately reflects the falsely displaced background in the SR after high-$z$cuts. MC simulations do not fully match the data in the $y0^{\mathrm{min}}$ and $\mathrm{z}_{\mathrm{vtx}}$ tails, and while the CR in 100 % of the data is unblinded, the event characteristics differ significantly from the lower momentum SR. Additionally, the 10 % data sample used to develop the selection provides limited statistics after all selections, resulting in large uncertainties in the expected background rate across search window masses.

To overcome the lack of a suitable control data set, the background in each search window is estimated using mass sidebands in the SR. This approach was evaluated using both the 10 % data sample and 100 % data in the CR. The CR is not used to estimate the expected background in the SR. Instead, it is used to validate the performance of the background estimation method by comparing the estimate results to the observed number of events. The background is continuous invariant mass and does not differ significantly between neighboring masses. Since the expected signal is centered in the search window, the mass sidebands are considered background-only and can be used to estimate the background in the search window. One straightforward approach is to average the number of observed events in mass sidebands on both sides of the search window, minimizing any correlation between mass and background rate. However, this simple averaging fails to account for statistical fluctuations and performs poorly.

A more accurate approach estimates the background using a variant of the "ABCD" background estimation method by extending the mass sideband into the $y0^{\mathrm{min}}$ parameter space. The background rate in a search window is estimated using six orthogonal regions defined in invariant mass versus $y0^{\mathrm{min}}$, as illustrated in Figure 6.1. The mass sidebands are required to be far enough from the center of the search window to have negligible signal contamination.

Region F is the signal region of the search window, including the $y0^{\mathrm{min}}$ selection. Regions A and E are defined as neighboring mass windows sharing the same $y0^{\mathrm{min}}$ selection

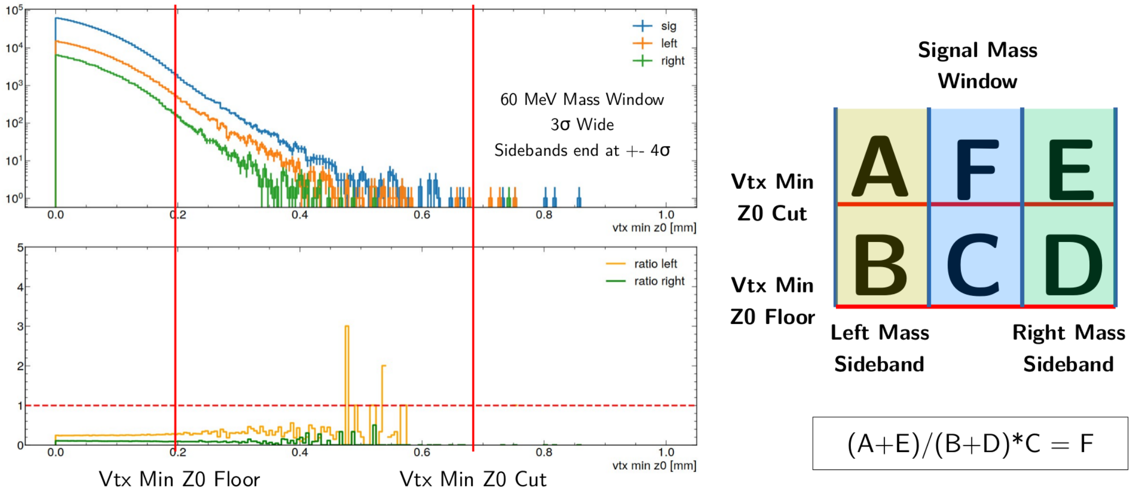
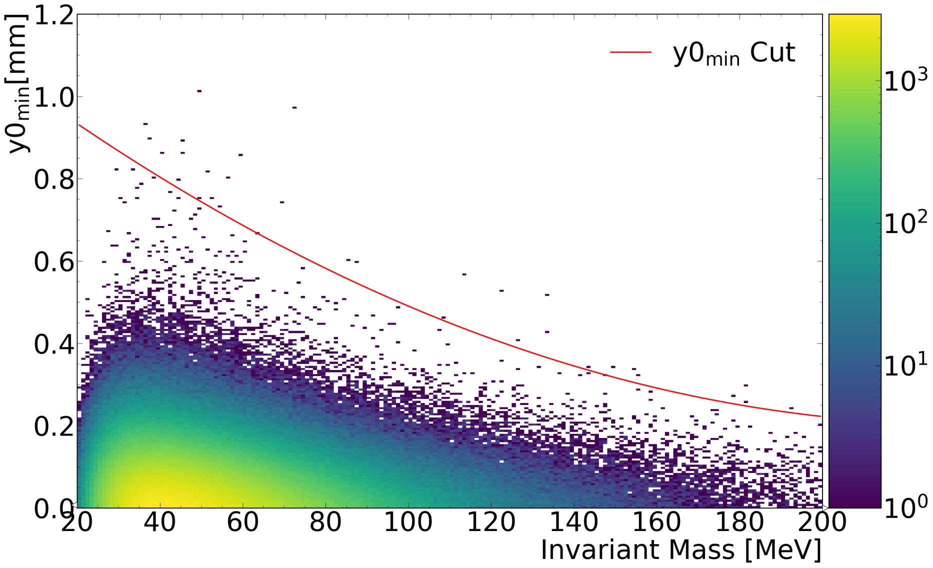


Figure 6.1: Illustration of the ABCD background estimation method. The plot shows histograms of $y0^{\min}$ for the 60 MeV mass search window and the two mass sidebands.

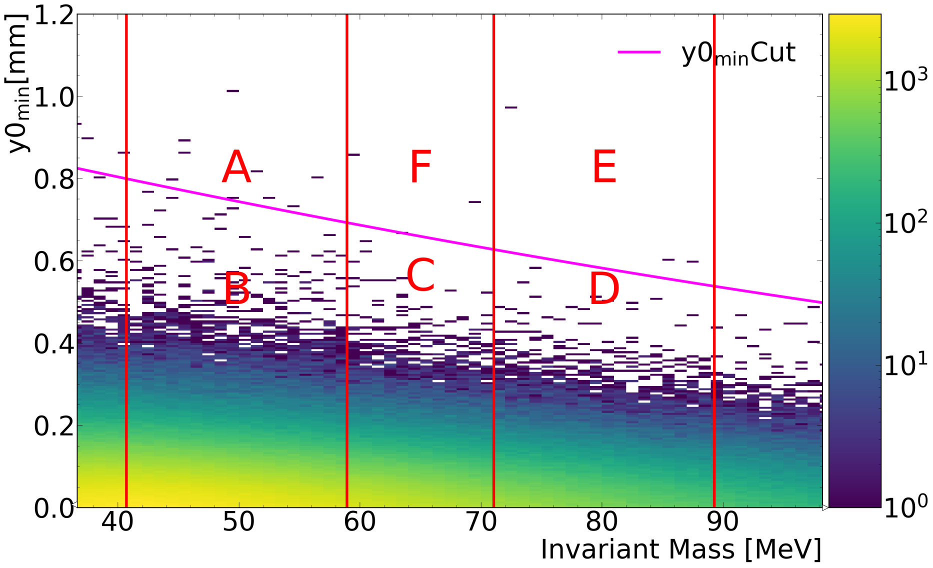
as F. Regions B, C, and D are extensions of A, F, and E respectively, into $y0^{\min}$. Assuming the fraction of background events passing the $y0^{\min}$ selection is similar in each mass window, the background rate (b) in F (the signal region) can be estimated using:

$$b = \left(\frac{A+E}{B+D}\right) C. \tag{6.1}$$

While $y0^{\min}$ and mass are correlated, the correlation is approximately linear within a narrow mass region, as shown in Figure 6.2. This linearity allows the correlation to cancel out when averaging.



(a) Preliminary $y0^{\text{min}}$ vs. Mass



(b) ABCD Regions

Figure 6.2: Invariant mass versus minimum track vertical impact parameter with all other cuts applied for 10 % data. (a) shows the preliminary $y0$ cut function. (b) illustrates the ABCD background estimation regions for the search window centered on 65 MeV. The mass region is approximately linear in $y0^{\text{min}}$.

Since the signal region is in the tail of the $y0^{\mathrm{min}}$ distribution, the background estimate should be driven by those events rather than the core of $y0^{\mathrm{min}}$. This is accomplished by restricting the $y0^{\mathrm{min}}$ parameter space in regions B, C, and D to $y0^{\mathrm{min}} > y0^{\mathrm{min}}_{\mathrm{floor}}$. It's necessary to mitigate signal contamination in C to avoid overestimating the background, so $y0^{\mathrm{min}}_{\mathrm{floor}}$ is defined, somewhat arbitrarily, as the $y0^{\mathrm{min}}$ value with at least 1,000 events in C. This ensures that the exponentially distributed background rate is at least an order of magnitude larger than the relatively flat signal and dominates the count in C. This threshold was validated by injecting a high rate of MC signal events into the data at each mass and for different lifetimes, ensuring that the background estimate remained independent of the injected signal.

There are two additional parameters involved in the background estimate: the mass sideband and search window sizes. The sideband width can significantly impact the background estimate quality because the validity of Equation (6.1) requires the mass sidebands to be close enough to center of the search window to maintain the linear correlation with $y0^{\mathrm{min}}$. If the sidebands are too large, they can extend beyond the linear region and systematically overestimate the background. Conversely, if the sidebands are too narrow, they fail to smooth out the natural statistical fluctuations in the background, resulting in an unreliable estimate. Additionally, the width of the search window both impacts the quality of the background estimate, and the search sensitivity. The search window must be large enough to maintain a high signal efficiency, but not so large that the background rate grows too quickly relative to signal.

In light of these considerations, the search window size was initially set to $\pm 2.0\sigma_m$, corresponding to roughly 90 % signal efficiency. The sidebands were set to $2\sigma_m$ wide, meaning they extend out to $6\sigma_m$ from the center of the search window. These preliminary values were used to test the background estimation method before optimizing them later in this section. The first test compared the observed number of events in each search window using the background-only 10 % data sample, shown in Figure 6.3. Each invariant mass bin in the plot represents an overlapping search window that is $\pm 2\sigma$ wide. The relatively flat ratio around 1 indicates good performance of the background estimation. While the estimate has large errors due to low statistics, it successfully describes the overall shape of the background.

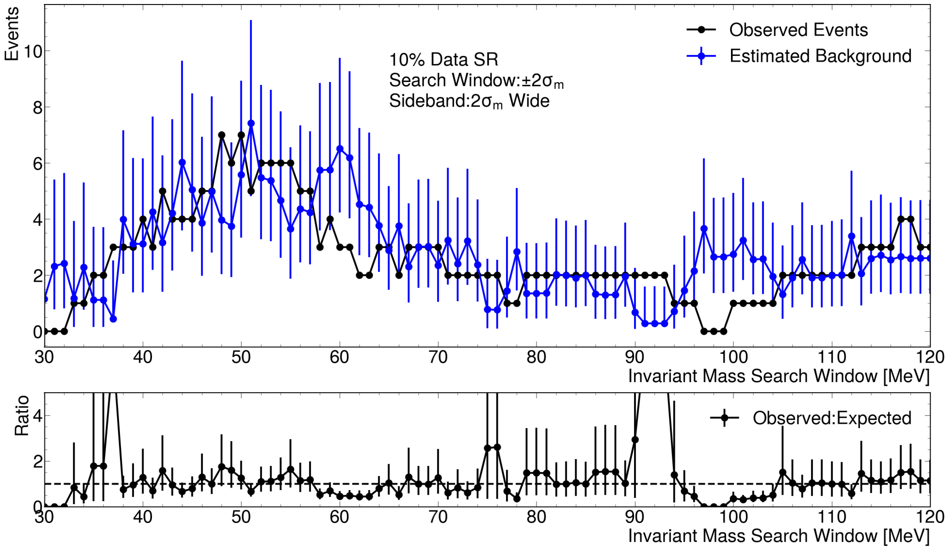


Figure 6.3: Preliminary tight selection events in 10 % Data. Top: The estimated background (blue) is compared to the observed number of events (black) as a function of the invariant mass search window. The search window is size $\pm 2\sigma_m$, with mass sidebands $2\sigma_m$ wide. Bottom: The ratio of the expected number of background to the observed number of events.

The background estimation was also tested using the 100 % data CR sample in Figure 6.3, which shows a similar degree of success in capturing the overall shape of the background. The CR selection used here is not optimized and does not represent the expected background in the SR. There is an apparent oscillation in the expected background relative to the number of observed events between 50 MeV to 20 MeV, however this is expected due to statistical fluctuations between search windows.

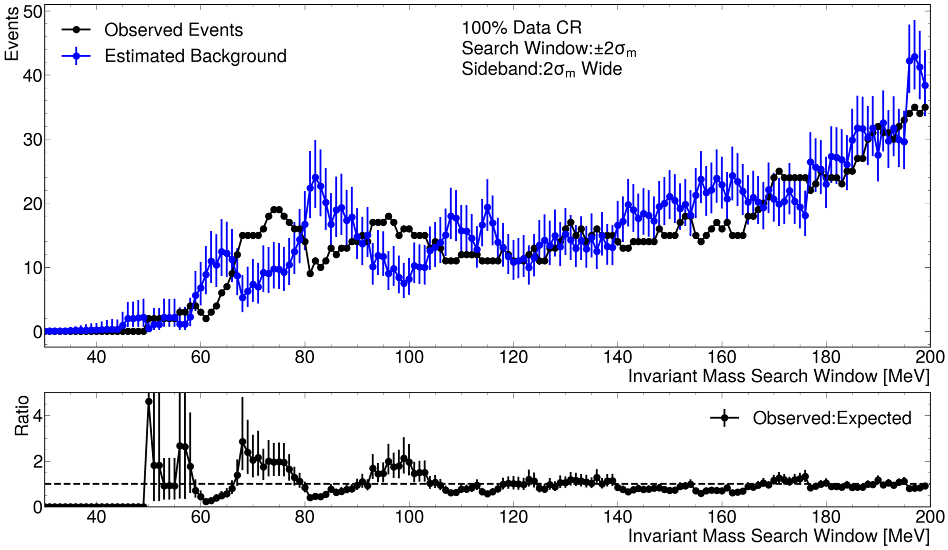


Figure 6.4: Preliminary tight selection events in the 100 % Data CR. The estimated background (blue) is compared to the observed number of events (black), as a function of the invariant mass search window. The search window is size $\pm 2\sigma_m$, with mass sidebands $2\sigma_m$ wide.

The mass sideband and search window parameterizations were studied further across a range of sizes to test the performance of the background estimation. The results showed that when the sidebands are very narrow the statistical fluctuations of the background are enhanced, leading to extreme oscillations in the expected background. Conversely, if the sidebands are too wide the background is typically overestimated due to the non-linear increase in events in lower-mass sideband (regions A and B). The performance was also loosely correlated to the size of the search window. In general, it was found that search window sizes between $\pm\,(1.5-2.5)\,\sigma_m$ with mass sideband widths between $(3-6)\,\sigma_m$ provide reasonable background estimates. The working point for the mass sideband width was selected as $4\sigma_m$, large enough to smooth out statistical fluctuations but narrow enough to maintain the required approximate linear correlation between mass and $y0^{\mathrm{min}}$.

With the mass sideband width fixed to $4\sigma_m$, the background estimation was again studied as a function of search window size. Figure 6.5 compares the estimated and observed background rate for different search windows using 10 % data in the SR, with the same results using 100 % CR data shown in Figure 6.6. These studies shown that the background is systematically overestimated, particularly at higher masses, when the search window is large (greater than $\pm 2\sigma_m$). In contrast, smaller windows result in both more accurate background estimations and a significantly decreased background rate. However, the search window size cannot be taken to arbitrarily small sizes because the signal efficiency decreases and the sideband contamination increases.

In order to account for both of these effects, the “optimal” search window size was chosen by injecting MC simulated signals into the data and calculating the significance of the injected data across a range of search windows and sizes. Before this study could be performed, the method for calculating the significance of the data had to be established, which is detailed in the following section before returning to the search window optimization. This study required choosing a method to calculated the significance of the data, which is detailed in the following section.

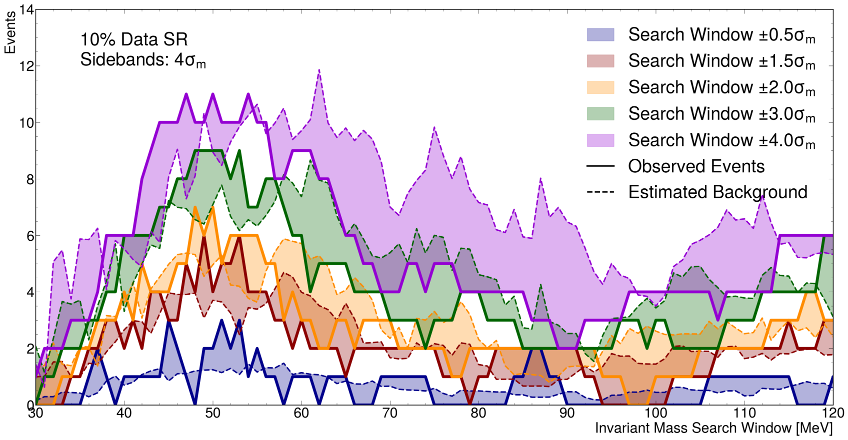


Figure 6.5: Preliminary tight selection 10 % data SR across a range of search window sizes. The observed number of events in search window (solid line) is compared to the estimated background (dashed line) for each search window size (color). The mass sideband width is $4\sigma_m$.

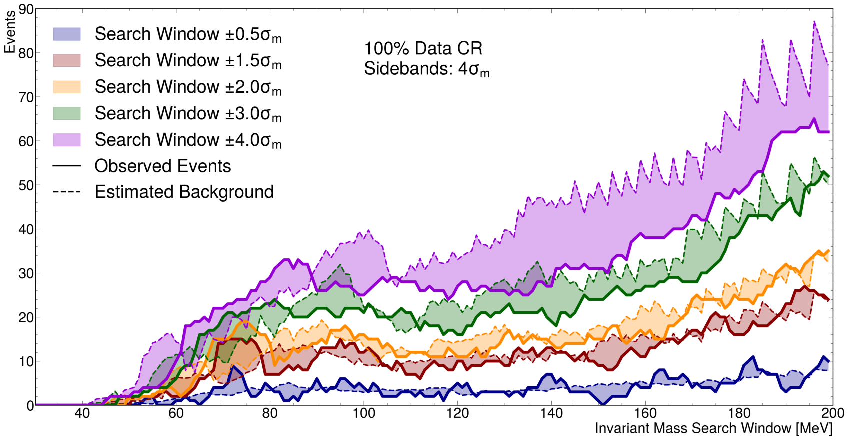


Figure 6.6: Preliminary tight selection 100 % data CR across a range of search window sizes. The observed number of events in search window (solid line) is compared to the estimated background (dashed line) for each search window size (color). The mass sideband width is $4\sigma_m$.

6.2 Data Significance Calculation

One approach to measure the data significance in a search window is by calculating the probability of observing at least n events, assuming the background follows a Poisson distribution with a mean of b:

$$P(n, b) = \sum_{k=n}^{\infty} \frac{b^k}{k!} e^{-b}. \tag{6.2}$$

However, this approach fails to account for the statistical uncertainty of the background estimate, resulting in an overestimate of the significance and potentially false evidence of signals. To account for this uncertainty and provide a more accurate significance measure, toy MC experiments are conducted. The parent distributions characterizing the ABCD regions used to estimate the background rate in Equation (6.1) are given by:

$$(B + D) \sim \mathcal{N}(\mu = B + D, \sigma = \sqrt{B + D}), \tag{6.3}$$

$$C \sim \mathcal{N}\left(\mu = C, \sigma = \sqrt{C}\right), \tag{6.4}$$

and

$$(A + E) \sim \text{Poisson}(\mu = A + E). \tag{6.5}$$

The statistical uncertainty is estimated by executing numerous toy MC experiments and sampling the three parent distributions to calculate b, where b represents the mean of the expected background distribution in the toy experiment, given by:

$$f(x) = \frac{b^k}{x!} e^{-b}. \tag{6.6}$$

The background-only test statistic t_0 is defined as the random sample of Equation (6.6) for each toy. The histogram of t_0 forms the background-only test statistic distribution $f(t_0)$

used to calculate the data significance by measuring the p-value given by:

$$\mathrm{p_{local}} = \int_{\mathrm{t_{obs}}}^{\infty} f(t_0)\,\mathrm{d}t_0, \tag{6.7}$$

where t_{obs} is the number of observed events in the search window. This method ensures that the significance calculation reflects the uncertainty of the background estimation, which is rather large given the low-statistics nature of the search, reducing the likelihood of false-positive signals.

The search result of each search window is summarized by the local p-value calculated using Equation (6.7), which corresponds to some local significance threshold $N\sigma_{\mathrm{local}}$. However, the true parameter of interest is the maximum global p-value across all of the overlapping search windows. The significance $N\sigma_{global}$ must be corrected to account for the additional statistical fluctuations introduced by searching over many independent search windows, which can increase the likelihood of observing a significant fluctuation. This is referred to as the Look Elsewhere Effect (LEE). For the search window mass range spanning 30 MeV to 124 MeV with an average mass resolution $\sigma_{\mathrm{m,avg}} = 3.63\,\mathrm{MeV}$, the approximate LEE correction is given by:

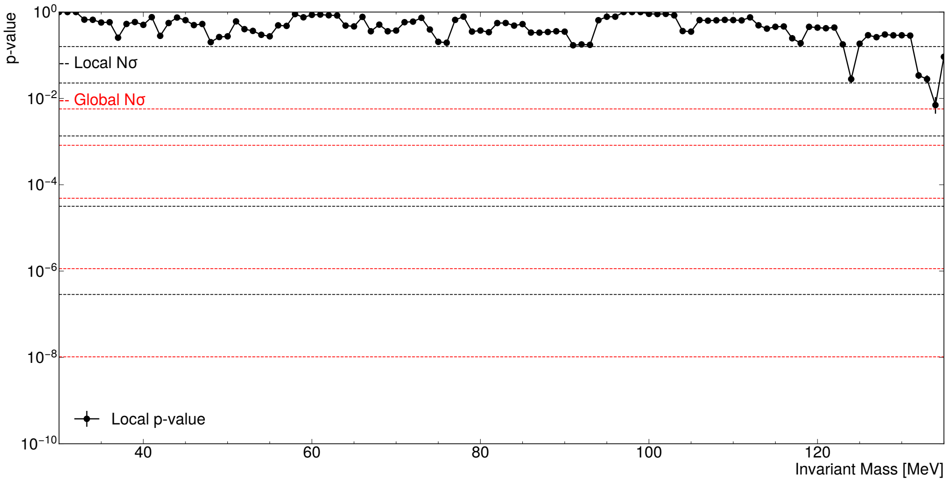
$$L_{\mathrm{corr}} \approx \frac{m_{\mathrm{max}} - m_{\mathrm{min}}}{\sigma_{\mathrm{m,avg}}} = 27.0, \tag{6.8}$$

where dividing by $\sigma_{\mathrm{m,avg}}$ accounts for the fact that each search window is not entirely independent.

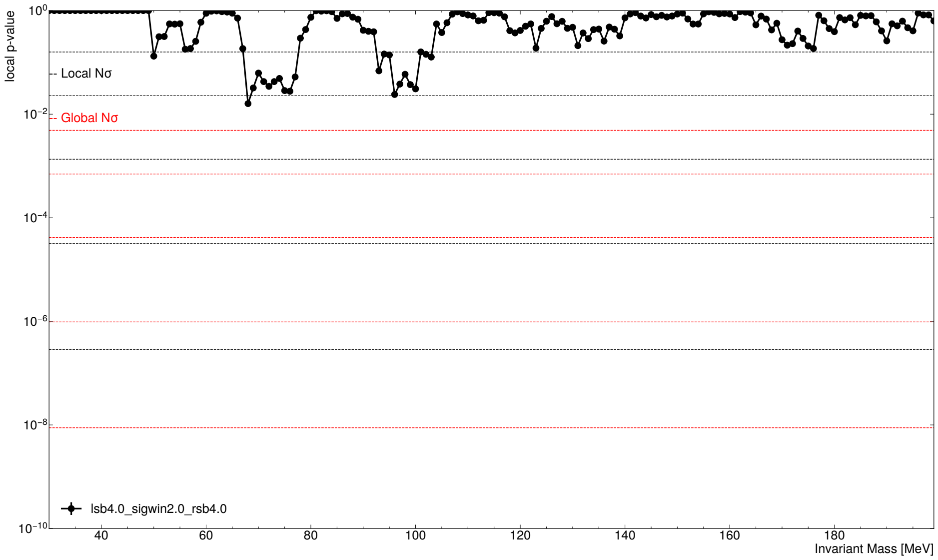
This significance calculation method was evaluated using the approximately "background-only" 10 %, shown in Figure 6.7(a), and the 100 % CR data sample in Figure 6.7(b). These preliminary results use a $\pm 2\sigma_m$ search window with $4\sigma_m$ mass sidebands. The local significance thresholds are denoted by the dashed black lines, descending in order of $N\sigma_{\mathrm{local}}$, while the red dashed lines denote the LEE-corrected $N\sigma_{\mathrm{global}}$ thresholds, defined by:

$$N\sigma_{\mathrm{global}} = L_{\mathrm{corr}} N\sigma_{\mathrm{local}}. \tag{6.9}$$

Given the lack of significant p-value fluctuations, there is no evidence of signal in either of the background-only data samples, as expected. This indicates that the significance calculation is sufficiently conservative to mitigate the risk of false signals. Additionally, the level of fluctuations are not so insignificant that the approach appears to be overly conservative, minimizing any potential signals. The next section tests the significance calculation by injecting the data with MC-generated signals, ensuring that the search remains sensitive.



(a) Preliminary 10 % data SR local p-values



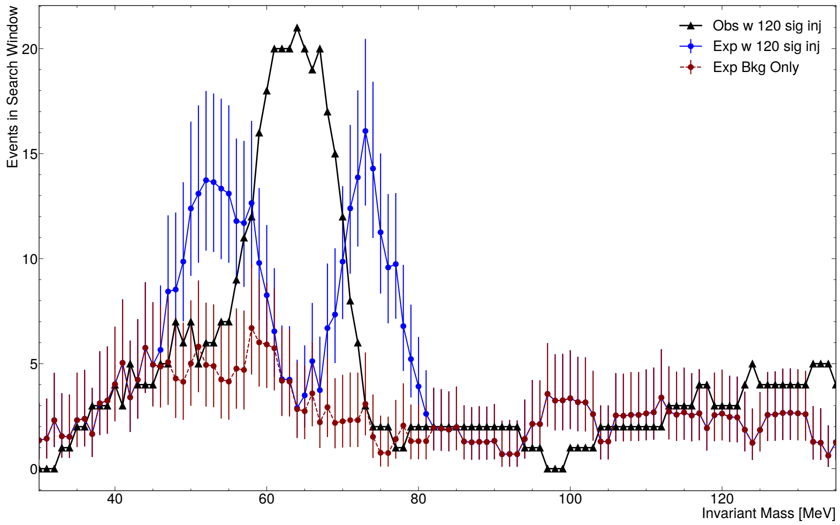
(b) Preliminary 100 % CR data local p-values

Figure 6.7: Demonstration of local p-value calculation results using control region data. Local significance thresholds are shown by black dashed lines, and global significance thresholds are shown in red. (a) shows preliminary 10 % data SR, and (b) shows 100 % CR data.

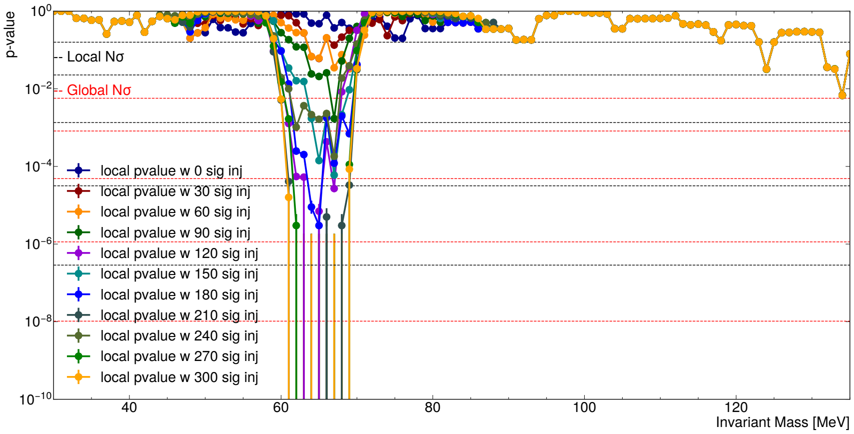
6.2.0.1 MC Signal Injection Studies

In order to validate the background estimation and significance calculation methods with signal present, MC simulated signals with a specified mass and ϵ were injected into the 10 % data SR, before applying the high-$z$cuts. The injected events were then reweighted according to the expected signal calculation given by Equation (5.17) to mimic real signal in the data. Figure 6.8(a) demonstrates the observed number of events and estimated background in each search window after injecting 120 MC-signal events at 64 MeV, with $\epsilon^2 = 6 \times 10^{-6}$ into 10 % The injected signal is observed as a Gaussian peak of approximately 21 events centered in the 64 MeV search window. Adjacent to the signal peak, the mass sideband background is significantly overestimated. However, this feature is inherent to the method, and is not a cause for concern. In the presence of real signal, the background in the signal adjacent sidebands is necessarily overestimated because the signal window with excess events serves as one of their sidebands. This feature is acceptable because it only occurs in the presence of a large signal, in which case the estimate in the sidebands is much less interesting.

Figure 6.8(b) shows the local p-values for various amounts of injected signal at the same mass and ϵ., where the significance of the p-value fluctuations increases with the number of signal as expected. The error bars in the figure are related to the number of toy MC experiments used to build the background-only test statistics distribution used to calculate the p-value in Equation (6.7). The result shown here is only meant to visually represent the expected response, thus a significantly smaller number of toys were thrown than are used in the final analysis, hence the large error bars.



(a) Expected Background and Observed Events with MC-Injected Signal.

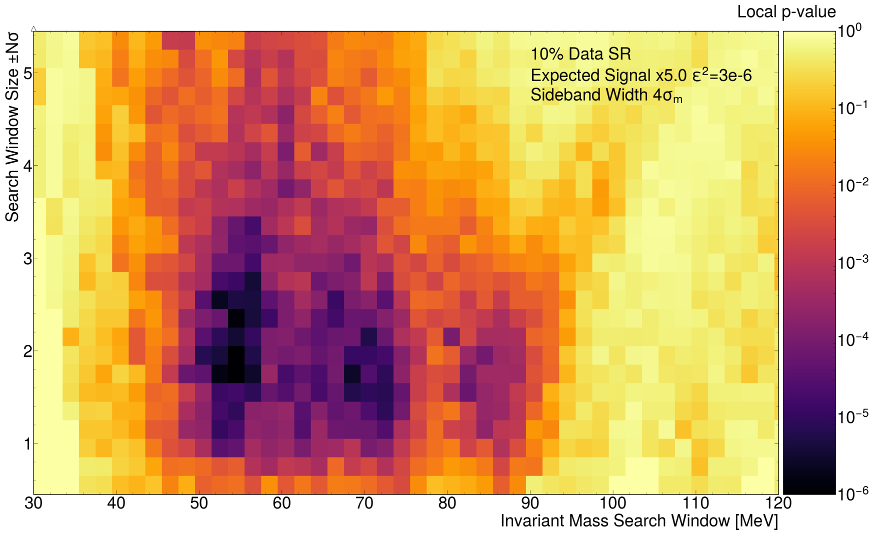


(b) P-value with MC-Injected Signals.

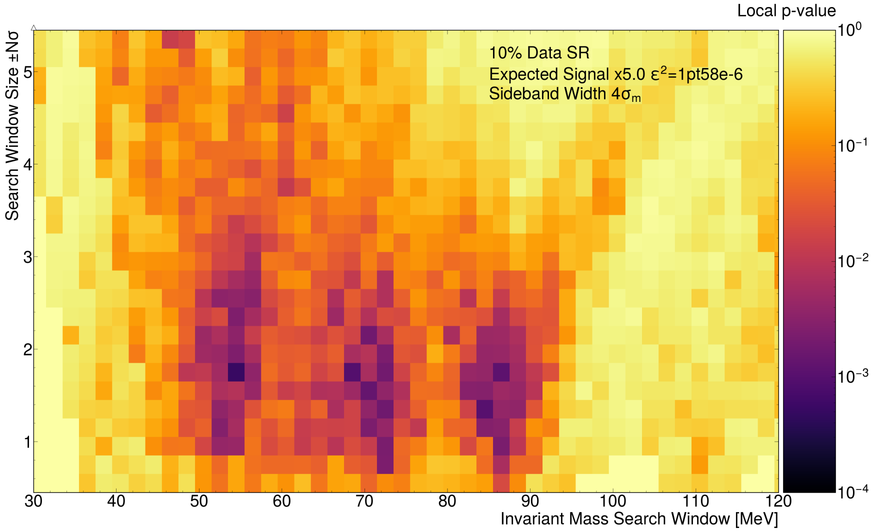
Figure 6.8: Expected background and local p-value for 10 % data SR with MC-injected signal. (a) Background estimate and observed events with 120 signal events injected at 64 MeV and $\epsilon^2 = 6 \times 10^{-6}$. (b) P-value for varying injected signal amounts, showing increasing significance with more injected signals. Error bars represent the number of toy MC trials.

6.2.0.2 Search Window Optimization Using MC-Signal Injections

The final search window size and tight selection were informed by studying their impact on the expected data significance with MC-signal injected. Figure 6.9 illustrates two examples of MC-signal injected local p-values for 10 % data as a function of the search window size. Each bin represents an independent test, where the signal was injected into the mass bin. The signal magnitude was scaled to five times the expected signal to illustrate a broad range of p-values. The significance remains relatively constant for search window sizes between $\pm\,(1-3)\,\sigma_m$, regardless of mass and ϵ. Based on these results and the studies of the background estimation, the search window size for the final analysis is chosen to be $\pm 1.5\sigma_m$, corresponding to an 86.6 % signal efficiency.



(a) Local p-values for MC-injected signal.



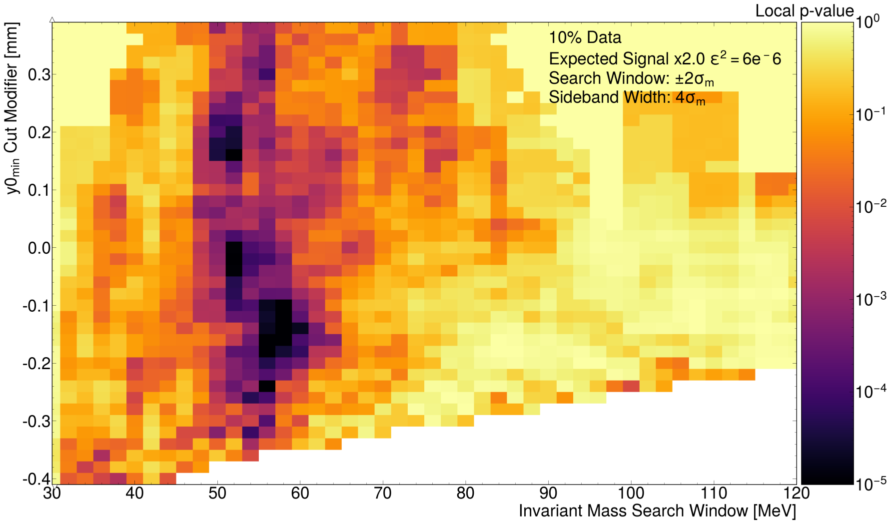
(b) Local p-values for MC-injected signal for different ϵ^2.

Figure 6.9: Local p-values for 10 % data with MC-injected signal, shown as a function of search window size, with the expected signal rate scaled up by a factor of five for better visibility. The signal is calculated for each mass independently using (a) $\epsilon^2 = 3e^{-6}$ and (b) $\epsilon^2 = 1.58e^{-6}$. The mass sideband width is fixed at $4\sigma_m$. The results show the effect of varying the search window size on local p-values.

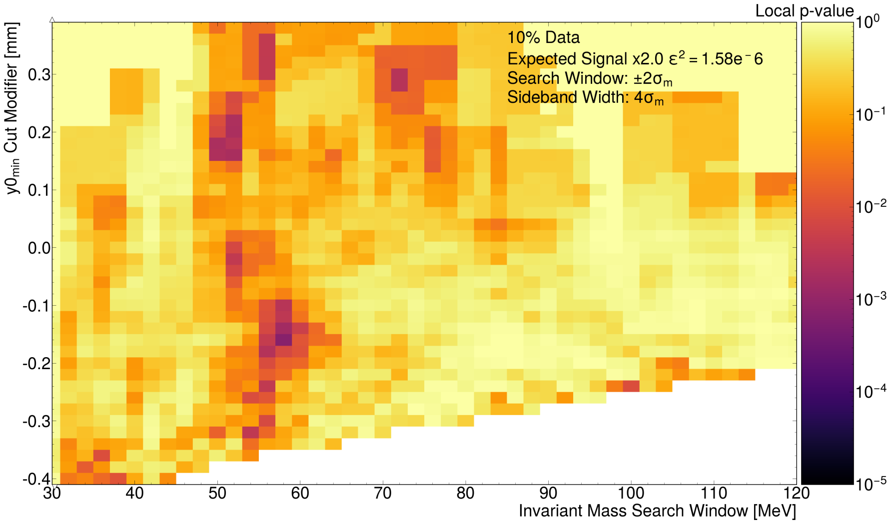
6.2.0.3 $y0^{\mathbf{min}}$ Optimization Using MC-Signal Injections

The MC-signal injection studies were also used to test the impact of tightening the $y0^{\min}$ cut on the expected significance to determine if the cut can be further tightened to protect against statistical fluctuations in the background that might harm the ability to exclude signal if none is found. The overall shape of the $y0^{\min}$ cut was previously optimized as a function of mass in Section 5.2.2.6. This shape is tightened/loosened by applying an additive "$y0^{\min}$ cut modifier" (in mm). Figure 6.10 shows the 10 % data MC-signal injected local p-values as a function of the $y0^{\min}$ cut modifier on the y-axis for two different values of ϵ^2. Again, the signal rate has been scaled up to produce a more dynamic result, this time by a factor of two. Loosening the cut ultimately reduces the signal sensitivity, particularly for masses greater than 75 MeV. However, slightly tightening the cut has a minimal impact on the significance.

Figure 5.35(b) demonstrates how tightening the $y0^{\min}$ cut by as little as +0.1 mm significantly reduces the background rate to 0-2 events in each search window (in 10 % data). This minimal tightening roughy maintains the nominal expected sensitivity while providing a buffer against large statistical fluctuations in the full data set, leading to a more conservative result. Thus, the ZBi optimized $y0^{\min}$ cut is tightened for all masses by +0.1 mm. This change was already introduced during the $y0^{\min}$ optimization description at the end of Section 5.2.2.6, and was included in the tight selection summary given in Table 5.5.



(a) Remaining background rate in 10 % data.



(b) Local p-value for MC-signal injected 10 % data.

Figure 6.10: Local p-values for 10 % data with MC-injected signal, shown as a function of $y0^{\mathrm{min}}$ cut modifier, with the expected signal rate scaled up by a factor of two for better visibility. The signal is calculated for each mass independently using (a) $\epsilon^2 = 6e^{-6}$ and (b) $\epsilon^2 = 1.58e^{-6}$. The mass sideband width is fixed at $4\sigma_m$. The results show that slightly tightening the cut does not significantly impact data significance.

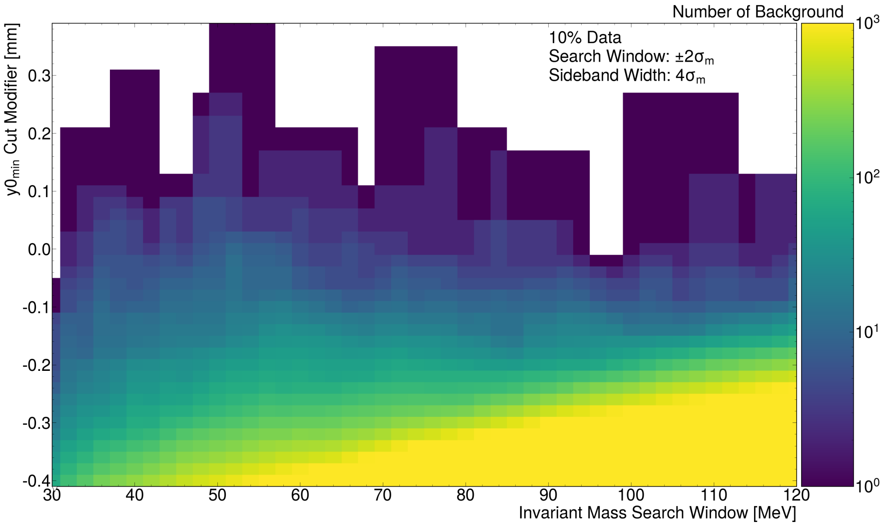


Figure 6.11: Background rate as a function of $y0^{\mathrm{min}}$ cut modifier for 10 % data. Slightly tightening the cut significantly reduces the expected background.

6.3 Setting Limits Using the Optimum Interval Method

If no evidence of signal is found in the data, the Optimum Interval Method (OIM) [69] can be used to set upper limits on the A' mass and coupling strength ϵ. The OIM is an extension of the "Maximum Gap Method" developed to set limits on dark matter experiments where the expected signal rate is small, the signal shape is known in one dimension, and there is a small unknown background source, well suited to this analysis. The signal shape for a specified mass in reconstructed vertex z is defined as:

$$S(m_{A'}, \epsilon, z) = \epsilon_{\text{bin}} \sum_i \left(\text{BR}(\rho_D) f_{\rho_D}(\epsilon, z_{\text{truth},i}) + \text{BR}(\phi_D) f_{\phi_D}(\epsilon, z_{\text{truth},i}) \right) \cdot \delta(z - z_i), \quad (6.10)$$

where:

- ϵ_{bin} represents the signal efficiency due to the search window size,
- $\text{BR}(\rho_D, \phi_D)$ is the branching ratio for $A' \to \rho_D, \phi_D + \pi_D$,
- $f_{\rho_D,\phi_D}(\epsilon, z_{\text{truth},i})$ is the acceptance times efficiency for ρ_D and ϕ_D (with all selections applied), as a function of the true vertex z, defined in Equation (5.16).

The Maximum Gap Method compares the data reconstructed vertex z distribution with the expected signal shape and searches for the maximum gap X between any two events in data, where

$$\text{X} = \max \left(\int_{z_i}^{z_{i+1}} S(m_{A'}, \epsilon, z)) \, dz \right). \quad (6.11)$$

The region between the two events that create the maximum gap corresponds to the zero-background signal region with the highest normalized expected signal rate. OIM uses the same maximum gap test, but performs the test k times, each time allowing for k events in the maximum gap, and sets the limit at the smallest value of μ that is rejected with 90 % confidence for all k. The data in reconstructed z is transformed to a normalized uniform distribution according to the signal shape by

$$x_i = \int_{-\infty}^{\infty} S(m_{A'}, \epsilon, z)\, dz - \int_{-z_i}^{\infty} S(m_{A'}, \epsilon, z)\, dz, \tag{6.12}$$

and the maximum gap size, with k events in the gap, is defined as

$$\mathrm{X}(k) = x_{i+k+1} - x_i. \tag{6.13}$$

After transforming the data, the confidence level for excluding the expected signal with mean μ can be calculated using toy MC simulations.

The expected distribution of maximum gaps, assuming the signal hypothesis μ is true, denoted as $f(\mathrm{X}(k)|\mu)$, is constructed by generating n ttoy experiments each with m events. Here, m is sampled from a Poisson with mean μ, and each event is uniformly distributed between 0 and 1. The maximum gap for each trial is computed using Equation (6.13). The signal rate μ is rejected with the maximum achieved confidence level for all values of k, given by:

$$C_k = \frac{\int_{\mathrm{X}(k)_{\mathrm{data}}}^{1} f(\mathrm{X}(k)|\mu)\, d\mathrm{X}(k)}{\int_0^1 f(\mathrm{X}(k)|\mu)\, d\mathrm{X}(k)} \tag{6.14}$$

6.4 Displaced Vertex Search Results

This section presents the results of the HPS SIMPs displaced vertex search, in the L1L1 hit-category, using the full 2016 Engineering Run data ($10\,703.81\,\mathrm{nb}^{-1}$). The top of Figure 6.12 compares the estimated background with the observed number of events in each search window, with all selections applied. The corresponding p-values are displayed in the bottom plot of the figure. Overall, the background estimation successfully describes the observed number of events, with a few insignificant fluctuations present across the search windows. Between $\sim 85\,\mathrm{MeV}$ to $110\,\mathrm{MeV}$ the background is systematically overestimated due to the poor mass resolution, approximately $4\,\mathrm{MeV}$, which results in the mass sidebands extending into regions of higher background. The maximum global significance of the data is approximately 0.9σ, within the $119\,\mathrm{MeV}$ search window. These results do not indicate any significant evidence of SIMP signal in the 2016 data set.

Given the lack of evidence for signal, OIM was used to set a $90\,\%$ confidence upper limit on the SIMP signal rate as a function of $m_{A'}$ (rather than m_{V_D}) and ϵ. The following results are presented before applying systematic uncertainties to the expected signal calculation. The final results with systematics included are presented in the next section. The expected signal rate is shown in Figure 6.13, along with the $90\,\%$ upper limit on the signal rate calculated using Optimum Interval Method (OIM). The $90\,\%$ upper limit exclusion contour in the $m_{A'}, \epsilon$ parameter space is presented in Figure 6.14. The SIMPs signal hypothesis is excluded with $90\,\%$ confidence where the ratio of the expected signal rate to the OIM-calculated upper limit is greater than or equal to 1.0, indicated by the red contour. Lastly, Figure 6.15 places the preliminary exclusion contour, without systematics

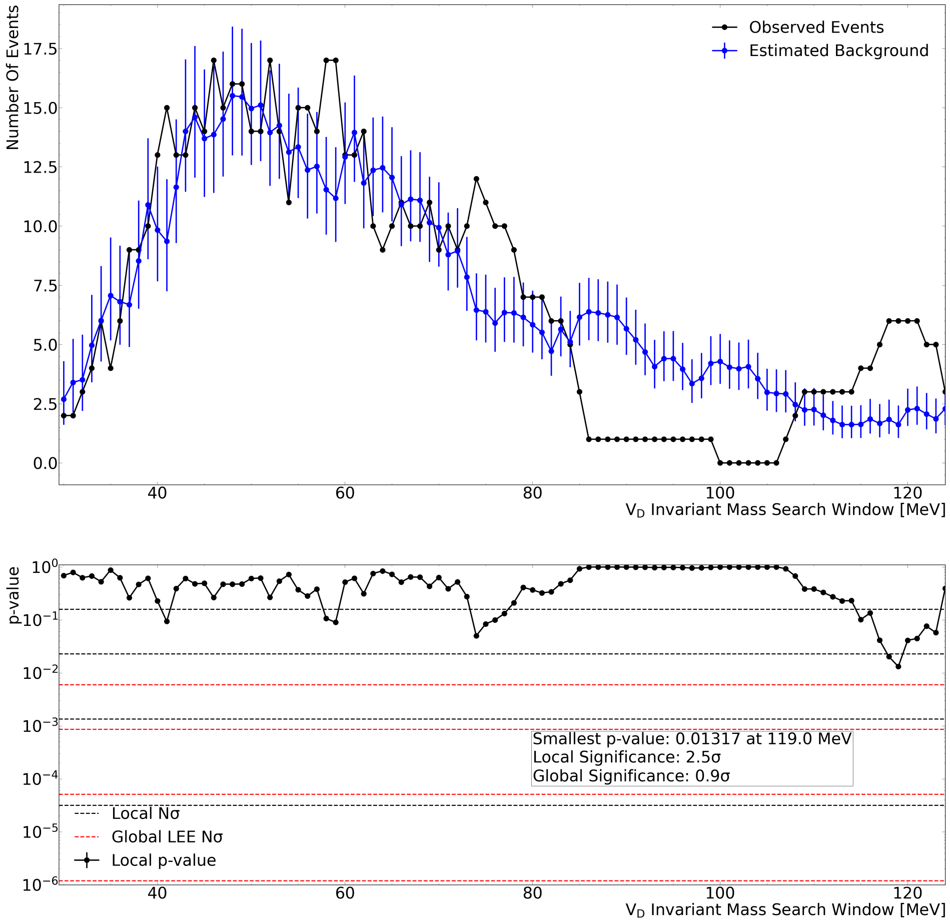
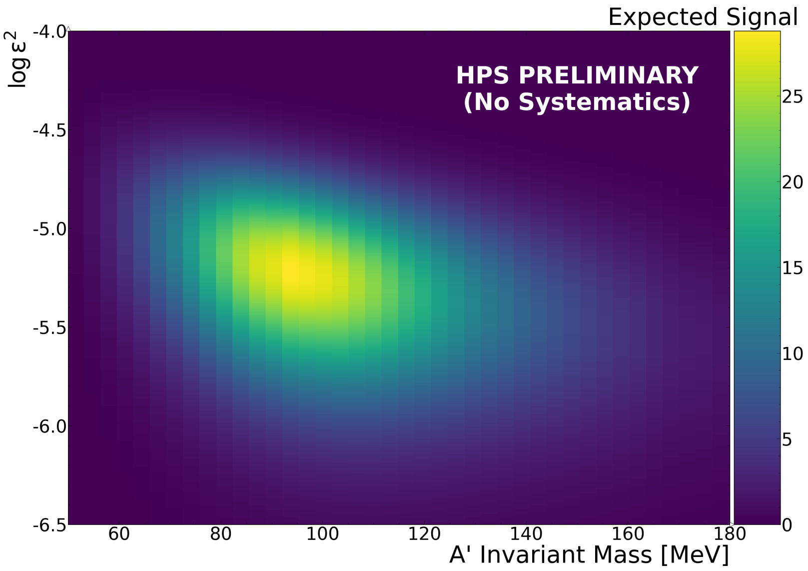
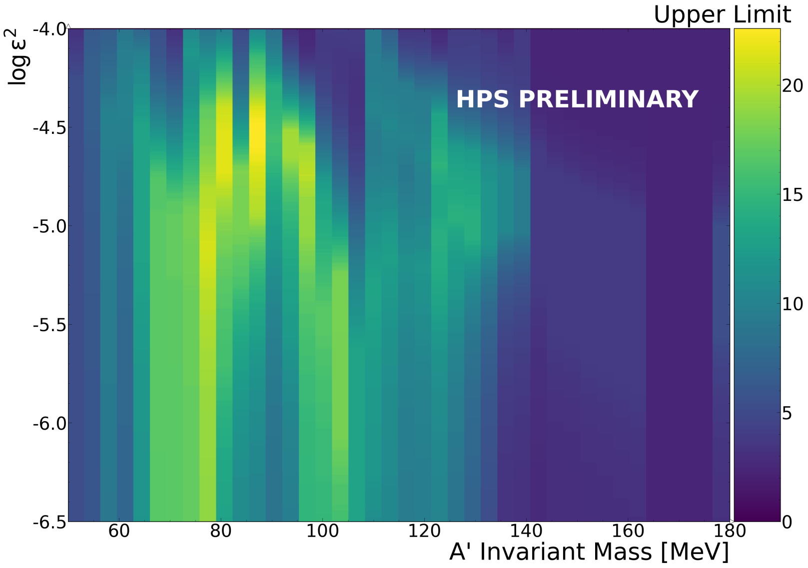


Figure 6.12: Search results for 100 % data in the SR after applying all final selections. (Top) Expected background (blue) and observed number of events (black) versus invariant mass search window size. (Bottom) Local p-values for each search window. The black dashed lines denote local significance thresholds, while the red dashed lines indicate global significance thresholds corrected for the LEE. The maximum global significance observed is 0.9σ at 119 MeV, with no significant evidence for signal found in the data.

included, on the plot of existing SIMP constraints for $m_{\pi_D}/f_{\pi_D} = 4\pi$ with dark vector visible decays restricted to two-body processes.



(a) Preliminary expected signal rate for 100 % data.



(b) 90 % upper limit on the signal rate.

Figure 6.13: Preliminary expected signal rate (top) and 90 % upper limit on the signal rate (bottom) for 100 % data. These results are presented before applying systematic uncertainties. (a) shows the expected signal rate, while (b) presents the 90 % confidence upper limit calculated using OIM.

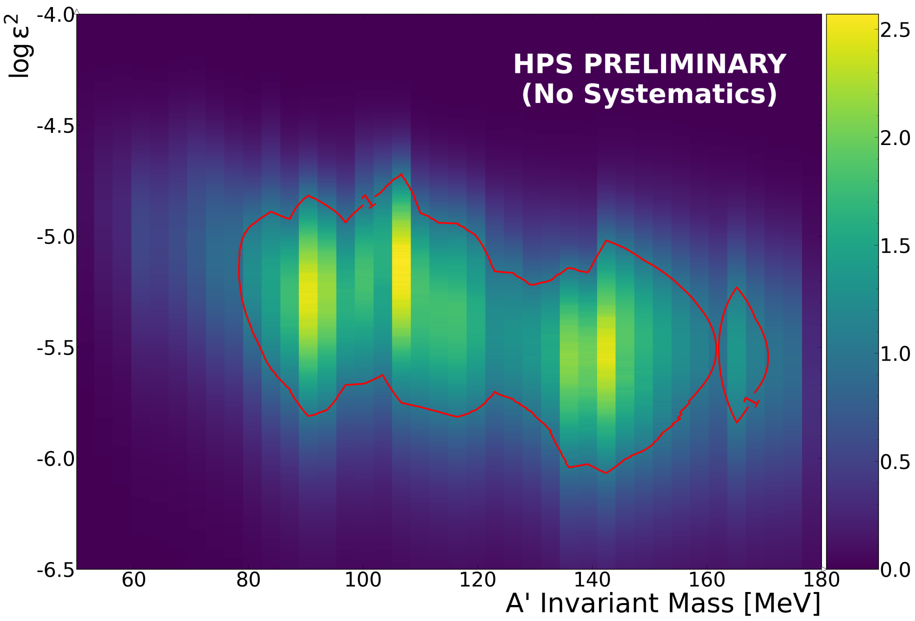


Figure 6.14: Preliminary (no systematics) 90 % confidence upper limit exclusion contour in the $m_{A'}$, ϵ parameter space for 100 % data. The exclusion contour (red) indicates regions where the ratio of the expected signal rate to the OIM-calculated upper limit is greater than or equal to 1.0, thus excluding the SIMPs signal hypothesis with 90 % confidence.

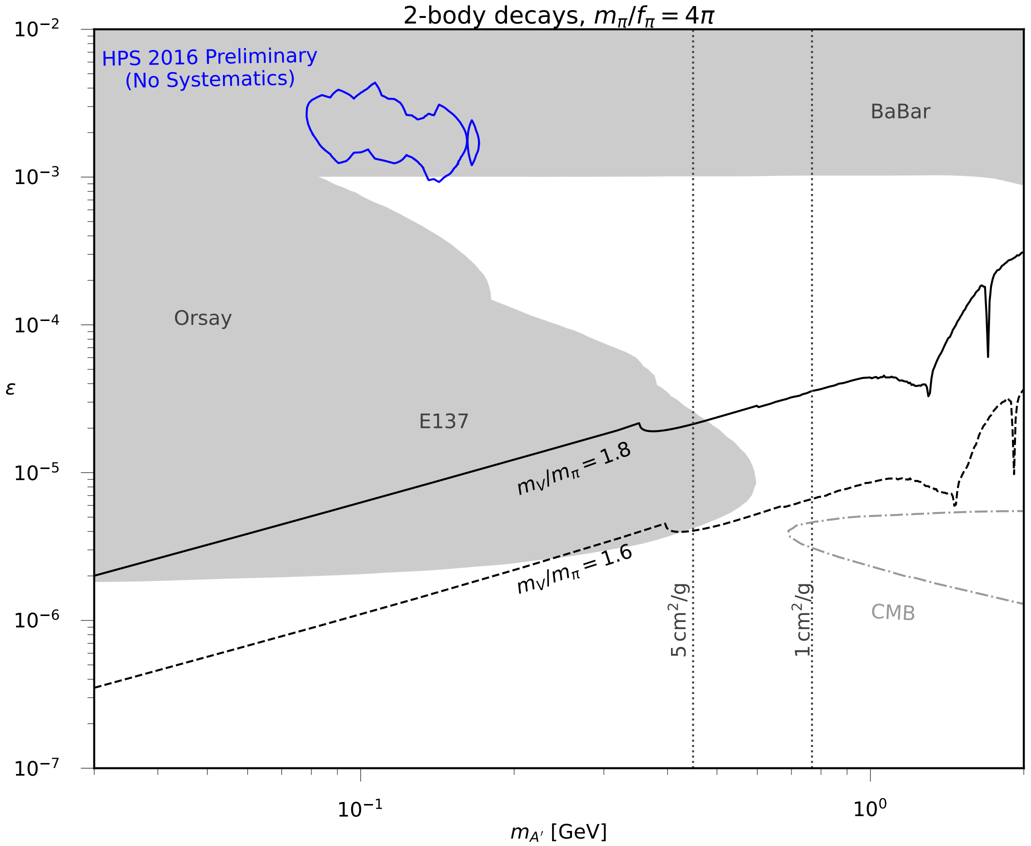


Figure 6.15: Preliminary 90 % confidence exclusion contour for 100 % data without including systematic uncertainties. The results are plotted against existing SIMP constraints for $m_{\pi_D}/f_{\pi_D} = 4\pi$, with dark vector visible decays restricted to two-body processes.

6.5 Systematic Uncertainties

The most significant expected contributions to the systematic uncertainty on the expected signal are discussed in this section.

6.5.1 MC Cross Section Uncertainty

The radiative fraction in Equation (5.8) is calculated purely using different MC samples, therefore the uncertainties in the MC cross sections will propagate into the radiative fraction. The radiative fraction in this analysis uses the same MC background generator level samples used in the 2016 A' analyses, so the cross section uncertainties are identical. The uncertainty in the radiative trident cross section is approximately 1 %, while the uncertainty in the tritrig cross section is significantly less than 1 %. In contrast, the uncertainty for the WAB cross section is approximately 19.8 % [68]. WABs constitute approximately 31.14 % of the background rate, which translates to a 6.2 % uncertainty on the radiative fraction denominator. If the radiative trident cross section uncertainty is taken to be 1 %, the total uncertainty on the radiative fraction is 6.28 %. This uncertainty on the radiative fraction is rounded up to 7 % . Additionally, the systematic uncertainty on the radiative acceptance due to the cross section uncertainty is taken to be 1 %.

6.5.2 Preselection Systematics

The preselection cuts summarized in Table 5.2 can result in different efficiencies between MC background and data, as illustrated in Table 5.3. The discrepancies in the cut efficiencies between MC and data reflect systematic uncertainties in the MC-based radiative acceptance term (Equation (5.10)) and signal acceptance-times-efficiency term

(Equation (5.14)), which impact the expected signal rate in Equation (5.17). The preselection efficiencies for MC tridents and MC signal are nearly identical to first order. Consequently, due to the inverse relationship between radiative acceptance and signal acceptance times efficiency, these systematic uncertainties largely cancel out. Additionally, the radiative fraction defined in Equation (5.8) is calculated using the ratio of preselection MC radiative tridents to preselection MC background, thus the preselection systematics also cancel for the radiative fraction term. The higher order systematic uncertainties are negligible compared to the uncertainty on the MC cross sections used to calculate the radiative fraction, and thus are not included.

6.5.3 Detector Misalignment

The MC simulations used to calculate the radiative trident and signal acceptance assume a perfectly aligned SVT detector. In reality, however, the alignment of the SVT, which consists of two separate volumes of six independently mounted silicon sensor modules, is not perfect. Thus, the discrepancy between the assumed detector alignment and the true physical alignment results in a systematic uncertainty on the detector acceptance in MC simulations, and must be accounted for in the final expected signal given by Equation (5.17). While the radiative fraction does not depend on the detector alignment because it is scaled to the reconstructed background rate in data, the radiative acceptance and the signal acceptance are impacted. The types of misalignments that primarily impact the acceptance are translations of the silicon sensors along the measurement direction (y in detector coordinates), translations along the beam direction (z), and rotations of the sensors about their normal. An initial mechanical survey on the SVT defines the sensor positions with a precision of 50 μm to 100 μm. This alignment is then refined offline using reconstructed tracks.

As described in Section 4.5, elastically scattered full-energy electrons and reconstructed Moller events provide standard candles for measuring the detector mass and momentum resolution. These events are also useful for validating the detector alignment by matching the reconstructed full-energy electron peak with the beam energy at 2.3 GeV, and aligning the reconstructed Moller peak with the predicted value of 48.45 MeV. The level of agreement between these predicted and reconstructed values, as shown in Section 4.5, demonstrates the proper alignment of the 2016 SVT detector.

In order to estimate the systematic uncertainty associated with reasonable uncertainties in this alignment, MC radiative tridents and SIMP signals, without electron beam background mixing, were produced using an intentionally misaligned version of the the detector geometry relative to the nominal geometry used in reconstruction. The silicon sensors were randomly misaligned via translations perpendicular to the silicon strips (Tu) and rotations around the sensor normal (Rw). The Tu (Rw) misalignments were generated by drawing random samples from a Gaussian distribution with $\mu = 0$ and $\sigma_{\mathrm{Tu}} = 10\,\mu\mathrm{m}$ ($\sigma_{\mathrm{Rw}} = 0.5\,\mathrm{mrad}$).

Figure 6.16 compares the radiative trident Psum distributions between the nominal detector geometry and the misaligned version. The observed shift in the radiative peak near 2.3 GeV is significantly larger than any reasonable misalignment would produce, indicating that the misalignment factors are too large. In addition, the track and vertex chi-squared numbers were significantly increased, leading to a massive artificial loss in selection efficiency. Based on these observations, this version of the misaligned detector is unrealistic and not valid. However, there was not enough time to generate new MC samples using more appropriate misalignment factors, so this study is completed here with this overly-misaligned detector in anticipation of forthcoming results.

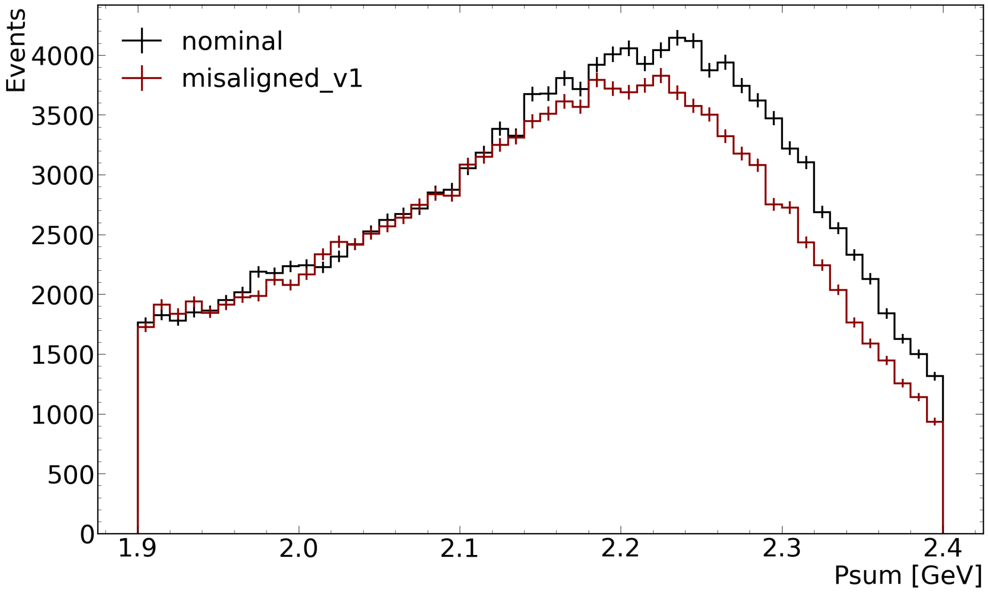


Figure 6.16: Comparison of radiative trident Psum distributions between the nominal detector geometry and a misaligned version. The significant shift in the radiative peak near 2.3 GeV, in addition to degraded track and vertex chi-squared values, indicates that the misalignment factors are too extreme.

The change in the radiative acceptance between the nominal and misaligned detector using radiative trident MC with no beam mixing is shown in Figure 6.17. If the assumed detector geometry used in the MC simulations has a smaller radiative trident acceptance than the real detector, the radiative acceptance term measured using MC is underestimated. Since this term is in the denominator of the expected signal calculation, this leads to overestimating the expected signal rate in data. However, the loss in acceptance also impacts the signal acceptance in the numerator of the expected signal calculation, so the alignment induced effects are expected to cancel to some degree.

Section 6.5.3 demonstrates the change in the expected signal rate associated with the alignment uncertainty, showing the ratio between the misaligned and nominal detector. The radiative acceptance used to calculate the expected signal in each case is taken from the corresponding 7th-order polynomial fits shown in Figure 6.16, with the MC signal simulated in 10 MeV intervals. Theoretically, the ratios presented would be applied to the final expected signal rate as a function of mass and ϵ. However, again, this misaligned detector is severely misaligned and does not reflect the expected systematic uncertainty, and an improved version is currently being processed.

6.5.4 Target Position Uncertainty

The target position, defined as $-4.3\,\mathrm{mm}$ along the $\hat{z}$ direction in detector coordinates, is known with an uncertainty of $\pm 0.5\,\mathrm{mm}$. This uncertainty contributes to the systematic uncertainty on the expected signal rate due to its effect on the acceptance for both signal and radiative tridents. To account for this, MC simulations of radiative tridents and signal events were generated for two alternative target positions at $-4.8\,\mathrm{mm}$ and $-3.8\,\mathrm{mm}$, corresponding to the uncertainty.

6.5.4.1 Radiative Acceptance

Section 6.5.4.1 shows the comparison between these two alternative target positions relative to the nominal position at $-4.3\,\mathrm{mm}$. For each case, the radiative acceptance was calculated and fitted with a 7-th order polynomial, as shown in the top part of the figure. The discrepancy between each off-nominal target position and the nominal position

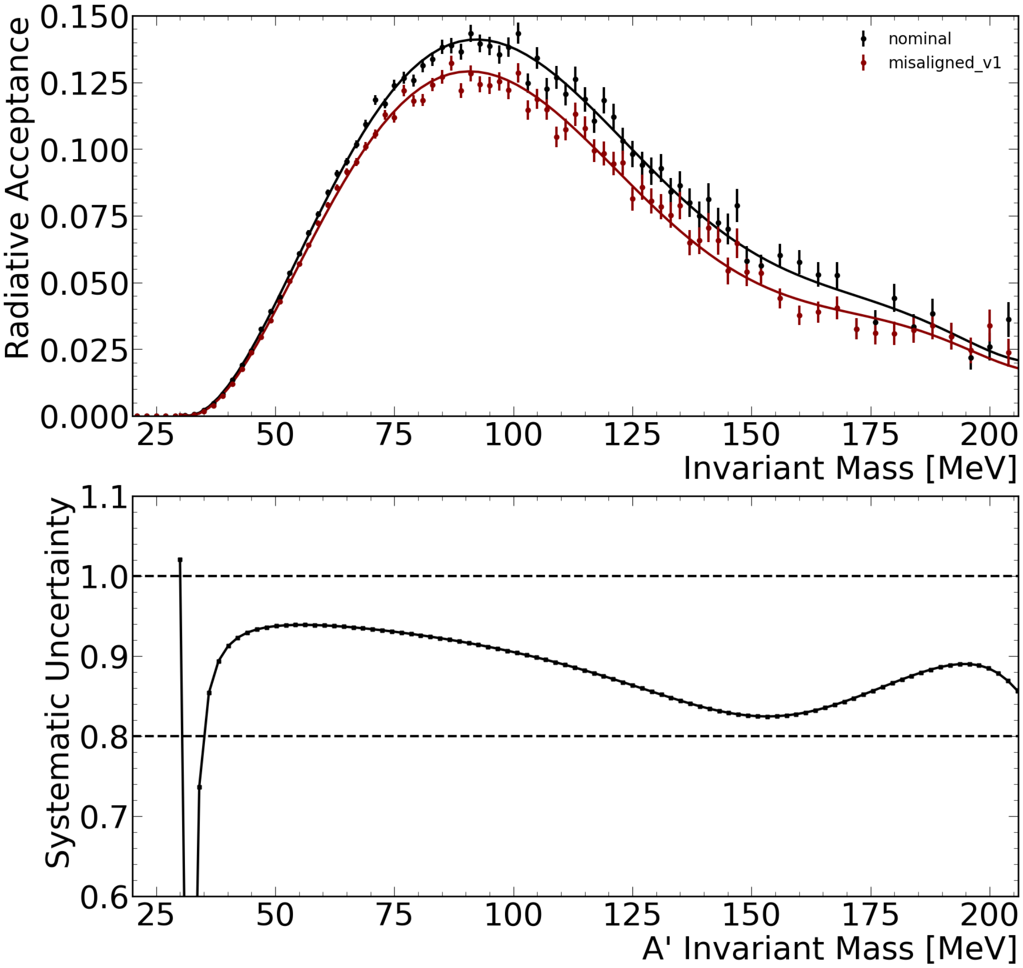


Figure 6.17: Comparison of radiative acceptance (no beam) between the nominal and misaligned detectors. (Top) Shows the radiative acceptance, fitted with a 7th-order polynomial, for both the nominal (black) and misaligned (red) detectors. (Bottom) Shows the ratio of radiative acceptance for the misaligned detector to the nominal detector. If this ratio is less than one, it indicates a need to apply a systematic uncertainty to the expected signal.

was evaluated by calculating the ratio of the off-nominal fit function to the nominal fit. Between 50 MeV to 175 MeV, the radiative acceptance is higher for the off-nominal target positions compare to the nominal. Beyond 175 MeV, the acceptance drops to around 85 %,

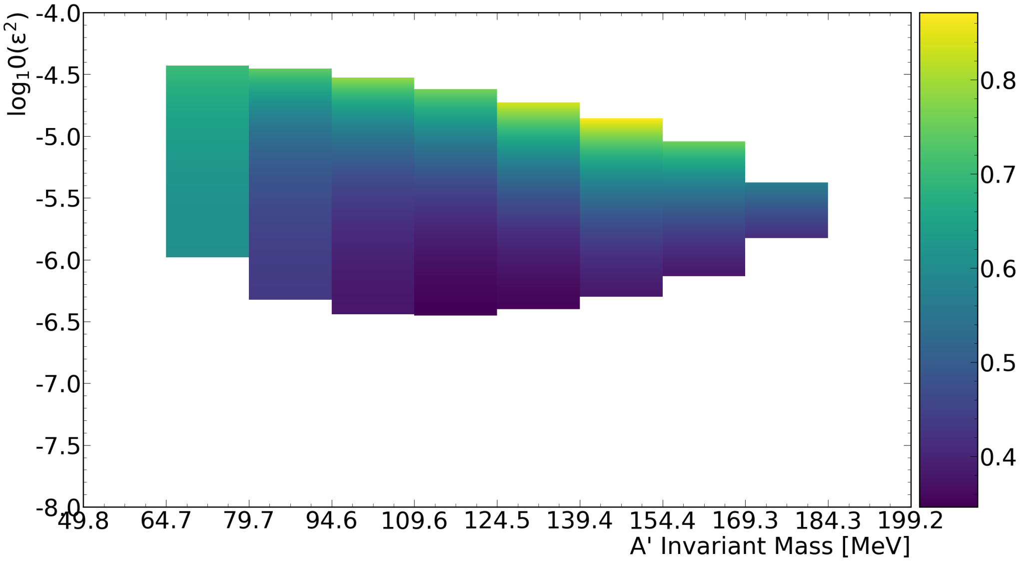


Figure 6.18: Ratios of the expected signal rate between the nominal and misaligned detectors, where the expected signal for each case is calculated using the radiative acceptance fits in Figure 6.17. The ratio represents the systematic uncertainty in each bin, where values less than one are applied to the expected signal. However, this version of the misaligned detector is known to be too extreme, and these results are not valid.

although the fits in this region are less reliable due to low statistics in the off-nominal simulations. If the true target position were one of the off-nominal values, the actual acceptance would be higher than assumed in the nominal radiative acceptance. Since the expected signal is inversely proportional to the radiative acceptance, this means that the nominal calculation would overestimate the signal rate. Thus, the mass-dependent ratio of the fitted off-nominal acceptance to the nominal acceptance is used to measure this overestimation. The maximum of the two ratios at each mass is used to calculate the systematic uncertainty, defined by:

$$\sigma_{\text{target z}}(m_{A'}) = 1 - \frac{1}{\max\big(r^{-}(m_{A'}), r^{+}(m_{A'})\big)} \tag{6.15}$$

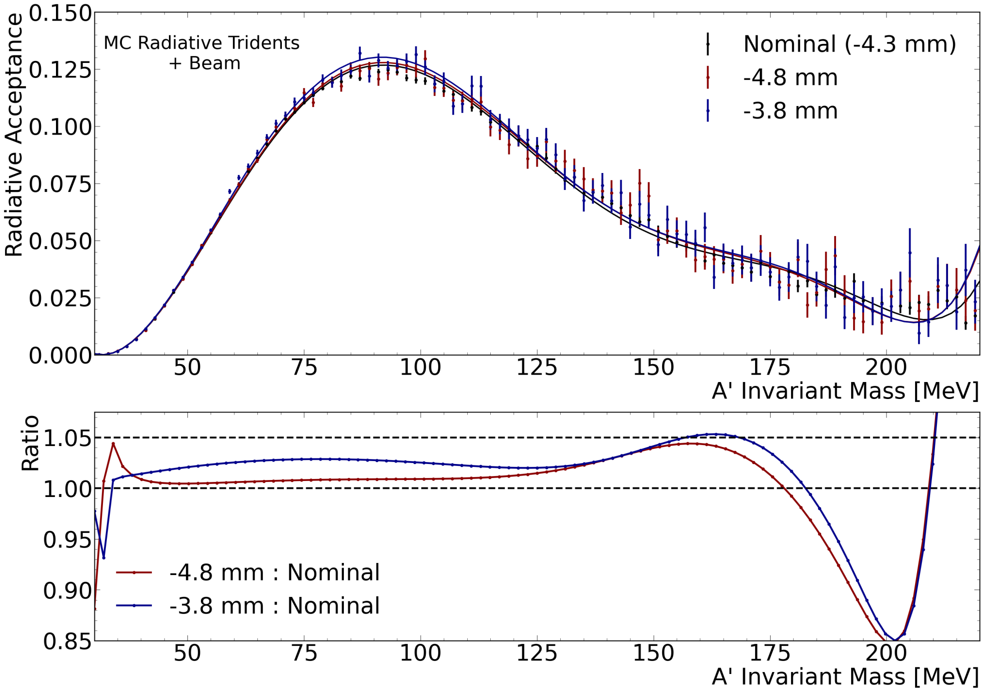
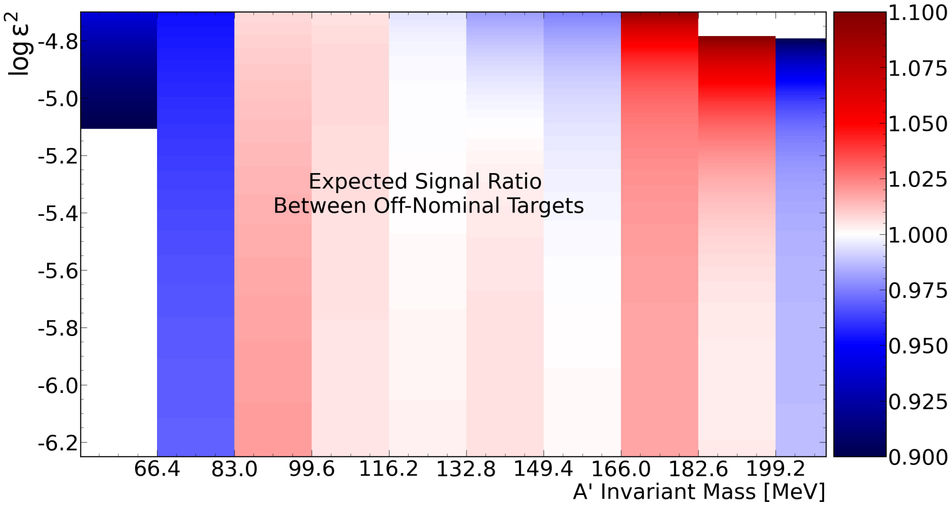


Figure 6.19: Comparison of the radiative acceptance (with beam) between the nominal and shifted target positions. (Top) Shows the radiative acceptance, fit with a 7th-order polynomial, for the nominal target position at $-4.3\,\mathrm{mm}$ (black), $-4.8\,\mathrm{mm}$ (red), and $-3.8\,\mathrm{mm}$ (blue). (Bottom) illustrates the ratios of the radiative acceptance between the off-nominal and nominal target positions. Ratios greater than one indicate the need to apply a systematic uncertainty.

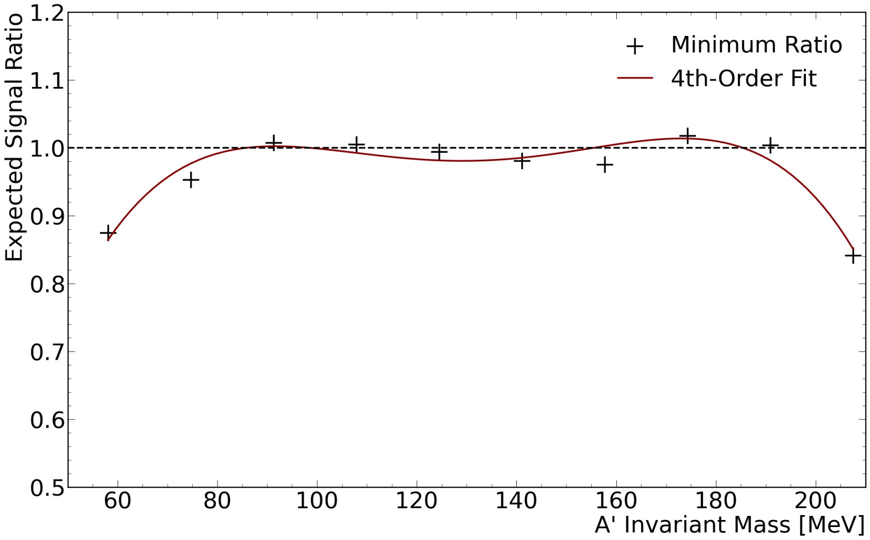
6.5.4.2 Signal Acceptance

Figure 6.20(a) shows the ratio of the expected signal rate between the two off-nominal targets as a function of $m_{A'}$ and $\log\epsilon^2$, representing a conservative measurement of the systematic uncertainty of the signal acceptance resulting from the target position uncertainty. The analysis only excludes SIMPs in the parameter space within $-6.2 < \log\epsilon^2 < -4.7$, so the systematic uncertainty measurement only considers the results in this region. Figure 6.20(b) shows the minimum ratios across all values of $\log\epsilon^2$ for each

MC-signal mass. The minimum ratio is taken to ensure a conservative estimate of the systematic uncertainty. The ratios are fitted with a 4th-order polynomial, parameterizing the uncertainty as a function of mass. If the fit value is greater than one, the systematic uncertainty is set to zero, since it represents underestimating the signal in the analysis.



(a) Ratios of expected signal between off-nominal targets.



(b) Minimum ratios fit with a 4th-order polynomial

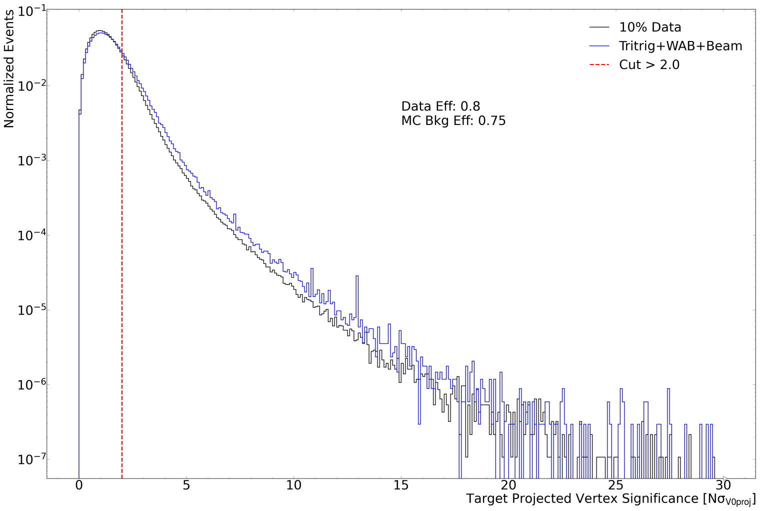
Figure 6.20: Signal acceptance systematic uncertainty due to the target position uncertainty. (a) Shows the ratio of the expected signal rate between the two off-nominal targets ($-4.8\,\mathrm{mm}$ and $-3.8\,\mathrm{mm}$. (b) Shows the minimum ratios for each mass bin in the $\log \epsilon^2$ range fitted with a 4th-order polynomial as a function of mass.

6.5.5 High-z Cuts

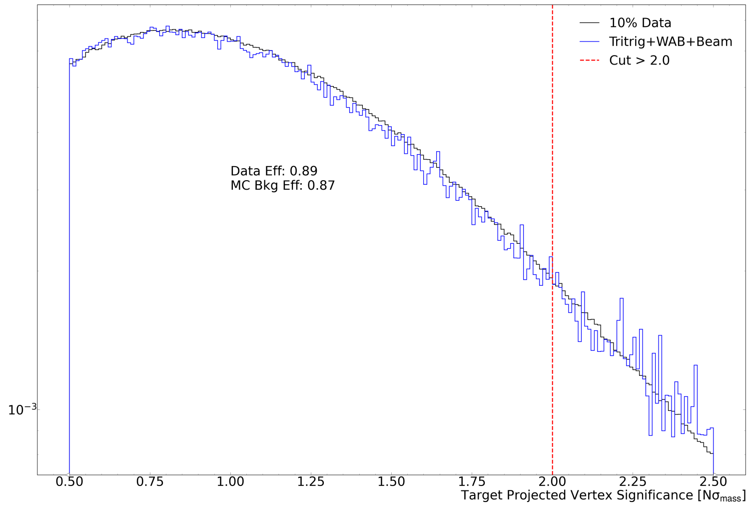
Since the signal selection efficiency is estimate using MC simulations, any differences in the high-$z$cut efficiencies between MC background and data can result in a systematic uncertainty on the expected signal.

6.5.5.1 Target Projected Vertex Significance Systematic ($\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$)

The first high-$z$cut, $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$, projects vertices back to the target and measures their level of significance relative to the fitted beamspot, which differs between MC and data. In particular, the MC beamspot is significantly narrower in both x and y compared to data, as shown in Section 5.2.2.4. The impact of these differences on the relative efficiency is minimized by cutting on the significance relative to the fit. However, since the vertices are projected back to the target location in $\mathrm{z}_{\mathrm{vtx}}$ using the reconstructed vertex momentum, differences are still observed. Figure 6.21(a) compares the MC background and data efficiency for the $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}} < 2.0$ selection, showing a 5 % lower efficiency in the MC background compared to data. Since the efficiency is lower in MC than data, this difference does not contribute to the systematic uncertainty on the expected signal rate, as the signal rate is already being underestimated using MC simulations. In case the $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ cut efficiency is a function of Psum, the MC background and data efficiencies were also compared in the CR at high Psum, shown in Figure 6.21(b). The cut efficiencies in this region agree to within 2 %, but the MC efficiency is still lower than data. Given these results, no systematic uncertainty associated with the $\mathrm{N}\sigma_{\mathrm{V0}_{\mathrm{proj}}}$ cut is applied to the expected signal rate.



(a) Low momentum signal region.



(b) High momentum control region.

Figure 6.21: Comparison of the target projected vertex significance distribution between MC background and data. (a) Shows the distribution in the signal region, and (b) shows the high momentum control region. The MC background cut efficiency is lower than data and does not introduce a systematic uncertainty.

6.5.5.2 Minimum Track Vertical Impact Parameter Systematic ($y0$)

There is a potential systematic uncertainty associated with the $y0^{\mathrm{min}}$ cut due to differences between the $y0$ distributions in MC simulated signal and data. There are no displaced events in data that can be used to compare with the displaced MC signal. Instead, the uncertainty is estimated by comparing data with MC background. Section 6.5.5.2 shows the normalized $y0$ distributions for data and MC background fitted with a Gaussian function to measure the $y0$ resolutions. The MC background resolution is almost 8 % narrower than data, which could result in an overestimate of the expected signal efficiency using MC simulations. This discrepancy is accounted for by calculating a $y0$ smearing factor that aligns the MC background resolution with data, calculated using:

$$\sigma_{\mathrm{smear}} = \sqrt{\sigma^2_{\mathrm{data}} - \sigma^2_{\mathrm{MC\ bkg}}}. \tag{6.16}$$

After applying the smearing factor, the MC background fitted resolution is within 2 % of the data, as shown in the bottom of Section 6.5.5.2, demonstrating the effectiveness of the smearing factor. Meanwhile, the signal efficiency changed by much less than 1 % for all masses. Therefore, the systematic uncertainty associated with the different $y0$ resolutions in MC simulations and data is negligible, and not included.

6.5.6 MC Mass Resolution

If the MC-derived signal mass resolution was narrower than it actually is in data, the signal efficiency in the invariant mass search windows would be overestimated, resulting in an inflated expected signal. The difference in the mass resolution between MC simulations and data was calculated by fitting the reconstructed Moller peak for using MC

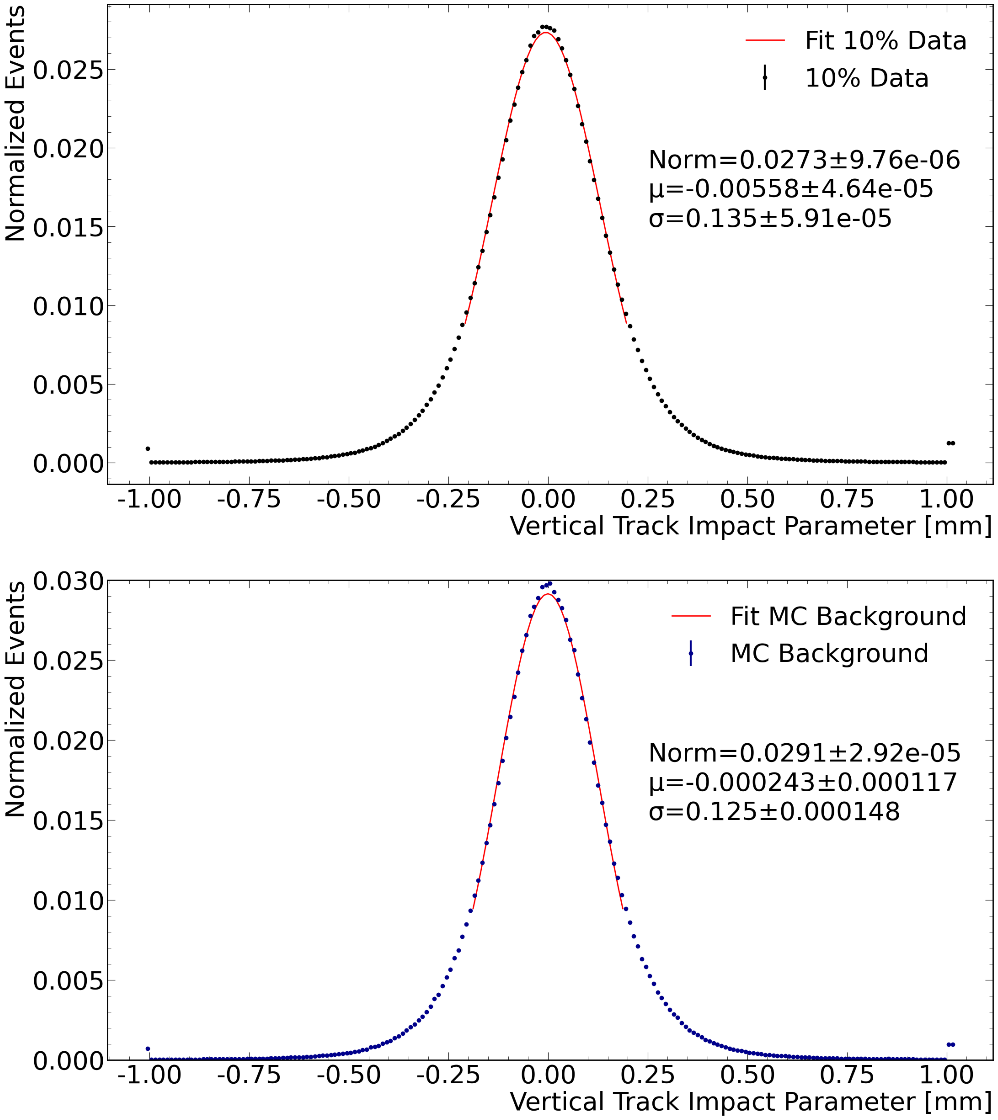


Figure 6.22: Comparison of the track vertical impact parameter distribution ($y0$) between MC background and data. (Top) 10 % Data. (Bottom) MC background. The distributions are fitted with Gaussian functions to compare $y0$ resolutions, slightly higher for the MC background. This introduces a systematic uncertainty on the $y0^{\text{min}}$ cut.

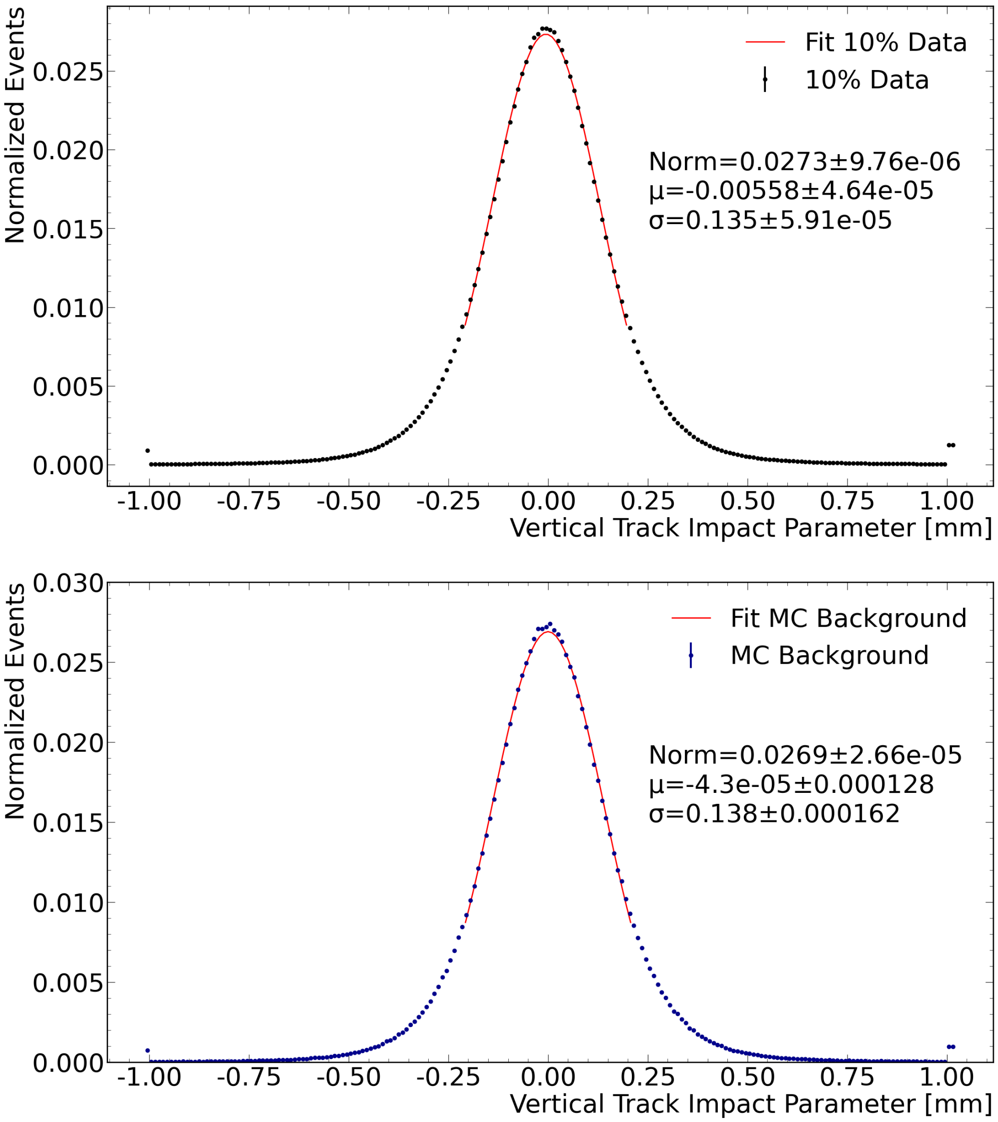


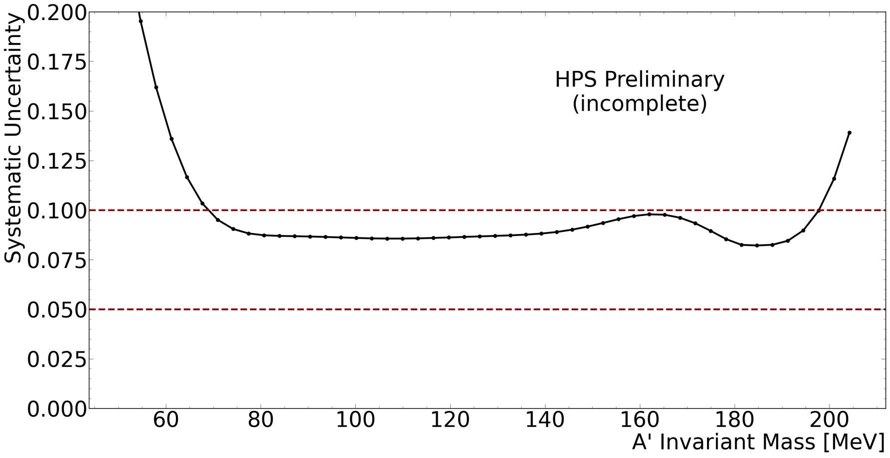
Figure 6.23: Comparison of the track vertical impact parameter distribution ($y0$) between MC background and data after smearing the MC to match data. (Top) 10 % Data. (Bottom) Smeared MC background.

and data, as shown in Section 4.5. The mass resolution for data Mollers was measured to be 2.54 MeV, compared to 2.32 MeV for MC, thus the MC mass resolution is 8.7 % higher (better) than data. This discrepancy is corrected for by calculating the change in the signal efficiency assuming the degraded mass resolution, resulting in a systematic uncertainty of 4.3 %.

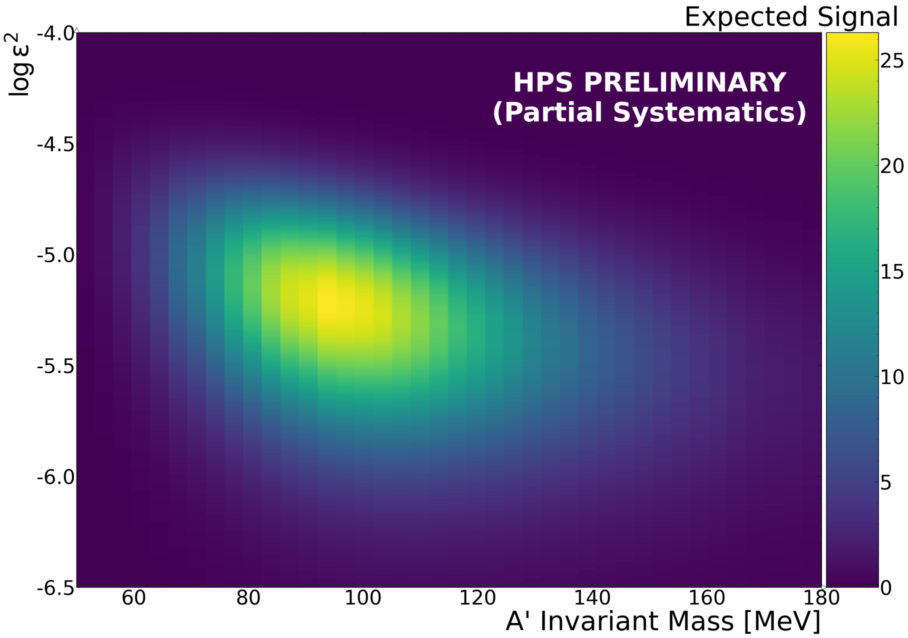
6.5.7 Result Including Systematics Uncertainties

Figure 6.24(a) presents a summary of the partial systematic uncertainties on the expected signal rate as a function of mass. Note that systematic uncertainties related to the detector misalignment are not yet included, but will be incorporated in the near future. Figure 6.24(b) shows the expected signal rate, reweighted to account for the systematics thus far included. The updated exclusion contour, derived by calculating the ratio between the reweighted signal and the OIM-calculated upper limits discussed in Section 6.4, is shown in Figure 6.25(a). The exclusion contours before and after accounting for systematic uncertainties are compared in Figure 6.25(b), demonstrating a slight reduction in the exclusion.

Finally, Figure 6.26 places the preliminary exclusion contour, with the partial set of systematic uncertainties included, on the plot of existing SIMP constraints for $m_{\pi_D}/f_{\pi_D} = 4\pi$ with dark vector visible decays restricted to two-body processes.

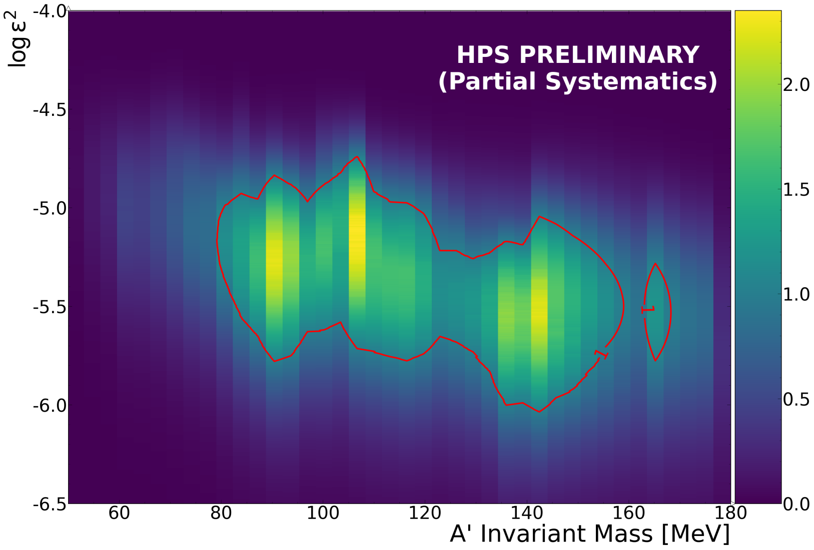


(a) Summary of systematic uncertainties on the expected signal rate.

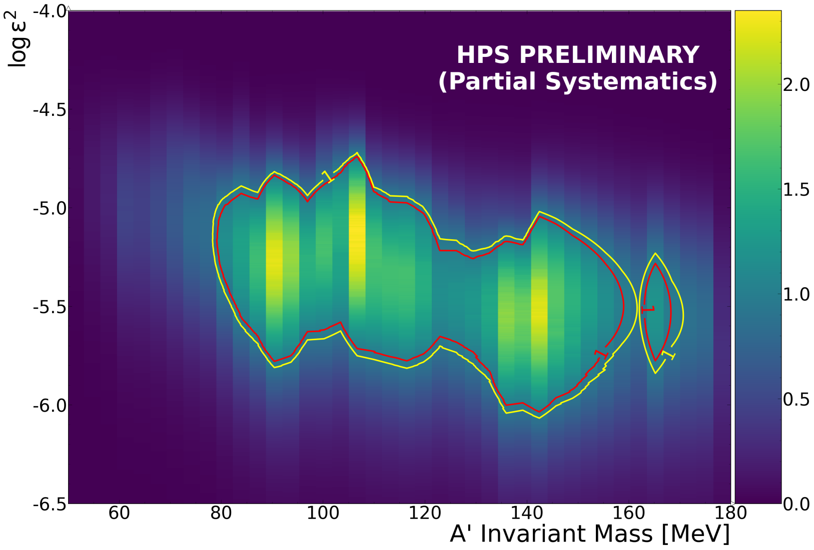


(b) Preliminary expected signal rate for 100 % data.

Figure 6.24: Overview of the partial systematic uncertainties and their impact on the expected signal rate. (A) Shows a summary of the uncertainties as a function of mass. (B) Shows the expected signal rate with the uncertainties included. The detector misalignment uncertainty is not yet included in the result.



(a)



(b) 90 % upper limit on the signal rate.

Figure 6.25: Preliminary 90 % confidence exclusion with partial systematic uncertainties on the expected signal rate included–the detector misalignment contribution is not yet included. (a) Shows the updated contour, while (b) compares the contours before (yellow) and after (red) including the uncertainties.

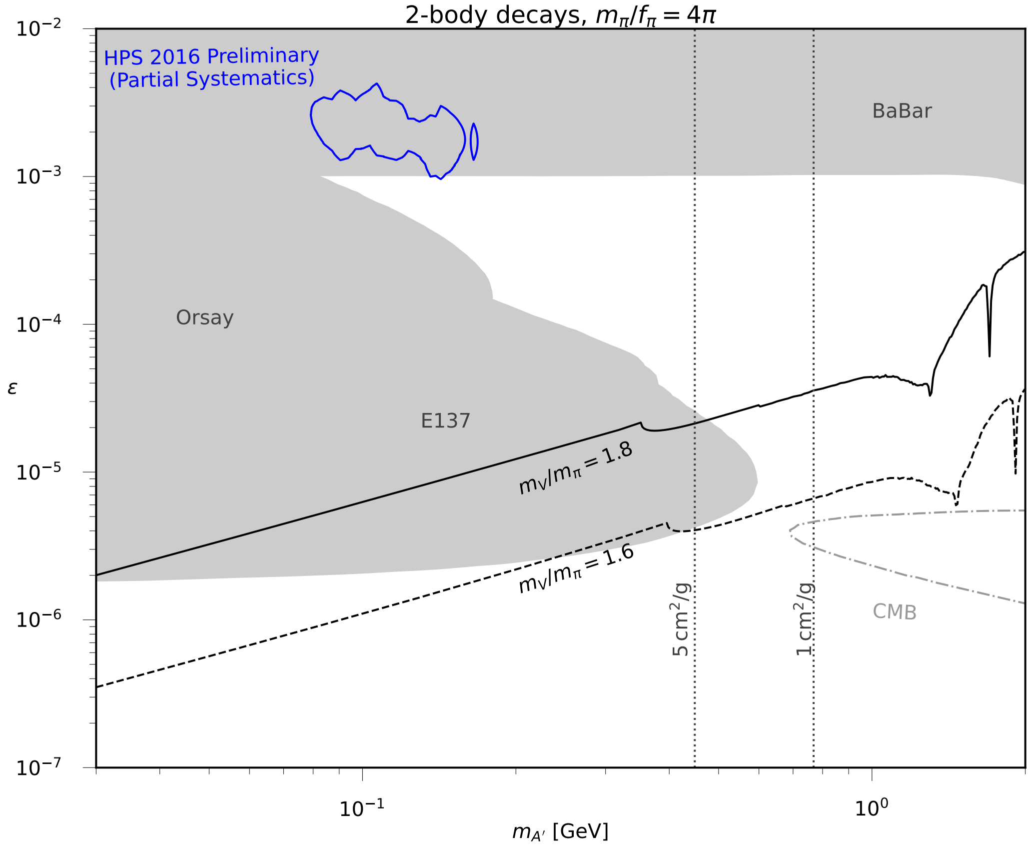


Figure 6.26: Preliminary 90 % confidence exclusion contour for 100 % data after accounting for most of the systematic uncertainties on the expected signal. The results are plotted against existing SIMP constraints for $m_{\pi_D}/f_{\pi_D} = 4\pi$, with dark vector visible decays restricted to two-body processes.

Chapter 7

Conclusion

The Heavy Photon Search (HPS) is a fixed target experiment hosted at the Thomas Jefferson National Lab designed to search for heavy photon (A') hidden sector (HS), also referred to as "dark" sectors. The A' arises from an additional $U(1)_D$ Abelian gauge symmetry, kinetically mixing with the SM photon and providing a portal into the HS. The A' framework accommodates sub-GeV thermal-relic Dark Matter (DM) and introduces new mechanisms allowing DM self-interactions compatible with cosmological and astrophysical observations. This dissertation focused on a specific extension of the minimal A' HS including Strongly-Interacting Massive Particles (SIMPs) arising from an additional QCD-like spontaneously broken $SU(3)_D$ symmetry. The SIMP HS contains dark vector mesons (V_D) and dark pions (π_D), analogous to the SM, where the dark pions make up dark matter. These HS particles introduce new thermal DM freeze-out mechanisms through $3\pi_D \rightarrow 2\pi_D$ annihilations and dark vector semi-annihilation to the SM through a virtual A'.

The HPS SIMP search utilizes a continuous electron beam incident on a thin tungsten foil to generate heavy photons at the target that can promptly decay to HS mesons through the process $A' \rightarrow \pi_D V_D$. While the π_D are invisible, the V_D are naturally long-lived and visibly decay to an e^+e^- pair (for $m_{V_D} < 2m_\mu$). HPS searches for these visible decays utilizing a compact, forward acceptance spectrometer for masses between 10 MeV to 220 MeV and kinetic mixing strength $\epsilon^2 \sim 10^{-2} - 10^{-10}$. For large values of ϵ^2, HPS can conduct a mass resonance search for V_Ds over the prompt QED background. For V_Ds with longer lifetimes (smaller values of ϵ^2), HPS uses a displaced vertex search to identify signal downstream of the target, beyond the prompt QED background.

This dissertation detailed the results of the SIMPs displaced vertex search using the 2016 Engineering Run luminosity ($10\,753\,\mathrm{nb}^{-1}$). The search focused on the V_D mass range between 30 MeV to 124 MeV, where the signal was expected to be most significant. Utilizing both MC simulations and data, a series of selection stages were developed to create a sample of high-quality reconstructed background events used to calibrate the analysis, optimize the signal region, and interpret the search results. The strongest evidence for signal was observed as a local p-value of 0.01317 at $m_A' = 119\,\mathrm{MeV}$, corresponding to a global significance of 0.9σ. Although the search did not find evidence of a signal, it successfully set a 90 % confidence upper limit across a broad range of the $m_{A'}, \epsilon$ parameter space. Not only does this analysis contribute to constraints on SIMP hidden sector models by reinforcing existing constraints derived from different signals, it utilized a new approach to the HPS displaced vertex search that could drive currently developing minimal A' searches using the 2019 and 2021 HPS data sets, making them even more competitive. These findings also motivate further investigation into the L1L2 and L2L2 analysis hit-categories in the 2016 Engineering Data, which could enhance sensitivity to longer-lived dark vectors and extend the limits established here into ground-breaking territory. Lastly, this analysis have motivated HPS to make SIMP searches a standard part of data analysis going forward, and these results will be greatly extended into new territory by the 2019 and 2021 data sets.

Bibliography

[1] Nathan Baltzell et al. The Heavy Photon Search Experimen, 2022. arXiv:2203.08234 [hep-ex], URL: `https://arxiv.org/abs/2203.08324`.

[2] A Candelori, R Bainbridge, Pierre Barrillon, Dario Bisello, Marcus French, G Hall, A Kaminski, V Khomenkov, E Noah, Marthin Raymond, L Stefanutti, and M Tessaro. Radiation hardness qualification of the APV25 chip production for the CMS experiment.

[3] I. Balossino et al. The HPS electromagnetic calorimeter. 854:89–99, 2017. URL: `https://www.sciencedirect.com/science/article/pii/S0168900217302620`.

[4] M. J. French et al. Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker. 466(2):359–365, 2001. 4th Int. Symp. on Development and Application of Semiconductor Tracking Detectors. URL: `https://www.sciencedirect.com/science/article/pii/S0168900201005897`.

[5] G. Arnison et al. Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV. *Phys. Lett. B*, 122:103–116, 1983.

[6] M. Banner et al. Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider. *Phys. Lett. B*, 122:476–485, 1983.

[7] F. J. J. Abe et al. Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab. *Physical Review Letters*, 74(14):2626–2631, Apr 1995.

[8] J. E. Augustin et al. Discovery of a Narrow Resonance in e^+e^- Annihilation. *Phys. Rev. Lett.*, 33:1406–1408, 1974.

[9] G. Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, Sep 2012.

[10] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, Sep 2012.

[11] David Galbraith. Ux standard model of the standard model, 2023. URL: `https://davidgalbraith.org/portfolio/ux-standard-model-of-the-standard-model/`.

[12] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helv. Phys. Acta*, 6:110–127, 1933.

[13] V. C. Rubin and Jr. Ford, W. K. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *ApJ*, 159:379, Feb 1970.

[14] V. C. Rubin. Dark matter in spiral galaxies. *Scientific American*, 248:96–106, Jun 1983.

[15] Clowe, D. and Bradač, M. and Gonzalez, A. H. and Markevitch, M. and others. A Direct Empirical Proof of the Existence of Dark Matter. *The Astrophysical Journal*, 648(2):L109–L113, Aug 2006.

[16] Aghanim, N. and others. Planck2018 results: VI. Cosmological parameters. *Astronomy Astrophysics*, 641:A6, Sep 2020.

[17] D. J. Eisenstein et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *The Astrophysical Journal*, 633(2):560–574, Nov 2005.

[18] D. S. Akerib et al. Results from a search for dark matter in the complete LUX exposure. *Physical Review Letters*, 118(2):021303, 2017.

[19] E. Aprile et al. Dark Matter Search Results from a One Ton-Year Exposure of XENON1T. *Physical Review Letters*, 121(11):111302, 2018.

[20] P. Agnes et al. First results from the DarkSide-50 dark matter experiment at Laboratori Nazionali del Gran Sasso. *Physics Letters B*, 743:456–466, 2015.

[21] M. Ackermann et al. Search for dark matter satellites in the Fermi-LAT source catalog. *The Astrophysical Journal*, 747(2):121, 2012.

[22] O. Adriani et al. An anomalous positron abundance in cosmic rays with energies 1.5-100 GeV. *Nature*, 458(7238):607–609, 2009.

[23] The Fermi-LAT Collaboration. Observation of the high-energy gamma-ray emission towards the Galactic center. *The Astrophysical Journal*, 840(1):43, 2017.

[24] G. Aad et al. Search for dark matter in events with a hadronically decaying vector boson and missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Journal of High Energy Physics*, 2018(1):126, 2018.

[25] V. Khachatryan et al. Search for dark matter and unparticles produced in proton–proton collisions at $\sqrt{s} = 8$ TeV using the missing transverse energy distribution. *Physics Letters B*, 755:102–122, 2016.

[26] A. M. Sirunyan et al. Search for dark matter and unparticles in events with a Z boson and missing transverse momentum in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics*, 2017(3):61, 2017.

[27] E. W. Kolb and M. S. Turner. *The Early Universe.* Addison-Wesley, 1990.

[28] K. M. Lee and S. Weinberg. Cosmological lower bound on the mass of thermal dark matter particles. *Physical Review Letters*, 39(8):165–168, 1977.

[29] Bob Holdom. Two U(1)'s and charge shifts. 166(2):196–198, 1986.

[30] James D. Bjorken, Rouven Essig, Philip Schuster, and Natalia Toro. New fixed-target experiments to search for dark gauge forces. *Physical Review D*, 80(7):075018, Oct 2009.

[31] M. Boylan-Kolchin, V. Springel, S. D. M. White, A. Jenkins, and G. Lemson. The large-scale structure of the Universe and the small-scale problems of Lambda cold dark matter. *Monthly Notices of the Royal Astronomical Society*, 415(1):40–57, 2011.

[32] R. A. Flores and J. R. Primack. Galaxy formation with an improved dark matter model. *Astrophysical Journal*, 427:L1–L4, 1994.

[33] B. Moore. Constraints on dark matter models from galaxy dynamics. *Nature*, 370(6488):629–631, 1994.

[34] Yonit Hochberg, Eric Kuflik, Tomer Volansky, and Jay G. Wacker. Mechanism for Thermal Relic Dark Matter of Strongly Interacting Massive Particles. *Physical Review Letters*, 113(17):171301, October 2014.

[35] Yonit Hochberg, Eric Kuflik, Hitoshi Murayama, Tomer Volansky, and Jay G. Wacker. Model for Thermal Relic Dark Matter of Strongly Interacting Massive Particles. *Physical Review Letters*, 115(2):021301, July 2015.

[36] Asher Berlin, Nikita Blinov, Stefania Gori, Philip Schuster, and Natalia Toro. Cosmology and accelerator tests of strongly interacting dark matter. *Physical Review D*, 97(5):055033, March 2018.

[37] WMAP Science Team. Nine year microwave sky, 2012. URL: `http://map.gsfc.nasa.gov/media/121238/index.html`.

[38] Planck Collaboration. Cosmic microwave background power spectrum, 2015. URL: `https://en.wikipedia.org/wiki/Cosmic_microwave_background#/media/File:PowerSpectrumExt.svg`.

[39] Asher Berlin, Dan Hooper, and Gordan Krnjaic. Thermal dark matter from a highly decoupled sector. *Physical Review D*, 94(9), November 2016.

[40] Brian Batell, Rouven Essig, and Ze'ev Surujon. Strong Constraints on Sub-GeV Dark Sectors from SLAC Beam Dump E137. *Physical Review Letters*, 113(17), October 2014.

[41] M. Davier and H. Nguyen Ngoc. An unambiguous search for a light Higgs boson. *Physics Letters B*, 229(1):150–155, 1989.

[42] Patrick deNiverville, Maxim Pospelov, and Adam Ritz. Observing a light dark matter beam with neutrino experiments. *Physical Review D*, 84(7), October 2011.

[43] P. H. Adrian et al. Searching for prompt and long-lived dark photons in electroproduced e^+e^- pairs with the heavy photon search experiment at JLab. *Physical Review D*, 108(1), jul 2023.

[44] Volker D. Burkert. The JLab 12GeV Upgrade and the Initial Science Program, 2012. URL: `https://arxiv.org/abs/1203.2373`, arXiv:1203.2373 [nucl-ex].

[45] A. Freyberger. Commissioning and Operation of 12 GeV CEBAF. In *Proceedings of the 6th International Particle Accelerator Conference (IPAC'15)*, International Particle Accelerator Conference, pages 1–5, Geneva, Switzerland, June 2015. JACoW.

[46] Reza Kazimi, Alicia Hofler, Joseph M. Grames, Yan Wang, John Hansknecht, Michael F. Spata, Tomasz E. Plawski, Matt Poelker, and Arne Freyberger. Operational results of simultaneous four-beam delivery at Jefferson Lab. *Proceedings of the IPAC*, 6 2019.

[47] Per Hansson, John Jaros, Takashi Maruyama, Omar Moreno, Tim Nelson, Marco Oriunno, Matt Solt, and Sho Uemura. *HPS SVT Operations Manual, v2.0.1.* HPS Collaboration, March 2016. User Manual, URL: `https://wiki.jlab.org/hps-run/images/3/3e/OPS_Manual_SVT.pdf`.

[48] Bruce Schumm. D0 Collaboration. Layer 2-5 silicon sensor specifications for D0 SMTII,

2003. Internal Note, URL: `https://scipp.ucsc.edu/~schumm/nlc/adr/cooper.pdf`.

[49] M. Raymond et al. The CMS tracker APV25 0.25-mu-m CMOS readout chip. *Conf. Proc. C*, 00091111:130–134, 2000.

[50] M. J. French et al. Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker. *Nucl. Instrum. Meth. A*, 466:359–365, 2001.

[51] Lawrence Jones. *APVMUX User Guide*, 2000. User Guide Version 1.0, May 9th, 2000. URL: `https://cds.cern.ch/record/1069892/files/cer-002725643.pdf`.

[52] Cameron Bravo. The Heavy Photon Search Experiment, 2019. arXiv:1910.04886 [hep-ex], URL: `https://arxiv.org/abs/1910.04886`.

[53] V. Fadeyev et al. Design and performance of silicon strip sensors with slim edges for HPS experiment. 969:163991, 2020.

[54] M. Battaglieri, S. Boyarinov, S. Bueltmann, V. Burkert, A. Celentano, G. Charles, W. Cooper, C. Cuevas, N. Dashyan, R. De Vita, C. Desnault, A. Deur, H. Egiyan, L. Elouadrhiri, R. Essig, V. Fadeyev, C. Field, A. Freyberger, Y. Gershtein, N. Gevorgyan, F.-X. Girod, N. Graf, M. Graham, K. Griffioen, A. Grillo, M. Guidal, G. Haller, A. Hansson, R. Herbst, M. Holtrop, J. Jaros, K. Kaneta, M. Khandaker, A. Kubarovsky, V. Kubarovsky, T. Maruyama, J. McCormick, K. Moffeit, O. Moreno, H. Neal, T. Nelson, S. Niccolai, M. Odian, M. Oriunno, R. Paremuzyan, R. Partridge, S. K. Phillips, E. Rauly, B. Raydo, J. Reichert, E. Rindel, P. Rosier, C. Salgado, P. Schuster, Y. Sharabian, D. Sokhan, S. Stepanyan, N. Toro, S. Uemura, M. Ungaro,

H. Voskanyan, D. Walz, L. B. Weinstein, and B. Wojtsekhowski. The Heavy Photon Search test detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 777:91–101, 2015.

[55] K. McCarty, V. Kubarovsky, and B. Raydo. Description and Tuning of the HPS Trigger. Technical report, Heavy Photon Search, October 2018. Internal Note, URL: `https://misportal.jlab.org/mis/physics/hps_notes/index.cfm?note_year=2018`.

[56] N. A. Graf. Event Reconstruction in Java. *Journal of Physics: Conference Series*, 331(3):032012, December 2011.

[57] Matt Solt et al. Search for Long-Lived Dark Photons with the Heavy Photon Search Experiment at JLab, 2020. Internal Note, URL: `https://confluence.slac.stanford.edu/display/hpsg/Physics+Analysis+Notes?preview=/146715820/341246632/Displaced_Vertex_Search_2016_March14_2022.pdf`.

[58] Linear Collider Detector. Linear Collider Simulation Seedtracker. Software packaged documentation, URL: `http://www.lcsim.org/sites/lcsim/apidocs/index.html?org/lcsim/recon/tracking/seedtracker/class-use/TrackCheck.html`.

[59] C. Kleinwort. General broken lines as advanced track fitting method. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 673:107–110, 2012.

[60] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nuclear*

Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 262:444–450, 1987.

[61] Robert P. Johnson. HPS Kalman Filter Documentation. Internal Note, URL: `https://confluence.slac.stanford.edu/download/attachments/146715820/KalmanDoc.pdf?api=v2`.

[62] P. Billoir and S. Qian. Fast vertex fitting with a local parametrization of tracks. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 311(1):139–150, 1992.

[63] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer. MadGraph/MadEvent v4: The New Web Generation. *Journal of High Energy Physics*, 2007(09):028, September 2007.

[64] H. Hirayama, Y. Namito, A. F. Bielajew, S. J. Wilderman, and W. R. Nelson. The EGS5 Code System. Technical Report SLAC-R-730, United States Department of Energy, Stanford, CA, 2005.

[65] FreeHEP Project. Freehep stdhep api documentation, 2007. Software package, URL: `https://java.freehep.org/freehep-stdhep/apidocs/index.html`.

[66] N. Graf and J. McCormick. Simulator for the Linear Collider (SLIC): A Tool for ILC Detector Simulations. *AIP Conference Proceedings*, 867(1):503–512, 2006.

[67] HPS Vertex Folks. Search for Long-Lived Dark Photons with the Heavy Photon Search Experiment at JLab. Internal Note, URL: `https://confluence.slac.`

`stanford.edu/download/attachments/146715820/Displaced_Vertex_Search_2016_March14_2022.pdf?version=1&modificationDate=1647276078000&api=v2`.

[68] HPS Bump Hunt Folks. Resonance search analysis of 2016 HPS spring run data. Internal Note, URL: `https://confluence.slac.stanford.edu/download/attachments/146715820/HPS_2016_Bump_Hunt_Internal_Note.pdf?version=1&modificationDate=1617814497144&api=v2`.

[69] S. Yellin. Finding an upper limit in the presence of an unknown background. *Physical Review D*, 66(3):032005, August 2002.

[70] R. D. Cousins and J. T. Linnemann. Evaluation of three methods for calculating statistical significance when incorporating a systematic uncertainty into a test of the background-only hypothesis for a Poisson process. 595(2):480–501, 2008.

Appendix A

Binomial Significance (ZBi)

The binomial significance of a given signal search window is calculated using the method described in [70], which performs a frequentist background-only hypothesis test for a Poisson process that can be separated into an "on" (signal+background) and "off" (background only) region. The result is only valid when the background in the signal region is small, but non-zero.

The "off" region consists of a reconstructed and selected background sample in the signal region, where n_{off} events are observed from a Poisson process with mean μ_{off}. The "on" region consists of the expected signal count n_{sig}, and the expected background n_{bkg} in the "on" region, where n_{on} events are observed from a Poisson process with mean $\mu_{\mathrm{on}} = \mu_{\mathrm{signal}} + \mu_{\mathrm{background}}$.

Under the null hypothesis H_0, when $\mu_{\mathrm{signal}} = 0$, the ratio of the expected Poisson means of the "off" and "on" regions is $\lambda = \mu_{\mathrm{off}}/\mu_{\mathrm{on}} = \mu_{\mathrm{off}}/\mu_{\mathrm{background}} = \tau$. The background rate in the "on" region is estimated using events in the "off" region, where τ is the ratio of the background between the two regions.

The null hypothesis test is constructed using the ratio of Poisson means in the "off" and "on" regions, where events in n_on and n_{off} are observed from two separate Poisson probabilities with unknown means μ_{on}, and μ_{off}, respectively.

The joint probability distribution for observing n_{on} and n_{off} events is the product of the individual Poisson probabilities,

$$P(n_{\text{on}}, n_{\text{off}}) = \frac{e^{-\mu_{\text{on}}} \mu_{\text{on}}^{n_{\text{on}}}}{n_{\text{on}}!} * \frac{e^{-\mu_{\text{off}}} \mu_{\text{off}}^{n_{\text{off}}}}{n_{\text{off}}!}. \tag{A.1}$$

This join probability distribution can be rewritten as the probability for observing n_{total} events,

$$P(n_{\text{on}}, n_{\text{off}}) = \frac{e^{-(\mu_{\text{on}}+\mu_{\text{off}})} (\mu_{\text{on}} + \mu_{\text{off}})^{n_{\text{tot}}}}{n_{\text{tot}}!}, \tag{A.2}$$

and the conditional binomial probability of observing n_{on} events, given n_{total} observations, where the probability of successfully observing an n_{on} event is $\rho = \lambda = \mu_{\text{on}}/\mu_{\text{tot}} = 1/(1+\lambda)$, is

$$\frac{n_{\text{tot}}!}{n_{\text{on}}!(n_{\text{tot}} - n_{\text{on}})!} (1/1 + \lambda))^{n_{\text{on}}} (1 - (1/1 + \lambda))^{(n_{\text{tot}} - n_{\text{on}})}, \tag{A.3}$$

The probability distribution used to test H_0 is therefore

$$P(n_{\text{on}}, n_{\text{off}}) = \frac{e^{-(\mu_{\text{on}}+\mu_{\text{off}})} (\mu_{\text{on}} + \mu_{\text{off}})^{n_{\text{tot}}}}{n_{\text{tot}}!} * \frac{n_{\text{tot}}!}{n_{\text{on}}!(n_{\text{tot}} - n_{\text{on}})!} (1/1+\lambda))^{n_{\text{on}}} (1-(1/1+\lambda))^{(n_{\text{tot}} - n_{\text{on}})}, \tag{A.4}$$

or more simply,

$$= P(n_{total}; \mu_{total}) P(n_{\text{on}} | n_{total}; 1/(1 + \lambda)). \tag{A.5}$$

All of the information about the ratio of Poisson means λ, and therefore the null hypothesis where $\lambda = \tau$, is contained in the conditional binomial probability term $P(n_{\text{on}} | n_{total}; 1/(1 + \lambda))$. H_0 is tested with a standard one-tailed binomial p-value test,

$p_{\rm Bi} = \sum_{j=n_{\rm on}}^{n_{\rm tot}} P(j|n_{\rm tot}; 1/(1+\tau)$, which is evaluated using the ratio of incomplete and complete beta functions [70]

$$p_{\rm Bi} = B(1/(1+\tau), n_{\rm on}, 1+n_{\rm off})/B(n_{\rm on}, 1+n_{\rm off}), \tag{A.6}$$

and then convert $p_{\rm Bi}$ to a binomial significance ($Z_{\rm Bi}$) using

$$Z_{\rm Bi} = \sqrt{2} erf^{-1}(1-2p_{\rm Bi}). \tag{A.7}$$

Appendix A

Preselection Cutflow

Figures A.1-A.6 show the impact of each individual preselection cut on the invariant mass for data and all MC samples. Of particular note, the vertex quality cut dramatically reduces the WABs rate in the bulk of its mass distribution. Figure A.11

Figures A.11-A.12 show the N-1 vertex z distributions, where again it is shown that the vertex quality cut contributes most to reducing the high-z tails in the background, specifically the WABs high-z tails, as shown in Figure A.9.

The invariant mass distributions for data and MC are shown as a function of the ordered preselection cuts in Figures A.13-A.16, the vertex z position distributions are shown in Figures A.17-A.20, and the track momentum sum distributions are shown in Figures A.21-A.24.

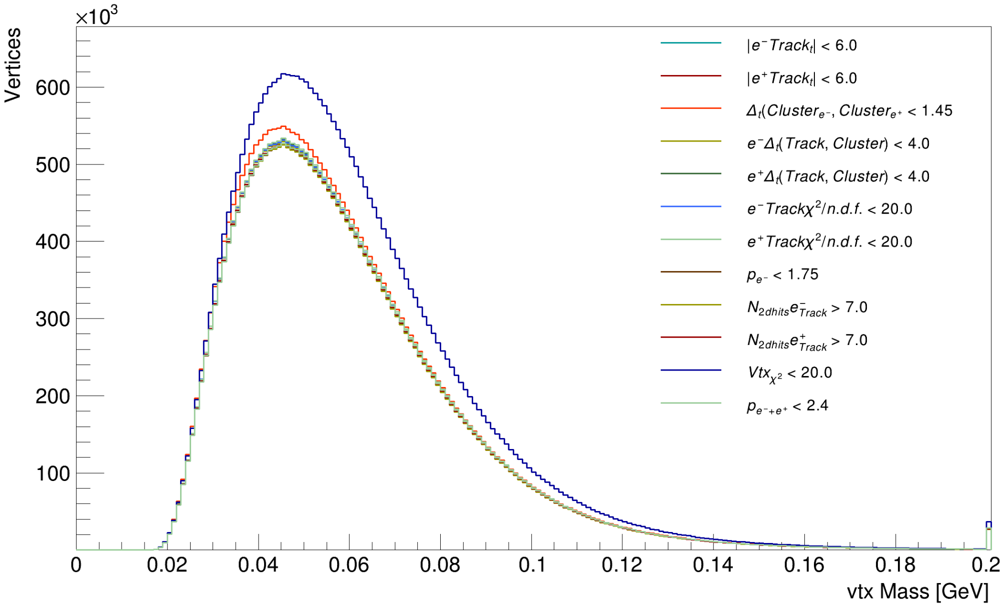


Figure A.1: 10 % Data Sample Invariant Mass distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

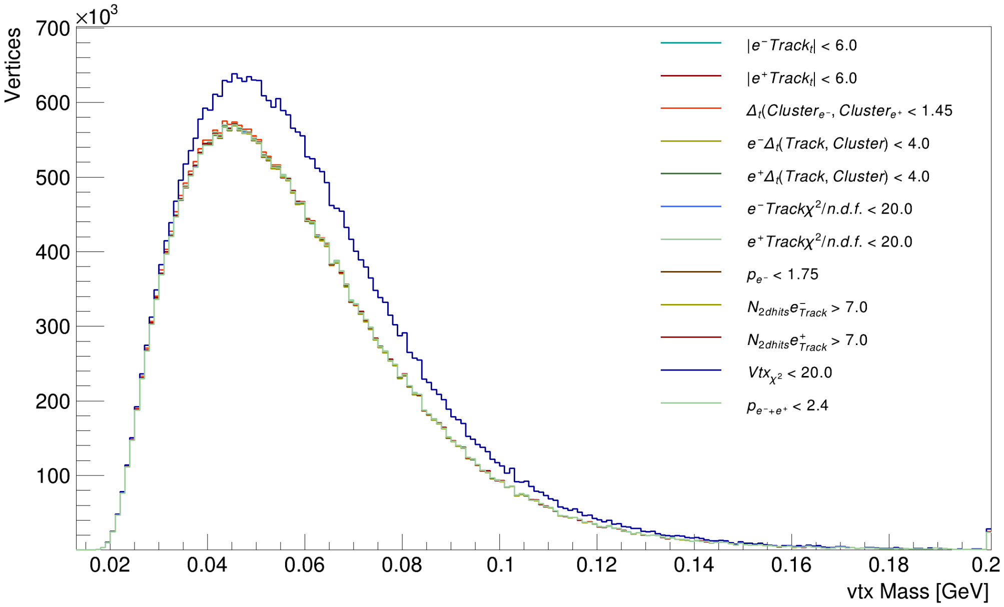


Figure A.2: Tritrig+WAB+Beam MC Invariant Mass distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

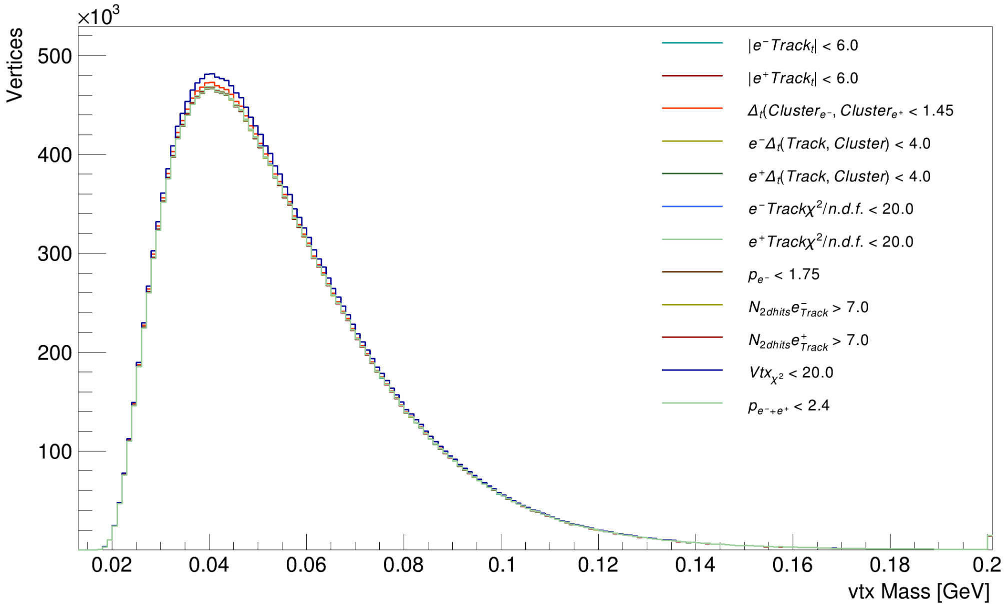


Figure A.3: Tritrig+Beam MC Invariant Mass distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

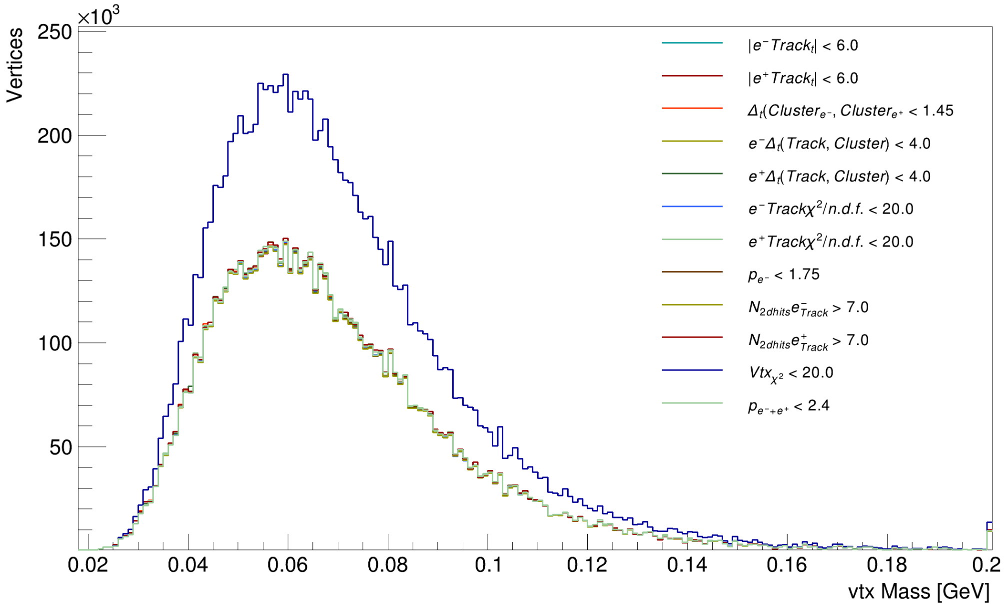


Figure A.4: WAB+Beam MC Invariant Mass distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

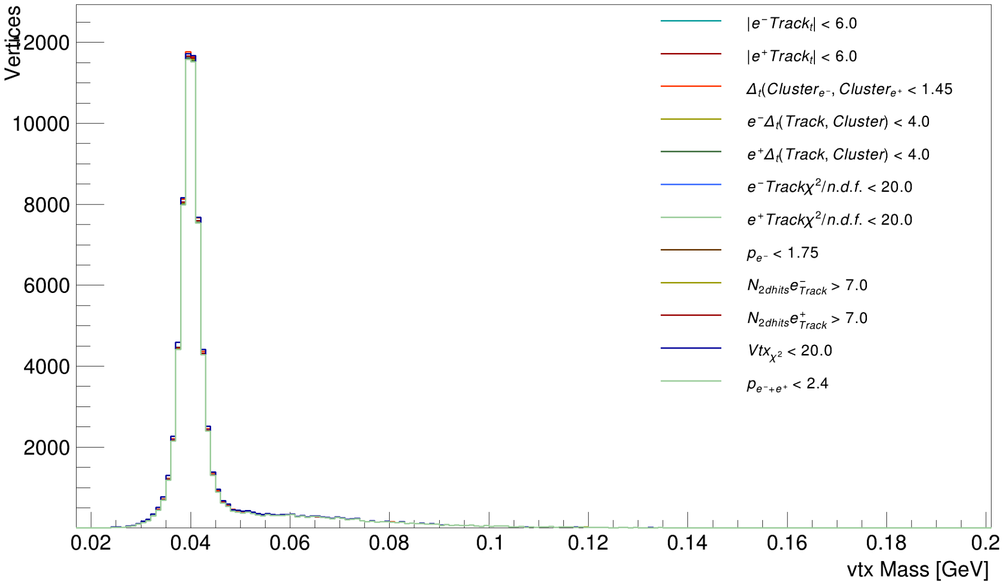


Figure A.5: 40 MeV Signal MC Invariant Mass distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

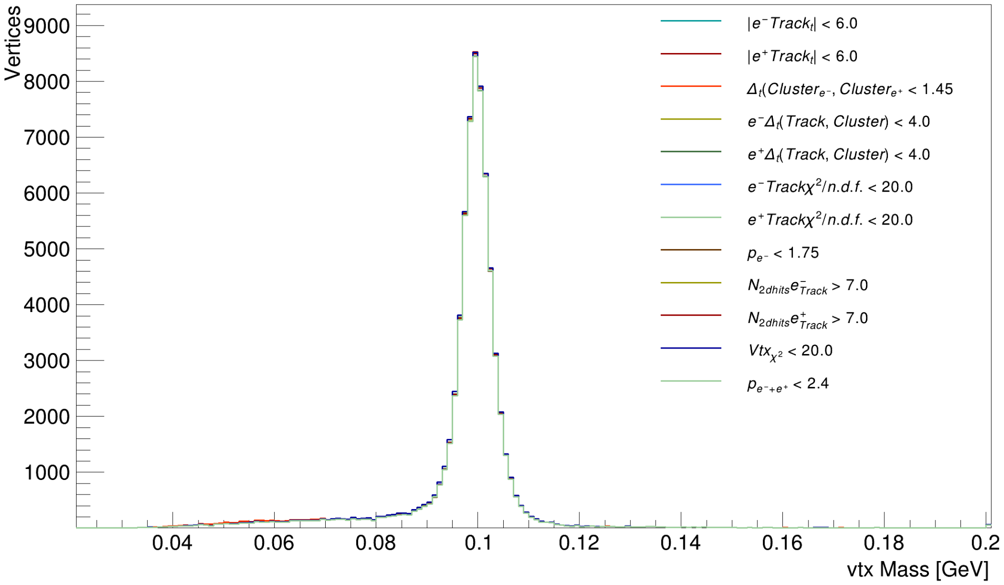


Figure A.6: 100 MeV Signal MC Invariant Mass distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

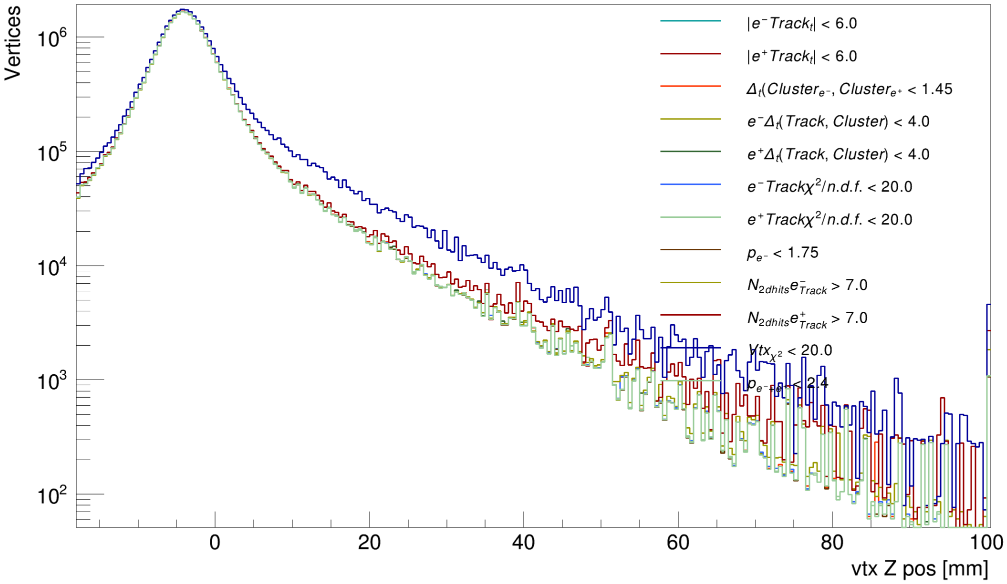


Figure A.7: Tritrig+WAB+Beam Vertex Z distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

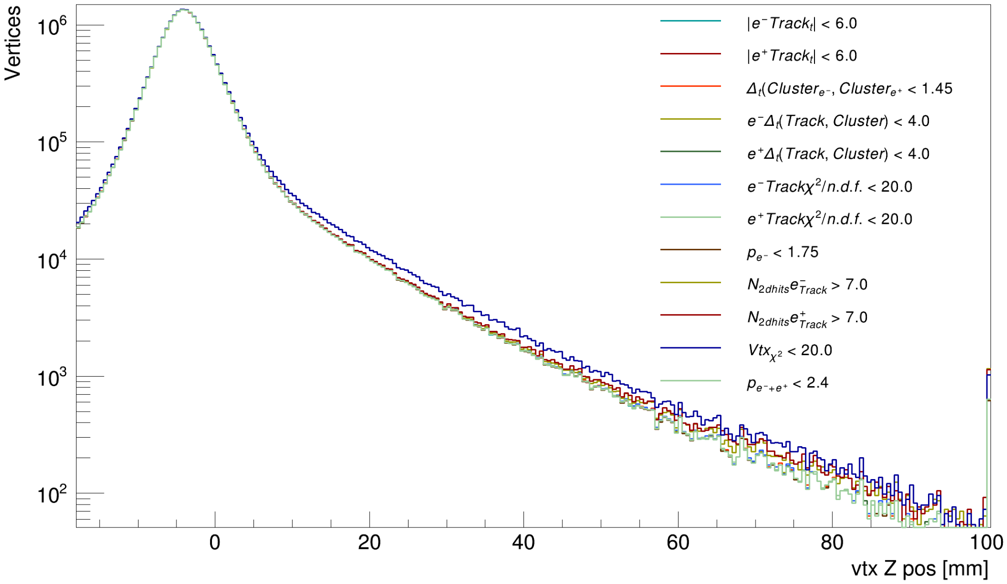


Figure A.8: Tritrig+Beam Vertex Z distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

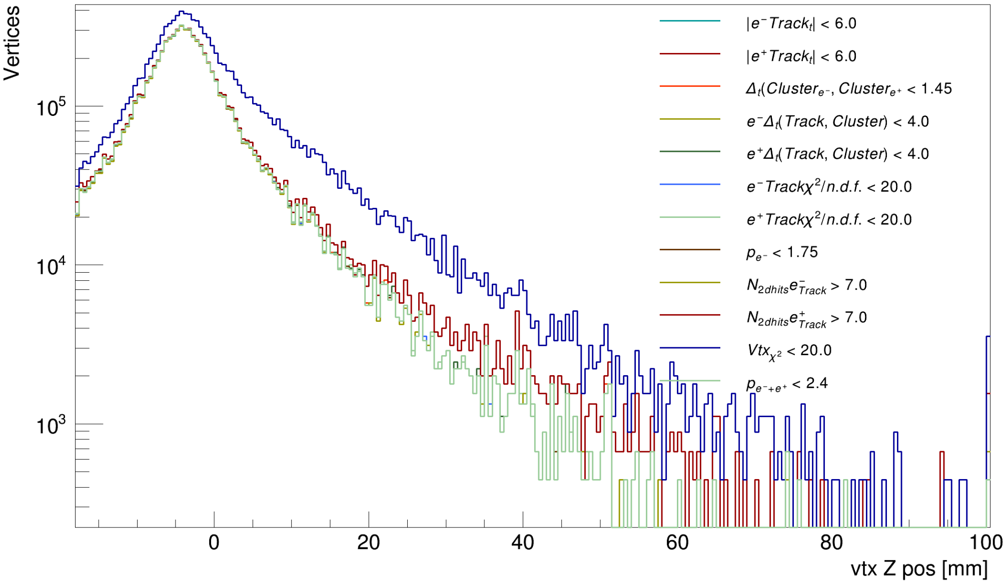


Figure A.9: WAB+Beam Vertex Z distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

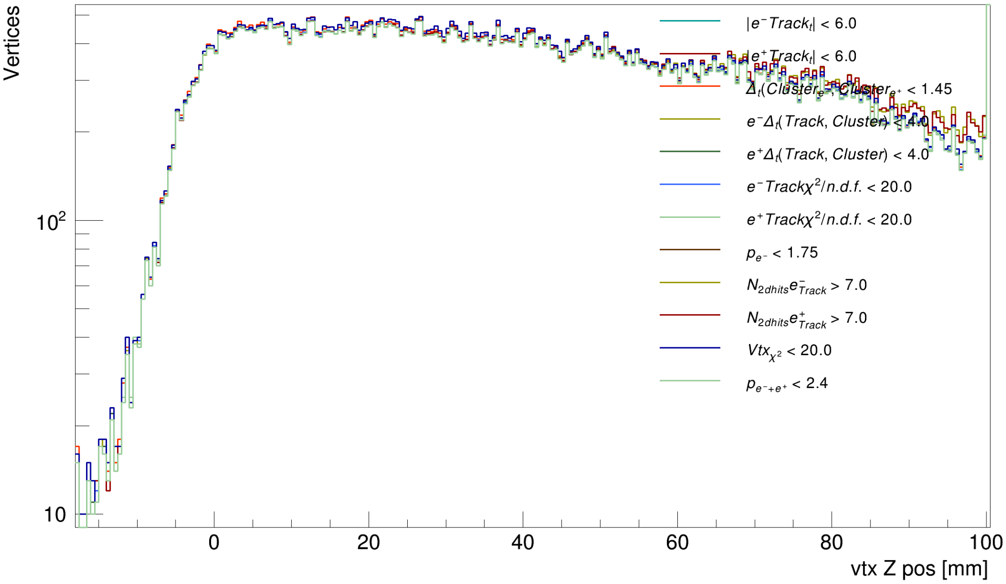


Figure A.10: 40 MeV Signal MC Vertex Z distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

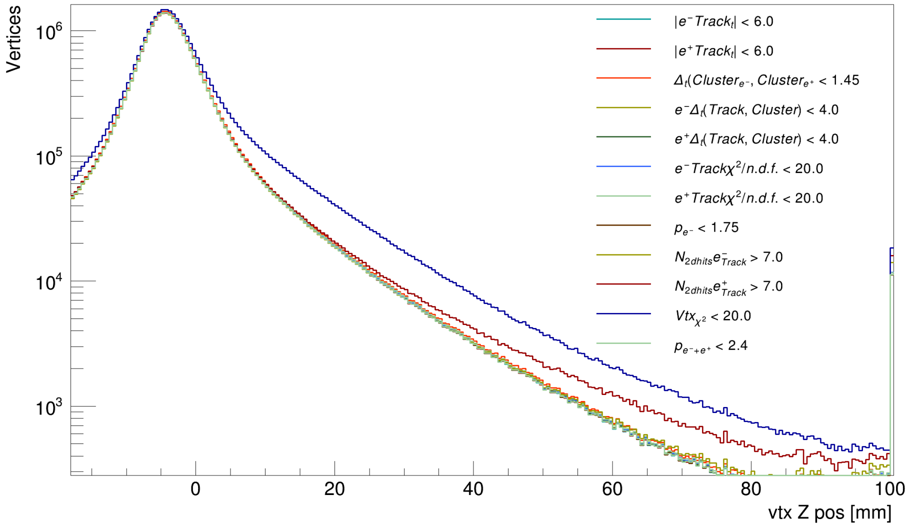


Figure A.11: Data Sample Vertex Z distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

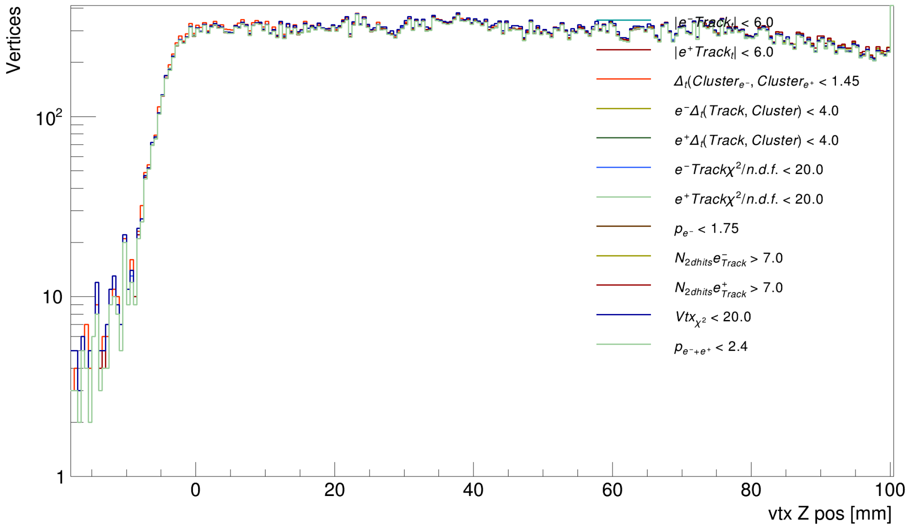


Figure A.12: Signal MC Vertex Z distribution as a function of preselection N-1 cuts. Each entry in the legend corresponds to all preselection cuts applied, except for the cut listed.

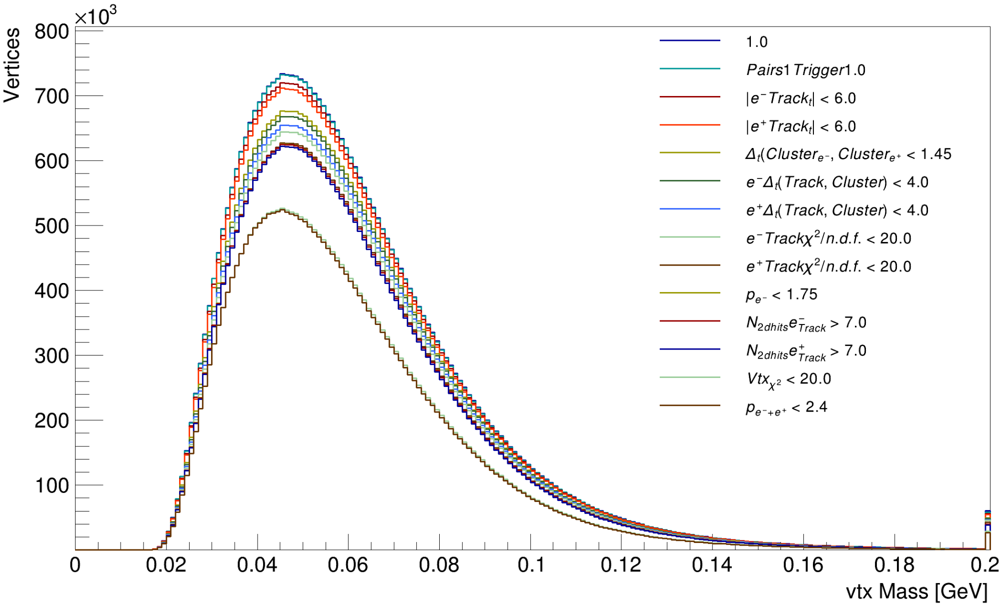


Figure A.13: Invariant mass preselection cutflow for the unblinded $\sim 10\%$ data sample.

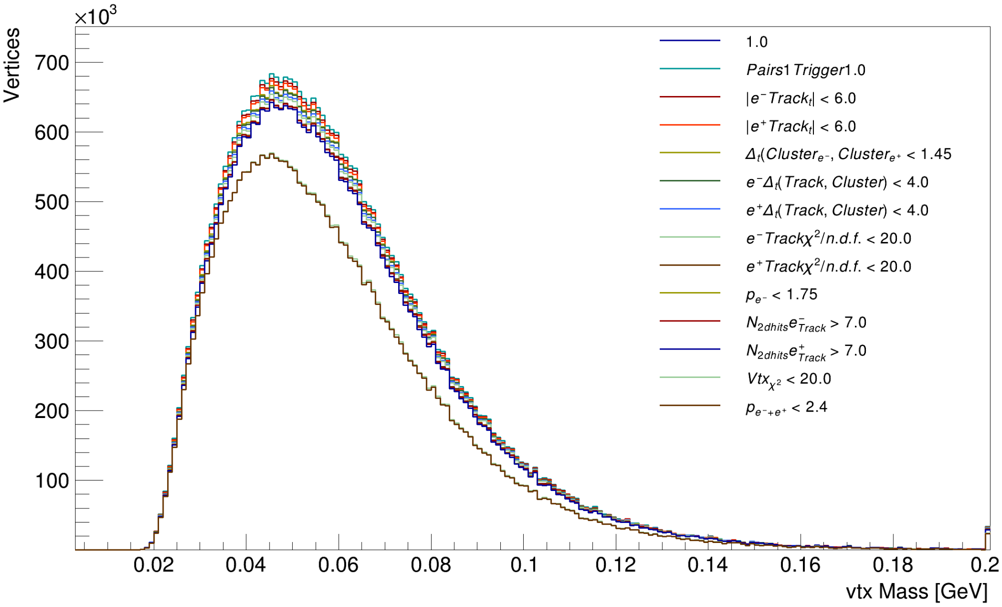


Figure A.14: Invariant mass preselection cutflow for Tritrig+WAB+Beam MC scaled to $\sim 10\,\%$ luminosity.

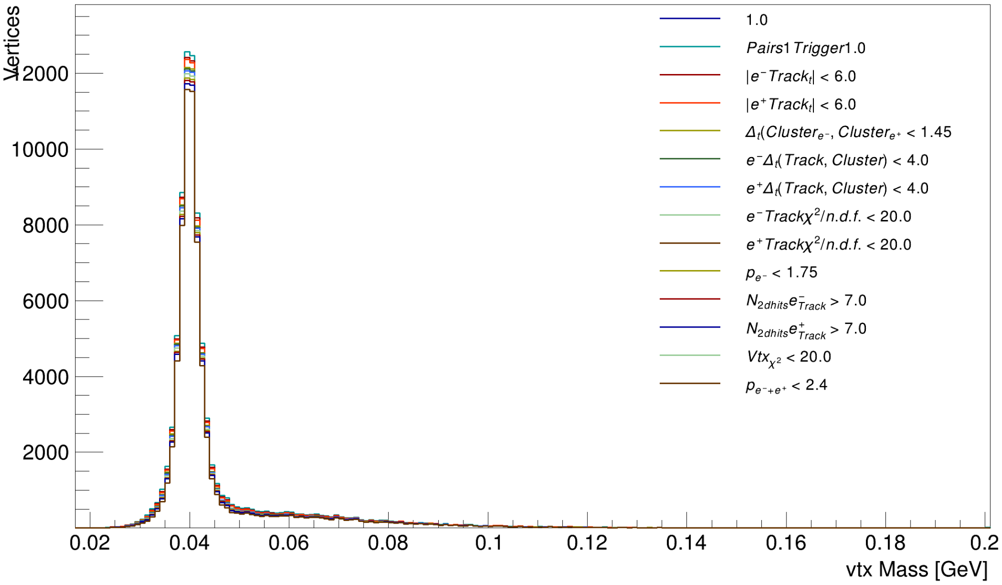


Figure A.15: Invariant mass preselection cutflow for 40 MeV Signal MC, unscaled.

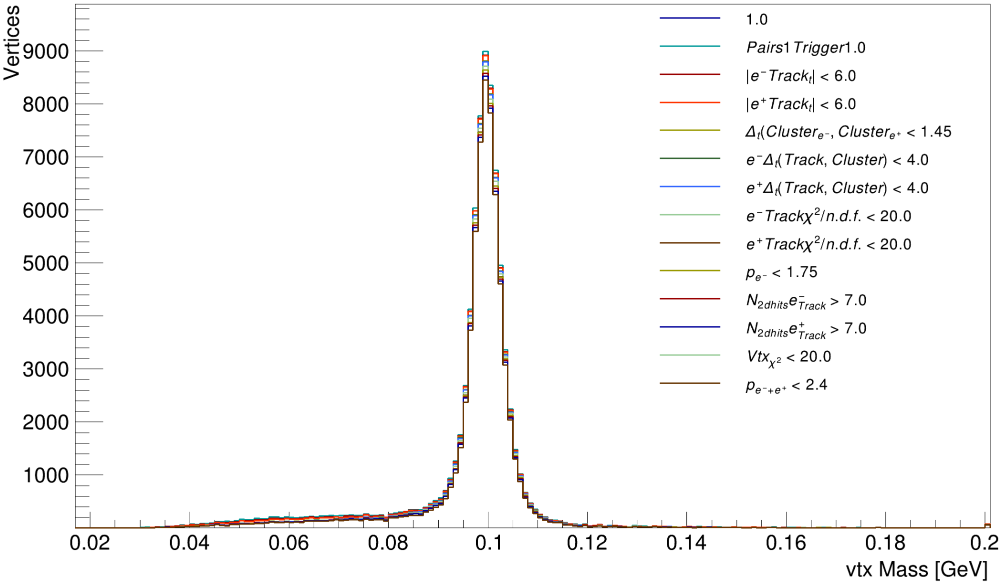


Figure A.16: Invariant mass preselection cutflow for 100 MeV Signal MC, unscaled.

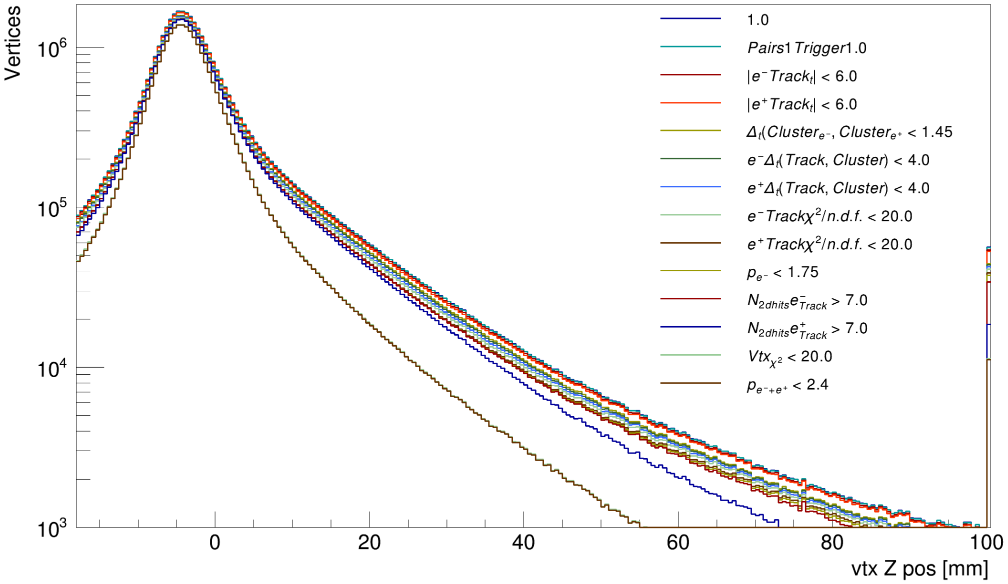


Figure A.17: Vertex z position preselection cutflow for unblinded $\sim 10\,\%$ data sample.

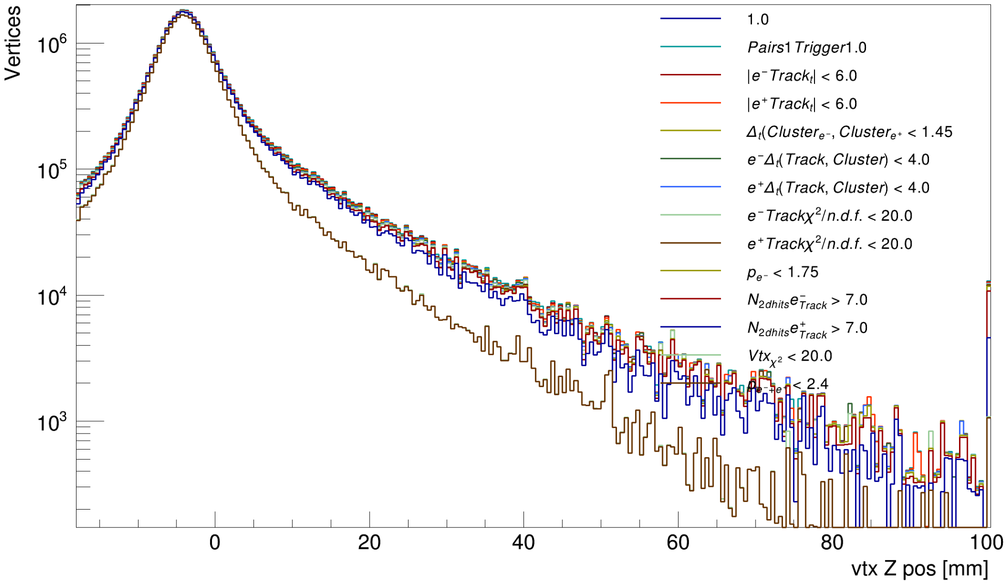


Figure A.18: Vertex z position preselection cutflow for Tritrig+WAB+Beam MC scaled to $\sim 10\,\%$ luminosity.

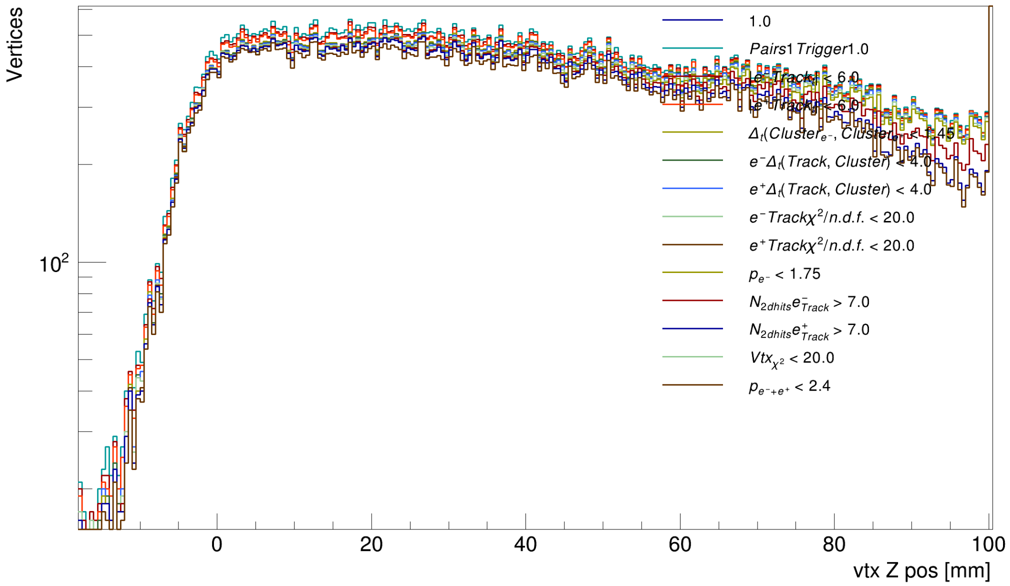


Figure A.19: Vertex z position preselection cutflow for 40 MeV Signal MC, unscaled.

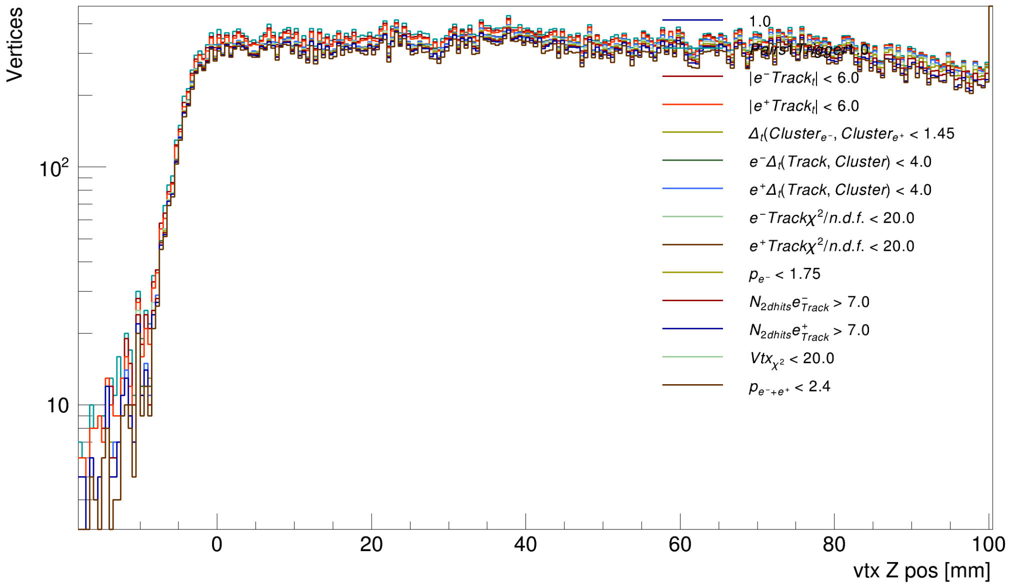


Figure A.20: Vertex z position preselection cutflow for 100 MeV Signal MC, unscaled.

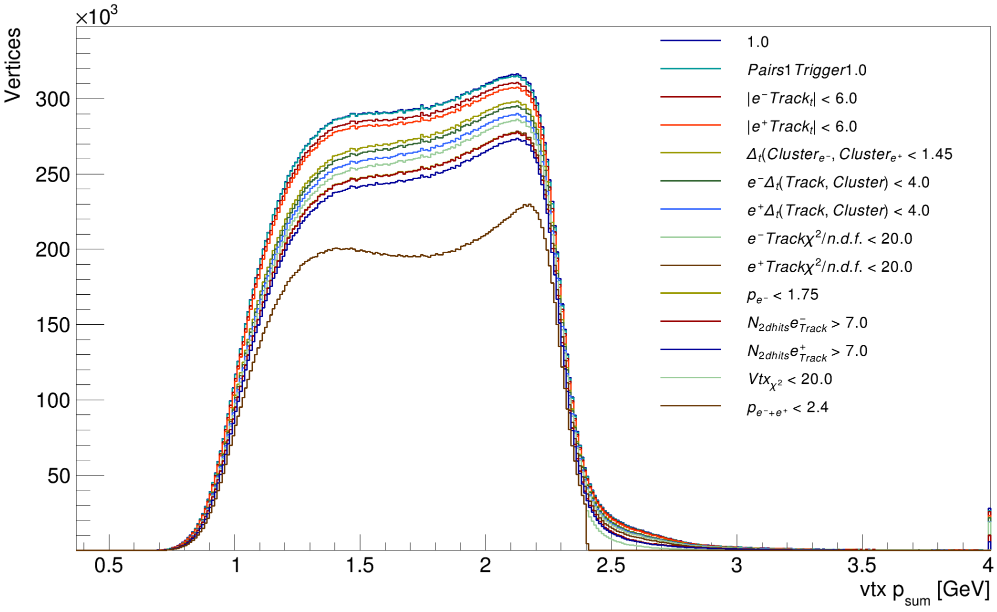


Figure A.21: Reconstructed vertex track momentum sum for unblinded $\sim 10\,\%$ data sample.

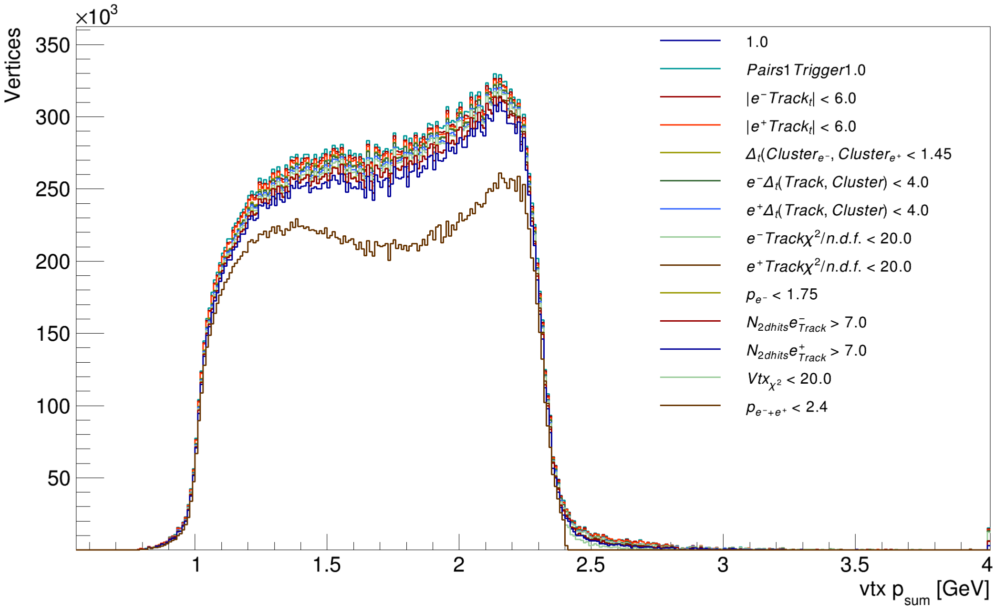


Figure A.22: Reconstructed vertex track momentum sum for Tritrig+WAB+Beam MC scaled to $\sim 10\,\%$ luminosity.

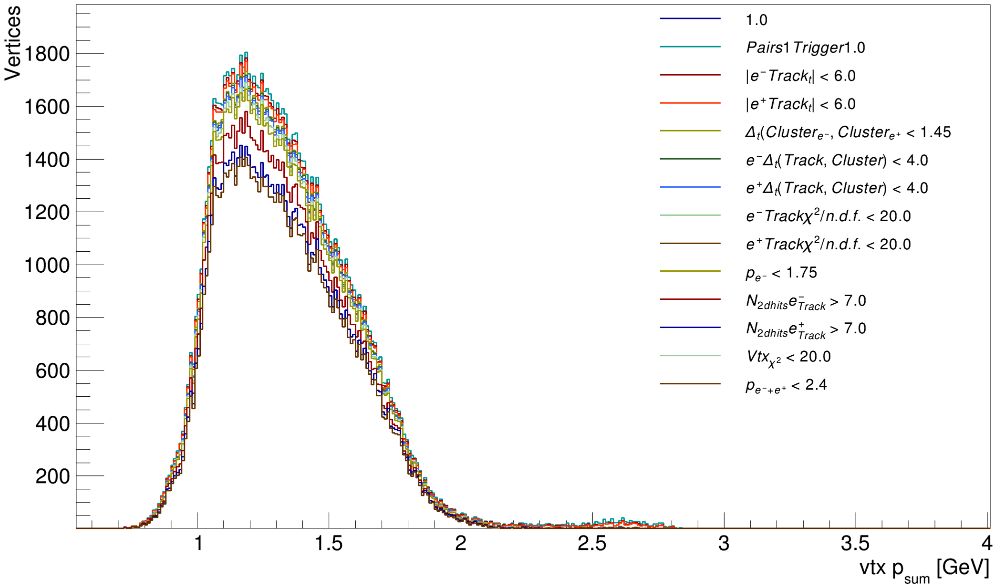


Figure A.23: Reconstructed vertex track momentum sum for 40 MeV Signal MC, unscaled.

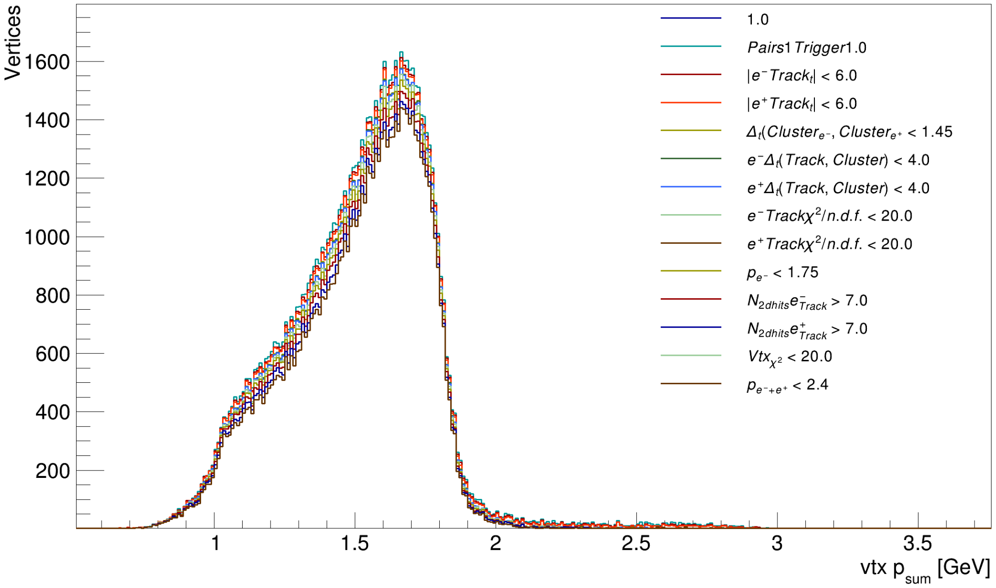


Figure A.24: Reconstructed vertex track momentum sum for 100 MeV Signal MC, unscaled.

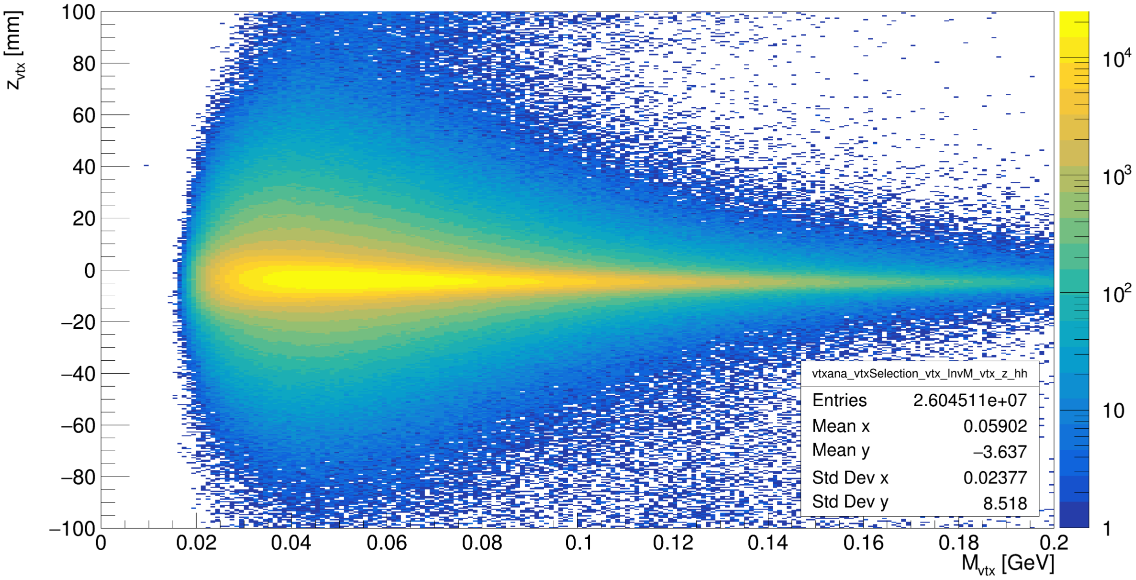


Figure A.25: Reconstructed vertex z position versus Invariant Mass for the $\sim 10\,\%$ Data sample.