

TRANSVERSE POSITION RECONSTRUCTION  
IN A LIQUID ARGON TIME PROJECTION  
CHAMBER USING PRINCIPAL COMPONENT  
ANALYSIS AND MULTI-DIMENSIONAL  
FITTING

---

A DISSERTATION  
SUBMITTED TO  
THE TEMPLE UNIVERSITY GRADUATE BOARD

---

IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE  
DOCTOR OF PHILOSOPHY

---

BY  
ANDREW WILLIAM WATSON

MAY 2017

EXAMINING COMMITTEE MEMBERS:  
C. J. MARTOFF, ADVISORY CHAIR, PHYSICS  
JIM NAPOLITANO, EXAMINING CHAIR, PHYSICS  
ANDREAS METZ, PHYSICS  
BERND SURROW, PHYSICS  
PETER MEYERS, PHYSICS, PRINCETON UNIVERSITY

©

Copyright

2017

by

Andrew William Watson

All Rights Reserved

# Abstract

One of the most enduring questions in modern physics is the dark matter problem. Measurements of galactic rotation curves taken in the middle of the twentieth century suggest that there are large spherical halos of unseen matter permeating and surrounding most galaxies, stretching far beyond their visible extents. Although some of this mass discrepancy can be attributed to sources like primordial black holes or Massive Astrophysical Compact Halo Objects (MACHOs), these theories can only explain a small percentage of this “missing matter”.

One approach which could account for the entirety of this missing mass is the theory of Weakly Interacting Massive Particles, or “WIMPs”. As their name suggests, WIMPs interact only through the weak nuclear force and gravity and are quite massive ( $100 \text{ GeV}/c^2$  to  $1 \text{ TeV}/c^2$ ). These particles have very small cross sections ( $\approx 10^{-39} \text{ cm}^2$ ) with nucleons and therefore interact only very rarely with “normal” baryonic matter.

To directly detect a dark matter particle, one needs to overcome this small cross-section barrier. In many experiments, this is achieved by utilizing detectors filled with liquid noble elements, which have excellent particle identification capabilities and are very low-background, allowing potential WIMP signals to be more easily distinguished from detector noise. These experiments also often apply uniform electric fields across their liquid volumes, turning the apparatus into Time Projection Chambers or “TPCs”. TPCs can accurately determine the location

of an interaction in the liquid volume (often simply called an “event”) along the direction of the electric field.

In DarkSide-50 (“DS-50” for short), the electric field is aligned antiparallel to the  $z$ -axis of the detector, and so the depth of an event can be determined to a considerable degree of accuracy by measuring the time between the first and second scintillation signals (“S1” and “S2”), which are generated at the interaction point itself and in a small gas pocket above the liquid region, respectively. One of the lingering challenges in this experiment, however, is the determination of an event’s position along the other two spatial dimensions, that is, its transverse or “ $xy$ ” position.

Some liquid noble element TPCs have achieved remarkably accurate event position reconstructions, typically using the relative amounts of S2 light collected by Photo-Multiplier Tubes (“PMTs”) as the input data to their reconstruction algorithms. This approach has been particularly challenging in DarkSide-50, partly due to unexpected asymmetries in the detector, and partly due to the design of the detector itself.

A variety of  $xy$ -Reconstruction methods (“ $xy$  methods” for short) have come and gone in DS-50, with only a few of them providing useful results. The  $xy$  method described in this dissertation is a two-step Principal Component Analysis / Multi-Dimensional Fit (PCAMDF) reconstruction. In a nutshell, this method develops a functional mapping from the 19-dimensional space of the signal received by the PMTs at the “top” (or the “anode” end) of the DarkSide-50 TPC to each of the transverse coordinates,  $x$  and  $y$ . PCAMDF is a low-level “machine learning” algorithm, and as such, needs to be “trained” with a sample of representative events; in this case, these are provided by the DarkSide geant4-based Monte Carlo, g4ds.

In this work, a thorough description of the PCAMDF  $xy$ -Reconstruction method

is provided along with an analysis of its performance on MC events and data. The method is applied to several classes of data events, including coincident decays, external gamma rays from calibration sources, and both atmospheric argon “AAr” and underground argon “UAr”. Discrepancies between the MC and data are explored, and fiducial volume cuts are calculated. Finally, a novel method is proposed for finding the accuracy of the PCAMDF reconstruction on data by using the asymmetry of the S2 light collected on the anode and cathode PMT arrays as a function of  $xy$ .

To all of the giants  
on whose shoulders  
I've stood.

# Acknowledgements

If it takes a village to raise a child, then it takes several villages (and lots of Skype credit, and more than a few trips to the bar) to bring a physics dissertation to maturity. I would be fooling myself if I thought I could give everyone the thanks they deserve, but I'll give it a go anyway. First and foremost, a tremendously sincere thank you to the Physics Department at Temple University, especially Ted Burkhardt and Jeff Martoff, for giving me the opportunity to succeed. Who knows where I would be if you hadn't taken a chance on me. A special thank you to Bernd Surov for your words of encouragement following a less-than-ideal qualifying exam result and to Jim Napolitano for helping me round out my final semester "on time and under budget". Thank you to Temple alumnus Dr. Price Kagey of KG Science Associates for your generous financial assistance. Thank you to John Tatarowicz and Christy Love for helping me to hit the ground running at Temple, and to Francesco Granato for giving me someone to talk to in the office again, more than two years after Christy graduated.

A big thank you to the Princeton physics department, especially Masayuki Wada and Jason Brodsky (now at LLNL), for helping me decipher the Voynich manuscript that is the CERN ROOT User's Guide, among other things. Thank you to Chengliang Zhu for having the patience to answer my constant barrage of Skype messages. Thank you to Shawn Westerdale (now at Snolab) and Stefano Davini (at GSSI) for your guidance and expertise during the neutron thermaliza-

tion signal search. And thank you to Peter Meyers for sponsoring my return trip to LNGS in December 2014 to help out with the extended neutron calibration campaign.

Speaking of LNGS, I cannot thank everyone in Assergi enough for their boundless hospitality. Thank you to Yura Suvorov for introducing us to the amazing Ristoro Mucciante. Thank you to Francesco Di Eusanio for the wonderful homemade wine and sausages. Thank you to Marco Carlini for welcoming us into your apartment for a fantastic dinner. Thank you to the Hawaii crew – Bernd, Erin, and Brianne – for your company and conversation in Assergi and Napoli alike. Thanks to the staff at Pollo Loco for the massive cheeseburgers and great music, and the owners of Il Parco for what is still the best cappuccino I’ve ever had. Finally, thanks to the physics department at Virginia Tech for letting me stay in your massive apartment during my time in Assergi (and the members of XENON, staying in the apartment below, for letting me steal your WiFi).

Thank you to everyone at UCLA, including Alden Fan (now at Berkeley Lab), Hanguo Wang, and Yi Wang, for making the month I spent there, prepping to move the ARIS detector, an enjoyable one. Thanks especially to Alden for helping me through some of the tougher aspects of coding the pulse-finder for ARIS. Thanks to APC and Davide Franco for financing my trip to Orsay to work on the experiment in person. Thanks also to the rest of the ARIS crew, including Quentin, Anyssa, Claudio, Sandro, Emilija, Ben, and Tessa for your company, your constructive criticism, and for introducing me to the quiche at La Pause Gourmande in Orsay, which was life-changing.

Thank you to Luca Grandi and Richard Saldanha at the University of Chicago for coordinating the analysis workshop I attended when I was still a bright-eyed, fresh-faced Master’s student who barely knew C++. Thanks to Stephen Pordes and Simone Rizzardini at FNAL for providing housing and transportation, re-

spectively, during my visits to the laboratory. Thanks to everyone at the general meeting in Milan for your company and everyone at the University of Notre Dame during SCENE beam time, for tolerating a first-year graduate student who didn't really have much business being there.

Thanks to everyone else I'm sure I missed for your expertise, your help, and your friendship. I truly could not have done any of this without you.

Finally, thank you to my close friends and family for your support over the past few years. I moved to Philly without much of a plan, other than working in a falafel shop until something better came along. Without a ton of help from my brother, Michael, and his wife, Jaime, I might still be there. Thanks for all of that. Thanks to my parents for my existence, but thanks also for your constant encouragement and support. Thanks to every teacher and professor who believed in me, even when I didn't believe in myself. Lastly, to the love of my life, Fliss, thank you for everything that you are and everything that you do.

This research was supported in part by the National Science Foundation through major research instrumentation grant number CNS-09-58854 as well as NSF grants PHY-1004054, PHY-1242611, and PHY-1314483. The ARIS Experiment is supported by Labex UnivEarthS grants ANR-10-LABX-0023 and ANR-11-IDEX-0005-02 under the JE2: Direct Dark Matter Search team, and by the France-Berkeley Fund.

# Table of Contents

	Page
ABSTRACT . . . . .	III
DEDICATION . . . . .	VI
ACKNOWLEDGEMENTS . . . . .	VII
LIST OF FIGURES . . . . .	XIII
LIST OF ACRONYMS AND SYMBOLS . . . . .	XVII
CHAPTER	
1 A BRIEF HISTORY OF DARK MATTER . . . . .	1
1.1 Matière Obscure . . . . .	2
1.2 Fritz Zwicky . . . . .	4
1.2.1 The Virial Theorem . . . . .	4
1.2.2 Mass-to-Light Ratios . . . . .	5
1.2.3 Gravitational Lensing (In Theory) . . . . .	6
1.3 Galactic Rotation Curves . . . . .	8
1.4 Gravitational Lensing (In Practice) . . . . .	12
1.5 Big Bang Nucleosynthesis . . . . .	15
1.6 The CMB and the $\Lambda$ CDM Model . . . . .	16
2 PAST AND PRESENT DARK MATTER CANDIDATES . . . . .	20
2.1 Baryonic Matter . . . . .	21
2.2 Black Holes . . . . .	22
2.3 Neutrinos . . . . .	23
2.4 Sterile Neutrinos . . . . .	25
2.5 Axions . . . . .	26
2.6 Supersymmetry . . . . .	28
2.6.1 The WIMP Miracle . . . . .	30
2.7 Other Candidates . . . . .	34
2.7.1 Gravitinos . . . . .	34
2.7.2 Kaluza-Klein Particles . . . . .	36

2.7.3	Axinos . . . . .	38
2.8	Detection of Dark Matter . . . . .	40
2.8.1	Production . . . . .	40
2.8.2	Indirect Detection . . . . .	41
2.8.3	Direct Detection . . . . .	45
2.8.4	State of the Field . . . . .	50
3	DARKSIDE . . . . .	54
3.1	Active Shielding . . . . .	55
3.2	Time-Projection Chamber . . . . .	59
3.2.1	Why $\ell$ Ar? . . . . .	63
3.2.2	AAr vs. UAr . . . . .	66
3.2.3	S1, S2, and S3 . . . . .	68
3.3	g4ds: The DarkSide Monte Carlo . . . . .	68
3.4	Data Acquisition and Event Reconstruction . . . . .	70
3.4.1	Event Reconstruction . . . . .	71
3.4.2	Calibrations and Corrections . . . . .	76
3.5	Pulse Shape Discrimination . . . . .	84
3.5.1	Particle Identification in $\ell$ Xe . . . . .	88
3.6	Backgrounds . . . . .	89
3.6.1	Intrinsic Backgrounds and $\ell$ Ar Purity . . . . .	89
3.6.2	Surface Backgrounds . . . . .	90
3.6.3	Fiducialization . . . . .	91
3.7	Cuts . . . . .	92
3.7.1	Data Quality Cuts . . . . .	92
3.7.2	Veto Cuts . . . . .	93
3.7.3	Physics Cuts . . . . .	96
3.7.4	Result of Cuts . . . . .	97
3.8	Exclusion Limit Plot . . . . .	98
3.8.1	SCENE . . . . .	100
3.8.2	$f_{90}$ . . . . .	101
3.8.3	Results . . . . .	102
4	PRINCIPAL COMPONENT ANALYSIS AND MULTI-DIMENSIONAL FITTING . . . . .	103
4.1	Principal Component Analysis . . . . .	105
4.1.1	The Covariance Matrix . . . . .	105
4.1.2	As Applied to DarkSide-50 . . . . .	111
4.2	Reduction of Dimensionality . . . . .	120
4.3	Multi-Dimensional Fitting . . . . .	121
5	PCAMDF $xy$ -RECONSTRUCTION . . . . .	127
5.1	Description of Event Sets . . . . .	127
5.1.1	Data . . . . .	128
5.1.2	Monte Carlo . . . . .	129

5.2	Reconstruction Code . . . . .	129
5.3	Performance on MC events . . . . .	135
5.4	Calculating a Fiducial Volume Cut . . . . .	148
5.5	Performance on Data . . . . .	151
5.5.1	PMT-centered S1 Distortions . . . . .	152
5.5.2	TPC-centered S1 Distortions . . . . .	152
5.5.3	Spatial Distributions of Reconstructed Events . . . . .	156
5.5.4	S2 <sub>TBA</sub> . . . . .	162
5.5.5	UAr Event Hole . . . . .	165
5.5.6	An S1-Based PCAMDF <i>xy</i> -Reconstruction . . . . .	170
5.5.7	S2 Corrections . . . . .	175
5.6	Quantifying <i>xy</i> Performance . . . . .	184
5.6.1	Resolution Without <i>xy</i> -Reconstruction . . . . .	186
5.6.2	Coincident Decays: BiPos and RnPos . . . . .	188
5.6.3	<sup>57</sup> Co Events . . . . .	196
5.6.4	S2 <sub>TBA</sub> Revisited . . . . .	202
6	CONCLUSION . . . . .	213
	BIBLIOGRAPHY . . . . .	216
	APPENDIX A: DERIVATION OF THE RELIC DENSITY EQUATION . . . . .	238
	APPENDIX B: DS-50 AUTHOR LIST . . . . .	247
	APPENDIX C: ARIS . . . . .	249
	APPENDIX D: NEUTRON THERMALIZATION IN DARKSIDE-50 . . . . .	258
	APPENDIX E: SAMPLE RECONSTRUCTION CODE . . . . .	265

# List of Figures

1.1	An excerpt from Einstein’s notebook musings on gravitational lensing	7
1.2	A few galactic rotation curves . . . . .	10
1.3	An image taken of the strong gravitational lens LRG 3-757 . . . . .	13
1.4	A composite visible / X-ray image of the Bullet Cluster . . . . .	14
1.5	Map of CMB temperature anisotropies . . . . .	18
2.1	Constraints on sterile neutrino dark matter . . . . .	25
2.2	Behavior of the relic density equation as a function of $T$ . . . . .	35
2.3	Dark matter energy density accounted for by particles of mass $m_\chi$ .	36
2.4	Current experimental limits on LKP photon and Z-boson masses . .	37
2.5	Interaction cross section vs. mass for selected DM candidates . . . .	39
2.6	Three possible modes of direct dark matter detection . . . . .	40
2.7	Artist’s conception of the KM3NeT detector . . . . .	43
2.8	Recoil energy of an argon nucleus . . . . .	46
2.9	Exclusion limits for most prominent dark matter searches . . . . .	51
2.10	DAMA/LIBRA single-hit scintillation event residuals . . . . .	52
2.11	Timeline of prominent direct dark matter detection claims . . . . .	53
3.1	Rendering of the DS-50 cryostat . . . . .	56
3.2	The effect of boron-loading the LSV scintillator . . . . .	58
3.3	A 3D model of the DarkSide-50 TPC . . . . .	61
3.4	DS-50 TPC PMT maps with channel indices . . . . .	64

3.5	S1 spectrum comparison between AAr and UAr data . . . . .	67
3.6	S1 <sub>late</sub> spectra from AAr and UAr DS-50 data . . . . .	69
3.7	S1 spectrum of $\alpha$ particles . . . . .	71
3.8	The DarkArt baseline-finder working in regions of baseline and signal	74
3.9	Example SER fit for a SCENE PMT . . . . .	77
3.10	SPE vs. time (run number) during an $^{241}\text{Am}^9\text{Be}$ campaign . . . . .	78
3.11	Un- $z$ -corrected S1 vs. $t_{\text{drift}}$ for a $^{83\text{m}}\text{Kr}$ calibration run . . . . .	79
3.12	S1 spectrum of a krypton calibration run . . . . .	81
3.13	Calibration data taken with an $^{241}\text{Am}^9\text{Be}$ source . . . . .	83
3.14	Scintillation in the LSV, showing the effect of PPO concentration .	84
3.15	The waveforms of typical ER and NR events in DS-50 . . . . .	86
3.16	Magnified waveforms from Figure 3.15 . . . . .	87
3.17	Full event set from the DS-50 70-day campaign with UAr . . . . .	98
3.18	Recent exclusion limit curves from a selection of DM experiments .	99
4.1	The covariance matrix of the first $10^6$ 50-Days AAr events . . . . .	113
4.2	The covariance matrix for $10^6$ uniformly-generated g4ds events . . .	114
4.3	The 50-Days AAr covariance matrix, with all CX analysis cuts applied	115
4.4	S2[] variance over the anode PMTs for the first $10^6$ AAr events . .	117
4.5	S2[] variance over the anode PMTs for $10^6$ g4ds events . . . . .	118
4.6	AAr S2[] variance over the anode PMTs with all CX analysis cuts .	119
4.7	Eigenspectrum plots for AAr data and g4ds. . . . .	122
5.1	Reconstruction accuracy for high and low statistics g4ds events. . .	137
5.2	Accuracy mapped onto $xy$ for high and low statistics g4ds events. .	138
5.3	Accuracy as a function of radius . . . . .	139
5.4	Accuracy as a function of S2 <sub>tot</sub> for low-S2 <sub>tot</sub> events . . . . .	140
5.5	Average accuracy as a function of S2 <sub>tot</sub> for low-S2 <sub>tot</sub> events . . . . .	141

5.6	Accuracy as a function of radius for low- $S2_{\text{tot}}$ events . . . . .	143
5.7	Average accuracy as a function of radius for low- $S2_{\text{tot}}$ events . . . . .	144
5.8	Radial bias in the PCAMDF reconstruction for MC events . . . . .	146
5.9	PMT-centered radial bias for reconstructed MC events . . . . .	147
5.10	Angular bias in the PCAMDF $xy$ -Reconstruction. . . . .	149
5.11	Fiducial cut radius as a function of the fraction of wall events cut. . . . .	151
5.12	Radial Energy Profiles for PMTs # 25, 30, and 35 using AAr data. . . . .	153
5.13	TPC-centered REPs for AAr and UAr data . . . . .	155
5.14	Radial distributions of reconstructed MC and data events. . . . .	157
5.15	AAr-subtracted UAr radial event distribution . . . . .	158
5.16	PCAMDF $xy$ distributions of AAr and UAr events . . . . .	160
5.17	Bottom (cathode) PMT array, prior to installation in DS-50 . . . . .	164
5.18	$S2_{\text{TBA}}$ maps for AAr and UAr data . . . . .	166
5.19	PMT # 2 “event hole” in UAr data . . . . .	167
5.20	UAr “event hole” $t_{\text{drift}}$ dependence . . . . .	168
5.21	S1 and S2-based $xy$ -Reconstructions for MC and data events . . . . .	172
5.22	“Near-cathode” events reconstructed with an S1-based $xy$ . . . . .	174
5.23	$S2_{\text{TBA}}$ maps with an S1- and an S2-based $xy$ . . . . .	176
5.24	$S2/S1$ as a function of radius for AAr and UAr data . . . . .	178
5.25	A crude radial S2 correction . . . . .	179
5.26	$S2_{\text{corr}}/S1$ mapped over $xy$ . . . . .	180
5.27	g4ds UV photon detection probability map . . . . .	181
5.28	$S2/S1$ residual maps after corrections . . . . .	182
5.29	Number of PE required for statistical distinguishability of events . . . . .	189
5.30	$^{226}\text{Ra}$ and $^{232}\text{Th}$ decay chains . . . . .	190
5.31	A typical “BiPo” event waveform . . . . .	191
5.32	$^{214}\text{Bi}$ - $^{214}\text{Po}$ decays in DarkSide-50 . . . . .	192

5.33	Reconstructed $xy$ positions of 38 BiPo events . . . . .	193
5.34	“RnPo” events in DS-50 . . . . .	194
5.35	Rn-Po distance as a function of radius . . . . .	196
5.36	Cartoon showing a few possible paths for a particle to enter the TPC	199
5.37	$^{57}\text{Co}$ events from a series of consecutive CALIS runs . . . . .	200
5.38	Locating the position of a $^{57}\text{Co}$ source used during a CALIS campaign	202
5.39	Fixing the discrepancy between the MC and data $S2_{\text{TBA}}$ spectra . . .	203
5.40	Events in the low- $S2_{\text{TBA}}$ peak seen in Figure 5.39 . . . . .	205
5.41	Events near the mean of the $S2_{\text{TBA}}$ spectrum seen in Figure 5.39 . .	206
5.42	True-position MC events in the $\theta = 0^\circ$ region . . . . .	207
5.43	Reconstructed-position MC events in the $\theta = 0^\circ$ region . . . . .	208
5.44	Reconstructed-position AAr events in the $\theta = 0^\circ$ region . . . . .	208
C.1	Before and after $z$ -correcting S1 in ARIS Run 101076 . . . . .	252
C.2	Before and after $z$ -correcting S2 (aka. $S2_{\text{tot}}$ ) in ARIS Run 101076 .	253
C.3	Finding the TPC tilt in ARIS. . . . .	256
C.4	$xyz$ -corrected S2/S1 map for ARIS Run 101076 . . . . .	257
D.1	LSV clusters in low-rate $^{241}\text{Am}^9\text{Be}$ data. . . . .	261
D.2	Removing events with 4.4 MeV $\gamma$ clusters in the LSV . . . . .	262
D.3	Energy spectrum for the late LSV clusters from Figure D.2 . . . . .	262
D.4	Time distribution of the clusters in the red box in Figure D.3 . . . .	263
D.5	Time and energy distributions of the neutron thermalization events	264

# List of Acronyms and Symbols

<b>AAr</b> Atmospheric Argon	<b>CX</b> DarkSide analysis cuts
<b>ADC</b> Analog-to-Digital Converter	<b>DAQ</b> Data AcQuisition system
<b>ARIS*</b> Argon Response Ionization and Scintillation	$d\vec{r}$ Distance from the true $xy$ -position of a MC event to its PCAMDF-reconstructed $xy$ -position
<b>BAO</b> Baryon Acoustic Oscillations	
<b>CALIS</b> CALibration Insertion System	<b>DS*</b> DarkSide (either the Collaboration or the experiment)
<b>CDM</b> Cold Dark Matter	
<b>CDMS*</b> The Cryogenic Dark Matter Search	<b>DS-10*</b> DarkSide-10, the 10 kg active volume prototype DarkSide detector
<b>CERN</b> <sup>†</sup> The European Organization for Nuclear Research (acronym derives from <i>Conseil Européen pour la Recherche Nucléaire</i> )	<b>DS-50*</b> DarkSide-50, the current-generation, 50 kg active volume DarkSide detector
<b>CMB</b> Cosmic Microwave Background	<b>DS-20k*</b> DarkSide-20k, the next-generation, 20 Mg active volume DarkSide detector
<b>CRH</b> Clean Room Hanoi	
<b>CTF</b> Counting Test Facility	<b>DM</b> Dark Matter

$\tau_{\text{ed}}$	Electron drift lifetime	<b>LHC*</b>	Large Hadron Collider
<b>EM</b>	Electromagnetic / Electromagnetism	<b>LNGS<sup>†</sup></b>	The National Laboratories at Gran Sasso (acronym derives from <i>Laboratori Nazionali del Gran Sasso</i> )
<b>ER</b>	Electron Recoil	<b>LKP</b>	Lightest Kaluza-Klein Partner
$f_{90}$	A PSD parameter which gives the fraction of PE detected in the first 90 ns of an S1 pulse	<b>LSP</b>	Lightest Supersymmetric Particle
<b>gAr</b>	gaseous argon	<b>LSV</b>	Liquid Scintillator Veto
<b>g4ds</b>	The Geant4-based DarkSide MC	$\ell$ <b>Xe</b>	liquid xenon
<b>HDM</b>	Hot Dark Matter	<b>LY</b>	Light Yield
<b>HHV</b>	High High Voltage	<b>MACHO*</b>	Massive Astrophysical Compact Halo Object
<b>IPNO<sup>†</sup></b>	The Institute of Nuclear Physics at Orsay (acronym derives from <i>L'Institut de Physique Nucléaire d'Orsay</i> )	<b>MC</b>	Monte Carlo
<b>ITO</b>	Indium Tin Oxide	<b>MDF</b>	Multi-Dimensional Fit
$\text{keV}_{\text{ee}}$	keV electron equivalent	<b>MSSM</b>	Minimal Supersymmetric Standard Model
$\text{keV}_{\text{nr}}$	keV nuclear recoil	<b>MWE</b>	Minimal Working Example
<b>KK</b>	Kaluza-Klein	<b>NR</b>	Nuclear Recoil
$\ell$ <b>Ar</b>	liquid argon	<b>PC</b>	Principal Component(s)
<b><math>\Lambda</math>CDM</b>	The standard dark energy ( $\Lambda$ ) / Cold Dark Matter of cosmology	<b>PCA</b>	Principal Component Analysis
		<b>PCAMDF</b>	Principal Component Analysis, Multi-Dimensional Fit

<b>PE</b> PhotoElectron(s)	$S1_{\text{bot}}^{\ddagger}$ The total S1 seen by the bottom (cathode) TPC PMTs
<b>PMT</b> Photo-Multiplier Tube	
<b>PPO</b> 2,5-diphenyloxazole (acronym derives from the chemical structure: two phenyl (P) rings attached to an oxazole (O) molecule)	$S1_{\text{corr}}^{\ddagger}$ $t_{\text{drift}}$ -corrected $S1_{\text{tot}}$ $S1_{\text{Mf}}^{\ddagger}$ The maximum value of $S1$ []
<b>PSD</b> Pulse Shape Discrimination	$S1_{\text{top}}^{\ddagger}$ The total S1 seen by the top (anode) TPC PMTs
<b>PTFE</b> PolyTetraFluoroEthylene (aka. Teflon)	$S1_{\text{tot}}^{\ddagger}$ The total S1 seen by all TPC PMTs, aka. S1
<b>QE</b> Quantum Efficiency	<b>SPE</b> Single PhotoElectron
<b>SCENE*</b> SCintillation Efficiency of Noble Elements	<b>S2</b> The second Scintillation signal seen during an event in an ionization / scintillation-detecting TPC
<b>SER</b> Single Electron Response	<b>S3</b> A third Scintillation signal, sometimes seen during an event in an ionization / scintillation-detecting TPC; aka. $S2_{\text{echo}}$
<b>SiPM</b> Silicon Photo-Multiplier	
<b>SLAD</b> SLim Analysis Data file	
<b>SM</b> Standard Model	<b>SUGRA</b> SUperGRAvity
<b>S1</b> The first Scintillation signal seen during an event in an ionization / scintillation-detecting TPC; value is given in PE	<b>SUSY</b> SUperSYmmetry / SUperSYmmetric $t_{\text{drift}}$ The time between the start of the S1 signal and the start of the S2 signal
$S1$ [] <sup>‡</sup> The S1 fraction seen by each TPC PMT	<b>TMB</b> TriMethyl Borate

<b>TPB</b> TetraPhenyl Butadiene	Particle
<b>TPC</b> Time-Projection Chamber	<b>WCD</b> Water Čerenkov Detector
<b>UAr</b> Underground Argon	<b>WDM</b> Warm Dark Matter
<b>“visible matter”</b> Matter which interacts via the EM force / couples to the EM field	<b>WLS</b> WaveLength Shifter
	$\vec{A}$ Notation for a vector
<b>WIMP</b> Weakly-Interacting Massive	$\hat{A}$ Notation for a matrix

\* = particle detector / particle physics apparatus

† = laboratory / research organization

‡ = similar variables are defined for S2

# Chapter 1

## A BRIEF HISTORY OF DARK MATTER

Physics, in the broadest sense, is the study of the properties and behavior of the universe. More often than not, these properties and behaviors are represented as purely mathematical expressions. Occasionally, in the course of evaluating one of these expressions, the terms on the left-hand side do not match those on the right-hand side, so to speak.

As was the case with the luminiferous æther, sometimes equations do not balance properly because an incorrect assumption has been made (for instance, assuming that light waves, like sound waves, require a physical medium through which to travel).<sup>1</sup> Oftentimes, some previously unknown physical aspect of the system simply has not yet been discovered, like when Albert Einstein's General Theory of Relativity explained the 43 arcsecond-per-century deviation from classical predictions of the precession of the perihelion of Mercury's orbit<sup>2</sup> though it had remained unexplained since its discovery 57 years earlier by French mathematician

---

<sup>1</sup> [Michelson \(1887\)](#)

<sup>2</sup> [Einstein \(1916\)](#)

Urbain Le Verrier<sup>3</sup>.

The history of dark matter is littered with stories like these; the justification of this theoretical form of matter is largely based on the balancing of mismatched equations. From the earliest application of the virial theorem to the motions in the Coma Cluster, to evidence based on classical Newtonian mechanics like galactic rotation curves, all the way to present-day studies of the temperature anisotropies in the cosmic microwave background, the evidence for dark matter is irrefutable, and continuously mounting. Although a much more complete history is available in [Bertone and Hooper \(2016\)](#), a summary of the major points is given in the following sections.

## 1.1 Matière Obscure

Most histories of dark matter begin in the 1930s. But long before then, astronomers and physicists were observing the night sky and wondering if they were really seeing everything that’s “out there”.

Urbain Le Verrier — even though he could not see it — calculated the position of a new planet in 1846,<sup>4</sup> because the observed motions of another planet (in this case, Uranus) were in disagreement with those predicted by Newtonian mechanics.<sup>5</sup> The first time he did this, he struck gold and discovered Neptune. The second time, he proposed the planet Vulcan,<sup>6</sup> inside the orbit of Mercury. This planet was meant to explain the infamous 43 arcsecond-per-century deviation of the precession of Mercury’s orbit which could not be explained by Newtonian mechanics. Unfortunately for Le Verrier, this planet was never discovered, and this deviation would have to wait to be solved a few generations later by a patent clerk

---

<sup>3</sup> [Le Verrier \(1859\)](#)

<sup>4</sup> [Locher \(2011\)](#)

<sup>5</sup> [Bouvard \(1821\)](#)

<sup>6</sup> Though he did not name this planet.

from Ulm.

During the interim, the phrase “dark matter” would have meant something quite different to an astronomer or a physicist than it does now. In the late 19th century, those searching for “dark matter” were looking primarily at the occasional patches of darkness which interrupted bright, dense fields of stars in the night sky. There was some debate over whether these were real dips in stellar density or whether they were just a phenomenon of perspective — perhaps there were dark clouds of gas between the Earth and those distant stars, blocking their light.<sup>7</sup>

Lord Kelvin was a proponent of the latter theory and applied his “theory of gases” to the Milky Way in an attempt to discover what fraction of the system’s mass was composed of — as he proposed — “extinct[,] dark stars”, and these clouds of gas, as opposed to regular, bright stellar matter. He estimated the volume of the galaxy and the mass it contained, and calculated an expected velocity dispersion which was in agreement with what was observed at the time.<sup>8</sup>

In 1906, Henri Poincaré came across Kelvin’s analysis and remarked that since the velocity dispersion calculated from the luminous bodies was in agreement with the observed dispersion, then the amount of “dark matter” must be equal to, or less than, the amount of visible matter. That same year, Poincaré published an article to this effect in *L’Astronomie*,<sup>9</sup> the bulletin of the Société Astronomique de France, which contains one of the first recorded usages of the phrase “dark matter”<sup>10</sup> to describe astronomical bodies which we can detect due to their interactions with other bodies, but which we cannot directly observe. For decades, this nomenclature was commonplace within the astronomy community: “dark matter” was some substance which we had reason to believe existed, but which we could

---

<sup>7</sup> Bertone and Hooper (2016)

<sup>8</sup> Kelvin (1904)

<sup>9</sup> Poincaré (1906)

<sup>10</sup> “*Matière obscure*” in the original French, though now “*matière noire*” is more common.

only detect indirectly, through its interaction with “visible” matter.

## 1.2 Fritz Zwicky

Twenty-seven years after Poincaré’s publication, Swiss astronomer Fritz Zwicky was studying the redshifts of galaxies within various galaxy clusters. His 1933 paper, “*Die Rotverschiebung von extragalaktischen Nebeln*”<sup>11</sup> is often cited as the source of the phrase “dark matter”,<sup>12</sup> though we now know that Zwicky was perfectly in step with his contemporaries in terms of nomenclature. He even shared his peers’ suspicions about the composition of this matter, a few years later wondering “to what extent the apparent luminosity of a given [galaxy] is diminished by the internal absorption of radiation because of the presence of dark matter”,<sup>13</sup> equating “dark matter” with dense clouds of gas and dust.

### 1.2.1 The Virial Theorem

Zwicky was, however, the first to use the virial theorem to calculate the mass of a galaxy cluster. Assuming the cluster is a gravitationally-bound system, we can use some simple kinematics and derive a relationship between the total gravitational potential energy and the kinetic energy of its member galaxies.

In its simplest form, the virial theorem can be written as

$$\langle T \rangle_t = \frac{n}{2} \langle U \rangle_t = -\frac{1}{2} \langle U \rangle_t \quad (1.1)$$

and states that the time-averaged kinetic energy,  $T$ , of a sufficiently mature, stable system, bound by a potential  $V(r)$  (which in this case is gravitational and

---

<sup>11</sup> Zwicky (1933)

<sup>12</sup> “*Dunkle Materie*” in Zwicky’s native tongue.

<sup>13</sup> Zwicky (1937)

thus goes as  $r^n = r^{-1}$ ), should be simply related to the time-averaged potential energy,  $U$ , of that same system.<sup>14</sup> At several billion years of age, the Coma cluster can be assumed to have reached a stable enough state such that these time averages hold.<sup>15</sup>

### 1.2.2 Mass-to-Light Ratios

Astronomers in Zwicky’s time, and now, use something called the mass-to-light ratio ( $M/L$ ) to estimate the amount of dark matter in galaxies and clusters. Using units of solar luminosity and solar mass, one can calculate how many Suns’ worth of mass are necessary to generate the observed gravitational effects. Then one can calculate how many Suns’ worth of luminosity are needed to account for the light emitted by the object.

Since we expect that the mass of a galaxy or cluster of galaxies is dominated by stars, the  $M/L$  ratio should be roughly of order 1; a Sun’s worth of mass contributes a Sun’s worth of light. (Of course, interstellar gas contributes a significant fraction of the mass of a galaxy or cluster — a few times as much as the stellar matter — and so any  $M/L$  ratio between about  $1 \rightarrow 10$  should be considered reasonable.<sup>16</sup>) A more precise estimate of the  $M/L$  ratio can be made if the mass-luminosity relation is used and the relative abundances of different types of stars are known. This ratio can also vary among bands of wavelengths and stellar spectral types, but on average, one should see a ratio of about 1–10 for galaxies and galaxy clusters.

A ratio much greater than 1–10 says that the mass of the cluster is much greater than the luminosity of the cluster, relative to the Sun’s mass and luminosity. This means that much of the mass of the cluster must be non-luminous (or “dark”)

---

<sup>14</sup> Collins (1978)

<sup>15</sup> Noonan (1969)

<sup>16</sup> Aleksić et al. (2010)

matter.

Zwicky found a stark difference between the kinetic energy of the Coma cluster, which he estimated using velocity measurements taken by Hubble a few years prior, and the potential energy of the cluster, which he estimated from the cluster’s total luminous (stellar) mass. In order for his redshift-measured galactic velocities to agree with the virial theorem, Zwicky concluded that “the average density in the Coma system would have to be at least 400 times larger than that derived on the grounds of observations of luminous matter.”<sup>17</sup> That is, Zwicky estimated the  $M/L$  ratio of the Coma cluster to be in excess of 400. Modern measurements have revised Zwicky’s estimates a bit, but the fact remains that the vast majority of the mass in this cluster cannot be accounted for by luminous matter alone.

Zwicky and his contemporaries tried to explain away the Coma cluster (and the high  $M/L$  ratios of other clusters) by writing off the highest-velocity galaxies within them as field nebulae, or by arguing that the virial theorem cannot be applied because clusters are unstable, rapidly expanding systems. Neither of these explanations held water and the community was stuck — they would need more evidence before they could determine the nature and extent of this missing mass.<sup>18</sup> This additional evidence would come in the form of galactic rotation curves.

### 1.2.3 Gravitational Lensing (In Theory)

Also in 1937, Zwicky proposed using gravitational lensing as a method for measuring the amount of dark matter contained within a galaxy or cluster of galaxies. Albert Einstein is credited as popularizing (see Figure 1.1) the idea that astronomical objects, “bending” space by virtue of their large masses, could act as “lenses”, allowing one to observe any luminous material behind them which would otherwise

---

<sup>17</sup> Zwicky (1933)

<sup>18</sup> Bertone and Hooper (2016)

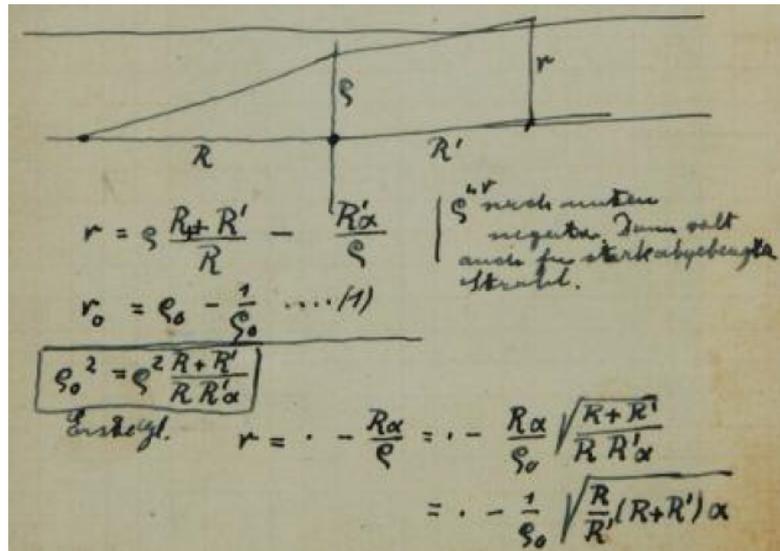


Figure 1.1: An excerpt from Einstein’s notebook musings on gravitational lensing from Sauer (2010). Published as part of the Collected Papers of Albert Einstein, Vol. 3, p. 585.

have been obscured.<sup>19</sup> The degree to which the light of the more distant object is bent is proportional to the amount of matter (whether visible or not) contained within the intervening body. The phenomenon of gravitational lensing directly follows from the theory of general relativity<sup>20</sup> and offers an alternative method for estimating the dark matter content of a galaxy or cluster, which doesn’t rely on measuring the velocities of their constituents.

Einstein abandoned his lensing theory on the grounds that the effect from even the largest of stars would not be significant enough to be measurable. In 1937, Zwicky picked up where Einstein left off and published an article which discussed the possibility of using far more massive galaxy clusters to observe gravitational lensing, and potentially calculate local dark matter densities. Unfortunately, his theory remained untested for many decades; Zwicky himself noted that “[u]ntil many plates of rich nebular fields taken under excellent conditions of seeing have been carefully examined it would be dangerous” to make assumptions about the

<sup>19</sup> Sauer (2010)

<sup>20</sup> Einstein (1936)

abundance of lensing clusters.<sup>21</sup>

Zwicky’s prerequisite was satisfied just over two decades later, when in 1958, George Abell published his now ubiquitous “Abell catalog of rich clusters of galaxies”<sup>22</sup> (though it originally only included northern hemisphere galaxies, the catalog was updated in 1989 to include a southern survey<sup>23</sup>). Despite this wealth of new information, it took an additional two decades until the first true gravitational lenses were observed.

### 1.3 Galactic Rotation Curves

After Zwicky’s gravitational lens proposal but before the first observation of said gravitational lenses, analysis of galactic rotation curves strengthened the case for the theory of dark matter. Additionally, since Zwicky was primarily concerned with clusters of galaxies, the study of galactic rotation curves allowed for the study of the effects of dark matter on much smaller scales.

Although the idea had been around since at least 1918 — when Francis Pease first measured the rotational velocity of Andromeda as a function of radius from its axis<sup>24</sup> — it wasn’t until the 1960’s and 70’s that the study of galactic rotation curves really took off. The idea is fairly simple: since randomly-oriented elliptical stellar orbits around the galactic center are, on average, circular, one can equate Newton’s Second Law of Motion with his Law of Universal Gravitation and substitute the centripetal acceleration ( $a = v^2/r$ ) for the “ $a$ ” in  $F = ma$ :

$$F = ma = m \frac{(v(r))^2}{r} = G \frac{M(r)m}{r^2} \quad (1.2)$$

---

<sup>21</sup> Zwicky (1937)

<sup>22</sup> Abell (1958)

<sup>23</sup> Abell et al. (1989)

<sup>24</sup> Pease (1918)

Here,  $G$  is the gravitational constant,  $m$  is the mass of a particular object orbiting the galactic center at a spherical radius  $r$ ,  $v(r)$  is the magnitude of the velocity of that object, and  $M(r)$  is the galactic mass contained within a sphere of radius  $r$ . (Recall that, for a spherically-symmetric mass distribution described by  $M(r)$ , only the mass within a sphere of radius  $r$  has a net effect on a mass orbiting at that same radius  $r$ . This is a result of the application of Gauss’s Law to the gravitational force.) Solving the two rightmost terms above for the velocity yields

$$v(r) = \sqrt{\frac{GM(r)}{r}} \quad \text{and} \quad \lim_{r \rightarrow r_L} v(r) = \frac{C}{\sqrt{r}} \quad (1.3)$$

In the second term,  $r$  approaches  $r_L$ , the maximum radius of the stellar (luminous) mass of the galaxy. If one assumes that most of the mass of the galaxy is luminous, then near or beyond the edge of the stellar disk, the mass,  $M(r)$ , should be more or less constant with increasing radius,  $r$ , and the velocity should drop as  $r^{-1/2}$ . In the visible part of the galaxy, one can measure the velocity from the redshift of the constituent stars; and beyond the visible disk, the 21 cm hydrogen line can be used.

In 1962, Vera Rubin, then at Georgetown College in Washington, D.C. (now known as Georgetown University), knew this expected relation, and “with [her] graduate students at Georgetown, made a study of the velocities of 1000 O and B stars” in the disk of the Milky Way in the vicinity of the Sun.<sup>25</sup> They recorded the radial velocities of these stars and created very detailed, very impressive rotation curves. These plots showed — without exception — that the velocities of the outer (high radius) stars did not decrease as would be expected for standard Keplerian (on average, circular) orbits. Rubin later noted that, unfortunately, this article seemed to have “apparently influenced no one and was ignored even by the senior

---

<sup>25</sup> [Rubin et al. \(1962\)](#)

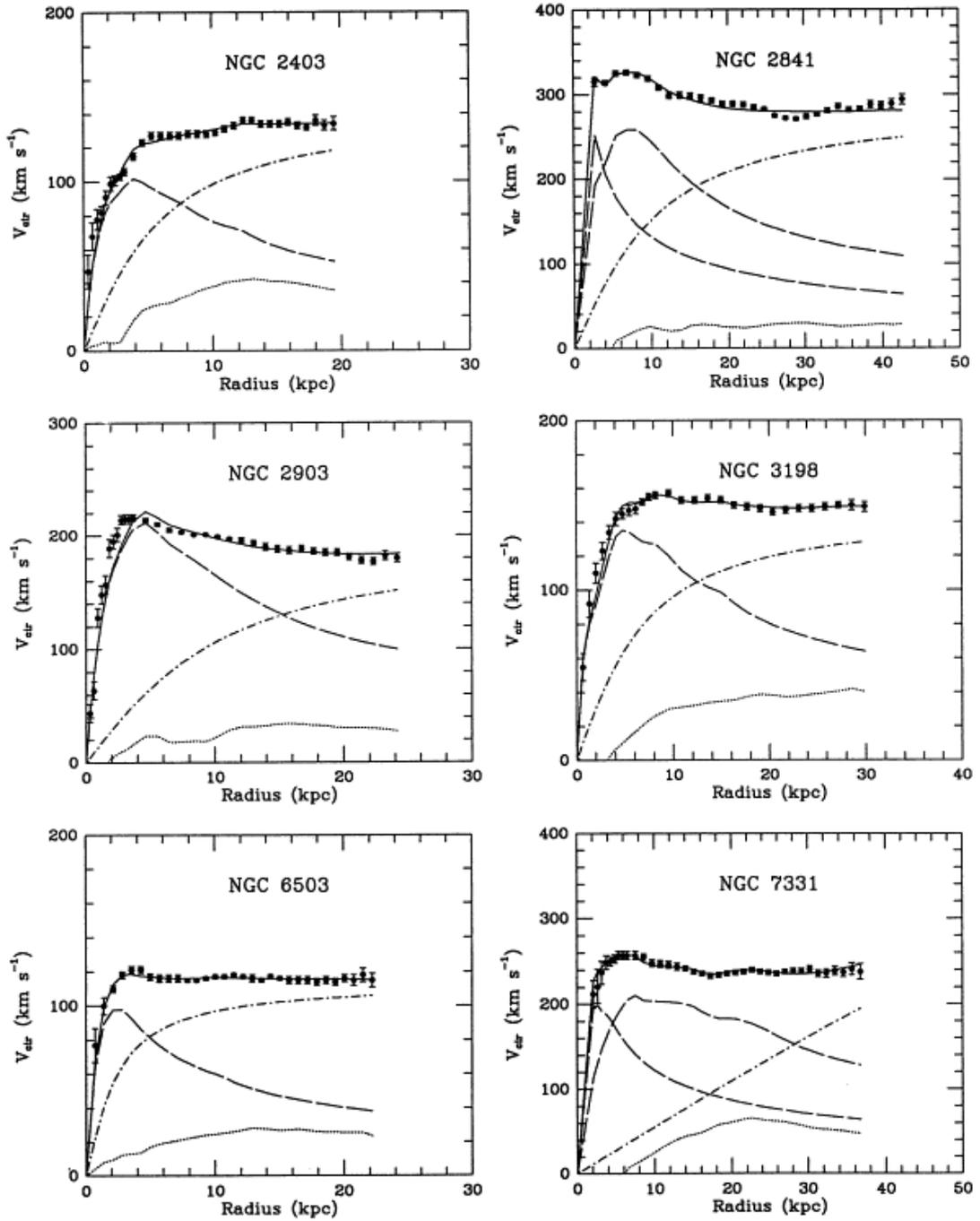


Figure 1.2: A few galactic rotation curves from [Begeman et al. \(1991\)](#). The data points are velocities of individual stars. The long-dashed curves are classical fits to the luminous matter in the galaxies. And the solid curves are fits that include a spherical halo of dark matter around the galaxies.

author [Rubin herself] when she returned to the problem of galaxy rotation a decade later”.<sup>26</sup>

The measurement of galactic rotation curves reached a fever pitch in the seventies. In 1970, Ken Freeman investigated the curves of M33 and NGC 300.<sup>27</sup> In 1972, Rogstad and Shostak measured NGC 2403, IC 342, M101, and NGC 6946.<sup>28</sup> In 1973, Roberts and Rots measured M31 and M81.<sup>29</sup> In 1978 Albert Bosma, for his Ph.D. thesis, published the rotation curves of 25 galaxies.<sup>30</sup> A few months later, Thonnard, Rubin, Ford, and Roberts published the 21 cm rotation curves for 183 galaxies.<sup>31</sup> In every one of these studies, for nearly every single spiral galaxy, flat or increasing rotation curves were observed out to the furthest measurable extent of the 21 cm line. It may have taken a few years for their impact to be felt in the scientific community, but it has been shown time and time again that the rotation curves for luminous galactic matter do not match classical predictions. (See, for instance, the curves in Figure 1.2.)

It was Ken Freeman, in 1970, who first made the claim that “if [the data] are correct, then there must be in these galaxies additional matter which is undetected, either optically or at 21 cm. Its mass must be at least as large as the mass of the detected galaxy, and its distribution must be quite different from the exponential distribution which holds for the optical galaxy.”<sup>32</sup> This was the first unequivocal proposal of a dark matter “halo” surrounding the observed galaxies.

Modern studies suggest that this halo is spherical, and can have a mass on the order of — or even larger than — the visible mass of the galaxy.<sup>33</sup> Some models

---

<sup>26</sup> Rubin (2000)

<sup>27</sup> Freeman (1970)

<sup>28</sup> Rogstad and Shostak (1972)

<sup>29</sup> Roberts and Rots (1973)

<sup>30</sup> Bosma (1978)

<sup>31</sup> Thonnard et al. (1978)

<sup>32</sup> Freeman (1970)

<sup>33</sup> Salucci and Borriello (2000)

suggest that a double power-law fit, where the halo has a core density which drops off as  $\sim 1/r$  and an outer density which goes like  $\sim 1/r^3$ .<sup>34</sup> While other radial distributions have been proposed which solve some issues and raise others,<sup>35,36,37</sup> the majority of studies show that adding a spherical dark matter halo to the galactic mass distribution improves agreement between data and model, and that these dark matter haloes are essential for structure formation in the universe.

## 1.4 Gravitational Lensing (In Practice)

In 1979, using the 2.1-meter telescope at Kitt Peak National Observatory in Tucson, Arizona, Dennis Walsh, Bob Carswell, and Ray Weymann observed what appeared to be a “twin quasar” (quasi-stellar radio source). Noting the small relative separation of the two objects (only 6 arcseconds), and their nearly identically-redshifted spectra, the team proposed that what they were seeing was “a single source which has been split into two images by a gravitational lens”.<sup>38</sup> This was the first identified gravitationally-lensed object.

Since 1979, experiments such as the Optical Gravitational Lensing Experiment (OGLE) have found many thousands of examples of “microlensing” (when there is no double-image or distortion in shape, but there is a measurable fluctuation in brightness over time) using advanced software and CCD cameras.<sup>39</sup> More dramatic examples of gravitational lensing are referred to as “strong” lensing and can be seen in the objects known as G2237+0305, J1000+0221, and LRG 3-757 (seen in Figure 1.3), to name a few. When the severity of a lensing effect indicates

---

<sup>34</sup> Cintio et al. (2014)

<sup>35</sup> Navarro et al. (1997)

<sup>36</sup> Merritt et al. (2006)

<sup>37</sup> McGaugh et al. (2007)

<sup>38</sup> Walsh et al. (1979)

<sup>39</sup> Wyrzykowski et al. (2015)

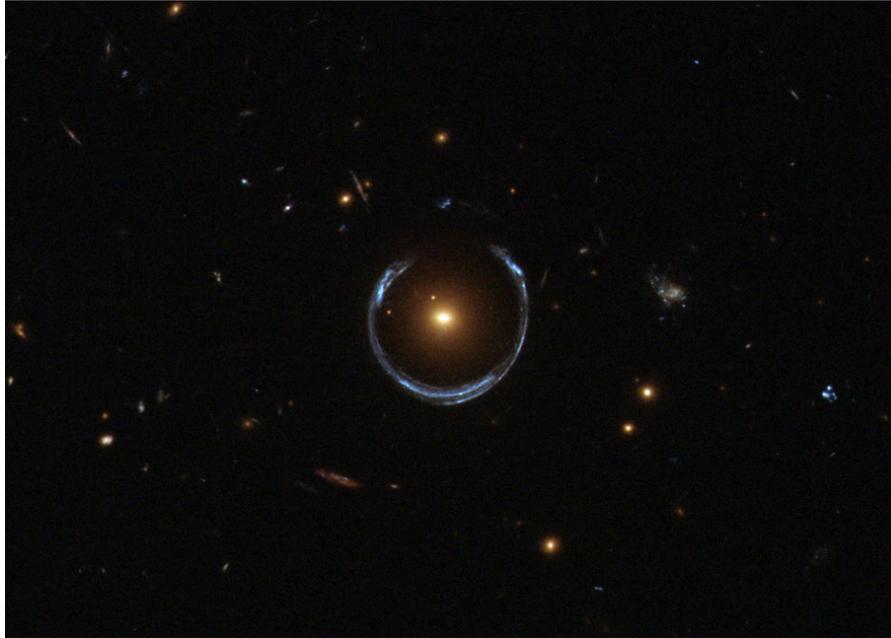


Figure 1.3: An image taken of the strong gravitational lens LRG 3-757 from the Hubble Space Telescope’s Wide Field 3 Camera. From [Nemiroff and Bonnell \(2011\)](#).

a mass density which is incompatible with the observed luminous matter, it can be concluded that dark matter must be present in large amounts in the lensing object.

Perhaps the best evidence for dark matter from gravitational lensing effects can be seen in the galaxy cluster designated “1E 0657-558”, more commonly known as the “Bullet Cluster”. Based on observations made by NASA’s/ESA’s Hubble Space Telescope and NASA’s Chandra X-Ray Observatory, the density of ordinary, visible matter can be mapped across the cluster (which is actually the aftermath of two galaxies which collided approximately 150 million years ago).<sup>40</sup> From the gravitational lensing which occurs, and at a statistical significance of  $8\sigma$ , it has been shown that the center of mass of the luminous matter does not align with the center of mass calculated via the lensing effects, and that this discrepancy

---

<sup>40</sup> [NASA / SAO / CXC / M. Markevitch et al. \(2009\)](#)

“cannot be explained with an alteration of the gravitational force law”.<sup>41</sup>

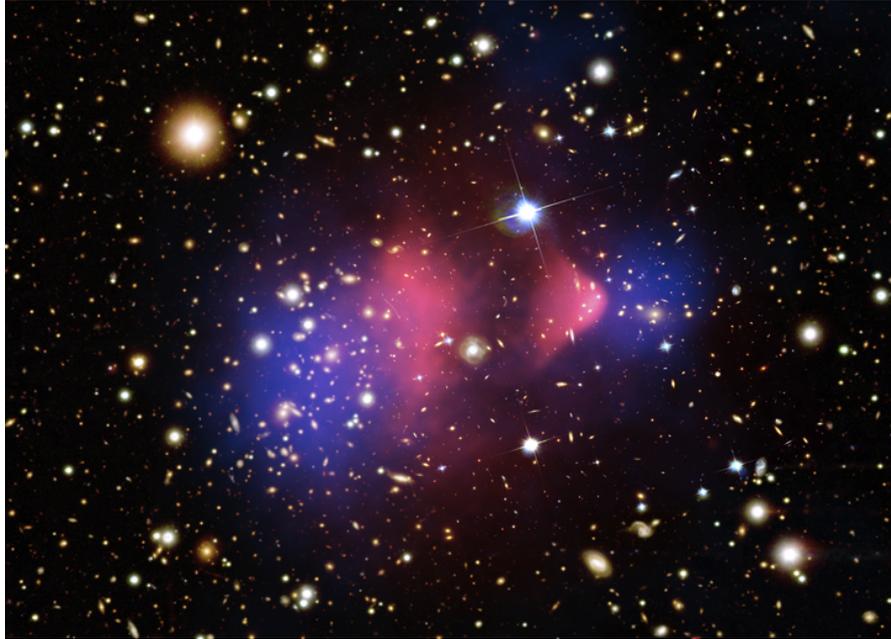


Figure 1.4: A composite visible/X-ray image of 1E 0657-56 (also known as the Bullet Cluster) from NASA’s/ESA’s Hubble Space Telescope and NASA’s Chandra X-Ray observatory. Luminous matter is falsely colored pink, dark matter is falsely colored blue. From [NASA / CXC / CfA / M. Markevitch et al. \(2009\)](#).

Another interesting aspect of the Bullet Cluster is the overall distribution of luminous versus dark matter. The luminous matter can be easily mapped, because it can be directly observed; and the dark matter distribution can be reconstructed from the lensing effects (this is just the reverse process of calculating the expected lensing of a given mass distribution). As can be seen in Figure 1.4, the ordinary matter (here falsely colored pink) is contained near the center of the cluster, while the dark matter is found far beyond the point of collision, as though it passed straight through, unaffected.<sup>42</sup> The reason for this unusual behavior will be discussed in later sections.

---

<sup>41</sup> [Clowe et al. \(2006\)](#)

<sup>42</sup> [NASA / CXC / CfA / M. Markevitch et al. \(2009\)](#)

## 1.5 Big Bang Nucleosynthesis

The process of synthesis of nucleons and light elements from the Big Bang is well-understood, though which products are created on which timescales (or, synonymously, at which energies) is a relatively complex process which can only be solved numerically.<sup>43</sup> To explain the significance of the results, only a cursory summary is needed, though a more detailed explanation of Big Bang Nucleosynthesis can be found in [Loer \(2011\)](#) or [Fan \(2016\)](#).

The relative fractions of light isotopes (mainly  $^4\text{He}$ ,  $^3\text{He}$ ,  $^2\text{H}$ , and  $^7\text{Li}$ ) which were present in the very early universe (after the average temperature of the universe fell below  $\sim 26$  GK or  $\sim 2.2$  MeV, the binding energy of deuterium,  $^2\text{H}$ ) can be measured by recording absorption spectra from distant galaxies. As their light passes through ancient interstellar gas — which we assume to be representative of the matter distribution of the primordial universe — the absorption spectra reveal the relative abundances of these simple isotopes.

With these data, as well as other known factors like the rate of condensation of light elements out of the early high-energy universe, a limit can be set on the fraction of the total universal mass-energy density which is contained in baryonic matter (matter composed of quarks, antiquarks, and their bound states, like protons and neutrons) in the universe. That fraction has been calculated as  $\sim 4 - 5\%$  of the critical density (see Subsection 2.6).<sup>44,45,46,47</sup> This implies that the majority of mass-energy in the universe must be contained in alternative, non-baryonic forms. This conclusion is found to be in agreement with analyses performed on the Cosmic Microwave Background (CMB), including recent Planck results, which

---

<sup>43</sup> [Steigman \(2007\)](#)

<sup>44</sup> [Burles et al. \(1999\)](#)

<sup>45</sup> [Planck Collaboration \(2016b\)](#)

<sup>46</sup> [Coc et al. \(2004\)](#)

<sup>47</sup> [Croton \(2013\)](#)

have constrained the dark energy density to  $\sim 68 - 69\%$  of the critical density, meaning that dark matter composes the remaining  $\sim 26 - 28\%$ .

## 1.6 The CMB and the $\Lambda$ CDM Model

The Cosmic Microwave Background, or CMB, has a long and storied history: Arno Penzias and Robert Wilson accidentally discovered this relic from the early universe while doing radio astronomy for Bell Labs in 1964.<sup>48</sup> Using a horn antenna originally built to detect radio waves reflected from balloon satellites launched into low earth orbit ( $\sim 1000$  mi altitude) for NASA’s Project Echo,<sup>49</sup> the pair wanted to identify and remove all sources of background noise from their data.

They accounted for radar and radio broadcasting and cooled the antenna receiver to 4 K using liquid helium. But a persistent 7.5 cm-wavelength noise remained. At all points in the sky, day and night, they saw this signal. Nothing they did — including clearing a nest of pigeons out of the antenna — could reduce this constant background. A preprint by a group of physicists at Princeton made its way to Penzias, and he realized the significance of his and Wilson’s discovery.<sup>50</sup>

The background they discovered is an imprint of the Big Bang. Once the universe expanded and cooled enough for free protons and electrons to form bound states<sup>51</sup> (primarily neutral hydrogen atoms), photons could move freely without being constantly scattered by these charged particles. After recombination, photons decoupled from matter and could stream freely through the universe.<sup>52</sup> The CMB, then, is a “snapshot” of the universe at the moment of photon decoupling,

---

<sup>48</sup> Penzias (2005)

<sup>49</sup> Hansen (2015)

<sup>50</sup> Wilson (1978)

<sup>51</sup> This is known as “recombination”, though since it’s the first time this ever happened, perhaps it should be known as “combination”.

<sup>52</sup> Their mean free paths became greater than the Hubble length, which in general sense, is related to the radius of the universal event horizon.

some 380,000 years after the Big Bang.

Over the next 14 billion years, this  $\sim 3300$  K background<sup>53</sup> was redshifted by a factor of about  $z = 1100$  to its current temperature of just under 3 K, a blackbody with a peak emission wavelength squarely in the microwave range, where Penzias and Wilson found it. Modern experiments like Planck have made incredibly detailed maps of the anisotropies of this background temperature (see Figure 1.5), showing that it varies by only  $\sim 10$  parts per million over the entire sky. This incredible uniformity is one of the pillars of the inflationary Big Bang model.<sup>54</sup>

And the scale of the anisotropies gives information about the mass content of the universe. Before recombination, the universe was so hot and dense that it's modeled as a photon-baryon fluid, where minute fluctuations in energy can give rise to small, dense regions of baryonic and dark matter. These gravitational wells pull in material until the repulsive electromagnetic force matches the attractive gravitational force.<sup>55</sup> When the former becomes greater than the latter, these densities “bounce back” and expand at a speed of about  $0.5c$ .<sup>56</sup> Since the dark matter doesn't interact electromagnetically, it doesn't rebound with the baryons and photons, and the equilibrium point of these oscillations stays fixed. This cosmological model is referred to as the “ $\Lambda$ CDM model”.  $\Lambda$ , the dominant component, represents the dark energy content of the universe, while “CDM” is short for “cold dark matter”.<sup>57</sup>

This process, known as baryon acoustic oscillation (BAO), results in pockets of over- and under-dense regions throughout the early universe. After photon

---

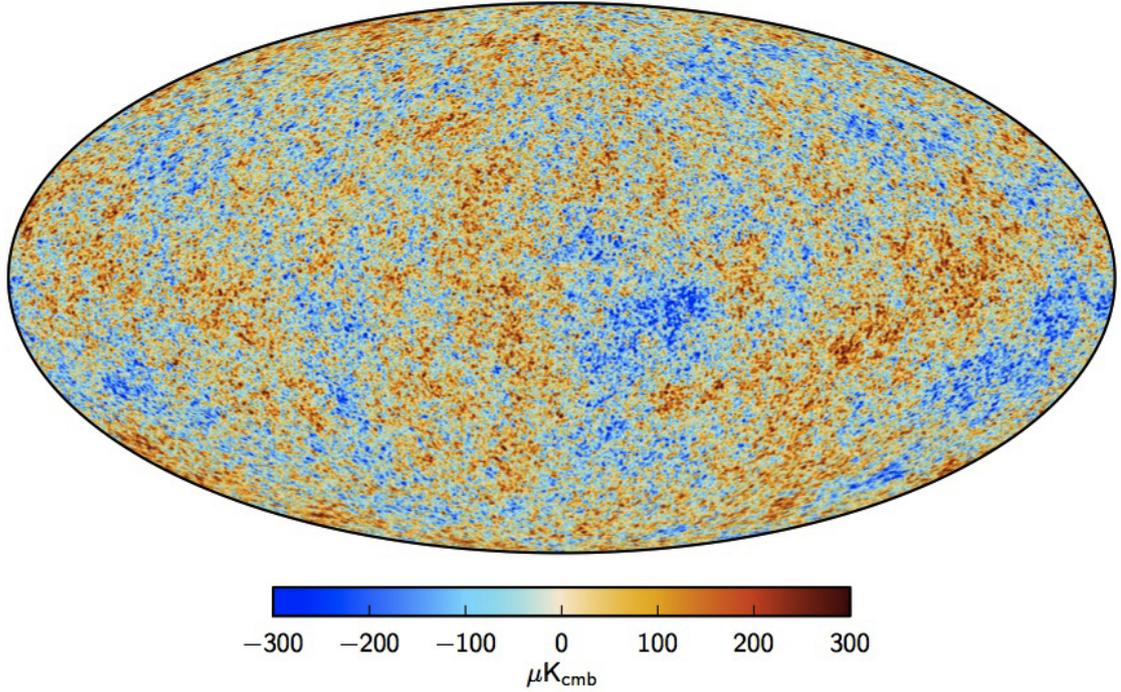
<sup>53</sup>  $\sim 0.3$  eV/ $k_B$ , from Wong (2008)

<sup>54</sup> Wong (2008)

<sup>55</sup> Fan (2016)

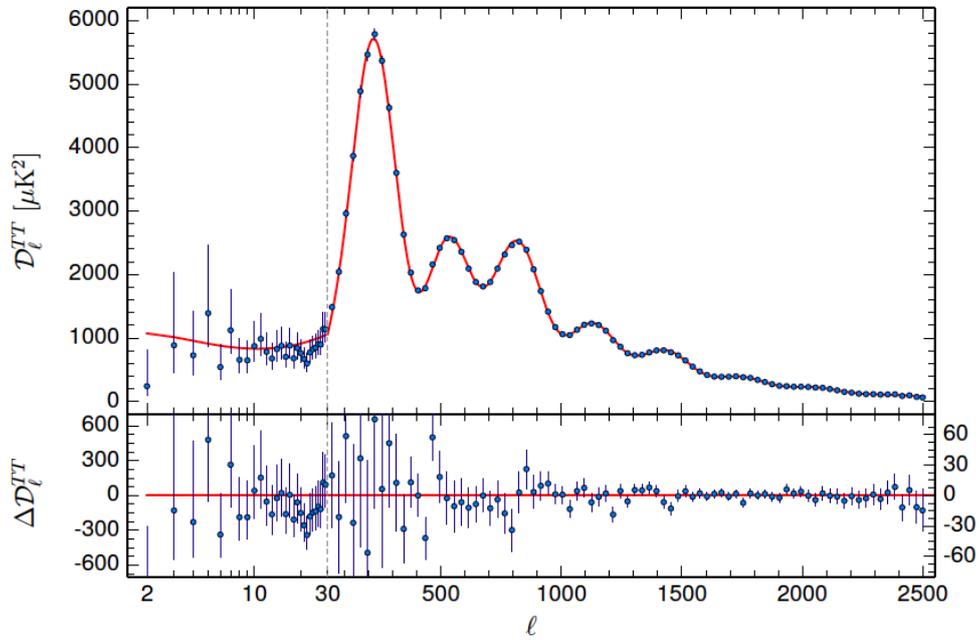
<sup>56</sup> White (2015)

<sup>57</sup> “Cold” in this case means non-relativistic at the time of decoupling, due to the large masses of the dark matter particles.



(a)

A&A 594, A1 (2016)



(b)

Figure 1.5: CMB temperature anisotropies as measured by [Planck Collaboration \(2016a\)](#). (a) Temperature fluctuations mapped across the sky. (b) Power spectrum of the temperature map, with 6-parameter  $\Lambda$ CDM fit and fit residuals.

decoupling, part of the outward pressure (the radiation pressure) was lost, and these over-densities of matter and dark matter became the seeds of structure formation in the early universe. So the “granularity” of the Planck CMB map gives us the characteristic amplitudes of these oscillations, which puts restrictions on the fraction of universal mass-energy which can be accounted for by baryons, radiation, dark matter, and dark energy.

Fitting the power spectrum of the CMB map with a six-parameter  $\Lambda$ CDM parameterization gives us the currently-accepted fractional densities of the baryons, dark matter, and dark energy of the universe, roughly 4%, 27% and 69%, respectively.<sup>58</sup>

---

<sup>58</sup> [Planck Collaboration \(2016b\)](#)

## Chapter 2

# PAST AND PRESENT DARK MATTER CANDIDATES

Many different kinds of astronomical observations provide clear evidence for the existence of dark matter. To summarize the previous chapter:

- The existence of dark matter was originally inferred from large-scale gravitational effects.
- Dark matter doesn't emit or absorb any electromagnetic radiation, and so it must be electrically neutral.
- Dark matter exists within most galaxies and clusters observed, oftentimes in concentrations tens to hundreds of times greater than the visible matter content, and it appears to be essential to structure formation in the universe.
- The study of baryon acoustic oscillations (BAO) and their effect on the cosmic microwave background (CMB) restricts the baryonic and dark matter concentrations in the early universe and now. This implies that — if most of the matter in the universe is dark matter, and most of the matter content of the universe is non-baryonic, then dark matter must be mostly non-baryonic.

- If the study of BAO is correct, then dark matter must also be stable: it must have a lifetime comparable to the age of the universe.

A variety of explanations have been put forward for dark matter since the turn of the twentieth century. In the following sections, some of the more prominent candidates from the past hundred years are explored. (Note that this list is by no means exhaustive; more complete reviews can be found in [Feng \(2010\)](#) or [Bertone and Hooper \(2016\)](#).)

## 2.1 Baryonic Matter

When it was first noticed that the visible matter content of some galaxies and galaxy clusters couldn't be reconciled with the velocities of their constituents, the initial explanation was that some portion of the mass of the galaxy or cluster wasn't being seen simply because it was "dark"; it didn't emit any light (and maybe blocked light coming from more distant objects). At the time, this argument satisfied Occam's razor; this "missing mass" should be just like any other matter, only we're having some difficulty detecting it.

In the early 1900s, dark matter candidates included "meteoric matter", "extinguished stars, dark clouds, meteors, comets", and any other variety of ordinary matter which may be difficult to see from lightyears away.<sup>1</sup> More recently, baryonic candidates like free hydrogen and ionized gases,<sup>2,3</sup> "M8 dwarf stars",<sup>4</sup> and Massive Astrophysical Compact Halo Objects (MACHOs)<sup>5</sup> have been proposed. All of these objects are composed of normal baryonic matter. These bodies all exist, but

---

<sup>1</sup> [Bertone and Hooper \(2016\)](#)

<sup>2</sup> [Penzias \(1961\)](#)

<sup>3</sup> [Meekins et al. \(1971\)](#)

<sup>4</sup> [Bertone and Hooper \(2016\)](#)

<sup>5</sup> MACHOs include objects like red, brown, and white dwarfs, rogue planets, and black holes.

not in large enough quantities to account for the observed large-scale gravitational effects. In addition, the recent Planck results restrict baryonic matter to compose no more than about 4% of universal mass-energy.<sup>6</sup>

## 2.2 Black Holes

Black holes are a favored dark matter candidate among the uninitiated, but since black holes generated from stellar collapse, including supermassive black holes found at the centers of galaxies, acquire most of their mass from baryonic sources, they fall into the same category as the objects outlined in the previous section. These black holes cannot contribute more than a few percent to the total mass-energy content of the universe.

However, primordial black holes — which were created when the universe was still a hot, dense plasma — could compose some larger fraction of the universal dark matter content. These objects would be created from the collapse of the over-dense regions discussed at the end of Section 1.6, which are largely composed of dark matter. (Though, this just seems to push the problem of determining the nature of dark matter itself back a step further.) These primordial black holes would also be stable over the age of the universe, because Hawking radiation is a very slow process.<sup>7</sup>

These black holes have been largely restricted to two mass ranges: “sub-lunar” black holes with masses about  $10^{-6} \cdot M_{Moon}$  to  $10^{-2} \cdot M_{Moon}$  and intermediate-mass black holes with masses of about  $M_{Sun}$  to  $1000 \cdot M_{Sun}$ .<sup>8</sup> (Interestingly, the colliding

---

<sup>6</sup> [Planck Collaboration \(2016b\)](#)

<sup>7</sup> According to [Page \(1976\)](#), a primordial black hole with a mass about one one-hundred-billionth the mass of the moon (or, the mass of a cubic kilometer of water) would be just finishing decaying at the present time. And black hole lifetimes are proportional to the mass cubed.

<sup>8</sup> [Carr et al. \(2016\)](#)

black holes detected by LIGO in 2016 each had a mass of  $\sim 30 \cdot M_{Sun}$ , consistent with primordial black hole masses.<sup>9)</sup> Primordial black holes are still a viable candidate for dark matter, with one study in particular showing that in addition to potentially solving the dark matter problem, they would also improve the fit of the  $\Lambda$ CDM model to the Cosmic Infrared Background.<sup>10</sup> With LIGO up and running, the universal abundance of sub-lunar and intermediate-mass black holes should be better constrained in the coming years, potentially clearing up the very foggy picture as to whether these are a significant component of the dark matter content of the universe.

## 2.3 Neutrinos

Standard Model neutrinos ( $\nu$ ) come in three “flavors” (electron  $e$ , muon  $\mu$ , and tau  $\tau$ ) and are massless, but a succession of experiments from the 1960s to present have provided evidence in favor of *neutrino oscillation*, the spontaneous changing of a neutrino from one flavor state to another.<sup>11,12,13</sup> Neutrino oscillation requires that each of the three neutrino flavors ( $\nu_e$ ,  $\nu_\mu$ , and  $\nu_\tau$ ) have some small mass. Due to measurement uncertainty, experiments to date have only been able to put a limit on the sum of the three masses,<sup>14</sup> and the squared difference in mass ( $\Delta m_{2,1}^2$ ,  $\Delta m_{3,1}^2$ , and  $\Delta m_{3,2}^2$ ) between each of the three mass states ( $\nu_1$ ,  $\nu_2$ , and  $\nu_3$ ), which do not necessarily correspond to the flavor states. Additionally, the squared difference in mass between  $\nu_1$  and  $\nu_2$  is thought to be about  $\Delta m_{2,1}^2 = 7.5 \times 10^{-5} \text{ (eV}/c^2)^2$ , while the difference between  $\nu_3$  and  $\nu_1$  (or  $\nu_3$  and  $\nu_2$ ) is about

---

<sup>9</sup> Bird et al. (2016)

<sup>10</sup> Kashlinsky (2016)

<sup>11</sup> Davis et al. (1968)

<sup>12</sup> Super-Kamiokande Collaboration (1998)

<sup>13</sup> F. P. An (2012)

<sup>14</sup> Giusarma et al. (2013)

$\Delta m_{3,(1 \text{ or } 2)}^2 = \pm 2.5 \times 10^{-3} \text{ (eV/c}^2\text{)}^2$ . This means that  $\nu_1$  and  $\nu_2$  have very similar masses, but  $\nu_3$  is either significantly heavier (known as “normal mass hierarchy”) or significantly lighter (“inverted mass hierarchy”) than  $\nu_1$  and  $\nu_2$ , and it is not clear at the moment which of these is the case.<sup>15</sup> This is known as the *neutrino hierarchy problem*.

While experiments like KATRIN, which aims to start collecting data in August of this year,<sup>16</sup> hope to clarify this mass hierarchy and discover the absolute masses of each of these three mass states to within  $0.2 \text{ eV/c}^2$ ,<sup>17</sup> the sum of the neutrino masses is all that is needed to determine whether or not they would be good dark matter candidates.<sup>18</sup> For neutrinos to account for the large-scale gravitational effects caused by dark matter, the universe must be “neutrino dominated”, where  $\Omega_\nu \approx 1$ , and  $m_\nu \approx 10 \text{ eV/c}^2$ .<sup>19</sup> It is now understood that this is not the case. Neutrinos less massive than  $\sim 1 \text{ MeV/c}^2$  would be relativistic at the time of decoupling and would hinder — rather than help — structure formation in the universe. Additionally, it has been shown that galaxy clusters, binary galaxies, and galactic haloes would require vastly different (and sometimes contradictory) neutrino masses to explain the observed gravitational effects. In short, neutrinos may contribute to some of the universal missing mass, but they certainly cannot be a dominant component of it. In 1979, near the end of the rotation curve boom, Tremaine and Gunn found these discrepancies and made the claim that “the most likely remaining possibilities [for dark matter] are small black holes or very low-mass stars, or perhaps some much heavier stable neutral particles.”<sup>20</sup>

---

<sup>15</sup> [Qian and Vogel \(2015\)](#)

<sup>16</sup> [Karlsruher Institut für Technologie \(2016\)](#)

<sup>17</sup> [KATRIN Collaboration \(2015\)](#)

<sup>18</sup> [Planck Collaboration \(2016b\)](#)

<sup>19</sup> [Zeldovich et al. \(1982\)](#)

<sup>20</sup> [Tremaine and Gunn \(1979\)](#)

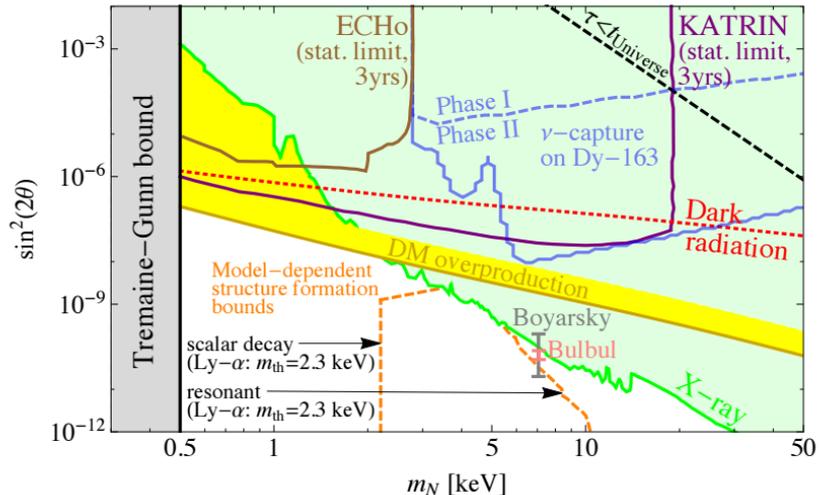


Figure 2.1: Experimental and observational constraints on sterile neutrino dark matter candidates with mass  $0.3 \text{ keV}/c^2$  to  $50 \text{ keV}/c^2$ . The total mixing angle between the three flavor states,  $\theta$ , is defined via  $\theta^2 \equiv \sum_{\alpha=e,\mu,\tau} |\theta_{\alpha 1}|^2$ . The green shaded area shows parameter space which has been excluded via non-observation of X-ray photons from sterile neutrino decay. The error bars on the mixing angle for a possible  $7.1 \text{ keV}/c^2$  sterile neutrino seen by [Boyarsky et al. \(2014\)](#) and [Bulbul et al. \(2014\)](#) are also shown. Figure from [Adhikari et al. \(2017\)](#).

## 2.4 Sterile Neutrinos

The fact that neutrinos have mass requires new physics beyond the Standard Model. Many models which include massive neutrinos also predict right-handed neutrinos which have zero electric, weak, or strong charge: sterile neutrinos.<sup>21</sup> These neutrinos would decay due to their (very small) mixing with the three standard neutrino flavors. This decay would produce a line in the X-ray band, which could be observed.<sup>22</sup> Searches performed by NASA’s Chandra X-ray Observatory<sup>23</sup> and others<sup>24,25</sup> have found hints of possible sterile neutrino decay lines at  $2.5 \text{ keV}$  and  $3.55 \text{ keV}$ , indicating neutrino masses of  $5 \text{ keV}/c^2$  and  $7.1 \text{ keV}/c^2$ ,

<sup>21</sup> [Canetti et al. \(2012\)](#)

<sup>22</sup> [Boyarsky et al. \(2009\)](#)

<sup>23</sup> [Loewenstein and Kusenko \(2010\)](#)

<sup>24</sup> [XQC Collaboration \(2015\)](#)

<sup>25</sup> [Bulbul et al. \(2014\)](#)

respectively, though these data are still being scrutinized.<sup>26</sup> Sterile neutrinos are an appealing dark matter candidate because they could also explain the baryon asymmetry of the universe, sometimes called the matter-antimatter asymmetry.<sup>27</sup> Sterile neutrinos could be a prominent dark matter candidate, but they have not yet been definitively detected nor ruled out completely. A graphical summary of sterile neutrino parameter space (mixing angle versus mass) can be seen in Figure 2.1.

## 2.5 Axions

If you record a video of a pendulum swinging, or take successive photographs of two binary stars as they orbit each other, or record the collision of two billiard balls, and you watch these images or videos in reverse, the physics you see in the reversed images is described by the same equations as when you play them in the correct order. This is because classical mechanics is *time-reversal-invariant*. These invariances under simple positive/negative involutions (functions which are their own inverses) are called “symmetry” laws.

Many symmetries other than time (“T”) symmetry exist in nature. Relevant to dark matter is what’s known as “CP symmetry”, which says that, if charge (“C”) and parity (“P”, also known as “handedness”, or “mirror symmetry”) are both inverted, then the same physics apply as in the original case. On their own, it seems logical that the same physics should describe the motions and interactions of particles and antiparticles (with the same mass, but the opposite charge), and that these laws shouldn’t change if we view those same motions and interactions in a mirror. For neutrinos, though, C or P symmetry individually do not hold:

---

<sup>26</sup> Note that, since these neutrinos are expected to be much more abundant than heavier DM candidates, they can have smaller masses and still account for the missing mass in galaxies and clusters.

<sup>27</sup> [Adhikari et al. \(2017\)](#)

right-handed neutrinos and left-handed antineutrinos have never been experimentally observed. The charge-parity inverse of left-handed neutrinos — right-handed antineutrinos — do exist, though, indicating CP symmetry.<sup>28</sup>

CP symmetry exists in electromagnetism, but is broken for some weak force phenomena, for example in the decay of kaons, K. And although CP symmetry breaking is not forbidden in strong interactions, it has never been observed experimentally, and measurements of the neutron dipole moment put very stringent limits on its likelihood. This lack of experimental evidence for CP symmetry breaking in quantum chromodynamics (QCD), in spite of it not being explicitly forbidden, is known as the strong CP problem.<sup>29</sup> In 1977, Peccei and Quinn proposed a light, neutral particle known as the “axion” which could solve this problem.<sup>30</sup> The axion is not technically a stable particle, but since its lifetime goes as  $m^5$ , even axions in the mass range  $10^{-6}$  eV to  $10^{-3}$  eV would have very long lifetimes, and could be viable dark matter candidates.

Although the very small mass of the axion seems to, like the neutrino, disqualify it as a cold dark matter candidate, mechanisms for axion production in the early universe can leave the axion with little or no kinetic energy.<sup>31</sup> Combined with their extremely weak couplings to ordinary matter,<sup>32</sup> axions could be a viable dark matter candidate. Several experimental efforts are currently underway to detect these particles,<sup>33,34</sup> with some claiming to have seen signatures of them in X-ray observations of the Sun<sup>35</sup> and, interestingly, Josephson junctions.<sup>36</sup>

---

<sup>28</sup> [DAEδALUS Experiment \(2014\)](#)

<sup>29</sup> [Fan \(2016\)](#)

<sup>30</sup> [Peccei and Quinn \(1977\)](#)

<sup>31</sup> [Erken et al. \(2012\)](#)

<sup>32</sup> [Kuster et al. \(2008\)](#)

<sup>33</sup> [CAST Collaboration \(2015\)](#)

<sup>34</sup> [Stern \(2017\)](#)

<sup>35</sup> [Fraser et al. \(2014\)](#)

<sup>36</sup> [Beck \(2015\)](#)

## 2.6 Supersymmetry

The Minimal Supersymmetric Standard Model (MSSM) pairs every standard model particle with a supersymmetric (SUSY) partner, which maintains all of the gauge interactions and quantum numbers as its Standard Model (SM) counterpart, but differs in spin by  $1/2$ . (So all SM fermions are paired with SUSY bosons and vice-versa.) SM fermions, like leptons and quarks, get scalar SUSY partners called sleptons and squarks. SM bosons, like the Higgs and gauge bosons, get fermionic SUSY partners called Higgsinos and gauginos.<sup>37</sup>

In SM interactions, baryon number and lepton number are conserved. For instance, although neutrons are stable when confined to atomic nuclei, free neutrons have a half-life of approximately  $\tau_{1/2} = 10$  minutes. That decay generally proceeds as:

$$n \rightarrow p + e + \bar{\nu}_e \tag{2.1}$$

where a neutron,  $n$ , decays into a proton,  $p$ , an electron,  $e$ , and an electron antineutrino,  $\bar{\nu}_e$ . The neutron and proton each have a baryon number of  $+1$ , while the electron has a lepton number  $+1$  and the electron antineutrino (since it is an antiparticle) has a lepton number  $-1$ . Thus, the sum of baryon numbers and the sum of the lepton numbers on either side of the equation are conserved. (Also, note that charge is conserved on either side of this decay.)

The MSSM allows for the breaking of this symmetry: lepton number and baryon number need not be individually balanced on either side of an interaction or a decay. Instead, a new symmetry is introduced:  $R$ -parity.  $R$ -parity requires that the  $R$ -number on either side of a decay or interaction is conserved, where  $R$  is defined as

$$R = (-1)^{3B+L+2s} \tag{2.2}$$

---

<sup>37</sup> Csáki (1996)

where  $B$  is the total baryon number,  $L$  is the total lepton number, and  $s$  is the total spin on either side of the interaction. SM particles have  $R = 1$ , while their SUSY partners have  $R = -1$  (since they have the same  $B$  and  $L$  numbers, but differ in spin,  $s$ , by  $1/2$ ). One consequence of  $R$ -parity is that the lightest supersymmetric particle (LSP) *must* be stable.<sup>38</sup>

In the MSSM, the bino,  $\tilde{B}$ ,<sup>39</sup> wino,  $\tilde{W}$ , and neutral Higgsinos,  $\tilde{H}_u^0$  and  $\tilde{H}_d^0$ , can mix, forming neutral particles called “neutralinos”, which only interact at the weak scale. Of the four neutralinos, the lightest one,  $\tilde{\chi}_1^0$ , has a mass on the order of  $100 \text{ GeV}/c^2$ . In the MSSM and most extensions to MSSM, the lightest neutralino is also the LSP. This makes  $\tilde{\chi}_1^0$  a weakly-interacting massive particle, or WIMP.<sup>40</sup>

SUSY superparticles (“sparticles”) are massive ( $\mathcal{O}(100 \text{ GeV} \rightarrow 10 \text{ TeV})$ ), and would have been created only in the early universe, when ambient temperatures were very high. As the universe cooled, all sparticles except for the LSP would have decayed away, leaving a relic density of WIMP matter. WIMPs are attractive dark matter candidates because, unlike many other proposed explanations, they can account for *all* of the missing mass required by observations like Planck’s maps of the CMB.<sup>41</sup>

---

<sup>38</sup> To balance the  $R$ -number of the decay of a SUSY particle, we must have  $R_{SUSY} \rightarrow \sum_n R_n$ , where  $R_a$  is the  $R$ -number of the particle, and there are  $n$  daughters from the decay. Since  $R_{LSP} \equiv -1$ , the minimal number of daughters is three, where  $R_{LSP} \rightarrow R_a + R_b + R_c$ . Here,  $R_a$  and  $R_b$  must both be  $-1$ , with  $R_c = +1$ , for the equation to balance. Since the LSP is the lightest SUSY particle, by definition it cannot decay into two new SUSY particles, and so the decay is forbidden by  $R$ -parity conservation.

<sup>39</sup> Pronounced “bee-no”. Similarly, wino is pronounced “wee-no”.

<sup>40</sup> Steffen (2009)

<sup>41</sup> Feng (2010)

## 2.6.1 The WIMP Miracle

To understand why WIMPs are such a prominent dark matter candidate, one needs to know how they are produced just after the Big Bang. (Note that a thorough, step-by-step explanation of this derivation, including all sources, can be found in [Appendix A](#), though the main points are covered here.) In the very early universe ( $t < 10^{-10}$  sec), the temperature at all points is so high that the heavy SUSY particles are in thermal equilibrium with ordinary SM particles. But as the universe expands and cools, the dark matter number density drops according to the Boltzmann equation, which describes the behavior of systems which are not in thermal equilibrium. The differential number density of particles in a 6-dimensional (3 spatial, 3 momentum) space is

$$dN = f(\vec{r}, \vec{p}, t) d^3r d^3p \quad (2.3)$$

where  $dN$  is the number of particles within a physical differential volume  $d^3r = dx dy dz$  about  $\vec{r} = (x, y, z)$ , and with momenta within a certain differential volume in momentum-space,  $d^3p = dp_x dp_y dp_z$  about  $\vec{p} = (p_x, p_y, p_z)$ .  $f(\vec{r}, \vec{p}, t)$  is a probability density function which depends on all seven variables of position, momentum, and time.

The probability density function,  $f$ , can have components which depend on collision rates, diffusion, and external forces, among other things. For SUSY particles, there are four terms:

$$\begin{aligned} \frac{dn_i}{dt} = & -3Hn_i - \sum_{j=1}^N \langle \sigma_{ij} v_{ij} \rangle (n_i n_j - n_i^{eq} n_j^{eq}) - \sum_{j \neq i} [\Gamma_{ij} (n_i - n_i^{eq}) - \Gamma_{ji} (n_j - n_j^{eq})] \\ & - \sum_{j \neq i} \left[ \langle \sigma'_{X_{ij}} v_{ij} \rangle (n_i n_X - n_i^{eq} n_X^{eq}) - \langle \sigma'_{X_{ji}} v_{ji} \rangle (n_j n_X - n_j^{eq} n_X^{eq}) \right] \quad (2.4) \end{aligned}$$

where  $n_i$  is the number density of the  $i$ -th SUSY species. Descriptions of these terms can be found in [Appendix A](#). We can greatly simplify this equation by noting that *all* SUSY particles eventually decay into the LSP, and so we can define a total number density,  $n$ , according to

$$\frac{dn}{dt} = \sum_i^N \frac{dn_i}{dt} \quad (2.5)$$

which simplifies Equation [2.4](#) to

$$\frac{dn}{dt} = -3Hn - \sum_{i,j=1}^N \langle \sigma_{ij} v_{ij} \rangle (n_i n_j - n_i^{eq} n_j^{eq}) \quad (2.6)$$

A few other simplifying assumptions yields

$$\frac{dn}{dt} = -3Hn - \langle \sigma_{eff} v \rangle (n^2 - n_{eq}^2) \quad (2.7)$$

where  $H$  is the hubble constant,  $\sigma_{eff}$  is an “effective annihilation cross section” of these SUSY particles,  $v$  is the “relative velocity” between the particles which collide and annihilate,  $n_{eq}$  is the number density of particles when they’re in thermal equilibrium, and angled brackets indicate a thermal average.

We can see that the first term on the right-hand side of Equation [2.7](#) dilutes the number density due to the expansion of the universe, while the second term increases or decreases the number density, relative to the equilibrium density, due to decreased or increased likelihood of SUSY annihilations, respectively. The second term, then, relaxes the number density toward the equilibrium density.

Next, we make a change of variables. Instead of  $dn/dt$ , the common convention is to work with  $dY/dx$ , where we define

$$Y(T) \equiv \frac{n}{s(T)} \quad \text{and} \quad x(T) \equiv \frac{m_\chi}{T} \quad (2.8)$$

where  $s$  is the entropy density,  $m_\chi$  is the mass of a particular SUSY particle, and  $T$  is the temperature. The entropy density in a radiation-dominated universe (like the early universe) goes as

$$s(T) = \frac{2\pi^2}{45} T^3 \cdot h_{eff}(T) \quad (2.9)$$

where  $h_{eff}$  gives the effective degrees of freedom for bosons and fermions.

Making this change of variables, and doing a bit of algebra (again, covered in detail in [Appendix A](#)), yields the following equation

$$\frac{dY}{dx} = -\lambda(T) \cdot \frac{Y^2(T) - Y_{eq}^2(T)}{x^2(T)} \quad (2.10)$$

where

$$\lambda(T) \equiv \sqrt{\frac{\pi}{45G}} \cdot g_\star^{1/2}(T) m_\chi \cdot \langle \sigma_{eff} v \rangle \quad (2.11)$$

$Y$  is proportional to the energy density as a fraction of the critical density,  $\Omega_\chi$ , of the WIMP, while  $x$  is proportional to the time,  $t$ , and inversely proportional to the temperature,  $T$ .  $g_\star^{1/2}(T)$  depends on the degrees of freedom and is defined in [Appendix A](#). As the temperature of the universe,  $T$ , drops below a particle's mass ( $\times k_B^{-1} c^{-2}$ ), production of that particle drops exponentially as  $e^{-m_\chi/T}$  (and so  $Y_{eq}$  falls exponentially) and higher-mass particles annihilate and produce lower-mass ones. However, as the universe expands ( $-3Hn$ ), the number density of these particles falls, and it becomes less and less likely that they can collide and annihilate. In essence, at a certain temperature  $T_f$ , the particle “freezes out”, and the exponential decrease of its energy density begins to stabilize.

At early times, then, a WIMP is in thermal equilibrium with the universe,  $Y$  is  $\approx Y_{eq}$ , and  $dY/dx$  is consistent with zero. As the temperature drops,  $Y_{eq}$  falls exponentially, and therefore so do  $dY/dx$  and  $Y$ . However,  $\lambda/x^2$  is also linearly

dependent on temperature, and so as the universe expands and cools,  $dY/dx$  is suppressed, and so  $Y$  asymptotically approaches some value,  $Y_0$ : the value of  $Y$  at present. The behavior of this equation can be seen in Figure 2.2.

To translate the value of  $Y$  to WIMP energy density,  $\Omega_\chi$ , we first take the limit of Equation 2.10 as  $T$  gets small (dropping the “(T)”s for brevity) and integrate from freeze-out to present day (where  $x \rightarrow \infty$  at present day)

$$\lim_{T \rightarrow T_f} \frac{dY}{dx} = -\lambda \cdot \frac{Y^2}{x^2} \quad \longrightarrow \quad \frac{dY}{Y^2} = -\lambda \cdot \frac{dx}{x^2} \quad \longrightarrow \quad -\frac{1}{Y_f} + \frac{1}{Y_0} = \frac{\lambda}{x_f} - \frac{\lambda}{\infty} \quad (2.12)$$

At freeze-out, the WIMP abundance  $n$  is much greater than present day (though it does quickly stabilize), so  $Y_f \gg Y_0$ , and we can write

$$Y_0 \simeq \frac{x_f}{\lambda} \quad (2.13)$$

Finally, we recognize that the energy density of dark matter as a fraction of the critical density of the universe,  $\Omega_\chi$  can be written

$$\Omega_\chi = \frac{\rho_{0,\chi}}{\rho_c} = \frac{n_0 m_\chi}{\rho_c} \quad (2.14)$$

where  $\rho_{0,\chi}$  is the WIMP mass density at present day,  $\rho_c$  is the critical density from the Friedmann equations, and  $n_0$  is the numerical density of WIMPs at present day. Note that recent Planck results find the curvature of the universe (TT+lensing+ext) to be  $\Omega_K = 0.000 \pm 0.005$ , consistent with zero.<sup>42</sup> This means that the total mass-energy density of the universe is essentially equal to the critical density. Throughout this dissertation, the two fractions are used interchangeably.

---

<sup>42</sup> [Planck Collaboration \(2016b\)](#)

Substituting Equation 2.8, then 2.13, then 2.11 yields:

$$\Omega_\chi = \frac{Y_0 s_0 m_\chi}{\rho_c} = \frac{x_f s_0 m_\chi}{\lambda \rho_c} = \sqrt{\frac{45G}{\pi}} \frac{x_f s_0}{\rho_c g_\star^{1/2} \langle \sigma_{eff} v \rangle} \quad (2.15)$$

and substituting numerical values yields<sup>43,44</sup>

$$\Omega_\chi \sim \frac{10^{-39} \text{ cm}^2}{\langle \sigma_{eff} \rangle} \sim \frac{10^{-27} \text{ cm}^3/\text{s}}{\langle \sigma_{eff} v \rangle} \quad (2.16)$$

This is the “WIMP Miracle”: the fact that a stable, electrically neutral particle predicted by SUSY, with an annihilation cross section of approximately  $10^{-39} \text{ cm}^2$  — on the order of the weak scale — has a predicted relic density which matches the observed dark matter relic density. Note that WIMP dark matter actually has a fairly large range of possible masses, as shown in Figure 2.3. Finding WIMP dark matter with masses from 100 GeV to 1 TeV is the goal of many direct-detection experiments, including the DarkSide dark matter search.

## 2.7 Other Candidates

### 2.7.1 Gravitinos

A slightly different SUSY extension to the SM, called Supergravity (SUGRA), emerges when SUSY is allowed to be spontaneously broken — that is, when SUSY is changed from a global to a local symmetry. In the MSSM, the gravitino is the massless superpartner of the graviton, the hypothetical gauge boson which mediates the gravitational force. In SUGRA, the gravitino acquires a mass and is a potential dark matter candidate.<sup>45</sup> Gravitinos are a class of particles known

---

<sup>43</sup> [Dodelson \(2003\)](#)

<sup>44</sup> [Belanger \(2012\)](#)

<sup>45</sup> [Steffen \(2006\)](#)

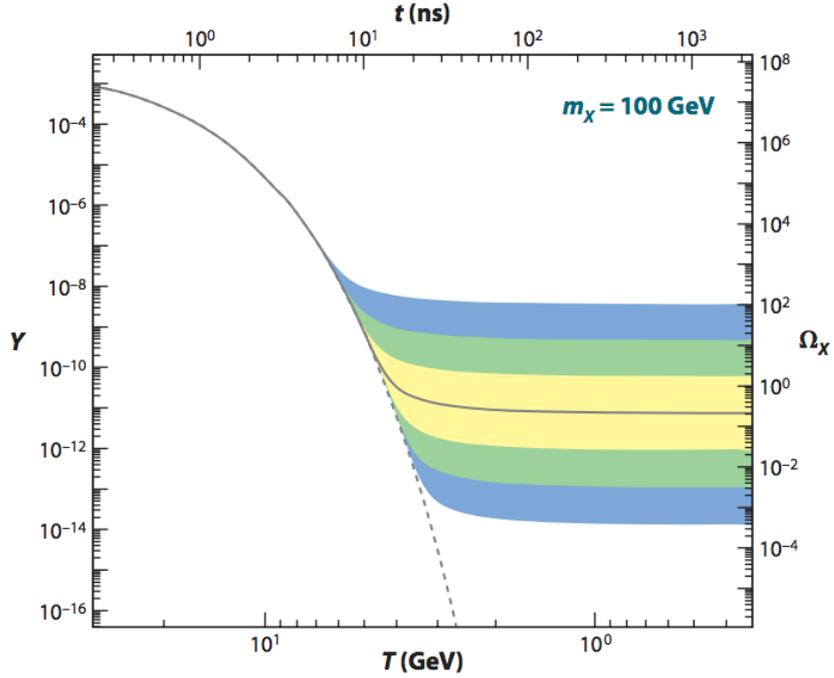


Figure 2.2: Behavior of the relic density equation as a function of  $T$ , in units of GeV. The dotted line shows the exponential fall of  $Y_{eq}$  as the temperature decreases, while the solid line shows the fall and later stabilization of  $Y$ , due to the suppression of  $\lambda/x^2$ . The solid line shows  $Y(\sigma_{eff}, m_\chi)$  where  $m_\chi$  is set at 100 GeV and  $\sigma_{eff}$  is chosen such that the correct present-day dark matter relic density is achieved. The shaded regions show  $Y$  where  $\sigma_{eff}$  is changed by factors of 10,  $10^2$ , and  $10^3$ . From [Feng \(2010\)](#).

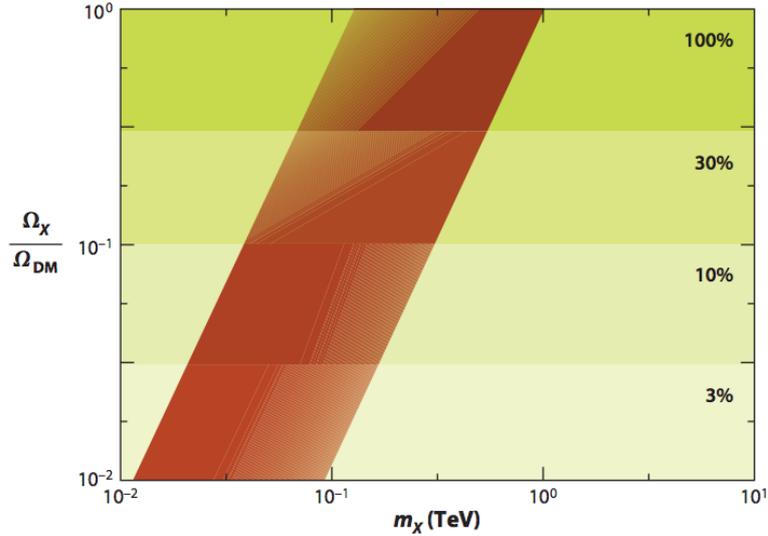


Figure 2.3: Percentage of total dark matter energy density which can be accounted for by particles of mass  $m_\chi$ . Note that a particle with any mass from roughly 100 GeV to 1 TeV can account for 100% of dark matter. From [Feng \(2010\)](#).

as superWIMPs, which obtain their relic density as the products of the decay of heavier WIMP-like particles, typically known as Next-to-Lightest Supersymmetric Particles (NLSPs). Kaluza-Klein particles are another example of a superWIMP.<sup>46</sup>

## 2.7.2 Kaluza-Klein Particles

Much like the MSSM LSP couples with SM particles and is stable due to R-parity, the Lightest Kaluza-Klein partner (LKP), predicted by extensions to the SM with Universal Extra Dimensions, couples to SM particles and is stable due to a Kaluza-Klein (KK) parity.<sup>47,48</sup> Probable LKP candidates include KK photons, KK neutrinos, and various KK bosons, and proposed methods of detection include the analysis of neutrino energy spectra from the Sun and the galactic center, as well as looking for specific TeV-scale gamma ray peaks from annihilation.<sup>49</sup> Kaluza-

<sup>46</sup> [Feng et al. \(2004\)](#)

<sup>47</sup> [Servanta and Tait \(2003\)](#)

<sup>48</sup> [Kakizaki et al. \(2005\)](#)

<sup>49</sup> [Tsuchida and Mori \(2017\)](#)

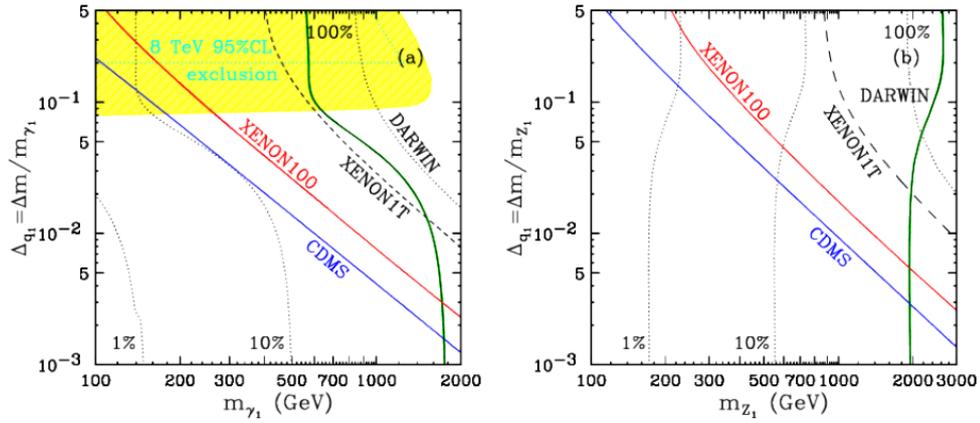


Figure 2.4: Current experimental limits on LKP photon (left) and Z-boson (right) masses, where the vertical axis is the mass splitting,  $\Delta q_1 = (m_{q_1} - m_{LKP}) / (m_{LKP})$ , and  $m_{q_1}$  is the KK quark mass. Points to the left of the dotted black lines marked “1%” (“10%”) could account for 1% (10%) of the dark matter relic density. The very narrow green band gives the limits from the Planck experiment, and the solid black line within it delineates where LKP particles could account for 100% of the relic dark matter density. The yellow shaded region shows the parameter space excluded by LHC results, while the solid red and blue lines show parameter space excluded (to the left) by the XENON100 and CDMS experiments, respectively. The dashed black line shows the expected limit from XENON1T, which could exclude nearly all of the parameter space for  $\gamma$  LKP dark matter. From [Arrenberg et al. \(2013\)](#).

Klein particles are currently being searched for in data gathered from the LHC and astronomical observations, though direct-detection experiments like DarkSide can put bounds on the LKP parameter space, as well.<sup>50</sup> The current remaining parameter space of two particular LKP candidates can be seen in Figure 2.4.

### 2.7.3 Axinos

Axinos,  $\tilde{a}$ , are the MSSM superpartners of axions, discussed in Section 2.5. Like WIMPs, they are stable or semi-stable on cosmological time scales, though they interact even more weakly with SM particles, and so they are classified as Extremely Weakly Interacting Massive Particles (EWIMPs). Axinos can be produced via two mechanisms: in thermal equilibrium with the early universe, prior to inflation; or after freeze-out of heavier SUSY particles, as one of their decay products. Accordingly, axinos can exist over a huge range of masses, from eV to multi-TeV scales. eV-scale axinos will act as hot dark matter (HDM), while GeV or TeV-scale axinos would behave like cold dark matter (CDM). Since both production mechanisms are allowed, a mixture of the two would result in axinos behaving like a warm dark matter (WDM) candidate. For axinos to be a prominent CDM candidate, they must have masses on the MeV/c<sup>2</sup> to GeV/c<sup>2</sup> scale. As can be seen in Figure 2.5, axinos have cross sections several orders of magnitude below those predicted for WIMP dark matter. Choi remarks that “the exceedingly weak strength of axino interactions makes axino detection in direct DM search experiments as well as at collider experiments rather hopeless”, though very meticulous searches can be performed to detect these particles in LHC data.<sup>51</sup>

---

<sup>50</sup> [Arrenberg et al. \(2013\)](#)

<sup>51</sup> [Choi and Kim \(2013\)](#)

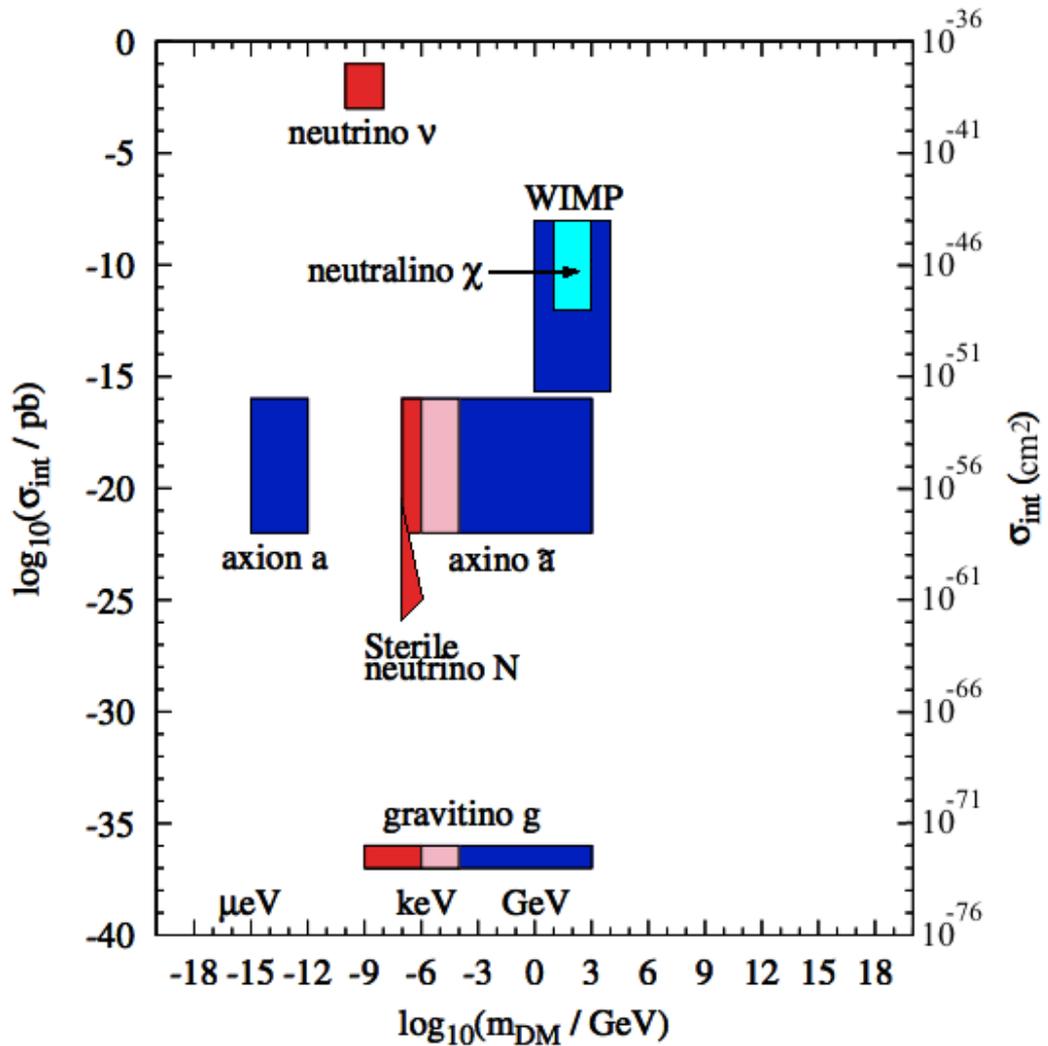


Figure 2.5: Interaction cross section vs. mass for all of the dark matter candidates discussed in this chapter. Blue (pink, red) areas indicate cold (warm, hot) dark matter candidates. Note that Kaluza-Klein photons and Z-bosons would occupy essentially the same parameter space as WIMPs on this diagram. From [Choi and Kim \(2013\)](#) (altered slightly for clarity).

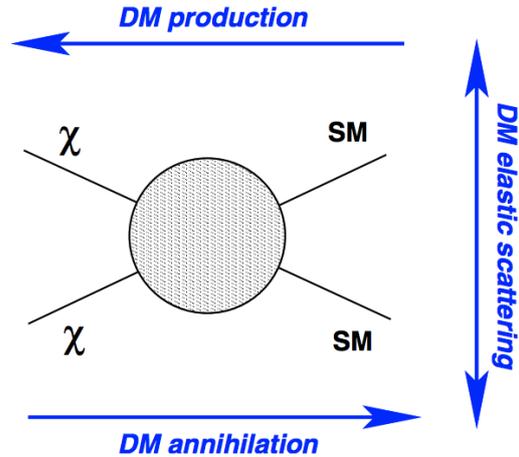


Figure 2.6: Schematic showing the three possible modes of direct dark matter (DM) detection: production in particle colliders, elastic scattering off of SM particles, or annihilation and subsequent photon production. From [Dienes et al. \(2015\)](#).

## 2.8 Detection of Dark Matter

To date, all observational evidence in support of dark matter has been indirect. We infer its existence primarily through its gravitational effects, and generally only on large (galactic, cluster, universal) scales. If the solution to the “missing mass” problem really is a particle, or class of particles, then we must be able to produce or directly detect these particles to confirm their existence. Production of WIMP dark matter could occur in large particle colliders like the Large Hadron Collider (LHC), while direct and indirect detection experiments utilize a variety of detection methods. Note that, if a signal seems to suggest the existence of a WIMP-like particle, it would have to be confirmed by another of these three methods. For that reason, I’ll cover each briefly.

### 2.8.1 Production

The LHC is currently operating at a proton-proton center of mass energy of  $\sqrt{s} = 13$  TeV and would be expected to produce lower-mass WIMPs (less than a few

hundred  $\text{GeV}/c^2$ ) if they existed. The method by which the ATLAS and CMS experiments look for these particles is by calculating the missing transverse energy (MET) in a collision.<sup>52</sup> The MET indicates the production of one or more particles which have escaped detection by the calorimeters and other apparatus. The LHC is able to put very strict limits on the WIMP-nucleon cross section,<sup>53</sup> and is complimentary to direct dark matter searches.

## 2.8.2 Indirect Detection

Indirect detection experiments are concerned primarily with looking for the products of dark matter annihilations and decays, usually photons, neutrinos, and charged antiparticles. These annihilations should produce peaks and lines in the energy spectra of these particles which would be difficult to explain by any other cosmic source. Since annihilation likelihood is enhanced in regions of space where DM is densest, indirect searches are typically focused on gravity wells like the Sun or the galactic center.<sup>54</sup>

The MAGIC gamma ray telescopes, located in the Canary Islands, are a pair of imaging atmospheric Čerenkov telescopes for detecting air showers from very high energy gamma rays, which lead to the generation of Čerenkov radiation upon interacting with particles in the Earth’s atmosphere.<sup>55</sup> The MAGIC Collaboration found no significant excess of gamma rays above 80 GeV when observing the Perseus Cluster of galaxies and NGC 1275 in 2008, restricting the boost factor<sup>56</sup> for photon emission from DM-DM annihilation.<sup>57</sup> A similar observation of the

---

<sup>52</sup> CMS Collaboration (2012)

<sup>53</sup> CMS and ATLAS Collaborations (2015)

<sup>54</sup> Conrad (2014)

<sup>55</sup> MAGIC Collaboration (2017)

<sup>56</sup> This is a “fudge factor” which increases the annihilation rate by assuming larger-than-expected dark matter densities, or invoking other obscure particle physics effects. Granger (2017)

<sup>57</sup> MAGIC Collaboration (2010)

dwarf spheroidal galaxy Draco by the same collaboration found no excess of events around 140 GeV, where a DM-DM annihilation peak was predicted by Monte Carlo simulations.<sup>58</sup> This constrained the parameter space of DM particles in minimal SUGRA (mSUGRA) models, but had little impact on MSSM predictions.

In 2012, the Fermi Large Area Telescope (LAT) made a  $3.2\sigma$  claim for the annihilation of a 130 GeV/ $c^2$  particle with cross section  $\langle\sigma v\rangle = 1.27 \times 10^{-27} \text{ cm}^3/\text{s}$ , in other words, a perfect DM candidate.<sup>59</sup> Although the observation caused a flurry of analyses to be published, the observation could not be replicated.<sup>60</sup> Gamma telescopes like Fermi LAT and MAGIC continue to search for the signals of dark matter annihilation.

The other major avenue for indirect detection of cosmic dark matter annihilation involves neutrino telescopes. These typically fall into one of three categories: underground, underwater, and under-ice. This natural shielding is required to block cosmic rays and cosmogenic isotopes, which would otherwise be insurmountable sources of background for these detectors.

## Underground Neutrino Detectors

Super-Kamiokande (Super-K) is the world's leading underground neutrino observatory, and arguably the world's most famous one. In 1998, Super-K observed atmospheric neutrino oscillations, explaining the solar neutrino problem and upending the standard model of particle physics.<sup>61,62</sup> Among many other studies, Super-K searched for neutrinos in coincidence with the first gravitational waves detected by LIGO in 2016<sup>63</sup> and severely restricted the allowed parameter space

---

<sup>58</sup> [Albert et al. \(2008\)](#)

<sup>59</sup> [Weniger \(2012\)](#)

<sup>60</sup> [Liang et al. \(2016\)](#)

<sup>61</sup> [Super-Kamiokande Collaboration \(1998\)](#)

<sup>62</sup> [Super-Kamiokande Collaboration \(2004\)](#)

<sup>63</sup> [Super-Kamiokande Collaboration \(2016\)](#)

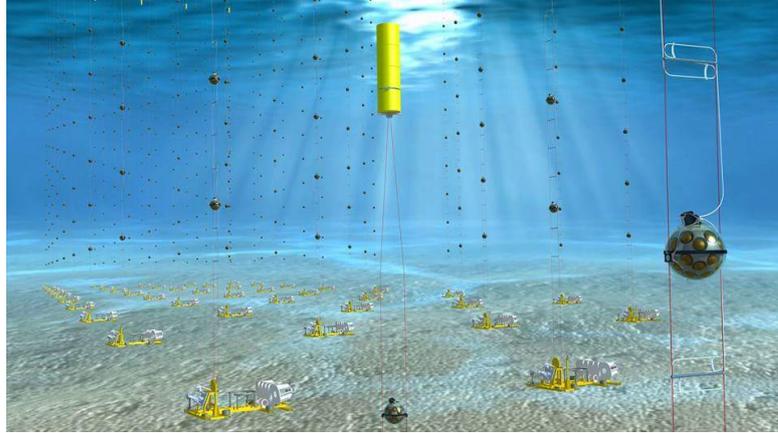


Figure 2.7: Artist’s conception of the KM3NeT detector. From [KM3NeT \(2017\)](#).

of massive magnetic monopoles, which are required by some grand unified theories (GUTs).<sup>64</sup> With regards to dark matter, Super-K has looked for neutrinos generated from WIMP annihilation in the Sun, and has been able to put very stringent limits on the spin-dependent WIMP-nucleon cross section. For spin-independent processes, Super-K has been able to rule out very low-mass candidates ( $\sim 7$  GeV), but is non-competitive with direct detection experiments at high WIMP masses.<sup>65,66,67</sup> Other underground neutrino telescopes include Borexino, LVD, and ICARUS (defunct) at the Laboratori Nazionali del Gran Sasso (LNGS) in Gran Sasso, Italy; the Baksan Neutrino Observatory near the Georgian border in southwestern Russia; and the Helium and Lead Observatory (HALO) at SNOLab in Ontario.

### Underwater Neutrino Detectors

ANTARES is the world’s foremost underwater neutrino telescope, which detects these particles when they lead to the generation of Čerenkov radiation upon entering the water target and exciting the atoms therein. ANTARES is composed

---

<sup>64</sup> [Super-Kamiokande Collaboration \(2012\)](#)

<sup>65</sup> [Super-Kamiokande Collaboration \(2011\)](#)

<sup>66</sup> [Kappl and Winkler \(2011\)](#)

<sup>67</sup> [Super-Kamiokande Collaboration \(2015\)](#)

of an array of lines secured to a seabed off the coast of southeast France. Each of the twelve lines has 75 attached photomultiplier tubes (PMTs), which detect the Čerenkov light generated when a neutrino enters the water from the atmosphere above.<sup>68</sup> The successor to ANTARES, KM3NeT, which is currently under construction, will be a series of seven arrays around the Mediterranean, similar to ANTARES, but consisting of over 350,000 PMTs in total. ANTARES is already the largest neutrino telescope in the northern hemisphere, but KM3NeT will be orders of magnitude larger and more sensitive.<sup>69,70</sup> Another notable underwater neutrino telescope is the Baikal Deep Underwater Neutrino Telescope (BDUNT), which operates using the same basic principles as ANTARES and is situated under Lake Baikal, in Russia.

### **Under-Ice Neutrino Detectors**

IceCube is the only large-scale under-ice neutrino detector in operation today. It works on the same principles as ANTARES, except the neutrino target material is ice, instead of water. Situated on the Antarctic continent, IceCube gets a unique view of the southern hemisphere sky, which complements the northern hemisphere observations of ANTARES, Super-K, and Borexino.

There are a few other searches which use ice in a less direct way, including NARC, the successor to RICE, which sits atop the Antarctic ice sheet and aims to detect radio-wavelength Čerenkov light; the Antarctic Impulse Transient Antenna (ANITA) which operates under the same principle as NARC; and the Advanced Thin Ionization Calorimeter (ATIC): essentially just a calorimeter attached to a high-altitude balloon which floats above the Antarctic and measures the energies of cosmic rays.

---

<sup>68</sup> [ANTARES \(2015\)](#)

<sup>69</sup> [KM3NeT \(2017\)](#)

<sup>70</sup> [Trovato et al. \(2014\)](#)

### 2.8.3 Direct Detection

Like indirect dark matter searches, direct dark matter searches employ a variety of techniques to record the low-energy signals expected from dark matter interactions. All direct detection methods share one thing in common: they look for some energy deposited in a target material when a WIMP recoils off of that target material. Modeling the recoil as an elastic scatter off a stationary target, the energy deposited in the target material by the recoiling nucleus following elastic WIMP scattering is

$$E = \mu^2 \frac{v^2}{m_T} (1 - \cos \theta) \quad (2.17)$$

where  $\mu = (m_\chi m_T)/(m_\chi + m_T)$  is the reduced mass of the system,  $m_\chi$  is the mass of the incident dark matter particle,  $m_T$  is the mass of the target,  $v$  is the velocity of the dark matter particle, and  $\theta$  is the center of mass angle of the colliding particles.

The dark matter velocity,  $v$ , is  $\mathcal{O}(10^2)$  km/s, and is a function of several factors, including the local dark matter density, the speed of the solar system as it moves through the galaxy, and the position of the Earth on its orbit about the sun. If a 100 GeV WIMP is incident on an argon nucleus ( $m_T = 39.948$  amu =  $37.21$  GeV/ $c^2$ ), the spin-independent recoil energy (assuming  $v = 360$  km/s)<sup>71</sup> is

$$E = (27.12 \text{ GeV}/c^2)^2 (3.9 \times 10^{-8} \text{ c}^4/\text{GeV})(1 - \cos \theta) = 29 \text{ keV} \cdot (1 - \cos \theta) \quad (2.18)$$

Note that the recoil energy is, of course, a function of the recoil angle  $\theta$ .  $\theta \sim 0$  when collisions are grazing, and  $\theta \rightarrow \pm\pi$  when collisions are nearly head-on. The recoil energy is also a function of the WIMP mass  $m_\chi$ . The relationship between recoil energy, WIMP mass, and recoil angle can be seen in Figure 2.8. Note that for a WIMP with mass  $100 \text{ GeV}/c^2 < m_\chi < 1000 \text{ GeV}/c^2$ , recoil energies are in

---

<sup>71</sup> Loer (2011)

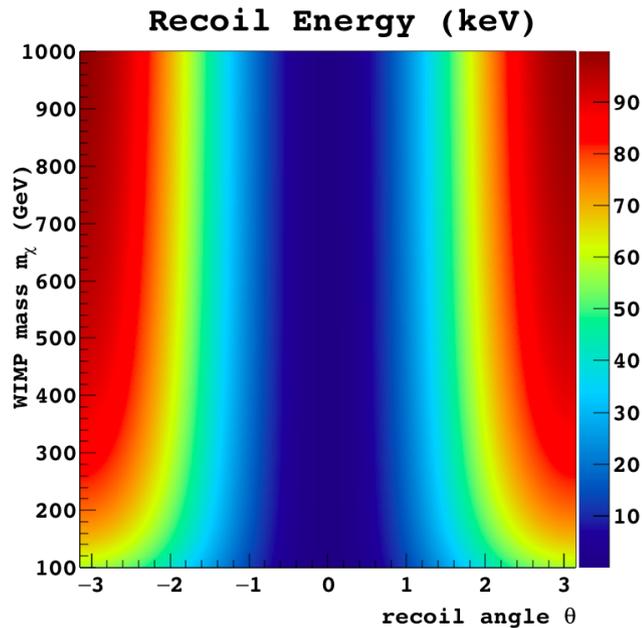


Figure 2.8: Recoil energy of an argon nucleus as a function of incident WIMP mass  $m_\chi$  and recoil angle  $\theta$ . The WIMP is assumed to have a velocity of 360 km/sec. The z-scale is in units of keV.

the range  $0 \rightarrow 100$  keV.

Unlike indirect detection experiments — which aim to observe the energy spectra of SM particles and seek out any deviations from theoretical predictions which might be indicative of WIMP annihilation — direct detection experiments are expected to have incredibly low rates, typically given in counts per kilogram of target per day. Expected spin-independent WIMP event rates with liquid noble element targets are on the order of  $10^{-3}$  /kg/day or lower.<sup>72</sup>

As both the expected event rates and the expected energy deposits of these interactions are miniscule, the apparatus meant to detect them must be very large and very sensitive. Direct detection experiments typically utilize phonons, ionization, or scintillation, and many use more than one of these.

---

<sup>72</sup> Poggiani (2017)

## Phonons

When the target material of a direct dark matter search is a crystal (usually germanium, silicon, cesium iodide, or sodium iodide), a WIMP elastic recoil should deposit detectable energy into the target in the form of phonons — vibrations in the crystal lattice. Of course, if the target material is at room temperature, the thermal noise within the crystal will overwhelm any potential WIMP phonon signal. For this reason, crystal lattice phonon experiments are held at extreme cryogenic temperatures (usually  $10 \text{ mK} < T < 100 \text{ mK}$ ) to reduce the thermal background. In addition, these crystals can only be manufactured so large, and so to achieve significant active masses, arrays of crystals are sometimes used.

The Super Cryogenic Dark Matter Search (SuperCDMS, successor to CDMS and CDMS II) is a phonon and ionization-based detector which uses germanium and silicon crystals. It is projected to reach a sensitivity of  $\sim 10^{-43}$  for small WIMP masses ( $\lesssim 5 \text{ GeV}/c^2$ ).<sup>73</sup> Other phonon-based experiments include CRESST<sup>74</sup> and EDELWEISS<sup>75</sup>. In addition to measuring phonons via bolometric methods, CRESST also detects scintillation light, and EDELWEISS also measures ionization of the target material.

## Ionization

Ionization is another channel through which the kinetic energy of a WIMP can be transferred to a target material. When a WIMP strikes an atom (or molecule) of the material, the kinetic energy transferred from the WIMP to the target atom can cause that atom to collide with further atoms and transfer kinetic energy to their electrons. If the kinetic energy transfer exceeds the ionization potential, then the latter atoms will release those electrons, and they can be detected. Typically

---

<sup>73</sup> [Agnese et al. \(2016\)](#)

<sup>74</sup> [cresst.de](#)

<sup>75</sup> [edelweiss.in2p3.fr](#)

these electrons are drifted toward wire arrays, charge amplifiers, or — in the case of dual-phase Time-Projection Chambers (TPCs) — a gaseous region, where they can generate electroluminescence.

Many experiments detect both free ionization electrons and scintillation, but a few others detect solely ionization. Of these, the more notable experiments are CoGeNT,<sup>76</sup> which uses a single 440 g high-purity Ge crystal as its target material; and DRIFT,<sup>77</sup> which is a directional dark matter detector, filled with gaseous carbon disulfide (CS<sub>2</sub>) and carbon tetrafluoride (CF<sub>4</sub>). The DRIFT experiment is currently the only large-scale direction-sensitive dark matter detector in operation, though several others are being planned.<sup>78</sup>

## Scintillation

Scintillation light is typically generated in two ways in a direct dark matter search: either directly, where an excited target atom or molecule de-excites by releasing photons; or indirectly, where an excited target atom or molecule releases ionization electrons, which are then drifted toward a gaseous region by an applied electric field, generating electroluminescence in that gaseous region. In situations where both of these de-excitation channels are desired, experimenters often opt to build Time-Projection Chambers (TPCs), which can detect both the initial scintillation (commonly called S1) and the secondary electroluminescence (S2).

Some of the better-known TPCs being used to detect dark matter today include XENON<sup>79</sup> (XENON10, XENON100, and now XENON1T); LZ,<sup>80</sup> which is the successor to both the ZEPLIN and LUX experiments; PandaX;<sup>81</sup> ArDM;<sup>82</sup> and

---

<sup>76</sup> [cogent.pnnl.gov](http://cogent.pnnl.gov)

<sup>77</sup> [Daw et al. \(2011\)](#)

<sup>78</sup> [Mayet et al. \(2016\)](#)

<sup>79</sup> [xenon.astro.columbia.edu](http://xenon.astro.columbia.edu)

<sup>80</sup> [lz.lbl.gov](http://lz.lbl.gov)

<sup>81</sup> [pandax.physics.sjtu.edu.cn](http://pandax.physics.sjtu.edu.cn)

<sup>82</sup> [ArDM Collaboration \(2016\)](#)

DarkSide<sup>83</sup> (DarkSide-10, DarkSide-50, and soon DarkSide-20k). Of these, all but DarkSide and ArDM use liquid xenon. DarkSide and ArDM use liquid argon as the target material. (In addition, a new xenon-based TPC has been proposed, DARWIN, though it is still very much in the planning stage.<sup>84</sup>)

Experiments which only utilize the scintillation channel include DEAP-3600<sup>85</sup> (successor to DEAP-1 and precursor to DEAP-50T), DAMA<sup>86</sup> (which includes both DAMA/NaI and DAMA/LIBRA), and the forthcoming SABRE experiment.<sup>87</sup> The target materials of these detectors are very different, though, as DAMA and SABRE detect the scintillation generated by particles interacting with NaI(Tl) crystals, while DEAP detects scintillation generated in a large liquid argon target.

Of particular interest is the forthcoming SABRE Experiment, which will be housed at the Stawell Underground Physics Laboratory (SUPL) in Stawell, Australia. SABRE is meant to be an independent test of the DAMA/LIBRA experiment, with twin detectors at SUPL in Australia and LNGS in Italy. The goal of SABRE is to provide a cross-check on a consistent annual modulation signature seen by the DAMA/LIBRA experiment (see Figure 2.10), which could be indicative of dark matter, but is currently in conflict with several other experiments. A timeline showing this discovery, along with other experimental results in favor of and against it, can be seen in Figure 2.11.

## Others

Direct detection experiments which don't utilize any of the three detection channels mentioned above are few and far between. Of the ones that don't, bubble chamber experiments seem to be the most common. These use superheated fluids

---

<sup>83</sup> [darkside.lngs.infn.it](http://darkside.lngs.infn.it)

<sup>84</sup> [Aalbers et al. \(2016\)](#)

<sup>85</sup> [deap3600.ca](http://deap3600.ca)

<sup>86</sup> [people.roma2.infn.it/ dama](http://people.roma2.infn.it/dama)

<sup>87</sup> [Froborg \(2015\)](#)

( $C_3F_8$ ,  $CF_3I$ ,  $C_2ClF_5$ ) which, when struck by an incident particle, nucleate around the interaction site, creating small bubbles. The bubbles can be detected via acoustic sensors or via direct imaging. The PICO<sup>88</sup> and COUPP<sup>89</sup> experiments both use large volumes of superheated liquid, while the SIMPLE experiment<sup>90</sup> uses droplets of superheated liquid suspended in a gel matrix.

#### 2.8.4 State of the Field

The search for dark matter, nearly a century after its discovery, is still a very active field. Figure 2.9 shows a (slightly outdated) snapshot of current and projected spin-independent cross section limits. This is also called an exclusion limit plot. In general, dual-phase noble element TPCs like XENON, LUX, and DarkSide are most sensitive to WIMPs ( $m_\chi \gtrsim 100 \text{ GeV}/c^2$ ), while experiments employing crystals as their targets are more sensitive to other, lower-mass dark matter candidates. The current best-limit for WIMP-mass particles comes from the LUX Collaboration, and is  $2.2 \times 10^{-46} \text{ cm}^2$  at  $m_\chi = 50 \text{ GeV}$ . XENON1T promises to be 10x more sensitive than LUX and has begun taking data this year. LZ, when finished, will be 10x more sensitive than that. With these advancements taken into consideration — along with the pending construction of the dual-continent SABRE experiment (which will either confirm or rule out DAMA’s controversial annual modulation signal) — all signs point toward significant progress on the “dark matter question” within the next decade.

---

<sup>88</sup> [picoexperiment.com](http://picoexperiment.com)

<sup>89</sup> [www-coupp.fnal.gov](http://www-coupp.fnal.gov)

<sup>90</sup> [sites.google.com/site/dm2011simple/](http://sites.google.com/site/dm2011simple/)

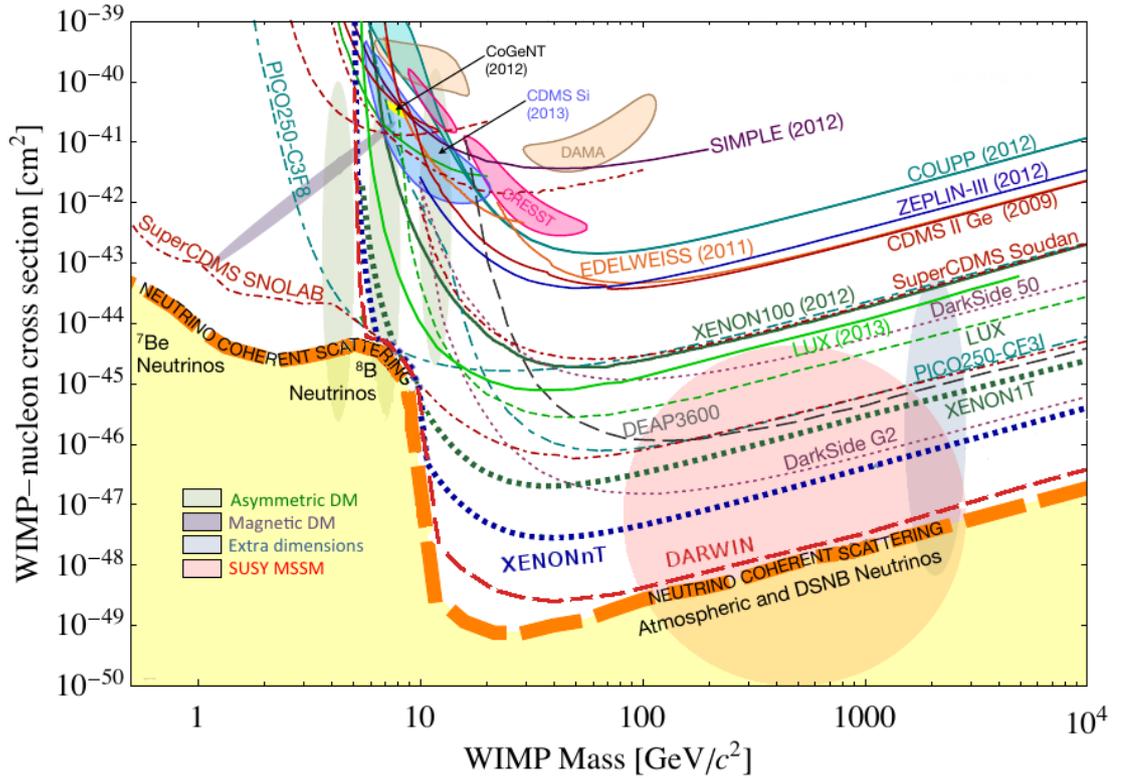


Figure 2.9: Exclusion limits for most prominent dark matter searches, up-to-date as of 2014. Solid lines are results and dashed or dotted lines are projections. Borderless shaded regions show allowed parameter spaces for several different theories, bordered shaded regions show regions where unconfirmed dark matter signals have been seen. The thick orange dashed line shows the “neutrino floor”, the region below which coherent solar and Diffuse Supernova Neutrino Background neutrino scattering becomes a significant, irreducible background. From [Klasen et al. \(2015\)](#).

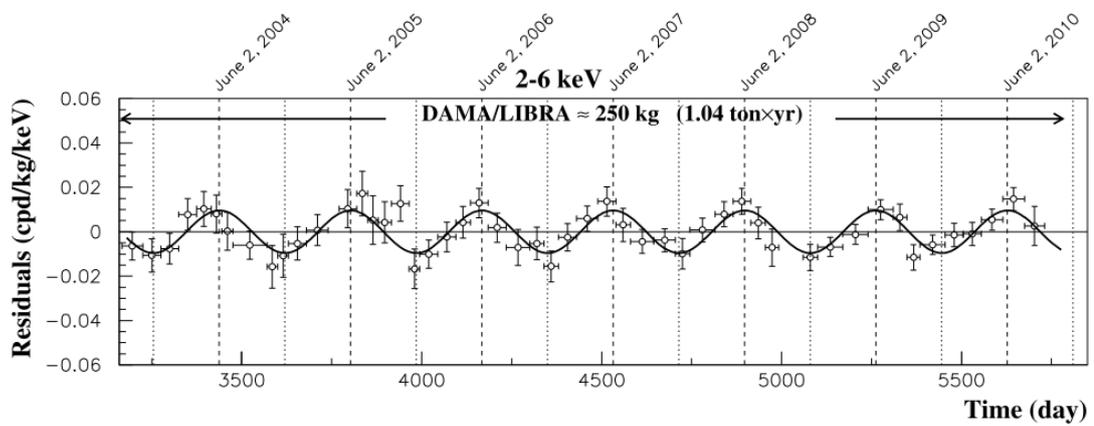


Figure 2.10: Single-hit scintillation event residuals in counts per (day · kg · keV) in the combined 1.04 ton · year exposure of DAMA/LIBRA. The period of oscillation is  $0.998 \pm 0.002$  year and peaks around June 2 each year. When combined with the DAMA/NaI observations, the annular oscillations are significant at the  $9.3\sigma$  level. From [Bernabei et al. \(2013\)](#).

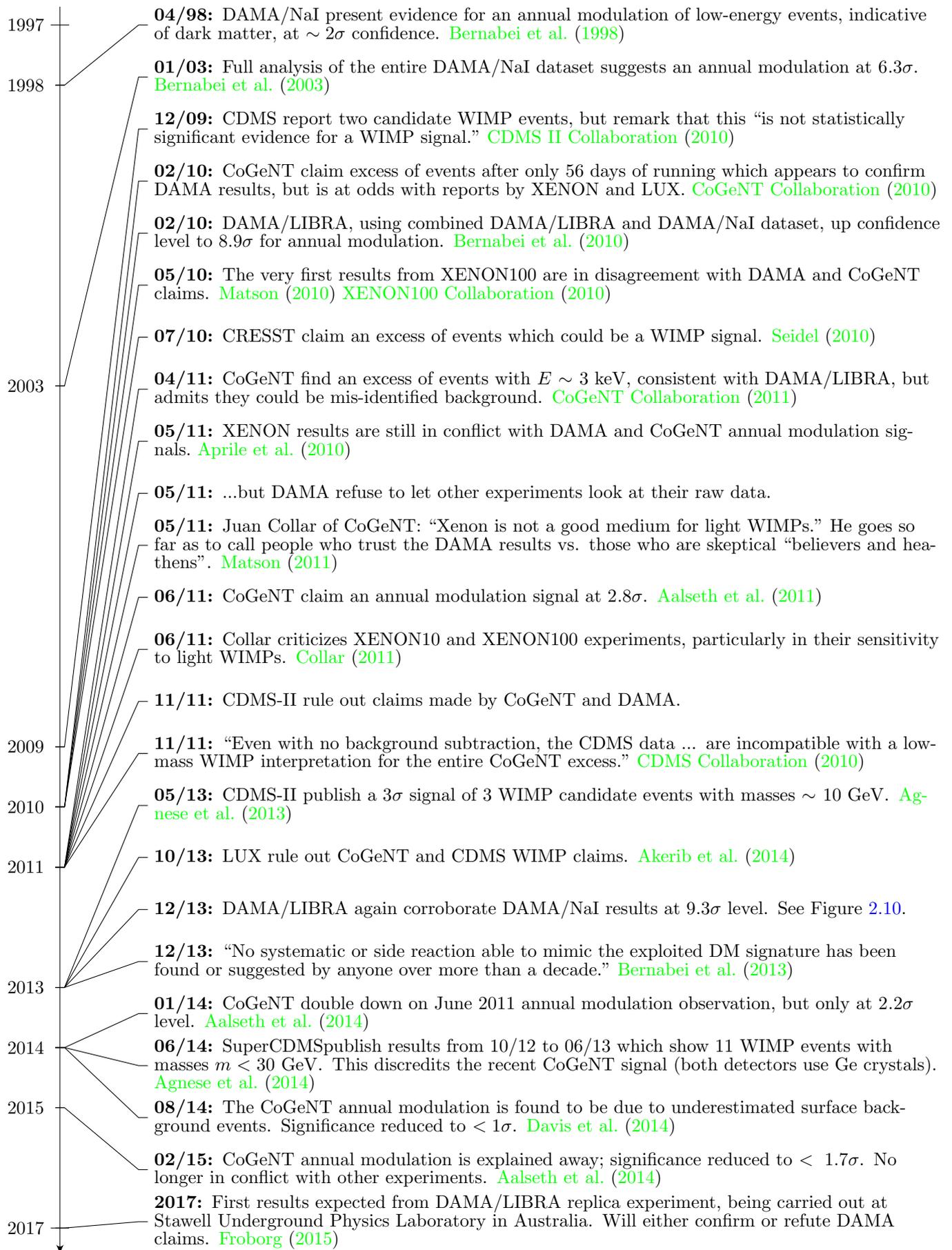


Figure 2.11: Timeline of prominent direct dark matter detection claims, 1997 to present.

# Chapter 3

## DARKSIDE

The DarkSide Collaboration was founded around 2009, shortly after a particular kind of argon (Underground Argon or UAr) was discovered in a CO<sub>2</sub> source in Cortez, Colorado. Originally, the Collaboration consisted of just a few members from a handful of American universities, but it has since grown into a group of researchers, students, and technicians representing nearly fifty different schools and organizations from over a dozen countries. An up-to-date author list can be found in [Appendix B](#). (Note that both the DarkSide Experiment and the DarkSide Collaboration are sometimes referred to as “DarkSide” or just “DS”, though it will be clear from context which is which.)

The DarkSide Experiment is a multi-generation direct dark matter detection effort. The prototype detector, DarkSide-10 (also “DS-10”), was a 10 kg active volume TPC, first turned on in 2010 and used to guide the development of the current-generation detector, DarkSide-50 (with a 50 kg active volume; also “DS-50”). DS-50 has been collecting data since October 2013. Though originally filled with high purity argon distilled out of the atmosphere (Atmospheric Argon or AAr), in March of 2015, DS-50 was emptied and on 1 April 2015 was filled with

UAr.<sup>1</sup> A second-generation detector with a 20 tonne active volume, DarkSide-20k (“DS-20k”) is currently being planned. DS-20k will use Silicon Photomultipliers (SiPMs) instead of the traditional Photomultiplier Tubes (PMTs) and is expected to go live around 2020.

DS-50 is currently housed in Hall C of the *Laboratori Nazionali del Gran Sasso* (the National Laboratories at Gran Sasso, LNGS) in the *Parco Nazionale del Gran Sasso e Monti della Laga*, within a stone’s throw of the town of Assergi, in central Italy. LNGS is the world’s largest underground laboratory, and is supremely accessible, as it’s built off of the A24 freeway which tunnels directly underneath the Gran Sasso massif.

LNGS is composed of both above-ground laboratories and offices, and underground experimental halls, which are almost directly underneath Corno Grande, the highest peak in the Apennines. The rock above the experimental halls provides  $\sim 3800$  meters water equivalent (m.w.e.) of passive shielding against cosmic rays and cosmogenic muons.<sup>2</sup> Passive shielding, of course, simply blocks some portion of the radiation from passing through by virtue of its attenuative properties, while active shielding generally emits light when a particle passes through it. This light can be detected and the particle can be characterized, allowing for certain events to be discarded. The active shielding in DarkSide consists of two apparatus, the Water Čerenkov Detector and the Liquid Scintillator Veto, both of which surround the active volume, the DarkSide-50 dual-phase argon TPC.

### 3.1 Active Shielding

The DarkSide detector has a “nesting doll”-type structure, where the innermost region, containing the active volume within the TPC, is surrounded by multi-

---

<sup>1</sup> Fan (2015)

<sup>2</sup> D’Angelo (2015)

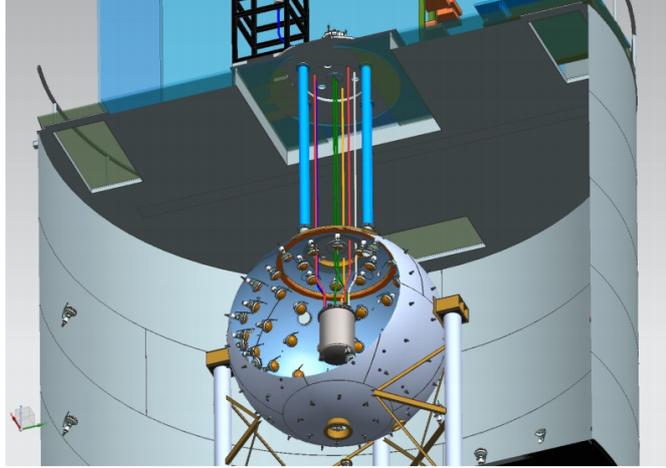


Figure 3.1: Rendering of the DS-50 cryostat, which contains the argon TPC, mounted within the spherical LSV and the cylindrical WCD. The entire apparatus is assembled beneath the radon-free clean room, CRH (Clean Room Hanoi). From [Agnes et al. \(2015\)](#).

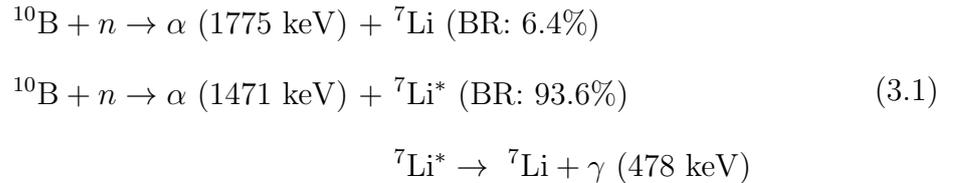
ple concentric spherical and cylindrical apparatus. The outermost “shell” of the detector is an 11 m diameter, 10 m tall cylindrical tank, filled with 1 k tonne of ultra-purified water. This Water Čerenkov Detector (WCD, also called the “Muon Detector” or “MD” in some publications) was originally part of the Borexino Counting Test Facility (CTF). The inside surface of the tank is covered with a laminated Tyvek-polyethylene-Tyvek reflector and is outfitted with 80 ETL 9351 8” PMTs, with an average 27% quantum efficiency (QE) at 420 nm.<sup>3</sup> As muons and other relativistic particles traverse the water inside the WCD, they generate Čerenkov radiation. This light can be detected, and these particles can be vetoed. Since WIMPs have no charge and do not generate Čerenkov light, they should pass through the WCD undetected. In addition, the WCD serves a second purpose as a passive shield against neutron and gamma radioactivity from the surrounding rocks.

The Liquid Scintillator Veto (LSV, also called the “Neutron Detector” or “ND” in some publications) is a 4 m diameter stainless steel sphere, filled with 30 tonnes

---

<sup>3</sup> [Agnes et al. \(2015\)](#)

of borated organic liquid scintillator. The purpose of the LSV is to tag cosmogenic neutrons created by cosmic ray spallation and radiogenic neutrons from the surrounding rock and the detector materials themselves. These particles can then be vetoed, though the LSV contains an independent event trigger, and so the LSV can also perform in-situ measurements of the neutron background. The LSV is lined with Lumirror reflecting foils and is outfitted with 110 Hamamatsu R5912 8" high-QE (37% at 408 nm) PMTs with low-radioactivity glass bulbs. The LSV was originally filled with a 1:1 mixture of the organic scintillator pseudocumene (PC) and trimethyl borate (TMB) (50% TMB),<sup>4</sup> though the ratio is now 19:1 (5% TMB). The PC is “boron-loaded” in this way to reduce the neutron thermalization and capture time. TMB ( $\text{B}(\text{OCH}_3)_3$ ) contains natural boron  $^{\text{nat}}\text{B}$ , with 20%  $^{10}\text{B}$  abundance;  $^{10}\text{B}$  has a large thermal neutron capture cross section (3840 barns). Without TMB, the neutron capture time is  $\sim 250 \mu\text{s}$ . With 50% (5%) TMB, the capture time is reduced by orders of magnitude to  $\sim 2.3 \mu\text{s}$  ( $\sim 20 \mu\text{s}$ ). See [Appendix D](#) for more information on the neutron capture time in the LSV. Neutron capture on  $^{10}\text{B}$  occurs through two channels:<sup>5</sup>



The LSV greatly reduces the cosmogenic neutron background in DS-50. [Wright et al. \(2011\)](#) note that, while “1 meter of scintillator veto is about as effective as 4 meters of water shielding in reducing backgrounds due to external cosmogenic neutrons”, borating that same volume of PC increases its rejection power. It’s “equivalent to increasing the depth of the experiment by more than 2 km.w.e.”.

---

<sup>4</sup> [D’Angelo \(2015\)](#)

<sup>5</sup> [Agnes et al. \(2016c\)](#)

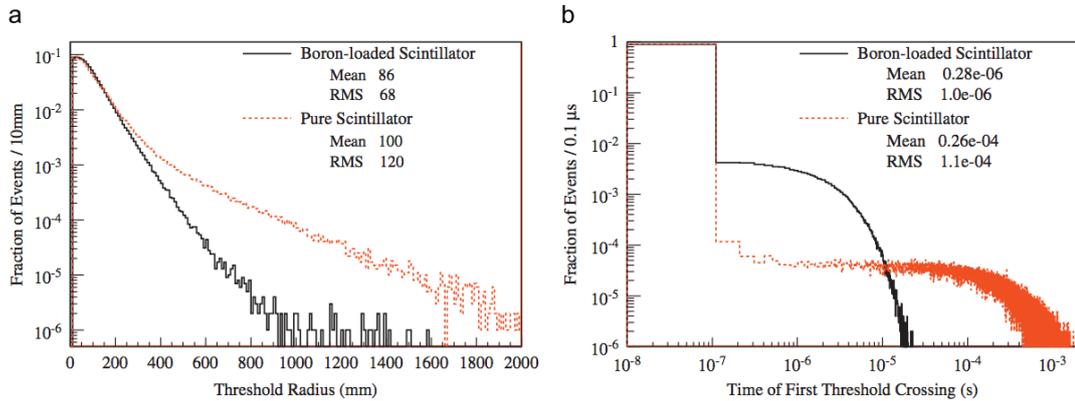


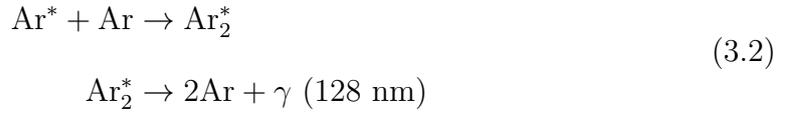
Figure 3.2: The effect of boron-loading the LSV scintillator: (a) radius in mm that a simulated neutron travels before being captured in pure scintillator (dotted red) and boron-loaded scintillator (solid black); (b) time in sec that a simulated neutron will travel before being captured. This simulation uses a 50%/50% by weight mixture of PC/TMB. From [Wright et al. \(2011\)](#).

The LSV also contains a small concentration of the wavelength shifter (WLS) 2,5-diphenyloxazole (PPO), typically at a concentration of a few g/L. (This concentration has been adjusted over time. See “CALIS” in Subsection 3.4.2.) The very short neutron capture time in the LSV, combined with its large volume, means that it is particularly effective at reducing this background. In the AAr exposure of DS-50, the vetoing efficiency of the LSV was estimated to be  $\sim 98.5 \pm 0.5\%$ . The TMB used for that campaign had an unexpectedly high concentration of the isotope  $^{14}\text{C}$ , which increased the dead time of the data acquisition (DAQ) system. For the UAr campaign, the LSV contained a 5% concentration by mass of a lower-activity TMB mixture ( $\sim 0.3$  kBq, compared to the  $\sim 150$  kBq AAr TMB), and the veto efficiency improved to  $> 99.1\%$ . The main inefficiency in LSV neutron detection is neutron capture on  $^1\text{H}$ , which generates a 2.2 MeV de-excitation gamma ray which may be absorbed in inactive materials rather than in the LSV.<sup>6</sup>

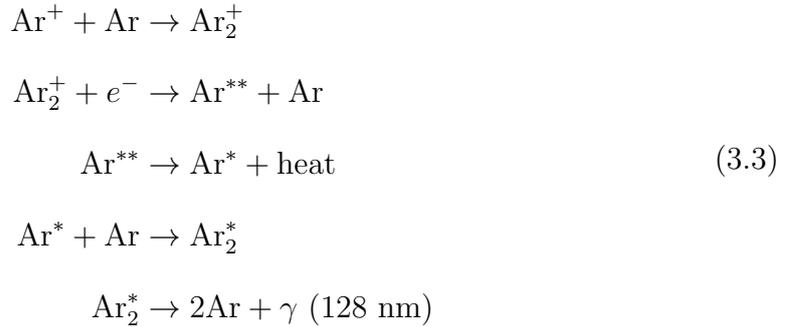
<sup>6</sup> [Agnes et al. \(2016c\)](#)

## 3.2 Time-Projection Chamber

When energy is deposited in a volume of liquid argon ( $\ell\text{Ar}$ ) — like when a particle collides with an argon nucleus — the  $\ell\text{Ar}$  typically re-emits that energy in two channels: scintillation and ionization. Excited argon atoms produce scintillation by forming a weakly-bound dimer (or a “self-trapped exciton”) and de-exciting:<sup>7</sup>



Ionized argon atoms can recombine and go through “non-radiative de-excitation” by transferring some of their energy to the surrounding atoms via thermal motion:



Note that the last two steps above are identical to the de-excitation channel for an excited argon nucleus. Argon ions and free electrons which recombine generate scintillation. The scintillation-to-ionization ratio is strongly influenced by the ionization density of the track, and by the presence of an external electric field,<sup>8</sup> and in general, the greater the field strength, the smaller this ratio. This is because if an external electric field is applied to the  $\ell\text{Ar}$  volume, free electrons will drift toward one end of the volume, reducing the recombination probability. There,

---

<sup>7</sup> Suzuki and Kubota (1979)

<sup>8</sup> Grandi (2005)

they can be detected directly via wire arrays or other apparatus (as is common in large  $\ell\text{Ar}$  neutrino detectors like MicroBooNE<sup>9</sup>, DUNE<sup>10</sup>, or SBND<sup>11</sup>). If the 128 nm scintillation light can be detected, as well as the ionization (also “drift”) electrons, which typically travel  $\mathcal{O}(\text{mm}/\mu\text{s})$ , then the time between the two signals gives the “depth” of the event in the direction along the electric field — hence “time projection chamber”.

The DarkSide-50 TPC is mounted inside a stainless steel cryostat and consists of a cylindrical active volume of  $\ell\text{Ar}$ ,  $\sim 35.5$  cm in diameter and  $\sim 34.5$  cm tall, with its axis perpendicular to the ground (so  $+z$  is “up”). The active volume is held at the boiling point of argon (89.1 K at the TPC pressure of 15.6 psi), and the argon is constantly recirculated through a gas purification loop at a rate of  $\sim 28.2$  slpm.<sup>12</sup> Directly above the  $\ell\text{Ar}$  region is a gaseous argon ( $g\text{Ar}$ ) “gas gap” with the same cross-sectional area as the liquid region, but a height of only  $\sim 1$  cm. There is a bit of uncertainty in these dimensions due to the shrinking of the detector materials at cryogenic temperatures, giving an active mass of  $46.4 \pm 0.7$  kg.

The active argon is bounded by a 1” thick PTFE (Teflon) cylinder around the circumference, “fabricated with a modified annealing cycle to increase its reflectivity”,<sup>13</sup> and by disk-shaped fused silica “windows” at the top and bottom. At the top of the active region, the fused silica window comprises most of the “diving bell”, which has a slightly larger radius than the rest of the active volume, with a “lip” around the circumference. This particular shape holds the gas gap in place, despite being surrounded on all sides by  $\ell\text{Ar}$ .

The fused silica windows are a major component of the HHV (High High Volt-

---

<sup>9</sup> <http://www.microboone.fnal.gov>

<sup>10</sup> <http://www.dunescience.org>

<sup>11</sup> <http://sbn-nd.fnal.gov>

<sup>12</sup> slpm = standard liters per minute

<sup>13</sup> Agnes et al. (2015)

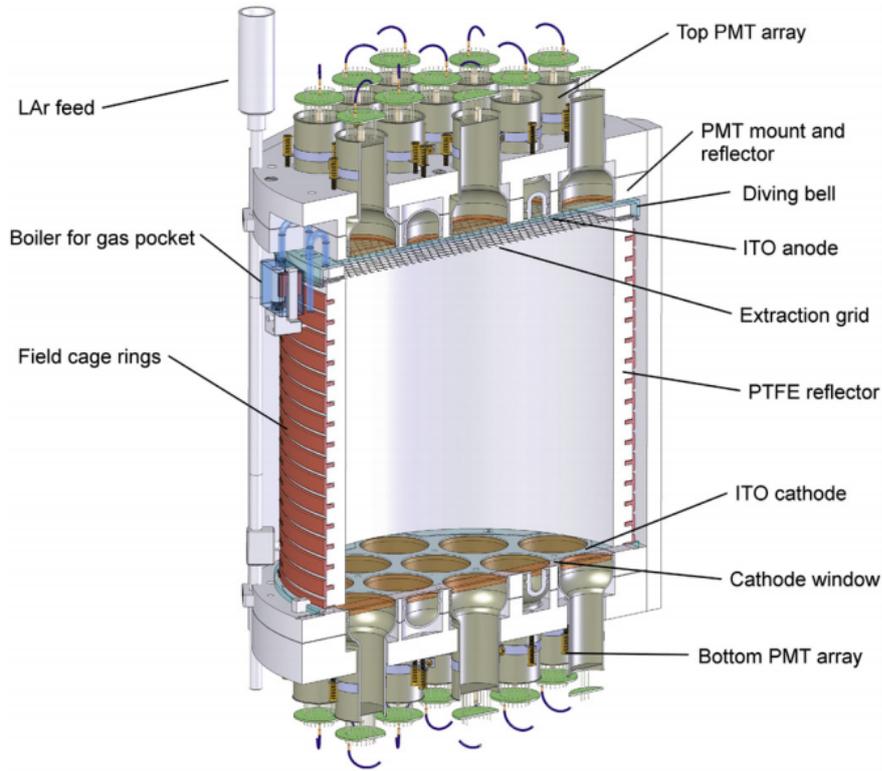


Figure 3.3: A 3D model of the DarkSide-50 TPC. From [Agnes et al. \(2015\)](#).

age) system, which maintains the electric fields in the active volume.<sup>14</sup> The windows are coated with indium tin oxide (ITO), a transparent conducting material. ITO has a complex index of refraction, so it absorbs some of the light incident on it, so DS-50 uses the thinnest coatings of ITO achievable by our vendor, at 15 nm. At this thickness, and at a wavelength of 420 nm, ITO absorbs only 2% of normally-incident light. In DS-50, the top ITO-coated fused silica window is grounded, while the bottom ITO-coated window is typically held at  $-12.7$  kV. Accordingly, the top window of the TPC is often called the “anode”, while the bottom window is called the “cathode”. In addition, a  $50\ \mu\text{m}$ -thick stainless steel hexagonal “grid” sits 5 mm below the  $g\text{Ar}/\ell\text{Ar}$  boundary. The grid is 95% transparent at normal incidence, with 2 mm pitch,<sup>15</sup> and is typically held at a potential

<sup>14</sup> [Love \(2013\)](#)

<sup>15</sup> From internal docdb entry # 564-v6.

of  $-5.6$  kV. Circular copper “field cage rings” surround the outside of the PTFE cylinder and are held at graded potentials to maintain the uniformity of the fields within the active volume.

These three voltages define three fields: the  $\mathcal{E}_d = 200$  V/cm “drift field” between the cathode and the grid, the  $\mathcal{E}_{ext} = 2.8$  kV/cm “extraction field” between the grid and the  $g\text{Ar}/\ell\text{Ar}$  boundary, and the  $\mathcal{E}_{elec} = 4.2$  kV/cm “electroluminescence field” (sometimes called the “multiplication field”) between the  $g\text{Ar}/\ell\text{Ar}$  boundary and the anode. Note that, even though the extraction field and the electroluminescence field are both between the grid and the anode, they have different field strengths due to the different dielectric constants of liquid and gaseous argon.

The drift field, as its name suggests, drifts the ionization electrons through the  $\ell\text{Ar}$  and toward the anode. The extraction field accelerates the drifting electrons out of the  $\ell\text{Ar}$  and into the  $g\text{Ar}$ , where they’re pulled toward the anode by the electroluminescence field and excite the gaseous argon atoms. The initial particle interaction in the  $\ell\text{Ar}$  generates a first scintillation signal, which we call S1. As the drift electrons are accelerated through the  $g\text{Ar}$ , the argon atoms they excite generate a second scintillation signal, S2, as they de-excite. The time between the S1 and S2 signals gives the drift time,  $t_{\text{drift}}$ , of the event, typically measured in  $\mu\text{s}$ . Since we know the drift speed of the ionization electrons in the  $\ell\text{Ar}$ , we can calculate the depth ( $z$ -position) of the initial interaction in the  $\ell\text{Ar}$ . And the distribution of the S2 light over the anode PMTs contains information about the  $xy$  position of an event, which must be extracted through further analysis.

The hard UV scintillation light (128 nm) generated by argon cannot be directly detected with commercial PMTs, so argon-based TPCs typically use wavelength-shifters like tetraphenyl butadiene (TPB) to move this to a more reasonable 420 nm. There is an optimal thickness of TPB, which is thick enough to absorb

most UV photons and convert them to visible wavelengths, but not so thick that it becomes opaque to those same visible photons. That thickness is  $\sim 200 \mu\text{g}/\text{cm}^2$ , and while DS-50 has come pretty close to this value, there are some variations over the surface area of the TPC. The entire inside surface of the active volume is coated in TPB — the cylindrical PTFE wall and both the cathode and anode windows. The centers of the windows have a TPB thickness of  $\sim 230 \pm 10 \mu\text{g}/\text{cm}^2$ , while near the edges, it’s closer to  $190 \pm 15 \mu\text{g}/\text{cm}^2$ . At the top and bottom of the PTFE cylinder, the thickness is  $224 \pm 27 \mu\text{g}/\text{cm}^2$ , while at half-height, it’s  $165 \pm 20 \mu\text{g}/\text{cm}^2$ . The TPB fluorescence decay time is  $\sim 1.8 \text{ ns}$ ,<sup>16</sup> much shorter than the fast component of the S1 light (see Section 3.5).

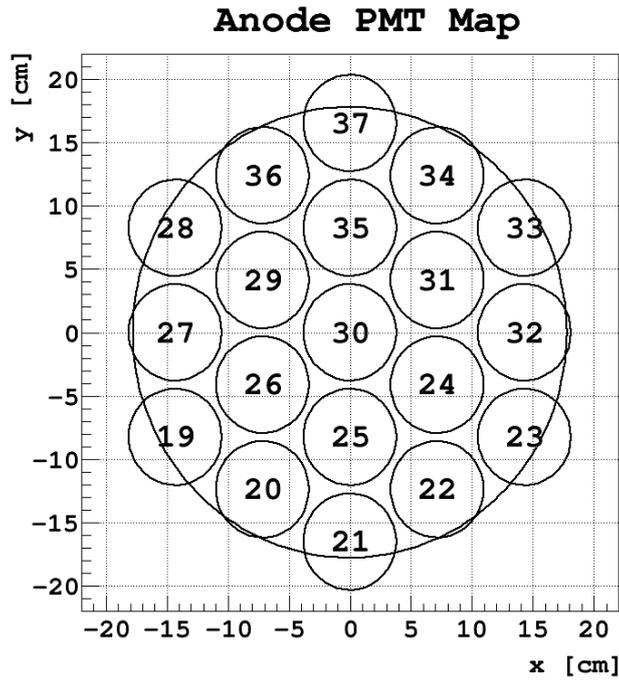
Once the scintillation light is wavelength-shifted by the TPB, it can be detected by the PMT arrays found behind the cathode and anode windows. Nineteen PMTs are arranged in a close-packed hexagonal pattern at the top and bottom of the active volume, for a total of thirty-eight TPC PMTs (see Figure 3.4). They are all 3” Hamamatsu R11065 low-radioactivity, high-QE ( $\sim 34\%$  QE at 420 nm) PMTs, though the QE varies slightly from PMT to PMT. A  $\sim 1 \text{ mm}$  thick layer of  $\ell\text{Ar}$  optically couples all of the PMTs to the fused silica windows. The cathode PMTs are sometimes referred to collectively as the “bottom” PMTs and the anode PMTs are referred to as the “top” PMTs; for instance, the sum of all of the S1 light seen by the cathode PMTs is called  $\text{S1}_{\text{bot}}$ .

### 3.2.1 Why $\ell\text{Ar}$ ?

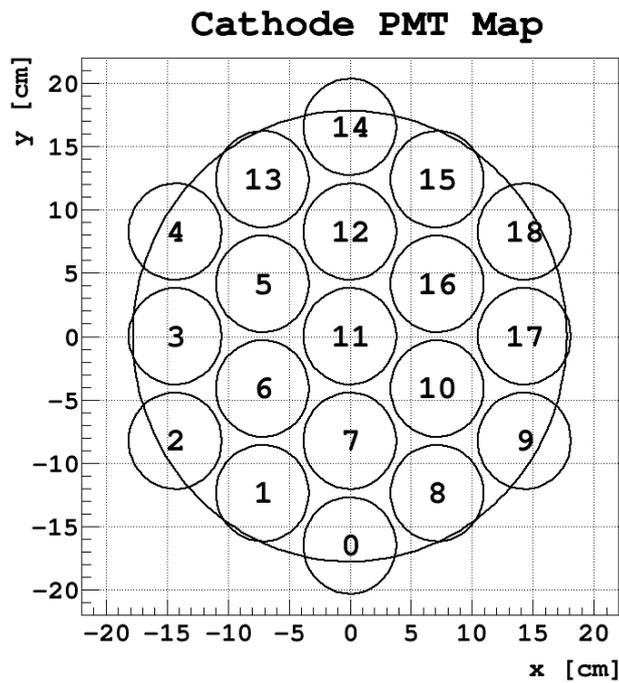
Liquid noble elements (primarily Xe and Ar, though sometimes Ne also) are the target of choice for large-scale direct dark matter detection experiments for several reasons. First, they have low boiling points ( $\sim 25 \rightarrow 165 \text{ K}$ ), but do not need to be operated at extreme cryogenic temperatures ( $\mathcal{O}(\text{mK})$ ) like crystal-based

---

<sup>16</sup> [Alexander et al. \(2013a\)](#)



(a) DS-50 anode PMT map.



(b) DS-50 cathode PMT map.

Figure 3.4: DS-50 TPC PMT maps with channel indices. Note that both of these maps show the PMT arrangement from a “top-down” view.

dark matter experiments. Xe and Ar, in particular, can be cooled to liquefaction with nothing more than a supply of liquid nitrogen. Second, Xe and Ar are both bright scintillators (46 and 41 photons/keV<sub>ee</sub><sup>17</sup> of energy deposited, respectively) and strong ionizers (64 and 42 electron-ion pairs/keV<sub>ee</sub>), and they're transparent to their own scintillation light. Third, relative to crystal-target detectors, liquid nobles are cheap and can be scaled up easily: multi-ton experiments are already being built and several more are being planned.<sup>18</sup>

In terms of cost, though, there is a clear winner among the liquid nobles: argon. Xe can be 100× as expensive as Ar, mainly due to its rarity (Ar is  $\sim 10^6\times$  as abundant as Xe in the atmosphere).<sup>19</sup> Additionally, particle identification is orders of magnitude more effective in Ar, as will be seen in Section 3.5. There are two major drawbacks to using atmospheric Ar, though: event rate and inherent background.

For spin-independent weak couplings, the WIMP-nucleus elastic scattering cross section scales as the square of the neutron number,  $(A - Z)^2$ .<sup>20</sup> Since the stable isotopes of Xe have around  $A - Z = 75$  neutrons while <sup>40</sup>Ar (the most common – 99.6% – isotope of argon) has only  $A - Z = 18$  neutrons, Xe has an event rate about one to two orders of magnitude larger than Ar at low recoil energies. (Though a form factor suppresses the interaction rate at higher recoil energies in Xe, due to a loss of coherence from the heavy nucleus.) Because of this, Xe experiments will always be able to probe smaller interaction cross-sections than

---

<sup>17</sup> keV<sub>ee</sub> stands for “keV electron equivalent”. As will be seen in Section 3.8, an electron recoil and a nuclear recoil of the same recoil energy will generate S1s of different sizes in  $\ell$ Ar, due to the quenching factor for nuclear recoils. This is explored in detail in Subsection 3.8.1. Similarly, keV<sub>nr</sub> stands for “keV nuclear recoil”.

<sup>18</sup> Schumann (2012)

<sup>19</sup> Editors of the *Encyclopædia Britannica* (1998)

<sup>20</sup> Savage et al. (2006). See also Jungman et al. (1996) and Lewin and Smith (1996).

Ar experiments of similar active mass and exposure time.

Additionally, AAr contains a naturally-occurring cosmogenic isotope,  $^{39}\text{Ar}$ , which is a  $\beta$ -emitter with an endpoint energy of 565 keV and a half life of  $\tau_{1/2} = 269$  years, encompassing all of the dark matter nuclear recoil (“WIMP-nucleon” recoil) energy range. The activity of  $^{39}\text{Ar}$  in liquefied AAr is  $\sim 1$  Bq/kg, making it a significant internal background. Xe has no such inherent background, though like Ar, atmospheric Xe contains  $^{85}\text{Kr}$ , but this can be removed via distillation.

With all of the drawbacks for Ar, it seems like Xe is the natural choice for WIMP detection experiments, but the DarkSide Collaboration has chosen to use  $\ell\text{Ar}$ . Why?

### 3.2.2 AAr vs. UAr

Since 2010, DarkSide has operated an Ar extraction facility at the Kinder Morgan Doe Canyon  $\text{CO}_2$  source in Cortez, Colorado.<sup>21</sup> Since  $^{39}\text{Ar}$  is a cosmogenic isotope, any underground source of argon (UAr) is naturally depleted, to some extent, in  $^{39}\text{Ar}$ . The UAr extracted from Doe Canyon has been shown to have an  $^{39}\text{Ar}$  concentration reduced by a factor of  $1.4 \pm 0.2 \times 10^3$ , relative to AAr (see Figure 3.5).<sup>22</sup> While the activity of  $^{39}\text{Ar}$  within AAr is  $\sim 1$  Bq/kg, in UAr, it’s  $\lesssim 6.5$  mBq/kg.<sup>23</sup> In addition to significantly reducing the  $\beta$  background in DS-50, UAr puts less strain on the DAQ: the trigger rate in DS-50 with AAr is  $\sim 50$  Hz, while with UAr, it’s only  $\sim 1.6$  Hz. DarkSide currently utilizes UAr in DS-50 and plans on using it in the next-generation DarkSide-20k detector.

---

<sup>21</sup> Wright (2011)

<sup>22</sup> Agnes et al. (2016c)

<sup>23</sup> D’Angelo (2015)

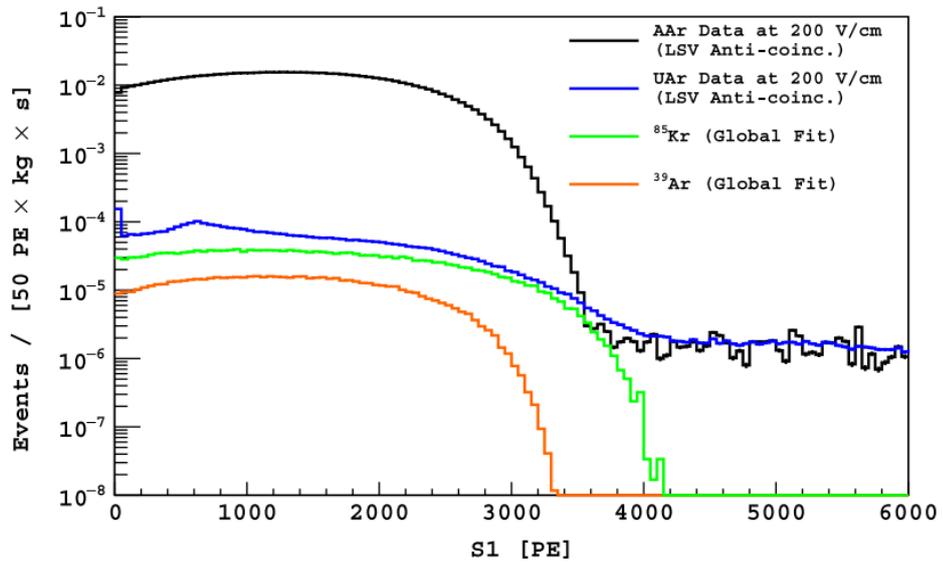


Figure 3.5: S1 spectrum comparison between AAr and UAr data, normalized to live time. The black line shows the AAr S1 spectrum, the blue line shows the UAr S1 spectrum, and the green and orange lines show the  $^{85}\text{Kr}$  and  $^{39}\text{Ar}$  components of the UAr spectrum, found from a MC fit. The peak in the lowest bin of the UAr spectrum is  $^{37}\text{Ar}$  and the peak around 600 PE is from  $\gamma$  Compton backscatters. The  $^{39}\text{Ar}$  concentration has clearly been reduced by orders of magnitude in the UAr data, relative to AAr. From [Agnes et al. \(2016c\)](#).

### 3.2.3 S1, S2, and S3

S1 and S2 are described in detail at the beginning of this section, but an additional scintillation signal, called S3 (also “S2 echo”) is sometimes seen as well. A third pulse has been observed in some events, where the time between S2 and this third pulse is always approximately equal to the maximum drift time in the detector. S3 also has approximately the same pulse shape as S2, though its integral is typically  $\sim 1000\times$  smaller.

While S1 is the scintillation caused in the  $\ell\text{Ar}$  by the initial event interaction, and S2 is the later scintillation via electroluminescence in the  $g\text{Ar}$ , it’s thought that S3 is an “echo” of S2. S2 is very bright: usually  $\sim 20\times$  as many PE are generated in the PMTs by S2 compared to S1. When this light strikes the cathode, it can free electrons which then drift to the anode, releasing a second S2-like pulse. In later sections, the analysis cuts will be discussed, which usually require “single-site” interactions where only S1 and S2 are present. Since S3 is thought to be caused by the same event which generated the preceding S1 and S2, events with S3 are also allowed to pass the “single-site” cut.

## 3.3 g4ds: The DarkSide Monte Carlo

To better understand detector backgrounds and to help guide the calibration campaigns and construction of future detectors, the Geant4-based DarkSide Monte Carlo simulation, g4ds (sometimes just called “the MC”), has been written from scratch by the DarkSide simulations group, headed by Davide Franco, at the *Laboratoire Astroparticule et Cosmologie* in Paris. Within g4ds, the geometries of all components of DS-50 are implemented, including the WCD, LSV, and TPC. Additional geometries have also been written to guide the development of DS-20k, and to analyze related experiments like ARIS (see [Appendix C](#)). An electronics

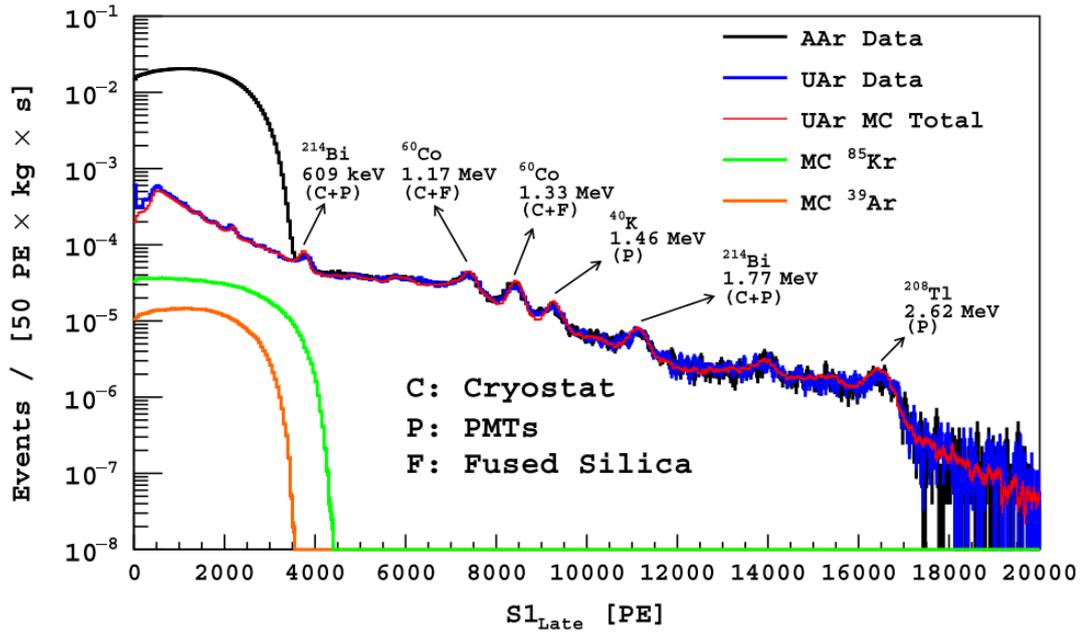


Figure 3.6:  $S1_{\text{late}}$  spectra from AAr (black) and UAr (blue) DS-50 data.  $S1_{\text{late}}$  is the integral of the  $S1$  pulse after the first 90 ns: it improves energy resolution by avoiding saturation in the early part of the  $S1$  pulse. The red line shows a fit to the data using `g4ds`, with the detector elements labeled which generate each peak in the spectra. The green and orange lines show the  $^{85}\text{Kr}$  and  $^{39}\text{Ar}$  components of the UAr spectrum, respectively. From [Agnes et al. \(2016c\)](#).

simulation add-on has also been written, which simulates the noise and energy resolution of the DAQ apparatus, as well as the biases and uncertainties of the low-level data reconstruction software, like the baseline-finder and pulse-finder algorithms (see the next Section).

`g4ds` performs a full ray-tracing optics simulation for all events, calculates recombination probabilities, and simulates the interactions of particles with all elements of the detector. The MC has been tuned on calibration data gathered with the help of the CALIS source insertion system (see the end of Subsection 3.4.2) and other experiments, and has been able to accurately reproduce many different aspects of real data gathered by DS-50. `g4ds` was used to study the light yield of the WCD and LSV vetoes and the  $^{39}\text{Ar}$  depletion of UAr, and is

used in this thesis to train the PCAMDF  $xy$ -Reconstruction.<sup>24</sup> A more thorough description of g4ds is given in [Agnes \(2016a\)](#).

The accuracy of the optics simulation is the most critical factor for the PCAMDF  $xy$ -Reconstruction outlined in this thesis. Detailed comparisons between MC and data will be explored in [Chapter 5](#).

### 3.4 Data Acquisition and Event Reconstruction

The hardware behind the DarkSide Data Acquisition (DAQ) system is described in detail in [DarkSide Collaboration \(2015\)](#) and [Agnes et al. \(2016b\)](#), though a few specific points are covered here. Of critical importance to DAQ in DS-50 are the cryogenic preamplifiers attached to each of the TPC PMTs. While they were manufactured to be high-QE and low-radioactivity, the Hamamatsu PMTs used by DS-50 have had sporadic, spontaneous light emission issues. Running these PMTs at lower gain ( $\sim 4 \times 10^5$ ) has dramatically reduced the frequency of these outbursts, but this would not be possible without the cryogenic preamplifiers attached to the PMT voltage dividers.<sup>25</sup>

Additionally, DarkSide makes use of two separate digitization systems. The TPC PMT signals were originally run through CAEN V1720 digitizers with a  $10\times$  amplification. Because this sometimes causes electronic saturation, these signals are now split and simultaneously sent through both the original CAEN V1720 digitizers with  $10\times$  amplification and new CAEN V1724 digitizers with no amplification. The V1724s can then be used to correct saturated signals received from the V1720s, as can be seen in [Figure 3.7](#).

The software for the DS-50 DAQ is partially written in LabVIEW and is flexible enough that it can be repurposed for use on the next-generation detector, DS-20k.

---

<sup>24</sup> [Agnes \(2016b\)](#)

<sup>25</sup> [Agnes et al. \(2015\)](#)

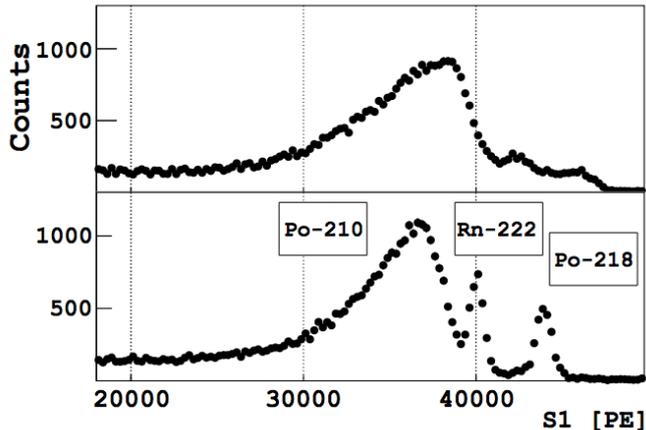


Figure 3.7: S1 spectrum of  $\alpha$  particles before (top) and after (bottom) incorporating data from the low-gain CAEN V1724 digitizers to correct for PMT saturation. From [Agnes et al. \(2017\)](#).

The user interface (also “The Run Controller”) is available through `ssh` and has a very shallow learning curve, so even new DarkSide collaborators find it relatively easy to control. The Run Controller allows users to adjust details like the lengths of acquisition windows and the PMT coincidence required to trigger an event, as well as more general run parameters like the number of events acquired or the length of a given run. During routine data taking, the TPC acquisition window is usually  $440 \mu\text{s}$  (maximum drift time is  $\sim 375 \mu\text{s}$ ) with an  $810 \mu\text{s}$  inhibit after each trigger, while the veto acquisition windows are usually  $< 200 \mu\text{s}$ , long enough to catch neutron thermalization signals (see [Appendix D](#)).

### 3.4.1 Event Reconstruction

The TPC event reconstruction code in DS-50, DarkArt, is based on the Fermi National Accelerator Laboratory’s *art* framework. An extremely thorough description of this software can be found in the Ph.D. thesis of one of its primary authors,<sup>26</sup> though a cursory description here will be sufficient. There are three

<sup>26</sup> [Fan \(2016\)](#)

main steps in the low-level event reconstruction code: (1) find the baseline of the waveform, (2) find any pulse(s) present in the waveform, (3) calculate a few basic parameters. After DarkArt event reconstruction, more variables are calculated and the data is saved in a user-friendly ROOT file:<sup>27</sup> this is the high-level reconstruction, SLAD.

## DarkArt

Raw data gathered from DS-50 is entirely *waveform* data: digitizer counts per time bin. Each TPC PMT (as well as each of the  $\sim 200$  veto PMTs) has a waveform for each event. The primary goal of DarkArt, which is written entirely in C++ and is available freely online,<sup>28</sup> is to take those waveforms and calculate a few critical parameters: S1, S2 (and S3, if present),  $t_{\text{drift}}$ , and  $f_{90}$ , as well as calculating the scintillation detected in the vetoes.  $f_{90}$  is the fraction of the S1 light within the first 90 ns of the pulse and is the critical parameter for Pulse Shape Discrimination (PSD) of events (see section 3.5).

The first task in event reconstruction is to calculate a baseline for a given waveform. Since each of the PMTs can have independent noise fluctuations, each PMT waveform has its baseline calculated independently. Additionally, since coherent noise across all PMTs is *also* a possibility, a simple interpolation from the beginning to the end of the waveform is not used, as this could result in biases due to accumulated noise when waveforms are summed together, as is the case with the TPC PMTs. The most appropriate baseline-finding algorithm for DS-50 is a moving average. The DarkArt baseline-finder categorizes each sample (with 4 ns sampling and 12-bit resolution in the TPC, 0.8 ns sampling and 10-bit resolution in the vetoes) as either “baseline” or “signal” based on its amplitude and

---

<sup>27</sup> <https://root.cern.ch/>

<sup>28</sup> <https://cdcv.s.fnal.gov/redmine/projects/darkart/repository>

relationship to the neighboring samples. If a sample is within a baseline region, the average of that sample and its 10 nearest neighbors forward and backward in time (for a total of 80 ns of waveform) is calculated and that value is assigned to the baseline at that sample time. Across signal regions of the waveform, the baseline is interpolated between the two baseline regions on either side of the signal. Example waveforms showing the performance of the baseline-finder across signal and baseline regions can be seen in Figure 3.8.

After the baseline is found and subtracted from the waveform, the next step is to find any pulses which may be present in the waveform. Since DS-50 typically requires coincident single PE signals across just a few PMTs in order to trigger the recording of an event, it's common for a given PMT to show no evidence of an event, ie. no obvious pulses. So while the baseline-finder works on individual channels, the pulse-finder works on the *sum channel* for the TPC PMTs. The individual PMT waveforms are zero-suppressed to reduce coherent noise in the sum channel waveform, a small time offset may be added to individual waveforms to properly align them, and then all channels (all PMTs) are scaled by a conversion factor which relates ADC counts (“Analog-to-Digital” counts, or digitizer counts) to PE (see Subsection 3.4.2).<sup>29</sup> The pulse-finder uses a “rough window/fine window” process, where a certain threshold of PE are required within a large time window to register a pulse. The large window is then broken up into smaller time windows to find the pulse start. The rough window looks for  $\geq 5$  PE within  $2 \mu\text{s}$  and the fine window looks for  $\geq 0.3$  PE within a single sample. Since S1 and S2 both have exponential tails, finding the end time of the pulse is not critical; any parameter which requires a pulse integration uses a fixed-width integral from the pulse start.

---

<sup>29</sup> Note that zero-suppression is used only to find the start and end times of the pulses, and is not applied to the waveforms when, for instance, calculating integrals of pulses.

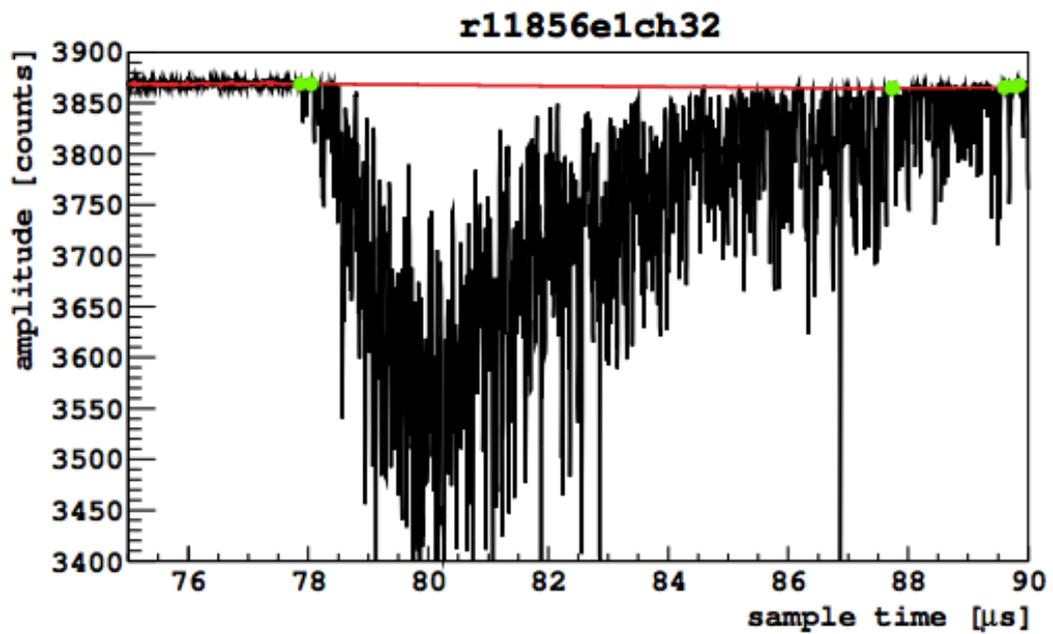
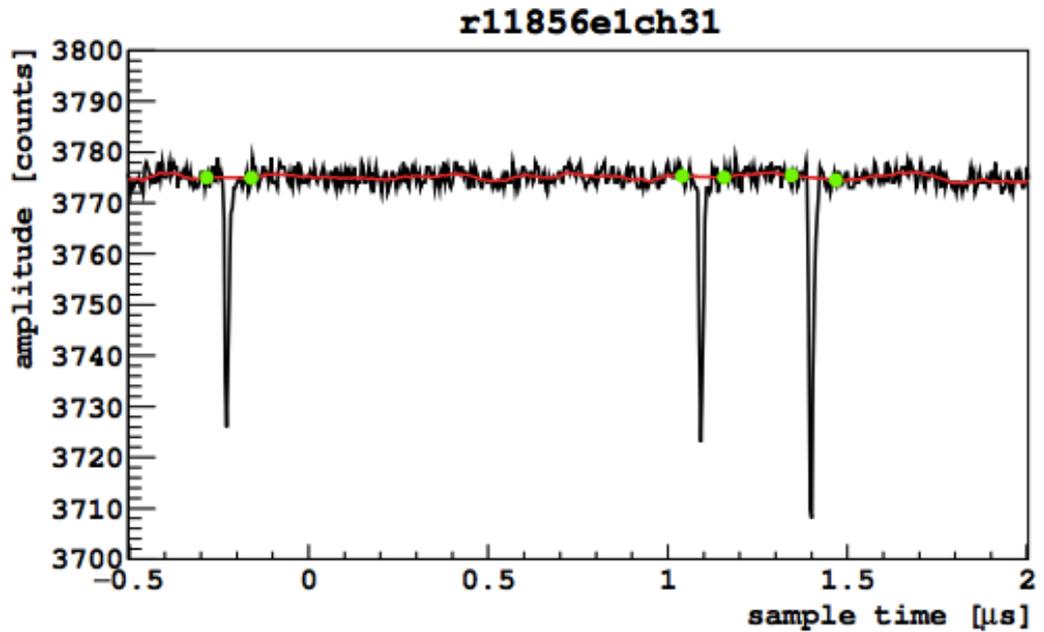


Figure 3.8: Example of the DarkArt baseline-finder working in regions of primarily baseline (top) and primarily signal (bottom). The black lines show event waveforms in ADC counts vs. sample time in  $\mu\text{s}$ , the red lines show the baseline, and the green dots show the endpoints of signal regions. From [Fan \(2016\)](#).

There are known inefficiencies and biases in both the baseline-finder and the pulse-finder. For instance, the linear interpolation the baseline finder performs in signal regions is only a crude approximation, especially where pulses are low-energy and the signal-to-noise ratio is small. The pulse-finder can have difficulty with *pile-up*, where two pulses fall very near each other. For this reason, the pulse-finder is not reliable for events with  $t_{\text{drift}} \lesssim 7 \mu\text{s}$ , where the tail of the S1 pulse can overlap with the S2 pulse. Both algorithms require some degree of hand-tuning, which can introduce biases. All of these different factors have been investigated in detail with g4ds, though, and introduce  $\lesssim 0.5\%$  error in the total integral for S1 and S2.

## SLAD

Once the channel baselines have been found and subtracted from the channel waveforms; the baseline-subtracted waveforms have been zero-suppressed, time-shifted as necessary, and scaled to PE; the sum channel has been constructed from the TPC channels; and the pulse start(s) have been found; then all of the relevant physics variables can be calculated. Though some of these variables are calculated within DarkArt and simply passed along to the high-level analysis software, many of them are calculated within that high-level software itself. SLim Analysis Data files (SLAD files) are ROOT files generated by the high-level software, and are the ones most commonly accessed by DarkSide analyzers. The most basic event variables are the ones which help us identify the type of particle which generated the given event and its position within the detector: S1, S2,  $t_{\text{drift}}$ , and  $f_{90}$ .

S1-like pulses have a shorter exponential tail than S2-like pulses, and so S1 is integrated from pulse start out to  $7 \mu\text{s}$ , while S2 is integrated from pulse start out to  $30 \mu\text{s}$ .  $t_{\text{drift}}$  is calculated as the time delay between the S1 start and the S2 start, and  $f_{90}$  is calculated as described above, by integrating the first 90 ns of

S1 and dividing by the total  $7 \mu\text{s}$  integral. Note that these integrals are actually performed at the channel level and then summed in a further attempt to reduce coherent noise, though the pulse start and end times are dictated by the times found by the pulse-finder, which works on the sum channel. Additionally, the individual channel integrals are saved in arrays, `S1[]` and `S2[]`, each of length 38 for the TPC. `S2[]` is particularly important for  $xy$ -position reconstruction, as will be seen in the next chapter.

Many other variables are calculated by SLAD and stored in its output files, but the ones relevant to this thesis will be introduced as necessary. A thorough description of all of the high-level variables can be found on the publicly-accessible DarkArt redmine page.<sup>30</sup>

### 3.4.2 Calibrations and Corrections

In order to convert the waveforms from ADC counts to PE, as mentioned in the previous Subsection, we need an ADC-to-PE conversion factor. The most common way to find this factor experimentally is to determine what the PMT response is to a single PE (SPE). This must be done for each PMT (as each PMT may have a slightly different gain, etc.) and it must be done regularly, as PMT response can and often does change over time.

After SPE calibration, and after the data is processed as described in the previous Section, a few low-level corrections must be made before the data can be used by analyzers. These are corrections to the S1 and S2 as a function of event depth (or  $z$ , or  $t_{\text{drift}}$ ).

---

<sup>30</sup> [https://cdcvns.fnal.gov/redmine/projects/darkart/wiki/SLAD\\_variables](https://cdcvns.fnal.gov/redmine/projects/darkart/wiki/SLAD_variables)

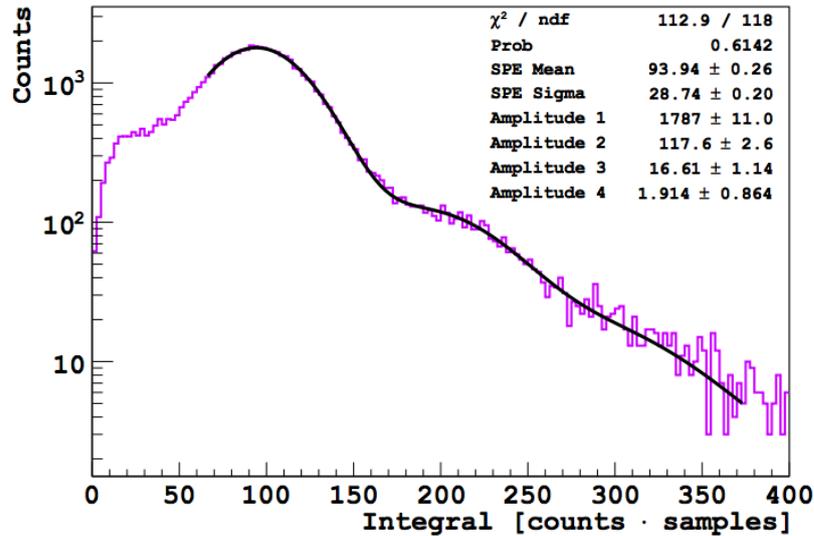


Figure 3.9: Example SER fit for a SCENE PMT. From [Cao \(2014\)](#).

### Single Photoelectron Calibration

During the SCENE Experiment (which will be discussed in the last Section of this Chapter), a “Single Electron Response” (SER) calibration was performed by looking for single PEs in the long exponential tails of S1 and S2. Integrating the pulse tails of the baseline-subtracted waveform over small, consistent intervals of time will yield a spectrum of digitizer counts, with a large peak surrounding the origin where only noise was integrated (the “pedestal”), then a large peak to the right of that, where a single PE was integrated, a smaller one further to the right where two PEs were integrated and so on. There will be noise and the PE peaks will overlap a bit, but they are clearly visible in the spectrum. See Figure 3.9 for an example SER fit where the pedestal has been suppressed, but 1PE, 2PE, and 3PE peaks can be seen.

A similar calibration, called an SPE calibration, is performed in DS-50. Instead of searching for PEs in the tails of pulses, a 405 nm blue LED injects a small amount of light into the TPC, while simultaneous triggers are sent to the DAQ, forcing it to record whatever signal the PMTs happen to see at the time. The

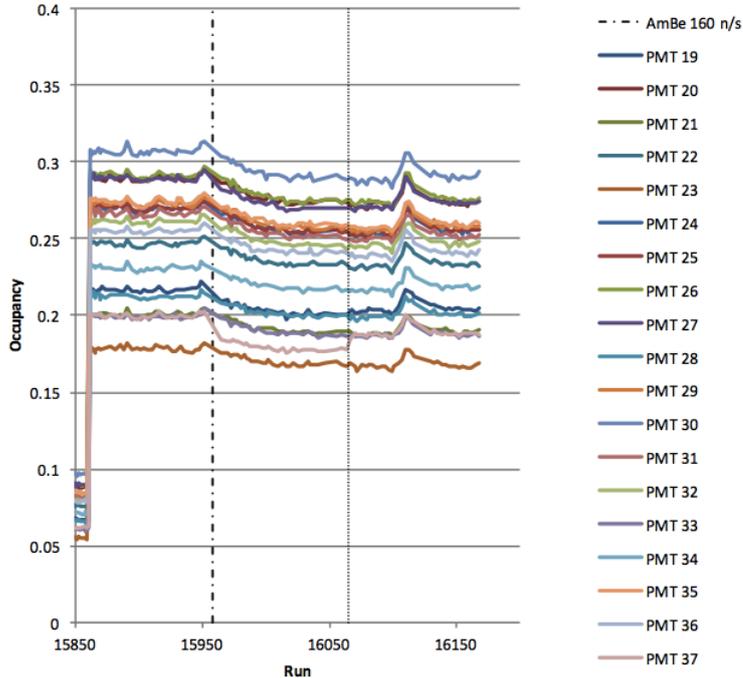


Figure 3.10: SPE vs. time (run number) during a CALIS  $^{241}\text{Am}^9\text{Be}$  calibration campaign. From internal docdb entry # 1560-v2.

laser signals themselves are  $\sim 60$  ps long and are sent at a rate of  $\sim 500$  Hz. (A similar procedure is performed for the veto PMTs, though in that case, each PMT has a dedicated optical fiber from a common external laser, which sits directly in front of the PMT window.) The spectrum is then fit with a pair of Gaussians — one for the pedestal and one for the single PE peak. This calibration is performed at least once every 24 hours and each raw data waveform is converted from ADC counts to PE according to the most recent SPE calibration run (also called “laser runs”). SPE calibrations allow us to perform this ADC-to-PE conversion, as well as track the SPE response of the PMTs over time (see Figure 3.10).

### S1 $z$ -Correction and Light Yield

As mentioned in Subsection 3.2.1,  $\ell\text{Ar}$  yields  $\sim 41$  photons and  $\sim 42$  electron-ion pairs per  $\text{keV}_{\text{ee}}$  of energy deposited. But the efficiency with which these photons

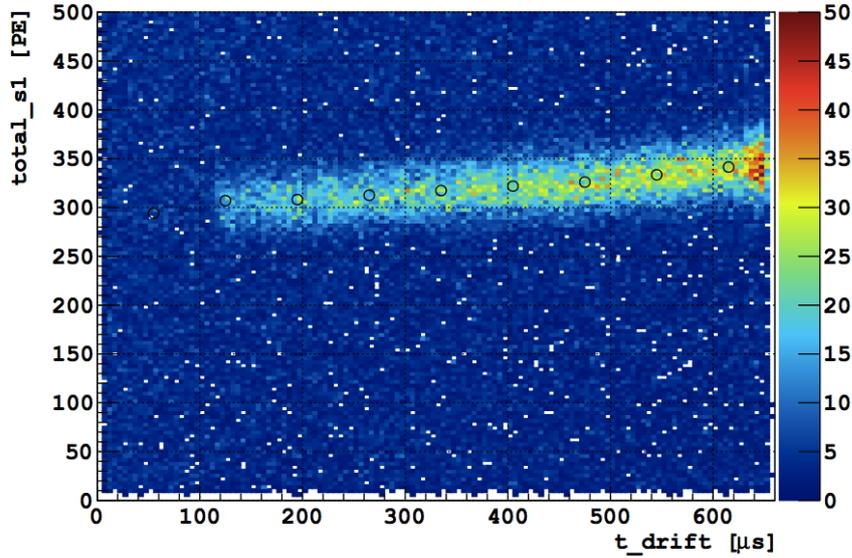


Figure 3.11: Un- $z$ -corrected S1 vs.  $t_{\text{drift}}$  for a  $^{83\text{m}}\text{Kr}$  calibration run with  $\mathcal{E}_d = 200$  V/cm. Open black circles are mean values of curve-fitted  $^{83\text{m}}\text{Kr}$  peaks. From internal docdb entry # 1611.

and electrons are detected is a function of multiple factors, principally among these are the geometry of the detector and the purity of the  $\ell\text{Ar}$  itself.

Since the DS-50 TPC has PMTs only at the top and bottom of the active volume, the integral of the S1 signal (the “size” of S1) is a function not just of energy, but also of  $t_{\text{drift}}$ . Events of a given recoil energy tend to have S1s which are  $\sim 19\%$  smaller (in integral) when they occur at the top of the  $\ell\text{Ar}$ , rather than at the bottom, due to partial internal reflection at the  $g\text{Ar}/\ell\text{Ar}$  boundary; this also causes the size of the S1 to be asymmetric about the half-height of the  $\ell\text{Ar}$ .<sup>31,32</sup> These two factors mean that S1 is often “ $z$ -corrected” and normalized throughout all  $t_{\text{drift}}$  to the value at the center of the active volume. (For the rest of this thesis, unless S1 is specifically noted to be un- $z$ -corrected — or if “S1” and “S1<sub>corr</sub>” are compared — “S1” refers to this  $z$ -corrected value.)

But how is the position dependence of S1 determined? To make the case that

<sup>31</sup> Agnes et al. (2017)

<sup>32</sup> Agnes et al. (2015)

the *light yield* ( $LY = \text{S1 generated [PE]}/\text{energy deposited [keV}_{ee}]$ ) is a function of  $t_{\text{drift}}$ , one must measure the number of S1 PE, but also *know* the amount of energy which was deposited in the  $\ell\text{Ar}$ . The easiest way to do this is with a radioactive source.  $^{22}\text{Na}$ ,  $^{57}\text{Co}$ , and  $^{137}\text{Cs}$  (among others) all have spectra with detectable peak energies (511 keV, 122 keV, and 662 keV, respectively). While it may be easy for smaller detectors to place and remove calibration sources, running a calibration source campaign in DS-50 is a time-consuming ordeal which requires several specially-trained operators. So these source runs are performed on occasion, but certainly not after every regular data run as calibrations.

A much easier way to accomplish this in DS-50 is with  $^{83\text{m}}\text{Kr}$ , a metastable isomer of Kr which decays to the ground state ( $\tau_{1/2} \approx 1.8$  hours) via two sequential transitions of 32.1 keV and 9.4 keV with an intermediate half-life of  $\tau_{1/2}^{\text{int.}} \approx 150$  ns.<sup>33</sup> These peaks are close enough together in time that they’re not easily resolvable in DS-50 and so we observe a single 41.5 keV peak.<sup>34</sup>  $^{83\text{m}}\text{Kr}$  is perfect for our purposes because it decays fast, has a sharp peak energy, and can be injected into the detector easily through the gas circulation system, quickly diffusing throughout the active volume. Making “slices” in  $t_{\text{drift}}$  and fitting the  $^{83\text{m}}\text{Kr}$  peak gives the number of S1 photoelectrons equivalent to 41.5 keV as a function of  $t_{\text{drift}}$ , in other words, the light yield as a function of  $z$  (see Figure 3.11).

A check on the light yield can be performed with another radioactive source distributed evenly throughout the active volume —  $^{39}\text{Ar}$ , which has an endpoint energy of 565 keV. By observing the  $^{39}\text{Ar}$  spectrum as a function of  $t_{\text{drift}}$ , a similar PE/energy calibration as with  $^{83\text{m}}\text{Kr}$  can be performed. Note that this endpoint energy is way beyond the region of interest for WIMP-nucleon interactions, and

---

<sup>33</sup> Zbořil et al. (2013)

<sup>34</sup> Agnes et al. (2015)

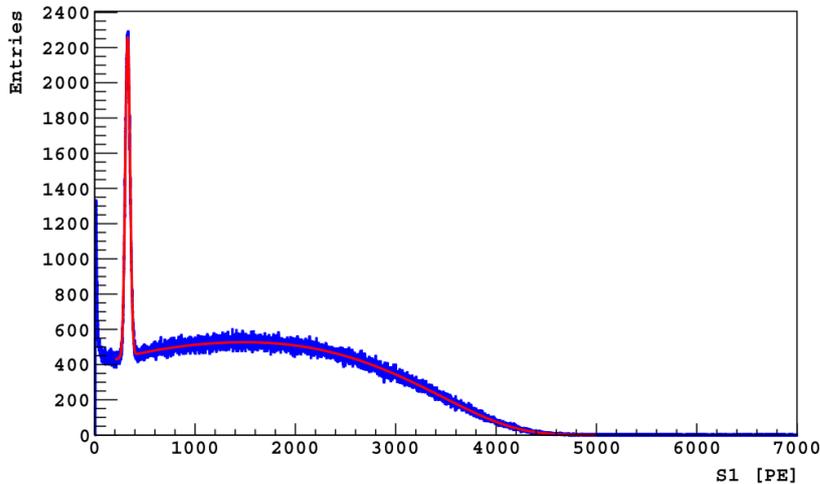


Figure 3.12: S1 spectrum of a krypton calibration run in AAr, where the DS-50 TPC was spiked with  $^{83\text{m}}\text{Kr}$ . In blue are the data and in red is the fit to the  $^{39}\text{Ar}$  spectrum, plus a narrow Gaussian for the 41.5 keV peak from  $^{83\text{m}}\text{Kr}$ . From [Agnes et al. \(2015\)](#).

so  $^{83\text{m}}\text{Kr}$  is generally the preferred calibration (though  $^{39}\text{Ar}$  is used for special investigations of high-energy particles like  $\alpha$ s and  $\mu$ s).

### Electron Drift Lifetime and S2 Position Corrections

The size of S1 can also be affected by impurities in the  $\ell\text{Ar}$ ; dissolved water, oxygen, and nitrogen can strongly affect its scintillation properties.<sup>35</sup> And the S2 light yield can be quenched by electropositive impurities in the  $\ell\text{Ar}$ : as the ionization electrons of an event drift upward, they can be captured by these impurities, reducing the size of S2. This process means that ionization electrons have an exponential lifetime as a function of  $t_{\text{drift}}$ : the *electron drift lifetime*,  $\tau_{\text{ed}}$ .

$\tau_{\text{ed}}$  can be found in a similar manner to the way the S1  $t_{\text{drift}}$  dependence is found: by looking at the average S2 size for events of a known S2 energy deposition. With S2 the process is a bit easier, even. Once S1 has been  $z$ -corrected, the mean S2/S1 for all  $^{83\text{m}}\text{Kr}$  events — or indeed, any electron recoil event (see Section

<sup>35</sup> [Alexander et al. \(2013a\)](#)

3.5) — should be constant as a function of  $t_{\text{drift}}$ . S2 is then  $z$ -corrected by simply plotting S2/S1 vs.  $t_{\text{drift}}$ , fitting the distribution, and dividing out the non-constant terms from S2, normalizing to the value of S2/S1 at the center of the detector volume. This process can be seen for ARIS data in [Appendix C](#).

Note that this S2  $z$ -correction is only necessary when there are significant electropositive impurities in the  $\ell\text{Ar}$ . With the first DS-50 campaign with AAr,  $\tau_{\text{ed}}$  was measured to be  $> 3.5$  ms,<sup>36</sup> giving a maximum 7% correction to the S2 yield as a function of event depth. Recent studies have shown that  $\tau_{\text{ed}}$  in DS-50 with UAr is  $> 5$  ms, roughly  $13\times$  longer than the maximum drift time at  $\mathcal{E}_d = 200$  V/cm. This is a testament to the state-of-the-art purification and distillation systems used by the DS Collaboration.  $\tau_{\text{ed}}$  is now so long in DS-50 that it cannot be accurately measured; because of this, no explicit  $z$  corrections are applied to S2.

S2 does have some  $xy$  dependencies, which will be explored in the next Chapter. S1 also has a very slight radial dependence ( $\sim 3\%$ ), but it's small enough that it can be disregarded.<sup>37</sup> The fully-corrected S1 LY at null field at the  $^{83\text{m}}\text{Kr}$  peak energy is  $8.1 \pm 0.2$  PE/keV<sub>ee</sub> with DS-50 filled with UAr.<sup>38</sup>

## CALIS

In September 2014, a CALibration Insertion System (CALIS) was installed in the clean room above the DarkSide Experiment.<sup>39</sup> This system allows small radioactive sources to be inserted into the LSV, right next to the stainless steel cryostat which surrounds the TPC. Several calibration campaigns during the fall of 2014 and the spring of 2015 which employed CALIS allowed for the measurement of

---

<sup>36</sup> [Agnes et al. \(2015\)](#)

<sup>37</sup> From internal docdb entry # 1611.

<sup>38</sup> [Agnes et al. \(2016c\)](#)

<sup>39</sup> [Agnes et al. \(2016a\)](#)

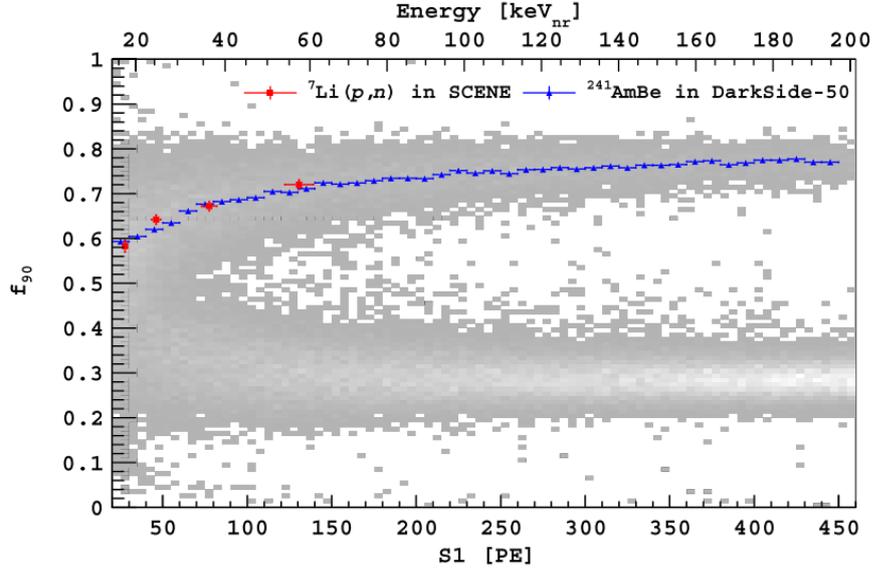


Figure 3.13: Calibration data taken with an  $^{241}\text{Am}^9\text{Be}$  source inserted with CALIS, exhibiting a strong neutron band at high  $f_{90}$ . Red ( $S1$ ,  $f_{90}$  median) points from SCENE are in very good agreement with the corresponding blue data points from AmBe. From [Agnes et al. \(2016a\)](#).

the detector response to a variety of radioactive sources, including  $^{57}\text{Co}$ ,  $^{133}\text{Ba}$ , and  $^{137}\text{Cs}$ . Data gathered from these sources provided crucial validation for g4ds, and data taken with an  $^{241}\text{Am}^9\text{Be}$  source “validated the nuclear recoil  $f_{90}$  response extrapolated from SCENE to DarkSide-50”<sup>40</sup> (see [Figure 3.13](#)).

During the calibration campaigns which used CALIS, the concentrations of TMB and PPO in the LSV were also adjusted to quantify their effect on the light yield of the LSV. Although a higher PPO concentration was found to reduce the LSV LY slightly ( $\sim 5\%$ , as can be seen in [Figure 3.14](#)), it also reduced  $\alpha$  scintillation quenching in the LSV allowing for easier detection and rejection of such events. A thorough description of CALIS and the CALIS calibration campaigns can be found in [Agnes et al. \(2016a\)](#).

<sup>40</sup> [Fan \(2015\)](#)

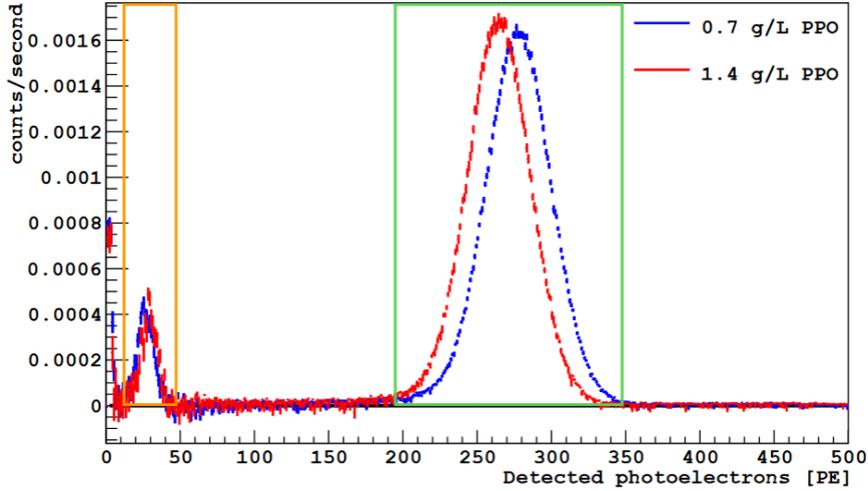


Figure 3.14: Scintillation signal in the LSV showing (orange box on the left) neutron capture leading to a 1775 keV  $\alpha$  (B.R. 6.4%) and (green box on the right) neutron capture leading to a 1471 keV  $\alpha$  plus a 478 keV  $\gamma$  (B.R. 93.6%). The peak shift to lower LY with higher PPO concentration is evident in the  $\gamma$  peak at  $\sim 270$  PE. From [Agnes et al. \(2016a\)](#).

### 3.5 Pulse Shape Discrimination

Although argon and xenon are both commonly used as target materials for direct dark matter searches, argon has one crucial advantage over xenon: particle identification. When a particle is incident on a  $\ell$ Ar target, there are two main types of collisions which can occur: Electron Recoils (ER) and Nuclear Recoils (NR). Typically, ERs occur when the incident particle interacts via the electromagnetic (EM) force with the electron cloud of an Ar atom, while NRs occur when a particle interacts — via the strong nuclear force — with the nucleus of an Ar atom. The gravitational force, of course, is so weak at these mass scales that it can effectively be ignored. The waveforms of typical ER and NR events in DS-50 can be seen in [Figure 3.15](#).

Electron ( $\beta$ ) and photon ( $\gamma$ ) collisions make up the bulk of the ERs in DS-50, though  $\alpha$  particle collisions have also been observed. Most of the NRs come from neutrons colliding with the Ar nuclei. WIMPs, since they do not interact

via the EM force, are NRs as well. In order to confirm a WIMP interaction in DS-50, then, one must be able to distinguish between ERs and NRs. This is done by examining the shape of the S1 pulse for these events, hence the name of this procedure, Pulse Shape Discrimination (PSD).

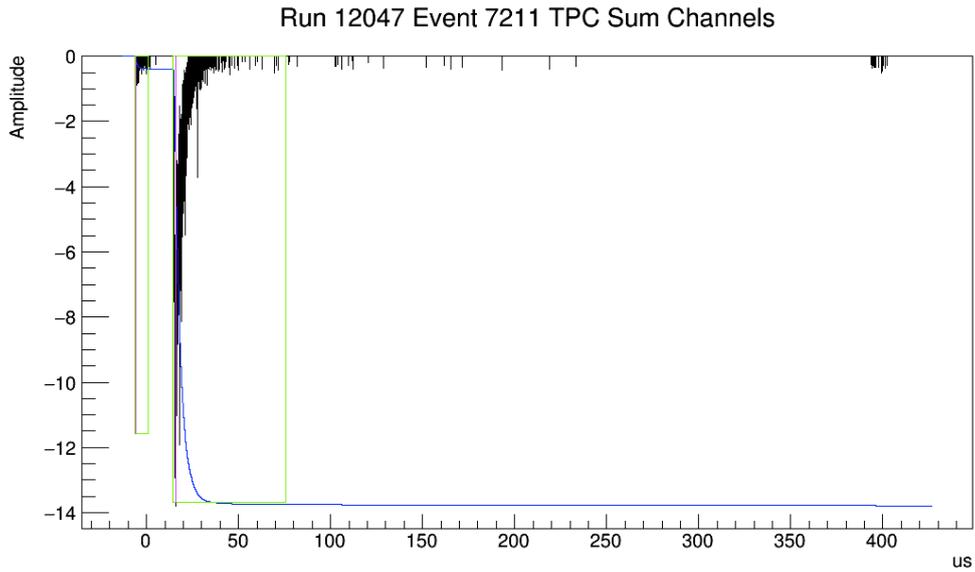
Recall Equations 3.2 and 3.3 from the beginning of Section 3.2. The weakly bound dimer  $\text{Ar}_2^*$ , or “exciton”, can be in one of two states: the singlet or the triplet. The singlet state has a  $\sim 7$  ns decay time, while the triplet has a much longer  $\sim 1.6$   $\mu\text{s}$  decay time,<sup>41</sup> though both states decay by emitting a single photon with wavelength  $\lambda = 128$  nm. Because the durations of scintillation from these two states are so distinctive, the shape of the S1 pulse depends on the singlet/triplet ratio. For ERs, this ratio is  $\sim 0.3$ , while for NRs, it’s closer to  $\sim 3.0$ . The exact mechanism behind *why* this is the case is still being investigated.

While the singlet/triplet ratio is not explicitly measured in DS-50, its effect on the shape of the S1 pulse is observed. The primary PSD parameter used is  $f_{90}$ : the fraction of the S1 pulse integral found within the first 90 ns. ERs have an  $f_{90}$  of  $\sim 0.3$ , while NRs have an  $f_{90}$  closer to  $\sim 0.75$  (see Figure 3.13). The shapes of typical ER and NR S1 pulses can be seen in Figure 3.16. “ERs” and “NRs”, then, aren’t strictly a measure of “interaction type”, but rather a measurement of the ionization density of the track a particle leaves behind as it interacts in the  $\ell\text{Ar}$ . Light particles leave long, diffuse tracks of ionization electrons and scintillation, while heavy particles leave behind short, dense ionization tracks. This is why  $\alpha$  particles, even though they are charged and interact primarily electromagnetically with the  $\ell\text{Ar}$ , have NR-like  $f_{90}$  values: it’s a consequence of the ionization density of the track.

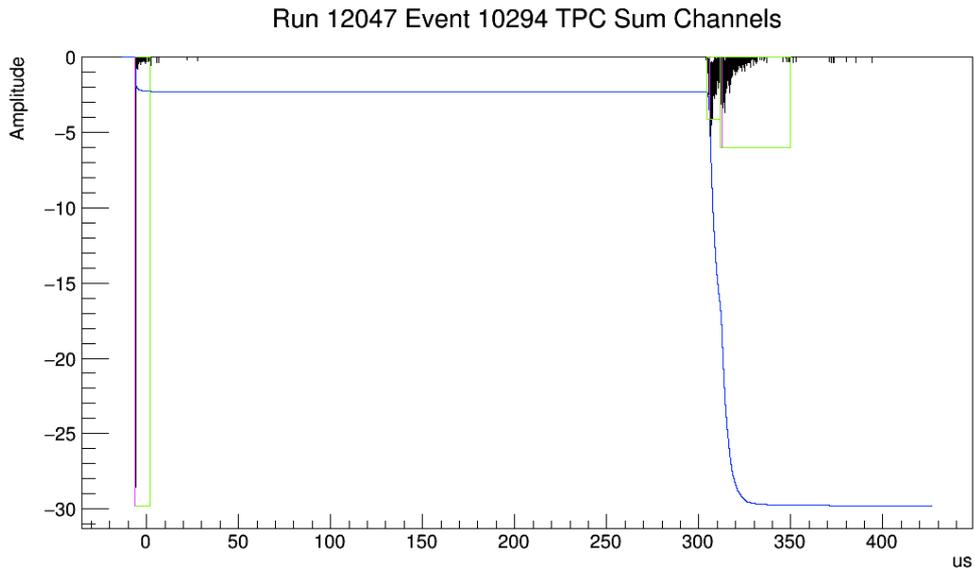
With AAr data in DS-50,  $f_{90}$  has proven to be a supremely useful parameter,

---

<sup>41</sup> [Alexander et al. \(2013a\)](#)

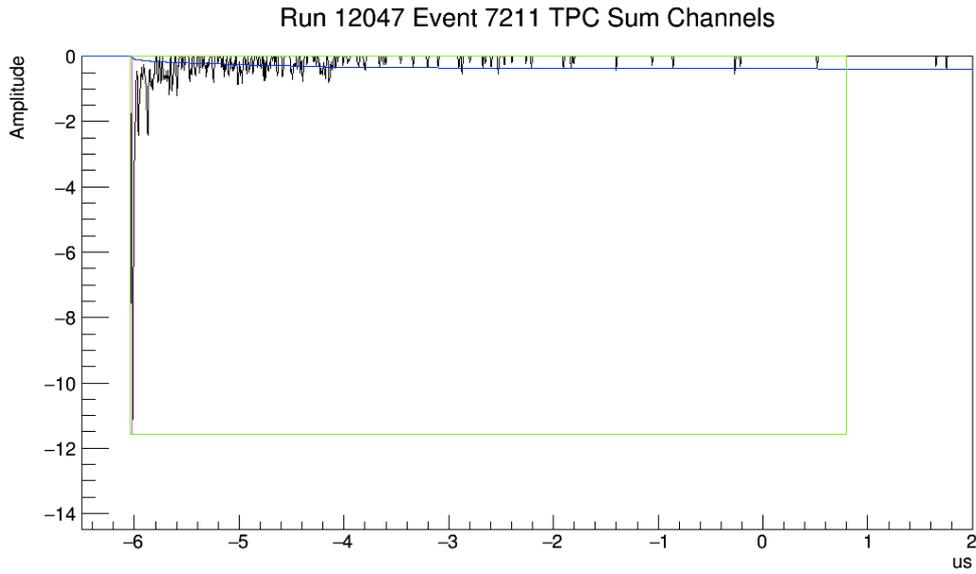


(a) ER event in DS-50.  $S_1 = 330.4$  PE,  $f_{90} = 0.284$

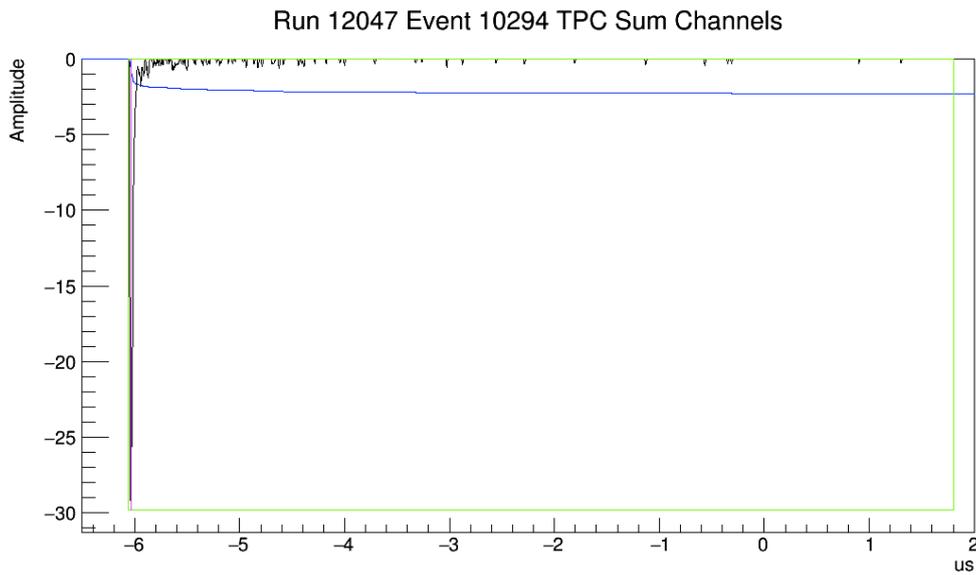


(b) NR event in DS-50.  $S_1 = 330.6$  PE,  $f_{90} = 0.708$

Figure 3.15: The waveforms of typical ER (a) and NR (b) events in DS-50. Black lines show the TPC sum channel waveforms, blue lines are the waveform integrals (scaled), green boxes show pulses found by the pulse-finder, and pink lines show the peaks of the pulses. The horizontal axes are in  $\mu\text{s}$  and the vertical axes are in PE.



(a) ER event in DS-50.  $S_1 = 330.4$  PE,  $f_{90} = 0.284$



(b) NR event in DS-50.  $S_1 = 330.6$  PE,  $f_{90} = 0.708$

Figure 3.16: The waveforms from Figure 3.15, magnified to see the shapes of the  $S_1$  pulses. Note that the NR event has a much sharper peak, as evidenced by the larger vertical scale, even though both events have  $S_1$ s of the same size ( $\sim 330$  PE).

exhibiting a discrimination power of at least 1 in  $1.5 \times 10^7$ .<sup>42</sup> (Larger data sets, which are currently being gathered, will allow stricter limits to be put on that figure, which is currently unbounded from above.) PSD in  $\ell\text{Ar}$  means that UAr exposures of at least 5.5 tonne · year can be made free of  $^{39}\text{Ar}$  background: DS-50 could run for two decades without mis-identifying an  $^{39}\text{Ar}$  ER as a NR.<sup>43,44</sup>

In addition to S1 PSD, the S2/S1 ratio can be used to provide additional rejection power. NR events should generate a higher density of electron-ion pairs, resulting in more recombination, and suppressing S2 relative to S1. ER events, conversely, should generate paths with a lower density of electron-ion pairs, with less recombination, and a larger S2/S1 ratio. This additional discrimination factor is currently not utilized in DS-50, though it is briefly investigated in [Agnes et al. \(2016c\)](#).

### 3.5.1 Particle Identification in $\ell\text{Xe}$

Particle identification in  $\ell\text{Xe}$  is quite different to the particle identification in  $\ell\text{Ar}$ . Qualitatively, the generation of scintillation light (S1) in  $\ell\text{Ar}$  and  $\ell\text{Xe}$  is similar: both have fast ( $\sim 7.0$  ns for Ar,  $\sim 4.3$  ns for Xe) and slow ( $\sim 1600$  ns,  $\sim 22$  ns) components from the decay of the singlet and triplet states of their excited dimers, respectively.<sup>45</sup> But the similarity of the decay times of these states in  $\ell\text{Xe}$  means that they are essentially indistinguishable, making PSD as it's performed in  $\ell\text{Ar}$  impossible.

Accordingly,  $\ell\text{Xe}$ -based experiments must make use of a different discrimination parameter to separate ERs from NRs in their data. The most common such parameter is the ratio S2/S1.  $\ell\text{Xe}$ -based experiments typically achieve rejection

---

<sup>42</sup> [Agnes et al. \(2016c\)](#)

<sup>43</sup> [Agnes et al. \(2015\)](#)

<sup>44</sup> [Agnes et al. \(2016c\)](#)

<sup>45</sup> [Hitachi et al. \(1983\)](#)

powers  $\mathcal{O}(10^3)$ ,<sup>46,47</sup> though the ZEPLIN-III experiment has reported a rejection power of  $10^4$  for very low recoil energies.<sup>48</sup> The  $> 1.5 \times 10^7$  discrimination power of DS-50 greatly exceeds the best quoted rejections of any  $\ell\text{Xe}$ -based experiment, and is one of the most attractive features of  $\ell\text{Ar}$  as a target for WIMP-detection experiments.

## 3.6 Backgrounds

WIMP searches are essentially exercises in background reduction. WIMP-nucleon collisions are expected to be extremely rare events, and so the validity of a WIMP detection claim rests on how well the backgrounds in a given experiment are understood. Employing active shielding allows backgrounds to be determined, compared to MC, and used to calculate the veto efficiency — something which cannot be done with passive shielding alone.<sup>49</sup> As outlined in [D'Angelo \(2015\)](#), DarkSide aims to be background-free by following three steps: (1) ensure very low intrinsic background levels, (2) discriminate ERs with PSD, and (3) actively suppress neutron backgrounds.

### 3.6.1 Intrinsic Backgrounds and $\ell\text{Ar}$ Purity

As mentioned previously, the primary background for any Ar-based direct dark matter search is the cosmogenic isotope,  $^{39}\text{Ar}$ .  $^{39}\text{Ar}$  is a  $\beta$ -emitter inherent to AAr, and so it can be distinguished from WIMP signals via PSD. But the concentration of  $^{39}\text{Ar}$  is also drastically reduced by using UAr, rather than AAr as the target material. Significant efforts have been made to use low-radioactivity materials in

---

<sup>46</sup> [Aprile et al. \(2012\)](#)

<sup>47</sup> [Akerib et al. \(2014\)](#)

<sup>48</sup> [Lebedenko et al. \(2009\)](#)

<sup>49</sup> [Wright \(2011\)](#)

the construction of DS-50, as well. Materials are radio-assayed with HPGe counting, mass spectrometry, and other techniques to characterize their radioactivity before being used in the experiment, and all assembly of detector materials is performed in a radon-free clean room to reduce the activity of radon daughter isotopes. The largest source of neutrons inherent to DS-50 are the Hamamatsu PMTs, though this background should be significantly reduced with the use of SiPMs in DS-20k.<sup>50</sup>

Finally, the chemical contamination of the  $\ell$ Ar can be determined with an exponential fit to the “slow component” ( $\sim 1.6 \mu\text{s}$  triplet state tail) of S1. Oxygen and nitrogen dimers, in particular, are known to reduce the scintillation output of  $\ell$ Ar by converting some of the deposited energy to heat (“non-radiative collisional de-excitation”).<sup>51</sup> While not an intrinsic radioactive background, these chemical contaminants do reduce the overall scintillation and ionization output of  $\ell$ Ar.

### 3.6.2 Surface Backgrounds

Neutrons and  $\alpha$  particles are the most dangerous backgrounds in an  $\ell$ Ar direct detection dark matter experiment, as these particles evade the usual PSD parameter cut, making them, at least naïvely, indistinguishable from WIMP-nucleon collisions. However, the possibility of a WIMP interacting more than once within the volume of the detector is vanishingly small, and so events which interact multiple times (perhaps in the vetoes and also in the TPC) are assumed to be neutrons or other neutral particles, and are discarded as WIMP candidates. Neutrons and  $\alpha$  particles which interact only once, though, are indistinguishable from WIMP interactions on an event-by-event basis.<sup>52</sup>

---

<sup>50</sup> [Agnes et al. \(2015\)](#)

<sup>51</sup> [Alexander et al. \(2013a\)](#)

<sup>52</sup> [Wright \(2011\)](#)

$\alpha$  particles (emitted by decays of isotopes like  $^{210}\text{Po}$  which is only found at the surfaces of the active volume) typically generate very large S1 signals ( $\mathcal{O}(10^4 \text{ PE})$ ) and are accordingly far outside of the DM energy range of interest in DS-50 ( $\sim 80 \rightarrow 460 \text{ PE}$ ),<sup>53</sup> but neutrons can easily fall within this range. (And degraded  $\alpha$  events can also occasionally fall within this energy range.) While cosmogenic neutrons are vetoed efficiently ( $\ll 1$  event expected in a multi-year DS exposure), radiogenic neutrons, emitted from the materials which comprise the walls of the TPC and the TPC PMTs are a concern, dealt with by (1) reducing the inherent neutron radioactivity of the materials by assembling the detector in a radon-free clean room, as mentioned above, and (2) fiducializing the detector.

### 3.6.3 Fiducialization

The “fiducial” volume of a detector is the volume which is used after position-dependent cuts have been made. For instance, in the UAr campaign of DarkSide-50, a  $t_{\text{drift}}$  cut is being used: ( $40 \mu\text{s} < t_{\text{drift}} < 334.6 \mu\text{s}$ ), equivalent to “shaving” 36.3 mm of height off of the top and bottom of the  $\ell\text{Ar}$  volume.<sup>54</sup> This eliminates a significant fraction of the NRs and ERs generated by the radioactivity inherent in the TPC PMTs as well as the anode and cathode windows, though there’s a corresponding loss of active volume and detector sensitivity. Currently, no radial fiducial volume cut is being used in the DarkSide-50 analysis chain, though this could be used to reduce surface WIMP-like backgrounds in future exposures. This thesis proposes radial fiducial volume cuts in [Chapter 4](#), investigating their efficiency and impact on NR acceptance.

---

<sup>53</sup> [Agnes et al. \(2017\)](#)

<sup>54</sup> [Agnes et al. \(2015\)](#)

## 3.7 Cuts

Several cuts are made to the data gathered by DS-50 before any conclusions can be made with regards to detector sensitivity. These cuts are broken into three groups: data quality cuts, veto cuts, and physics cuts. All standard analysis cuts begin with **CX** and are defined in internal docdb entry # 937-v7. All quoted text in the following Subsections is taken from this docdb entry.

### 3.7.1 Data Quality Cuts

**CX01:** This cut requires that all 38 TPC PMT channels were active during the acquisition of the given event. This is especially important for  $xy$  position reconstruction, which requires all 19 anode channels to be active.

**CX02:** This cut requires that the baseline is properly recognized in the TPC PMT sum channel. If the baseline isn't successfully found, there will be issues with determining the S1 and S2 integrals.

**CX03:** **CX03** is defined so that the physical start time of a given event is at least 1.35 ms after the physical start time of the preceding event. This removes piled-up events and events which may have triggered on the S2 tail of the preceding event.

**CX04:** **CX04** was used only during AAr running and required that the lifetime of an event be  $< 1$  s. Events with lengths in excess of this usually indicated stalls in the DAQ, which were reduced significantly when moving to UAr, where the TPC trigger rate was  $\sim 30\times$  smaller.

### 3.7.2 Veto Cuts

The veto cuts have changed significantly between the AAr campaign and the UAr campaign. Below (after CX05, which is common to both AAr and UAr data), the AAr cuts are given first, followed by the UAr cuts.

**CX05:** This cut requires the boolean `veto_present` to be `true`. For some runs, the veto was turned off completely, but for others, individual events may be missing veto information due to data corruption or other errors.

#### AAr Veto Cuts

**CX06:** This cut removes events with  $\geq 10$  PE within  $-10 \text{ ns} < t < 200 \text{ ns}$  (relative to the TPC trigger time) in the LSV. CX06 aims to cut neutron events which leave a detectable signal in the LSV over a few neutron thermalization times. In [Appendix D](#), I show how the LSV neutron thermalization time was found to be  $\sim 20 \text{ ns}$  with the current TMB concentration (5%).

**CX07:** CX07 is a combination of three veto cuts: the delayed, late, and CTF cuts.

The delayed cut is designed to detect the gamma ray from neutron capture on  $^{10}\text{B}$ ; this cut finds the maximum charge seen in any continuous 300 ns window within the first  $X \mu\text{s}$  after the TPC trigger, where  $X$  is approximately four times the neutron capture time in the LSV. The integral over this 300 ns region is required to be  $< 80$  PE. With early campaigns, a 50/50 PC/TMB ratio meant that the capture time was  $\sim 2.2 \mu\text{s}$ , while in later campaigns, the TMB concentration was reduced to  $\sim 5\%$  and the capture time rose to  $\sim 20 \mu\text{s}$ . See [Appendix D](#) for more details.

The late cut is designed to detect the neutron capture in stainless steel, which can produce a gamma ray in the LSV much later than the  $^{10}\text{B}$  capture does. The late cut removes events with more than 110 PE in the entire veto window.

The CTF (Counting Test Facility, aka WCD) cut is designed to cut events which contain the signature of muons crossing the WCD. This cuts events with “an integral of the CTF sum waveform in a 300 ns window with more than 200 PE in the entire veto record length”.

### UAr Veto Cuts

In UAr, CX06 and CX07 aren't as clearly defined. Instead, there are five named veto cuts, described in detail in internal docdb entry # 1657:

**The Prompt Cut:** This cut aims to remove  $\gamma$ s and neutron TPC events with prompt thermalization signals in the LSV. This cut requires the LSV charge collected during  $[-50 \text{ ns}, 250 \text{ ns}]$ , relative to the S1 time, to be  $\leq 1$  PE.

**The Delayed Cut:** The delayed cut is designed to veto neutrons which exhibit a capture signal in the LSV. This cut finds the 500 ns acquisition sub-window (which starts any time after the start of S1) with the greatest charge in the LSV and vetoes the event if the charge in that window is  $> 3$  PE ( $> 6$  PE) for veto run numbers  $< 12638$  ( $\geq 12638$ ). The size of the veto acquisition window was extended from  $140 \mu\text{s}$  to  $200 \mu\text{s}$  at run 12638, which may lead to more background appearing in the window; hence the larger threshold.

**The Pre-Prompt Cut:** The pre-prompt cut vetoes external neutron and gamma events, which give a signal in the LSV before the start of the S1 pulse. Similar to the delayed cut, the pre-prompt cut has a sliding 500 ns window. But the pre-prompt cut window must start and end prior to the S1 start time, while the delayed cut window must start and end after the S1 start time. An event is rejected if the integral of the pre-prompt window is anywhere  $> 3$  PE.

**The Muon Cut:** The muon cut aims to remove muons and other high-energy cosmogenic backgrounds which occur before the S1 start time and travel through the WCD or the LSV. An event is rejected by this cut if the integral of the entire LSV window exceeds 2000 PE or if the integral of the entire WCD window exceeds 400 PE.

**The Cosmogenic Cut:** The final UAr veto cut, the cosmogenic cut, vetoes delayed cosmogenic-like backgrounds. Muons can produce heavy nuclei with decay lifetimes  $\mathcal{O}(100 \text{ ms})$  which produce high-energy neutrons. The cosmogenic cut is designed to avoid those neutrons, by vetoing any events that occur within 2 sec after a muon (as defined by the muon cut, above).

For the remainder of this dissertation, “all CX cuts” or “all fifteen cuts” includes (for the veto) CX05, CX06, and CX07 for AAr, and CX05, the Prompt Cut, the Delayed Cut, the Pre-Prompt Cut, the Muon Cut, and the Cosmogenic Cut, for UAr.

### 3.7.3 Physics Cuts

**CX08:** This cut requires events to have two pulses (S1 and S2) but also allows events with three pulses if the third pulse is S3. S3 must occur  $\sim 375 \mu\text{s}$  (the maximum  $t_{\text{drift}}$ ) after S2. This cut ensures that only single-sited events are selected, eliminating some neutron and gamma ray backgrounds.<sup>55</sup>

**CX09:** This cut only passes events which have S1 start times at the trigger time  $\pm 50 \text{ ns}$ . This removes events which trigger on the tails of pulses or which trigger on pileup. The pulse-finder is independent of the hardware trigger, and so it is possible that the first pulse of an event is 50+ ns away from the trigger time.

**CX10:** CX10 requires that the S1 signal isn't saturated. Within the WIMP energy range, this cut removes essentially zero events.

**CX11:** This is the “S1 max fraction” (S1Mf) cut, designed to remove Čerenkov events, where a particle passing through one of the fused silica windows (or one of the PMT windows) generates Čerenkov radiation. This radiation, because it is usually generated within, or right next to, a PMT, gives a very large fraction of the event's light to a single PMT. This cut is designed to remove events with the highest 5% (1%) of S1Mfs within bins of 5 PE for UAr (AAr). Čerenkov + ER coincidence events are a dangerous background because they can mimic the  $f_{90}$  of a NR if the Čerenkov light is detected in coincidence with the start of the S1 pulse. This cut reduces this background.

**CX12:** This cut removes events with S2  $f_{90\text{s}}$  of  $> 0.2$ . S2  $f_{90}$  is usually  $\lesssim 0.05$ , and so in general, this cut removes events where the second pulse in an event was

---

<sup>55</sup> [Agnes et al. \(2015\)](#)

an S1 pile-up within the acquisition window.

**CX13:** This cut removes events where  $S2 \leq 100$  PE. S2 is typically  $\sim 20\times$  larger than S1 for ER events in the WIMP collision energy range, and so S2 is usually  $\mathcal{O}(10^3 - 10^4$  PE).

**CX14:** This cut passes AAr events with  $80 \text{ PE} < S1 < 460 \text{ PE}$ . These are events with  $35 \text{ keV}_{\text{nr}} \lesssim S1 \lesssim 205 \text{ keV}_{\text{nr}}$ . Events with energies below this range must be removed due to a loss of PSD efficiency from reduced PE statistics. Events with energies above this range are outside the WIMP region of interest. Note that this range was widened a bit for UAr running (see Figure 3.17).

**CX15:** This is a cut on  $t_{\text{drift}}$  which removes events which occur within 36.3 mm of the cathode or anode windows (ignoring the gas pocket at the top of the active volume). The maximum  $t_{\text{drift}}$  is measured occasionally, so this cut changes a bit over the run database, but in general it's fairly close to  $40 \mu\text{s} \lesssim t_{\text{drift}} \lesssim 334.6 \mu\text{s}$  when  $\mathcal{E}_d = 200 \text{ V/cm}$ .

### 3.7.4 Result of Cuts

In the AAr campaign, the fifteen cuts described above remove  $18_{-1}^{+4}\%$  of events from the initial  $\sim 74$  M event sample, leaving  $47.1 \pm 0.2$  days of run time with a  $36.9 \pm 0.6$  kg active volume for a total  $1422 \pm 67$  kg · days of exposure. After all CX cuts are applied, no events remain in the region of parameter space expected for DM candidates, defined in the  $f_{90}$  vs. S1 plane (seen for UAr in Figure 3.17).

Note that, while  $xy$  was originally planned to be used to provide a fiducial volume cut to reduce surface backgrounds, no fiducial cut was used in the above-described exposure. PSD, the veto cuts, and the  $t_{\text{drift}}$  cut are more than adequate

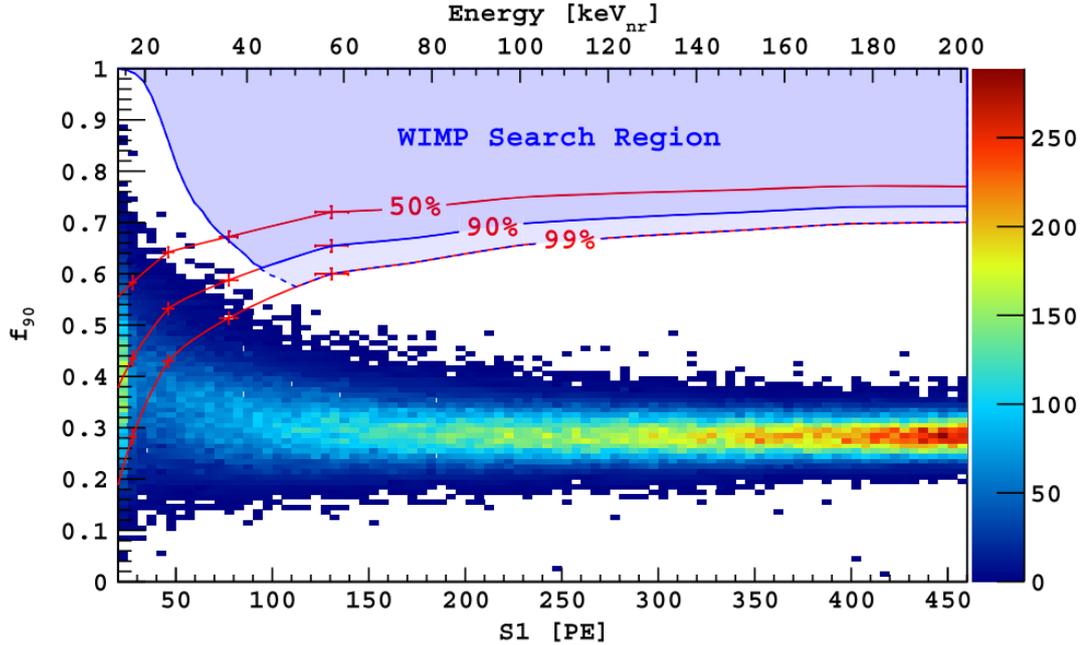


Figure 3.17: Full event set from the DS-50 70-day campaign with UAr. All CX cuts from Section 3.7 have been applied and no WIMP candidates can be seen. Red crosses show SCENE NR data points and error bars, red lines show  $^{241}\text{Am}^9\text{Be}$  data for 50%, 90%, and 99% NR acceptance in the given 5 PE S1 bin. The concentration of events with  $f_{90} \sim 0.4$  in the lowest energy bin are from  $^{37}\text{Ar}$ . From Agnes et al. (2016c).

to remove the surface  $\gamma$ ,  $\beta$ , and neutron backgrounds without the aid of fiducial volume cuts or the additional potential discrimination power provided by a cut on S2/S1.<sup>56</sup> Fiducial volume cuts could become necessary in longer exposures, though, and so this thesis does discuss them in Chapter 4.<sup>57</sup>

### 3.8 Exclusion Limit Plot

For direct dark matter searches, the common point of contact among all experiments is the exclusion limit plot, first introduced here in Subsection 2.8.4. This plot shows the common parameter space of WIMP-nucleon cross section on the

<sup>56</sup> Agnes et al. (2016c)

<sup>57</sup> Agnes et al. (2015)

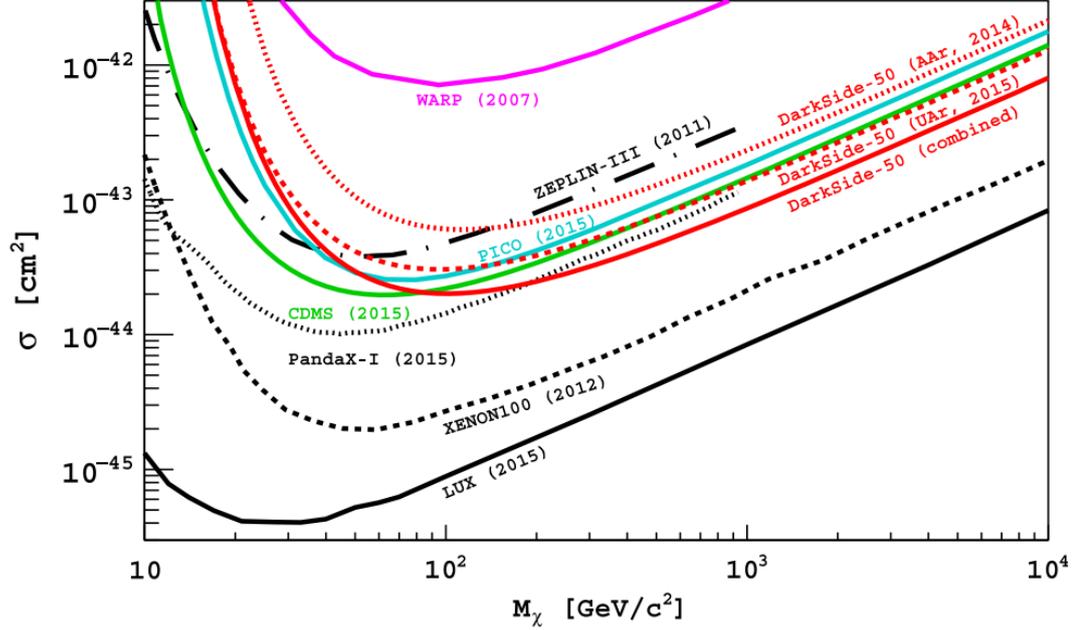


Figure 3.18: Recent exclusion limit curves (as of October 2015) from a selection of DM experiments. All of the experiments featured are mentioned in Subsection 2.8.3 except WArP (the “WIMP Argon Programme”), an Ar-based predecessor to DarkSide. From [Agnes et al. \(2016c\)](#).

vertical axis vs. WIMP mass on the horizontal axis. Usually, the cross section is presented in  $\text{cm}^2$  and the WIMP mass in  $\text{GeV}/c^2$ . The standard procedure through which these plots are generated was established in [Lewin and Smith \(1996\)](#), with a velocity correction as introduced in [Savage et al. \(2006\)](#). For DarkSide, we assume a standard isothermal WIMP halo with a WIMP galactic escape velocity  $v_{esc} = 544 \text{ km/s}$ , a galactic rotation velocity of  $v_0 = 220 \text{ km/s}$  at the Earth’s position, a revolution velocity of the Earth about the Sun of  $v_{Earth} = 232 \text{ km/s}$ , and a local DM density of  $\rho_{dm} = 0.3 \text{ GeV}/(c^2 \cdot \text{cm}^3)$ .

The exclusion limit plot depends on seven key factors, as outlined in [Cao \(2014\)](#): (1) the energy spectrum of WIMP-induced nuclear recoils, (2) the S1 and S2 LY of the detector, (3) the scintillation and ionization efficiencies of nuclear recoils, (4) the S1 and S2 resolution functions, (5) the DAQ trigger efficiency as a function of S1, (6) the expected distributions of the discrimination parameters ( $f_{90}$

and S1), and (7) the WIMP acceptance regions in the discrimination parameter space. A full review of this procedure is outside the scope of this work, but [Fan \(2016\)](#) gives a very clear step-by-step walkthrough as applied to DarkSide. Two points in particular are worth discussing, though.

### 3.8.1 SCENE

During June and October of 2013, the SCENE (“SCintillation Efficiency of Noble Elements”) Experiment took place at the University of Notre Dame Institute for Structure and Nuclear Astrophysics, in Notre Dame, Indiana. SCENE measured the S1 and S2 LYs of recoiling nuclei in  $\ell$ Ar as a function of  $\mathcal{E}_d$ . Like DS, SCENE used a two-phase argon TPC where all interior surfaces were coated with TPB, employed Hamamatsu R11065 PMTs, had fused silica windows coated in ITO, and performed  $^{83\text{m}}\text{Kr}$  calibrations to track the LY.<sup>58</sup>

Unlike DS-50, the SCENE TPC is small and unshielded, and so source calibrations which cannot be performed in DS-50 without CALIS can be performed quite easily with SCENE. SCENE’s primary goal, though, was to measure the  $\mathcal{E}_d$ -dependence of the S1 LY, as earlier studies with  $\ell\text{Xe}$  suggested a  $< 10\%$  effect on the NR LY, independent of energy. By placing SCENE in a narrowband neutron beam and capturing the recoiling neutrons with EJ301 liquid scintillator-based neutron detectors, the energy that the neutron deposited in the  $\ell$ Ar could be calculated analytically. SCENE showed that the NR LY (and, accordingly, the NR quenching,  $\mathcal{L}_{eff}$ ) depends not only on the energy of the NR, but on  $\mathcal{E}_d$  as well. For NRs between  $\sim 10$  and  $\sim 50$  keV<sub>nr</sub>, “a significant dependence (up to 32%) on the drift field of liquid argon scintillation” was found.<sup>59</sup> (SCENE also used less chemically pure  $\ell$ Ar, relative to DS-50, and thus had to make the S2  $z$ -correction

---

<sup>58</sup> [Cao et al. \(2015\)](#)

<sup>59</sup> [Alexander et al. \(2013b\)](#)

mentioned in Subsection 3.4.2.)

$\mathcal{L}_{eff}$  is defined (as  $\mathcal{L}_{eff,^{83m}\text{Kr}}$ ) in Cao et al. (2015) as the scintillation efficiency of NRs relative to that of ERs from  $^{83m}\text{Kr}$  at zero field

$$\mathcal{L}_{eff}(E_{\text{NR}}, \mathcal{E}_d) = \frac{S1_{\text{NR}}(E_{\text{NR}}, \mathcal{E}_d)/E_{\text{NR}}}{S1_{\text{Kr}}(\mathcal{E}_d = 0)/E_{\text{Kr}}} \quad (3.4)$$

where  $E_{\text{Kr}}$  is 41.5 keV,  $E_{\text{NR}}$  is the NR energy and  $\mathcal{E}_d$  is the drift field.  $\mathcal{L}_{eff}$  is crucial to determining the relationship between NR energy and S1 in an  $\ell\text{Ar}$ -based TPC. The presence of an external electric field means that electron-ion pairs are separated more easily when a particle interacts in the  $\ell\text{Ar}$ , where lower-density tracks are affected more. Since the density of these tracks differs between ERs and NRs (see Section 3.5), the likelihood of recombination (and hence the size of S1) is dependent on the type of interaction. In short, SCENE was able to determine an energy-dependent, drift-field-dependent conversion factor between S1 and NR energy. To convert the S1 measurements in DarkSide to  $\text{keV}_{\text{nr}}$ , a conversion factor like this is necessary.

### 3.8.2 $f_{90}$

Point (6), above, requires that we know “the expected distributions of the discrimination parameters” to be able to generate the exclusion limit plot. In DarkSide-50, the discrimination parameters are S1 (NR energy) and  $f_{90}$  (the particle identification parameter). Point (1), above, which can be calculated by making a few assumptions about the local WIMP density, interaction cross section, and mass as outlined in Cao (2014), gives the expected distribution of S1, but calculating the expected distribution of  $f_{90}$  has been significantly more difficult in DS-50.

$f_{90}$  was defined and first studied in detail in Lippincott et al. (2008), using an analytical model introduced by Hinkley (1969), and has been scrutinized by the

DS analysis team for several years.  $f_{90}$  models seem to over- or under-estimate the tails of the distribution, though recent results, including new models developed by DarkSide collaborators, seem to match the AAr  $f_{90}$  data quite nicely. With the expected distributions of S1 and  $f_{90}$  in hand, the expected leakage (the number of events which make it past all CX cuts and could be mistaken for NR events) can be calculated, and the DS-50 exclusion limit plot can be constructed.

### 3.8.3 Results

For the full 70-day UAr exposure, the events passing all cuts discussed above are shown in the  $f_{90}$  vs. S1 plane in Figure 3.17. In this plane, one can define a “leakage curve” by finding the  $f_{90}$  value at which — for a given 5 PE bin in S1 — the  $f_{90}$  model predicts a 0.01 ER leakage into the NR band. The WIMP search area is defined by finding the intersection of this event leakage curve and the 90% NR acceptance line from SCENE (red crosses) and  $^{241}\text{Am}^9\text{Be}$  campaign data (red line), and choosing whichever curve is highest for a given S1 bin. This procedure predicts a total of  $< 0.1$  leakage events over the entire UAr exposure. Combined with the previous AAr campaign results, a total of 1 WIMP-like event remains in the expected parameter space after all cuts have been applied, in agreement with the MC prediction of  $2 \pm 2$  such events. Thus, no event consistent with WIMP detection occurs in this data set.

DarkSide-50 is the leading DM direct detection experiment with an  $\ell\text{Ar}$  target. The combined AAr-UAr exposure gives a WIMP-nucleon cross section upper limit of  $2.0 \times 10^{-44} \text{ cm}^2$  ( $8.6 \times 10^{-44} \text{ cm}^2$ ,  $8.0 \times 10^{-43} \text{ cm}^2$ ) for a WIMP mass of  $100 \text{ GeV}/c^2$  ( $1 \text{ TeV}/c^2$ ,  $10 \text{ TeV}/c^2$ ) — the most restrictive limit ever found using an argon TPC.

## Chapter 4

# PRINCIPAL COMPONENT ANALYSIS AND MULTI-DIMENSIONAL FITTING

Principal Component Analysis (PCA), in a general sense, reveals the underlying structure or patterns in a large set of multivariate data. This dataset is composed of vectors of the same dimensionality, measured in the same units.<sup>1</sup> For the DS-50 PCAMDF *xy*-Reconstruction, these vectors are 19-dimensional, and

---

<sup>1</sup> For instance, a set of 3-dimensional position vectors is fine, but a set of 4-dimensional spacetime position vectors is not. PCA is sensitive to the scaling of the variables used.

consist of the mean-centered logarithms of the anode S2 fractions:

$$\vec{X}_i = \begin{bmatrix} L_i^{(1)} \\ L_i^{(2)} \\ L_i^{(3)} \\ \dots \\ L_i^{(19)} \end{bmatrix} \quad (4.1)$$

$\vec{X}_i$  is the  $i$ -th event vector, consisting of

$$L_i^{(j)} = \ln \left( \frac{S2_i^{(j)} [\text{PE}]}{S2_i^{\text{top}} [\text{PE}]} \right) - \left\langle \ln \left( \frac{S2_i^{(j)} [\text{PE}]}{S2_i^{\text{top}} [\text{PE}]} \right) \right\rangle = \ell_i^{(j)} - \langle \ell^{(j)} \rangle \quad (4.2)$$

which are the mean-centered logarithms of the anode S2 light fractions. Here, angled brackets give the average over all events (over all  $i$ ),  $S2_i^{(j)}$  is the S2 light (in PE) seen by the  $j$ -th anode PMT for the  $i$ -th event,<sup>2</sup> and  $S2_i^{\text{top}}$  is the sum of the S2 light (in PE) seen by all 19 anode PMTs for the  $i$ -th event. For brevity in future expressions, also defined are

$$\ell_i^{(j)} = \ln \left( \frac{S2_i^{(j)} [\text{PE}]}{S2_i^{\text{top}} [\text{PE}]} \right) \quad \text{and} \quad \langle \ell^{(j)} \rangle = \left\langle \ln \left( \frac{S2_i^{(j)} [\text{PE}]}{S2_i^{\text{top}} [\text{PE}]} \right) \right\rangle \quad (4.3)$$

which are the logarithm of the fraction of the anode S2 light seen by the  $j$ -th anode PMT, and its average over all  $i$  events, respectively. Since  $S2_i^{(j)}$  and  $S2_i^{\text{top}}$  are both measured in PE,  $\ell_i^{(j)}$  is unitless.  $i$  runs from 1 to  $N_E$ , where  $N_E$  is the total number of events in the dataset.  $j$  runs from 1 to  $N_D$ , where  $N_D$  is the dimensionality of the dataset (in this case,  $N_D = 19$ ). Logarithms are used because they improve

---

<sup>2</sup> Figure 3.4 shows that the anode PMTs have indices 19–37. The  $j$ -th PMT throughout this chapter refers to the  $j$ -th *anode* PMT. So  $j = 1$  refers to PMT 19,  $j = 2$  refers to PMT 20, and so on. No bottom PMT information is used in the PCAMDF  $xy$ -Reconstruction.

the accuracy of the  $xy$ -Reconstruction method on simulated events.

## 4.1 Principal Component Analysis

### 4.1.1 The Covariance Matrix

The philosophy behind PCA is that there is some “best way” to express any set of data, which stores as much information about the data in as few dimensions as possible. In a nutshell, PCA is a transformation from the original multidimensional space in which the measurements were made, to a new “optimal” space (equivalently, to a new “basis”), where the fewest number of dimensions see the largest possible amount of variance in the dataset. PCA is also a *non-parametric* algorithm. It’s a step-by-step procedure with a unique solution (modulo trivial degeneracies like multiplying eigenvectors by  $-1$ ), where nothing needs to be “tuned” by the user.<sup>3</sup> How does a PCA determine the optimal basis? If one were to write a step-by-step procedure for finding the new basis, it might look like:

1. Find the direction in the  $N_D$ -dimensional measurement space over which the data has the maximum variance. Normalize this vector and save it as  $\vec{p}_1$ .
2. Find a new direction, orthogonal to  $\vec{p}_1$ , over which the data has the maximum remaining variance (it will be  $\leq$  the variance along the  $\vec{p}_1$  direction). Normalize this vector and save it as  $\vec{p}_2$ .
3. Continue step 2 until all  $\vec{p}_j$  (where  $j = 1 \rightarrow N_D$ ) have been found. Each new  $\vec{p}_j$  must be orthogonal to all previous  $\vec{p}_j$  found, and each new  $\vec{p}_j$  will be

---

<sup>3</sup> This is both a positive and a negative: positive because it means the process is algorithmic and straightforward, negative because it requires that the user has some intuition as to how their data are related. If there is a non-linear relationship between the different measurements, the user needs to first parameterize the data before using a PCA.

along a direction where the data exhibit equal or lesser variance than along any previously found  $\vec{p}_j$ .

In a general sense, this is what a PCA does; it finds the basis vectors  $\vec{p}_j$ .

The unit vectors found by this procedure are all orthogonal, and so they form an orthonormal basis which spans the  $N_D$ -dimensional vector space within which the original data lies. To find these vectors, start with the undiagonalized covariance matrix,  $\hat{C}_X$ :

$$\hat{C}_X \equiv \frac{1}{N_E} \hat{X} \hat{X}^T \quad (4.4)$$

where  $\hat{X}$  is a matrix containing all event vectors  $\vec{X}_i$ , which form the columns of  $\hat{X}$ , and  $^T$  indicates a transpose. This means that  $\hat{X}$  is an  $N_D \times N_E$  matrix, so  $\hat{X} \hat{X}^T$  is a symmetric, square  $N_D \times N_D$  matrix. The  $j$ -th diagonal element of  $\hat{X} \hat{X}^T$  can be expressed as

$$L_1^{(j)} \cdot L_1^{(j)} + L_2^{(j)} \cdot L_2^{(j)} + \dots + L_{N_E}^{(j)} \cdot L_{N_E}^{(j)} = \sum_i^{N_E} \left( L_i^{(j)} \right)^2 \quad (4.5)$$

which means that the  $j$ -th diagonal element of  $\hat{C}_X$  is the variance of  $\ell_i^{(j)}$ :

$$\frac{1}{N_E} \sum_i^{N_E} \left( L_i^{(j)} \right)^2 = \frac{1}{N_E} \sum_i^{N_E} \left( \ell_i^{(j)} - \langle \ell^{(j)} \rangle \right)^2 = Var \left( \ell_i^{(j)} \right) \quad (4.6)$$

For off-diagonal elements of  $\hat{C}_X$ , we have (for the element in the  $a$ -th row and the  $b$ -th column):

$$C_{X;a,b} = \frac{1}{N_E} \left( L_1^{(a)} \cdot L_1^{(b)} + L_2^{(a)} \cdot L_2^{(b)} + \dots + L_{N_E}^{(a)} \cdot L_{N_E}^{(b)} \right) = \frac{1}{N_E} \sum_i^{N_E} \left( L_i^{(a)} \cdot L_i^{(b)} \right) \quad (4.7)$$

which is equal to the covariance between the  $a$ -th and  $b$ -th elements of  $\vec{X}_i$  over all events  $i$ :

$$\begin{aligned}
C_{X;a,b} &= \frac{1}{N_E} \sum_i^{N_E} \left( L_i^{(a)} \cdot L_i^{(b)} \right) \\
&= \frac{1}{N_E} \sum_i^{N_E} \left( \left( \ell_i^{(a)} - \langle \ell^{(a)} \rangle \right) \cdot \left( \ell_i^{(b)} - \langle \ell^{(b)} \rangle \right) \right) = \text{Covar} \left( \ell_i^{(a)}, \ell_i^{(b)} \right)
\end{aligned} \tag{4.8}$$

hence the name ‘‘covariance matrix’’ for  $\hat{C}_X$ .<sup>4</sup>  $\hat{C}_X$  encodes information about the relationships between the elements of the data vector across all events. If two measurements are strongly correlated (for instance, the S2[] values for two PMTs which are physically near each other in the detector), the covariance of those measurements will be large and positive. If one of those two PMTs has a large S2[] value, the other will tend to also have a large S2[] value, due to their proximity.

If two measurements are strongly negatively correlated (or ‘‘anti-correlated’’; for instance, the S2[] values for two PMTs which are on opposite sides of the anode), the covariance of those measurements will be large and negative. Covariance is nearly zero when there is little to no correlation between two measurements over the data set (for instance, if one of the PMTs was having frequent and random light emission issues).

The variance is large when a given anode PMT sees a large spread of S2[] values over the dataset, and small when a given anode PMT sees only a narrow range of S2[] values over the dataset. Variance is always non-negative, due to the squaring of the  $L_i^{(j)}$  term, above. High-variance PMTs convey more information, while PMTs with low variance may have hindered performance due to the

---

<sup>4</sup>  $\hat{C}_X$  is also sometimes called the ‘‘variance-covariance matrix’’, particularly when being compared to a  $\hat{C}_X$  which has had its diagonal ‘‘zeroed’’. In that case, the second matrix would be called the ‘‘covariance’’ matrix, since it only stores covariance information and no variance information.

geometry of the detector or detector defects.

To summarize, large diagonal elements correspond to PMTs with large variances over the dataset. PMTs with large variances contain some information about the event set, while PMTs with small variances could be recording mostly noise. In some sense, a large variance for a PMT corresponds to a high signal-to-noise ratio (“*SNR*”; assuming the noise is comparable across all PMTs). Large (whether positive or negative) off-diagonal elements correspond to pairs of PMTs with large covariances. A pair of PMTs with a small covariance indicates little correlation between their measurements. High covariance, then, indicates some redundancy in the measurements.

Summarizing what we know about  $\hat{C}_X$ :<sup>5</sup>

1. It’s an  $N_D$ -dimensional, symmetric square matrix.
2. The  $j$ -th diagonal element of  $\hat{C}_X$  gives the variance of the  $j$ -th measurement (in DarkSide-50, this would be the variance of the logarithm of the anode S2 light fraction on a particular anode PMT).
3. The off-diagonal elements give the covariance between different measurements (PMTs).
4. Large (small) diagonal elements correspond to “interesting dynamics” (noise).
5. Large (small) off-diagonal elements correspond to high (low) redundancy.

To maximize the *SNR* while minimizing redundancy, then, the solution is obvious: we want to diagonalize  $\hat{C}_X$ . The transformation which does this is

$$\hat{Y} = \hat{P}\hat{X} \tag{4.9}$$

---

<sup>5</sup> [Shlens \(2014\)](#)

where  $\hat{X}$  is the data matrix in the “feature” or “measurement” space, and  $\hat{Y}$  is  $\hat{X}$  translated into “principal component” space or “pattern” space.  $\hat{P}$  is a square  $N_D$ -dimensional translation / rotation matrix (called the “transformation matrix”) which relates the feature space to the pattern space.

To find  $\vec{P}$ , following the procedure outlined in [Richardson \(2009\)](#), one starts by forming the covariance matrix of the transformed data matrix,  $\hat{Y}$ :

$$\hat{C}_Y \equiv \frac{1}{N_E} \hat{Y} \hat{Y}^T \quad (4.10)$$

with some substitution and matrix manipulation, this becomes:

$$\hat{C}_Y = \frac{(\hat{P}\hat{X})(\hat{P}\hat{X})^T}{N_E} = \frac{(\hat{P}\hat{X})(\hat{X}^T\hat{P}^T)}{N_E} = \frac{\hat{P}(\hat{X}\hat{X}^T)\hat{P}^T}{N_E} = \frac{\hat{P}\hat{S}\hat{P}^T}{N_E} \quad (4.11)$$

where

$$\hat{S} = \hat{X}\hat{X}^T \quad (4.12)$$

Note that, since  $\hat{S}^T = (\hat{X}\hat{X}^T)^T = (\hat{X}^T)^T\hat{X}^T = \hat{X}\hat{X}^T = \hat{S}$ ,  $\hat{S}$  is symmetric. Since  $\hat{S}$  is also square ( $N_D \times N_D$ ), as a symmetric square matrix, it must be orthogonally diagonalizable.<sup>6</sup> Let

$$\hat{S} = \hat{E}\hat{D}\hat{E}^T \quad (4.13)$$

where, as [Richardson \(2009\)](#) defines  $\hat{E}$  (with my notation convention) as an

“ $N_D \times N_D$  orthonormal matrix whose columns are the orthonormal eigenvectors of  $\hat{S}$ , and  $\hat{D}$  [as] a diagonal matrix which has the eigenvalues of  $\hat{S}$  as its (diagonal) entries.”

---

<sup>6</sup> As shown in [Shlens \(2014\)](#).

Note that this procedure is not specific to [Richardson \(2009\)](#), but is common throughout the literature.<sup>7,8</sup> Finally, we can define our transformation matrix,  $\hat{P}$  such that the rows of  $\hat{P}$  comprise the eigenvectors of  $\hat{S}$ . The effect of this is that

$$\hat{C}_Y = \frac{\hat{P}\hat{S}\hat{P}^T}{N_E} = \frac{\hat{E}^T(\hat{E}\hat{D}\hat{E}^T)\hat{E}}{N_E} = \frac{(\hat{E}^T\hat{E})\hat{D}(\hat{E}^T\hat{E})}{N_E} = \frac{\hat{D}}{N_E} \quad (4.14)$$

$\hat{C}_Y$  is diagonalized! Just by defining the transformation matrix,  $\hat{P}$ , in this particular way, we're able to transform  $\hat{X}$  to  $\hat{Y}$  in such a way that the covariance matrix,  $\hat{C}_Y$  is diagonalized. This minimizes the off-diagonal elements, reducing redundancy, and maximizes the diagonal elements, concentrating the “information” encoded in the principal components.

Step-by-step, to perform the PCA, one must:

1. calculate the eigenvectors of  $\hat{S} = \hat{X}\hat{X}^T$ , and arrange them in order of decreasing eigenvalue
2. set the eigenvector with the highest eigenvalue as the first column of  $\hat{E}$ , the eigenvector with the second-highest eigenvalue as the second column of  $\hat{E}$ , and so on
3. set the diagonal elements of  $\hat{D}$  to (from top-left to bottom-right) the eigenvalues of  $\hat{S}$ , in decreasing order, such that the  $j$ -th diagonal element of  $\hat{D}$  corresponds to the  $j$ -th column of  $\hat{E}$ . Since  $\hat{D}$  is a diagonal matrix, every off-diagonal element of  $\hat{D}$  is 0.
4. finally, construct  $\hat{P} = \hat{E}^T$

Now, for a given set of  $N_E$  measurements of dimensionality  $N_D$ , one can easily define  $\hat{X}$  as an  $N_D \times N_E$  matrix, where each column corresponds to a single event,

---

<sup>7</sup> [Shlens \(2014\)](#)

<sup>8</sup> [Smith \(2002\)](#)

and each row corresponds to a set of simultaneous measurements of the system.<sup>9</sup> One can then convert from feature space to pattern space with the transform  $\hat{Y} = \hat{P}\hat{X}$ , where  $\hat{Y}$  is also  $N_D \times N_E$ . And the diagonalized covariance matrix  $\hat{C}_Y$  can be calculated using Equation 4.10.

Transforming  $\hat{X}$  to  $\hat{Y}$  according to Equation 4.9 presents the data in a new  $N_D$ -dimensional space, where the variance is greatest along the first coordinate axis and smallest along the  $N_D$ -th axis (since the largest eigenvalue of  $\hat{S}$  is chosen to be the first column of  $\hat{E}$  and the smallest eigenvalue of  $\hat{S}$  is chosen to be the last column of  $\hat{E}$ ). And thus, we have accomplished our goal: representing the greatest amount of “interesting information” (ie. variance) of a large set of multi-dimensional data in the fewest dimensions of a given basis. In the next Subsection, this procedure is applied to real DS-50 data.

### 4.1.2 As Applied to DarkSide-50

For an  $xy$ -Reconstruction in DS-50, the input data are points in a 19-dimensional space, where the magnitude of the  $j$ -th component of the  $i$ -th event vector,  $\vec{X}_i$ , is usually given as the mean-centered logarithm of the anode S2 light fraction (sometimes the total S2 light fraction) seen by the  $j$ -th anode PMT,  $L_i^{(j)}$ , for  $j = 1..N_D$ ,  $N_D = 19$ , and  $i = 1..N_E$ . So each TPC event in the detector gives a 19-dimensional vector  $\vec{X}_i$ , as shown in Equation 4.1.

Following the procedure outlined in the previous Subsection, one can calculate  $\hat{X}$ , where  $\vec{X}_i$  is the  $i$ -th column of  $\hat{X}$ , and then find  $\hat{S} = \hat{X}\hat{X}^T$  in a very straightforward matrix calculation. Then, the eigenvectors and eigenvalues of  $\hat{S}$  can be calculated,<sup>10</sup> and arranged as prescribed to form  $\hat{D}$  and  $\hat{E}$ . Once this is done, the

---

<sup>9</sup> Or, similarly, each column corresponds to a given measurement vector, and each row corresponds to a given element of that vector.

<sup>10</sup> Though I won't go into the details of calculating the eigenvalues of a 19-dimensional square matrix here.

transformation matrix  $\hat{P} = \hat{E}^T$ , which converts the feature space vectors  $\vec{X}_i$  into pattern space vectors  $\vec{Y}_i$ , can be defined.<sup>11</sup>

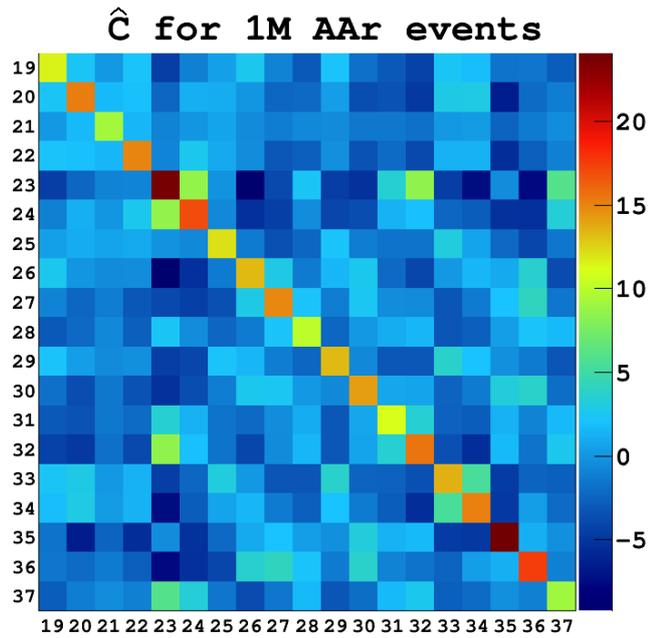
Note that once  $\hat{P}$  has been calculated, we can extend our analysis to include events which weren't a part of the original set. The set of events used to create  $\hat{P}$  is called the “training sample”, and any additional sets of events not included in the training are called “operative samples”. It is assumed that the relationships between the elements of the training events also hold for the operative events — that is, the training and operative sets are both samples of some larger population. This assumption is critical when performing the MDF (see Section 4.3). Note that, while not strictly necessary, I use the same sample of events to construct the matrix  $\hat{C}_Y$  as I do to calculate the MDF; accordingly, I consider both of these steps as parts of the “training”.

A graphical representation of the un-diagonalized covariance matrix,  $\hat{C}_X$ , for the first one million AAr events, can be seen in Figure 4.1 (a). Note that the elements along the diagonal (that is, the variances) overwhelm the other colors of the plot, so the diagonal is set to zero in Figure 4.1 (b). These matrices look quite different from the corresponding ones for one million g4ds events in Figure 4.2, but once the CX analysis cuts are applied (Figure 4.3), the matrices look quite similar. The g4ds events in the plots in this Chapter are one million uniformly-distributed (in  $xyz$ , throughout the  $\ell$ Ar) ER events with S1 values roughly within the DM range (about  $0 < S1 \text{ [PE]} < 600$ ).

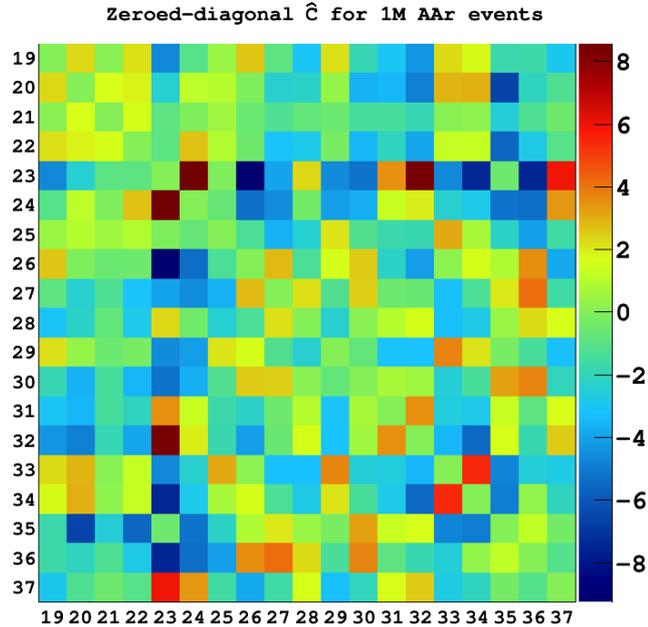
Recall the meaning of large (strictly non-negative) elements along the diagonal and off the diagonal (large negative values or large positive values). Large diagonal elements mean that there are “interesting dynamics”<sup>12</sup> occurring within the system

<sup>11</sup> Note that multiplying an  $N_D \times N_E$  matrix by its transpose — where  $N_E$  may be millions of events — is very memory intensive. Rather than store all of this in RAM, the covariance matrix is often calculated “on the fly” through various methods. See <https://root.cern.ch/doc/master/classTPrincipal.html> for instance.

<sup>12</sup> Richardson (2009)

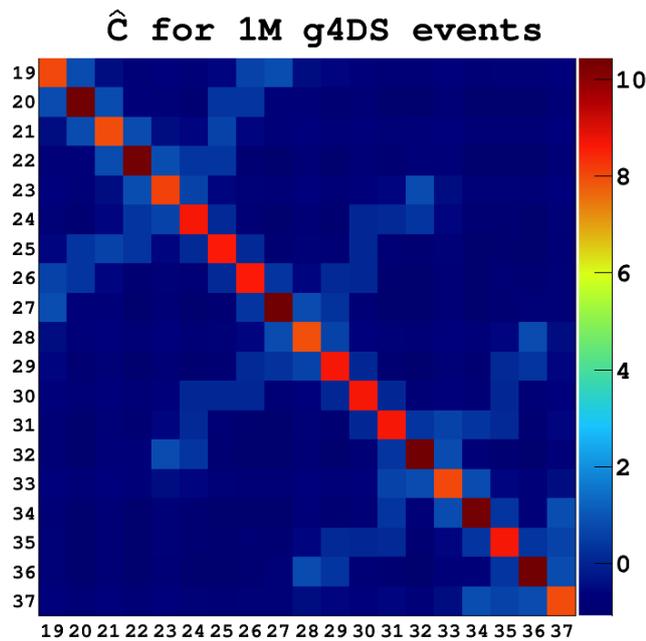


(a) Color-coded un-diagonalized covariance matrix  $\hat{C}$  for the first one million 50-Days AAr events. The  $z$ -axis gives the numerical variance or covariance and has been scaled by 1000 for readability. Redder cells indicate high variance (high information content) along the diagonal and high covariance (high redundancy) for the off-diagonal elements.

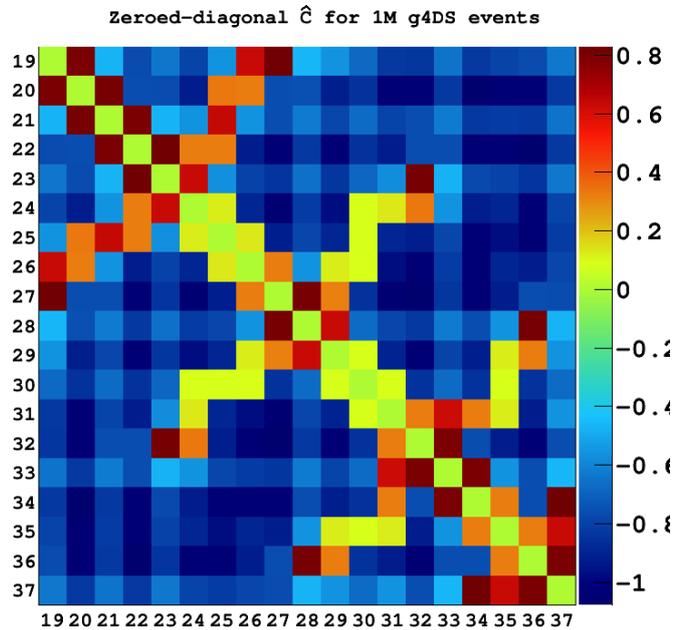


(b) The same covariance matrix as in (a), but with the diagonal elements set to zero to visually enhance the changes in covariance throughout the matrix.

Figure 4.1: The covariance matrix of the first one million 50-Days AAr events. Color figure available online.

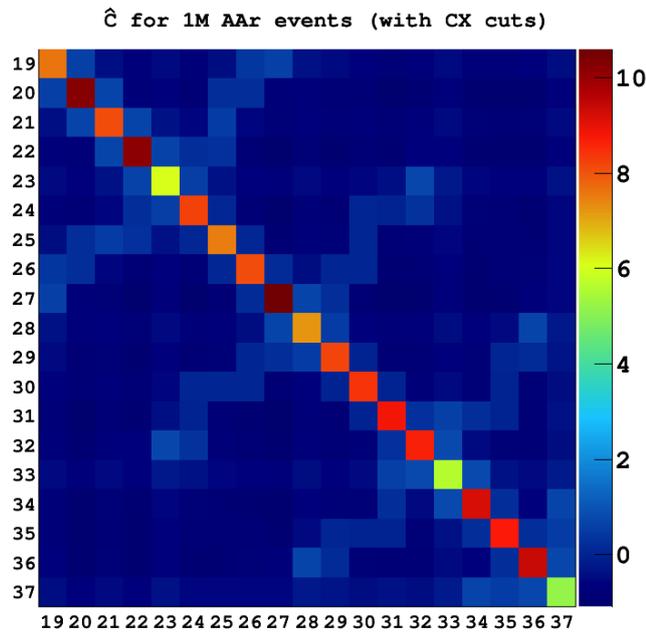


(a) Color-coded un-diagonalized covariance matrix  $\hat{C}$  for one million uniformly-generated g4ds events with S1 values in the DM range. The  $z$ -axis gives the numerical variance or covariance and has been scaled by 1000 for readability. Redder cells indicate high variance (high information content) along the diagonal and high covariance (high redundancy) for the off-diagonal elements.

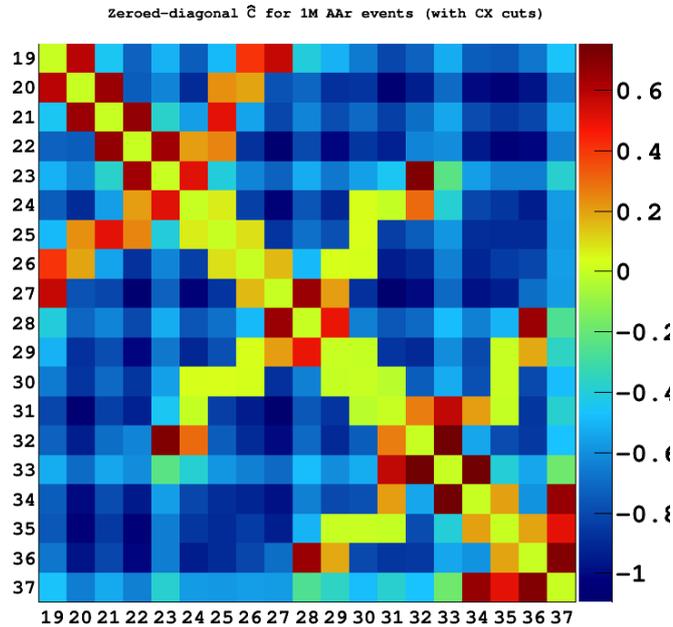


(b) The same covariance matrix as in (a), but with the diagonal elements set to zero to visually enhance the changes in covariance throughout the matrix.

Figure 4.2: The covariance matrix for one million uniformly-generated g4ds events with S1 values in the DM range. Color figure available online.



(a) Color-coded un-diagonalized covariance matrix  $\hat{C}$  for the first one million 50-Days AAr events, with all CX analysis cuts applied. The  $z$ -axis gives the numerical variance or covariance and has been scaled by 1000 for readability. Redder cells indicate high variance (high information content) along the diagonal and high covariance (high redundancy) for the off-diagonal elements.



(b) The same covariance matrix as in (a), but with the diagonal elements set to zero to visually enhance the changes in covariance throughout the matrix.

Figure 4.3: The covariance matrix of the first one million 50-Days AAr events, with all CX analysis cuts applied. Color figure available online.

for those measurements. All of our measurements in DS-50 are S2 light fractions seen by the anode PMTs. PCA is telling us that, in g4ds (Figure 4.2 (a)), all of the PMTs are providing some amount of information, but in data (Figure 4.3 (a)), PMTs 23, 33, and 37 (the 5th, 15th, and 19th diagonal elements) have smaller variances and could be providing less useful information than the other anode PMTs, assuming that the noise is more or less constant across all PMTs. These diagonal elements are green-ish, meaning that the  $S2[\ ]$  variance on these PMTs is correspondingly smaller. This is very interesting, considering that, before CX cuts were applied, PMT 23 had the largest variance of any anode PMT (Figure 4.1 (a)).

Large off-diagonal elements mean we have redundancies in our measurements.<sup>13</sup> Look at the top row of Figure 4.3 (b): the reddish-orange off-diagonal elements correspond to PMTs 20, 26, and 27. These are all PMTs which surround PMT 19 (see Figure 3.4). The same goes for the second row, where the large off diagonal elements correspond to PMTs 19, 21, 25, and 26: all PMTs which surround PMT 20. These measurements are slightly redundant because these PMTs are near each other in the detector; if an event occurs close to one, it occurs close to the others, and so all of these PMTs are likely to see large fractions of the S2 light. If an event occurs far away from one of these PMTs, it occurs far away from the others, and so all see little S2 light. In this sense, these measurements are correlated and there is some dimensional redundancy in our measurements.

A helpful way of visualizing the variance and covariance of S2 fractions between

---

<sup>13</sup> Not in the sense of noise or useless information, but rather in the sense of redundant dimensions. If PMT “A” and PMT “B” are on opposite sides of the detector, localizing an event along the line connecting A and B requires knowledge of the S2 fractions of both A and B. If  $S2^A$  and  $S2^B$  are roughly equal, then the event should be about halfway between the two PMTs. If  $S2^A$  is larger, the event should be nearer PMT A and vice versa. However, the ratio of  $S2^A/S2^B$  provides the same information and reduces the dimensionality of the problem by 1.

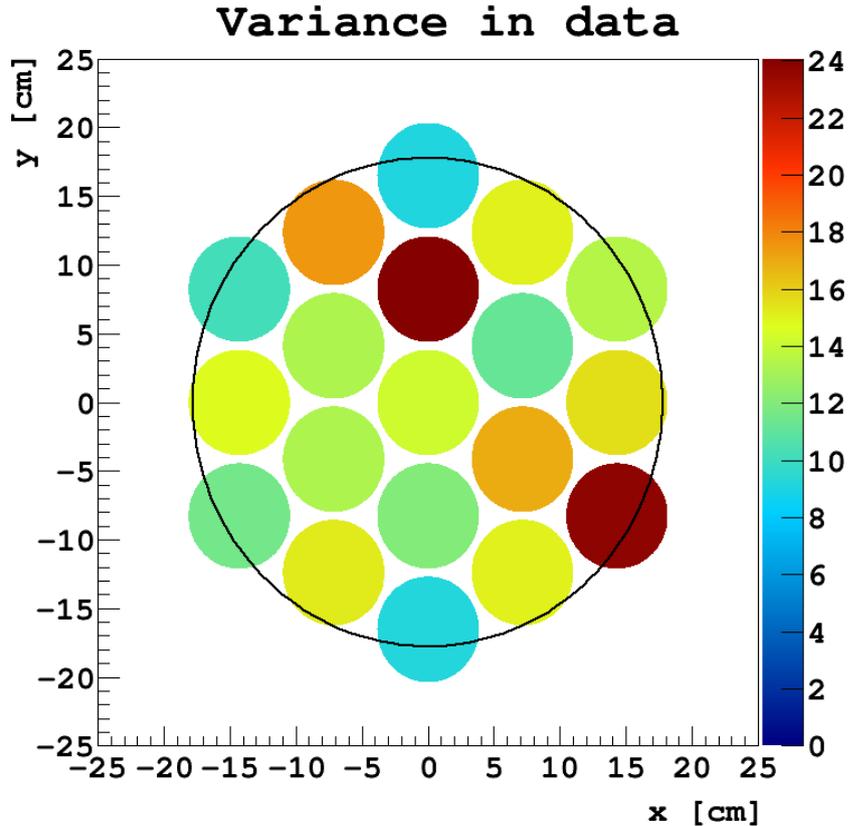


Figure 4.4: Color-coded  $S2[\ ]$  variance over the anode PMTs for the first one million AAr events. The  $z$ -axis gives the variance and has been scaled by 1000 for readability. Redder PMTs have higher  $S2[\ ]$  variance. Color figure available online.

the anode PMTs is to plot them as shown in Figure 4.4. Here, we see the variance in  $S2$  fractions on each PMT plotted on top of that PMT's physical location in the detector. We can compare this to the same plot for  $g4ds$  events in Figure 4.5 and the first one million AAr events with all  $CX$  cuts applied in Figure 4.6.

Note that, before analysis cuts, there is a huge discrepancy between data (Figure 4.4) and  $g4ds$  (Figure 4.5). PMTs 35 (just above the central PMT) and 23 (at about 4 o'clock) have much larger variances than any other anode PMT. The MC, however, seems much more predictable: the only thing which seems to affect the variance of the  $S2$  fractions on a given PMT in  $g4ds$  is that PMT's relative position within the detector. The central PMTs all have very similar variances,

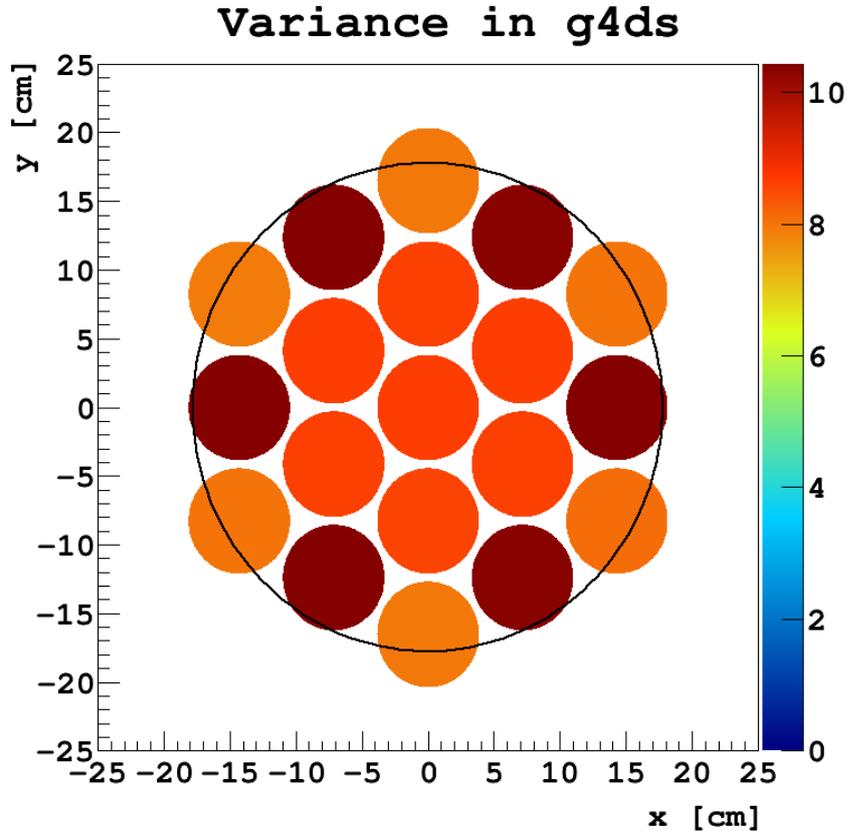


Figure 4.5: Color-coded  $S2[]$  variance over the anode PMTs for one million g4ds events. The  $z$ -axis gives the variance and has been scaled by 1000 for readability. Redder PMTs have higher  $S2[]$  variance. Color figure available online.

while the outer-ring “corner” PMTs (19, 21, 23, 28, 33, and 37) all seem to have a slightly lower variance and the outer-ring “side” PMTs (20, 22, 27, 32, 34, and 36) all have a slightly higher variance.

Applying the CX cuts to data (Figure 4.6) gives us a PMT variance map which seems to look more with g4ds, but still has some disagreements. In particular, all of the outer-ring PMTs at the “northeast” end of the detector have smaller variances than in g4ds. This suggests that there is something different about the response of these PMTs to the S2 light of an event. This discrepancy will be investigated further in the next Chapter.

Finally, a note about nomenclature: while the dimensions of the original fea-

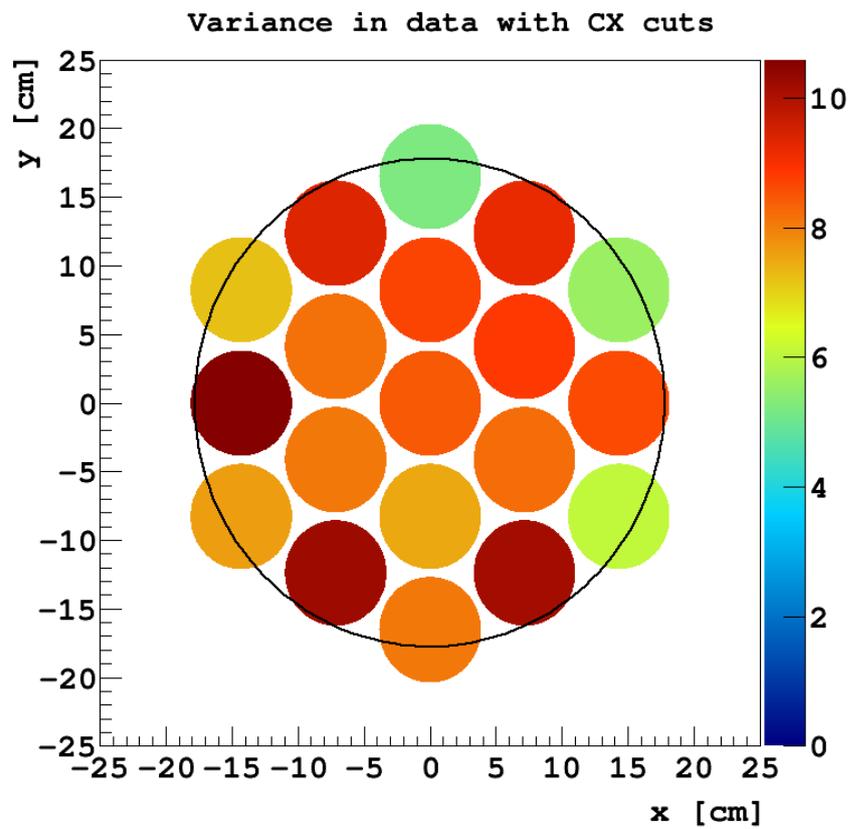


Figure 4.6: Color-coded  $S2[]$  variance over the anode PMTs for the first one million AAr events with all CX analysis cuts applied. The  $z$ -axis gives the variance and has been scaled by 1000 for readability. Redder PMTs have higher  $S2[]$  variance. Color figure available online.

ture space have no special names, the dimensions of the pattern space, as the eigenvectors of  $\hat{S}$ , are sometimes called simply “the eigenvectors” or the “principal components” (the PCs). These names will be used interchangeably in the coming sections.

## 4.2 Reduction of Dimensionality

When a significant amount of the variance lies along just a few principal components, the dimensionality of the dataset can be reduced with a minimal loss of information (variance). In DS, however, such a reduction in dimensionality is not feasible.

The eigenspectrum plot in Figure 4.7 (a) quantifies the variance contained along all 19 PCs in both AAr and g4ds. Generally, 70% to 80% of the variance must be accounted for to locate outliers in a reduced-dimensionality dataset. That is, if a few of the smaller-eigenvalue principal components are ignored to reduce the dimensionality of the data, the remaining PCs must contain at least 70% to 80% of the variance from the original dataset in order for one to reliably find outliers. Figure 4.7 (b) shows that, in g4ds and AAr data with all CX cuts applied, at least 11 PCs are necessary for finding outliers.<sup>14</sup>

However, image compression using PCA typically requires that  $\gg 95\%$  of the variance be accounted for to generate a suitable image.<sup>15</sup> Since reconstructing  $xy$  event positions is, in some sense, an imaging process, Figure 4.7 (b) suggests that we can lose no more than a few dimensions before distortions start to become evident in the reconstruction. That being said, PCA still concentrates the maximum amount of variance in the fewest number of dimensions possible, making it ideal

---

<sup>14</sup> The eigenvector numbering starts at 0, so eigenvector #10 is the 11th eigenvector.

<sup>15</sup> [Richardson \(2009\)](#)

for coupling with a multi-dimensional fit.<sup>16</sup>

### 4.3 Multi-Dimensional Fitting

Finding outliers and characterizing trends in data is useful, but the goal of this dissertation is to reconstruct the  $xy$  positions of events in DS-50. How is this accomplished with PCAMDF?

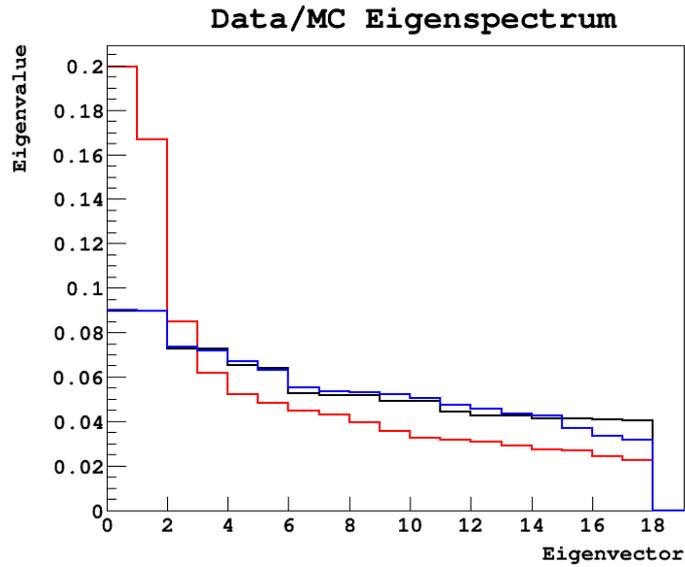
There are four steps in this process:

1. generate the transformation matrix  $\hat{P}$  by constructing  $\hat{C}_Y$  with simulated MC events
2. using MC events, develop a multi-dimensional fit (MDF) between the dependent variables (the  $x$  and  $y$  positions) and the independent variables (the pattern space variables)
3. transform the data vectors from “feature space” to “pattern space” using  $\hat{P}$
4. use the MDF constructed in (2) to calculate the dependent variables ( $x$  and  $y$ ) from the independent pattern-space variables calculated in (3)

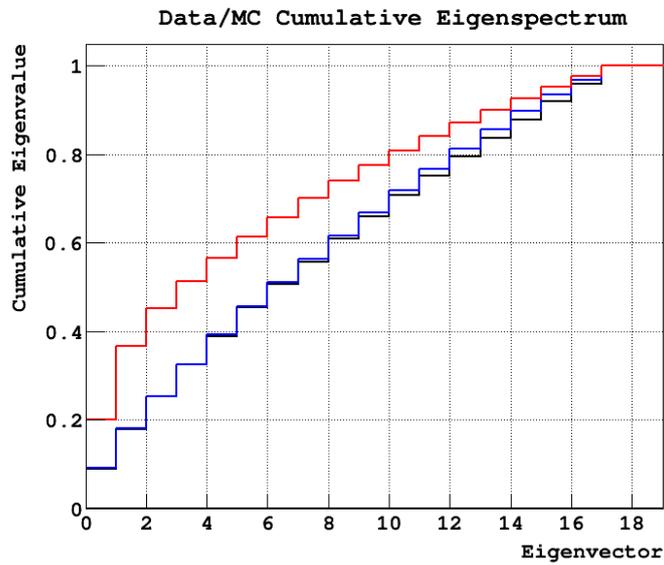
Steps (1) and (3) are discussed in the first Section of this Chapter, and step (2) is straightforward. In low-dimensionality spaces, fitting a dependent variable to independent variables is easy. Given a scatterplot of coordinate pairs in the Cartesian plane, for instance, one chooses a suitable fit function, then varies the parameters of that function until some convergence condition is met (usually that

---

<sup>16</sup> Additionally, it has been suggested that performing a separate PCAMDF for each of the 19 anode PMTs (where the events included in the training and operative samples for each PCAMDF must have their maximum  $S2$  on the given PMT) could lead to a greater variance in the magnitudes of the eigenvalues, which in turn might lead to better outlier rejection. This is currently under investigation.



(a) Eigenvalue as a function of eigenvector index for AAr data (red), AAr data with all CX cuts applied (blue) and MC (black). Note how applying the standard analysis cuts brings the data into much better agreement with the MC.



(b) Cumulative eigenvalue as a function of eigenvector index for AAr data (red), AAr data with all CX cuts applied (blue) and MC (black). Neither data nor MC hit 70% until at least eigenvector # 10.

Figure 4.7: Eigenspectrum plots for AAr data and g4ds.

the fit doesn't improve — reduce the  $\chi^2$  — significantly with a change of parameters). This is a significant part of the “training” step of the PCAMDF and it is necessarily performed on simulated MC events, where the true event positions are known.

There are a few sticking points here, however. First, it is assumed that a suitable functional form for the fit is known. With complicated data, and especially in higher-dimensional space, it's non-trivial to figure out what form a fit function should take. Second, *overfitting* of data must also be avoided. Although 10 points in two dimensions can be perfectly fitted with a 9th order polynomial, it's likely that a much lower order fit would follow the data nearly as well, and be generalizable to new data not contained within the “training” set. These concepts will be touched on in the following Chapter.

For this work, the ROOT classes `TPrincipal`<sup>17</sup> for the PCA and `TMultiDimFit`<sup>18</sup> for the MDF have been used. A complete explanation of the least squares fitting performed by `TMultiDimFit` is beyond the scope of this text, but can be found on the ROOT Reference Guide website, <https://root.cern.ch/>.

Lastly, note that, while PCA is non-parametric, MDF does require some “hand-tuning” by the user. MDF uses either monomials, Legendre polynomials, or Chebyshev polynomials to find the least-squares fit to a dataset.<sup>19</sup> For a given (independent variable, dependent variables(s)) pair

$$(x_i, \vec{X}_i) \tag{4.15}$$

---

<sup>17</sup> <https://root.cern.ch/doc/master/classTPrincipal.html>

<sup>18</sup> <https://root.cern.ch/doc/master/classTMultiDimFit.html>

<sup>19</sup> In principle, these should all work equally well, as they each form an orthonormal basis and can be used to approximate any continuous function with arbitrarily high accuracy. As shown in [Appendix E](#), Chebyshev polynomials are used for the PCAMDF described in this dissertation, because they appeared to perform best in tests on MC events.

where  $x_i$  is the dependent quantity (either the  $x$  or  $y$  position of the MC training event), `TMultiDimFit` attempts to minimize the sum

$$S \equiv \sum_{i=1}^{N_E} \left( x_i - x'_i \left( \vec{X}_i \right) \right)^2 \quad (4.16)$$

where  $x_i$  is the true coordinate of the event and  $x'_i$  is the fitted position.  $x'_i$  is defined according to the parameterization

$$x'_i \left( \vec{X}_i \right) = \sum_{m=1}^M c_m \prod_{j=1}^{N_D} p_{mj} \left( L_i^{(j)} \right) = \sum_{m=1}^M c_m F_m \left( \vec{X}_i \right) \quad (4.17)$$

where  $c_m$  are constants;  $p_{mj}$  are the monomials, Legendre polynomials, or Chebyshev polynomials;  $L_i^{(j)}$  are the elements of the  $i$ -th  $N_D$ -dimensional independent variable vector,  $\vec{X}_i$ ; and of course, we've defined

$$F_m \left( \vec{X}_i \right) = \prod_{j=1}^{N_D} p_{mj} \left( L_i^{(j)} \right) \quad (4.18)$$

As the ROOT User's Guide notes, it's unlikely that  $S$  will ever become exactly zero as a result of the fitting procedure, and so we decide our fit is sufficient when

$$R = \frac{S}{\sum_{i=1}^{N_E} x_i^2} < \epsilon \quad (4.19)$$

where  $\epsilon$  is a user-defined parameter. In effect, this allows the fitting procedure to quit when the fit achieves some desired fractional error. So the fitting procedure is parametric and does require user tuning. In particular, the following `TMultiDimFit`-namespace routines are called:

**SetMaxFunctions** This sets the maximum number of  $p_{mj}$  which can be included in the fit.

**SetMaxStudy** This sets the maximum number of  $p_{mj}$  to be "studied". If they

reduce the residual, they're kept; if not, they're skipped.

**SetMaxTerms** This sets  $M$  in Equation 4.17 to a user-defined  $M_{max}$ , limiting the number of polynomials or monomials which can be used in the fit.

**SetMaxPowers** This sets the maximum power,  $P_{max,j}$ , to which each variable,  $L_i^{(j)}$  can be raised during the minimization. This helps to avoid over-fitting the data.

**SetPowerLimit** This sets the “power control limit”  $Q$ , where  $Q$  is defined as

$$Q_m \equiv \sum_{j=1}^{N_D} \frac{P_{mj}}{P_{max,j}} < Q \quad (4.20)$$

where  $Q_m$  is the value calculated for the  $m$ -th function included in the fit, summed over all independent variables  $j$  (of  $N_D$  variables). This drastically reduces the number of functions,  $F_m$ , which can be used in the fit.

**SetMinRelativeError** This sets  $\epsilon$  in Equation 4.19 to a user defined value.

Determining the best parameters ( $M_{max}$ ,  $P_{max,j}$ ,  $Q$ , etc.) for these fits is non-trivial, and partially done by trial-and-error for a 19-dimensional data set. The parameters chosen are the ones which give the best performance on MC events, out of several hundred trials. These are:  $Q_x = 5$ ,  $Q_y = 2$ ,

$$P_{max,x} = \{3, 9, 5, 1, 9, 9, 4, 6, 2, 6, 8, 10, 2, 3, 6, 2, 9, 9, 9\} \quad (4.21)$$

and

$$P_{max,y} = \{8, 4, 1, 8, 6, 1, 9, 5, 1, 3, 4, 10, 2, 8, 8, 2, 9, 2, 10\} \quad (4.22)$$

The maximum number of functions, studies, and terms are all set to 5000 for both  $x$  and  $y$ . This increases the computation time, but allows the MDF

to find the best possible fit. Finally,  $\epsilon$  is set to 0.00001 (1/1000 %) for both  $x$  and  $y$ , again to leave the MDF essentially unbounded to find the best possible fit. (Note that 1/1000 % accuracy is never actually achieved, because different termination conditions are met — usually using the maximum number of terms — before that sort of accuracy can be achieved.) The following Chapter examines the performance of this particular PCAMDF tuning on both MC and data events.

# Chapter 5

## PCAMDF

### *xy*-RECONSTRUCTION

#### 5.1 Description of Event Sets

PCAMDF *xy*-Reconstruction requires at least two sets of events: a “training” set to calculate the transformation matrix  $\hat{P}$  and determine the fit function via MDF; and an “operative” set, which is taken from the same population of events as the training set, but is not used to calculate  $\hat{P}$  or determine the fit function. Reconstructing the *xy* positions of events in an operative MC set allows us to characterize the accuracy of the reconstruction itself, while by definition, all data (that is, detector) events are operative because their true *xy* positions are unknown. The biggest assumption made is that the training MC event set and the operative data event set are from the same population, that is, that the two sets of data “look alike”. This assumption will be investigated in the coming Sections. For the studies performed in this thesis, five DS-50 data event sets and two g4ds event sets are used.

### 5.1.1 Data

DS-50 has, to date, published the results of two DM search campaigns: a 50-day exposure with atmospheric argon<sup>1</sup> (called the “50-Days” AAr dataset) and a 70-day exposure with underground argon<sup>2</sup> (the “70-Days” UAr dataset). Before (after) the standard CX analysis cuts are applied, the AAr event set contains  $7.37 \times 10^7$  ( $1.33 \times 10^7$ ) events, and the UAr event set contains  $9.93 \times 10^6$  ( $1.14 \times 10^5$ ) events. Note that even though the exposure of the UAr campaign ( $2616 \pm 43$  kg · day) is much larger than that of the AAr campaign ( $1422 \pm 67$  kg · day), the UAr event set has less than 1% of the number of events as the AAr event set, after cuts. This is a testament to the extremely low  $^{39}\text{Ar}$  background afforded by the use of UAr (see Subsection 3.2.2).

Because AAr is dominated by the  $^{39}\text{Ar}$  background, which is spread evenly throughout the  $\ell\text{Ar}$  volume, one might expect that when a large sample of AAr events are reconstructed with the PCAMDF, the  $xy$  distribution should be “flat” — that is, uniformly distributed in the  $xy$ -plane. This assumption is tested in Subsection 5.5.3.

Additional data sets are used to judge the performance of the PCAMDF  $xy$ -Reconstruction on data. Coincident decays can be used to quantify the precision of the  $xy$ -Reconstruction; there are two such data sets currently available in DS-50, both from the  $^{226}\text{Ra}$  decay chain:  $^{222}\text{Rn}^{218}\text{Po}$  decays (“RnPos”) and  $^{214}\text{Bi}^{214}\text{Po}$  decays (“BiPos”). Both of these types of coincident decays are investigated in Subsection 5.6.2. Finally, the fifth set of data events used is from a series of source calibration runs with a  $^{57}\text{Co}$  source, and is used to judge the angular accuracy of the PCAMDF reconstruction near the edge of the detector. This data set will be discussed further in Subsection 5.6.3.

---

<sup>1</sup> Agnes et al. (2015)

<sup>2</sup> Agnes et al. (2016c)

### 5.1.2 Monte Carlo

Two sets of g4ds MC events are used for this dissertation, a “high PE statistics” (“high stats”) set and a “low PE statistics” (“low stats”) set. Both consist of uniformly generated (throughout the  $\ell$ Ar) ER events, though the high stats ones are given an unphysically high 1800 keV of kinetic energy (with  $S1_{\text{tot}}$ s distributed mostly between 9000 PE and 12000 PE)<sup>3</sup>, while the low stats ones are uniformly distributed between 0 keV and 100 keV (with  $S1_{\text{tot}}$ s mostly between 0 PE and 600 PE). The energy range of the high stats MC event set was chosen in order to judge the accuracy of the PCAMDF  $xy$ -Reconstruction under optimal conditions (extremely high PE statistics), while the energy range of the low stats set was chosen to be roughly in agreement with the DS-50 DM S1 range. Data events are reconstructed using a PCAMDF which was trained on this low PE statistics MC event set.

Minimal cuts are made to the MC events, and these are mainly “data quality” cuts: ( $S1_{\text{tot}} > 0$ ), ( $S2_{\text{tot}} > 0$ ),  $! \text{IsNaN}(S2[n])$  (which checks for data corruption in the S2 fractions), etc. Any cuts applied beyond these will be specified in the text. Before any cuts are applied, the high statistics set contains  $4.41 \times 10^5$  events and the low statistics set contains  $1.76 \times 10^7$  events.

## 5.2 Reconstruction Code

When implemented in C++, using the ROOT data analysis framework, PCAMDF is quite straightforward. The entire PCAMDF algorithm consists of just a few main steps:

1. loop over all MC training events and fill the covariance matrix

---

<sup>3</sup> About 90% of these events have  $S2_{\text{tot}} > 1 \times 10^6$  PE, with a peak around  $1.09 \times 10^6$  PE.

2. perform the PCA on those MC training events
3. loop over the same events as step (1) (or different events from the same population) and add (dependent, independent) variable pairs,  $(x_i, \vec{Y}_i)$ , to the `TMultiDimFit` object (one for  $x$  and one for  $y$ ), again, for MC events
4. construct the MDFs (again, one for  $x$  and one for  $y$ )
5. use the PCAMDF generated in steps (1) through (4) to reconstruct data events

A minimal working example (MWE), which reads in an external ROOT g4ds event file, can be found in its entirety (without line numbers, for easy copying-and-pasting) in [Appendix E](#). The entire PCAMDF can be performed in about 200 lines of C++ code. I've broken it into steps, below, to explain how the code works:

```

17  #include "TFile.h"
18  #include "TTree.h"
19  #include "TPrincipal.h"
20  #include "TMultiDimFit.h"
21  #include "TMath.h"
22
23  TPrincipal pca(19, "ND");
24
25  TMultiDimFit* MDFx
26    = new TMultiDimFit(19, TMultiDimFit::kChebyshev, "V");
27
28  TMultiDimFit* MDFy
29    = new TMultiDimFit(19, TMultiDimFit::kChebyshev, "V");
30
31  float x, y, S2tot, S2top, S2[38];
32  double Xi[19], Yi[19];
33  float xprime, yprime;
34  int ii(0), ll;
35
36  void setUpMDFs();
37  bool applyCuts();
38  void constructXi();

```

Lines 17–21 are ROOT header files which allow for, respectively: reading in an external ROOT file, reading the elements from a `TTree` data structure stored in that file, performing the PCA, performing the MDF, and executing some simple

mathematics (namely, logarithms). Lines 23–29 define the `TPrincipal` object, used for performing the PCA, and the two `TMultiDimFit` objects, used to calculate the fit functions  $x'_i(\vec{Y}_i)$  and  $y'_i(\vec{Y}_i)$ . And lines 31–34 are some self-explanatory variable definitions (`ii` and `ll` are just counters).

Lines 36–38 define a few subroutines which make the code a bit more reusable. `setUpMDFs()` is a function which configures the two `TMultiDimFit` objects. `applyCuts()` is a function which returns `true` if the current event passes all data quality cuts and returns `false` otherwise. And `constructXi()` is a function which constructs the vector  $\vec{X}_i$  from the variables read in from the external data file. All event variables are defined globally for speed and ease of use.

```
47  int NE = 10000;
```

Line 47 simply sets the number of events,  $N_E$ , that we'd like to use to construct the covariance matrix. Note that this is not necessarily equal to the number of events given to the `TMultiDimFit` object to calculate the fit functions, though this is how the MWE is coded.

```
64  TFile* mcfile = new TFile("external_file.root");
65  TTree* mctree = (TTree*)(mcfile->Get("dstree"));
66  int nentries = mctree->GetEntries();
67
68  mctree->SetBranchAddress("x",      &x);
69  mctree->SetBranchAddress("y",      &y);
70  mctree->SetBranchAddress("S2tot",  &S2tot);
71  mctree->SetBranchAddress("S2top",  &S2top);
72  mctree->SetBranchAddress("S2",     S2);
```

Lines 64 and 65 read in the external ROOT file and the `TTree` object contained within it, and line 66 gets the total number of entries (MC events) written to the `TTree`. Lines 68–72 link the variables from line 31 to their corresponding elements in the `TTree`, so that when `TTree::GetEntry(i)` is called, they're populated with their correct values for event `i`. Note that `S2` is purposefully not preceded by an ampersand, because it's an array, while the other variables are not.

```

78 TFile* outfile = new TFile("minimal.root", "RECREATE");
79 TTree* outtree = new TTree("outtree", "minimal.C_output_TTree");
80
81 outtree->Branch("x",      &x,      "x/F");
82 outtree->Branch("y",      &y,      "y/F");
83 outtree->Branch("S2tot",  &S2tot,  "S2tot/F");
84 outtree->Branch("S2top",  &S2top,  "S2top/F");
85 outtree->Branch("S2",     S2,       "S2[38]/F");
86
87 outtree->Branch("xprime", &xprime, "xprime/F");
88 outtree->Branch("yprime", &yprime, "yprime/F");
89 outtree->Branch("Xi",     Xi,       "Xi[19]/D");
90 outtree->Branch("Yi",     Yi,       "Yi[19]/D");

```

Lines 78 and 79 define an output ROOT file, where the true  $(x_i, y_i)$  positions of the reconstructed MC events will be written, as well as their reconstructed positions  $(x'_i, y'_i)$ , in addition to some other variables. Lines 81–85 simply link the same variables to the output file which we read from the input file, while lines 87–90 link the reconstructed event positions, the independent variable vectors  $\vec{X}_i$  and those same vectors transformed to pattern space,  $\vec{Y}_i$ . Note that a /F (/D) at the end of one of these lines indicates that the variable being linked is of type float (double), and that — as on line 72 — S2, Xi, and Yi do not get ampersands, because they are arrays of values.

```

96 for(int i = 0; i < nentries; ++i){
97     if(ii >= NE) break; mctree->GetEntry(i);
98     if(!applyCuts()) continue;
99     constructXi();
100    pca.AddRow(Xi);
101    ++ii; ll = i;
102 } ii = 0;
103
104 // make P matrix, print results of PCA to screen
105 pca.MakePrincipals();
106 pca.Print();

```

Lines 96–102 are the first loop over the events in the MC file. First, in line 97, we check to make sure we still need events, and if we don't, we quit. If we do, we get the next event from the TTree and see if the event passes the cuts defined in applyCuts(). If it doesn't, we move to the next event, if it does, we construct  $\vec{X}_i$  by calling constructXi(). This whole procedure is performed for each of the

three event loops. After we calculate  $\vec{X}_i$ , we add it to the PCA. The first event loop breaks when `ii==NE`. After the first event loop, we perform the PCA on lines 105 and 106 and print the eigenvalues to the screen for the user.

```

112     for(int i = 0; i < nentries; ++i){
113         if(ii >= NE) break;  mctree->GetEntry(i);
114         if(!applyCuts()) continue;
115         constructXi();
116         pca.X2P(Xi, Yi);
117         MDFx->AddRow(Yi, x);
118         MDFy->AddRow(Yi, y);
119         ++ii;
120     } ii = 0;
121
122     // find fit functions for x and y
123     MDFx->FindParameterization();
124     MDFy->FindParameterization();

```

The second event loop (lines 112–120) is the same as the first up to and including line 115. On line 116, we call `TPrincipal::X2P()` to convert the measurement vector from feature space to pattern space ( $\vec{X}_i$  to  $\vec{Y}_i$  or `Xi` to `Yi`). This function was generated by ROOT when `TPrincipal::MakePrincipals()` was called on line 105. On lines 117 and 118, we add the pattern-space independent variable vector ( $\vec{Y}_i$ ) and the dependent variables ( $x$  and  $y$ ) to their respective `TMultiDimFit` objects (`MDFx` and `MDFy`). After the second event loop, the MDFs are constructed on lines 123 and 124, finding the best functions to map the independent variables to the dependent variables.

```

130     for(int i = nentries; i > 11; --i){
131         if(ii >= NE) break;  mctree->GetEntry(i);
132         if(!applyCuts()) continue;
133         constructXi();
134         pca.X2P(Xi, Yi);
135         xprime = MDFx->Eval(Yi);
136         yprime = MDFy->Eval(Yi);
137         outtree->Fill();
138         ++ii;
139     }
140
141     outfile->Write();

```

The final event loop (lines 130–139) is the same as the second up to and including line 134, except for the first line (130), which now loops *backward* over the events in the external file. This is to avoid using the same events in the training and the operative samples.<sup>4</sup> On lines 135 and 136, we calculate the reconstructed  $xy$  position by calling `TMultiDimFit::Eval()`, which evaluates the function generated by the MDF using given vector of independent variables (mapped to pattern space),  $Y_i$ , as input. Our job completed, we fill the output tree with the data from this event on line 137. Once the final event loop is finished, we write the output tree to the output file and quit. The additional subroutines declared on lines 36–38 are straightforward and can be found in [Appendix E](#).

The training procedure takes quite a bit of time when `TMultiDimFit::SetMaxTerms(size)`, `TMultiDimFit::SetMaxFunctions(size)`, and `TMultiDimFit::SetMaxStudy(size)` are given large values for `size` (greater than a few hundred), though of course, the more terms and functions we allow in the fit, and the more “studies” of functions we allow `TMultiDimFit` to perform, the better the final result will be, for the most part. Running the MWE with `NE = 10000` and `size = 8000` takes a little over an hour on a single 2.8 MHz AMD processor.

Note, however, that the training only needs to be performed once. Using the functions `TPrincipal::MakeCode()` and `TMultiDimFit::MakeMethod("MDF†")` (where  $\dagger$  is either  $x$  or  $y$ ), the results of the PCA and MDF are written to external files, where they can simply be read in for future event sets.<sup>5</sup> After the training has been performed once, event loops 1 and 2 can be skipped in their entirety for

---

<sup>4</sup> Note that this does not guarantee that overlap is avoided. The user must be careful in setting `NE`; possibly adding a check to make sure no overlap occurs.

<sup>5</sup> Because they can be thousands of lines long, the ROOT-generated files, `pca.C`, `MDFxMDF.cxx`, and `MDFyMDF.cxx` are not included in this thesis. However, when `TMultiDimFit` generates the MDF output files, it assumes the user has put together the corresponding header file. Those header files are included in [Appendix E](#).

future  $xy$ -Reconstructions. This makes a PCAMDF-based  $xy$ -Reconstruction *very* fast. Using the second script included in [Appendix E](#), `recononly.C`, ten million MC events are reconstructed in 4619.270 CPU seconds on the machine described above, working out to just over 0.46 ms/event.<sup>6</sup>

PCAMDF is simple to implement in C++ and very fast, after the initial training has been completed. In the next Section, the accuracy of the method on MC events is investigated.

### 5.3 Performance on MC events

To compare the performance of the PCAMDF on MC events with the high PE statistics and low PE statistics event sets, both sets were trained using the same number of events and the same MDF parameters (see the end of the previous Chapter).  $10^5$  events were then reconstructed from the end of each of these files to avoid overlap between the training and operative samples.

There are several ways to judge the performance of an  $xy$ -Reconstruction on a MC event set. The first, and most obvious, is to look at the distance from the true event position  $(x_i, y_i)$  to the reconstructed position  $(x'_i, y'_i)$ . In my code, I call this variable `drvec` or  $d\vec{r}$ , and it's defined as

$$d\vec{r} = \text{drvec} = \text{sqrt}(\text{pow}(x - \text{xprime}, 2) + \text{pow}(y - \text{yprime}, 2)) \quad (5.1)$$

or, equivalently

$$d\vec{r}_i = \sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2} \quad (5.2)$$

where `xprime` and `yprime` are the the reconstructed  $x$  and  $y$  positions,  $x'_i$  and  $y'_i$ , and `sqrt()` and `pow()` are the standard C++ functions.

---

<sup>6</sup>Note that this can include up to 6 `TMath` function calls per event (see code in [Appendix E](#)), and so this is a high estimate on the time required per event.

Figure 5.1 shows the overall  $d\vec{r}$  for the high statistics (a) and low statistics (b) g4ds event sets. The only cut made is that  $r_{\text{prime}} < 17.78$ , where  $r_{\text{prime}}$  is the reconstructed radius of the event and 17.78 is the radius of the DS-50 active volume, in centimeters. This is to remove outlying events, which are sometimes reconstructed very far from the active volume. Note that 99.1% of high statistics MC events reconstruct within the active volume, compared to 98.1% of low statistics events. The high statistics reconstruction clearly performs better overall, with a mean of 0.35 cm (and an RMS of 0.71 cm) compared to the low statistics reconstruction, with a mean of 0.68 cm (and an RMS of 1.14 cm).

If we want to use PCAMDF to calculate a fiducial volume cut in DS-50 we need to know how the reconstruction performs over  $xy$ , and as a function of radius. The performance of the high and low statistics reconstructions over  $xy$  (with  $r_{\text{prime}} < 17.78$ ) can be seen in Figure 5.2. There are two things to note here. First, the overall reconstruction is much better with the high statistics data set; the  $z$ -scale is the same on both plots, so they can be compared by just observing the difference in color between the two of them. Second, The low statistics plot performs worse overall, but it performs particularly poorly near the edge. Or rather, under the outer-ring PMTs.

This is particularly evident when looking at the performance as a function of radius, as in Figure 5.3. In subfigure (a), we see that the high statistics reconstruction performs well, independent of the value of  $r$ . The low statistics reconstruction, though, has a performance which is strongly radius-dependent: it's worst near the outside edge of the TPC (near  $r=17.78$ ), but performs almost as well as the high statistics reconstruction between about  $r=4$  and  $r=10$ .

PCAMDF is sensitive to the scaling of the variables. Since all energy measurements in DS are made in PE, this doesn't seem like it would be a problem, but the statistical fluctuations on the elements of the S2 fraction array  $S2[]$  are also

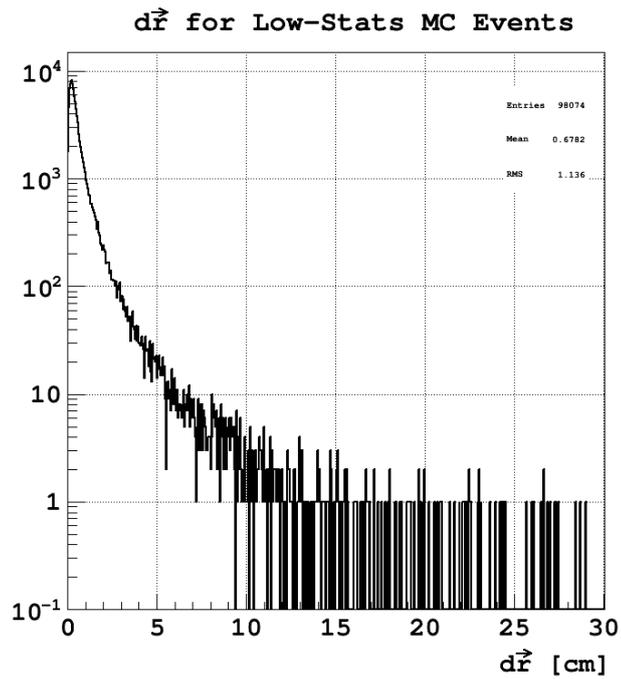
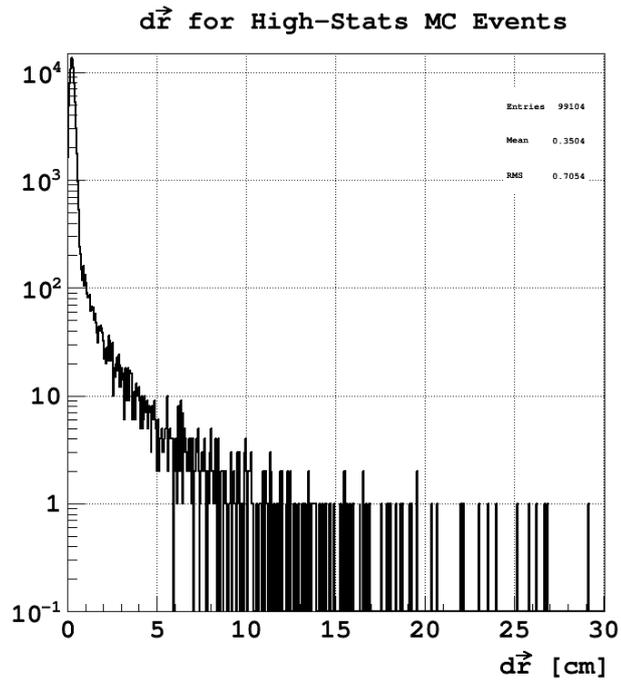
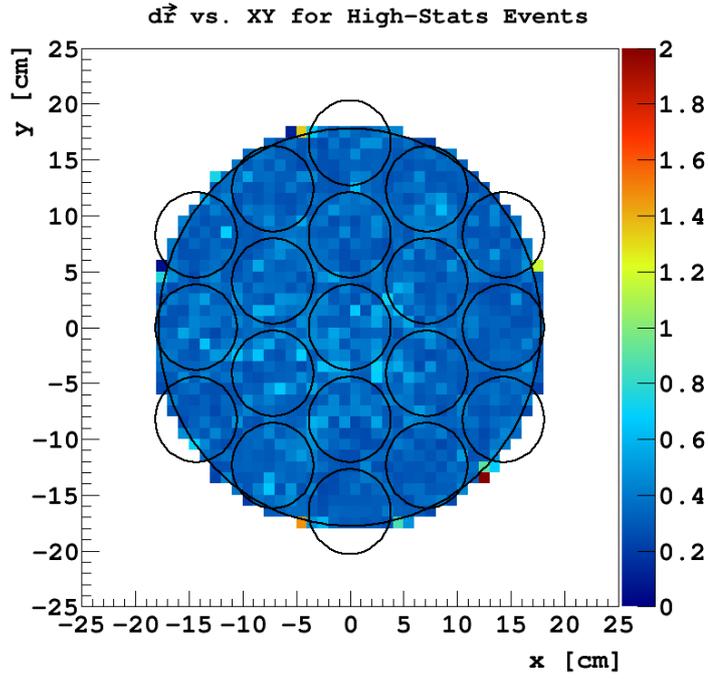
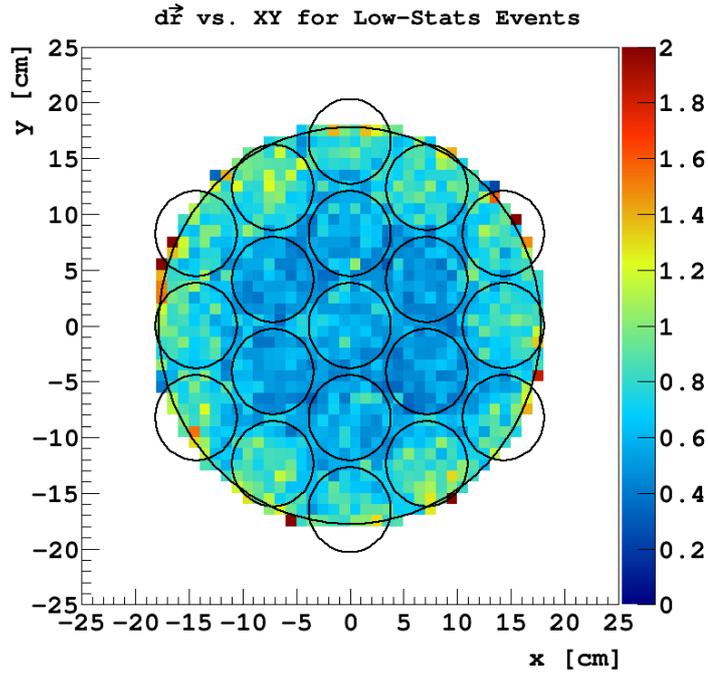


Figure 5.1: PCAMDF reconstruction accuracy for high (a) and low (b) PE statistics g4ds events.  $d\vec{r}$  is the distance from the true  $xy$  position of the MC event to its reconstructed position. The high statistics reconstruction has a mean  $d\vec{r}$  of 0.35 cm (with an RMS of 0.71 cm), and the low statistics reconstruction has a mean  $d\vec{r}$  of 0.68 cm (and an RMS of 1.14 cm).

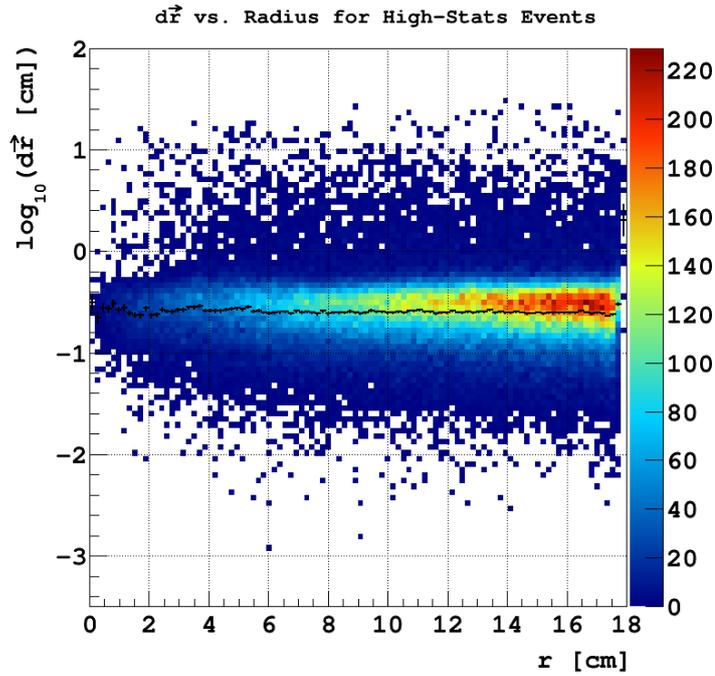


(a)

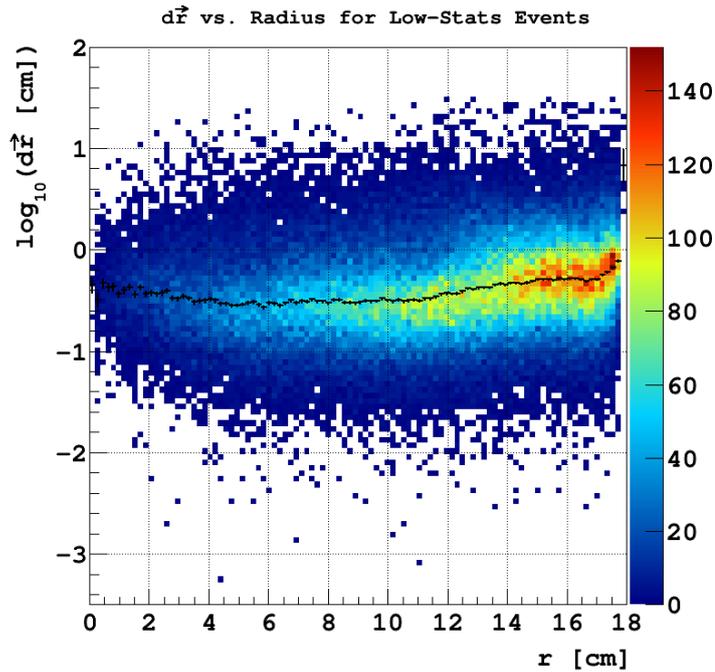


(b)

Figure 5.2: PCAMDF  $xy$ -Reconstruction accuracy mapped onto  $xy$  for high (a) and low (b) PE statistics g4ds events. The  $z$ -value of each  $xy$  bin is the average reconstruction error,  $\langle d\vec{r} \rangle$ , for events in that bin. Note that the high statistics MC event set performs better overall, and that the low statistics set performs worse near the edge of the TPC, compared to the center.

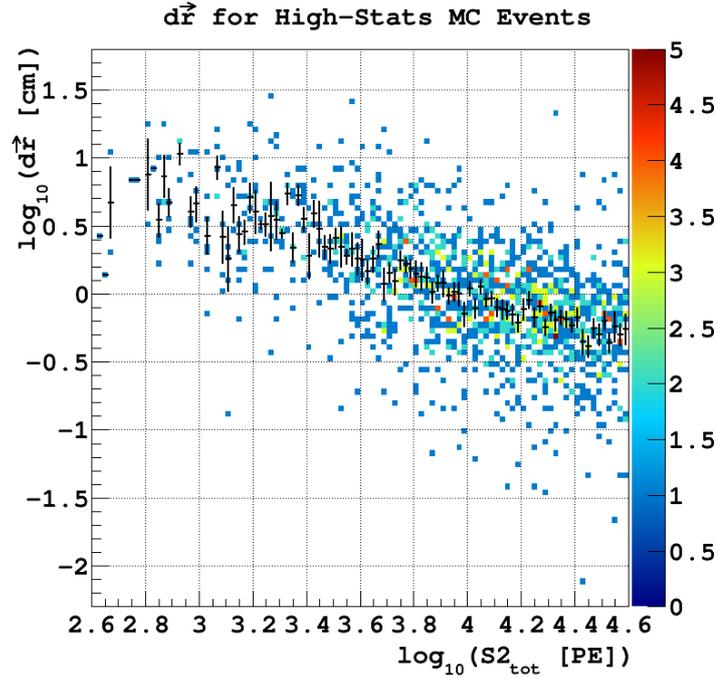


(a)

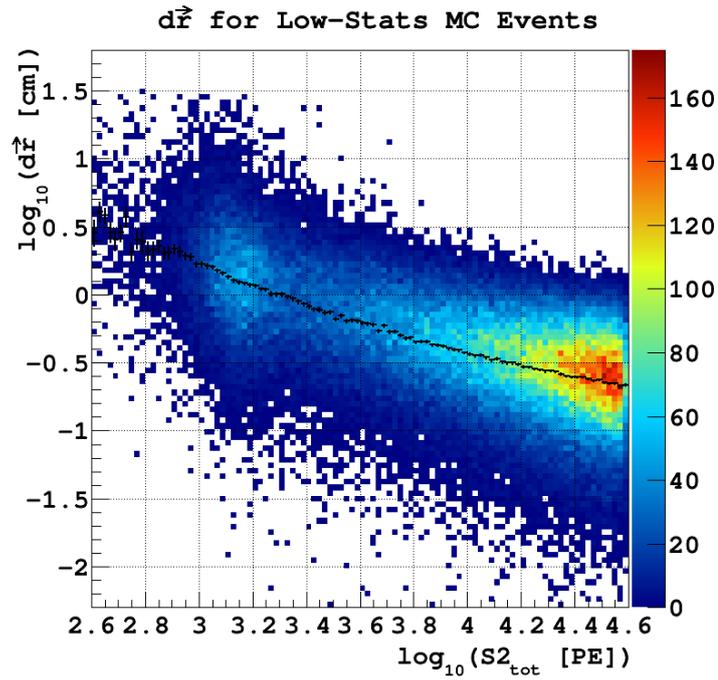


(b)

Figure 5.3: PCAMDF reconstruction accuracy as a function of radius for high (a) and low (b) PE statistics g4ds events. The high statistics MC event set performs equally well at all radii, while the low statistics set shows a distinct radial dependence: it performs best between  $r \approx 4$  cm and  $r \approx 10$  cm, and worst near the edge of the TPC ( $r = 17.78$  cm).



(a)



(b)

Figure 5.4: PCAMDF reconstruction accuracy as a function of  $S2_{\text{tot}}$  for high (a) and low (b) PE statistics MC events with  $S2_{\text{tot}} < 3.98 \times 10^4$  PE. Both event sets show worse performance at low  $S2_{\text{tot}}$  values.

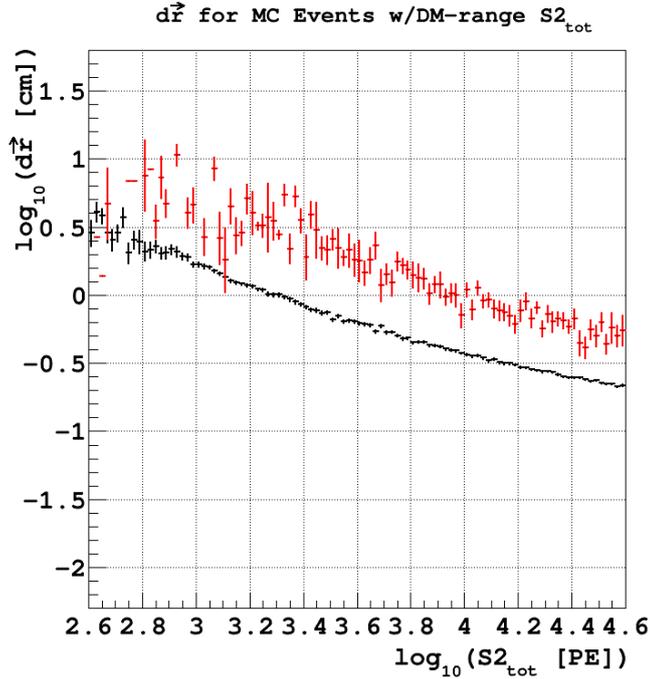


Figure 5.5: Average PCAMDF reconstruction accuracy as a function of  $S2_{\text{tot}}$  for MC events with  $S2_{\text{tot}} < 3.98 \times 10^4$  PE. The red (black) crosses show the vertical means of the horizontal bins for the high PE statistics (low PE statistics) g4ds event set. The vertical error bars on the means are statistical errors and the horizontal error bars just give the width of the horizontal bins. Note that the low statistics event set performs better than the high statistics event set over essentially this entire range of  $S2_{\text{tot}}$  values.

dependent on  $S2_{\text{tot}}$ , which means that the  $S2$  fraction spectra change shape with  $S2_{\text{tot}}$  — in other words, the PCA is sensitive to the size of  $S2$ , even when it's only given  $S2$ . Though the majority of the events in the high statistics event set have  $S2$ s in excess of 1 M PE, there are a few hundred events with  $S2_{\text{tot}} \lesssim 40$  k PE, consistent with ERs in the DM  $S1$  range ( $40 \text{ PE} < S1_{\text{tot}} < 460 \text{ PE}$ ). This allows us to compare the performance of the high statistics and low statistics reconstructions over a similar sample of events.

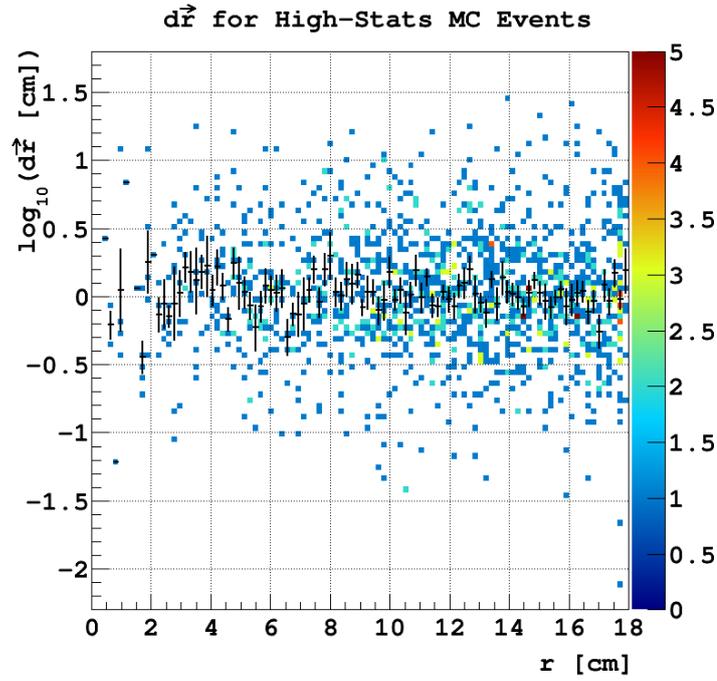
Figure 5.4 shows the performance of the high and low statistics reconstructions on MC events with  $S2_{\text{tot}} < 3.98 \times 10^4$  PE and  $r' < 17.78$  cm. Because the mean (the black crosses) can be difficult to see on the high statistics plot, I've overlaid them in Figure 5.5 (where red is the high statistics reconstruction and black is

the low statistics reconstruction). Here, we can easily see that the low statistics reconstruction performs better across this entire regime. At high- $S_{2\text{tot}}$  values, the low statistics reconstruction is, on average, several millimeters more accurate than the high statistics reconstruction. At low- $S_{2\text{tot}}$  values ( $\sim 2000$  PE), this increases to multiple *centimeters*. The high statistics MC PCAMDF was not trained on events with S2s this low, and so it performs poorly where we need it to perform the best: in the DM S1 range.

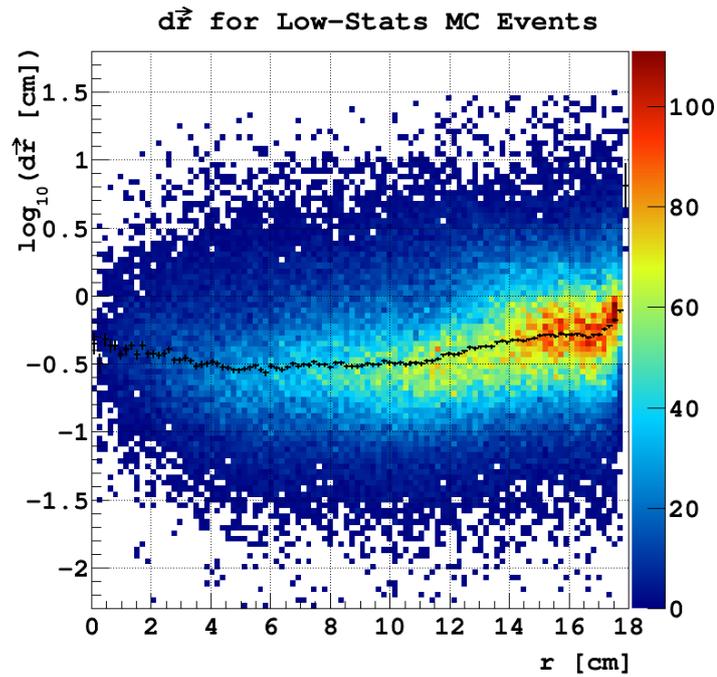
We see a similar trend when we look at these events as a function of radius, as in Figure 5.6, or Figure 5.7, where again, the mean profiles have been overlaid for clarity. We can see that the low statistics reconstruction (again, in black) performs, on average, several mm better than the high statistics set (in red), from the center of the detector out to just a few mm from the edge, where the low statistics reconstruction has a spike in inaccuracy. All of these factors taken into account, since we want to reconstruct data events which have S1s in the DM range and S2s less than a few tens of thousands of PE, the best option is to use the low statistics MC event set for training. Throughout the remainder of this thesis, “the MC event set” refers to the low statistics event set.

In summary, we know that the PCAMDF — when trained and executed on MC events — has a performance which depends on several factors:

1. The accuracy of the reconstruction depends on radius: it’s worst near the edge of the TPC ( $r \approx 17.78$  cm) and under the outer-ring PMTs, and best between  $r \approx 4$  cm and  $r \approx 10$  cm.
2. When trained on DM-S1-range g4ds events, the PCAMDF performs best near the high end of this energy range; the accuracy of the method depends roughly linearly on the size of S2 between  $\sim 1$  k PE and  $\sim 40$  k PE.
3. PCAMDF is sensitive to the  $S_{2\text{tot}}$  of the events used during training, even if



(a)



(b)

Figure 5.6: PCAMDF reconstruction accuracy as a function of radius for high (a) and low (b) PE statistics MC events with  $S2_{\text{tot}} < 3.98 \times 10^4$  PE. The black crosses show the vertical means of the horizontal bins. The vertical error bars on the means are statistical errors and the horizontal error bars just give the width of the horizontal bins. The accuracy of the high statistics event set still seems to be independent of radius.

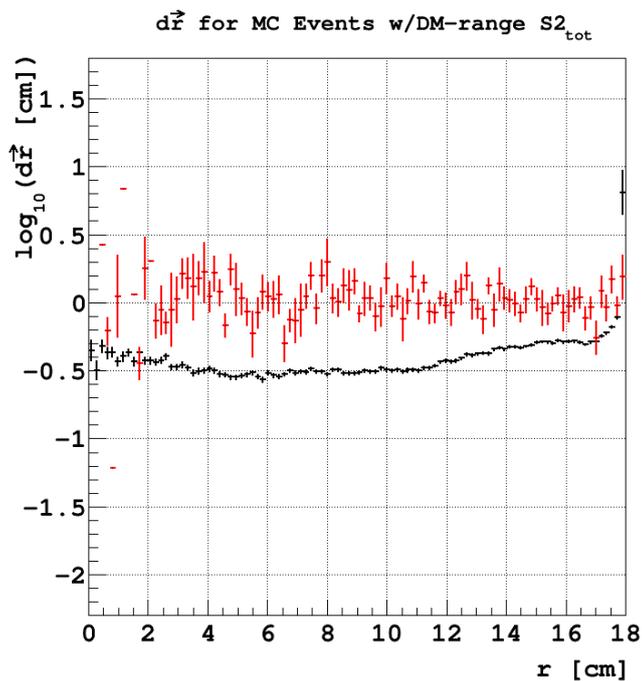


Figure 5.7: Average PCAMDF reconstruction accuracy as a function of radius for events with  $S2_{\text{tot}} < 3.98 \times 10^4$  PE. The red (black) crosses show the vertical means of the horizontal bins for the high PE statistics (low PE statistics) g4ds event set. The vertical error bars on the means are statistical errors and the horizontal error bars just give the width of the horizontal bins. The low statistics event set performs better than the high statistics event set over essentially this entire range of radius values.

it's only given the S2 fractions,  $S2[\ ]$ . Accordingly, it's best to reconstruct data with a PCAMDF which has been trained on MC events with energies in the range of interest (in our case, in the DM S1 range).

We know the overall accuracy of the method when performed on MC events, as well as the accuracy as a function of  $r$  and  $S2_{\text{tot}}$ , but does the PCAMDF have any biases? Does it tend to reconstruct events at higher or lower radii than their true radii, for instance? Figure 5.8 suggests that there is a radial bias: events with radii between  $\sim 10$  cm and  $\sim 16$  cm are “pushed” toward the TPC wall (their reconstructed radii are, on average, larger than their true radii). Between  $r \approx 16$  cm and the wall of the TPC, events are “pulled” inward (their reconstructed radii tend to be smaller than their true radii). There is also a slight tendency for events to be pushed away from the center PMT, as can be seen around  $0 < r \lesssim 1$  cm in Figure 5.8. Is there a similar “repulsive” effect for all PMTs?

Figure 5.9 suggests that this is not the case. This figure shows, for the seven central PMTs, the radial bias in reconstructed position, with respect to the center of each PMT. To be explicit, the vertical axis is the radial bias of the reconstructed positions, relative to the center of a given PMT: (true radius from center of given PMT minus reconstructed radius from center of same PMT). If an event has a high vertical position on this plot, it means it was “pulled in” toward the PMT center, since the true radius would be higher than the reconstructed radius. The horizontal axis is the distance of an event from the center of the given PMT, squared and divided by 4.1275 cm (half of the PMT-to-PMT spacing). Figure 5.9 shows that the PCAMDF does not have a tendency to pull events toward PMT centers, or push them away from PMT centers. This point will be crucial when discussing the reconstructed positions of data events later in this Chapter.

Even though there is a strong radial bias with the PCAMDF method, there

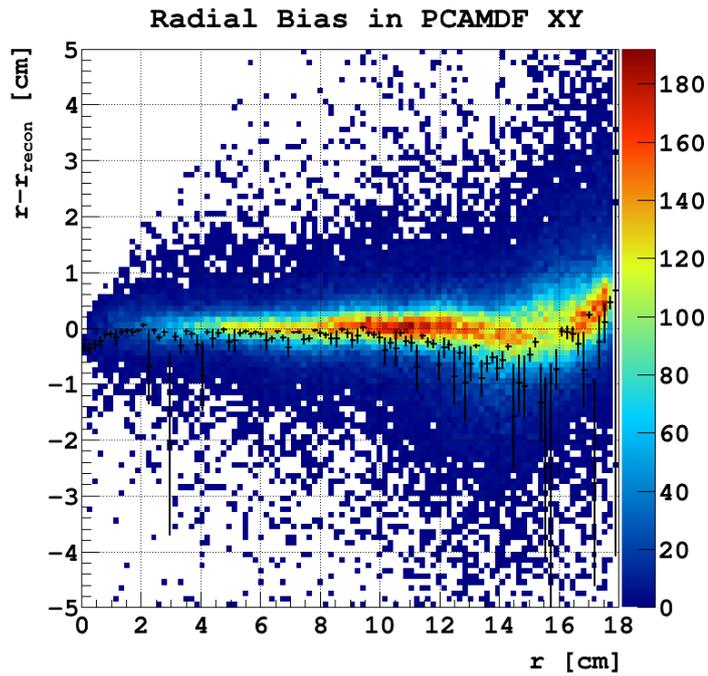


Figure 5.8: Radial bias in the PCAMDF reconstruction for MC events. Since the vertical axis is  $r - r'$ , a positive value means that an event was “pulled in” toward the detector center and a negative value means it was “pushed out”, away from the detector center. Note that there is a strong tendency for the PCAMDF to push out events with  $10 \text{ cm} \lesssim r \lesssim 17 \text{ cm}$  and to pull in events with  $r \gtrsim 17 \text{ cm}$ . The black crosses show the vertical means of the horizontal bins. The vertical error bars on the means are statistical errors and the horizontal error bars just give the width of the horizontal bins.

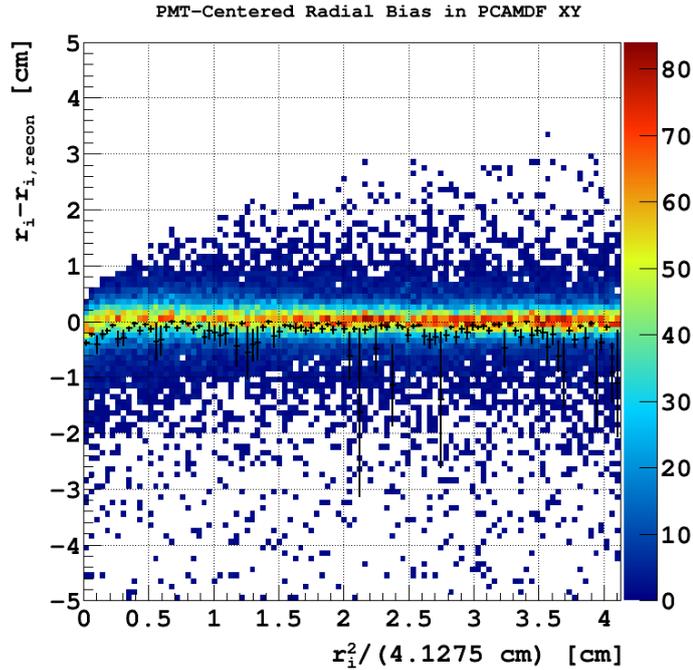


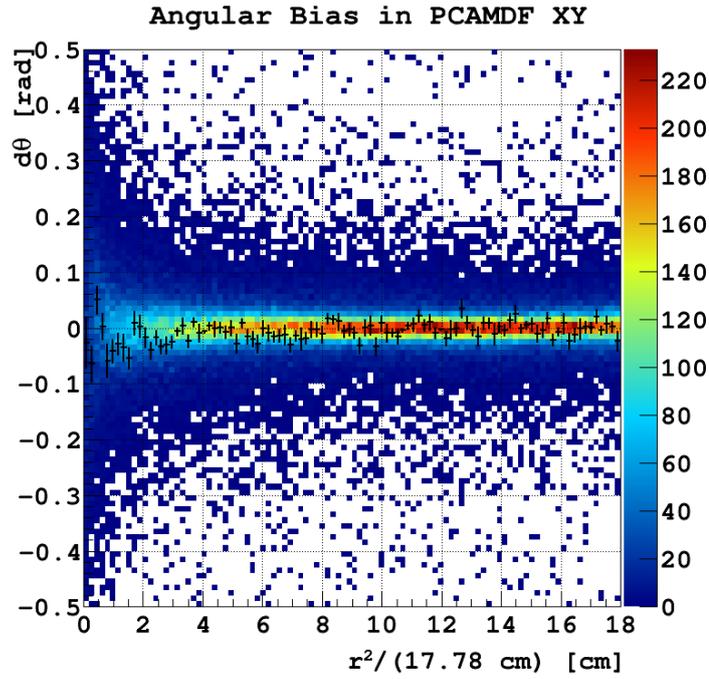
Figure 5.9: PMT-centered radial bias in the PCAMDF reconstruction for MC events. *On this plot only*,  $r_i$  measures the radius of an event from the center of the  $i$ -th PMT, where the  $i$ -th PMT is the nearest PMT to that event, and only the seven inner anode PMTs are considered.  $r'_i$  is the reconstructed event radius from the center of the  $i$ -th PMT. Although the statistical error bars on some of the horizontal bins are quite large, there is no overall tendency for events to be pulled toward, or pushed away from, the center of a PMT. The horizontal axis shows  $r_i^2$  instead of  $r_i$  so that the horizontal bins contain equal volumes of the detector, sharing the statistical error as evenly as possible across all bins. The black crosses show the vertical means of the horizontal bins. The vertical error bars on the means are statistical errors and the horizontal error bars just give the width of the horizontal bins.

seems to be no angular bias. In fact, the PCAMDF reconstructs the angular positions of MC events astonishingly well, as can be seen in Figure 5.10. The PCAMDF  $xy$  shows no tendency to reconstruct events toward higher or lower angular position ( $\theta$ ) overall, though the accuracy of the angular reconstruction is poor for events with  $r \lesssim 4$  cm (see subfigure (a)). Note that this is approximately the radius of the central PMT, so we might expect the angular reconstruction to be a bit poorer here. Subfigure (b) shows high means at  $\pm\pi$ , where a slight change in position yields a large change in  $\theta$ , due to the discontinuity at  $\pm 180^\circ$ . Selecting events with  $|\langle\theta\rangle| < 2$  rad and  $r > 4$  cm gives an average reconstruction error in  $\theta$  of  $\langle d\theta \rangle < 2.03^\circ$  overall. Note that, when looking over restricted ranges of  $\theta$ , there is an extremely slight preference for events to be pulled toward PMT centers, though the average bias is  $< 1^\circ$  overall. Overall, the PCAMDF has very good angular accuracy, reconstructing MC events within  $\pm 2^\circ$  of their true positions, outside of the central PMT.

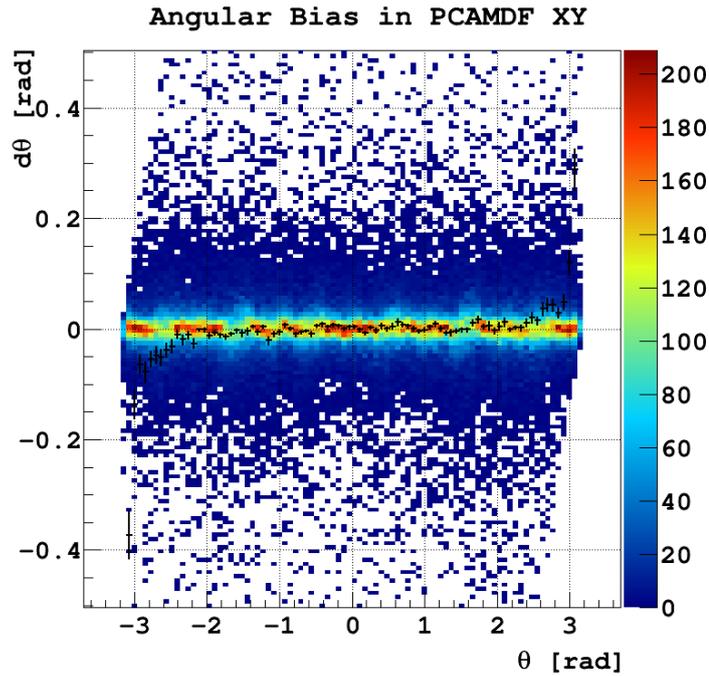
## 5.4 Calculating a Fiducial Volume Cut

One of the main reasons for developing an  $xy$ -Reconstruction method in DS-50 is to use it to generate a fiducial volume cut. DS publications to date haven't needed such a cut simply because no events remained in the expected WIMP parameter space after all CX cuts had been applied. Larger exposures in DS-50, or the larger detector volume of DS-20k could necessitate a fiducial volume cut in the future, though, and so I go through the steps of finding one here.

First, we know that the PCAMDF has strong radial biases near the edge of the TPC – where we'd like to make our fiducial volume cut. “Wall events” (with  $r \geq 16.78$  cm) can be reconstructed much closer to the center of the detector than where they actually occurred. Whatever metric we use to generate a fiducial



(a)



(b)

Figure 5.10: Angular bias as a function of radius (a) and azimuthal angle (b) in the PCAMDF  $xy$ -Reconstruction for events with  $S2_{\text{tot}} < 3.98 \times 10^4$  PE. Note that there are no strong angular biases either as a function of radius or azimuthal angle,  $\theta$ . What appear to be biases at  $\pm\pi$  are just due to the discontinuity of the value of  $\theta$  at  $\theta = \pm\pi$ . The black crosses show the vertical means of the horizontal bins. The vertical error bars on the means are statistical errors and the horizontal error bars just give the width of the horizontal bins.

%ile	$r_{cut}$	%ile	$r_{cut}$								
1	18.4448	18	17.4535	35	17.194	52	16.9544	69	16.6765	86	16.2088
2	18.1549	19	17.4353	36	17.1801	53	16.9398	70	16.6572	87	16.1622
3	18.0141	20	17.4179	37	17.1662	54	16.9248	71	16.6374	88	16.1119
4	17.9213	21	17.4009	38	17.1523	55	16.9097	72	16.6174	89	16.054
5	17.8504	22	17.3847	39	17.1384	56	16.8946	73	16.5961	90	15.9909
6	17.7951	23	17.3685	40	17.1246	57	16.8793	74	16.5746	91	15.9205
7	17.7482	24	17.3529	41	17.1107	58	16.864	75	16.5522	92	15.84
8	17.7073	25	17.3376	42	17.0968	59	16.8484	76	16.5287	93	15.7459
9	17.672	26	17.3225	43	17.0828	60	16.8325	77	16.5041	94	15.6308
10	17.6396	27	17.3078	44	17.0686	61	16.8165	78	16.4784	95	15.4933
11	17.6103	28	17.2932	45	17.0545	62	16.8	79	16.452	96	15.3247
12	17.5833	29	17.2787	46	17.0404	63	16.7835	80	16.4236	97	15.0939
13	17.5585	30	17.2644	47	17.0261	64	16.7664	81	16.3935	98	14.7701
14	17.5348	31	17.2501	48	17.0119	65	16.7493	82	16.3614	99	14.2046
15	17.5129	32	17.236	49	16.9977	66	16.7316	83	16.327	99.5	13.6488
16	17.4921	33	17.222	50	16.9833	67	16.7138	84	16.2908	99.9	12.0189
17	17.4724	34	17.208	51	16.9689	68	16.6954	85	16.2516	99.99	5.90486

Table 5.1: Fiducial cut percentiles for different values of  $r_{cut}$ . Note that making a cut at the detector radius ( $r_{cut} = 17.78$  cm) already removes about 5-6% of wall events. These events are incorrectly reconstructed outside of the detector volume. A 1 cm fiducial cut ( $r_{cut} = 16.78$  cm) removes 63-64% of all wall events, and a 2 cm fiducial cut ( $r_{cut} = 15.78$  cm) removes 92-93% of all wall events. There are  $6.41 \times 10^5$  wall events in this study, so a 99.99% cut would pass only 64 wall events.

volume cut needs to take this into account.

The simplest method is to calculate at what radius,  $r_{cut}$ , we need to cut to remove a given percentage of wall events. For instance, due to the radial bias in the reconstruction, setting  $r_{cut} = 16.78$  cm actually removes only about 63% of MC wall events. We can calculate the percentage of wall events removed for any value of  $r_{cut}$  and leave it up to future analyzers to decide where to make the cut to provide the best balance of background reduction and conservation of fiducial volume. The script `fiducial.C` in [Appendix E](#) can calculate  $r_{cut}$  for an arbitrary percentile, but all integer percentiles are included in [Table 5.1](#). A plot of the data in [Table 5.1](#) can be seen in [Figure 5.11](#).

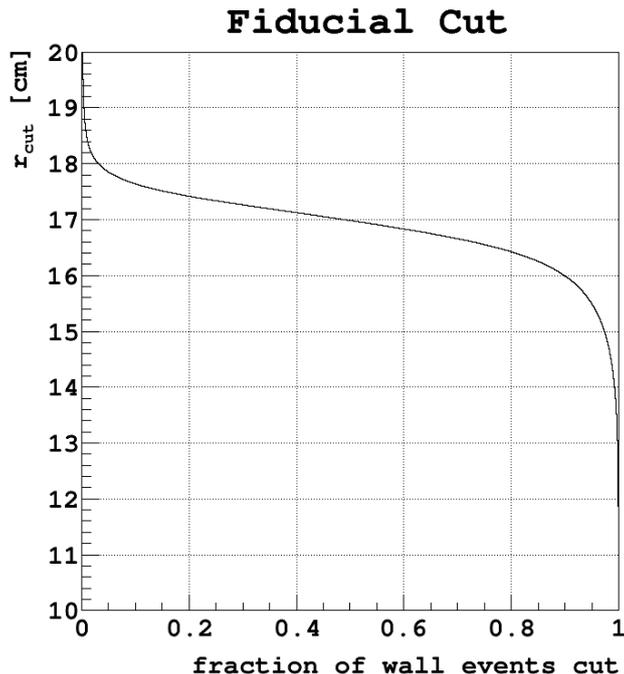


Figure 5.11: Fiducial cut radius as a function of the fraction of wall events cut. The values in Table 5.1 are taken from this plot. For fiducial efficiencies  $\gtrsim 90\%$ , the cut radius  $r_{cut}$  drops dramatically.

## 5.5 Performance on Data

Previous DS-50  $xy$ -Reconstructions — even though their event distributions seemed to agree with our uniformity assumption for  $^{39}\text{Ar}$  — suffered from severe distortions of the S1 spectrum across  $xy$ . Two problems of note are PMT-centered and TPC-centered S1 spectrum distortions. In Subsections 5.5.1 and 5.5.2, it will be shown that the PCAMDF  $xy$  suffers from no such distortions. In Subsection 5.5.3, the spatial distributions (radial and over  $xy$ ) of data events reconstructed with the PCAMDF for both AAr and UAr are investigated. In Subsection 5.5.4, the S2 Top-Bottom Asymmetry ( $S2_{\text{TBA}}$ ) map is introduced, though it will be covered in more detail in the next Section. In Subsection 5.5.5, an overview of an unusual “hole” of events near PMT # 2, on the cathode, is given; this is investigated further with an S1-based PCAMDF  $xy$ -Reconstruction in Subsection 5.5.6. And

in Subsection 5.5.7, the general procedure for finding  $xy$ -dependent S2 corrections in DS-50 is discussed.

### 5.5.1 PMT-centered S1 Distortions

Standard  $xy$ -Reconstructions in DS-50 are not given any S1 information. The PCAMDF uses the logarithm of the anode S2 fractions ( $\log(S2[n]/S2_{\text{top}})$ ) and an older  $xy$ -Reconstruction, called “XYlocator”<sup>7</sup>, used the logarithm of the total S2 fractions ( $\log(S2[n]/S2_{\text{tot}})$ ) as well as the fraction of light seen by the bottom PMTs ( $S2_{\text{bot}}/S2_{\text{tot}}$ ). Since the reconstructions don’t know about S1, it can be a useful crosscheck on these methods.

Figure 5.12 shows (what I’ve dubbed) “Radial Energy Profiles” (REPs) for three anode PMTs (PMTs # 25, 30, and 35) using AAr data. An REP shows how the S1 spectrum changes with distance from the center of a given PMT. In each of the six panels of Figure 5.12, the black (red, green, blue, yellow, violet) line gives the S1 spectrum for all events reconstructed within 1 cm (2 cm, 3 cm, 4 cm, 5 cm, 6 cm) of the center of the PMT in question.<sup>8</sup> The PCAMDF shows no dependence of the S1 spectrum upon reconstructed radius under the centers of these PMTs, while XYlocator shows severe distortions, which are worst near the PMT center and decrease in severity as the radius increases. With the PCAMDF  $xy$ , no anode PMT suffers the S1 spectrum distortions evident when using XYlocator, which shows at least mild distortions on 16 of the 19 anode PMTs.

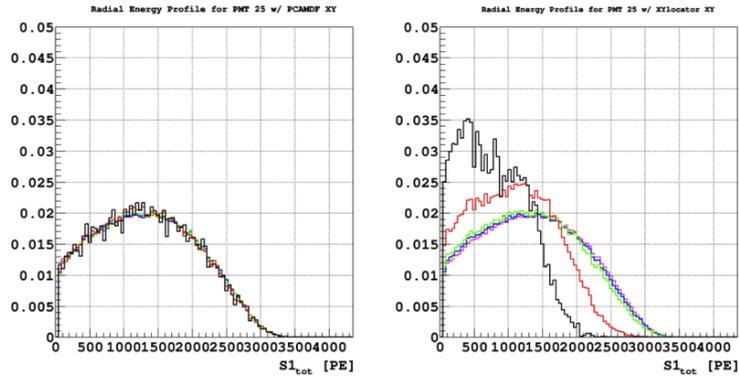
### 5.5.2 TPC-centered S1 Distortions

Since PMT # 30 is the central anode PMT, Figure 5.12 (b) shows the behavior of the TPC-centered S1 spectrum with both the PCAMDF  $xy$  and XYlocator near

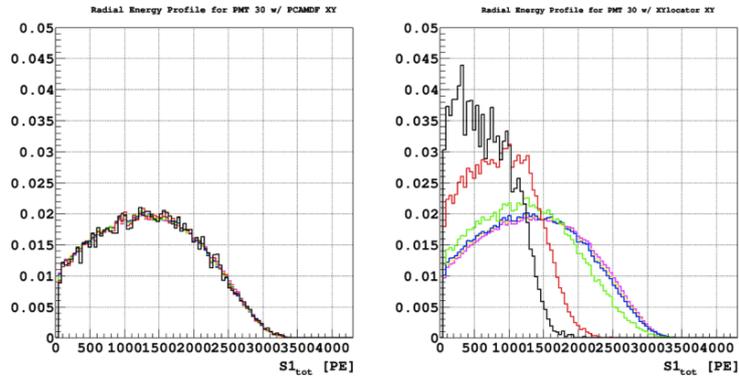
---

<sup>7</sup> Brodsky (2015)

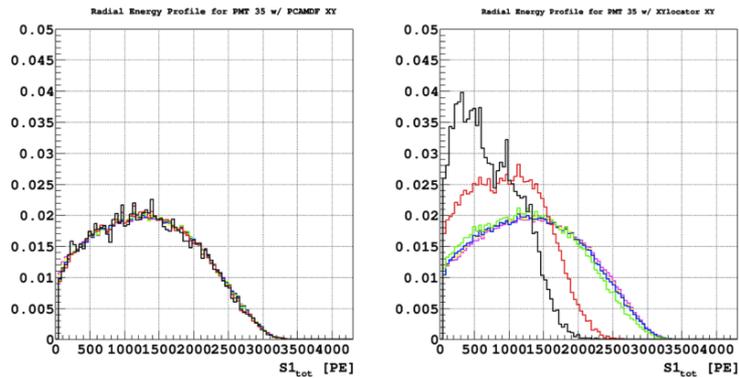
<sup>8</sup> Note that these are overlapping disks, not annuli.



(a)



(b)



(c)

Figure 5.12: Radial Energy Profiles for PMTs # 25 (a), 30 (b), and 35 (c) using AAr data. The PCAMDF (left) shows no dependence of the S1 spectrum under the PMT centers. The same cannot be said for the previously-used DS-50  $xy$ -Reconstruction,  $XYlocator$  (right).

the axis of the TPC. Near the outside wall, though, we need to perform a slightly different study. Figure 5.13 shows TPC-centered REPs for AAr (a) and UAr (b), with some caveats. First, though the PMT-centered REPs are cumulative (that is, the black line is the integral from  $0 < r_{PMT_i} < 1$  cm, the red line is the integral from  $0 < r_{PMT_i} < 2$  cm, etc.), the TPC-centered REPs are not (the black line is the integral from  $0 < r < 2$  cm, the red line is the integral from  $2 < r < 4$  cm, etc.). Second, I've added more contours for the TPC-centered REPs (all with 2 cm spacing, instead of 1 cm), where the color scheme is the same as the PMT REPs, and the three new contours are cyan ( $12 < r < 14$  cm), leaf green ( $14 < r < 16$  cm), and periwinkle ( $16 < r < 18$  cm).<sup>9</sup> Finally, the TPC-centered REPs are scaled so that their behavior at low- $S1_{tot}$  is clear: the integral of each profile from  $1500 \text{ PE} < S1_{tot} < 2000 \text{ PE}$  is normalized to 1.

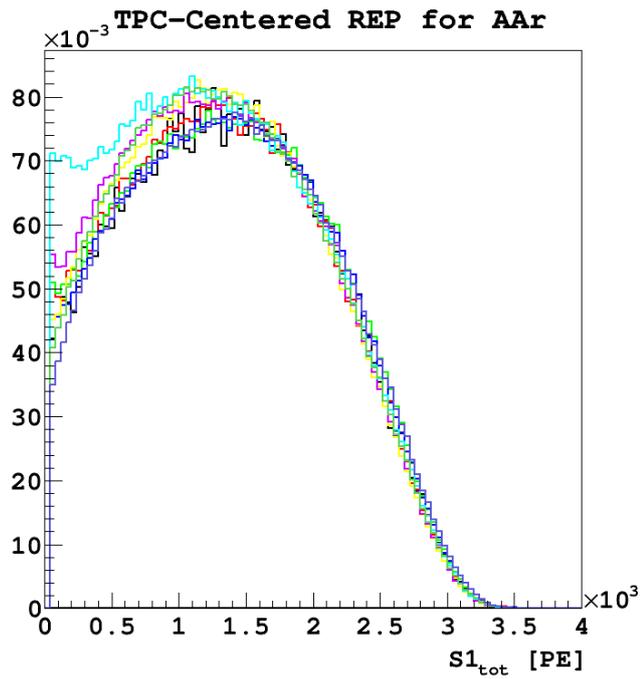
Figure 5.13 (a) shows a consistent  $^{39}\text{Ar}$  spectrum out to the cyan line ( $12 < r < 14$  cm), where it gains a low- $S1_{tot}$  shoulder and then loses it again for the leaf green and periwinkle lines. Figure 5.13 (b) shows the  $^{39}\text{Ar}$  spectrum with a large peak at 600 PE, caused by  $\gamma$ -ray Compton backscatters<sup>10</sup>. Since this peak is enhanced by external  $\gamma$ 's entering the TPC around the cylindrical wall, we would expect the peak height to strictly increase in size from the TPC core out to the wall; this is indeed what we observe. Additionally, we see the same low-energy cyan peak in UAr data as we do in AAr data; a smaller such peak is visible in the violet line ( $10 < r < 12$  cm).

Although there are  $S1_{tot}$  dependences of the energy spectrum upon reconstructed radius, these mostly seem to be due to real, physical effects. The peak at 600 PE is due to Compton backscatters, which increase in likelihood the nearer we get to the wall. This peak is obvious in UAr data, but is washed out in AAr data.

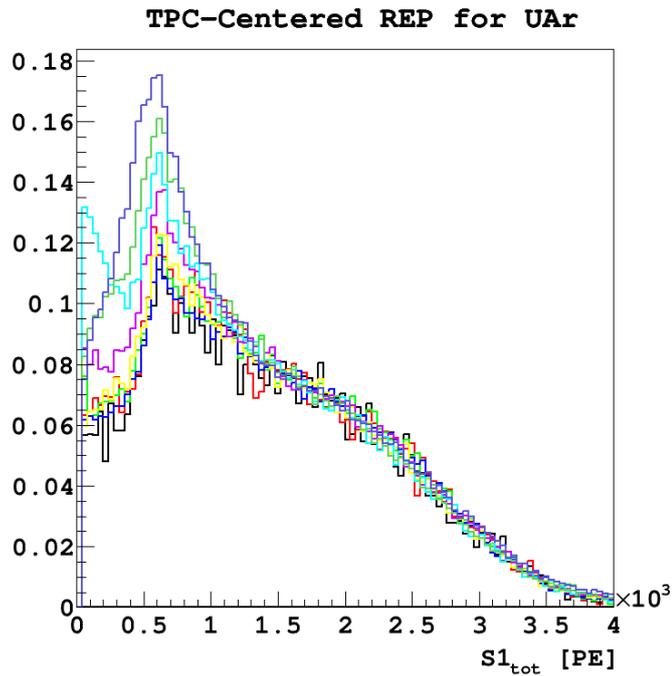
---

<sup>9</sup> In Figure 5.13 (b), at the peak at  $\sim 600$  PE, the contours from the highest (vertical) one down are colored periwinkle, leaf green, and cyan, to be explicit.

<sup>10</sup> Agnes et al. (2016c)



(a)



(b)

Figure 5.13: TPC-centered REPs for AAr (a) and UAr (b) data, using the PCAMDF  $xy$ -Reconstruction. Both plots are scaled such that the integral from  $1500 \text{ PE} < S1_{tot} < 2000 \text{ PE}$  is normalized to 1, to show the behavior at low- $S1_{tot}$ . Further description in the text.

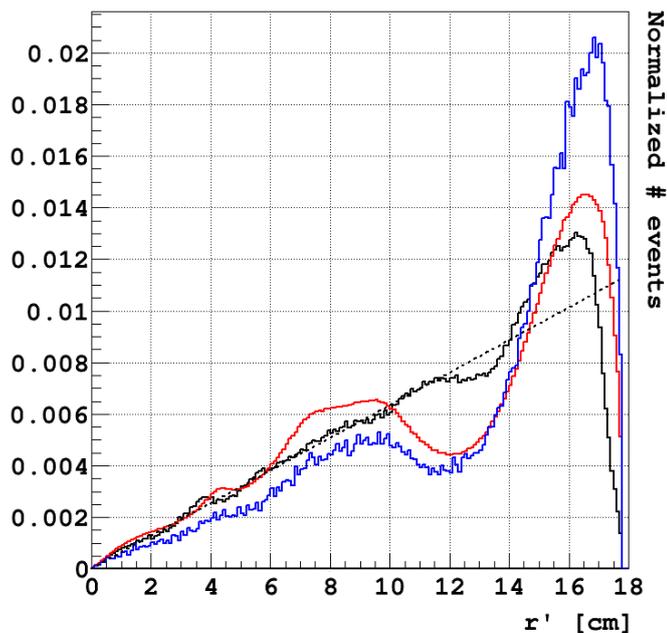
Its presence is noticeable in AAr, though, in the shift of the peak of the distribution towards lower  $S1_{\text{tot}}$  with increasing radius. The cyan and violet shoulders are quite interesting, too: their radii ( $\sim 10 \text{ cm} \rightarrow 14 \text{ cm}$ ) place them between the centers of the inner-ring and outer-ring PMTs. If this were an effect specific to this inter-PMT region of the detector, we would expect to see a similar distortion between the center PMT ( $\# 30$ ) and the inner-ring PMTs. The green line in Figure 5.13 (b) shows a bump in the lowest bin ( $40 < S1_{\text{tot}} [\text{PE}] < 80$ ), but nothing else. This feature is still being investigated at the time of this writing.

### 5.5.3 Spatial Distributions of Reconstructed Events

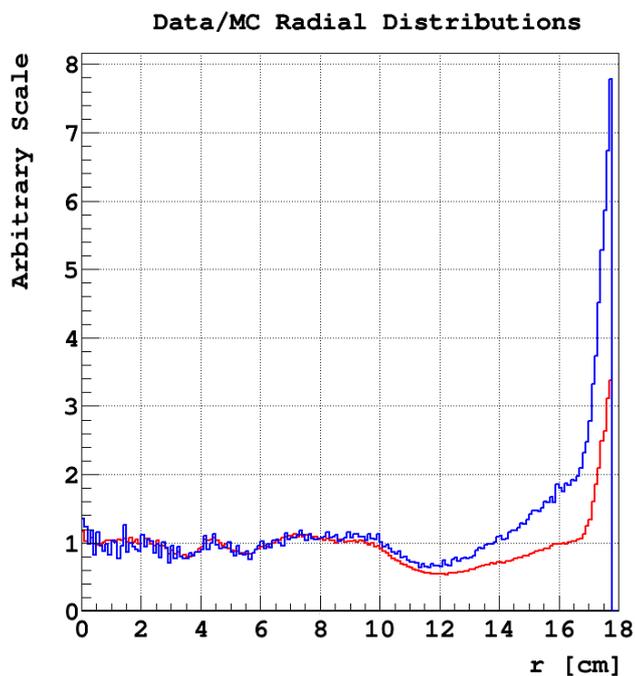
Since the true event positions in data aren't known, we can't perform most of the studies that we did for MC events in Section 5.3. But we can look at the radial and transverse ( $xy$ ) distributions of reconstructed events. Figure 5.14 (a) shows the normalized radial distributions of g4ds (black), AAr (red), and UAr (blue) events with PCAMDF-reconstructed  $xy$  positions. The dotted black line shows the radial distribution for an event set with a perfectly uniform  $xy$  distribution, where the number of events increases linearly with increasing  $r$  (due to the cylindrical geometry of the detector). As was shown earlier (Figure 5.8), the MC events have a tendency to be “pushed” outward, toward the edge of the TPC, when their true positions are between  $10 \lesssim r \text{ cm} \lesssim 17 \text{ cm}$  and “pulled” in, away from the edge of the TPC, when their true positions are greater than about  $r = 17 \text{ cm}$ . The result is an over-density of events just within the TPC radius, with a peak around  $r = 16 \text{ cm}$ . This peak is evident in both AAr and UAr data.

Normalizing the data histograms by the MC one gives a clearer picture of what's going on, as can be seen in Figure 5.14 (b). Dividing out the MC line and vertically scaling shows that the UAr and AAr event distributions are nearly identical in shape out to about  $r = 10 \text{ cm}$ , after which, the AAr and UAr drop

### Radial Nonuniformity



(a) Normalized radial distributions of g4ds (black), AAr (red), and UAr (blue) events with PCAMDF-reconstructed  $xy$  positions. The dotted black line shows the radial distribution for an event set with a perfectly uniform  $xy$  distribution.



(b) MC-normalized event distributions for AAr (red) and UAr (blue) data. These two datasets have nearly identical event distributions out to  $r \approx 10$  cm, after which, a wall event spike is evident.

Figure 5.14: Radial distributions of reconstructed MC and data events.

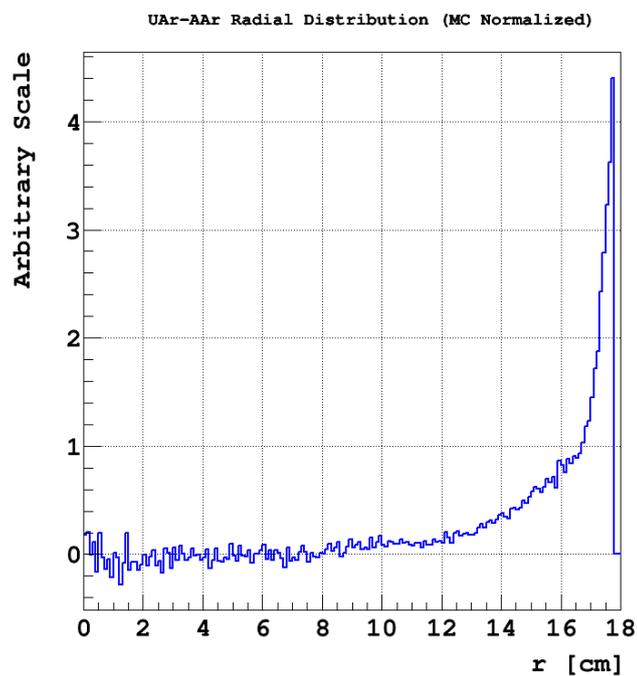


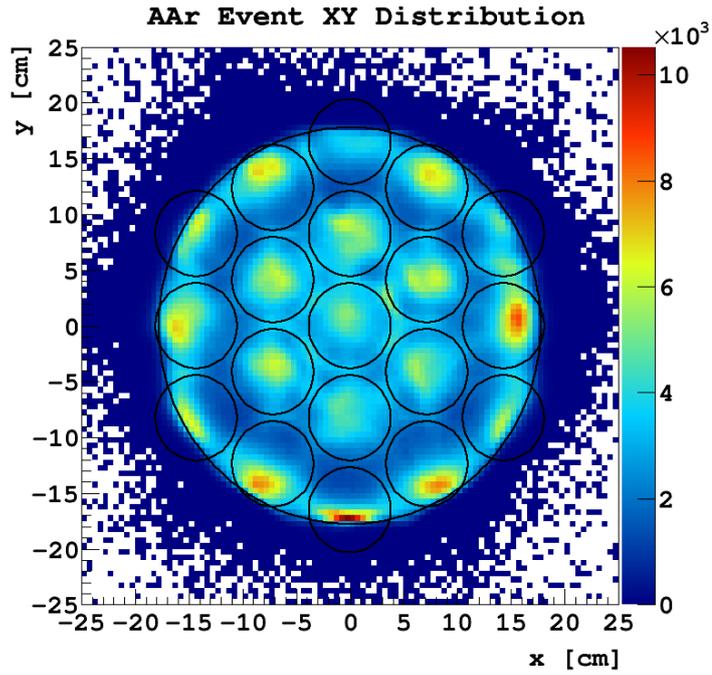
Figure 5.15: The UAr distribution from Figure 5.14 (b) with the AAr distribution subtracted. The exponential drop in the prevalence of wall events with increasing distance from the wall is evident. No attempt has been made to fit the distribution as it is complicated by the cylindrical geometry of the detector.

and then slowly increase out to  $r \approx 17$  cm. This slow drop and rise cannot be attributed to the TPC-centered radial bias of the PCAMDF method, since this should have been normalized away by dividing by the MC distribution. After dividing, both data distributions show a sharp peak as  $r \rightarrow 17.78$  cm. This spike is from external gamma rays and wall events which were not simulated for the MC event set used in this dissertation. Their presence can also be seen in the shift of the peak at high radii in Figure 5.14 (a), between MC and data. The exponential drop in the prevalence of these events as the distance from the wall increases is shown in Figure 5.15, where the AAr distribution in Figure 5.14 (b) has been subtracted from the UAr distribution.

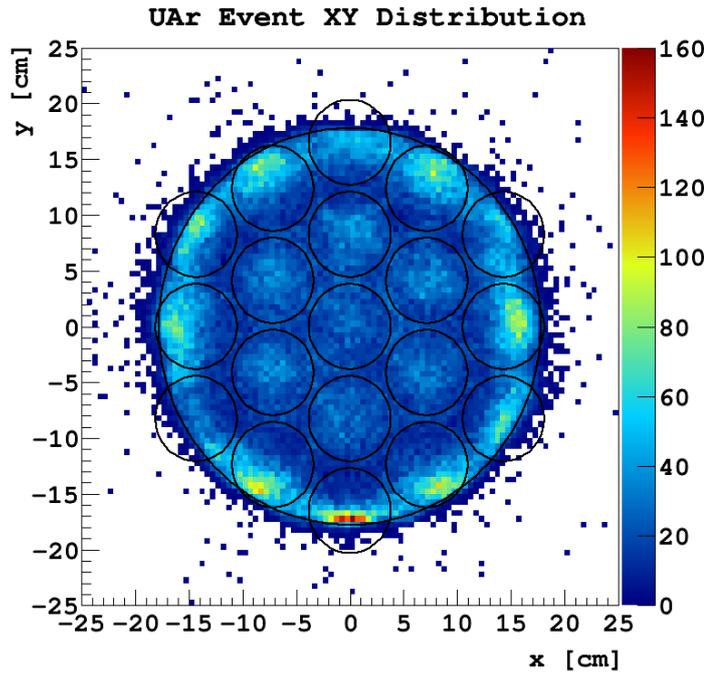
The  $xy$  distributions of the reconstructed events are also very interesting. Figure 5.16 shows the  $xy$  distributions of the events which pass all CX cuts and are roughly in the DM S1 range for AAr (a) and UAr (b). There are several significant features to note here:

**The PCAMDF “finds” the edge of the detector.** The PCAMDF is not explicitly “told” where the edge of the active volume is, nor is it restricted to only reconstruct events to a radius  $r < 17.78$  cm. The PCAMDF generates a function which receives S2 fractions as input, and returns an  $x$  or  $y$  position. If the input data map to an  $xy$  position which is outside of the detector volume, then the event is assigned an unphysical position. This is in contrast to previous  $xy$ -Reconstructions in DS-50, which explicitly required events to be reconstructed within the active volume.

**Events can be reconstructed outside the detector.** Even though the PCAMDF *mostly* reconstructs events within the detector volume, some can be reconstructed outside of it. Typically, these events are discarded when performing studies which require event  $xy$  positions. In AAr, 97.61% of events which



(a) PCAMDF  $xy$  distribution of AAr events which pass all CX cuts, and have  $0 < S1_{\text{tot}} < 500$  PE.



(b) PCAMDF  $xy$  distribution of UAr events which pass all CX cuts, and have  $0 < S1_{\text{tot}} < 500$  PE.

Figure 5.16: PCAMDF  $xy$  distribution of AAr (a) and UAr (b) events which pass all CX cuts, and have  $0 < S1_{\text{tot}} < 500$  PE. Note how the PCAMDF “finds” the edge of the detector, even though it was given no explicit instruction to do so during training.

pass all CX cuts and have  $0 < S1_{\text{tot}} [\text{PE}] < 500$  are reconstructed within the detector volume; in UAr, it's 96.03% (presumably because there is a larger fraction of wall events in UAr, which are more likely to be reconstructed outside of the detector, due to their close proximity to the wall).

**In UAr, there is a significant population of near-wall events.** The  $z$ -scales of the plots in Figure 5.16 give us an idea as to where events are concentrated in  $xy$  for each of these data sets. While both plots show yellow, orange, and red bins near the detector edge, only AAr shows bright yellow bins under the inner-ring PMTs. This suggests that, while it's about equally likely for an event to occur underneath any given anode PMT in AAr data, it's much more likely for an event to occur under an outer-ring PMT in UAr data. This implies that wall events and external gamma rays make up a greater proportion of total events in UAr than in AAr, a notion which is supported by the relative event distributions shown in Figure 5.14 (b).

**Events are “clumped” under PMTs.** In Section 5.3, the biases of the PCAMDF reconstruction on MC were investigated, and it was shown (in Figure 5.9) that there is no PMT-centered radial bias in the reconstruction. That is, for MC events at least, there is no tendency for events to be “pulled in” or “pushed away from” PMT centers. However, especially in the AAr data set, there is a very clear  $xy$  dependence on the density of events. This is less evident in UAr data. A previously-used DS-50  $xy$ -Reconstruction, XYlocator, applied a “uniformity pressure” when reconstructing AAr events (which are assumed to be uniformly distributed throughout the volume, since the event set is dominated by  $^{39}\text{Ar}$  ERs). While this results in a uniform  $xy$  distribution of events overall, it causes geometric artifacts to appear in the distribution of some variables across  $xy$ , and it distorts the S1 spectrum

underneath PMT centers.<sup>11</sup>

It should be noted that the reason *why* there appear to be more events underneath the PMTs in DS-50 is still being investigated. PCAMDF tests on the MC suggest that there is no tendency for the method to pull events underneath PMTs, but a mechanism for *detecting* more events under PMTs (as opposed to underneath the PTFE reflector in which the anode PMTs are mounted) is still being debated. Biases in the TPC hardware trigger, or the DarkArt pulse-finder seem unlikely, but they’ve yet to be investigated fully. There may also be a radial dependence of the QE of the PMT photocathodes, though toy MC studies which investigated this effect were inconclusive.

#### 5.5.4 S2<sub>TBA</sub>

Perhaps the most convincing evidence for the efficacy of the PCAMDF *xy*-Reconstruction is the S2 Top-Bottom Asymmetry (S2<sub>TBA</sub>) map. Since DS-50 has two arrays of PMTs: a “top” one at the anode and a “bottom” one at the cathode, a given event in the detector has some portion of its S2 light detected by the top PMTs (S2<sub>top</sub>), and the rest is detected by the bottom PMTs (S2<sub>bot</sub>). Though at first it might seem odd, the amount detected at each end of the detector is a strong function of the *xy* position of an event.

After a particle interacts at a given *xyz* position in the  $\ell$ Ar, its ionization electrons drift straight upward and generate S2 light along a vertical path through the *g*Ar. Many electrostatics simulations show that the electric fields in DS-50 are uniform and perpendicular to the *g*Ar/ $\ell$ Ar boundary (neglecting a small detector tilt).<sup>12</sup> This means that the S2 light of an event is generated directly above that

---

<sup>11</sup> [Brodsky \(2015\)](#)

<sup>12</sup> [Love \(2013\)](#)

event's  $xy$  position in the  $\ell\text{Ar}$ .<sup>13</sup> But the S2 light is only *generated* along a straight line through the gas gap. Obviously, since we detect S2 light across all 38 PMTs in DS-50, the light must be getting diffused: there are two major factors which affect this.

First, S2 light is generated via electroluminescence. As the drift electrons are accelerated through the  $g\text{Ar}$ , they collide with argon atoms and excite them. The atoms de-excite via the process described in Section 3.2, emitting hard UV photons isotropically with a peak wavelength of  $\sim 128$  nm. This light cannot be directly detected, because the QE of our Hamamatsu R11065 PMTs is zero below about 150 nm.<sup>14</sup> This is why we cover the inside of the entire detector with the wavelength-shifter TPB, which absorbs these UV photons and emits new photons with a peak wavelength of 420 nm. Once the light has been wavelength-shifted, it can finally be detected by the PMTs. After the argon de-excites, the rest of the above process happens in a matter of just a few nanoseconds.

Essentially every step in the preceding paragraph works to diffuse the S2 light: the excited argon atoms emit their UV photons isotropically, the TPB absorbs these UV photons and re-emits new photons pseudo-isotropically, the photons may cross the  $g\text{Ar}/\ell\text{Ar}$  boundary and change direction, and so on. So why does the  $xy$  position of the event matter at all? Because the gas gap is so small (1 cm thick), the S2 light is “proximity-focused”, meaning it's detected by the anode PMTs mostly in a small region of  $xy$ , simply because most of the solid angle it “sees” is occupied by the anode PMT array. If an event occurs directly underneath an anode PMT, most of its S2 light will be detected by that PMT — this is the working principle behind any S2-based  $xy$ -Reconstruction.

But what happens when an event occurs at an  $xy$  position which is *between*

---

<sup>13</sup> This is, in fact, why the S2 fractions contain any position information at all.

<sup>14</sup> <https://www.hamamatsu.com>

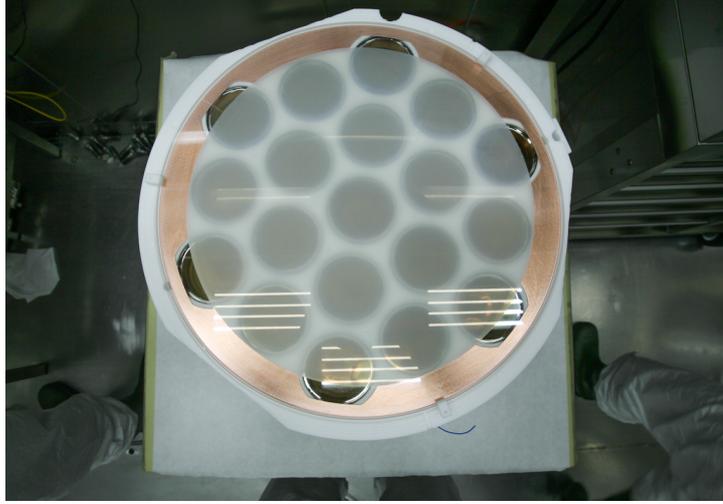


Figure 5.17: Bottom (cathode) PMT array, prior to installation in DS-50. The PTFE reflector mount holds the PMTs in place, and the entire array is covered by the cathode window, which is coated in conductive ITO and the wavelength shifter TPB. From internal docdb entry # 1590. Note that the horizontal white bars are reflections of the fluorescent lights in the room in which the picture was taken. Photo by Yury Suvorov.

anode PMTs? Then, most of its S2 light hits the PTFE reflector within which the anode PMTs are mounted. This means more of the S2 light is reflected down toward the cathode PMTs. Overall, then, we expect that when an event occurs underneath an anode PMT, the ratio  $S2_{\text{top}}/S2_{\text{bot}}$  will be higher, on average, than when an event occurs between anode PMTs. This is the basis behind the  $S2_{\text{TBA}}$  map. The cathode PMTs, mounted in PTFE reflector and covered with the cathode window (with a coating of ITO and TPB) can be seen in Figure 5.17.

Note that the ratio  $S2_{\text{top}}/S2_{\text{bot}}$  is asymmetric and singular when  $S2_{\text{bot}} \rightarrow 0$ , so instead, we generally look at the  $S2_{\text{TBA}}$ , which is defined as

$$S2_{\text{TBA}} = \frac{S2_{\text{top}} - S2_{\text{bot}}}{S2_{\text{top}} + S2_{\text{bot}}} \quad (5.3)$$

$S2_{\text{TBA}}$  maps (2D histograms showing the average value of  $S2_{\text{TBA}}$  for each  $xy$  bin) for AAr (a) and UAr (b) events which pass all **CX** cuts can be seen in Figure 5.18. There are clear peaks underneath the PMTs and valleys between them.

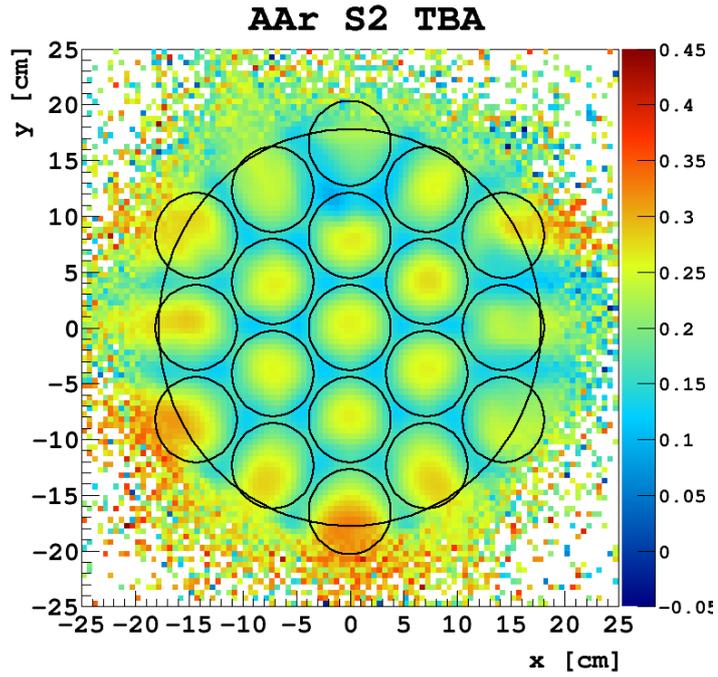
This a strong validation of the PCAMDF reconstruction, *since it is not given any bottom PMT information*. A quantitative comparison between  $S2_{\text{TBA}}$  in MC and data will be presented in Subsection 5.6.4.

### 5.5.5 UAr Event Hole

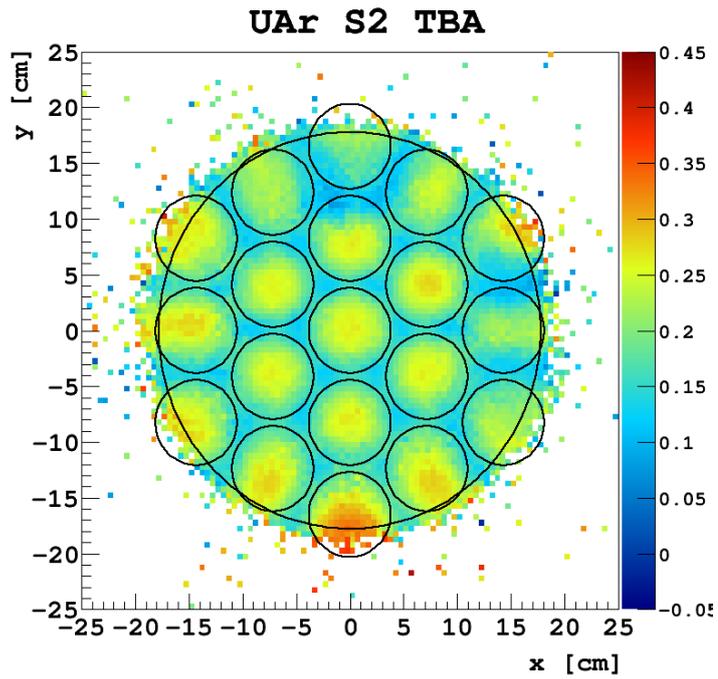
In DarkSide-50,  $xy$ -Reconstruction methods have primarily been used as detector probes, since a fiducial volume cut has not yet been necessary.  $xy$ -Reconstructions like PCAMDF have helped us learn quite a bit about our detector. Mostly, these have fallen into two categories: unusual features in event distributions, and properties of the detector which we can use to fine-tune the MC. This Subsection gives an example from the former category, and Subsection 5.5.7 gives an example of the latter.

In the data gathered during the UAr campaign, there is a rather unusual feature in the distribution of events within the detector. Figure 5.19 shows AAr (a) and UAr (b) events which pass CX cuts 1–14 and have  $t_{\text{drift}} > 300 \mu\text{s}$ . The black circle drawn on each plot shows the  $xy$ -position of PMT # 19 (on the anode) and also PMT # 2 (on the cathode). We see that there is a dearth of events near PMT # 2, relative to the other outer-ring PMTs. Rotating the detector  $-30^\circ$  and plotting  $-t_{\text{drift}}$  versus  $x_{\text{rot}}$  (where  $x_{\text{rot}}$  is the  $x$ -coordinate in the rotated coordinate system) for a small “slice” of  $y_{\text{rot}}$  gives us the plots in Figure 5.20.

Figure 5.20 shows that the density of events is dependent on all three spatial coordinates: in UAr data, there are more events at very high and very low  $t_{\text{drift}}$ , due to the radioactivity of the cathode and anode windows; and in both UAr and AAr, there are “streaks” of events throughout the detector. It seems prudent to attribute these to a bias in the PCAMDF reconstruction, but more investigation is needed. (Section 5.3 suggests that there is no such PMT-centered bias inherent in the PCAMDF  $xy$ -Reconstruction; these streaks could be caused

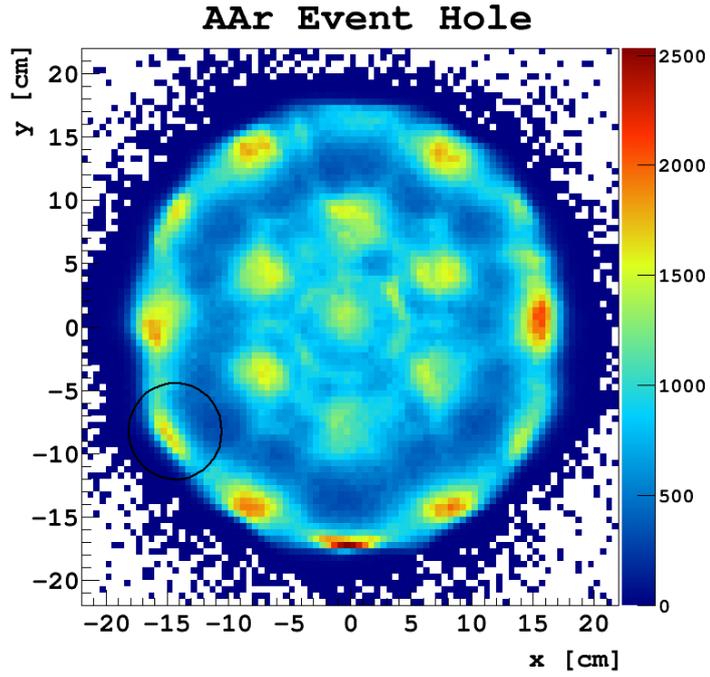


(a)

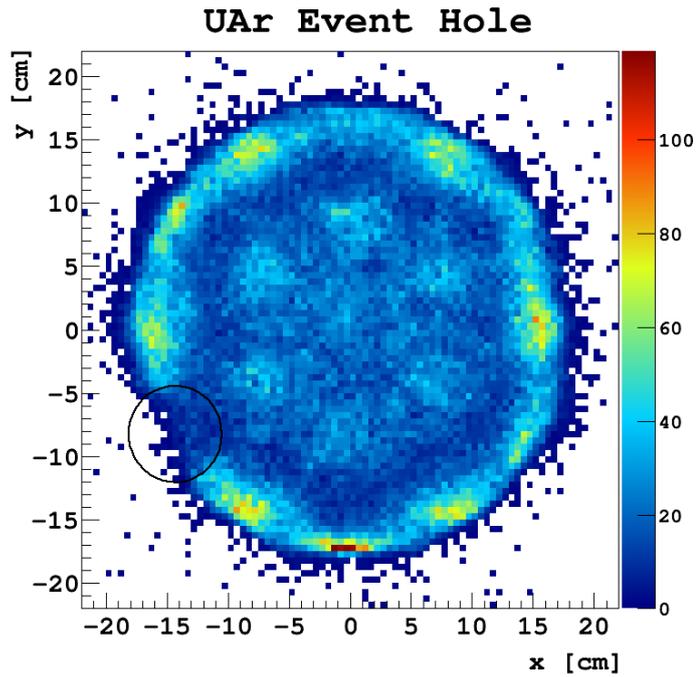


(b)

Figure 5.18: The average value of  $S2_{TBA}$  as a function of  $xy$  for AAr (a) and UAr (b) events which pass all CX cuts and have  $0 < S1_{tot} [\text{PE}] < 500$ . Note that these maps show the expected variations in  $S2_{TBA}$  described in the text, despite the fact that the position reconstruction was given no bottom PMT information.

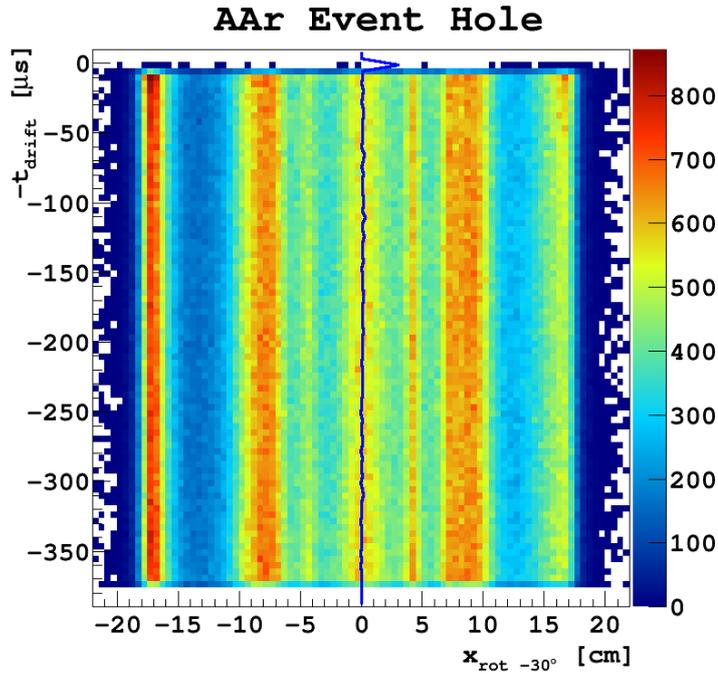


(a)

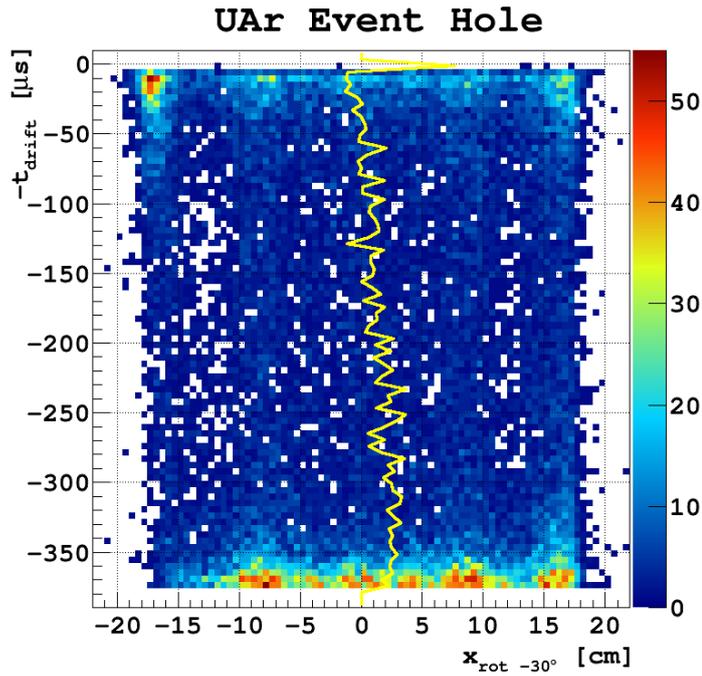


(b)

Figure 5.19: AAr (a) and UAr (b) events which pass CX cuts 1–14 and have  $t_{\text{drift}} > 300 \mu\text{s}$ . The black circle drawn on each plot shows the  $xy$ -position of PMT # 19 (on the anode) and, directly underneath it, PMT # 2 (on the cathode). Note the distinct lack of events near PMT # 2 in UAr data.



(a)



(b)

Figure 5.20: AAr (a) and UAr (b) events which pass CX cuts 1–14 and lie along the detector diameter which passes through the center of PMT # 2. The vertical axis is  $-t_{\text{drift}}$  (proportional to event depth in the  $\ell\text{Ar}$ ) and the horizontal axis is the  $x$ -position of the event after the detector has been rotated  $-30^\circ$ . The blue line (a) and the yellow line (b) show the mean  $x$ -position of the events in each  $t_{\text{drift}}$  bin. Events are constrained to lie within  $\pm 1$  cm of the  $x$ -axis of the rotated coordinate system.

by  $xy$ -dependencies of the hardware trigger, or of the DarkArt-level algorithms like the pulse-finder or baseline-finder.)

It is obvious, though, that the density of events underneath PMT # 19 (or, equivalently, the density of events above PMT # 2) is dependent on  $t_{\text{drift}}$ . As  $t_{\text{drift}}$  increases, and the depth of the event in the detector increases, it becomes less likely for that event to be reconstructed in the “PMT 19/2 column”. The average  $x_{\text{rot}}$  position as a function of  $-t_{\text{drift}}$  is shown as a blue (yellow) line for the AAr (UAr) data. We can see that  $\langle x_{\text{rot}} \rangle$  has a strong  $t_{\text{drift}}$  dependence in UAr data, but no noticeable dependence in AAr.

This feature has been investigated thoroughly and its main characteristics (aside from those given above) are as follows:

1. There is no “lack” of events near PMT # 2. The number of events with `S1Mc==2` (where `S1Mc`, “S1 Max channel”, is the channel index of the PMT which sees the largest fraction of S1 light) is roughly equivalent to the number for any other PMT. This suggests that these events aren’t “missing”, but just mis-reconstructed by the  $xy$  for some reason.
2. The size of S2/S1 is not distorted for events with `S1Mc==2`, suggesting that their S2 light is not “lost”, it’s just being distorted somehow.
3. An S1-based PCAMDF  $xy$  reconstructs near-cathode events inside the hole. This (a) means that the S1 of these events is not distorted, and (b) supports point (1) that there are indeed events occurring in this volume, but they’re being mis-reconstructed by the S2-based  $xy$ ’s. (See the next Subsection.)

All evidence to date suggests an electrostatic distortion. If the drift field lines in the PMT 19/2 column are distorted, ionization electrons could be dispersed, spreading out the S2 signal and confusing the  $xy$ -Reconstruction. The S1-based  $xy$  discovered that these events also have  $S2_{\text{TBA}}$  values which are similar to events

which occur *between* the anode PMTs: this is another piece of evidence for a distortion of the drift field lines.

SIMION<sup>®</sup> Ion and Electron Optics Simulator software was used to simulate the paths of drift electrons within the TPC when electric field distortions are present. With a grounded cryostat and field cage rings held at stepped potentials to maintain a constant drift field throughout the active volume, positive (+600 V) and negative (−600 V) field distortions on the cathode were simulated.<sup>15</sup> The positive distortion at the cathode gives a clean hole of events at the anode, while the negative distortion gives a hole with scattered events. Additionally, the positive hole shows no distortion for events which start just 50 mm above the cathode. Since Figure 5.20 shows event distribution distortions throughout all  $t_{\text{drift}}$ , this study suggests that the event hole in UAr is due to a negatively-charged field distortion on the cathode.

But these events are not yet fully understood. The size of the event hole remained stable throughout the 70-Days campaign and then shrank dramatically during a 3-day <sup>22</sup>Na calibration run, but has again remained stable in the ten months since. The size of the hole also depends on the strength of the drift field: as the field strength decreases, the hole size increases. There is still much to investigate about these events, but it seems unlikely that they would have been discovered if not for the *xy*-Reconstruction: they do not seem to stand out dramatically in any sense except in their unusual *xy* positions.

### 5.5.6 An S1-Based PCAMDF *xy*-Reconstruction

During the investigation of the events covered in the previous Subsection, it was thought that the implementation of an S1-based *xy* might be beneficial. It was hypothesized that there is a class of events which occur in a specific area near the

---

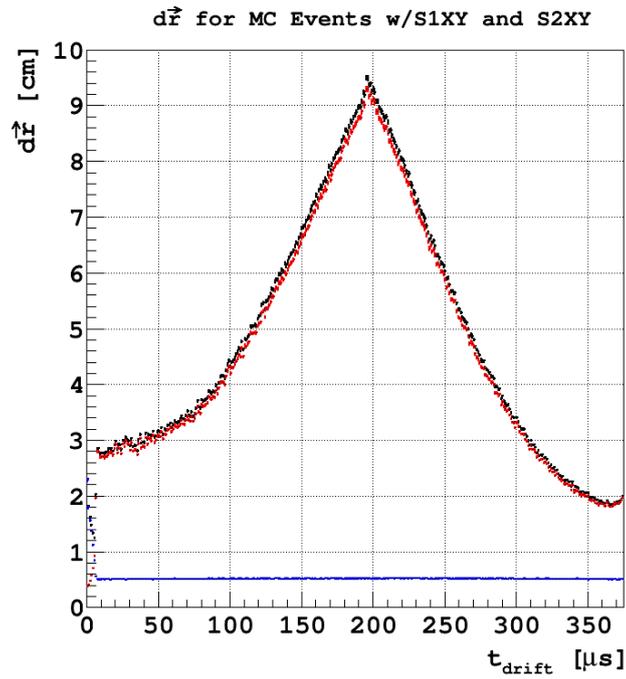
<sup>15</sup> See internal docdb entry # 1608.

cathode, but which have their drift electrons dispersed due to some electrostatic effect on the way to the gas pocket. Since an S1-based  $xy$  would be unaffected by a detector flaw like this, it could provide a useful cross-check on these event positions.

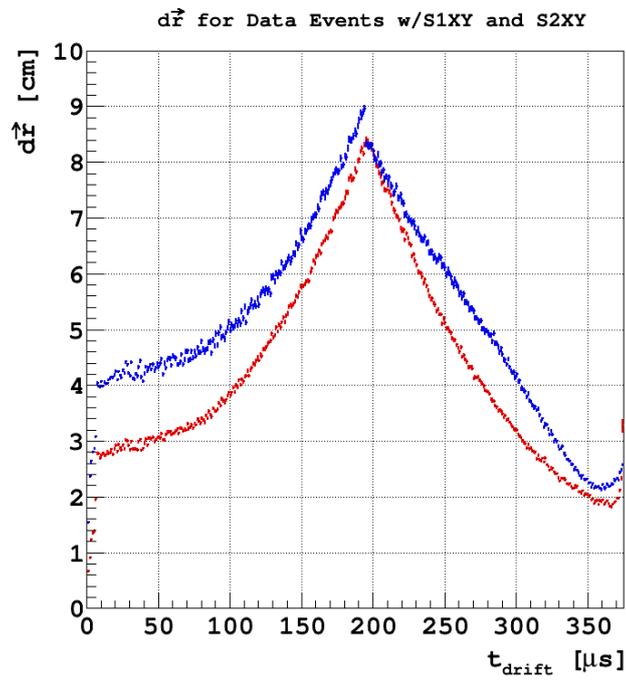
The S1-based  $xy$ -Reconstruction is trained just like the standard S2-based  $xy$ , except — because the amount of position information contained within the S1 fractions depends strongly on the  $z$ -position of an event — the detector is broken up into many “slices” of  $t_{\text{drift}}$ . An independent PCAMDF training is performed on each of these slices. Because the low PE statistics MC set has more than enough events to provide a 10,000-event training sample for each integral value of  $t_{\text{drift}}$ , this is precisely what was done: 375 independent S1-based PCAMDF  $xy$ -Reconstructions were tuned, one for every  $\mu\text{s}$  of  $t_{\text{drift}}$  in the detector. The result for MC (a) and data (b) can be seen in Figure 5.21.

The S1-based PCAMDF reconstruction performs much worse on MC events, relative to the S2-based reconstruction, as can be seen in Figure 5.21 (a). Here, the blue line gives the average distance (for each  $1 \mu\text{s}$  bin in  $t_{\text{drift}}$ ) between an event’s true MC position and its position when reconstructed with the S2-based PCAMDF. The red line gives the same quantity, but with the S1-based PCAMDF. And the black line gives the average distance (again, as a function of  $t_{\text{drift}}$ ) between the S2-based and S1-based positions. Note that the black line closely follows the red line, since the S1-based reconstruction is much less accurate than the S2-based reconstruction.

At its best, the S1-based  $xy$  gives an average reconstruction error of just under 2 cm, for MC events near the cathode (high  $t_{\text{drift}}$ ). At its worst, near the center of the detector, the S1-based  $xy$  reconstructs events — on average — about 9 cm away from their true positions. This is larger than the PMT-PMT spacing. Accordingly, the S1-based  $xy$ -Reconstruction is most useful near the cathode and the anode.



(a)



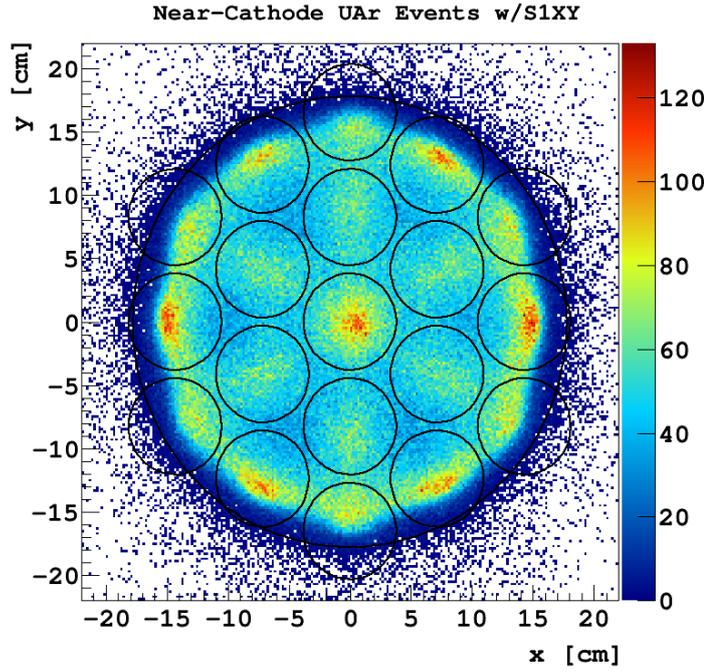
(b)

Figure 5.21: Comparison of S1 and S2-based PCAMDF  $xy$ -Reconstructions for MC (a) and data (b) events. Further description in text.

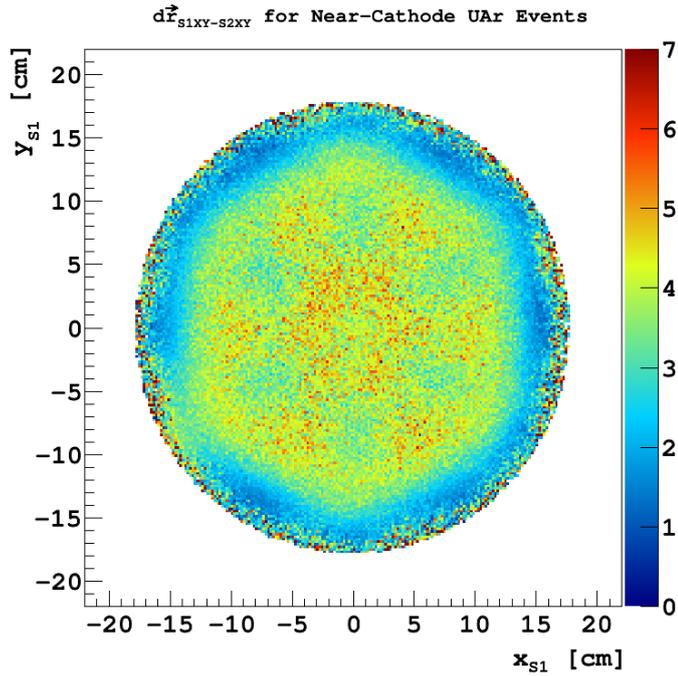
Figure 5.21 (b) shows the average distance between the positions of data events reconstructed with the S2-based  $xy$ , relative to their positions with the S1-based  $xy$  (similar to the black line in Figure 5.21 (a)). On this plot, the blue line represents UAr data, while the red line represents AAr data. Note that the AAr curve is very similar in shape to the black MC curve from Figure 5.21 (a), though the average value near the center of the detector is slightly smaller in data. There is a discontinuity in the blue (UAr) curve due to the way the S1-based  $xy$  was trained on MC events. The MC training sample is entirely DM S1-range ERs, and since the AAr data set is dominated by the  $^{39}\text{Ar}$  background, it “looks like” the training sample more than the UAr data set, which has a higher proportion of “unusual events” like wall events, external gamma ray events, multi-sited events, etc. The peak around  $t_{\text{drift}} = 200 \mu\text{s}$  is due to the S1-based reconstruction switching from using the anode PMTs to the cathode PMTs at  $t_{\text{drift}} = 195 \mu\text{s}$ .

The  $xy$  positions of all “near-cathode” ( $t_{\text{drift}} > 350 \mu\text{s}$ ) events from the UAr data set, reconstructed with both of these methods, can be seen in Figure 5.22 (a). The hole which appears in Figure 5.19 with the S2-based  $xy$ -Reconstruction is completely “filled in” with the S1-based reconstruction. Figure 5.22 (b) shows the average distance between the S1-based reconstruction and the S2-based reconstruction over each  $xy$  bin. Note that there is only a single strong asymmetrical feature, near PMT # 2. This shows that the S1  $xy$  and the S2  $xy$  disagree with each other to a greater extent in that region of the detector than they do, for instance, near any of the other outer-ring PMTs.

Finally, the S1-based  $xy$  also recreates the  $\text{S2}_{\text{TBA}}$  map, as can be seen in Figure 5.23, although it has noticeably poorer resolution than the same map made with the S2-based  $xy$ . In spite of this, the general pattern of the map can still be seen, as well as the “TPB dead spots” (discussed in the next Subsection) at the top and to the right. The S1-based  $xy$  is still being developed and refined, though it



(a) “Near-cathode” ( $t_{\text{drift}} > 350 \mu\text{s}$ ) events from UAr data, reconstructed with an S1-based PCAMDF  $xy$ .



(b) The average distance between the S1-based reconstruction and the S2-based reconstruction for the “near-cathode” events shown in Figure 5.22 (a).

Figure 5.22: “Near-cathode” ( $t_{\text{drift}} > 350 \mu\text{s}$ ) events from UAr data, reconstructed with an S1-based PCAMDF  $xy$ .

promises to be a very helpful investigatory tool for near-cathode and near-anode events.

### 5.5.7 S2 Corrections

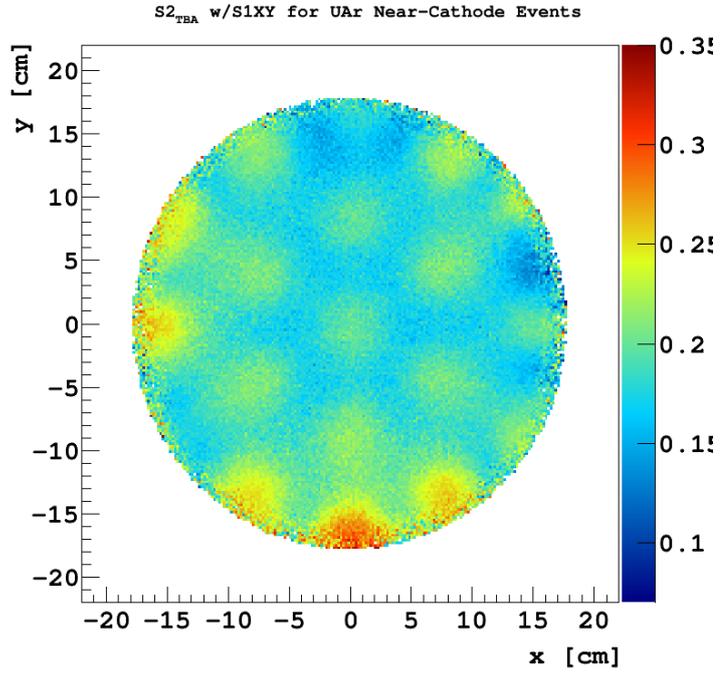
$xy$ -Reconstruction methods have also discovered features of the DS-50 detector which have been used to better tune g4ds to match the data. One of the most significant discoveries is the dependence of  $S2_{\text{tot}}$  on the  $xy$  position of an event in the detector. [Appendix C](#) shows, step by step, how the detector tilt is found in ARIS data.<sup>16</sup> A slight TPC tilt increases the thickness of the gas pocket on one side of the detector and decreases it on the other, resulting in larger average S2's on the side with the thicker gas pocket.

Other position dependencies of S2 have been found in DarkSide-50 including, most notably, a radial dependence of the size of S2. This S2 radial dependence is believed to be caused by a deformation of the materials which maintain the extraction and electroluminescence fields — namely, the extraction grid and/or the diving bell (see [Figure 3.3](#)). It results in a substantially higher S2/S1 ratio ( $\sim 4\times$  higher) ratio for events near the detector axis, relative to events near the edge (the cylindrical wall) of the TPC. The dependence of the extraction field strength (and the ratio of S2/S1) on the height of the gas pocket has been modeled by Chengliang Zhu of Princeton University and will surely appear in an upcoming publication.<sup>17</sup> His study suggests that, to achieve a  $\sim 4\times$  higher S2/S1 ratio near the center of the TPC relative to the edge, the difference in gas pocket height would need to be about 6 mm between the center of the TPC and the edge. This is not an unreasonable deformation of the materials involved, though a much more

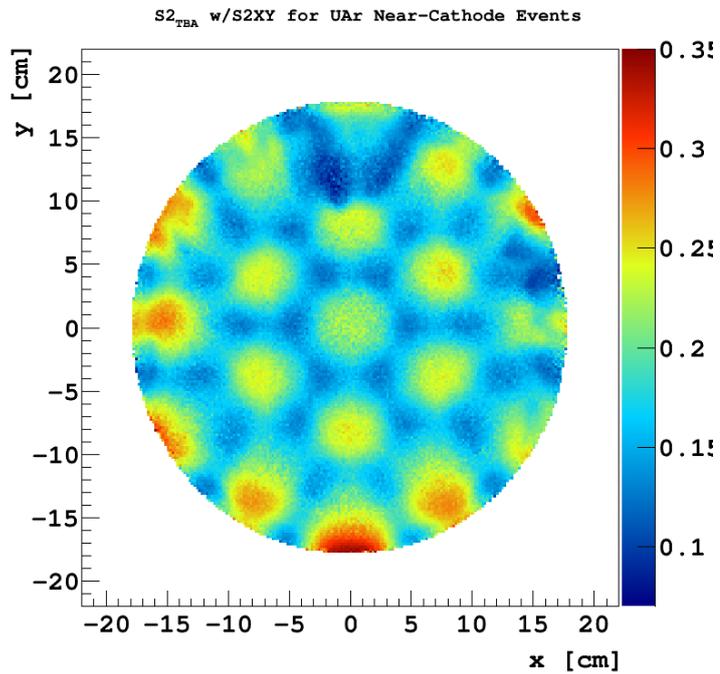
---

<sup>16</sup> ARIS data is used to illustrate how to normalize  $t_{\text{drift}}$ -dependent S2 (as well as S1) dependencies, which aren't significant in DS-50, due to the very high electropositive purity of the  $\ell\text{Ar}$ .

<sup>17</sup> For now, see internal docdb entry # 1287.



(a)



(b)

Figure 5.23:  $S2_{TBA}$  maps for UAr data with an S1-based (a) and an S2-based (b) PCAMDF  $xy$ . “Near-cathode” events only.

thorough analysis of this will be available soon.

This behavior can be seen in Figure 5.24 which shows  $S2_{\text{tot}}/S1_{\text{corr}}$  ( $S1$  with the standard  $z$ -correction) as a function of the event radius, normalized for the geometry of the detector. The black crosses which show the vertical means of the horizontal bins are almost entirely hidden by the blue line showing a 2nd-order polynomial fit.<sup>18</sup> The fit parameters from the AAr (a) and UAr (b) curves agree quite well, suggesting that this isn't a transient phenomenon in the detector. We can normalize this radial dependence by dividing out the curve fit; the result can be seen in Figure 5.25.

The radially-corrected  $S2_{\text{rcorr}}$  can be seen for AAr (a) and UAr (b) in Figure 5.25. Note that there are rather prominent “bumps” in the mean profile still; these bumps occur underneath the PMTs, as shown in Figure 5.26.

These under-PMT  $S2/S1$  “bumps” are at least partially caused by the  $xy$  dependence of the detection probability for a UV photon generated in the gas pocket. Figure 5.27 shows the detection probability for MC UV photons generated uniformly throughout the  $g\text{Ar}$ ; it shows that photons generated underneath PMTs (especially near the edge of the TPC) are more likely to be detected than photons generated elsewhere. This not only affects the  $S2/S1$  distribution across  $xy$ , but also lends credence to the theory that non-uniformities in the density of events reconstructed across  $xy$  are due to an  $xy$ -dependent event detection probability, rather than biases in the PCAMDF reconstruction. However, this is still an area of active research. Dividing out the UV photon detection probability gives us a

---

<sup>18</sup> Note that this is only an approximation: the true functional form of the radial dependence requires (1) the shape of the deformation of the diving bell, (2) the electric field dependence as a function of that shape, and (3) the  $S2$  LY as a function of the electroluminescence field strength. This analysis requires enough work to fill an entire dissertation by itself and will not be covered here, though it will surely be published soon in some form by Chengliang Zhu of Princeton University.

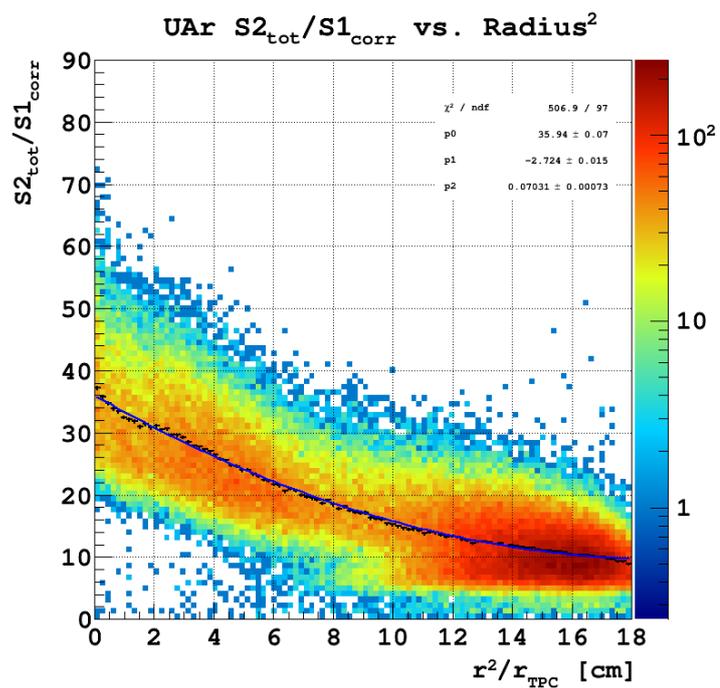
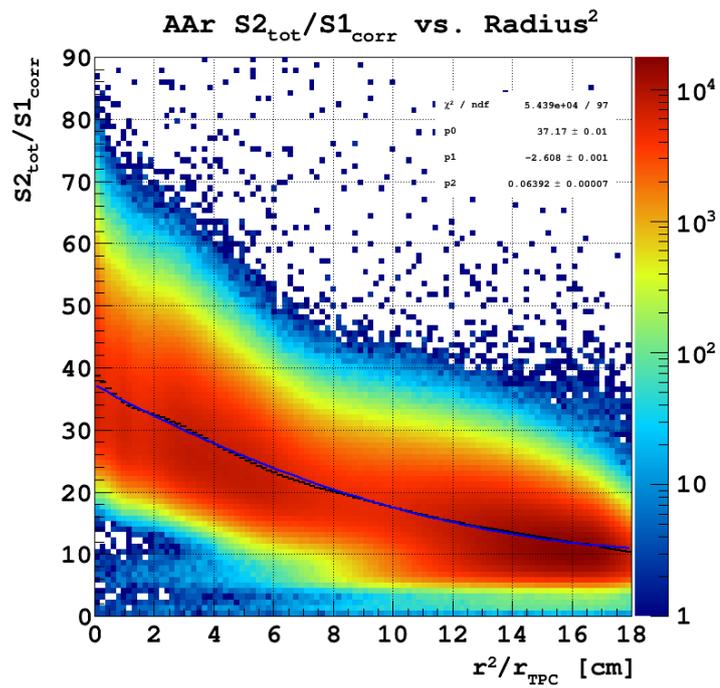
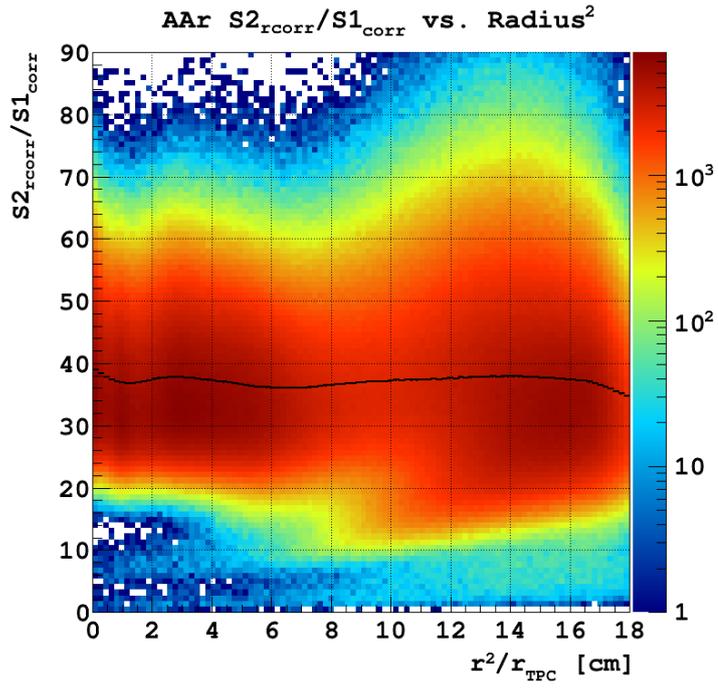
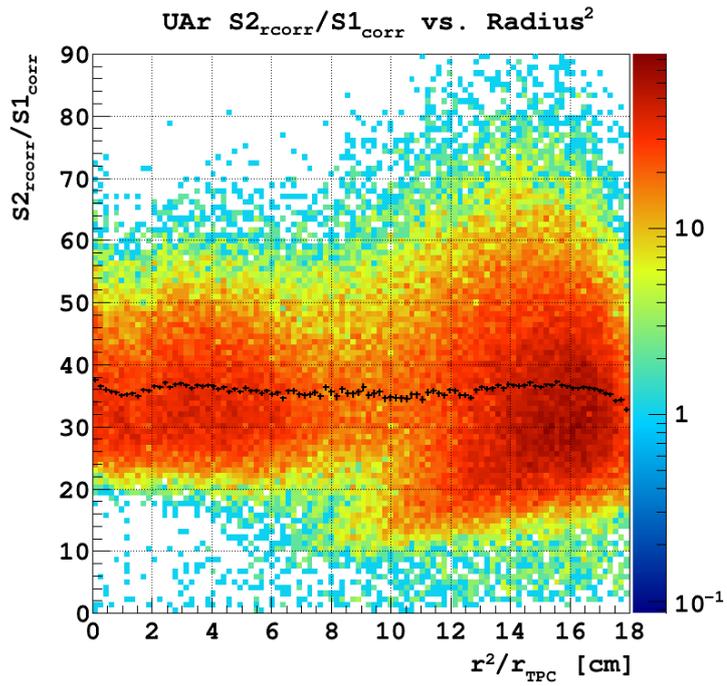


Figure 5.24:  $S2/S1$  as a function of radius for AAr (a) and UAr (b) data. The blue 2nd order polynomial curve fits almost completely obscure the black crosses, which show the vertical means of the horizontal bins.

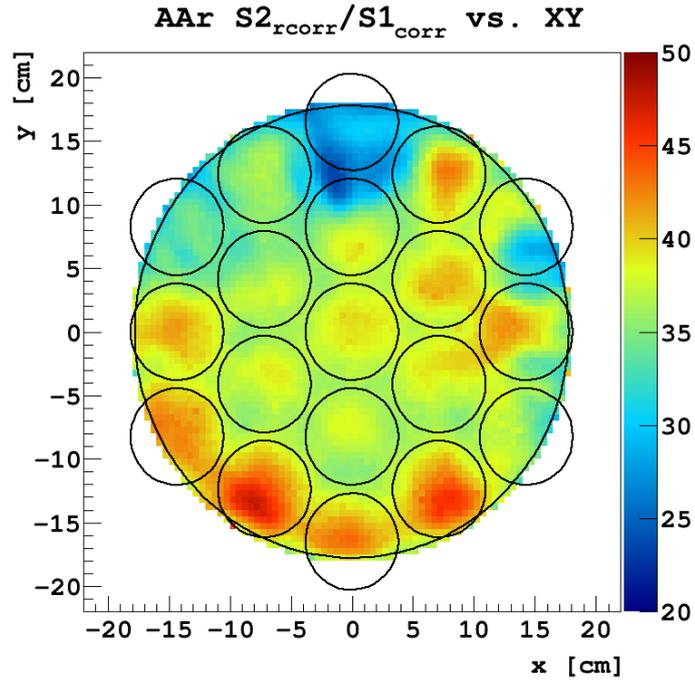


(a)  $S2_{\text{corr}}/S1$  as a function of radius for AAr data, with a crude radial correction on S2.

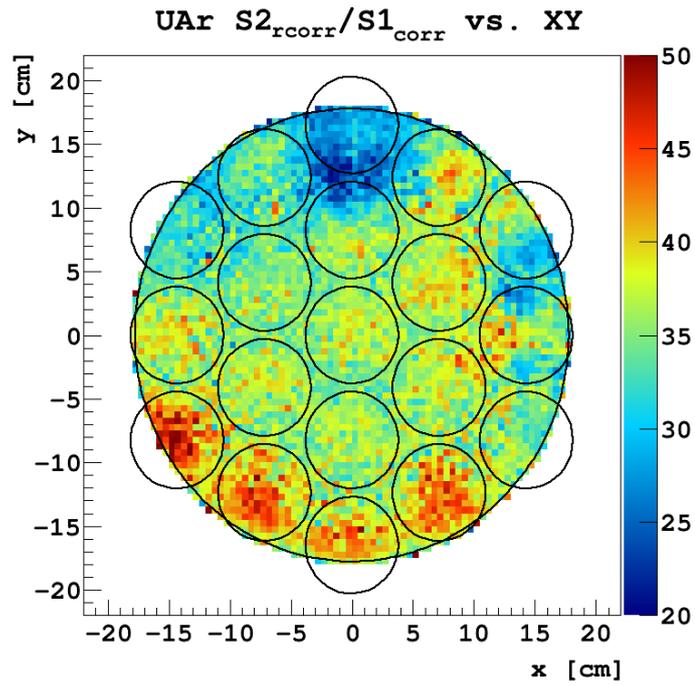


(b)  $S2_{\text{corr}}/S1$  as a function of radius for UAr data, with a crude radial correction on S2.

Figure 5.25: A crude radial S2 correction for AAr (a) and UAr (b) data.  $S2_{\text{corr}}$  is radially-corrected by dividing out the curve fits shown in Figure 5.24.



(a)  $S2_{\text{rcorr}}/S1$  for AAr data mapped onto  $xy$ , with the  $S2$  radial correction from Figure 5.24 (a).



(b)  $S2_{\text{rcorr}}/S1$  for UAr data mapped onto  $xy$ , with the  $S2$  radial correction from Figure 5.24 (b).

Figure 5.26:  $S2_{\text{rcorr}}/S1$  for AAr (a) and UAr (b) data mapped onto  $xy$ , with the  $S2$  radial corrections from Figure 5.24. The  $z$ -value of each bin gives the average value of  $S2_{\text{rcorr}}/S1$  for all events in that bin. Note the similarity of the low- $S2_{\text{rcorr}}/S1$  regions between AAr and UAr.

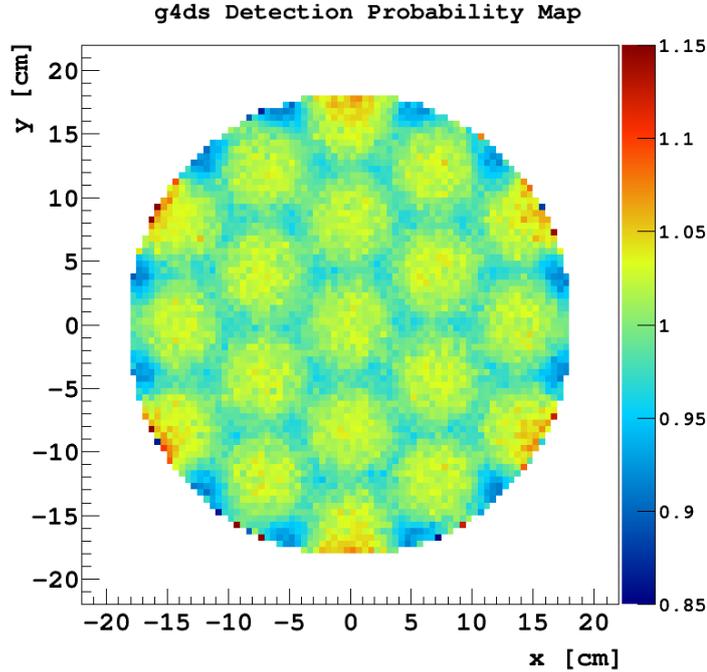


Figure 5.27: g4ds detection probability map for  $\sim 2.4 \times 10^8$  UV photons generated uniformly throughout the *g*Ar. Blue areas show where UV photons generated in the gas pocket (e.g. S2 electroluminescence photons) are less likely to be detected, and red areas show where they are more likely to be detected.

near-final S2/S1 map, which can be seen in Figure 5.28.

Figure 5.28 shows the average (per  $xy$  bin) S2/S1 as a function of  $xy$ , after an approximate radial correction and a UV detection probability correction. Note that a tilt correction could also be made, but it has only a minor effect. We can note a few things about these maps. First, by and large, the AAr and UAr maps are qualitatively very similar. The statistics in the UAr set are much lower and so the shapes of the “blobs” in the map are blurrier, but there is a big blue spot at the top in the center, and a dark red band near the southwest side of the map.

Blue areas on these maps are places where the S2 is small, or the S1 is large, or both. And the red areas are regions where the S2 is large, or the S1 is small, or both.  $^{83\text{m}}\text{Kr}$  calibrations (see Subsection 3.4.2) have shown that the S1  $xy$ -dependence is negligible, so we can assume that any variation here is a result

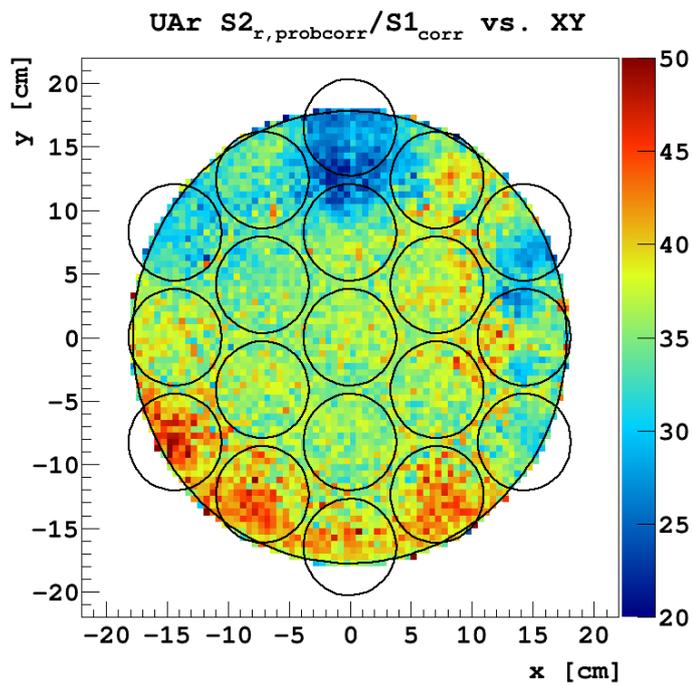
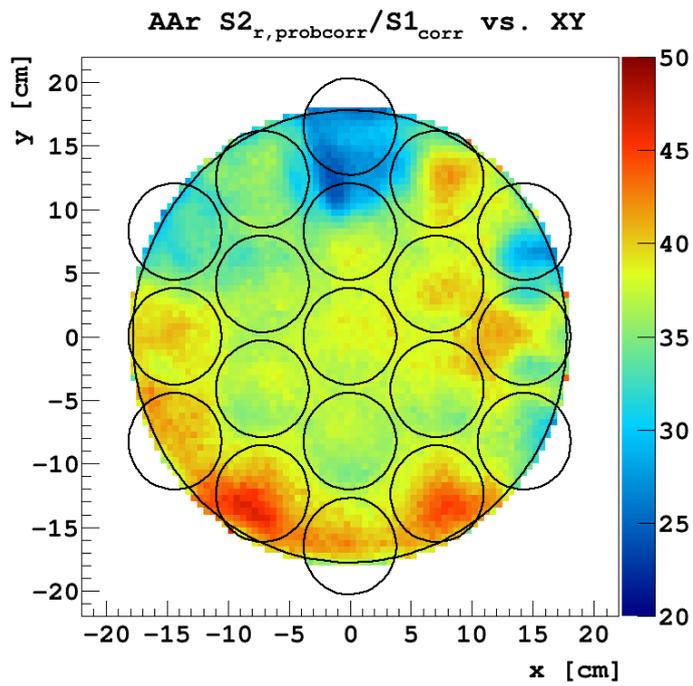


Figure 5.28:  $S2/S1$  “residual maps” for AAr (a) and UAr (b) data after radial  $S2$  corrections and g4ds probability map corrections. Note the distinct blue areas at the top and right, which are assumed to be due to reduced TPB thicknesses in those areas.

of fluctuations in the average value of S2 across  $xy$ . But no obvious electrostatic effect should produce such drastic variations across  $xy$ , sometimes with very sharp changes across very short distances.

The current theory is that these variations in S2/S1 are due to non-uniformities in the TPB coating across the anode window. As noted in Section 3.2, the changes in TPB thickness across the windows are a known phenomenon, but the thickness has not been mapped out as a function of  $xy$ . It seems likely that there could be small areas with a thinner or thicker coating, or that some of the coating may have been accidentally rubbed off during handling when the detector was being commissioned.

Note that while a detector tilt or a radial dependence of the S2/S1 light production should only affect the size (that is, the integral in PE) of  $S2_{\text{tot}}$ , it should not affect the distribution of S2 light across the anode PMTs. Since the S2 light is generated in a column through the gas pocket, at only a single  $xy$  position, any phenomena which affect the *generation* of S2 should have a negligible effect on the S2[] fractions. Non-uniformities of the TPB across  $xy$ , though, would affect the *detection* of S2<sup>19</sup> and so could affect the S2[] fractions of an event generated at any position in  $xy$ .

This leads to a chicken-and-egg-type problem, where we need to implement the TPB non-uniformities in the MC in order to generate events to train the PCAMDF, but the S2 residual map (Figure 5.28) generated from those reconstructed data events needs to be implemented in g4ds in order to generate events with accurate S2[] fractions. The S2[] fractions need to be accurate in order to

---

<sup>19</sup> The S2 light is *generated* at a single  $xy$  position, but bounces around the detector after it's created. If one PMT has a thinner coating of TPB on the window in front of it, relative to another PMT, then the former PMT will see fewer UV photons be wavelength-shifted by the TPB in front of it. This means fewer wavelength-shifted photons will be generated directly in front of this PMT, where it has the highest probability of detecting them.

reconstruct the data events, and so on. Unfortunately, an iterative approach, like the one described above, did not perform well during tests.<sup>20</sup> There is still much work to be done on this front.

## 5.6 Quantifying $xy$ Performance

For as long as  $xy$ -Reconstructions have been used on data in DarkSide-50, we’ve been asking ourselves, “how can we validate it?” This is really a two-pronged question: (1) how can we judge the *accuracy* of the method, and (2) how can we judge its *precision*? “Accuracy” implies that we *know* the true position of an event and can calculate how close the PCAMDF  $xy$  comes to reconstructing the event at that position. Because of the “nesting doll” structure of DS-50, we cannot generate a sample of events with known positions, for example by moving an LED to a predetermined spot within the detector. We also cannot perform a “flood exposure”,<sup>21</sup> which is a common technique for finding distortions in imaging systems like the DS-50 PMT arrays. Although we can perform accuracy tests on MC events with reconstructed  $xy$  positions, we cannot be sure that these accurately reflect the performance of the PCAMDF on data events, which may differ from MC events in ways which we haven’t yet thought of.

Initially, there were two main tests for the “goodness” of an  $xy$ -Reconstruction: (1) how well it performed on MC events, and (2) how uniformly it reconstructed AAr events. The first test is straightforward: we can quantify how well an  $xy$  is performing when we run it on events with known positions. But, as mentioned earlier, we can’t be sure that this translates to accuracy in reconstructing data events. Tests like these for PCAMDF were covered in Section 5.3.

---

<sup>20</sup> This approach led to events being reconstructed unphysically far away from the walls, being pulled in very strongly toward the axis of the detector.

<sup>21</sup> Uniformly and brightly illuminating an entire cross-section of the detector to find where there are distortions in the image.

The second test is a bit more nuanced. While similar in spirit to a flood exposure-type measurement, the expected result is not as clear cut.  $^{39}\text{Ar}$  should be uniformly distributed throughout the  $\ell\text{Ar}$ , but do we expect to *detect* it uniformly? Figure 5.27 suggests that there are  $xy$ -dependent photon detection probabilities, which could be amplified by as-yet-unknown biases in the lower-level event reconstruction algorithms, leading to an  $xy$ -dependent event detection probability. Enforcing uniformity over  $xy$  of reconstructed events also seems to lead to geometric artifacts in the final event distribution: “weaker” events with larger position uncertainties end up “clustered” into small regions of the detector, sometimes into tiny  $\text{mm}^2$ -scale regions which only become apparent when quality cuts are made.<sup>22</sup>

There are several analyses which can judge the precision of a given  $xy$ -Reconstruction. If we have two events in the detector with similar  $\text{S2}[\ ]$  values — that is, similar  $\text{S2}$  distributions over the anode — we would naturally expect those events to be reconstructed very near to each other in  $xy$ , and vice versa; there should be little, if any, degeneracy in the transformation from  $\text{S2}[\ ]$ -space to  $xy$  space due to the geometry of the detector. The most compelling of this type of analysis is the reconstruction of coincident decays (“BiPos” and “RnPos”), for which there is a physical reason why we expect these pairs of events to be reconstructed near each other in  $xy$ ; these events will be covered in Subsection 5.6.2. An additional  $xy$  validation, using  $^{57}\text{Co}$  events from a CALIS calibration run, is also proposed. While this test serves as a proof of concept, the event statistics are too low to actually use it for validation of the reconstruction method as intended. However, further study of these  $^{57}\text{Co}$  events does yield a measurement of reconstruction accuracy for high-radius events. These studies are covered in detail in Subsection 5.6.3. Finally, as  $\text{S2}_{\text{TBA}}$  is an emergent phenomenon of the  $\text{S2}$ -based PCAMDF  $xy$ -Reconstruction, which only uses anode PMT information, the distribution of

---

<sup>22</sup> From internal docdb entry # 1570.

events within specific ranges of  $S2_{\text{TBA}}$  across  $xy$  can be used to calculate preliminary figures for the accuracy of the reconstruction on data over limited areas in  $xy$ . This study is presented in Subsection 5.6.4. First, however, it's worth discussing statistical limits on the accuracy of *any*  $xy$ -Reconstruction in DarkSide-50.

### 5.6.1 Resolution Without $xy$ -Reconstruction

PE statistics necessarily put a limit on the precision of any  $xy$ -Reconstruction: the fewer PEs, the less information is available to the reconstruction algorithm. Events which occur very near to each other in  $xy$  will give very similar S2 fractions over the anode PMTs, and below some threshold number of PEs, the PE number distributions over the anode PMTs will be statistically indistinguishable.

An optical MC was used to simulate groups (1000 events in each group) of very high PE-statistics events ( $\sim 10^6$  M PE) along the  $x$ -axis of the TPC, with either 1 cm or 2.5 mm spacing. These MC events have negligible statistical fluctuations, so we can use them to test the hypothesis that the S2 light fractions on the anode PMTs are *uniquely* determined by the  $xy$  position of a given event.

To be able to resolve two nearby events (call them event  $i$  and event  $i + 1$ ), the  $\chi^2$  difference between their S2[] patterns must be large enough that we can rule out statistical fluctuations as the cause of any differences in PE number:

$$\chi_{i,i+1}^2 = \sum_{j=1}^{N_D} \frac{\left(N_{PE} \cdot S2_i^{(j)} - N_{PE} \cdot S2_{i+1}^{(j)}\right)^2}{N_{PE} \cdot S2_i^{(j)}} \quad (5.4)$$

where  $N_D$  is the number of dimensions (PMTs) used in the study,  $N_{PE}$  is the number of PEs seen by all PMTs considered (and is roughly the same for both events  $i$  and  $i + 1$ ), and  $S2_i^{(j)}$  is the S2[] value on the  $j$ -th PMT for the  $i$ -th event (recall that this is this PMT's fraction of the total S2, so it is unitless). For this particular study, the 19 anode PMTs were used, plus the cathode PMTs were

combined into a single “mega PMT”, meaning  $N_D = 20$  and the number of degrees of freedom is 20. In that case:

$$\chi_{i,i+1}^2 \approx N_{PE} \cdot \left( \sum_{j=19}^{37} \frac{(S2_i^{(j)} - S2_{i+1}^{(j)})^2}{S2_i^{(j)}} + \frac{(S2_i^{\text{bot}} - S2_{i+1}^{\text{bot}})^2}{S2_i^{\text{bot}}} \right) \quad (5.5)$$

where the value is approximate as  $N_{PE}$  may be different for event  $i$  and event  $i+1$ , though for the events used, the difference is usually  $<$  a few %. (For this study, the average value of the two events was used.)

For a  $p$ -value of 0.05 (a 95% probability that the differences in  $S2[]$  values are *not* due to statistical fluctuations), and 20 degrees of freedom, the  $\chi_{i,i+1}^2$  must be greater than 31.4, or

$$N_{PE} \cdot \left( \sum_{j=19}^{37} \frac{(S2_i^{(j)} - S2_{i+1}^{(j)})^2}{S2_i^{(j)}} + \frac{(S2_i^{\text{bot}} - S2_{i+1}^{\text{bot}})^2}{S2_i^{\text{bot}}} \right) \geq 31.4 \quad (5.6)$$

Note that, even though including the bottom PMTs increases the left side of the above equation, improving “distinguishability” of events (and presumably improving the performance of a given  $xy$ -Reconstruction), the bottom PMTs are left out of the PCAMDF reconstruction in order to assess the performance of the method using the  $S2_{\text{TBA}}$ , as will be described at the end of this Chapter. The above equation can be rewritten as

$$N_{PE} \geq \frac{31.4}{X}, \quad \text{where } X = \left( \sum_{j=19}^{37} \frac{(S2_i^{(j)} - S2_{i+1}^{(j)})^2}{S2_i^{(j)}} + \frac{(S2_i^{\text{bot}} - S2_{i+1}^{\text{bot}})^2}{S2_i^{\text{bot}}} \right) \quad (5.7)$$

For two events to be statistically distinguishable at 95% certainty, where each event has  $S2_{\text{tot}} \approx N_{PE}$ ,  $N_{PE}$  must satisfy the above equation. This gives a limit on the minimum number of  $S2_{\text{tot}}$  PE required for “distinguishability” of events, hence

the name of this study, “Resolution Without  $xy$ -Reconstruction”. The result of this analysis for events generated in a line along the  $x$ -axis of the detector can be seen in Figure 5.29.

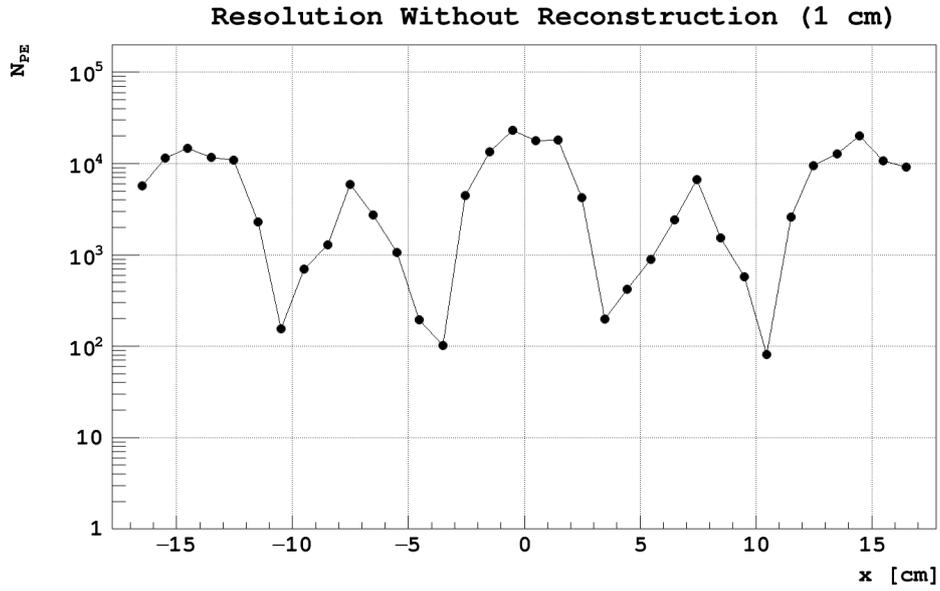
Both subfigures of Figure 5.29 show that more PEs are required for statistical distinguishability underneath the anode PMTs and near the wall of the TPC, while fewer PEs are required for distinguishability *between* PMTs. Figure 5.29 also shows that, as one might expect, for better resolution, more PEs are required: the average number of  $S2_{\text{tot}}$  PEs required to reconstruct events with 2.5 mm resolution underneath the leftmost PMT (PMT# 27, see Figure 3.4) is about  $7 \times 10^4$  PE  $\rightarrow 8 \times 10^4$  PE, while at 1 cm resolution, it’s only about  $1 \times 10^4$  PE. ER events at the high end of the DM S1 range have  $S2_{\text{tot}}$  values  $\lesssim 4 \times 10^4$  PE. We can conclude from this study, then, that  $xy$  resolution should be worst under the centers of PMTs and near the wall of the TPC, and that the best average resolution we can hope for, near the edge of the TPC, is likely  $\sim 1$  cm. However, around the circumferences of PMTs, for events with  $S2_{\text{tot}}$  values larger than  $\sim 1$  k PE, we can expect resolution better than  $\sim 2.5$  mm. These values are in agreement with the studies presented in the following Subsections.

### 5.6.2 Coincident Decays: BiPos and RnPos

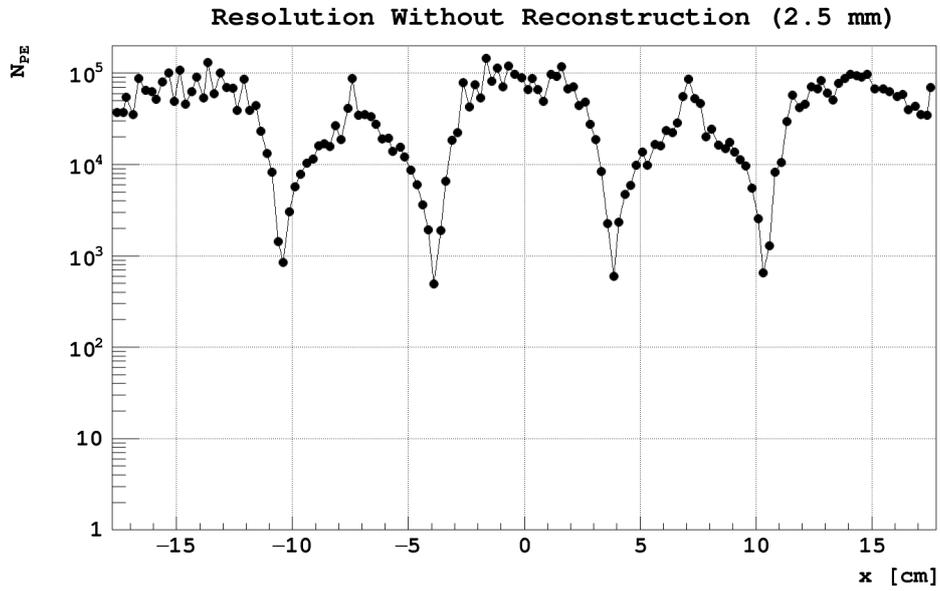
In the upper half of the  $^{226}\text{Ra}$  decay chain (see Figure 5.30 (a)<sup>23</sup>),  $^{214}\text{Bi}$   $\beta$  decays to  $^{214}\text{Po}$  with a half-life of 19.9 min. That  $^{214}\text{Po}$  then  $\alpha$  decays to  $^{210}\text{Pb}$  with a half-life of 164.3  $\mu\text{s}$ . Even though DarkSide-50 was assembled in a radon-free clean room, there is some residual  $^{226}\text{Ra}$  in the system which leads to this decay chain. The consecutive  $^{214}\text{Bi}$  and  $^{214}\text{Po}$  (“BiPo”) decays are particularly useful for  $xy$  validation in DS-50.

---

<sup>23</sup> Licensed under a Creative Commons Attribution 3.0 Unported License. By Tosaka. No modifications have been made to the original image. <http://bit.do/226Ra>



(a) Number of PE required for statistical distinguishability of events, as a function of the position of an event along the  $x$ -axis of the detector at 1 cm event spacing.



(b) Number of PE required for statistical distinguishability of events, as a function of the position of an event along the  $x$ -axis of the detector at 2.5 mm event spacing.

Figure 5.29: Number of PE required for statistical distinguishability of events, as a function of the position of an event along the  $x$ -axis of the detector at (a) 1 cm event spacing and (b) 2.5 mm event spacing.



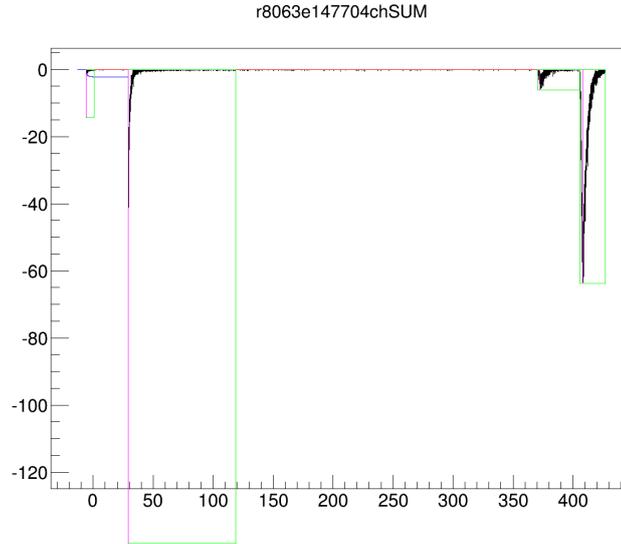


Figure 5.31: A typical “BiPo” event waveform. Note the bismuth S1-S2 pair at the beginning of the acquisition window, and the polonium pair at the end. The drift time between each S1 and its corresponding S2 is roughly the same.

decays via electron emission, we require the  $f_{90}$  of the earliest S1 pulse to be ER-like. Finally, as we expect these events to occur at roughly the same  $t_{\text{drift}}$ , we require their  $t_{\text{drift}}$ ’s to match within some margin of error (90% have the same  $t_{\text{drift}}$  within  $\pm 1.5 \mu\text{s}$ ). A typical BiPo event waveform can be seen in Figure 5.31. Note that a similar analysis of these events was conducted in Brodsky (2015), while the event selection was performed by Chris Stanford at Princeton University.

The results of the BiPo analysis with the PCAMDF can be seen in Figure 5.32, which shows the distance between each bismuth decay and its corresponding polonium decay. The histogram has a mean of 7.5 mm with a 7.3 mm RMS. The distances appear to be normally distributed, and a Gaussian curve fit to them gives a standard deviation of  $8.2 \pm 2.3$  mm. Large samples of normally distributed data exhibit the relationship  $\langle \text{drvec} \rangle = \sqrt{2/\pi} \cdot \sigma$ ,<sup>24</sup> where  $\langle \text{drvec} \rangle$  is the mean deviation (the average distance between true and reconstructed position). If we assume that there is no systematic shift which would bias the reconstruction in

<sup>24</sup> <http://mathworld.wolfram.com/MeanDeviation.html>

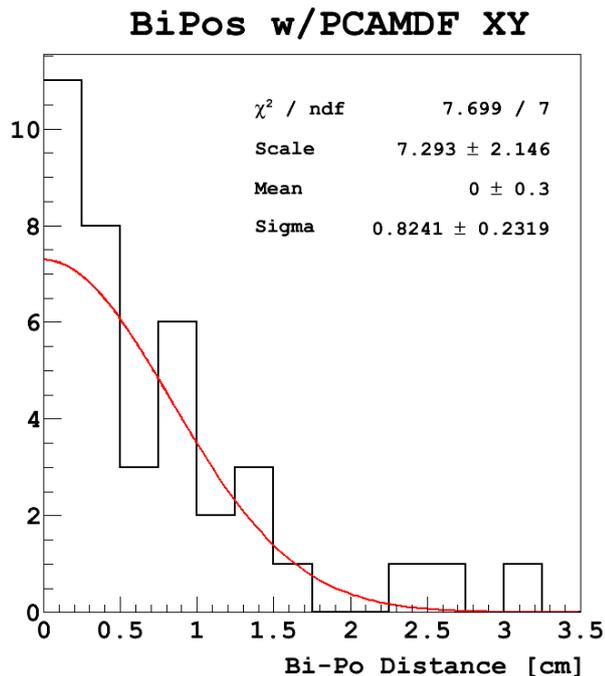


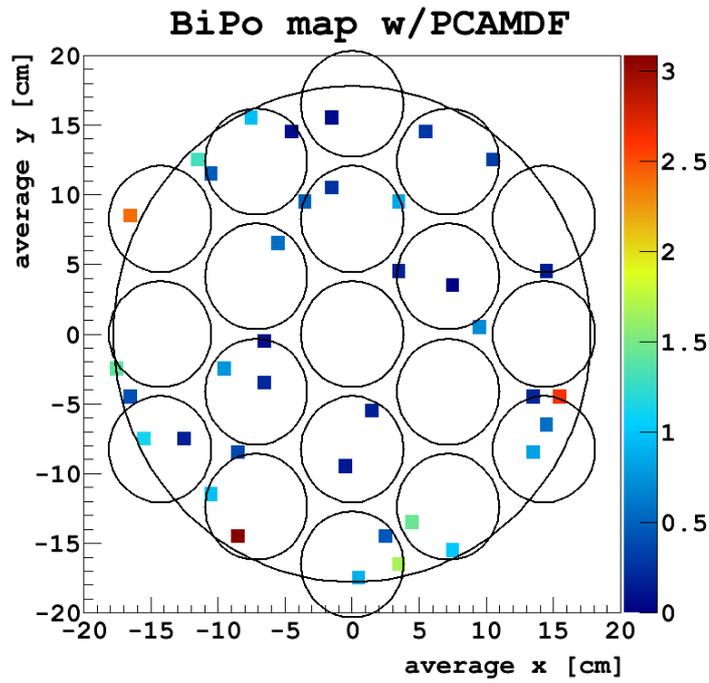
Figure 5.32: The distance between the locations of the coincident  $^{214}\text{Bi}^{214}\text{Po}$  decays.

any particular direction, then we can calculate the mean deviation of these BiPo events as 6.6 mm, though this requires a rather liberal definition of the word “large”.

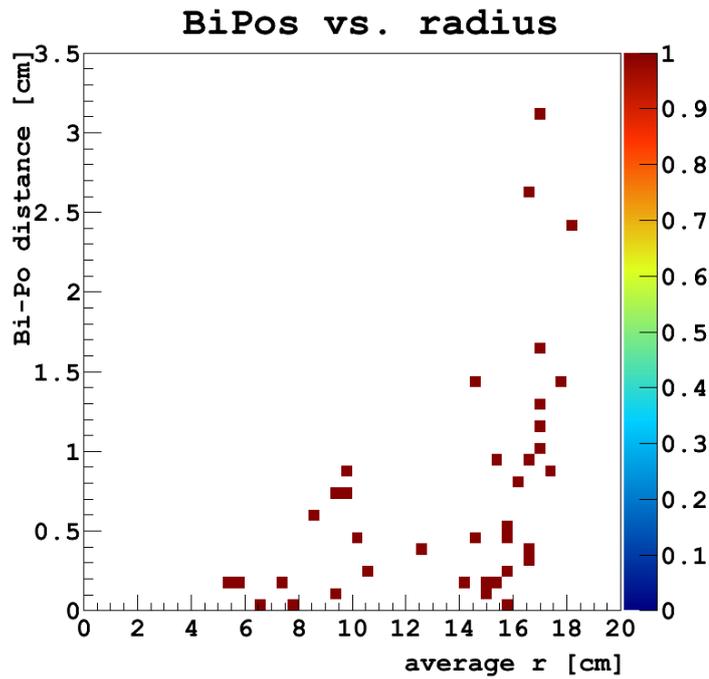
However, this precision appears to be position-dependent, as can be seen in Figure 5.33. For both plots in Figure 5.33, the  $xy$  position of the event is taken as the average of the Bi and the Po  $xy$  positions, and the  $z$ -value of subfigure (a) is the distance between the two positions. Note that the events are spread out fairly uniformly in  $xy$ , and that the average reconstructed distance between the Bi’s and the Po’s increases with radius from the detector axis.

These BiPo events give the PCAMDF an average reconstruction error of 6.6 mm overall, and show a dramatic deterioration in reconstruction accuracy for events near the edge of the detector. These results are in agreement with the precision prediction of Resolution Without  $xy$ -Reconstruction.

An additional coincident decay has been identified in DS-50 data: “RnPos”,



(a) Average reconstructed  $xy$  positions of 38 BiPo events. The  $z$ -scale shows the distance between the “Bi” and the “Po” decays in each event.



(b) The distance between each “Bi” and “Po” decay as a function of radius. Note how the position uncertainty increases with increasing radius.

Figure 5.33: The distance between the “Bi” and “Po” decays of 38 identified BiPo events over  $xy$  (a) and as a function of radius (b). The position uncertainty shows a marked increase as radius increases.

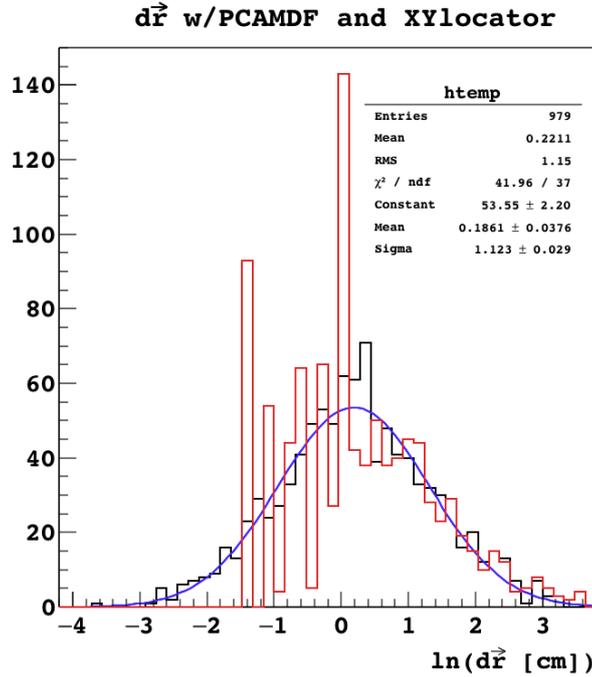


Figure 5.34: “RnPo” events in DS-50. Black shows the logarithm of the distance between the reconstructed positions of the Rn and the Po decays with PCAMDF, and red shows this same distance with XYlocator. The large red spikes are due to XYlocator quantizing reconstructed positions on a  $2.5 \text{ mm} \times 2.5 \text{ mm}$  grid.

where  $^{222}\text{Rn} \alpha$  decays into  $^{218}\text{Po}$  with a half-life of  $\sim 3.8$  days, and then that  $^{218}\text{Po} \alpha$  decays into  $^{214}\text{Pb}$ , with a half-life of  $\sim 3.1$  min. This decay chain can also be seen in Figure 5.30 (a).<sup>25</sup> Reconstructing these events with both XYlocator and PCAMDF gives the plot in Figure 5.34.

Figure 5.34 shows (in black) the distance between the reconstructed position of the “Rn” decay and the reconstructed position of the “Po” decay with the PCAMDF reconstruction. In red is this same quantity for XYlocator. The large red spikes are due to XYlocator quantizing reconstructed positions on a  $2.5 \text{ mm} \times 2.5 \text{ mm}$  grid, while PCAMDF lets reconstructed positions vary continuously. This quantization causes the XYlocator and PCAMDF histograms to

<sup>25</sup>Note that another “RnPo” decay ( $^{22}\text{Rn} \rightarrow \alpha + ^{216}\text{Po}$ ,  $\tau_{1/2} \approx 55$  sec, then  $^{216}\text{Po} \rightarrow \alpha + ^{212}\text{Pb}$ ,  $\tau_{1/2} \approx 140$  ms) used in Brodsky (2015) and seen in Figure 5.30 (b), is not considered in this work.

differ quite substantially to the left of the origin, though they agree fairly well to the right of the origin. A Gaussian fit to the PCAMDF histogram yields a mean of 0.1861 and a standard deviation of 1.123, as shown in Figure 5.34. This is a “log-normal” distribution, where the logarithm of some original variable (in this case,  $d\vec{r}$ ) is normally distributed due to the original variable being exponentially distributed. To convert the mean  $\mu$ , and the standard deviation  $\sigma$ , of this distribution to the  $\mu$  and  $\sigma$  of the non-logarithm data ( $d\vec{r}$ ), we use the relations:

$$\mu_{d\vec{r}} = \exp\{\mu_{\ln(d\vec{r})} + \sigma_{\ln(d\vec{r})}^2/2\} \quad (5.8)$$

and

$$\sigma_{d\vec{r}} = \exp\{2 \cdot \mu_{\ln(d\vec{r})} + \sigma_{\ln(d\vec{r})}^2\} \cdot (\exp\{\sigma_{\ln(d\vec{r})}^2\} - 1)$$

where  $\mu_{\ln(d\vec{r})}$  and  $\sigma_{\ln(d\vec{r})}$  are the mean and standard deviation of the logarithm data distribution (the normal distribution) and  $\mu_{d\vec{r}}$  and  $\sigma_{d\vec{r}}$  are the mean and standard deviation of the non-logarithm data. Using these equations, we can calculate the mean and standard deviation of  $d\vec{r}$  for the PCAMDF-reconstructed data as 2.26 cm and 3.60 cm, respectively. Since the right-hand edge of the XYlocator histogram more or less matches the right-hand edge of the PCAMDF histogram, we can assume its mean and standard deviation are similar, suggesting that the reconstruction here is dominated by statistics, and not the systematics of the two reconstruction methods.

This mean value is surprisingly large, considering the apparent precision of the PCAMDF method on BiPos; Figure 5.35 shows why this might be the case. Chris Stanford, who constructed these event sets, notes that he “purposefully left the cuts a little loose for this search in in the interest of providing greater statistics. So there may be a couple accidental coincidences that made it past the cuts.” Those coincidences are apparent in Figure 5.35, where many RnPos are reconstructed

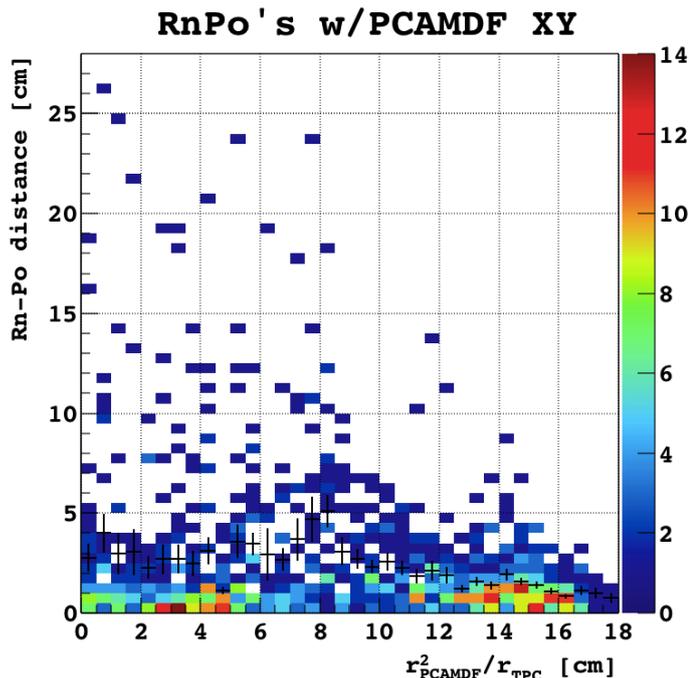


Figure 5.35: The distance between the “Rn” and “Po” decays of 979 identified RnPo decays as a function of radius. Outliers (likely accidentals) are apparent, especially at low radii.

> 10 cm from each other. Removing these coincidences by restricting the range of decay times between the Rn and the Po, and by requiring them to have similar values of  $t_{\text{drift}}$ , is the next step in this analysis, which unfortunately isn’t yet finished at the time of this writing.

### 5.6.3 $^{57}\text{Co}$ Events

Although we cannot put radioactive sources *inside* the active volume for calibration tests, CALIS allows us to put them right next to the cryostat which surrounds the DS-50 TPC. And although there is approximately 17.7 cm of space between the outside of the cryostat and the outside edge of the active volume when the detector is cold,<sup>26</sup> about 9.4 cm of this is the cryostat vacuum, which does not attenuate the source radioactivity at all. However, the copper field cage rings which

<sup>26</sup> From internal docdb entry # 1167.

surround the active volume strongly attenuate the 122 keV photons emitted by a  $^{57}\text{Co}$  source. These  $\gamma$ 's have interaction mean free paths of 3.33 cm in PTFE, 4.76 cm in  $\ell\text{Ar}$ , and 0.46 cm in stainless steel type 304L. Therefore, they pass through the materials of the cryostat and the TPC wall relatively easily. Through the OFHC copper field cage rings, these gamma rays have a mean free path of just 0.40 cm, comparable to the stainless steel.

The field cage rings are  $\sim 3.2$  mm thick, but are electrically insulated from each other by a  $\sim 3$  mm “gap” of  $\ell\text{Ar}$ . These gamma rays therefore have a different attenuation when they follow a path which travels between two field cage rings, as opposed to when they travel straight through the material of one of the rings themselves, as shown in Figure 5.36 for the yellow and red paths, respectively. This difference in material along these two paths translates into a *radiographic contrast*, which gives a different density of events in the detector volume relative to whether the photons travelled through the additional material of the field cage rings or not. Contrast is calculated according to the formula

$$\Delta\Phi = \frac{|\Phi_2 - \Phi_1|}{\Phi_1} \quad (5.9)$$

where  $\Phi_1$  and  $\Phi_2$  are defined according to

$$\Phi(x) = \Phi_0 \cdot e^{-\mu(E)x} \quad (5.10)$$

In the above equations,  $\Phi(x)$  is the radiation flux at a depth  $x$  in the detector medium (measured from the wall of the cryostat in, toward the TPC axis),  $\Phi_0$  is the initial flux (calculated from the activity of the source), and  $\mu$  is the linear attenuation coefficient of the material (which is a function of energy,  $E$ ).  $\Phi_1$  and  $\Phi_2$  are the attenuations along two different paths, 1 and 2. Finally, when traveling

through multiple materials, contrasts combine multiplicatively:

$$\Phi(x) = \Phi_0 \cdot \prod_{n=1}^N e^{\mu_n(E)x_n} \quad (5.11)$$

The simplest comparison is between two paths which travel perpendicular through the cryostat, the liquid argon between the inside wall of the cryostat and the TPC, and the materials which comprise the TPC wall, as shown in the cartoon in Figure 5.36. The red and yellow paths can be chosen to fall arbitrarily close to each other, while restricting the red path to pass through the material of a field cage ring, while the yellow path does not. Using the linear attenuation coefficients of these materials, given above, and the thicknesses of these materials at  $\ell$ Ar temperatures,<sup>27,28</sup> the contrast between the red and yellow paths is calculated to be about 0.51.

This means that approximately twice as many photons pass through the gaps in the field cage rings (traveling along the yellow path) as pass through the field cage rings themselves (traveling along the red path). In data, this should look like “sunbeams” coming through the wall of the detector. Comparing the shapes of these sunbeams (and the overall event distribution) should provide us with another validation of the PCAMDF  $xy$ -Reconstruction. Since the physical extent of the source is small ( $\mathcal{O}(\text{mm})$ ), these sunbeams should appear to all radiate from the same point, and should give relatively flat planes of events in the TPC for  $t_{\text{drift}}$  values near the  $z$ -position of the source, but inclined planes of events for  $t_{\text{drift}}$  positions far from the source, with the incline increasing with increasing

---

<sup>27</sup> Cooling the detector to the boiling point of argon causes thermal contraction of all materials within the detector.

<sup>28</sup> At 87.35 K, the two cryostat walls are each 4.016 mm thick, the  $\ell$ Ar outside of the TPC is 49.8 mm thick from the inside of the inner cryostat wall to the outside of the PTFE, the field cage rings are 3.16 mm thick, and the PTFE is 24.68 mm thick.

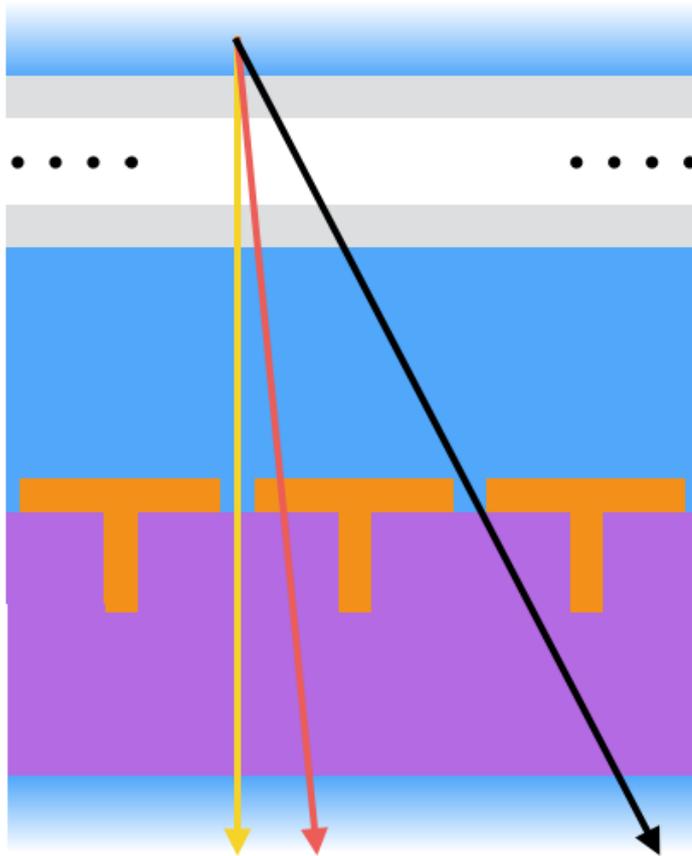


Figure 5.36: Cartoon showing a few possible paths for a particle to enter the TPC. The grey areas show the stainless steel of the cryostat wall (with the vacuum reduced in thickness for clarity), the blue areas show regions of  $\ell$ Ar, the purple region is the PTFE of the TPC wall, and the orange T-shaped areas are the copper field cage rings. The radiographic contrast for a  $^{57}\text{Co}$  source between the yellow and red paths is large, due to the short mean free path of 122 keV through OFHC copper.

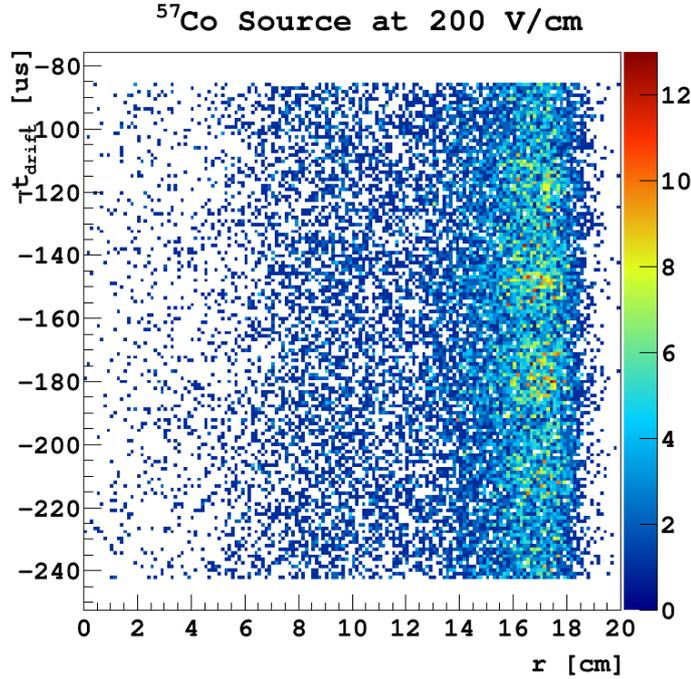


Figure 5.37:  $^{57}\text{Co}$  events plotted over  $-t_{\text{drift}}$  versus radius from a series of consecutive CALIS runs. Cuts are described in the text. The three or four concentrations of events at high radius (near the TPC wall) are the “sunbeams” of events, described in the text, due to high radiographic contrast of the materials comprising the TPC wall.

difference in  $t_{\text{drift}}$ . As the difference in  $t_{\text{drift}}$  increases between the point of interest in the detector and the  $z$ -position of the source, the radiographic contrast will also decrease, since the “red” and “yellow” paths will both have to travel through much more  $\ell\text{Ar}$  to enter the TPC at such a steep angle. As such, the regions of interest are those near the  $z$ -position of the source.

Figure 5.37 shows events from a series of consecutive CALIS calibration runs, with a  $^{57}\text{Co}$  source at the same position for each run. The events shown pass CX cuts 1–4, 9, and 12; have exactly two pulses; and pass a few other quality cuts ( $f_{90} \geq 0.05$  and  $S2_{\text{tot}} \geq 30$  PE). Additionally, since we want to select only the 122 keV  $\gamma$ 's which were emitted directly from the source and had their first interaction in the  $\ell\text{Ar}$  of the active volume, we choose events which have an  $S1_{\text{tot}}$  within  $\pm 50$  PE of the peak energy. This gives events in which the 122 keV  $\gamma$  has

its first and only interaction in the  $\ell$ Ar, depositing all of its energy, and avoids partial energy deposition events like Compton scatters. A final cut on  $t_{\text{drift}}$  focuses the plot on only the events of interest. The “200 V/cm” in the title of the plot indicates that the TPC was running at the standard drift field voltage during this campaign.

In Figure 5.37, we can just make out the sunbeams that we expect to see in this set of data. Ideally, we would like to make “slices” of  $t_{\text{drift}}$  and compare the patterns seen in reconstructed data to theory, but unfortunately we do not have enough events to perform this analysis. I have estimated that it would take approximately five to ten times as much data (about a week) at a single source position to gather enough events to perform a detailed study. One thing we can do, though, is calculate the positions of the different sources placed outside of the cryostat and compare these to their true positions. We know the true positions of the sources to within about a degree of accuracy ( $139^\circ \pm 1.2^\circ$ ) from tests performed on the CALIS source positioning arm.<sup>29</sup> The standard configuration was for sources to be placed at the level of the center of the detector (in  $z$ ) and the same angle,  $\theta$ , for each run.

To calculate the azimuthal angle of the source, we make the same cuts as described for Figure 5.37, and we add a radial cut, to focus only on events within a few cm of the wall. The result for one of the sources of interest can be seen in Figure 5.38. This study was performed with a  $^{57}\text{Co}$  source at 200 V/cm and 100 V/cm and with a  $^{133}\text{Ba}$  source at 150 V/cm and 100 V/cm. The average source angle calculated by the PCAMDF  $xy$  was  $141.3^\circ \pm 0.7^\circ$ . At the outside wall of the cryostat,  $1^\circ$  is equivalent to  $\sim 6.2$  mm of arc, so a  $2.3^\circ$  difference between the known source position and the reconstructed position corresponds to a  $\sim 1.4$  cm error in the reconstruction of the angular position of the source.<sup>30</sup>

---

<sup>29</sup> [Agnes et al. \(2016a\)](#)

<sup>30</sup> Coincidentally, this is also a  $\sim 1.4\%$  error in the source position, if we consider

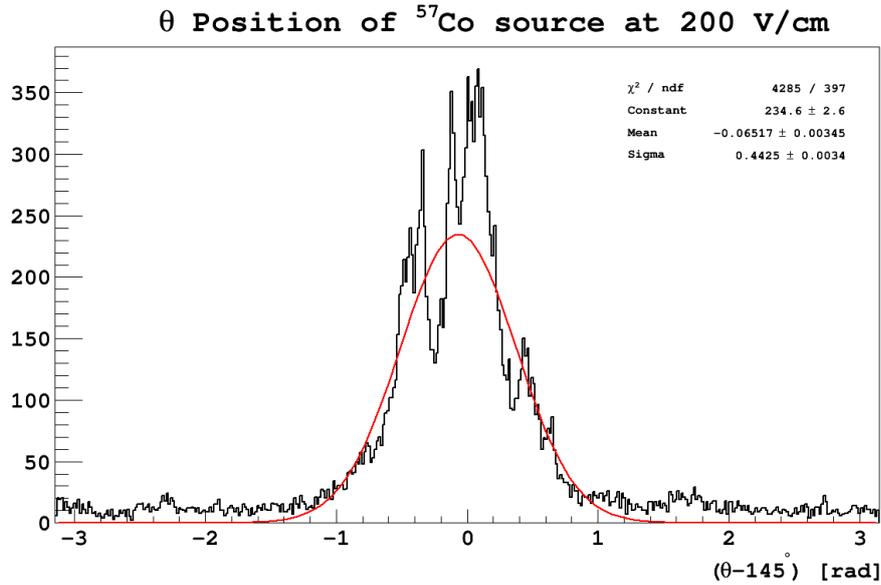
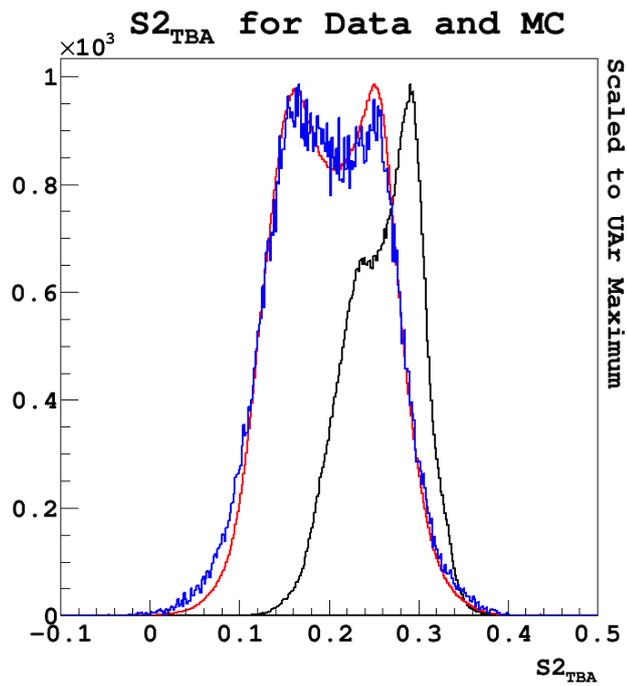


Figure 5.38: Locating the azimuthal position of a  $^{57}\text{Co}$  source used during a CALIS calibration campaign.

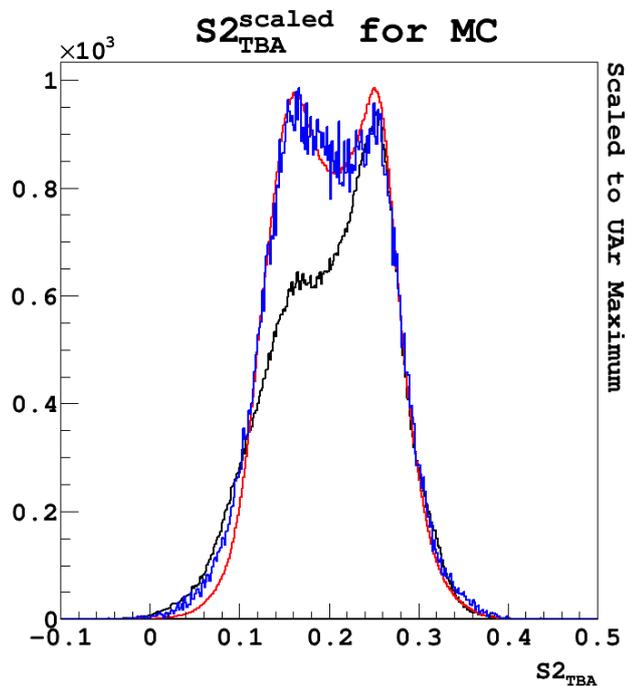
#### 5.6.4 $S2_{\text{TBA}}$ Revisited

The last validation study covered in this thesis uses the S2 Top-Bottom Asymmetry ( $S2_{\text{TBA}}$ ), first covered here in Subsection 5.5.4. Because the PCAMDF is given no bottom PMT information, the  $S2_{\text{TBA}}$  map is an “emergent phenomenon” of position-reconstructed data events. The distributions of events over restricted ranges of  $S2_{\text{TBA}}$  can give the reconstruction accuracy for data, something previously unavailable. To perform this study, we first need to fix a discrepancy between the MC and the data: namely, the shape of the  $S2_{\text{TBA}}$  distribution in g4ds.

Figure 5.39 (a) shows the overlaid  $S2_{\text{TBA}}$  spectra for AAr (red), UAr (blue), and MC (black), where the areas of the MC and AAr histograms have been scaled to fit on the same plot with the UAr. The MC distribution is obviously mean-shifted and narrowed, compared to the data distributions, which seem to agree fairly well with each other. To make the comparison between data and MC easier, we “resize” the MC distribution to match the data. This is done by defining a new variable 180° around the cryostat from the true position as “100% error”.



(a)  $S2_{TBA}$  spectra for MC (black), AAr (red), and UAr (blue). Notice how the MC spectrum is narrower and has a higher mean than the data spectra.



(b)  $S2_{TBA}$  spectra for MC (black), AAr (red), and UAr (blue). The MC  $S2_{TBA}$  has been scaled to match the UAr.

Figure 5.39: Manually fixing the discrepancy between the MC and data  $S2_{TBA}$  spectra.

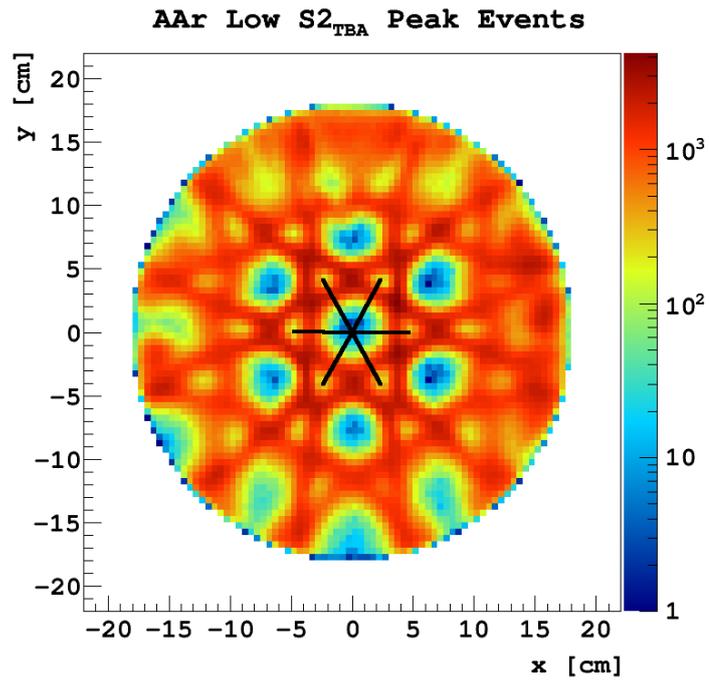
for the MC data,  $S2_{\text{TBA}}^{\text{scaled}} = c_1 \cdot S2_{\text{TBA}} + c_2$ , which is simply scaled and mean-shifted along the horizontal axis of Figure 5.39 (a). The result of this process can be seen in Figure 5.39 (b). The best-fit parameters are  $c_1 = 1.60021$  and  $c_2 = -0.212974$ , and were found by minimizing the sum of the squared differences between the MC and UAr histograms from the mean of the high peak out to  $S2_{\text{TBA}} = 0.5$ . The peak means and sigmas are found by fitting the AAr histogram with two independent Gaussians and are found to be  $m_1 = 0.1632 \pm 0.0002$ ,  $\sigma_1 = 0.04124 \pm 0.00070$ ,  $m_2 = 0.2509 \pm 0.0001$ , and  $\sigma_2 = 0.03161 \pm 0.00092$ . (Of course, a scaling factor is also applied to the area of the curves, since each of these event sets have different numbers of events. The best-fit “area scaling factor” is  $c_3 = 0.129$  for MC.)

Next, we investigate two particular classes of events: “low peak” events, and “average” events. Low peak events have  $S2_{\text{TBA}}$  values within  $\pm 0.5 \cdot \sigma_1$  of the mean of the low peak,  $m_1$ . Average events have  $S2_{\text{TBA}}$  values within  $\pm 0.3 \cdot \text{RMS}$  of the mean value of the AAr histogram, where RMS is the RMS of the AAr histogram. Unfortunately, the UAr data set has too few events in these regions to study, so the remainder of this analysis is performed with the AAr event set and the MC event set only.

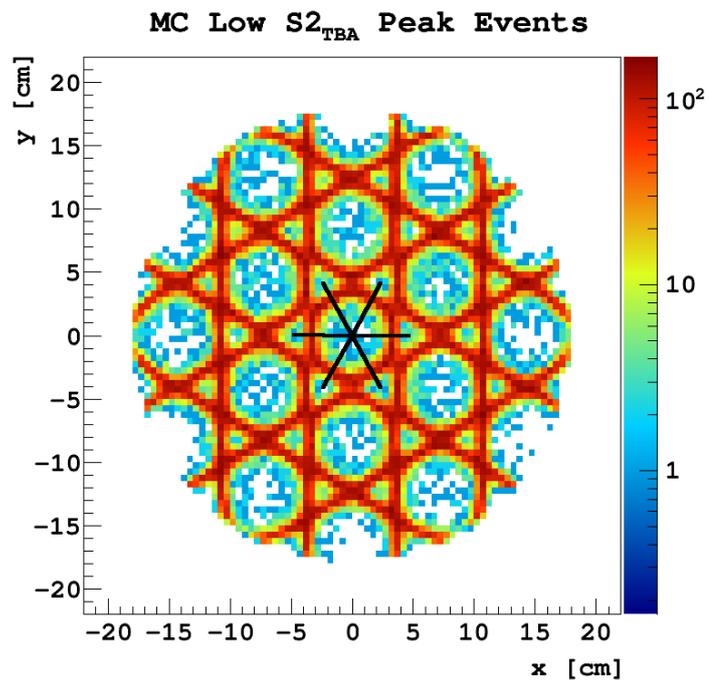
Low peak events and average events exhibit narrow structures over  $xy$  (see Figure 5.40 and Figure 5.41). Our ability to resolve these structures quantifies the precision of our reconstruction method,<sup>31</sup> and the (dis)agreement of the means of these structures between MC and data quantifies the accuracy of our reconstruction method. The black lines in Figure 5.40 and Figure 5.41 show the regions of interest. Each of these 12 regions (6 from the low peak events and 6 from the average events) contains a single “structure” (a single concentration of events)

---

<sup>31</sup> Similar to Solovov et al. (2012), though the wires of our grid are  $100\times$  thinner than in ZEPLIN-III (theirs have 5 mm thickness), so we cannot use our grid to judge the accuracy of our reconstruction, because our resolution is not even of that order of magnitude.

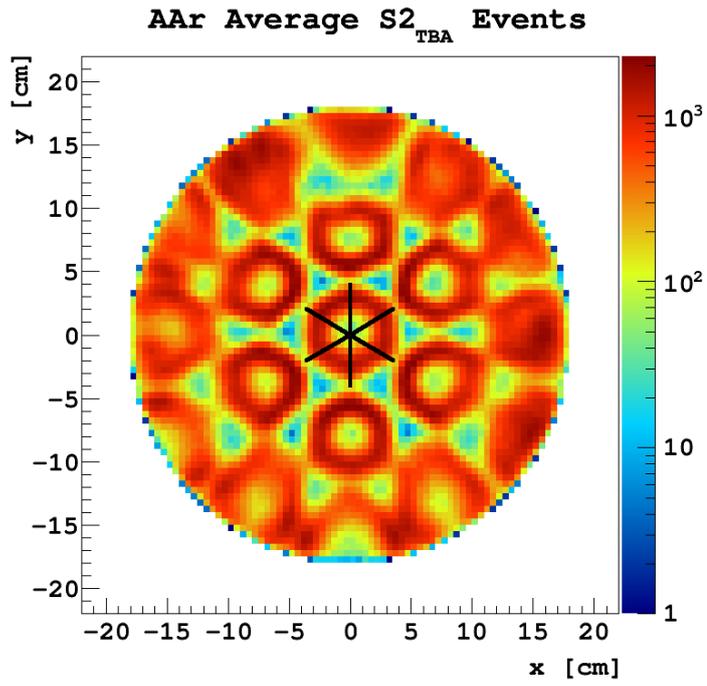


(a) AAr events in the low- $S2_{TBA}$  peak seen in Figure 5.39.

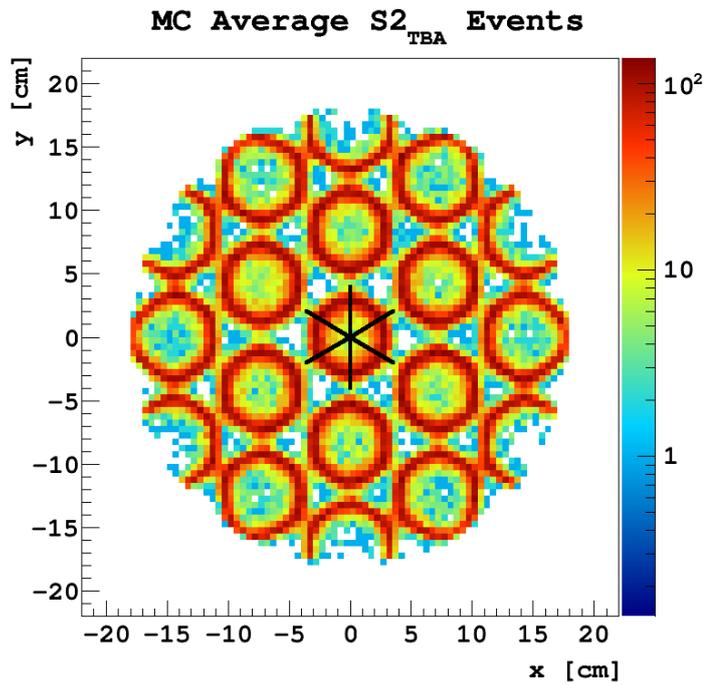


(b) MC events in the low- $S2_{TBA}$  peak seen in Figure 5.39.

Figure 5.40: Events in the low- $S2_{TBA}$  peak seen in Figure 5.39. The black lines indicate the regions over which the MC and AAr event distributions are compared.



(a) AAr events around the mean of the  $S2_{TBA}$  spectrum seen in Figure 5.39.



(b) MC events around the mean of the AAr  $S2_{TBA}$  spectrum seen in Figure 5.39.

Figure 5.41: Events around the mean of the  $S2_{TBA}$  spectrum seen in Figure 5.39. The black lines indicate the regions over which the MC and AAr event distributions are compared.

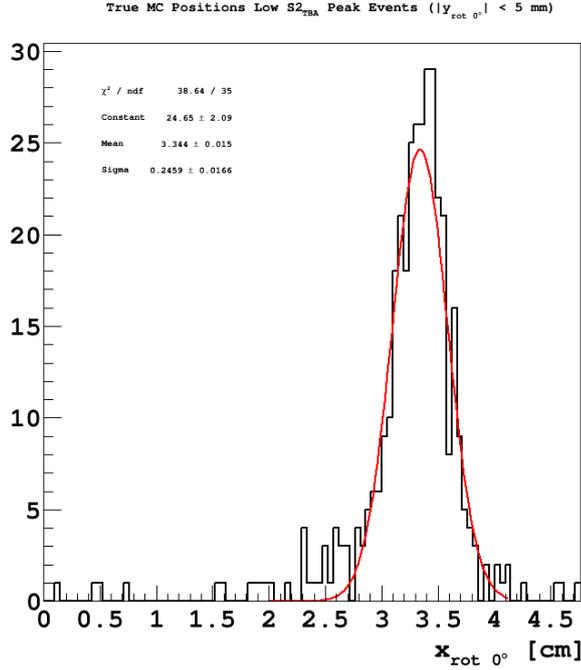


Figure 5.42: True-position MC events in the region which lies along the  $+x$ -axis of Figure 5.40 (b).

which we can compare between MC and data. To compare each of these regions, we must rotate the  $xy$  positions of all events in multiples of  $30^\circ$ , so we define a new coordinate system for each region  $(x_{\text{rot } \theta^\circ}, y_{\text{rot } \theta^\circ})$ , where  $\theta$  is the angle through which we rotate the coordinate axes (so  $(x_{\text{rot } 90^\circ}, y_{\text{rot } 90^\circ})$  has the positive  $x$ -axis pointing “up” and the positive  $y$ -axis pointing “left”). For the low peak regions,<sup>32</sup> we look at events with  $0 < x_{\text{rot } \theta^\circ} [\text{cm}] < 8.255/\sqrt{3}$  and for the average regions,<sup>33</sup> we look at events with  $0 < x_{\text{rot } \theta^\circ} [\text{cm}] < 8.255/2$ . For all regions, we require  $|y_{\text{rot } \theta^\circ} [\text{cm}]| < 1/2$ . Example histograms with Gaussian fits to the peaks are shown for MC with true positions (Figure 5.42), MC with reconstructed positions (Figure 5.43), and AAr with reconstructed positions (Figure 5.44).

Figure 5.42 shows the events in the  $\theta = 0^\circ$  region for MC events with true  $xy$  positions. The mean of the Gaussian fit is  $m_a = 3.344 \pm 0.015$  and the standard

<sup>32</sup>  $\theta = 0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ$

<sup>33</sup>  $\theta = 30^\circ, 90^\circ, 150^\circ, 210^\circ, 270^\circ, 330^\circ$

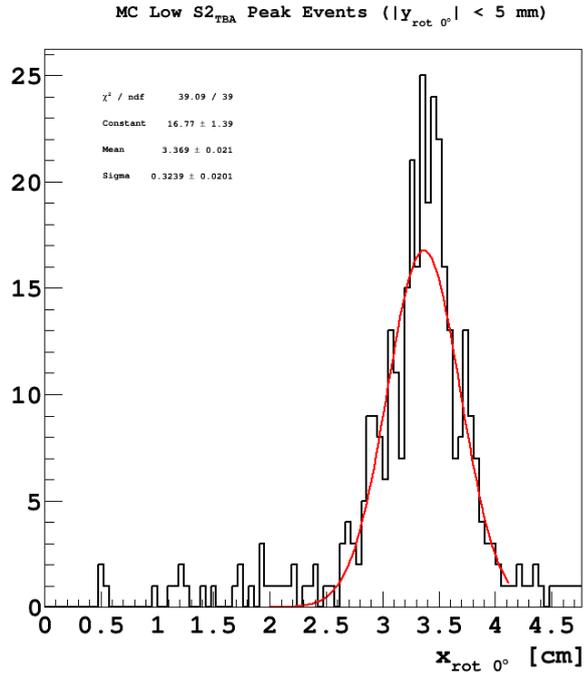


Figure 5.43: Reconstructed-position MC events in the region which lies along the  $+x$ -axis of Figure 5.40 (b).

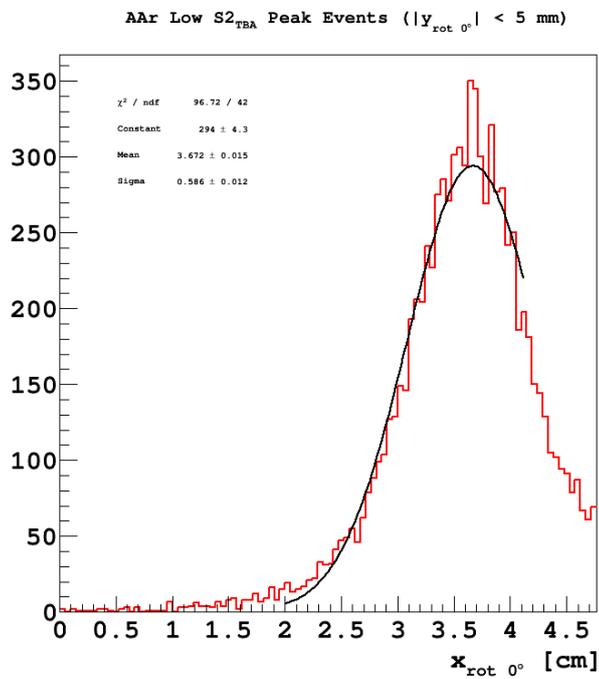


Figure 5.44: Reconstructed-position AAr events in the region which lies along the  $+x$ -axis of Figure 5.40 (a).

deviation is  $\sigma_a = 0.2459 \pm 0.0166$ . Figure 5.43 shows the events in the  $\theta = 0^\circ$  region for MC events with reconstructed  $xy$  positions. The mean of the Gaussian fit is  $m_b = 3.369 \pm 0.021$  and the standard deviation is  $\sigma_b = 0.3239 \pm 0.0201$ . Figure 5.44 shows the events in the  $\theta = 0^\circ$  region for AAr events with reconstructed  $xy$  positions. The mean of the Gaussian fit is  $m_c = 3.672 \pm 0.015$  and the standard deviation is  $\sigma_c = 0.586 \pm 0.012$ .

The absolute difference of the data mean and the true-position MC mean is  $|m_c - m_a| = 0.328$  cm. This tells us that our reconstruction accuracy in this region is about 3.3 mm. If we assume that the reconstructed-position MC distribution is simply the true-position MC distribution convolved with a Gaussian blur due to the PCAMDF method, then we can calculate the standard deviation of that blur, due to a particular property of convolved normal distributions: their standard deviations add in quadrature. This means that

$$\sigma_{\text{PCAMDF}} = \sqrt{\sigma_b^2 - \sigma_a^2} \quad (5.12)$$

where  $\sigma_{\text{PCAMDF}}$  is the standard deviation of the Gaussian blur due to the PCAMDF reconstruction. The difference between the width of the data distribution and the width of the true-position MC distribution should then be due to this PCAMDF-induced blurring, plus some additional smearing due to discrepancies between the MC and the data,  $\sigma_{\Delta(\text{data}, \text{MC})}$

$$\sigma_{\Delta(\text{data}, \text{MC})} = \sqrt{\sigma_c^2 - \sigma_b^2} \quad (5.13)$$

So the total Gaussian smearing is

$$\sigma_{\text{tot}} = \sqrt{\sigma_{\text{PCAMDF}}^2 + \sigma_{\Delta(\text{data}, \text{MC})}^2} \quad (5.14)$$

and the standard deviation of the data is the inherent width of the feature in the MC,  $\sigma_a$  (with some smearing due to PE statistics), plus this  $\sigma_{\text{tot}}$ :

$$\sigma_c = \sqrt{\sigma_a^2 + \sigma_{\text{tot}}^2} \quad (5.15)$$

Separating out the imprecision like this shows us that most of our reconstruction uncertainty is due to differences in the data versus the MC (see Table 5.2). The discrepancies between the data and the MC account for 71% of the total imprecision ( $\sigma_{\text{tot}}$ ) in the reconstruction on data ( $0.399^2/0.473^2$ ), while the reconstruction error from the PCAMDF method accounts for only the remaining 29%. The average means of the fits to the peaks of the distributions in true-position MC and data are quite close (3.116 cm and 3.173 cm, a difference of just 0.57 mm), but the average error in the position of the peaks ( $\langle |m_c - m_a| \rangle = 0.203$  cm) tells us that, on average, the mean value of a given peak in MC is about 2 mm away from the mean value of the same peak in data. This seems paradoxical, but the first number, 0.57 mm, is just telling us that there is little to no bias in the *direction* in which the data peak will be centered, relative to the MC. The second number, 2 mm, tells us that, on average, the mean of the given data peak is 2 mm away from the mean of the corresponding MC peak; the small bias just means that the data peak is equally as likely to be reconstructed 2 mm closer to the center of the central PMT as it is to be reconstructed 2 mm further away from the center of the central PMT.

Studying features of known dimension with imaging systems like the DS-50 PMT array is common practice for understanding the precision and biases inherent in the imaging system. Although there are no physical features large enough for our  $xy$ -Reconstruction methods to resolve, the high-reflectivity PTFE in which the anode PMT array is mounted generates a Top-Bottom Asymmetry in S2 for events

at differing transverse ( $xy$ ) positions within the DarkSide-50 detector. The limited spatial extent of these patterns over restricted ranges of  $S2_{\text{TBA}}$  gives us features wide enough to resolve with our  $xy$ -Reconstruction methods, yet narrow enough to provide us with information about the accuracy — as well as the precision — of our reconstruction methods. This analysis supports the conclusion reached in Section 5.3, using MC events, that there is no PMT-centered bias in the PCAMDF reconstruction,<sup>34</sup> it suggests that the reconstruction accuracy around the perimeter of the central PMT is approximately 2 mm, and it suggests that the majority of the reconstruction uncertainty is due to discrepancies between data and the MC, and not due to the reconstruction method itself.

---

<sup>34</sup> And no TPC-centered bias near  $r \approx 4$  cm, the radius of the central PMT.

$\theta$ [°]	0	30	60	90	120	150	180	210	240	270	300	330	Avg.
$m_a$	3.344	2.773	3.479	2.789	3.58	2.791	3.314	2.838	3.366	2.783	3.59	2.747	3.116
$\sigma_a$	0.245	0.378	0.363	0.305	0.352	0.353	0.266	0.344	0.287	0.345	0.392	0.326	0.330
$m_b$	3.369	2.716	3.463	2.732	3.544	2.774	3.293	2.856	3.305	2.777	3.603	2.741	3.097
$\sigma_b$	0.323	0.463	0.377	0.467	0.421	0.445	0.359	0.440	0.411	0.432	0.465	0.443	0.420
$m_c$	3.672	2.856	3.656	2.272	3.718	3.003	3.467	2.571	3.374	2.937	3.903	2.657	3.173
$\sigma_c$	0.586	0.494	0.642	0.590	0.701	0.534	0.519	0.579	0.674	0.535	0.645	0.558	0.588
$ m_c - m_a $	0.328	0.082	0.177	0.517	0.138	0.212	0.153	0.267	0.008	0.154	0.313	0.089	0.203
$\sigma_{\text{PCAMDF}}$	0.210	0.268	0.099	0.353	0.230	0.270	0.241	0.274	0.294	0.259	0.250	0.299	0.254
$\sigma_{\Delta(\text{data}, \text{MC})}$	0.488	0.172	0.520	0.361	0.560	0.295	0.374	0.377	0.534	0.317	0.447	0.339	0.399
$\sigma_{\text{tot}}$	0.531	0.319	0.530	0.505	0.606	0.400	0.445	0.466	0.609	0.409	0.512	0.453	0.473

Table 5.2: Measured and calculated variables for the S2<sub>TBA</sub> accuracy study.  $\theta$  gives the angle of rotation of the coordinate system;  $m_a$  and  $\sigma_a$  are the mean and sigma of the Gaussian fitted to the true-position MC peak at each angle;  $m_b$  and  $\sigma_b$  are the mean and sigma of the Gaussian fitted to the reconstructed-position MC peak at each angle; and  $m_c$  and  $\sigma_c$  are the mean and sigma of the Gaussian fitted to the reconstructed-position AAr peak at each angle.  $|m_c - m_a|$  gives the error in the peak position between MC and AAr;  $\sigma_{\text{PCAMDF}} = \sqrt{\sigma_b^2 - \sigma_a^2}$  is the standard deviation of the Gaussian blur imparted to the reconstructed data by using the PCAMDF reconstruction;  $\sigma_{\Delta(\text{data}, \text{MC})} = \sqrt{\sigma_c^2 - \sigma_b^2}$  is the standard deviation of the Gaussian blur due to discrepancies between the MC and the data; and  $\sigma_{\text{tot}}$  is the quadrature sum of  $\sigma_{\text{PCAMDF}}$  and  $\sigma_{\Delta(\text{data}, \text{MC})}$ , the total position uncertainty on top of the irreducible uncertainty from PE statistics. Further description in the text.

# Chapter 6

## CONCLUSION

Transverse position reconstruction in DarkSide-50 has been, and continues to be, a challenge. As we have no sample of events with known positions with which to judge the accuracy of an  $xy$ -Reconstruction, a PCAMDF-based algorithm relies on one main assumption: that the Monte Carlo events resemble the events in real data. If the MC events and the data events are not samples of the same population, then the reconstruction fails.

By several measures, the PCAMDF reconstruction excels: 97.61% of AAr events and 96.03% of UAr events are reconstructed within the radius of the TPC, without the method being given any constraints or information about the TPC geometry; the method generates an S2 Top-Bottom Asymmetry map, despite being given no information about the bottom PMTs; there are no distortions of the  $^{39}\text{Ar}$  spectrum under PMT centers; and the event density increases dramatically near the wall of the TPC, as expected.

MC tests suggest that there is no PMT-centered bias in the reconstruction, and studies of the  $S2_{\text{TBA}}$  map support this assertion. The MC also suggests that the PCAMDF is particularly adept at reconstructing the angular positions of events, and studies of  $^{57}\text{Co}$  in the detector confirm this, putting the angular accuracy at

$\pm \sim 2.3^\circ$  at the wall of the TPC, equivalent to  $\sim 7.1$  mm of error in the angular position near the wall. Studies of coincident “BiPo” decays in the detector suggest a resolution of about 6.6 mm on average, and analysis of the S2<sub>TBA</sub> map purports a reconstruction accuracy around 2 mm near the circumference of the center PMT. In total, these studies suggest a sub-cm average accuracy overall, which is best near the axis of the TPC and worst near the wall.

MC-based accuracy studies are slightly more mixed. They suggest a maximum average reconstruction error at the low end of the DM S1 range of  $\sim 3.16$  cm, but a minimum average error of about 2 mm at the high end of that range. These same studies suggest that the reconstruction accuracy is strongly radius-dependent. Events within the inner 12 cm or so of the detector have an average reconstruction error of  $\sim 3.2$  mm, and this accuracy worsens slightly (to  $\sim 6.3$  mm) out to about 17 cm, where it deteriorates dramatically. According to the MC, the average reconstruction error right next to the wall can be several centimeters. The MC also shows that the PCAMDF exhibits a TPC-centered radial bias: tending to “push out” events with true radii of about 10 cm to 17 cm and “pull in” events with true radii greater than about 17 cm. The average reconstruction error of MC events overall is about 6.8 mm.

Taken together, these results suggest that the PCAMDF *xy*-Reconstruction is reliable... to a point. It certainly has sub-cm accuracy out to about a 10 cm or 12 cm radius, but outside of that range, it gets more difficult to pin a number on its performance. To complicate matters further, TPB “dead spots” appear at high radii, concentrated at the “north” and “east” ends of the detector, making position reconstruction less reliable. The fiducial volume cuts given in the final Chapter of this thesis are based on MC events, not data events, and so should be taken with a grain of salt. If a larger fiducial volume cut can be taken, it should be taken. Studies of the NR acceptance loss due to these fiducial volume cuts are

ongoing.

The largest source of imprecision in the PCAMDF  $xy$ -Reconstruction is the disagreement between the MC and the data. Great care has been taken by the Paris-based DarkSide simulations team to bring the MC into agreement with the data, and for the most part, g4ds performs exceedingly well. PCAMDF is extremely sensitive to changes in the independent variables given to it, though, and more work is required to bring S2 generation and distribution throughout the detector into agreement with data. Several people are involved in this effort and it will surely make great strides in the coming months and years.

The transverse position reconstruction performed by the PCAMDF method is the primary tool for analyzing the spatial distributions of events in the DS-50 detector. In spite of its shortcomings at high radii, the PCAMDF has been, and certainly will continue to be, a very useful algorithm for DS analyzers. The PCAMDF  $xy$ , along with the other  $xy$ -reconstruction algorithms used by DarkSide over the past few years, have discovered many aspects of our detector which were previously unknown. The TPB defects in the S2<sub>TBA</sub> map, the “missing” events near PMT # 2, and several other classes of unusual and unexpected events would likely still be undetected, if not for these  $xy$ -Reconstructions. There is still more work to be done, but the performance of the PCAMDF  $xy$ -Reconstruction on MC and data events in DS-50 is now largely understood.

# Bibliography

Aalbers, J. et al. (2016, Jun). DARWIN: towards the ultimate dark matter detector. [arXiv: 1606.07001](#).

Aalseth, C. E., P. S. Barbeau, J. Colaresi, J. I. Collar, et al. (2011, Jun). Search for an annual modulation in a p-type point contact germanium dark matter detector. [arXiv: 1106.0650v2](#).

Aalseth, C. E., P. S. Barbeau, J. Colaresi, J. I. Collar, et al. (2014, Jan). Search for an annual modulation in three years of CoGeNT dark matter detector data. [arXiv: 1401.3295v1](#).

Aalseth, C. E., P. S. Barbeau, J. D. Leon, J. E. Fast, et al. (2014, Jan). Maximum likelihood signal extraction method applied to 3.4 years of CoGeNT data. [arXiv: 1401.6234](#).

Abell, G. O. (1958, May). The distribution of rich clusters of galaxies. *Astrophysical Journal Supplement Series 3*, 211–288. [DOI: 10.1086/190036](#).

Abell, G. O., J. Harold G. Corwin, and R. P. Olowin (1989, May). A catalog of rich clusters of galaxies. *Astrophysical Journal Supplement Series 70*, 1–138. [DOI: 10.1086/191333](#).

Adhikari, R., M. Agostini, N. A. Ky, T. Araki, et al. (2017, Jan). A white paper

- on keV sterile neutrino dark matter. *Journal of Cosmology and Astroparticle Physics* 2017. DOI: [10.1088/1475-7516/2017/01/025](https://doi.org/10.1088/1475-7516/2017/01/025).
- Agnes, P. (2016a). *Direct Search for Dark Matter with the DarkSide Experiment*. Ph. D. thesis, Université Paris Diderot. <http://hal.in2p3.fr/tel-01497505/document>.
- Agnes, P. (2016b, Feb). g4ds, the DarkSide simulation tool. In *UCLA Dark Matter, Upcoming detectors, calibration, technological issues II*. UCLA Dark Matter 2016: Sources and Detection of Dark Matter and Dark Energy in the Universe. <https://conferences.pa.ucla.edu/dm16>.
- Agnes, P. et al. (2015, Apr). First results from the DarkSide-50 dark matter experiment at Laboratori Nazionali del Gran Sasso. *Physics Letters B* 743, 456–466. DOI: [10.1016/j.physletb.2015.03.012](https://doi.org/10.1016/j.physletb.2015.03.012).
- Agnes, P. et al. (2016a, Nov). CALIS - a CALibration Insertion System for the DarkSide-50 dark matter search experiment. [arXiv: 1611.02750v2](https://arxiv.org/abs/1611.02750v2).
- Agnes, P. et al. (2016b, Dec). The electronics and data acquisition system for the DarkSide-50 veto detectors. *Journal of Instrumentation* 11(P12007). DOI: [10.1088/1748-0221/11/12/P12007](https://doi.org/10.1088/1748-0221/11/12/P12007).
- Agnes, P. et al. (2016c, Apr). Results from the first use of low radioactivity argon in a dark matter search. *Physical Review D* 93(081101). DOI: [10.1103/PhysRevD.93.081101](https://doi.org/10.1103/PhysRevD.93.081101).
- Agnes, P. et al. (2017, Jan). Effect of low electric fields on alpha scintillation light yield in liquid argon. *Journal of Instrumentation* 12(P01021). DOI: [10.1088/1748-0221/12/01/P01021](https://doi.org/10.1088/1748-0221/12/01/P01021).

- Agnese, R. et al. (2013, Dec). Silicon detector dark matter results from the final exposure of CDMS II. *Physical Review Letters* 111(251301). DOI: [10.1103/PhysRevLett.111.251301](https://doi.org/10.1103/PhysRevLett.111.251301).
- Agnese, R. et al. (2014, Feb). Search for low-mass WIMPs with SuperCDMS. arXiv: [1402.7137v2](https://arxiv.org/abs/1402.7137v2).
- Agnese, R. et al. (2016, Sep). Projected sensitivity of the SuperCDMS SNOLAB experiment. arXiv: [1610.00006v1](https://arxiv.org/abs/1610.00006v1).
- Akerib, D. S. et al. (2014, Mar). First results from the LUX dark matter experiment at the Sanford Underground Research Facility. *Physical Review Letters* 112(091303). DOI: [10.1103/PhysRevLett.112.091303](https://doi.org/10.1103/PhysRevLett.112.091303).
- Albert, J. et al. (2008, May). Upper limit for  $\gamma$ -ray emission above 140 GeV from the dwarf spheroidal galaxy Draco. *The Astrophysical Journal* 679(1), 428–431. DOI: [10.1086/529135](https://doi.org/10.1086/529135).
- Aleksić, J., L. A. Antonelli, P. Antoranz, M. Backes, et al. (2010, Jan). MAGIC gamma-ray telescope observation of the Perseus cluster of galaxies: implications for cosmic rays, dark matter, and NGC 1275. *The Astrophysical Journal* 710(1), 634–647. DOI: [10.1088/0004-637X/710/1/634](https://doi.org/10.1088/0004-637X/710/1/634).
- Alexander, T. et al. (2013a, Sep). Light yield in DarkSide-10: A prototype two-phase argon TPC for dark matter searches. *Astroparticle Physics* 49, 44–51. DOI: [10.1016/j.astropartphys.2013.08.004](https://doi.org/10.1016/j.astropartphys.2013.08.004).
- Alexander, T. et al. (2013b, Nov). Observation of the dependence on drift field of scintillation from nuclear recoils in liquid argon. *Physical Review D* 88(092006). DOI: [10.1103/PhysRevD.88.092006](https://doi.org/10.1103/PhysRevD.88.092006).

- ANTARES (2015, Oct). Astronomy with a Neutrino Telescope and Abyss environmental RESearch. <http://antares.in2p3.fr/>.
- Aprile, E. et al. (2010, Sep). First dark matter results from the xenon100 experiment. *Physical Review Letters* 105(131302). DOI: [10.1103/PhysRevLett.105.131302](https://doi.org/10.1103/PhysRevLett.105.131302).
- Aprile, E. et al. (2012, Nov). Dark matter results from 225 live days of XENON100 data. *Physical Review Letters* 109(181301). DOI: [10.1103/PhysRevLett.109.181301](https://doi.org/10.1103/PhysRevLett.109.181301).
- ArDM Collaboration (2016, Dec). The ArDM liquid argon time projection chamber at the Canfranc Underground Laboratory: a ton-scale detector for dark matter searches. [arXiv: 1612.06375](https://arxiv.org/abs/1612.06375).
- Arrenberg, S., L. Baudis, K. Kong, et al. (2013, Aug). Kaluza-Klein dark matter: Direct detection vis-a-vis LHC (2013 update). [arXiv: 1307.6581](https://arxiv.org/abs/1307.6581).
- Bassett, B. A. and R. Hlozek (2009, Oct). Baryon acoustic oscillations. [arXiv: 0910.5224](https://arxiv.org/abs/0910.5224).
- Beck, C. (2015, Mar). Axion mass estimates from resonant Josephson junctions. *Physics of the Dark Universe* 7–8, 6–11. DOI: [10.1016/j.dark.2015.03.002](https://doi.org/10.1016/j.dark.2015.03.002).
- Begeman, K. G., A. H. Broeils, and R. H. Sanders (1991, Apr). Extended rotation curves of spiral galaxies: dark haloes and modified dynamics. In *Monthly Notices of the Royal Astronomical Society*, Volume 249, pp. 523–537. DOI: [10.1093/mnras/249.3.523](https://doi.org/10.1093/mnras/249.3.523).
- Belanger, G. (2012, Aug). Dark matter and physics beyond the standard model. In *XVIII Vietnam School of Physics (VSOP-18)*, Quy Nhon, Binh Dinh, Vietnam. <http://www.iop.vast.ac.vn/theor/conferences/vsop/18>.

- Bernabei, R., P. Belli, F. Cappella, V. Caracciolo, et al. (2013, Dec). Final model independent result of DAMA/LIBRA–phase1. *The European Physical Journal C* 73(2648). DOI: [10.1140/epjc/s10052-013-2648-7](https://doi.org/10.1140/epjc/s10052-013-2648-7).
- Bernabei, R., P. Belli, F. Cappella, R. Cerulli, et al. (2003, Jul). Dark matter search. arXiv: [astro-ph/0307403](https://arxiv.org/abs/astro-ph/0307403).
- Bernabei, R., P. Belli, F. Cappella, R. Cerulli, et al. (2010, May). New results from DAMA/LIBRA. *The European Physical Journal C* 67(1), 39–49. DOI: [10.1140/epjc/s10052-010-1303-9](https://doi.org/10.1140/epjc/s10052-010-1303-9).
- Bernabei, R., P. Belli, F. Montecchia, W. D. Nicolantonio, et al. (1998, Apr). Searching for WIMPs by the annual modulation signature. *Physics Letters B* 424(1–2), 195–201. DOI: [10.1016/S0370-2693\(98\)00172-5](https://doi.org/10.1016/S0370-2693(98)00172-5).
- Bertone, G. and D. Hooper (2016). A history of dark matter. arXiv: [1605.04909](https://arxiv.org/abs/1605.04909).
- Bird, S., I. Cholis, J. B. Muñoz, Y. Ali-Haïmoud, et al. (2016, May). Did LIGO detect dark matter? *Physical Review Letters* 116(201301). DOI: [10.1103/PhysRevLett.116.201301](https://doi.org/10.1103/PhysRevLett.116.201301).
- Bosma, A. (1978). *The Distribution and Kinematics of Neutral Hydrogen in Spiral Galaxies of Various Morphological Types*. Ph. D. thesis, Groningen University. <http://bit.do/bosma78>.
- Bouvard, A. (1821). *Tables astronomiques publiées par le Bureau des longitudes de France, contenant les tables de Jupiter, de Saturne et d’Uranus, construites d’après la théorie de la mécanique céleste*, Volume 49. Paris: Bachelier et Huzard. <http://gallica.bnf.fr/ark:/12148/bpt6k9668766k>.
- Boyarsky, A., O. Ruchayskiy, D. Iakubovskyi, and J. Franse (2014, Dec). Uniden-

- tified line in X-ray spectra of the Andromeda Galaxy and Perseus galaxy cluster. *Physical Review Letters* 113(251301). DOI: [10.1103/PhysRevLett.113.251301](https://doi.org/10.1103/PhysRevLett.113.251301).
- Boyarsky, A., O. Ruchayskiy, and M. Shaposhnikov (2009, Nov). The role of sterile neutrinos in cosmology and astrophysics. *Annual Review of Nuclear and Particle Science* 59, 191–214. DOI: [10.1146/annurev.nucl.010909.083654](https://doi.org/10.1146/annurev.nucl.010909.083654).
- Brodsky, J. P. (2015, Sep). *xy-Position Reconstruction in DarkSide-50*. Ph. D. thesis, Princeton University. <http://arks.princeton.edu/ark:/88435/dsp01c534fr32w>.
- Bulbul, E., M. Markevitch, A. Foster, R. K. Smith, et al. (2014, Jul). Detection of an unidentified emission line in the stacked X-ray spectrum of galaxy clusters. *The Astrophysical Journal* 789(1), 13–35. DOI: [10.1088/0004-637X/789/1/13](https://doi.org/10.1088/0004-637X/789/1/13).
- Burles, S., K. M. Nollett, and M. S. Turner (1999, Mar). Big-bang nucleosynthesis: Linking inner space and outer space. arXiv: [astro-ph/9903300](https://arxiv.org/abs/astro-ph/9903300).
- Canetti, L., M. Drewes, and M. Shaposhnikov (2012, Sep). Matter and antimatter in the universe. *New Journal of Physics* 14. DOI: [10.1088/1367-2630/14/9/095012](https://doi.org/10.1088/1367-2630/14/9/095012).
- Cao, H. (2014, Nov). *A Study of Nuclear Recoils in Liquid Argon Time Projection Chamber for the Direct Detection of WIMP Dark Matter*. Ph. D. thesis, Princeton University. <http://arks.princeton.edu/ark:/88435/dsp0108612q761>.
- Cao, H. et al. (2015, May). Measurement of scintillation and ionization yield and scintillation pulse shape from nuclear recoils in liquid argon. *Physical Review D* 91(092007). DOI: [10.1103/PhysRevD.91.092007](https://doi.org/10.1103/PhysRevD.91.092007).
- Carr, B., F. Kühnel, and M. Sandstad (2016, Oct). Primordial black holes as dark matter. *Physical Review D* 94(083504). DOI: [10.1103/PhysRevD.94.083504](https://doi.org/10.1103/PhysRevD.94.083504).

- CAST Collaboration (2015, Jul). New solar axion search using the CERN Axion Solar Telescope with  $^4\text{He}$  filling. *Physical Review D* 92(021101). DOI: [10.1103/PhysRevD.92.021101](https://doi.org/10.1103/PhysRevD.92.021101).
- CDMS Collaboration (2010, Nov). Results from a low-energy analysis of the CDMS II germanium data. arXiv: [1011.2482v3](https://arxiv.org/abs/1011.2482v3).
- CDMS II Collaboration (2010, Mar). Dark matter search results from the CDMS II experiment. *Science* 327(5973), 1619–1621. DOI: [10.1126/science.1186112](https://doi.org/10.1126/science.1186112).
- Choi, K.-Y. and J. E. Kim (2013, Nov). Review of axino dark matter. *Journal of the Korean Physical Society* 63(9), 1685–1695. DOI: [10.3938/jkps.63.1685](https://doi.org/10.3938/jkps.63.1685).
- Cintio, A. D., C. B. Brook, A. A. Dutton, A. V. Macciò, et al. (2014, May). A mass-dependent density profile for dark matter haloes including the influence of galaxy formation. *Monthly Notices of the Royal Astronomical Society* 441(4), 2986–2995. DOI: [10.1093/mnras/stu729](https://doi.org/10.1093/mnras/stu729).
- Clowe, D., M. Bradač, A. H. Gonzalez, M. Markevitch, et al. (2006, Sep). A direct empirical proof of the existence of dark matter. *The Astrophysical Journal Letters* 648(2), L109–L113. DOI: [10.1086/508162](https://doi.org/10.1086/508162).
- CMS and ATLAS Collaborations (2015, Oct). LHC results for dark matter from ATLAS and CMS. arXiv: [1510.01516v1](https://arxiv.org/abs/1510.01516v1).
- CMS Collaboration (2012, Jan). Search for dark matter at the LHC using missing transverse energy. arXiv: [1206.0753](https://arxiv.org/abs/1206.0753).
- Coc, A., E. Vangioni-Flam, P. Descouvemont, et al. (2004, Jan). Updated big bang nucleosynthesis compared with Wilkinson Microwave Anisotropy Probe observations and the abundance of light elements. *The Astrophysical Journal* 600(2), 544–552. DOI: [10.1086/380121](https://doi.org/10.1086/380121).

- CoGeNT Collaboration (2010, Feb). Results from a search for light-mass dark matter with a p-type point contact germanium detector. [arXiv: 1002.4703](#).
- CoGeNT Collaboration (2011, Mar). Results from a search for light-mass dark matter with a p-type point contact germanium detector. *Physical Review Letters* 106. DOI: [10.1103/PhysRevLett.106.131301](#).
- Collar, J. I. (2011, Jun). A realistic assessment of the sensitivity of XENON10 and XENON100 to light-mass WIMPs. [arXiv: 1106.0653v2](#).
- Collins, II, G. W. (1978). *The virial theorem in stellar astrophysics*, Volume 7 of *Astronomy and Astrophysics Series*. Tucson, AZ: Pachart Publishing House. <http://adsabs.harvard.edu/abs/1978vtsa.book.....C>.
- Conrad, J. (2014, Nov). Indirect detection of WIMP dark matter: a compact review. [arXiv: 1411.1925v1](#).
- Croton, D. J. (2013, Oct). Damn you, little h! (or, real-world applications of the Hubble constant using observed and simulated data). *Publications of the Astronomical Society of Australia* 30. DOI: [10.1017/pasa.2013.31](#).
- Csáki, C. (1996, Mar). The minimal supersymmetric standard model. *Modern Physics Letters A* 11(8). DOI: [10.1142/S021773239600062X](#).
- DAE $\delta$ ALUS Experiment (2014, Nov). CP violation. <https://www.nevis.columbia.edu/daedalus/motiv/cp.html>.
- D'Angelo, D. (2015, Jan). DarkSide-50: results from first argon run. [arXiv: 1501.03541](#).
- DarkSide Collaboration (2015, Jan). The electronics and data acquisition system of the DarkSide dark matter search. [arXiv: 1412.2969](#).

- Davis, J. H., C. McCabe, and C. B. hm (2014, Aug). Quantifying the evidence for dark matter in CoGeNT data. *Journal of Cosmology and Astroparticle Physics*. DOI: [10.1088/1475-7516/2014/08/014](https://doi.org/10.1088/1475-7516/2014/08/014).
- Davis, R., D. S. Harmer, and K. C. Hoffman (1968, May). Search for neutrinos from the Sun. *Physical Review Letters* 20(21), 1205–1209. DOI: [10.1103/PhysRevLett.20.1205](https://doi.org/10.1103/PhysRevLett.20.1205).
- Daw, E. et al. (2011, Oct). The DRIFT dark matter experiments. [arXiv: 1110.0222](https://arxiv.org/abs/1110.0222).
- Dienes, K. R., J. Kumar, B. Thomas, and D. Yaylali (2015, Feb). Dark-matter decay as a complementary probe of multicomponent dark sectors. *Physical Review Letters* 114. DOI: [10.1103/PhysRevLett.114.051301](https://doi.org/10.1103/PhysRevLett.114.051301).
- Dodelson, S. (2003, Mar). *Modern Cosmology* (1 ed.), pp. 73–77. Academic Press. <http://bit.do/dodelson>.
- Editors of the Encyclopædia Britannica (1998, Jul). *Air*. Encyclopædia Britannica, inc. <http://www.britannica.com/science/air>.
- Edsjö, J. and P. Gondolo (1997, Aug). Neutralino relic density including coannihilations. *Physical Review D* 56(4), 1879–1894. DOI: [10.1103/PhysRevD.56.1879](https://doi.org/10.1103/PhysRevD.56.1879).
- Einstein, A. (1916). Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik* 354(7), 769–822. DOI: [10.1002/andp.19163540702](https://doi.org/10.1002/andp.19163540702).
- Einstein, A. (1936, Dec). Lens-like action of a star by the deviation of light in the gravitational field. *Science* 84, 506–507. DOI: [10.1126/science.84.2188.506](https://doi.org/10.1126/science.84.2188.506).
- Erken, O., P. Sikivie, H. Tam, and Q. Yang (2012, Feb). Axion dark matter and cosmological parameters. *Physical Review Letters* 108(061304). DOI: [10.1103/PhysRevLett.108.061304](https://doi.org/10.1103/PhysRevLett.108.061304).

- F. P. An, e. a. (2012, Apr). Observation of electron-antineutrino disappearance at Daya Bay. *Physical Review Letters* 108(171803). DOI: [10.1103/PhysRevLett.108.171803](https://doi.org/10.1103/PhysRevLett.108.171803).
- Fan, A. (2015, Oct). Status and results from DarkSide-50. [arXiv: 1511.00676v1](https://arxiv.org/abs/1511.00676v1).
- Fan, A. (2016). *Results from the DarkSide-50 Dark Matter Experiment*. Ph. D. thesis, University of California, Los Angeles. <http://search.proquest.com/docview/1823196595>.
- Feng, J. L. (2010, May). Dark matter candidates from particle physics and methods of detection. *Annual Review of Astronomy and Astrophysics* 48, 495–545. DOI: [10.1146/annurev-astro-082708-101659](https://doi.org/10.1146/annurev-astro-082708-101659).
- Feng, J. L., S. Su, and F. Takayama (2004, Sep). SuperWIMP gravitino dark matter from slepton and sneutrino decays. *Physical Review D* 70(6). DOI: [10.1103/PhysRevD.70.063514](https://doi.org/10.1103/PhysRevD.70.063514).
- Fraser, G. W. et al. (2014, Mar). Potential solar axion signatures in X-ray observations with the XMM-Newton observatory. [arXiv: 1403.2436](https://arxiv.org/abs/1403.2436).
- Freeman, K. C. (1970, Jun). On the disks of spiral and S0 galaxies. *Astrophysical Journal* 160, 811–830. DOI: [10.1086/150474](https://doi.org/10.1086/150474).
- Froborg, F. (2015, Sep). SABRE: WIMP modulation detection in the northern and southern hemisphere. In *Journal of Physics: Conference Series*, Volume 718 of *Dark Matter*. DOI: [10.1088/1742-6596/718/4/042021](https://doi.org/10.1088/1742-6596/718/4/042021).
- Garrett, K. and G. Duda (2011). Dark matter: A primer. *Advances in Astronomy* 2011. DOI: [10.1155/2011/968283](https://doi.org/10.1155/2011/968283).
- Giusarma, E., R. de Putter, S. Ho, and O. Mena (2013, Sep). Constraints on

- neutrino masses from Planck and galaxy clustering data. *Physical Review D* 88(063515). DOI: [10.1103/PhysRevD.88.063515](https://doi.org/10.1103/PhysRevD.88.063515).
- Grace, E. and J. A. Nikkel (2016, Feb). Index of refraction, Rayleigh scattering length, and Sellmeier coefficients in solid and liquid argon and xenon. [arXiv: 1502.04213](https://arxiv.org/abs/1502.04213).
- Grandi, L. (2005). *WARP: An argon double phase technique for dark matter search*. Ph. D. thesis, University of Pavia.
- Granger, D. (2017, Jan). CALET-USA: CALorimetric Electron Telescope. <http://calet.phys.lsu.edu/Science/DMS.php>.
- Hansen, J. R. (2015, Aug). *Spaceflight Revolution*. Andesite Press.
- Hinkley, D. V. (1969, Dec). On the ratio of two correlated normal random variables. *Biometrika* 56(3), 635–639. DOI: [10.1093/biomet/56.3.635](https://doi.org/10.1093/biomet/56.3.635).
- Hitachi, A., T. Takahashi, N. Funayama, K. Masuda, et al. (1983, May). Effect of ionization density on the time dependence of luminescence from liquid argon and xenon. *Physical Review B* 27(5279). DOI: [10.1103/PhysRevB.27.5279](https://doi.org/10.1103/PhysRevB.27.5279).
- Huang, K. (1987). *Statistical Mechanics*, pp. 52–71. John Wiley and Sons. <http://bit.do/huangbook>.
- Jungman, G., M. Kamionkowski, and K. Griest (1996, Mar). Supersymmetric dark matter. *Physics Reports* 267(5–6), 195–373. DOI: [10.1016/0370-1573\(95\)00058-5](https://doi.org/10.1016/0370-1573(95)00058-5).
- Kakizaki, M., S. Matsumoto, Y. Sato, and M. Senami (2005, Jun). Significant effects of second Kaluza-Klein particles on dark matter physics. *Physical Review D* 71(123522). DOI: [10.1103/PhysRevD.71.123522](https://doi.org/10.1103/PhysRevD.71.123522).

- Kappl, R. and M. W. Winkler (2011, Sep). New limits on dark matter from Super-Kamiokande. *Nuclear Physics B* 850(3), 505–521.  
DOI: [10.1016/j.nuclphysb.2011.05.006](https://doi.org/10.1016/j.nuclphysb.2011.05.006).
- Karlsruher Institut für Technologie (2016, Oct). Neutrinowaage KATRIN feiert “first light”. <http://www.kit.edu/kit/20624.php>.
- Kashlinsky, A. (2016, May). LIGO gravitational wave detection, primordial black holes, and the near-IR cosmic infrared background anisotropies. *The Astrophysical Journal Letters* 823. DOI: [10.3847/2041-8205/823/2/L25](https://doi.org/10.3847/2041-8205/823/2/L25).
- Katrin Collaboration (2015). Status of the KATRIN experiment and prospects to search for keV-mass sterile neutrinos in tritium  $\beta$ -decay. In *Physics Procedia*, Volume 61, pp. 267–273. 13th International Conference on Topics in Astroparticle and Underground Physics, TAUP 2013. DOI: [10.1016/j.phpro.2014.12.043](https://doi.org/10.1016/j.phpro.2014.12.043).
- Kelvin, W. T. (1904). *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*. London: C. J. Clay and Sons. <http://bit.do/wtkelvin>.
- Klasen, M., M. Pohl, and G. Sigl (2015, Nov). Indirect and direct search for dark matter. *Progress in Particle and Nuclear Physics* 85, 1–32.  
DOI: [10.1016/j.pnpnp.2015.07.001](https://doi.org/10.1016/j.pnpnp.2015.07.001).
- KM3NeT (2017, Jan). KM3NeT. <https://www.km3net.org/>.
- Kuster, M., G. Raffelt, B. Beltrán, and P. Sikivie (2008). *Axions*, Volume 741 of *Lecture Notes in Physics*, pp. 19–50. Springer Berlin Heidelberg.  
DOI: [10.1007/978-3-540-73518-2\\_2](https://doi.org/10.1007/978-3-540-73518-2_2).
- Le Verrier, U. J. J. (1859). *Comptes rendus hebdomadaires des séances de l’Académie des sciences*, Volume 49, pp. 379–383. L’Académie des sciences.  
<http://bit.do/leverrier>.

- Lebedenko, V. N. et al. (2009, Sep). Results from the first science run of the ZEPLIN-III dark matter search experiment. *Physical Review D* 80(052010). DOI: [10.1103/PhysRevD.80.052010](https://doi.org/10.1103/PhysRevD.80.052010).
- Lewin, J. D. and P. F. Smith (1996, Dec). Review of mathematics, numerical factors, and corrections for dark matter experiments based on elastic nuclear recoil. *Astroparticle Physics* 6(1), 87–112. DOI: [10.1016/S0927-6505\(96\)00047-3](https://doi.org/10.1016/S0927-6505(96)00047-3).
- Liang, Y.-F., Z.-Q. Shen, X. Li, Y.-Z. Fan, et al. (2016, May). Search for a gamma-ray line feature from a group of nearby galaxy clusters with Fermi LAT pass 8 data. *Physical Review D* 93. DOI: [10.1103/PhysRevD.93.103525](https://doi.org/10.1103/PhysRevD.93.103525).
- Lippincott, W. H. et al. (2008, Sep). Scintillation time dependence and pulse shape discrimination in liquid argon. *Physical Review C* 78(035801). DOI: [10.1103/PhysRevC.78.035801](https://doi.org/10.1103/PhysRevC.78.035801).
- Locher, F. (2011). Urbain Le Verrier. <http://bit.do/adfulv>.
- Loer, B. M. (2011, Nov). *Towards a Depleted Argon Time Projection Chamber WIMP Search: DarkSide Prototype Analysis and Predicted Sensitivity*. Ph. D. thesis, Princeton University. <http://adsabs.harvard.edu/abs/2011PhDT.....137L>.
- Loewenstein, M. and A. Kusenko (2010, Apr). Dark matter search using Chandra observations of Willman 1 and a special feature consistent with a decay line of a 5 keV sterile neutrino. *The Astrophysical Journal* 714, 652–662. DOI: [10.1088/0004-637X/714/1/652](https://doi.org/10.1088/0004-637X/714/1/652).
- Love, C. E. (2013, Jan). *Design and Analysis for the DarkSide-10 Two-Phase Argon Time Projection Chamber*. Ph. D. thesis, Temple University. <http://adsabs.harvard.edu/abs/2013PhDT.....18L>.

- MAGIC Collaboration (2010, Jan). MAGIC gamma-ray telescope observation of the Perseus Cluster of galaxies: Implications for cosmic rays, dark matter, and NGC 1275. *The Astrophysical Journal* 710(1), 634–647.  
DOI: [10.1088/0004-637X/710/1/634](https://doi.org/10.1088/0004-637X/710/1/634).
- MAGIC Collaboration (2017). The MAGIC telescopes.  
<https://magic.mpp.mpg.de/>.
- Matson, J. (2010, May). Early results from large dark matter detector cast doubt on earlier claims. *Scientific American*.  
<https://www.scientificamerican.com/article/xenon-100-dark-matter/>.
- Matson, J. (2011, May). WIMP wars: Astronomers and physicists remain skeptical of long-standing dark matter claim. *Scientific American*.  
<https://www.scientificamerican.com/article/dama-dark-matter/>.
- Mayet, F. et al. (2016, Apr). A review of the discovery reach of directional dark matter detection. *Physics Reports* 627, 1–49.  
DOI: [10.1016/j.physrep.2016.02.007](https://doi.org/10.1016/j.physrep.2016.02.007).
- McGaugh, S. S., W. J. G. de Blok, J. M. Schombert, et al. (2007, Apr). The rotation velocity attributable to dark matter at intermediate radii in disk galaxies. *The Astrophysical Journal* 659(1), 149–161. DOI: [10.1086/511807](https://doi.org/10.1086/511807).
- Meekins, J. F., G. Fritz, T. A. Chubb, et al. (1971, May). Physical sciences: X-rays from the Coma Cluster of galaxies. *Nature* 231, 107–108.  
DOI: [10.1038/231107a0](https://doi.org/10.1038/231107a0).
- Merritt, D., A. W. Graham, B. Moore, et al. (2006, Dec). Empirical models for dark matter halos. I. nonparametric construction of density profiles and comparison with parametric models. *The Astronomical Journal* 132(6), 2685–2700. DOI: [10.1086/508988](https://doi.org/10.1086/508988).

- Michelson, Albert A.; Morley, E. W. (1887, Nov). On the relative motion of the Earth and of the luminiferous ether. *Sidereal Messenger* 6, 306–310.  
<http://adsabs.harvard.edu/abs/1887SidM....6..306M>.
- NASA / CXC / CfA / M. Markevitch et al. (2009, Apr). 1E 0657-56: NASA finds direct proof of dark matter. <http://chandra.harvard.edu/photo/2006/1e0657/>.
- NASA / SAO / CXC / M. Markevitch et al. (2009, Feb). 1E 0657-56: A bow shock in a merging galaxy cluster. <http://chandra.harvard.edu/photo/2002/0066/>.
- Navarro, J. F., C. S. Frenk, and S. D. M. White (1997, Dec). A universal density profile from hierarchical clustering. *The Astrophysical Journal* 490(2), 493–508.  
DOI: [10.1086/304888](https://doi.org/10.1086/304888).
- Nemiroff, R. and J. Bonnell (2011, Dec). Astronomy picture of the day: A horse-shoe Einstein ring from Hubble. <http://apod.nasa.gov/apod/ap111221.html>.
- Noonan, T. W. (1969, Jun). Age of the Coma Cluster of galaxies from tidal-limited galaxian sizes. In *Bulletin of the American Astronomical Society*, Volume 1, pp. 255. <http://adsabs.harvard.edu/abs/1969BAAS....1..255N>.
- Page, D. N. (1976, Jan). Particle emission rates from a black hole: Massless particles from an uncharged, nonrotating hole. *Physical Review D* 13, 198–206.  
DOI: [10.1103/PhysRevD.13.198](https://doi.org/10.1103/PhysRevD.13.198).
- Pease, F. G. (1918, Jan). The rotation and radial velocity of the central part of the Andromeda nebula. In *Proceedings of the National Academy of Sciences of the United States of America*, Volume 4, pp. 21–24. DOI: [10.1073/pnas.4.1.21](https://doi.org/10.1073/pnas.4.1.21).
- Peccei, R. D. and H. R. Quinn (1977, Jun). CP conservation in the presence of pseudoparticles. *Physical Review Letters* 38, 1440–1443.  
DOI: [10.1103/PhysRevLett.38.1440](https://doi.org/10.1103/PhysRevLett.38.1440).

- Penzias, A. (2005, Jun). Biographical. 1978 Nobel Prize Biography.
- Penzias, A. A. (1961, Mar). Free hydrogen in the Pegasus I cluster of galaxies. *Astronomical Journal* 66, 293. DOI: [10.1086/108555](https://doi.org/10.1086/108555).
- Planck Collaboration (2016a, Oct). Planck 2015 results: I. overview of products and scientific results. *Astronomy and Astrophysics* 594. DOI: [10.1051/0004-6361/201527101](https://doi.org/10.1051/0004-6361/201527101).
- Planck Collaboration (2016b, Oct). Planck 2015 results: XIII. cosmological parameters. *Astronomy and Astrophysics* 594, A13.1–A13.63. DOI: [10.1051/0004-6361/201525830](https://doi.org/10.1051/0004-6361/201525830).
- Poggiani, R. (2017). *High Energy Astrophysical Techniques*, pp. 142–143. UNITEXT for Physics. Springer International Publishing. DOI: [10.1007/978-3-319-44729-2](https://doi.org/10.1007/978-3-319-44729-2).
- Poincaré, H. (1906). La Voie lactée et la théorie des gaz. In *Bulletin de la société astronomique de France*, Volume 20, pp. 153–165. <http://bit.do/poincare1906>.
- Qian, X. and P. Vogel (2015, Jul). Neutrino mass hierarchy. *Progress in Particle and Nuclear Physics* 83, 1–30. DOI: [10.1016/j.pnpnp.2015.05.002](https://doi.org/10.1016/j.pnpnp.2015.05.002).
- Richardson, M. (2009, May). Principal component analysis. <http://www.dsc.ufcg.edu.br/~hmg/disciplinas/posgraduacao/rn-copin-2014.3/material/SignalProcPCA.pdf>.
- Roberts, M. S. and A. H. Rots (1973). Comparison of rotation curves of different galaxy types. *Astronomy and Astrophysics* 26, 483–485. <http://adsabs.harvard.edu/abs/1973A&A....26..483R>.
- Rogstad, D. H. and G. S. Shostak (1972). Gross properties of five Scd galaxies as

- determined from 21-centimeter observations. *Astrophysical Journal* 176, 315–321. DOI: [10.1086/151636](https://doi.org/10.1086/151636).
- Rubin, V. C. (2000, Jun). One hundred years of rotating galaxies. In *Publications of the Astronomical Society of the Pacific*, Volume 112, pp. 747–750.
- Rubin, V. C., J. Burley, A. Kiasatpoor, B. Klock, et al. (1962, Oct). Kinematic studies of early-type stars. I. photometric survey, space motions, and comparison with radio observations. *The Astronomical Journal* 67, 491–531.  
DOI: [10.1086/108758](https://doi.org/10.1086/108758).
- Salucci, P. and A. Borriello (2000, Nov). The distribution of dark matter in galaxies: Constant-density dark halos envelop the stellar disks.  
arXiv: [astro-ph/0011079](https://arxiv.org/abs/astro-ph/0011079).
- Sauer, T. (2010). A brief history of gravitational lensing.  
[http://www.einstein-online.info/spotlights/grav\\_lensing\\_history](http://www.einstein-online.info/spotlights/grav_lensing_history).
- Savage, C., K. Freese, and P. Gondolo (2006, Aug). Annual modulation of dark matter in the presence of streams. *Physical Review D* 74(043531).  
DOI: [10.1103/PhysRevD.74.043531](https://doi.org/10.1103/PhysRevD.74.043531).
- Schelke, M. (2004). *Supersymmetric Dark Matter – aspects of sfermion coannihilations*. Ph. D. thesis, Stockholm University. <http://bit.do/schelke>.
- Scherrer, R. J. and M. S. Turner (1986, Mar). On the relic, cosmic abundance of stable, weakly interacting massive particles. *Physical Review D* 33(6), 1585–1589. DOI: [10.1103/PhysRevD.33.1585](https://doi.org/10.1103/PhysRevD.33.1585).
- Schumann, M. (2012, Jun). Dark matter search with liquid noble gases.  
arXiv: [1206.2169](https://arxiv.org/abs/1206.2169).

- Seidel, W. (2010, Jul). CRESST-II. In *IDM 2010*, Volume 8 of *International Workshop on Identification of Dark Matter*. <http://bit.do/CRESSTII>.
- Servanta, G. and T. M. P. Tait (2003, Feb). Is the lightest Kaluza–Klein particle a viable dark matter candidate? *Nuclear Physics B* 650(1–2), 391–419.  
DOI: [10.1016/S0550-3213\(02\)01012-X](https://doi.org/10.1016/S0550-3213(02)01012-X).
- Shlens, J. (2014, Apr). A tutorial on principal component analysis.  
arXiv: [1404.1100](https://arxiv.org/abs/1404.1100).
- Smith, L. I. (2002, Feb). A tutorial on principal components analysis.  
[http://www.cs.otago.ac.nz/cosc453/student\\_tutorials/principal\\_components.pdf](http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf).
- Solovov, V. N., V. A. Belov, D. Y. Akimov, H. M. Araújo, et al. (2012, Dec). Position reconstruction in a dual phase xenon scintillation detector. *IEEE Transactions on Nuclear Science* 59(6), 3286–3293. DOI: [10.1109/TNS.2012.2221742](https://doi.org/10.1109/TNS.2012.2221742).
- Steffen, F. D. (2006, Sep). Gravitino dark matter and cosmological constraints. *Journal of Cosmology and Astroparticle Physics* 2006.  
DOI: [10.1088/1475-7516/2006/09/001](https://doi.org/10.1088/1475-7516/2006/09/001).
- Steffen, F. D. (2009, Jan). Dark-matter candidates: axions, neutralinos, gravitinos, and axinos. *The European Physical Journal C* 59(2), 557–588.  
DOI: [10.1140/epjc/s10052-008-0830-0](https://doi.org/10.1140/epjc/s10052-008-0830-0).
- Steigman, G. (1979, Dec). Cosmology confronts particle physics. *Annual Review of Nuclear and Particle Science* 29, 313–338.  
DOI: [10.1146/annurev.ns.29.120179.001525](https://doi.org/10.1146/annurev.ns.29.120179.001525).
- Steigman, G. (2007, Nov). Primordial nucleosynthesis in the precision cosmology era. *Annual Review of Nuclear and Particle Science* 57, 463–491.  
DOI: [10.1146/annurev.nucl.56.080805.140437](https://doi.org/10.1146/annurev.nucl.56.080805.140437).

Stern, I. (2017, Jan). ADMX status. [arXiv: 1612.08296](#).

Super-Kamiokande Collaboration (1998, Aug). Evidence for oscillation of atmospheric neutrinos. *Physical Review Letters* 81(8), 1562–1567.

DOI: [10.1103/PhysRevLett.81.1562](#).

Super-Kamiokande Collaboration (2004, Jan). Precise measurement of the solar neutrino day-night and seasonal variation in Super-Kamiokande-I. *Physical Review D* 69. DOI: [10.1103/PhysRevD.69.011104](#).

Super-Kamiokande Collaboration (2011, Nov). An indirect search for weakly interacting massive particles in the Sun using 3109.6 days of upward-going muons in Super-Kamiokande. *The Astrophysical Journal* 742(2), 742–748.

DOI: [10.1088/0004-637X/742/2/78](#).

Super-Kamiokande Collaboration (2012, Aug). Search for GUT monopoles at Super-Kamiokande. *Astroparticle Physics* 36(1), 131–136.

DOI: [10.1016/j.astropartphys.2012.05.008](#).

Super-Kamiokande Collaboration (2015, Mar). Search for neutrinos from annihilation of captured low-mass dark matter particles in the Sun by Super-Kamiokande. [arXiv: 1503.04858v1](#).

Super-Kamiokande Collaboration (2016, Oct). Search for neutrinos in Super-Kamiokande associated with gravitational-wave events GW150914 and GW151226. *The Astrophysical Journal Letters* 830(1).

DOI: [10.3847/2041-8205/830/1/L11](#).

Suzuki, M. and S. Kubota (1979, Aug). Mechanism of proportional scintillation in argon, krypton and xenon. *Nuclear Instruments and Methods* 164(1), 197–199.

DOI: [10.1016/0029-554X\(79\)90453-1](#).

- Thonnard, N., V. C. Rubin, J. Ford, W. K., and M. S. Roberts (1978). Radial velocities of spiral galaxies determined from 21-cm neutral hydrogen observations. *Astronomical Journal* 83, 1564–1565. DOI: [10.1086/112364](https://doi.org/10.1086/112364).
- Tremaine, S. and J. E. Gunn (1979, Feb). Dynamical role of light neutral leptons in cosmology. *Physical Review Letters* 42, 407–410.  
DOI: [10.1103/PhysRevLett.42.407](https://doi.org/10.1103/PhysRevLett.42.407).
- Trovato, A., P. Koojiman, R. Coniglione, P. Sapienza, and The KM3Net Collaboration (2014). Sensitivity of the MK3NeT detector to neutrino fluxes from galactic point-like sources. In *6th International Workshop on Very Large Volume Neutrino Telescopes*, Volume 1630 of *AIP Conference Proceedings*, pp. 62–65. AIP Publishing LLC. DOI: [10.1063/1.4902772](https://doi.org/10.1063/1.4902772).
- Tsuchida, S. and M. Mori (2017, Jan). Observability of gamma-ray spectral feature from Kaluza-Klein dark matter. In *AIP Conference Proceedings*, Volume 1792 of *AIP Conference Proceedings*. American Institute of Physics: American Institute of Physics. DOI: [10.1063/1.4968994](https://doi.org/10.1063/1.4968994).
- Undagoitia, T. M. and L. Rauch (2015, Dec). Dark matter direct-detection experiments. *Journal of Physics G: Nuclear and Particle Physics* 43(1).  
DOI: [10.1088/0954-3899/43/1/013001](https://doi.org/10.1088/0954-3899/43/1/013001).
- Upton, G. and I. Cook (2014). *A Dictionary of Statistics* (3rd Ed. ed.). Oxford University Press. DOI: [10.1093/acref/9780199679188.001.0001](https://doi.org/10.1093/acref/9780199679188.001.0001).
- de Vaucouleurs, G. H. (1959). General physical properties of external galaxies. *Handbuch der Physik* 53, 311–372.  
<http://adsabs.harvard.edu/abs/1959HDP....53..311D>.
- Walsh, D., R. F. Carswell, and R. J. Weymann (1979, May). 0957 + 561 A, B:

- twin quasistellar objects or gravitational lens? *Nature* 279, 381–384.  
DOI: [10.1038/279381a0](https://doi.org/10.1038/279381a0).
- Weniger, C. (2012, Aug). A tentative gamma-ray line from dark matter annihilation at the Fermi Large Area Telescope. *Journal of Cosmology and Astroparticle Physics* 2012. DOI: [10.1088/1475-7516/2012/08/007](https://doi.org/10.1088/1475-7516/2012/08/007).
- Westerdale, S. S. (2016, Nov). *A Study of Nuclear Recoil Backgrounds in Dark Matter Detectors*. Ph. D. thesis, Princeton University.  
<http://arks.princeton.edu/ark:/88435/dsp01hx11xh73n>.
- White, M. (2015, Apr). Baryon acoustic oscillations and dark energy.  
<http://w.astro.berkeley.edu/~mwhite/bao/>.
- Wilson, R. W. (1978, Dec). The cosmic microwave background radiation. 1978 Nobel Lecture.
- Wong, W. Y. (2008, Aug). *Cosmological Recombination*. Ph. D. thesis, The University of British Columbia. [arXiv: 0811.2826](https://arxiv.org/abs/0811.2826).
- Wright, A. (2011, Sep). The DarkSide program at LNGS. [arXiv: 1109.2979v1](https://arxiv.org/abs/1109.2979v1).
- Wright, A., P. Mosteiro, B. Loer, and F. Calaprice (2011, Jul). A highly efficient neutron veto for dark matter experiments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 644(1), 18–26. DOI: [10.1016/j.nima.2011.04.009](https://doi.org/10.1016/j.nima.2011.04.009).
- Wyrzykowski, L., A. E. Rynkiewicz, J. Skowron, S. Kozłowski, et al. (2015, Jan). OGLE-III microlensing events and the structure of the galactic bulge. *The Astrophysical Journal Supplement Series* 216(1).  
DOI: [10.1088/0067-0049/216/1/12](https://doi.org/10.1088/0067-0049/216/1/12).

- XENON100 Collaboration (2010, May). First dark matter results from the XENON100 experiment. [arXiv: 1005.0380](#).
- XQC Collaboration (2015, Nov). Search for keV sterile neutrino dark matter with X-ray microcalorimeter sounding rockets. *The Astrophysical Journal* 814(1), 82–93. DOI: [10.1088/0004-637X/814/1/82](#).
- Zbořil, M. et al. (2013, Mar). Ultra-stable implanted  $^{83}\text{Rb}/^{83\text{m}}\text{Kr}$  electron sources for the energy scale monitoring in the KATRIN experiment. *Journal of Instrumentation* 8(P03009). DOI: [10.1088/1748-0221/8/03/P03009](#).
- Zeldovich, Y. B., J. Einasto, and S. F. Shandarin (1982, Dec). Giant voids in the universe. *Nature* 300, 407–413. DOI: [10.1038/300407a0](#).
- Zwicky, F. (1933). Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta* 6, 110–127. <http://adsabs.harvard.edu/abs/1933AcHPh...6..110Z>.
- Zwicky, F. (1937, Oct). On the masses of nebulae and of clusters of nebulae. *The Astrophysical Journal* 86, 217–246. DOI: [10.1086/143864](#).

# APPENDIX A: DERIVATION OF THE RELIC DENSITY EQUATION

The fact that the relic density of a  $\sim 100$  GeV particle with a weak-scale annihilation cross section matches the relic density of dark matter, combined with the fact that the Minimal Supersymmetric Standard Model predicts a stable “Lightest Supersymmetric Particle” with a mass of  $\mathcal{O}(100 \text{ GeV})$  *and* the fact that the astronomical evidence suggests a massive weakly-interacting particle to account for large-scale gravitational anomalies is known as the “WIMP Miracle”.

To check if the predicted relic density of WIMPdark matter and the known relic density of dark matter WIMPs are in agreement, we must first calculate the predicted relic density using the Boltzmann equation. (The following discussion is based largely on [Schelke \(2004\)](#), [Huang \(1987\)](#), and [Garrett and Duda \(2011\)](#), though some of the finer points of the derivation are outlined in [Steigman \(1979\)](#), [Feng \(2010\)](#), and [Scherrer and Turner \(1986\)](#).)

To calculate the theoretically predicted WIMP relic density, the starting point is:

$$dN = f(\vec{r}, \vec{p}, t) d^3r d^3p \quad (6.1)$$

Here,  $dN$  is the number of particles within a physical differential volume

$d^3r = dx dy dz$  about  $\vec{r} = (x, y, z)$ , and with momenta within a certain differential volume in momentum-space,  $d^3p = dp_x dp_y dp_z$  about  $\vec{p} = (p_x, p_y, p_z)$ .  $f(\vec{r}, \vec{p}, t)$  is a probability density function which depends on all seven variables of position, momentum, and time. In general, we're interested in the number density,  $n \equiv N/V$ , of WIMPs as a function of time, so we can rewrite the above equation as

$$\frac{dn}{dt} = \frac{d(dN/dV)}{dt} = \frac{d}{dt} \left( dN \frac{1}{d^3r} \right) = \frac{d}{dt} (f(\vec{r}, \vec{p}, t) d^3p) \quad \longrightarrow \quad \frac{dn}{dt} = \frac{df}{dt} d^3p \quad (6.2)$$

where  $dV \equiv d^3r$  is the differential volume, and  $f(\vec{r}, \vec{p}, t)$  has been shortened in the final term to  $f$ , for ease of use. The above equation says that the number density changes over time if and only if the function,  $f$ , changes with respect to time. In general,  $df/dt$  can have multiple components. Often, these are terms which deal with diffusion, collisions, or some external force:<sup>1</sup>

$$\frac{df}{dt} = \left( \frac{df}{dt} \right)_{\text{diff}} + \left( \frac{df}{dt} \right)_{\text{coll}} + \left( \frac{df}{dt} \right)_{\text{ext}} + \dots \quad (6.3)$$

For SUSY particles, there are four terms,<sup>2</sup> which yield a number density  $n_i$  for the  $i$ -th species according to the following equation:

$$\begin{aligned} \frac{dn_i}{dt} = & -3Hn_i - \sum_{j=1}^N \langle \sigma_{ij} v_{ij} \rangle (n_i n_j - n_i^{eq} n_j^{eq}) - \sum_{j \neq i} [\Gamma_{ij} (n_i - n_i^{eq}) - \Gamma_{ji} (n_j - n_j^{eq})] \\ & - \sum_{j \neq i} \left[ \langle \sigma'_{X_{ij}} v_{ij} \rangle (n_i n_X - n_i^{eq} n_X^{eq}) - \langle \sigma'_{X_{ji}} v_{ji} \rangle (n_j n_X - n_j^{eq} n_X^{eq}) \right] \quad (6.4) \end{aligned}$$

The first term on the right-hand side, which depends on the Hubble constant,  $H$ , decreases the number density due to the expansion of the universe. The second term is a coannihilation term, which decreases the SUSY number density when

---

<sup>1</sup> [Huang \(1987\)](#)

<sup>2</sup> [Garrett and Duda \(2011\)](#)

two SUSY particles collide and create SM particles, and increases it during the reverse process. The third term decreases the number density when the given SUSY particle decays. And the fourth term describes the scattering of SUSY particles off of the universal thermal background. Throughout this Appendix, angled brackets (“ $\langle$ ”, “ $\rangle$ ”) indicate a thermal average, and the subscripts  $i$  and  $j$  indicate SUSY particles. A more thorough explanation of this equation can be found in [Edsjö and Gondolo \(1997\)](#), but I will describe its constituent variables, for completeness. In the second term,

$$\sigma_{ij} = \sum_X \sigma(\chi_i \chi_j \rightarrow X) \quad (6.5)$$

gives the annihilation cross section of two given SUSY particles,  $\chi_i$  and  $\chi_j$ , which yield a set of SM particles  $X$ ;<sup>3</sup>

$$v_{ij} = \frac{\sqrt{(\mathbf{P}_i \cdot \mathbf{P}_j)^2 - m_i^2 m_j^2}}{E_i E_j} \quad (6.6)$$

is the “relative velocity”, where  $\mathbf{P}_i$  is the 4-momentum of the  $i$ -th species,  $E_i$  is its energy, and  $m_i$  is its mass; and

$$n_i^{eq} = \frac{g_i}{(2\pi)^3} \int f_i d^3p \quad (6.7)$$

is the number density of species  $i$  when it’s in thermal equilibrium. The term  $g_i$  is simply a statistical weight on the  $i$ -th species, also called the “degrees of freedom” of the species in some derivations, and  $f_i$  is the equilibrium distribution function of the  $i$ -th species, in this case, given according to the Maxwell-Boltzmann

---

<sup>3</sup> Note that in order for  $R$ -parity to be conserved, an even number of SM particles must be created in the annihilation.

approximation:

$$f_i = e^{E_i/T} \quad (6.8)$$

where  $T$  is the temperature. Finally,

$$\Gamma_{ij} = \sum_X \Gamma(\chi_i \rightarrow \chi_j X) \quad (6.9)$$

describes the decay rate of the  $i$ -th SUSY species into the  $j$ -th one.

Since all SUSY particles eventually decay into the LSP, we can assume that the combined original number density of all SUSY particles contributes to the final number density of the LSP. In that case, we can make the simplification that  $n_{LSP} = n \equiv n_1 + n_2 + \dots + n_N$ . This introduces a second summation index into Equation 6.4, such that

$$\frac{dn}{dt} = \sum_i^N \frac{dn_i}{dt} \quad (6.10)$$

With this summation, the third and fourth terms of Equation 6.4 sum to zero, and we end up with

$$\frac{dn}{dt} = -3Hn - \sum_{i,j=1}^N \langle \sigma_{ij} v_{ij} \rangle (n_i n_j - n_i^{eq} n_j^{eq}) \quad (6.11)$$

[Edsjö and Gondolo \(1997\)](#) note that, because the number density of SM particles,  $n_X$  is greater than the number density of any SUSY species,  $n_i$ , the scattering rate of SUSY particles is much greater than their annihilation rate. Because of this, the SUSY species remain in thermal equilibrium, and in particular, the fractional number density of a given SUSY species,  $i$ , is asymptotically equal to the equilibrium fractional number density of that same species

$$\frac{n_i}{n} \simeq \frac{n_i^{eq}}{n^{eq}} \quad (6.12)$$

Using this, we can define an “effective” annihilation cross section,  $\sigma_{eff}$ , which satisfies:

$$\langle \sigma_{eff} v \rangle = \sum_{i=1}^n \sum_{j=1}^n \langle \sigma_{ij} v_{ij} \rangle \frac{n_i^{eq} n_j^{eq}}{n^{eq} n^{eq}} \quad (6.13)$$

and rewrite Equation 6.11 as

$$\frac{dn}{dt} = -3Hn - \langle \sigma_{eff} v \rangle (n^2 - n_{eq}^2) \quad (6.14)$$

where we maintain the Hubble expansion term and the coannihilation term, which now takes a slightly different form.

Next, since we’re mainly interested in the “co-moving” number density of WIMPs (that is, the number density, ignoring the expansion of space itself), we introduce the entropy density,  $s$ :

$$s = \frac{S}{a^3} \quad (6.15)$$

where  $S$  is the entropy, and  $a$  is the scale factor in the definition of the Hubble constant

$$H \equiv \frac{\dot{a}}{a} = \frac{da}{dt} \cdot \frac{1}{a} \quad (6.16)$$

The entropy density in the radiation-dominated early universe is given by

$$s(T) = \frac{2\pi^2}{45} T^3 \cdot h_{eff}(T) \quad (6.17)$$

where

$$h_{eff}(T) = \sum_{i=bosons} g_i \left( \frac{T_i}{T} \right)^3 + \frac{7}{8} \sum_{i=fermions} g_i \left( \frac{T_i}{T} \right)^3 \quad (6.18)$$

All of this is necessary because the common method of presenting the dark matter relic density is not as a numerical density as a function of time, but rather as

an entropy-normalized numerical density as a function of temperature. Or rather, mass over temperature. To find this, we need to introduce two new variables

$$Y \equiv \frac{n}{s} \quad \text{and} \quad x \equiv \frac{m_\chi}{T} \quad (6.19)$$

where  $m_\chi$  is the mass of a particular SUSY particle.

We begin by finding the time rate of change of  $Y$

$$\frac{dY}{dt} = \frac{dn}{dt} \cdot \frac{1}{s} + n \cdot \frac{d}{dt} \left( \frac{1}{s} \right) = \frac{dn}{dt} \cdot \frac{1}{s} - n \cdot \left( \frac{1}{s^2} \right) \frac{ds}{dt} \quad (6.20)$$

and substituting  $dn/dt$  from Equation 6.14

$$\frac{dY}{dt} = -3H \frac{n}{s} - \frac{1}{s} \langle \sigma_{eff} v \rangle (n^2 - n_{eq}^2) - n \cdot \left( \frac{1}{s^2} \right) \frac{ds}{dt} \quad (6.21)$$

Next, note that

$$\frac{ds}{dt} = \frac{d}{dt} \left( \frac{S}{a^3} \right) = \frac{dS}{dt} \cdot \frac{1}{a^3} + S \cdot \frac{d}{dt} \left( \frac{1}{a^3} \right) = \frac{dS}{dt} \cdot \frac{1}{a^3} - 3S \cdot \left( \frac{1}{a^4} \right) \frac{da}{dt} \quad (6.22)$$

When a system is in thermal equilibrium, there is no change in the entropy over time, so the term  $dS/dt \rightarrow 0$ , and the above equation becomes

$$\frac{ds}{dt} = -3S \cdot \left( \frac{1}{a^4} \right) \frac{da}{dt} = -3 \frac{S}{a^3} \frac{\dot{a}}{a} = -3Hs \quad (6.23)$$

and 6.21 becomes

$$\frac{dY}{dt} = -3H \frac{n}{s} - \frac{1}{s} \langle \sigma_{eff} v \rangle (n^2 - n_{eq}^2) - n \cdot \left( \frac{1}{s^2} \right) (-3Hs) \quad (6.24)$$

or, since the first and third terms on the right cancel, just

$$\frac{dY}{dt} = -\frac{1}{s} \langle \sigma_{eff} v \rangle (n^2 - n_{eq}^2) \quad (6.25)$$

And finally, we can put those last “ $n$ ”s in terms of  $Y$ :

$$\frac{dY}{dt} = -s \langle \sigma_{eff} v \rangle (Y^2 - Y_{eq}^2) \quad (6.26)$$

The last step — and it’s a big one — is moving from  $dY/dt$  to  $dY/dx$ . To do this, we need to find

$$\frac{dx}{dt} = \frac{d}{dt} \left( \frac{m_\chi}{T} \right) = -\frac{m_\chi}{T^2} \frac{dT}{dt} = -\frac{m_\chi}{T^2} \frac{dT}{ds} \frac{ds}{dt} = \frac{3m_\chi H s}{T^2} \frac{dT}{ds} \quad (6.27)$$

Luckily,  $x$  is a function only of  $t$  (through  $T$ ), and  $s$  is a function only of  $T$ , so we can easily invert this equation

$$\frac{dt}{dx} = \frac{T^2}{3m_\chi H s} \frac{ds}{dT} \quad (6.28)$$

where we find  $ds/dT$  from Equation 6.17

$$\frac{ds}{dT} = \frac{d}{dT} \left( \frac{2\pi^2}{45} T^3 \cdot h_{eff}(T) \right) = \frac{6\pi^2}{45} T^2 \cdot h_{eff}(T) + \frac{2\pi^2}{45} T^3 \cdot \frac{dh_{eff}}{dT} \quad (6.29)$$

or

$$\frac{ds}{dT} = \frac{6\pi^2}{45} T^2 \cdot h_{eff}(T) \left( 1 + \frac{T}{3h_{eff}(T)} \cdot \frac{dh_{eff}}{dT} \right) \quad (6.30)$$

Next, we get the Hubble constant from the Friedmann equations

$$H = \sqrt{\frac{8\pi G\rho}{3}} \quad \text{where} \quad \rho = \frac{\pi^2}{30} T^4 g_{eff}(T) \quad (6.31)$$

and, similar to Equation 6.18,

$$g_{eff}(T) = \sum_{i=bosons} g_i \left( \frac{T_i}{T} \right)^4 + \frac{7}{8} \sum_{i=fermions} g_i \left( \frac{T_i}{T} \right)^4 \quad (6.32)$$

Bringing it all together — substituting Equation 6.30 and  $H(T)$  from Equation 6.31 into Equation 6.28, we get

$$\frac{dt}{dx} = \frac{T^2}{3m_\chi s} \sqrt{\frac{3 \cdot 30}{8\pi G \pi^2 T^4 g_{eff}(T)}} \cdot \frac{6\pi^2}{45} T^2 \cdot h_{eff}(T) \left( 1 + \frac{T}{3h_{eff}(T)} \cdot \frac{dh_{eff}}{dT} \right) \quad (6.33)$$

and, after some algebraic simplification,

$$\frac{dt}{dx} = \frac{T^2}{m_\chi s} \sqrt{\frac{\pi}{45G g_{eff}(T)}} \cdot h_{eff}(T) \left( 1 + \frac{T}{3h_{eff}(T)} \cdot \frac{dh_{eff}}{dT} \right) \quad (6.34)$$

We clean up the notation a bit by introducing a new variable,

$$g_\star^{1/2}(T) \equiv \frac{h_{eff}(T)}{\sqrt{g_{eff}(T)}} \left( 1 + \frac{T}{3h_{eff}(T)} \frac{dh_{eff}}{dT} \right) \quad (6.35)$$

so that

$$\frac{dt}{dx} = \frac{T^2}{m_\chi s} \sqrt{\frac{\pi}{45G}} \cdot g_\star^{1/2}(T) = \frac{m_\chi}{x^2 s} \sqrt{\frac{\pi}{45G}} \cdot g_\star^{1/2}(T) \quad (6.36)$$

Finally, we can use Equation 6.36 to write

$$\frac{dY}{dx} = \frac{dY}{dt} \frac{dt}{dx} = -s \langle \sigma_{eff} v \rangle (Y^2 - Y_{eq}^2) \frac{m}{x^2 s} \sqrt{\frac{\pi}{45G}} \cdot g_\star^{1/2}(T) \quad (6.37)$$

or, cleaning it up a bit,

$$\frac{dY}{dx} = -\sqrt{\frac{\pi}{45G}} \cdot \frac{g_\star^{1/2}(T) m_\chi}{x^2} \cdot \langle \sigma_{eff} v \rangle \cdot (Y^2 - Y_{eq}^2) \quad (6.38)$$

and defining the dimensionless variable,  $\lambda$ , where

$$\lambda = \sqrt{\frac{\pi}{45G}} \cdot g_{\star}^{1/2}(T) m_{\chi} \cdot \langle \sigma_{eff} v \rangle \quad (6.39)$$

we arrive at the final form of the relic density equation:

$$\frac{dY}{dx} = -\lambda \cdot \frac{Y^2 - Y_{eq}^2}{x^2} \quad (6.40)$$

which is utilized in Subsection [2.6.1](#).

# APPENDIX B: AUTHOR LIST

[ Current as of 5 February 2017. ]

P. Agnes,<sup>1</sup> I. F. M. Albuquerque,<sup>2</sup> T. Alexander,<sup>3</sup> A. K. Alton,<sup>4</sup> D. M. Asner,<sup>3</sup> H. O. Back,<sup>3</sup> B. Baldin,<sup>5</sup> K. Biery,<sup>5</sup> V. Bocci,<sup>6</sup> G. Bonfini,<sup>7</sup> W. Bonivento,<sup>8</sup> M. Bossa,<sup>9,7</sup> B. Bottino,<sup>10,11</sup> A. Brigatti,<sup>12</sup> J. Brodsky,<sup>13</sup> F. Budano,<sup>14,15</sup> S. Bussino,<sup>14,15</sup> M. Cadeddu,<sup>16,8</sup> M. Cadoni,<sup>16,8</sup> F. Calaprice,<sup>13</sup> N. Canci,<sup>17,7</sup> A. Candela,<sup>7</sup> M. Caravati,<sup>16,8</sup> M. Cariello,<sup>11</sup> M. Carlini,<sup>7</sup> S. Catalanotti,<sup>18,19</sup> P. Cavalcante,<sup>20,7</sup> A. Chepurinov,<sup>21</sup> C. Cicalò,<sup>8</sup> A. G. Cocco,<sup>19</sup> G. Covone,<sup>18,19</sup> D. D'Angelo,<sup>22,12</sup> M. D'Incecco,<sup>7</sup> S. Davini,<sup>9,7,11</sup> S. De Cecco,<sup>23</sup> M. De Deo,<sup>7</sup> M. De Vincenzi,<sup>14,15</sup> A. Derbin,<sup>24</sup> A. Devoto,<sup>16,8</sup> F. Di Eusanio,<sup>13</sup> G. Di Pietro,<sup>7,12</sup> C. Dionisi,<sup>6,25</sup> E. Edkins,<sup>26</sup> A. Empl,<sup>17</sup> A. Fan,<sup>27</sup> G. Fiorillo,<sup>18,19</sup> K. Fomenko,<sup>28</sup> G. Forster,<sup>29,5</sup> D. Franco,<sup>1</sup> F. Gabriele,<sup>7</sup> C. Galbiati,<sup>13,12</sup> S. Giagu,<sup>6,25</sup> C. Giganti,<sup>23</sup> G. K. Giovanetti,<sup>13</sup> A. M. Goretti,<sup>7</sup> F. Granato,<sup>30</sup> M. Gromov,<sup>21</sup> M. Guan,<sup>31</sup> Y. Guardincerri,<sup>5,a</sup> B. R. Hackett,<sup>26</sup> K. Herner,<sup>5</sup> D. Hughes,<sup>13</sup> P. Humble,<sup>3</sup> E. V. Hungerford,<sup>17</sup> A. Ianni,<sup>13,7</sup> I. James,<sup>14,15</sup> T. N. Johnson,<sup>32</sup> C. Jollet,<sup>33</sup> K. Keeter,<sup>34</sup> C. L. Kendziora,<sup>5</sup> G. Koh,<sup>13</sup> D. Korablev,<sup>28</sup> G. Korga,<sup>17,7</sup> A. Kubankin,<sup>35</sup> X. Li,<sup>13</sup> M. Lissia,<sup>8</sup> B. Loer,<sup>3</sup> P. Lombardi,<sup>12</sup> G. Longo,<sup>18,19</sup> Y. Ma,<sup>31</sup> I. N. Machulin,<sup>36,37</sup> A. Mandarano,<sup>9,7</sup> S. M. Mari,<sup>14,15</sup> J. Maricic,<sup>26</sup> L. Marini,<sup>10,11</sup> C. J. Martoff,<sup>30</sup> A. Meregaglia,<sup>33</sup> P. D. Meyers,<sup>13</sup> R. Milincic,<sup>26</sup> J. D. Miller,<sup>17</sup> D. Montanari,<sup>5</sup> A. Monte,<sup>29</sup> B. J. Mount,<sup>34</sup> V. N. Muratova,<sup>24</sup> P. Musico,<sup>11</sup> J. Napolitano,<sup>30</sup> A. Navrer Agasson,<sup>23</sup> S. Odrowski,<sup>7</sup> A. Oleinik,<sup>35</sup> M. Orsini,<sup>7</sup> F. Ortica,<sup>38,39</sup> L. Pagani,<sup>10,11</sup> M. Pallavicini,<sup>10,11</sup> E. Pantic,<sup>32</sup> S. Parmeggiano,<sup>12</sup> K. Pelczar,<sup>40</sup> N. Pelliccia,<sup>38,39</sup> A. Pocar,<sup>29</sup> S. Pordes,<sup>5</sup> D. A. Pugachev,<sup>36</sup> H. Qian,<sup>13</sup> K. Randle,<sup>13</sup> G. Ranucci,<sup>12</sup> M. Razeti,<sup>8</sup> A. Razeto,<sup>7,13</sup> B. Reinhold,<sup>26</sup> A. L. Renshaw,<sup>17</sup> M. Rescigno,<sup>6</sup> Q. Riffard,<sup>1</sup> A. Romani,<sup>38,39</sup> B. Rossi,<sup>19,13</sup> N. Rossi,<sup>7</sup> D. Rountree,<sup>20</sup> D. Sablone,<sup>13,7</sup> P. Saggese,<sup>12</sup> W. Sands,<sup>13</sup> C. Savarese,<sup>9,7</sup> B. Schlitzer,<sup>32</sup> E. Segreto,<sup>41</sup> D. A. Semenov,<sup>24</sup> E. Shields,<sup>13</sup> P. N. Singh,<sup>17</sup> M. D. Skorokhvatov,<sup>36,37</sup> O. Smirnov,<sup>28</sup> A. Sotnikov,<sup>28</sup> C. Stanford,<sup>13</sup> Y. Suvorov,<sup>27,7,36</sup> R. Tartaglia,<sup>7</sup> J. Tatarowicz,<sup>30</sup> G. Testera,<sup>11</sup> A. Tonazzo,<sup>1</sup> P. Trinchese,<sup>18,19</sup> E. V. Unzhakov,<sup>24</sup> M. Verducci,<sup>6,25</sup> A. Vishneva,<sup>28</sup> B. Vogelaar,<sup>20</sup> M. Wada,<sup>13</sup> S. Walker,<sup>18,19</sup> H. Wang,<sup>27</sup> Y. Wang,<sup>31,27</sup> A. W. Watson,<sup>30</sup> S. Westerdale,<sup>13</sup> J. Wilhelmi,<sup>30</sup> M. M. Wojcik,<sup>40</sup> X. Xiang,<sup>13</sup> X. Xiao,<sup>27</sup> J. Xu,<sup>13</sup> C. Yang,<sup>31</sup> W. Zhong,<sup>31</sup> C. Zhu,<sup>13</sup> and G. Zuzel<sup>40</sup>

The DarkSide Collaboration

---

<sup>a</sup>Deceased

- <sup>1</sup>APC, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Obs de Paris, USPC, Paris 75205, France
- <sup>2</sup>Instituto de Física, Universidade de São Paulo, São Paulo 05508-090, Brazil
- <sup>3</sup>Pacific Northwest National Laboratory, Richland, WA 99352, USA
- <sup>4</sup>Physics Department, Augustana University, Sioux Falls, SD 57197, USA
- <sup>5</sup>Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
- <sup>6</sup>INFN Sezione di Roma, Roma 00185, Italy
- <sup>7</sup>INFN Laboratori Nazionali del Gran Sasso, Assergi (AQ) 67100, Italy
- <sup>8</sup>INFN Cagliari, Cagliari 09042, Italy
- <sup>9</sup>Gran Sasso Science Institute, L'Aquila 67100, Italy
- <sup>10</sup>Physics Department, Università degli Studi di Genova, Genova 16146, Italy
- <sup>11</sup>INFN Genova, Genova 16146, Italy
- <sup>12</sup>INFN Milano, Milano 20133, Italy
- <sup>13</sup>Physics Department, Princeton University, Princeton, NJ 08544, USA
- <sup>14</sup>INFN Roma Tre, Roma 00146, Italy
- <sup>15</sup>Mathematics and Physics Department, Università degli Studi Roma Tre, Roma 00146, Italy
- <sup>16</sup>Physics Department, Università degli Studi di Cagliari, Cagliari 09042, Italy
- <sup>17</sup>Department of Physics, University of Houston, Houston, TX 77204, USA
- <sup>18</sup>Physics Department, Università degli Studi "Federico II" di Napoli, Napoli 80126, Italy
- <sup>19</sup>INFN Napoli, Napoli 80126, Italy
- <sup>20</sup>Virginia Tech, Blacksburg, VA 24061, USA
- <sup>21</sup>Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow 119991, Russia
- <sup>22</sup>Physics Department, Università degli Studi di Milano, Milano 20133, Italy
- <sup>23</sup>LPNHE, Université Pierre et Marie Curie, CNRS/IN2P3, Sorbonne Universités, Paris 75252, France
- <sup>24</sup>Saint Petersburg Nuclear Physics Institute, Gatchina 188350, Russia
- <sup>25</sup>Physics Department, Sapienza Università di Roma, Roma 00185, Italy
- <sup>26</sup>Department of Physics and Astronomy, University of Hawai'i, Honolulu, HI 96822, USA
- <sup>27</sup>Physics and Astronomy Department, University of California, Los Angeles, CA 90095, USA
- <sup>28</sup>Joint Institute for Nuclear Research, Dubna 141980, Russia
- <sup>29</sup>Amherst Center for Fundamental Interactions and Physics Department, University of Massachusetts, Amherst, MA 01003, USA
- <sup>30</sup>Physics Department, Temple University, Philadelphia, PA 19122, USA
- <sup>31</sup>Institute of High Energy Physics, Beijing 100049, China
- <sup>32</sup>Department of Physics, University of California, Davis, CA 95616, USA
- <sup>33</sup>IPHC, Université de Strasbourg, CNRS/IN2P3, Strasbourg 67037, France
- <sup>34</sup>School of Natural Sciences, Black Hills State University, Spearfish, SD 57799, USA
- <sup>35</sup>Radiation Physics Laboratory, Belgorod National Research University, Belgorod 308007, Russia
- <sup>36</sup>National Research Centre Kurchatov Institute, Moscow 123182, Russia
- <sup>37</sup>National Research Nuclear University MEPhI, Moscow 115409, Russia
- <sup>38</sup>Chemistry, Biology and Biotechnology Department, Università degli Studi di Perugia, Perugia 06123, Italy
- <sup>39</sup>INFN Perugia, Perugia 06123, Italy
- <sup>40</sup>M. Smoluchowski Institute of Physics, Jagiellonian University, 30-348 Krakow, Poland
- <sup>41</sup>Physics Institute, Universidade Estadual de Campinas, Campinas 13083, Brazil

# APPENDIX C: ARIS

The Argon Response Ionization and Scintillation (ARIS) experiment was planned during the spring and summer of 2016 and performed at *L'Institut de Physique Nucléaire d'Orsay* (IPNO) in Orsay, just outside of Paris, France, during the fall of 2016. This facility has a high-flux directional neutron beam, LICORNE,<sup>1</sup> which can be used for studies of nuclear recoils in, among other things, argon TPCs.

Like DarkSide-50, ARIS was a two-phase (liquid/gas) argon TPC. ARIS mainly used smaller, 1" square PMTs instead of DS-50's 3" diameter ones, and used fewer of them – an array of eight 1" PMTs covered the anode (top) of the active volume, while a single 3" PMT covered the cathode (bottom) of the active volume. The working principles behind ARIS and DS-50 are the same, however: scintillation in the liquid region gives S1, electroluminescence in the gas gap gives S2, an applied electric field drifts ionization electrons from the  $\ell\text{Ar}$  to the  $g\text{Ar}$ ,  $f_{90}$  is the primary particle identifier, and so on.

Like SCENE, though, ARIS's utility lies in the fact that it was small, mobile, and easy to work with. The ARIS Experiment has four main goals:<sup>2</sup>

1. Measure the scintillation efficiency of NRs for energies 16 keV to 130 keV
2. Map  $f_{90}$  as a function of energy and  $\mathcal{E}_d$
3. Measure the recombination probability as a function of energy and  $\mathcal{E}_d$

---

<sup>1</sup> <http://ipnwww.in2p3.fr/LICORNE>

<sup>2</sup> <http://aris.in2p3.fr/>

#### 4. Measurement of NR directionality

The first bullet point is precisely what the SCENE Experiment did, but for higher recoil energies. ARIS aims to confirm the results seen by SCENE, while extending that analysis to smaller NR energies. The second point ties in to Subsection 3.8.2: mapping the behavior of  $f_{90}$  for low-energy NRs is something which has never been done before. Gaining a greater understanding of this crucial PSD parameter would be invaluable to DS and other argon-based direct detection experiments.

The third and fourth points, unfortunately, have not been achieved at this time. During beam time at IPNO, there were unresolved problems with S2 generation in ARIS. Along with some difficulties with the beam itself, this led to points 3 and 4, above, becoming less significant in the analysis. However, a few early runs do have S2 signals (though they are small) which can be used for a pedagogical exercise on performing the position corrections of S1 and S2, similar to those performed in DarkSide-50.

As mentioned in Subsection 3.4.2, S1 and S2 are often “ $z$ -corrected” to account for the detector geometry and impurities in the  $\ell$ Ar. This is typically done in steps: first, S1 is  $z$ -corrected, then S2, and then either or both are corrected for radial or  $xy$  dependencies. The plots on the following pages all use a single ARIS Run (101076) and have cuts applied on  $t_{\text{drift}}$  (10, 120), S1  $f_{90}$  (0.2, 1), and S2  $f_{90}$  (0, 0.05).

The first step is to correct S1 as a function of  $t_{\text{drift}}$  ( $z$ ). As can be seen in Figure C.1, this is performed by simply plotting S1 (aka.  $S1_{\text{tot}}$ ) as a function of  $t_{\text{drift}}$  and curve-fitting the means of the horizontal bins. Since we don’t expect  $S1(t_{\text{drift}})$  to have any particular functional form, a 2nd order polynomial fit is usually sufficient. Curve-fitting and subtracting out all of the non-constant terms gives us a  $z$ -corrected  $S1_{\text{corr}}$  which is independent of  $t_{\text{drift}}$ . In addition,  $S1_{\text{corr}}$  is

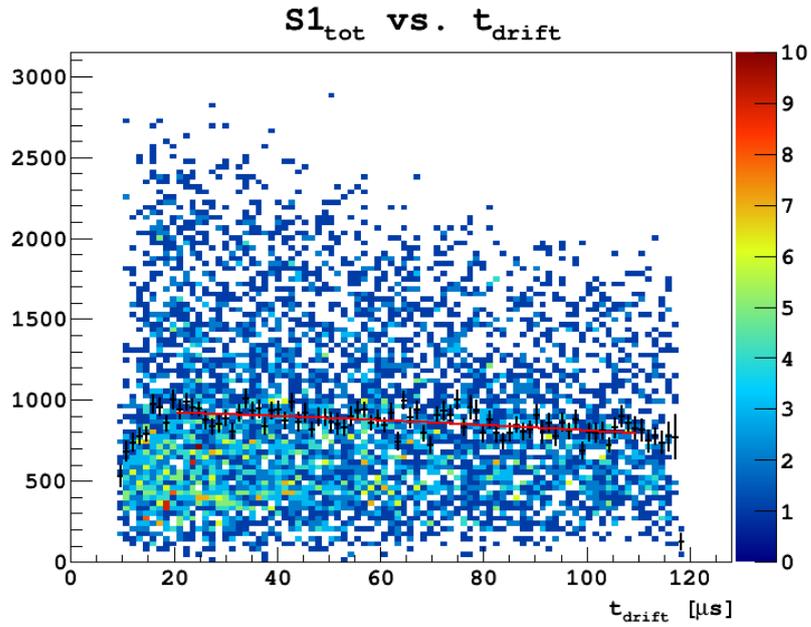
usually set to the value of S1 at the center of the detector.

Once S1 is  $z$ -corrected, S2 can be  $z$ -corrected, as well. In DarkSide-50, this step is not necessary, because there are so few electropositive impurities in the DS-50  $\ell$ Ar that the electron drift lifetime,  $\tau_{\text{ed}}$ , is too long to be measurable. In ARIS, this is unfortunately not the case. As can be seen in Figure C.2, the electron drift lifetime  $\tau_{\text{ed}}$ , is not much greater than the maximum  $t_{\text{drift}}$  in the detector, and so S2 is significantly reduced for events with high  $t_{\text{drift}}$ .

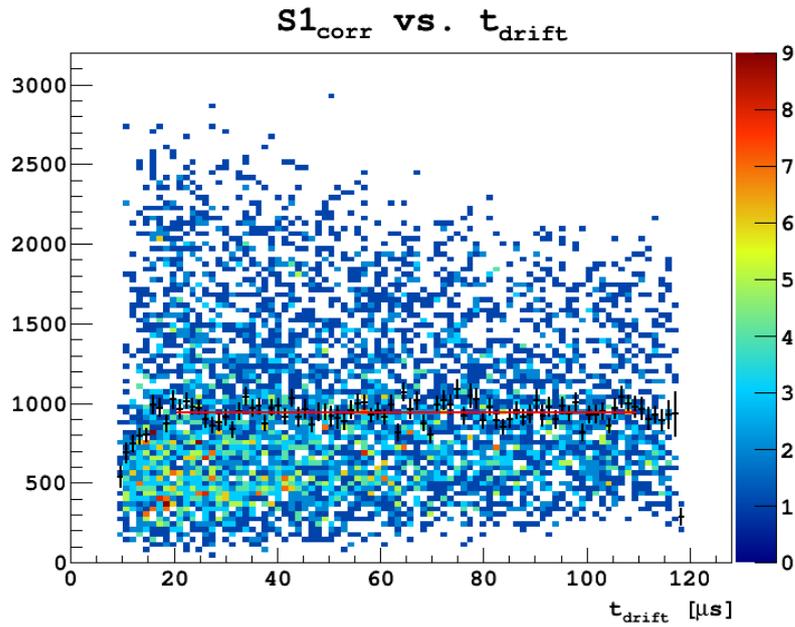
Once S1 and S2 have both been  $z$ -corrected, additional position dependencies can be investigated. In DarkSide-50, there is only a slight radial dependence of S1, but a significant  $xy$  dependence of S2. For instance, it has been shown that a slight tilt of the TPC results in higher average S2 for events which occur on one side of the detector. Additionally, S2 has a strong radial dependence in DS-50, which has been attributed to a slight “sagging” of the anode window. After all of these effects are accounted for, a “residual map” tells us that there are spots of reduced TPB coverage over the anode. DS-50 has many S2 position dependencies, and all must be accounted for.

ARIS is a bit easier. The only position-dependent S2 variation found to date is a detector tilt. While it’s possible that there could be radial effects or residuals, the size of the TPC works against the likelihood of these. Because the TPC is so small (8 cm diameter), it’s more likely that the TPB has been evenly coated on the inner surfaces, and less likely that any DS-50-like “sagging” would occur. However, neither of these effects have been investigated fully, because the MC does not match the data well enough to give reasonable event positions with the PCAMDF method. The detector tilt can be investigated with the barycenter method.

As can be seen in Figure C.3, there is a preference for high- $S2_{\text{corr}}/S1_{\text{corr}}$  events to occur to the right of (roughly) the line  $y = -x$ , and for low- $S2_{\text{corr}}/S1_{\text{corr}}$  events

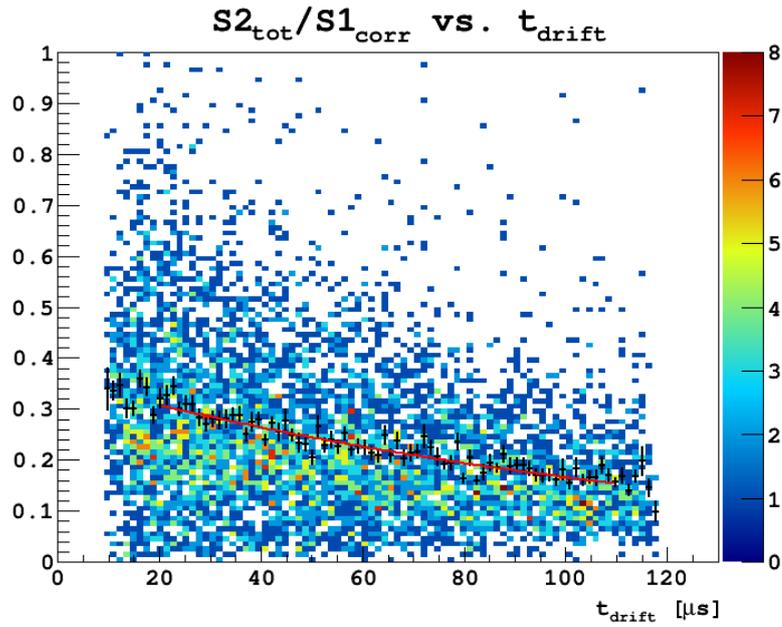


(a) Uncorrected S1 (aka.  $S1_{\text{tot}}$ ) vs.  $t_{\text{drift}}$  for ARIS Run 101076. The vertical axis is in PE. The black crosses show the vertical means of the horizontal bins and the red line is the 2nd-order polynomial curve fit. The curve fit function is  $S1_{\text{tot}}(t_{\text{drift}}) = -0.00451357 \cdot t_{\text{drift}}^2 - 0.850101 \cdot t_{\text{drift}} + 942.351$ .

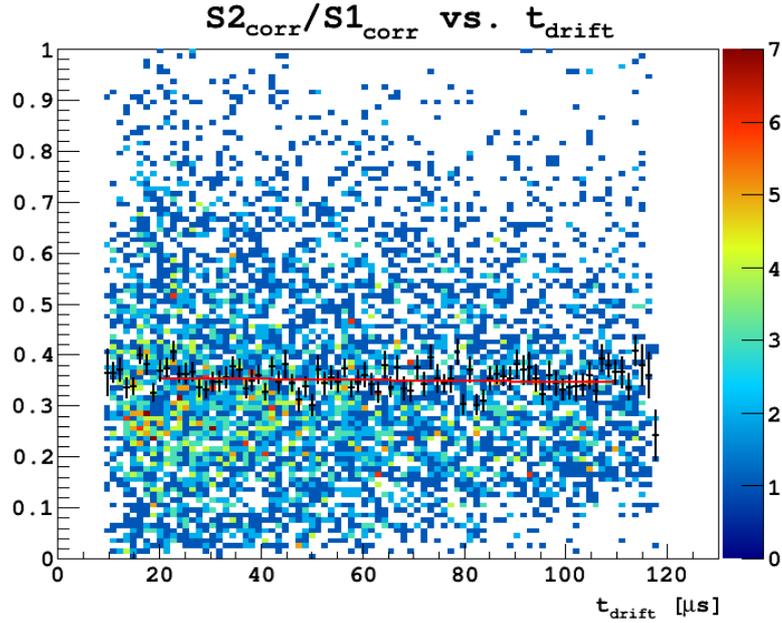


(b)  $z$ -corrected S1 vs.  $t_{\text{drift}}$  for ARIS Run 101076. The vertical axis is in PE. The black crosses show the vertical means of the horizontal bins and the red line is the 2nd-order polynomial curve fit. The curve fit function is  $S1_{\text{corr}}(t_{\text{drift}}) = -0.000413621 \cdot t_{\text{drift}}^2 + 0.0512403 \cdot t_{\text{drift}} + 940.638$ .

Figure C.1: Before (a) and after (b)  $z$ -correcting S1 in ARIS Run 101076. The detector geometry does not have much influence on the light yield in ARIS, so the errors on the non-constant curve fit parameters are consistent with zero both before and after S1 correction.



(a)  $S2_{\text{tot}}/S1_{\text{corr}}$  vs.  $t_{\text{drift}}$  for ARIS Run 101076. The vertical axis is unitless. The black crosses show the vertical means of the horizontal bins and the red line is the exponential curve fit. The curve fit function is  $S2_{\text{tot}}/S1_{\text{corr}}(t_{\text{drift}}) = \exp(-1.02138 - 0.00772852 \cdot t_{\text{drift}})$ .



(b)  $S2_{\text{corr}}/S1_{\text{corr}}$  vs.  $t_{\text{drift}}$  for ARIS Run 101076. The vertical axis is unitless. The black crosses show the vertical means of the horizontal bins and the red line is the exponential curve fit. The curve fit function is  $S2_{\text{corr}}/S1_{\text{corr}}(t_{\text{drift}}) = \exp(-1.03328 - 2.62446 \times 10^{-4} \cdot t_{\text{drift}})$ .

Figure C.2: Before (a) and after (b)  $z$ -correcting  $S2$  (aka.  $S2_{\text{tot}}$ ) in ARIS Run 101076. ARIS's  $\ell\text{Ar}$  is not as electropositively pure as DarkSide-50's, and so there is a noticeable  $t_{\text{drift}}$  dependence of  $S2_{\text{tot}}/S1_{\text{corr}}$ . The second term in the first exponential fit, 0.00772852, is the inverse of the electron drift lifetime,  $\tau_{\text{ed}} = 129.4 \mu\text{s}$ . (Compare to DS-50's  $> 5$  ms.)

to occur to the left of this line. One possible explanation for this behavior is that the detector is tilted such that these higher- $S2_{\text{corr}}/S1_{\text{corr}}$  events occur in a region where the  $g\text{Ar}$  gas gap is thicker, and the lower- $S2_{\text{corr}}/S1_{\text{corr}}$  events occur in a region where it is thinner. A thicker gas gap means more S2 is produced for an event with a given S1. This behavior has also been observed in DS-50.

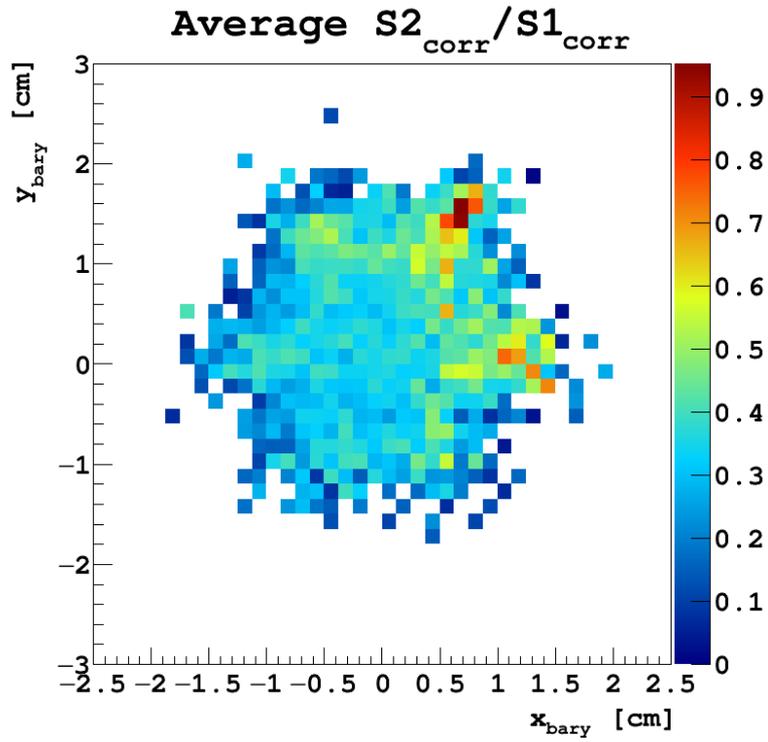
To calculate the tilt angle in DS-50, a function is fit to the distribution which also takes into account the radial dependence of S2 in that detector. But this function depends on events being reconstructed to their true (or nearly true) positions within the detector. That is not possible in ARIS, where only the barycenter  $xy$  is available. While a barycenter  $xy$  reconstruction is not particularly adept at reconstructing radial positions near the outer edge of a TPC, it does reconstruct the angular position fairly well, and so is sufficient for our purpose here.

Figure C.3 shows, in subfigure (b), the average  $S2_{\text{corr}}/S1_{\text{corr}}$  for ARIS Run 101076 as a function of the azimuthal position,  $\theta$ . The black crosses show the vertical means of the horizontal bins and the red line is a curve fit. Fitting this data is complicated by the fact that  $S2_{\text{corr}}/S1_{\text{corr}}$  shows peaks under each of the individual PMTs. MC simulations have shown that this behavior in DS-50 is due to a position-dependent likelihood of detecting the electroluminescence photons generated in the gas gap, and the same physics is likely at work here. To fit the data, then, we require an independent Gaussian for each of the six outer anode PMTs, plus a sinusoidal background — this background gives us the detector tilt. The best-fit function to this data is:

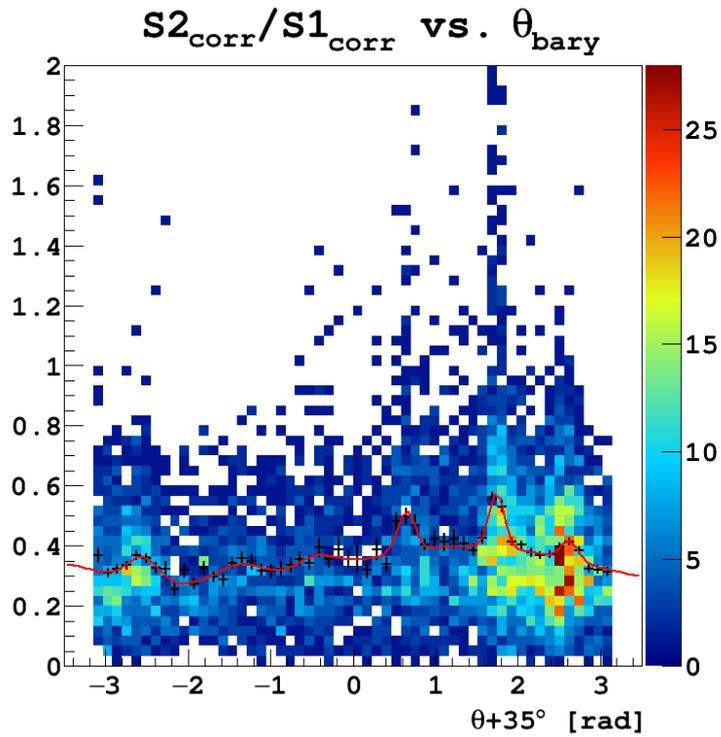
$$\begin{aligned}
f(x) = & 0.0786344 \cdot \exp\left(-\frac{(x + 2.54841)^2}{2 \cdot (0.181627)^2}\right) + 0.0679917 \cdot \exp\left(-\frac{(x + 1.37408)^2}{2 \cdot (0.300000)^2}\right) \\
& + 0.0519989 \cdot \exp\left(-\frac{(x + 0.462299)^2}{2 \cdot (0.246105)^2}\right) + 0.132270 \cdot \exp\left(-\frac{(x - 0.642349)^2}{2 \cdot (0.105138)^2}\right) \\
& + 0.185742 \cdot \exp\left(-\frac{(x - 1.74209)^2}{2 \cdot (0.0895908)^2}\right) + 0.0629997 \cdot \exp\left(-\frac{(x - 2.61911)^2}{2 \cdot (0.102794)^2}\right) \\
& + 0.0662158 \cdot \cos(x - 1.37027) + 0.332718 \quad (8.41)
\end{aligned}$$

The term we're interested in is the one inside the cosine, 1.37027 rad. This (minus the  $35^\circ$  added to the horizontal axis of the plot so that no PMT would be split at  $\pm\pi$ ), is the tilt angle of the ARIS TPC, about  $43.5^\circ$ . Subtracting this tilt from the  $S2_{\text{corr}}/S1_{\text{corr}}$   $xy$  map in Figure C.3 gives us Figure C.4, which still has “hot spots” on the two PMTs in the first quadrant, but is noticeably “flatter” in  $z$  than the map in Figure C.3. Note that this is the *azimuthal* angle of the tilt, not the polar angle, and so this says nothing about the *degree* to which the TPC is tilted, just that it is tilted.

ARIS just finished gathering data a few months ago and that data is still being analyzed. At the time of this writing, the ARIS Experiment has not yet published any findings, though we hope to in the near future. Future publications will surely be found on the experiment's website, <http://aris.in2p3.fr/>.



(a)



(b)

Figure C.3: Average  $S2_{\text{corr}}/S1_{\text{corr}}$  vs.  $xy$  (a) and vs.  $\theta$  (b) for ARIS Run 101076. The vertical axes are unitless. The black crosses show the vertical means of the horizontal bins and the red line in (b) is the curve fit described in the text. The preference for blue-green bins at the bottom-left of (a) and yellow-orange bins at the top-right of (a) indicates a detector tilt along that axis.

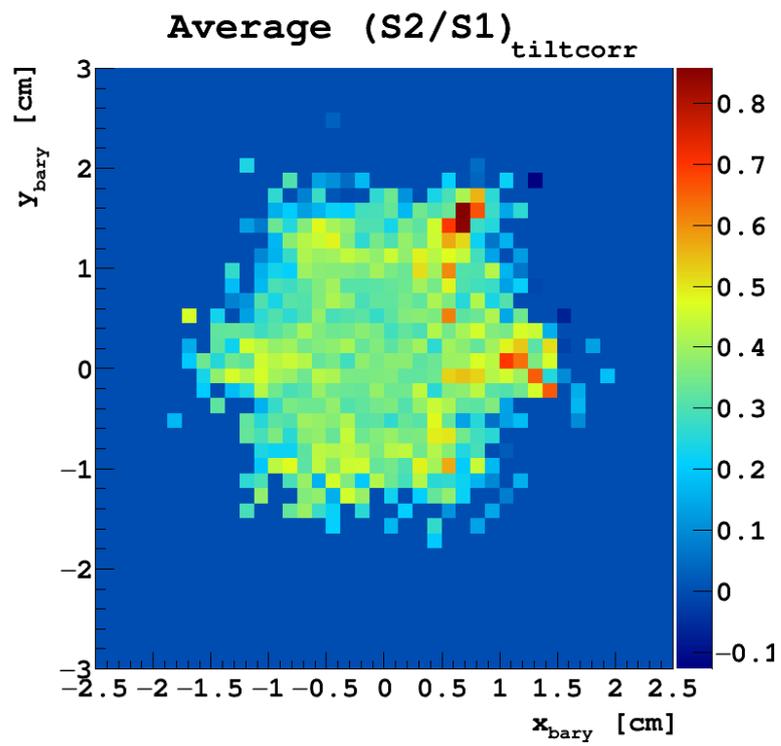


Figure C.4:  $xyz$ -corrected  $S2/S1$  map for ARIS Run 101076. Notice that the  $z$ -axis has a maximum of  $< 1$ : this speaks to how small the  $S2$  signals are in ARIS data.

# APPENDIX D: NEUTRON THERMALIZATION IN DarkSide-50

As noted in [Chapter 3](#), the performance of the LSV can be adjusted by changing the PPO or TMB concentrations in the PC organic liquid scintillator which fills the veto volume. Changing the PPO concentration changes the light yield, and also the  $\alpha$  scintillation quenching, as discussed in Subsection [3.4.2](#) (“CALIS”). Changing the TMB concentration affects the neutron capture time in the LSV.

In Subsection [3.7.2](#), it was mentioned that the LSV neutron capture time was  $\sim 2 \mu\text{s}$  when the PC/TMB ratio was 50/50 (50% TMB). When the ratio was changed to 19/1 (5% TMB), the capture time increased by a factor of ten. The procedure for measuring the neutron capture time in the LSV is outlined below.

The DarkSide-50 runs used in this study are numbers 11131–48, excluding 11136, 11140, 11145, and 11147. These are all low-rate  $^{241}\text{Am}^9\text{Be}$  source runs. The cuts applied are CX cuts (see Section [3.7](#)) 1–4, 8–13, and 15. Additionally, a rather liberal S1 cut is applied (80, 5000), and a cut is made to select events within the DM search box (see Figure [3.17](#)), of course with the high S1 boundary removed.<sup>1</sup>

---

<sup>1</sup> See internal docdb entry # 1414-v2.

The data structure for the vetoes is a bit different than for the TPC. While TPC pulses can be found and identified as S1 or S2 based on their  $f_{90}$ , “pulses” in the LSV are typically grouped together as “clusters”. The cluster-finding algorithm works quite differently than the TPC pulse-finding algorithm, and is described in detail in [Westerdale \(2016\)](#). The end result is that, for each event, there can be any number of clusters in the LSV. Each cluster has a timestamp (relative to the time of the S1 pulse in the TPC, either in  $\mu\text{s}$  or ns) and a charge (measured in PE). The CX cuts in Section 3.7 describe how LSV information is used for particle identification and event rejection. To study the neutron thermalization signal and the TMB neutron capture time in the LSV, though, we can entirely disregard the TPC.

Figure D.1 shows all of the LSV clusters for the DS-50 runs listed above. Both axes are logarithmic so that all classes of clusters can be identified. An offset of 200 ns is added to the horizontal (time) axis so that clusters around the S1 start time (the solid black line in subfigure (b)) can also be seen.

In subfigure (b), the brick red box at the top-left shows high-energy prompt clusters from the 4.4 MeV  $\gamma$  emitted by the  $^{241}\text{Am}^9\text{Be}$  source. The orange box below shows degraded gammas from the red box and high-energy neutrons emitted by the source, many of which interacted in the LSV prior to interacting in the TPC (the S1 start time is the solid black line). The green box below that shows low-energy neutrons which first interacted in the TPC and then thermalized in the LSV.

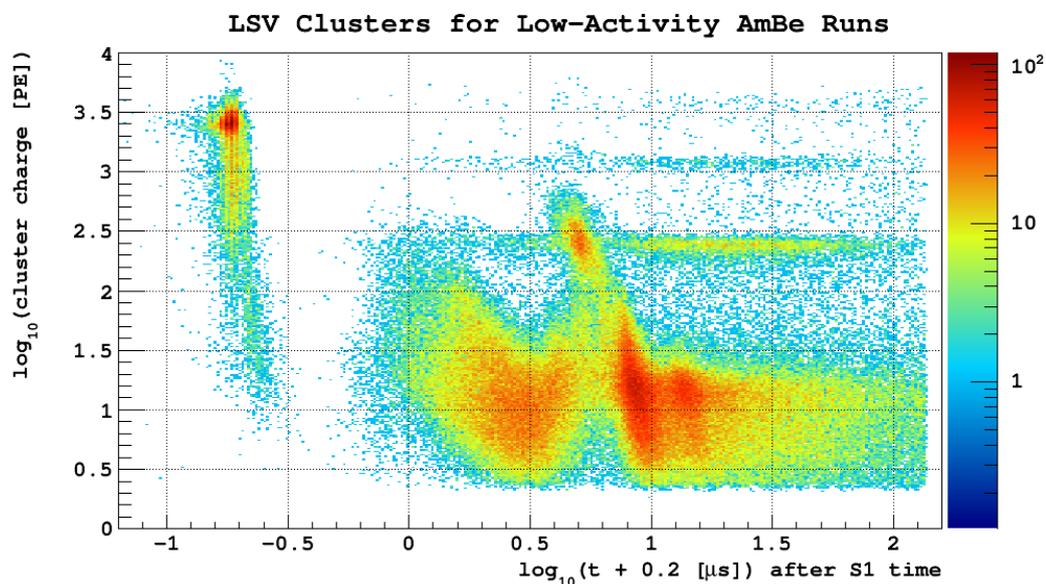
On the right-hand side of subfigure (b), at the top we see the signal from neutron capture on  $^{12}\text{C}$ , surrounded by the grey box. Below that is the neutron capture on  $^1\text{H}$ , surrounded by the pink box, then the neutron capture on  $^{10}\text{B}$ , surrounded by the blue box. Sitting on top of the blue distribution — which is the one we want to investigate — is a group of after-pulses, attributed to the

prompt 4.4 MeV  $\gamma$  events. Finally, at the bottom, in the brown boxes, we have several additional groups of low-energy after-pulses.

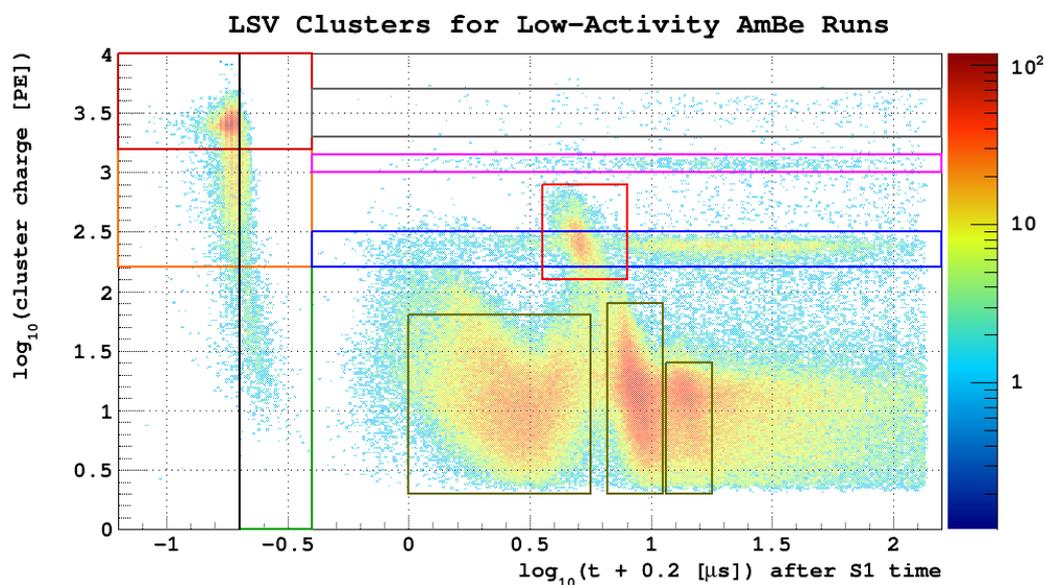
To study the  $^{10}\text{B}$  signal (the TMB capture signal), we must remove the after-pulses in the red box. Luckily, this is easily achieved since these after-pulses are correlated with the prompt 4.4 MeV  $\gamma$  clusters. As can be seen in Figure D.2, removing events with clusters in the brick red box at the top-left also removes clusters in the bright red box near the center of the plot. This also shrinks the extent of the brown-boxed classes of after-pulses, pulling their high-energy extents away from the  $^{10}\text{B}$  band. Once removed, we can measure the TMB capture time (for events in the blue box in Figure D.1 (b)), as well as the neutron thermalization signal (events in the green box in the same subfigure).

Looking only at late clusters ( $t > 200$  ns), we can easily identify the peaks in the energy spectrum of Figure D.3. From left-to-right, they are: after-pulses ( $\langle E \rangle \sim 14$  PE), neutron capture on  $^{10}\text{B}$  ( $\langle E \rangle \sim 240$  PE), neutron capture on  $^1\text{H}$  ( $\langle E \rangle \sim 1150$  PE), and two small peaks which are likely neutron capture on  $^{12}\text{C}$  ( $\langle E \rangle \sim 2500$  PE) and on the iron in the stainless steel of the cryostat ( $\langle E \rangle \sim 3800$  PE). The clusters we're interested in are shown in the red box of Figure D.3, with energies around 250 PE. The time distribution of these clusters can be seen in Figure D.4. The best-fit function gives a neutron capture time of  $\sim 18$   $\mu\text{s}$  for a 5% TMB concentration in the LSV.

Finally, we can look at the clusters in the green box in Figure D.1 (b). These are low-energy clusters which occur after the S1 start time — this is where we would expect to find neutrons which were emitted from the source and first interacted in the TPC before they deposited any energy in the LSV. Simply plotting the time and energy distributions of these clusters, as we do in Figure D.5, allows us to characterize them. The mean cluster time gives the mean travel time from the point of the neutron's S1 interaction in the TPC to the point of its interaction in



(a) LSV clusters for the DS-50 run numbers given in the text. The vertical axis gives the cluster charge and the horizontal axis gives the cluster time, with a small offset. Several groups of clusters can be seen.



(b) The same LSV clusters as seen in subfigure (a), but grouped into classes. A full description of all classes of clusters can be found in the text.

Figure D.1: LSV clusters in low-rate  $^{241}\text{Am}^9\text{Be}$  data.

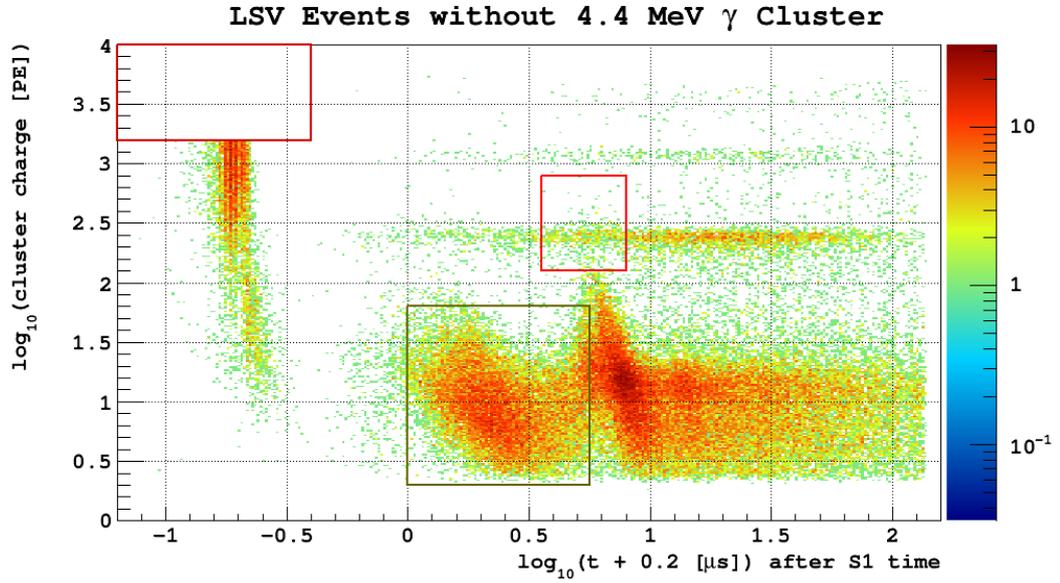


Figure D.2: Removing events with 4.4 MeV  $\gamma$  clusters (the brick red box at the top-left) also removes their after-pulses (the bright red box near the center). The high-energy extent of the first of the brown-boxed after-pulse classes is also reduced, significantly cleaning the  $^{10}\text{B}$  band around 250 PE.

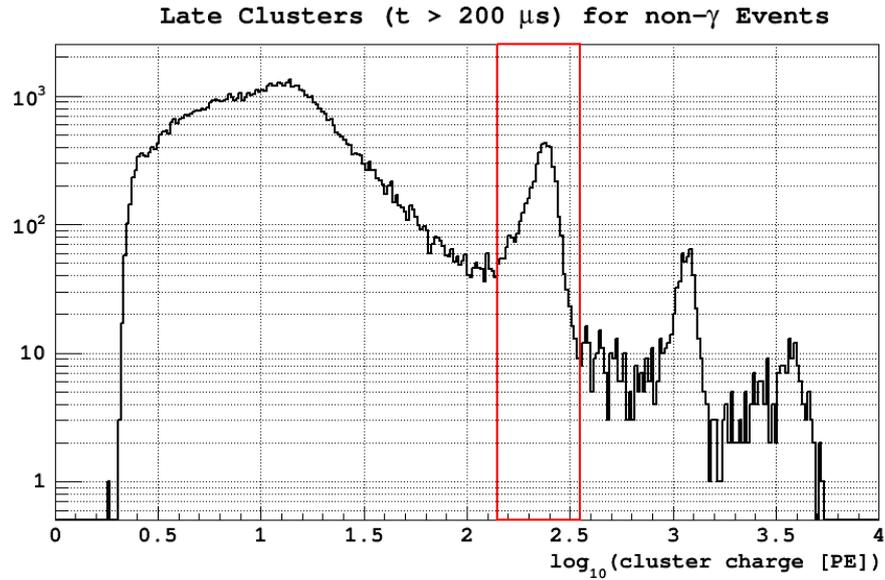


Figure D.3: Energy spectrum for the late ( $t > 200$  ns) LSV clusters from Figure D.2. The peaks are described in the text.

### $^{10}\text{B}$ Capture-like Clusters

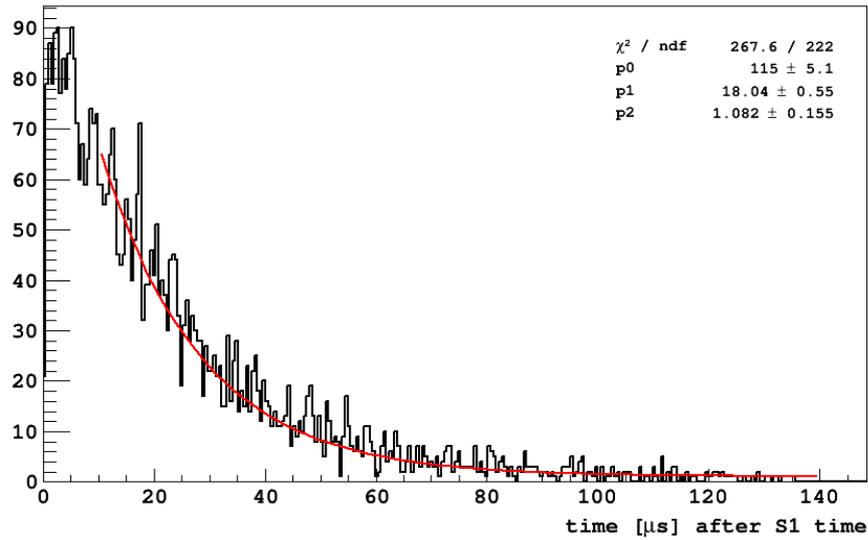
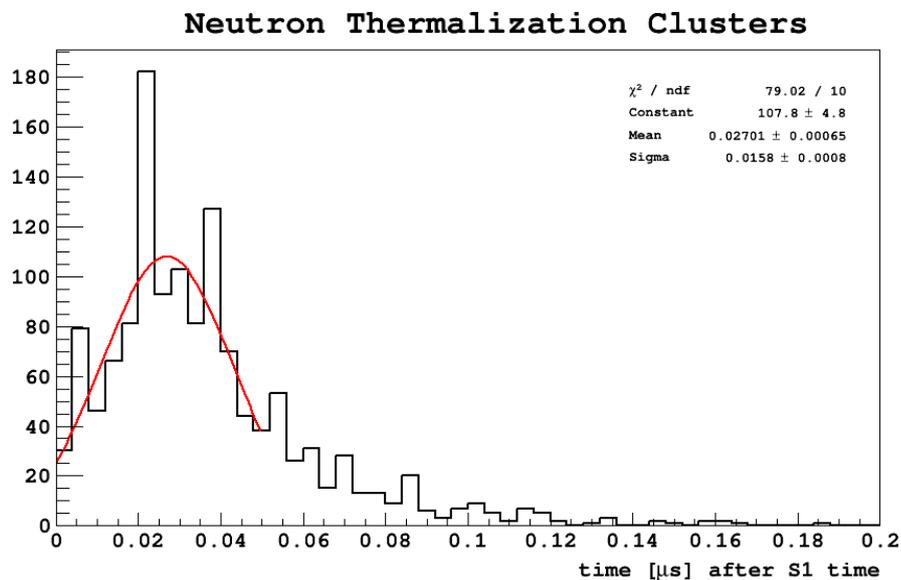


Figure D.4: Time distribution of the clusters in the red box in Figure D.3. The applied energy cut is (150 PE  $\rightarrow$  300 PE). The fit function is  $f(x) = p_0 \cdot \exp(-x/p_1) + p_2$ , giving a best-fit neutron capture time of  $\tau = 18.04 \pm 0.55 \mu\text{s}$  for a 5% TMB concentration in the LSV. Note that changing the energy range or the fit range can vary the lifetime by a few  $\mu\text{s}$ , so there are some systematic effects to take into account.

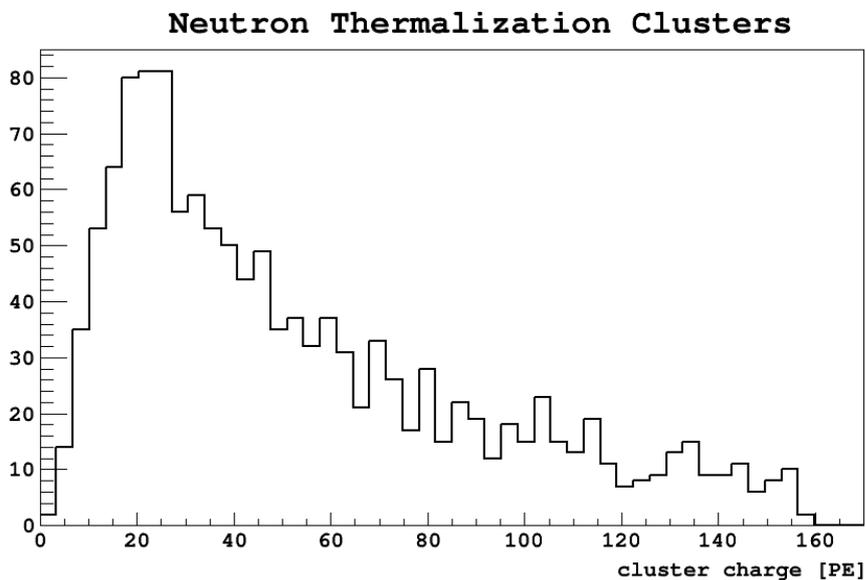
the LSV. At  $\sim 27 \text{ ns}$ , this time is consistent with the travel time of a  $\sim 3 \text{ MeV}$  neutron traveling the radius of the DS-50 cryostat ( $\sim 32 \text{ cm}$ ) plus a few mean free paths<sup>2</sup> — further evidence for these being neutron thermalization clusters.

---

<sup>2</sup> Traveling with a speed of  $\sim 2.5 \text{ cm/ns}$ .



(a) Time distribution of the clusters in the green box in Figure D.1 (b). The mean of the Gaussian fit gives the average neutron travel time from the point of the S1 interaction in the TPC to the point of its interaction in the LSV:  $27.0 \text{ ns} \pm 0.7 \text{ ns}$ .



(b) Energy distribution of the clusters in the green box in Figure D.1 (b). The peak value is around 20 PE.

Figure D.5: Time (a) and energy (b) distributions of the neutron thermalization events in the green box of Figure D.1 (b).

# APPENDIX E: SAMPLE RECONSTRUCTION CODE

minimal.C

---

```
///=====
///
/// minimal.C | Andrew W Watson | 9 March 2017
///
/// Minimal working PCAMDF example for MC events. The script
/// reads in S2tot, S2top, x, y, and S2 fractions from an ext-
/// ernal ROOT file, performs the PCA and the MDF, then recon-
/// structs events from the END of the file, to avoid giving an
/// erroneously good reconstruction. Full description in text.
///
///=====

///=====
///
/// includes, subroutines, and non-user-controlled parameters

#include "TFile.h"
#include "TTree.h"
#include "TPrincipal.h"
#include "TMultiDimFit.h"
#include "TMath.h"

TPrincipal pca(19, "ND");

TMultiDimFit* MDFx
    = new TMultiDimFit(19, TMultiDimFit::kChebyshev, "V");

TMultiDimFit* MDFy
    = new TMultiDimFit(19, TMultiDimFit::kChebyshev, "V");

float x, y, S2tot, S2top, S2[38];
double Xi[19], Yi[19];
float xprime, yprime;
int ii(0), ll;

void setUpMDFs();
```

```

bool applyCuts();
void constructXi();

///
///=====

//=====
//
// user-controlled parameters

int NE = 10000;

//
//=====

////////////////////////////////////
/// int main()
////////////////////////////////////

int main(){

  setUpMDFs(); // move this to a subroutine to keep main() neat

  ///-----
  /// read data from input file
  ///-----

  TFile* mcfile = new TFile("external_file.root");
  TTree* mctree = (TTree*)(mcfile->Get("dstree"));
  int nentries = mctree->GetEntries();

  mctree->SetBranchAddress("x", &x);
  mctree->SetBranchAddress("y", &y);
  mctree->SetBranchAddress("S2tot", &S2tot);
  mctree->SetBranchAddress("S2top", &S2top);
  mctree->SetBranchAddress("S2", S2);

  ///-----
  /// set up output file
  ///-----

  TFile* outfile = new TFile("minimal.root", "RECREATE");
  TTree* outtree = new TTree("outtree", "minimal.C_output_TTree");

  outtree->Branch("x", &x, "x/F");
  outtree->Branch("y", &y, "y/F");
  outtree->Branch("S2tot", &S2tot, "S2tot/F");
  outtree->Branch("S2top", &S2top, "S2top/F");
  outtree->Branch("S2", S2, "S2[38]/F");

  outtree->Branch("xprime", &xprime, "xprime/F");
  outtree->Branch("yprime", &yprime, "yprime/F");
  outtree->Branch("Xi", Xi, "Xi[19]/D");
  outtree->Branch("Yi", Yi, "Yi[19]/D");

  ///-----
  /// event loop #1: construct covariance matrix
  ///-----

```

```

for(int i = 0; i < nentries; ++i){
    if(ii >= NE) break; mctree->GetEntry(i);
    if(!applyCuts()) continue;
    constructXi();
    pca.AddRow(Xi);
    ++ii; ll = i;
} ii = 0;

// make P matrix, print results of PCA to screen
pca.MakePrincipals();
pca.Print();

///-----
/// event loop #2: add data to MDF object
///-----

for(int i = 0; i < nentries; ++i){
    if(ii >= NE) break; mctree->GetEntry(i);
    if(!applyCuts()) continue;
    constructXi();
    pca.X2P(Xi, Yi);
    MDFx->AddRow(Yi, x);
    MDFy->AddRow(Yi, y);
    ++ii;
} ii = 0;

// find fit functions for x and y
MDFx->FindParameterization();
MDFy->FindParameterization();

///-----
/// event loop #3: reconstruct events from end of file
///-----

for(int i = nentries; i > ll; --i){
    if(ii >= NE) break; mctree->GetEntry(i);
    if(!applyCuts()) continue;
    constructXi();
    pca.X2P(Xi, Yi);
    xprime = MDFx->Eval(Yi);
    yprime = MDFy->Eval(Yi);
    outtree->Fill();
    ++ii;
}

outfile->Write();
return 0;
}

////////////////////////////////////
/// void setUpMDFs()
////////////////////////////////////

void setUpMDFs(){

    int size = NE*0.8;

    // _PowLimit and _MaxPowers must be hand-tuned!
    int xPowLimit(2), yPowLimit(2);

```

```

int xMaxPowers[19] = {
    7, 7, 6, 6, 5, 5, 4, 4, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2 };

int yMaxPowers[19] = {
    7, 7, 6, 6, 5, 5, 4, 4, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2 };

MDFx->SetMaxFunctions      ( size );
MDFx->SetMaxStudy          ( size );
MDFx->SetMaxTerms          ( size );

MDFy->SetMaxFunctions      ( size );
MDFy->SetMaxStudy          ( size );
MDFy->SetMaxTerms          ( size );

MDFx->SetMaxPowers         ( xMaxPowers );
MDFx->SetPowerLimit        ( xPowLimit );
MDFx->SetMinRelativeError ( 0.000001 );

MDFy->SetMaxPowers         ( yMaxPowers );
MDFy->SetPowerLimit        ( yPowLimit );
MDFy->SetMinRelativeError ( 0.000001 );

return;
}

////////////////////////////////////
///  bool  applyCuts()
////////////////////////////////////

bool applyCuts(){

    bool pass = true;

    // check the S2 fractions to make sure they're ok
    for(int pmt = 0; pmt < 38; ++pmt){
        if(TMATH::IsNaN(S2[pmt])||
            !TMATH::Finite(S2[pmt])||S2[pmt]<=0){
            pass = false; break;
        } } if(!pass) return pass;

    // check S2tot and S2top to make sure they're ok
    if(TMATH::IsNaN(S2tot)||TMATH::IsNaN(S2top)||
        !TMATH::Finite(S2tot)||!TMATH::Finite(S2top)||
        S2tot<=0||S2top<=0)
        return false;

    // drop events with S2tot<10, something went wrong
    if(S2tot<10) return false;

    return true;
}

////////////////////////////////////
///  void  constructXi()
////////////////////////////////////

void constructXi(){

```

```
// this is the standard Xi vector I construct
for(int nn = 0; nn < 19; ++nn)
    Xi[nn] = -log(S2[nn+19]*S2tot/S2top);
return;
}
```

# recononly.C

---

```
///=====
///
/// recononly.C | Andrew W Watson | 9 March 2017
///
/// Accompanying script to minimal.C. Reads in PCAMDF config-
/// uration from external files, instead of generating it. This
/// makes the reconstruction much faster. A full description is
/// given in the text.
///
///=====
///=====
///
/// includes, subroutines, and non-user-controlled parameters

#include "TFile.h"
#include "TTree.h"
#include "TPrincipal.h"
#include "TMultiDimFit.h"
#include "TMath.h"

#include "pca.C"
#include "MDFxMDF.cxx"
#include "MDFyMDF.cxx"

MDFx mdfx;
MDFy mdfy;

float x, y, S2tot, S2top, S2[38];
double Xi[19], Yi[19];
float xprime, yprime;
int ii(0);

bool applyCuts();
void constructXi();

///
///=====
//=====
//
// user-controlled parameters

int NE = 1000000;

//
//=====
////////////////////////////////////
/// int main()
////////////////////////////////////

int main(){

  ///-----
```

```

/// read data from input file
///-----

TFile* mcfile = new TFile("external_file.root");
TTree* mctree = (TTree*)(mcfile->Get("dstree"));
int nentries = mctree->GetEntries();

mctree->SetBranchAddress("x",      &x);
mctree->SetBranchAddress("y",      &y);
mctree->SetBranchAddress("S2tot",  &S2tot);
mctree->SetBranchAddress("S2top",  &S2top);
mctree->SetBranchAddress("S2",    S2);

///-----
/// set up output file
///-----

TFile* outfile = new TFile("recononly.root", "RECREATE");
TTree* outtree = new TTree("outtree", "recononly.C_output_TTree");

outtree->Branch("x",      &x,      "x/F");
outtree->Branch("y",      &y,      "y/F");
outtree->Branch("S2tot",  &S2tot,  "S2tot/F");
outtree->Branch("S2top",  &S2top,  "S2top/F");
outtree->Branch("S2",    S2,      "S2[38]/F");

outtree->Branch("xprime", &xprime, "xprime/F");
outtree->Branch("yprime", &yprime, "yprime/F");
outtree->Branch("Xi",    Xi,      "Xi[19]/D");
outtree->Branch("Yi",    Yi,      "Yi[19]/D");

///-----
/// reconstruct events from end of file
///-----

for(int i = nentries; i >= 0; --i){
    if(ii >= NE) break; mctree->GetEntry(i);
    if(!applyCuts()) continue;
    constructXi();
    X2P(Xi, Yi);
    xprime = mdfx.MDF(Yi);
    yprime = mdfy.MDF(Yi);
    outtree->Fill();
    ++ii;
}

outfile->Write();
return 0;
}

bool applyCuts(){
    // same as in minimal.C
}

void constructXi(){
    // same as in minimal.C
}

```

## MDFx.h

---

```
class MDFx {  
  
    public:  
  
        static int gNVariables; // Number of variables  
        static int gNCoefficients; // Number of terms  
        static double gDMean; // Mean from training sample  
        static double gXMean[]; // Mean from training sample  
        static double gXMin[]; // Min from training sample  
        static double gXMax[]; // Max from training sample  
        static double gCoefficient[]; // Coefficients  
        static double gCoefficientRMS[];  
        static int gPower[]; // Function powers  
  
        static double MDF(double *x);  
  
};
```

## MDFy.h

---

```
class MDFy {  
    // same as in MDFx.h  
};
```

## fiducial.C

---

```
///=====
///
/// fiducial.C | Andrew W Watson | 10 March 2017
///
/// Script to generate fiducial volume cuts, using processed MC
/// events as input. MC events must have already been recon-
/// structed with PCAMDF XY. This script needs true and recon-
/// structed XY positions to run.
///
///=====

///=====
///
/// includes, subroutines, and non-user-controlled parameters

#include <iostream>
#include <map>

using namespace std;

#include "TROOT.h"
#include "TFile.h"
#include "TTree.h"
#include "TMath.h"
#include "TGraph.h"
#include "TAxis.h"
#include "TH1.h"
#include "TCanvas.h"

///
///=====

////////////////////////////////////
/// int main()
////////////////////////////////////

int main(){

    // comment out if unwanted: sets graphics preferences
    gROOT->ProcessLine(".x_~/rootstart.C");

    // load file and read in TTree; get # entries in TTree
    TFile* infile = new TFile("recononly.root");
    TTree* intree = (TTree*)(infile->Get("outtree"));
    int nentries = intree->GetEntries();

    // declare some parameters and link them to the TTree
    float x, y, r, xprime, yprime, rprime;

    intree->SetBranchAddress("x", &x);
    intree->SetBranchAddress("y", &y);

    intree->SetBranchAddress("xprime", &xprime);
    intree->SetBranchAddress("yprime", &yprime);

    // create a std::map for easy (automatic) sorting
```

```

std::map<float,float> map_rprime_v_r;

// loop over tree and fill map
for(int ee = 0; ee < nentries; ++ee){
    intree->GetEntry(ee);
    r = sqrt(x*x+y*y);
    if(r < 16.77) continue; // wall events only!
    rprime = sqrt(xprime*xprime+yprime*yprime);

    // fill map: sorted by key (rprime) value automatically
    map_rprime_v_r[rprime] = r;
}

// get length of map and create map iterator
int len = map_rprime_v_r.size();
std::map<float,float>::iterator it;

// tell the user about the percentiles
cout << "\n#\ "wall events\ "(r>=16.77cm): " << len
      << "\n\n%ile_index\t%ile_value\t"
      << "%ile_entry#\t%ile_radius" << endl;

// set specific percentiles that you want radii for

const int npctiles = 5;
float pctiles[npctiles] = { 50, 90, 95, 99, 99.9 };

bool pct_found[npctiles]; // have we found this %ile radius yet?
int  pct_entry[npctiles]; // entry in map of given pctile
int  pct_count; // counter

for(int pc = 0; pc < npctiles; ++pc)
    pct_entry[pc] = (1.0-pctiles[pc]/100.0)*len;

float radius[npctiles]; // fid cut radius

// reset counter and flags
pct_count = 0;

for(int ff = 0; ff < npctiles; ++ff)
    pct_found[ff] = false;

// search map
for(it=map_rprime_v_r.begin();it!=map_rprime_v_r.end();++it){
    for(int pct = 0; pct < npctiles; ++pct)
        if(pct_count>=pct_entry[pct] && !pct_found[pct]){
            radius[pct] = it->first; pct_found[pct] = true; }
        ++pct_count;
}

// alert user
for(int pct = 0; pct < npctiles; ++pct)
    cout << pct << "\t\t" << pctiles[pct] << "\t\t"
         << pct_entry[pct] << "\t\t" << radius[pct] << endl;

// declare TGraph object to draw plot
TGraph* tgraph_rprime = new TGraph();

pct_count = 0;

```

```

for(it=map_rprime_v_r.end();it!=map_rprime_v_r.begin();--it){
    tgraph_rprime->SetPoint(tgraph_rprime->GetN(),
                           (float)(pct_count)/len, it->first);
    ++pct_count;
}

// create a TCanvas on which to draw the plot
TCanvas* can = new TCanvas("can", "can", 800, 800);

// we need to Draw() before we can change the axis extents
tgraph_rprime->Draw("AC");
tgraph_rprime->GetXaxis()->SetRangeUser(0,1);
tgraph_rprime->GetYaxis()->SetRangeUser(10,20);

// set the gridlines
gPad->SetGridx(); gPad->SetGridy();

// set the title and save!
tgraph_rprime->GetHistogram()->SetTitle(
    "Fiducial_Cut;fraction_of_wall_events_cut;r_{cut}_[cm]");
can->SaveAs("fiducial.png");

return 0;
}

```