

A search for inelastic dark matter at the Belle II Experiment

Tommy Lam

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Physics

Leo E. Piilonen, Chair

Thomas O'Donnell

Ian M. Shoemaker

Uwe C. Täuber

May 1, 2026

Blacksburg, Virginia

Keywords: Belle II, particle physics, dark matter, inelastic dark matter

Copyright 2026, Tommy Lam

A search for inelastic dark matter at the Belle II Experiment

Tommy Lam

(ABSTRACT)

The Standard Model (SM) of particle physics is a remarkably successful and highly predictive theory of fundamental particles and interactions. However, it is an incomplete description of nature, as it fails to account for several observed phenomena, most notably the existence of dark matter. Observations ranging from galactic rotation curves to anisotropies in the Cosmic Microwave Background (CMB) provide compelling evidence for a non-baryonic gravitational source that lies entirely beyond the scope of the SM.

One viable extension to the SM is the inelastic dark matter (iDM) model. This model introduces a dark sector containing two dark fermions, χ_1 and χ_2 , with a mass splitting ($\Delta m_\chi = m_{\chi_2} - m_{\chi_1}$). These particles couple to the SM through a massive dark photon (A) portal via kinetic mixing with the electroweak $U(1)_Y$ hypercharge component. In this work, we search for iDM production in e^+e^- collisions through the process $e^+e^- \rightarrow A'\gamma_{ISR} \rightarrow \chi_1\chi_2\gamma_{ISR}$. The signature is characterized by a high-energy initial-state radiation (ISR) photon and a displaced vertex arising from the decay $\chi_2 \rightarrow \chi_1 A'^* (\rightarrow l^+l^-)$ for $l \in [e, \mu]$.

This analysis is performed using data collected by the Belle II detector at the SuperKEKB asymmetric-energy electron-positron collider in Tsukuba, Japan. We outline reconstruction and selection strategies optimized to isolate these rare signatures from dominant SM backgrounds, such as Bhabha scattering, two-photon processes, and continuum $q\bar{q}$ production. To maximize signal sensitivity across a broad mass range ($0.05 \text{ GeV}/c^2 \leq m_{\chi_1} \leq 3.2 \text{ GeV}/c^2$), we employ an ensemble of Gradient-Boosted Decision Trees. These models are trained on

specific mass windows and particle identification categories to effectively discriminate signal from background using features such as vertex displacement, track-cluster isolation, and missing energy.

The statistical analysis uses a frequentist approach to determine Belle II's sensitivity. We perform a single-bin counting experiment using sideband data to estimate the background level within a defined signal window. In the absence of an observed excess, we set median upper limits on the signal yield at a 95% confidence level (CL) using the $\tilde{q}$ test statistic and toy-based calculations.

A search for inelastic dark matter at the Belle II Experiment

Tommy Lam

(GENERAL AUDIENCE ABSTRACT)

One of the most profound realizations in modern physics is that everything we have observed accounts for only a small fraction of the universe. In fact, the “ordinary” matter described by our best theories only makes up about 15% of the total matter density. The remaining 85% is a mysterious, invisible substance known as dark matter. While we can see its gravitational “signatures” across the cosmos, such as how it pulls on stars in galaxies or bends the path of distant light, dark matter itself remains one of the greatest unsolved puzzles in physics because it does not emit or reflect any light.

This thesis investigates a theory called inelastic dark matter. In this model, dark matter comprises a family of particles that can undergo state changes upon interaction. To search for this, I used data from the Belle II experiment in Japan, which smashes subatomic particles together at nearly the speed of light. If this theory is correct, these collisions could produce a “heavy” dark matter particle that travels a short distance before transforming into a “light” version. This transformation leaves a unique signal, called a displaced vertex, which is essentially a delayed “flash” of energy that appears to come from nowhere.

Detecting these rare “flashes” is an immense challenge because they are buried among billions of ordinary events. To find them, I used the Belle II detector, which acts like a massive, three-storey digital camera that captures the aftermath of every collision event. Because the data is so vast, I employ machine learning methods to act as a sophisticated filter. This software is trained to distinguish subtle differences between common background noise and the specific patterns expected from dark matter. Finally, I use statistical tools to assess

whether any remaining signals are genuine discoveries or merely random fluctuations. By combining high-energy physics with modern data science, this work provides a new approach to testing the frontiers of our universe and searching for the invisible constituents that hold it together.

Dedication

To those who have borne us aloft: This treatise we lay before you, in acknowledgment that what far sight we claim is owed to the height of your stature. Whatever luminescence we cast is but a reflection of the fire you first kindled.

Acknowledgments

As the adage suggests, it takes a village to raise a Ph.D. I am profoundly indebted to the community that has sustained me through this long and arduous journey.

I wish to begin by extending my deepest appreciation to my advisor, Leo E. Piilonen. For eight years, I have been fortunate to draw upon his immense technical expertise, reliable acumen, and sagacious perspective. From the countless opportunities to develop as an experimentalist to the essential encouragement needed to establish my own footing as an independent researcher within the Belle II collaboration, his aegis has been the cornerstone of my professional growth. I hope to reflect the high standard of scientific integrity and dedication he models as I move into the next phase of my career.

I offer my sincere thanks to my committee members, Profs. Thomas O'Donnell, Uwe C. Täuber, Ian M. Shoemaker, and Shunsaku Horiuchi, for their critical feedback and guidance. I am likewise grateful to my colleagues in the Piilonen group: Haruki Kindo, Vipin Gaur, Zachary Stottler, Taylor Kimmel, Logan Benninghoff, Santi Naylor, and Ryan Mueller, whose collective insight proved invaluable.

A critical community that has been vital to this thesis is the large number of Belle II collaborators I have met throughout my tenure. For their pivotal mentorship and guidance, I express my appreciation to Savino Longo, Mikihiro Nakao, Christopher Hearty, Luigi Corona, Giacomo de Pietro, Seokhee Park, Takuto Kunigo, Kay Kinoshita, Alan Schwartz, Soeren Prell, Chunhui Chen, Peter Lewis, Jake Bennett, Daniel Greenwald, Diego Tonelli, and the late Gary Varner. For their direct assistance with the physics analysis and for making the atmosphere of the dark sector group so engaging, I am indebted to Miho Wakai, Daniel Crook, Alexander Heidelbach, Jonas Epplet, and Patrick Ecker. For their technical exper-

tise and for being partners in crime throughout Ibaraki and Tokyo, I thank Björn Spruck, Dima Levit, Chia-Ling Hsu, Priyanka Cheema, Quinn Campagna, Noah Brenny, Michele Veronesi, Tristan Fillinger, Cyrille Praz, Mattéo Maushart, Naveen Baghel, Jaeyoung Kim, Alex Gale, Jeffery Chin, Riccardo Manfredi, Boyang Zhang, Raynette van Tonder, Matt Barrett, and Christopher Ketter. I also thank the administrative staff at KEK, Tomoya Yokoyama, Sachiko Tanaka, Heasook Han, Megumi Yaoshima, and Junko Kohri, as well as those in Japan who made my time there unforgettable: Mai-san of Idouji, Mokkun Coffee, the Konno family, Roberta Bellini, Kanjana Worarado, and Anchalee Vongto.

For the community at Virginia Tech, I am grateful for the early mentorship of Giti Khodaparast, Tatsu Takeuchi, Eric Sharpe, and the late Kyungwha Park. To the comrades-in-arms of Robeson Hall who kept my spirits high during the pandemic: Ho Lung Tang, Joseph Camilleri, Anri Karanovich, Eva Takou, Luis A. Delgadillo, Hong Yao, David Smith, Tonya Pruitt, Arian Vezvaei, Karla Tellez-Giron-Flores, Sarah Healy, Daniel Valmassei, Nick Ekanger, Sean Heston, Andrew Gustafson, James Stidham, Nicholas Smith, Rachel Maizel, Sunhaeng Hur, Nathan J. Beets, James V. Haag IV, Benjamin E. Dillinger, Sheri Singerling, Alan Strom, and Keyu Ding; thank you for making those corridors and Blacksburg a home. A special mention goes to Betty Wilkins and Katrina Lockhart-Elfeky. Whether they were performing administrative miracles, providing insight on the inner workings of the university, or navigating the reimbursement of a certain New York City taxi ride, their support will be sorely missed.

To my ‘ohana, Hubert Szczygieł, Alyssa Lau, Victoria Bui, Gabor Paczolay, Kelsey Angel Baehrens, Christina Huynh, Tyler Trent, Makana Silva, and Ka‘imi Kahihihikolo: thank you for being the vital anchor that kept me whole and grounded; I am profoundly grateful for a decade of friendship that has served as an enduring beacon of warmth and light.

Finally, to my family. To my partner, Do Kieu Trang Thoi: your steadfast presence and

boundless love have been the quiet cornerstones of my resilience. I reached this finish line only because you were there to support me. To my parents: thank you for thirty years of unwavering and unconditional love.

Contents

List of Figures	xvi
-----------------	-----

List of Tables	xxxii
----------------	-------

1	Introduction	1
1.1	The Standard Model of Particle Physics	2
1.1.1	Constituents of the Standard Model	2
1.1.2	Going Beyond the Standard Model	4
1.2	Evidence for Dark Matter	6
1.2.1	Galaxy Rotation Curves	7
1.2.2	Lensing of Galaxy and Galaxy Clusters	9
1.2.3	Density Fluctuations and Structure Formation	11
1.2.4	Cosmic Microwave Background	14
1.2.5	Lyman-alpha Forests	15
1.2.6	Big Bang Nucleosynthesis	16
1.3	Alternatives to Dark Matter	18
1.4	Searches for Dark Matter	18
1.4.1	Weakly-Interacting Massive Particles	20

1.4.2	Axion-Like Particles	23
1.4.3	Primordial Black Holes	24
1.5	Inelastic Dark Matter	25
2	The Belle II Experiment	32
2.1	SuperKEKB	33
2.1.1	IR Design	33
2.2	Requirements of the Belle II Detector	36
2.3	Pixel Vertex Detector	38
2.4	Silicon Vertex Detector	40
2.5	Central Drift Chamber	41
2.6	Time-Of-Propagation Counter	44
2.7	Aerogel Ring-Imaging Cerenkov Detector	46
2.8	Electromagnetic Calorimeter	48
2.8.1	Upgrades from Belle	51
2.9	Solenoid	52
2.10	K_L^0 and Muon Detector	53
2.11	Trigger System	59
2.12	Data Acquisition System	61
2.13	Software	65

2.13.1	Tracking	66
2.13.2	Clustering	69
2.13.3	Vertexing	70
2.13.4	Particle Identification	71
2.13.5	Simulation and Calibration	71
3	Analysis	73
3.1	Tools and Strategy	74
3.2	Reconstruction and Selection	85
3.2.1	Background Generation	85
3.2.2	Reconstruction	87
3.2.3	Background Modeling	102
3.2.4	Event Selection	104
3.2.5	Signal Extraction	108
3.2.6	Application of Event Selection	112
3.3	Systematics	115
3.3.1	Trigger Efficiencies	115
3.3.2	BDT Systematics	118
3.3.3	Background Shape Parameters	120
3.3.4	Validating the Upper Limit Scheme	122

3.3.5	Linearity Test	124
4	Results and Discussion	133
4.1	Data and Monte-Carlo Agreement	135
4.2	Validating the cut on number of CDC Hits	138
4.3	Validating Background Shapes	140
4.4	Excess of ~ 1 GeV Photons	141
4.5	Validating BDT performance	145
5	Conclusions	149
	Bibliography	152
	Appendices	193
	Appendix A Belle II Alarm Server	194
	Appendix B Technical Details of KLM's Data Quality Monitoring	200
B.1	Motivation	200
B.2	Framework Features	201
B.2.1	Comparisons with References	202
B.2.2	Delta Histograms	203
B.2.3	MiraBelle	204

B.2.4	EPICS Integration and Alarming	205
B.3	KLM Monitored Quantities	206
B.3.1	Occupancies	206
B.3.2	Layer Efficiencies	207
B.3.3	Background Hit-Rates	210
B.4	Conclusion	212
Appendix C Generative Adversarial Network for KLM Backgrounds		214
C.1	Introduction	214
C.2	Generative Adversarial Network	215
C.2.1	Wasserstein GANs	217
C.2.2	Architecture	218
C.3	Procedure	219
C.3.1	Priors	220
C.3.2	Training and Development	221
C.3.3	Figure of Merit	222
C.4	Results	225
C.4.1	Progression of Training	226
C.4.2	Dependence on d_{latent}	228
C.4.3	Dependence on n_{critic}	229

C.4.4	Evaluation of Deployed WGANs	231
Appendix D	An analytic model for supernova neutrinos	237
D.1	Introduction	237
D.2	Formulation	240
D.2.1	Müller’s Analytic Model	240
D.2.2	Adding Neutrinos To The Model	249
D.3	Results	255
D.3.1	Calibration	255
D.3.2	Comparisons	257
D.4	Discussion and Conclusions	259

List of Figures

1.1	Rotation curves of the Triangle Nebula (M33) galaxy, suggesting the presence of (dark) matter that has no optical presence [24, Fig. 4]. The dashed curve is the prediction for $R_G \approx 3\text{kpc}$; the solid curve is the observation.	8
1.2	(a) An illustration of gravitational lensing. (b) An image from the Hubble Space Telescope of an almost ideal Einstein ring, a particular case of strong gravitational lensing. [24, Fig. 16]	9
1.3	Illustration of the different ways weak lensing impacts an object. [31, Fig. 3]	10
1.4	An image showing two clusters merging. The green contours are from weak lensing and outline the clusters' mass distributions. The blue/red shows the X-ray emissions from the clusters' plasma clouds. [40, Fig. 1b]	11
1.5	Illustration of structure formation without (left) and with (right) the presence of dark matter, and its role in the CMB [23, Fig. 1]	13
1.6	ΛCDM 's theoretical predictions of anisotropies in the CMB through the l 'th multipole moment for different $\Omega_b h^2$ [44, Fig. 4] on the left, and the power spectrum of both the best fit ΛCDM model and observations on the right [3, Fig. 29.2].	16

1.7	Using the 95% CL ranges (yellow boxes) of primordial abundances of helium-4 (top), deuterium (middle), helium-3 (middle), and lithium-7 (bottom), and predictions from the standard model of BBN (colored bands), the subsequent results are compared to the CMB measurement of $\Omega_{\text{baryon}}h^2$ (pink/blue band) [3, Fig. 24.1].	19
1.8	Illustrating the Ω_{DM} and co-moving number density $Y = n_{DM}/s$ (number density vs. entropy ratio) vs. temperature or time. The solid line shows the theoretical prediction for the expected thermal relic abundance, while the surrounding color bands show deviations of this abundance by a factor of 10/100/1000 for the yellow/green/blue bands [68, Fig. 1].	21
1.9	Current limits from large TPC experiments on WIMP searches [3, Ch. 27]. .	22
1.10	Experimental limits of axion-photon (left) and axion-electron (right) couplings, with theoretical (KSVZ or DFSZ) expectations shown. [3, Fig. 90.1-90.2].	24
1.11	The expected cross section of our baseline iDM model, obtained from MadGraph. The model parameters used are $m_{A'} = 3m_{\chi_1}$, $\Delta m_\chi = 0.4m_{\chi_1}$, and $\alpha_D = 0.1$	28
2.1	Schematic of a e^+e^- crossing in SuperKEKB.	34
2.2	3D illustrations of the e^+e^- crossings in KEKB and SuperKEKB, highlighting the nano-beam scheme [91].	35
2.3	An illustration of the QCS magnet locations. The four magnets are QCS1RE, QCS1RP, QCS1LE, QCS1LP, QCS1RE, QCS1RP, QCS1LE, and QCS1LP. Their locations span 0.935m to 2.925m [97, Fig. 5.1]	36

2.4	A top view design drawing of the Belle II detector [98, Fig. 1.9].	37
2.5	A technical drawing of the PXD, illustrating the arrangement of the pixel cells and electronics, the Data Handling Processor (DHP) and Drain Current Digitizer (DCD) [98, 102].	38
2.6	Left) An illustration of a DEPFET cross-section. Right) View of a PXD module, containing the DEPFET pixel cells, switchers, DCDs, and DHPs. [101, Fig. 1, 2].	39
2.7	Illustrations of the SVD layers, depicted as light blue and pink, and PXD layers as black. The left image shows a cross-sectional view, while the right image is a view along the z axis. All dimensions are in mm.	40
2.8	Measurement of SVD's position resolution of the P (left) and N (right) side of layer 3 using $e^+e^- \rightarrow \mu\mu$ data [107, Fig. 15].	41
2.9	An illustration of two neighboring CDC cells, with the orange dots representing the sense wires and the white dots representing the field wires [108, Fig. 2].	43
2.10	Diagram illustrating the CDC wire configuration, alternating between axial and stereo configurations [107, Fig. 17].	43
2.11	Diagram illustrating position resolution (left), transverse momentum resolution (middle), and CDC dE/dx (left) performance using collision and cosmic data [107, Fig. 20-22].	44
2.12	Illustration of how the TOP detector measures a charged particle's Cherenkov angle. The solid arrow is the light that generates the prompt signature, while the dotted arrow is the delayed signal.	45

2.13	Using $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays in both MC simulation (blue) and data (red), we determine the K -ID efficiency and π mis-ID rate. Left: momentum dependence. Right: polar-angular dependence [107, Fig. 29].	45
2.14	Illustration of the principle behind ARICH's measurement of θ_{ch} (left) and the detector schematic with its main components (right) [107, Fig. 30]. . . .	46
2.15	ARICH's π/K separation performance from a $D^{*+}(D^0 \rightarrow K^-\pi^+)\pi^+$ control sample [107, Fig. 37]. The left plot shows K -ID efficiency vs. π fake rate (i.e., the ROC curve for K -ID). The right plot shows K -ID efficiency and π mis-ID as a function of track momentum.	48
2.16	An image of the CsI(Tl) crystal with the photodiode and preamplifier used in both Belle and Belle II [114].	49
2.17	An illustration summarizing the fitting procedure done by Belle II ECL's shaper module [127]. The left plot is for the ideal case and is used to highlight various features of the pulse-fitting procedure (T_S , P , A , t_0). The right plot is a typical plot from real data.	50
2.18	Plots showing a comparison in performance between Belle vs. Belle II electronics. Left: Pile-up noise dependence on polar angle for (old) Belle electronics and the (new) Belle II prototypes. Right: Time resolution as a function of single-crystal hit energy.	52
2.19	Left) A top view of a barrel RPC, showing the internal spaces and service ports for gas and HV. Right) An isometric view of an octant superlayer in the barrel [98, Fig. 10.4].	56

2.20	An illustration of an RPC superlayer of the KLM [134]. The two gas gaps and orthogonal readout planes, streamers left by a K_L^0 or muons will be read out.	57
2.21	Left) The endcap consists of one layer of scintillator strips. Right) An illustration of scintillator light detection in the strip with the WLS fibers embedded. Source: [98, Fig. 10.20]	58
2.22	Illustrating KLM performance for μ ID (left) and K_L^0 -ID from data [107, Fig. 51, 52].	58
2.23	A simplified illustration of Belle II's DAQ system.	62
2.24	Illustration of the steps in Belle II's tracking algorithm.	67
2.25	KLM efficiency monitoring from live data-taking on 13 March 2026.	68
2.26	Showcasing cluster shape analyses using PCA on simulated samples of K_L^0 using a particle gun, $e^+e^- \rightarrow s\bar{s}$, $e^+e^- \rightarrow B^0\bar{B}^0$, and $e^+e^- \rightarrow \mu^+\mu^-$. Rather than showing all the relevant variances, taking their ratio highlights this analysis tool's discriminative capabilities better.	70
3.1	Given a data set (illustrated by stars) and a model's prediction (which is represented by the central box), one can classify the performance of a given model by determining how many true/false signals were caught relative to the number of true/false background events which were labeled.	78

3.2	Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the number of CDC hits of the negatively charged track (with an electron mass hypothesis) for each model of a given m_{χ_1} . The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.	88
3.3	Showcasing the pointing angle ($\ln(1 - \cos \angle_{PA})$) distribution of background and light signal simulation.	89
3.4	Binning signal models by m_{χ_1} , we show missing momentum’s polar angle. The lower boundaries of the CDC and ECL, along with the $\sim 1^\circ$ gap between the forward and barrel ECL, are also indicated. While the left and right plots are of the same distribution, the left is a representation using box-whiskers, while the right is a violin plot.	90
3.5	In MC background simulation, the vertex mass using $\pi^\pm$ tracks is shown near the K_s^0 mass. Using a Gaussian plus linear fit, the width is measured to be $\sigma \sim 2 \text{ MeV}$	92
3.6	Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the number of CDC hits of the negatively charged track (with an electron mass hypothesis) for each model of a given m_{χ_1} . The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data in a violin plot. The K_s^0 mass window is also outlined, where the gray band indicates a $\pm 10\sigma$ width.	93
3.7	In MC background simulation, the vertex mass using $e^\pm$ tracks is shown near zero.	94

3.8	Illustration of the discriminating variable ΔR in the transverse plane for the two tracks that form a vertex.	94
3.9	Illustration of the discriminating variable ΔR for expected background samples.	95
3.10	Illustration of the discriminating variable ΔR for simulated signal grouped by m_{χ_1} and visualized by box-whisker plots. The left shows the full view, while the right shows an expanded view near $\Delta R = 0$ to better visualize the lighter iDM models.	95
3.11	For light and medium iDM models, optimizing a cut on ΔR based on the Punzi figure-of-merit.	96
3.12	Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the range of momenta p for each model with a given m_{χ_1} . The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.	97
3.13	Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the range of transverse momenta p_T for each model with a given m_{χ_1} . The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.	97
3.14	Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the range of momenta p for each model with a given $-\log(\epsilon)$. The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.	98

3.15	Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the range of vertex distances r_{V_0} for each model with a given $-\log \epsilon$. The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.	99
3.16	Given a light, medium, and heavy iDM model, we show the minimum muon ID of both tracks vs. the minimum electron ID of both tracks.	100
3.17	Given a light, medium, and heavy iDM model, we show the maximum μ ID of both tracks vs. the maximum e ID of both tracks.	100
3.18	From MC simulation, illustrating $\cos \angle_{V_0, \gamma}$, or how back-to-back the displaced vertex and photon are between heavy signal models and background.	101
3.19	Total extra event energy in an event for signal and background MC simulation.	103
3.20	For all the background scaled to $L = 364 fb^{-1}$, we fit this distribution with two crystal balls and an exponential tail. The background MC shown is the events that passed the criteria discussed in Sec. 3.2.2.	105
3.21	F-Score of the ensemble of BDTs used for background suppression. The results shown are the average over five trials for each category, with error bars indicating the standard deviation.	109
3.22	Punzi Figure-of-merit distributions for the BDT output distribution. Since each mass window contains several iDM models, we present the distribution of the n’t h quartile to assess how a cut will affect the majority of models.	110
3.23	Over the entire m_{χ_1} mass range, retention by the BDT ensembles (trained on certain ranges) of signal events post-reconstruction.	111

3.24	Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the range of vertex mass $m_{V_0 \rightarrow e^+e^-}$ for each model of a given m_{χ_1} . The left uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.	111
3.25	Examples of reconstructed signal models’ m_{recoil}^2 distribution.	112
3.26	Signal Window Interpolation at a confidence level of 90%.	113
3.27	Illustrating the impact of the pre-selection on the background MC simulation.	114
3.28	For the baseline model, we show the signal efficiencies at the various stages of the analysis. Left is post-reconstruction with preselection. Middle is of the reconstructed events, and how many signal events are retained by the BDT and other signal cuts. Right is the full reconstruction efficiency.	114
3.29	Impact of the BDT selection in the top row and the BDT selection and $\Delta R_{\gamma \rightarrow e^+e^-}$ selection on the bottom row.	115
3.30	Over the entire signal MC, the trigger efficiencies with respect to the L1 Trigger are illustrated as a function of m_{χ_1} , as the kinematics of the ISR photon and tracks are strongly dependent on m_{χ_1}	117
3.31	Over the entire signal MC, the trigger efficiencies with respect to the L1 Trigger are illustrated as a function of $-\log \epsilon$, to highlight the importance of the χ_2 lifetime.	118
3.32	Showcasing for each BDT category how the response is on a random selection of signal and background events. The x-axis shows each BDT’s response on the i ’th event minus the averaged response.	119

3.33	For various signal windows for a corresponding model with some $m_{A'}$, we integrate $\tilde{p}(x \tilde{\theta})$ within said signal window, and show the result from 300 trials.	120
3.34	For the various background components, we show the length of the error bars on Fig. 3.33 vs. $m_{A'}$, and their interpolations.	122
3.35	For a subset of models, we show our signal extraction scheme using the uniform PDF scheme and signal shape scheme on a toy dataset devoid of signal. The left shows the toy dataset with a given signal window, with the fit results written on the side. The right shows the upper limits from the left distribution and compares it to $N_{th} = \sigma \mathcal{L}_{run 1} \eta$.	127
3.36	From the toy shown in Fig. 3.35, we scan through our models and show the corresponding exclusion plot for the uniform (blue) and shape schemes (black outline).	128
3.37	For the average upper limit determined by our two schemes for each background expectation, we outline the Z-score vs. m_{χ_1} .	129
3.38	For each toy dataset, we compare the upper limit set by our uniform vs. shape and plot it as a function of the number of events in the signal window. For clarity, we split the distributions apart by mass.	129
3.39	For all our background scenarios (baseline, baseline $\times 3$, exponential inflated, and exponential inflated $\times 3$) averaged over 200 trials, we show their respective exclusion contours.	130
3.40	Linearity test with the exponential enhanced background levels $\times 3$	131
4.1	A Data/MC comparison of the minimum cluster to track distance using all our control samples.	135

4.2	A Data/MC comparison of the ZernikeMVA output using all our control samples.	135
4.3	A Data/MC comparison of the cluster energy-momentum ratio using the relevant control samples.	136
4.4	A Data/MC comparison of our back-to-back control sample of the E_{cm} and M of the displaced vertex.	137
4.5	A Data/MC comparison of our K_S^0 sample of displaced vertex mass using different track hypotheses.	138
4.6	A Data/MC comparison of $\Delta R_{\gamma \rightarrow e^+ e^-}$ using all our control samples.	138
4.7	A Data/MC comparison of the opening angle (where $\cos \angle_{open} = \hat{p}_{t_0} \cdot \hat{p}_{t_1}$ using our control samples.	139
4.8	A Data/MC comparison of our control samples illustrating the pointing angle where $\cos \angle_{PA} = \hat{p}_{V_0} \cdot \hat{dr}_{V_0}$	140
4.9	A Data/MC comparison of $\cos \angle_{V_0, \gamma}$ using our control samples.	141
4.10	A Data/MC comparison of the number of CDC hits for the negatively charged track using our control samples.	141
4.11	A Data/MC comparison of the number of CDC hits for the positively charged track using our control samples.	142
4.12	Back-to-back control sample data for the positively charged track using our control samples, with a cut on the number of CDC hits on the negatively charged track.	142

4.13 Floating only the yields, we fit our control samples using our background PDFs defined in Sec. 3.2.3.	143
4.14 A Data/MC comparison of E_{cms} using all our control samples.	144
4.15 A Data/MC comparison of recoil (from the ISR photon) squared mass (m_{recoil}^2) distribution using all our control samples.	144
4.16 A Data/MC comparison of total extra energy in an event using all our control samples.	145
4.17 A Data/MC comparison of missing energy in an event using all our control samples.	145
4.18 Illustrating the excess of events with $E_{cluster}/pc \sim 1$ for our converted photon and pointing angle sample.	146
4.19 A Data/MC comparison of θ_{miss}^{lab} in an event using all our control samples. .	146
4.20 A Data/MC comparison of m_{recoil}^2 in an event using all our control samples.	147
4.21 Applying the medium and heavy BDT and ΔR cut on the K_s^0 sample. . . .	147
4.22 Applying the light and medium BDT and ΔR cut on the $\cos \angle_{PA}$ sample. . .	148
4.23 Applying all the BDTs and ΔR cut on the converted photon sample.	148
A.1 Example diagram of a KAFKA cluster	195
A.2 Layout of alarm framework used at Belle II	196
A.3 Main panel GUI (bottom) and CDC GUI components (top)	197
A.4 KLM Detector Safety GUI	198

B.1	The efficiency of the upper backward octant of the barrel KLM as a function of run number.	201
B.2	Flow diagram of the DQM histogram analysis framework.	202
B.3	Diagram of how extrapolated hits from muon-like tracks are matched and used for KLM efficiencies.	208
B.4	An example of KLM’s efficiency ratio plot quickly highlighting an extremely inefficient layer, taken on 2026-03-30.	209
B.5	For a nominal physics run, the RPC timing distribution. In particular, the timing of the arrival time from the data concentrator electronics with respect to the L1 trigger.	211
C.1	The architecture of the critic (on the left) and generator (on the right) used in this work.	219
C.2	The left image is the raw image where $N = 60$ points are chosen from a 2D Gaussian centered with $\sigma = 1$. The right image is the digitized version of the left image, where if a point falls in a 1×1 bin, that is considered a hit. . . .	220
C.3	An overlay of 1000 images stacked for each prior.	221

C.4	An illustration of how a GMM would characterize prior “d.” Despite the digitization process causing the GMM to obtain standard deviations that are larger than the initial input, it still offers a baseline of what an ideal WGAN should reproduce. (a) A histogram of the square root of the eigenvalues of the covariance matrix using GMM on 3000 images. Because the 1σ curve forms an ellipse, the larger (smaller) eigenvalue is associated with the major (minor) axis of the ellipse. (b) A sample image from prior “d.” (c) Applying GMM on the sample image.	224
C.5	Given an image from a successfully trained WGAN using prior “d,” we illustrate the effects of the thresholding procedure for different values.	225
C.6	A comparison of the $g(t)$ of a GAN learning prior “b” with hyper-parameters $\{d_{\text{latent}}, n_{\text{critic}}\} = \{85, 5\}$. The maximum epoch (i.e., when $g(t) = 400$) is $t \approx 950$	226
C.7	A comparison of the $g(t)$ of a GAN learning prior “c” with hyper-parameters $\{d_{\text{latent}}, n_{\text{critic}}\} = \{85, 5\}$. The maximum epoch (i.e., when $g(t) = 400$) is $t \approx 920$	227
C.8	A comparison of the $g(t)$ of a GAN learning prior “b” with hyper-parameters $\{d_{\text{latent}}, n_{\text{critic}}\} = \{86, 5\}$. The maximum epoch (i.e., when $g(t) = 400$) is $t \approx 3200$	228
C.9	Multiple runs with $d_{\text{latent}} = 86$	229
C.10	Multiple runs with $d_{\text{latent}} = 87$	229
C.11	Multiple runs with $d_{\text{latent}} = 90$	230
C.12	A side-by-side comparison of $g(t)$ with varying n_{critic} and fixed $d_{\text{latent}} = 87$. . .	231

C.13 Samples of using DPGMM to determine the number of clusters from a generator with $d_{\text{latent}} = 85$ and $n_{\text{critic}} = 4$.	233
C.14 Using a data set of 250 generated images with target prior b , the left plot compares the number of clusters determined by the DPGMM with one determined visually. The right plot panel shows the number of clusters determined visually minus the number determined by the DPGMM. From this, we see that the DPGMM has an accuracy of $82.8 \pm 5.8\%$ while being within ± 1 clusters $98.8 \pm 6.3\%$ of the time with a slight bias towards underestimating the number of clusters.	234
C.15 A histogram of the eigenvalues of the covariance matrix, sorted by the major and minor values. The prior used here was prior “b.” When we compare this to Fig. C.4a, we see that while generated images look reasonable individually, the FoM illustrates that the generator’s approximation of the prior is not faithful.	235
C.16 Using prior “d” and a successfully trained GAN using $d_{\text{latent}} = 85$ and n_{critic} , we see that removing the complication of translational invariance improves the GAN’s ability to approximate the prior.	236
D.1 Mass accretion rate $\dot{M}$ vs. post-bounce time from M16 in solid red, our modified model in dashed blue, and numerical simulation in dashed-dotted black, all for the core collapse of the s20 progenitor of WH07 [327].	251

D.2	Parabolic fits to the total neutrino emission energies from Table 1 of Nakazato et al. (2013) [328], shown in solid colored lines (red, blue, black for ν_e , $\bar{\nu}_e$, and ν_x , respectively). The star symbols are obtained from our analytic model computed using the progenitor set of [319] and plotted at the corresponding final remnant baryonic mass.	252
D.3	Same as Fig. D.2 but for neutrino mean energies.	254
D.4	Neutrino luminosity L_{ν_e} versus post-bounce time from M16 (red), our modified model (blue), and the simulation (black). All are computed based on the s20 progenitor from WH07 [327]. The modified analytic model does significantly better in matching simulation results. Note the time-axis switch at $t = 0$	255
D.5	Comparison of neutrino luminosities predicted by our modified analytic model (dashed lines) to numerical simulations (dot-dashed lines) for three different progenitors, from left to right: n8.8 [325], z9.6 [326], and s11.2 [326]. All these predictions from the analytic models use parameters calibrated to the numerical simulation of the core collapse of the s20 progenitor (see text).	256
D.6	Comparison of the predicted bounce time (relative to the initial collapse) vs. $\xi_{2.5}$ with modified model in blue and [341] in black. The solid lines in the bottom plot represent the difference in bounce times between our model and Nakamura (i.e., $\Delta t_b = t_{b,\text{mod}} - t_{b,\text{Nakamura}}$).	259
D.7	Comparison of the predicted neutrino emission during the first 1 second post-bounce. The differences between our model and Nakamura's prediction are illustrated in the middle and bottom plots for $\nu = \nu_e + \bar{\nu}_e$ and ν_x , respectively.	260

List of Tables

1.1	Model configurations used in this analysis. The benchmark model is “Grid 10.”	27
1.2	The list of runs whose conditions we use to simulate signal. As these runs are representative of a given bucket, the number of events simulated in each run is determined by scaling the 10,000 events by $\mathcal{L}_{bucket}/\mathcal{L}_{\text{Run 1}}$	29
2.1	Machine Parameters of SuperKEKB [90, Table 2.2]. Note that parentheses values indicate the parameter’s value at zero beam current.	34
2.2	Summarizing the specifications of Belle and Belle II’s CDC [98, Table 6.1].	42
2.3	Main parameters of the solenoid coil [128, Table 11.2]	53
2.4	Main parameters of the iron structure [128, Table 11.1].	54
2.5	Cross sections and expected rates of the main interaction processes of $\Upsilon(4S)$ -resonance beam cross section at the target luminosity $\mathcal{L} = 6 \times 10^{35} \text{cm}^{-2} \text{s}^{-1}$ [107, Table 8].	59
2.6	Main trigger conditions and rates in the 2022 data-taking runs with $\mathcal{L} = 4.5 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ [107, Table 10].	62

3.1	Generic background samples, their primary physics generators, and the amount of data (in units of integrated luminosity) simulated for each sample. Along with the primary generators, secondary generators are also used to handle the decay of τ (TAUOLA [183]), the fragmentation of quarks (PYTHIA8 [184]), and the QED radiative corrections (PHOTOS [185, 186]).	86
3.2	Cuts on the ISR photon candidate	87
3.3	Cuts on the tracks used to form the displaced vertex (V_0).	88
3.4	Cuts for the formed displaced vertex.	90
3.5	Overall cuts on the signal event itself.	91
3.6	Model dependent cuts	108
3.7	Trigger lines of interest	116
3.8	Linearity results for baseline background scenario.	126
3.9	Linearity results for baseline background $\times 3$ scenario.	126
3.10	Linearity results for exponential background scenario.	128
3.11	Linearity results for exponential enhanced background $\times 3$ scenario the number of events.	132
C.1	Comparing the average epoch t where $g(t) = 400$ ($\overline{\text{Epoch}}_{max}$) with $d_{\text{latent}} \in [85, 88]$ for different n_{critic} . We see that there is a clear trend between $\overline{\text{Epoch}}_{max}$ and n_{critic} , implying that n_{critic} affects the rate of the discriminator learning and delays the point where overtraining occurs.	232
D.1	The adjustable parameters of M16's analytic model	243

D.2 Additional adjustable parameters in the modified analytic model	250
---	-----

List of Abbreviations

K_L^0 K-Long, or the long-lived neutral kaon

K_S^0 K-Short, or the short-lived neutral kaon

ALP Axion-Like Particle

ARICH Aerogel Ring Imaging Cherenkov (Detector)

AUC Area under the curve

basf2 Belle II Analysis Software Framework

BBN Big Bang Nucleosynthesis

BDT (Gradient-)Boosted Decision Tree

CDC Central Drift Chamber

CL Confidence Level, or the probability that a confidence interval will contain a proportion of a population

CMB Cosmic Microwave Background

CP Charge Parity

DAQ Data Acquisition (System)

DM Dark Matter

DT Decision Tree

EM Electromagnetism

HLT High Level Trigger
 iDM Inelastic Dark Matter
 KEK High Energy Accelerator Research Organization
 KLM K_L^0 and Muon (Detector)
 L1 Level-1
 LID Lepton ID (electron and muon)
 MC Monte Carlo
 PID Particle ID
 PXD Pixel Detector
 ROC Receiver Operating Characteristic
 SM Standard Model
 SVD Silicon Vertex Detector
 TOP Time of Propagation Detector
 TRG Trigger (System)
 UL Upper Limit, or the highest value of the confidence interval
 VXD Vertex Detector
 WIMP Weakly-Interacting Massive Particle

Chapter 1

Introduction

In this chapter, after introducing some fundamentals of the Standard Model of Particle Physics as they apply to ordinary matter, I summarize our current understanding of the enigmatic dark matter that lies outside the Standard Model’s scope, including evidence for its existence and ongoing searches. To conclude the chapter, I focus on a specific model that is the main focus of this thesis.

1.1 The Standard Model of Particle Physics

One of the crowning achievements of the 20th century is the development of the Standard Model of Particle Physics [1]. In this section, I describe the basic constituents of the Standard Model (SM), their interactions, and the limitations of the SM.

1.1.1 Constituents of the Standard Model

The fundamental particles in the SM are quarks, leptons, force carriers, and the Higgs [2].

Quarks and leptons are fermions (spin-1/2 elementary particles) that come in three generations of matter. As fermions, their properties are described via the Dirac equation¹, with a Lagrangian term of the form $\mathcal{L}_{\text{fermion}} \ni \bar{\psi}(\not{\partial} - m)\psi$. Properties that distinguish these fermions from one another are their electric charge, color charge, mass, and flavor. Leptons are the set of fermions with zero color charge and integer multiples of the unit e of electric charge (i.e., $0, \pm e$). Quarks are the set of fermions with non-zero color charge and fractional electric charge (either $\pm e/3$ or $\pm 2e/3$).

Quantum chromodynamics, a part of the Standard Model, describes how quarks interact with the strong nuclear force; this force confines quarks into colorless bound states called

¹Natural units are used: $c = 1$.

hadrons, providing a mechanism for why free-propagating quarks have not been observed. Hadrons with two (valence) quarks are known as mesons, with common examples being kaons (ex. $K^- = s\bar{u}$, $K^0 = \bar{s}d$, $\bar{K}^0 = CP(K^0) = s\bar{d}$, $K_L^0 = (K^0 + \bar{K}^0) + \varepsilon(K^0 - \bar{K}^0)/\sqrt{2(1 + |\varepsilon|^2)}$ with a small parameter ε providing a measure of CP violation) and pions (ex. $\pi^+ = u\bar{d}$, $\pi^- = CP(\pi^+) = d\bar{u}$, $\pi^0 = (u\bar{u} - d\bar{d})/\sqrt{2}$). Hadrons with three (valence) quarks are known as baryons, such as the nucleons ($p = uud$, $n = ddu$).

The force carriers are spin-1 bosons that represent three of the four fundamental forces of nature, which are electromagnetic (EM), strong nuclear, and weak nuclear². The gauge bosons, which carry the mentioned fundamental forces, are the photons (γ), gluons (g), and weak bosons ($W^\pm$ and Z).

The specifics of how we describe the interactions between the gauge bosons and the matter fermions manifest as Lie groups (groups that describe the symmetry of a system). In particular, the properties of these groups' generators — the Lie algebras — are used to capture how the SM fermions interact with gauge bosons via Noether's theorem. For EM, there is only one charge ($\pm e$), so the symmetry is $U(1)$ (i.e., symmetries of a circle) with only one generator. For the weak force, there is a weak isospin for interactions with $W^\pm$ and a weak hypercharge for interactions with Z , which can be captured via $SU(2)$ (i.e., isomorphic to the group of unit quaternions). Lastly, for the strong force, there are color charges that mediate the strong interaction, which are well-described by $SU(3)$ [3, Ch. 47, 48].

The properties of these interactions are encoded in the SM Lagrangian by Yang-Mills theories [4], which take the form $\mathcal{L}_{\text{gauge boson}} \ni \frac{1}{4}F^{a,\mu\nu}F_{\mu\nu}^a$ with $F_{\mu\nu}^a = D_\mu A_\nu^a - D_\nu A_\mu^a$, $D_\mu = \partial_\mu - igT^a A_\mu^a$, T^a representing the generators of the associated Lie algebra (given a matrix representation), coupling constant g , and A_μ^a as the vector potential. Using said vector potentials, we write interaction terms between fermions ψ_i and gauge boson A_μ^a as $A_\mu^a \bar{\psi}_1 X^\mu \psi_2$

²Gravity cannot be incorporated into the SM; string theory is needed to accommodate this force.

(the fermion species might be distinct). Here, X^μ is an arbitrary operator meant to indicate specific properties of said interaction. For instance, if the interaction with two fermions is of a ‘vector’ nature, then $X^\mu = \gamma^\mu$, where γ^μ are the gamma matrices; if an interaction with two fermions is a ‘vector-axial’ nature, then $X^\mu = \gamma^\mu(1 - \gamma^5)$ (e.g., weak interactions), where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ is a matrix that is used to distinguish a particle’s chirality.

The last piece of the SM is the spin-0 Higgs boson, which is associated with the Higgs field (ϕ) that gives mass to other fundamental particles (e.g., $\mathcal{L}_{\text{fermion } \psi} = \bar{\psi}\gamma^\mu D_\mu\psi + G_\psi\bar{\psi}\phi\psi$, where G_ψ is the fermion’s Yukawa coupling and $m_\psi \equiv G_\psi\langle\phi\rangle$). This mass-imparting mechanism is coined spontaneous symmetry breaking. Because the Higgs field leads to a nonzero vacuum expectation value ($\langle\phi\rangle$), this breaks the weak isospin symmetry and gives mass to the weak bosons.

1.1.2 Going Beyond the Standard Model

While many aspects of the SM have endured many precision tests [5–7], there are several phenomena or observations that the SM cannot explain. We list a few open questions below [8–10]:

- Why is gravity so much weaker than the other forces? (i.e. the Hierarchy Problem [11, 12]) [8, Ch. 2.3.1]
 - When comparing the coupling strength of matter with gravity vs. the other fundamental forces of nature, it is *many* orders of magnitude smaller (ex., the ratio between the Fermi constant and Newton’s constant of gravitation is 10^{24}). There are ensembles of theories (ex. extra dimensions [3, Ch. 85] and supersymmetry [3, Ch. 88, 89]) that try to address this. However, evidence supporting a particular theory has not yet been found.

- Why are there three generations of quarks and leptons? [13]
 - In the SM, both quarks and leptons seem to have only three generations. While there are theories with an additional ‘sterile’ neutrino [10], measurements of the CMB and limits from BBN have constrained the number of neutrino flavors to three and, therefore, the number of fermion generations to three.
- What is the neutrino mass, and where does it come from? [8, Ch. 2.3.2]
 - From neutrino flavor oscillations observed at SuperK [14] and SNOLAB [15], we know that neutrinos must have mass, a fact which is not predicted in the SM [3, Ch. 14]. Because we know neutrinos have mass, there are ongoing experiments that are trying to determine the differences between pairs of neutrino masses and, more ambitiously, the values of each of the neutrino masses [16, 17]. Additionally, understanding the mechanism that gives neutrinos mass is of interest. The SM sets their masses to zero by fiat, but the Higgs mechanism that gives rise to the masses of the other fermions is considered implausible for the extremely light neutrinos.
- Why is there so much more matter than antimatter, when the SM predicts matter and antimatter should be created in (almost) equal parts? [18, 19]
 - From the SM, the mechanism by which matter can behave differently than antimatter (i.e., CP violation) is the Cabbibo-Kobayashi-Maskawa (CKM) matrix for quarks, or the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix for neutrinos. While experiments like DUNE are aimed at measuring components of the PMNS matrix, the Jarlskog invariant of the CKM matrix [3, Ch. 12] is too small to account for all the matter seen in the observed universe. A related problem to

this is the Flavor Puzzle [8, Ch. 2.3.3], which does reference the strength of flavor oscillation from the CKM vs. PMNS matrix³.

- Why is there no CP violation in the strong sector? (Strong CP problem) [20, 21]
 - In the SM Lagrangian, there are no terms that forbid CP violation in the strong sector. However, this has not yet been observed in experiments. To suppress CP violation in the strong sector, some theorists propose the introduction of a new global anomalous symmetry, which is spontaneously broken at low energies and gives rise to a pseudo-Goldstone boson called the axion [3, Ch. 90].

The most important open question in this thesis is to understand the nature and properties of dark matter, which lie entirely beyond the scope of the SM.

1.2 Evidence for Dark Matter

While the Standard Model (SM) has been incredibly successful, we know this is not the end of the story. Aside from several open puzzles (see Ch. 1.1.2), we know the SM is not all-encompassing. In particular, we observe from astronomical and cosmological observations the presence of a gravitational source that affects several scales (from galactic to cosmological) and appears to emit no electromagnetic signature; this unseen source of gravity is referred to as dark matter (DM). These observations, which stand as our strongest evidence for DM, are the following: galactic rotation curves, gravitational lensing of bullet clusters, anisotropies in the cosmic microwave background (CMB), and Big Bang nucleosynthesis [22, 23].

³The flavor puzzle also references the mass ratios of the quarks, but that is irrelevant to this bullet point.

1.2.1 Galaxy Rotation Curves

The first major observation that led physicists to suspect the existence of DM was the observation of galaxy rotation curves.

When determining the velocities of stars throughout the galaxy, we can perform the following two calculations to obtain a rough estimate of the velocity distribution as a function of the distance r of a star from the galactic center. Near the center of the galaxy, luminosity curves and pictures of typical spiral galaxies suggest that the mass distribution is roughly uniform with density $\rho_G \approx 3M_G/4\pi R_G^3$ where M_G is the galaxy's visible mass and R_G is the galaxy's optical radius. For a far that is sufficiently far from the center and have encapsulated most of the visible mass M_G , we need only this mass when computing the star's velocity. The assumptions above in these two regimes provide the following velocity distribution curves:

$$v^2(r) = \begin{cases} \frac{G}{r} \int_0^r \rho_G dV' \approx \frac{4G\pi\rho_G}{3} r^2 & \text{when } r \ll R_G, \\ \frac{GM_G}{r} & \text{when } r \gg R_G. \end{cases} \quad (1.1)$$

As early as 1970, measurements of galaxy rotation curves [24–26] contradict this expectation outlined in Eq. 1.1. Illustrated in Fig. 1.1, the prediction (dashed line) for the $r \ll R_G$ regime is linear with v but underestimates the velocity; the prediction in the $r \gg R_G$ regime suggests the mass distribution is inconsistent with optical observations. While there exist examples of galaxies whose velocity curves agree with Eq. 1.1 [27, 28], more than 95% of galaxies' rotation curves have trends consistent with Fig. 1.1 [22]. A typical explanation for the observed galactic rotation-curve distributions we see is via a 'dark matter halo' that extends much farther than R_G (in Fig. 1.1, beyond $5R_G$).

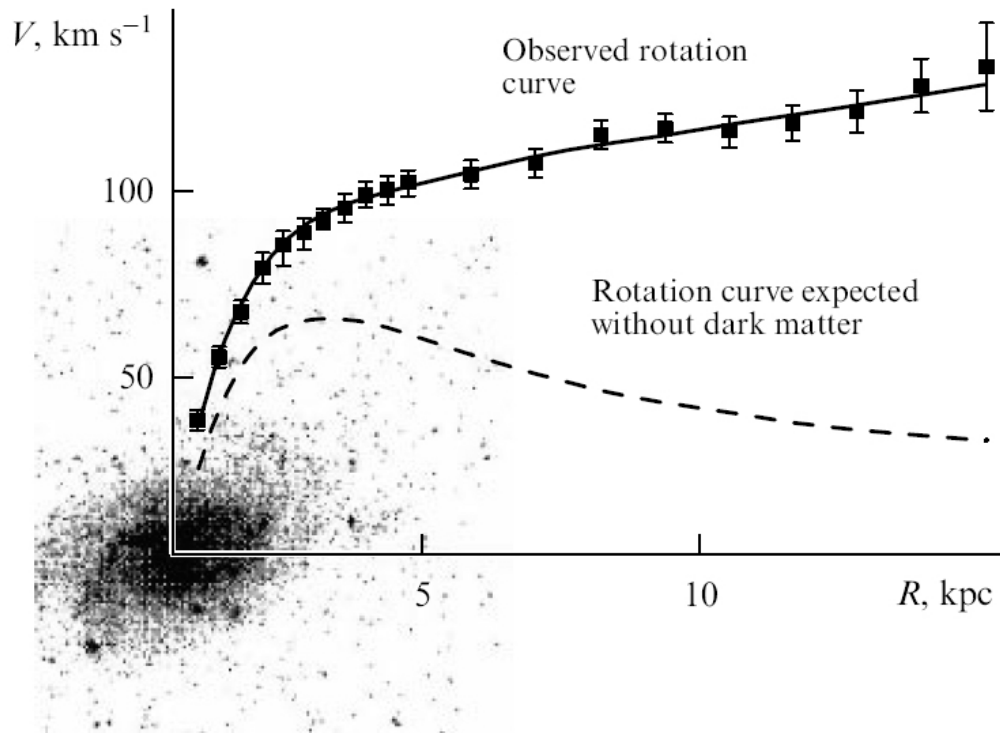


Figure 1.1: Rotation curves of the Triangle Nebula (M33) galaxy, suggesting the presence of (dark) matter that has no optical presence [24, Fig. 4]. The dashed curve is the prediction for $R_G \approx 3\text{kpc}$; the solid curve is the observation.

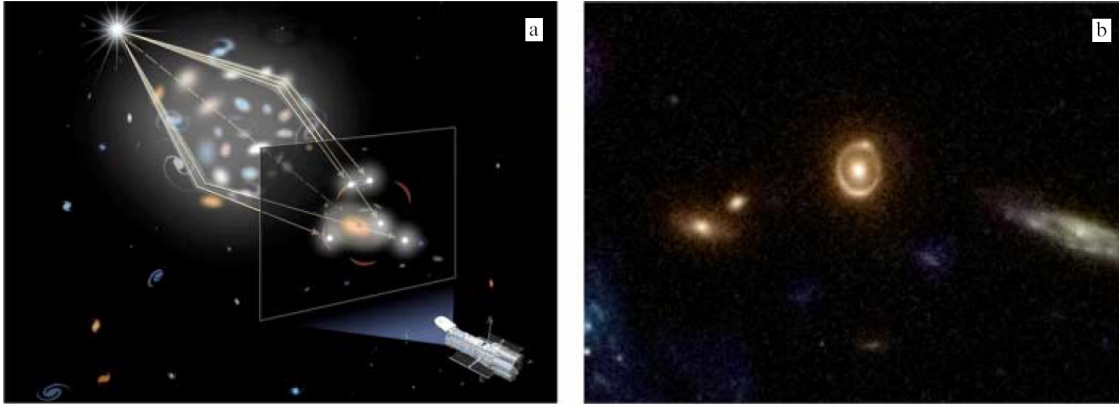


Figure 1.2: (a) An illustration of gravitational lensing. (b) An image from the Hubble Space Telescope of an almost ideal Einstein ring, a particular case of strong gravitational lensing. [24, Fig. 16]

1.2.2 Lensing of Galaxy and Galaxy Clusters

Another collection of observations that provides evidence for DM is from gravitational lensing. As the first test of Einstein’s General Relativity [29, 30], the bending of light due to celestial bodies has been used by astronomers to study intermediate-distance bodies that might be faint or emit no light (e.g., dim galaxies, dark matter). Fig. 1.2 illustrates this phenomenon and an example of an observation from the Hubble telescope [24]. The three types of this effect are “micro-lensing,” “strong lensing,” and “weak lensing” [24, 31].

Micro-lensing occurs when the gravitational field of an intermediate-mass object (e.g., a star or planet) bends the light from a distant star, analogous to how a lens would bend light. While this technique would be useful for finding DM within our solar system [32] or within others [33], it is of no use in measuring galactic-scale DM structures (e.g., DM halos).

Strong gravitational lensing happens when an intermediate massive object’s gravity bends light from a more distant light source, producing multiple images (e.g., Einstein ring [34], Einstein Cross [35]) [36, 37]. This observation technique is a powerful tool for astronomers to place constraints on cosmological constants [36] or as additional galactic-scale probes of

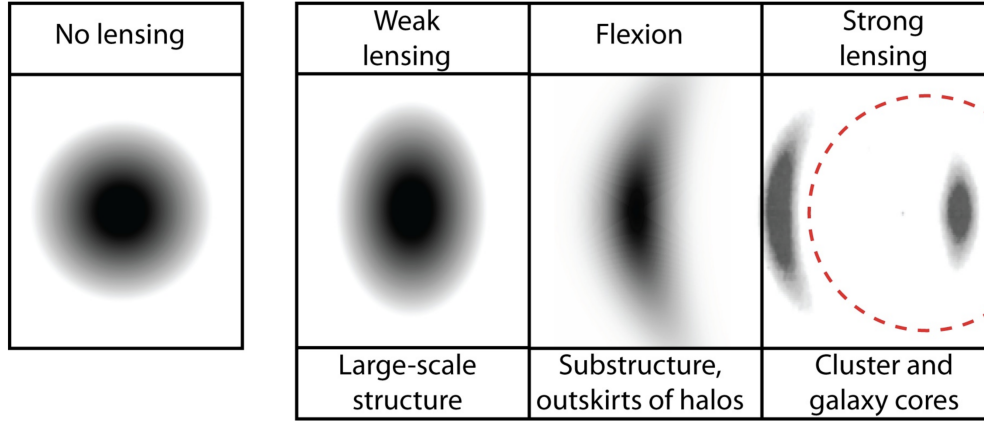


Figure 1.3: Illustration of the different ways weak lensing impacts an object. [31, Fig. 3]

DM [38].

Weak gravitational lensing happens when a massive object bends the light slightly but systematically distorts (or *shears*) the light from a collection of more distant objects [39]. Fig. 1.3 illustrates how weak lensing impacts various objects and how it differs from strong lensing. For large-scale structures, such as galaxy clusters, it provides an outline of the mass distribution based on the level of distortion.

Weak lensing, when combined with X-ray observations, provides conclusive evidence for the existence of DM. In the context of merging clusters, these provide two different observations. For example, merging cluster 1E0657-558 is depicted in Fig. 1.4. X-rays, coming from ionized gas in the clusters' intracluster medium, show hot plasma permeating throughout both clusters as they interact with one another in the central region of the image. On the other hand, weak lensing data imply that most of the matter content of the merging clusters is at the far left or right of the image.

When observations are coupled, they imply either the existence of an unseen (via electromagnetic radiation) but massive component (i.e., DM) in our universe or the necessity to significantly modify our understanding of gravity.

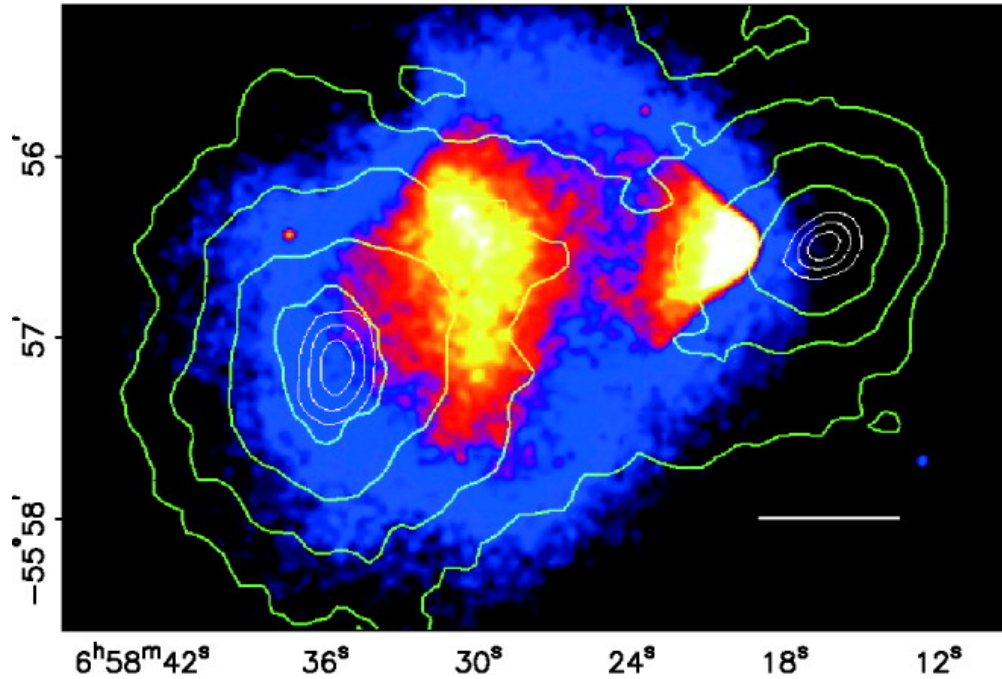


Figure 1.4: An image showing two clusters merging. The green contours are from weak lensing and outline the clusters' mass distributions. The blue/red shows the X-ray emissions from the clusters' plasma clouds. [40, Fig. 1b]

1.2.3 Density Fluctuations and Structure Formation

The next set of observational signatures of DM is from cosmology. Before delving into the observational evidence, I briefly summarize the relevant theory.

Minutes after the Big Bang, the universe had expanded enough that the fundamental forces were active and protons and electrons co-existed with photons in a photo-baryonic fluid.

To describe the role matter and energy play in this period, the Friedmann–Lemaître equations are used [41], namely

$$\left(\frac{\ddot{a}}{a}\right)^2 = -\frac{4\pi G}{3}(\rho + 3P/c^3) \quad (1.2)$$

where a is the cosmic scale parameter (roughly, the size of the universe $R(t)$ normalized by the current size today or R_0), $\ddot{a}$ is the acceleration rate of the universe, G is the gravitational

constant, ρ is the matter density of the photo-baryonic fluid, and P is the associated pressure.

For the energy-matter density, we note that the average density can be rewritten as

$$\rho = \sum_{n=-\infty}^{\infty} \rho_n a^{-n}, \quad (1.3)$$

where $n = 0$ corresponds to the cosmological constant, $n = 3$ corresponds to matter densities, and $n = 4$ corresponds to radiation.

Along with the assumption that the photo-baryonic fluid is a (near) perfect fluid such that $\omega = P/\rho c^2 = n/3 - 1$, then we can rewrite Eq. 1.2 as

$$\left(\frac{\ddot{a}}{a}\right)^2 = H_0^2 \sum_n \left(1 - \frac{n}{2}\right) \Omega_n a^{-n}, \quad (1.4)$$

where H_0 is the Hubble constant and Ω_n is a normalized matter density (in units of $\rho_c = 3H_0^2/8\pi G$).

About 378,000 years after the Big Bang, the universe had cooled enough for protons and electrons to finally combine to form hydrogen. As a result of this recombination process, photons are no longer coupled with matter, allowing them to propagate freely and making the universe transparent. The light emitted at the end of this recombination era is then redshifted after 13.7 billion years to form what we see today as the Cosmic Microwave Background (CMB).

Before recombination, since the existing electrons, protons, and photons co-exist within the photon-baryonic fluid, the density would be mostly flat across the universe. To describe the level of anisotropy in this fluid, we define the local density fluctuation $\delta = \rho(x, t)/\bar{\rho} - 1$. From Eq. 1.4 and the relevant equation of state, we note $\delta \propto a^{-3(1+\omega)}(t)$, where $\delta \propto a(t)$ in a matter-dominated universe [3, Ch. 22.4].

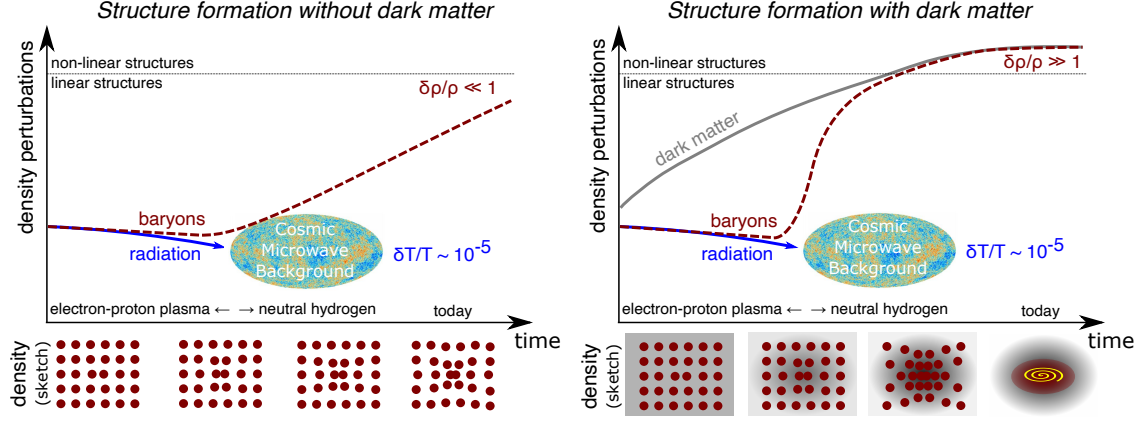


Figure 1.5: Illustration of structure formation without (left) and with (right) the presence of dark matter, and its role in the CMB [23, Fig. 1]

Using this and fluctuations in the CMB (see 1.2.4 for details), we can estimate the scale factors and their fluctuations. From anisotropies in the CMB, we see $\Delta T/T \sim 10^{-5}$ across the sky, which roughly translates to $\delta(t_{CMB}) \sim 10^{-5}$. Additionally, we know the current temperature of CMB relic photons to be ~ 0.2 meV, relative to the ~ 0.2 eV at their emission long ago, which yields $a(t_{\text{today}})/a(t_{CMB}) \sim 10^3$ (with an expanding universe, we expect this ratio to be > 1). Combining these two, we get $\delta(t_{\text{today}}) \sim \delta(t_{CMB})(a(t_{\text{today}})/a(t_{CMB})) \sim 10^{-2}$, which implies the universe should be largely homogeneous [23]. Considering that the universe has rich structure (e.g., galaxy structures, stars), we would expect $\delta(t_{\text{today}}) \sim 1$, implying that the density fluctuations at the time of the CMB are substantially underestimated. The conclusions of this argument are also highlighted in Fig. 1.5. Rigorous simulations [42, 43] further support these conclusions.

Another measurement that probes the nature of the early universe is the abundance of light elements (hydrogen to lithium). From our understanding of Big Bang nucleosynthesis (see 1.2.6), the relative abundance of light elements is fixed roughly at 180s after the Big Bang. Since there are no other production mechanisms for these light elements, the ratio of deuterium to hydrogen (D/H) or helium to hydrogen ($He-3/H$) provides a strong probe of

the blackbody photon density of the early universe ($\eta \propto \Omega_{\text{baryon}} h^2$) [3, Ch. 24.2-24.3].

If $\Omega_{\text{matter}} = \Omega_{\text{baryon}} + \Omega_{\text{DM}}$, then the additional DM component must play a role in observations — we will highlight DM’s role in the anisotropies seen in the CMB. In addition, there are independent methods to determine Ω_{baryon} , for example, via the Lyman α forest and through Big Bang nucleosynthesis.

1.2.4 Cosmic Microwave Background

As the first signature of the Big Bang, the CMB has been pivotal to our understanding of the early universe, as it is the remnant light from the Big Bang (in particular, post-recombination) that’s been redshifted to the microwave spectrum.

Anisotropies in this mostly flat microwave distribution play a vital role in understanding the matter distribution in this period. The anisotropies in the CMB can be described by taking angular-dependent CMB observations and modeling them using spherical harmonics:

$$T(\theta, \phi) = \sum_{l,m} a_{l,m} Y_{l,m}(\theta, \phi) \quad (1.5)$$

where $Y_{l,m}(\theta, \phi)$ are the spherical harmonics, $T_{l,m}$ is the temperature profile of the CMB, a_{lm} is the contribution of each harmonic, non-negative integer l , and integer $m \in [-l, l]$. To quantify it across the sky, the power spectrum is defined to be $D_l = l(l+1) \langle |a_{lm}^2| \rangle_m / 2\pi$. An illustration of the power spectrum from both theory and observation is illustrated in Fig. 1.6.

For the low moments, the dominant contributions to the anisotropy come from the Solar System’s relative motion relative to a nearly isotropic blackbody field for $l = 1$, the fluctuations of matter/energy density during recombination for $l \lesssim 100$, and the dominance of

dark energy at low redshift of said anisotropies from the last scattering (i.e., the integrated Sachs-Wolfe effect) for $l \lesssim 10$.

The relevant contributions for this discussion on dark matter is the baryon acoustic peaks at $100 \lesssim l \lesssim 1000$ and the damping tail at $l \gtrsim 1000$, where the rich structure seen at these moments are a consequence of gravity-driven acoustic oscillations that occur before the atoms in the universe become neutral [44]. In the period after the Big Bang, as the universe cools, the photo-baryon fluid decouples and recombines, producing photons, and the fluctuations in the matter density correspond to acoustic oscillations in the power spectrum. Since the recombination process is not instantaneous, there should be a dampening effect as $l \gtrsim 1000$.

The features of the acoustic oscillations are strongly tied to the baryon density $\Omega_{\text{baryon}} h^2$ relative to the total matter density $\Omega_{\text{matter}} h^2$. Within the photon-baryon fluid, gravitational forces from the fluid act as the driving force, while photon pressures resist these gravitational compressions. [44, 45]

Observations of the power spectrum have been made by WMAP [46] for $l \lesssim 1000$ and PLANCK [47] for $l \lesssim 2500$. When compared with theoretical predictions from baryon acoustic oscillations [44, 45] and other effects [3, Ch. 29], the models with $\Omega_{\text{baryon}} h^2 = 0.0224 \pm 0.0001$ and an additional (cold dark) matter component of $\Omega_{DM} h^2 = 0.120 \pm 0.001$ [47] are necessary. This comparison is illustrated in Fig. 1.6.

1.2.5 Lyman-alpha Forests

The Lyman- α forest is a region in space where a ‘forest’ of absorption lines is observed. Each absorption line mostly comes from neutral hydrogen gas (both hydrogen and deuterium), which allows one to obtain information from multiple redshifts ($z \sim 0 - 6$) [48, 49]. If one focuses on $z \sim 2 - 3$, the absorption lines in this regime correspond to spatial fluctuations

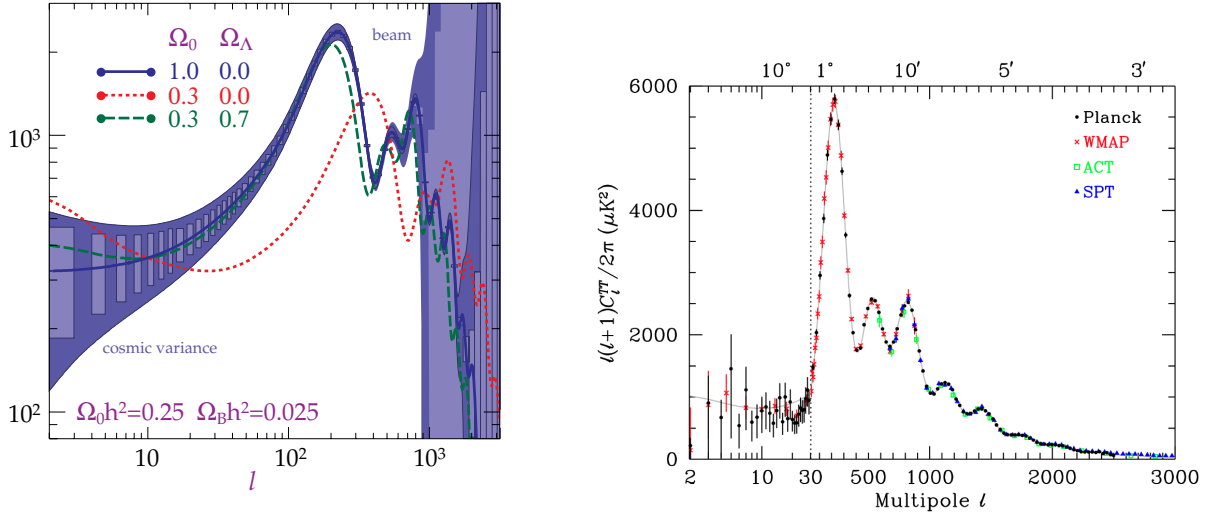


Figure 1.6: Λ CDM’s theoretical predictions of anisotropies in the CMB through the l ’th multipole moment for different $\Omega_b h^2$ [44, Fig. 4] on the left, and the power spectrum of both the best fit Λ CDM model and observations on the right [3, Fig. 29.2].

in the intergalactic medium and, thus, a measurement of the matter density at early times. Observational surveys, such as SDSS [50] and DESI [51], are able to constrain $\Omega_{\text{baryon}} h^2 = 0.02218 \pm 0.00055$ (for standard Λ CDM), which is consistent with the CMB measurements.

1.2.6 Big Bang Nucleosynthesis

While the CMB anisotropies provide a measurement of the baryon density Ω_{baryon} , it would be critical to our understanding to have another approach to confirm the CMB anisotropy measurements.

During the era of Big Bang Nucleosynthesis (BBN), electrons and other baryons separate from the Big Bang’s photons and eventually form light neutral atoms (ex., ^2H , ^3H , ^3He , ^4He , and ^7Li). The relative production rate of said light neutral atoms can be determined, given the SM and its interaction strengths. In particular, since the universe thermalizes sufficiently where baryons are no longer being produced, comparing the number density

of baryons (n_b) to the remnant photon density (n_γ) should provide estimates of Ω_{baryon} at the time of recombination. In particular, standard BBN relates the baryon-photon number density ratio η_{10} to Ω_{baryon} by

$$\eta \equiv n_b/n_\gamma = \rho_b/\rho_c = \left(\frac{3H_0^2}{8\pi G}\right) \Omega_{\text{baryon}} \approx 2.74 \times 10^{-8} \Omega_{\text{baryon}} h^2, \quad (1.6)$$

where G is Newton's gravitational constant and ρ_c is the critical matter density at present times [52].

Using various analytic [53, 54] and Monte-Carlo based methods [55, Sec. IV.A], one can relate measurements, such as the helium-4 mass fraction Y_p or primordial light-element abundance ratios (e.g., D/H , ${}^3\text{He}/H$), to η and other parameters (e.g., the number of neutrinos, the gravitational constant, the neutron lifetime). By measuring the primordial abundances of light elements (e.g., helium and deuterium), this can be used to provide an estimate of η .

Since D is difficult to find (as it is recycled into stars), one source to estimate its primordial abundance comes from distant and chemically unprocessed matter (i.e., regions of low metallicity). Datasets from large sky surveys [50, 56–58], such as the Sloan Digital Sky Survey (SDSS) [59, 60], near-infrared measurements of these low-metallic regions [61], and Lyman- α measurements [62] provide an measurement of $D/H = (25.47 \pm 0.25) \times 10^{-6}$.

For helium-4, similar spectra measurements from SDSS. The typical methodology is to determine the helium abundance of a system as a function of the abundance of a heavy element (e.g., O/H). By tracking this abundance across SDSS's collection of low-metallicity systems, one can determine that $Y_p = 0.245 \pm 0.003$.

With these measurements, the weighted average limits of η are $5.8 \leq \eta \times 10^{10} \leq 6.5$ at 95% CL, implying that the limit on the baryon density is $0.021 \leq \Omega_{\text{baryon}} h^2 \leq 0.024$ at 95% CL. Note that this is in line with the CMB measurement of $\Omega_{\text{baryon}} h^2$, further supporting the

claim that only a small fraction of the universe is baryonic matter, with dark matter making up the largest component of the matter density⁴. A summary of how these values relate to $\Omega_{\text{baryon}}h^2$ is illustrated in Fig. 1.7.

1.3 Alternatives to Dark Matter

While DM is a very attractive model in describing the aforementioned cosmological observations, Modified Newtonian Dynamics (MOND) serves as an alternative model in which structure formation and galactic velocity curves can be described [63–65]. While there are ongoing tests of the validity of MOND [66, 67], further discussions are outside the scope of this work.

1.4 Searches for Dark Matter

After reviewing evidence for DM from astronomy and cosmology, I now turn our attention to the range of terrestrial particle physics searches for DM. As a result, a major assumption that will be made is that all considered models of particle dark matter must have some coupling to SM (i.e., a ‘SM portal’). The goal for these particle DM searches is to extend the range of measurements of said couplings as far as possible (i.e., before ambient SM processes such as the neutrino floor, thus obscuring such signals).

Since there are various models of DM with corresponding search methods, this section is partitioned by model family: weakly interacting massive particles (WIMPs), axion-like particles (ALPs), and primordial black holes.

⁴Observant readers will notice that statements about the lithium abundance are absent. The reason for this, as shown in Fig. 1.7, is that its prediction of η would underestimate the baryon density. This discrepancy is known as the Lithium Problem and is still an active question in cosmology.

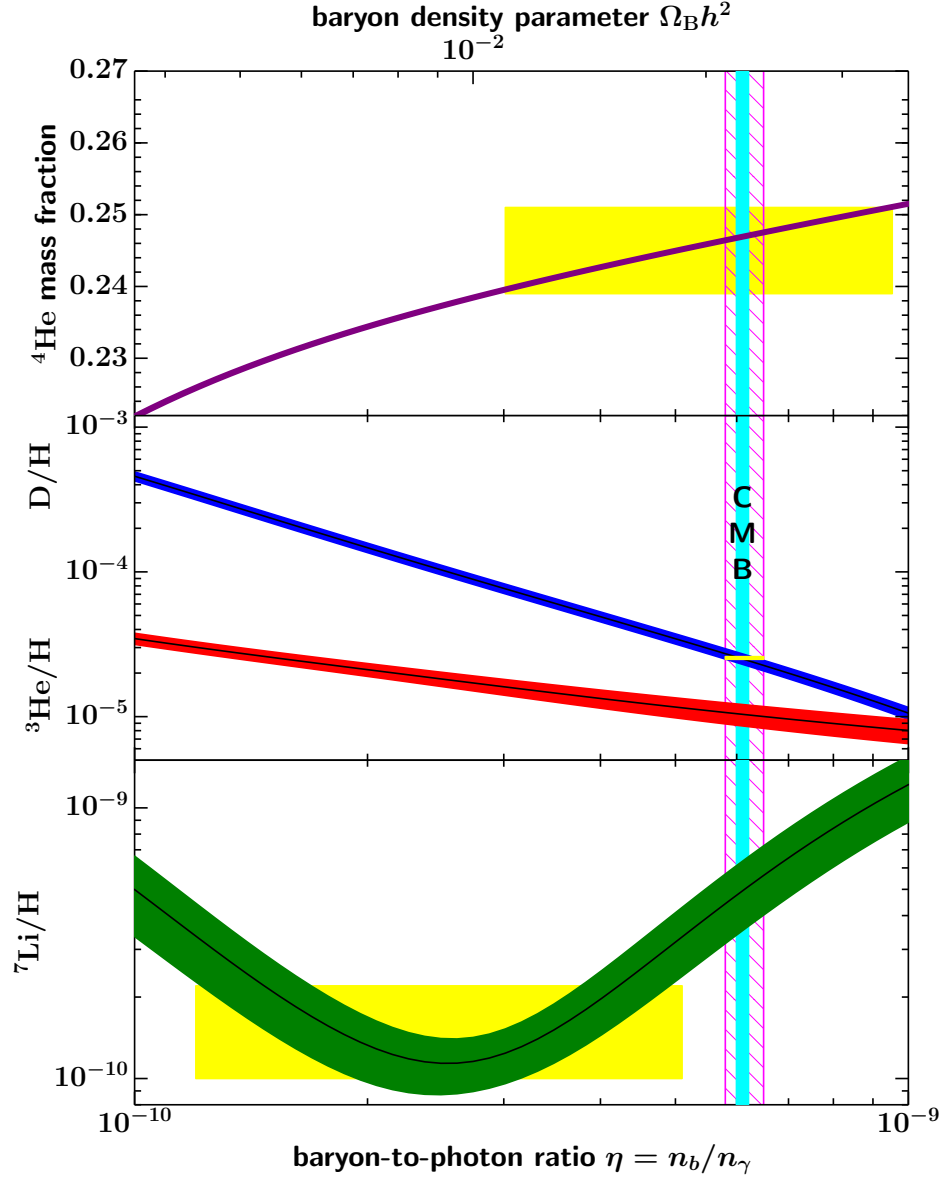


Figure 1.7: Using the 95% CL ranges (yellow boxes) of primordial abundances of helium-4 (top), deuterium (middle), helium-3 (middle), and lithium-7 (bottom), and predictions from the standard model of BBN (colored bands), the subsequent results are compared to the CMB measurement of $\Omega_{\text{baryon}} h^2$ (pink/blue band) [3, Fig. 24.1].

1.4.1 Weakly-Interacting Massive Particles

One of the most appealing models of DM is the weakly interacting massive particle (WIMP). In most cases (relevant to particle physicists), weakly interacting implies a non-zero coupling to the SM weak force rather than a characterization of its interaction with gravity. Considering the scenario where $m_{DM} \sim 100 \text{ GeV}/c^2$ (near $m_{W^\pm}$) and $g_{DM}/g_{\text{weak}} \sim 1$ and the number density of DM follows a Boltzmann distribution (diffusion coming from the expansion of the universe and annihilation coming from $X_{DM} + X_{DM} \rightarrow SM$), one obtains the thermal abundance of DM to be

$$\Omega_{\text{WIMP}} = \frac{m_{DM} n_{\text{today}}}{\rho_c} \sim \frac{m_{DM} T_0^3}{T_f M_{pl}} \langle \sigma_A v \rangle^{-1} \quad (1.7)$$

where $\rho_c = 3H^2/8\pi G$ is the critical density, T_f is the temperature of the universe at the time of DM freeze out, $T_0 = 2.7K$ or the current-day temperature of the universe, $M_{pl} = \sqrt{hc/G}$ is the Planck mass and $\langle \sigma_A v \rangle$ is the cross section of the aforementioned annihilation process. As illustrated in Fig. 1.8, these assumptions miraculously predict the expected thermal relic abundance seen in the CMB (see Ch. 1.2.4) [68]. For this reason, this remarkable coincidence is typically referred to as the WIMP miracle.

Motivated by the above concurrence, many detectors have been developed to search for WIMPs. The most popular type is a large liquid noble gas (e.g., Ar, Xe) detector where a WIMP would recoil (very infrequently) with a noble atom nucleus. The recoil of said nucleus would ionize or excite the surrounding atoms, thus creating a visible-light signature from the subsequent de-excitation that is easily detectable. In addition, the electrons from these ionizations would drift toward a sensor (assuming an ambient electric field to accelerate them), producing a complementary delayed signal. Examples of such detectors are LUX [69], LZ [70], PANDAX [71], XENON1T [72], DarkSide [73], and SuperCDMS [74]. Their

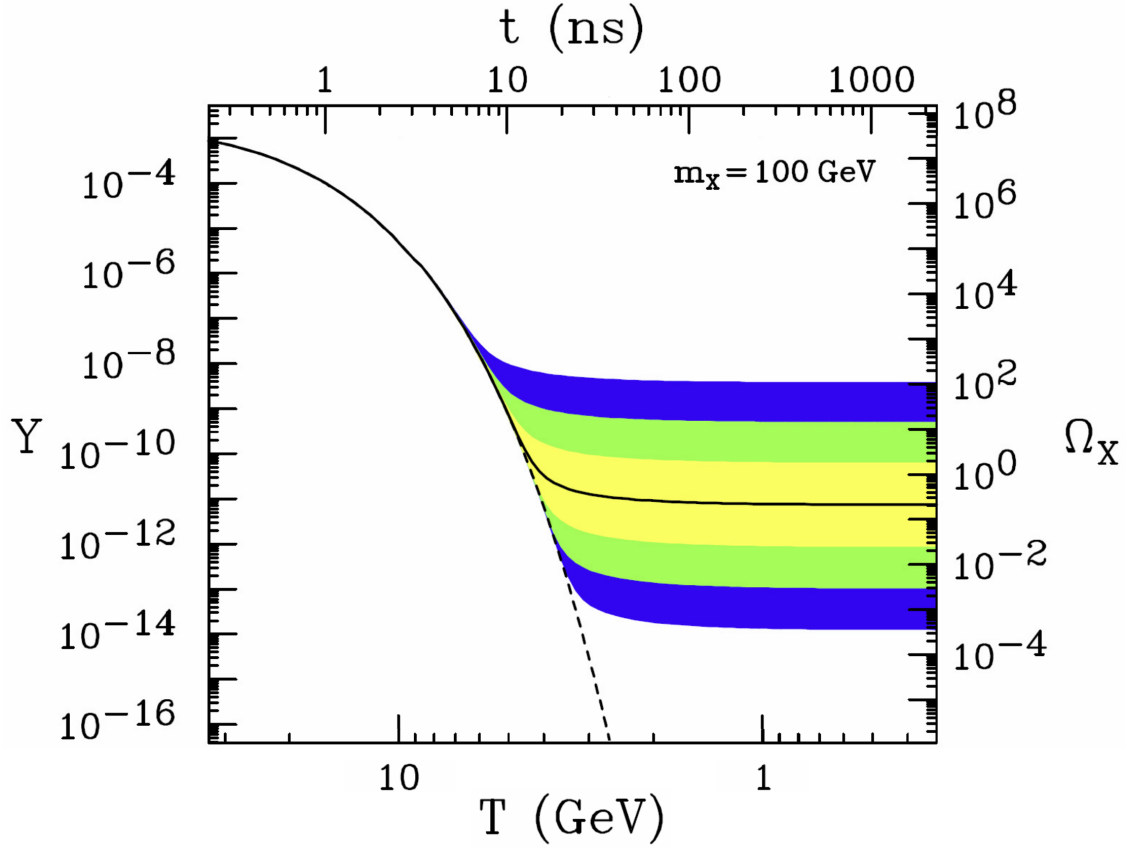


Figure 1.8: Illustrating the Ω_{DM} and co-moving number density $Y = n_{DM}/s$ (number density vs. entropy ratio) vs. temperature or time. The solid line shows the theoretical prediction for the expected thermal relic abundance, while the surrounding color bands show deviations of this abundance by a factor of 10/100/1000 for the yellow/green/blue bands [68, Fig. 1].

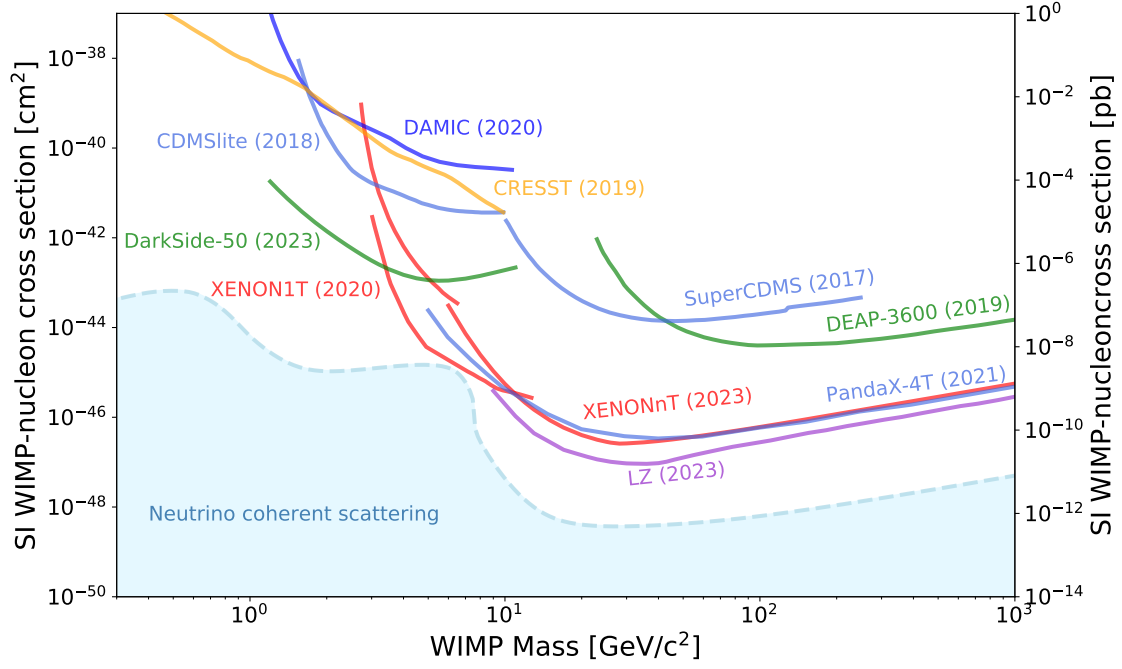


Figure 1.9: Current limits from large TPC experiments on WIMP searches [3, Ch. 27].

contributions to WIMP searches are illustrated in Fig. 1.9, ruling out large portions of the parameter space and approaching the regime where diffuse neutrinos will pose as backgrounds [3, Ch. 27].

Complementary searches of ambient WIMPs in our solar system (i.e., a halo of WIMPs) have also been performed. The idea is that if the WIMP distribution is centered on the center of our galaxy, then, as the Earth orbits the Sun, any interactions with WIMPs would exhibit annual fluctuations. Instead of liquid xenon/argon, this type of detector uses scintillating crystals (e.g., NaI(Tl)) whose nuclei recoil from DM interactions and emit scintillation light. To suppress backgrounds from cosmic rays or ambient radioactive material, these crystals are shielded (e.g., in copper boxes and tons of concrete) and placed underground. While DAMA/LIBRA is the most famous, results from COSINE-100 and ANAIS are consistent with those from liquid xenon/argon detectors.

1.4.2 Axion-Like Particles

As a solution to the Strong CP problem (see Sec. 1.1.2), axions were proposed as a pseudo-scalar that suppresses CP violation in the strong sector [20]. Motivated by the initial inspiration of said models, axion-like particles (ALPs) are a proposed DM candidate that couples to photons via the additional Lagrangian term $\mathcal{L}_a \ni -\frac{1}{4}g_{a\gamma\gamma}aF_{\mu\nu}F^{\mu\nu}$, where a is the ALP, $F_{\mu\nu}$ is the electromagnetic field-strength tensor, and $g_{a\gamma\gamma}$ is the coupling of the axion to two photons. In this context, the interaction mediating the SM portal is $a \leftrightarrow \gamma\gamma$. Thus, experiments searching for axions are looking for di-photon resonances. Examples of experiments that directly search for the axion are CAST and ADMX.

CAST, or the CERN Axion Solar Telescope, is an astroparticle physics experiment that is searching for axions originating from the sun [75]. The main instrument of CAST is a helioscope that uses a dipole magnet directed toward the Sun. Solar axions, produced through various processes (e.g., thermal photons produced deep in the sun via the Primakoff production mechanism [76]), would be converted to X-rays in the dipole magnet, where the pair of X-rays would be measured by an array of detectors. Each sub-detector is tailored for a particular energy range. Time projection chambers (TPCs) are used for low-intensity X-ray 1-10 keV signals. Finally, charged-coupled device (CCD) chips are used for more energetic X-rays. To estimate the background, a secondary helioscope is used that points away from the Sun.

Similarly, the Axion Dark Matter Experiment (ADMX) [77, 78] operates on a principle similar to that of CAST. Supplying a strong static magnetic field in a large cavity, virtual photons from said magnetic field allow galactic axions to convert into real photons, which then can be read out by an antenna within the cavity itself. Depending on the cavity resonant frequency, this would tune the energy of the available virtual photons, thereby changing the

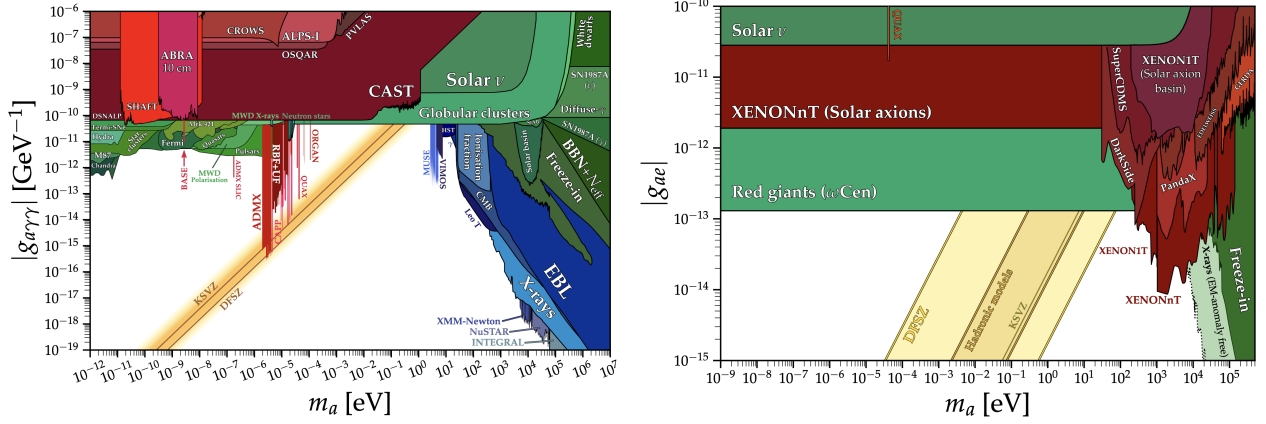


Figure 1.10: Experimental limits of axion-photon (left) and axion-electron (right) couplings, with theoretical (KSVZ or DFSZ) expectations shown. [3, Fig. 90.1-90.2].

sensitivity range for axion masses.

In addition to the axion searches relying on the axion's photon coupling, searches for axions can also rely on an electron-electron coupling. Similar to the photon coupling, the form of the Lagrangian for this interaction is $\mathcal{L}_{aee} \propto -\partial_\mu a \bar{\Psi}_e \gamma^\mu \gamma_5 \Psi_e$. While star formation can be used to set limits, the previously mentioned liquid noble gas detectors (i.e., XENON1T [79], SuperCDMS [80], PandaX [81]) can be used to search for $a + e \rightarrow e$ as well.

1.4.3 Primordial Black Holes

As opposed to the particle nature of DM described above, another viable and attractive candidate of DM is the existence of primordial black holes (PBH). Considering the size of the universe at periods near the Big Bang, the density of matter needed to form black holes would allow for a wide range of PBH masses, ranging from the Planck mass to even a solar mass (where $1M_\odot \approx 2 \times 10^{30} \text{kg}$). The wide range of PBH masses makes it an attractive model for describing scale-varying phenomena in DM.

However, since this type of model is quite the departure from previously mentioned models

and the rest of this work, as they eschew a distinct particle nature for DM beyond the SM content, please refer to reviews on the topic [23, 82]. For example, rather than elementary-particle detectors (which complement approaches from astronomy/cosmology), searches for PBH rely on gravitational-wave detectors [83].

1.5 Inelastic Dark Matter

One particular particle-like dark matter search that is relevant for accelerators is the search for inelastic DM (iDM) production [84]. As this will be the focus of this body of work, I elaborate here on the mechanism by which iDM models can produce an experimentally accessible signature of a displaced vertex and an initial state radiated (ISR) photon.

The Lagrangians used to describe this model are:

$$\begin{aligned}\mathcal{L}_\psi &= i\bar{\chi}_1\cancel{\partial}\chi_1 + i\bar{\chi}_2\cancel{\partial}\chi_2 + \frac{i}{2}g_X\hat{X}_\mu\bar{\chi}_2\gamma^\mu\chi_1 - \frac{i}{2}g_X\hat{X}_\mu\bar{\chi}_1\gamma^\mu\chi_2 - \frac{1}{2}m_{\chi_1}\bar{\chi}_1\chi_1 - \frac{1}{2}m_{\chi_2}\bar{\chi}_2\chi_2 \\ \mathcal{L}_{DM\leftrightarrow SM} &= -\frac{1}{4}\hat{X}_{\mu\nu}\hat{X}^{\mu\nu} + \frac{1}{2}m_{A'}^2\hat{X}_\mu\hat{X}^\mu - \frac{\epsilon}{2c_W}\hat{X}_{\mu\nu}\hat{B}^{\mu\nu}.\end{aligned}$$

The first part $\mathcal{L}_\psi$ describes two dark fermions with masses m_{χ_1} and m_{χ_2} . By convention, we stipulate $m_{\chi_2} > m_{\chi_1}$ and define $\Delta m_\chi = m_{\chi_2} - m_{\chi_1}$.

The second part $\mathcal{L}_{DM\leftrightarrow SM}$ describes a massive dark photon A' with mass $m_{A'}$, which provides a portal to the standard model (SM) via a coupling with $\hat{B}^{\mu\nu}$, or the component pertaining to the $U(1)_Y$ (weak hypercharge) component of the electroweak force. This kinetic mixing term describing the iDM and SM coupling is scaled by ϵ , normalized by $c_W \equiv \cos \theta_W$, where θ_W is the Weinberg angle. For the proposed model, we assume that there is no mass mixing between the dark photon and the standard model $\hat{Z}$ boson, which could arise from the scenario where either the SM Higgs is charged under the dark $U(1)_X$ or there a new scalar

field (e.g., dark Higgs) that is charged under both the SM and dark gauge group.

The dark photon interacts with both dark fermions (i.e. $\chi_1\chi_2 \leftrightarrow A'$ and $\chi_2 \rightarrow \chi_1 A'$) via a three-point interaction ($\hat{X}_\mu \bar{\chi}_2 \gamma^\mu \chi_1$ and its conjugate), scaled by g_X . For convenience, we define $\alpha_D = g_X^2/4\pi$.

The primary motivation for this dark matter model aims to explain the thermal relic abundance of dark matter inferred from anisotropies in the Cosmic Microwave Background (CMB) [47]. To be insensitive to the initial conditions of the early universe but allow for sufficient early-stage annihilations to be consistent with the observed abundance ($\Omega h^2 = 0.12$), this iDM model satisfies these requirements by (1) having a standard model coupling via the kinetic mixing between the standard model photon and dark photon A' , and (2) having the dominant annihilation channel be $\chi_1\chi_2 \rightarrow A' \rightarrow SM$. If the heavier χ_2 is also unstable with a short lifetime, the aforementioned channel scarcely occurs after recombination since χ_2 's become scarce; the DM in the late universe then consists solely of χ_1 's.

In addition to constraints from the observed abundance, there are other limits on the parameters of this model. The model requires $m_{A'} > m_{\chi_1} + m_{\chi_2}$ to suppress $\chi_1\chi_1 \rightarrow A'A'$ as a viable annihilation channel, such that the dark matter abundance is constant in the post-recombination era universe. For Δm_χ , experimental limits from the Large Electron-Positron (LEP) collider experiments set $\Delta m_\chi \lesssim 0.5m_{\chi_1}$ while observational limits from Big Bang Nucleosynthesis (BBN) on the χ_2 lifetime sets $\Delta m_\chi \lesssim 1\text{MeV}$. For m_{χ_1} , bounds on photon and neutrino temperatures from BBN and the CMB set $m_{\chi_1} \geq 10\text{ MeV}$ [85, 86]. For ϵ and α_D , LEP bounds $\epsilon \lesssim 3 \times 10^{-2}$ [63], and SuperCDMS bounds $\alpha_D \epsilon^2 \sim 10^{-3}$ for $m_\chi \sim 1\text{ GeV}/c^2$.

Belle II can search for this model by looking for events where $e^+e^- \rightarrow \gamma_{ISR} A' \rightarrow \chi_1\chi_2$ occurs, and the $\chi_2 \rightarrow \chi_1 A'^*$ decays within the Belle II volume with the virtual $A'^* \rightarrow \ell^+\ell^-$ forming a displaced vertex. The theoretical cross-section of this process, derived from the MadGraph

files used by Ref. [84], yields the cross-section table visualized in Fig. 1.11. We will use this as our baseline for the rest of the analysis.

Signal Generation

Following the model description above, MadGraph [87] is employed to generate events. The benchmark model we will focus on is outlined in Table 1.1 as “Grid10.” However, to avoid being completely reliant on a single model configuration, we also generate events using the other model configurations outlined in Table 1.1, which will be used to train the background-suppression boosted decision tree.

Model Nickname	$m_{A'}/m_{\chi_1}$	α_D	$\Delta m_{\chi}/m_{\chi_1}$	m_{χ_1} (GeV/ c^2)	$-\log \epsilon$ values
“Grid10”	3	0.1	0.4	[0.05, 0.08, 0.13, 0.2, 0.32, 0.5, 0.79, 1.3, 2.0, 3.2]	2.2 to 4.0 in increments of 0.2
“Grid11”	3	0.1	0.2	[0.13, 0.2, 0.32, 0.5, 0.79, 1.3, 2.0, 3.2]	2.2 to 4.0 in increments of 0.2

Table 1.1: Model configurations used in this analysis. The benchmark model is “Grid 10.”

The corresponding settings needed to produce the required signal MDSTs are:

- **BASF2 RELEASE-06-00-11** [88] (to be consistent with MC15rd background MC)
- 40,000 generated events in total (for more parallelization, generated in four 10,000-event chunks).
- Distribution of signal events across all the experiments and buckets (grouping of runs) in the Run 1 period by taking the ratio of the bucket’s fractional luminosity as summarized in Table 1.2.
- run-dependent beam background from the specific experiment and run

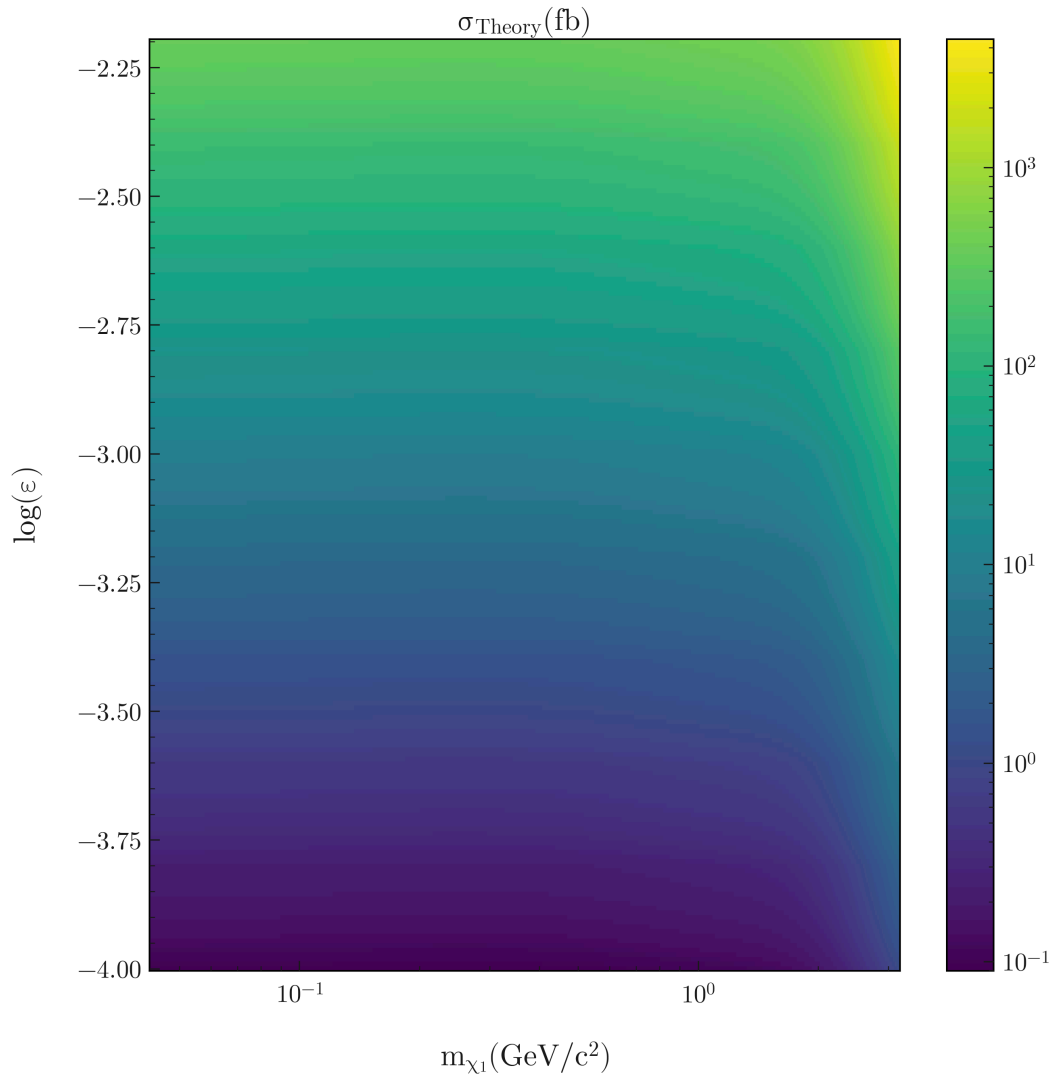


Figure 1.11: The expected cross section of our baseline iDM model, obtained from MadGraph. The model parameters used are $m_{A'} = 3m_{\chi_1}$, $\Delta m_\chi = 0.4m_{\chi_1}$, and $\alpha_D = 0.1$.

- Out of the events above, we choose $\Upsilon(4S)$ on-resonance runs only.
- conditions global tags `mc_production_mc15rd_a_<exp>_<bucket>`, `data_reprocessing_prompt`, and `online` when producing the MDSTs
- conditions global tags `neutrals_2024` (for photon systematics) and the standard analysis tags when applying our pre-selection
- Incorporation of A' , χ_1 , and χ_2 via the `PDG` module
- boosting the particles from the center-of-momentum system to the laboratory frame using the `BOOSTMCPARTICLES` module
- Use of the `GENERATEDVERTEXDISPLACER` module to displace the A' and χ_2 .
- Use of the `SMEARPRIMARYVERTEX` module to smear the primary vertex.
- Use of the `ACTIVATEPXDPIXELMASKER` and `ACTIVATEPXDGAINCALIBRATOR` to simulate the dead/masked pixels and detector gain of the PXD in simulation.

Due to the end-of-life of `LINUX CENTOS07`, `APPTAINER` was deployed to use pre-compiled versions of `BASF2` in producing MDSTs on KEKCC.

Table 1.2: The list of runs whose conditions we use to simulate signal. As these runs are representative of a given bucket, the number of events simulated in each run is determined by scaling the 10,000 events by $\mathcal{L}_{bucket}/\mathcal{L}_{\text{Run 1}}$.

Exp	Run	Luminosity (fb^{-1})	Ratio	# of Events
10	3869	8.628617	0.020130	201
12	1809	2.764822	0.006450	65
12	2731	10.321492	0.024079	241
12	3402	12.664138	0.029544	295

Exp	Run	Luminosity (fb^{-1})	Ratio	# of Events
12	4074	17.985799	0.041959	420
12	4309	5.034816	0.011746	117
12	5213	9.929873	0.023165	232
12	5388	13.265239	0.030946	309
14	1162	10.836159	0.025279	253
14	2014	5.664119	0.013214	132
16	708	10.294246	0.024015	240
17	171	10.715471	0.024998	250
18	181	8.921205	0.020812	208
18	751	9.087033	0.021199	212
18	893	8.788578	0.020503	205
18	1335	17.731445	0.041365	414
18	1815	18.131480	0.042299	423
18	2300	11.439058	0.026686	267
18	2692	15.801665	0.036863	369
20	644	3.787874	0.008837	88
21	264	19.788423	0.046164	462
22	364	11.867127	0.027685	277
22	449	20.193127	0.047108	471
24	866	20.620154	0.048104	481
24	1184	17.533473	0.040903	409
24	1571	22.974185	0.053596	536
24	2183	24.514591	0.057190	572
25	343	24.575142	0.057331	573

Exp	Run	Luminosity (fb^{-1})	Ratio	# of Events
26	1190	35.848111	0.083629	836
26	1430	18.947345	0.044202	442

We group signal models by mass in the following categories: light mass models ($m_{\chi_1} < 0.32 \text{ GeV}/c^2$), medium mass models ($0.32 \text{ GeV}/c^2 \leq m_{\chi_1} < 1.3 \text{ GeV}/c^2$), and heavy mass models ($m_{\chi_1} \geq 1.3 \text{ GeV}/c^2$). These boundaries were chosen because they are kinematically alike and to ensure sufficient statistics for MVA training.

Chapter 2

The Belle II Experiment

In this chapter, I provide an overview of the SuperKEKB accelerator and the Belle II experiment, outlining the goals of each subsystem and how said goals are realized.

2.1 SuperKEKB

SuperKEKB [89] is a particle accelerator, located in Tsukuba, Japan at the High Energy Accelerator Research Organization (KEK), that produces 7 GeV electrons and 4 GeV positrons, colliding them with a center-of momentum $\sqrt{s}$ near the $\Upsilon(4S)$ resonance, making it a second-generation B -factory for Belle II.

The machine parameters for the SuperKEKB system are summarized in Table 2.1. After producing electrons and positrons at their respective energies via a linear accelerator, they are stored in the high-energy ring (HER) and the low-energy ring (LER), respectively. While some quantities are necessarily the same between the two storage rings (e.g., circumference, crossing angle between the two beams, e^+/e^- number of bunches), various quantities are tuned for each ring's optimal running efficiency.

2.1.1 IR Design

To improve upon the luminosity, or event rate, seen by the first generation B-Factories KEKB [92] and PEP-II [93], a ‘nano-beam scheme’ [94] is used to focus the electron/positron beams down to a smaller interaction region via focusing magnets. Figs. 2.1 and 2.2 are illustrations of the beam crossing in this scheme.

The luminosity $\mathcal{L}$ of a collider is defined as:

$$\mathcal{L} = \frac{\gamma_{\pm}}{2er_e} \left(1 + \frac{\sigma_y^*}{\sigma_x^*} \right) \frac{I_{\pm} \xi_{y\pm}}{\beta_y^*} \frac{R_L}{R_{\xi_y}}, \quad (2.1)$$

	LER (e^+)	HER (e^-)	units
Beam Energy (E)	4.000	7.007	GeV
Circumference (C)		3016.315	m
Half Crossing Angle (ϕ)		41.5	mrاد
Emittance (ε_x)	3.2(1.9)	4.6(4.4)	nm
Emittance ratio ($\varepsilon_y/\varepsilon_x$)	0.27	0.28	%
Beta Function at IP (β_x^*/β_y^*)	32/0.27	25/0.30	mm
Horizontal Beam Size (σ_x^*)	10	11	μm
Vertical Beam Size (σ_y^*)	48	62	nm
Betatron tune (ν_x/ν_y)	44.53/46.57	45.53/43.57	
Momentum Compaction (α_c)	3.2×10^{-4}	4.55×10^{-4}	
Energy Spread (σ_ε)	$7.92(7.53) \times 10^{-4}$	$6.37(6.30) \times 10^{-4}$	
Beam Current (I)	3.6	2.6	A
Number of Bunches/rings (n_b)		2500	
Energy Loss/Turn (U_0)	1.76	2.43	MeV
Total Cavity Voltage	9.4	15	MV
Harmonic Number (h)		51200	
Synchrotron Tune (ν_s)	-0.0245	-0.028	
Bunch Length (σ_x)	6.0(4.7)	5.0(4.9)	mm
Beam-Beam Parameter (ξ_y)	0.0881	0.0807	
Luminosity (L)		8×10^{35}	$\text{cm}^{-2}\text{s}^{-1}$

Table 2.1: Machine Parameters of SuperKEKB [90, Table 2.2]. Note that parentheses values indicate the parameter's value at zero beam current.

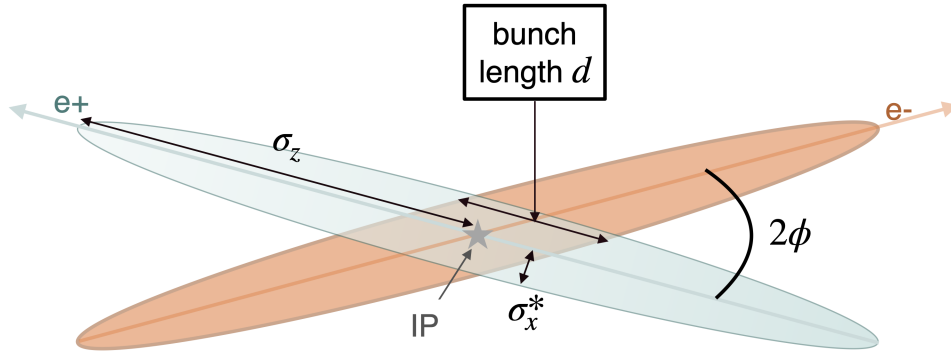


Figure 2.1: Schematic of a e^+e^- crossing in SuperKEKB.

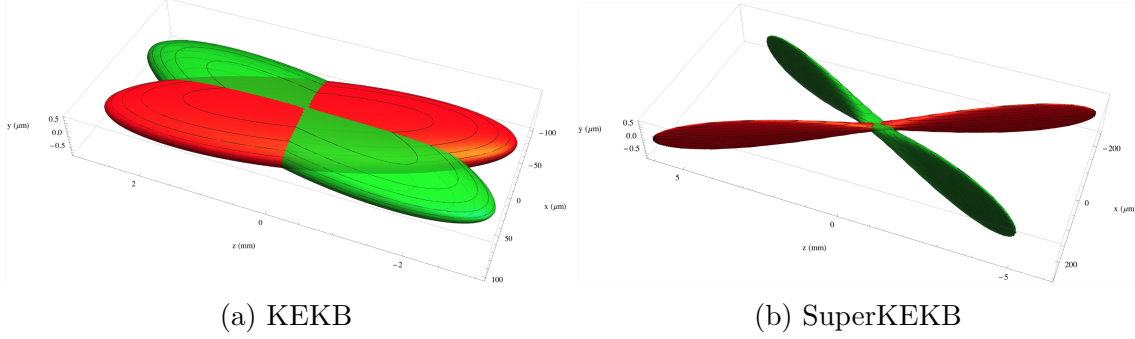


Figure 2.2: 3D illustrations of the e^+e^- crossings in KEKB and SuperKEKB, highlighting the nano-beam scheme [91].

where $\gamma_{\pm}$ is the Lorentz factor of the $e^{\pm}$, e is the electron charge, r_e is the classical radius of the electron, $\sigma_{x/y}^*$ is the horizontal/vertical beam size at the IP, $I_{\pm}$ is the current of the $e^{\pm}$ beam, ξ_y is the beam-beam parameter, which quantifies the strength of the beam-beam interaction, β_y^* is the vertical optical amplitude at the interaction point, R_L (R_{ξ_y}) is the reduction factor for the luminosity (vertical beam-beam parameter) which arises from the crossing angle (IP shape) [89, 95].

In the nano-beam scheme, the crossing angle ϕ is rather large (83 mrad at SuperKEKB vs. 22 mrad at KEKB), allowing the bunch length $d = \sigma_x^*/\phi$ and β_y^* to be smaller. Since Super-KEKB can keep the other parameters (e.g., ξ_y , I , R_L , R_{ξ_y}) mostly unchanged from KEKB, reducing d and β_y^* are the main avenues in improving upon KEKB's luminosity. The technology that enables the focusing of the $e^{\pm}$ beams is a superconducting quadrupole magnet (QCS) for each beam, which squeezes the beams to $\beta_y^* = 0.8$ mm [96], corresponding to a bunch height of only 50 nm in the vertical direction. The locations of the QCS magnets are illustrated in Fig. 2.3. Comparing this to KEKB's $\beta_y^* = 10$ mm, SuperKEKB is expected to gain a factor of 40 increase in instantaneous luminosity.

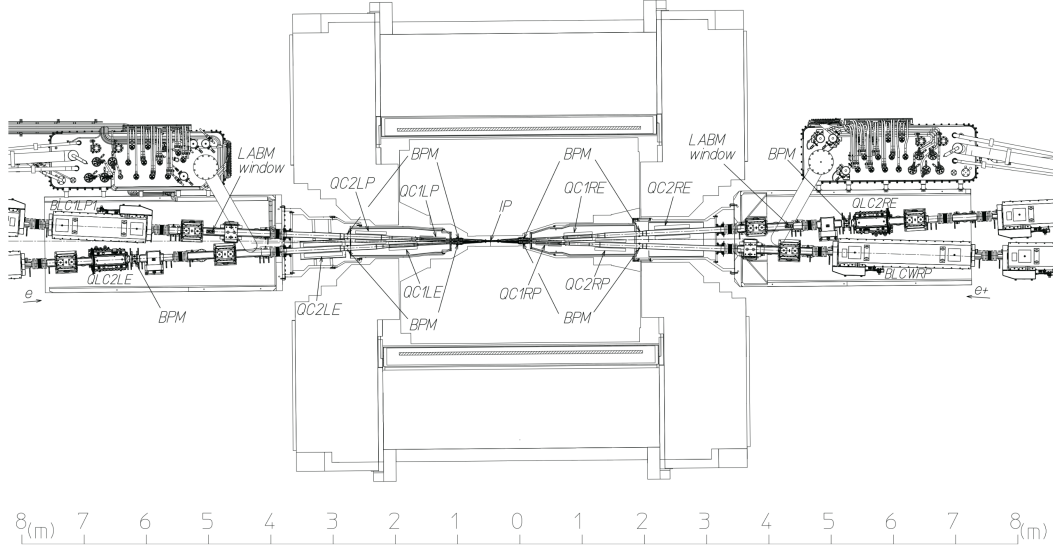


Figure 2.3: An illustration of the QCS magnet locations. The four magnets are QCS1RE, QCS1RP, QCS1LE, QCS1LP, QCS1RE, QCS1RP, QCS1LE, and QCS1LP. Their locations span 0.935m to 2.925m [97, Fig. 5.1]

2.2 Requirements of the Belle II Detector

To observe the e^+e^- collisions at SuperKEKB, the Belle II detector [98], illustrated in Fig. 2.4, was built to observe the outgoing particles with enough information to reconstruct corresponding particle decays. The following ingredients are needed for any particle collider experiment: a tracking system, particle identification (PID), and calorimetry. [99]. For the broader physics goals of Belle II [100], the tracking system, comprised of the Pixel Detector (PXD), Silicon Vertex Detector (SVD), and Central Drift Chamber (CDC), needs excellent vertex resolution at the IP (i.e. $\sim 50\mu\text{m}$), very high momentum resolution for charged particles with momenta greater than a few hundred MeV/c , and high momentum resolution across the remaining kinematic range of the experiment (i.e. up to $\sim 8\text{GeV}/c$). The Belle II PID system, made up of the Time-of-Propagation (TOP), Aerogel Ring-Imaging Cherenkov (ARICH), Electromagnetic Calorimeter (ECL), and K_L^0 and muon detector (KLM), needs high efficiency for identifying pions, protons, kaons, electrons, and muons over the full kine-

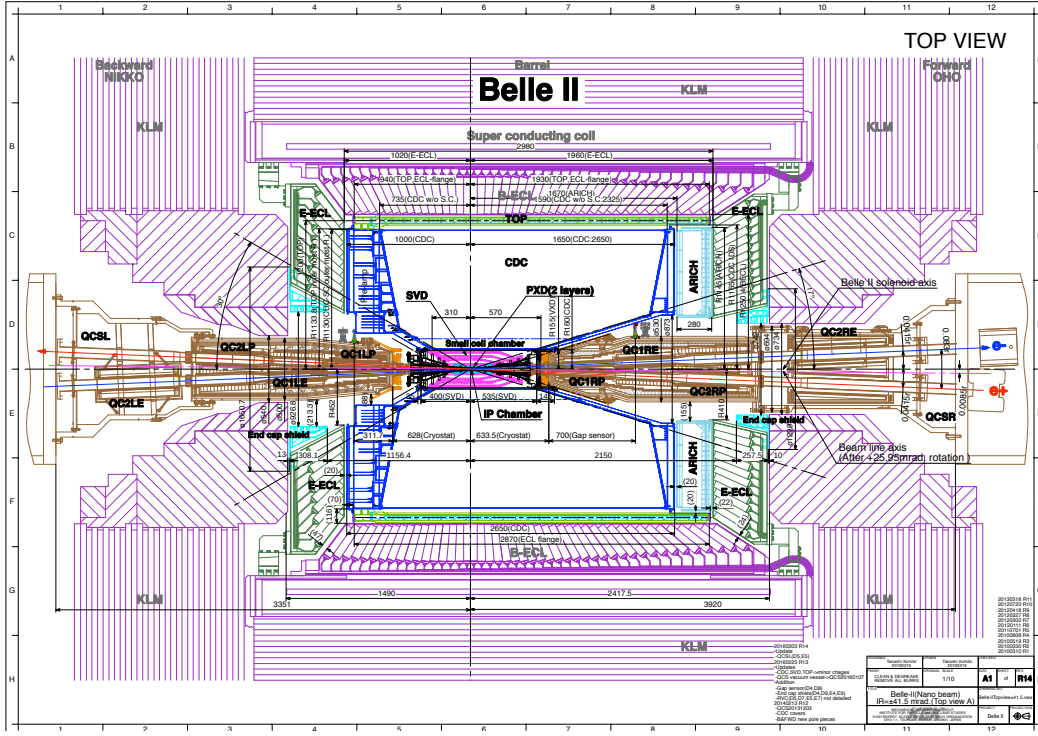


Figure 2.4: A top view design drawing of the Belle II detector [98, Fig. 1.9].

matic range (up to $4 - 5 \text{ GeV}/c$). Finally, the ECL needs precise photon energy and direction measurements from $\mathcal{O}(10 \text{ MeV})$ to $\mathcal{O}(8 \text{ GeV})$. Subsequent sections will outline each subsystem and how it meets these requirements.

Along with these requirements, Belle II is at the forefront of the luminosity frontier, so high event rates with a plethora of backgrounds are of concern. A brief description of Belle II's trigger system and data acquisition system will be provided as well to outline how Belle II addresses these challenges.

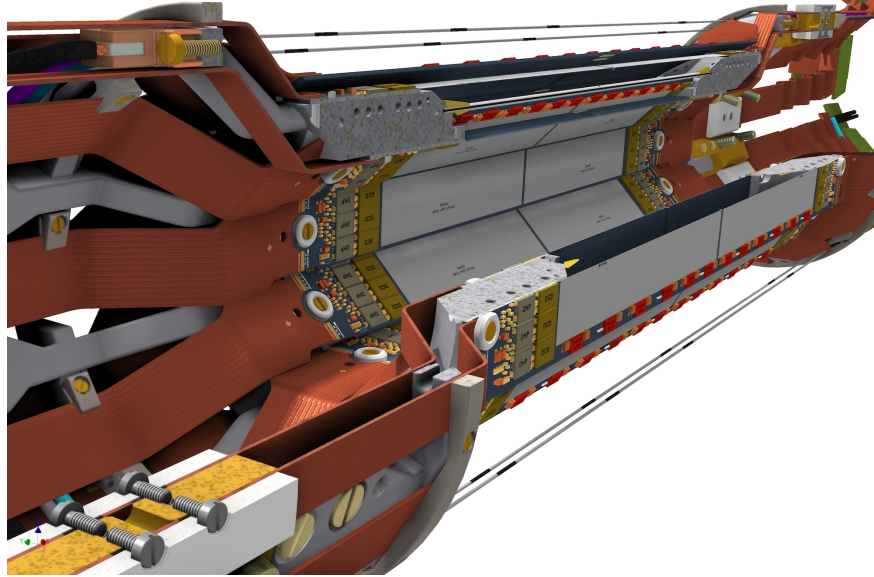


Figure 2.5: A technical drawing of the PXD, illustrating the arrangement of the pixel cells and electronics, the Data Handling Processor (DHP) and Drain Current Digitizer (DCD) [98, 102].

2.3 Pixel Vertex Detector

The PXD, illustrated in Fig. 2.5, is a two-layer pixel detector installed at radii of 14 and 22 mm from the z axis (between the two beam lines). The original design of the inner (outer) layer contains 8 (12) ladders, each a pair of mirrored modules glued face-to-face. A module, illustrated in Fig. 2.6, consists of a collection of Depleted P-Channel Field Effect Transistor (DEPFET) pixels, which is read out with a set of switchers, Drain Current Digitizers (DCDs), and Data Handling Processors (DHPs) [98, 101].

DEPFET pixels, whose cross-section is illustrated in Fig. 2.6 (left), contain a p-channel MOSFET. The MOSFET is integrated onto a silicon detector substrate, acting as the bulk material. The initial state of the DEPFET is that the silicon bulk is fully depleted by applying a large reverse-bias voltage to the backside p^+ contact on the bottom, which forms a potential minimum—referred to as the internal gate—located roughly 1 μm beneath the

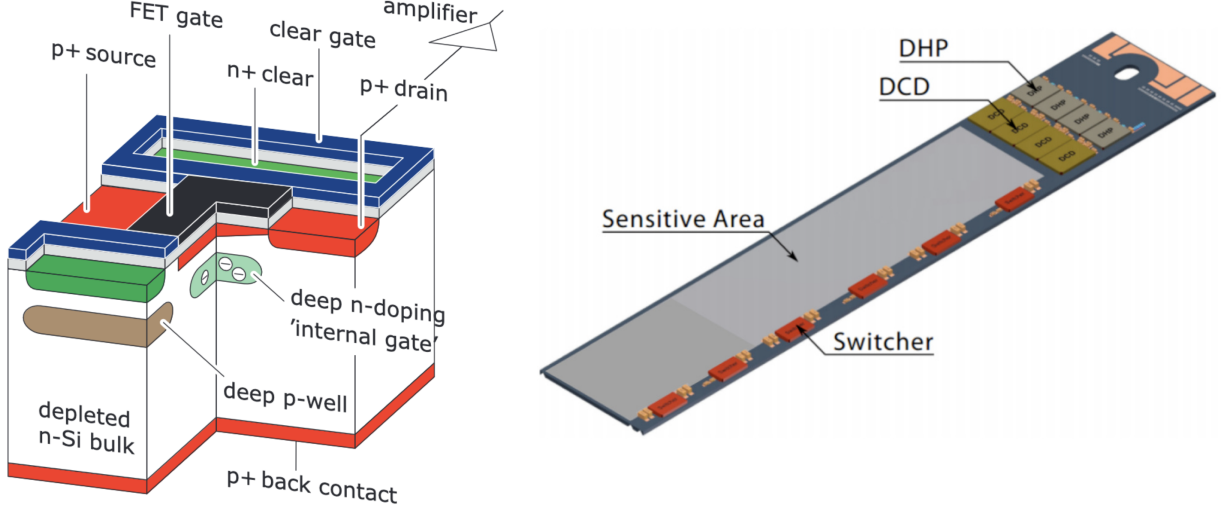


Figure 2.6: Left) An illustration of a DEPFET cross-section. Right) View of a PXD module, containing the DEPFET pixel cells, switches, DCDs, and DHPs. [101, Fig. 1, 2].

transistor channel. As an incident particle passes through the bulk, electron-hole pairs are created; holes drift to the back contact, and electrons accumulate towards the internal gate. The accumulated charge in the internal gate modulates the transistor channel current. When the DEPFET receives a ‘clear’ signal, the transistor lowers the potential barrier, allowing the stored electrons to flow to the clear contact, thereby reading out the charge as a hit and resetting the pixel. Note that a deep p^+ well is located beneath the clear implant; this well prevents electrons injected from the clear region from back into the internal gate. Finally, a n^+ ‘clear’ contact helps remove any remaining trapped charges.

DEPFET pixels are used at Belle II for the following reasons. An important advantage of DEPFET pixels is the very small capacitance of the internal gate, which leads to low electronic noise and therefore a high signal-to-noise ratio. Additionally, the clear mechanism is non-destructive, allowing the pixel to be reset repeatedly without degrading the signal measurement. On a PXD module, illustrated in Fig. 2.6 (right), the readout area is a matrix of 250×768 DEPFET pixels covering an area of approximately 8 cm^2 . The board’s

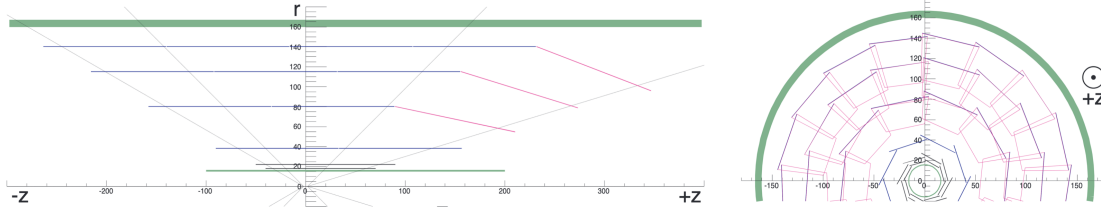


Figure 2.7: Illustrations of the SVD layers, depicted as light blue and pink, and PXD layers as black. The left image shows a cross-sectional view, while the right image is a view along the z axis. All dimensions are in mm.

electronics components are the following: switchers, along the side of the board, which control the DEPFET gate and clear signals; DCDs, which digitize the drain current from DEPFET pixels; and DHPs, which perform zero suppression, control the steering signals to other readout components, and manage operational modes and monitoring.

From Belle II's first data taking period (up to 2019), an analysis of the transverse impact parameter (d_0) of $e^+e^- \rightarrow \mu^+\mu^-$ scattering events found the resolution of said impact parameter to be $14.1 \pm 0.1 \mu\text{m}$ [102], which allows Belle II to obtain precise lifetime measurements twice as good as Belle's (e.g., the D_s^+ lifetime [103]).

2.4 Silicon Vertex Detector

The SVD, the next four layers of Belle II's VXD, is a collection of double-sided silicon strip detectors (DSSDs). As illustrated in Fig. 2.7, the SVD covers Belle II's angular acceptance of $17^\circ < \theta < 150^\circ$ at radii of 39, 80, 103, and 135 mm from the z axis [104, 105].

The physics requirements of the SVD are as follows: provides data to help extrapolate tracks reconstructed in the CDC inward to the PXD with high efficiency for vertexing; reconstruct low-momentum tracks (i.e., $\mathcal{O}(10 \text{ MeV}/c)$); reconstruct K_s^0 that decay outside the PXD; operate with low dead-time in high beam-background environments expected from

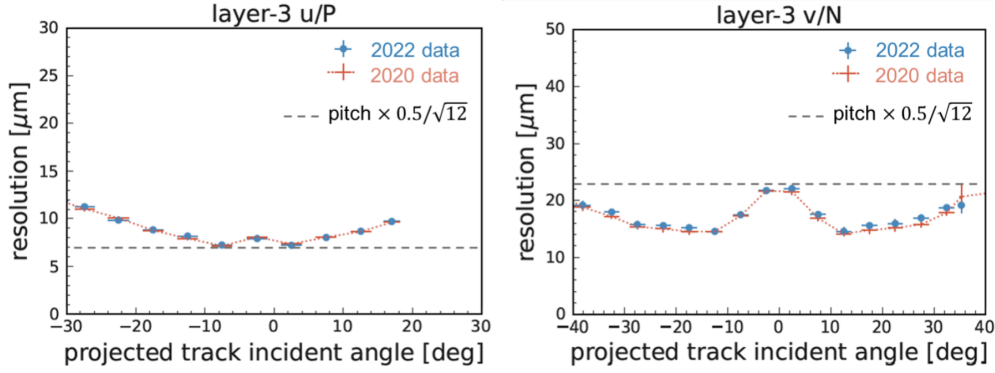


Figure 2.8: Measurement of SVD's position resolution of the P (left) and N (right) side of layer 3 using $e^+e^- \rightarrow \mu\mu$ data [107, Fig. 15].

SuperKEKB's full luminosity (i.e., up to 30 kHz trigger rate) [98, Ch. 5].

To achieve this, DSSDs are used. The principle behind DSSDs is analogous to that of a reverse-biased diode. A charge particle passing through the silicon strip will generate electron-hole pairs along its trajectory. If the strip is biased, electron-hole pairs would drift toward opposite electrodes. Stacking perpendicular strips, thereby making the detector double-sided, enables 2D hit information [106].

Collision data using $e^+e^- \rightarrow \mu^+\mu^-$ events were used to determine the position resolution of the SVD, as illustrated in Fig. 2.8. In addition to the $\mathcal{O}(10)\mu\text{m}$ position resolution, the corresponding timing resolution is < 3 ns, allowing for a 50% removal of off-time background hits not from the e^+e^- collision.

2.5 Central Drift Chamber

The next Belle II subdetector is the central drift chamber (CDC). The main roles of the CDC are reconstructing charged tracks and measuring their corresponding momenta; providing PID information using energy loss (dE/dx) within its gas volume; and sending quick,

Variable Name	Belle	Belle II
Inner cylindrical radius (mm)	77	160
Outer cylindrical radius (mm)	880	1130
Innermost sense wire radius (mm)	88	168
Outermost sense wire radius (mm)	863	1111.4
Number of layers	50	56
Number of sense wires	8,400	14,336
Gas	He:C ₂ H ₆	He:C ₂ H ₆
Sense Wire Diameter (μm)	30	30

Table 2.2: Summarizing the specifications of Belle and Belle II’s CDC [98, Table 6.1].

reliable trigger signals for charged particles. Although the general structure follows that of its predecessor in Belle, the readout electronics and trigger systems were improved to accommodate the expected event rates.

The general structure of the CDC is summarized in Table 2.2. The CDC covers a cylindrical radius of 160 – 1130 mm from the z axis with a length of $\Delta z \approx 2.65$ m (0.735 m in the backward region, and 1.59 m in the forward region). The volume is partitioned into a collection of 14,336 cells with a gold-plated tungsten wire at the center of the cell that senses the presence of charged particles and a collection of aluminum wires at the cell’s perimeter; the potential difference between the anode (tungsten wire) and cathode (aluminum wire) generates an electric field within the cell. The layout of a CDC cell is illustrated in Fig. 2.9. As a charged particle passes through a cell and ionizes nearby gas molecules, the liberated atomic electrons drift towards the sense wires, generating a signal where drift times and dE/dx can be measured, while the positive ions drift much more slowly towards the field wires.

The wires are arranged into two different orientations: axial (parallel to the z axis) and stereo (skewed by 40-70 mrad relative to the z axis). As illustrated in Fig. 2.10, the wires alternate between these two orientations in 9 different ‘super-layers.’ While the drift times in all the layers can be used for position measurements in the r – ϕ plane, the alternating wire

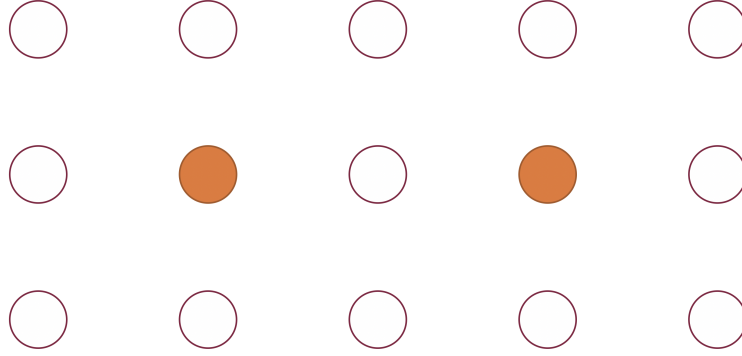


Figure 2.9: An illustration of two neighboring CDC cells, with the orange dots representing the sense wires and the white dots representing the field wires [108, Fig. 2].

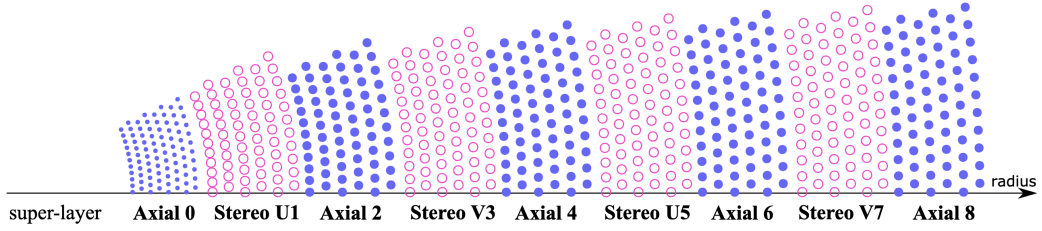


Figure 2.10: Diagram illustrating the CDC wire configuration, alternating between axial and stereo configurations [107, Fig. 17].

orientations allow for precise z -position measurements.

A demonstration of the CDC's performance is illustrated in Fig. 2.11. For the position resolution, tracks from B decay events in e^+e^- collision events were used to measure a resolution of $100 \mu m$ at the middle part of the cell. For the transverse momentum resolution, both cosmic muons and $e^+e^- \rightarrow \mu^+\mu^-$ events were used. Finally, for the dE/dx measurements, various control samples from e^+e^- collision data were used to demonstrate that measurements of all the charged final state particles seen at Belle II (p , $e^\pm$, $\mu^\pm$, $K^\pm$, $\pi^\pm$) align with predictions made via the Bethe-Bloch formula.

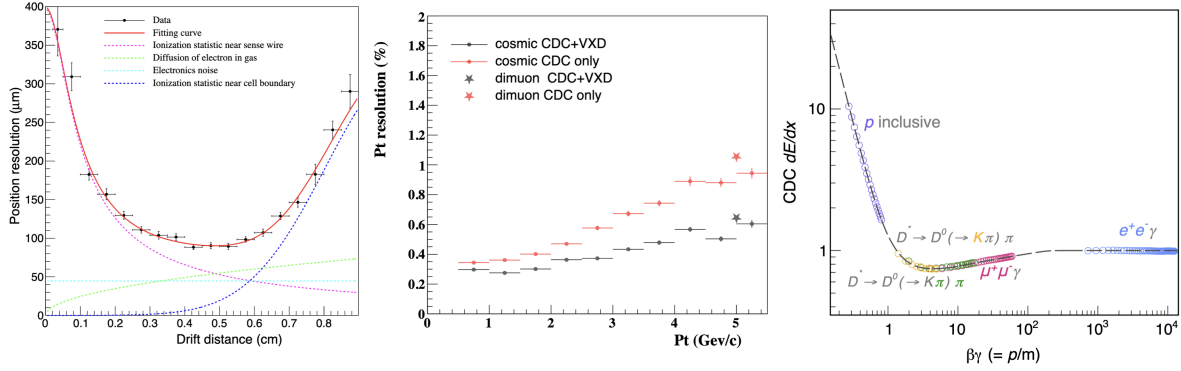


Figure 2.11: Diagram illustrating position resolution (left), transverse momentum resolution (middle), and CDC dE/dx (left) performance using collision and cosmic data [107, Fig. 20-22].

2.6 Time-Of-Propagation Counter

The detector that aids with particle identification (ID) in the barrel is the Time of Propagation (TOP) detector [107, Ch. 6] [109]. It is an array of 16 fused silica quartz bars (with index of refraction $n = 1.47$ at wavelengths of 400 nm) with dimensions of $270 \times 45 \times 2 \text{ cm}^3$. At one end of the quartz bar is an array of photo-multiplier tubes (PMTs); the other end is a spherical mirror (focal length of 3.25 m) that reflects light towards the PMTs to increase the amount of collected light. Importantly, the Cherenkov angle of the visible light generated by a throughgoing fast charged particle is preserved in the multiple bounces of this totally internally reflected light along the quartz bar.

The main goal of TOP is to clearly distinguish between charged pions ($m_{\pi^\pm} \sim 139.6 \text{ MeV}$) and kaons ($m_{K^\pm} \sim 493.7 \text{ MeV}$). While the tracking detector provides momentum information, additional properties of the charged particles are needed to perform this distinction. Belle II does this by measuring the velocity $\beta = v/c$ of charged particles via Cherenkov light.

As a charged particle passes through the quartz bar with a speed greater than c/n , as illustrated in Fig. 2.12, it emits Cherenkov radiation at an angle $\theta_{ch} = \cos^{-1}(1/n\beta)$ and

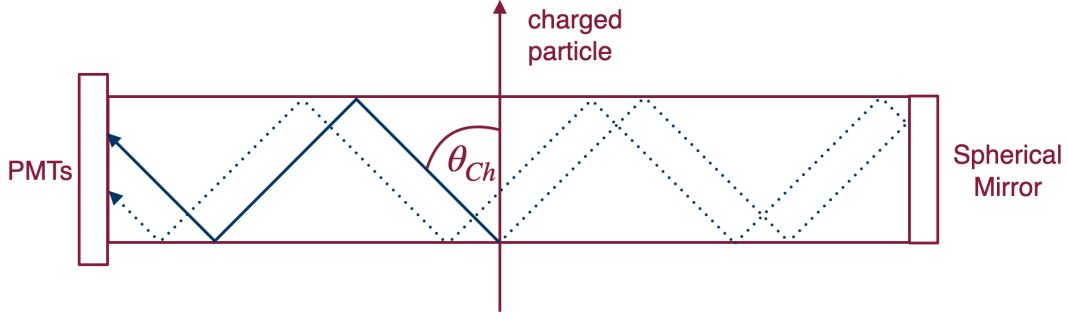


Figure 2.12: Illustration of how the TOP detector measures a charged particle's Cherenkov angle. The solid arrow is the light that generates the prompt signature, while the dotted arrow is the delayed signal.

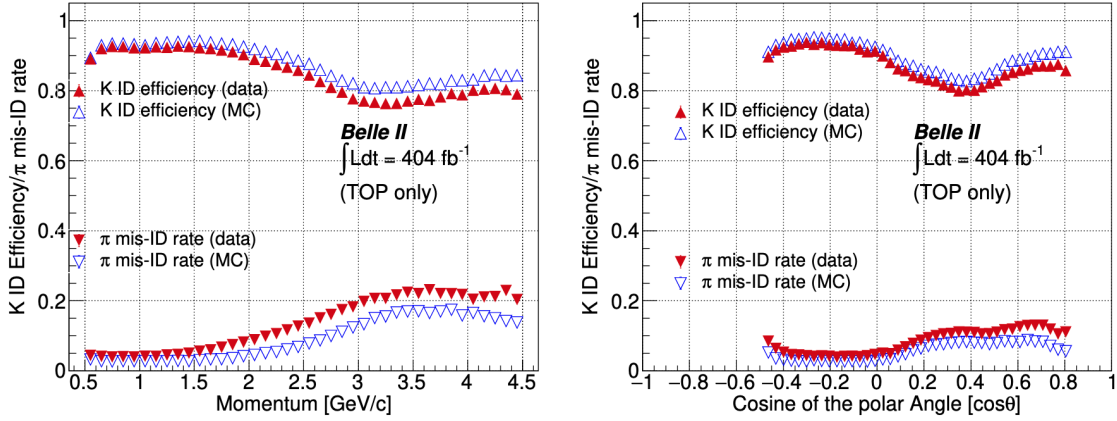


Figure 2.13: Using $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays in both MC simulation (blue) and data (red), we determine the K -ID efficiency and π mis-ID rate. Left: momentum dependence. Right: polar-angular dependence [107, Fig. 29].

internally reflects along the walls of the quartz bars. The light emitted towards the PMTs provides a prompt signal, while the light emitted away from the PMTs is reflected by the spherical mirror and provides a delayed signal. Based on the locations of the Cherenkov-light hits along the PMT array and their arrival times, one can deduce θ_{ch} ; from the corresponding velocity combined with the CDC-measured momentum, we can estimate the mass of the charged particle [110].

For π/K separation in the $p \sim 1 - 5$ GeV range (assuming $n = 1.47$), the ranges of expected Cherenkov angles are about 46.6° - 47.1° (40.6° - 46.8°) for pions (kaons) [111]. With photo-

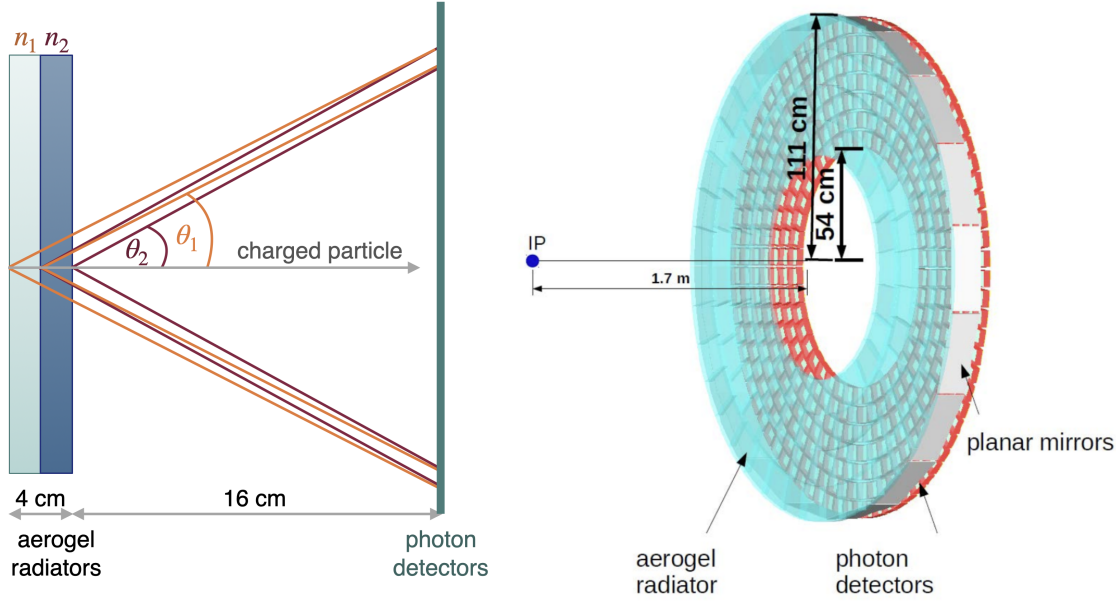


Figure 2.14: Illustration of the principle behind ARICH's measurement of θ_{ch} (left) and the detector schematic with its main components (right) [107, Fig. 30].

cathodes of $23 \times 23 \text{ mm}^2$ and a time resolution of 50 ps, we obtain the K efficiency/ π mis-ID rate illustrated in Fig. 2.13. Looking at both simulation and data of $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$ decays, we have high TOP standalone K -ID efficiency of $\gtrsim 90\%$ for $p \in [0.5, 2.5] \text{ GeV}/c$ or $\gtrsim 80\%$ for $p \in [2.5, 4.5] \text{ GeV}/c$.

2.7 Aerogel Ring-Imaging Cerenkov Detector

Similar to the TOP detector, the goal of the Aerogel Ring Imaging Cerenkov detector (ARICH) is to provide particle ID, with a particular focus on $\pi^\pm/K^\pm$ separation in the forward endcap region¹ by determining the β of the charged tracks via the Cherenkov angle.

The main principle behind ARICH is illustrated in Fig. 2.14. To begin, consider a scenario with a single aerogel radiator of width d and a photon detection plane. When a charge

¹Due to the boost of the e^+e^- collision, charged hadrons in the backward endcap have low momentum and so can be distinguished by dE/dx in the CDC.

particle goes through aerogel, it emits Cherenkov light at a Cherenkov angle $\cos \theta_{ch} = 1/n\beta$ along the entire distance d , forming a ring of width $d \tan \theta_{ch}$. Decreasing d would improve upon our spatial resolution of θ_{ch} but generate less Cherenkov light.

The way to overcome this limitation is to use two aerogel radiators of width $d/2$ with different refractive indices n_i . The total amount of radiating material is fixed at d , so the amount of Cherenkov light remains the same. However, because the radiators have two different n 's, the projected Cherenkov rings would have a difference of the inner radii Δr to be

$$\Delta r \approx (L + d/2) \tan \theta_1 - L \tan \theta_2. \quad (2.2)$$

We can use this constraint to determine what n_2 should be such that the two Cherenkov rings completely overlap (i.e., $\Delta r = 0$). To have dependence on the charged track mass and momentum, note $1/(\gamma\beta)^2 = (mc/p)^2 = n^2 \cos^2 \theta_{ch} - 1$. In addition, we can roughly approximate $\tan \theta_2 - \tan \theta_1 \approx \Delta \theta_{ch} = (n_2 - n_1)/n_1 \tan \theta_1$. Combining everything, we have

$$n_2 - n_1 \approx \frac{dn_1}{2L} \tan \theta_1^2 \approx \frac{dn_1}{2L} \sin \theta_1^2 = \frac{d}{2n_1 L} \left(n_1^2 - 1 - \frac{m^2 c^2}{p^2} \right). \quad (2.3)$$

In Belle II, we optimize for pions with $p \in (1, 4)$ GeV/ c , yielding $n_2 = 1.055$ [112, 113].

The resolution of ARICH's photo detector (i.e., the ARICH hybrid avalanche photodetector's pixel size) is $4.9 \times 4.9 \text{ mm}^2$, which is sufficient to see the few $mrad$ difference between a Cherenkov ring generated from pions and kaons. This is illustrated in Fig. 2.15.

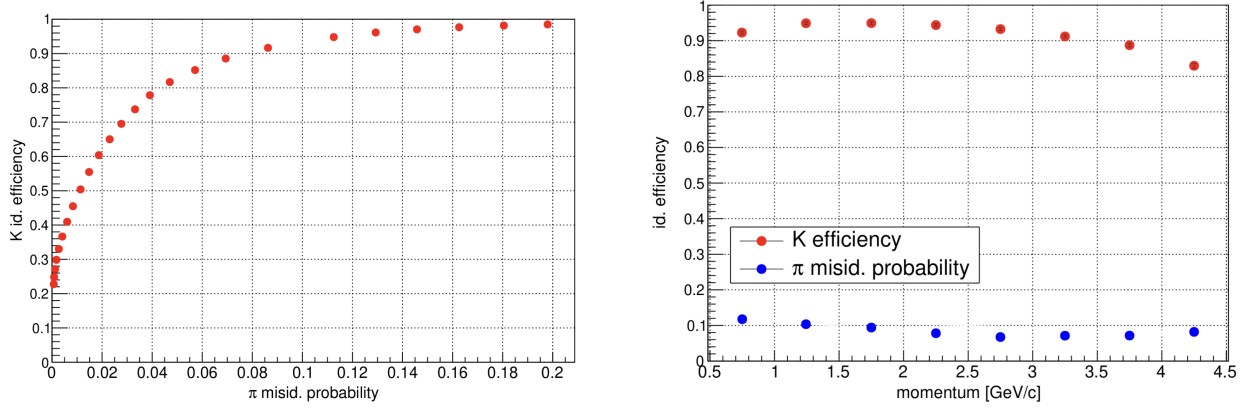


Figure 2.15: ARICH’s π/K separation performance from a $D^{*+}(D^0 \rightarrow K^-\pi^+)\pi^+$ control sample [107, Fig. 37]. The left plot shows K -ID efficiency vs. π fake rate (i.e., the ROC curve for K -ID). The right plot shows K -ID efficiency and π mis-ID as a function of track momentum.

2.8 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECL) [114] is tasked with the following critical tasks: detect photons with high efficiency across a reasonable energy range, and determine their energies and directions; identify electrons and other charged particles; assist with K_L^0 detection; measure luminosities (both during and after data collection).

The ECL’s active detector elements are CsI(Tl) crystals, which were originally used in Belle’s calorimeter [115, 116]. Depicted in Fig. 2.16, CsI(Tl) was chosen due to its high light yield (especially due to the doped thallium [117, 118]), short radiation length $X_0 = 1.86$ cm, and cost-effectiveness. To cover a polar angle region of $12.4^\circ < \theta < 155.1^\circ$, except for two $\sim 1^\circ$ wide gaps between the barrel and endcaps, an ensemble of 6624 CsI(Tl) crystals is used in the barrel and 2112 crystals in the endcaps. The nominal dimensions of the CsI(Tl) crystals are 6×6 cm² in cross-section and 30 cm (or $16X_0$) in length to minimize shower leakage.

In detailed studies, the CsI(Tl) crystals used from Belle [119] were found to have an energy resolution of 1.4% to 4% resolution for photon energies $E_\gamma \in [20 \text{ MeV}, 5 \text{ GeV}]$ and posi-

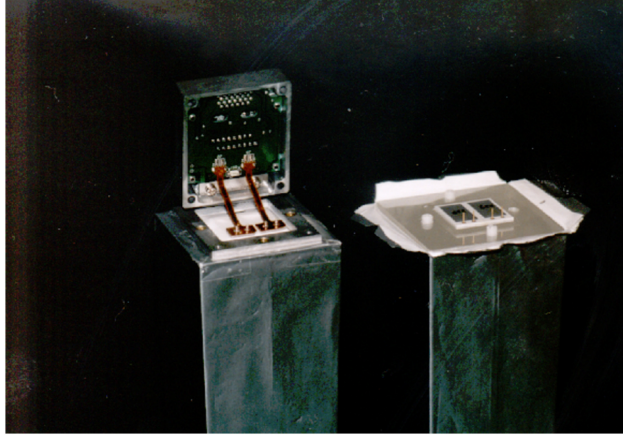


Figure 2.16: An image of the CsI(Tl) crystal with the photodiode and preamplifier used in both Belle and Belle II [114].

tion² resolutions of ~ 8 mm for 470 MeV photons; this is still sufficient for Belle II. The sources for the variation in energy resolution, which is modeled in the form of $\sigma_E^2/E^2 = (a/E^{0.25})^2 + (b/E)^2 + c$, are due to effects such as: electronic noise contribution, shower leakage, temperature effects, energy deposited in inactive material between crystals or in front of the calorimeter.

Scintillating crystals such as CsI(Tl) measure photon energies via the photoelectric effect, Compton scattering, and pair production [120, Ch. 13]. If an incoming photon's energy is identical to the energy gap between two electron orbitals, it can be absorbed and remitted via the photoelectric effect. This creates a characteristic detector response, which can be measured by a photo-detector. However, since the energies in which this phenomenon occurs are between 100 eV to 100 keV, this process is not of interest for Belle/Belle II. For a wider range of photon energies, crystals typically measure the energy of an incident photon when it scatters atomic electrons via Compton scattering; the scattered electrons then ionize the crystal's atoms, which release visible light upon de-excitation [121]. For photon energies larger than twice the electron mass ($E_\gamma \geq 1.022$ MeV) near a crystal's nucleus, an

²The position of the ECL cluster, when combined with the assumption of the incident photon's direction at the IP, gives the photon's direction.

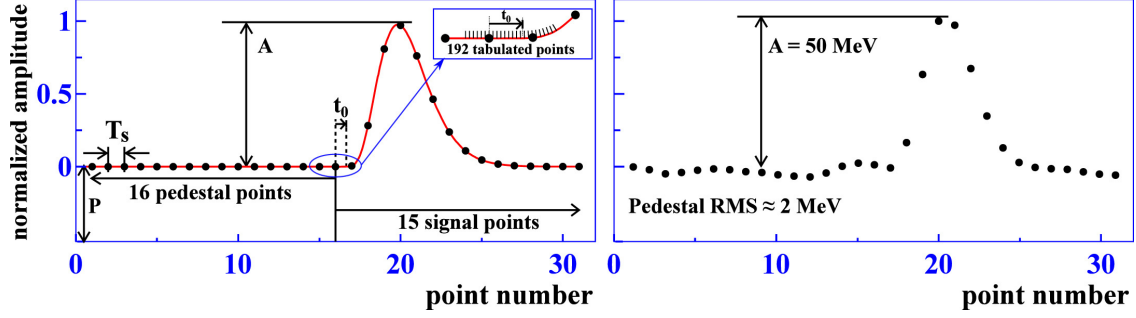


Figure 2.17: An illustration summarizing the fitting procedure done by Belle II ECL’s shaper module [127]. The left plot is for the ideal case and is used to highlight various features of the pulse-fitting procedure (T_s , P , A , t_0). The right plot is a typical plot from real data.

electron-positron pair is produced. The emitted electron and positron ionize the crystal’s atoms, producing more visible light upon de-excitation. When the emitted positron loses all of its kinetic energy through this repeated scattering, it annihilates with a neighboring electron, creating a pair of 511 keV photons, which can then liberate more atomic electrons via Compton scattering. The photodiodes attached to the crystal collect all the visible light generated by the atomic de-excitations; this is found to be proportional to the energy of the incident photon.

Along with the photon signatures described above, CsI(Tl) crystals have also been found to distinguish between different particles [122–125]. For one, the properties of electromagnetic and hadronic showers are quite different due to the different successive stages and energy distributions across a collection of crystals. The light output measured from a single crystal can also be used. Along with the inherit “slow” ($\tau_{slow} \sim 5.8\mu s$) and “fast” ($\tau_{fast} \sim .85\mu s$) components of the CsI(Tl) scintillation, hadronic interactions also provides a characteristic light output with $\tau_{hadron} \sim 0.63\mu s$ [124, 126].

2.8.1 Upgrades from Belle

A crucial component of the ECL upgraded from Belle is its electronics, given the expected elevated background levels, although the photo-detectors and preamplifiers remain unchanged [115]. The high beam backgrounds of Belle II would produce high-energy photon backgrounds (> 20 MeV) which could yield substantial combinatorial backgrounds during event reconstruction, and soft background photons (< 1 MeV) which would generate pile-up noise, causing fluctuations in the number of arriving photons and their timing [98, Ch. 9.4].

One major component that was upgraded was the modules that digitize the light-to-current signal from the photo-detectors (ShaperDSP, where DSP stands for digital shape processing) [114, 127]. To maintain the expected 30 kHz event rate, an 18-bit analog-to-digital converter with a 1.76 MHz clock is used to quickly sample pulses from the crystal. Sampling every $T_S = 1/1.76 \text{ MHz} \approx 568 \text{ ns}$, the ShaperDSP module takes 31 different sampling points (16 points for the baseline and 15 points for the pulse shape) and computes a pedestal/baseline P , amplitude A , and an initial start to the pulse t_0 . An illustration of the fitting procedure is given in Fig. 2.17, which nominally takes $\sim 5\mu\text{s}$.

The difference between the amplitude A and the pedestal P is used for hit selection in the ECL. If $A - P \lesssim 1 \text{ MeV}$ (to minimize the impact of pile-up noise), the hit information is discarded. Events with $1 \text{ MeV} \lesssim A - P \lesssim 2 \text{ MeV}$, it is assumed that t_0 is equal to the trigger time and fit only A and P . Finally, if $A - P > 50 \text{ MeV}$, the waveform itself is sent along with the fit parameters, enabling pulse-discrimination techniques for particle ID. In addition, $|t_0 - t_{trg}| < 5000 \text{ ns}/(E/\text{MeV})$ can be used to suppress the number of fake (background) photons ($E_\gamma > 20 \text{ MeV}$) by a factor of ~ 7 with a true photon retention rate of 93%.

In some initial tests [98, Ch. 9.4.3], a comparison between Belle electronics with Belle II prototypes demonstrated a decrease in pile-up noise, where the region closest to the beampipe

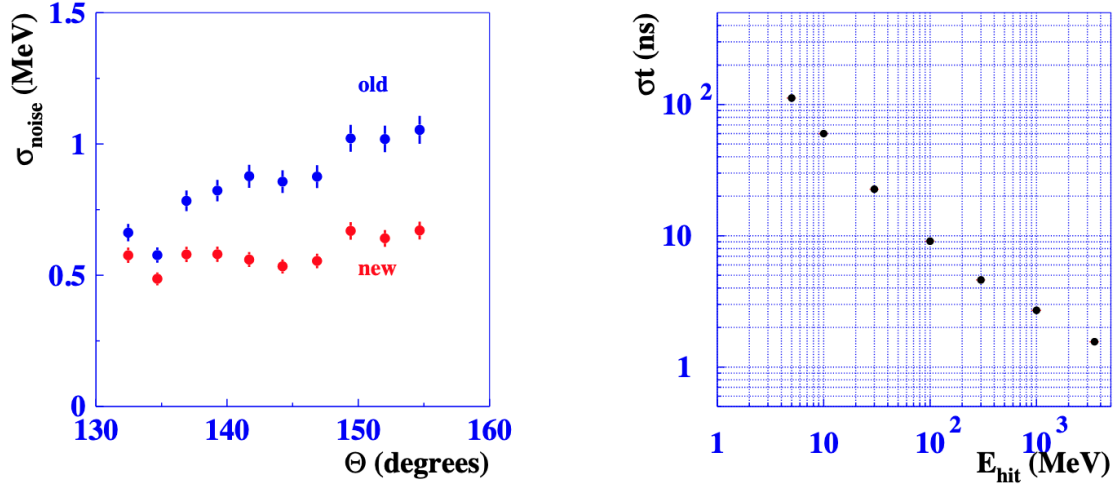


Figure 2.18: Plots showing a comparison in performance between Belle vs. Belle II electronics. Left: Pile-up noise dependence on polar angle for (old) Belle electronics and the (new) Belle II prototypes. Right: Time resolution as a function of single-crystal hit energy.

had a suppression factor of 1.5. The time resolution also ranged from 100 ns for 5 MeV cluster energies to 3 ns for 1 GeV cluster energies, compared to Belle’s ~ 20 ns resolution for energies above 100 MeV [119, Ch. 4.8.7]. The results of these tests are illustrated in Fig. 2.18.

2.9 Solenoid

An important component of the Belle II detector outside the ECL is a cylindrical superconducting solenoid, 3.4 m in diameter and 4.4 m in length. The 1.5 T magnetic field produced by this solenoid is critical for distinguishing positive and negative tracks, as it causes charged particles to spiral through the inner detectors. The solenoid’s specifications are outlined in Table 2.3.

Belle II’s iron structure serves as the magnetic flux return path for the solenoid, while also supporting all the detector components. The structure consists of eight blocks weighing 608 tons, with 200 mm-thick iron plates surrounding the outermost KLM layers, and movable

Table 2.3: Main parameters of the solenoid coil [128, Table 11.2]

Items	Parameters
Cryostat	
Radius: outer/inner	2.00 m/1.70 m
Central field	1.5 T
Total weight	23 t
Effective cold mass	~ 6 t
Length	4.41 m
Coil	
Effective radius	1.8 m
Length	3.92 m
Conductor dimensions	3×33 mm ²
Superconductor	NbTi/Cu
Stabilizer	99.99 % aluminum
Nominal current	4400 A
Inductance	3.6 H
Stored energy	35 MJ
Typical charging time	0.5 h
Liquid helium cryogenics	Forced flow two-phase
Cool down time	≤ 6 days
Quench recovery time	≤ 1 day

end-caps, each weighing 262 tons. Details of these iron yokes are outlined in Table 2.4.

2.10 K_L^0 and Muon Detector

The K_L^0 and muon detector (KLM), outside the solenoid, is the outermost component of the Belle II detector, covering a polar range from 45° to 125° from an octagonal barrel and 20° to 155° from the endcaps. It consists of thin (4-cm) large-area detector panels interleaved with the iron plates of the solenoid's magnetic flux return. There are 15 detector layers in the barrel, 14 in the forward endcap, and 12 in the backward endcap. Since these particles are long-lived and either neutral or free of strong interactions (respectively), they pass through the inner detectors with little to no trace. To see K_L^0 , the flux return's alternating layers of

Table 2.4: Main parameters of the iron structure [128, Table 11.1].

Items	Parameters
Belle Iron Yoke	
Height	9.57 m
Beam level	5.72 m
Total Weight	11740 kN
Barrel yoke	
Shapes	octagonal
Material	S10C iron
Height	7.7 m
Width	7.7 m
Length	4.4 m
Total Weight	6240 kN
Number of iron plates	15
Thickness of iron plate	47 mm
Thickness of gap	44 mm
End Yoke	
Material	S10C iron
Height	7.7 m
Width	7.7 m
Length in beam direction	1321 mm
Total Weight	5254 kN
Number of iron plates	15
Thickness of iron plate	47 mm
Thickness of gap	44 mm

4.7 cm thick iron plates and active detector elements outside the superconducting solenoid, providing an additional 3.9 hadronic interaction lengths, along with the initial 0.8 interaction lengths of the ECL. This implies that most of the produced K_L^0 s will interact hadronically inside the ECL or KLM, producing an observable hadronic shower of charged kaons and pions [98, Ch. 10]. Energetic muons ($\gtrsim 1$ GeV), on the other hand, will pass through the ECL and KLM's iron layers, gradually losing energy via atomic ionization (dE/dx); this energy loss within each detector panel provides a measurable signal.

Inherited from Belle [116, 129], most of the active detector elements in the barrel are glass-

electrode resistive plate chambers (RPC) [130–132], with a superlayer shown in Fig. 2.19. The basic layout of a single RPC is two parallel electrode planes separated by a gas gap (comprised of 60% HFC-135a, 30% argon, and 8% butane [98, Ch. 10.2]) such that an incoming charged particle passing through would ionize some of the gas molecules. With a high enough voltage difference between the two plates, the electric field is strong enough to prevent recombination of ionized gas molecules, so that electrons drift to the cathode plate. As the field strength increases, the initially ionized gas molecules are accelerated sufficiently to ionize neighboring gas molecules (i.e., secondary ions). In this regime (i.e., the proportional regime), the number of accelerated secondaries is still proportional to the number of initial (or primary) ions, so the charge readout at the end of this Townsend avalanche is still linearly related to the energy of the ionizing particle. If one increases the field strength further beyond, the linearity between the number of secondary and primary ionized particles is broken; the Geiger-Müller region is reached, where, due to the build-up of charge near the cathode plane, the avalanches will eventually stop [133]. For Belle and Belle II, KLM’s RPCs are somewhere in between the proportional and Geiger-Müller region where streamers, or an avalanche whose corresponding electric field becomes of a comparable magnitude to the external field (i.e. ~ 4.3 kV/mm), are used to detect incoming charged particles. In addition to the very high discharge from the streamer, the presence of butane is used to absorb stray UV rays and prevent tertiary avalanches from forming far away from the initial streamer.

As illustrated in Fig. 2.20, a single superlayer of the KLM (in the barrel) consists of two RPCs with orthogonal 5 cm strips of cathode planes, separated from each other by a dielectric foam and ground planes. The cathode plane system forms a transmission line with an impedance of approximately 50Ω that transmits pulses to a discriminator via twisted-pair cables. Pulses, generated from events like streamers in either RPC, is imaged on both readout

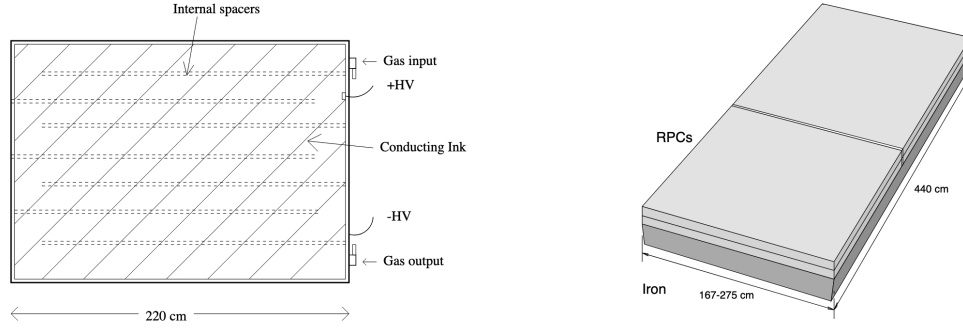


Figure 2.19: Left) A top view of a barrel RPC, showing the internal spaces and service ports for gas and HV. Right) An isometric view of an octant superlayer in the barrel [98, Fig. 10.4].

planes corresponding to a hit in the KLM when the discriminator gets a signal from both planes due to the RPC material (73% silicon dioxide, 14% sodium oxide, 9% calcium oxide, and 4% un-categorized compounds) being electrically transparent to fast signals (a FWHM of < 50 ns). Since the pulse resulting from a streamer is typically around 100 mV, the discriminator is set to near 40 mV. With this system, the KLM has a spatial resolution of $\sigma_{\text{STD}} \sim 1.1$ cm when either one or two adjacent cathode strips fire or $\sigma_{\text{STD}} \sim 1.7$ cm for the rare case of three strips firing (due to UV-photon-induced after-pulsing of the neighboring electrode area) [98, Ch. 10.2] and a timing resolution of a few ns [116, Ch. 9.3].

Although the timing resolution of the RPCs is good, the intrinsic dead time associated with recovering the RPC's electric field near a streamer is a concern. Events coming from low-energy neutrons or bremsstrahlung photons originating upstream along the beamline are major contributors to the ambient hit rate, which also scales with luminosity. Because the sources of these backgrounds are from the beamline, the hit rates in the most inner layers are as high as a magnitude greater than the most outer layers of the barrel (e.g., ambient rates of 1.07 vs. 0.09 Hz/cm² in the barrel). Belle II operates at much higher luminosities, so if one extrapolates these ambient hit rates to the target luminosity and examines the muon detection efficiency, the two inner layers in the barrel drop from 91% and 95% to 70%

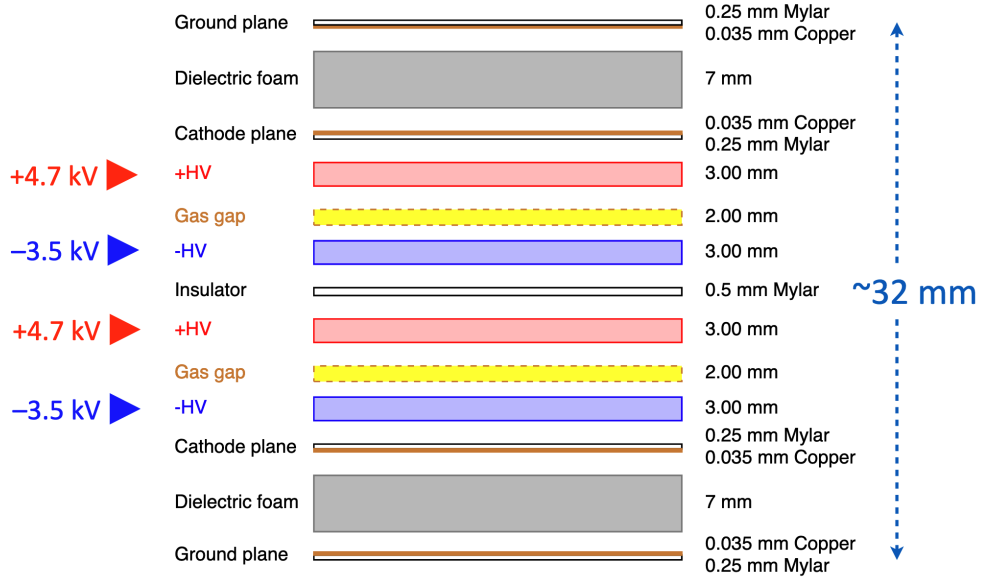


Figure 2.20: An illustration of an RPC superlayer of the KLM [134]. The two gas gaps and orthogonal readout planes, streamers left by a K_L^0 or muons will be read out.

and 81%. If one repeats this exercise assuming the endcaps were comprised of RPCs, the efficiency would drop to zero [98, Table 10.2]. Thus, to maintain a high muon efficiency, an alternative was required [98, 10.4].

The other active detector element for the endcaps and the first two layers of the barrel is scintillators, which are comprised of materials (polystyrene doped with PTP (1.5%) and POPOP (0.01%)) that detect charged particles through continuous interactions with the medium's electrons via Coulomb interactions, resulting in atomic excitations or ionization. In the case of neutral particles like K_L^0 , scintillation occurs from the hadronic-shower charged byproducts [135]. The scintillation light is trapped in embedded wavelength-shifting (WLS) fibers where the signal is read out by silicon photo-multipliers (SiPM) [98, Ch. 10.4] [136].

In the 14 (12) superlayers in the forward (backward) endcaps, there are 16,800 scintillator strips with each strip accommodating the WLS readout via grooves at their center, and each superlayer being formed by two orthogonal strips. As illustrated in Fig. 2.21, the dimension

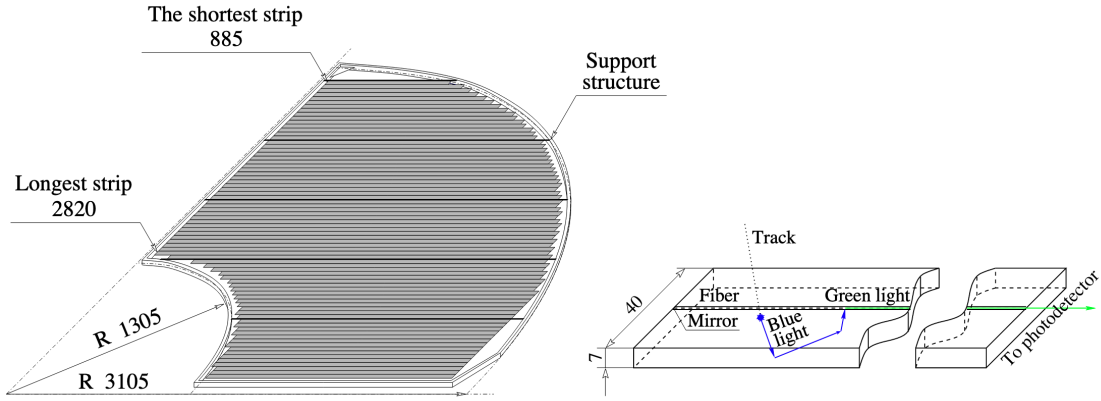


Figure 2.21: Left) The endcap consists of one layer of scintillator strips. Right) An illustration of scintillator light detection in the strip with the WLS fibers embedded. Source: [98, Fig. 10.20]

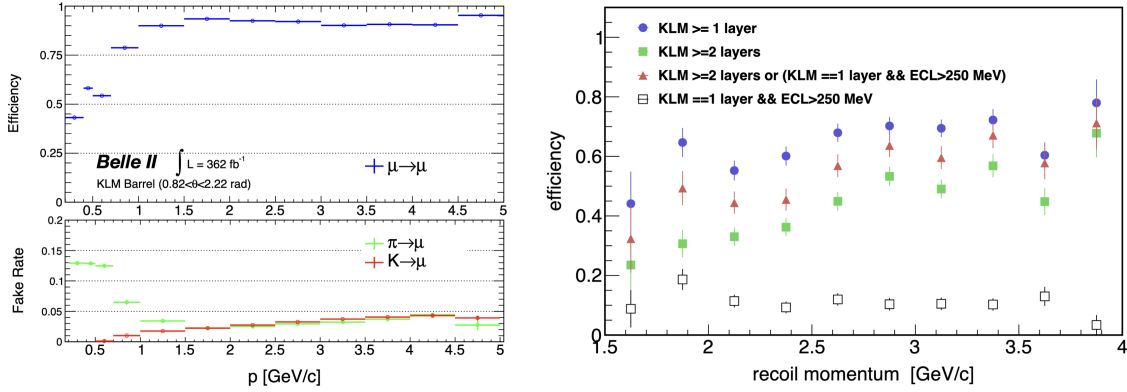


Figure 2.22: Illustrating KLM performance for μ ID (left) and K_L^0 -ID from data [107, Fig. 51, 52].

of a strip is $(7-10)\text{mm} \times 40\text{ mm} \times \sim 2.8\text{ m}$, providing a spatial resolution comparable to that seen in the RPCs. The time resolution (as obtained from cosmics) was $\sigma_t \approx 0.7\text{ ns}$ [98, Ch. 10.4].

The readout electronics used for the scintillators are described in [137].

The KLM performance is illustrated in Fig. 2.22. For μ -ID, the number of KLM layers activated is used to determine the viability of a track being a muon. For standalone K_L^0 -ID, a localized collection of KLM hits is used as K_L^0 candidates, assuming it is at least 10 cm

Physics Process	Cross section (nb)	Rate (Hz)
$e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$	1.2	720
$e^+e^- \rightarrow q\bar{q}, q = u, d, s, c$	2.8	1700
$e^+e^- \rightarrow \mu^+\mu^-$	0.8	480
$e^+e^- \rightarrow \tau^+\tau^-$	0.8	480
(radiative) Bhabha	125	75,000
$e^+e^- \rightarrow \gamma\gamma$ ($\theta_{lab} \geq 17^\circ$)	2.4	1400
Two-photon particle production ($\theta_{lab} \geq 17^\circ, p_t \geq 0.1 \text{ GeV}/c$)	~ 80	$\sim 48,000$

Table 2.5: Cross sections and expected rates of the main interaction processes of $\Upsilon(4S)$ -resonance beam cross section at the target luminosity $\mathcal{L} = 6 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ [107, Table 8].

away from any tracks.

2.11 Trigger System

While I’ve outlined how Belle II detects various final state particles, not all events seen from e^+e^- collisions are interesting. As shown in Table 2.5, rates from $e^+e^- \rightarrow B\bar{B}$ make up less than 1% of the events produced. A robust triggering system is needed to ensure that events of interest are retained while reducing the number of Bhabha and two-photon particle production events. Additional requirements for this trigger system are a fixed latency of $4.4 \mu\text{s}$ (due to the SVD buffer size), a maximum rate of 30 kHz (from our data acquisition system), and no dead time.

This is accomplished via a multi-layered trigger system that consists of a level-1 (L1) hardware trigger [107, Ch. 10] and a subsequent software trigger (HLT) that will be described in the next section. The goal of the L1 trigger is to use fast-digitized information from the CDC (for tracking), ECL (for photons and other clusters), TOP (for timing), and KLM (for muons and K_L^0 s) to determine when an event occurs and when to record it. The global reconstruction logic (GRL) combines the information from the aforementioned subsystems,

while the global decision logic (GDL) makes the final decision on the desirability of an event. The CDC trigger looks for hit patterns in each “super-layer” that could correspond to a track segment candidate [107, Ch. 10.2]. Track segments are formed if enough hits are strung together and within a 512 ns window of one another. The CDC trigger also examines the candidate, compares it against a lookup table of track segments, and uses the segments to form tracks. Since there are two wire configurations (axial and stereo), the trigger first makes 2D tracks for a fixed configuration before combining them into 3D tracks via a Neural Network 3D Tracker [138].

The types of events the ECL is tasked with triggering are charged tracks, neutrals that decay in the ECL, Bhabhas, and $\gamma\gamma$ [107, 10.2]. While tracks and neutrals are of primary interest, Bhabhas and $\gamma\gamma$ monitoring rates are still monitored for online luminosity measurements.

Aggregated information from the ECL’s ShaperDSP searches for energy deposits in 4×4 groupings that are $> 0.1 \text{ GeV}$. If the total energy deposit across the ECL is $> 1 \text{ GeV}$ or there is a sufficient number of clusters, then information about the event topology, number of isolated clusters, and total event ECL energy is considered. If two high-energy clusters are back-to-back (in the CM frame), they’re flagged as Bhabha events and vetoed about 80% of the time. Otherwise, events with high cluster energies (e.g., single photons with less than 2 GeV in the endcap; $\sum E > 1 \text{ GeV}$) or interesting topologies (e.g., back-to-back clusters in the lab frame) will be flagged.

The TOP’s contribution, due to its timing resolution of 2 ns (the PMT digitization time), is to help determine the collision time. For many $B\bar{B}/q\bar{q}/\tau\tau$ events, a large number of hadronic tracks are produced, enabling TOP to provide a prompt trigger for these events.

Last but certainly not least, we want a trigger for muon events. The simplest muon trigger would just be counting the number of KLM layers fired in a given event; if 8+ layers are

fired, there was probably a muon. However, cosmic rays would also trigger the KLM. To address this, a straight-line fitter is implemented to determine whether muons come from the interaction point (ex., $e^+e^- \rightarrow \mu^+\mu^-$) or not.

From here, the GRL interprets and processes all information to match tracks in the CDC and relate them to KLM and ECL information. With an estimate of the number of charged tracks and neutral particle decays (ex., γ , π_0), Belle II's GDL then determines whether events are to be kept or not. One condition the GDL considers is whether an event needs to be 'prescaled,' i.e., whether we want to reduce the number of undesirable events by some factor. For instance, Bhahba scatterings, $\gamma\gamma$, and other two-photon productions are prescaled by a factor of 10-100. Another condition the GDL considers is whether SuperKEKB injects more $e^\pm$ into the storage ring near the trigger inputs. The GDL receives signals from SuperKEKB's injection kickers indicating that background events will be elevated and that the GDL should completely veto events during this interval.

A summary of the major trigger conditions is outlined in Table 2.6. From our 2022 data-taking run, the trigger efficiency for $B\bar{B}$ ($\tau^+\tau^-$) events is more than 99% (95%), while maintaining a total trigger rate of 11.5 kHz.

2.12 Data Acquisition System

The Data Acquisition (DAQ) faces the enormous challenge of reading in data from the other sub-systems and recording everything within a $30\mu s$ time window ($\sim 1/30$ kHz). The challenge arises from the PXD data rates, which far exceed those of the other subsystems (7.2 gigabytes/s vs. $O(10 - 100)$ megabytes/s). To manage this, the system illustrated in Fig. 2.23 was developed to accomplish this [98, Ch. 13.2].

Trigger Condition	Raw Trigger Rate (kHz)	Physics Target
CDC (3 tracks)	2.18	$B\bar{B}$
CDC back-to-back 2 tracks	1.77	$B\bar{B}, \tau^+\tau^-$
CDC single track, $p > 0.7 \text{ GeV}/c$	2.93	$\tau^+\tau^-$
ECL four clusters	0.47	$B\bar{B}$
ECL $\sum E > 1 \text{ GeV}$	2.02	$B\bar{B}, \tau^+\tau^-$, photon
ECL cluster energy $> 2 \text{ GeV}$ at endcap	1.00	Single Photon
ECL back-to-back 2 clusters	0.60	ISR Photon, ALP
KLM 2 clusters	1.59	$\mu^+\mu^-$, ALP
CDC-KLM matched single track	0.51	$\mu^+\mu^-$
ECL-KLM matched single cluster	0.61	μ at Barrel
Sampling of $e^+e^- \rightarrow e^+e^-, \gamma\gamma, e^+e^-l^+l^-$	1.35	Calibration, Systematic
Beam background sampling	0.01	Calibration, Systematic
Total	11.5	

Table 2.6: Main trigger conditions and rates in the 2022 data-taking runs with $\mathcal{L} = 4.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ [107, Table 10].

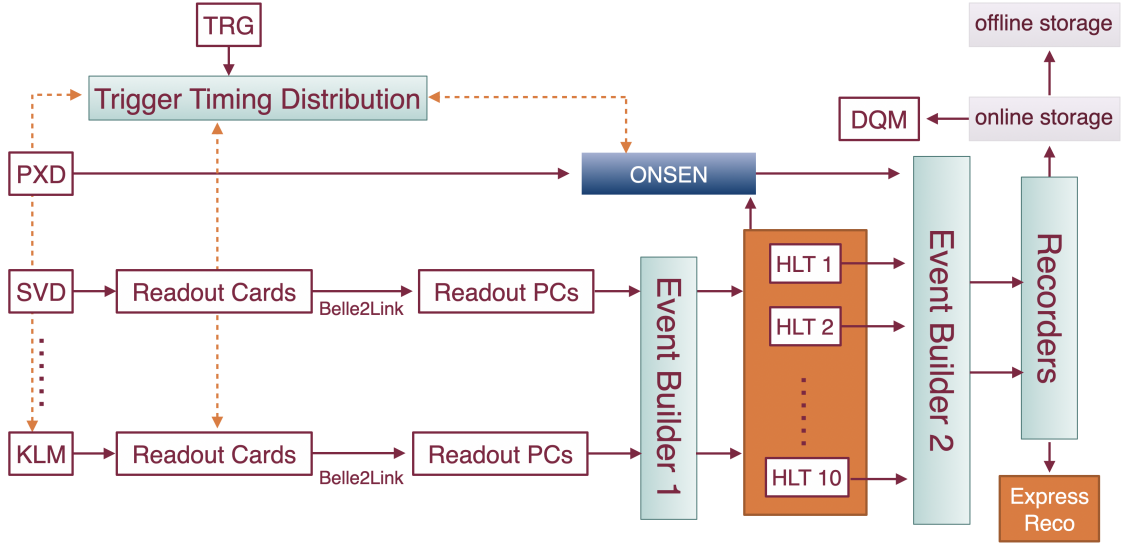


Figure 2.23: A simplified illustration of Belle II's DAQ system.

The major components of the DAQ system are the Trigger-Timing Distribution system (TTD), the readout cards, the event builders, and the High Level Triggers (HLT). These components work in tandem with the other subsystems to process raw detector information, quickly reconstruct events for online data quality monitoring (DQM), and store events.

The next main component is TTD, which manages timing across the subsystems and the accelerator [139]. A system-wide clock of 127 MHz (based on SuperKEKB radio-frequency 509 MHz clock³) is generated, in which the L1 trigger is synchronized to. TTD then distributes timing from this system-wide clock, along with trigger distribution to/from the various subsystems.

The readout card's task is to provide a unified interface across the subsystems (sans PXD) that interface with Belle II's timing distribution system, to digitize and reduce the detector information, and to send it to the readout PC via high-speed optical links (Belle2Link). The readout PC then combines event fragments into a single packet, performs additional data checks, and sends this information to the first event builder [107, Ch. 11.3]. Initially used from 2018 to 2020, the Common Pipeline Platforms for Electronics Readout (COPPER) boards [140] were used for the Belle II readout cards with a data-taking efficiency of 84.2%. However, the COPPER system impedes high-speed data transfers due to the 1 Gb/s Ethernet output and limits the raw data available from the subsystems (if subsystems wish to move online data processing away from their readout electronics). Additionally, there is a serious concern about the availability of parts within the next decade. Thus, Belle II upgraded the readout cards to a PCI express (PCIe40) system [141, 142], initially developed for LHCb [143]. With the PCIe40's specifications of 4.8 GB/s per board, data transfer rates are no longer a bottleneck. In addition, the PCIe40 system reduced the number of readout cards from 203 COPPER boards to 21 PCIe40 boards, further improving the DAQ's

³The 508 MHz comes from a 100kHz beam revolution cycle (speed of light / SuperKEKB's 3 km circumference) and 5120 beam bunches stored at any given time.

maintainability.

The next step of the DAQ pipeline is the event builder. The event builders create tags/events and combine all the event fragments from the subsystems into a single event packet [144]. Due to the size of PXD data, a second event builder, after the HLT units, combines relevant PXD information with the rest of the event.

After an event is packaged, Event Builder 1 passes the event containing all the raw detector information into a ring buffer system, where an available HLT unit processes said event to provide additional software triggers and produce histograms for DQM by using the Belle II’s offline reconstruction software (BASF2, see Ch. 2.13). Each HLT unit hosts about 500 cores, enabling us to process many events simultaneously [145].

With the various tracking algorithms in BASF2 (see Ch. 2.13.1), tracks formed by the SVD/CDC can be extrapolated towards the PXD to find “regions of interest.” PXD information, saved by the ONSSEN system, is then used in these regions, which are then digitized and sent to Event Builder 2 [146]. For PXD and tracking DQM, a small subset of events post-Event Builder 2 is sent to an express reconstruction (ExpressReco) unit, which operates similarly to an HLT unit.

Finally, events are stored at KEK’s computing center and stored offline. This is also the step where either HLT or ExpressReco histograms are broadcast for monitoring.

Data Quality Monitoring

At the HLT and ExpressReco stages, a large number of physics observables are computed for many of the various software triggers. We leverage the availability of these quantities to assess the real-time performance of the Belle II detectors. Whether it is using tracking algorithms to assess detector efficiencies or vertexing algorithms to reconstruct K_s^0 to assess PID detector

performance, the relevant information is stored as histograms during the data-taking stage and viewed by detector experts and operators to continuously assess Belle II's detector performance. For additional processing and fitting of histograms produced by the **HLT**s and ExpressRecos, a dedicated DQM server is used to perform various analyses. The range of tasks performed at this stage includes arithmetic operations on histograms to define more meaningful quantities (e.g., efficiencies, hit rates), fitting histograms to extract variances and means, aesthetic modifications, and distributing quantities across the Belle II network. Histograms and quantities collected during an experimental run are displayed to experts and operators using JSROOT [147]. To collect specific detector quantities and arrange them across multiple experimental runs, an in-house tool called MiraBelle was developed to store and display said quantities. Relevant quantities for subsystems and the accelerator groups are broadcast on Belle II's **EPICS** [148] network. We want to ensure these services are run with minimal downtime, so **SYSTEMD** [149] in **LINUX** is used to maintain the DQM analysis services, histogram transfers, and **EPICS** broadcasts. Additional details are provided in Appendix B.

2.13 Software

The Belle II Software Analysis Framework (**BASF2**) [150, 151] is tasked with fulfilling all of Belle II's software needs, which includes reconstructing events from raw detector information, performing a plethora of tasks (e.g., simulating the Belle II detector via **GEANT4** [152] and using KLM hit patterns to better identify low-momentum muons via deep-learning [153]). The framework is divided into (mostly) **C++**-based modules for detector and other operational needs. This enables the timely reconstruction of software events on the **HLT**s.

For the user's ease (for both detector operations and physics analyses), the modules are

invoked via `PYTHON` scripts. For a comprehensive list of modules, please see [Belle II's software documentation page](#).

For the remainder of this subsection, I provide an overview of the aforementioned tasks that are relevant to Ch. 1.5 or 4. In particular, I summarize how the following works: tracking [154, 155], clustering, vertexing [156], and particle identification [157].

2.13.1 Tracking

A summary of the Belle II tracking algorithm is illustrated in Fig. 2.24.

Belle II's track finding algorithm [155] begins with the largest tracking detector – the CDC. Hits in one CDC drift cell correspond to a given wire fired, and provide the drift length from said wire. For an ensemble of hits that correspond to a track, Belle II uses both a Legendre-based algorithm and a Cellular automaton to find whether a collection of CDC hits belongs to some number of tracks. The Legendre-based algorithm [158] transforms the sense field wire locations and the drift length into a sinogram via Legendre transforms, where intersections in this transformed basis typically correspond to hits from a signal track. The Cellular automaton method [159], however, starts from a small number of hits and builds a track by adding neighboring hits.

Next, we examine SVD hits to either attach them to existing CDC track candidates or build standalone SVD hits. In the standalone case, the same cellular automaton method is used, starting from the SVD outermost layer and proceeding inward. In the combined case, a Combinatorial Kalman Filter (CKF) [160] is used to add hits that are in the vicinity of the extrapolated CDC track, layer-by-layer.

The CDC and SVD track candidates are combined into a single list with the CKF, passed through the PXD, and added to neighboring PXD hits. With all information from the

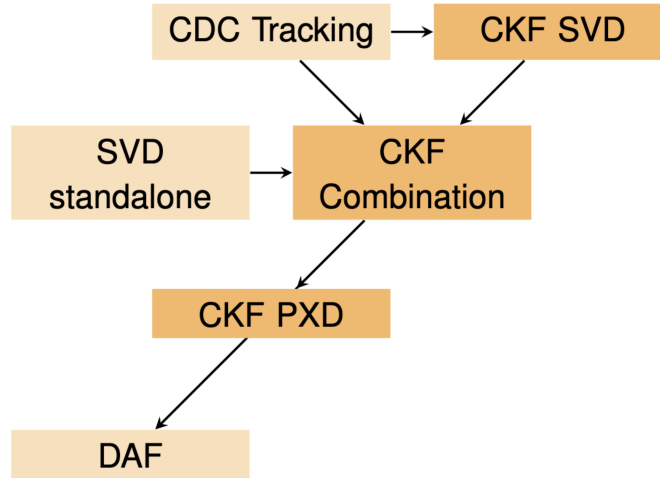


Figure 2.24: Illustration of the steps in Belle II's tracking algorithm.

subsystems included, a final algorithm (Determinist Annealing Filter [161]) is used to clean the track and finalize the expected track trajectory (via helices).

Using the best-fit track helix parameters, `GEANT4` [162] is used to extrapolate the track trajectory to the outer detectors. Two applications of this extrapolation are particle identification using energetic signatures in Belle II's calorimeter and efficiency measurements of the KLM.

For the first application, given an energetic shower or cluster (see Sec. 2.13.2) in the ECL, we need to determine what type of particle produced said cluster. Since the characteristic time of ECL pulses is used to distinguish hadronic and electromagnetic showers, let's focus on the electromagnetic case first; the easiest way to distinguish between clusters originating from photons, electrons, and muons is to determine whether there is a nearby track (via the point of closest approach) and whether that track has KLM hits along its trajectory. It is a similar case for hadrons, where the information from the ARICH and TOP can be used in conjunction with the presence of a track to identify the hadron.

For the second application, we wish to assess the performance of the KLM. As highly ener-

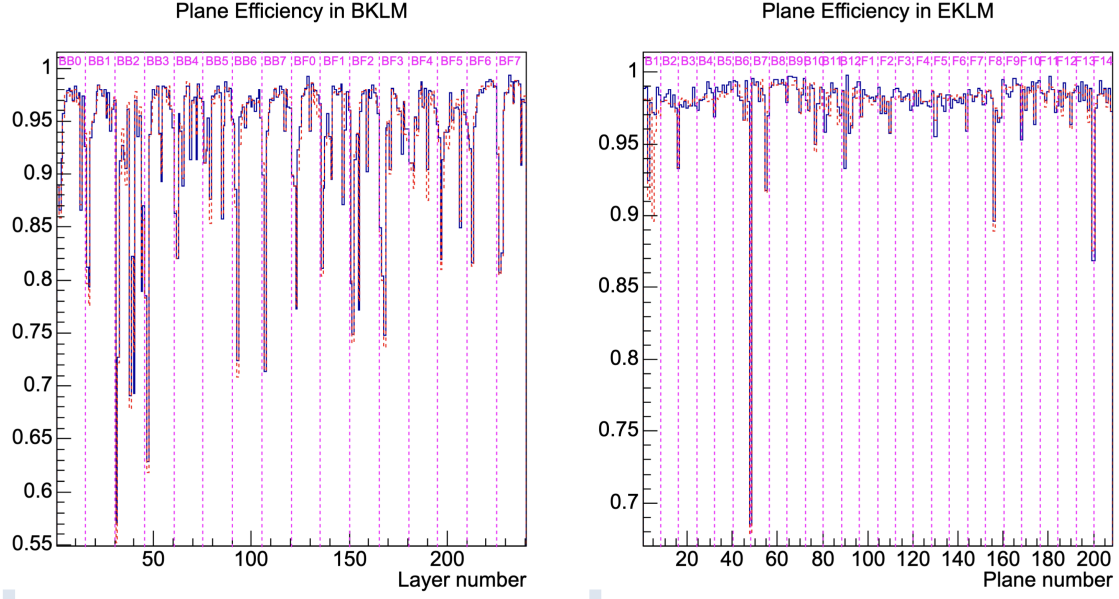


Figure 2.25: KLM efficiency monitoring from live data-taking on 13 March 2026.

getic (> 1 GeV) muons are likely to penetrate all the layers of the KLM, we use this fact with the information from the tracking detectors to determine how often an individual strip or layer of the KLM is activated. The extrapolated trajectory of a given muon-like track is used to ‘swim’ through the KLM and search for nearby KLM hits. If there is a KLM hit (typically within a few strips from the track extrapolated hit), we say that we’ve matched a hit to said track and determine that the KLM layer was effective in said event. Over tens of thousands of collisions, we can assess the efficiency of a given layer by taking the ratio of matched hits to extrapolated hits. An example is illustrated in Fig. 2.25, where the layer efficiencies during an actual live run are plotted for each layer of the barrel and endcap KLM. Along with the benefit of an observable that is resistant to the expected increase in beam backgrounds, we can also use this efficiency information to ensure that Belle II’s Monte Carlo simulations use detector conditions representative of the KLM (and other detectors).

2.13.2 Clustering

Complementary to tracking is clustering. For clarity, I am specifically referring to both ECL deposits coming from some number of particles (e.g., photons, hadrons, leptons) and a collection of hits found in the KLM coming from neutral particles (photons and K_L^0) or muons.

The basic concept is the following: for a collection of neighboring hits with corresponding properties (e.g., energy depositions, timing, hit positions), how can one tell whether they came from the same particle or not?

While there are more complicated situations (e.g., clustering in the ECL with varying energy depositions), I will outline the simplest case for clarity: when one has only coordinate and timing information (i.e., forming KLM clusters). For a given event, there is a list of hits with x, y, z -coordinates, with hits removed if they are out of time of the e^+e^- collision. Using the hit closest to the average hit position (i.e., i such that $d(\vec{r}_i, \vec{r}_{cm})$ is minimized) as our initial cluster seed, one can ask whether any of the remaining hits (r_j) have a distance (or angle relative to the IP) greater than some minimum threshold (i.e., $\vec{r}_i \cdot \vec{r}_j / r_i r_j < \theta_{\min}$). For all hits that pass this criterion, one can label them as a single cluster and compute their properties, including their position and energy.

Some (KLM-)cluster properties highlighted include its shape, calculated via Principal Component Analysis [163]. We can calculate a covariance matrix using the distances of hits relative to the cluster center (i.e., $M_{ij} = \sum_n ((\vec{r}_n - \vec{r}_{cm}) \cdot \hat{i})((\vec{r}_n - \vec{r}_{cm}) \cdot \hat{j}) / n$ with $i, j \in \{x, y, z\}$), diagonalize it, and determine the variance along the primary/secondary/tertiary axes (i.e., $\sigma_1, \sigma_2, \sigma_3$). Using Monte-Carlo simulations of a K_L^0 -rich sample (utilizing a particle gun), $e^+e^- \rightarrow s\bar{s}$ continuum sample, $e^+e^- \rightarrow B^0\bar{B}^0$ sample, and $e^+e^- \rightarrow \mu^+\mu^-$ sample, we can show how these features can help discriminate between K_L^0 s and other clusters found in the

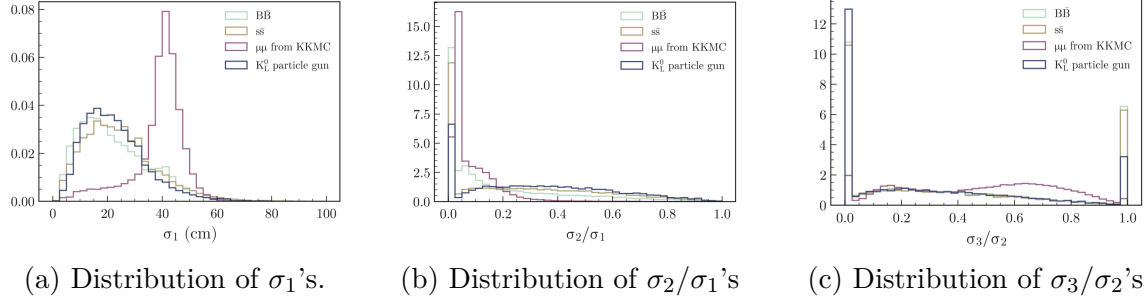


Figure 2.26: Showcasing cluster shape analyses using PCA on simulated samples of K_L^0 using a particle gun, $e^+e^- \rightarrow s\bar{s}$, $e^+e^- \rightarrow B^0\bar{B}^0$, and $e^+e^- \rightarrow \mu^+\mu^-$. Rather than showing all the relevant variances, taking their ratio highlights this analysis tool's discriminative capabilities better.

KLM. This is showcased in Fig. 2.26.

2.13.3 Vertexing

Given a collection of neutral clusters and tracks from the previous algorithms, a dedicated TreeFit algorithm [156] has been developed for vertexing in Belle II. For simplicity (and relevance), consider the scenario in which two tracks originate from a particle located at a distance d from the interaction point. Given that they are fit using helices, we can extrapolate the tracks and determine their point of closest approach. Given said geometric information (from the helices), kinematic information (i.e., 4-momenta of the two tracks), timing, and additional optional constraints (e.g., is there an expected vertex mass (constraining $p^\mu p_\mu = m^2$); does the vertex come from the IP). For the corresponding parent or vertex, the most naive parameters to give it are the point at which its decay occurs, at the two tracks' point-of-closest approach, and its 4-momenta to be $p_{V_0}^\mu = p_{t_0}^\mu + p_{t_1}^\mu$. To incorporate constraints, a Kalman Filter is applied to re-optimize the vertex parameters by minimizing χ^2 .

2.13.4 Particle Identification

After events are reconstructed, we need a way to use detector information to infer the nature of the reconstructed particles, given the tracks and clusters. Since SVD, CDC, TOP, ARICH, ECL, and KLM all provide particle identification (PID) information, the data from each subsystem must be considered. To do this, likelihood estimations are provided in the following fashion [157].

For each particle species, the subsystem will provide a likelihood prediction; these likelihoods are then combined to form an overall likelihood of a given particle being of species X . In other words, for a given final state particle (FSP) X , the likelihood of a particle being X is defined as $L_X = \prod_i^{det.} L_{i,X}$.

To combine these likelihoods into a PID-like quantity, we define $PID_X = L_X / \sum_Y^{FSP} L_Y$, where the list of (detectable) FSPs we expect are $\pi^\pm$, $K^\pm$, proton, deuteron, $e^\pm$, and $\mu^\pm$.

2.13.5 Simulation and Calibration

As it is standard protocol to conduct blind data analyses in Belle II, it is critical that our Monte Carlo-based simulations are representative of the detector’s real conditions⁴.

To assess and store the detector conditions for a given run, subsets of the data where the underlying physics is well-understood (e.g., $e^+e^- \rightarrow \Upsilon(4S) \rightarrow \mu^+\mu^-$, $e^+e^- \rightarrow \phi\gamma$) are used to assess the performance of each Belle II subsystem for each final-state particle of interest in the momentum and angular ranges of the physics process (“signal”) under consideration. This ranges from the alignment and efficiency of the various components to the combined

⁴There are exceptions where such high statistics are required for some analyses where run-dependent detector conditions are unfeasible. In such cases, run-independent simulations are generated, assuming the detector is 100% efficient; it is then the analyst’s responsibility to carefully validate the agreement between data and MC.

performance of the tracking algorithm. The information is then saved to a ‘payload’, which is used during data reprocessing or simulation generation to incorporate any detector inefficiencies or settings. For example, in some recent data-taking, the voltages of the CDC have been reduced due to high background conditions, resulting in a drop in efficiency; the CDC payloads need to be updated to ensure that real bona fide tracks are not mis-reconstructed as fragmented tracks or ignored altogether.

For simulations, an ensemble of physics generators is used to generate the underlying physics processes (see Sec. 3 for the specific generators used for this work). The final-state particles from said processes are then propagated, where their interactions with the Belle II detector are simulated with **GEANT4** [162]. The simulated detector responses are then ‘digitized’ in a format that emulates the expected precision of the raw data, and we apply the reconstruction techniques in the subsections above in an analogous methodology to the data.

Chapter 3

Analysis

In this chapter, I describe the details of our search for iDM. This includes explanations of crucial tools used throughout the analysis (Sec. 3.1), programs to simulate them going through the Belle II detector (Sec. 3.2); details of how iDM events are selected over SM backgrounds 3.2.4; and outlines of various systematics with their resulting impact on our measurement (Sec. 3.3).

3.1 Tools and Strategy

In this discussion, we summarize the analysis strategy, including the tools used.

Boosted Decision Trees

To assist with event selection, we use boosted decisions formed from various physical observables (see Sec. 3.2.4).

Decision trees (DTs) [164] are supervised machine learning models that partition datasets using a sequential decision structure. We focus on classification, as it is relevant to this analysis.

Like other classification problems, we start with some dataset $X = \{\vec{x}_i \text{ for } i \in [1, N_d]\}$ with N_d data-points, m features (e.g. $\vec{x} \in \mathbb{R}^m$), and some loss function $G(X)$. True labels y_i are provided for each $\vec{x}_i$ during supervised model training.

Consider a DT $T(X)$ with D decisions, where the d 'th node makes a decision. The goal of each decision is to split the dataset X using a threshold t_d on the d 'th node that best separates the data according to y_i . The optimized value of t_d is determined via

$$t_d^{\text{optimized}} = \arg \min_{t_n=d} (f_{<t_d} G(X|\vec{x}_i^j < t_d) + f_{\geq t_d} G(X|\vec{x}_i^j \geq t_d)) \quad (3.1)$$

where x_i^j is the j 'th feature of data-point i , and the fraction $f_{<t_d}$ ($f_{\geq t_d}$) of events that pass $x_i^j < t_d$ ($x_i^j \geq t_d$). The results of the cut are fed to node $d + 1$, where another feature is chosen and the process is repeated unless some stopping criteria are met (e.g., tree length reaches a predefined maximum; number of events in a subset produced by a cut is too small). Nodes that have met certain stopping criteria are typically referred to as *leaves*.

Loss functions typically separate data by using a probability p_l of choosing an event with label l as input. Examples of said functions are: Gini impurity ($G = \sum^L p_l(1 - p_l)$), Shannon entropy ($S = -\sum^L p_l \log(p_l)$), or information gain ($I = S_{d-1} - \sum^{\text{cuts}(t_d)} S_d$). Typically, Gini impurity or entropy is used to determine the threshold t_d while information gain is used to determine which feature is selected at node d .

To evaluate the classification of a DT, a few methods can be used. The simplest approach is to define some objective function and determine the best configuration that minimizes this. More robust methods include confusion matrices and receiver operating characteristic (ROC) curves, which quantify true/false positives/negatives.

Ensembles

While a single DT can be effective for classifying linearly separable datasets, the main challenge with single DTs is the trade-off between the number of decisions D (i.e., model complexity) and overfitting. The most straightforward way to address this is to use an ensemble of DTs. The simplest ensemble is a random forest, in which the dataset X is randomly sampled M times (i.e., bagging) to generate M unique DTs; the outputs of these M trees can be aggregated via majority voting or averaging. This is a straightforward way to increase model complexity but keep the depth of the trees at a minimum.

The other popular ensemble approach is boosted decision trees (BDT), in which boosting

refers to the use of one DT's output to inform the structure of subsequent trees [165]. The specific package used for this work is **XGBoost** [166]. We summarize how boosting is implemented in **XGBoost**.

For the t 'th tree in this approach, let us define the objective function as:

$$O(y_i, T_i^t) = \sum_i^{N_d} F(y_i, T_i^t) + \gamma N_{leaf} + \lambda \sum_j^{N_{leaf}} T^{t,k}, \quad (3.2)$$

where F is some training loss (we will use mean-square error for this analysis), $T_i^t \equiv T^t(\vec{x}_i) = \sum_j^{N_{leaf}} T^{t,j}(\vec{x}_i)$ is the prediction score of the t 'th tree on $\vec{x}_i$ (which comes from the leaf with index j), N_{leaf} is the number of leaves at the bottom of the tree, and $T^{t,j} = \sum_i^{N_d} T^{t,j}(\vec{x}_i)$ is the total score of the training data set at the j 'th leaf. Here, λ and γ are arbitrary scaling constants (or hyper-parameters).

Assuming the training loss is reasonably smooth, we Taylor expand the training loss F up to second order to obtain

$$O(y_i, T_i^t) \simeq \sum_{i=1}^{N_d} \left(F(y_i, T_i^{t-1}) + g_i T_i^t + \frac{1}{2} h_i (T_i^t)^2 \right) + \frac{\lambda}{2} \sum_{j=1}^{N_{leaf}} (T^{t,j})^2 + \gamma T \quad (3.3)$$

$$= \sum_{j=1}^{N_{leaf}} \left(G_j T^{t,j} + \frac{1}{2} (H_j + \lambda) (T^{t,j})^2 \right) + \gamma T + \text{constants} \quad (3.4)$$

where $g_i = \partial_{T_i^{t-1}} F(y_i, T^t)$, $h_i = (\partial_{T_i^{t-1}})^2 F$, $G_j = \sum_i^N g_i$, $H_j = \sum_i^N h_i$, and the one-to-one correspondence of T_i^t with a corresponding leaf j lets us use a single summation index. By

taking derivatives with respect to T_j^t , the optimal value of the objective function is:

$$T_j^{t,\text{optimal}} = -\frac{G_j}{H_j + \lambda} \quad (3.5)$$

$$O(y_i, T_i^t)^{\text{optimal}} = -\frac{1}{2} \sum_j^T \frac{G_j^2}{H_j + \lambda} + \gamma T. \quad (3.6)$$

From the above, the key takeaway regarding boosting as an ensemble model is that properties and information from the previous tree are encoded via the gradient terms g_i and h_i and inform the properties of subsequent trees.

Evaluating Performances

We aim to evaluate the performance of the BDT using various methods. The simplest method is to measure the following quantities: accuracy, true positive (TP), true negative (TN), false positive (FP), and false negative. Illustrated in Fig. 3.1, the amount of true (false) positives is determined by a given model's ability to correctly (incorrectly) label signal as signal; the converse also applies for false (true) negatives for labeling background as signal.

With this defined, it is straightforward to define the following standard metrics for a given model: precision = $TP/(FN+TN)$, recall = $TP/(FN+TP)$, accuracy = $(TP+TN)/(TP+TN+FP+FN)$, true positive rate $TPR = 1 - TP/(TP+FN)$, and false positive rate $FPR = FP/(FP+TN)$.

Since the BDT provides a signal probability, we can set a threshold value that determines whether the BDT labels an event as signal. This ability to vary the threshold probability, thereby changing the number of TP/FPs, allows us to see how TPR varies with FPR. This corresponding distribution is also known as the Receiver Operator Characteristic (ROC) curve [167, 168], where the area under the ROC curve (AUC) provides a measure of how

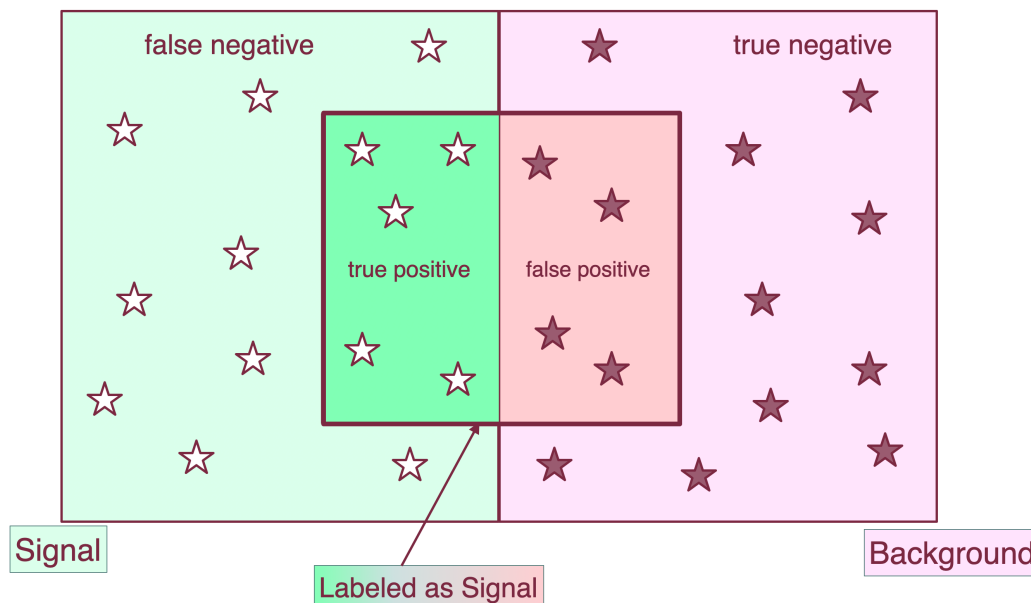


Figure 3.1: Given a data set (illustrated by stars) and a model’s prediction (which is represented by the central box), one can classify the performance of a given model by determining how many true/false signals were caught relative to the number of true/false background events which were labeled.

well-trained a model is. The greater the deviation of the AUC from 1, the more likely a model is either poorly trained or overtrained.

Since DTs are particularly vulnerable to overfitting, we use ROC AUC as our primary evaluation metric during hyperparameter tuning and training. With ROC curves, ensemble learning, and shallow tree depths, this should be enough to circumvent this concern.

Sidebands

To characterize the background shape during signal extraction, it is advantageous to define a signal window and its sideband. For the signal window, we choose a region that contains 90% of our signal events (see Sec. 3.2.5); we use the remaining region as a sideband to estimate background levels within the signal window.

In the case where we see enough events to take advantage of the structure seen in our backgrounds, we assume that background component i (see Sec. 3.2.3 for details) is represented by a probability density function $p_i(x)$ (where $x \mapsto m_{\text{recoil}}^2$ is our fitting variable).

Given a simulated or real background dataset with $N_{\text{tot}} = \sum N_i = \sum (N_{sb,i} + N_{sw,i})$ data points, there will be $N_{sb} = \sum N_{sb,i}$ counts in the sidebands and $N_{sw} = \sum N_{sw,i}$ in a given signal window. We predict $N_{sw,i}$ by fitting $p = \sum \alpha_i p_i$ (where $\alpha_i = N_{sb,i}/N_{sb}$) to the data in the sidebands via **ZFIT**'s implementation of unbinned negative log-likelihood estimation [169].

Assuming the results of fitting p to the sidebands are reasonable, we estimate the number of events from the i 'th background component in the signal window to be

$$N_{sw,i} = N_{sb,i} \frac{\int_{sw} p_i(x) dx}{\int_{sb} p_i(x) dx} \equiv N_{sb,i} \tau_{sw}. \quad (3.7)$$

In the case where we observe no events in our sidebands, we drop the background-related terms from our likelihood. In the scenario where we see too few events (i.e., under 5 events), we assume a single uniform distribution to describe our background. This simplifies our estimation of the background in the signal window to be $N_{sw} = N_{sb} \Delta_{sw}/\Delta_{sb} \equiv N_{sb} f$, where Δ_{sw} and Δ_{sb} are the widths of the signal and sideband regions, respectively. Details of how these quantities are incorporated in the likelihood function are discussed in Sec. 3.1. The effectiveness of the two different scenarios is outlined in Sec. 3.3.4.

Counting Experiment

As elaborated in Secs. 3.1 and 3.2.4, we will have removed enough events to perform a single-bin counting experiment to determine the numbers of signal and background events. In this subsection, we outline our procedure [170].

For a given collection of cuts and a pre-determined signal window with N_{sw} number of events, we expect to have $N_{sw} \geq 0$ but need to distinguish between signal and background events. We do this as outlined in Sec. 3.1, using events in the sideband to estimate background levels in the signal window. In this way, we lessen our reliance on Monte Carlo simulation and require less modeling in the post event selection stage.

To begin, let's take the simplest case where our sidebands are empty (i.e., $N_{sb} = 0$). In this case, our background expectation is set to zero, so any events in the signal window should directly correspond to signal events, yielding our likelihood to be:

$$L(s|N_{sw}, N_{sb} = 0) = P_{\circlearrowleft}(s|N_{sw}) f(\mathcal{L}) f(\eta) \quad (3.8)$$

where $P_{\circlearrowleft}(s|N_{sw})$ is a Poissonian distribution¹ with an expectation of s and N_{sw} number of occurrences, while $f(\mathcal{L})$ and $f(\eta)$ are Gaussian constraints of $s = \eta \mathcal{L} \sigma$ for uncertainties related to luminosity $\mathcal{L}$ and signal efficiency η , respectively. For the Belle II luminosity, $\mathcal{L} \equiv 364 \pm 1.84 \text{fb}^{-1}$, we incorporate the luminosity's uncertainty as a constraint term via a normal distribution with $\sigma = 1.84/364$.

For the efficiency term $f(\eta)$, we note that this is a model-dependent term that relies on ϵ , m_{χ_1} , and $m_{A'}$. To simplify this for a given m_{χ_1} , we scan through our range of ϵ and find the model configuration with the largest signal reconstruction efficiency (η). Since this is the model we will most likely find, we use its $\delta\eta/\eta$ as the definition of the one-standard-deviation variation for the corresponding normal constraint.

In the case where $N_{sb} > 0$, we assume N_{sw} and N_{sb} both come from a Poissonian distribution.

¹ $P_{\circlearrowleft}(\mu|N) = e^{-\mu} \mu^N / N!$

We wish to maximize the following likelihood distribution L :

$$L(s, b | \tau, N_{sw}, N_{sb}) = P_{\circlearrowleft}(s + b | N_{sw}) P_{\circlearrowleft}(\tau b | N_{sb}) f(\mathcal{L}) f(\eta). \quad (3.9)$$

The modification from the $N_{sw} = 0$ case is that we assume that the background contribution is uniform, so we relate the events in the sideband to the background contribution in the signal window via $N_{sw} \sim N_{sb}/\tau_{sw}$ where $\tau_{sw} = \Delta_{SW}/\Delta_{SB}$ describes the ratio of areas in the sideband vs. signal window. Otherwise, all other terms remain the same.

As a side note, if we were to take a Bayesian approach, we would need to consider the additional case where $N_{sw} = 0$. When marginalizing over the full likelihood, the background prior decouples from the signal prior, reverting to the likelihood in the $N_{sb} = 0$ case. However, because we're performing a frequentist analysis and considering the profile likelihood, we do not have the same luxury.

In the event that we want to consider multiple background sources (ex. $e^+e^- \rightarrow \gamma\gamma + e^+e^- \rightarrow K_L^0 K_S^0 \gamma$, continuum, and others), we can generalize the likelihood function to the following:

$$L(s, b_i | N_{sw}, N_{sb}, N_i) = P_{\circlearrowleft} \left(s + \sum_i b_i \middle| N_{sw} \right) f(\mathcal{L}) \prod_i f(N_i) f(\tau_i).$$

Rather than a simple scale relation to map the N_{sb} to b , we prescribe a probability distribution function p_i to each contribution and define each component's contribution as

$$b_i = N_i \int_{sw} p_i(x | \theta_i) dx / \int_{sb} p_i(x | \theta_i) dx \equiv N_i \tau_{sw},$$

where N_i corresponds to the normalization of each background component from fitting the background in the sideband. When we perform an extended unbinned negative log-likelihood

analysis in our sidebands, the uncertainties of the yields provide the σ_{N_i} , allowing us to use correlated shape systematics for N_i rather than a Poissonian distribution as is done in the uniform background case.

However, if event rates are very low, attempting to fit PDFs will inflate the uncertainties on N_i ; in this situation, we assume a uniform background contribution to avoid this inflation.

To initialize the Poissonian distributions, we minimize the likelihood in Eq. 3.9 and determine σ , using **PyHF** [171, 172] with **iminuit** [173] as the backend for the optimization.

Tools for Statistical Analysis

In Sec. 3.1, we define a likelihood with constraints based on uncertainties given by fitting. In this section, we describe the tools used to estimate the number of background events in a given signal window, construct the likelihood function, extremize it, and incorporate uncertainties via constraints or modifiers.

Following the success of tools such as **HISTFACTORY** [174] for statistical analyses in **ROOT**, Pythonic tools such as **PyHF** [171, 172] and **scikit-HEP** [175] were developed to streamline binned statistical analyses. While tools such as **ZFIT** [169] can also perform counting experiments, the general convenience and accessibility of the likelihood functions make it appealing. In addition, as shown in other analyses [176], both frequentist and Bayesian approaches to upper limits are possible. We adopt the frequentist approach [170, 177], which I outline below.

PyHF uses dictionaries to create the likelihood function. Given the counting experiment we are trying to perform (for clarity, assume Eq. 3.9), we define two channels to describe the likelihoods for the signal window and sideband. The sideband sample, which constructs $P_{\infty}(\tau b | N_{sb})$, has one background channel with a free parameter b , which is scaled by a

factor $\tau = \Delta_{sb}/\Delta_{sw}$, and is fitted to the observable N_{sb} . For the signal window sample, which constructs $P_{\odot}(s + b|N_{sw})$, two free parameters s and b are defined and fitted to the observable N_{sw} .

Constraint terms on both s and b are incorporated via ‘modifiers.’ For quantities where a central value and uncertainties are provided (e.g., luminosity, signal reconstruction efficiency, background PDF uncertainties), multiplicative modifiers are used. Given a variable x , a nominal value $\langle x \rangle$, and uncertainty $\delta_x^{\pm}$, **PyHF** implements a piecewise linear interpolation between a standard normal distribution $\mathcal{N}(\alpha)$ and normalization term $1 - \delta_x^{\pm}/\langle x \rangle$. The piecewise linear interpolation maps $\alpha = 0$ to $f(x) = 1$ and $\alpha = \pm 1$ (or 1σ) to $f(x) = 1 \pm \delta_x^{\pm}/\langle x \rangle$. During the profile likelihood tests, α should vary by $\pm 2\sigma$ to incorporate the impact of systematic uncertainties into our upper limit calculations.

Upper Limit Setting

Assuming we don’t observe an excess of signal, we will perform an upper-limit calculation at a 95% confidence level (CL).

Given a likelihood function² $L(s, \theta|N_{sw}, N_{sb})$ with particular values of N_{sw} and N_{sb} , we want to determine what value of s is still likely relative to a null hypothesis (i.e., $s \equiv 0$) via a profile likelihood ratio $\lambda(s) = L(s, \hat{\hat{\theta}})/L(\hat{s}, \hat{\theta})$. The numerator is the likelihood function of a given value of s with other parameters $\hat{\hat{\theta}}$ maximized via a conditional maximum-likelihood (ML) estimation. The denominator is the likelihood of the best-fit values of both the signal yield and the other parameters. As our parameter of interest must satisfy $s \geq 0$, we elect to

²For brevity, I will suppress the explicit dependence on N_{sw} and N_{sb} .

use the $\tilde{q}$ test statistic for upper limits, which is defined to be

$$\tilde{q}_s = \begin{cases} -2 \ln(L(s, \hat{\hat{\theta}}(s))/L(0, \hat{\hat{\theta}}(0))) & \hat{s} < 0 \\ -2 \ln(L(s, \hat{\hat{\theta}}(s))/L(\hat{s}, \hat{\hat{\theta}}(0))) & 0 \leq \hat{s} < s \\ 0 & \text{otherwise} \end{cases} \quad (3.10)$$

By sampling 2000 toy datasets³, we construct the PDF that describes q_s , given s , as $f(\tilde{q}_s|s)$ and define the null-hypothesis p -value as $p_s = \int_{q_{s,obs}}^{\infty} f(q_s|s)dq_s$, where the value of s that yields a p -value of 0.05 provides a median upper limit on s at a 95% confidence level. We use `HEPSTATS` [175]’s implementation of a toy-based upper limits calculator, as that interfaces with `PYHF`.

³Due to the expected limited statistics, the likelihood ratio cannot be approximated as a Gaussian.

3.2 Reconstruction and Selection

In this section, we discuss how we reconstruct signal events and determine the selection criteria to reject background. The criteria used to select the ISR photon, the two tracks, the displaced vertex formed from said tracks, and the event topology itself are described, respectively, in Tables 3.2, 3.3, 3.4, and 3.5. Many of the selection criteria for the quantities used in the aforementioned tables are recommended by the relevant Belle II performance groups (e.g., neutrals and PID). Thus, we expand on the unique cuts and selections tailored for this analysis in the subsequent subsections.

3.2.1 Background Generation

Before outlining our reconstruction, we enumerate the range of possible processes that are considered as background. The background data collections used in this analysis are summarized in Table 3.1. For these generic background samples, we employ an ensemble of generators to simulate various physics processes.

For processes with two particles (fermions or bosons) in the final state ($e^+e^- \rightarrow f\bar{f}$), we employ **BABAYAGA@NLO** [178] for large angle Bhabha scattering ($e^+e^- \rightarrow e^+e^-(\gamma)$) and exclusive photon final states, and **KKMC** [179] for $\mu^+\mu^-$, $\tau^+\tau^-$, $q\bar{q}$ production. An additional sample of an extensive large-angle Bhabha scattering was also produced using **BHWIDE** [180] and included to improve our characterization of these events. In all three cases, the generator also generates the initial- and final-state radiated photons. For processes with four charged particles (fermions, mesons) in the final state, **AAFH** [181] is used to handle the purely lepton final states ($e^+e^-e^+e^-$, $e^+e^-\mu^+\mu^-$, $e^+e^-\tau^+\tau^-$, $\mu^+\mu^-\mu^+\mu^-$, and $\mu^+\mu^-\tau^+\tau^-$) and **TREPS** is used to handle a two virtual photon radiative state, where one photon decays to e^+e^- and the other decays to $\pi^+\pi^-$, K^+K^- , or $p\bar{p}$. For the various B physics processes, **EVTGEN** [182]

Process	Generator	$\int \mathcal{L} dt \text{ (fb}^{-1}\text{)}$	$N_{streams}$
$e^+e^- \rightarrow \tau\tau(\gamma)$	KKMC	1458.959	$4 \times \mathcal{L}_{run\ 1}$
$e^+e^- \rightarrow q\bar{q} \text{ (} q \in \{u, d, s, c\}\text{)}$	KKMC	1458.959	$4 \times \mathcal{L}_{run\ 1}$
$e^+e^- \rightarrow B\bar{B}$	EVTGEN	1458.959	$4 \times \mathcal{L}_{run\ 1}$
$e^+e^- \rightarrow \gamma\gamma$	BABAYAGA@NLO	729.479	$2 \times \mathcal{L}_{run\ 1}$
$e^+e^- \rightarrow \mu^+\mu^-(\gamma)$	KKMC	1458.959	$4 \times \mathcal{L}_{run\ 1}$
$e^+e^- \rightarrow e^+e^-e^+e^-$	AAFH	364.74	$1 \times \mathcal{L}_{run\ 1}$
$e^+e^- \rightarrow e^+e^-\mu^+\mu^-$	AAFH	364.74	$1 \times \mathcal{L}_{run\ 1}$
$e^+e^- \rightarrow llXX$ $e^+e^- \rightarrow e^+e^-\tau^+\tau^-$ $e^+e^- \rightarrow e^+e^-\pi^+\pi^-$ $e^+e^- \rightarrow e^+e^-K^+K^-$ $e^+e^- \rightarrow e^+e^-p\bar{p}$ $e^+e^- \rightarrow \mu^+\mu^-\mu^+\mu^-$ $e^+e^- \rightarrow \mu^+\mu^-\tau^+\tau^-$	AAFH TREPS TREPS TREPS AAFH AAFH	364.74	$1 \times \mathcal{L}_{run\ 1}$
$e^+e^- \rightarrow hh\gamma_{ISR}$ $e^+e^- \rightarrow \pi^+\pi^-\gamma$ $e^+e^- \rightarrow \pi^+\pi^-\pi^0\gamma$ $e^+e^- \rightarrow K_s^0 K_L^0 \gamma$ $e^+e^- \rightarrow K^+K^-\gamma$	PHOKHARA	364.74	$1 \times \mathcal{L}_{run\ 1}$
$e^+e^- \rightarrow e^+e^-(\gamma)$	BABAYAGA@NLO BHWIDE	36.474 345	$0.1 \times \mathcal{L}_{run\ 1}$ $\sim 1 \times \mathcal{L}_{run\ 1}$

Table 3.1: Generic background samples, their primary physics generators, and the amount of data (in units of integrated luminosity) simulated for each sample. Along with the primary generators, secondary generators are also used to handle the decay of τ (TAUOLA [183]), the fragmentation of quarks (PYTHIA8 [184]), and the QED radiative corrections (PHOTOS [185, 186]).

is used to handle the production of $e^+e^- \rightarrow B\bar{B}$ (both charged and neutral modes).

Along with the primary generators mentioned above, secondary generators are called to handle the decays of some of the particles mentioned above. To deal with the fragmentation of quarks (e.g., in $e^+e^- \rightarrow q\bar{q}$), PYTHIA8 [184] is used. TAUOLA [183] is used to deal with the decays of $\tau^\pm$ (e.g., in $e^+e^- \rightarrow \tau^+\tau^-$, $e^+e^- \rightarrow \ell^+\ell^-\tau^+\tau^-$). Finally, PHOTOS [185, 186] is used to compute QED radiative corrections in various decays (e.g., B decays from EVTGEN).

Cut on ISR Photon	Variable Description
$E_{cms} > 1 \text{ GeV}$	ISR CM frame energy
clusterReg = 2	Selecting ECL clusters in the barrel
$ t_{ISR} < 200 \text{ GeV}$	Cluster Timing
ZernikeMVA > 0.7	MVA output trained on Zernike Moments
ClusterPulseShapeMVA > 0.01	MVA output trained on ECL pulse shapes to discriminate between electromagnetic (1) vs. hadronic (0) showers

Table 3.2: Cuts on the ISR photon candidate

3.2.2 Reconstruction

Due to the high levels of backgrounds in the endcaps of the ECL, we predominantly use photons that are found in the barrel ECL. Furthermore, we demand that they have high energy and do not appear to arise from K_L^0 via cuts in Table 3.2. The cut on the photon CM energy is determined from studies for the single-photon analysis. One cut useful for this analysis, if there is isolation between the ISR photon and the displaced vertex, is the ZernikeMVA output; this cluster-shape variable helps separate neutral non-photon clusters from photon clusters. (We will highlight cluster-track isolation later.) Additionally, a modest cut on the ClusterPulseShapeDiscriminationMVA is applied to suppress ISR photon candidates that are obviously hadronic (K_L^0 or neutron). To reiterate a point made in Sec. 1.5, we apply a photon-efficiency correction to the pre-LS1 data with respect to MC15rd, although many of the weights of the photon correction are near unity.

The standard Belle II TreeFit algorithm [187] is used to reconstruct the displaced vertex. The inputs to the vertexing algorithm are the two charged tracks; we keep all vertices that are successfully reconstructed. For simplicity, we use the $e\pm$:ALL reconstructed-particle lists and apply the cuts outlined in Table 3.3. The dr cut removes tracks that originate directly from the IP. The cut on the number of CDC hit removes low-momentum curling tracks that originate from beam-gas interactions or two-photon processes. Fig. 3.2 also suggests that

Cut on V_0 daughters	Variable Description
$dr > 0.1$ cm	distance ($ d_0 $ for tracks) w.r.t IP
nCDCHits > 4	Number of CDC hits associated to the tracks
$E_{ECL} < 2$ GeV	ECL Cluster Energy

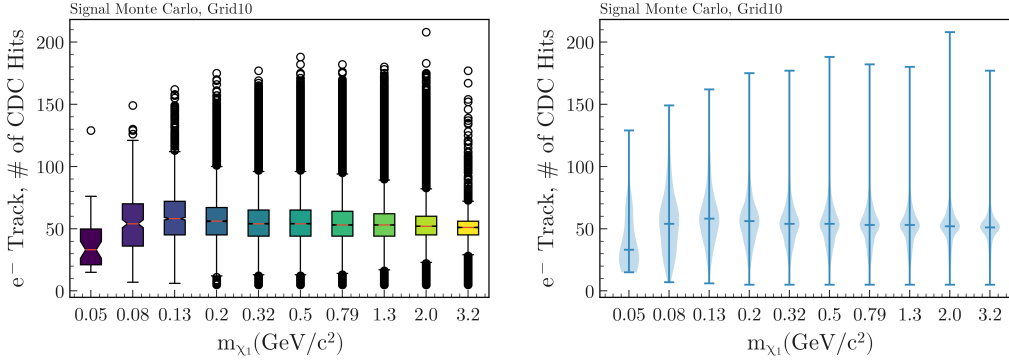
Table 3.3: Cuts on the tracks used to form the displaced vertex (V_0).

Figure 3.2: Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the number of CDC hits of the negatively charged track (with an electron mass hypothesis) for each model of a given m_{χ_1} . The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.

tracks originating from displaced vertices typically have a large number of CDC hits, so the loose cut we place should be sufficient for removing these low-momentum backgrounds while having a minuscule effect on signal efficiency. However, tracking systematics studies claim that cuts on the number of CDC hits introduce a charge asymmetry, so applications of cuts on CDC hits should be used with caution and be as loose as possible. We confirm the necessity of this cut in Sec. 3.3, as we relax this cut in control samples.

The cuts applied to the displaced vertices formed from the tracks are outlined in Table 3.4. These cuts perform preliminary selections to ensure that the vertices do not originate from the IP or exhibit unphysical properties. We suppress two dominant sources of displaced vertices ($K_s^0 \rightarrow \pi^+\pi^-$, and $\gamma \rightarrow e^+e^-$) by applying two additional mass cuts; these selections are outlined in Secs. 3.2.2 and 3.2.2. For the pointing angle ($\cos \angle_{PA} \equiv \hat{p} \cdot d\hat{r}$), we find that many continuum, $\gamma\gamma$, and light-signal models have events with collinear pointing angles. We

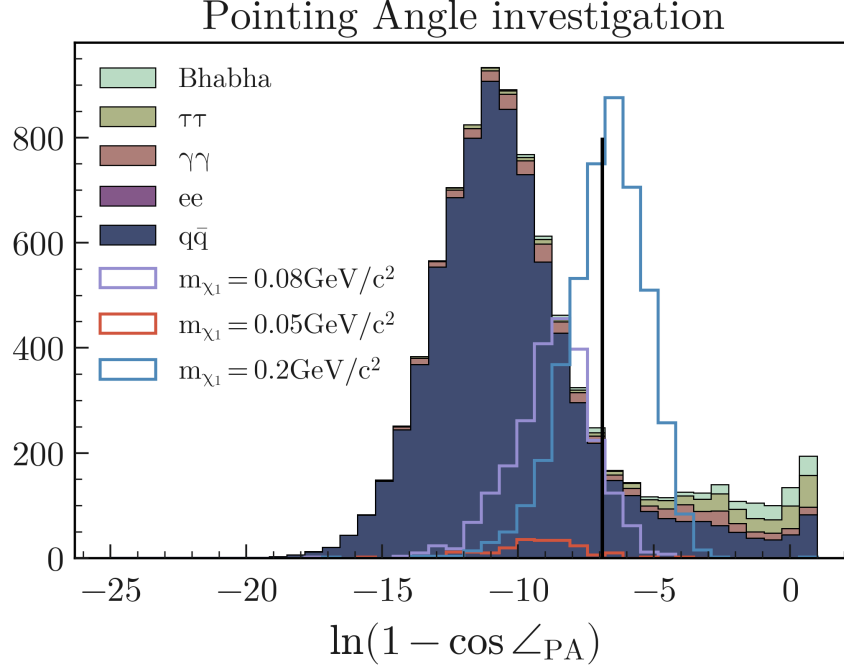


Figure 3.3: Showcasing the pointing angle ($\ln(1 - \cos \angle_{PA})$) distribution of background and light signal simulation.

transform the pointing angle by using $\ln(1 - \cos \angle_{PA})$ to help us distinguish between the different event types near $\cos \angle_{PA} \sim 1$. From Fig. 3.3, where $\ln(1 - 0.999) = -6.9$, we see that $q\bar{q}$ and $\gamma\gamma$ peak near $\ln(1 - \cos \angle_{PA}) \sim -12$. As the background contributions begin to diminish near $\ln(1 - 0.999) = -6.9$, we elect to place a cut on the pointing angle there.

Although the background samples are strongly suppressed, Fig. 3.3 also shows that this cut adversely reduces the signal efficiency for our lighter models ($m_{\chi_1} = 0.05, 0.08$); however, the expected sensitivity in Fig. 5 of Ref. [84] suggests that these models are also ruled out by other experiments (NuCal and MiniBooNE), so we prioritize models with $m_{\chi_1} > 0.2 \text{ GeV}/c^2$ and therefore accept the reduced efficiency of these lighter models.

Finally, we have Table 3.5 to outline cuts that consider properties of the event as a whole. The angle between the ISR photon and the vertex is considered to ensure the isolation of the ISR photon and suppress $\tau\tau$ backgrounds, as Fig. 3.4 suggests most of our signal's

Cut on V_0	Variable Description
$p > 0.1 \text{ GeV}/c$	Vertex's Momentum
$p_{cm} < 2.0 \text{ GeV}/c$	Vertex's momentum in CM frame
$d\cos\theta \in (-0.6, 0.95)$	Polar angle of vertex w.r.t. IP
$dr > 0.5\text{cm}$	Transverse distance with respect to IP for a vertex
$d_0 < 0.5 \text{ cm}$	d_0 , distance from IP in $r - \phi$ plane, of each daughter
$z_0 < 1 \text{ cm}$	d_0 , or z distance from IP, of each daughter
$\vec{p} \cdot d\vec{r}/p dr < 0.999$	Displaced vertex's pointing vector
$m_{\gamma \rightarrow e^+e^-} > 0.01 \text{ GeV}/c^2$	Vertex mass, assuming it comes from photo-conversions
$ m_{V_0 \rightarrow \pi^+\pi^-} - m_{K^0} > 20 \text{ MeV}/c^2$	Vertex mass, assuming the vertex daughters are pions

Table 3.4: Cuts for the formed displaced vertex.

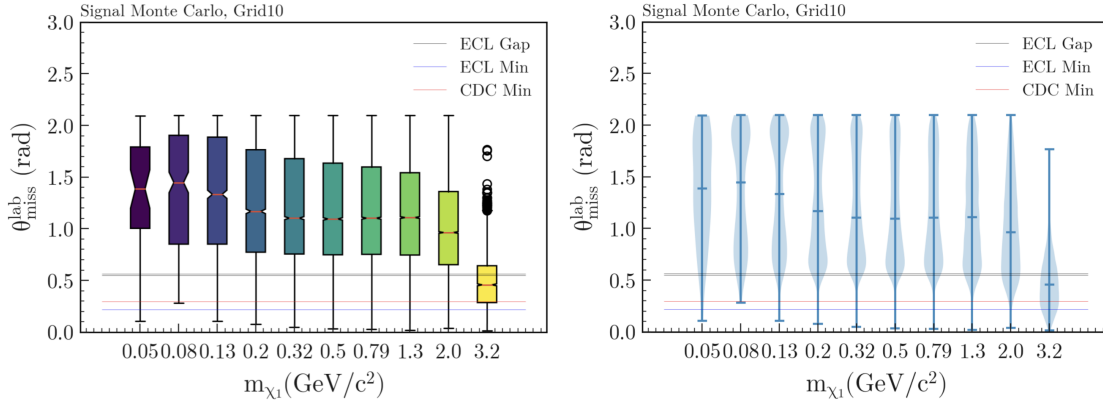


Figure 3.4: Binning signal models by m_{χ_1} , we show missing momentum's polar angle. The lower boundaries of the CDC and ECL, along with the $\sim 1^\circ$ gap between the forward and barrel ECL, are also indicated. While the left and right plots are of the same distribution, the left is a representation using box-whiskers, while the right is a violin plot.

missing momentum is within the barrel or forward regions. We also ensure that the missing momentum vector does not point into dead regions of the detector, nor the backward region. Finally, we demand events of interest to have only one ISR photon candidate and two tracks.

With the general outline of how we conduct track selection, we will take a brief interlude to discuss relevant observables relevant to other sections of this dissertation.

Cut on Event	Variable Description
$\cos \angle_{V_0, \gamma} < 0.88$	Angle between ISR momentum and V_0 position
$\theta_{ISR, e^\pm} > 0.5$ rad	Angle between ISR and V_0 daughters' momenta
$\theta_{miss}^{lab} \in (0.297, 2.1)$ & $\theta_{miss}^{lab} \notin (0.54, 0.57)$	Polar angle (rad) of the missing momentum
$n_{ISR \text{ Photon}} = 1$	Number of ISR photons in the event
$n_{Tracks} = 2$	Number of tracks in the event

Table 3.5: Overall cuts on the signal event itself.

K_s^0 Mass

A dominant background we expect is $K_s^0 \rightarrow \pi^+ \pi^-$, which produces displaced vertices. To account for this, we make a 10σ cut near $m_{K_s^0}$ to suppress them. Firstly, we redo vertexing assuming a pion hypothesis for the two tracks. From Fig. 3.5, we see that fitting our background simulation, rich in continuum, to a Gaussian plus a linear background yields $\sigma \sim 2 \text{ MeV}/c^2$, setting our K_s^0 mass veto to be $\pm 20 \text{ MeV}/c^2$.

For signal models where $\Delta m_\chi \lesssim m_{K_s^0}$, this cut has little to no effect due to kinematics, as the vertex mass is bounded above by Δm_χ . For heavier signal models, this would affect our signal efficiency, but the distributions are broad across the entire vertex-mass range; the impact on signal reconstruction efficiency is marginal, as shown in Fig. 3.6.

Converted Photons

Another dominant background arises from photons converting into a displaced e^+e^- vertex. The predominant source of these displaced vertices is photon interactions with the Belle II detector medium, resulting in e^+e^- pair production. To account for this, we look at MC samples $e^+e^- \rightarrow \gamma\gamma(\gamma)$ and Bhabha scattering (along with our other samples for completeness), as these provide an abundant source of converted photons.

Fig. 3.7 compares the quantity $m_{\gamma \rightarrow e^+e^-}$, or the mass of the displaced vertex under the

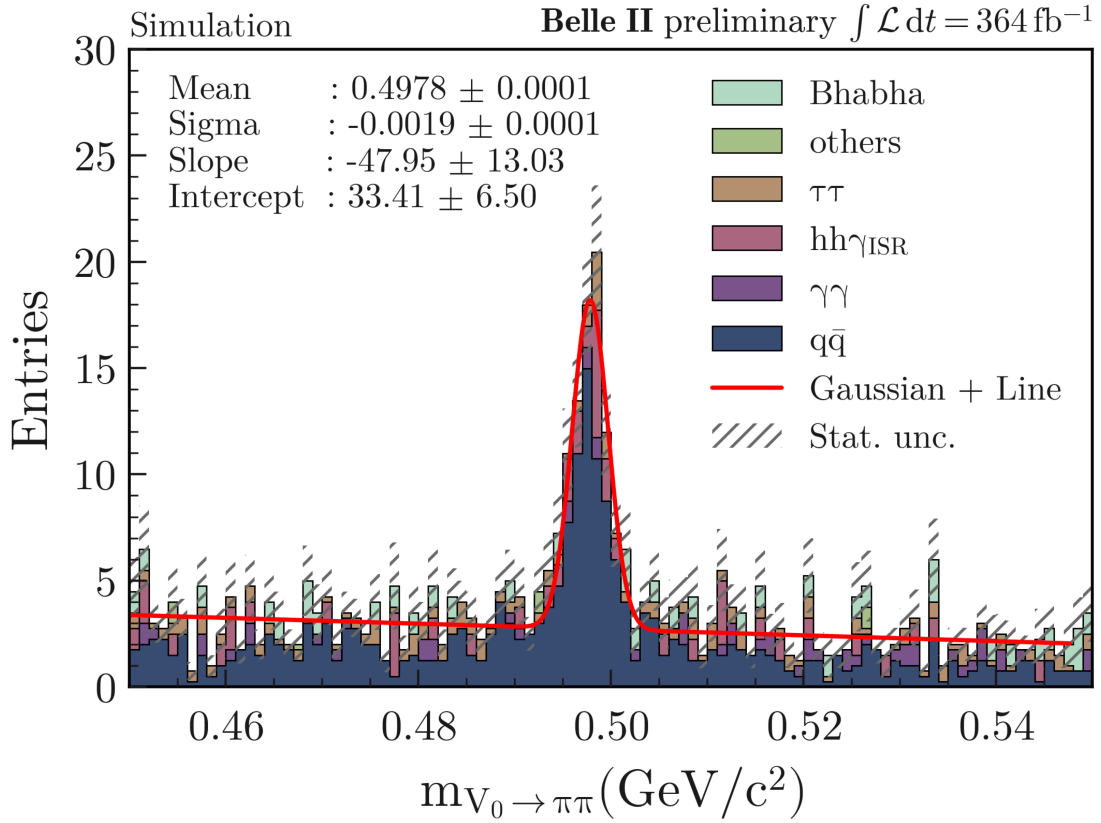


Figure 3.5: In MC background simulation, the vertex mass using $\pi^\pm$ tracks is shown near the K_s^0 mass. Using a Gaussian plus linear fit, the width is measured to be $\sigma \sim 2 \text{ MeV}$.

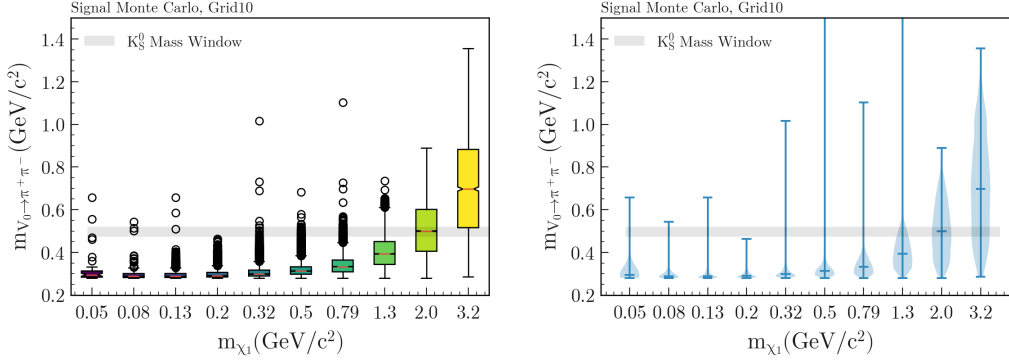


Figure 3.6: Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the number of CDC hits of the negatively charged track (with an electron mass hypothesis) for each model of a given m_{χ_1} . The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data in a violin plot. The K_S^0 mass window is also outlined, where the gray band indicates a $\pm 10\sigma$ width.

kinematic assumption that the angle between the two electron tracks is zero, between these converted photon-rich MC samples and some light mass signal models. Comparing this with a few lighter signal models, we obtain Fig. 3.7.

The Bhabha and $\gamma\gamma$ distribution sharply peaks near $m_{\gamma \rightarrow e^+e^-} = 0$ and sharply decreases, flatlining near $m_{\gamma \rightarrow e^+e^-} = 0.01 \text{ GeV}/c^2$, while our lightest signal models are much broader in comparison. Thus, we find that this cut at this value is sufficient to suppress photon conversions.

In addition, we use geometric considerations of the track helices to reject more background. We define $\Delta R = |r_1| + |r_2| - |\bar{C}_1 - \bar{C}_2|$, the difference between the sum of the circle radii and the center-to-center distance, to characterize the intersection of the two tracks in the transverse (r - ϕ) plane. We reject unphysical vertices with $\Delta R > 250 \text{ cm}$. A very powerful variable in selecting or suppressing converted photon-like events whose tracks emerge near zero opening angle, i.e., with $\Delta R \simeq 0$.

We therefore bound $\Delta R_{\gamma \rightarrow e^+e^-}$ above and below, and optimize on the Punzi Figure-of-merit ($\eta_{recon}/(1.5 + \sqrt{B})$) [188]. For the light ($m_{\chi_1} \leq 0.32 \text{ GeV}/c^2$) and medium models

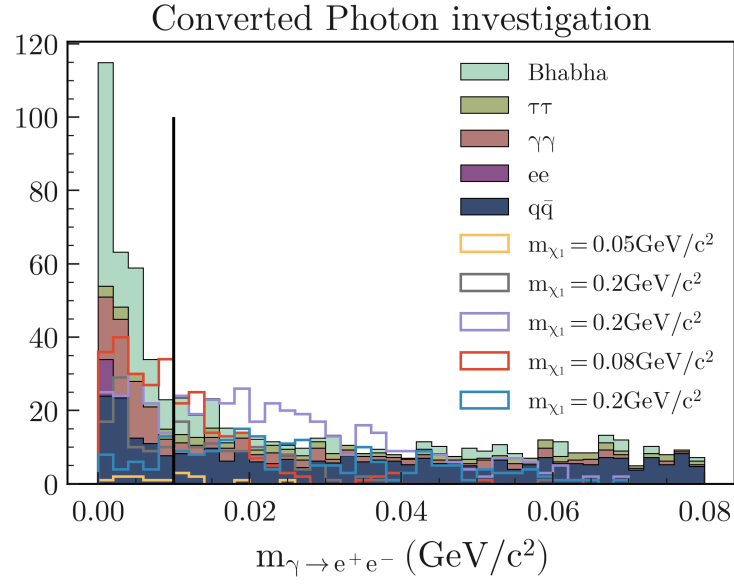


Figure 3.7: In MC background simulation, the vertex mass using $e^\pm$ tracks is shown near zero.

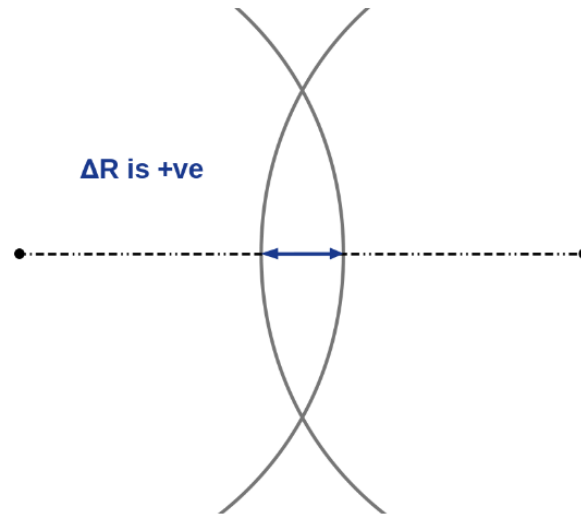


Figure 3.8: Illustration of the discriminating variable ΔR in the transverse plane for the two tracks that form a vertex.

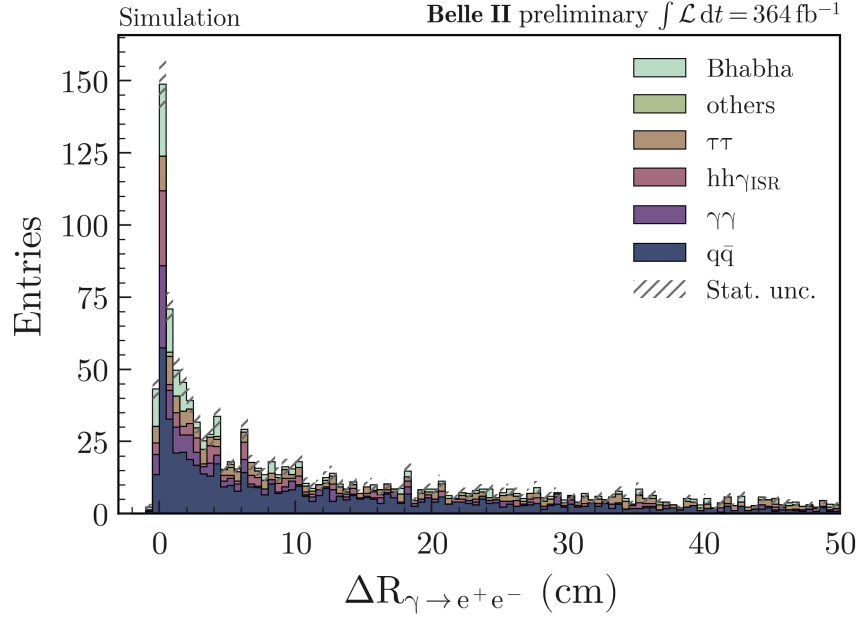


Figure 3.9: Illustration of the discriminating variable ΔR for expected background samples.

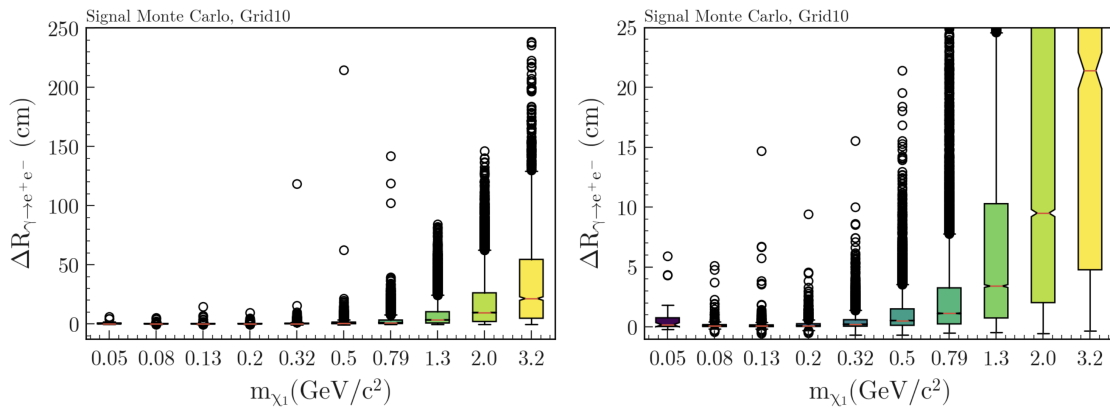


Figure 3.10: Illustration of the discriminating variable ΔR for simulated signal grouped by m_{χ_1} and visualized by box-whisker plots. The left shows the full view, while the right shows an expanded view near $\Delta R = 0$ to better visualize the lighter iDM models.

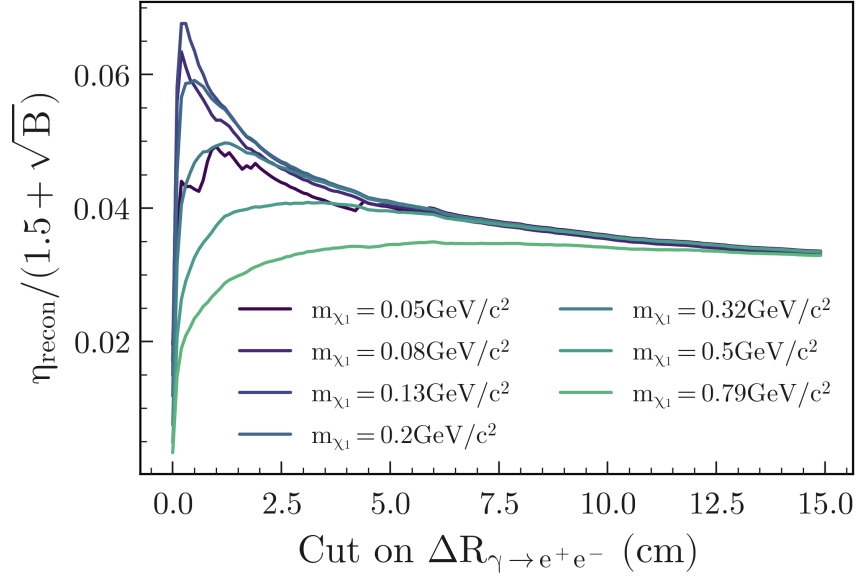


Figure 3.11: For light and medium iDM models, optimizing a cut on ΔR based on the Punzi figure-of-merit.

($0.32 < m_{\chi_1} < 1.3 \text{ GeV}/c^2$), these events are geometrically similar to converted photons so using the results shown in Fig. 3.11, $\Delta R < 1 \text{ cm}$ and $\Delta R < 6 \text{ cm}$, respectively reject most converted-photon backgrounds. Conversely, as we approach heavier models, this variable is no longer helpful, since the distributions peak near $\Delta R = 0$ and then decrease exponentially.

Track Momenta

For documentation purposes, we include the momenta and transverse momenta of the 0th track in Fig. 3.12 and 3.13, respectively. Aside from the lightest model, which has a low reconstruction efficiency due to its similarity to converted photons (i.e., $\gamma \rightarrow e^+e^-$), the general trend is that our range of momenta increases as we approach heavier models due to the increase in available energy.

One track class to highlight is curlers. For a transverse momentum p_T of $100 \text{ MeV}/c$, one calculates a radius of curvature to be $r = p_T/eB \sim 22 \text{ cm}$. The CDC ranges from ~ 16 to

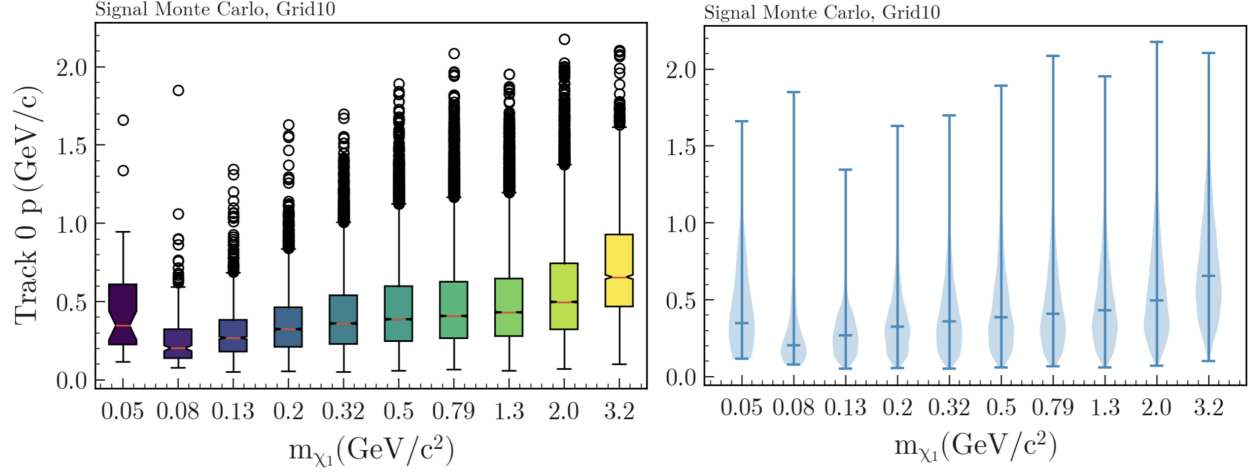


Figure 3.12: Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the range of momenta p for each model with a given m_{χ_1} . The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.

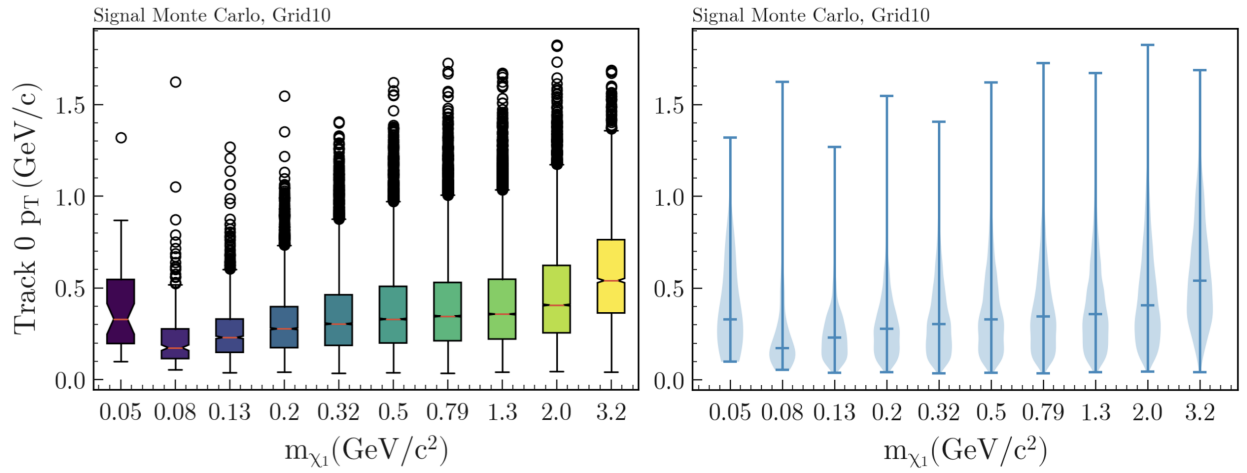


Figure 3.13: Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the range of transverse momenta p_T for each model with a given m_{χ_1} . The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.

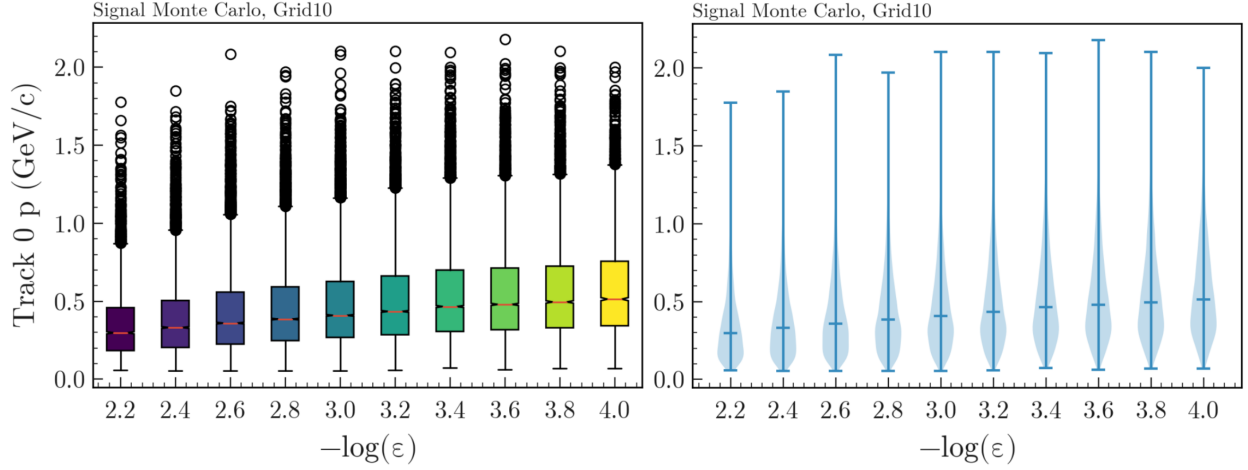


Figure 3.14: Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the range of momenta p for each model with a given $-\log(\epsilon)$. The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.

~ 110 cm from the IP, so low-momentum tracks will definitely curl within the CDC volume.

With cuts relating to the number of CDC hits, one might consider placing cuts on similar detector information from the VXD. The most natural cut to consider is the transverse distance in the r - ϕ plane of the vertex relative to the IP (dr_{V_0}) and the innermost layer that an associated track hit with transverse distance $dr_{in} \pm \delta dr_{in}$. Naively, cutting on vertices in which $dr_{V_0} > dr_{in} - \delta dr_{in}$ should remove many beam-background events without adversely impacting our signal reconstruction efficiency. However, as many of our iDM models produce curlers, cuts on dr_{in} decrease our background and signal reconstruction efficiency at near equal rates. We therefore refrain from utilizing dr_{in} .

Particle Identification

To determine how particle identification can be helpful in this analysis, three models of different m_{χ_1} values are selected to provide a representative picture.

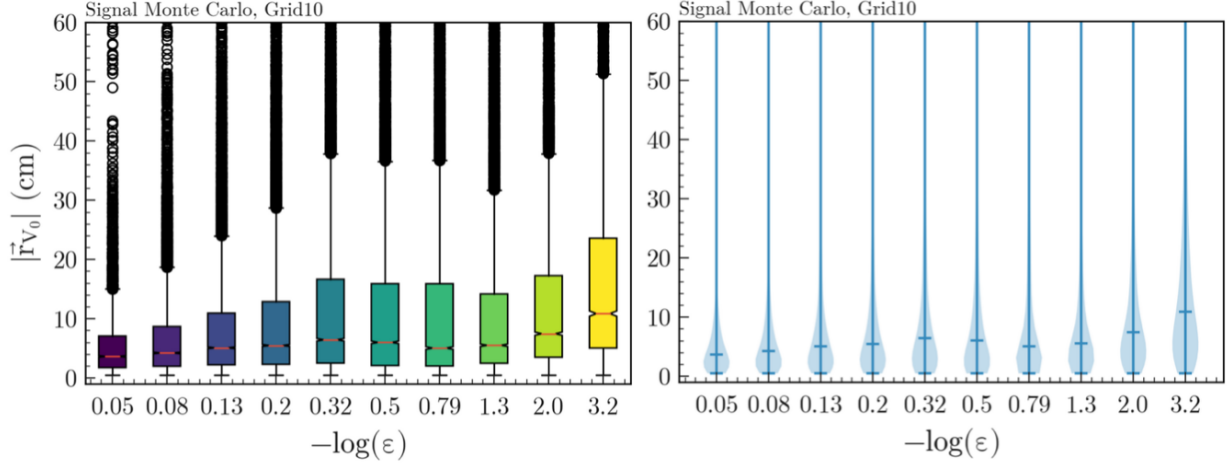


Figure 3.15: Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the range of vertex distances r_{V_0} for each model with a given $-\log \epsilon$. The left panel uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.

As we assume $A^{*} \rightarrow \ell^{+}\ell^{-}$ with $\ell \in [e, \mu]$, we expect the e ID and μ ID to be the most impactful. In addition, a disadvantage of working with displaced vertices is the unreliability of TOP in providing accurate hadron ID (viz., pion versus kaon discrimination), as the current TOP PID algorithm assumes that tracks originate from the IP. Fig. 3.15 highlights the range of displacements we see in our signal MC. As such, we will only highlight the lepton IDs.

First, our light model representative, iDM models with $m_{\chi_1} = 0.08 \text{ GeV}/c^2$ are energetically forbidden to produce muon tracks, so we primarily expect to rely on electron ID. As we move to heavier models ($\Delta m_{\chi} < 2m_{\mu}$), the muon decay channel becomes kinematically available, so the heaviest models should see a mix of muon and electron tracks. Finally, given the momentum ranges shown in Sec. 3.2.2 or in Fig. 3.12, not all muons will reach the KLM; this would worsen Belle II’s ability to easily distinguish between muons and pions, and we would expect several signal tracks to have low muon and pion ID. To highlight the above features, Fig. 3.16 and 3.17 show the 2D correlation plots between the minimum e - vs. μ -ID

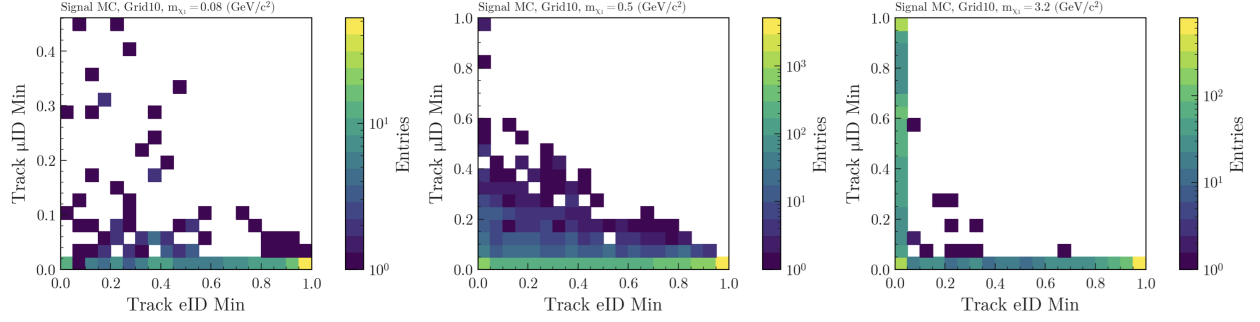


Figure 3.16: Given a light, medium, and heavy iDM model, we show the minimum muon ID of both tracks vs. the minimum electron ID of both tracks.

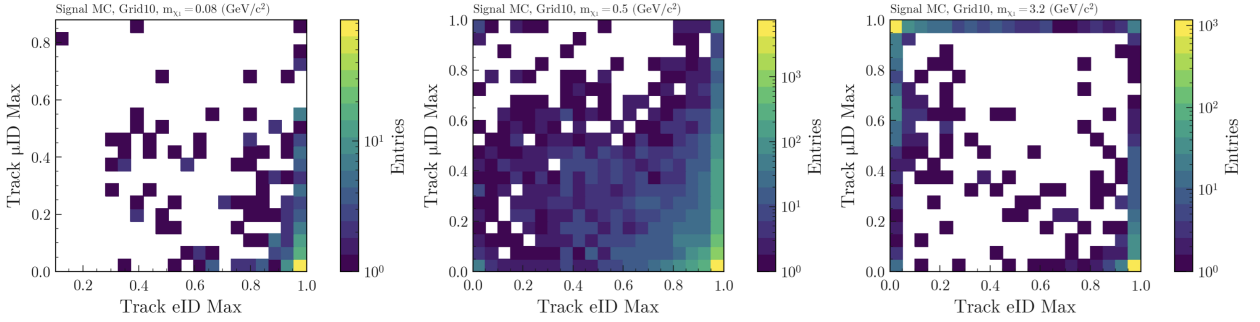


Figure 3.17: Given a light, medium, and heavy iDM model, we show the maximum μ ID of both tracks vs. the maximum e ID of both tracks.

and maximum e^- vs. μ -ID tracks, respectively, in signal simulation.

For $m_{\chi_1} = 0.08 \text{ GeV}/c^2$, we see both tracks typically have $eID \sim 1$ and very low μID . For $m_{\chi_1} = 3.2 \text{ GeV}/c^2$ (the heaviest signal model studied here), we see that, typically, one of the tracks is well-identified as either an electron or a muon. However, there is an abundance of tracks with low μID and eID , which most likely correspond to the low-momentum muons mentioned previously. Lastly, the $m_{\chi_1} = 0.5 \text{ GeV}/c^2$ model lies in the middle ground, where tracks are predominantly electrons, but there remains an abundance of signal tracks with at least one track exhibiting poor eID and μID .

Rather than exploiting this information directly by applying lepton-ID cuts at the outset, we use these variables as inputs in the construction of our BDT scheme.

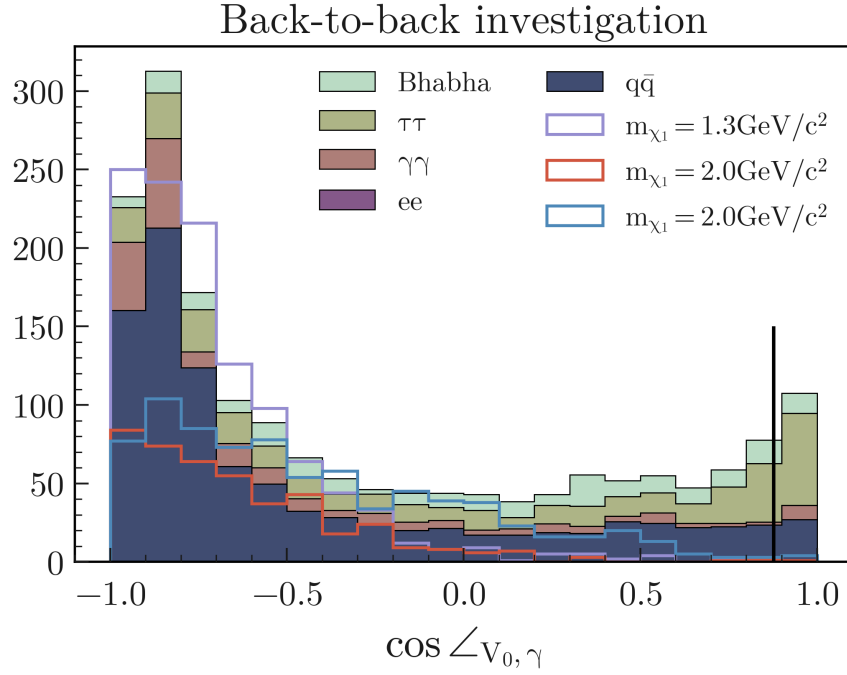


Figure 3.18: From MC simulation, illustrating $\cos \angle_{V_0, \gamma}$, or how back-to-back the displaced vertex and photon are between heavy signal models and background.

Back-to-backness

One useful quantity to highlight is $\cos \angle_{V_0, \gamma} = \hat{p}_\gamma \cdot \hat{dr}_{V_0}$, or the cosine of the angle between the displaced vertex and the ISR photon. As this describes how closely the two constituents are back-to-back, we refer to this quantity as the ‘back-to-backness’ of the event.

One important reason to incorporate this is to ensure that the ISR photon is isolated from the rest of the event, as the ZernikeMVA requires it. In addition, from Fig. 3.18, we see that $\tau\tau$ backgrounds (i.e., both τ ’s undergo one-prong decays) are also abundant when γ_{ISR} and the displaced vertex directions are nearly collinear. The heavy iDM signal models benefit the most from their separation from these tau-pair backgrounds. We find that a requirement of $\cos \angle_{V_0, \gamma} < 0.88$ is optimal.

Total Event Extra Energy

While we constrain events to only contain one energetic (≥ 1 GeV) photon and two tracks, there are still energy depositions in the event from low-energy photon candidates or beam background that can be used to help suppress additional backgrounds.

From our photon list ($\gamma : \text{ALL} + |t_\gamma| < 200$ ns from the EventT0 + 1.5 cluster hits), we rank the photon candidates by momentum and select the three most energetic photons in each of the barrel, forward endcap, and backward endcaps, excluding the ISR photon candidate. The sum of the energy of these non-ISR energetic photons is defined to be $E_{extra} = \sum_{region} (\sum_{i \neq \gamma_{ISR}} E_{\gamma_i, region}^{lab})$. In principle, if there were any additional tracks, we would rank them by momentum and compute the energy they contribute to the event. However, because we require only two tracks, there is no additional energy contribution from the tracking detectors. The distribution of E_{extra} is illustrated in Fig. 4.16, where we showcase a range of signal-relevant and background contributions.

Given that similar beam background conditions are used for both signal and background simulation, the excess in E_{extra} in background must arise from soft photons coming from various background processes. While a cut on $E_{extra} < 0.7$ GeV would be sufficient, Fig. 4.16 also shows that some signal models exhibit mass-dependent shape differences that could be further exploited. To capitalize on the potential discriminating power, we incorporate E_{extra} into our MVA-based event selection.

3.2.3 Background Modeling

The main sources of background from the generators mentioned in Sec. 3.2.1 that would pass signal event pre-selections are:

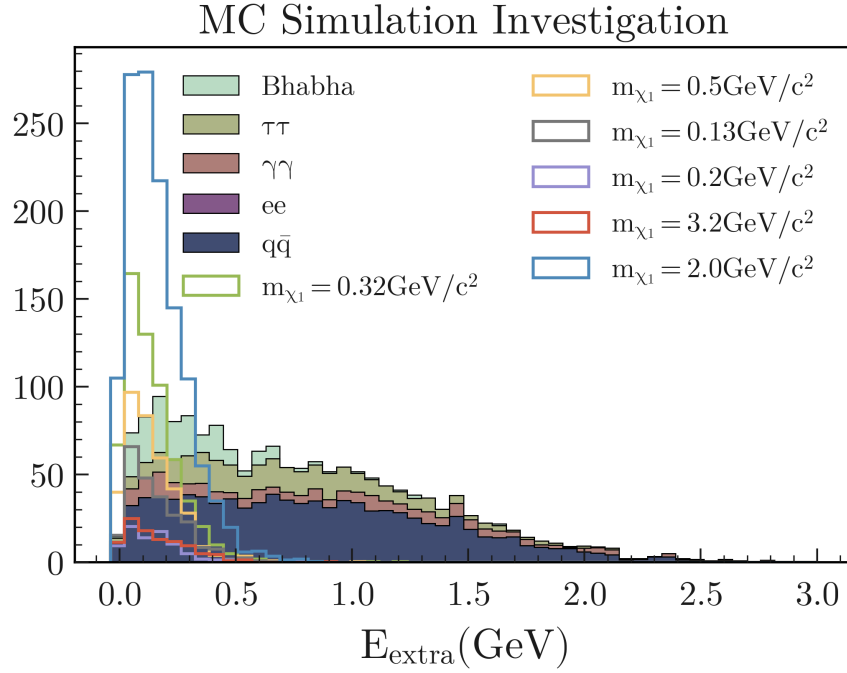


Figure 3.19: Total extra event energy in an event for signal and background MC simulation.

- $e^+e^- \rightarrow K_s^0 K_L^0 \gamma_{ISR}$ where $K_s^0 \rightarrow \pi^+ \pi^-$
- $e^+e^- \rightarrow \gamma\gamma$ where $\gamma \rightarrow e^+e^-$
- $e^+e^- \rightarrow q\bar{q}$ where $q \in \{u, d, s, c\}$ with $K_s^0 \rightarrow \pi^+ \pi^-$
- $e^+e^- \rightarrow \ell^+ \ell^-$ where $\ell \in \{e, \mu, \tau\}$

From recoil mass distribution, we notice there are two peaks in the low recoil mass region, in which $\gamma\gamma$, $K_L^0 K_s^0(\gamma_{ISR})$, and $q\bar{q}$ contribute.

To describe this, we use two separate Crystal Balls (CBs) distributions [189, 190] with one for $q\bar{q}$ and another for $\gamma\gamma/K_L^0 K_s^0(\gamma_{ISR})$ in the lower mass region. Although the two CBs can be described by a single global function, some cuts (e.g., the converted-photon mass) affect certain background components more than others, so we elect to adopt a more flexible scheme. The parameters for the two crystal balls are:

- $(\gamma\gamma)$ and $K_s^0 K_L^0 \gamma_{ISR}$'s CB: $\mu = 1.58_{-0.48}^{+0.50}$, $\sigma = 1.60_{-0.28}^{+0.31}$, $\alpha = -0.36_{-0.18}^{+0.13}$, $n = 0.97_{-0.19}^{+0.28}$
- $q\bar{q}$'s CB: $\mu = 9.00_{-0.46}^{+0.49}$, $\sigma = 3.96_{-0.36}^{+0.42}$, $\alpha = -0.63_{-0.17}^{+0.15}$, $n = 0.79_{-0.19}^{+0.28}$.

For the high recoil-mass region, which has contributions from multiple backgrounds (especially $e^+e^- \rightarrow \ell^+\ell^-$ for $\ell \in [e^\pm, \mu^\pm, \tau^\pm]$), $e^+e^- \rightarrow q\bar{q}$, $e^+e^- \rightarrow \gamma\gamma$), the shape can modeled with the following function:

$$y_{expo}(x|\lambda, y_0, L, H) = A \left(\lambda \exp(\lambda x) / (e^{\lambda H} - 1) - y_0 \times \Theta(L < x < H) \right) \quad (3.11)$$

with a rate of $\lambda = 0.09 \pm 0.01$, a Heavyside step function with normalization $y_0 = 0.055 \pm 0.02$. Because of some signal-dependent cuts on the missing energy in an event, the non-zero range of this component ranges from lower bound $L = 13.50 \pm 0.13 \text{ (GeV}/c^2\text{)}^2$ and upper bound $H = 90 \text{ (GeV}/c^2\text{)}^2$ based on the broadest recoil range.

Finally, we allow the normalizations of the individual components to float during signal extraction, as they vary with the iDM model used.

3.2.4 Event Selection

In addition to the cuts outlined in the previous section, we use a forest of (gradient-)boosted decision trees (BDTs) to better separate background from signal (see Sec. 3.1 for details). In particular, we use **XGBoost** [191] with the following input features:

- Total extra event energy (E_{extra}): Sum over extra ECL cluster energy
- Opening angle: $-\ln(1 - \cos(\angle_{\text{OA}}))$ where $\cos(\angle_{\text{OA}}) = \hat{p}_{t-} \cdot \hat{p}_{t+}$
- Pointing angle: $-\ln(1 - \cos(\angle_{\text{PA}}))$ where $\cos(\angle_{\text{PA}}) = \hat{p}_{V_0} \cdot \hat{dr}_{V_0}$

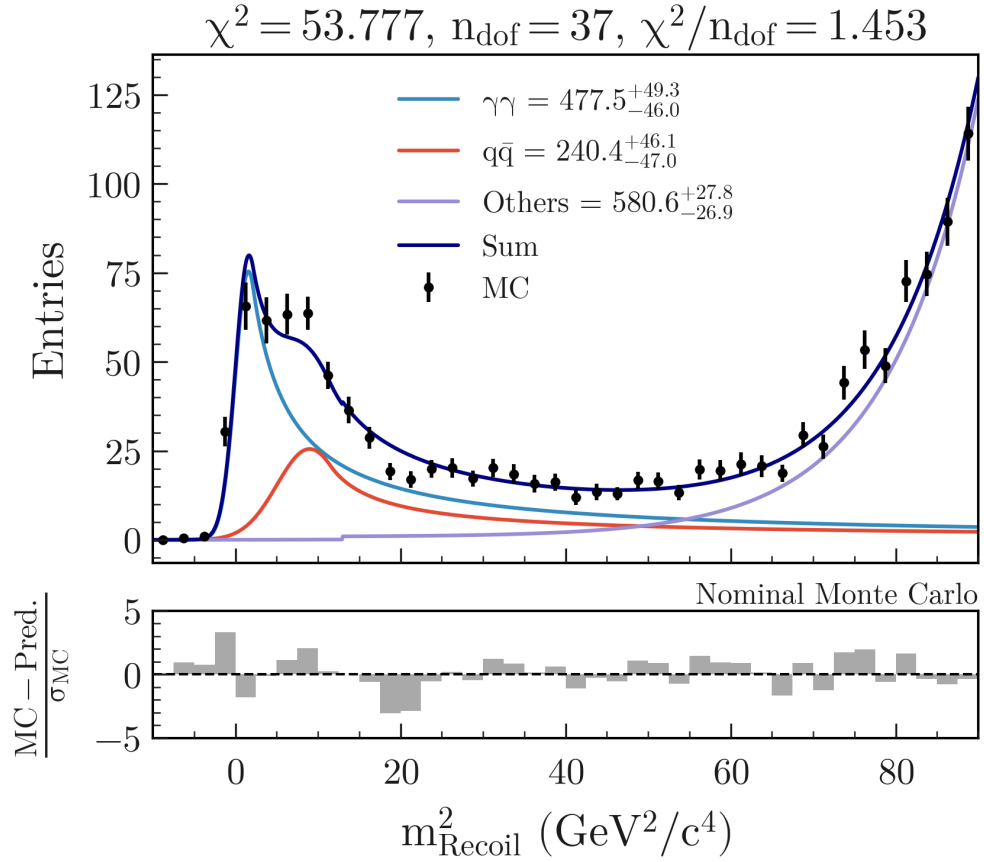


Figure 3.20: For all the background scaled to $L = 364\text{fb}^{-1}$, we fit this distribution with two crystal balls and an exponential tail. The background MC shown is the events that passed the criteria discussed in Sec. 3.2.2.

- Back-to-backness: $\cos \angle_{V_0, \gamma} = \hat{p}_{ISR} \cdot \hat{dr}_{V_0}$
- Electron ID of each track (eID_{t_-} and eID_{t_+})
- Muon ID of each track (μID_{t_-} and μID_{t_+})

Because some of the listed features' separation power depends on the available phase space of each iDM model, we define three different mass windows in which we use an ensemble of BDTs to classify. Based on the number of events in both signal and background, we define the light/medium/heavy m_{χ_1} intervals to be $[0, 0.32]$ $[0.32, 1.3]$ $[1.3, 3.2]$ GeV/ c^2 .

In addition, the kinematics depend on the decay mode of the displaced vertex, so we further partition our dataset by the electron and muon IDs of the tracks. To minimize scenarios where one track is leptonic and the other hadronic, we say a vertex has sufficient lepton ID if both tracks have $eID > 10^{-2}$ and $\mu ID > 10^{-2}$. For vertices that pass this criterion, we then look at the maximum eID and μID between the two tracks. If $\max([eID_{t_i}]) > \max([\mu ID_{t_i}])$, we then apply the dedicated electron BDT for background suppression. We have the analogous scenario to use the muon BDT if $\max([eID_{t_i}]) \leq \max([\mu ID_{t_i}])$. For events that do not satisfy the sufficient lepton ID criteria, we apply the hadronic-like BDT to better separate low-momentum muons and pions. Altogether, we have nine categories of BDTs used in tandem to suppress backgrounds according to the lepton ID values of the two tracks.

Finally, the models' training data assume all the cuts outlined in the previous sections, except the K_s^0 mass cut. The omission of the K_s^0 mass cut in our training data is to ensure that the BDT can learn on a continuum, so we can use the K_s^0 sample to validate the BDTs' performance.

This is highlighted in Fig. 3.21, which shows F-score = $\frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$ of each model's features and how their importance varies across mass windows. To reinforce that the BDTs' responses

are similar at the transition points (i.e., $m_{\chi_1} = 0.32 \text{ GeV}/c^2$ and $1.3 \text{ GeV}/c^2$), we include these points in both samples (e.g., m_{χ_1} models appear in both light and medium mass models).

The size of our training dataset is $f_t \times 2L$, where f_t is some sub-sampling fraction and L is the minimum dataset size between run-dependent background and signal simulations. We sample $f_t \times L$ events from both signal and background (adding up to $2L$) to give an equal representation of signal and background. In addition, we consider the luminosity of each background sample (e.g., $q\bar{q}$ containing $4\mathcal{L}_{\text{run } 1}$ implies a weight of $1/4$) during the BDT's training.

We use a subsampling fraction of $f_t = 0.9$ and set aside the testing dataset of size $0.1 \times 2L$ to independently evaluate the performance of our models.

For hyper-parameter tuning, we do a Grid Search with 3-Fold cross-validation [192]. The grid of hyperparameters used for the search is:

- Depth of a given tree: 3-6
- Learning Rate: 0.1, 0.2, 0.3
- Sub-sampling fraction for a given tree's training: 0.5, 0.6, 0.7, 0.8
- L1 regularization term on the BDT's weights: 0, 1, 5
- Minimum sum of instance weight (hessian) needed in creating a leaf: 0.5, 1, 2
- Evaluation Metric: RMS, ROC AUC

Once we determine the optimal hyperparameters for the BDT, we use this configuration to train new models. To obtain a measure of uncertainty, we bootstrap [193] five models with different sub-samplings of our run-dependent simulations. To introduce some variation

among these models, we subsample the training dataset with a fraction $f_v = 0.8$ (so the bootstrapped training dataset has $f_v f_t \times 2L$ events).

To evaluate the performance of the final BDTs, we inspect accuracy and ROC curves on the entire simulation dataset (which is particularly important for the signal, allowing us to assess whether the BDTs preferentially select some models over others).

After optimizing each mass-range ensemble based on the Punzi FoM, which is shown in Fig. 3.22, the number of signal events retained post-reconstruction is illustrated in Fig. 3.23, showing that, at the transition points, the signal retention efficiency is within $\sim 1\sigma$ from each other. The BDT cuts used for the light/medium/heavy models are 0.96401/0.97907/0.97987.

3.2.5 Signal Extraction

Having selected a description of our backgrounds and performed our event selection, we now outline how we consider the specific properties of each iDM model (i.e., values of $\{m_{A'}, \epsilon, \alpha_D, m_{\chi_1}, \Delta m_\chi\}$). and extract them from a given dataset.

Signal Dependent Cuts	Variable Description
$m_{V_0} \leq \max(0.1 \text{ GeV}, \Delta m_\chi)$	Invariant mass of the displaced vertex, bounded by $\Delta\chi$
$E_{\text{miss}} \geq 2m_{\chi_1}$	Missing Energy ($\sqrt{s} - E_{\gamma,cm} - E_{V_0,cm}$), bounded by $2m_{\chi_1}$

Table 3.6: Model dependent cuts

The main way we use the properties of each iDM model is by making model-dependent cuts to reject more backgrounds while retaining as much signal as possible. The relevant quantities, which are also found in Table 3.6, are the missing energy of the system (E_{miss}) and the displaced vertex mass (m_{V_0}).

For $E_{\text{miss}} = \sqrt{s} - E_{\gamma,cm} - E_{V_0,cm}$, we notice that this is also equal to the energy carried away by the two χ_1 's in a signal event: E_{miss} is bounded below by $2m_{\chi_1}c^2$. From our signal MC

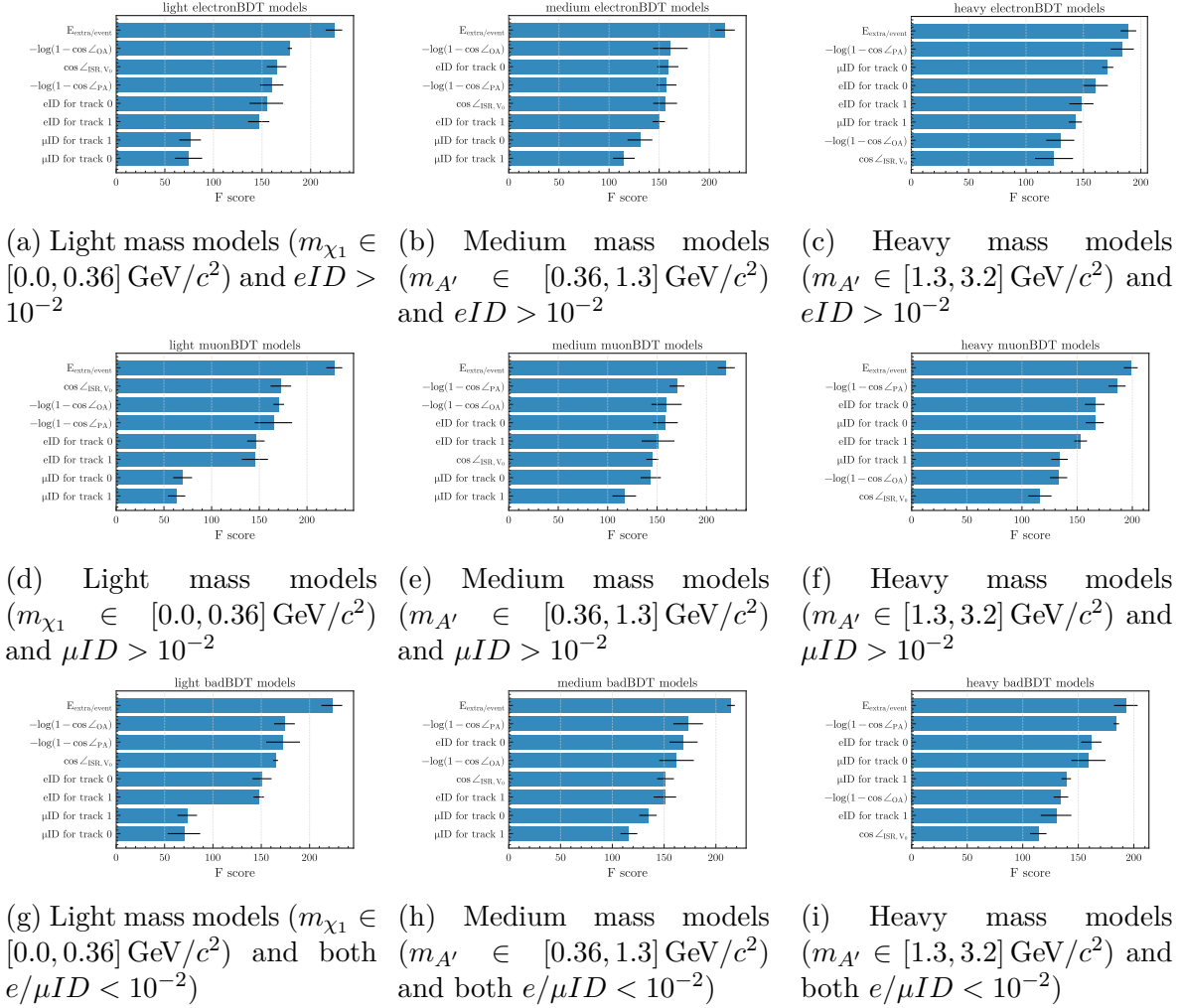


Figure 3.21: F-Score of the ensemble of BDTs used for background suppression. The results shown are the average over five trials for each category, with error bars indicating the standard deviation.

dataset, we find that this cut is more useful for models with heavier m_{χ_1} since the momentum of the system provides a more stringent maximum lower bound (infimum) for lighter models.

For the displaced vertex, the available energy to produce it comes from the mass splitting Δm_χ , providing an upper bound for the displaced vertex mass (m_{V_0}). Due to detector resolution effects, we have some number of reconstructed signal events with $m_{V_0} > \Delta m_\chi$ (which is especially an issue for $\Delta m_\chi < 0.1 \text{ GeV}$). To circumvent this, we define a supremum (min-

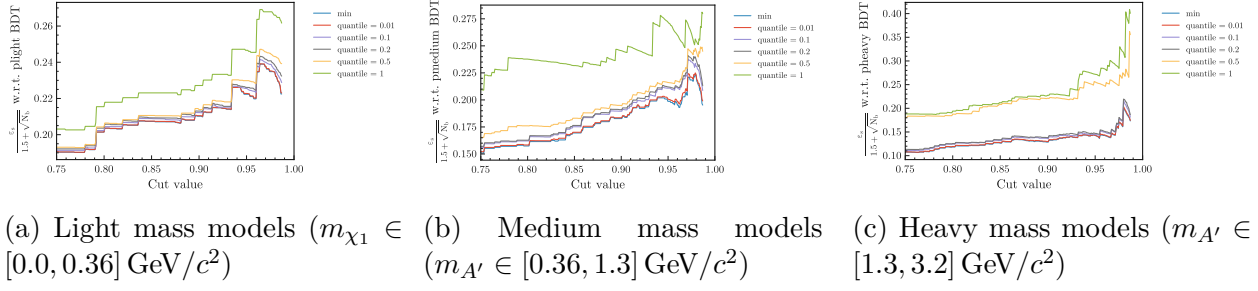


Figure 3.22: Punzi Figure-of-merit distributions for the BDT output distribution. Since each mass window contains several iDM models, we present the distribution of the n'th quartile to assess how a cut will affect the majority of models.

imal upper bound) threshold value, which sets a minimum m_{V_0} mass cut used during signal extraction. To determine this value, we examine the cut efficiency ($\frac{\text{events post vertex mass cut}}{\text{events post analysis cuts}}$), as illustrated in Fig. 3.24. For models from $m_{\chi_1} \in [0.08, 0.2] \text{ GeV}$, $m_{V_0 \rightarrow e^+ e^-} < 0.08 \text{ GeV}$ would only cut events outside the inter-quartile range of said models. For $m_{\chi_1} = 0.05 \text{ GeV}$, the limited statistics of the lighter model yield the broader distribution shown. To be conservative, the supremum of the m_{V_0} cut is reported in Table 3.6 and is determined based on preserving 75% of our lightest iDM model while still suppressing most of our backgrounds in this regime.

Signal Description

We now outline how we extract signal events for our collection of models, where an iDM model denotes a particular set of model parameters.

For a given model MC sample, we note that the signal's m_{recoil}^2 distribution can be fitted using a double-sided Crystal Ball (DSCB) [189, 190, 194]; this is sensible as the recoil mass is with respect to the ISR photon. Some examples are illustrated in Fig. 3.25.

However, since we're doing a counting experiment, and the fact that some (heavy) models have part of their distribution truncated from a minimal γ_{ISR} energy requirement of

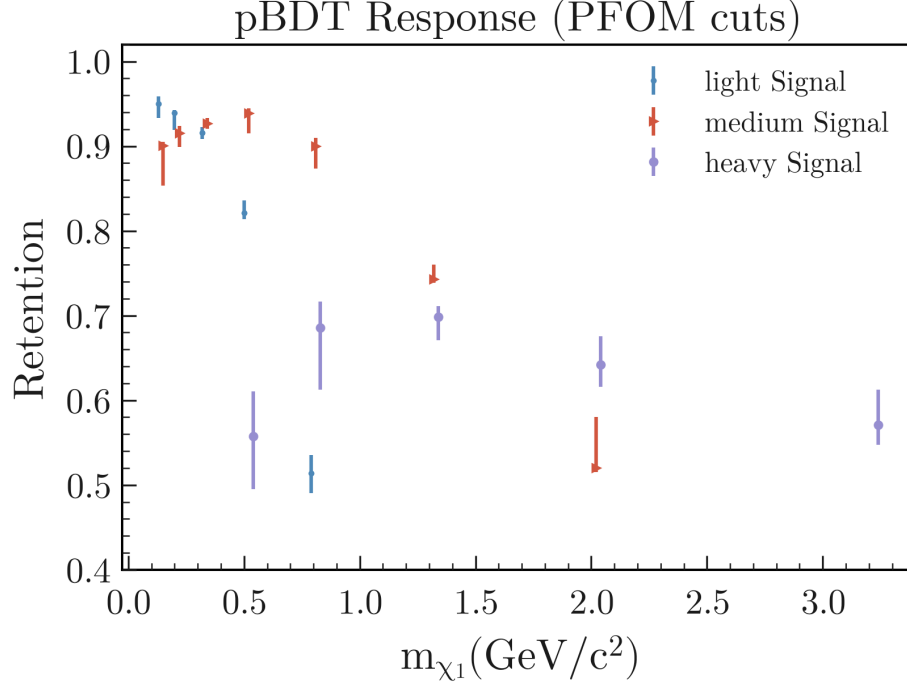


Figure 3.23: Over the entire m_{χ_1} mass range, retention by the BDT ensembles (trained on certain ranges) of signal events post-reconstruction.

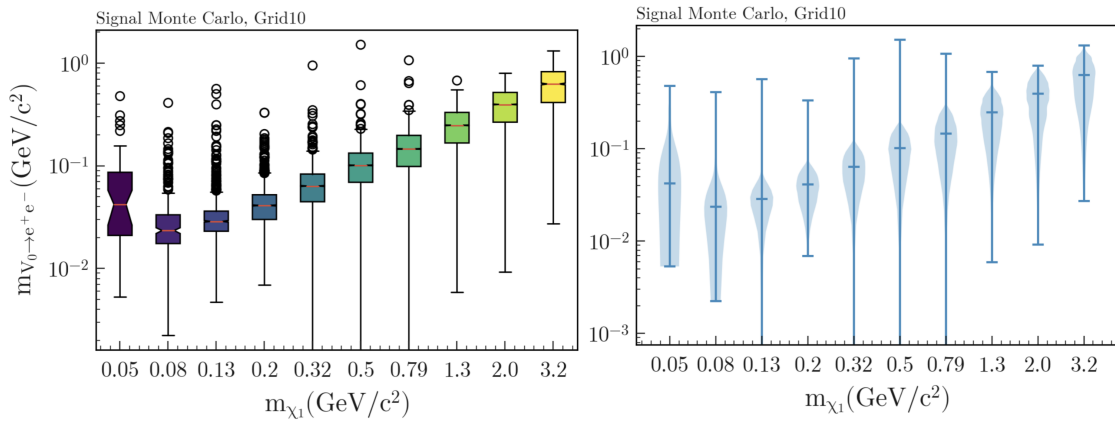


Figure 3.24: Using the signal configuration labeled ‘Grid10’ outlined in Table 1.1, we show the range of vertex mass $m_{V_0 \rightarrow e^+ e^-}$ for each model of a given m_{χ_1} . The left uses box-whiskers with a box length of 1.5 times the inter-quartile range, whereas the right shows the same data using a violin plot.

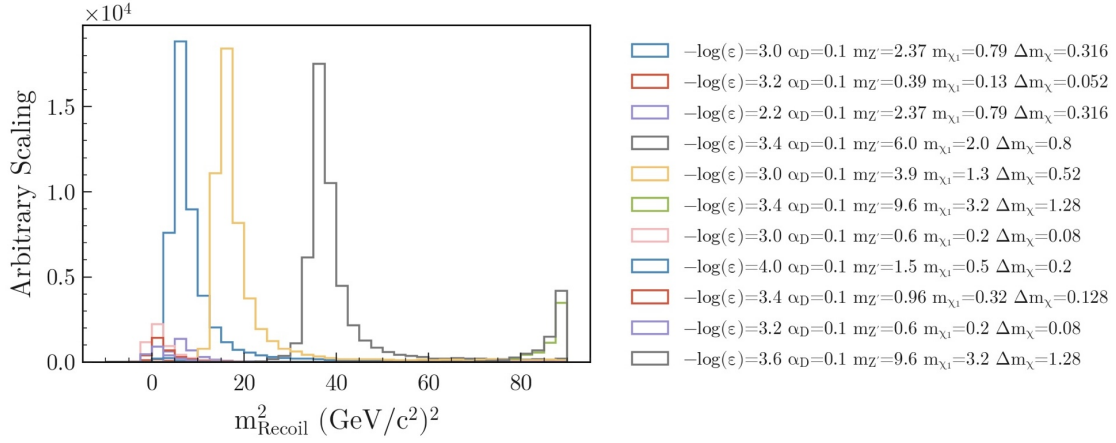


Figure 3.25: Examples of reconstructed signal models' m_{recoil}^2 distribution.

$\min(E_{ISR}) = 1 \text{ GeV}$ (setting an upper-bound for the recoil mass distribution), we categorize each signal shape by defining a confidence level via a lower and higher edge. As the ISR photon is recoiling against the A' , the median of m_{recoil}^2 corresponds to $m_{A'}^2$. Thus, in defining our signal window, we center it at the median of the signal distributions. In defining the lower and upper edges, we set the confidence level to 90% and require equal areas on either side of the median. After determining the lower/upper edges of the signal window for our collection of signal models, we extrapolate these values using a cubic, as illustrated in Fig. 3.26. While one might expect these trends to be parabolic, edge effects shift the boundaries for heavier A' .

3.2.6 Application of Event Selection

With the strategy defined above, we show the recoil mass distribution as we apply the cuts outlined in Table 3.2, 3.3, 3.4, 3.5, then the BDT cuts, then the other signal dependent cuts.

In Fig. 3.28, we show the signal efficiency throughout the various stages of our pipeline. The left plot illustrates the efficiency of reconstructing our signal events, and the preselection outlined in Tables 3.2, 3.3, 3.4, and 3.5. The middle plot shows the efficiency of said

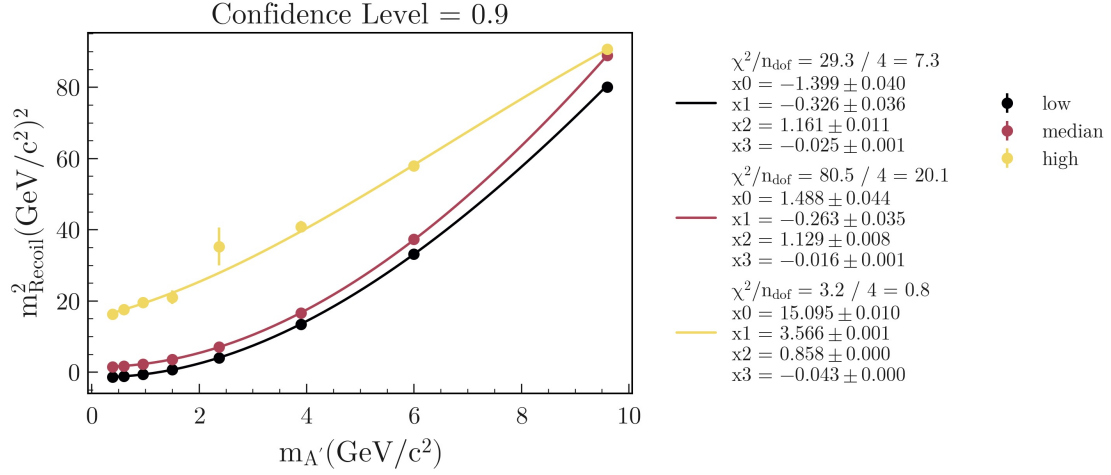


Figure 3.26: Signal Window Interpolation at a confidence level of 90%.

reconstructed events passing the BDT cut, signal window, and other signal-dependent cuts ($m_{V_0 \rightarrow e^+e^-}$, E_{miss} , $\Delta R_{\gamma \rightarrow e^+e^-}$). The right plot shows the final reconstruction efficiency that we use for the exclusion regions of our parameter space.

In addition to the signal efficiencies, we also highlight the background distributions. Figs. 3.27 and 3.29 show the appearance of the background distribution through different stages of our analysis pipeline, similar to Fig. 3.28. Fig. 3.27 shows the selection applied after reconstruction, which is used to fit the background PDFs. The top row of Fig. 3.29 illustrates the impact of the PFOM optimized BDT scheme. The bottom row illustrates the impact of the PFOM optimized BDT scheme and the $\Delta R_{\gamma \rightarrow e^+e^-}$ cuts. As mentioned in Sec. 3.1, our BDT scheme reduces our background levels to $\mathcal{O}(1)$.

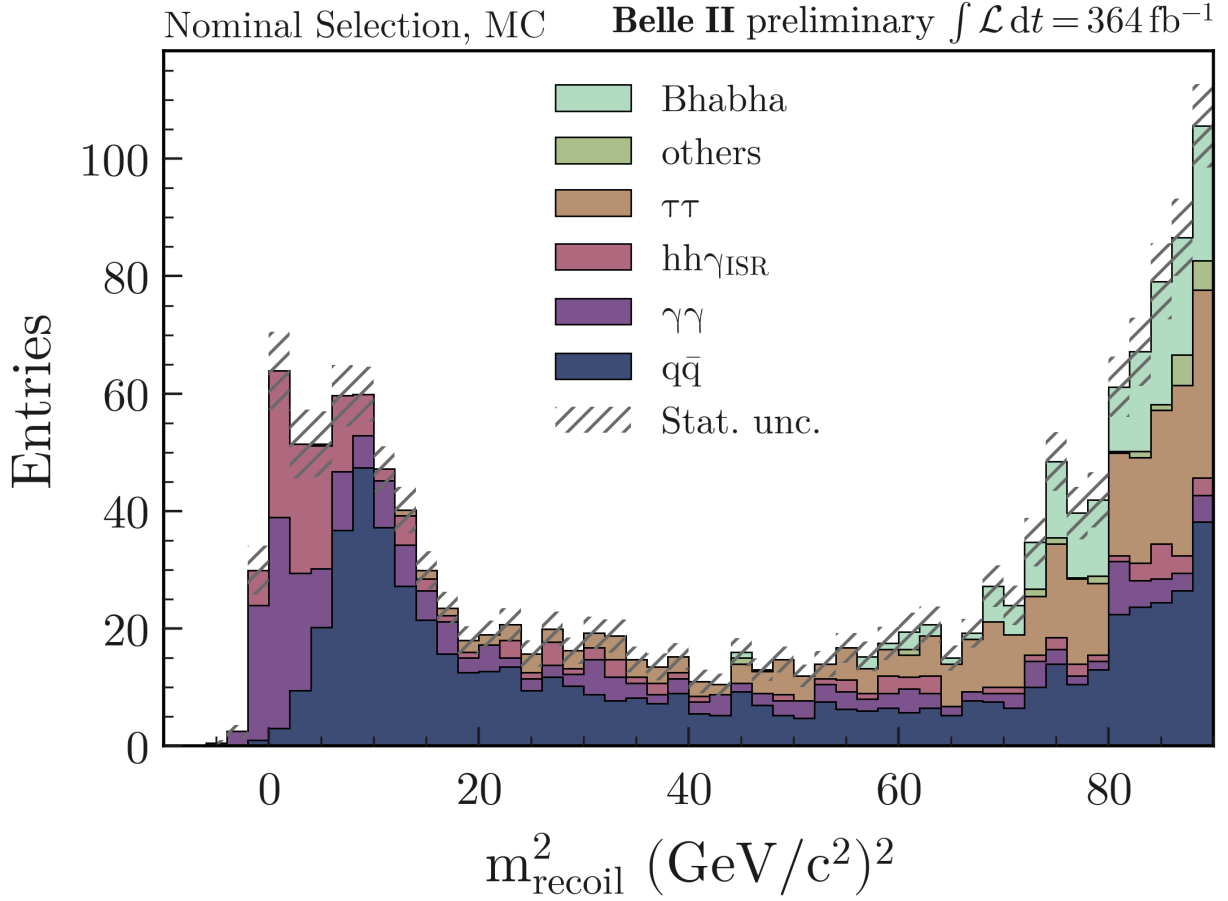


Figure 3.27: Illustrating the impact of the pre-selection on the background MC simulation.

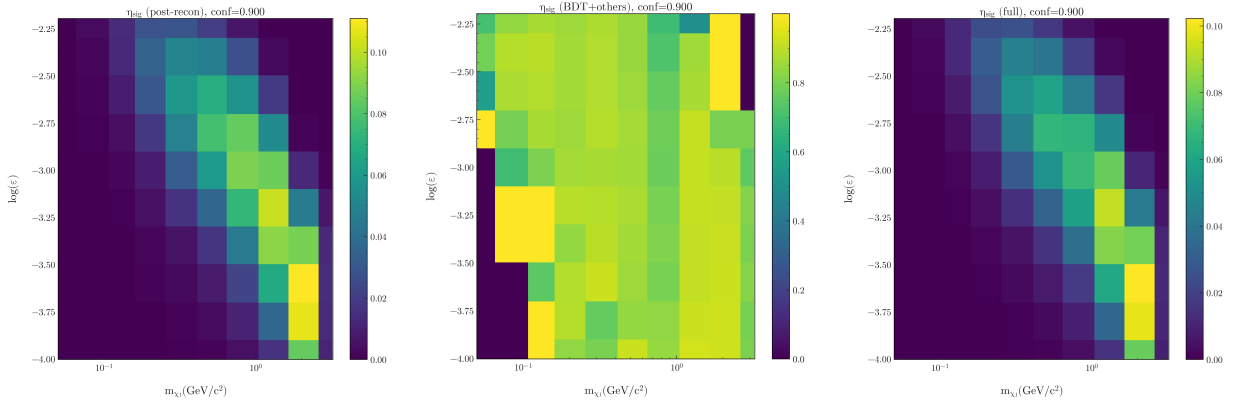


Figure 3.28: For the baseline model, we show the signal efficiencies at the various stages of the analysis. Left is post-reconstruction with preselection. Middle is of the reconstructed events, and how many signal events are retained by the BDT and other signal cuts. Right is the full reconstruction efficiency.

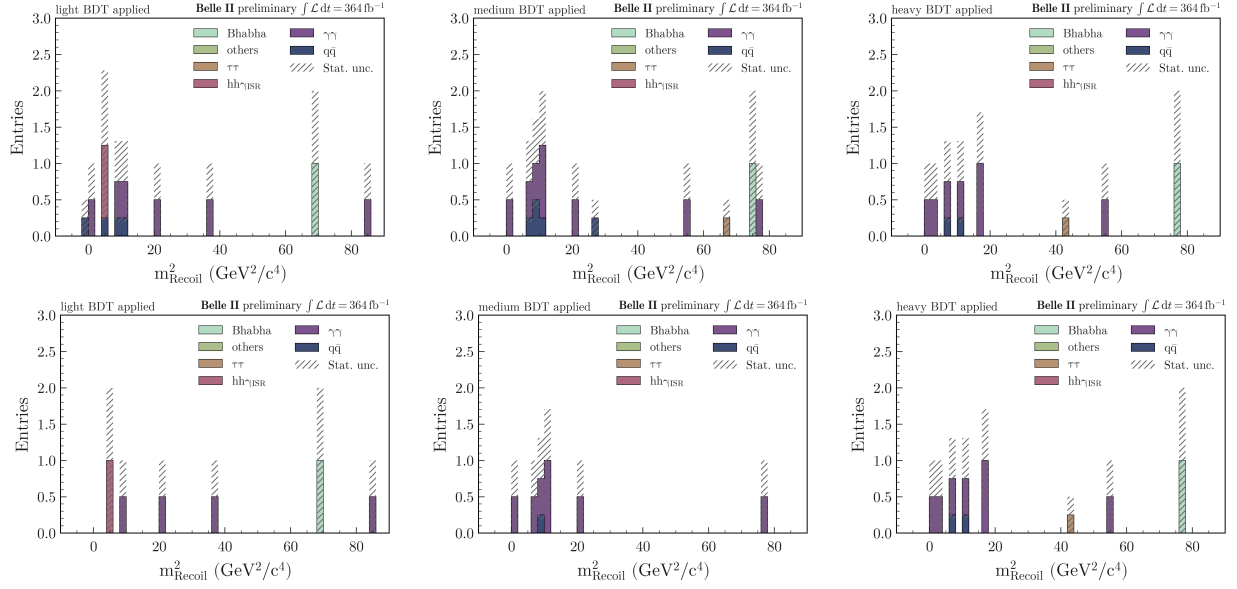


Figure 3.29: Impact of the BDT selection in the top row and the BDT selection and $\Delta R_{\gamma \rightarrow e^+e^-}$ selection on the bottom row.

3.3 Systematics

3.3.1 Trigger Efficiencies

In this subsection, we describe the influence of the L1 trigger and the High Level Trigger (HLT) on the signal efficiency in MC (using TSIM). However, since we found that all L1 triggers are selected by the HLT trigger, we emphasize here the L1 trigger performance.

From the L1 trigger, the most important track line is the high energy (HIE) line, which triggers on events containing ECL energy depositions above 1 GeV ($\sum_j E_j \geq 1 \text{ GeV}$). Since we also consider models in which $m_{A'} \sim 9.5 \text{ GeV}$, implying $\gamma_{ISR} \sim 1 \text{ GeV}$, we expect there to be a slight efficiency drop due to resolution effects. To account for this, we consider the following three trigger lines that help supplement the HIE trigger in this region.

The first one is the single-track trigger (STR), which fires when at least one track with $p \geq 0.7 \text{ GeV}/c$ and $|z| < 15 \text{ cm}$ is present in the event. The next one is the 0th low-multiplicity

line (**LML0**), which triggers when there are three ECL clusters within the range $\theta_{ID} \in [1, 17]$ (or $12.4^\circ < \theta^{lab} < 154.7^\circ$) with at least one cluster containing $E_{cl}^{lab} \geq 0.3 \text{ GeV}$ and not an ECL Bhabha event. The final one is the 12th low-multiplicity line (**LML12**), which triggers when there are three ECL clusters with at least one cluster containing $E_{cl}^{lab} \geq 0.5 \text{ GeV}$, and $\theta_{ID}^{ECL} \in [2, 16]$ (approximately $18.5^\circ < \theta^{lab} < 139.3^\circ$). These two work synergistically to help trigger on signal-like events, as heavier models yield lower-energy photons with more energetic tracks, and vice versa for lighter models. One complication is the presence (or absence of) a given trigger and of trigger prescaling. The **STT** line was introduced at the start of exp. 16. For **LML0**, the prescale factor was changed from 1 to 10 in exp. 20 run 916, so its role is diminished for later runs. To address this, we consider only the pre-scaled and masked (PSNM) trigger information, which should manifest in a reduced efficiency relative to trigger lines with prescales of 1, such as **HIE** and **LML12**. The information on the trigger lines is summarized in Table 3.7.

Trigger	Definition	Prescales	Fraction of events with trigger
HIE	$(\sum_{ECL} E > 1 \text{ GeV}) \ \& \ !(Bhabha)$	1	100%
STT	at least one 3D track w/ $p > 0.7 \text{ GeV}/c \ \& \ z < 15$	$0 \rightarrow 1$ in Exp. 16 Run 666	74.9%
LML0	(3+ ECL clusters in $\theta_{ID} \in [1, 17]$) & (one cluster w/ $E_{cl}^{lab} > 0.3 \text{ GeV}$) & !(Bhabha)	$1 \rightarrow 10$ in Exp. 20 Run 916	53.7%
LML12	(3+ ECL clusters in $\theta_{ID} \in [1, 17]$) & (one cluster w/ $E_{cl}^{lab} > 0.5 \text{ GeV}$) & !(Bhabha)	1	100%

Table 3.7: Trigger lines of interest

Fig. 3.30 shows the dependence of m_{χ_1} on the efficiencies of the two triggers, relative to the **L1** trigger. The **HIE** distribution clearly shows that we have a high trigger efficiency ($> 99\%$) for all but the heaviest iDM models, for which $E_{\gamma_{ISR}} \sim 1 \text{ GeV}$ that results in a minor efficiency drop to 97%. Conversely, for the **STT** and the two **LML** lines, the trigger

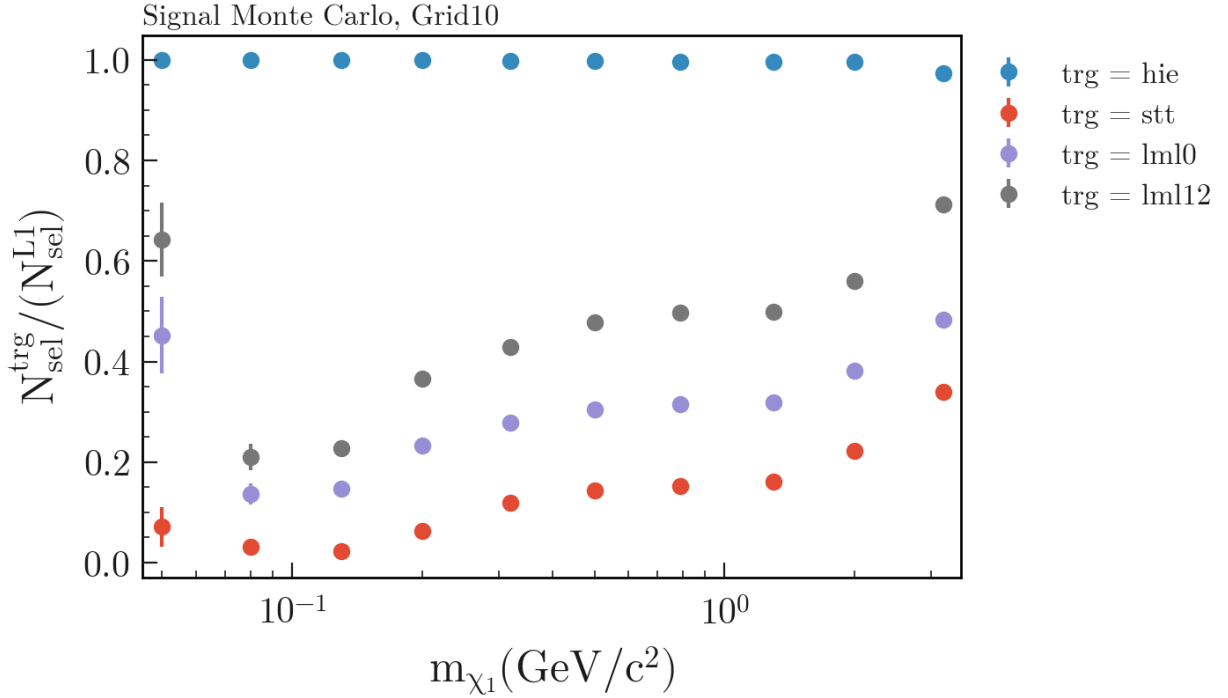


Figure 3.30: Over the entire signal MC, the trigger efficiencies with respect to the L1 Trigger are illustrated as a function of m_{χ_1} , as the kinematics of the ISR photon and tracks are strongly dependent on m_{χ_1} .

efficiency is quite low for $m_{\chi_1} \lesssim 0.1$ GeV because most of the available energy in the event is carried by the ISR photon. However, heavier models have more kinematic phase space for more energetic tracks.

Fig. 3.31 shows the dependence of $-\log(\epsilon)$ on the efficiencies of the two triggers, relative to the L1 trigger. For the **HIE** distribution, trigger efficiencies remain relatively high ($> 99\%$), confirming that m_{χ_1} is a more prevalent factor in the event kinematics. For the **STT** distribution, a single prevalent effect appears to cause the peaking-like structure observed. One subdominant effect is that, as $-\log \epsilon$ increases, the lifetime also increases, displacing vertices enough to violate the $|z|$ requirement. The dominant effect affecting lower-mass models is a correlation between the average track momenta and $-\log \epsilon$, as depicted in Fig. 3.14, which rejects many low-momentum tracks for $-\log \epsilon \in [2, 3]$. The interplay between

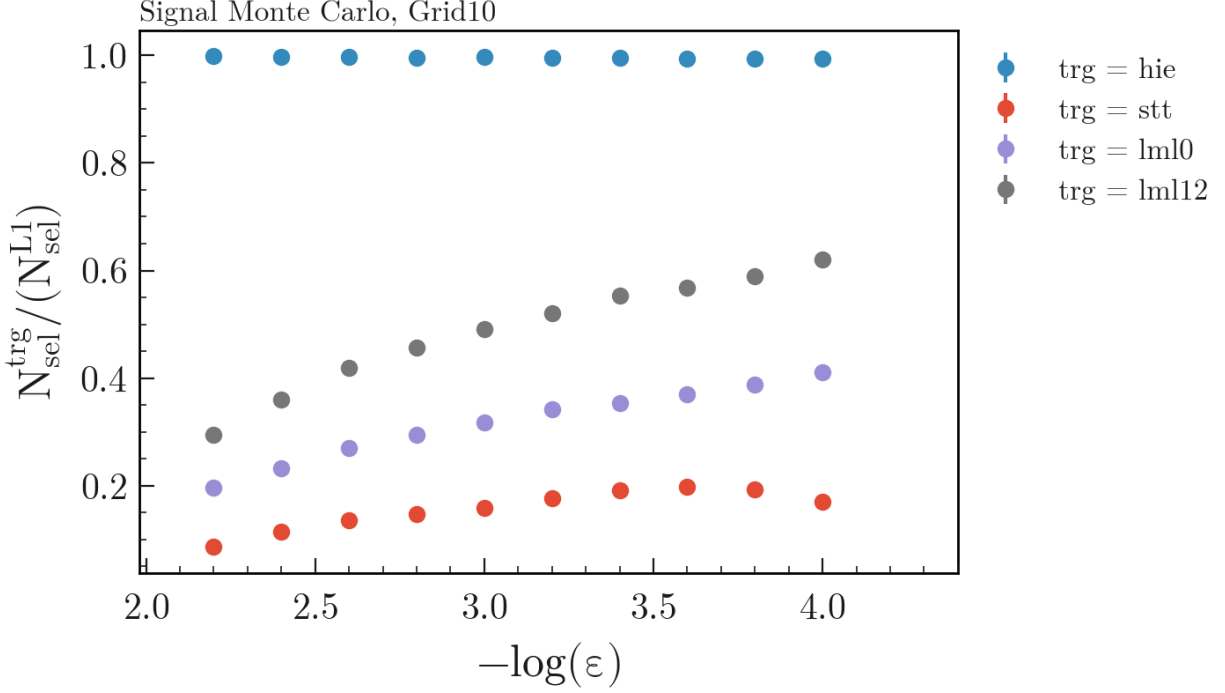


Figure 3.31: Over the entire signal MC, the trigger efficiencies with respect to the L1 Trigger are illustrated as a function of $-\log \epsilon$, to highlight the importance of the χ_2 lifetime.

these two effects would produce the observed peaking-like structure. For the other two trigger lines, the increase with respect to $-\log \epsilon$ seems to be consistent with $-\log \epsilon$'s correlation with the track momenta.

3.3.2 BDT Systematics

To measure the uncertainty in the BDT used for background suppression, we train five BDTs per category, using the average BDT response at the time of event selection. To quantify the systematic uncertainty of the BDT response, we look at the difference between each BDT's response and the averaged response across our entire training/validation datasets. When doing so, we see that the standard deviation of the BDT score was 0.03-0.04. The distributions and values of said deviation are illustrated in Fig. 3.32. Because the BDT

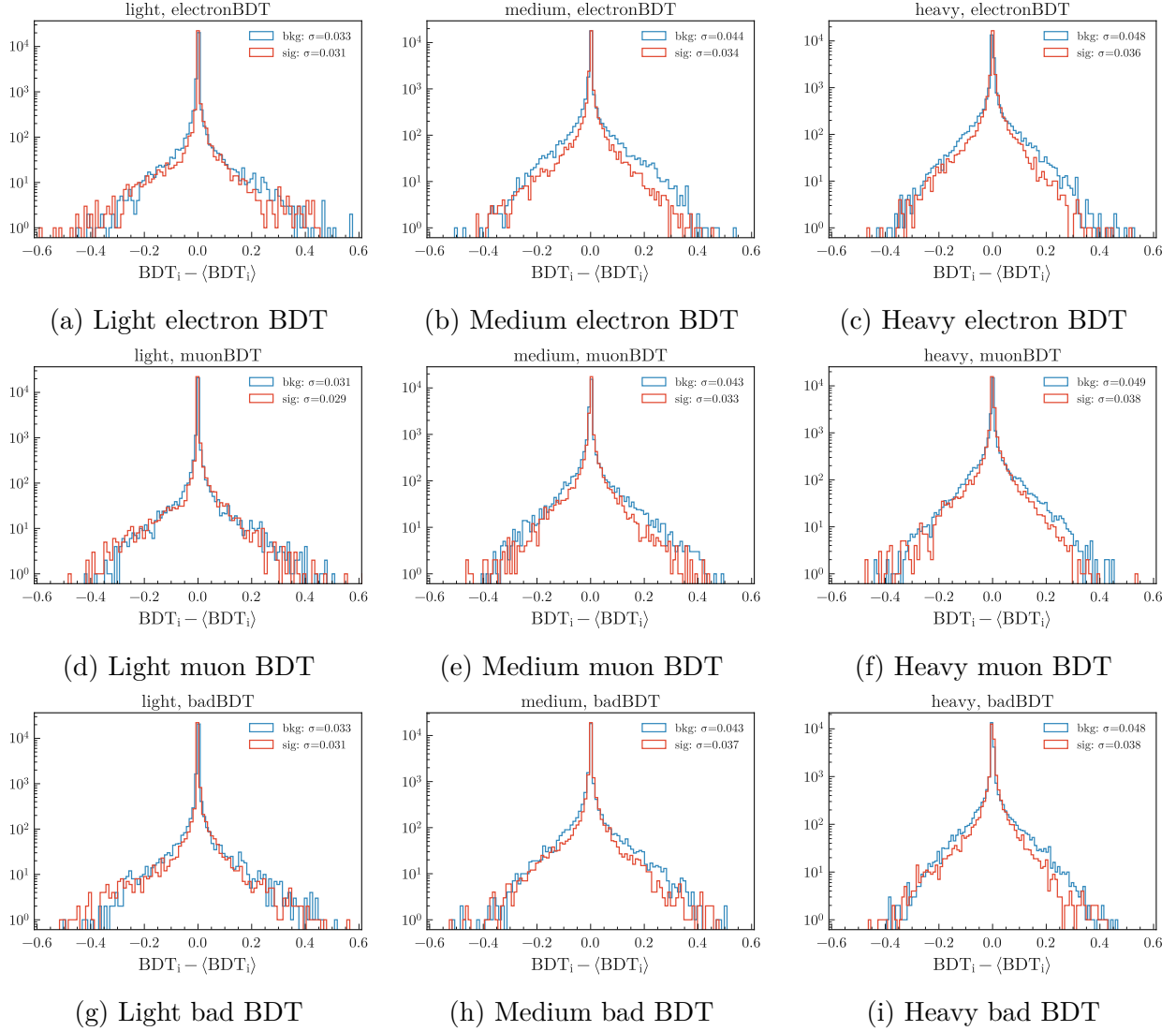


Figure 3.32: Showcasing for each BDT category how the response is on a random selection of signal and background events. The x-axis shows each BDT's response on the i 'th event minus the averaged response.

variations are all negligible, we conclude that their union is also negligible.

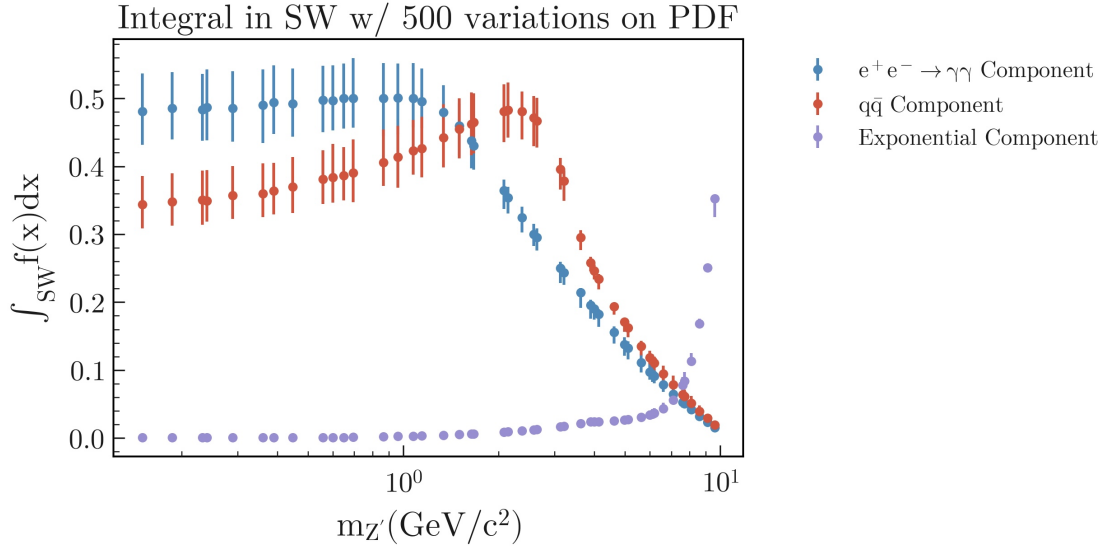


Figure 3.33: For various signal windows for a corresponding model with some $m_{A'}$, we integrate $\tilde{p}(x|\tilde{\theta})$ within said signal window, and show the result from 300 trials.

3.3.3 Background Shape Parameters

In Sec. 3.2.3, the best fit parameters $\theta_{i,j}$ of probability density p_i of background component i (Crystal Ball parameters for $(\gamma\gamma + K_s^0 K_L^0 \gamma_{ISR})$, another Crystal Ball for continuum, and an exponential function for the rest) are provided with uncertainties. Here, we show how we capture and encode these shape systematics.

We begin by conducting the following toy study. We generate $\tilde{p}_i$ with randomly sampled parameters $\tilde{\theta}_{i,j}$. As many of the uncertainties for parameter $\theta_{i,j}$ (for the i 'th component with j 'th parameter) are asymmetric in nature, the distribution we sample from is a split-normal distribution with σ_- and σ_+ of said normal distribution coming from the lower and upper bound of a given uncertainty. Using $\tilde{p}_i$, we integrate over a given signal window. This is repeated 500 times for a range of $m_{A'}$ (and its corresponding signal window), with Fig. 3.33 as the output of said toy study.

While the integral in region R , or $I_R = \int_R p(x) dx$, is easy to calculate in situ, saving a large

table of values from this fine scan, or re-performing the above toy study seems suboptimal. Instead, we encode these systematics by fitting an n 'th order polynomial as a function of $m_{A'}$. We use `IMINUIT` [173] to determine the parameters of these polynomials $\vec{\rho} = (\rho_0, \dots, \rho_n)$ for polynomial $\sum_i^n x^i \rho_i$ are listed below:

- $\gamma\gamma$, lower bound: $\vec{\rho} = [-3.935 \cdot 10^{-2}, -4.697 \cdot 10^{-2}, 6.583 \cdot 10^{-2}, -3.083 \cdot 10^{-2}, 7.127 \cdot 10^{-3}, -8.772 \cdot 10^{-4}, 5.518 \cdot 10^{-5}, -1.398 \cdot 10^{-6}]$
- $\gamma\gamma$, upper bound: $\vec{\rho} = [5.076 \cdot 10^{-2}, 2.771 \cdot 10^{-2}, -3.855 \cdot 10^{-2}, 2.381 \cdot 10^{-2}, -5.093 \cdot 10^{-3}, 5.779 \cdot 10^{-4}, -3.373 \cdot 10^{-5}, 7.990 \cdot 10^{-7}]$
- $q\bar{q}$, lower bound: $\vec{\rho} = [-3.430 \cdot 10^{-2}, 7.368 \cdot 10^{-3}, -3.302 \cdot 10^{-2}, 2.471 \cdot 10^{-2}, -7.337 \cdot 10^{-3}, 1.063 \cdot 10^{-3}, -7.528 \cdot 10^{-5}, 2.088 \cdot 10^{-6}]$
- $q\bar{q}$, upper bound: $\vec{\rho} = [4.133 \cdot 10^{-2}, 2.258 \cdot 10^{-3}, 1.618 \cdot 10^{-2}, -1.749 \cdot 10^{-2}, 5.832 \cdot 10^{-3}, -8.874 \cdot 10^{-4}, 6.418 \cdot 10^{-5}, -1.795 \cdot 10^{-6}]$
- others, lower bound: $\vec{\rho} = [-1.837 \cdot 10^{-4}, 1.414 \cdot 10^{-3}, -4.428 \cdot 10^{-3}, 5.767 \cdot 10^{-3}, -4.066 \cdot 10^{-3}, 1.711 \cdot 10^{-3}, -4.486 \cdot 10^{-4}, 7.378 \cdot 10^{-5}, -7.395 \cdot 10^{-6}, 4.126 \cdot 10^{-7}, -9.818 \cdot 10^{-9}]$
- others, upper bound: $\vec{\rho} = [5.501 \cdot 10^{-4}, 1.582 \cdot 10^{-4}, -1.046 \cdot 10^{-3}, 2.110 \cdot 10^{-3}, -1.711 \cdot 10^{-3}, 7.270 \cdot 10^{-4}, -1.731 \cdot 10^{-4}, 2.320 \cdot 10^{-5}, -1.628 \cdot 10^{-6}, 4.637 \cdot 10^{-8}]$

The comparison of the length of the error bars in Fig. 3.33 and the result of this encoding is illustrated in Fig. 3.34, showing that it captures the general trends and values (up to $\lesssim \pm 0.01$). The variations shown in Fig. 3.34 are a result of the number of toy studies done.

One quantity that depends on this integral is the scaling factor $\tau_{sw} = I_{SW}/I_{SB} \equiv I_{SW}/(1 - I_{SW})$, which is necessary when scaling background yields expected from the sidebands to the signal window. When we wish to incorporate the shape systematic in this area calculation,

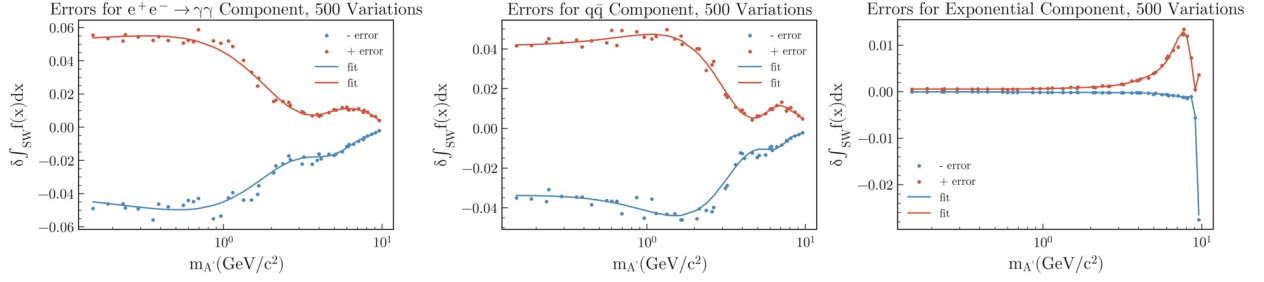


Figure 3.34: For the various background components, we show the length of the error bars on Fig. 3.33 vs. $m_{A'}$, and their interpolations.

then $\delta\tau_{SW} = \delta I_{SW} |I_{SW}(1 - I_{SW})|^{-1}$. If we consider uniform background components, there is no uncertainty in the shape of the background as there are no fitting parameters (sans normalization), so $\delta\tau_{SW} = 0$ in this scenario.

3.3.4 Validating the Upper Limit Scheme

Lastly, we aim to validate our upper-limit scheme. In particular, we want to compare the use of a uniform background PDF vs the nominal background shapes illustrated in Fig. 3.20.

To do this, we need toys that are representative of the dataset we expect to see. From applying our BDT ensemble on background MC, we observe there to be ~ 2.5 $\gamma\gamma$ -like events, 0.5 $q\bar{q}$ events, and ~ 2 Bhabha/ $\tau\tau$ -like events that remain. This is shown in Fig. 3.29, where we take the medium mass model expectation as the worst case scenario. From Sec. 4, we expect an excess in Bhabha/ $\tau\tau$ -like events, so we also consider the scenario where we inflate their contribution by a factor of 3. In addition, we aren't sure of the overall normalization, so we inflate our baseline and our Bhabha/ $\tau\tau$ -enhanced dataset by a factor of three as well.

Given one of these background expectations, we construct a toy dataset by drawing the number of each background component from MC and sampling from its corresponding PDF using a Poisson distribution. Given this toy, we apply our signal-extraction method outlined

in Secs. 3.1 and 3.1.

For an initial trial, we look at one toy dataset with the high Bhabha/ $\tau\tau$ -enhanced baseline further inflated by a factor of three (i.e., $(\gamma\gamma, q\bar{q}, \tau\tau) = (P_{\infty}(\mu = 7.5), P_{\infty}(\mu = 1.5), P_{\infty}(\mu = 18))$) in depth to illustrate the steps in our signal extraction process, before averaging over all the toy datasets and background expectations. As a caveat, we assume when fitting our backgrounds that the yield of events contributing to the $q\bar{q}$ PDF is zero. As the $\gamma\gamma$ and $q\bar{q}$ peaks are close together, and the event yield is low, we obtain reasonable fits while also lowering strongly correlated uncertainties when fitting the overlapping Crystal Ball PDFs.

Given the dataset shown in Fig. 3.35, we first inspect the different signal-mass windows. For brevity, we explicitly show the results of fitting $m_{\chi_1} = 0.05, 0.32, 6.0$, and $9.6 \text{ GeV}/c^2$ using both schemes. We observe that, in the heaviest models, the uniform background model cannot accommodate the exponential excess, and conclude that there must be a few background events. In the low- and high-recoil-mass regions, the background-shape scheme consistently predicts the correct number of background events within ± 1 . However, due to the large uncertainties in the yield and the additional contribution to the shape systematics, the upper limit is comparable to that from the uniform PDF methodology. This is also reflected in Fig. 3.36; the exclusion regions are comparable.

As we move from this trial to the ensemble of models, we repeat this procedure for all our toys and all our background scenarios. For each background expectation, we compute the average and standard deviation for the two schemes, then compute the difference between the averages, divided by the uncertainty. The result is illustrated in Fig. 3.37, showing that, on average, the upper limits computed by both schemes are consistent with one another, with the largest systematic bias at $m_{\chi_1} = 3.2 \text{ GeV}/c^2$.

Another way to confirm this bias is to use a toy dataset: for a given toy dataset, we com-

pute the ratio of the two upper limits, and bin the number of signal events and m_{χ_1} . The results of this are shown in Fig. 3.38, where they are consistent with one another, with the exception of $m_{\chi_1} = 3.2 \text{ GeV}/c^2$, where the uniform PDF scheme consistently overestimates the background abundance for scenarios where $n_{sw} \gg 1$.

Lastly, we show the exclusion contour plots for each background scenario in Fig. 3.39, where the most notable deviation is in the Bhabha/ $\tau\tau$ -enhanced scenario – this aligns with the previously mentioned findings.

To summarize the results of the study, we generate 200 toys across different background scenarios and find that, when we apply our signal extraction pipeline, the two schemes are largely consistent, except for the lowest and highest of m_{χ_1} models. Because of the backgrounds' contributions ($\tau\tau$, $q\bar{q}$, Bhabhas) to the exponential rise at high recoil mass and the potential peaking background near low recoil mass, we prefer the scheme that employs the background PDF shapes.

3.3.5 Linearity Test

In addition to validating the upper-limit scheme on simulated samples devoid of iDM signal events, we should verify that bona fide signal events can be extracted. To do this, we take the previously mentioned toy datasets and randomly sample $N_{sig,in}$ number of signal events from signal MC. We then apply our signal-extraction technique using the background shape scheme and obtain N_{sig}^{PyHF} . We repeat this for two hundred random samplings (with random toys for background) for different mass points and $N_{sig,in} \in \{0, 1, 2, \dots, 12, 13\}$.

As our signal extraction scheme uses a 90% confidence interval for the signal window, the expected slope of a linearity test would be near 0.9.

For the sake of brevity, Fig. 3.40 only illustrates the results from the exponential-enhanced

background levels with three times the number of background events ($\text{expo} \times 3$). The results of all the background scenarios we consider are summarized in Tables 3.8, 3.9, 3.10, and 3.11.

For the most part, the slopes of both the background shape scheme and the uniform scheme range from ~ 0.7 to ~ 0.9 , which is potentially an artifact of our signal window interpolation (see Fig. 3.26). For the intercept for the uniform shape scheme, we consistently see that the biases are within a few events. In the scenarios where we've tripled the number of expected background events, the regions where we expect peaking background events (from $e^+e^- \rightarrow \gamma\gamma$, $e^+e^- \rightarrow q\bar{q}$, and the exponential background), the uniform background scheme begins to overestimate the number of extracted signal events. For the shape scheme, the biases consistently overestimate the number of background events (hence a negative signal yield); this is expected, as the background shape scheme was engineered to expect more background events. Despite the large biases, the upper limits produced in Sec. 3.3.4 are comparable to those from the uniform scheme, suggesting that the large uncertainties in the yields and additional systematics compensate for the overestimation of the number of background events. One additional effect from injected signal events that fall outside of the signal window further inflates the background expectation. With these considerations in mind, we proceed with our current schemes and conclude that our signal extraction scheme can reasonably extract signal events, especially in scenarios with low background levels.

Table 3.8: Linearity results for baseline background scenario.

	Background Shape Scheme		Uniform Shape Scheme	
$m_{A'}$ (GeV/ c^2)	Slope	Intercept	Slope	Intercept
0.15	0.834 ± 0.008	-1.06 ± 0.07	0.924 ± 0.004	0.789 ± 0.034
0.24	0.836 ± 0.009	-1.18 ± 0.07	0.923 ± 0.005	0.761 ± 0.035
0.39	0.772 ± 0.010	-1.14 ± 0.07	0.892 ± 0.005	0.807 ± 0.036
0.60	0.766 ± 0.010	-1.28 ± 0.08	0.890 ± 0.005	0.778 ± 0.037
0.96	0.771 ± 0.010	-1.24 ± 0.08	0.877 ± 0.005	0.824 ± 0.038
1.50	0.772 ± 0.010	-0.89 ± 0.07	0.879 ± 0.005	0.788 ± 0.037
2.37	0.781 ± 0.008	-0.42 ± 0.06	0.812 ± 0.006	0.196 ± 0.041
3.90	0.872 ± 0.005	-0.01 ± 0.04	0.868 ± 0.005	-0.431 ± 0.036
6.00	0.908 ± 0.005	0.06 ± 0.04	0.889 ± 0.005	-0.772 ± 0.035
9.60	0.743 ± 0.010	-1.29 ± 0.08	0.801 ± 0.005	0.268 ± 0.042

Table 3.9: Linearity results for baseline background $\times 3$ scenario.

	Background Shape Scheme		Uniform Shape Scheme	
$m_{A'}$ (GeV/ c^2)	Slope	Intercept	Slope	Intercept
0.15	0.816 ± 0.011	-3.93 ± 0.09	0.907 ± 0.006	2.69 ± 0.05
0.24	0.812 ± 0.012	-4.05 ± 0.09	0.902 ± 0.006	2.70 ± 0.05
0.39	0.750 ± 0.013	-4.19 ± 0.09	0.876 ± 0.006	2.71 ± 0.05
0.60	0.739 ± 0.013	-4.33 ± 0.10	0.872 ± 0.007	2.77 ± 0.05
0.96	0.737 ± 0.013	-4.23 ± 0.10	0.855 ± 0.007	2.89 ± 0.05
1.50	0.748 ± 0.013	-3.24 ± 0.10	0.853 ± 0.007	2.83 ± 0.05
2.37	0.758 ± 0.010	-0.69 ± 0.07	0.760 ± 0.007	1.72 ± 0.05
3.90	0.875 ± 0.007	0.29 ± 0.05	0.801 ± 0.006	-0.08 ± 0.04
6.00	0.903 ± 0.006	0.65 ± 0.05	0.783 ± 0.007	-0.72 ± 0.05
9.60	0.698 ± 0.012	-2.63 ± 0.09	0.781 ± 0.006	1.11 ± 0.04

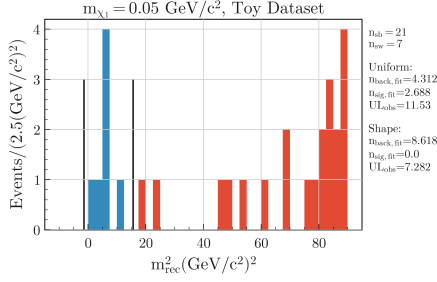
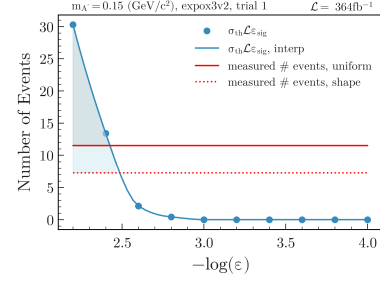
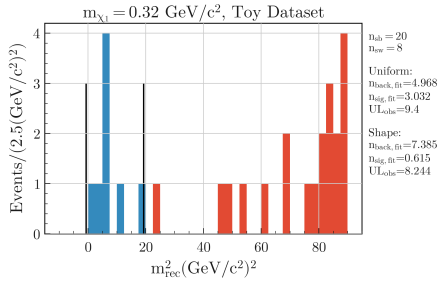
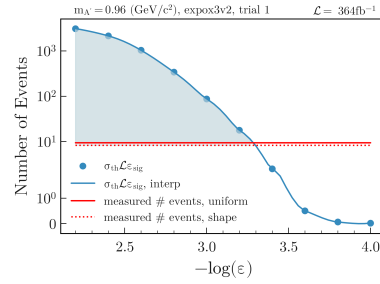
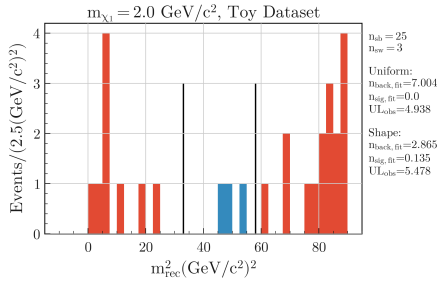
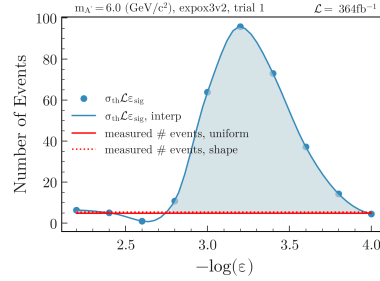
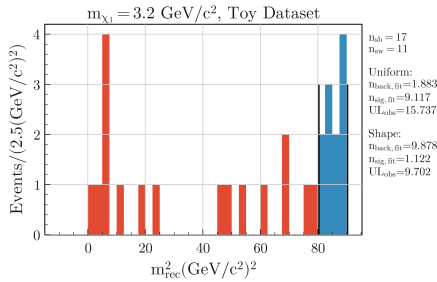
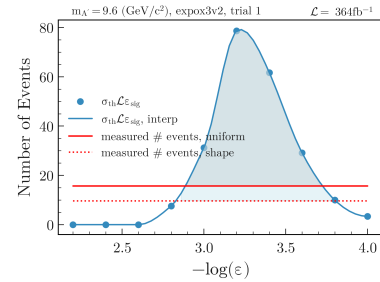
(a) $m_{\chi_1} = 0.05 \text{ GeV}/c^2$, observable(b) $m_{\chi_1} = 0.05 \text{ GeV}/c^2$, comparing with σ_{th} (c) $m_{\chi_1} = 0.32 \text{ GeV}/c^2$, observable(d) $m_{\chi_1} = 0.32 \text{ GeV}/c^2$, comparing with σ_{th} (e) $m_{\chi_1} = 2.0 \text{ GeV}/c^2$, observable(f) $m_{\chi_1} = 2.0 \text{ GeV}/c^2$, comparing with σ_{th} (g) $m_{\chi_1} = 3.2 \text{ GeV}/c^2$, observable(h) $m_{\chi_1} = 3.2 \text{ GeV}/c^2$, comparing with σ_{th}

Figure 3.35: For a subset of models, we show our signal extraction scheme using the uniform PDF scheme and signal shape scheme on a toy dataset devoid of signal. The left shows the toy dataset with a given signal window, with the fit results written on the side. The right shows the upper limits from the left distribution and compares it to $N_{th} = \sigma \mathcal{L}_{run} \eta$.

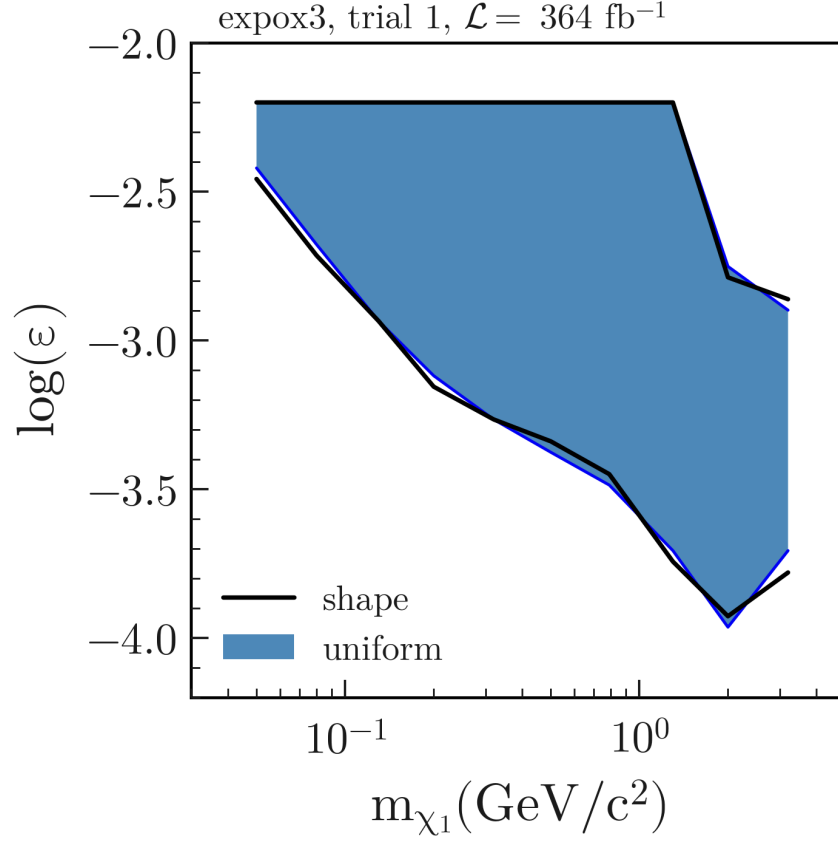


Figure 3.36: From the toy shown in Fig. 3.35, we scan through our models and show the corresponding exclusion plot for the uniform (blue) and shape schemes (black outline).

Table 3.10: Linearity results for exponential background scenario.

	Background Shape Scheme		Uniform Shape Scheme	
$m_{A'} (\text{GeV}/c^2)$	Slope	Intercept	Slope	Intercept
0.15	0.823 ± 0.011	-3.53 ± 0.09	0.914 ± 0.004	-0.04 ± 0.03
0.24	0.821 ± 0.012	-3.64 ± 0.09	0.915 ± 0.005	-0.06 ± 0.04
0.39	0.776 ± 0.013	-3.83 ± 0.10	0.892 ± 0.005	-0.06 ± 0.04
0.60	0.726 ± 0.013	-3.83 ± 0.10	0.870 ± 0.005	0.01 ± 0.04
0.96	0.749 ± 0.013	-4.07 ± 0.10	0.868 ± 0.005	-0.10 ± 0.04
1.50	0.751 ± 0.012	-3.33 ± 0.09	0.858 ± 0.005	-0.10 ± 0.04
2.37	0.767 ± 0.009	-1.09 ± 0.07	0.784 ± 0.006	-0.49 ± 0.04
3.90	0.877 ± 0.006	0.30 ± 0.05	0.838 ± 0.005	-0.74 ± 0.04
6.00	0.912 ± 0.006	0.58 ± 0.05	0.855 ± 0.006	-0.69 ± 0.04
9.60	0.704 ± 0.011	-2.36 ± 0.08	0.777 ± 0.005	1.67 ± 0.04

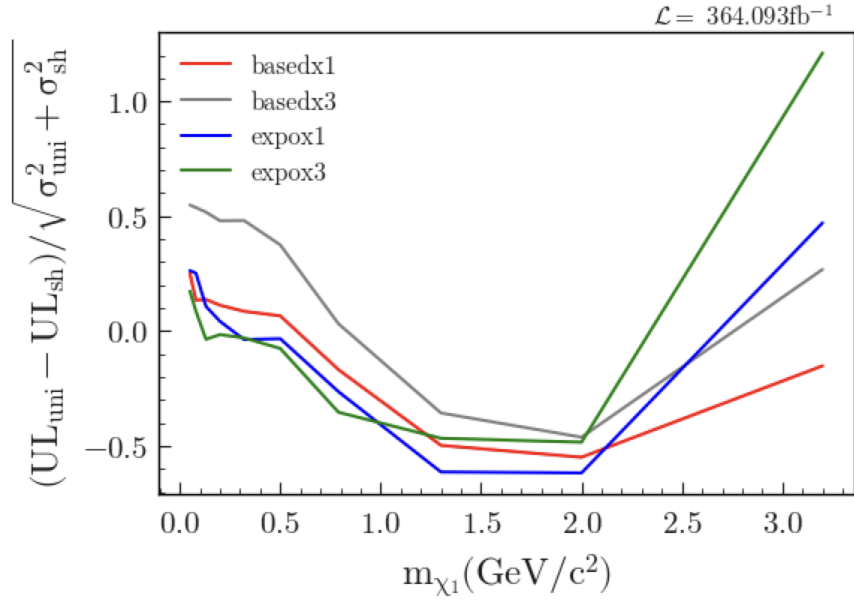


Figure 3.37: For the average upper limit determined by our two schemes for each background expectation, we outline the Z-score vs. m_{χ_1} .

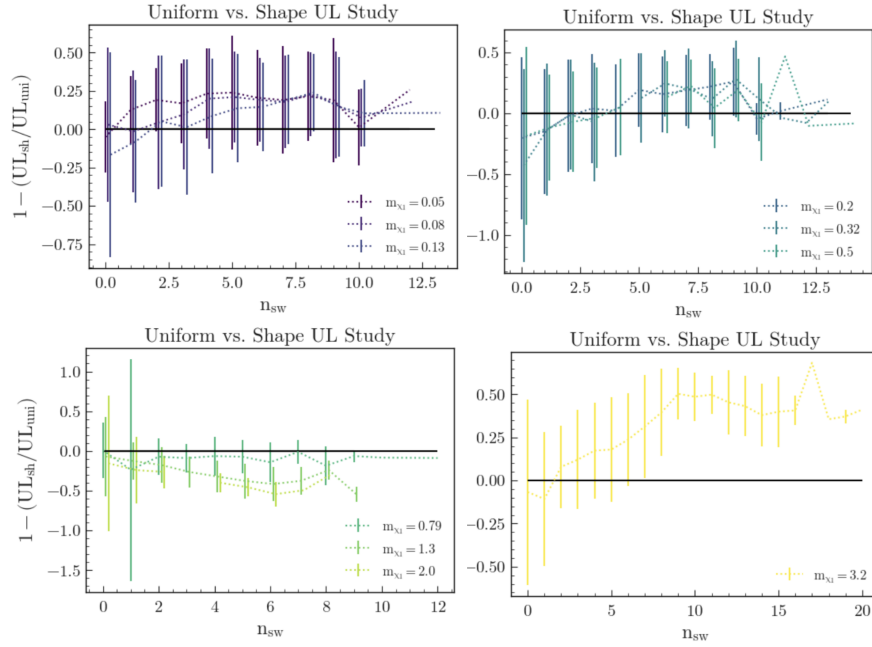


Figure 3.38: For each toy dataset, we compare the upper limit set by our uniform vs. shape and plot it as a function of the number of events in the signal window. For clarity, we split the distributions apart by mass.

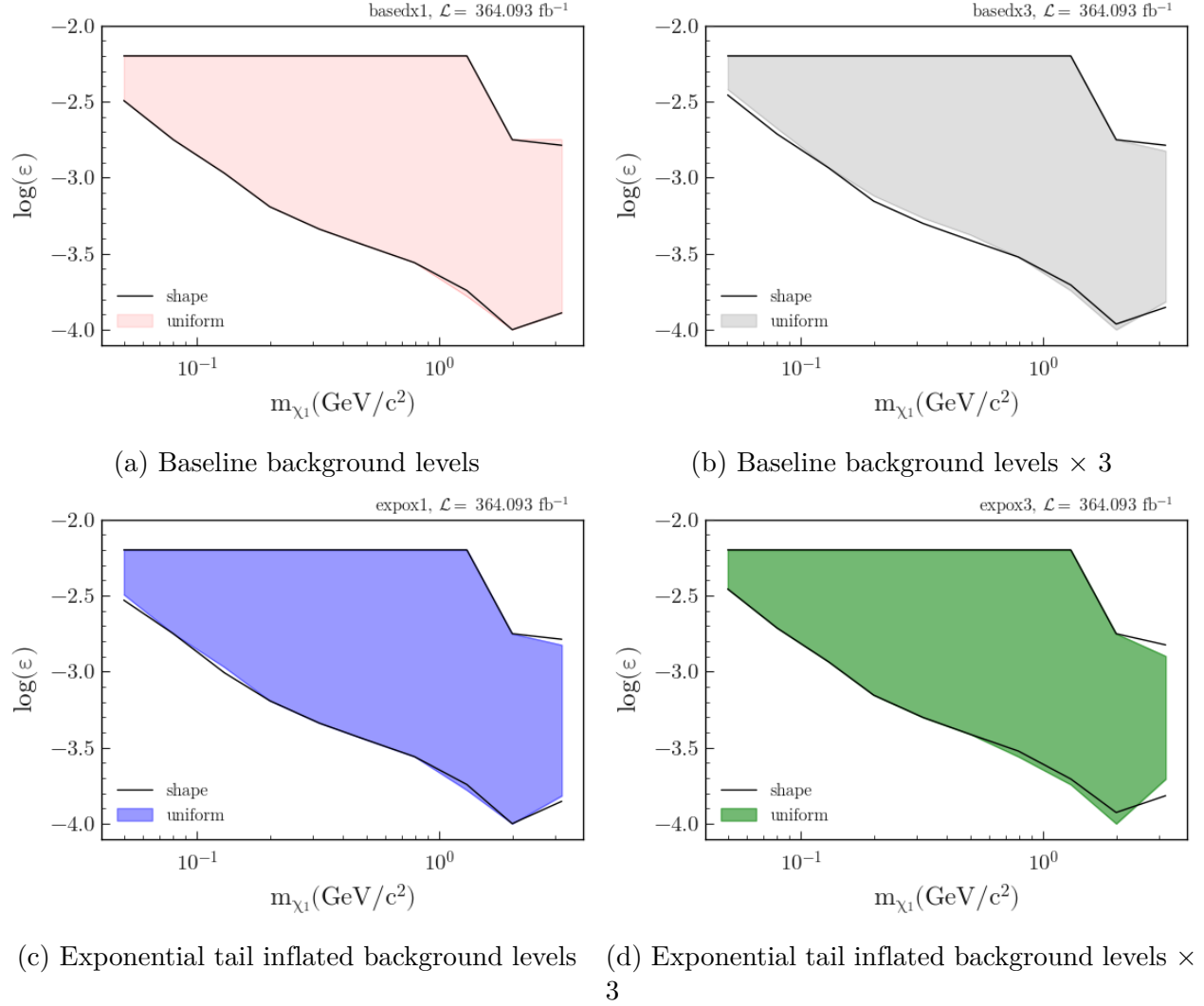
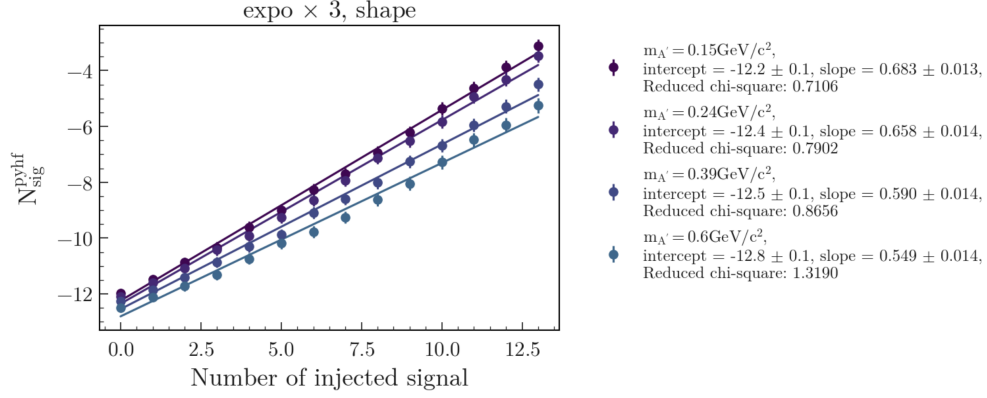
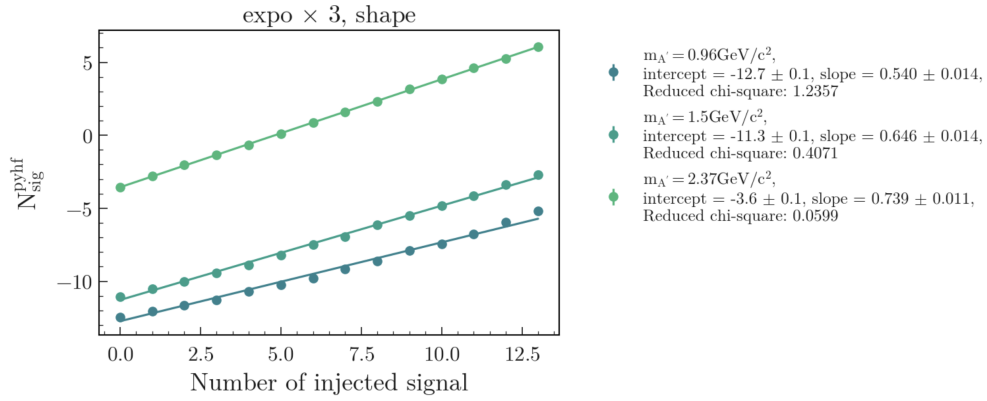


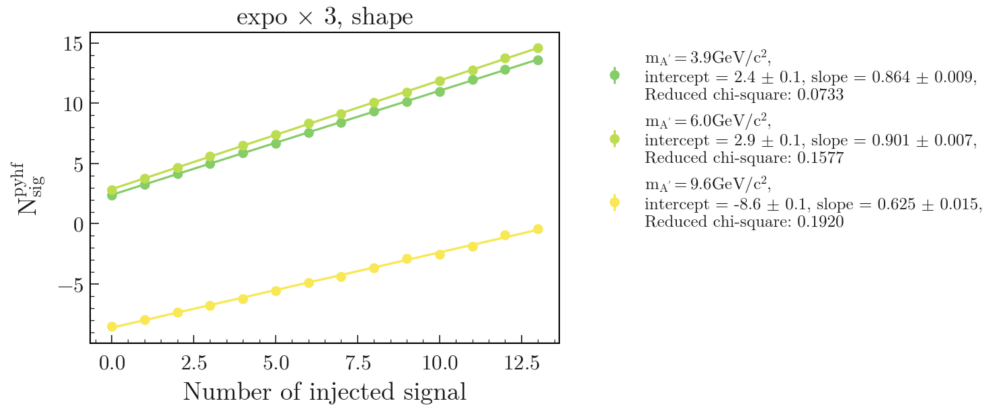
Figure 3.39: For all our background scenarios (baseline, baseline $\times 3$, exponential inflated, and exponential inflated $\times 3$) averaged over 200 trials, we show their respective exclusion contours.



(a) Light iDM mass models



(b) Medium iDM mass models



(c) Heavy iDM mass models

Figure 3.40: Linearity test with the exponential enhanced background levels $\times 3$

Table 3.11: Linearity results for exponential enhanced background $\times 3$ scenario the number of events.

	Background Shape Scheme		Uniform Shape Scheme	
$m_{A'}$ (GeV/ c^2)	Slope	Intercept	Slope	Intercept
0.15	0.776 ± 0.015	-13.33 ± 0.11	0.856 ± 0.006	0.97 ± 0.05
0.24	0.757 ± 0.016	-13.55 ± 0.12	0.841 ± 0.007	1.03 ± 0.05
0.39	0.699 ± 0.016	-13.90 ± 0.12	0.814 ± 0.007	1.08 ± 0.05
0.60	0.673 ± 0.017	-14.42 ± 0.12	0.803 ± 0.007	1.08 ± 0.05
0.96	0.665 ± 0.017	-14.43 ± 0.12	0.791 ± 0.007	1.23 ± 0.05
1.50	0.721 ± 0.016	-12.19 ± 0.12	0.789 ± 0.007	1.24 ± 0.05
2.37	0.739 ± 0.011	-3.56 ± 0.09	0.652 ± 0.007	0.79 ± 0.05
3.90	0.865 ± 0.009	2.38 ± 0.07	0.678 ± 0.008	1.09 ± 0.06
6.00	0.902 ± 0.008	2.85 ± 0.06	0.689 ± 0.008	0.33 ± 0.05
9.60	0.654 ± 0.016	-9.00 ± 0.12	0.762 ± 0.007	2.98 ± 0.05

Chapter 4

Results and Discussion

Using the details and strategies outlined in Ch. 3, this chapter outlines how our event selection performs on control samples, namely subsets of the Belle II dataset in which the signal contribution is highly suppressed relative to SM backgrounds.

Sec. 3.1 outlines how we estimate the number of background events in a given signal window using sidebands, allowing us to minimize our reliance on Monte-Carlo simulations.

A critical detail to mention is that we aim to validate the shape of the variables using a set of control samples. While overall normalization is not crucial, the BDT and background estimation methods depend on accurate estimates of feature distributions. In addition, if we want to implement a K_S^0 veto, we should check a channel rich in K_S^0 s to check for any systematic biases.

With this in mind, we propose inverting the following four signal-enhancement cuts to obtain a sample rich in continuum, photo-conversions, and tau pairs background. Firstly, the K_S^0 mass window is an obvious choice. For continuum and photo-conversions, we note that inverting the pointing-angle cut ($\cos \angle_{PA} < 0.999$) yields backgrounds rich in photo-conversions and continuum. An additional cut that helps enhance or suppress predominantly photo-conversions is the photo-converted vertex mass ($m_{\gamma \rightarrow e^+e^-} < 0.01 \text{ GeV}/c^2$). Finally, we check the back-to-backness of the ISR photon and the displaced vertex ($\hat{p}_{ISR} \cdot \hat{dr}_{V_0} < 0.88$) in samples with multiple neutrinos (e.g., tau pairs).

The following distributions also do not include the cut on $\theta_{miss}^{lab} > 0.216 \text{ rad}$, as this was introduced after reviewing the features in our control samples.

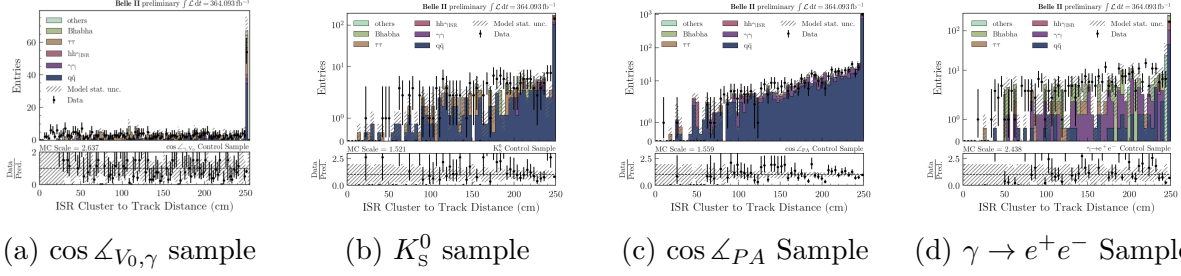


Figure 4.1: A Data/MC comparison of the minimum cluster to track distance using all our control samples.

4.1 Data and Monte-Carlo Agreement

Using our control samples, we first check the properties of the displaced tracks and the ISR photon.

For the properties of the ISR photon, all our samples provide meaningful insights. One variable of potential concern is the Zernike MVA output, as track-cluster isolation needs to be considered. Fig. 4.1 illustrates the minimum cluster to track distance for the ISR photon for all our control samples, showing that most of the events have fairly isolated photons. When we also examine the ZernikeMVA output in Fig. 4.2, our MC simulations agree quite well with the data, providing confidence that we can keep this cut.

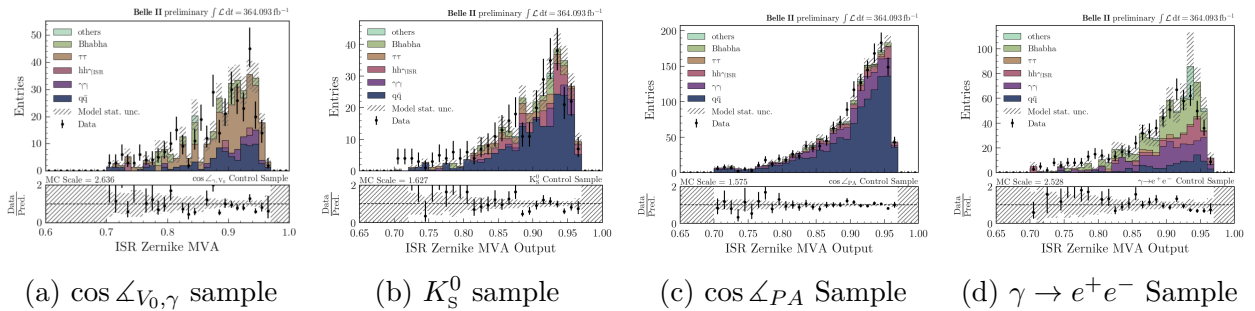


Figure 4.2: A Data/MC comparison of the ZernikeMVA output using all our control samples.

Next, we highlight the agreement between the data/MC for our tracks. For simplicity, we show only one of the tracks. For this sample, the K_S^0 and back-to-back samples are the most

representative, so we will focus on them. The main quantity we aim to establish, as it is relevant to particle identification, is the cluster energy-momentum ratio. This is illustrated in Fig. 4.3, where data/MC agreement here suggests particle ID is well-modeled.

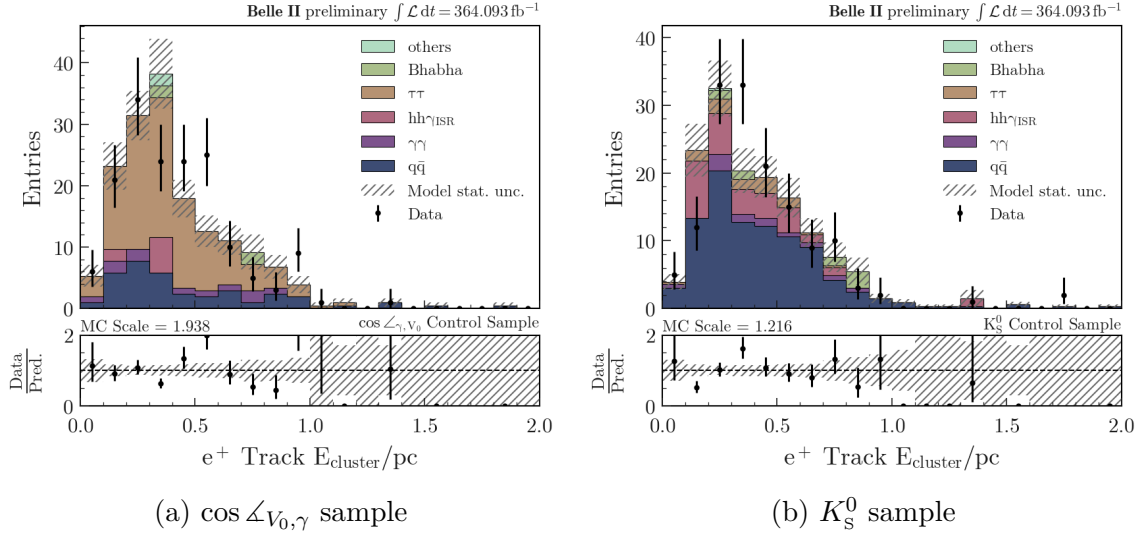


Figure 4.3: A Data/MC comparison of the cluster energy-momentum ratio using the relevant control samples.

The control samples relevant for confirming properties of the displaced vertex are mainly the back-to-back sample, though the K_s^0 sample has value for quantities that are independent of mass. As we assume displaced vertices originate from e^+ and e^- tracks, we expect some systematic shifts in the vertex's mass and energy (in the center-of-momentum system) and should therefore ensure that the simulation models them well. Fig. 4.4 illustrates agreement between data and MC simulation in our back-to-back sample, which is critical for our signal kinematic cuts.

Fig. 4.5 also shows the performance of our vertexing algorithm of K_s^0 . Whether we use $e^\pm$ or $\pi^\pm$ tracks, the data and MC simulation agreement confirms that the vertexing algorithm does perform as expected.

We also should check geometric properties of the vertex. A relevant variable is ΔR , assuming

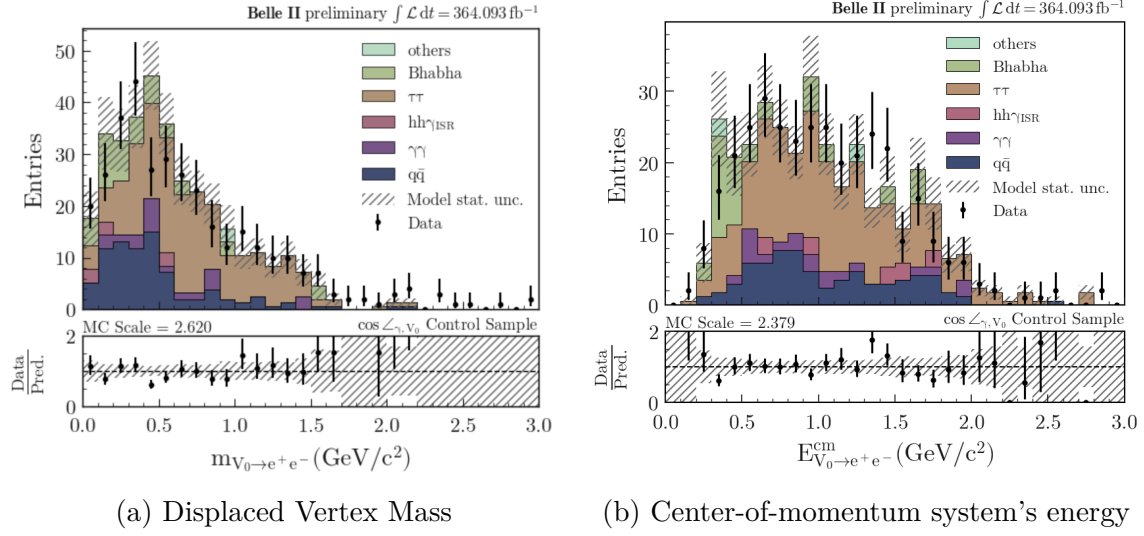


Figure 4.4: A Data/MC comparison of our back-to-back control sample of the E_{cm} and M of the displaced vertex.

the vertex arises from a photon conversion. Fig. 4.6 shows the ΔR distributions on all our control channels. From the K_s^0 control samples, the agreement for vertices $\Delta R \gg 0$ suggests our medium and heavy signal models should be reasonably constructed. From the pointing angle and converted photon sample, the reasonable agreement for events near zero suggests light signal models should also be reconstructed in a reasonable way.

To check distributions for the BDT-dependent variables, Fig. 4.7 and 4.8 show reasonable agreement in the relevant control samples.

Finally, for the overall topology of the event, we look at all but the back-to-back sample to look at $\cos \angle_{V_0, \gamma}$, as it is a BDT input variable, where in Fig. 4.9 shows reasonable data/MC agreement.

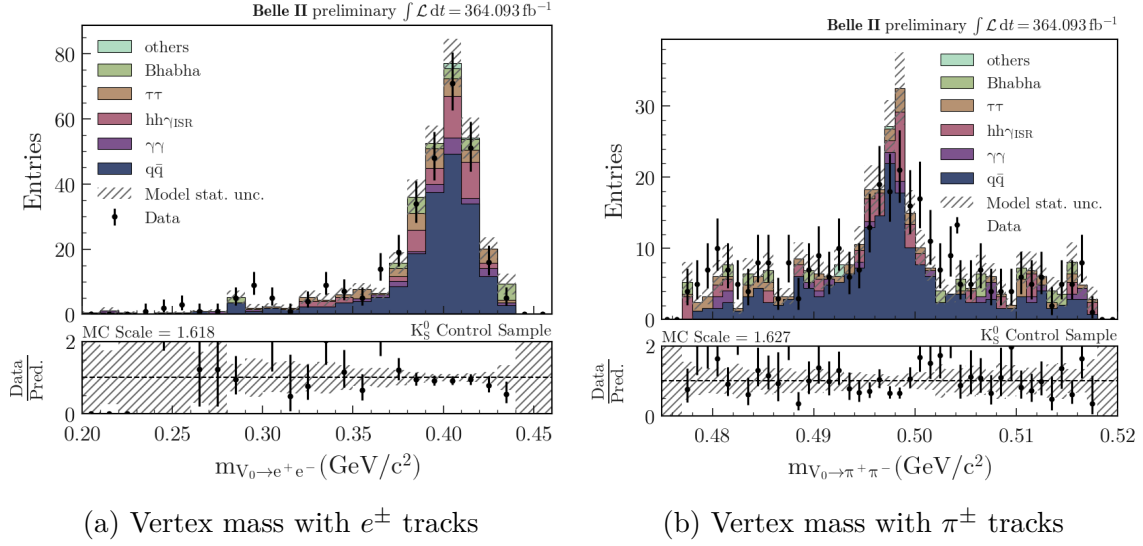


Figure 4.5: A Data/MC comparison of our K_S^0 sample of displaced vertex mass using different track hypotheses.

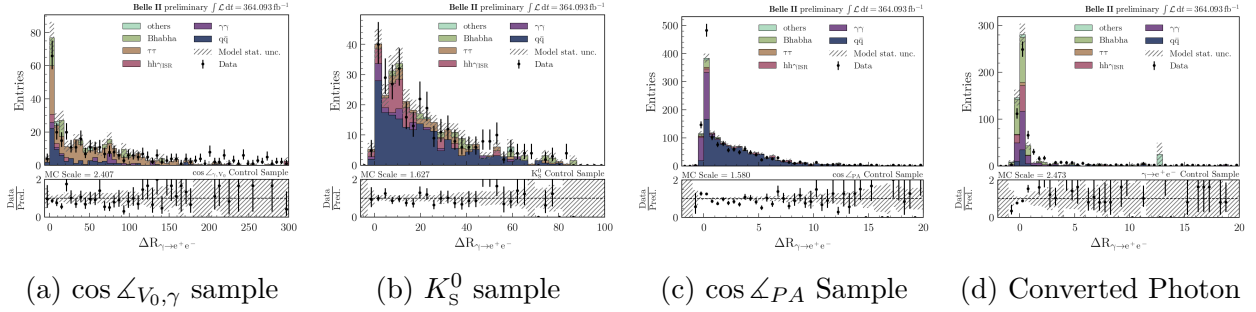


Figure 4.6: A Data/MC comparison of $\Delta R_{\gamma \rightarrow e^+ e^-}$ using all our control samples.

4.2 Validating the cut on number of CDC Hits

As emphasized in Sec. 3.2.2, we introduced a cut on the number of CDC hits (nCDCHits), despite the performance group's recommendations. We use our control samples to validate the need for this cut by relaxing this cut. Fig. 4.10/4.11 shows nCDCHits of the e^-/e^+ track for all our control samples, with MC included for comparison. For the K_S^0 , $\cos \angle_{PA}$, and converted photon sample, there are no tracks below nCDCHits < 4, so the proposed cut has no effect. Also note that despite there being no cut on nCDCHits, an inherent

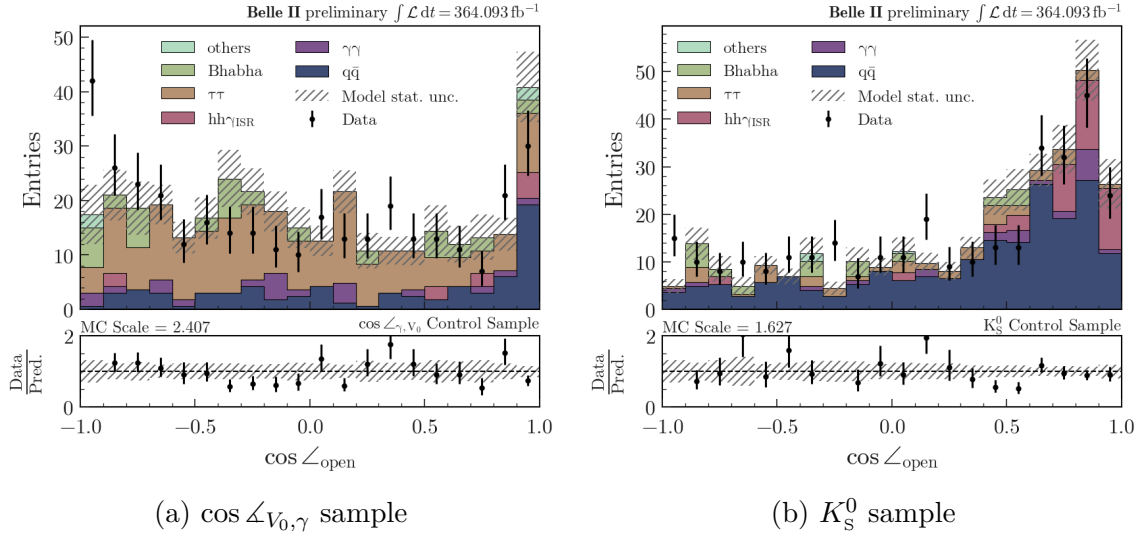


Figure 4.7: A Data/MC comparison of the opening angle (where $\cos \angle_{open} = \hat{p}_{t_0} \cdot \hat{p}_{t_1}$ using our control samples.

charge asymmetry still persists, so the loose cut suggested here does not change this. For the $\cos \angle_{V_0, \gamma} > 0.88$ sample, we see a massive excess of tracks with $nCDCHits < 4$.

To understand the properties of this excess, Fig. 4.12 shows relevant properties of interest, where we look at the events where $nCDCHits \leq 4$ of just the negatively charged track and its impact on the whole effect. We see that this excess in tracks are very soft curlers within the VXD volume which occasionally produce a CDC hit. The relative light vertex mass and the e^+ track also having a low $nCDCHits$ suggests these are unmodeled low-energy two-photon events that produce curlers and look quite different from our expected signal. In order to suppress these unmodeled backgrounds, we do indeed need $nCDCHits > 4$, as Table 3.3 suggests.

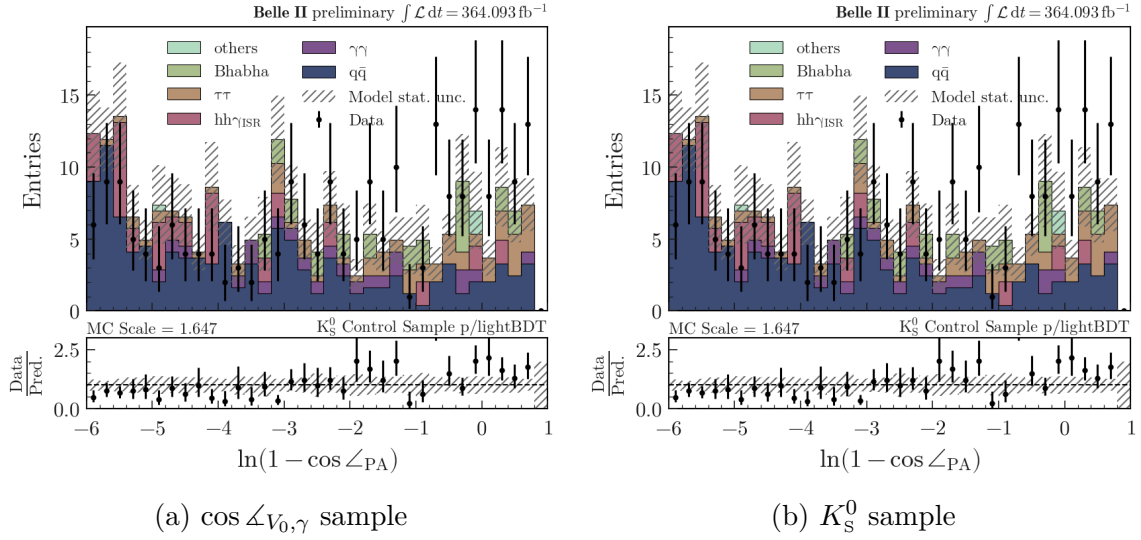


Figure 4.8: A Data/MC comparison of our control samples illustrating the pointing angle where $\cos \angle_{PA} = \hat{p}_{V_0} \cdot \hat{d}r_{V_0}$.

4.3 Validating Background Shapes

We validate the background PDF shapes defined in Sec. 3.2.3 by testing them on our control samples. As they should have varying contributions of $\gamma\gamma$ -like, $q\bar{q}$, and Bhabha-like events, we float only the yields of our three PDFs and report the $\chi^2/\text{d.o.f.}$ The results of the fits are shown in Fig. 4.13.

The $\cos \angle_{V_0, \gamma}$ sample has a reasonable $\chi^2/\text{d.o.f.}$, as the contribution from the two CBs are minimal. As we look at the other samples, the large fluctuations in the high-recoil-mass distribution are the main contributors to the larger $\chi^2/\text{d.o.f.}$

Despite some regions of large fluctuations, the general trends are captured, and the $\gamma\gamma$ -like peak and the continuum peak are well-modeled; we thus claim that the background PDFs have valid shapes.

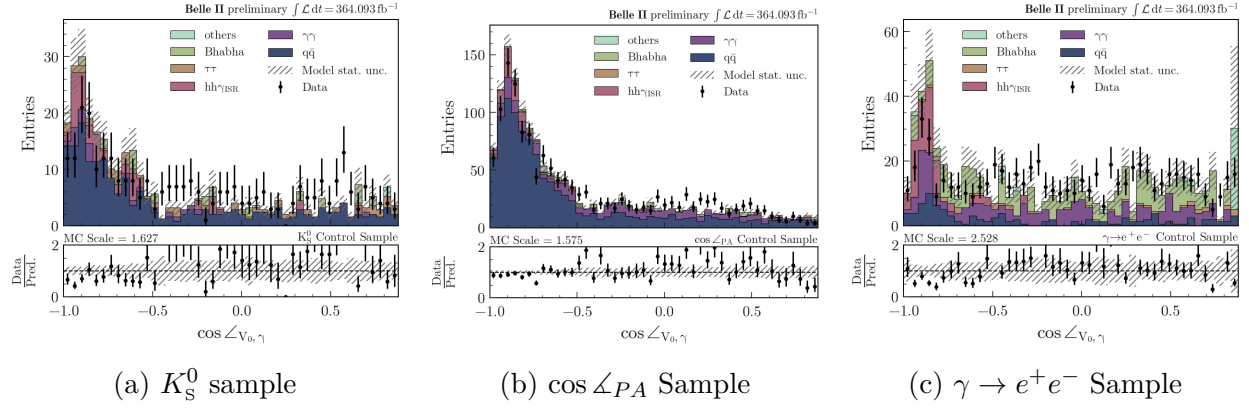
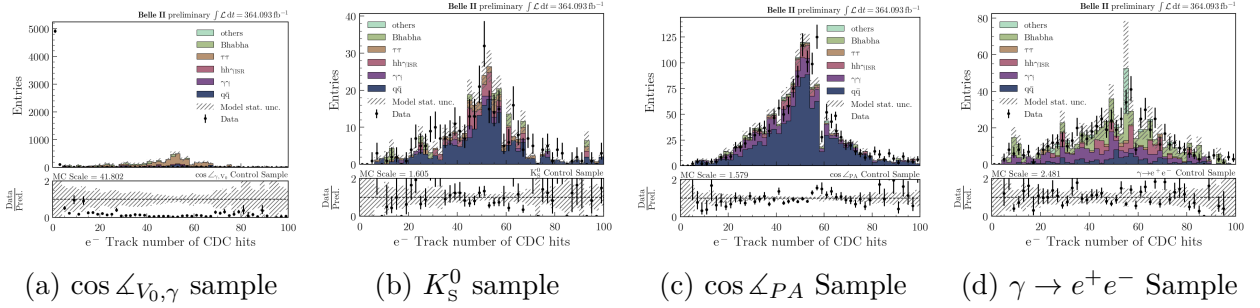
Figure 4.9: A Data/MC comparison of $\cos \Delta_{V_0, \gamma}$ using our control samples.

Figure 4.10: A Data/MC comparison of the number of CDC hits for the negatively charged track using our control samples.

4.4 Excess of ~ 1 GeV Photons

While many distributions show reasonable agreement between the data and the MC simulation, the main discrepancy we observe, aside from the event count, is an excess of low-energy photons. As shown in Fig. 4.14¹, we have a sharp excess of events near $E_{cm} \sim 1$ GeV. This impacts distributions that depend on the ISR photon's kinematics, such as the recoil mass distribution (see Fig. 4.15), total extra energy in an event, and missing energy.

Regarding the excess of events we observe across our four samples, we note the following. Many of the distributions involving the tracks have strong MC/data agreement. One vari-

¹The following distributions in this subsection do not include a cut on $\theta_{miss}^{lab} > 0.297$ rad, as this section will motivate it.

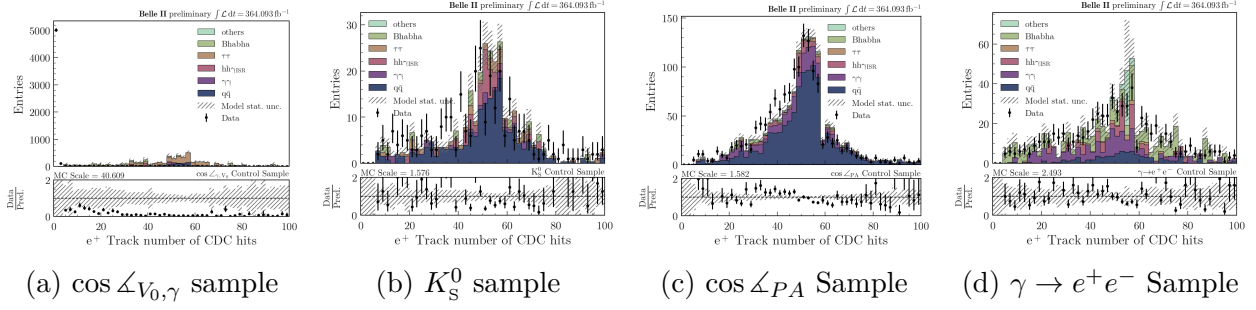


Figure 4.11: A Data/MC comparison of the number of CDC hits for the positively charged track using our control samples.

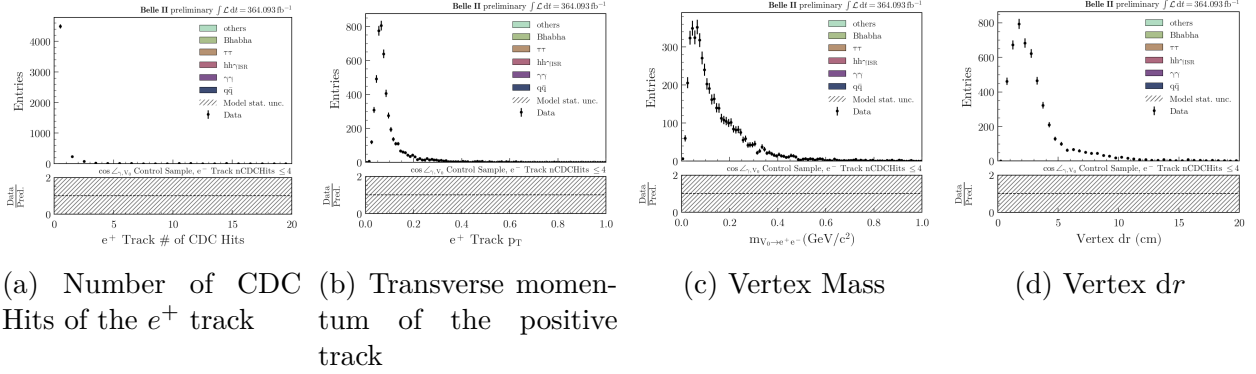


Figure 4.12: Back-to-back control sample data for the positively charged track using our control samples, with a cut on the number of CDC hits on the negatively charged track.

able that shows some discrepancy is the missing energy in a given event (E_{miss}), which is determined by taking the total energy (~ 10.5 GeV) and subtracting the energies of the ISR photon, the displaced vertex, and any low energy neutral clusters found in the ECL (E_{extra}). The E_{extra} and E_{miss} distributions are illustrated in Fig. 4.16 and 4.17, respectively.

From inspecting the MC simulations and their relation to our control samples, we notice that the distributions that peak in a similar manner to our control samples are the Bhabha samples and two-photon processes. This suggests that the excess may arise from either two-photon processes or radiative Bhabha events that are not included in the simulations. Fig. 4.18 also indicates that events that look similar to two-photon and Bhabha events are likely the culprit. If they really are bona fide two-photon or Bhabha events, then we expect much

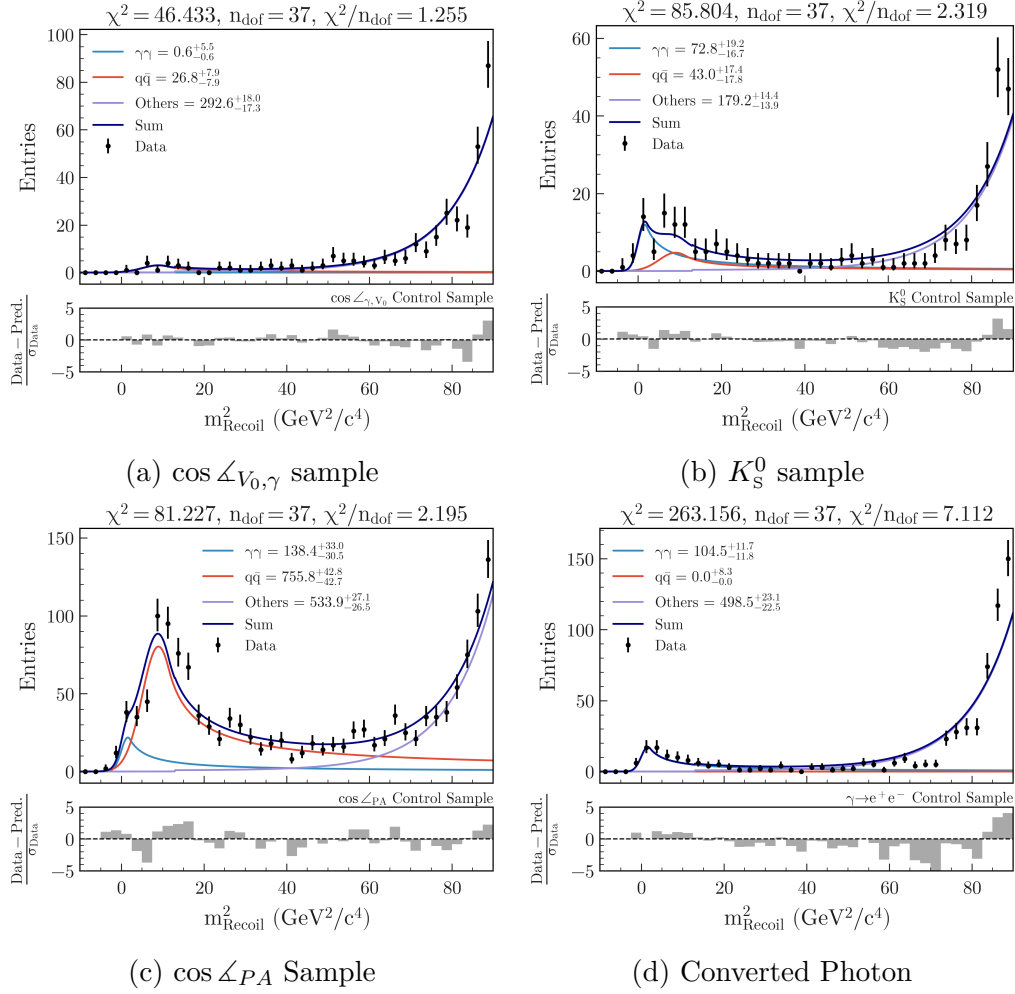
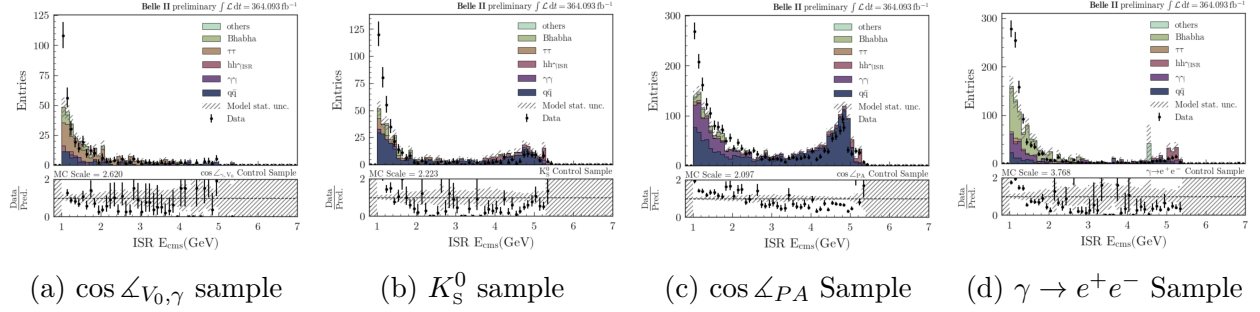
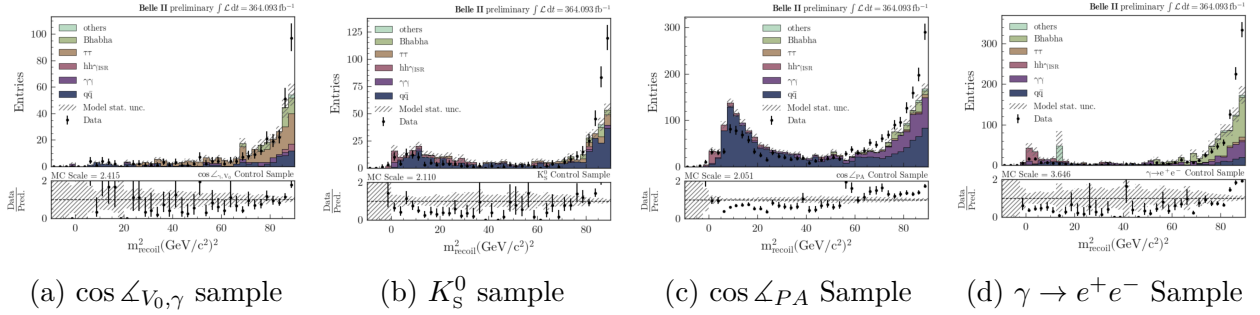


Figure 4.13: Floating only the yields, we fit our control samples using our background PDFs defined in Sec. 3.2.3.

of this excess to still be suppressed after applying our BDTs to our control samples.

As additional confirmation, we examine the missing momentum direction. To easily interpret it and its relation to the Belle II detector, Fig. 4.19 illustrated $\theta_{\text{miss}}^{\text{lab}} = \cos^{-1}(\hat{p}_{\text{miss},z})$, where $\vec{p}_{\text{miss}} = \vec{p}_{\text{beams}} - \sum_i^{FSP} p_i$. Since the excess is predominantly in the beamline direction, this further supports the hypothesis that our excess is due to two-photon physics, with tracks and neutrals traveling undetected near the beamline and faking the iDM signal of interest.

To address this, we inspect the signal distributions, which are illustrated in Fig. 4.19. If we

Figure 4.14: A Data/MC comparison of E_{cms} using all our control samples.Figure 4.15: A Data/MC comparison of recoil (from the ISR photon) squared mass (m_{recoil}^2) distribution using all our control samples.

were to propose a physically motivated cut to suppress these two-photon events, it would either have to be at the CDC acceptance of $\theta_{miss}^{lab} = 0.297$ rad or the ECL acceptance of $\theta_{miss}^{lab} = 0.216$ rad. Based on the box-and-whisker plot of the heaviest model (see Fig. 3.4), we elect to cut on the CDC acceptance, as shown in Table 3.5 and applied retroactively. In addition, we also see an excess of events at the ECL gap at ~ 0.55 rad and so also constrain $\theta_{miss}^{lab} \notin [0.54, 0.57]$ rad.

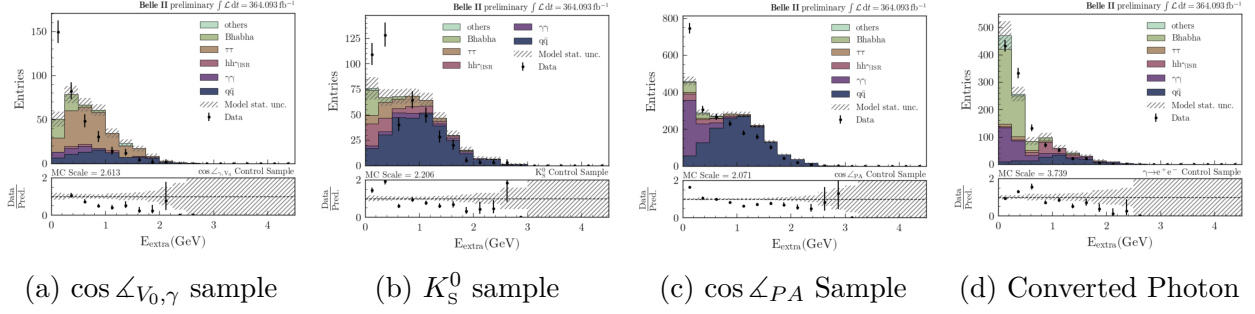


Figure 4.16: A Data/MC comparison of total extra energy in an event using all our control samples.

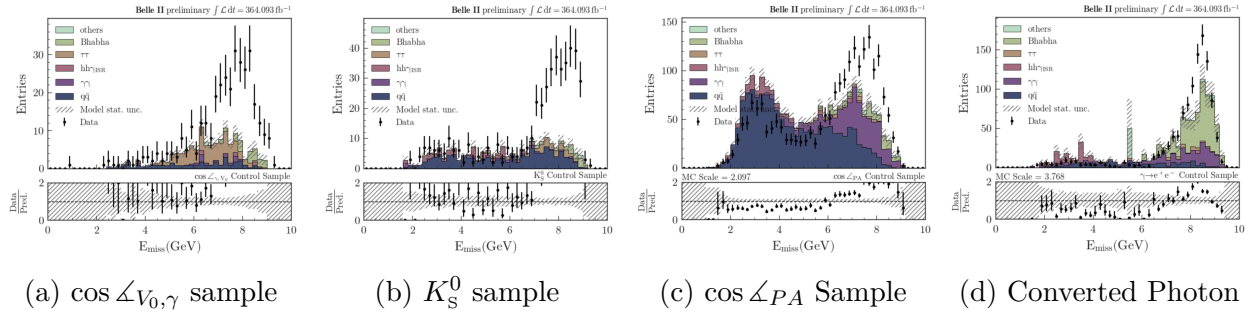


Figure 4.17: A Data/MC comparison of missing energy in an event using all our control samples.

4.5 Validating BDT performance

Except for this excess of unmodeled 1 GeV photon events, it seems that most of our simulation modeling matches our expectation, so we apply the P.F.O.M optimized selection on our control sample to assess how well it performs on background-enriched data. Additionally, we note that the θ_{miss}^{lab} and the number of CDC hit cuts have been applied here.

The modified recoil mass distributions with this fix are illustrated in Fig. 4.20, where the θ_{miss}^{lab} cut helps improve our data-MC agreement (i.e., closer to a ratio of 1).

To avoid showing 3×4 plots every time we show a distribution, we spotlight the most relevant BDT for the corresponding sample. The K_S^0 and back-to-back samples have the highest signal efficiency for heavier iDM models, so the results of applying the BDT model

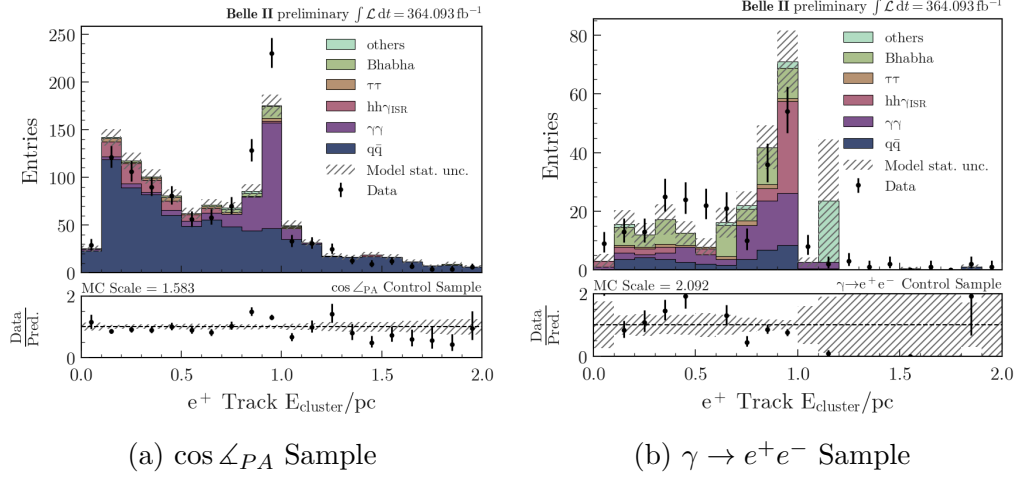


Figure 4.18: Illustrating the excess of events with $E_{\text{cluster}}/pc \sim 1$ for our converted photon and pointing angle sample.

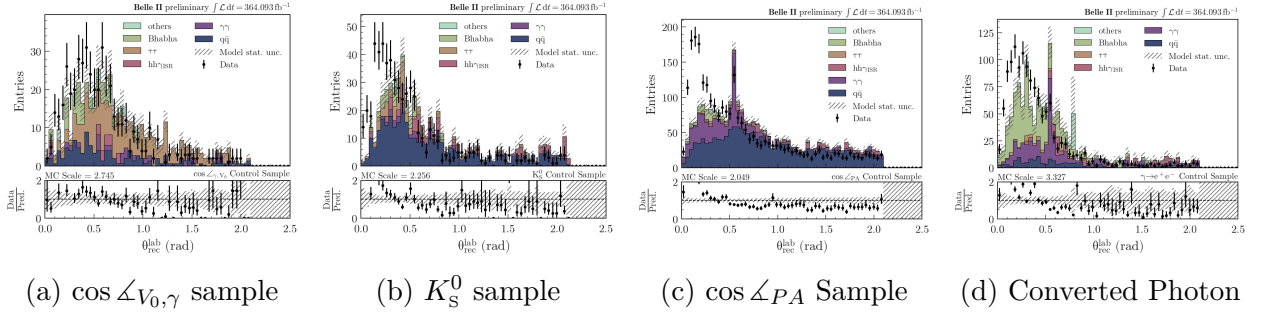


Figure 4.19: A Data/MC comparison of $\theta_{\text{miss}}^{\text{lab}}$ in an event using all our control samples.

trained on heavy models are shown in Fig. 4.21. The back-to-back sample post-BDTs are all empty, so their plots are excluded. The pointing angle and converted-photon samples have the highest signal efficiency for lighter iDM models, so we present the light and medium BDT models in Fig. 4.22 and 4.23, respectively. As these samples have the largest number of events that pass in both data and MC, we use them to assess our BDT performance on data.

To characterize this, let us clarify what the normalization factor is applied to in the simulation to match the number of events observed in the control samples. In the pre-BDT plots in Fig. 4.15, the normalizations for the pointing angle and converted photon samples are 1.558 and

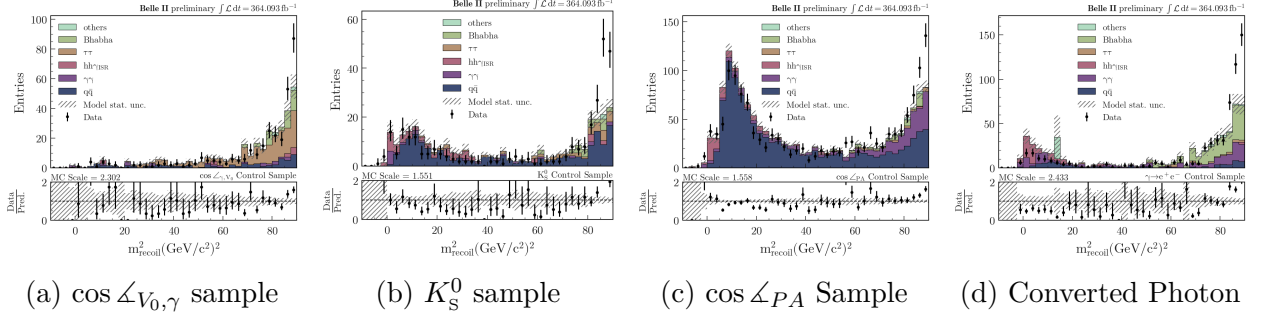


Figure 4.20: A Data/MC comparison of m_{recoil}^2 in an event using all our control samples.

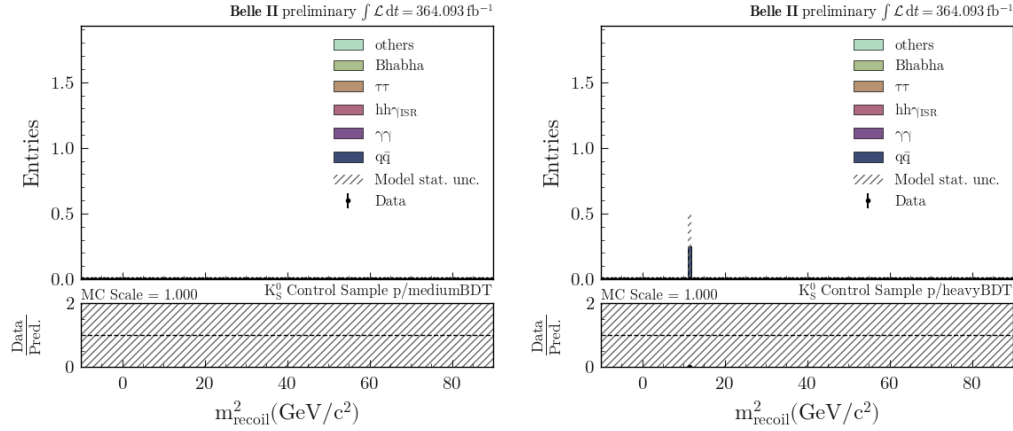


Figure 4.21: Applying the medium and heavy BDT and ΔR cut on the K_S^0 sample.

2.443, respectively. After applying the light BDT schemes, the normalization factors become 2.42 and 1.40. After applying the medium BDT schemes, the normalization factors become 2.50 and 1.56. While the normalizations have fluctuated, the statistics in the MC samples are low, and the fluctuations are less than a factor of 3, which still suggests our BDTs will retain their $\mathcal{O}(1)$ performance during unblinding.

Additionally, the distributions are consistent in shape, suggesting that the relative abundances of the different background components are comparable.

Although there is a minor discrepancy in the expected background, the general trends from the BDT cuts strongly suggest we can achieve similar performance in our nominal selection.

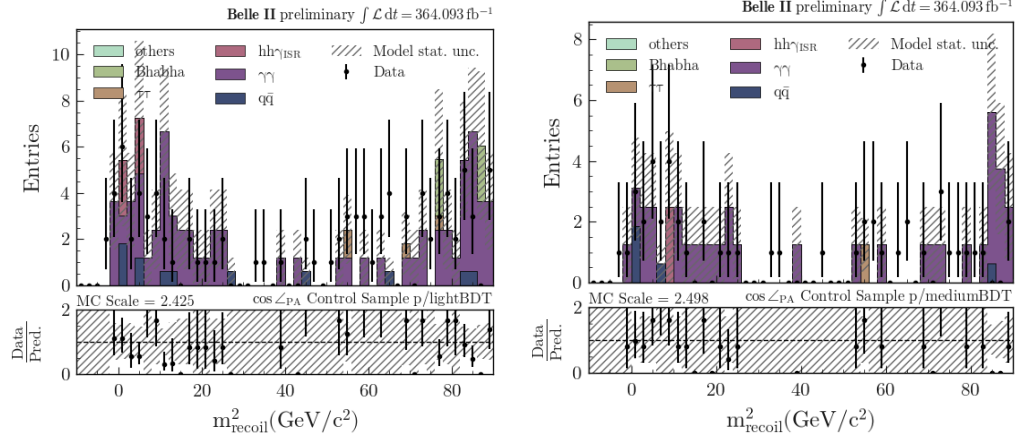


Figure 4.22: Applying the light and medium BDT and ΔR cut on the $\cos \angle_{PA}$ sample.

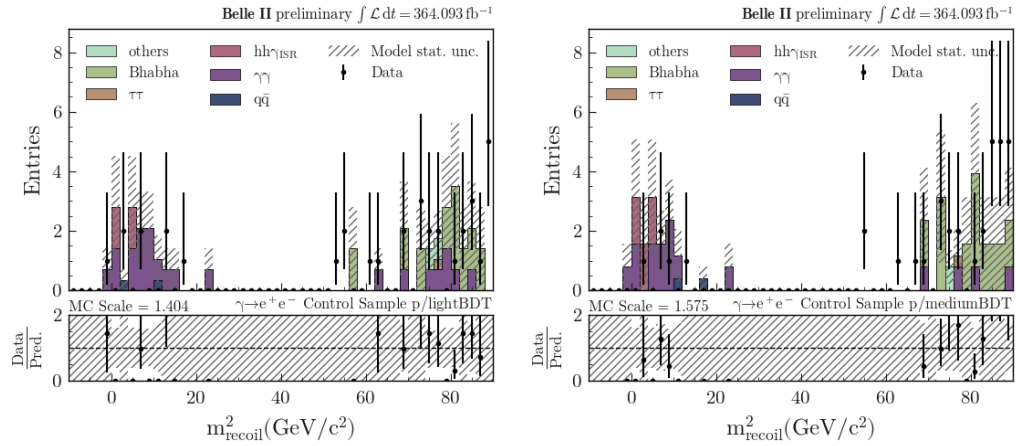


Figure 4.23: Applying all the BDTs and ΔR cut on the converted photon sample.

Chapter 5

Conclusions

As we continue to work through the mandatory stages of Belle II internal protocols and reviews, the work we present here is the final analysis methodology to be applied to the real data once permission is granted to “open the box” for setting world-first limits for inelastic dark matter using our Run 1 dataset of an integrated luminosity of 364fb^{-1} . Inelastic dark matter is an attractive extension of the Standard Model as it provides a mechanism for the abundance of dark matter to be thermally frozen in the post-recombination era. This model contains a massive dark photon A' that acts as our Standard Model portal via kinetic mixing, a stable dark fermion χ_1 , and an unstable χ_2 whose main decay channel is $\chi_2 \rightarrow A'\chi_1$, where a virtual A' could decay within the Belle II volume and produce a displaced vertex via electron or muon tracks. In addition, the initial-state radiated photons would provide important kinematic information for measuring the A' mass (using the known four-momentum of the initial e^+e^- collision) and for the event trigger. We apply some minimal pre-selections to work with bona fide photons and well-reconstructed displaced vertices. To isolate these rare signals from Standard Model backgrounds, a sophisticated selection pipeline is developed. An ensemble of forty-five Gradient-Boosted Decision Trees is employed, leveraging kinematic properties and particle-identification variables to enhance signal significance. Given the expected low background levels, we performed a counting experiment for each mass hypothesis, in which the sidebands are fitted to provide a background estimate in the corresponding signal window. Upper limits are computed at a 95% confidence level to determine Belle II’s sensitivity.

While this work has not yet unblinded the signal region, the established limits and selection techniques are poised to explore significant new regions of the inelastic dark matter parameter space, potentially setting world-first constraints on inelastic dark matter couplings and complementing the searches conducted by other experiments in high-energy particle physics. As luminosity continues to increase, the combination of high-precision tracking and advanced

machine learning filters will remain essential for probing the frontiers of the Standard Model.

Bibliography

- [1] Mary K. Gaillard, Paul D. Grannis, and Frank J. Sciulli. The standard model of particle physics. *Reviews of Modern Physics*, 71(2):S96–S111, March 1999.
- [2] Alessandro Bettini. *Introduction to elementary particle physics*. Cambridge Univ. Press, Cambridge, 2008.
- [3] S. Navas et al. Review of particle physics. *Phys. Rev. D*, 110(3):030001, 2024.
- [4] C. N. Yang and R. L. Mills. Conservation of Isotopic Spin and Isotopic Gauge Invariance. *Phys. Rev.*, 96:191–195, Oct 1954.
- [5] Ayres Freitas. TASI 2020 Lectures on Precision Tests of the Standard Model, 2020.
- [6] Guido Altarelli. The Higgs and the Excessive Success of the Standard Model, 2014.
- [7] Tomio Kobayashi and Takaaki Kajita. Experimental verification of the standard model of particle physics. *Proceedings of the Japan Academy, Series B. Physical and Biological Sciences (Online)*, 97(5), December 2024.
- [8] Yann Gouttenoire. Beyond the Standard Model Cocktail, 2022.
- [9] Martin Rosenlyst. Addressing Six Standard Model Problems with Technically Natural Higgs Models, 2024.
- [10] Dhananjay Saikumar. Exploring the Frontiers: Challenges and Theories Beyond the Standard Model, 2024.
- [11] Nima Arkani-Hamed, Savas Dimopoulos, and Gia Dvali. The hierarchy problem and new dimensions at a millimeter. *Physics Letters B*, 429(3–4):263–272, June 1998.

- [12] Seth Koren. The Hierarchy Problem: From the Fundamentals to the Frontiers, 2020.
- [13] Maurice Goldhaber. A closer look at the elementary fermions. *Proceedings of the National Academy of Sciences*, 99(1):33–36, January 2002.
- [14] Y. Fukuda et al. Evidence for Oscillation of Atmospheric Neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998.
- [15] Q. R. Ahmad et al. Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89:011301, Jun 2002.
- [16] Larisa A. Thorne. Neutrino mass experiments: current and future, 2024.
- [17] KATRIN Collaboration[†]. Direct neutrino-mass measurement based on 259 days of KATRIN data. *Science*, 388(6743):180–185, 2025.
- [18] Laurent Canetti, Marco Drewes, and Mikhail Shaposhnikov. Matter and antimatter in the universe. *New Journal of Physics*, 14(9):095012, September 2012.
- [19] M. Bauer. Matter-antimatter Asymmetry, CP Violation and the Time Operator in Relativistic Quantum Mechanics, 2022.
- [20] R. D. Peccei and Helen R. Quinn. CP Conservation in the Presence of Pseudoparticles. *Phys. Rev. Lett.*, 38:1440–1443, Jun 1977.
- [21] Dan-di Wu. A Brief Introduction to the Strong CP Problem. *Z. Naturforsch. A*, 52:179–181, 1997.
- [22] K. Freese. Review of Observational Evidence for Dark Matter in the Universe and in upcoming searches for Dark Stars. *EAS Publications Series*, 36:113–126, 2009.

- [23] Csaba Balazs, Torsten Bringmann, Felix Kahlhoefer, and Martin White. A Primer on Dark Matter, 2024.
- [24] A V Zasov, A S Saburova, A V Khoperskov, and S A Khoperskov. Dark matter in galaxies. *Physics-Uspekhi*, 60(1):3–39, April 2017.
- [25] V C Rubin and W K Ford, Jr. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *Astrophys. J.* 159: 379-403(Feb 1970)., Jan 1970.
- [26] Daeun Jeong, Ho Seong Hwang, Haeun Chung, and Yongmin Yoon. Diverse Rotation Curves of Galaxies in a Simulated Universe: The Observed Dependence on Stellar Mass and Morphology Reproduced. *The Astrophysical Journal*, 982(1):11, March 2025.
- [27] S. D. Ryder, A. V. Zasov, V. J. McIntyre, W. Walsh, and O. K. Sil’chenko. The peculiar rotation curve of ngc 157. *Monthly Notices of the Royal Astronomical Society*, 293(4):411–428, February 1998.
- [28] W. J. G. de Blok, F. Walter, E. Brinks, C. Trachternach, S. H. Oh, and R. C. Kennicutt, Jr. High-Resolution Rotation Curves and Galaxy Mass Models from THINGS. , 136(6):2648–2719, December 2008.
- [29] Steven Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, chapter Classic Tests of Einstein’s Theory. John Wiley and Sons, New York, 1972.
- [30] R. d’Inverno. *Introducing Einstein’s relativity*. 1992.
- [31] Richard Massey, Thomas Kitching, and Johan Richard. The dark matter of gravitational lensing. *Reports on Progress in Physics*, 73(8):086901, July 2010.

- [32] S. Lin, W. Luo, Y. F. Cai, Q. Guo, L. Wei, B. Wang, Q. Li, C. Su, and A. Rodriguez. The density profile of milky way dark matter halo constrained from the ogle microlensing sky map. *Astronomy & Astrophysics*, 693:L8, January 2025.
- [33] Tom Broadhurst et al. Dark Matter Distinguished by Skewed Microlensing in the “Dragon Arc”. *Astrophys. J. Lett.*, 978(1):L5, 2025.
- [34] O’Riordan et al. Euclid: A complete einstein ring in ngc 6505. *Astronomy & Astrophysics*, 694:A145, February 2025.
- [35] Daniela Bettoni, Renato Falomo, Riccardo Scarpa, Mattia Negrello, Alessandro Omizolo, Romano L. M. Corradi, Daniel Reverte, and Benedetta Vulcani. A new einstein cross gravitational lens of a lyman-break galaxy. *The Astrophysical Journal Letters*, 873(2):L14, March 2019.
- [36] Priyamvada Natarajan, Liliya L. Williams, Marusa Bradac, Claudio Grillo, Agniva Ghosh, Keren Sharon, and Jenny Wagner. Strong lensing by galaxy clusters, 2024.
- [37] Prasenjit Saha, Dominique Sluse, Jenny Wagner, and Liliya L. R. Williams. Essentials of strong gravitational lensing, 2024.
- [38] S. Vegetti, S. Birrer, G. Despali, C. D. Fassnacht, D. Gilman, Y. Hezaveh, L. Perreault Levasseur, J. P. McKean, D. M. Powell, C. M. O’Riordan, and G. Vernardos. Strong gravitational lensing as a probe of dark matter, 2023.
- [39] Matthias Bartelmann and Matteo Maturi. Weak gravitational lensing, 2016.
- [40] Douglas Clowe, Maruša Bradač, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. A Direct Empirical Proof of the Existence of Dark Matter. *The Astrophysical Journal*, 648(2):L109–L113, August 2006.

- [41] Robert J. Nemiroff and Bijunath Patla. Adventures in friedmann cosmology: A detailed expansion of the cosmological friedmann equations. *American Journal of Physics*, 76(3):265–276, March 2008.
- [42] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophysical Journal*, 292:371–394, May 1985.
- [43] Mark Vogelsberger, Federico Marinacci, Paul Torrey, and Ewald Puchwein. Cosmological simulations of galaxy formation, 2019.
- [44] Wayne Hu, Naoshi Sugiyama, and Joseph Silk. The physics of microwave background anisotropies. *Nature*, 386(6620):37–43, March 1997.
- [45] Wayne Hu and Martin White. Acoustic signatures in the cosmic microwave background. *The Astrophysical Journal*, 471(1):30–51, November 1996.
- [46] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright. Nine-year wilkinson microwave anisotropy probe (wmap) observations: Final maps and results. *The Astrophysical Journal Supplement Series*, 208(2):20, September 2013.
- [47] N. Aghanim et al. Planck2018 results: VI. Cosmological parameters. *Astronomy & Astrophysics*, 641:A6, September 2020.
- [48] Lars Hernquist, Neal Katz, David H. Weinberg, and Jordi Miralda-Escudé. The lyman-alpha forest in the cold dark matter model. *The Astrophysical Journal*, 457(2), February 1996.

- [49] Paul Murdin. Encyclopedia of Astronomy and Astrophysics. chapter L, pages 148–155. 2002.
- [50] Alex Woodfinden, Will J Percival, Seshadri Nadathur, Hans A Winther, T S Fraser, Elena Massara, Enrique Paillas, and Slađana Radinović. Cosmological measurements from void-galaxy and galaxy-galaxy clustering in the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 523(4):6360–6370, June 2023.
- [51] A.G. Adame et al. Desi 2024 iv: Baryon acoustic oscillations from the lyman alpha forest. *Journal of Cosmology and Astroparticle Physics*, 2025(01):124, January 2025.
- [52] Gary Steigman. Primordial nucleosynthesis: The predicted and observed abundances and their consequences, 2010.
- [53] Jeremy Bernstein, Lowell S. Brown, and Gerald Feinberg. Cosmological helium production simplified. *Rev. Mod. Phys.*, 61:25–39, Jan 1989.
- [54] V. Mukhanov. Nucleosynthesis without computer. *International Journal of Theoretical Physics*, 43(3):669–693, March 2004.
- [55] Richard H. Cyburt, Brian D. Fields, Keith A. Olive, and Tsung-Han Yeh. Big bang nucleosynthesis: Present status. *Reviews of Modern Physics*, 88(1), February 2016.
- [56] Ryan J. Cooke, Max Pettini, Regina A. Jorgenson, Michael T. Murphy, and Charles C. Steidel. Precision measures of the primordial abundance of deuterium. *The Astrophysical Journal*, 781(1):31, January 2014.
- [57] T M C Abbott et al. Dark Energy Survey Year 1 Results: A Precise H0 Estimate from DES Y1, BAO, and D/H Data. *Monthly Notices of the Royal Astronomical Society*, 480(3):3879–3888, July 2018.

- [58] A.G. Adame et al. DESI 2024 VI: cosmological constraints from the measurements of baryon acoustic oscillations. *Journal of Cosmology and Astroparticle Physics*, 2025(02):021, February 2025.
- [59] Kevork N. Abazajian et al. The Seventh Data Release of the Sloan Digital Sky Survey. *The Astrophysical Journal Supplement Series*, 182(2):543–558, May 2009.
- [60] Andrés Almeida et al. The eighteenth data release of the sloan digital sky surveys: Targeting and first spectra from sdss-v. *The Astrophysical Journal Supplement Series*, 267(2):44, August 2023.
- [61] Y. I. Izotov, T. X. Thuan, and N. G. Guseva. A new determination of the primordial He abundance using the He-I 10830 Å emission line: cosmological implications. *Monthly Notices of the Royal Astronomical Society*, 445(1):778–793, September 2014.
- [62] S. A. Balashev, E. O. Zavarygin, A. V. Ivanchik, K. N. Telikova, and D. A. Varshalovich. The primordial deuterium abundance: subdla system atzabs = 2.437 towards the qso j 1444+2919. *Monthly Notices of the Royal Astronomical Society*, 458(2):2188–2198, February 2016.
- [63] Anson Hook, Eder Izaguirre, and Jay G. Wacker. Model independent bounds on kinetic mixing, 2010.
- [64] Benoît Famaey and Stacy S. McGaugh. Modified Newtonian Dynamics (MOND): Observational Phenomenology and Relativistic Extensions. *Living Reviews in Relativity*, 15(1), September 2012.
- [65] Benoit Famaey and Amel Durakovic. Modified Newtonian Dynamics (MOND), 2025.
- [66] João Magueijo and Ali Mozaffari. Case for testing modified newtonian dynamics using lisa pathfinder. *Physical Review D*, 85(4), February 2012.

- [67] A.Yu. Ignatiev. Testing mond on earth. *Canadian Journal of Physics*, 93(2):166–168, February 2015.
- [68] Jonathan L. Feng. The wimp paradigm: Theme and variations. *SciPost Physics Lecture Notes*, June 2023.
- [69] D.S. Akerib et al. The Large Underground Xenon (LUX) experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 704:111–126, March 2013.
- [70] D.S. Akerib et al. The lux-zeplin (lz) experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 953:163047, 2020.
- [71] Li Zhao and Jianglai Liu. PandaX: A deep underground dark matter search experiment in China using liquid xenon. *Modern Physics Letters A*, 33(30):1830013, September 2018.
- [72] E. Aprile et al. The XENON1T dark matter experiment. *The European Physical Journal C*, 77(12), December 2017.
- [73] Andrea Zani. The DarkSide-20k experiment, 2024.
- [74] SuperCDMS Collaboration. A Strategy for Low-Mass Dark Matter Searches with Cryogenic Detectors in the SuperCDMS SNOLAB Facility, 2023.
- [75] K. Altenmüller et al. New Upper Limit on the Axion-Photon Coupling with an Extended CAST Run with a Xe-Based Micromegas Detector. *Physical Review Letters*, 133(22), November 2024.

- [76] Quan-feng Wu and Xun-Jie Xu. A comprehensive calculation of the Primakoff process and the solar axion flux. *Journal of Cosmology and Astroparticle Physics*, 2024(07):013, July 2024.
- [77] R. Khatiwada et al. Axion Dark Matter Experiment: Detailed design and operations. *Review of Scientific Instruments*, 92(12), December 2021.
- [78] ADMX Collaboration. Search for Axion Dark Matter from 1.1 to 1.3 GHz with ADMX, 2025.
- [79] Peter Athron, Csaba Balázs, Ankit Beniwal, J. Eliel Camargo-Molina, Andrew Fowlie, Tomás E. Gonzalo, Sebastian Hoof, Felix Kahlhoefer, David J. E. Marsh, Markus Tobias Prim, Andre Scaffidi, Pat Scott, Wei Su, Martin White, Lei Wu, and Yang Zhang. Global fits of axion-like particles to XENON1T and astrophysical data. *Journal of High Energy Physics*, 2021(5), May 2021.
- [80] M. F. Albakry et al. Light dark matter constraints from SuperCDMS HVeV detectors operated underground with an anticoincidence event selection. *Physical Review D*, 111(1), January 2025.
- [81] Xinning Zeng et al. Exploring New Physics with PandaX-4T Low Energy Electronic Recoil Data. *Physical Review Letters*, 134(4), January 2025.
- [82] Bernard Carr and Florian Kuhnel. Primordial black holes, 2025.
- [83] Nicolas Herman, André Füzfa, Léonard Lehoucq, and Sébastien Clesse. Detecting planetary-mass primordial black holes with resonant electromagnetic gravitational-wave detectors. *Physical Review D*, 104(2), July 2021.
- [84] Michael Duerr, Torben Ferber, Christopher Hearty, Felix Kahlhoefer, Kai Schmidt-

- Hoberg, and Patrick Tunney. Invisible and displaced dark matter signatures at Belle II. *JHEP*, 02:039, 2020.
- [85] Céline Boehm, Matthew J Dolan, and Christopher McCabe. Increasingneffwith particles in thermal equilibrium with neutrinos. *Journal of Cosmology and Astroparticle Physics*, 2012(12):027–027, December 2012.
- [86] Paul Frederik Depa, Marco Hufnagel, Kai Schmidt-Hoberg, and Sebastian Wild. Bbn constraints on the annihilation of mev-scale dark matter. *Journal of Cosmology and Astroparticle Physics*, 2019(04):029–029, April 2019.
- [87] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), July 2014.
- [88] Belle II collaboration. Belle II Analysis Software Framework (basf2). <https://doi.org/10.5281/zenodo.5574115>.
- [89] Kazunori Akai, Kazuro Furukawa, and Haruyo Koiso. Superkekb collider. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 907:188–199, November 2018.
- [90] K. Furukawa Y. Funakoshi. SuperKEKB Design Report: Overview, September 2019.
- [91] Eugenio Paoloni. Status of superkekb and belle ii. Belle II Collaboration Talk, BELLE2-TALK-CONF-2016-024, June 2016. Slides 11–12 referenced.
- [92] S. Kurokawa and Eiji Kikutani. Overview of the KEKB accelerators. *Nucl. Instrum. Meth. A*, 499:1–7, 2003.

- [93] M S Zisman. The pep-ii project: Design status and r&d results. Lawrence Berkeley Lab., CA (United States); Stanford Linear Accelerator Center, Menlo Park, CA (United States); Lawrence Livermore National Lab., CA (United States), 07 1993.
- [94] SuperB Collaboration. Superb: A high-luminosity asymmetric $e^+ e^-$ super flavor factory. conceptual design report, 2007.
- [95] W Herr and T Pieloni. Beam-Beam Effects, 2014.
- [96] N. Ohuchi, Y. Arimoto, K. Akai, K. Aoki, N. Higashi, K. Kanazawa, M. Kawai, T. Kawamoto, H. Koiso, Y. Kondou, M. Masuzawa, A. Morita, S. Nakamura, Y. Ohnishi, Y. Ohsawa, K. Oide, N. Okada, T. Oki, H. Sugimoto, M. Tawada, K. Tsuchiya, R. Ueki, X. Wang, H. Yamaoka, Z. Zong, B. Parker, M. Anerella, J. Escallier, A. Ghosh, H. Hocker, A. Jain, A. Marone, P. Wanderer, J. DiMarco, J. Nogiec, M. Tartaglia, and G. Velev. SuperKEKB beam final focus superconducting magnet system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1021:165930, 2022.
- [97] K. Furukawa Y. Funakoshi. SuperKEKB Design Report: IR Design, October 2020.
- [98] T. Abe et al. Belle II Technical Design Report. 11 2010.
- [99] How a detector works: Just as hunters can identify animals from tracks in mud or snow, physicists identify subatomic particles from the traces they leave in detectors. <https://home.cern/science/experiments/how-detector-works>. Accessed: 2022-10-19.
- [100] E Kou, P Urquijo, et al. The Belle II Physics Book. *Progress of Theoretical and Experimental Physics*, 2019(12), Dec 2019.

- [101] H. Ye et al. Commissioning and performance of the Belle II pixel detector. *Nucl. Instrum. Meth. A*, 987:164875, 2021.
- [102] PXD: Belle II Detector. <https://belle2.jp/pxd/>. Accessed: 2024-07-31.
- [103] I. Adachi et al. Precise Measurement of the D_s^+ Lifetime at Belle II. *Phys. Rev. Lett.*, 131:171803, Oct 2023.
- [104] Y. Uematsu et al. The Silicon Vertex Detector of the Belle II experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1033:166688, 2022.
- [105] K. Adamczyk et al. The design, construction, operation and performance of the Belle II silicon vertex detector. *Journal of Instrumentation*, 17(11):P11042, November 2022.
- [106] G. Lutz and R. Klanner. *Solid State Detectors*, chapter 5, pages 137–200. Springer International Publishing, Cham, 2020.
- [107] C. Marinas et al. The Belle II Detector at the SuperKEKB Flavor Factory. *JINST*, PLACEHOLDER 2025.
- [108] N. Taniguchi. Central drift chamber for belle-ii. *Journal of Instrumentation*, 12(06):C06014, jun 2017.
- [109] Hulya Atmacan et al. The Imaging Time-of-Propagation Detector at Belle II, 2025.
- [110] M. Starič. The top counter and determination of bunch-crossing time at belle ii. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1055:168474, October 2023.
- [111] I. H. Malitson. Interspecimen Comparison of the Refractive Index of Fused Silica*,†. *Journal of the Optical Society of America (1917-1983)*, 55(10):1205, October 1965.

- [112] T. Iijima, S. Korpar, I. Adachi, S. Fratina, T. Fukushima, A. Gorišek, H. Kawai, H. Konishi, Y. Kozakai, P. Križan, T. Matsumoto, Y. Mazuka, S. Nishida, S. Ogawa, S. Ohtake, R. Pestotnik, S. Saitoh, T. Seki, T. Sumiyoshi, Y. Uchida, Y. Unno, and S. Yamamoto. A novel type of proximity focusing RICH counter with multiple refractive index aerogel radiator. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 548(3):383–390, August 2005.
- [113] Peter Križan, Samo Korpar, and Toru Iijima. Study of a nonhomogeneous aerogel radiator in a proximity focusing RICH detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 565(2):457–462, September 2006.
- [114] A. Kuzmin. Electromagnetic calorimeter of Belle II. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 958:162235, 2020. Proceedings of the Vienna Conference on Instrumentation 2019.
- [115] Kenkichi Miyabayashi. Belle electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 494(1):298–302, 2002. Proceedings of the 8th International Conference on Instrumentation for Colliding Beam Physics.
- [116] A. Abashian et al. The Belle detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 479(1):117–232, 2002. Detectors for Asymmetric B-factories.
- [117] N. Balamurugan, A. Arulchakkaravarthi, S. Selvakumar, M. Lenin, Rakesh Kumar, S. Muralithar, K. Sivaji, and P. Ramasamy. Growth and characterization of un-

- doped and thallium doped cesium iodide single crystals. *Journal of Crystal Growth*, 286(2):294–299, 2006.
- [118] V. B. Mikhailik, V. Kapustyanyk, V. Tsybulskyi, V. Rudyk, and H. Kraus. Luminescence and scintillation properties of CsI: A potential cryogenic scintillator. *physica status solidi (b)*, 252(4):804–810, January 2015.
- [119] H. Ikeda, A. Satpathy, B.S. Ahn, V.M. Aulchenko, A.E. Bondar, B.G. Cheon, T.H. Chu, M. Fukushima, A.Yu. Garmash, H. Hayashii, H.K. Jang, K. Kaneyuki, K. Kazui, G.Ya. Kezerashvili, H.J. Kim, E.V. Kremyanskaya, A.S. Kuzmin, M.H. Lee, K. Miyabayashi, N.Yu. Muchnoi, A.I. Naumenkov, S. Noguchi, C.S. Park, V.G. Popov, I.Ya. Protopopov, Yu.A. Pogorelov, L.V. Romanov, N.I. Root, H. Sagawa, D.N. Shatilov, A.V. Sidorov, E.A. Simonov, B.A. Shwartz, K. Tamai, T. Tsukamoto, Yu.V. Usov, Y. Watanabe, and V.N. Zhilich. A detailed test of the CsI(Tl) calorimeter for BELLE with photon beams of energy between 20 MeV and 5.4 GeV. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 441(3):401–426, 2000.
- [120] Miles N. Wernick and John N. Aarsvold. *Emission Tomography: The Fundamentals of PET and SPECT*. Elsevier Academic Press, Amsterdam, Boston, 2004.
- [121] Takayuki YANAGIDA. Inorganic scintillating materials and scintillation detectors. *Proceedings of the Japan Academy*, 94(2):75–97, 2018.
- [122] R S Storey, W Jack, and A Ward. The Fluorescent Decay of CsI(Tl) for Particles of Different Ionization Density. *Proceedings of the Physical Society*, 72(1):1, jul 1958.
- [123] W. Skulski and M. Momayezi. Particle identification in csi(tl) using digital pulse shape analysis. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 458(3):759–771, 2001.

- [124] S. Longo and J. M. Roney. Hadronic vs. electromagnetic pulse shape discrimination in csi(tl) for high energy physics experiments. *Journal of Instrumentation*, 13(03):P03018, mar 2018.
- [125] S. Longo, J.M. Roney, C. Cecchi, S. Cunliffe, T. Ferber, H. Hayashii, C. Hearty, A. Hershenhorn, A. Kuzmin, E. Manoni, F. Meier, K. Miyabayashi, I. Nakamura, M. Remnev, A. Sibidanov, Y. Unno, Y. Usov, and V. Zhulanov. CsI(Tl) pulse shape discrimination with the Belle II electromagnetic calorimeter as a novel method to improve particle identification at electron–positron colliders. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 982:164562, 2020.
- [126] John David Valentine. *Temperature dependence of CsI(Tl) gamma-ray excited scintillation characteristics*. PhD thesis, 01 1993.
- [127] V. Aulchenko, A. Bobrov, B.G. Cheon, A. Kuzmin, S. Longo, D. Matvienko, K. Miyabayashi, I. Nakamura, M. Nakao, M. Remnev, J.M. Roney, V. Shebalin, B. Shwartz, Y. Unno, Yu.V. Usov, A. Vinokurova, S. Yamada, and V. Zhulanov. Development of data acquisition system for Belle II electromagnetic calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1030:166468, 2022.
- [128] T. Abe. Belle II technical design report. 2010.
- [129] A. Abashian et al. The KL/ μ detector subsystem for the BELLE experiment at the KEK B-factory. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 449(1):112–124, 2000.
- [130] V.V. Parkhomchuck, Yu.N. Pestov, and N.V. Petrovykh. A spark counter with large area. *Nuclear Instruments and Methods*, 93(2):269–270, 1971.

- [131] R. Cardarelli, R. Santonico, A. Di Biagio, and A. Lucci. Progress in resistive plate counters. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 263(1):20–25, 1988.
- [132] Yu.N. Pestov. Status and future developments of spark counters with a localized discharge. *Nuclear Instruments and Methods in Physics Research*, 196(1):45–47, 1982.
- [133] Georg Steinhauser and Karl Buchtela. *Gas ionization detectors*, pages 245–305. 01 2020.
- [134] L. E. Piilonen. Belle II KLM. Meeting of the APS Division of Particles and Fields, 2015.
- [135] P. Lecoq. *Scintillation Detectors for Charged Particles and Photons*, chapter 3, pages 45–89. Springer, Cham, 2020.
- [136] T. Aushev, D.Z. Besson, K. Chilikin, R. Chistov, M. Danilov, P. Katrenko, R. Mizuk, G. Pakhlova, P. Pakhlov, V. Rusinov, E. Solovieva, E. Tarkovsky, I. Tikhomirov, and T. Uglov. A scintillator based endcap K_L^0 and muon detector for the Belle II experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 789:134–142, 2015.
- [137] C. Ketter et al. Design and Commissioning of Readout Electronics for a K_L^0 and Detector at the Belle II Experiment, 2025.
- [138] S. Bähr, H. Bae, J. Becker, M. Bertemes, M. Campajola, T. Ferber, G. Inguglia, Y. Iwasaki, T. Jülg, C. Kiesling, Y. T. Lai, Y. Liu, A. Knoll, T. Koga, A. Lenz, F. Meggendorfer, H. Nakazawa, M. Neu, J. Schieck, E. Schmidt, J. G. Shiu, S. Skambraks, K. Unger, and J. Yin. The Neural Network First-Level Hardware Track Trigger of the Belle II Experiment, 2024.

- [139] M Nakao. Timing distribution for the Belle II data acquisition system. *Journal of Instrumentation*, 7(01):C01028, jan 2012.
- [140] S. Yamada, S. Y. Suzuki, R. Itoh, K. Nakamura, M. Nakao, T. Konno, and T. Higuchi. Global data processing system with onboard CPU cards for the Belle II experiment. In *2014 19th IEEE-NPSS Real Time Conference*, pages 1–2, 2014.
- [141] Q. D. Zhou, S. Yamada, P. Robbe, D. Charlet, R. Itoh, M. Nakao, S. Y. Suzuki, T. Kunigo, E. Jules, E. Plaige, M. Taurigna, H. Purwar, O. Hartbrich, M. Bessner, K. Nishimura, G. Varner, Y.-T. Lai, T. Higuchi, R. Sugiura, D. Biswas, and P. Kapusta. PCI-Express Based High-Speed Readout for the Belle II DAQ Upgrade. *IEEE Transactions on Nuclear Science*, 68(8):1818–1825, August 2021.
- [142] D. Levit, M. Bessner, D. Biswas, D. Charlet, O. Hartbrich, T. Higuchi, R. Itoh, E. Jules, P. Kapusta, T. Kunigo, Y.-T. Lai, T. S. Lau, M. Nakao, K. Nishimura, S.-H. Park, E. Plaige, H. Purwar, P. Robbe, R. Sugiura, S. Y. Suzuki, M. Taurigna, G. Varner, S. Yamada, and Q.-D. Zhou. Trigger Timing Interface for the Read-Out Upgrade of the Belle II DAQ. *IEEE Transactions on Nuclear Science*, 70(6):941–948, 2023.
- [143] J.P. Cachemiche, P.Y. Duval, F. Hachon, R. Le Gac, and F. Réthoré. The PCIe-based readout system for the LHCb experiment. *Journal of Instrumentation*, 11(02):P02013, feb 2016.
- [144] D. Levit et al. Software-Assisted Event Builder for Belle II Experiment. *IEEE Trans. Nucl. Sci.*, 72(3):287–294, 2025.
- [145] Itoh, Ryosuke, Baur, Anselm, Park, Seokhee, Nakao, Mikihiro, Yamada, Satoru, Suzuki, Soh Y., Kunigo, Takuto, and Levit, Dmytro. Improved HLT Framework for Belle II Experiment. *EPJ Web of Conf.*, 295:02016, 2024.

- [146] Thomas Gesler, Wolfgang Kuhn, Jens Soren Lange, Zhenr'An Liu, David Munchow, Bjorn Spruck, and Jingzhou Zhao. The onsen data reduction system for the belle ii pixel detector. *IEEE Transactions on Nuclear Science*, 62(3):1149–1154, June 2015.
- [147] ROOT Collaboration et al. JSROOT - root data analysis in javascript.
url<https://root.cern>, [**Year of access or version year, please specify**].
- [148] EPICS Documentation. <https://docs.epics-controls.org/en/latest/>.
- [149] systemd Developers. systemd system and service manager, 2024. Accessed: 2024-03-16.
- [150] T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth, and N. Braun. The Belle II Core Software. *Comput. Softw. Big Sci.*, 3(1):1, 2019.
- [151] Thomas Kuhr, Fabrizio Bianchi, Giacomo De Pietro, David Jaffe, Frank Meier, Martin Sevier, and Philipp Urquijo. Publication of the Belle II Software. *EPJ Web Conf.*, 295:05018, 2024.
- [152] D Y Kim, M Ritter, T Bilka, A Bobrov, G Casarosa, K Chilikin, T Ferber, R Godang, I Jaegle, J Kandra, P Kodys, T Kuhr, P Kvasnicka, H Nakayama, L Piilonen, C Pulvermacher, L Santelj, B Schwenker, A Sibidanov, Y Soloviev, M Starič, and T Uglov. The simulation library of the belle ii software system. *Journal of Physics: Conference Series*, 898(4):042043, oct 2017.
- [153] Zihan Wang et al. Muon identification with Deep Neural Network in the Belle II K-Long and Muon detector. 3 2025.
- [154] Nils Lennart Braun. *Combinatorial Kalman Filter and High Level Trigger Reconstruction for the Belle II Experiment*. PhD thesis, KIT, Karlsruhe, 2018.
- [155] Valerio Bertacchi et al. Track finding at Belle II. *Computer Physics Communications*, 259:107610, February 2021.

- [156] J.-F. Krohn, F. Tenchini, P. Urquijo, F. Abudinén, S. Cunliffe, T. Ferber, M. Gelb, J. Gemmler, P. Goldenzweig, T. Keck, I. Komarov, T. Kuhr, L. Ligioi, M. Lubej, F. Meier, F. Metzner, C. Pulvermacher, M. Ritter, U. Tamponi, and A. Zupanc. Global decay chain vertex fitting at belle ii. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 976:164269, 2020.
- [157] Belle II Collaboration. Charged-hadron identification at Belle II, 2025.
- [158] T. Alexopoulos, M. Bachtis, E. Gazis, and G. Tsipolitis. Implementation of the legendre transform for track segment reconstruction in drift tube chambers. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 592(3):456–462, 2008.
- [159] D Funke, T Hauth, V Innocente, G Quast, P Sanders, and D Schieferdecker. Parallel track reconstruction in cms using the cellular automaton approach. *Journal of Physics: Conference Series*, 513(5):052010, jun 2014.
- [160] Nils Braun. *Combinatorial Kalman Filter*, pages 117–174. Springer International Publishing, Cham, 2019.
- [161] F. Didierjean, G. Duchêne, and A. Lopez-Martens. The deterministic annealing filter: A new clustering method for -ray tracking algorithms. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 615(2):188–200, 2010.
- [162] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003.
- [163] Felipe L. Gewers, Gustavo R. Ferreira, Henrique F. De Arruda, Filipi N. Silva, Cesar H.

- Comin, Diego R. Amancio, and Luciano Da F. Costa. Principal component analysis: A natural approach to data exploration. *ACM Computing Surveys*, 54(4):1–34, May 2021.
- [164] L. Breiman, Jerome H. Friedman, Richard A. Olshen, and C. J. Stone. Classification and Regression Trees. 1984.
- [165] Jerome H. Friedman. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5):1189–1232, 2001.
- [166] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794. ACM, August 2016.
- [167] Andrew P. Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145–1159, 1997.
- [168] Mark R. J. Junge and Joseph R. Dettori. ROC Solid: Receiver Operator Characteristic (ROC) Curves as a Foundation for Better Diagnostic Tests. *Global Spine Journal*, 8(4):424–429, 2018. PMID: 29977728.
- [169] Jonas Eschle, Albert Puig Navarro, Rafael Silva Coutinho, and Nicola Serra. zfit: scalable pythonic fitting. 10 2019.
- [170] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), February 2011.
- [171] Lukas Heinrich, Matthew Feickert, and Giordon Stark. pyhf: v0.7.6. <https://github.com/scikit-hep/pyhf/releases/tag/v0.7.6>.

- [172] Lukas Heinrich, Matthew Feickert, Giordon Stark, and Kyle Cranmer. pyhf: pure-python implementation of histfactory statistical models. *Journal of Open Source Software*, 6(58):2823, 2021.
- [173] J. Pivarski et al. iminuit. <https://zenodo.org/records/13923658>.
- [174] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. Technical report, New York U., New York, 2012.
- [175] Eduardo Rodrigues et al. The Scikit HEP Project – overview and prospects. *EPJ Web Conf.*, 245:06028, 2020.
- [176] Belle II Collaboration. A model-agnostic likelihood for the reinterpretation of the $B^+ \rightarrow K^+ \nu \bar{\nu}$ measurement at Belle II, 2025.
- [177] Gary J. Feldman and Robert D. Cousins. Unified approach to the classical statistical analysis of small signals. *Physical Review D*, 57(7):3873–3889, April 1998.
- [178] G. Balossini, C. Bignamini, C.M. Carloni Calame, G. Montagna, O. Nicrosini, and F. Piccinini. Photon pair production at flavour factories with per mille accuracy. *Physics Letters B*, 663(3):209–213, May 2008.
- [179] S. Jadach, B. F. L. Ward, and Z. Was. The precision Monte Carlo event generator KK for two-fermion final states in e^+e^- collisions. *Comput. Phys. Commun.*, 130:260, 2000.
- [180] S. Jadach, W. Płaczek, and B.F.L. Ward. BHWIDE 1.00: $O(\alpha_s^2)$ YFS exponentiated Monte Carlo for Bhabha scattering at wide angles for LEP1/SLC and LEP2. *Physics Letters B*, 390(1–4):298–308, January 1997.

- [181] F.A. Berends, R. Pittau, and R. Kleiss. All electroweak four-fermion processes in electron-positron collisions. *Nuclear Physics B*, 424(2):308–342, 1994.
- [182] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth.*, A462:152, 2001.
- [183] S. Jadach, J. H. Kuhn, and Z. Was. TAUOLA: A library of Monte Carlo programs to simulate decays of polarized tau leptons. *Comput. Phys. Commun.*, 64:275, 1990.
- [184] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An Introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159–177, 2015.
- [185] Elisabetta Barberio, Bob van Eijk, and Zbigniew Was. PHOTOS: A universal Monte Carlo for QED radiative corrections in decays. *Comput. Phys. Commun.*, 66:115, 1991.
- [186] Elisabetta Barberio and Zbigniew Was. PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0. *Comput. Phys. Commun.*, 79:291, 1994.
- [187] J.-F. Krohn et al. Global decay chain vertex fitting at Belle II. *Nucl. Instrum. Meth.*, A976:164269, 2020.
- [188] Giovanni Punzi. Sensitivity of searches for new signals and its optimization. *eConf*, C030908:MODT002, 2003.
- [189] John Gaiser. *Charmonium spectroscopy from radiative decays of the J/ψ and ψ'* . PhD thesis, Stanford University, 1982.
- [190] Tomasz Skwarnicki. *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. PhD thesis, Cracow, INP, 1986.

- [191] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 785–794. ACM, 8 2016.
- [192] Tuning the hyper-parameters of an estimator. https://scikit-learn.org/stable/modules/grid_search.html. Accessed: 2025-02-26.
- [193] Leo Breiman. Bagging Predictors. *Machine Learning*, 24:123–140, 1996.
- [194] M. Oreglia. A Study of the reactions $\psi' \rightarrow \gamma\gamma\psi$. Other thesis, 12 1980.
- [195] K. Kasemir, X. Chen, and E.N Danilova. The Best Ever Alarm System Toolkit, 2009. <https://accelconf.web.cern.ch/icalepcs2009/papers/tua001.pdf>.
- [196] Phoebus Alarm Server. <https://github.com/ControlSystemStudio/phoebus/tree/master/services/alarm-server>.
- [197] Introduction: Everything you need to know about Kafka in 10 minutes. <https://kafka.apache.org/intro>.
- [198] Apache ZooKeeper. <https://zookeeper.apache.org/>.
- [199] J. Carder. How (and Why) To use Kafka With ZooKeeper, 2022. <https://www.openlogic.com/blog/using-kafka-zookeeper>.
- [200] Ansible playbooks. https://docs.ansible.com/ansible/latest/playbook_guide/playbooks_intro.html.
- [201] Highsoft AS. Highcharts: Interactive JavaScript charts for your web pages, 2026. Accessed: 2026-03-30.
- [202] J. Allison et al. Recent Developments in Geant4. *Nucl. Instrum. Meth. A*, 835:186–225, 2016.

- [203] Harold Erbin and Sven Krippendorff. GANs for Generating EFT models. *ArXiv*, abs/1809.02612, 2018.
- [204] Callum R. Brodie, Andrei Constantin, Rehan Deen, and André Lukas. Machine Learning Line Bundle Cohomology. *arXiv: High Energy Physics - Theory*, 2019.
- [205] Tom Rudelius. Learning to inflate. A gradient ascent approach to random inflation. *Journal of Cosmology and Astroparticle Physics*, 2019:044–044, 2019.
- [206] Luke de Oliveira, Michela Paganini, and Benjamin Nachman. Learning Particle Physics by Example: Location-Aware Generative Adversarial Networks for Physics Synthesis. *Computing and Software for Big Science*, 1:1–24, 2017.
- [207] Joshua Yao-Yu Lin, Wahid Bhimji, and Benjamin Nachman. Machine learning templates for QCD factorization in the search for physics beyond the standard model. *Journal of High Energy Physics*, 2019:1–16, 2019.
- [208] Riccardo Di Sipio, Michele Fauci Giannelli, Sana Ketabchi Haghighat, and Serena Palazzo. DijetGAN: a Generative-Adversarial Network approach for the simulation of QCD dijet events at the LHC. *Journal of High Energy Physics*, 2019:110, 2019.
- [209] Luke de Oliveira, Michela Paganini, and Benjamin Nachman. Controlling Physical Attributes in GAN-Accelerated Simulation of Electromagnetic Calorimeters. *ArXiv*, abs/1711.08813, 2017.
- [210] Michela Paganini, Luke de Oliveira, and Benjamin Nachman. CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks. *ArXiv*, abs/1712.10321, 2017.
- [211] Michela Paganini, Luke de Oliveira, and Benjamin Nachman. Accelerating Science

- with Generative Adversarial Networks: An Application to 3D Particle Showers in Multilayer Calorimeters. *Physical review letters*, 120 4:042003, 2018.
- [212] Martin Erdmann, Lukas Geiger, Jonas Glombitza, and David D. Schmidt. Generating and Refining Particle Detector Simulations Using the Wasserstein Distance in Adversarial Networks. *Computing and Software for Big Science*, 2:1–9, 2018.
- [213] Bobak Hashemi, Nick Amin, Kaustuv Datta, Dominick Olivito, and Maurizio Pierini. LHC analysis-specific datasets with Generative Adversarial Networks. *ArXiv*, abs/1901.05282, 2019.
- [214] Sydney Otten, Sascha Caron, Wieske de Swart, Melissa van Beekveld, Luc Hendriks, Caspar van Leeuwen, Damian Podareanu, R. Ruiz de Austri, and Rob Verheyen. Event Generation and Statistical Sampling for Physics with Deep Generative Models and a Density Information Buffer. 2019.
- [215] Matej Srebre, Pascal Schmolz, Hosein Hashemi, Martin Ritter, and Thomas Kuhr. Generation of Belle II Pixel Detector Background Data with a GAN. 2019. CHEP 2019.
- [216] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron C. Courville, and Yoshua Bengio. Generative Adversarial Networks. *ArXiv*, abs/1406.2661, 2014.
- [217] Martín Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein GAN. *ArXiv*, abs/1701.07875, 2017.
- [218] Luke Metz, Ben Poole, David Pfau, and Jascha Sohl-Dickstein. Unrolled Generative Adversarial Networks. *ArXiv*, abs/1611.02163, 2017.

- [219] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 1 edition, 2004.
- [220] Naveen Kodali, Jacob D. Abernethy, James Hays, and Zolt Kira. On Convergence and Stability of GANs. *ArXiv*, abs/1705.07215, 2017.
- [221] Mark A. Peletier. Variational modelling : energies, gradient flows, and large deviations. 2014.
- [222] Ishaan Gulrajani, Faruk Ahmed, Martín Arjovsky, Vincent Dumoulin, and Aaron C. Courville. Improved Training of Wasserstein GANs. *ArXiv*, abs/1704.00028, 2017. https://github.com/eriklindernoren/Keras-GAN/blob/master/wgan_gp/wgan_gp.py.
- [223] Farzan Farnia and David Tse. A Convex Duality Framework for GANs. In *NeurIPS*, 2018.
- [224] Christopher Bishop. *Pattern Recognition and Machine Learning*. Springer-Verlag New York, 1 edition, 2006.
- [225] Xinhua Zhang. A Very Gentle Note on the Construction of Dirichlet Process, September 2008.
- [226] Samuel J. Gershman and David M. Blei. A Tutorial on Bayesian Nonparametric Models. *Journal of Mathematical Psychology*, 56:1–12, 2011.
- [227] Abhijoy Saha and Sebastian Kurtek. Geometric Sensitivity Measures for Bayesian Nonparametric Density Estimation Models. 2019.
- [228] Dilian Görür and Carl Edward Rasmussen. Dirichlet Process Gaussian Mixture Models: Choice of the Base Distribution. *Journal of Computer Science and Technology*, 25(4):653–664, 2009.

- [229] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [230] Sara Sabour, Nicholas Frosst, and Geoffrey E. Hinton. Dynamic Routing Between Capsules. *ArXiv*, abs/1710.09829, 2017.
- [231] S. E. Woosley and Thomas A. Weaver. The Physics of Supernova Explosions. *Annual Review of Astronomy and Astrophysics*, 24(1):205–253, 1986.
- [232] Bernhard Müller. The Core-Collapse Supernova Explosion Mechanism. *Proceedings of the International Astronomical Union*, 12(S329):17–24, Nov 2016.
- [233] A. Burrows and D. Vartanyan. Core-collapse supernova explosion theory. *Nature*, 589(7840):29–39, Jan 2021.
- [234] Demos Kazanas, Rabindra N. Mohapatra, Shmuel Nussinov, Vigdor L. Teplitz, and Yongchao Zhang. Supernova bounds on the dark photon using its electromagnetic decay. *Nuclear Physics B*, 890:17–29, Jan 2015.
- [235] Shun Zhou. Supernova bounds on kev-mass sterile neutrinos. *International Journal of Modern Physics A*, 30(13):1530033, May 2015.
- [236] William DeRocco, Peter W. Graham, Daniel Kasen, Gustavo Marques-Tavares, and Surjeet Rajendran. Observable signatures of dark photons from supernovae. *JHEP*, 02:171, 2019.
- [237] Anna M. Suliga, Irene Tamborra, and Meng-Ru Wu. Lifting the core-collapse supernova bounds on kev-mass sterile neutrinos. *Journal of Cosmology and Astroparticle Physics*, 2020(08):018–018, Aug 2020.

- [238] Anson Hook, Gustavo Marques-Tavares, and Clayton Ristow. Supernova constraints on an axion-photon-dark photon interaction, 2021.
- [239] H A Bethe. Nuclear physics needed for the theory of supernovae. *Annual Review of Nuclear and Particle Science*, 38(1):1–29, 1988.
- [240] K. Langanke and G. Martínez-Pinedo. The role of electron capture in core-collapse supernovae. *Nuclear Physics A*, 928:305–312, 2014. Special Issue Dedicated to the Memory of Gerald E Brown (1926-2013).
- [241] Janka, H.-Th., Müller, B., Kitaura, F. S., and Buras, R. Dynamics of shock propagation and nucleosynthesis conditions in o-ne-mg core supernovae. *A&A*, 485(1):199–208, 2008.
- [242] Rodrigo L. Fernandez. Hydrodynamics of core-collapse supernovae at the transition to explosion. i. spherical symmetry. *The Astrophysical Journal*, 749:142, 2011.
- [243] Thierry Foglizzo, Rémi Kazeroni, Jérôme Guilet, Frédéric Masset, Matthias González, Brendan K. Krueger, Jérôme Novak, Micaela Oertel, Jérôme Margueron, Julien Faure, Noël Martin, Patrick Blottiau, Bruno Peres, and Gilles Durand. The Explosion Mechanism of Core-Collapse Supernovae: Progress in Supernova Theory and Experiments. *PASA*, 32:e009, March 2015.
- [244] David Radice, Ernazar Abdikamalov, Christian D Ott, Philipp Mösta, Sean M Couch, and Luke F Roberts. Turbulence in core-collapse supernovae. *Journal of Physics G: Nuclear and Particle Physics*, 45(5):053003, apr 2018.
- [245] Bernhard Müller. Hydrodynamics of core-collapse supernovae and their progenitors. *Living Reviews in Computational Astrophysics*, 6(1), Jun 2020.

- [246] S. W. Bruenn, K. R. De Nisco, and A. Mezzacappa. General relativistic effects in the core collapse supernova mechanism. *The Astrophysical Journal*, 560(1):326–338, Oct 2001.
- [247] Evan P. O’Connor and Sean M. Couch. Two-dimensional core-collapse supernova explosions aided by general relativity with multidimensional neutrino transport. *The Astrophysical Journal*, 854(1):63, Feb 2018.
- [248] Wakana Iwakami, Kei Kotake, Naofumi Ohnishi, Shoichi Yamada, and Keisuke Sawada. Three-dimensional simulations of standing accretion shock instability in core-collapse supernovae. *The Astrophysical Journal*, 678(2):1207–1222, May 2008.
- [249] Florian Hanke, Andreas Marek, Bernhard Muller, and Hans-Thomas Janka. Is Strong SASI Activity the Key to Successful Neutrino-Driven Supernova Explosions? *Astrophys. J.*, 755:138, 2012.
- [250] Florian Hanke, Bernhard Müller, Annop Wongwathanarat, Andreas Marek, and Hans-Thomas Janka. Sasi activity in three-dimensional neutrino-hydrodynamics simulations of supernova cores. *The Astrophysical Journal*, 770(1):66, May 2013.
- [251] Sean M. Couch, Emmanouil Chatzopoulos, W. David Arnett, and F. X. Timmes. The Three Dimensional Evolution to Core Collapse of a Massive Star. *Astrophys. J. Lett.*, 808(1):L21, 2015.
- [252] Jade Powell and Bernhard Müller. Three-dimensional core-collapse supernova simulations of massive and rotating progenitors. *Mon. Not. Roy. Astron. Soc.*, 494(4):4665–4675, 2020.
- [253] Takashi Yoshida, Tomoya Takiwaki, David R. Aguilera-Dena, Kei Kotake, Koh Takahashi, Ko Nakamura, Hideyuki Umeda, and Norbert Langer. A three-dimensional hy-

- drodynamics simulation of oxygen-shell burning in the final evolution of a fast-rotating massive star. *Mon. Not. Roy. Astron. Soc.*, 506(1):L20–L25, 2021.
- [254] A. Marek and H. Th. Janka. Delayed Neutrino-Driven Supernova Explosions Aided by the Standing Accretion-Shock Instability. *Astrophysical Journal*, 694(1):664–696, March 2009.
- [255] B. Müller. The dynamics of neutrino-driven supernova explosions after shock revival in 2D and 3D. *Monthly Notices of the Royal Astronomical Society*, 453(1):287–310, October 2015.
- [256] Hans-Thomas Janka, Tobias Melson, and Alexander Summa. Physics of Core-Collapse Supernovae in Three Dimensions: A Sneak Preview. *Annual Review of Nuclear and Particle Science*, 66(1):341–375, Oct 2016.
- [257] Kate Scholberg. Supernova Neutrino Detection. *Annual Review of Nuclear and Particle Science*, 62(1):81–103, Nov 2012.
- [258] Francesco Vissani. Comparative analysis of SN1987A antineutrino fluence. *J. Phys. G*, 42:013001, 2015.
- [259] A. Mirizzi, I. Tamborra, H. Janka, N. Saviano, K. Scholberg, R. Bollig, Lorenz Hudepohl, and S. Chakraborty. Supernova Neutrinos: Production, Oscillations and Detection. *arXiv: High Energy Astrophysical Phenomena*, 2015.
- [260] Shunsaku Horiuchi and James P Kneller. What can be learned from a future supernova neutrino detection? *Journal of Physics G: Nuclear and Particle Physics*, 45(4):043002, Mar 2018.
- [261] Kate Scholberg. SNEWS: The supernova early warning system. *AIP Conference Proceedings*, 2000.

- [262] Pietro Antonioli, Richard Tresch Fienberg, Fabrice Fleurot, Yoshiyuki Fukuda, Walter Fulgione, Alec Habig, Jaret Heise, Arthur B McDonald, Corrinne Mills, Toshio Namba, and et al. SNEWS: the SuperNova Early Warning System. *New Journal of Physics*, 6:114–114, Sep 2004.
- [263] S Al Kharusi, S Y BenZvi, J S Bobowski, W Bonivento, V Brdar, T Brunner, E Caden, M Clark, A Coleiro, M Colomer-Molla, and et al. SNEWS 2.0: a next-generation supernova early warning system for multi-messenger astronomy. *New Journal of Physics*, 23(3):031201, Mar 2021.
- [264] A. L. Baxter et al. Agile scrum development in an ad hoc software collaboration, 2021.
- [265] Amanda L. Baxter et al. SNEWPY: A Data Pipeline from Supernova Simulations to Neutrino Signals. 9 2021.
- [266] Ko Nakamura, Shunsaku Horiuchi, Masaomi Tanaka, Kazuhiro Hayama, Tomoya Takiwaki, and Kei Kotake. Multimessenger signals of long-term core-collapse supernova simulations: synergetic observation strategies. *Mon. Not. Roy. Astron. Soc.*, 461(3):3296–3313, 2016.
- [267] MacKenzie L. Warren, Sean M. Couch, Evan P. O’Connor, and Viktoriya Morozova. Constraining properties of the next nearby core-collapse supernova with multimessenger signals. *The Astrophysical Journal*, 898(2):139, Jul 2020.
- [268] Tomoya Takiwaki, Kei Kotake, and Yudai Suwa. A Comparison of Two- and Three-dimensional Neutrino-hydrodynamics simulations of Core-collapse Supernovae. *Astrophys. J.*, 786:83, 2014.
- [269] Bernhard Müller, Tobias Melson, Alexander Heger, and Hans-Thomas Janka. Supernova simulations from a 3D progenitor model – Impact of perturbations and evo-

- lution of explosion properties. *Monthly Notices of the Royal Astronomical Society*, 472(1):491–513, Aug 2017.
- [270] E. J. Lentz, S. W. Bruenn, W. R. Hix, A. Mezzacappa, O. E. B. Messer, E. Endeve, J. M. Blondin, J. A. Harris, P. Marronetti, and K. N. Yakunin. Three-dimensional Core-collapse Supernova Simulated Using a $15 M_{\odot}$ Progenitor. *Astrophys. J.*, 807:L31, July 2015.
- [271] Alexander Summa, Hans-Thomas Janka, Tobias Melson, and Andreas Marek. Rotation-supported Neutrino-driven Supernova Explosions in Three Dimensions and the Critical Luminosity Condition. *The Astrophysical Journal*, 852(1):28, Jan 2018.
- [272] Evan P. O’Connor and Sean M. Couch. Exploring fundamentally three-dimensional phenomena in high-fidelity simulations of core-collapse supernovae. *The Astrophysical Journal*, 865(2):81, Sep 2018.
- [273] K.-C. Pan, M. Liebendörfer, S. M. Couch, and F.-K. Thielemann. Equation of State Dependent Dynamics and Multi-messenger Signals from Stellar-mass Black Hole Formation. *Astrophysical Journal*, 857:13, April 2018.
- [274] Hiroki Nagakura, Adam Burrows, David Radice, and David Vartanyan. Towards an understanding of the resolution dependence of Core-Collapse Supernova simulations. *Monthly Notices of the Royal Astronomical Society*, 490(4):4622–4637, Dec 2019.
- [275] Hiroki Nagakura, Adam Burrows, David Radice, and David Vartanyan. A systematic study of proto-neutron star convection in three-dimensional core-collapse supernova simulations. *Monthly Notices of the Royal Astronomical Society*, 492(4):5764–5779, March 2020.

- [276] Ko Nakamura, Tomoya Takiwaki, and Kei Kotake. Long-term simulations of multi-dimensional core-collapse supernovae: Implications for neutron star kicks. *PASJ*, 71(5):98, October 2019.
- [277] Ko Nakamura, Tomoya Takiwaki, and Kei Kotake. Three-dimensional simulation of a core-collapse supernova for a binary star progenitor of SN 1987A. *arXiv e-prints*, page arXiv:2202.06295, February 2022.
- [278] Tobias Melson, Daniel Kresse, and Hans-Thomas Janka. Resolution Study for Three-dimensional Supernova Simulations with the Prometheus-Vertex Code. *APJ*, 891(1):27, March 2020.
- [279] Robert Glas, H.-Thomas Janka, Tobias Melson, Georg Stockinger, and Oliver Just. Effects of LESA in Three-dimensional Supernova Simulations with Multidimensional and Ray-by-ray-plus Neutrino Transport. *The Astrophysical Journal*, 881(1):36, Aug 2019.
- [280] Bernhard Müller and Vishnu Varma. A 3d simulation of a neutrino-driven supernova explosion aided by convection and magnetic fields. *Monthly Notices of the Royal Astronomical Society: Letters*, 498(1):L109–L113, Aug 2020.
- [281] David Vartanyan, Adam Burrows, and David Radice. Temporal and angular variations of 3D core-collapse supernova emissions and their physical correlations. *Monthly Notices of the Royal Astronomical Society*, 489(2):2227–2246, Aug 2019.
- [282] Jade Powell and Bernhard Müller. Three-dimensional core-collapse supernova simulations of massive and rotating progenitors. *Monthly Notices of the Royal Astronomical Society*, 494(4):4665–4675, Apr 2020.

- [283] Jade Powell, Bernhard Müller, and Alexander Heger. The final core collapse of pulsational pair instability supernovae. *Monthly Notices of the Royal Astronomical Society*, 503(2):2108–2122, Mar 2021.
- [284] Evan O’Connor and Christian D. Ott. Black hole formation in failing core-collapse supernovae. *The Astrophysical Journal*, 730(2):70, Mar 2011.
- [285] Marcella Ugliano, Hans-Thomas Janka, Andreas Marek, and Almudena Arcones. Progenitor-explosion connection and remnant birth masses for neutrino-driven supernovae of iron-core progenitors. *The Astrophysical Journal*, 757(1):69, Sep 2012.
- [286] Tuguldur Sukhbold, T. Ertl, S. E. Woosley, Justin M. Brown, and H.-T. Janka. Core-collapse supernovae from 9 to 120 solar masses based on neutrino-powered explosions. *The Astrophysical Journal*, 821(1):38, Apr 2016.
- [287] A. Perego, M. Hempel, C. Fröhlich, K. Ebinger, M. Eichler, J. Casanova, M. Liebendörfer, and F. K. Thielemann. PUSHing Core-collapse Supernovae to Explosions in Spherical Symmetry I: the Model and the Case of SN 1987A. *Astrophys. J.*, 806(2):275, June 2015.
- [288] Kevin Ebinger, Sanjana Curtis, Carla Fröhlich, Matthias Hempel, Albino Perego, Matthias Liebendörfer, and Friedrich-Karl Thielemann. PUSHing Core-collapse Supernovae to Explosions in Spherical Symmetry. II. Explodability and Remnant Properties. *Astrophys. J.*, 870(1):1, January 2019.
- [289] Sanjana Curtis, Kevin Ebinger, Carla Fröhlich, Matthias Hempel, Albino Perego, Matthias Liebendörfer, and Friedrich-Karl Thielemann. PUSHing Core-collapse Supernovae to Explosions in Spherical Symmetry. III. Nucleosynthesis Yields. *Astrophys. J.*, 870(1):2, January 2019.

- [290] Kevin Ebinger, Sanjana Curtis, Somdutta Ghosh, Carla Fröhlich, Matthias Hempel, Albino Perego, Matthias Liebendörfer, and Friedrich-Karl Thielemann. PUSHing Core-collapse Supernovae to Explosions in Spherical Symmetry. IV. Explodability, Remnant Properties, and Nucleosynthesis Yields of Low-metallicity Stars. *Astrophys. J.*, 888(2):91, January 2020.
- [291] Sanjana Curtis, Noah Wolfe, Carla Fröhlich, Jonah M. Miller, Ryan Wollaeger, and Kevin Ebinger. Core-collapse Supernovae: From Neutrino-driven 1D Explosions to Light Curves and Spectra. *Astrophys. J.*, 921(2):143, November 2021.
- [292] Somdutta Ghosh, Noah Wolfe, and Carla Fröhlich. PUSHing Core-collapse Supernovae to Explosions in Spherical Symmetry. V. Equation of State Dependency of Explosion Properties, Nucleosynthesis Yields, and Compact Remnants. *Astrophys. J.*, 929(1):43, April 2022.
- [293] Luca Boccioli, Grant J. Mathews, In-Saeng Suh, and Evan P. O’Connor. Effect of the Nuclear Equation of State on Relativistic Turbulence-induced Core-collapse Supernovae. *Astrophys. J.*, 926(2):147, February 2022.
- [294] Bernhard Müller, Alexander Heger, David Liptai, and Joshua B. Cameron. A simple approach to the supernova progenitor–explosion connection. *Monthly Notices of the Royal Astronomical Society*, 460(1):742–764, 05 2016.
- [295] David R. Aguilera-Dena, Norbert Langer, John Antoniadis, and Bernhard Müller. Pre-collapse Properties of Superluminous Supernovae and Long Gamma-Ray Burst Progenitor Models. *The Astrophysical Journal*, 901(2):114, October 2020.
- [296] F. R. N. Schneider, Ph. Podsiadlowski, and B. Müller. Pre-supernova evolution, compact object masses and explosion properties of stripped binary stars. *Astron. Astrophys.*, 645:A5, 2021.

- [297] Yudai Suwa, Akira Harada, Ken'ichiro Nakazato, and Kohsuke Sumiyoshi. Analytic solutions for neutrino-light curves of core-collapse supernovae, 2020.
- [298] S. E. Woosley and Alexander Heger. ERRATUM: “LONG GAMMA-RAY TRANSIENTS FROM COLLAPSARS” (2012, *Astrophys. J.*, 752, 32). *Astrophys. J.*, 806(1):145, jun 2015.
- [299] Bernhard Müller, Hans-Thomas Janka, and Harald Dimmelmeyer. A new multi-dimensional general relativistic neutrino hydrodynamic code for core-collapse supernova. i. method and code tests in spherical symmetry. *The Astrophysical Journal Supplement Series*, 189(1):104–133, Jun 2010.
- [300] B. Müller and H.-Th. Janka. Non-radial instabilities and progenitor asphericities in core-collapse supernovae. *Monthly Notices of the Royal Astronomical Society*, 448(3):2141–2174, 02 2015.
- [301] T. Fischer, S. C. Whitehouse, A. Mezzacappa, F.-K. Thielemann, and M. Liebendörfer. The neutrino signal from protoneutron star accretion and black hole formation. *Astronomy & Astrophysics*, 499(1):1–15, Mar 2009.
- [302] Bernhard Müller and Hans-Thomas Janka. A NEW MULTI-DIMENSIONAL GENERAL RELATIVISTIC NEUTRINO HYDRODYNAMICS CODE FOR CORE-COLLAPSE SUPERNOVAE. IV. THE NEUTRINO SIGNAL. *Astrophys. J.*, 788(1):82, may 2014.
- [303] James M. Lattimer and A. Yahil. Analysis of the Neutrino Events from Supernova 1987A. *Astrophysical Journal*, 340:426, May 1989.
- [304] J. M. Lattimer and M. Prakash. Neutron star structure and the equation of state. *The Astrophysical Journal*, 550(1):426–442, mar 2001.

- [305] L. Hüpdepohl. Neutrinos from the formation, cooling, and black hole collapse of neutron stars. 2014.
- [306] H. Thomas Janka. Conditions for shock revival by neutrino heating in core collapse supernovae. *Astron. Astrophys.*, 2001. [Astron. Astrophys.368,527(2001)].
- [307] Jeremiah W. Murphy, Adam Burrows Steward Observatory, University of Arizona, Astronomy Department, University of Washington, Department of Astrophysical Sciences, and Princeton University. Criteria for core-collapse supernova explosions by the neutrino mechanism. *The Astrophysical Journal*, 688:1159–1175, 2008.
- [308] David Radice, Ernazar Abdikamalov, Christian D. Ott, Philipp Mösta, Sean M. Couch, and Luke F. Roberts. Turbulence in core-collapse supernovae. *Journal of Physics G Nuclear Physics*, 45(5):053003, May 2018.
- [309] David Vartanyan, Matthew S. B. Coleman, and Adam Burrows. The collapse and three-dimensional explosion of three-dimensional massive-star supernova progenitor models. *MNRAS*, 510(4):4689–4705, March 2022.
- [310] Hans-Thomas Janka. Explosion mechanisms of core-collapse supernovae. *Annual Review of Nuclear and Particle Science*, 62(1):407–451, 2012.
- [311] Alexander Summa, Florian Hanke, Hans-Thomas Janka, Tobias Melson, Andreas Marek, and Bernhard Müller. Progenitor-dependent explosion dynamics in self-consistent, axisymmetric simulations of neutrino-driven core-collapse supernovae. *The Astrophysical Journal*, 825(1):6, Jun 2016.
- [312] Kitaura, F. S., Janka, H.-Th., and Hillebrandt, W. Explosions of o-ne-mg cores, the crab supernova, and subluminescent type ii-p supernovae. *A&A*, 450(1):345–350, 2006.

- [313] T. Fischer, S. C. Whitehouse, A. Mezzacappa, F. K. Thielemann, and M. Liebendörfer. Protoneutron star evolution and the neutrino-driven wind in general relativistic neutrino radiation hydrodynamics simulations. *Astronomy and Astrophysics*, 517:A80, July 2010.
- [314] Tobias Melson, Hans-Thomas Janka, and Andreas Marek. Neutrino-driven Supernova of a Low-mass Iron-core Progenitor Boosted by Three-dimensional Turbulent Convection. *Astrophys. J.*, 801(2):L24, March 2015.
- [315] A. Marek and H. Th. Janka. Delayed neutrino-driven supernova explosions aided by the standing accretion-shock instability. *Astrophys. J.*, 694:664–696, 2009.
- [316] Christopher D. Matzner and Christopher F. McKee. The Expulsion of Stellar Envelopes in Core-Collapse Supernovae. *Astrophysical Journal*, 510(1):379–403, January 1999.
- [317] Leonhard Scheck, K. Kifonidis, H. Th. Janka, and E. Mueller. Multidimensional supernova simulations with approximative neutrino transport. 1. neutron star kicks and the anisotropy of neutrino-driven explosions in two spatial dimensions. *Astron. Astrophys.*, 457:963, 2006.
- [318] Bernhard Müller, Hans-Thomas Janka, and Andreas Marek. A New Multi-dimensional General Relativistic Neutrino Hydrodynamics Code for Core-collapse Supernovae. II. Relativistic Explosion Models of Core-collapse Supernovae. *Astrophysical Journal*, 756(1):84, September 2012.
- [319] S. E. Woosley, A. Heger, and T. A. Weaver. The evolution and explosion of massive stars. *Rev. Mod. Phys.*, 74:1015–1071, Nov 2002.
- [320] Markus Rampp and H. Thomas Janka. Radiation hydrodynamics with neutrinos:

- Variable Eddington factor method for core collapse supernova simulations. *Astron. Astrophys.*, 396:361, 2002.
- [321] Stuart L. Shapiro and Saul A. Teukolsky. Black holes, white dwarfs, and neutron stars : the physics of compact objects, 1983.
- [322] G.G. Raffelt. Stars as laboratories for fundamental physics: The astrophysics of neutrinos, axions, and other weakly interacting particles, 5 1996.
- [323] Tobias Fischer, Meng-Ru Wu, Benjamin Wehmeyer, Niels-Uwe F. Bastian, Gabriel Martínez-Pinedo, and Friedrich-Karl Thielemann. Core-collapse supernova explosions driven by the hadron-quark phase transition as rare r process site. 3 2020.
- [324] Kei Kotake, Tomoya Takiwaki, Tobias Fischer, Ko Nakamura, and Gabriel Martínez-Pinedo. Impact of neutrino opacities in core-collapse supernova simulations. 2018.
- [325] Hirokazu Sasaki, Tomoya Takiwaki, Shio Kawagoe, Shunsaku Horiuchi, and Koji Ishidoshiro. Detectability of collective neutrino oscillation signatures in the supernova explosion of a $8.8M_{\odot}$ star. *Physical Review D*, 101(6), Mar 2020.
- [326] Masamichi Zaizen, John F. Cherry, Tomoya Takiwaki, Shunsaku Horiuchi, Kei Kotake, Hideyuki Umeda, and Takashi Yoshida. Neutrino halo effect on collective neutrino oscillation in iron core-collapse supernova model of a $9.6M_{\odot}$ star. *Journal of Cosmology and Astroparticle Physics*, 2020(06):011–011, Jun 2020.
- [327] S.E. Woosley and A. Heger. Nucleosynthesis and remnants in massive stars of solar metallicity. *Physics Reports*, 442(1):269 – 283, 2007. The Hans Bethe Centennial Volume 1906-2006.
- [328] Ken'ichiro Nakazato, Kohsuke Sumiyoshi, Hideyuki Suzuki, Tomonori Totani, Hideyuki Umeda, and Shoichi Yamada. SUPERNOVA NEUTRINO LIGHT CURVES

- AND SPECTRA FOR VARIOUS PROGENITOR STARS: FROM CORE COLLAPSE TO PROTO-NEUTRON STAR COOLING. *The Astrophysical Journal Supplement Series*, 205(1):2, feb 2013.
- [329] M. Kachelrieß, R. Tomàs, R. Buras, H.-Th. Janka, A. Marek, and M. Rampp. Exploiting the neutronization burst of a galactic supernova. *Physical Review D*, 71(6), Mar 2005.
- [330] Evan O’Connor and Christian D. Ott. THE PROGENITOR DEPENDENCE OF THE PRE-EXPLOSION NEUTRINO EMISSION IN CORE-COLLAPSE SUPERNOVAE. *The Astrophysical Journal*, 762(2):126, dec 2012.
- [331] Tom Fischer, Gabriel Martínez-Pinedo, Maja Hempel, Lutz Huther, Gerd Ropke, Stefan Typel, and Andreas Lohs. Expected impact from weak reactions with light nuclei in corecollapse supernova simulations. *arXiv: High Energy Astrophysical Phenomena*, 109:06002, 2016.
- [332] H. Th. Janka. Neutrino emission from supernovae. *arXiv: High Energy Astrophysical Phenomena*, pages 1575–1604, 2017.
- [333] André Sieverding, Gabriel Martínez-Pinedo, Karlheinz Langanke, Robert Bollig, Hans-Thomas Janka, and Alex Heger. The ν -process with fully time-dependent supernova neutrino emission spectra. *The Astrophysical Journal*, 876:151, 2019.
- [334] Yudai Suwa, Kohsuke Sumiyoshi, Ken’ichiro Nakazato, Yasufumi Takahira, Yusuke Koshio, Masamitsu Mori, and Roger A. Wendell. Observing Supernova Neutrino Light Curves with Super-Kamiokande: Expected Event Number over 10 s. *The Astrophysical Journal*, 881(2):139, Aug 2019.
- [335] Tomoya Takiwaki, Kei Kotake, and Yudai Suwa. Three-dimensional simulations of

- rapidly rotating core-collapse supernovae: finding a neutrino-powered explosion aided by non-axisymmetric flows. *Mon. Not. Roy. Astron. Soc.*, 461(1):L112–L116, 2016.
- [336] M. Liebendoerfer, S. C. Whitehouse, and T. Fischer. The isotropic diffusion source approximation for supernova neutrino transport. *Astrophys. J.*, 698:1174–1190, 2009.
- [337] Evan O’Connor et al. Global Comparison of Core-Collapse Supernova Simulations in Spherical Symmetry. *J. Phys. G*, 45(10):104001, 2018.
- [338] H. Togashi, K. Nakazato, Y. Takehara, S. Yamamuro, H. Suzuki, and M. Takano. Nuclear equation of state for core-collapse supernova simulations with realistic nuclear forces. *Nucl. Phys. A*, 961:78–105, 2017.
- [339] Hajime Sotani and Tomoya Takiwaki. Dimension dependence of numerical simulations on gravitational waves from protoneutron stars. *Phys. Rev. D*, 102(2):023028, 2020.
- [340] Hajime Sotani and Tomoya Takiwaki. Avoided crossing in gravitational wave spectra from protoneutron star. *Mon. Not. Roy. Astron. Soc.*, 498(3):3503–3512, 2020.
- [341] Ko Nakamura, Tomoya Takiwaki, Takami Kuroda, and Kei Kotake. Systematic features of axisymmetric neutrino-driven core-collapse supernova models in multiple progenitors. *Publications of the Astronomical Society of Japan*, 67(6):107, Oct 2015.
- [342] Hiroki Nagakura and David Vartanyan. Efficient method for estimating the time evolution of the proto-neutron star mass and radius from a supernova neutrino signal. *Monthly Notices of the Royal Astronomical Society*, 512(2):2806–2816, May 2022.
- [343] Masamitsu Mori, Yudai Suwa, Ken’ichiro Nakazato, Kohsuke Sumiyoshi, Masayuki Harada, Akira Harada, Yusuke Koshio, and Roger A. Wendell. Developing an end-to-end simulation framework of supernova neutrino detection. *PTEP*, 2021(2):023E01, 2021.

Appendices

Appendix A

Belle II Alarm Server

In this section, I discuss my activities at KEK under the Ozaki Exchange Program, including the development and maintenance of Belle II’s alarm and monitoring system. The work was conducted in collaboration with T. Kunigo, M. Nakao, J. Kim, S. Park, C.H. Kim, and L. Piilonen.

The first major part of the alarm system is the framework itself, which I set up and maintained. Since Belle II and SuperKEKB utilize **EPICS** [148] as a major component of their network framework, we adopted the alarm server functionalities from **PHOEBUS** [195, 196] to streamline the processing of detector-monitored values, which come via **EPICS** process variables (PV) from input/output controllers (IOCs). The three major components of the framework are **KAFKA**, **ZOOKEEPER**, and the alarm server.

KAFKA [197] is an event streaming platform that reads events, processes them into a standard format, and broadcasts them to some client (here, the alarm server). An illustration of a **KAFKA** cluster is illustrated in Fig. A.1. Each ‘**KAFKA** node’ sorts events into topics – those related to data acquisition (DAQ), data quality monitoring (DQM), or detector control and safety (DCS) – and places them in a partition. If multiple nodes have a shared topic, the **KAFKA** cluster replicates the processed event across nodes as a form of fault tolerance.

ZOOKEEPER [198] synchronizes, names, and configures the **KAFKA** cluster. During the initialization of a **KAFKA** cluster, **ZOOKEEPER** maintains the list of topics and determines how a

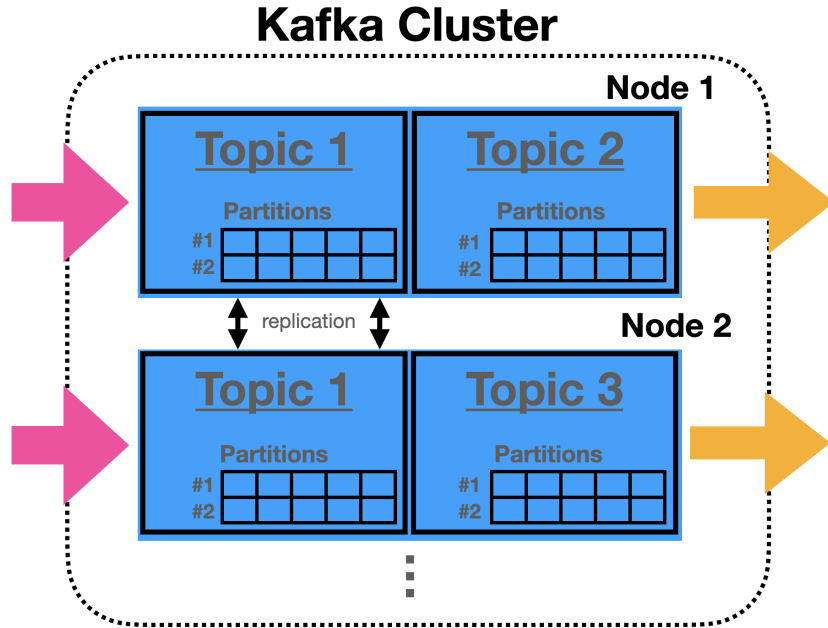


Figure A.1: Example diagram of a **KAFKA** cluster

topic's partitions are distributed across the cluster. **ZOOKEEPER** also manages the entry/exit of a **KAFKA** node and maintains the timestamps used while **KAFKA** processes events [199]. Similarly to **KAFKA**, a cluster of **ZOOKEEPER** nodes can be used for fault tolerance.

The specific layout used for Belle II's alarm system is illustrated in Fig. A.2. Subsystems' IOCs distribute detector information via **EPICS** PVs on Belle II's DAQ network; changes in the status of the **EPICS** PVs are processed by **KAFKA** as the 'events.' For fault tolerance, we set up a three-node system, where a **KAFKA** and **ZOOKEEPER** service are set up on a virtual machine. After extensive testing and parameter tuning (memory allocation, swap memory settings, number of partitions, network management, etc.), we established a stable framework that allows Belle II to monitor our detector with minimal downtime: months-long uninterrupted operation without intervention. **KAFKA** and **ZOOKEEPER** services are run via Linux's **SYSTEMD** [149].

Remaining improvements for the framework involve using server-grade PCs for the host ma-

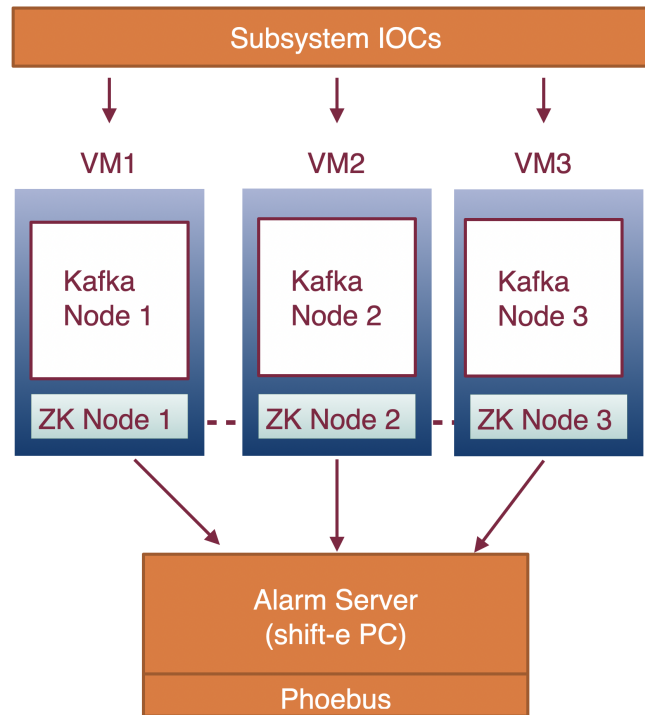


Figure A.2: Layout of alarm framework used at Belle II

chines, using an Ansible playbook [200] for an automated way to set up a **KAFKA-ZOOKEEPER** node in case of hard restarts, and better automated methods to maintain the system.

The second major part is the alarm server and how PVs are configured and represented. Across the three major topics (DCS, DAQ, DQM), we organize the PVs by their relevant subsystem. An example of said organization is depicted in Fig. A.3, which shows how a shifter would access and visualize CDC's DAQ and DCS status. My initial responsibility was the inclusion and management of the following subsystems: CDC, KLM, PXD, and SVD. Since PXD and SVD have their own alarm servers and were thus easy to include in the global alarm server, I focus here on my contributions to KLM and CDC.

The CDC DCS monitors are beam background via leak current ratios, gas-related monitors (pressure, humidity, and composition), temperature of electronics, and HV monitors. While the CDC group maintains the IOCs, their PVs were not equipped with any alarm conditions

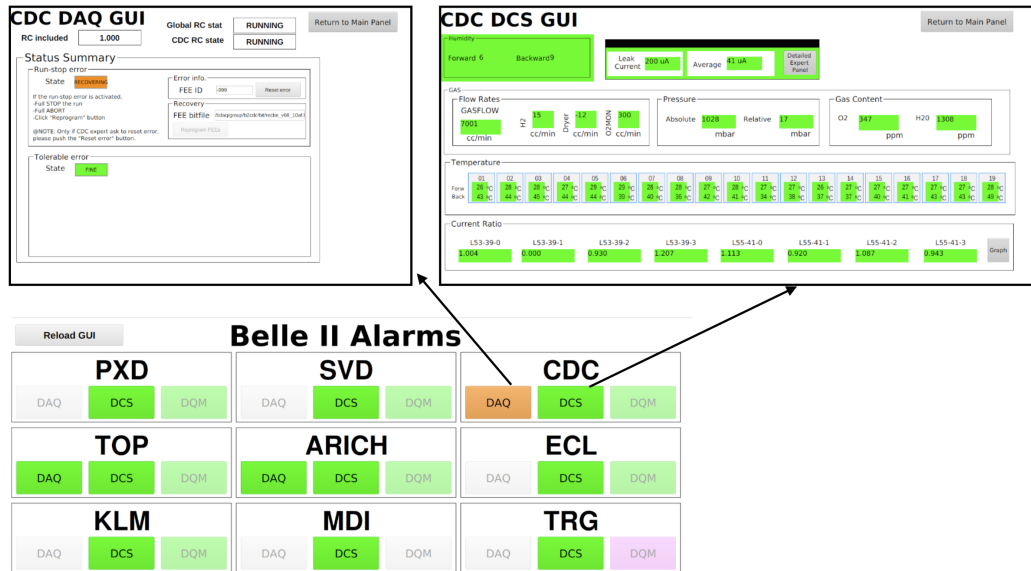


Figure A.3: Main panel GUI (bottom) and CDC GUI components (top)

or functionalities. I worked with the CDC experts to either define new alarm thresholds for PVs or take existing PVs and add said functionalities. I also designed the CDC DCS graphical user interface (GUI) displayed in Fig. A.3, which summarizes the aforementioned DCS monitors to make it comprehensible to control room (CR) shifters and useful to experts. The management of this GUI is now under the CDC group, though I still provide technical support when needed.

KLM's DCS monitors the background hit rates via dark current, input gas flow of the resistive plate chambers (RPCs) system, and the gas flow rate of a given RPC plane via KLM's bubbler system. While many of the PVs were already present, I updated them to include alarm conditions via percent/absolute differences and calibration offsets. I also designed the KLM GUI to summarize the aforementioned alarms, as seen in Fig. A.4. As a member of the KLM group through my affiliation with Virginia Tech, I continue to oversee its maintenance and development.

In addition to the CDC and KLM work described above, I also accomplished the following

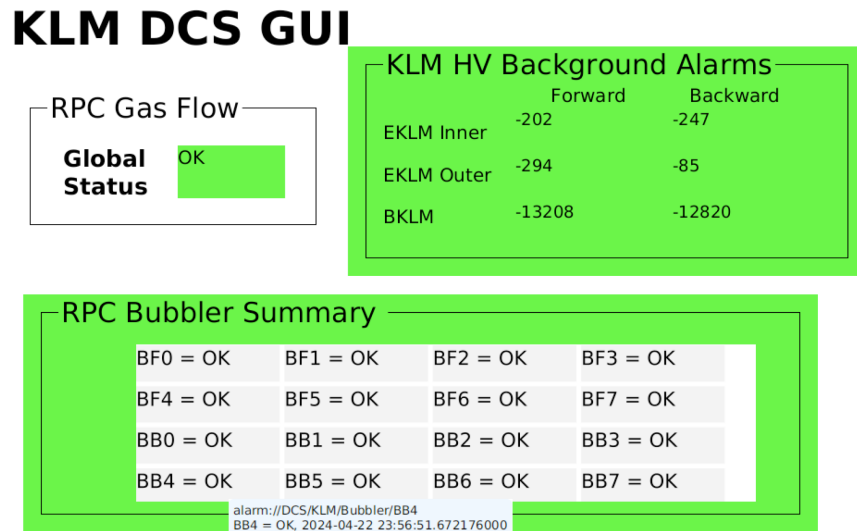


Figure A.4: KLM Detector Safety GUI

during my tenure at KEK: work with the DQM developers to seamlessly incorporate DQM PVs into the Belle II alarm system, develop widgets/templates used in several subsystem GUIs, serve as the point of contact in case of trouble with the alarm server, write documentation for DAQ shifters to eventually take over the system's maintenance throughout future run periods, include additional monitors from the accelerator group, and provide feedback to the other alarm monitoring members while taking CR shifts.

While we were able to deliver a stable alarm system, there are still several foreseeable improvements that I've started investigating. First, in order to deploy the system without lag, I had to remove half of the ARICH's PVs and change the GUI initial update rate. Nevertheless, additional ARICH monitors will be needed in the future, so we need to resolve these lag and capacity problems. The SVD group suggested we remove some hard-coded values in **PHOEBUS** parameters. Additionally, there was a request for **PHOEBUS** to use the computer's default voice rather than a hard-coded JSON voice file; this requires a local version of **PHOEBUS** that needs to be compiled and built manually.

In addition to the above, I list here other minor requests made throughout my tenure: additional beam loss monitors for CR shifters; finer granularity for the RPC bubbler system; new KLM environmental monitors when they become available; and improved CR response to alarms via guidance features and automated alerts.

Appendix B

Technical Details of KLM's Data Quality Monitoring

In this section, I elaborate on the details outlined in Ch. 2.12, including features developed and implemented that enhance KLM's monitoring capabilities. As Belle II is at the forefront of the luminosity frontier, it is imperative that we uncover and resolve problems as quickly as possible. Much of the work in this section was conducted in collaboration with N. Baghel, B. Spurk, L. Santelj, M. Garcia-Hernandez, M. Veronesi, C. Ketter, H. Kindo, and L. Piilonen.

B.1 Motivation

As we developed a more sophisticated monitor of the K_L^0 and muon detector (KLM) (see Sec. B.3.2), we quickly discovered that the backward upper octant of the barrel KLM (BB2) had a drastically lower efficiency than the rest of the detector. After extensive investigations and cross-checking against Belle II's offline calibrations, whose results are shown in Fig. B.1, we found that something had been plaguing BB2 since the summer of 2021. Experts at KEK soon discovered that the gas flow had ceased and that melted ice flowed from the solenoid's chimney and leaked into the chamber. Along with two years of remediation efforts, there was also a strong push to ensure problems of similar magnitudes will not go unnoticed.

In this section, we build upon the information outlined in Sec. 2.12 and 2.13.1 by describing

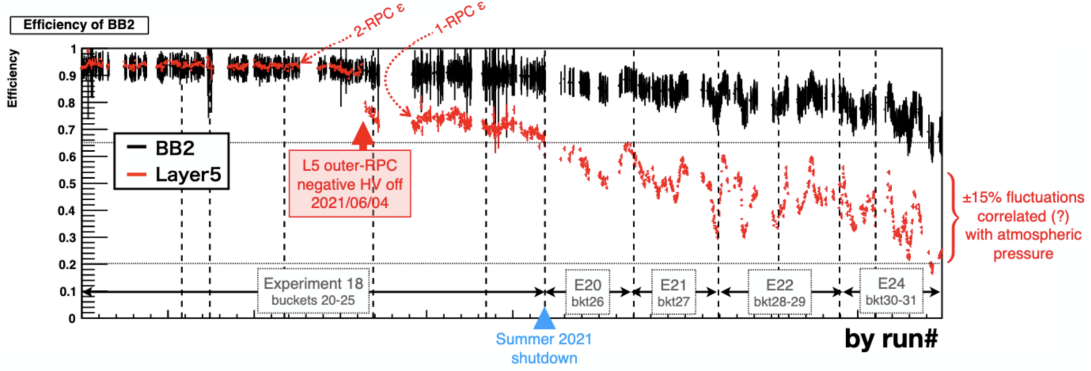


Figure B.1: The efficiency of the upper backward octant of the barrel KLM as a function of run number.

the efforts undertaken by the KLM and DQM developers, along with the motivation and use cases considered; this body of work is provided so that future experiments may enhance their data quality monitoring capabilities by drawing on the experience of Belle II. We begin by reviewing the Belle II's overall DQM framework before focusing on how KLM exploits it.

B.2 Framework Features

To facilitate dynamic data processing and analysis, the DQM server leverages the modular, C++-based architecture of `BASF2`. The monitoring framework executes a sequential pipeline of dedicated `BASF2` analysis modules during each server update cycle. This modularity empowers individual subsystems, such as the KLM, to develop and deploy highly tailored data-processing algorithms (e.g., custom efficiency calculations, background hit-rate estimates) without risking interference with other subsystems' operations. Consequently, this decoupled, `BASF2`-integrated architecture provides the flexible software foundation necessary to implement experiment-wide diagnostic features detailed below and illustrated in Fig. B.2.

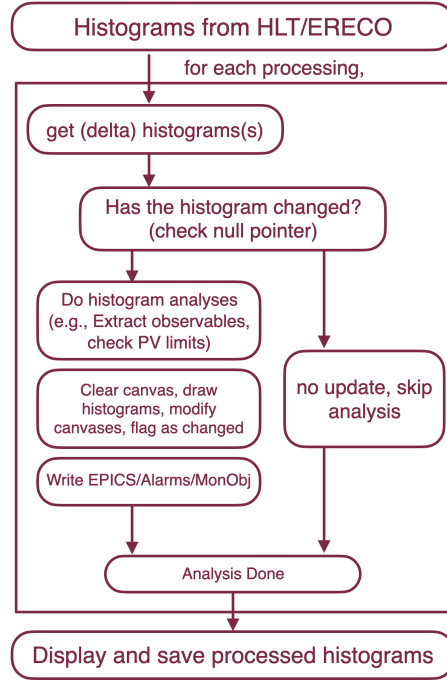


Figure B.2: Flow diagram of the DQM histogram analysis framework.

B.2.1 Comparisons with References

As the state of each subsystem can vary significantly, even within a given run period, we need to clearly communicate to shifters the expected features to minimize false alarms and prevent the waste of crucial beam time. This is achieved by having each subsystem select a set of reference histograms, which are overlaid on the collected real-time histograms, with each reference histogram scaled to match the real-time histogram's normalization. To account for the different physics expected in collision, cosmic, and single-beam runs, different references that are run-type dependent are also incorporated.

In the past, subsystem experts manually compiled histograms in a **Root** file and distributed them to the DQM librarians, who merged them all together. While this is acceptable for basic monitoring (e.g., detector occupancies, timing), these references were inaccessible for direct comparisons for more sophisticated quantities. The necessity of manual interventions

across multiple parties also introduced numerous points of failure that often disoriented shifters and introduced unnecessary complications. In addition, the process of extracting histograms from the DQM server, synthesizing them, and reintroducing them to said server introduces opportunities for unnecessary human error.

The DQM framework was overhauled to address the aforementioned concerns and deficiencies. Given the list of reference runs for each subsystem that should be used for collision/-cosmic/beam/other run-types, daemons pull the subsystem’s histograms from the stored run on the DQM server and compile them to a master reference **Root** file. Subsystem experts provide the aforementioned list of runs by modifying a JSON configuration file, which is checked at the start of every run. As this process is now automated and standardized, it is possible to construct a dedicated module to map and store reference histograms within the DQM analysis framework, allowing subsystems to reliably call upon them.

B.2.2 Delta Histograms

One inconvenience for DQM with Belle II’s current data collection scheme is that problems can be obscured during long, stable runs ($\lesssim 8$ hours) because the **integrated** histograms are processed and displayed by the DAQ and DQM framework. If there are potential silent detector complications (e.g., power supplies suddenly turning off, electronic problems, or data corruption) that occur in the middle of the run, it would take considerable statistics or other issues to compound the problem and require a restart before said problem is discovered. Depending on the severity of the complication, this could force us to reprocess or discard quality data, which would be a waste of resources either way.

This inconvenience was addressed by the creation of ‘delta histograms.’ While the data-collecting framework manages and processes the full set of integrated histograms, the DQM

framework can temporarily save histograms h_i at time t_i . At time t_{i+1} , where $\Delta t_i = t_{i+1} - t_i \sim 10$ minutes, $\Delta h_i = h_{i+1} - h_i$ provides a reasonable proxy of the detector's performance within the last Δt_i period¹.

An important consideration is the choice of criteria for updating these delta histograms. While time is often intuitive, it is a suboptimal metric for us due to synchronization problems, making it tricky to implement. Simpler metrics that are used in its place are the number of entries and the number of events processed². The update condition for variable N_e and some threshold N_{th} is as follows: if $N_e > N_{th}$ from $h_{i+1} - h_i$, update the stored Δh_i and h_i ; else, keep the stored Δh_i and h_i and continue to wait until the next DQM server update.

B.2.3 MiraBelle

It is often desirable to illustrate how detector quantities vary across a given run period or experiment, whether it is to determine how detector performance depends on accelerator conditions or identify problems spanning across dozens of runs (ex. Fig. B.1).

The information stored by MiraBelle is computed at the end of the run, where all the stored histograms are processed to extract quantities of interest and store them in a **Root** file, along with the experiment and run number from which the histograms came. A dedicated daemon on the DQM server then broadcasts the collected information and displays it using **HIGHCHARTS** [201]. Information from Belle II's run database (rundb.belle2.org) is used to extract additional details about said run; this can be used to filter out short or anomalous runs.

¹Note that the minimum time difference is set by how often the DQM server updates, not the rate of e^+e^- collisions.

²A histogram's underflow/overflow can also be utilized as a proxy for the number of events, but only one subsystem utilizes this functionality.

B.2.4 EPICS Integration and Alarming

As the DAQ and SuperKEKB rely on **EPICS**, it would be beneficial to use **EPICS** functionality to assist with monitoring and changing variables/settings throughout a run. Additional details about how **EPICS** is used by the DAQ group are outlined in Sec. [A](#).

Subsystems can distribute alarms or quantities extracted from histograms (e.g., averages, medians, variances) via **EPICS** PVs and either explicitly define alarm conditions via enumerated states, or set alarm thresholds that the histogram-extracted PVs should not exceed. The DQM server takes the list of PVs mid-run and broadcasts it to the Belle II-wide **EPICS** network, where it is accessible to subsystem/SuperKEKB experts. For instance, the $\Upsilon(4S)$ mass from a pair of $\ell^+\ell^-$ from the IP is monitored by the accelerator group to confirm they are operating at the desired center-of-mass energy for the colliding beams.

To help tune alarm conditions mid-operation, subsystems can also request that PVs be modifiable, with the thresholds or the PV values themselves used to modify alarm conditions. The expected use case is to quickly adjust alarm settings and visual indicators (e.g., the histogram canvas' color) when subsystem conditions vary and require extended periods to properly address (e.g., modifying high-voltage settings or masking detector components). While the C++ interface to **EPICS** is straightforward, it would be nontrivial to modify histogram and **Root** settings mid-run. Thus, additional tools were developed that check JSON files at the start of every run, update the PVs settings for the alarm server and histogram analysis modules, and broadcast the PVs themselves.

B.3 KLM Monitored Quantities

The performance and reliability of the K_L^0 and muon detector (KLM) are critical for ensuring high-quality physics data at the luminosity frontier. This work presents the development and deployment of a comprehensive data quality monitoring (DQM) framework tailored to the KLM subsystem. We describe key monitoring observables designed to provide robust and interpretable diagnostics under varying accelerator conditions. We also outline how the features outlined in Sec. B.2 are utilized to reduce false-positive rates, enhance our anomaly-detection capabilities, and promptly respond to detector issues (e.g., mid-run high-voltage trips, single-event upsets, and prolonged gas-system inefficiencies).

B.3.1 Occupancies

The base element of the KLM is a strip, whose timing and hit information are stored in a data structure dubbed `KLMDIGIT`. The simplest monitor available then is the number of hits, or digits, of a given strip. Along with the rudimentary check of ensuring ensembles of strips have no recorded hits (typically suggesting an issue with high voltage), we can also look at the distribution of hits across the detector, where deviations with respect to some reference can be a strong indicator that the front-end electronics components have problems (e.g., single event upsets). The aggregate timing distributions can also provide meaningful diagnosis of KLM's firmware performance, so this is also monitored.

While the aforementioned scheme is straightforward to implement, there are a few pitfalls. The biggest pitfall is that the reference strip hit distribution changes drastically with background conditions, as backgrounds (e.g., low-angle Bhabhas scattering off detector materials, fast neutrons) in the forward endcaps and inner layers of the barrel vary with beam currents. In addition, requesting KLM shifters to go through and carefully check the large number of

strips while minimizing downtime is nontrivial. Because of this, we elect to use muon-like tracks to perform efficiency-based monitoring.

B.3.2 Layer Efficiencies

For the intuition of how the tracking is utilized to assess the performance of a given layer, refer to Ch. 2.13.1. To summarize, we extrapolate tracks from the CDC and VXD to the KLM and determine a layer's efficiency if there exists a KLM hit within the vicinity of where we would expect the extrapolated hit to be.

To ensure we use genuine $\mu^\pm$ tracks, as introducing energetic $\pi^\pm$ would unnecessarily deflate the monitored values, we apply the following criteria:

- Only accept hits as matched if they are 8 strip widths away from the extrapolated hit
- Only accept tracks that contain at least 14 matched hits (effectively, ~ 7 layers were active)
- For monitoring during collisions, we only accept hits that pass the **MUMU-TIGHT** criteria
 $(E_{\text{track } i}^{ECL} < 0.5 \text{ GeV}, p_{\text{track } i} > 0.5 \text{ GeV}, E_{\mu\mu} < 2 \text{ GeV}, \text{ number of tracks} = 2, |\Delta\phi_{\mu^+\mu^-}^{p_{cm}} - 180| < 10^\circ, |(\theta_{\text{track } 1}^{p_{cm}} + \theta_{\text{track } 2}^{p_{cm}}) - 180| < 10^\circ)$
- For monitoring during cosmic ray tests, we accept any cosmic trigger.

In addition, we consider criteria for whether extrapolated/matched hits should be included in our efficiency calculation of a given layer. The said criteria are:

- If there are 4+ hits in layers higher than the layer in question, we accept the extrapolated and potentially matching hit for sure.

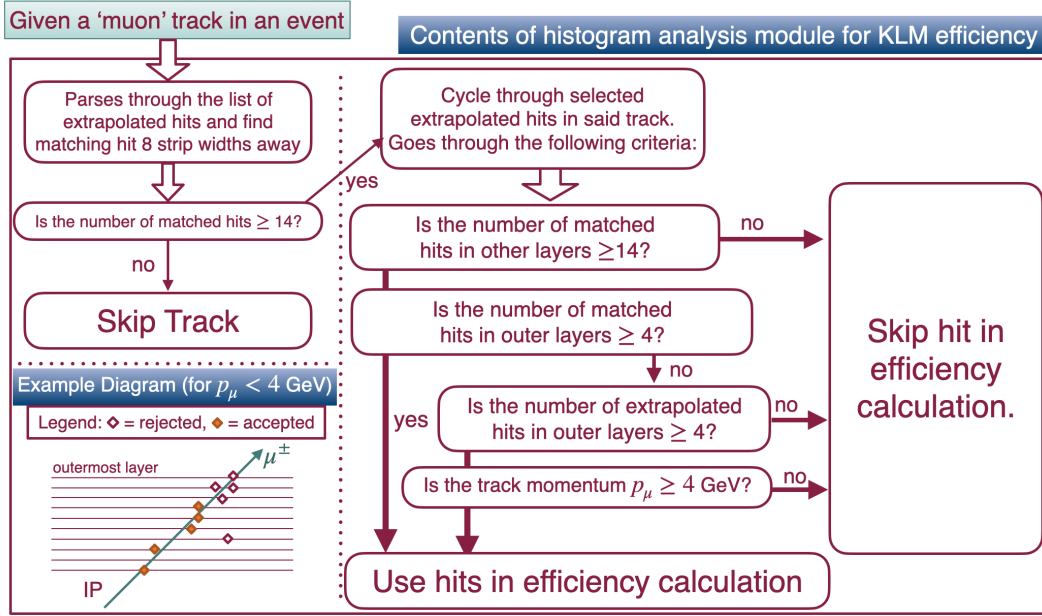


Figure B.3: Diagram of how extrapolated hits from muon-like tracks are matched and used for KLM efficiencies.

- If the first criterion fails, but the measured track momentum satisfies $p \geq 4$ GeV, then the hits in question are still accepted, as the tracks should be energetic enough to penetrate through all the layers of the KLM.
- From the tracking detectors, the measured momentum satisfies $p > 1.5$ GeV/ c
- If we are taking data with collisions, the transverse impact parameter and z-coordinate of the point-of-closest approach satisfy $|d_0| < 2$ cm and $|z_0| < 4$ cm, respectively

Using these criteria, which are also illustrated in Fig. B.3, we monitor the $e^+e^- \rightarrow \mu^+\mu^-$ process to accurately assess the KLM's performance in both online monitoring and offline calibration.

Because of the updates on reference handling outlined in Sec. B.2, this enables KLM to visualize our layer efficiency plots as a collection of 2D efficiency ratios, as illustrated in Fig. B.4. While Fig. 2.25 encodes all the information a KLM expert could ever need to

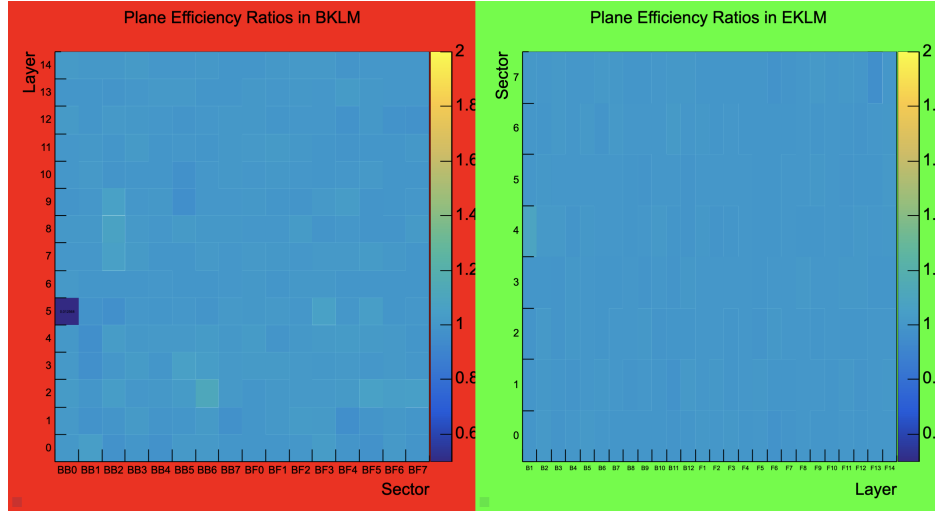


Figure B.4: An example of KLM’s efficiency ratio plot quickly highlighting an extremely inefficient layer, taken on 2026-03-30.

assess KLM’s performance, it is quite dense and provides a barrier to entry for new shifters to diagnose. Given the ability to leverage the efficiencies of the reference histograms, we compute ratios (or pulls) to detect deviations from one (or zero). To fully utilize the canvas, a heat map with efficiency ratios on the z-axis highlights unexpected drops in efficiency, as shown in Fig. B.4, including an issue with layer 5 in octant 0 of the backward barrel (BB0).

To inform shifters about what requires immediate attention and what does not, the canvas’s color highlights the severity of the problem. The conditions we consider are the number of inefficient layers and the severity of a layer’s inefficiency. MiraBelle is also used for inter-run comparisons, as additional sanity checks for KLM shifters.

In addition, KLM relies on the **EPICS** integration to tune the following values:

- condition of what an inefficient layer is (relative to the reference, calculate the z-score)
- the number of events needed to generate alarms on efficiency-based quantities.

In addition, we use **EPICS** to distribute as **EPICS** PVs and create alarm values on:

- number of inefficient layers in the barrel
- number of inefficient layers in the endcaps
- number of strips whose occupancy is 100x the mean occupancy of any sector
- number of dead modules.

B.3.3 Background Hit-Rates

While this efficiency-based monitoring is more comprehensible and resilient to beam backgrounds than the legacy hit occupancy plots, it is not completely impervious. One contribution that an elevated background level introduces is that a random background hit might, by coincidence, strike a strip near a track and fall within a reasonable time window, inflating our measured efficiency. The other contribution is detector dead time, in which elevated background levels temporarily blind a given strip or chamber while the active detector element recovers — this is especially an issue in the RPCs.

As SuperKEKB reaches higher luminosities, the RPC dead time due to elevated backgrounds will play a larger role; we thus track background levels in the KLM by adding the following quantities to the aforementioned suite of detector variables. The simplest quantity is to have a measure of a given layer's ‘hit rate.’ The hit rate is defined to be the number of hits of said layer in a given run, normalized by the total number of hits, the look-back timing window ($\sim 10.4\mu s$), and the layer's area. In other words,

$$R_{\text{layer } i} = \frac{N_{\text{layer } i}}{A_{\text{layer } i} \times \Delta t_{\text{look-back}} \times \sum_i N_{\text{layer } i}}. \quad (\text{B.1})$$

Eq. [B.1](#) is a clear quantity that depends on the beam currents and the layer-by-layer specific information allows us to hone in on problematic layers.

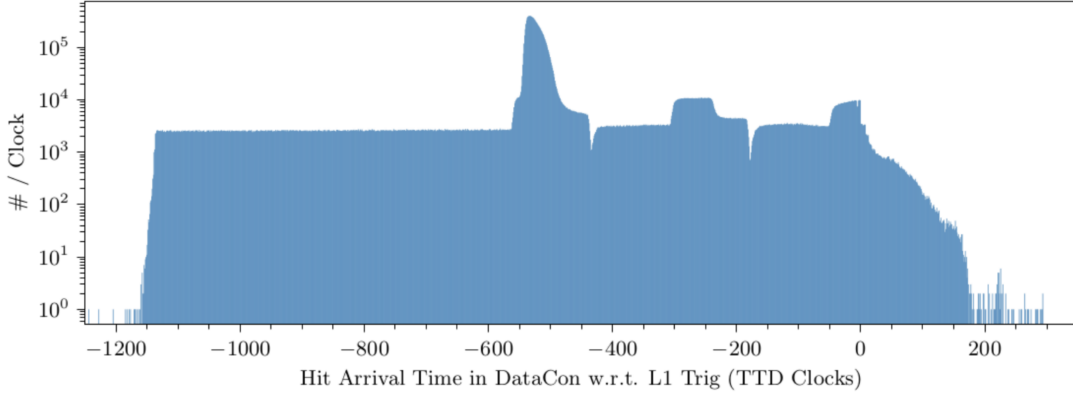


Figure B.5: For a nominal physics run, the RPC timing distribution. In particular, the timing of the arrival time from the data concentrator electronics with respect to the L1 trigger.

However, it would behoove us to consider removing contributions from genuine physics events, as this would systematically shift and obfuscate the measurement of background levels. The simplest way to do this is to look at time intervals known to be devoid of background. For the KLM scintillators, many of the backgrounds are suppressed at the firmware level, so an alternative strategy (e.g., fitting) would be needed. For the RPC scintillators, the timing distribution shown in Fig. B.5 suggests that using events with RPC hit times below the -600 ns peak to estimate background levels or hit rates would be straightforward. In addition, we can simply fit the physics peaks (around -600, -300, and 0 ns) to subtract them, then calculate the background hit rates. However, as the shapes at these time intervals are non-standard and susceptible to variation, we do not prioritize this approach.

Another way we remove the contribution of physics collisions is by considering the delayed Bhabha physics trigger (DPHY); the DPHY trigger saves events that occur $\sim 50\mu s$ after a Bhabha trigger was fired. As this is not typically in time with the e^+e^- collisions/crossings, it is a useful trigger to get a reasonable background rate³. Thus, we simply modify Eq. B.1

³While there are also Poisson or random triggers we can exploit, there are often too few events in the KLM look-back window for said triggers to have any meaningful contribution beyond what is available using the delayed Bhabha trigger.

by using events triggered by DPHY.

As Belle II continues to run and the stability of SuperKEKB increases, the validity of said hit-rate strategies will become clearer and provide a good measure for the Belle II upgrade group on how KLM should prepare for SuperKEKB's design luminosity. For their convenience, the hit rates are accessible via MiraBelle.

B.4 Conclusion

We have developed and deployed a comprehensive data quality monitoring framework, aimed at improving the speed, clarity, and reliability with which detector issues can be identified during data taking, with particular emphasis on KLM. By combining physically robust observables, such as muon-based layer efficiencies and background-sensitive hit rates, with a suite of analysis and visualization tools, the framework provides an accessible and effective interface for both experts and non-expert shifters.

A central outcome of this work is the significant reduction in the time required to identify and diagnose detector problems. Features such as automated reference comparisons, efficiency-ratio heat maps, and delta histograms have enabled shifters to recognize deviations from nominal behavior much earlier in a run, minimizing the impact of transient or evolving issues. The integration with EPICS further enables timely, configurable alarm handling, ensuring that critical conditions are promptly flagged.

Equally important, the framework reflects an iterative process informed by several years of Belle II operational experience. The prolonged gas system failure in the BB2 octant, which catalyzed the comprehensive overhaul of our monitoring strategy, serves as a stark reminder of the necessity for proactive diagnostics. While KLM has not experienced a

severe, prolonged anomaly of similar magnitude since then, the current DQM framework has repeatedly demonstrated its operational value. For instance, DQM's integration of EPICS alarming functionality has successfully detected sudden, silent high-voltage system trips mid-run, enabling immediate intervention. Furthermore, the continuous, automated monitoring of strip occupancy has enabled the team to promptly identify and recover from single-event upsets in KLM's front-end electronics before they severely impact data quality over long periods.

As a result, the current system is more responsive and resilient to the complexities introduced by increasing luminosity and beam backgrounds. These developments have strengthened the overall data quality assurance of the KLM subsystem and provide a robust model for future enhancements, both within Belle II and in other high-luminosity experiments where rapid and reliable monitoring is essential.

Appendix C

Generative Adversarial Network for KLM Backgrounds

In this section, I include work done to assess the feasibility of using a generative adversarial network (GAN) from convolutional neural networks (CNNs) to produce background templates for Belle II’s KLM.

C.1 Introduction

Large, detailed background studies are important for high-energy physics experiments, such as Belle II and those at CERN’s Large Hadron Collider (LHC), that are searching for rare physics processes and beyond the standard model (BSM) physics. While Monte Carlo (MC) methods based on the GEANT4 toolkit [202] are the standard, Generative Adversarial Networks (GANs) have become a popular alternative in recent years. Along with some novel uses to investigate potential new physics scenarios [203–205], GANs have been studied heavily at LHC in their use as fast event generators for various physics processes (e.g., jets [206–208], particle showers in calorimeters [209–212], and particle collisions [213, 214]) with reasonable agreement to MC events. In Belle II, GANs could be used to dynamically generate background events as an alternative to the existing catalog of recorded background events with

minimal concerns of statistical biases¹. This exercise has been done at Belle II using GANs containing convolutional neural networks (CNNs) for a single layer of Belle II’s pixel vertex detector (PXD) with some success [215]. While the physics processes and event rates differ, one might expect that a similar architecture could be used to repeat this exercise for other subdetectors, such as the KLM.

Since the PXD is the closest subdetector to the interaction point, there are a large number of detector hits per sector per event, providing a relatively complex prior. In contrast, the KLM is the farthest away from the interaction point; the background images typically have $\mathcal{O}(1)$ hits per sector per event, coming from either stray neutrons, out-of-time charged-particle tracks from the beam line, or cosmic rays. The biggest difference between neutrons and the other backgrounds for the KLM is layer dependence. Neutrons collide and ionize a gas molecule inside the detector material and surrounding material; the resulting ionization is highly localized to a hit in a few pixels on a single layer of the KLM. Out-of-time tracks and cosmic rays would pass through multiple layers of RPCs, creating hits across the various layers of the KLM along the particle’s path. Because the neutron backgrounds are simpler, being layer independent², we use them as an initial test to determine whether a GAN similar to that used for the PXD [215] is appropriate for the KLM. In this report, we summarize our findings on the usage of GANs for KLM neutron background generation.

C.2 Generative Adversarial Network

Generative Adversarial Networks (GANs) are a type of deep-learning generative model that utilizes a generator and a discriminator in an adversarial fashion [216]. The generator G is

¹The size of the background catalog cannot be expanded without limit because of the disk space limitations on the grid-computing nodes.

²While the rate of neutron events might be layer dependent, a hit from a neutron on one layer should not create a hit on another layer.

some model that transforms noise sampled from a uniform or normal distribution in some latent space with dimension d_{latent} into an output similar to the priors. The discriminator D samples from both G and the training data set and attempts to distinguish the two distributions. During the training process, the loss function of D is based on D 's accuracy (e.g., minmax loss or binary-cross entropy). The loss function of G is based on whether D misidentified the generated image, setting the two networks as adversaries. D and G would then be updated by some protocol (e.g., stochastic or batch gradient descent) during the training phase. A traditional GAN is successfully trained when the discriminator loss converges to $1/2$, signaling that D cannot discriminate between the two distributions.

Despite the successes of GANs in high energy physics [203, 206, 208–211], GANs are notorious for their instability in training [212, 217, 218]. In one scenario, the discriminator learns much faster than G , which would cause D 's loss to be minimal and the gradient of G 's loss to be near zero; this inhibits the GAN's ability to further approximate the prior. A less extreme scenario is mode collapse, wherein G collapses early on in training and only emulates a subset of the prior. Finally, the training of the GAN might not converge.

These instabilities in training GANs can be understood as the optimization problem not being convex [218]. Convex sets are simply-connected sets where the path between any two points are a line [219, Ch. 2]. A function whose domain is a convex set is itself convex, which will lead to the fact that convex functions will have at most one minimum [219, Ch. 3]. In the context of GANs, the loss functions are not constrained to be convex; considering the complexity of GANs, ensuring that any fixed points are the only fixed points would be one way to mitigate the main concerns of GANs [218]. Another way to interpret this is mode collapse and vanishing gradients are due to sharp gradients during training, causing D to associate a subset of our latent space to a single output [220].

C.2.1 Wasserstein GANs

One prescription to address the latter point is by incorporating the Earth-Mover (EM) or Wasserstein distance as the loss function. Given random variables X and Y sampled marginally from two probability distributions P_1 and P_2 , respectively, the formal definition of the EM distance is

$$W(P_1, P_2) = \inf_{\gamma} \int \|x - y\| d\gamma(x, y) = \inf_{\gamma} \mathbb{E}(\|X - Y\|), \quad (\text{C.1})$$

where γ is a joint distribution between P_1 and P_2 [217]. To gain some intuition, let's consider the discrete case, where $\rho_i = \sum_{j=1}^N m_j \delta(x - x_i)$ for $i \in \{1, 2\}$. The EM distance in this case would be

$$W(\rho_1, \rho_2) = \inf_{\sigma \in S_n} \sum_{j=1}^N m_j |x_j - y_{\sigma_j}|,$$

where $\sigma \in S_n$ is the set of permutations of a fixed x_i and varying y_j . Intuitively, the expression above describes the most efficient way to map the N particles from ρ_1 into ρ_2 [221]. Returning to the continuous case, $\gamma(x, y)$ can then be interpreted as some set of paths mapping the masses dm in P_1 to P_2 and the optimal γ corresponds to the EM distance.

In both the discrete or the continuous definition, the EM distance is continuous if x is continuous, and differentiable almost everywhere if x is also Lipschitz continuous (i.e., $\|f(x) - f(y)\| \leq \|x - y\|$). Other commonly used metrics (e.g., Jensen-Shannon divergence and the Kullback-Leibler divergence [216]) are not guaranteed to be continuous everywhere [217, Theorem 1], motivating the use of the EM distance as a more sensible cost function. Constraining our function space to only Lipschitz continuous functions also implies

$$\|\nabla f\| = \left\| \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \right\| \leq \left\| \lim_{h \rightarrow 0} \frac{x+h - x}{h} \right\| = 1.$$

This ensures that we will not have sharp gradients, minimizing the chance of mode collapse [220].

To implement the EM distance and establish the Wasserstein GAN (WGAN), the Kantorovich-Rubinstein duality is used to modify Eq. C.1 to

$$W(P_1, P_2) = \sup_{\|\Delta f\| \leq \|\Delta x\|} [\mathbb{E}_{x \sim P_1}(f(x)) - \mathbb{E}_{x \sim P_2}(f(x))], \quad (\text{C.2})$$

with the Lipschitz continuous f being the GAN’s “discriminator,” P_1 being the prior, and P_2 being G ’s probability distribution [217]. Because the label’s accuracy is not explicitly used in Eq. C.2, the “discriminator” is replaced with a “critic.” To further ensure that f is Lipschitz continuous, a Lagrange multiplier can be added in the form of $\lambda(\|\nabla f\| - 1)^2$ where $\lambda = 10$ as a default [222]. This constraint has been shown to also improve the performance of more traditional GANs [223].

The hyper-parameters of the WGAN are d_{latent} and n_{critic} . As common to all GANs, the latent dimension d_{latent} is the length of the noise vector we feed into G where each component is sampled from a normal or uniform distribution [216]. Specific to WGANs, n_{critic} is the number of iterations the critic has to train per generator iteration [217]. The reason for this is to obtain accurate $\nabla W(P_1, P_2)$, which is used in updating G . While there are other hyper-parameters, such as the length of the CNNs and batch size, we have found that the aforementioned hyper-parameters to be more sensitive for a stable and successful training.

C.2.2 Architecture

The critic and generator architectures, as illustrated in Fig. C.1, is heavily based off of [222]. Comparing to the architecture used for the PXD [223], the proposed architecture has the same activation functions and number of layers for the CNNs. One major difference is the

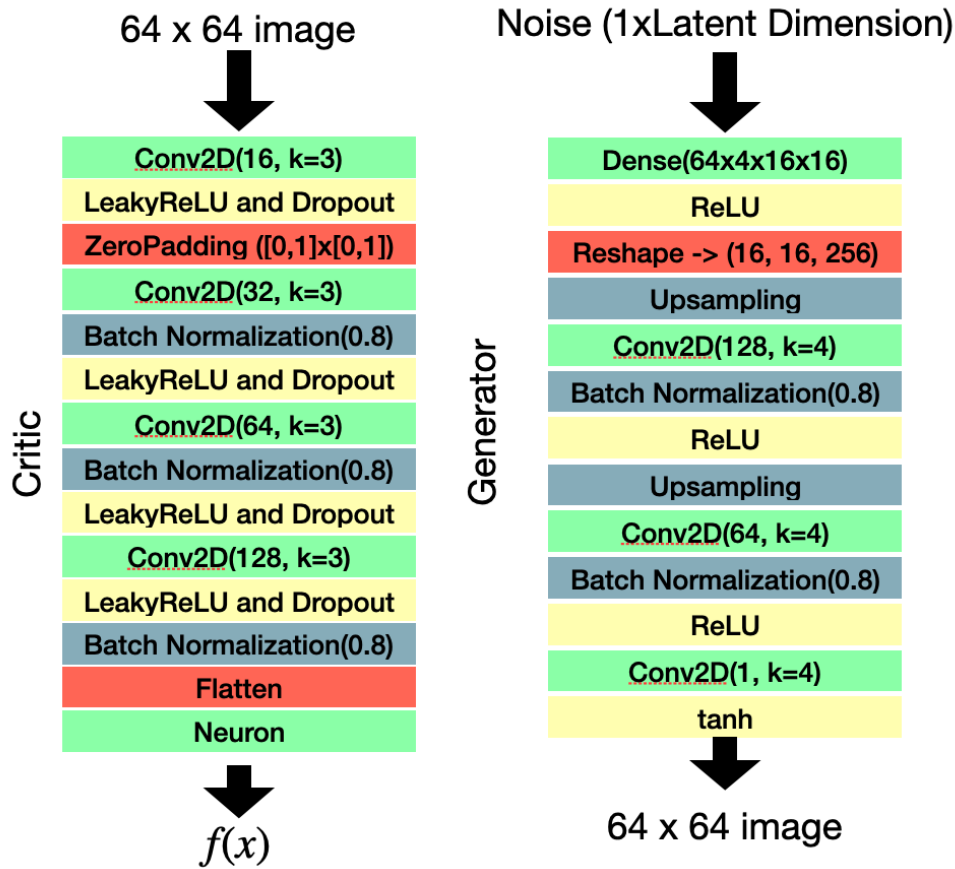


Figure C.1: The architecture of the critic (on the left) and generator (on the right) used in this work.

depth of the CNNs. We decided to use shallower CNNs for our WGAN since our prior is expected to be far less complex than the PXD's.

C.3 Procedure

In this section, we design an experiment to determine if the WGAN can learn a prior with a similar level of complexity to that seen by the KLM. We also prescribe a figure of merit to evaluate the performance of the WGAN.

C.3.1 Priors

Rather than use neutron backgrounds, we want to construct a prior that is well-understood but of a similar level of complexity in order to understand what the WGAN can learn. We start our construction of a neutron-like hit on a 64×64 quadrant in $\mathbb{R}^2$, which is comparable to the number of pixels in a single KLM panel (one layer of one sector), and have some procedure to determine the centroid's location. The neutron-like hit is generated by sampling $N = 60$ points from a 2D Gaussian distribution with $\sigma_r = 1$, serving as the base of our prior. This is meant to emulate a streamer in the KLM's resistive plate chamber (RPC), creating a large enough potential difference between the resistive plates to correspond to a "pixel hit." In the digitization process, which is illustrated in Fig. C.2, we histogram our 60 points with binning 1×1 squares. If there are no hits in the bin, the pixel remains 0. Otherwise, the pixel is set to 1.

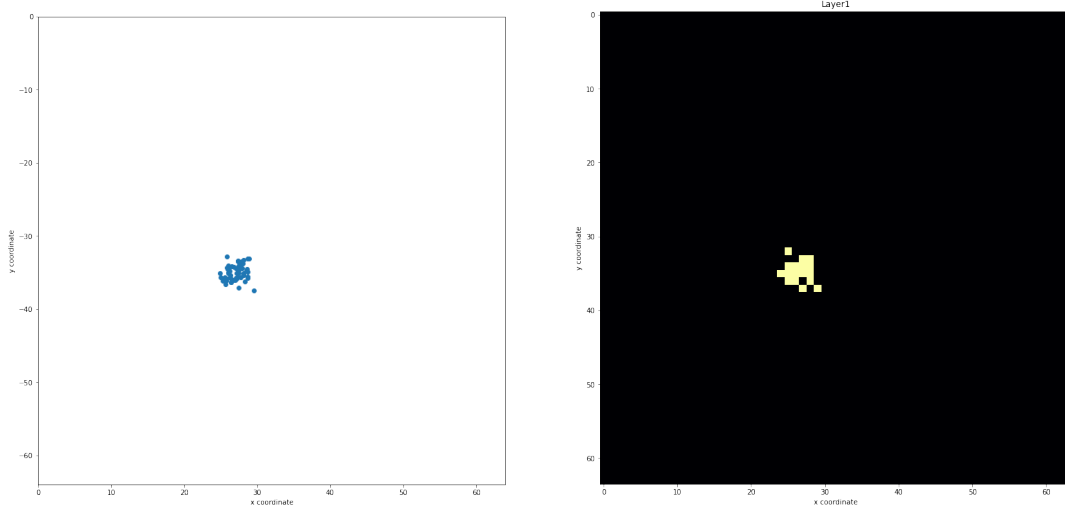


Figure C.2: The left image is the raw image where $N = 60$ points are chosen from a 2D Gaussian centered with $\sigma = 1$. The right image is the digitized version of the left image, where if a point falls in a 1×1 bin, that is considered a hit.

Various priors can be constructed by changing the number of hits in a given image and the distribution of the centroids. Throughout the report, we will focus on three different

priors and discuss a fourth one towards the end³, as illustrated in Fig. C.3. Priors “a” and “b” have a single hit with the centroid sampled uniformly across $[0, 64] \times [0, 64]$ and middle $[16, 48] \times [16, 48]$, respectively. Prior “c” has two hits, one with a centroid sampled uniformly in the upper left quadrant and the other sampled uniformly in the bottom right quadrant. The final prior “d” has a single hit with the centroid fixed at coordinate $(x, y) = (32, 32)$.

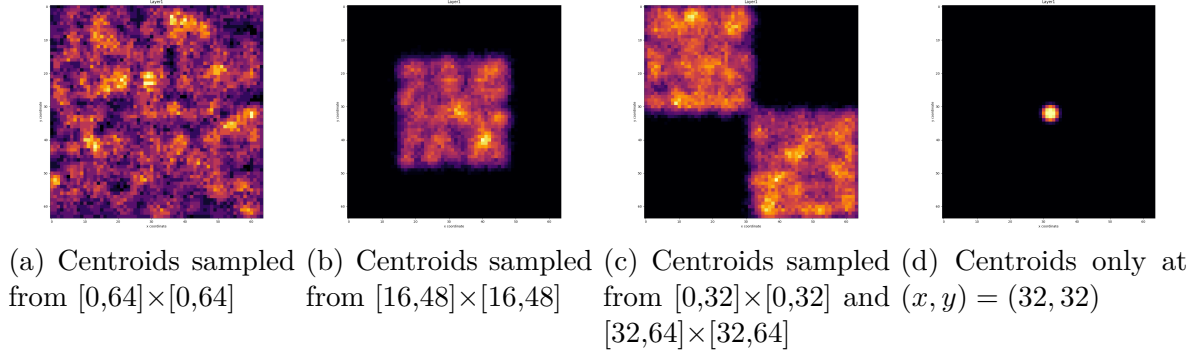


Figure C.3: An overlay of 1000 images stacked for each prior.

C.3.2 Training and Development

The actual training of a single WGAN with fixed hyper-parameters $\{d_{\text{latent}}, n_{\text{critic}}, \text{batch size} = 32\}$ consists of a data set of 12,000 images from one of the priors illustrated in Fig. C.3 and an upper bound of 15,000 training epochs. Due to the instability of GANs, we found that different GANs with the same hyperparameters can have drastically varying results. One way to circumvent this, aside from having multiple copies of the same hyperparameters, is to vary d_{latent} by ± 2 for some fixed choice. We show in Sec. D.3 that varying the d_{latent} by $\delta d_{\text{latent}} \sim 1$ can potentially yield similar results.

One way to evaluate the training of the GAN is by inspecting the generator’s loss function vs. training epoch, which we indicate as $g(t)$. One advantage of using the Wasserstein metric as

³All four priors have the N and σ for the cluster’s prior fixed.

the generator’s loss function is that it provides useful information during the progression of the training phase. For instance, over-training is indicated when $g(t)$ exponentially increases. Additionally, negative values of g can potentially indicate that the GAN is constraining itself to Lipschitz continuous functions. Because $\lambda = 10$, we note that it will dominate the training until $(\|\nabla f\| - 1)^2$ becomes sufficiently small. With this information, we have a prescription on how to assess the GAN’s training. For a positive and exponentially increasing $g(t)$, we consider the GAN has over-trained and end the training early. When $g(t) < 0$ for a prolonged range of epochs t , we assume the GAN is Lipschitz continuous and inspect various instances of the generator for $g(t) \sim 0$.

While the aforementioned prescription is useful in a quick assessment, it comes with concerns. One concern is if we inspect $g(t)$ for some range of epochs and $g(t)$ is approximately a constant (ex. $\propto e^{-x}$ for $x \gg 1$), it will be difficult to quantitatively determine which instance of the generator is best to deploy for background generation. Another concern is the ambiguity on how to compare the generator for different WGANs, assuming the “best” instance of each generator was chosen. In order to resolve this, we are interested in finding a way to evaluate the deployed generator’s performance that is independent of the WGAN’s training procedure.

C.3.3 Figure of Merit

Two natural figures of merit (FoMs), considering our priors, would be based on the number of clusters in a given image and the standard deviation of each cluster. We use a Dirichlet Process Gaussian Mixture Model (DPGMM) to efficiently achieve both simultaneously.

Gaussian Mixture Model (GMM) with k components is a specific clustering algorithm that can also determine the spread of a given cluster from a given data set, assuming the under-

lying distribution has a fixed number of k Gaussian distributions. The joint probability of observing a data point $\vec{x}$ in some configuration of k Gaussian clusters is

$$p(\vec{x}|\{\vec{\mu}_i, \Sigma_i\}) = \sum_{j=1}^K \pi_j \mathcal{N}(\vec{x}|\vec{\mu}_j, \Sigma_j) \quad (\text{C.3})$$

where $\mathcal{N}(\vec{\mu}_j, \Sigma_j)$ is the j 'th normalized 2D-Gaussian and π_j is weighting of j 'th with respect to the others clusters such that $\sum_j^k \pi_j = 1$. The optimal GMM would maximize $\Pi_j^N p(\vec{x}_j|\{\vec{\mu}_i, \Sigma_i\})$ by determining the optimal $\{\vec{\mu}_i, \Sigma_i\}$ for all N data points simultaneously. This can be computed by maximizing

$$F = \ln(\Pi_j^N p(\vec{x}_j|\{\vec{\mu}_i, \Sigma_i\})) = \sum_i^N \ln(p(\vec{x}_j|\{\vec{\mu}_i, \Sigma_i\})), \quad (\text{C.4})$$

which yields

$$\sum_{j=1}^n \frac{\pi_k \mathcal{N}(\vec{x}_j|\vec{\mu}_i, \Sigma_i)}{p(\vec{x}_j|\{\mu_l, \Sigma_l\})} \left(\frac{\vec{x}_j - \vec{\mu}_i}{\Sigma_i^2} \right) \equiv \sum_{j=1}^n \gamma_j \frac{\vec{x}_j - \vec{\mu}_i}{\Sigma_i^2} = 0. \quad (\text{C.5})$$

The optimal GMM can be approximated then for some initialization of $\vec{\mu}_i$ and Σ_i (ex. K-means) by iteratively solving for γ_i and $\vec{\mu}_i$ until some convergence criterion is reached [224].

In order to generalize the GMM to where k is not fixed, a Dirichlet Process can be used. A Dirichlet Process is a family of stochastic processes whose observed values are probabilities. A common representation of a Dirichlet Process is the Chinese Restaurant Problem (CRP). In order to seat N customers, the number of tables needed would be determined by the following procedure and a choice of a concentration parameter α . For the N 'th person entering the restaurant, the probability of them joining a table with n people is $n/(\alpha + N - 1)$ while the probability of joining a new table is $\alpha/(\alpha + N - 1)$. The CRP is a Dirichlet Process because the stochastic process determines whether the N 'th customer will be seated at a new table while the observed value, the number of tables, has an underlying probabilistic distribution

[225–228].

To implement the DPGMM as a FoM, we use `SCIKIT-LEARN`’s [229] Bayesian Gaussian Mixture class. We set an upper limit of clusters to be six and determine the effective number of clusters by looking at the cluster’s weight. If the weight of the cluster is less than 10^{-4} , we say this to be effectively zero and discard it. The upper limit was chosen, expecting the generator to produce only a few clusters. We also expect the standard deviation of the distribution to be $\mathcal{O}(1)$ so if the standard deviation of a distribution is greater than 5, we assume DPGMM underestimated the number of clusters by 1. In order to determine the standard deviations, we use another GMM where the number of clusters is fixed by the first DPGMM. When characterizing the prior, we skip the first analysis and just assume the correct number of clusters for simplicity. As shown in Fig. C.4, we see that despite the initial prior being spherical with $\sigma = 1$, the digitization of the image causes the GMM to see some variation.

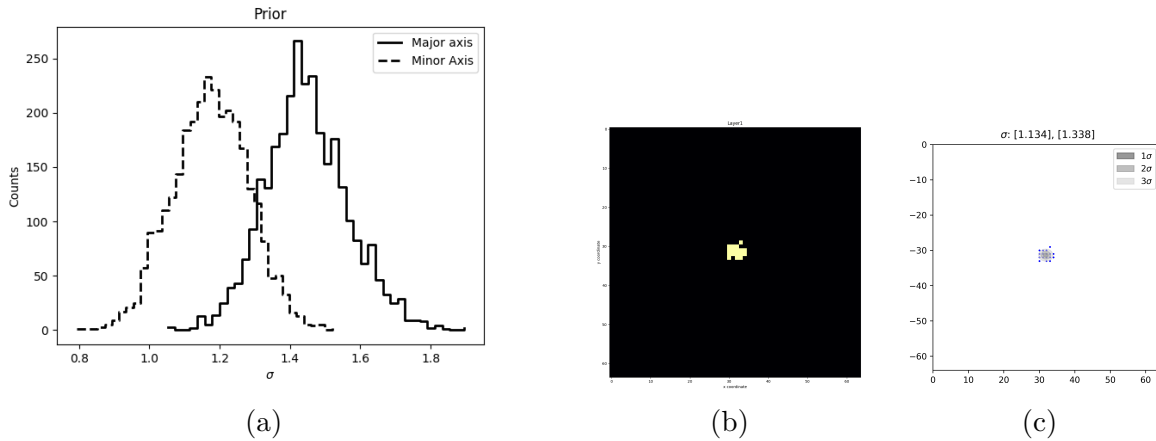


Figure C.4: An illustration of how a GMM would characterize prior “d.” Despite the digitization process causing the GMM to obtain standard deviations that are larger than the initial input, it still offers a baseline of what an ideal WGAN should reproduce. (a) A histogram of the square root of the eigenvalues of the covariance matrix using GMM on 3000 images. Because the 1σ curve forms an ellipse, the larger (smaller) eigenvalue is associated with the major (minor) axis of the ellipse. (b) A sample image from prior “d.” (c) Applying GMM on the sample image.

In order to perform a similar analysis for the GAN generated image, image preprocessing is required since the generator’s output ranges from -1 to 1. From inspection of successfully trained WGANs, we apply a Heaviside-step function $H(x - 0.7)$ where x is the “intensity” of the pixel. This digitization process is depicted in Fig. C.5. From the raw image, we see that the cluster we expect to see is present as the “brightest” pixels. For values less than 0.7, we see that the digitized image picks up too much noise, which would adversely affect the DPGMM. Threshold values near 0.7 highly suppress most (if not, all) the noise.

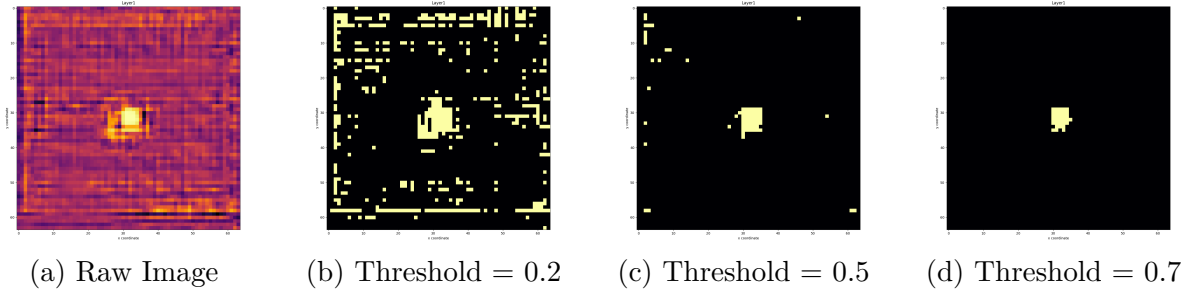


Figure C.5: Given an image from a successfully trained WGAN using prior “d,” we illustrate the effects of the thresholding procedure for different values.

C.4 Results

This section begins with a qualitative examination of the GAN’s training stages, examining how one can assess the training from an inspection of the generator loss vs. epochs ($g(t)$) by looking at a number of examples. A more general discussion then focuses on the dependence of $g(t)$ based on d_{latent} and n_{critic} , where the default values of $d_{\text{latent}} \in [85, 88]$ and $n_{\text{critic}} = 4$ were determined.

C.4.1 Progression of Training

As a base case, a GAN trained with prior “b,” $d_{\text{latent}} = 85$, and $n_{\text{critic}} = 5$ is chosen. The loss function and sample images at various epochs t are depicted in Fig. C.6. At early epochs, the initial weights were normalized so nearly random noise is expected for Fig. C.6b, as depicted. As the GAN continued to learn, the GAN seemed to be picking up that the middle quadrant is relevant in the learning procedure, while trying to minimize the surrounding noise. However, the GAN seems to fail to learn much about the prior by epoch 900 since $g(t)$ began to increase exponentially, making it evident that the GAN has clearly failed training.

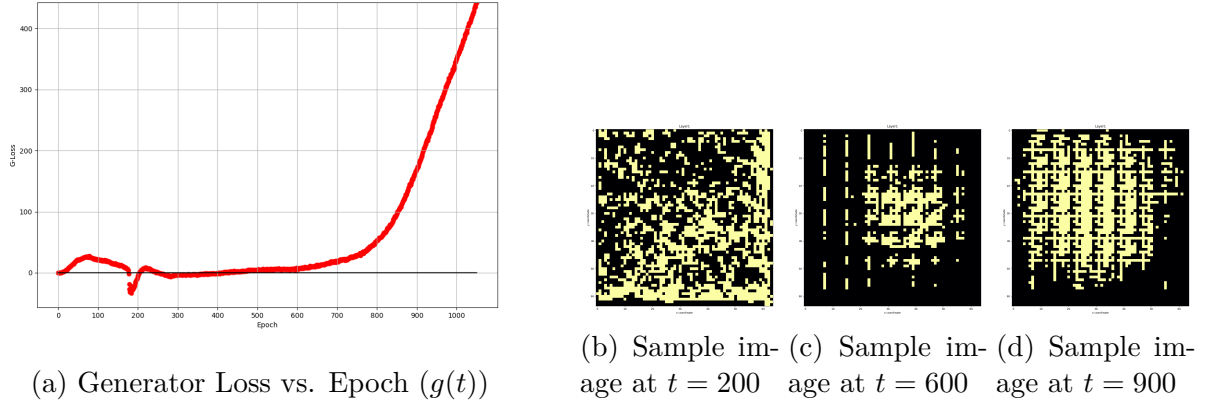
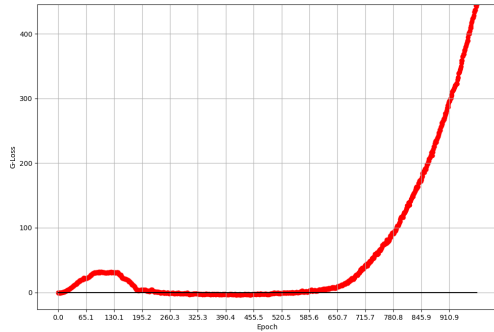


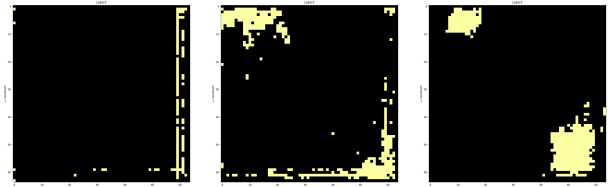
Figure C.6: A comparison of the $g(t)$ of a GAN learning prior “b” with hyper-parameters $\{d_{\text{latent}}, n_{\text{critic}}\} = \{85, 5\}$. The maximum epoch (i.e., when $g(t) = 400$) is $t \approx 950$.

As a second example, a GAN trained with prior “c,” $d_{\text{latent}} = 85$, and $n_{\text{critic}} = 5$ is chosen. The loss function and sample images at various epochs t are depicted in Fig. C.7. The previous example showed that the first couple of hundred epochs should be ignored since only noise is outputted early on. By epoch 400, depicted in Fig. C.7b, the image has mostly quieted down, aside from some anomalies at the edge of the image. The GAN then seems to be transitioning during epoch 500 from a mostly quiet image to recognizing that

hits happening in the upper left and lower right quadrants are desirable, where this is fully achieved by epoch 700. At this point, however, $g(t)$ seems to be increasing exponentially, signaling that the GAN will not improve past this point.



(a) Generator Loss vs. Epoch ($g(t)$)



(b) Sample im- (c) Sample im- (d) Sample im-
age at $t = 400$ age at $t = 500$ age at $t = 700$

Figure C.7: A comparison of the $g(t)$ of a GAN learning prior “c” with hyper-parameters $\{d_{\text{latent}} n_{\text{critic}}\} = \{85, 5\}$. The maximum epoch (i.e., when $g(t) = 400$) is $t \approx 920$.

The final example is a GAN trained with prior “b” $d_{\text{latent}} = 86$, and $n_{\text{critic}} = 5$. While one might expect a similar trend to the aforementioned examples, the loss function depicted in Fig. C.8a suggests a distinction. Fig. C.8b shows that the GAN has started to recognize the middle quadrant’s relevance much earlier than initially expected. As training progressed, the outputs of the GAN began to become less noisy, and clusters of points appeared in the middle quadrant. The GAN seems to then take this to the extreme, producing smaller clusters or even blank images as training reaches ~ 2000 epochs.

When we compare this example to the previous two, the shape of the loss function $g(t)$ does indeed provide meaningful information about the WGAN’s training, as advertised in Sec. C.3.2. The initial bump at early epochs appears to indicate the region where the GAN is learning to minimize the noise in the desired region. A telltale sign that training is improving is the noticeable decrease in $g(t)$ and the smooth return to zero. The loss function, however, does not seem to provide detailed information on when ideal images are being generated, as

indicated by the GAN providing blank images despite $g(t) \sim 0$.

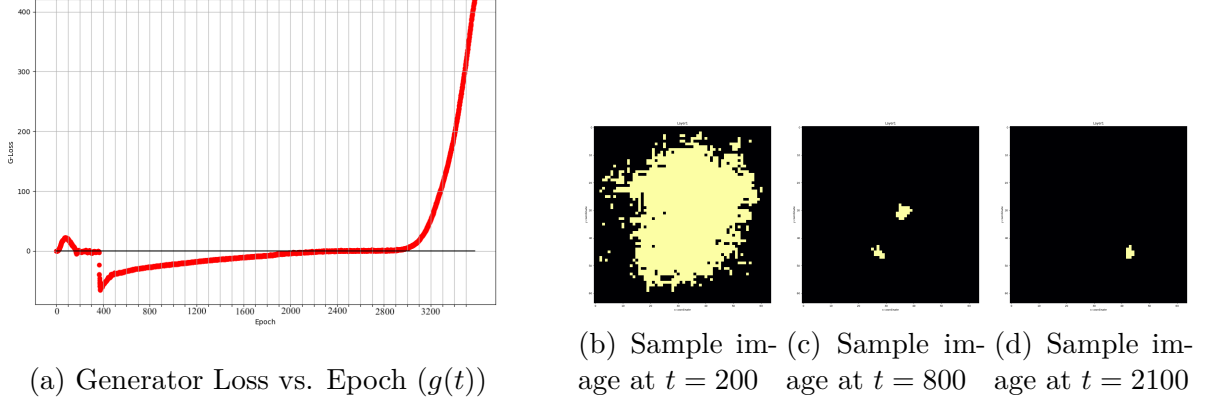
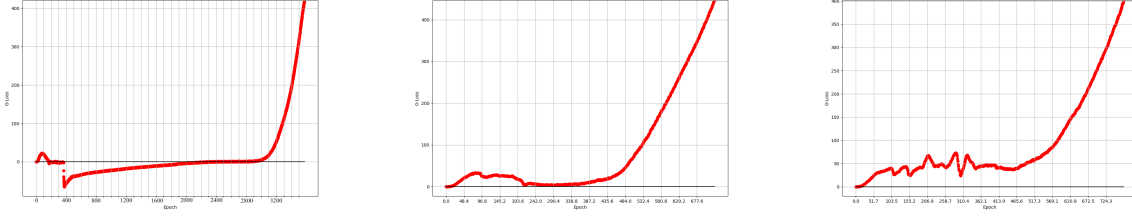
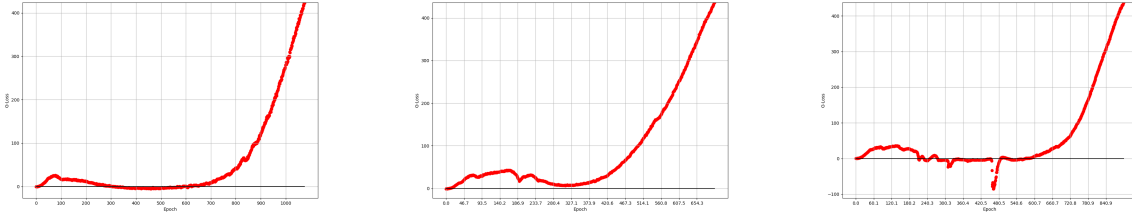


Figure C.8: A comparison of the $g(t)$ of a GAN learning prior “b” with hyper-parameters $\{d_{\text{latent}} n_{\text{critic}}\} = \{86, 5\}$. The maximum epoch (i.e., when $g(t) = 400$) is $t \approx 3200$.

C.4.2 Dependence on d_{latent}

Continuing the previous analysis, we inspect multiple runs of a fixed d_{latent} for $n_{\text{critic}} = 5$ but varying priors (“a,” “b,” and “c”). We notice when inspecting Fig. C.9 and C.10 that the differences in $g(t)$ between GANs with the same hyperparameters are quite substantial, to the point where the variation between $d_{\text{latent}} = 86$ and $d_{\text{latent}} = 87$ seems comparable. Comparing Fig. C.9 with Fig. C.11 and Fig. C.6a/C.7a, we might be able to qualitatively claim that the sharper fluctuations in $g(t)$ for $d_{\text{latent}} = 90$ are more prevalent and noticeable relative to $d_{\text{latent}} = 86$. Though not included, $d_{\text{latent}} = 70$ and $d_{\text{latent}} = 75$ trials appeared to never plateau near zero, so we rule out smaller values of d_{latent} . Larger values of $d_{\text{latent}} \approx 100$ appeared to never yield desirable results after many runs so we also rule out larger values of d_{latent} . From this qualitative analysis, we conclude that a reasonable range of d_{latent} to consider is relatively near $d_{\text{latent}} = 86$. We define relatively near to be ± 2 since $d_{\text{latent}} = 85$ and $d_{\text{latent}} = 87$ are definitely nearly indistinguishable while $d_{\text{latent}} = 85$ and $d_{\text{latent}} = 90$ have some minute qualitative differences.

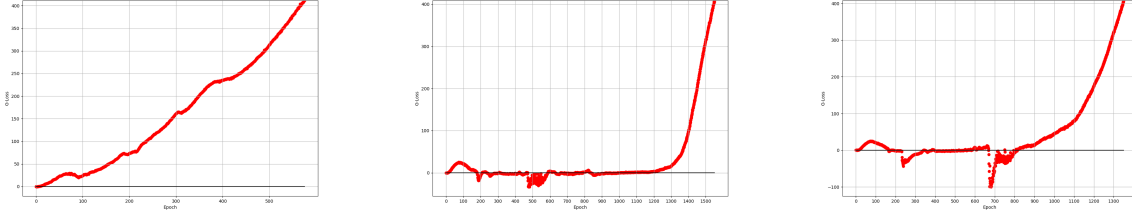
Figure C.9: Multiple runs with $d_{\text{latent}} = 86$ Figure C.10: Multiple runs with $d_{\text{latent}} = 87$

C.4.3 Dependence on n_{critic}

The initial motivation in investigating n_{critic} is that varying d_{latent} clearly did improve the stability of the GAN’s training. Despite an optimal range of d_{latent} being determined, the rate of successfully trained GANs was too low. With no other way to modify the generator explicitly⁴, the next hyper-parameter to investigate was n_{critic} .

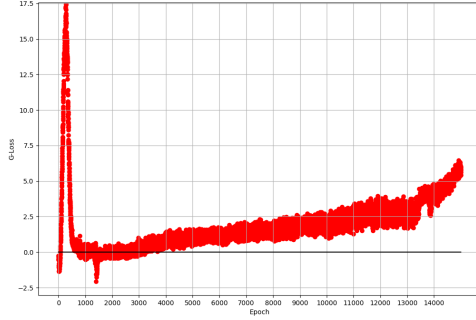
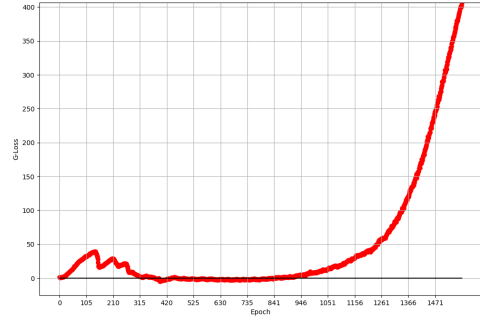
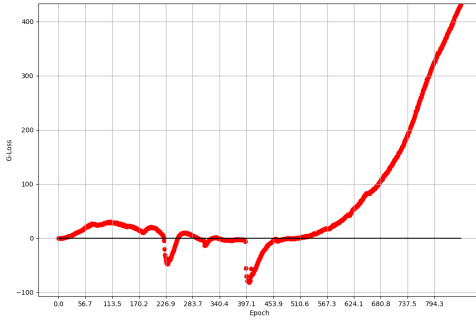
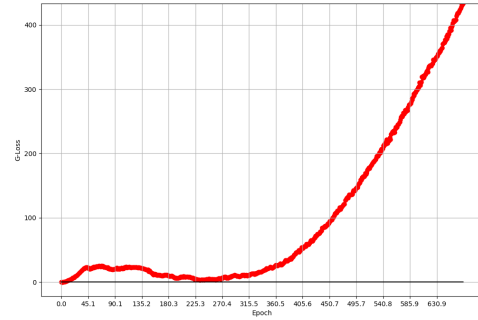
In the analysis of a WGAN’s dependence on n_{critic} , we first investigate qualitative trends while holding d_{latent} fixed (i.e., $d_{\text{latent}} = 87$). As illustrated in Fig. C.12, the difference between $n_{\text{critic}} = 1$ and $n_{\text{critic}} = 7$ is apparent. For small values of n_{critic} , we find that $g(t)$ remains fairly small for $t \in (0, 15000)$, so there is minimal concern of the WGAN over-training. The explanation for smaller values of n_{critic} is that the gradient information is not accurate enough, causing the update of the generator’s weights to be suppressed. For larger values of n_{critic} , $g(t)$ tends to explode much earlier. The intuitive explanation for this is

⁴We found that modifying the CNN’s depth and increasing the data set by $\mathcal{O}(10)$ had minimal effects.

Figure C.11: Multiple runs with $d_{\text{latent}} = 90$

related to the adversarial nature of the GANs; if the critic has too many opportunities to train, relative to the generator, then the WGAN would suffer from the critic penalizing the generator too harshly to the point where the generator cannot recover.

We validate this hypothesis by scanning over various n_{critic} multiple times. As a way to save time and artificially increase the available data set, we used our compiled list of trained GANs with $d_{\text{latent}} \in [85, 88]$, assuming a variation of ± 2 has less of an impact than the instability of the GAN's adversarial nature. With our available data set, we repeat the exercise above: record the final epoch before we terminate training ($\text{Epoch}_{\text{max}}$) and take an average over fixed n_{critic} . In this data set, the number of WGANs we have ranges from $N = 3$ (for $n_{\text{critic}} = 1, 3$) to $N = 11$ (for $n_{\text{critic}} = 5$). For the smaller data sets, the difference between unsuccessful and successful training is reflected in a larger standard deviation. The exception to this is for $n_{\text{critic}} = 1$, where all 3 accounts reached the maximum 15000 epochs so there is no reported variance. The results of this exercise are consolidated in Table. C.1. We tried to account for the varying sizes of the data sets by reporting both the standard deviation (σ_{max}) and the standard error ($\sigma_{\text{max}}/\sqrt{N}$). On the whole, while it is hard to statistically distinguish between $n_{\text{critic}} = 3$ and $n_{\text{critic}} = 4$, we can still see a trend with n_{critic} and $\overline{\text{Epoch}_{\text{max}}}$, providing some validity to our claim above. From this analysis, we find $n_{\text{critic}} = 4$ seems to be a reasonable default value.

(a) $n_{\text{critic}} = 1$, $\text{Epoch}_{\text{max}} = 15000$ (b) $n_{\text{critic}} = 3$, $\text{Epoch}_{\text{max}} = 1500$ (c) $n_{\text{critic}} = 5$, $\text{Epoch}_{\text{max}} = 800$ (d) $n_{\text{critic}} = 7$, $\text{Epoch}_{\text{max}} = 650$ Figure C.12: A side-by-side comparison of $g(t)$ with varying n_{critic} and fixed $d_{\text{latent}} = 87$.

C.4.4 Evaluation of Deployed WGANs

With our hyper-parameters chosen and well-trained GANs obtained, we evaluate the deployed generator's performance to see if the prior has been faithfully approximated.

Using a successfully trained WGAN with $(d_{\text{latent}}, n_{\text{critic}}) = (85, 4)$, we generate 250 images and use the DPGMM to count the number of clusters present in an image and validate its performance by eye. Fig. C.13 illustrates that the DPGMM isn't perfect when counting the number of clusters. Fig. C.14a is obtained by repeating this exercise for all 250 images. While the distributions between what was determined visually and the DPGMM are similar⁵,

⁵Only the zero clusters bins are identical since recognizing a blank image doesn't require a DPGMM

n_{critic}	$\overline{\text{Epoch}}_{\text{max}}$	σ_{Epoch}	$\sigma_{\text{Epoch}}/\sqrt{n}$
1	15000	N/A	N/A
3	2042	1116	645
4	1963	2227	908
5	1171	809	244
7	558	128	64

Table C.1: Comparing the average epoch t where $g(t) = 400 (\overline{\text{Epoch}}_{\text{max}})$ with $d_{\text{latent}} \in [85, 88]$ for different n_{critic} . We see that there is a clear trend between $\overline{\text{Epoch}}_{\text{max}}$ and n_{critic} , implying that n_{critic} affects the rate of the discriminator learning and delays the point where overtraining occurs.

Fig. C.14b informs us that the DPGMM has an accuracy of $(82.8 \pm 5.8)\%$ while $\|n_{\text{cluster}}^{\text{DPGMM}} - n_{\text{cluster}}^{\text{truth}}\| \leq 1$ $(98.8 \pm 6.3)\%$ of the time⁶. We also notice that the DPGMM has a slight bias of underestimating the number of clusters in a given image. One way we correct for this is by inspecting the square root of the eigenvalues of the covariance matrices $\vec{\sigma} = (\sigma_{\text{Major}}, \sigma_{\text{Minor}})$. If the larger eigenvalue is over some threshold (i.e., $\sigma_{\text{Major}} > 5$), we can safely assume that the DPGMM has most likely underestimated the number of clusters by at least 1 and record the image containing one more cluster than initially predicted.

With the number of clusters determined from the aforementioned procedure, we can use the GMM one more time to get a more precise determination of a cluster’s covariance matrix. Using the same generator as above, we use a data set of 3000 images and histogram σ_{Major} and σ_{Minor} separately. Fig. C.15 illustrates that the average value for the major and minor axes⁷ around 0.9 and 1.1, respectively. The distributions are also significantly wider than Fig. C.4a, which is an artifact of the GAN underestimating the number of clusters.

Fig. C.14 and C.15 illustrate that, despite the generator of a well-trained GAN producing qualitatively reasonable images, the generator is not faithful based on our FoM, which in-

⁶The percent errors were determined from scaling the Poisson error by the total number of generated images.

⁷Since cross sections of 2D Gaussians are ellipses, major (minor) is used to label the larger (smaller) σ ’s.

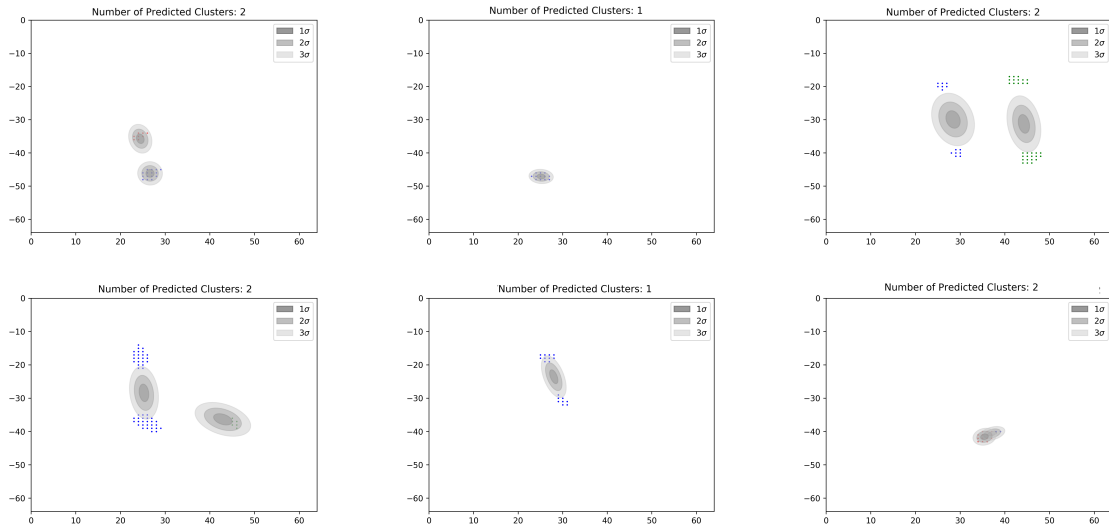


Figure C.13: Samples of using DPGMM to determine the number of clusters from a generator with $d_{\text{latent}} = 85$ and $n_{\text{critic}} = 4$.

cludes priors “a” through “c.” Not only are the number of clusters not an invariant number, but the distribution of the cluster sizes are also in disagreement. In the case of PXD, their prior has a larger and varying number of clusters per image [215]. Furthermore, although the shape of the clusters in their generator appears to be slightly different to the priors, the approximate size of the clusters are comparable between the two. However, our prior has a different level of complexity; we can qualitatively make these distinctions much easier, forcing us to be more cautious when approaching these differences.

A hypothesis as to what could cause this comes from the motivation behind capsule networks [230]. The authors claim that CNN, the max/average pooling in particular, has major issues with variations in images, including translational and rotational invariance. If this is indeed the case, then one would expect priors similar to “a,” “b,” and “c” to be difficult for the critic/discriminator to interpret easily since $\vec{\sigma}$ should be mostly invariant under translation and rotation. To test our claim, we repeated the previous exercise with prior “d.” Because the centroid of the cluster is always at $(x, y) = (32, 32)$, the GAN’s job is easier since it will

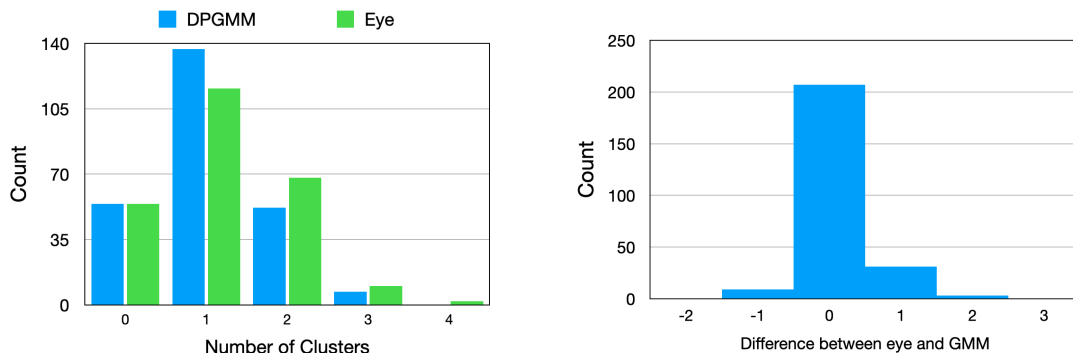


Figure C.14: Using a data set of 250 generated images with target prior b , the left plot compares the number of clusters determined by the DPGMM with one determined visually. The right plot panel shows the number of clusters determined visually minus the number determined by the DPGMM. From this, we see that the DPGMM has an accuracy of $82.8 \pm 5.8\%$ while being within ± 1 clusters $98.8 \pm 6.3\%$ of the time with a slight bias towards underestimating the number of clusters.

not be concerned with translations. Fig. C.16 clearly illustrates that, while the mean of the distribution seems to be slightly less than the one in Fig. C.4a, the tails of the distributions are in agreement. This leads us to the conclusion that GANs are not optimal in learning both the distribution of the centroids and the cluster’s underlying distribution simultaneously due to the critic/discriminator’s difficulty with translational invariance.

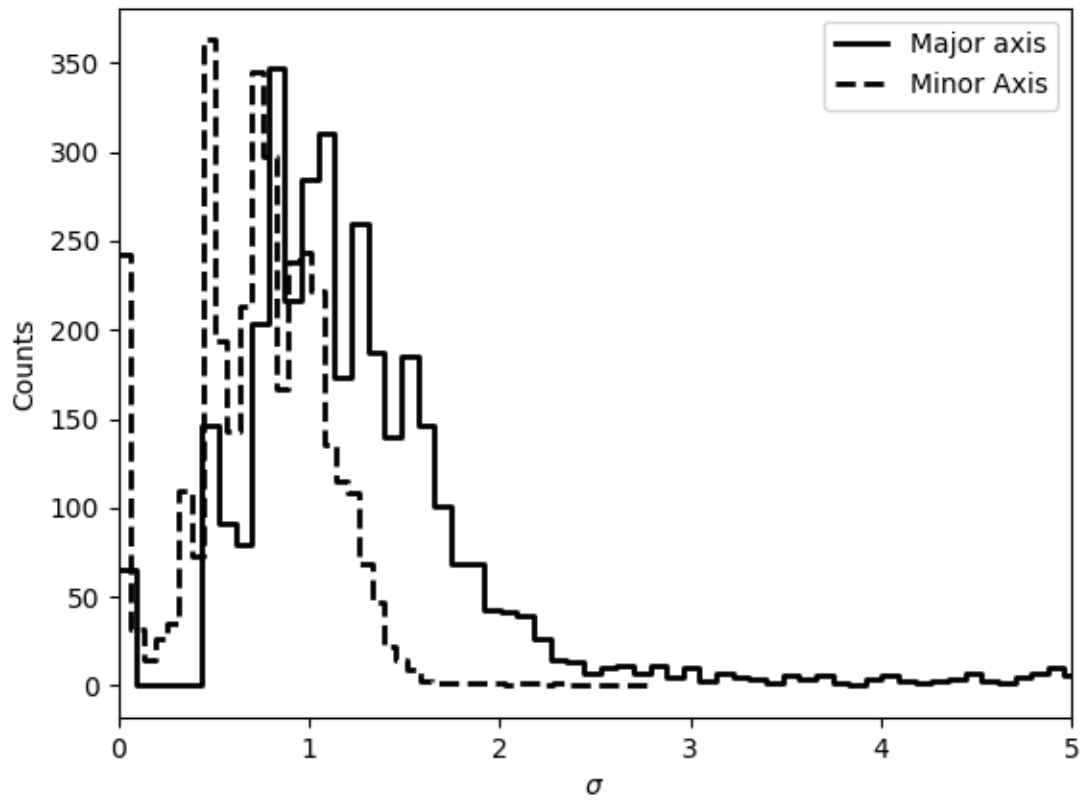


Figure C.15: A histogram of the eigenvalues of the covariance matrix, sorted by the major and minor values. The prior used here was prior “b.” When we compare this to Fig. C.4a, we see that while generated images look reasonable individually, the FoM illustrates that the generator’s approximation of the prior is not faithful.

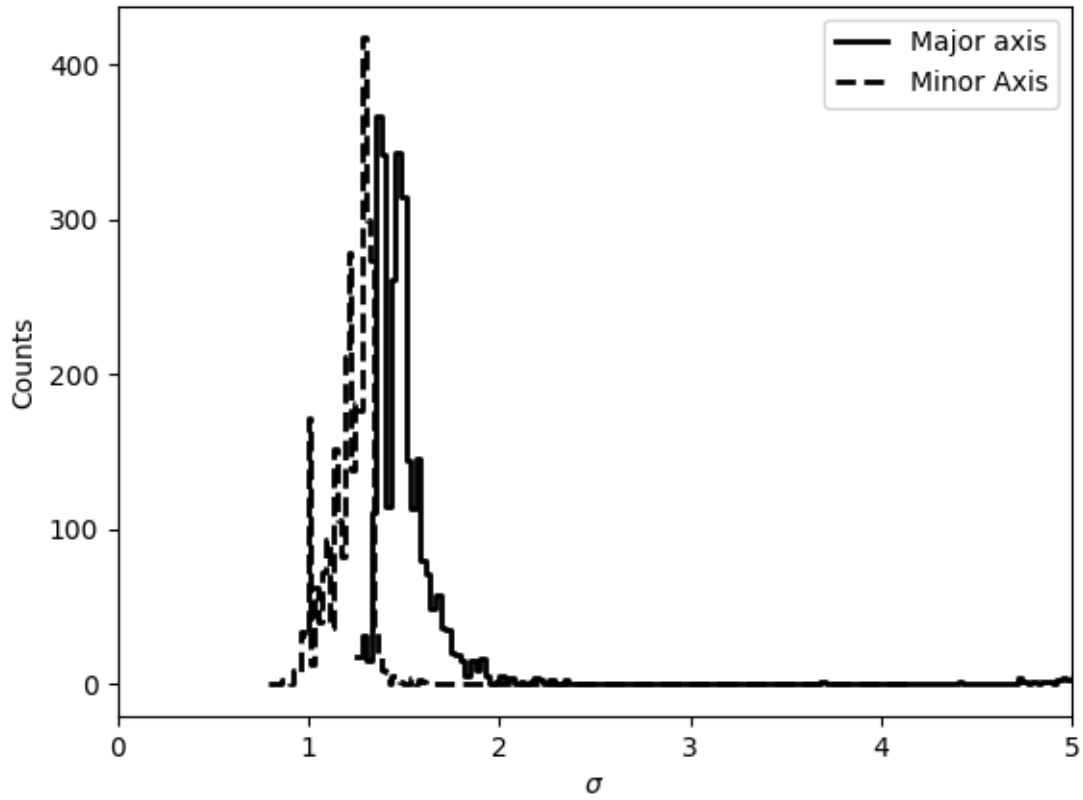


Figure C.16: Using prior “d” and a successfully trained GAN using $d_{\text{latent}} = 85$ and n_{critic} , we see that removing the complication of translational invariance improves the GAN’s ability to approximate the prior.

Appendix D

An analytic model for supernova neutrinos

In this chapter, I outline the work I have done on the development of an analytic model for neutrino output from core-collapse supernova. The work was conducted in collaboration with Prof. Shunsaku Horiuchi, Prof. Tomoya Takiwaki, and Prof. James P. Kneller.

D.1 Introduction

Core-collapse supernovae (CCSNe) are magnificent explosions of Nature marking the end of a massive star's life and the birth of neutron stars and black holes [231–233]. In addition to compact objects, CCSNe are also important contributors to the element enrichment of the Universe as well as the energy and cosmic rays injected into the interstellar medium. They are also natural laboratories for a broad range of physics well beyond terrestrial probes, including beyond Standard Model physics [234–238]. Given the important roles played by CCSNe in a variety of contexts, a significant effort has been made to understand the underlying explosion mechanisms and, amongst others, the importance of nuclear physics [239, 240], hydrodynamics [241–245], gravity [246, 247], fluid instabilities [248–250], and their progenitors [251–253].

The majority of the energy liberated by the gravitational collapse of a massive star's core

is emitted in the form of neutrinos. Therefore, neutrinos and neutrino transport are key pieces of the puzzle in explaining why massive stars explode [254–256]. Observationally, the release of a huge number of neutrinos was cemented by neutrino detections from SN1987A, and those neutrinos continue to guide and inform research thirty years later today, e.g., [257, 258]. Simulations have advanced substantially with state-of-the-art three-dimensional numerical simulations incorporating detailed neutrino transport (for reviews, see [241, 259]) and the identification of many opportunities to test hypotheses [260]. On the experimental side, efforts have been made to prepare for the next core collapse in the Milky Way galaxy, including not only next generation neutrino detectors but also the formation of a coalition of neutrino detectors [261–263] and software integration to incorporate signals from multiple detectors as well as other multi-messenger sources [264, 265].

While it has long been recognized that the next supernova in the Galaxy will greatly advance our understanding of how stars explode, the neutrino, nuclear physics under extreme conditions and much more, before we can extract these information from observation of a nearby supernova, we must first obtain the data. The data we obtain from electromagnetic telescopes would be enhanced if we could quickly extract information about the progenitor and basic explosion properties from the neutrino signal [266, 267]. The most straightforward way to extract this information from the neutrinos is via the comparison of observation and theory, and this comparison would be easiest if we have a collection of simulations covering a broad range of inputs physics and progenitors for comparison. However, while the number of high-fidelity multi-dimensional simulations are increasing [255, 268–283], obtaining a large enough set of computationally expensive in-depth simulations is challenging. An alternative way to address this is to utilize large collections of computationally cheaper spherically symmetric or axisymmetric simulations, which have been performed to study systematic trends, e.g., [284–293]. This ‘template matching’ of a neutrino signal relies upon the landscape of

simulations encompassing all possible signals, something which is difficult to ensure and a potential weakness of the approach.

Here, we explore yet another strategy of comparing theory and observation: analytic models with adjustable parameters. While analytical models cannot compete with individual numerical simulations, they can be used to sample quickly over a vast range of possibilities; once narrowed down, simulations can take over. Analytic and semi-analytic models of CC-SNs have been developed in the literature with this and other considerations in mind. One of them is Müller et al. (2016) [294] (hereinafter, M16), where the authors developed an analytic model which takes as an input the progenitor mass density structure, and predicts from it the mass accretion and properties of neutrino-driven explosions, e.g., neutrino energetics, the explodability (how easy it is to explode), and baryonic masses of the remnant compact star. While hydrodynamical simulations (1D and multi-dimensional) were needed to obtain scaling relations, the appeal of M16’s model is that it only requires scaling relations and simple differential equations with a handful of parameters (see also [295, 296] for the recent application of M16). In parallel, analytic models have also been developed focusing on the late-time cooling phase of supernova neutrino emissions [297]. The application of an analytic model in the context of early supernova neutrino emission would extend previous studies to the mass accretion phase of a CCSN and the explosion mechanism. Furthermore, analytic prescriptions for neutrino oscillations are now also being implemented [265].

In this appendix, we construct an analytic model covering the first ~ 1 second of the neutrino emission from CCSN. Until now, analytic models have focused more on the initial collapse and late-time cooling, with the intermediate phase largely absent. To this end, we start with the analytic model of M16 and confirm its main predictability, and then make improvements to the neutrino emission. We then confirm its main features are consistent with multi-dimensional simulations, as well as a larger sample of numerical simulations of core collapse.

This appendix is organized as follows. In Section D.1, we begin by discussing the model of M16, before delving into our modifications. We focus on improving the model to predict the core bounce times and neutrino emissions for a generic massive star. We then compare this analytic model to simulations to test its validity in Section D.2, and conclude in Section D.3. We have also prepared a Jupyter notebook demonstrating our analytic model, which is publicly available at <https://github.com/SNEWS2/snewpy>.

D.2 Formulation

We first begin by briefly summarizing the analytic model of M16. We note places where we have encountered typos in the original manuscript. Then, we discuss modifications to model neutrino emissions.

D.2.1 Müller’s Analytic Model

One of the main goals of M16 is to determine when shock revival occurs. M16 starts with the infall time t , i.e., the time it takes for a mass shell to reach the neutron star, in order to derive the mass accretion rate and how certain aspects of the collapse, such as the shock front or regions of neutrino heating, form and evolve. In order to see how the shock front evolves, the properties of the shock front are related to physical quantities such as the mass accretion rate ($\dot{M}$) and the neutrino luminosity L_ν . In turn, these physical quantities can either be calculated from first principles or through scaling relations obtained through calibrations to simulations.

The infall time t is computed by assuming it is related to the free-fall time by some constant factor C . The free-fall time of a given mass shell depends on the average density contained

within the shell (i.e., $\bar{\rho} = \frac{1}{V}(\int_0^{r(M)} dM)$). With $C = \sqrt{8/3}$ [294] we get

$$t = \sqrt{\frac{\pi}{4G\bar{\rho}}}. \quad (\text{D.1})$$

The mass accretion rate $\dot{M}$, is then defined to be [298],

$$\dot{M} \equiv \left(\frac{dt}{dM}\right)^{-1} = \frac{-2\bar{\rho}}{t} \left(\frac{d\rho}{dM}\right)^{-1} = \frac{2M}{t} \frac{\rho}{\bar{\rho} - \rho}, \quad (\text{D.2})$$

where $\rho = dM/dV$ is the density of a given mass shell.

With $\dot{M}$ determined, we move on to calculate the neutrino luminosity L_ν . We denote by L_ν the sum of the electron and anti-electron neutrino luminosities. The main contributions to L_ν are from accretion (L_{acc}) and diffusion (L_{diff}). The accretion component comes from the conversion of a mass shell's potential energy into neutrino emission, which can be modeled by the incoming mass through the gain region. The gain radius r_g is found from the equation

$$r_g = {}^3\sqrt{r_1^3 \left(\frac{\dot{M}}{M_\odot s^{-1}}\right) \left(\frac{M}{M_\odot}\right)^{-3} + r_0^3}, \quad (\text{D.3})$$

where $r_1 = 120$ km and $r_0 = 12$ km in order to match simulations from the VERTEXT-CoCoNut [299] and CoCoNuT-FMT [300] codes. Using r_g , the accretion luminosity is defined to be

$$L_{\text{acc}} = \zeta \frac{GM\dot{M}}{r_g}, \quad (\text{D.4})$$

where ζ is a parameter describing the efficiency of the conversion of the accretion energy in the gain region to electron-type neutrino luminosity. In M16, the default value is $\zeta = 0.8$.

The second component of the luminosity, L_{diff} , originates from the inner layers of the proto-neutron star [301, 302]. It depends on the binding energy of the cold neutron star E_{bind}

[303, 304] and the characteristic cooling time-scale τ_{cool} . Respectively, E_{bind} and τ_{cool} are given by

$$E_{\text{bind}} = 0.084 M_{\odot} c^2 (M/M_{\odot})^2, \quad (\text{D.5})$$

$$\tau_{\text{cool}} = \tau_{1.5} \times \left(\frac{M}{1.5 M_{\odot}} \right)^{5/3}, \quad (\text{D.6})$$

where $\tau_{1.5} = 1.2$ sec was determined based on [305]. L_{diff} is then defined as

$$L_{\text{diff}} = 0.3 (E_{\text{bind}}/\tau_{\text{cool}}) e^{-t/\tau_{\text{cool}}} \quad (\text{D.7})$$

where the pre-factor of 0.3 accounts for the fact that about a third of E_{bind} is emitted as electron-type neutrinos and antineutrinos (the remainder going to heavy-lepton neutrino flavors).

Thus, the combined L_{ν} is

$$L_{\nu} = (L_{\text{acc}} + L_{\text{diff}}). \quad (\text{D.8})$$

To account for general relativistic effects since the neutrino emission is occurring around the proto-neutron star (PNS), L_{ν} can be scaled by a redshift factor α ,

$$\alpha = \sqrt{1 - \frac{2GM}{r_{\text{PNS}} c^2}}, \quad (\text{D.9})$$

where $r_{\text{PNS}} \approx (5/7)r_g$ is the radius of the PNS. We use the post-application of α to keep a better track of factors of α and to more easily compare with simulation results.

Table D.1: The adjustable parameters of M16’s analytic model

Parameter	Explanation	Default Value
α_{out}	outflow’s volume fraction	0.5
α_{turb}	shock expansion from turbulence	1.18
β_{expl}	shock compression ratio	4
ζ	accretion to L_{acc} efficiency	0.8
$\tau_{1.5}$	$1.5M_{\odot}$ neutron star cooling time	1.2 s

Shock Revival

With the quantities $(t, \dot{M}, r_g, L_{\nu})$ determined by the model, the shock revival time can now be calculated. To do this, we recall that in the initial collapse of the core, the density of the PNS continually increases due to the accreting matter, until the PNS reaches a critical density due to nucleon repulsion. The rebound of the over-compressed PNS leads to the formation of a shock which then propagates into the infalling matter. This is the so-called prompt bounce shock. This prompt bounce shock stalls due to the energy loss associated with nuclear-disassociation and neutrino emission, and must be energetically revived in order to trigger a supernova explosion. In the neutrino mechanism, the shock revival is achieved through a combination of neutrino heating and fluid turbulence [242, 306–309].

To determine whether neutrinos can revive the shock, two competing effects are taken into consideration. The first is advection, or the transfer of heat away from the shock interface; the second is heating, or the energy transfer between emitted neutrinos from the PNS and the shock region. The shock revival time t_{sr} is then determined based on when the neutrino heating overtakes advection. This can be done by comparing the time-scales of these competing processes.

To this end, M16 starts by using the scaling relations observed in [310] to define the radius

of the shock front r_{sh} as¹

$$r_{\text{sh}} = \alpha_{\text{turb}} \times 0.55 \text{ km} \times \left(\frac{\alpha^3 L_\nu}{10^{52} \text{ erg s}^{-1}} \right)^{4/9} \times \left(\frac{M}{M_\odot} \right)^{5/9} \left(\frac{r_g}{10 \text{ km}} \right)^{16/9} \left(\frac{\dot{M}}{M_\odot \text{ s}^{-1}} \right)^{-2/3}, \quad (\text{D.10})$$

where the turbulence parameter α_{turb} is set to 0.86 based on results from simulations [241, 300, 311–314]. Also, simulation results from [300] show that the advection time scale, $\tau_{\text{adv}} \equiv M_g / \dot{M}$ (where M_g is the mass of the gain region), can be modeled as

$$\tau_{\text{adv}} \approx 18 \text{ ms} \left(\frac{r_{\text{sh}}}{100 \text{ km}} \right)^{3/2} \left(\frac{M}{M_\odot} \right)^{-1/2} \ln \left(\frac{r_{\text{sh}}}{r_g} \right). \quad (\text{D.11})$$

The neutrino heating time scale is characterized by the average mass-specific neutrino heating rate $\dot{q}_\nu$ and the average net binding energy e_g . Scaling laws from [300, 306, 310] relate $\dot{q}_\nu$ to L_ν , $E_\nu \propto M$, and r_g through the following equation:

$$\dot{q}_\nu \propto \frac{L_\nu E_\nu^2}{r_g^2}. \quad (\text{D.12})$$

For the average binding energy, M16 uses [255, 300, 310] to model it as

$$e_g = \epsilon_{\text{therm}} - \frac{GM}{\max(r_g, r_{\text{sh}})}, \quad (\text{D.13})$$

where the thermal energy per unit mass ϵ_{therm} is related to the dissociation energy ϵ_{diss} , or the energy lost by the shock-wave from breaking up nuclei per unit mass, by Bernoulli's

¹An extra factor of α was required to match M16.

theorem,

$$\epsilon_{\text{therm}} = \frac{3}{4} \left(\frac{GM}{\max(r_g, r_{\text{sh}})} - \epsilon_{\text{diss}} \right). \quad (\text{D.14})$$

Hence the average net binding energy $|e_g|$ is²

$$|e_g| = \frac{3}{4} \epsilon_{\text{diss}} + \frac{GM}{4 \max(r_g, r_{\text{sh}})}. \quad (\text{D.15})$$

M16 then sets

$$\begin{aligned} \tau_{\text{heat}} = 150 \text{ms} \times & \left(\frac{|e_g|}{10^{19} \text{erg g}^{-1}} \right) \left(\frac{r_g}{100 \text{km}} \right)^2 \\ & \times \left(\frac{\alpha L_\nu}{10^{52} \text{ergs s}^{-1}} \right)^{-1} \left(\frac{M}{M_\odot} \right)^{-2} (\alpha)^{-2} \end{aligned} \quad (\text{D.16})$$

With the two characteristic time-scales determined, the ratio between τ_{adv} and τ_{heat} can be used as a proxy for $\dot{q}_\nu/\dot{q}_{\text{adv}}$ to determine the shock revival time t_{sr} . At early times, advection is the dominate effect so $\tau_{\text{adv}}/\tau_{\text{heat}} < 1$. At later times, if shock revival has occurred, heating is expected to overtake advection and become dominant, so $\tau_{\text{adv}}/\tau_{\text{heat}} > 1$. Therefore, the infall time corresponding to the moment when $\tau_{\text{adv}}/\tau_{\text{heat}} = 1$, is the shock revival time. For progenitors that never reach $\tau_{\text{adv}}/\tau_{\text{heat}} = 1$, M16 assumes shock revival never occurs and a black hole is formed instead.

Finally, M16 defines an efficiency parameter $\eta_{\text{acc}} \equiv \dot{Q}_\nu/\dot{M} = \dot{q}_\nu M_g/\dot{M}$ for convenience. Using the scaling relations of $\dot{q}_\nu$ in Eq. (D.12) and the definition of τ_{adv} , it can be shown that³

$$\eta_{\text{acc}} \equiv \frac{\dot{Q}_\nu}{\dot{M}} = |e_g| \frac{\tau_{\text{adv}}}{\tau_{\text{heat}}}. \quad (\text{D.17})$$

²Note that we also set $r_{sh} \rightarrow \max(r_g, r_{sh})$ to account for the entire gain region.

³Alternatively, one can use [315] to determine a reasonable constant of proportionality for Eq. (D.12); either approach yields similar results for important quantities, such as t_{sr} .

Explosion Phase

For progenitors which have undergone shock revival, M16 then proceeds to calculate the explosion properties, in particular, the “diagnostic explosion energy” E_{diag} , which is the total energy of the unbounded material at various instances of the explosion.

The explosion phase is split into two sub-phases. Phase I is characterized by ongoing mass accretion even after shock revival, which powers the explosion’s energy through neutrino-heated outflows driven by high L_{acc} . Phase II is when the residual mass accretion ceases, and the explosion energy’s dominant contribution comes from neutrino cooling of the PNS. The transition between these two sub-phases is estimated by finding the moment when the speed of the shocked material v_{sh} becomes equal to the local escape velocity [255]. The criterion for accretion subsiding is

$$\frac{\beta_{\text{expl}} - 1}{\beta_{\text{expl}}} v_{\text{sh}} = \sqrt{\frac{2GM}{r}}, \quad (\text{D.18})$$

where β_{expl} is a parameter that accounts for the fluid compression that occurs at the shock front. Since M16 assumes spherical symmetry, v_{sh} can be calculated using the following result from [255, 316]:

$$v_{\text{sh}} = 0.794 \left(\frac{E_{\text{imm}}}{M - M_{\text{ini}}} \right)^{1/2} \left(\frac{M - M_{\text{ini}}}{\rho r^3} \right)^{0.19}, \quad (\text{D.19})$$

where M_{ini} is the initial mass cut defined as the mass coordinate associated with t_{sr} , E_{imm} is diagnostic energy of a mass shell as the shock passes by, and density ρ and radius r are in reference to the initial progenitor model (i.e., the same as used in the shock revival calculation).

M16 deduces using [317, 318] that in phase I, E_{diag} and E_{imm} can be related to $\dot{M}$ and the

recombination energy $\epsilon_{\text{rec}} \approx 8.8 \text{ MeV}$ [255] by

$$\frac{dE_{\text{diag}}}{dM_{\text{sh}}} = \frac{(1 - \alpha_{\text{out}})\epsilon_{\text{rec}}\eta_{\text{acc}}}{|e_g|} + \alpha_{\text{out}}(\epsilon_{\text{bind}} + \epsilon_{\text{burn}}) \quad (\text{D.20})$$

$$\begin{aligned} \frac{dE_{\text{imm}}}{dM_{\text{sh}}} &= \frac{\epsilon_{\text{rec}}\eta_{\text{acc}}}{|e_g|} \min\left(1, \frac{\dot{M}}{4\pi r^2 v_{\text{sh}} \rho}\right) \\ &+ \alpha_{\text{out}}(\epsilon_{\text{bind}} + \epsilon_{\text{burn}}), \end{aligned} \quad (\text{D.21})$$

where $\alpha_{\text{out}} = 0.5$ is the surface fraction occupied by neutrino-driven outflows, $\epsilon_{\text{bind}} = \epsilon_{\text{in}} - \frac{GM}{r}$ is the binding energy per unit mass of the unshocked material⁴, and ϵ_{burn} is the energy released from nuclear burning.

Using the initial and final mass fractions X_i and X'_i , respectively, ϵ_{burn} at a given moment is defined as

$$\epsilon_{\text{burn}} = \sum_i (X_i - X'_i) \epsilon_{\text{rm},i}, \quad (\text{D.22})$$

where $\epsilon_{\text{rm},i}$ is the rest-mass contributions per unit mass for nucleus⁵ i . To determine X_i , M16 applies the “flashing” method of [320], where different burning processes (C-, O-, Si-burning) occurs instantaneously at different ignition temperatures; determining the post-shock temperature at a given mass shell T_{sh} would then dictate what burning process is occurring at said mass shell. Derived from the Rankine-Hugoniot jump conditions, assuming the pre-shock region’s density ρ can be related to the post-shock density ρ_{sj} by a factor of β_{expl} , T_{sh} is obtained to be

$$T_{\text{sh}}^4 = \frac{3\beta_{\text{expl}} - 1}{a\beta_{\text{expl}}} \rho v_{\text{sh}}^2, \quad (\text{D.23})$$

where a is the radiation constant. X'_i is then determined based on T_{sh} in the following prescription:

⁴The internal energy ϵ_{in} is a quantity provided in the progenitor model.

⁵The list of elements used for this work are the elements given in the [319]’s progenitors.

- For $T_{\text{sh}} \in [2.5, 3.5) \times 10^9 \text{ K}$, elements lighter than O burn to $i' = \text{O-16}$.
- For $T_{\text{sh}} \in [3.5, 5) \times 10^9 \text{ K}$, elements lighter than Si burn to $i' = \text{Si-28}$
- For $T_{\text{sh}} \in [5 \times 10^9, T_\alpha) \text{ K}$, all elements burn to $i' = \text{Ni-56}$.

where T_α is a density-dependent temperature for which the mass fraction of α -particles reaches 0.5 in nuclear statistical equilibrium [320, 321] and is given by

$$\log_{10} \rho = 11.62 + 1.5 \log_{10} \left(\frac{T_\alpha}{10^9 \text{ K}} \right) - 39.17 \left(\frac{T_\alpha}{10^9 \text{ K}} \right)^{-1}. \quad (\text{D.24})$$

Eq. (D.22) can be rewritten based on the assumptions of the flashing method in the following fashion:

$$\epsilon_{\text{burn}}(T_{\text{sh}}) = \sum_{i < i'} X_i (\epsilon_{\text{rm},i} - \epsilon_{\text{rm},i'}(T_{\text{sh}})) \quad (\text{D.25})$$

Using Eq. (D.21, D.19, D.23, D.22), v_{sh} and E_{diag} can now be solved iteratively to determine the transition from Phase I to II of the explosion using Eq. (D.18). Since the mass accretion has diminished during phase II, Eq. (D.20) is

$$\frac{dE_{\text{diag}}}{dM_{\text{sh}}} = \epsilon_{\text{bind}} + \epsilon_{\text{burn}}$$

for the remainder of the explosion.

Finally, the baryonic mass of the PNS M_{by} can be modeled through the explosion as

$$\frac{dM_{\text{by}}}{dM_{\text{sh}}} = (1 - \alpha_{\text{out}})(1 - \eta_{\text{acc}}/|e_g|) \quad (\text{D.26})$$

up to Phase I. During Phase II, M_{by} remains constant since mass accretion has ceased.

In Table D.1, we summarize the parameters of the analytic model of M16, along with their

physical description and default values set by matching to results of simulations.

D.2.2 Adding Neutrinos To The Model

Building off the success of M16’s analytic model, we now add the physics to better describe the neutrino emission. We first investigate the bounce time (t_b), then modify $\dot{M}$ and L_ν to better agree with simulations. We discuss our calibration and predictions for these additions in Section D.3.

Bounce Time

One of the important physical quantities when accounting for the neutronization contribution to the neutrino luminosity is the bounce of the inner core. During the initial collapse of the star, the density of the inner core continually increases until it reaches a saturation density ($\rho_0 \approx 3 \times 10^{14} \text{ g cm}^{-3}$) at which point the equation of state of the core stiffens due to nucleon repulsion. This stiffening creates a sonic point near the inner-core, which causes the accreting matter to rebound or “bounce” off the core [322, Ch. 11]. The bounce time t_b is then defined to be the moment in time when the inner-core reaches this saturation density.

As seen from [323, 324], the time at which the inner-core reaches the saturation density is also when the PNS radius (r_{PNS}), temperature, and electron fraction change dramatically. Thus, rather than modeling the baryon density of the inner core, we use the PNS radius as defined in M16’s model, $r_{\text{PNS}} \approx (5/7)r_g$, as a proxy to predict the bounce time. The bounce time is then defined to be the time at which r_{PNS} reaches some “critical radius” (r_{crit}), which can be calibrated to match simulations.

Table D.2: Additional adjustable parameters in the modified analytic model

Parameter	Explanation	Default Value
r_{crit}	critical PNS radius	66 km
τ_d	decay time	0.05 s
r_{thres}	threshold radius for normalization	500 km
L_{NB}	neutronization burst amplitude	4 foe/s
σ_{NB}	length of neutronization burst	5 ms
p	late-time exponential power	0.4
τ_{av}	time constant for $\langle E_{\text{av}} \rangle$	0.02 s

Mass Accretion Rate ($\dot{M}$)

While M16's definition of mass accretion rate $\dot{M}$ is successful in predicting quantities such as the shock revival time and PNS baryonic mass, its definition causes an artificially high accretion rate in the early phase of the collapse. For example, from Eq. (D.2), $\dot{M}$ is singular for $t \ll 1$. However, when comparing to simulations [324–326], accretion should start at $\dot{M}/M_\odot \sim 0$. Additionally, the $\dot{M}$ calculated in simulations actually peaks near the bounce time and then decreases.

To account for these features, we modify Eq. (D.2) by defining

$$\dot{M}_{\text{mod}} = \frac{2M}{D(t)} \frac{\rho}{\bar{\rho} - \rho} \quad (\text{D.27})$$

where

$$D(t) = (1 \text{ sec}) \ln(1 + e^{-(t-\tau_b)/\tau_d}) + (t - \tau_b) + t_{\text{off}}. \quad (\text{D.28})$$

In Eq. (D.28), $\tau_d = 0.04$ sec controls how quickly the modified mass accretion rate $\dot{M}_{\text{mod}}$ converges to M16's definition; $\tau_b = t_b - \tau_d \ln(\frac{1-\tau_d}{\tau_d})$ is determined so the modified mass accretion peaks at the bounce time; and t_{off} is chosen to conserve the mass defined to be $M_{\text{conser}} = \int_{t(r_{\text{thres}})}^{t_{\text{sr}}} \dot{M} dt$, where r_{thres} is radius of the mass shell we start integrating from and Eq. D.2 was used for $\dot{M}$. The reason for the inclusion of r_{thres} is because in Eq. (D.2), $\dot{M}$

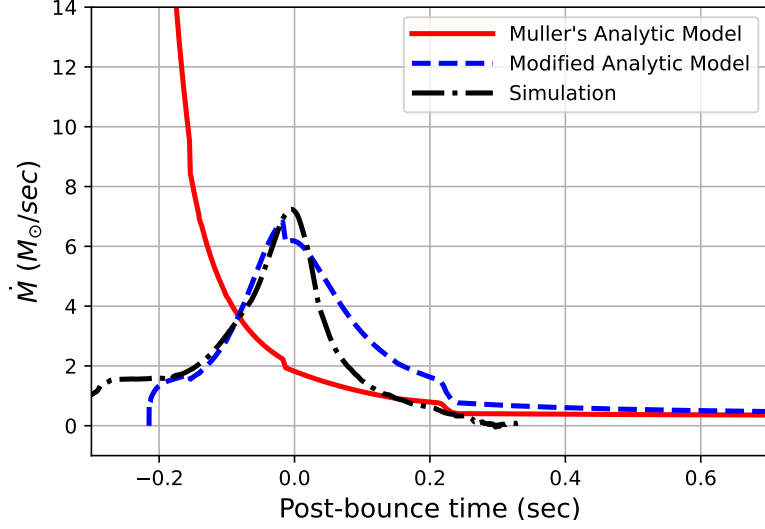


Figure D.1: Mass accretion rate $\dot{M}$ vs. post-bounce time from M16 in solid red, our modified model in dashed blue, and numerical simulation in dashed-dotted black, all for the core collapse of the s20 progenitor of WH07 [327].

is singular for $t_{\text{ff}} \sim 0$. Thus, the prescription conserves the total mass accreted and only redistributes it in time. Figure D.1 compares the mass accretion rate $\dot{M}$ from M16, i.e., from Eq. (D.2), and our new modified model, i.e., Eq. (D.27), to a numerical simulation (see Section D.3.1 for details).

Neutrino Energetics

Our modifications to the neutrino energetics (i.e., L_ν and related quantities) include the addition of the neutronization burst, propagation of the modification to the mass accretion rate described in Sec. D.2.2, and modeling the late-time luminosity.

The first modification is to include the neutronization burst into the electron neutrino luminosity, L_{ν_e} . Since we have determined the bounce time, the neutronization burst is defined to be $L_{NB} = (4 \times 10^{51}) \text{ erg s}^{-1} \times \mathcal{N}(\mu = t_b, \sigma_{NB} = 5\text{ms})$ where $\mathcal{N}(\mu, \sigma)$ is a normalized

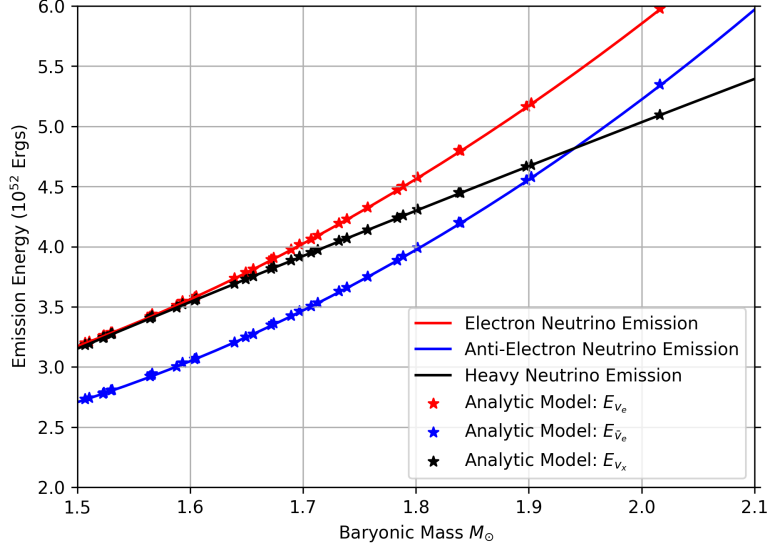


Figure D.2: Parabolic fits to the total neutrino emission energies from Table 1 of Nakazato et al. (2013) [328], shown in solid colored lines (red, blue, black for ν_e , $\bar{\nu}_e$, and ν_x , respectively). The star symbols are obtained from our analytic model computed using the progenitor set of [319] and plotted at the corresponding final remnant baryonic mass.

Gaussian. The features of the neutronization burst appear to be largely independent of the progenitor [324–326, 328, 329], so the normalization and σ are taken to be the same for all progenitors. Thus, we define

$$L_{\nu_e} = L_{\nu}/2 + L_{NB} = (L_{\text{acc}} + L_{\text{diff}})/2 + L_{NB} \quad (\text{D.29})$$

$$L_{\bar{\nu}_e} = L_{\nu}/2 = (L_{\text{acc}} + L_{\text{diff}})/2 \quad (\text{D.30})$$

$$L_{\nu_x} = L_{\text{diff}}/2, \quad (\text{D.31})$$

where L_{ν_e} and $L_{\bar{\nu}_e}$ is defined to remain consistent with M16’s definition and $L_{\nu_x} = (L_{\nu_\mu} + L_{\bar{\nu}_\mu} + L_{\nu_\tau} + L_{\bar{\nu}_\tau})/4$ should only depend on diffusion [330].

The second modification we make to L_{ν_i} is to propagate $\dot{M}_{\text{mod}}$, which amounts to modifying r_g and $\dot{M}$ in Eq. (D.4). This has the effect of slightly suppressing L_{ν_i} for $t < t_{\text{bounce}}$ (due to

the lower mass accretion) and also decreasing $E_{\nu_i, \text{total}}(t) = \int_0^t L_{\nu_i}(t') dt'$ for $t < t_{\text{sr}}$.

The final modification is related to the late-time cooling luminosity. Simulations [323, 328, 331–333] show that L_{ν_i} for $\nu_i \in \{\nu_e, \bar{\nu}_e, \nu_x\}$ can be approximately modeled from the shock revival until ~ 20 sec by the following:

$$L_{\nu_i}^{\text{cool}}(t > t_{\text{sr}}) = L_0 \exp\left(-\left[\frac{\Delta t}{\tau_{\text{LC}}}\right]^p\right) \quad (\text{D.32})$$

where L_0 is the luminosity at shock revival, $L_0 \equiv L_{\nu_i, \text{mod}}(t_{\text{sr}})$, $\Delta t \equiv t - t_{\text{sr}}$, the index p is a fixed non-zero number, and τ_{LC} is the late-time cooling time scale. Many simulations, including [324–326, 328, 334], indicate that p is approximately progenitor independent, so we adopt $p = 0.4$ as the default value across progenitors. The timescale τ_{LC} should be derived from the neutrino-cooling PNS, but it can also be determined by fixing the integrated neutrino luminosity $E_{\nu_i} \equiv \int L_{\nu_i} dt$ to match the total gravitational binding energy liberated. For a given target integrated luminosity E_{ν_i} and fixed value of p , this means τ_{LC} is uniquely set,

$$\tau_{\text{LC}} = \frac{p \left(E_{\nu_i} - \int_0^{t_{\text{sr}}} L_{\nu_i} dt \right)}{L_0 \cdot \Gamma(1/p)}, \quad (\text{D.33})$$

where $\Gamma(x)$ is the gamma function and the luminosity given by our analytic model (Eqs. D.29–D.31) are used for the pre-cooling phase. As shown in [328, Table 1], E_{ν_i} is strongly dependent on M_{by} . This is because more massive neutron-star masses lead to larger release of gravitational binding energy, the majority of which should be emitted as neutrinos. We approximate the relation between E_{ν_i} and M_{by} by a polynomial. We use a parabola of the form $p(M_{\text{by}}) = a + b(M_{\text{by}}/M_{\odot}) + c(M_{\text{by}}/M_{\odot})^2$ where a , b , and c for the electron neutrino emission energy interpolation are 17.844, -12.846, and 4.7229, respectively. Similarly, the same constants for the anti-electron neutrino emission energies are 20.566, -14.252, and 5.2278 and the constants for the heavy neutrino emission energy are 13.357, -3.1527, and 1.3748. The

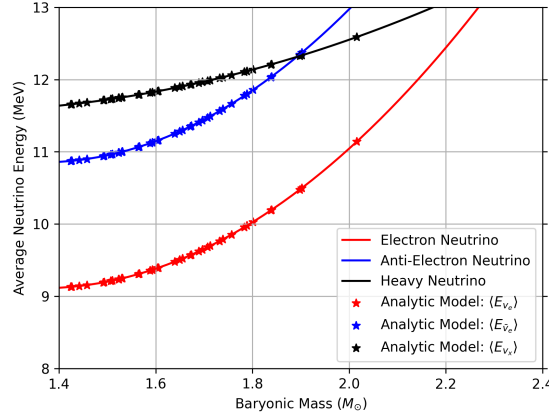


Figure D.3: Same as Fig. D.2 but for neutrino mean energies.

resulting correlations, separately for ν_e , $\bar{\nu}_e$, and ν_x , are shown in Fig. D.2 as solid lines. The star symbols in Fig. D.2 are the result of applying the modified M16 model to the progenitor set of Ref. [319] and plotting the resulting M_{by} vs. E_{ν_i} . Since we fixed τ_{LC} to match the prediction E_{ν_i} dependence on M_{by} in [328], which is illustrated in as the solid lines in Fig. D.2, we expect the our model to follow the same trend, which the figure confirms.

The last addition we include to our modified analytic model is the mean neutrino energy $\langle E_{\nu_i} \rangle$. Using [328] to determine the dependence of $\langle E_{\nu_i} \rangle$ as a function of M_{by} , we approximate its time dependence with the following:

$$\langle E_{\nu_i} \rangle = \langle E_{\nu_i}(M_{\text{by}}) \rangle \cdot \left(1 + \exp \left(-\frac{t - t_b}{\tau_{\text{av}}} \right) \right)^{-1} \quad (\text{D.34})$$

with $\tau_{\text{av}} = 0.02$ s. Fig. D.3 shows the fit.

Our modifications to M16 introduce several additional adjustable parameters to the analytic model, summarized in Table D.2 along with their nominal values.

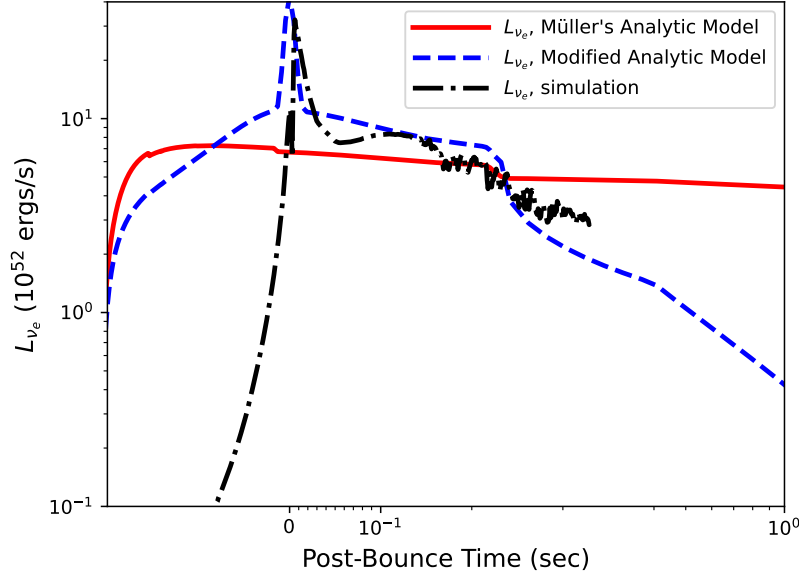


Figure D.4: Neutrino luminosity L_{ν_e} versus post-bounce time from M16 (red), our modified model (blue), and the simulation (black). All are computed based on the s20 progenitor from WH07 [327]. The modified analytic model does significantly better in matching simulation results. Note the time-axis switch at $t = 0$.

D.3 Results

In this section, we begin calibrating the modified analytic model with a numerical simulation of core collapse. After calibration, we apply our modified analytic model to larger sets of progenitors to study global behaviors and assess the model's ability to make reliable predictions.

D.3.1 Calibration

We perform a comparison of our modified analytic model to the original M16 model and numerical simulation of core collapse by the 3DnSNE code [335] using the $20M_{\odot}$ progenitor from WH07 [327].

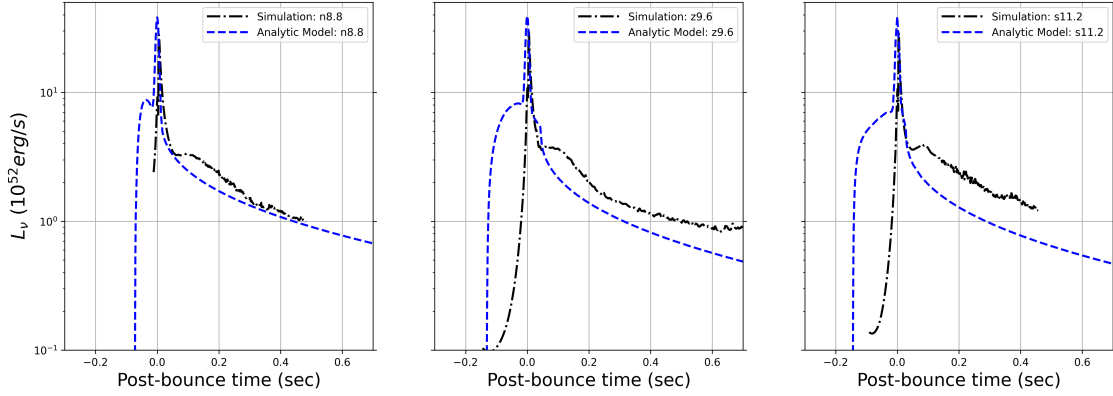


Figure D.5: Comparison of neutrino luminosities predicted by our modified analytic model (dashed lines) to numerical simulations (dot-dashed lines) for three different progenitors, from left to right: n8.8 [325], z9.6 [326], and s11.2 [326]. All these predictions from the analytic models use parameters calibrated to the numerical simulation of the core collapse of the s20 progenitor (see text).

The numerical setup is similar to [324]. The resolution of the spherical polar grid is $512 (r) \times 128 (\theta)$, which cover 0–5000 km and $0-\pi$, respectively. Neutrino transport is solved by using the isotropic diffusion source approximation [268, 336]. The employed set of neutrino reactions is the same as that in [337]. We use an EOS derived with a variational method [338]. The gravitational wave emission of this model is discussed in [339, 340].

Figure D.1 compares the mass accretion rate $\dot{M}$ from M16, i.e., Eq. (D.2) and our modified model, i.e., Eq. (D.27), compared to the numerical simulation. The values for new parameters in the modified model, $r_{\text{crit}} = 66$ km, $r_{\text{thres}} = 500$ km and $\tau_d = 0.05$ s, were chosen to closely match the simulation’s profile at early times (i.e., before the bounce time). After bounce, the modified model (dashed blue curve) is somewhat larger than the simulation (dot-dashed black curve), but is a significant improvement over the M16 model (red curve). The deviation from the simulation can be attributed to our choice of delay function $D(t)$ converging to $1/t$ too slowly. However, by construction, the modified model converges to what is expected at times around shock revival.

Figure D.4 shows a comparison of the resulting electron neutrino luminosity L_{ν_e} , using Eq. (D.32) for the late-time luminosity. The modified model's prediction of L_{ν_e} near the bounce time and shock revival time is similar to the numerical simulations. Although having $\dot{M}_{\text{mod}}(t \sim 0)/M_{\odot} \sim 0$ was not enough to suppress $L_{\nu_i}(t < t_b)$, it is a significant improvement over the M16 model (red line). The shock revival times are very similar, being 0.25 seconds post-bounce in numerical simulation versus 0.23 seconds in the M16 model, highlighting the good predictability of the analytic model at this epoch.

D.3.2 Comparisons

We now compare our modified analytic model, calibrated to s20 as described in the previous section, with simulations of other progenitors. The aim is to test whether the modified model is able to reproduce the neutrino emission for other progenitors.

Figure D.5 shows the comparison with numerical simulations of the core collapse of four other progenitors: n8.8 [325], z9.6 [326], and s11.2 [326]. We refer the reader to the relevant publications for details of the simulations. Here, we select lighter progenitors with density profiles different from our fiducial s20. We notice that similar to Fig. D.4, the pre-bounce features are somewhat over-estimated, but from the bounce time onward, the general trends with our default modified neutrino luminosities given by Eqs. (D.29)–(D.31) and value of p used in Eq. (D.32) are consistent with simulations. Also, the simulation's bounce-time for the n8.8, z9.6, and s11.2 progenitors are 0.053 sec, 0.186 sec, and 0.132 sec. For the analytic models, the bounce-times are 0.0893 sec, 0.181 sec, and 0.174 sec. The average time difference between the simulation and analytic model, $|\overline{\Delta t_b}| \sim 0.03$ sec illustrates the success of the analytic model. Similarly, the shock revival times are 0.27, 0.17, and 0.23 seconds post bounce for the n8.8, z9.6, and s11.2 simulations, respectively, while the analytic model

values are 0.012, 0.04, and 0.028 seconds post bounce. Thus, the average time difference is a mere 0.17 seconds.

We now move from one-on-one progenitor comparisons to global trends predicted by the analytic model. For this, progenitors from the WHW02 suite [319] were used. Specifically, we compare the bounce times found by the systematic simulations study by [341]. In [341], the progenitors were quantified according to their core compactness, defined

$$\xi_{M/M_\odot} = \frac{M/M_\odot}{r(M)/1000\text{km}} \quad (\text{D.35})$$

with $M = 2.5M_\odot$. In Fig. D.6, we therefore compare our bounce times (blue stars) to those found by [341] (black triangles), all plotted against the core compactness $\xi_{2.5}$. We find that the analytic model is able to predict the bounce time well. As shown in the bottom panel, the difference is typically within $\sim 10\%$, and furthermore, the model reproduces the overall trend of the dependence on the progenitor core compactness.

Next, we compare the energy liberated in the form of neutrinos during the first 1 second, i.e., during the accretion phase and in many cases slightly beyond. We perform comparisons separately for the electron-type neutrinos and heavy-lepton neutrinos, and show the results in Fig. D.7. Overall, the analytic model is able to capture the hierarchy of the neutrino flavors and the rise with progenitor compactness seen in [341], but tends to overestimate the neutrino emission by $\sim 10 - 20\%$ for the electron type neutrinos, and a slightly larger error for the heavy-lepton neutrinos.

The correlation between the compactness and the energy liberated by the neutrinos has been pointed out in the literature (e.g., see Fig. 4 of [330]). From the total energy, $2E_\nu + 4E_{\nu_x}$, one can estimate the PNS mass [342], and other correlations of observables with supernova properties are also studied in [267]. Our method could be a powerful tool to obtain an

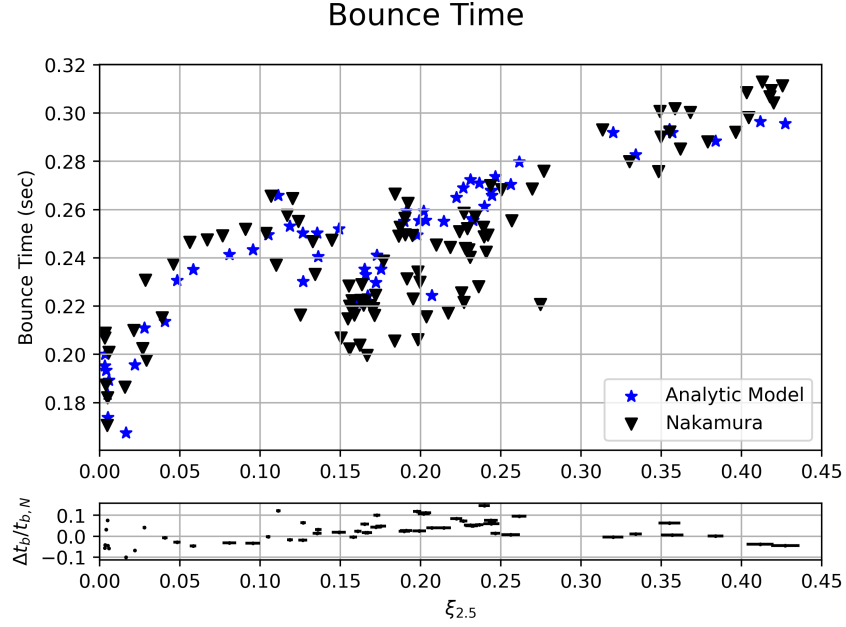


Figure D.6: Comparison of the predicted bounce time (relative to the initial collapse) vs. $\xi_{2.5}$ with modified model in blue and [341] in black. The solid lines in the bottom plot represent the difference in bounce times between our model and Nakamura (i.e., $\Delta t_b = t_{b,\text{mod}} - t_{b,\text{Nakamura}}$).

empirical relation between observable and progenitor structure from a number of models.

D.4 Discussion and Conclusions

In this section, we developed an analytic model of CCSN neutrino emission with a focus on the first few seconds post-collapse for a given progenitor profile. We base our model on the previous work of M16 [294] and modified / extended it to better match the emission seen in multi-dimensional CCSN simulations. Specifically, we model the bounce time, which allows us to attach a neutronization burst at the appropriate time. We also modify the mass accretion rate, which allows us to improve the modeling of the accretion neutrino emission. Further, we combine a progenitor-dependent model for the late-time cooling neutrino

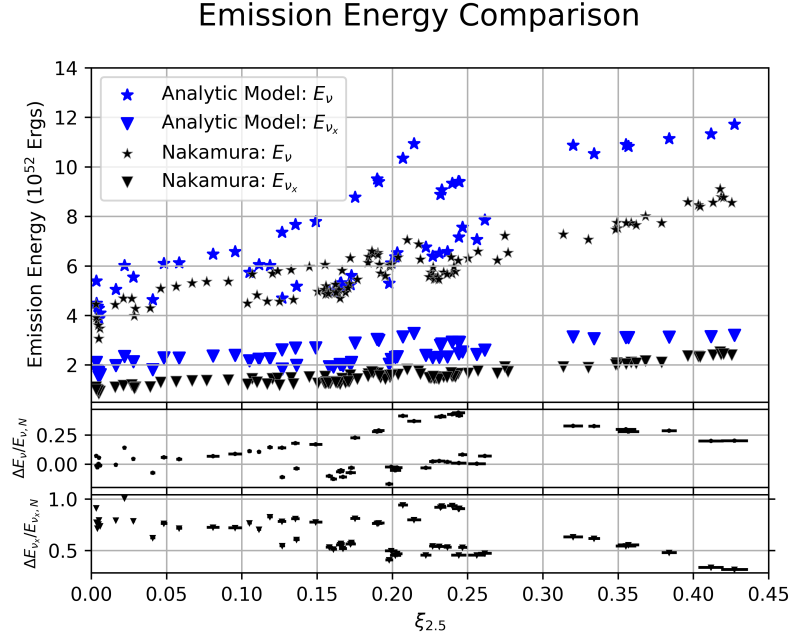


Figure D.7: Comparison of the predicted neutrino emission during the first 1 second post-bounce. The differences between our model and Nakamura’s prediction are illustrated in the middle and bottom plots for $\nu = \nu_e + \bar{\nu}_e$ and ν_x , respectively.

emission. Our modifications introduce new parameters which we calibrate with a 3D model of the collapse of a $20M_\odot$ star. We then test the progenitor dependence of our extended analytic model and confirm that it reproduces the trends seen in population studies using approximately 100 progenitors.

While our comparisons to simulations have provided some validation for our analytic model, we must caution that some discrepancies remain, indicating potential next steps for improvements.

For modeling mass accretion, the main discrepancy occurs between the bounce time and shock revival (see Fig. D.1); the mass accretion rate by the analytic model tends to be overestimated within this region. This discrepancy likely stems from our choice of delay function $D(t)$ (Eq. D.28) converging towards $1/(t - t_b)$ rather than $1/t$. However, this discrepancy does not adversely affect the neutrino luminosity relative to simulations (as seen

in, e.g., Figs. D.4 and D.5). For modeling the neutrino emission, the discrepancy is largest prior to bounce, $t < t_b$. We see that despite the delay function $D(t)$ and our normalization procedure conserving mass and suppressing $\dot{M}$, it is not enough to prevent overestimating the neutrino luminosity at early times. The main cause for the overestimation is the diffusion luminosity component prior to the shock revival remaining unchanged. This discrepancy also impacts the late-time luminosity, being underestimated since the late-time cooling time scale τ_{LC} depends on conserving the time-integrated neutrino luminosity. Nevertheless, the comparison to simulation results illustrated in Figs. D.4 and D.5 shows that the modified model still agrees reasonably well with simulations.

One way this discrepancy may be addressed is by appending extra energy to the neutrino emission, e.g., through an additional parameter $A > 1$ that scales the curves in Fig. D.2 higher, which would allow the cooling time scale τ_{LC} to be slightly larger. One physical motivation for needing such a renormalization A parameter is the nuclear equation of state (EOS). The EOS adopted in the simulations we used to construct the cooling-phase neutrino emission [328] is different from the EOS used in our calibration of the accretion and explosion phase. In particular, other simulations of the cooling phase with EOSs more similar to our calibration simulations show more neutrino emission compared to [328]. While the precise value of the normalization factor A is not yet robustly determined, it would help reduce the discrepancy we see.

While the bounce time is observed to follow a very similar trend with respect to the progenitor compactness as seen in the suite of 100 axisymmetric simulations of [341], there are still small deviations, e.g., near $\xi_{2.5} \in (\sim 0.11, 0.21)$. Though determining the bounce time using r_{PNS} is physically motivated and provides reasonable results, simulations [323, 324] show that other variables could have similar predictive power (e.g., density, shock radius, temperature).

In the future, it would be desirable to extend the analytic model to the late-time PNS cooling

phase. For example, the neutrino light curve is characterized by a parameter p which sets whether the long-term decay is super- or sub-exponential; at present we keep p fixed but physically this parameter should depend on the mass of the PNS and thus the progenitor. In parallel, analytic solutions to the cooling phase have recently been investigated by various authors [297, 343], offering potential improvements. Finally, our analytic model contains some dozen model parameters, both from the original M16 and from our modifications. The analytic late-time cooling solutions also have model parameters [297, 343]. In the future, it would be possible to fit the analytic model to simulation sets and obtain best-fit values and uncertainties for model parameters.