Measurements of the Neutron Magnetic Form Factor and the Two-Photon Exchange Contribution to the Electron-Neutron Elastic Scattering Cross Section

John Alan Boyd III

A Dissertation presented to the Graduate Faculty of the University of Virginia in Candidacy for the Degree of Doctor of Philosophy

Department of Physics

University of Virginia

April, 2024

Abstract

The Super BigBite Spectrometer (SBS) program in Experimental Hall A of Jefferson Lab is a series of high-precision, large Q^2, electromagnetic form factor experiments. GMn and nTPE were the first two experiments of the SBS program and ran from October 2021 through February 2022. Both experiments detect neutron-tagged $d(e, e'n)$ and proton-tagged $d(e, e'p)$ quasi-elastic electrons scattering from a deuterium target. The GMn experiment is a measurement of the magnetic form factor of the neutron, G_M^n, via the ratio method over a Q^2 range of 3.5 to 13.5 $(\mathrm{GeV/c})^2$. The nTPE experiment ran in-parallel with the GMn experiment and shared an overlapping kinematic point with GMn. nTPE extends the GMn experiment to assess the differences between Rosenbluth Slop and the single-photon exchange approximation. The nTPE experiment had two kinematic settings corresponding to different electron scatting angles at a fixed Q^2 of 4.5 $(\mathrm{GeV/c})^c$. Comparisons of nucleon form factor ratio measurements acquired through polarization transfer and LT separation reveal discrepancies as Q^2 increases. Two-photon exchange is strongly favored as the reason for this discrepancy. G_M^n has been extracted at two kinematic settings corresponding to Q^2 values of 3.0 and 4.5 $(\mathrm{GeV/c})^2$. Additionally, I will present an assessment of the differences between Rosenbluth Separation and single-photon exchange measurements.

Dedicated to my wife Victoria who always believed in my stubbornness and foolish tenacity in seeing this thing through...

...and Pi for helping me write it.

Acknowledgements

This thesis would not be possible without the guidance and support of so many people. It would be impossible to list every individual person and effort that bolstered this work, but I will try my best. Any omissions are due to my own oversight, lack of focus, and poor memory (blame it on my old age).

First, I would like to thank my advisor Nilanga Liyanage. Since day one, he has been a pillar of support and confidence. He always trusted my strengths and challenged my weaknesses (for the better). He also fostered so many of the academic connections that I've developed over the past several years; most of the people on this list are connected through Dr. Liyanage. I can't express the amount of thanks and gratitude that I feel, and know, are deserved. I am proud to have called you my advisor, and there is no one else that I would have wanted to take this journey with.

I would also like to thank my co-advisor Huong Nguyen for her un-wavering support and friendship throughout the entire time I've known her. Her positive attitude and drive was always an inspiration. I'd also like to extend my gratitude to Kondo Gnanvo for ushering me into so much of my early research.

I want to thank all of my colleagues within Dr. Liyanage's research group, especially Anuruddha Rathnayake and Xinzhan Bai. Anu made it through boot camp... I mean UVa Graduate Physics first year and quals... with me. We spent countless hours working over homework, qual prep, and complaining about TA-ing (we still do). And then there were the many, many, hours of GEM, experimental, and analysis work that we shared at JLab. I will always have cherished memories of those experiences. Xinzhan provided me with so much help when we both knew that it was just going to lead to me asking for more help. I appreciated every bit of it and will always be thankful for all of his guidance. Truly every student in Dr. Liyanage's group has provided so much support and enthusiasm: Sean Jeffas, Salina Ali, Bhashitha Dharmasena, Vimukthi Gamage, John Matter, Siyu Jian, and Jacob McMurty.

I have so many colleagues at JLab to thank. Because of the sweat and tears that we've shared I'll start with the graduate students: Sebastian Seeds, Provakar Datta, Zeke Wertz, Maria Satnik, Catherine Feldman, Manjukrishna Suresh, and Saru Dhital.

The entire SBS Collaboration deserves my deepest gratitude. I would like to extend a special thanks to Brad Sawatzky for all of the invaluable guidance that they provided. His advice truly influenced my path and outlook, and I am extremely grateful for that. I would also like to extend a thanks to Eric Fuchey, Andrew Puckett,

Bogdan Wojtsekhowski, Thir Gautam, Ben Raydo, Marc McMullen, Bryan Moffit, Arun Tadepalli, and Holly Szumila-Vance.

I would not have made it here were it not for the tunes that I hear in my head and my outlets to get them out! That's right, all of you musicians who helped me skank my way along this path. Fahad Bishara has been my musical Gibraltar: steadfast, immovable, and true. He has been a musical bloodline from day one. All of the projects that we participated in were special because of the magic we share. This goes the same for the other Larry's, Zane and Joe — the leisure suit store called and they're running out of you! I will also never forget the escape and friendship I found up on the mount by North Garden with Jeremy Layel and The Charlatones. I also can't say thank you enough to Clay Ames and the Rootz Equation foundation (Chuck Laney and Ed Hood). I enjoyed every last (one) drop of it. Thank you all so much for literally providing the music of my life.

I would also like to thank my thesis committee members for their time, support, and patience. Additionally, the entire UVa Physics department has provided such a foundation for my learning experience and opportunities. Similarly, the UVa Physics Bridge program is probably the only reason why I was even given a chance. Therefore, I offer special thanks to Chris Neu, Olivier Pfister, and Nilanga for their hand in procuring the program and offering me special guidance through that process. I know there are many more involved in that program and they all deserve a big thanks.

A huge loving thanks to my entire — yes all you — family. Especially, my parents Sam and Kim, and my siblings James, Sara, Anne, and Olivia (and Ripley too). They have all offered so much support, guidance, smiles, and laughter. They have always been willing to lend a loving ear, offer their time, provide nothing but love, and just be there for me.

Lastly, and certainly not leastly, *je tiens à remercier ma merveilleuse épouse, Victoria.* I can't express how important you are to me and how much of a wonderful influence you provide. Hopefully, I am able to show that to you through my actions each and every day. We have started and finished so many great — and perhaps not so great — adventures, and each one has always made us stronger and surer in the process. You are the reason *why* so many things are the way they are in my life, my head, and my heart. These two words say it all:

"It's you." - Jonathan Richman

Preface

In presenting this thesis, I feel compelled to offer a caveat to the reader. The work contained within these pages are the result of a rigorous exploration into electromagnetic form factor measurements. It is imperative to acknowledge that the nature of this research is primarily exploratory.

Throughout the course of this research, my aim has been to delve as deep as I could into these topics. I have attempted to do so with an open mind while also embracing the uncertainty and complexity inherent in such endeavors. While every effort has been made to ensure the accuracy, reliability, and validity of the data and analysis presented herein, it is essential to recognize the inherent limitations of exploratory research.

The conclusions drawn from these investigations are by no means intended to be definitive or conclusive. Instead, they can hopefully serve as a starting point for further inquiry and discussion within the academic community. It is my hope that this thesis will assist future researchers in building upon the foundations explored here.

Contents

Preface **v**

1 Introduction **1**

1.1 The Structure of the Nucleon 3

1.2 Electromagnetic Form Factors 4

1.3 Probing the Structure of the Nucleon 4

1.4 The Single Photon Exchange Approximation 6

1.5 Extracting G_M^n, The Ratio Method 9

1.6 The Rosenbluth Technique 10

1.7 The Two Photon Exchange Approximation 12

1.8 Previous Measurements 13

1.8.1 Previous G_M^n Measurements 13

1.8.2 Previous nTPE Measurements 15

1.9 The Form Factor Ratio Puzzle 15

2 Experimental Setup **19**

2.1 Overview 19

2.2 Accelerator and Beamline 20

2.2.1 CEBAF 20

2.2.2 Experimental Hall A 21

2.3 Experimental Configuration and Layout 23

2.4 Spectrometers 26

2.4.1 BigBite Spectrometer 26

2.4.2 Super BigBite Spectrometer 28
2.5 Targets 31
2.5.1 Liquid Hydrogen Target 32
2.5.2 Liquid Deuterium Target 33
2.6 Detector Packages 34
2.6.1 BigBite Calorimeter 34
2.6.2 HCal 37
2.6.3 Momentum and Optics From BB and SBS 39
2.6.4 GEM Detectors 44
2.6.5 GRINCH 45
2.6.6 Timing Hodoscope 47
2.7 GEM Data-Acquisition System 49
2.7.1 APV25 Readout Card 50
2.7.2 MPDs 51
2.7.3 VTPs 52
2.7.4 CODA Software Platform 53

3 Gaseous Electron Multipliers **55**
3.1 An Overview of GEM Detectors 56
3.1.1 Anatomy of a GEM Foil 56
3.1.2 Anatomy of a GEM Chamber 58
3.1.3 The GEM Readout Board 59
3.2 Interactions of Charged Particles With Matter 63
3.2.1 Heavy charged Particles 63
3.2.2 Light Charged Particles 66
3.2.3 Neutral Particles in Matter 67
3.2.4 Interactions and Gas Mixtures 68
3.3 Ionization, Gain, and Charge Transfer 70
3.3.1 Electric Fields in GEM Foil Holes 70
3.3.2 Ionization in the Gas Detector 72

3.3.3 Optical and Electrical Transparency . . . 73
3.3.4 Extraction Efficiency . . . 75
3.3.5 Gain . . . 77
3.3.6 Multiple GEM Foils in Succession . . . 78
3.4 Physical Design and Principles . . . 80
3.4.1 GEM Chamber General Design Layout . . . 80
3.4.2 Foil Layers Within a GEM Detector . . . 82
3.4.3 GEM Frames . . . 94
3.4.4 GEM Chamber Gas Distribution . . . 99
3.4.5 Aluminized Polyimide Entrance Window . . . 105
3.4.6 Two Types of UVa GEMs: XY and UV . . . 107
3.5 GEM Detector Fabrication and Testing . . . 111
3.5.1 Frame Preparation . . . 112
3.5.2 Foil Preparation . . . 113
3.5.3 Foil Stretching . . . 113
3.5.4 Assembling the Detector . . . 115
3.5.5 High-Voltage Testing . . . 116
3.5.6 Cosmic Testing . . . 119
3.6 Experimental GEM Setup in Hall A . . . 122
3.6.1 GEMs on the BigBite Stack . . . 124
3.6.2 Gas Distribution System for the GEMs . . . 127
3.6.3 GEM High Voltage Power Supply . . . 128
3.6.4 GEM Low Voltage Power Supply . . . 130

4 Experimental Technique **131**
4.1 Experimental Kinematics and Configurations . . . 131
4.2 Particle Deflection on the Scattering Arms . . . 134
4.3 HCal Observables: dx, dy, and dxdy . . . 135
4.3.1 HCal Coordinates . . . 135
4.3.2 The dxdy Plot . . . 137

4.4 GMn and nTPE Simulations . . . 141
4.4.1 g4sbs and SIMC . . . 142
4.4.2 Simulation dx histograms and plots . . . 146
4.4.3 Simulation Event Weighting . . . 147
4.4.4 Scaling MC to Data . . . 150
4.5 Extracting GMn Experimentally . . . 152
4.6 Extracting nTPE Experimentally . . . 157
4.7 A Note on World Data and Parameterizations . . . 162

5 Exploratory Analysis 165
5.1 Analysis Software . . . 165
5.1.1 SBS Offline and SBS Replay . . . 166
5.1.2 ROOT . . . 167
5.2 Initial Calibrations and Replay . . . 168
5.3 Elastic Event Selection and Cuts . . . 169
5.3.1 BBCal PreShower (Pion) Cut . . . 171
5.3.2 Invariant Mass Cuts . . . 172
5.3.3 Timing Cut . . . 175
5.3.4 HCal Cluster Selection . . . 175
5.3.5 HCal Acceptance Cut . . . 178
5.3.6 dy cuts . . . 179
5.3.7 Fiducial Cut . . . 180
5.3.8 Background Subtraction . . . 185
5.4 Fitting Methods and Techniques . . . 193
5.5 HCal Detection Efficiency . . . 194
5.6 HCal Inefficiencies and Data Issues . . . 201
5.6.1 Data Quality for SBS9 . . . 201
5.6.2 Corrections for SBS9 Data . . . 204

6 Exploratory Experimental Results 211
6.1 Exploratory GMn Projections . . . 211

6.1.1 GMn Observables 211
6.1.2 Exploratory GMn Results At A Glance 212
6.1.3 A Note on World Data Parameterizations 214
6.2 Exploratory nTPE Projections 215
6.2.1 nTPE Observables 215
6.2.2 Exploratory nTPE Results At A Glance 215
6.2.3 A note on nTPE Error 218
6.3 Error Analysis 219
6.3.1 GMn Statistical Uncertainties 220
6.3.2 GMn Systematic Uncertainties 222
6.3.3 nTPE Statistical Uncertainties 227
6.3.4 nTPE Systematic Uncertainties 227

7 Conclusion **229**

7.0.1 Summary of GMn Analysis and Findings 229
7.0.2 Summary of nTPE Analysis and Findings 231
7.0.3 Future Work 235
7.0.4 Final Remarks 239

Bibliography **252**

List of Figures

1-1 Standard Model of Particle Physics 2

1-2 The nature of interactions between the virtual photons exchanged during electron-nucleon scattering at various momentum-transfer ranges . 5

1-3 Feynman diagram for elastic electron-nucleon scattering in the single photon exchange, or Born, approximation 6

1-4 A plot of neutron reduced cross-section from parameterization vs. epsilon 10

1-5 Feynman diagrams for two photon exchange 12

1-6 The form factor ratio puzzle 16

1-7 Plot of the proton reduced cross-section vs. epsilon for LT Separation and polarization transfer methods 18

2-1 Continuous Electron Beam Accelerator Facility with the 12 GeV upgrade 21

2-2 Aerial photograph of JLab's accelerator campus 22

2-3 Panoramic photo of the inside of Hall A 23

2-4 Drawing of the experimental layout of the first run of the SBS Program 24

2-5 Experimental setup of GMn and nTPE in Hall A 25

2-6 The BigBite Spectrometer 27

2-7 The BigBite Magnet 28

2-8 The Super BigBite Spectrometer 29

2-9 The SBS Dipole Magnet 30

2-10 Hall A cryotargets 31

2-11 Anatomy of the GMn and nTPE target loop 32

2-12 LH2 and LD2 dx plots 33

2-13 BBCal on the BigBite Arm 34
2-14 Drawing of the to sub-calorimeters that make up the BigBite Calorimeter: the PreShower and the Shower calorimeters 36
2-15 Calibration of BBCal for the energy over momentum plot of the scattered electron 37
2-16 The Hadron Arm 38
2-17 The module layout and configuration of the Hadron Calorimeter . . . 39
2-18 The design and configuration of the individual blocks of HCal 40
2-19 Upstream view of the SBS magnet, exit beamline, scattering chamber, and BigBite Spectrometer 41
2-20 The BigBite magnet deflects a charged particle (scattered electron) as it passes through the yoke 42
2-21 Sieve plate for optics and momentum calibrations 43
2-22 Sieve plate dimensional drawing and corresponding GEM hitmap . . 44
2-23 Trajectory reconstruction using layers of GEM detectors 45
2-24 The GRINCH detector on the BigBite detector stack 46
2-25 The GRINCH is filled with a heavy gas which allows for Cherenkov radiation for electrons but not for pions 47
2-26 The BigBite Timing Hodoscope is on the BigBite Spectrometer between the PreShower and Shower layers of BBCal 48
2-27 A typical GEM Data-acquisition system 49
2-28 A bank of four APVs as used on the UVa GEMs 50
2-29 JLab's Multi-Purpose Digitizers (MPDs) 51
2-30 An array of VXS Trigger Processors (VTPs) 52
2-31 Screenshot of the CODA computer interface for data-taking 53
2-32 Schematic of the basic flow scheme for a simple CODA implementation 53

3-1 A single ionizing particle develops an avalanche through the layers of a Triple-GEM and the amplified signal reaches the readout plane . . . 56
3-2 GEM foil geometry 57

3-3 "Single GEM" detector 58
3-4 Configuration of layers for multi-layer GEMS 60
3-5 GEM readout planes 61
3-6 Two different readout plane configurations 61
3-7 Readout from an APV showing the ADC signals read out from various strips on a GEM detector 62
3-8 Plot of Bethe-Bloch formula for positive muons in copper as a function of muon momentum, represented here by $\beta\gamma$ 66
3-9 Computed electric fields within the perforated regions of a GEM foil . 71
3-10 Electric field computed along a line through the center of the holes, for different hole diameters 72
3-11 Efficiency vs. Applied GEM voltage plot 77
3-12 Drawings of the three GEM gain factors 78
3-13 Triple GEMs 79
3-14 Exploded view of an SBS UV GEM chamber 81
3-15 A typical voltage distribution across the foils of a GEM detector with an overall voltage setting of 4.1 kV 83
3-16 UV Drift Cathode foil photo and schematic 83
3-17 UV GEM foil photo and schematic 85
3-18 Close-up of three independent UV GEM foil sectors with their copper traces 86
3-19 Close-up of the first UV GEM foil sector with its trace to the connection array 87
3-20 All 120 connector pads and final traces for the top and bottom sectors of the UV GEM foil 88
3-21 Two different GEM foil electrode sector layouts 89
3-22 GEM hit map showing two chambers that each have a single non-functioning sector 91
3-23 UV Readout Foil photo and schematic 92
3-24 UV APV25 connections and cards 94

3-25 High-voltage dividers and protective resistor boards 95
3-26 Triple GEM internal structure and layout 96
3-27 GEM frames 98
3-28 Cut-away drawing of gas flow through the UV GEMs 100
3-29 Exploded layer view of UV gas flow simulation 102
3-30 Top view of the gas flow simulation on each layer 103
3-31 Effects of different frame designs on gas flow and circulation 104
3-32 GEM chamber with non-aluminized polyimide window that has collapsed 106
3-33 Drawing of a non-aluminized polyimide Entrance Window and how it can collapse onto the Drift Cathode foil below it 107
3-34 Drawing of the aluminized polyimide Entrance Window and how it mitigates collapse 107
3-35 An XY GEM layer 108
3-36 Hitmaps on an XY GEM layer 109
3-37 Frame layout and design of a typical XY GEM detector 109
3-38 A UV GEM with shielding installed, ready for installation in Experimental Hall A 110
3-39 GEM frames hanging to dry in the clean room 112
3-40 Using a static-cling type duster on a stretched Entrance Window . . . 113
3-41 GEM foil stretching table that was custom-built for the UV foils . . . 114
3-42 Clamps of the stretching table that are fitted with force measuring load cells 115
3-43 Multiple framed foils stack on a UV GEM readout assembly 116
3-44 Testing a raw UV foil in the nitrogen-filled high-voltage bo 118
3-45 Connections to the GEM foil in the dry nitrogen HV test box 119
3-46 UVa UV GEM cosmic test setup 121
3-47 The first UV GEM cosmic measurements 122
3-48 Experimental setup for GMn and nTPE with the three primary arms of the experimental highlighted 123

3-49 The BigBite detector stack and the location of the Front and Rear GEM trackers 124
3-50 GEM layer positions on the BigBite stack 126
3-51 The five GEM layers of the BigBite stack 126
3-52 Visualization of track reconstruction across the multiple GEM layers of the BigBite detector stack 127
3-53 Schematic of the gas distribution system for GEM detectors on the BigBite stack 128
3-54 Schematic of the gas manifold for routing the Ar/CO2 mixture to each GEM at its required flow rate 129

4-1 Downstream regions of the experimental setup 134
4-2 A basic HCal dxdy plot 135
4-3 Photo of the SBS HCal in Experimental Hall A 136
4-4 Measured hits on HCal 137
4-5 Basic e-N scattering kinematics and projections of expected hadron positions on HCal 138
4-6 A typical dx vs. dy, or dxdy, HCal plot 139
4-7 Projecting a 2D dxdy plot onto the y-axis and retrieving a dx plot . . 140
4-8 X and Y axes projections of an HCal dxdy plot 141
4-9 Comparisons of data and simulations 142
4-10 The GMn and nTPE experimental layout in the g4sbs software platform 146
4-11 A dx plot for simulated data 147
4-12 Unweighted MC dx plot 149
4-13 Comparisons between pre-scaled and post-scaled MC dx peaks with experimental data 151
4-14 LD2 dx plot with labelled nucleon peaks and background approximation 153
4-15 Neutron reduced cross-section from parameterization vs. virtual photon polarization, ϵ 159
4-16 Global data and their parametric fits by Kelly for G_E^p, G_M^p, G_E^n, and G_M^n 162

5-1 The evolution of the SBS8 dx plots as more cuts are applied 170
5-2 BigBite Calorimeter PreShower cluster energy without and with the pion-rejection cut. 172
5-3 Elastic and inelastic scattering of an electron and target proton . . . 173
5-4 Invariant mass, W^2, for SBS4 with an SBS field scale of 30% and an LD2 target 174
5-5 ADC time for highest energy block in largest HCal cluster 175
5-6 Simulated scattering event and cluster on HCal 176
5-7 Comparison of cluster selection methods for LH2 177
5-8 The HCal acceptance cut 178
5-9 The dy plot without and with cuts 179
5-10 Comparison of a dxdy plot with and without dy cuts 180
5-11 The fiducial cut 181
5-12 Comparisons between W^2 with and without a tight, sharp cut 185
5-13 A total dx fit comprised of three functions 186
5-14 Initial total dx fit with a proton Gaussian, neutron Gaussian, and a third order polynomial for the background 188
5-15 A dxdy plot with a 3-sigma dy anti-cut 189
5-16 Inelastic background with polynomial fit 189
5-17 Comparison of total dx fits with and without background from a dy anti-cut projection 191
5-18 Comparison of data and simulation before and after background subtraction 191
5-19 Comparisons of data and simulation after background subtraction . . 192
5-20 Four MC HCal momentum bin selections 195
5-21 Monte Carlo HCal proton and neutron detection efficiencies 196
5-22 The various plots used to calculate the experimental HCal proton detection efficiency for SBS4 200
5-23 SBS9 dx p8lot highlighting the effects of HCal inefficiencies on the data and resulting extractions 202

5-24 HCal inefficiencies for SBS9 204
5-25 Map of HCal blocks and their operational status 205
5-26 SBS9 data-based HCal expected x and y efficiency map 206
5-27 SBS9 simulated HCal hits with and without data-based HCal efficiency map 207
5-28 SBS9 dx plot after applying HCal efficiency map corrections 208

6-1 SBS4, SBS8, and SBS9 dx plots with Q^2, Scale Factor Ratio, and G_M^n values 213
6-2 Experimental neutron Reduced C.S. vs. epsilon for nTPE 216
6-3 A dx plot made from the Y-Projection of a W^2 slice as part of a systematic study 224

7-1 Experimental G_M^n values overlaid with the Kelly parameterization . . 231
7-2 Experimental G_M^n values overlaid with world G_M^n data 232
7-3 Experimental G_M^n values overlaid with world G_M^n data in the lower Q^2 range (0 - 5 $(\mathrm{GeV/c})^2$) 232
7-4 Experimental Rosenbluth Slope vs. Q^2 for SBS8 and SBS9 234
7-5 World data and parameterizations for G_M^n with experimental values for SBS4, SBS8, and SBS9 237
7-6 Neutron reduced cross-section from parameterization vs. epsilon with nominal values of SBS8 and SBS9 238

List of Tables

1.1 A table of neutron form factor experiments 14

4.1 Nominal experimental configurations for the GMn and nTPE experiments 133

4.2 Experimental measurements for each kinematic configuration of the GMn and nTPE experiments 133

6.1 Summary of Q^2, $F_{n/p}$, and G_M^n for SBS4, SBS8, and SBS9 212

6.2 Experimental nTPE values for Q^2, τ, and ϵ 216

6.3 Normalized Rosenbluth Slope values extracted from Figure 6-2 217

6.4 Statistical errors associated with $F_{n/p}$, $R_{n/p}$, and G_M^n 222

6.5 Systematic uncertainty contributions to $R_{n/p}$, $F_{n/p}$, and G_M^n from elastic cuts for SBS4, SBS8, and SBS9 226

7.1 Summary of experimental G_M^n extractions for SBS4, SBS8, and SBS9 230

7.2 Summary of Kelly G_M^n parameterizations for SBS4, SBS8, and SBS9 . 230

Chapter 1

Introduction

Nucleons are the pillars on which the nucleus is built and make up most of the matter we see around us. The characterization of nucleons has been a fundamental challenge for experimental and theoretical physicists since their discovery in 1933; Rutherford discovered the nucleus by scattering alpha particles from gold foil [111]. This led to the conclusion that the atom contained a very small, dense, and positively charged nucleus at its core. This nucleus is made up by nucleons (protons and neutrons). Protons and neutrons were once themselves considered fundamental particles. Further experimentation would prove this concept to be incorrect. In 1933, Otto Stern measured the magnetic moment of the proton and the deuteron, and showed that his results were not directly aligned with what would be expected for point-like spin-1/2 particles [52]. This implied that previous notions of a "point" proton or neutron were incorrect. It is now well-accepted that protons and neutrons are not fundamental particles, but rather, composite in form, with internal structure and constituents.

The Standard Model explains the interplay of the electromagnetic, weak, and strong forces, and describes the twelve fundamental building blocks of all matter. Figure 1-1 shows how the fundamental particles in the Standard Model are organized into fermions (matter) on the outer ring and bosons (force carriers) on the inner rings. Fermions can be either quarks or leptons, and the bosons are either gauge bosons or the Higgs boson. The quarks and leptons each have three generations which

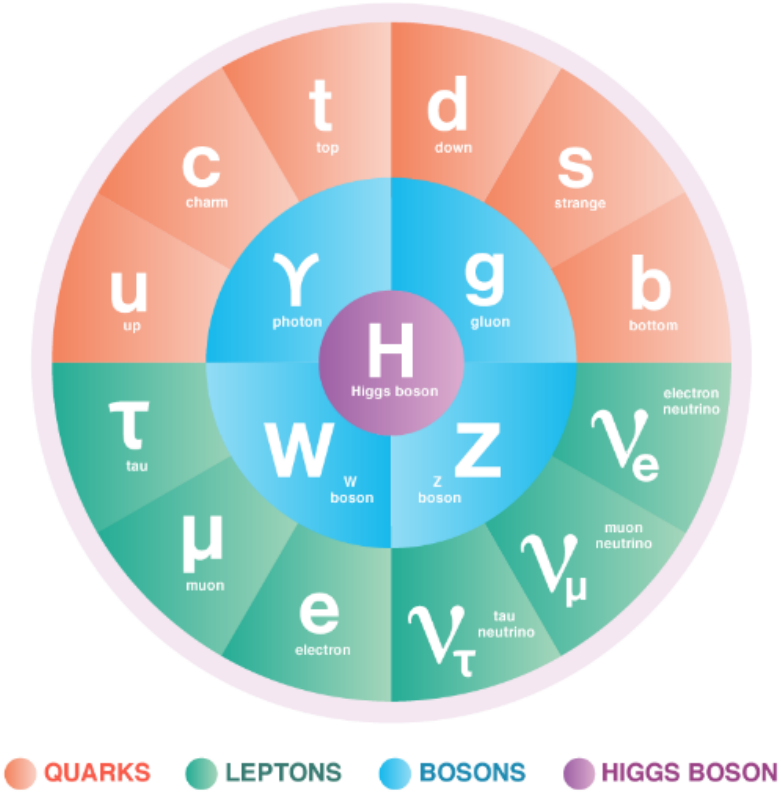


Figure 1-1: Standard Model of Particle Physics. Image source: [87].

are governed by the strong, weak, electromagnetic, and gravitational interactions. Each of these interactions, apart from the gravitational, are mediated by different force carriers (bosons). The gluon is the mediator of the strong interaction, which binds protons and neutrons together. The weak interaction that is responsible for some nuclear processes and β-decays, is mediated by the W and Z bosons. The electromagnetic interaction that binds electrons to the nucleus through electrostatic attraction, is mediated by the photon.

In the Standard Model, protons and neutrons are each made up of three elementary fermions called quarks. Quarks are bound together by the strong force acting between their inherent color charges. There are six flavors of quarks: up (u), down (d), charm (c), strange (s), top (t), and bottom (b). Quarks have one of two electric charges: $Q = +\frac{2}{3}e$ or $Q = -\frac{1}{3}e$ and in general, the masses of the quarks increase from the first generation to the third.

1.1 The Structure of the Nucleon

The magnetic moment of a point-like spin $^1/_2$ particle is given by:

$$\mu = \frac{ge'}{2e}\mu_N \tag{1.1}$$

where g is the Landé g-factor, e' is the nucleon's charge, e is the elementary charge of the electron, and μ_N is the nuclear magneton. The nuclear magneton is defined in terms of Planck's constant, $\hbar$, the electron charge e, and the nucleon mass, m:

$$\mu_N = \frac{e\hbar}{2m}. \tag{1.2}$$

If the proton and neutron were point-like Dirac particles, we would expect to find:

$$\mu_p = \mu_N \quad \text{and} \quad \mu_n = 0 \tag{1.3}$$

where p and n denote the proton and neutron, respectively. In 1933, Otto Stern showed that this was not the case when they measured [52]:

$$\mu_p = 2.79\mu_N, \quad \mu_n = -1.91\mu_N, \quad \text{and} \quad \frac{\mu_n}{\mu_p} \approx -0.685. \tag{1.4}$$

This suggested that previous notions were incorrect and that nucleons are indeed not point-like structures. In the 1950's, Hofstadter related the ep cross-section in elastic scattering at a given angle and energy to the Mott cross-section [31]. The Mott cross-section describes scattering of a high energy electron beam from a point-like object and is defined as [133]:

$$\sigma_{Mott} = \frac{\alpha^2}{4E_e^2 sin^4(\frac{\theta_e}{2})}\frac{E'_e}{E_e}\cos^2(\frac{\theta_e}{2}). \tag{1.5}$$

In this equation, α is the fine-structure constant, E_e is the electron beam energy, and E'_e and θ_e are the scattered electron's energy and angle, respectively. Hofstadter found that the cross-section was not just the Mott cross-section, but rather, the Mott cross-section multiplied by some *correction factor*. Hofstadter established that the

proton has an extended charge distribution and subsequently measured its size [31].

1.2 Electromagnetic Form Factors

If we consider the finite extent of the proton's charge distribution, we can introduce a *form factor* which accounts for phase differences in the scattered wave across the charge distribution [133]. These form factors are now called the Sach's form factors and are denoted: G_E (electric form factor) and G_M (magnetic form factor). These form factors describe the spatial distributions of electric charge and magnetization. They are both functions of momentum-transfer Q^2 ($G_E(Q^2)$ and $G_M(Q^2)$). These form factors are extremely important observables for the characterization of nucleons. From $G_E(Q^2 = 0)$ and $G_M(Q^2 = 0)$ we can obtain the electric charge and magnetic moments, respectively. Also, from the slopes at $G_E(Q^2 = 0)$ and $G_M(Q^2 = 0)$, we can obtain the electric and magnetic radii.

1.3 Probing the Structure of the Nucleon

In modern particle physics, the exchange of force-carrying gauge bosons mediates the interactions between particles. In Quantum Electrodynamics (QED), the gauge boson between charged particles is the *virtual photon*. In QED, a *virtual particle* represents summing over all time-ordered Feynman diagrams and if applicable, all polarization states of the exchanged particle [133]. The implementation of a virtual photon in QED allows us to describe the exchange of forces between charged particles in terms of these virtual photons instead of some "spooky action at a distance".

In electron-nucleon scattering, virtual photons are exchanged between the electron and the nucleon. The nature of the interaction between the virtual photon and the nucleon depends strongly on its wavelength. Figure 1-2 shows four Q^2 ranges and depicts how an incoming electron interacts with a nucleon target. For very low Q^2, we consider the limit as $Q^2 \to 0$. In this case, the virtual photon has a relatively large wavelength and because of this, it is difficult to detect any internal structure.

At this energy range, the scattering is equivalent to that from a point-like, spin-less object.

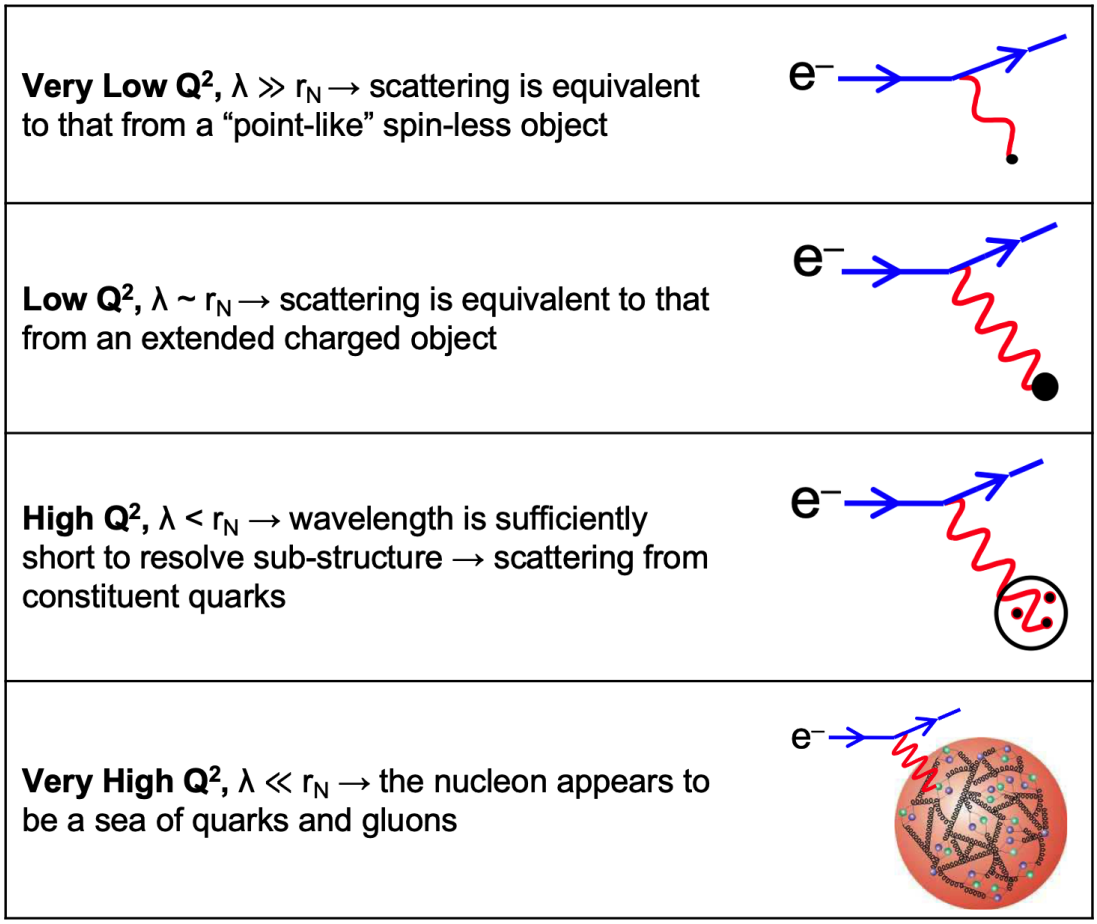


Figure 1-2: The nature of interactions between the virtual photons exchanged during electron-nucleon scattering at various momentum-transfer ranges. As Q^2 increases, the wavelength of the virtual photon decreases and can probe deeper and deeper until it interacts with the quarks and gluons within. Plot reproduced from [132].

As Q^2 increases to the range where the virtual photon's wavelength is on the order of the nucleon radius, the proton and neutron recoil is negligible, and the scattering is equivalent to that from an extended charged object. As we get to relatively higher ranges of Q^2, we move into the regime where the wavelength is now shorter than the nucleon radius. At this point, the wavelength is short enough to start resolving the sub-structure due to scattering from constituent quarks. In the upper limits of Q^2, the wavelength is much smaller than the radius of the nucleon. Here, the nucleon appears as a sea of quarks and gluons. The GMn and nTPE experiments operate

within the third regime shown, where scattering is from the constituent quarks and we are able to resolve the sub-structure.

1.4 The Single Photon Exchange Approximation

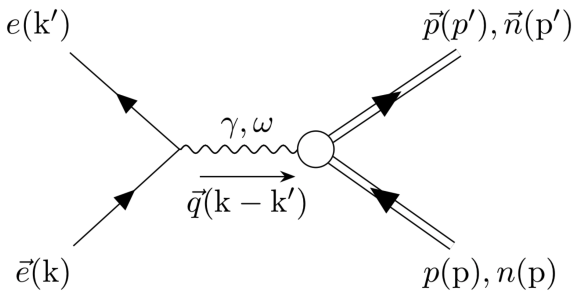


Figure 1-3: Feynman diagram for elastic electron-nucleon scattering in the single photon exchange, or Born, approximation.

Figure 1-3 shows the Feynman diagram for the lowest order approximation for elastic electron-nucleon scattering. This is the *single photon exchange approximation*, or the *Born term*. In Figure 1-3 we have an electron, e, with initial and final momenta, k and k', scattering from a nucleon with initial (final) momentum p (p'). The q-vector is the four-momentum transferred via the virtual photon in the reaction ($q = k - k'$). Following the Feynman rules for *quantum electrodynamics* (QED), we can obtain the amplitude for electron-nucleon scattering from the product of the leptonic and hadronic currents:

$$i\mathcal{M} = \frac{-i}{q_\mu^2} \ell_\mu \mathcal{J}^\mu. \tag{1.6}$$

For the leptonic current, ℓ_μ, we are simply referring to the electron. As of yet, the electron is still considered a fundamental, point-like, particle with no constituent components. We therefore know its form exactly within the standard model. For the hadrons, however, there is still much to uncover and learn as we actively probe their structures in hopes of filling in the blanks. As this thesis is dedicated to the structure of the neutron, we can not get off so easy as we did with the electron. Therefore, we

must take this as our entry point for exploring the nucleon structure. The hadronic current, written in terms of the Dirac (F_1) and Pauli (F_2) form factors is:

$$\mathcal{J}^{\mu} = i\bar{e}\nu(p')[\gamma^{\mu}F_1(Q^2) + \frac{i\sigma^{\mu\nu}q_{\nu}}{2M}\kappa_j F_2(Q^2)]\nu(p). \tag{1.7}$$

In Equation 1.7, M is the mass of the nucleon and κ is the nucleon anomalous magnetic moment. This formulation requires Lorentz invariance, current conservation, and parity conservation [133]. The Dirac and Pauli form factors describe the spatial distribution of electric charge and current inside the nucleon.

In the lab frame the cross-section is:

$$\left(\frac{d\sigma}{d\Omega}\right)_0 = \frac{|\overline{\mathcal{M}}|^2}{64\pi^2}\left(\frac{E_e'}{E_e}\right)^2\frac{1}{M} \tag{1.8}$$

where the 0-subscript denotes this as the one-photon approximation cross-section and

$$|\overline{\mathcal{M}}|^2 = \frac{1}{(Q^2)^2}|\overline{\ell\cdot\mathcal{J}}|^2. \tag{1.9}$$

Equation 1.8 can be written in terms of the Mott cross-section and the Dirac and Pauli form factors as:

$$\left(\frac{d\sigma}{d\Omega}\right)_0 = \left(\frac{d\sigma}{d\Omega}\right)_{Mott}\frac{E_e}{E_{Beam}}\left\{F_1^2 + \tau\left[F_2^2 + 2(F_1+F_2)^2\tan^2(\frac{\theta_e}{2})\right]\right\} \tag{1.10}$$

where the Mott cross-section is defined in Eq. 1.5 and $\tau = {}^{Q^2}/_{4M^2}$.

The experimental cross-section is more easily analyzed in terms of *Sach's form factors*. These are G_E for the electric form factor, and G_M for the magnetic form factor. The Sach's form factors are defined as linear combinations of F_1 and F_2:

$$G_E = F_1 - \tau F_2 \qquad\qquad G_M = F_1 + F_2. \tag{1.11}$$

The scattering cross-section (Eq. 1.10) can be written in a simple form in terms of the Sach's form factors (Eq. 1.11):

$$\left(\frac{d\sigma}{d\Omega}\right)_0 = \sigma_{Mott}\left(\frac{1}{1+\tau}\right)(G_E^2 + \frac{\tau}{\epsilon}G_M^2). \tag{1.12}$$

In Eq. 1.12 we have introduced ϵ, the polarization of the virtual photon[1]:

$$\epsilon = \left[1 + 2(1+\tau)\tan^2(\frac{\theta_e}{2})\right]^{-1}. \tag{1.13}$$

Early form factor measurements showed that G_E and G_M appeared to be well-parameterized by the dipole form factor [21]

$$G_D = \left(1 + \frac{Q^2}{\Delta^2}\right)^2 \tag{1.14}$$

where $\Delta^2 = 0.71$ (GeV)^2. Three out of four of the form factors still seem reasonably well-modeled by this parameterization: G_E^p, $\frac{G_M^p}{\mu_p}$, and $\frac{G_M^n}{\mu_n}$ [22]. Notice that for the magnetic form factors we have incorporated the magnetic moment for each nucleon ($\mu_p \approx 2.79\mu_N$ and $\mu_n \approx -1.91\mu_N$). The electric form factor of the neutron, G_E^n, only fits this parameterization up to a certain Q^2 point. Below $Q^2 = 1$ GeV^2 it is well-fit by the dipole form factor however, at higher Q^2, G_E^n is dampened by the $1/\tau$ term [3, 21]. From Eq. 1.12 for the neutron, it is clear that as Q^2 increases, the term becomes dominated by G_M^n. To a certain extent, a better fitting parameterization for G_E^n is the Galster fit. It has the general form:

$$G_E^n(Q^2)_{\text{Galster}} = \frac{A\tau}{(1+B\tau)}G_D(Q^2) \tag{1.15}$$

where A and B are fitted parameters [61]. There are many parameterizations which are "Galster-like" and in the same form as the generalized Galster formula shown in Eq. 1.15 [3, 21]. The analysis described in this thesis will reference multiple variations of the Galster fit for G_E^n. Depending on the particular situation at hand, different versions of the parameterization are used across various scripts, simulations, and calculations.

[1] ϵ, or rather, $1/\epsilon$, by definition must always be greater than unity. In effect, it describes how much the interaction probes the magnetic structure of the nucleon relative to the electric structure.

1.5 Extracting G_M^n, The Ratio Method

The form factor G_M^n will be extracted via the so-called *Ratio Method* (or Durand Technique). This method requires taking simultaneous yield measurements of quasi-elastically scattered neutrons and protons from our deuterium target. From those two quantities, we can form the *Elastic Scattering Ratio*, R:

$$R = \frac{\#\text{ of neutrons}}{\#\text{ of protons}} = \frac{\left(\frac{d\sigma}{d\Omega}\right)_{n(e,e')}}{\left(\frac{d\sigma}{d\Omega}\right)_{p(e,e')}}. \tag{1.16}$$

We then determine the differential cross-section for e-N elastic scattering via the Born term:

$$\frac{d\sigma}{d\Omega_e} = \left(\frac{d\sigma}{d\Omega}\right)_{Mott}\left(\frac{1}{1+\tau}\right)\left(G_E^2 + \frac{\tau}{\epsilon}G_M^2\right). \tag{1.17}$$

Equation 1.17 can then be substituted into Equation 1.16. Note that subscripts for the proton (p) and neutron (n) will now be introduced where applicable.

$$R = \frac{\left(\frac{d\sigma}{d\Omega}\right)_{Mott}\left(\frac{1}{1+\tau}\right)\left((G_E^n)^2 + \frac{\tau}{\epsilon}(G_M^n)^2\right)}{\left(\frac{d\sigma}{d\Omega}\right)_{p(e,e')}}. \tag{1.18}$$

In Eq. 1.18, we have the term G_M^n. This is precisely the value we are looking for. With a little bit of algebra we can re-arrange variables and solve for G_M^n directly:

$$\boxed{(G_M^n)^2 = \frac{\epsilon}{\tau}\left(R\left(\frac{d\sigma}{d\Omega}\right)_{p(e,e')}\left(\frac{1+\tau}{\left(\frac{d\sigma}{d\Omega}\right)_{Mott}}\right)\right) - (G_E^n)^2.} \tag{1.19}$$

The ratio method is an effective technique that offers many advantages. For one, in forming the ratio, R (Eq. 1.16), many systematic uncertainties are cancelled. Also, this technique is not sensitive to parameters such as electron acceptance, reconstruction efficiency, luminosity, or trigger efficiency [80]. The ratio method requires accurate knowledge of the nucleon detection efficiencies. These detection efficiencies will be used to *correct* the experimental quasi-elastic yields. Many experiments have successfully implemented the ratio method in extracting G_M^n (see Table 1.1). It is an

effective technique for measuring the neutron form factor even though there are no free neutron targets.

1.6 The Rosenbluth Technique

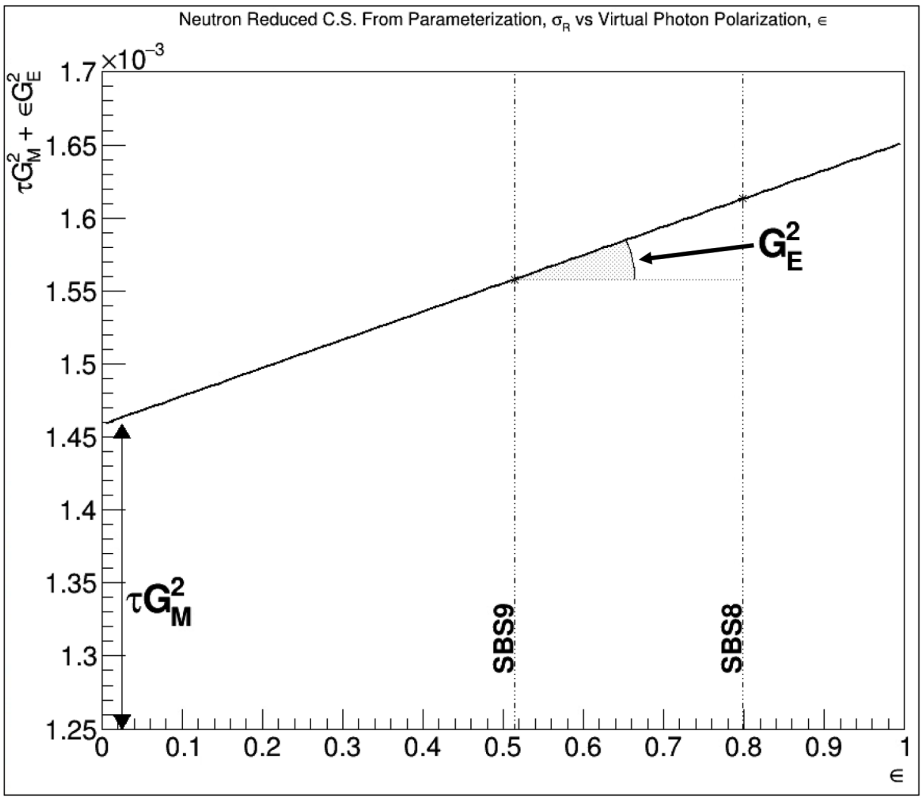


Figure 1-4: A plot of neutron reduced cross-section from parameterization vs. epsilon. The reduced cross-section of the neutron is calculated at the nominal epsilon points for SBS8 and SBS9 using the Ye parameterization ([140]). For the two epsilon points shown, a linear extrapolation can be made. The y-intercept yields τG_M^2 and the slope provides G_E^2.

The extraction of nTPE is in a way, an *extension* of the GMn experiment. From two related G_M^n measurements (SBS8 and SBS9) we can form the so-called *Rosenbluth Slope* via a technique called *Rosenbluth Separation*, or *LT Separation*. Rosenbluth Separation is used to extract the *Form Factor Ratio* (*FFR*), G_E/G_M. The technique exploits the linear-dependence in epsilon of the *reduced cross-section*:

$$\sigma_R = \epsilon G_E^2 + \tau G_M^2. \tag{1.20}$$

Figure 1-4 shows a plot of the neutron reduced cross-section versus epsilon. The neutron reduced cross-section is calculated using Ye's parameterization ([140]) at the nominal epsilon values for SBS8 and SBS9. Two epsilon points (both at the same Q^2 value) are shown ($\epsilon_{\text{SBS8}} \approx 0.79$ and $\epsilon_{\text{SBS9}} \approx 0.51$) along with a linear extrapolation between them. We can extract some valuable information from the reduced cross-section if we operate on it a bit. If we look at σ_R for epsilon $= 0$ we find:

$$\sigma_R(\epsilon = 0) = \tau G_M^2. \tag{1.21}$$

This is the transverse component of the cross-section, σ_T. Also, if we take the derivative of σ_R with respect to epsilon we get,

$$\frac{d\sigma_R}{d\epsilon} = G_E^2, \tag{1.22}$$

which is just the longitudinal component of the cross-section, σ_L. We therefore can summarily define the longitudinal and transverse components of the cross-section:

$$\text{Longitudinal:} \quad \sigma_L = G_E^2 \qquad \text{Transverse:} \quad \sigma_T = \tau G_M^2. \tag{1.23}$$

Substituting σ_L and σ_T from Eq. 1.23 into Eq. 1.20 provides:

$$\sigma_R = \quad \epsilon G_E^2 + \tau G_M^2 \quad = \quad \epsilon \sigma_L + \sigma_T. \tag{1.24}$$

Now, referring back to Figure 1-4 we can see that from the y-intercept of the linear extrapolation we directly get τG_M^2 and from the slope we find G_E^2. From these two values we can then combine them to determine the Rosenbluth slope:

$$\text{Rosenbluth Slope: RS or } S = \frac{G_E^2}{\tau G_M^2} = \frac{\sigma_L}{\sigma_T}. \tag{1.25}$$

We can also re-write Equation 1.12 in terms of the reduced cross-section, σ_R, as

$$\left(\frac{d\sigma}{d\Omega}\right)_0 = \sigma_{Mott}\frac{\sigma_R}{\epsilon(1+\tau)}. \tag{1.26}$$

1.7 The Two Photon Exchange Approximation

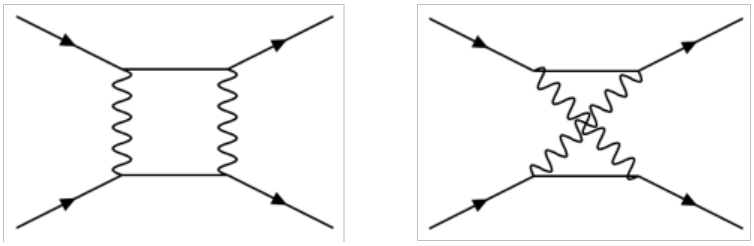


Figure 1-5: Feynman diagrams for two photon exchange. These are second order diagrams which account for higher order corrections arising from the consideration of two virtual photons.

Up until this point, we have been discussing the single photon approximation, or Born term. In that approximation, higher order considerations are truncated. This approximation therefore fails to account for radiative corrections to the cross-section, one-loop virtual corrections (vacuum polarization, $e-\gamma$ and $p-\gamma$ corrections, etc.), and inelastic brehmsstrahlung for real photon emission [141, 19, 4]. Figure 1-5 shows the higher order Feynman diagrams for the two photon exchange in elastic electron-nucleon scattering. We follow a similar formalism as discussed in Section 1.4, but now include the higher order corrections. In doing so, we find that the interference terms between the one and two photon exchange amplitudes result in a relative TPE correction to σ_R (to order α) and can be written as:

$$\delta_{\text{TPE}} = \frac{2Re\{\mathcal{M}^{\dagger}\mathcal{M}^{2\gamma}\}}{|\mathcal{M}|^2}. \tag{1.27}$$

Here, $\mathcal{M}$ is the one-photon exchange (OPE) amplitude and 2γ denotes the two-photon exchange(TPE) amplitude. These higher order TPE corrections effectively modify the cross-section to order α [141, 19, 4],

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega}\right)_0 d\Omega(1+\delta_{\text{TPE}}) \tag{1.28}$$

where δ_{TPE} is the correction factor for the two photon exchange contribution,

$$\delta_{\text{TPE}} = \delta_{\text{virtual}} + \delta_{\text{brem}} \tag{1.29}$$

and $(d\sigma/d\Omega)_0$ is the Born term from Eq. 1.12. In Equation 1.29, $\delta_{virtual}$ includes the contributions from the second virtual photon ("hard" and "soft" components) and δ_{brem} has the contributions from inelastic bremsstrahlung processes.

We can now re-formulate Eq. 1.26 such that the two photon exchange contribution can be attributed as a modification in order α to the cross-section:

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega}\right)_0 d\Omega(1 + \delta_{\text{TPE}}) \tag{1.30}$$

where $(d\sigma/d\Omega)_0$ is the single photon approximation Born term and δ is the measured TPE correction (Eq. 1.29). We can now define this in terms of the reduced cross-sections,

$$\delta_{\text{TPE}} = \frac{\sigma_{R,\text{ TPE}}}{\sigma_{R,\text{ OPE}}} - 1. \tag{1.31}$$

This relation and subsequent extraction is discussed further in Chapter 4.

1.8 Previous Measurements

1.8.1 Previous G_M^n Measurements

The SBS G_M^n experiment was the first of its kind to make measurements at Q^2 values greater than 4.5 $(\text{GeV/c})^2$ [107]. It is exciting to be able to push the momentum transfer bounds of these equations further and further as technology evolves. None of this would be possible without the many ground-breaking experiments that paved the way for G_M^n measurements over the past several decades. Table 1.1 showcases electric and magnetic form factor experiments over the last 60 years. Given that G_M^n is a measurement on the neutron, there is a common difficulty that must be addressed: there are no free neutron targets. Therefore, the majority of previous experiments have employed a deuterium target. For each entry we find the form

factor(s) measured, the type of extraction, the momentum-transfer Q^2 range, and a reference to the primary document.

Form Factor	**Type**	**Q^2 (GeV/c)^2**	**Reference**
G_M^n	Elastic	0.116 - 0.195	J.I. Friedman et al. PR 120, 992 (1960)
G_E^n	Elastic	0.012 - 0.085	Drickey and Hand, PRL 9, 1774 (1962)
G_E^n, G_M^n	Elastic	0.116 - 0.195	P. Benaksas et al., PRL 13, 1774 (1964))
G_E^n, G_E^n	Quasi-elastic	0.04 - 1.2	E.B. Hughes et al. PR 139, B458 (1965)
G_E^n,G_M^n	Quasi-elastic	0.11 - 0.16	B. Grossette et al., PR 141, 1435 (1966)
G_E^n, G_M^n	Quasi-elastic	0.006 - 0.3	E.B. Hughes et al. PR 146, 973 (1966)
G_E^n, G_M^n	Ratio	0.22 - 0.58	P. Stein et al. PRl 16, 592 (1966)
G_M^n	Quasi-elastic	0.04 - 0.2	D. Braess et al. Zeit Phys. 198, 527 (1967)
G_E^n, G_E^n	Anti-coincidence	0.27 - 4.47	R.J. Budnitz et al. PR 173, 1357 (1968)
G_E^n, G_E^n	Ratio	0.39 - 0.78	W. Bartel et al. PL 30B, 285 (1969)
G_E^n	Elastic	0.004 - 0.032	F.A. Bumiller et al., PRL 25, 1774 (1970)
G_E^n	Elastic	0.2 - 0.56	S. Galster et al. NP B32, 221 (1971)
G_E^n, G_M^n	Ratio	1.0 - 1.53	W. Bartel et al. PL 39B, 407(1972)
G_E^n, G_M^n	Quasi-elastic	0.39 - 1.5	W. Bartel et al. NP B58, 429(1973)
G_E^n, G_M^n	Anti-coincidence	0.28 - 1.8	K.M. Hanson et al. PR D8, 753 (1973)
G_E^n	Elastic	0.002 - 0.016	G.G. Simon, et al., NP A364, 285 (1981)
G_M^n	Quasi-elastic	2.5 - 10.0	S. Rock et al. PRL 49, 1139 (1982)
G_M^n	Quasi-elastic	0.48 - 0.83	A.S. Esaulov et al. Sov. J. NP 45, 258 (1987)
G_M^n	Quasi-elastic	0.75 - 2.57	R.G. Arnold et al. PRL 61, 806 (1988)
G_E^n	Elastic	0.04 - 0.72	S. Platchov et al. NP A510, 740 (1990)
G_E^n, G_M^n	Quasi-elastic	1.75 - 4.0	A. Lung et al. PRL 70, 718 (1993)
G_M^n	Coincidence	0.109 - 0.255	P. Markowitz et al. PR C48, R5 (1993)
G_M^n	Ratio	0.11	H.A. Anklin et al., PL B336, 313 (1994)
G_M^n	Ratio	0.125 - 0.605	E.E. Bruins et al. PRL 75, 21 (1995)
G_M^n	Ratio	0.2 - 0.9	H.A. Anklin et al., PL B428, 248 (1998)
G_M^n	—	0.2 - 0.9	G. Kubon et al., PL B524, 26 (2002)
G_M^n	Quasi-elastic	1.0 - 4.8	Lachniet et al. PRL 102, 192001 (2009)

Table 1.1: A table of neutron form factor experiments.[2]

[2]This list is by no means exhaustive and may not include all experiments.

1.8.2 Previous nTPE Measurements

In short, there are no previous measurements of the two photon exchange contribution in electron-neutron scattering [4]. Similarly, this experiment is the first employment of the Rosenbluth Separation method to experimentally measure G_E^n [4]. There are numerous experiments for the two-photon exchange for the proton but none for the neutron. Some recent proton TPE experiments include GMP12 (2022) [32], OLYMPUS (2017) [68], and CLAS (2015) [1]. The shift of focus from electron-proton to electron-neutron comes after further discoveries continue to highlight the *Form Factor Ratio Puzzle (FFRP)* [4, 32, 140, 19].

1.9 The Form Factor Ratio Puzzle

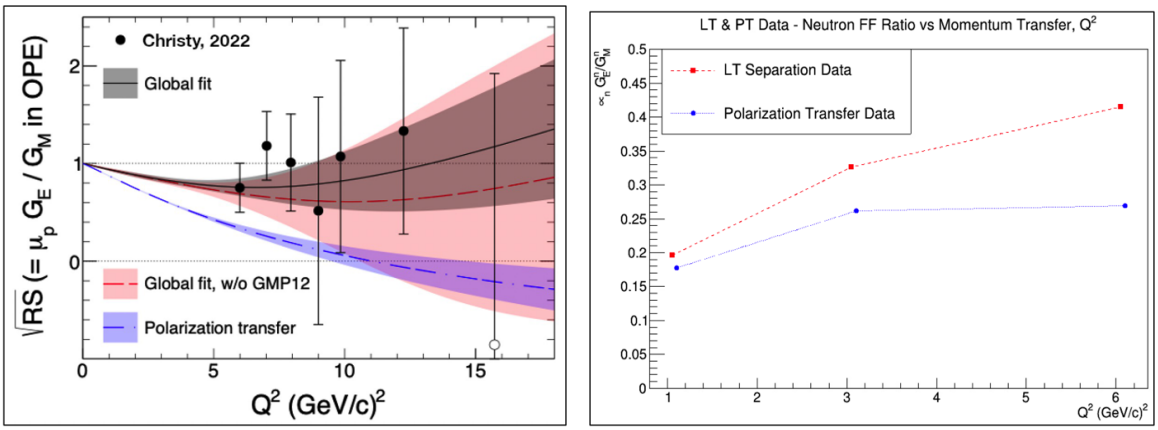
The two primary measurement techniques for the electromagnetic form factors are the Rosenbluth Separation and polarization transfer methods [19, 32]. Both experimental techniques utilize the ratio of form factors, R, in their extraction method. The ratio of form factors is related to a term known as the Rosenbluth Slope, S, which is the ratio of longitudinal and transverse cross-section components[3]. These are:

$$\begin{aligned} \text{Form factor ratio:} \quad & R = \frac{G_E}{G_M} \\ \text{Rosenbluth Slope:} \quad & S = \frac{G_E^2}{\tau G_M^2} = \frac{\sigma_L}{\sigma_T}. \end{aligned} \tag{1.32}$$

Several decades of form factor measurements have highlighted an unexpected deviation between the results of the two methods. As the momentum-transfer Q^2 increases, the form factor ratio as measured via polarization transfer decreases drastically while R, as measured via Rosenbluth Separation, does not [4, 32, 109]. Figure 1-6 presents these differences. Fig. 1-6a shows world data for proton form factor

[3]The variable used to represent the Rosenbluth Slope varies from author to author. In this thesis, I will primarily use S to denote the Rosenbluth Slope. This can be super-scripted for either the proton, S^p, or neutron S^n. At times, I may reference the Rosenbluth Slope as RS.

ratio measurements made by either Rosenbluth Separation or polarization transfer [32]. The red dashed line represents data for the Rosenbluth Slope method while the blue line is for polarization transfer. A separation between the two measurement techniques can be seen as Q^2 increases. This was an unexpected discovery.



(a) Proton data [32] **(b)** Neutron predictions [2]

Figure 1-6: The form factor ratio puzzle. **(a)** World data for proton form factor ratio shows a deviation between the Rosenbluth Separation and polarization transfer methods [32] **(b)** Calculations using parameterizations for the neutron form factors predict a similar discrepancy for the neutron. Plot re-produced from [2].

Figure 1-6b shows predictions for the form factor ratio of the neutron. The nTPE experiment is the first measurement of the neutron form factor ratio and therefore no world data exists [4]. The plot of Fig. 1-6b was produced by calculating the form factors via widely-accepted parameterizations [2, 140, 22]. The red dashed line is for Rosenbluth (or LT) Separation calculations while the blue dashed line is from parameterizations strictly for polarization transfer data. From these predictions, we expect to see a similar deviation between the two measurement techniques for the neutron as well. As Q^2 increases in Fig. 1-6b, the two datasets diverge, with the Rosenbluth Separation points increasing more rapidly than those of the polarization transfer technique.

The following discussion is in regards to the *proton form factor ratio puzzle*. A similar discussion should apply for the neutron, but this discussion will only focus on the proton since there is experimental precedence therein. An early explanation for this disagreement was that large systematic errors could be associated with the

Rosenbluth Separation method, especially as the experimental Q^2 increased [109]. This was a reasonable concern. However, measurement precision has evolved greatly since the first appearance of the FFRP, and Rosenbluth Separation systematic errors are now on par with similar polarization transfer method measurements [73, 106, 19].

Another possible cause of the form factor ratio discrepancy was thought to be an uncertainty in the relative normalization of the cross-section for L-T (Rosenbluth) separation [32, 109]. This explanation has fallen out of favor, as several re-analyses of global data have been able to correct for these normalizations. Through the efforts of those like Ye, Arrington, and others, the relative normalization across the global data was updated and yet, the same discrepancies continue to exist [8, 140, 22].

The next (and presumably last) possible cause to consider is that there are higher order effects and corrections which are not included in our current formulation. Up to the single photon exchange approximation of the Born term, we are already including radiative corrections, except for *hard two photon exchange* contributions [32]. Therefore, in terms of the Born approximation these considerations are complete. However, if we look beyond the single photon approximation and consider the next higher order effects, the largest contributor (and primary suspect) is the contribution from two-photon exchange [19, 10, 32, 109].

Figure 1-7 is similar to Fig. 1-6b except that it is for the proton and highlights the slope between the two measurement techniques. This plot was taken from [10]. The measurements from polarization transfer only restrict the slope of the cross-section and not its actual value. It has been normalized arbitrarily to align with the Rosenbluth results at $\epsilon = 0$ [9]. If we assume that the Born approximation is completely correct, then the true slope is represented by the red line. This would then include all of the possible contributions from the correct electric form factor [8]. The remaining slope is therefore attributed to unaccounted for corrections; namely, the two-photon exchange. This remaining slope should be nearly linear. The linearity of this remaining slope is crucial to maintaining the expected linearity of the Rosenbluth formula, as the reduced cross-section aligns with a linear ϵ-dependence within our accepted uncertainty range [10].

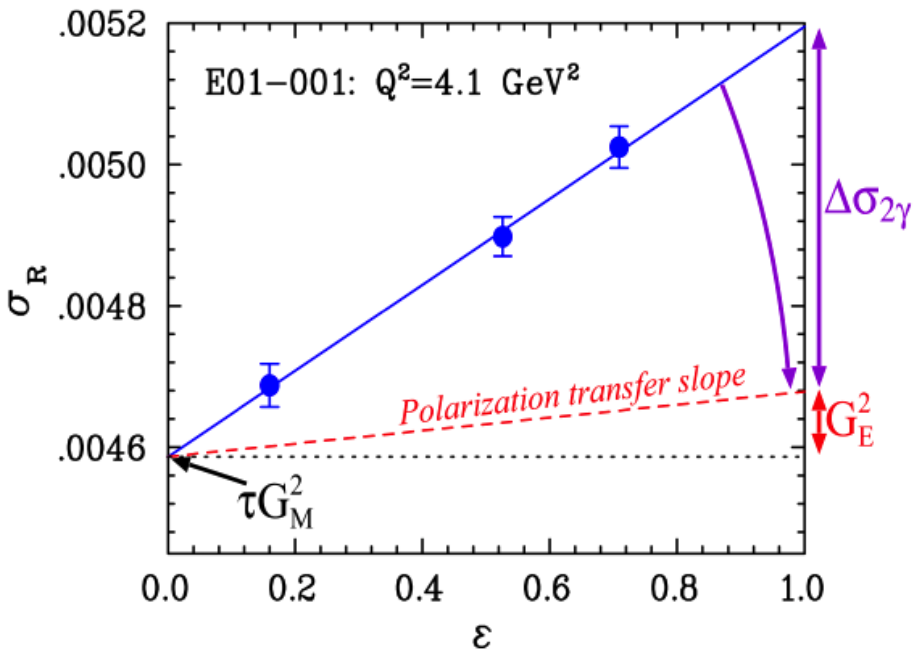


Figure 1-7: Plot of the proton reduced cross-section, σ_R, vs. epsilon for LT Separation and polarization transfer methods. The blue data points represent LT Separation measurements along with their linear fit. The red dashed line corresponds to data taken via polarization transfer methods. The discrepancy between the two is shown as $\Delta\sigma_{2\gamma}$ [10, 106]. Plot credit: [10].

It is important to state that two photon exchange measurements are only reliable at reasonably high Q^2 values [10]. Up to this point, all discussion has been under the assumption that the impact from TPE on polarization transfer measurements is negligible. This is a reasonable enough assumption at high Q^2 values where LT separation is highly sensitive to TPE [19, 109]. At lower Q^2 values, the effects of TPE on both measurement techniques are either very small or similar to each other [9].

The Form Factor Ratio Puzzle is a central component of why we have made the measurements of G_M^n at kinematic points SBS8 and SBS9. These measurements allow us to attempt an extraction of nTPE that can be correlated to the discrepancies observed and discussed as the FFRP. A value extracted and attributed to the nTPE contribution will hopefully correspond to the calculations and projections shown in Figure 1-6b. In any case, this is a truly important premiere measurement and first step towards solving this puzzle.

Chapter 2

Experimental Setup

2.1 Overview

The Super BigBite Spectrometer (SBS) Program is a series of ground-breaking experiments for measuring nucleon elastic electromagnetic form factor ratios. The SBS Program is part of the so-called "12 GeV Era" of physics at Thomas Jefferson National Accelerator Laboratory (Jefferson Lab or JLab) in Newport News, Virginia. The 12 GeV upgrade at Jefferson Lab doubled the energy of the electron beam to 12 GeV, added an additional experimental hall to the site (Hall D), and included the construction of updated detector hardware. The original line-up of experiments for the SBS Program consisted of GMn, GEn-RP, nTPE, GEn-II, and GEp-V. The program was scheduled as three separate runs with some configuration changes in-between. The first run of experiments included GMn and nTPE, the second run was the GEn-II experiment, and the third run — the last of the program — will be GEn-RP, K_{LL}, and GEp-V.

Due to various constraints, the first run of experiments only included GMn and nTPE. These two experiments ran in parallel with each other and shared experimental hardware and detector setups. The only differences between the two were the kinematic setups required for each specific measurement. The nTPE experiment in a sense ran *parasitically* to the GMn experiment and so, often, the experimental setup of nTPE may be referenced as the GMn setup, or simply, GMn. In GMn there were

six different kinematic settings: SBS4, SBS7, SBS8, SBS9, SBS11, and SBS14. The two settings which correspond to nTPE are SBS8 and SBS9. Technically, SBS9 is both an nTPE and GMn measurement point whereas SBS8 is only for nTPE. However, GMn, can be extracted from all of the kinematic points listed. The finer details of why SBS8 and SBS9 overlap for the nTPE experiment are discussed further in Chapter 4.

2.2 Accelerator and Beamline

2.2.1 CEBAF

At the heart of all experiments at Jefferson Lab is the Continuous Electron Beam Accelerator Facility (CEBAF). Since 1995, it has provided a high-duty factor continuous electron beam to each experimental hall. The original design had an energy specification of 4 GeV but as the experiments evolved through their successes, greater and greater beam energies were needed [91]. The CEBAF recently underwent a major upgrade (completed in October 2017), which now allows the beam to reach energies up to 12 GeV. This upgrade provides a pathway to exploring such scientific frontiers as the quark sub-structure of nuclei, searching for new physics through high-precision tests of the Standard Model, and probing into the physical origins of quark confinement [67].

Figure 2-1 shows a drawing of the CEBAF after the 12 GeV upgrade. The CEBAF works on the premise of a recirculating *linear accelerator* (LINAC or linac) where charged particles acquire energy while moving on a series of circulating linear paths. The CEBAF has a polarizable electron beam, two superconducting LINACs (North and South) connected by two sets of recirculating arcs, and various extraction elements. In order to produce a polarized electron beam which can be injected into the accelerator loop, a 1497 MHz diode laser is placed on a strained gallium arsenide cathode. Each recirculating arc of CEBAF contains both quadrupole and dipole magnets in separated beamlines which allow for the electron beam to be accelerated up

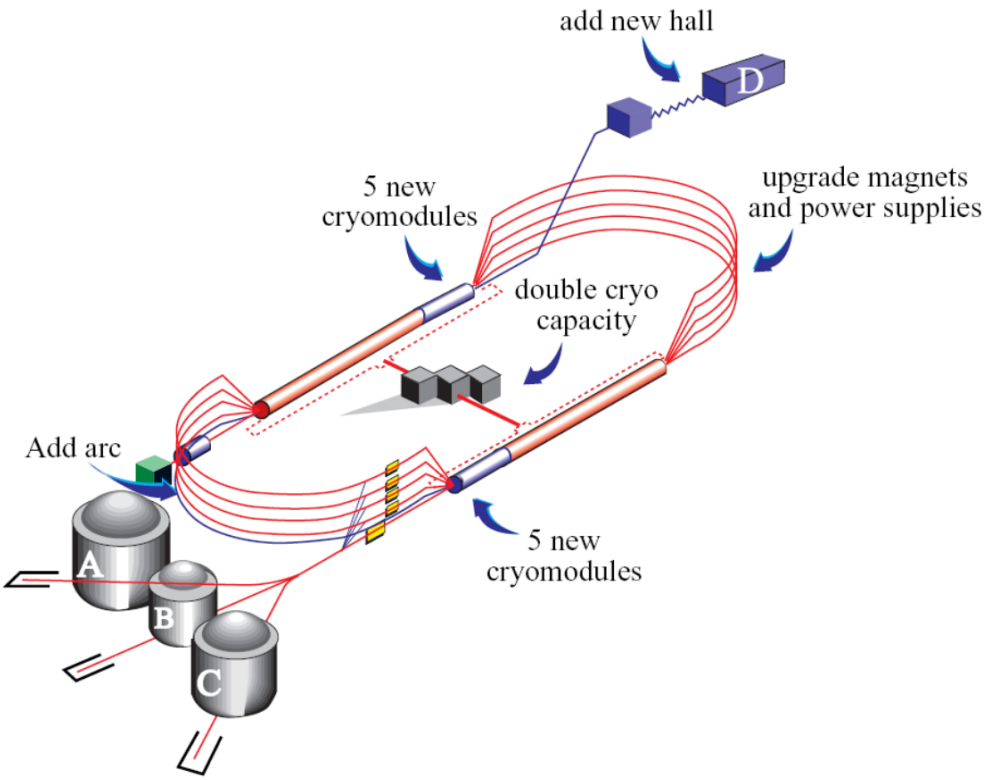


Figure 2-1: Continuous Electron Beam Accelerator Facility with the 12 GeV upgrade.

to five times through each linac. Each linac section on the CEBAF consists of 25 cryomodules, each of which has eight superconducting radio frequency (RF) cavities. Each cryomodule is capable of providing an average acceleration of 100 MeV. This provides 1.1 GeV through each linac section and allows each recirculated beam to reach 2.2 GeV [91]. Extraction units are located at the end of the South linac. This allows CEBAF to provide an 11 GeV beam to Halls A, B, and C. The inclusion of one additional recirculating arc on the North linac will allow Hall D to receive the full 12 GeV beam [67].

2.2.2 Experimental Hall A

The SBS Program is running in Experimental Hall A of Jefferson Lab. In total, there are four experimental halls at Jefferson Lab (A, B, C, and D). An aerial shot of the accelerator campus is shown in Figure 2-2. The domed ground coverings

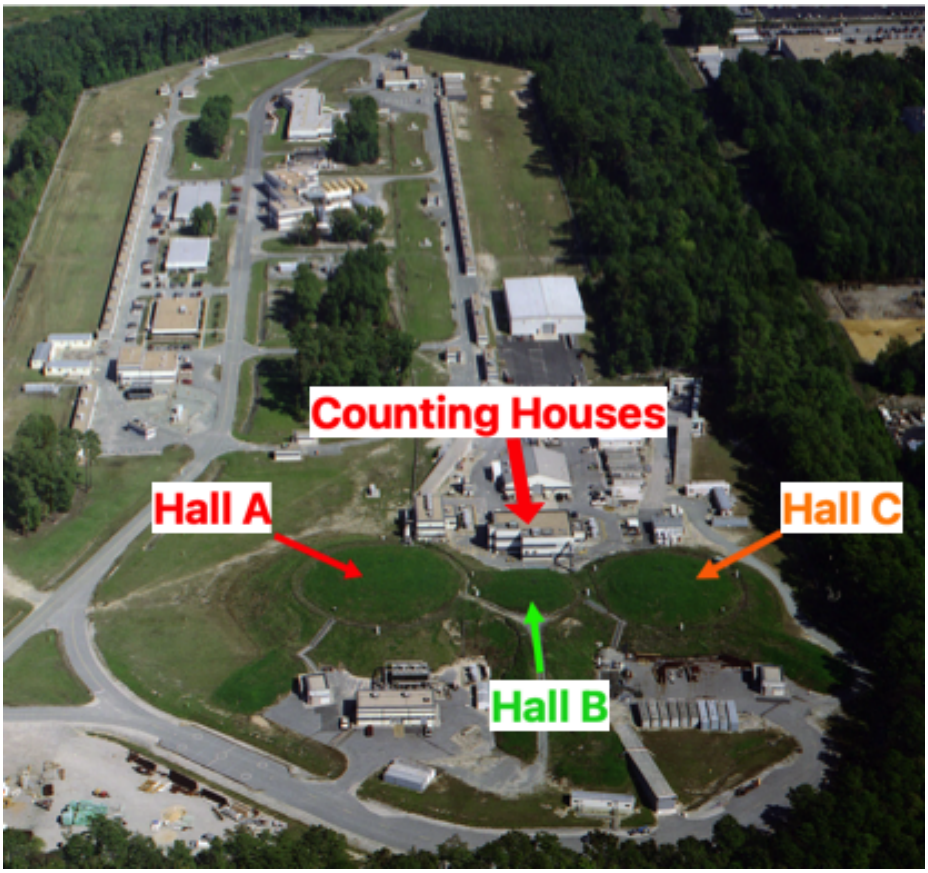


Figure 2-2: Aerial photograph of JLab's accelerator campus. The domed ground covers above Experimental Hall A, B, and C are shown.

for Halls A, B, and C are labelled. Each hall, just like the entire accelerator, is situated underground for radiation shielding and for operational, security, and safety reasons. Each experimental hall can receive the electron beam from the CEBAF, but only some fraction of the total beam energy is available for a given hall during multi-hall operation. This limitation comes from the total power that the accelerator can deliver, as given by the familiar electric power equation, $P = VI$; in this case, V corresponds to the beam energy. The electron beam is parsed out amongst the shared halls accordingly. Therefore, if one Hall requires highest beam energy at a high current, the others are restricted to the remainder of power leftover. Apart from some monitoring devices, electronics, and gas systems, the entire experimental setup for the SBS Program resides underground in Hall A. This includes detector sub-systems such as the BigBite Calorimeter (BBCal), the Hadron Calorimeter (HCal), the GEM

trackers, the BigBite Magnet, the Super BigBite Magnet, the Cherenkov detector (GRINCH), the Hodoscope, various DAQ components, and the computing systems. The primary ground-level access point to the experimental halls is the counting house. Each hall has a corresponding counting house. For instance, Counting House A is for Hall A (Fig. 2-2). This is the main station where the experimental runs are operated and managed by the Shift Crews. This houses much of the sub-system monitoring equipment and other various computational resources as well.

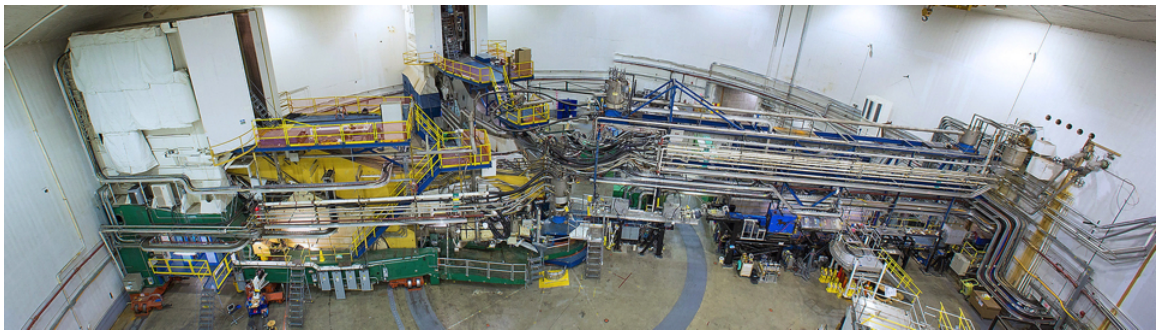


Figure 2-3: Panoramic photo of the inside of Hall A. The scattering chamber can be seen near the center of the photo. Note, this photo is not of the SBS Program and is only used to provide a reference for the interior space of Experimental Hall A.

Figure 2-3 shows a panoramic shot of the inside of Experimental Hall A. This was not the configuration used during the GMn or nTPE experiments; however, this does provide a relative idea of the inner space of the hall. The beam line runs horizontally in this picture at nearly the vertical midway point of the photo. The scattering chamber at the target can be seen near the center of photo. More detailed photos will be shown to reference the various sub-systems that are uniquely configured for the SBS Program within the circular hall shown in Figure 2-3.

2.3 Experimental Configuration and Layout

The experimental layout of the nTPE experiment (and GMn alike) consists of three primary *arms*. These arms are the upstream incoming beam line and the two downstream scattering arms for the electron and the nucleon(s). Figure 2-4 is an overhead drawing of this experimental setup. The three arms of the experiment are

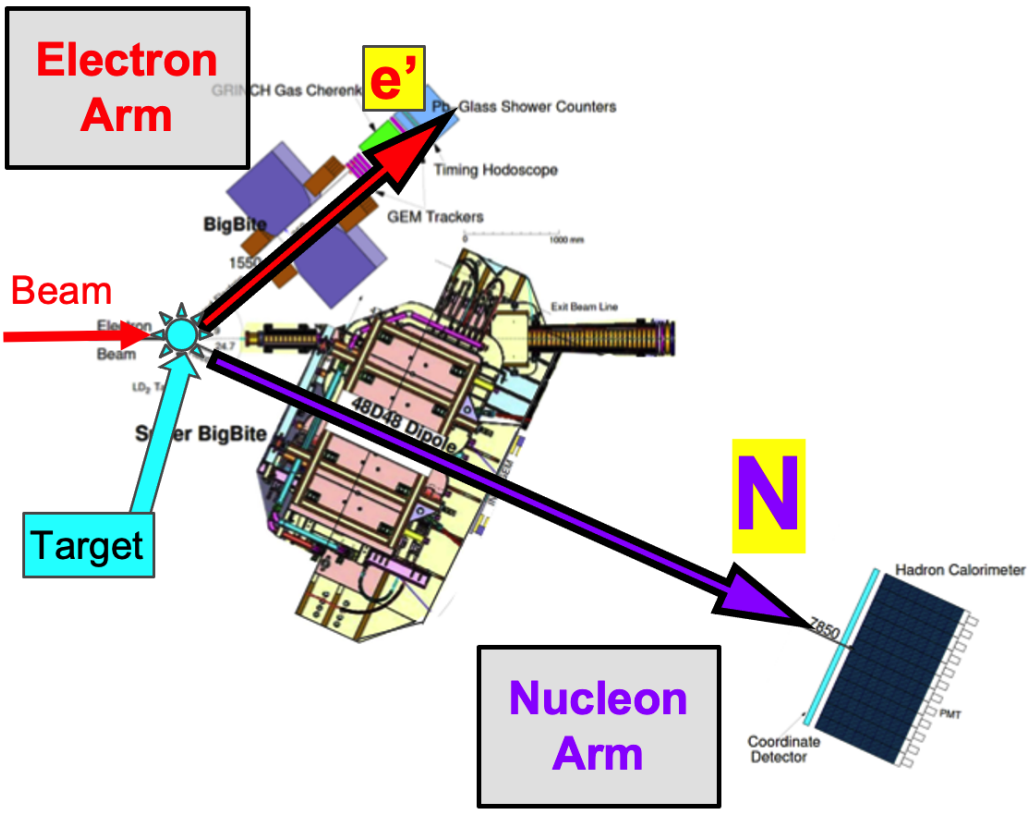


Figure 2-4: Drawing of the experimental layout of the first run of the SBS Program. Shown are the three primary arms of the setup: incoming beam, scattered electron arm, and scattered nucleon arm.

labelled, along with the target. The incoming beam line corresponds to the electron beam that is delivered from CEBAF. It enters the hall and is incident on the target which is enclosed in a scattering chamber. The vast majority of the incoming electrons do not interact in the target, and continue down the exit beam pipe. Downstream of the target, on either side of the exit beam pipe, are the two scattering arms, which each contain their own corresponding spectrometer. The Electron Arm corresponds to the path of the scattered electron and has the BigBite Spectrometer. It consists of the BigBite magnet, GEM detectors, the GRINCH, the BigBite Calorimeter, and the Timing Hodoscope. This arm, or spectrometer, is referenced by many terms, such as BigBite, BB arm, or e-arm. The scattered hadrons travel down the Nucleon Arm. This arm is the SuperBigBite Spectrometer and consists of the Super BigBite Magnet and the Hadron Calorimeter. It is often referenced as the SBS Arm, the Hadron Arm,

or simply SBS.

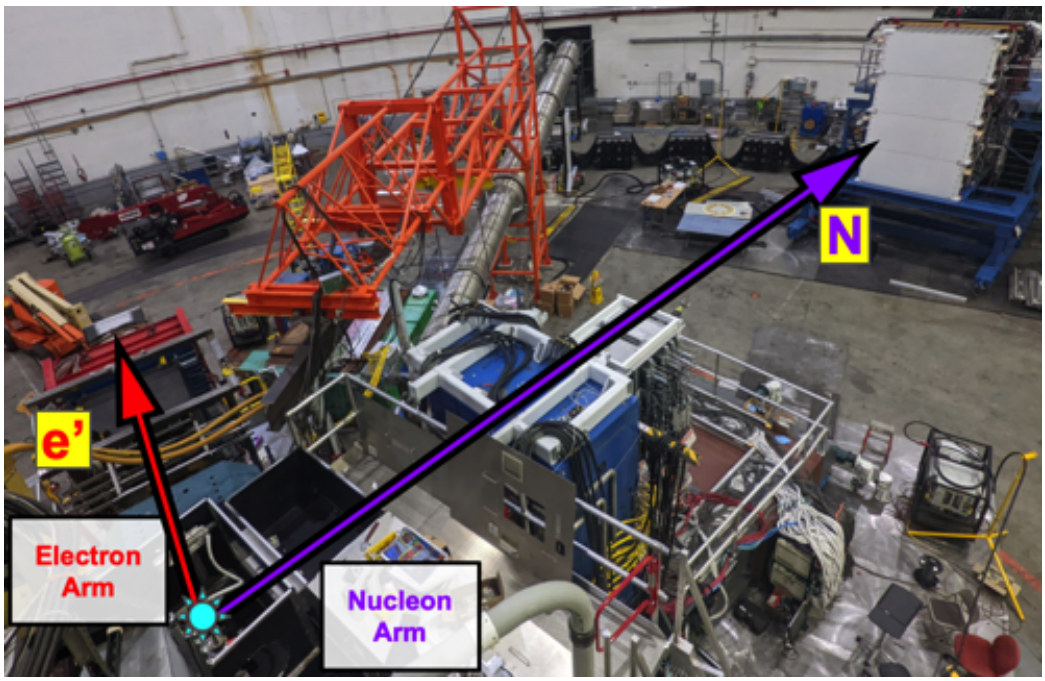


Figure 2-5: Experimental setup of GMn and nTPE in Hall A. Shown are the two downstream arms. On the left is the scattered electron arm and on the right is the Hadron Arm. HCal can be seen in the upper right hand corner of the image.

Figure 2-5 shows the two downstream spectrometer arms of the experimental setup for GMn and nTPE in Hall A. After the incoming electron beam interacts with the target, the electron scatters to the left (as shown in the picture and looking downstream from the target). It then travels through the BigBite detector stack and *lands* on the BigBite Calorimeter (BBCal). The scattered hadron (proton or neutron) scatters off to the right (as shown in the picture), passes through the SBS Magnet, and lands on the Hadron Calorimeter (HCal).

2.4 Spectrometers

2.4.1 BigBite Spectrometer

BigBite Observables

The BigBite spectrometer provides most of the measurement channels of the experiment and is used to measure and define all of the parameters associated with the scattered electron. The BB Spectrometer defines the q-vector, the scattered electron momentum (p_e), the trajectory of e', the reaction vertex for the selection of quasi-elastic scattering, and the trigger time correlation. BBCal and GRINCH also provide separate measurements, which can be used for pion identification (pion rejection). This is by no means a complete list of extractable observables from the BigBite stack, but rather just the major parameters defined by this arm. The measurements from the scattered electron not only provide direct information about the electron, but they are also used to calculate projections of the elastically scattered hadrons along the SBS arm as well [47].

Design and Layout

The BigBite Spectrometer is comprised of all of the components along the scattered electron arm. This spectrometer has a relatively large acceptance and is a medium resolution spectrometer. The scattered electron first passes through the BigBite dipole magnet and then through the BigBite detector stack. The BigBite dipole field deflects the electron's trajectory for momentum analyses. This detector stack consists of multiple sub-detector systems. The BigBite stack has two sets of tracking GEM detectors: a front tracker and back tracker. The front tracker is the first sub-detector of the BigBite stack. These GEMs are followed by the particle-identifying GRINCH Cherenkov detector. After the GRINCH is the back GEM layer, the BigBite PreShower Calorimeter, Timing Hodoscope, and the BigBite Shower Calorimeter.

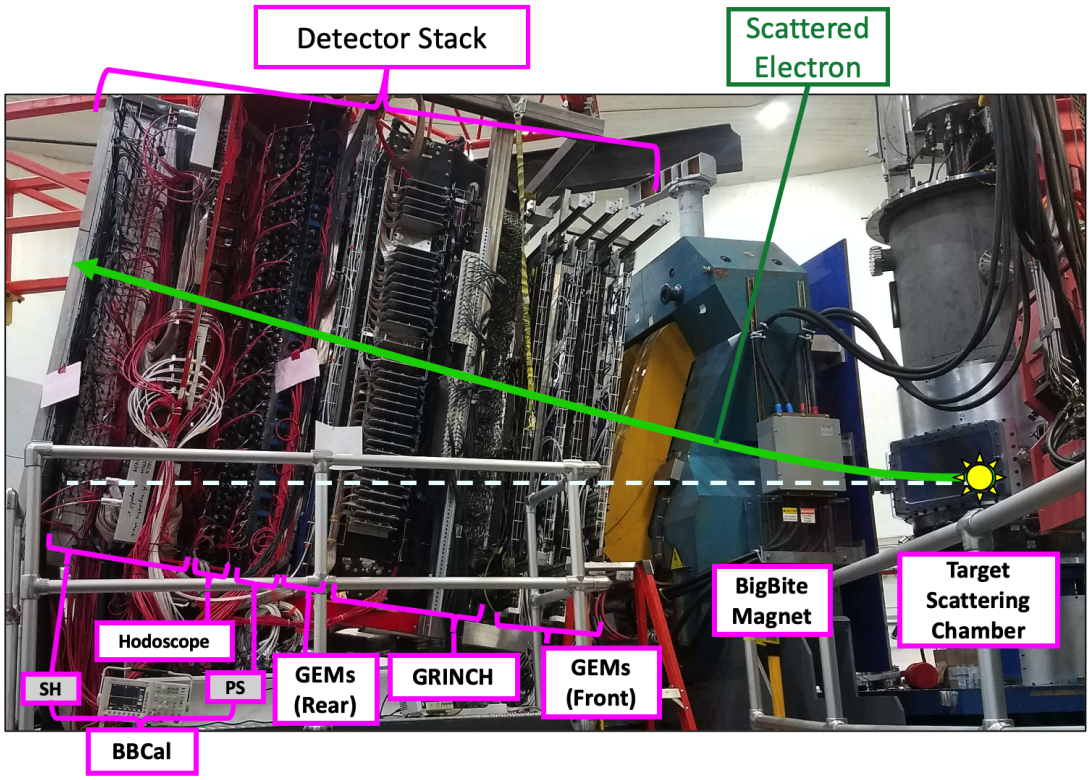
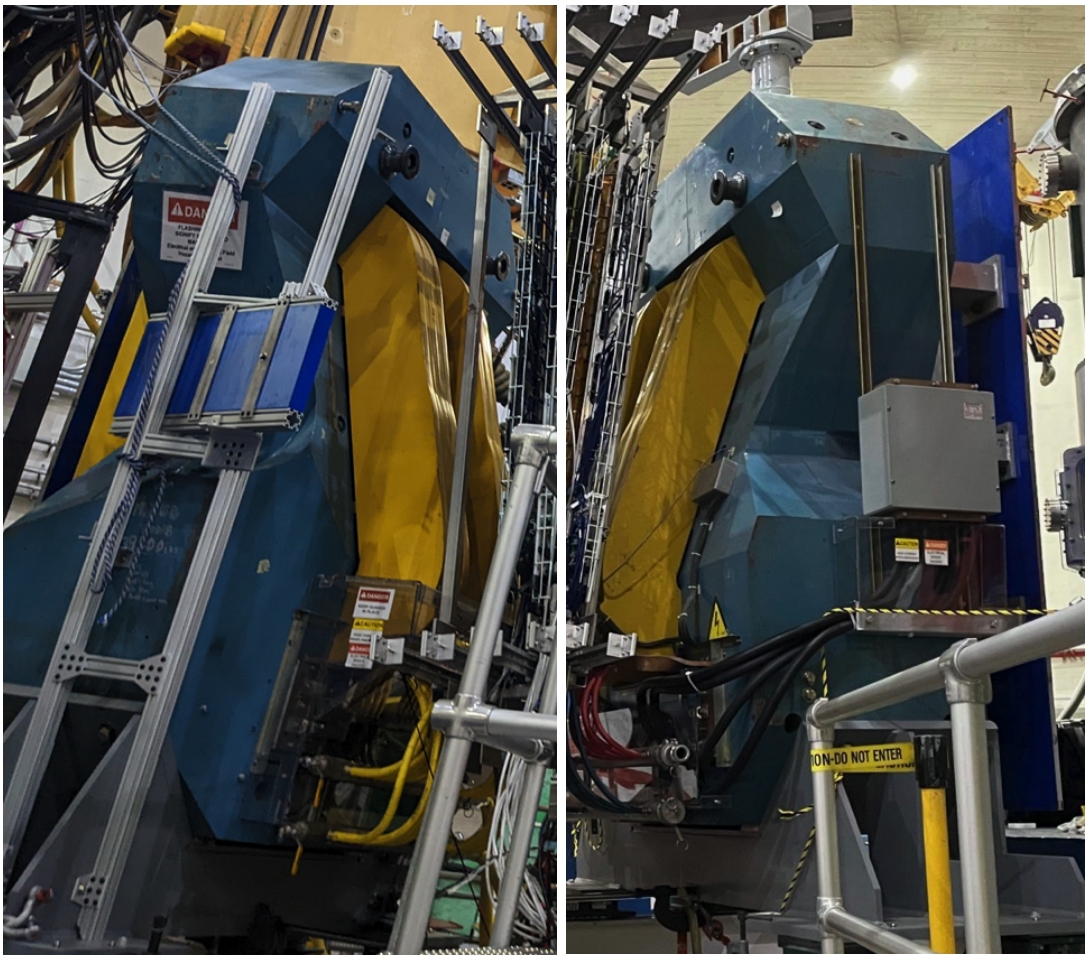


Figure 2-6: The BigBite Spectrometer. The BB Spectrometer is along the scattered electron arm and contains the BigBite Magnet, front and rear GEM trackers, the GRINCH, the BBCal PreShower, the Timing Hodoscope, and the BBCal Shower.

The BigBite Magnet

The BigBite Magnet (also referred to as the BB magnet or dipole) is the large dipole magnet that sits at the front of the Electron Arm, just downstream of the target chamber. The magnet has an opening, or yoke, through its center, which allows the scattered electrons to pass through. As they pass through the magnet, the electrons are deflected. This allows for accurate measurement of the electron's momentum: $p_e\theta_e \approx 0.3 \int B \cdot dl$. Here, p_e and θ_e are the scattered electron's momentum and the out-of-plane angle, respectively. The right-hand-side of the equation is the scaled BB dipole field integral. From this we can solve for the momentum given the known quantities on the R.H.S. and the measured values for θ_e. Further discussion on the optics and momentum calculations can be found in Section 2.6.3. The maximum field integral of the BigBite dipole magnet is approximately 0.9 T-m. The field strength is



(a) **(b)**

Figure 2-7: The BigBite Magnet. **(a)** Downstream "right" side of the BigBite magnet on the BB detector stack. **(b)** Downstream "left" side of the BB magnet.

adjustable for optimization and tuning purposes, but once the field was set it remained at that setting for the entirety of the experiment [105].

2.4.2 Super BigBite Spectrometer

Super BigBite Observables

Though the Super BigBite Spectrometer arm consists of only a couple of sub-detectors, it provides extremely important quantities: all proton and neutron measurements. The field from the SBS dipole magnet deflects the protons and provides a means for nucleon identification. The Hadron Calorimeter also provides the scattered

nucleon momentum (p_p and p_n), as well as their directions and final positions. This arm also provides the time-of-flight and HCal cluster energies that can be used to veto background particles and assist in elastic event selection [47].

Design and Layout

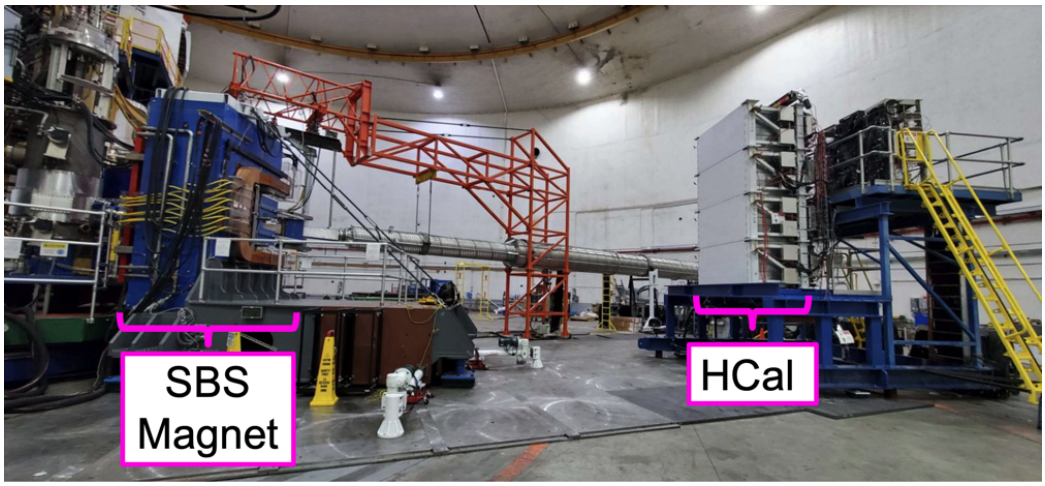


Figure 2-8: The Super BigBite Spectrometer. The SBS Spectrometer is along the scattered hadron arm and contains the Super BigBite Magnet and the Hadron Calorimeter.

The Super BigBite Spectrometer is along the scattered hadron arm and is comprised of the Super BigBite Magnet and the Hadron Calorimeter. The scattered hadron(s) first pass through the Super BigBite Magnet and then land on the Hadron Calorimeter at the *end* of the arm. The SBS Magnet deflects the charged protons that pass through it. Because of this, when the SBS magnetic field is *on*, the protons and neutrons will land at different spots on the HCal. The neutrons fly straight through without deflection while the protons are deflected upwards (away from the ground).

The Super BigBite Magnet

The Super BigBite Magnet (also referred to as the SBS magnet, 48D48 magnet, and the SBS dipole magnet) is the large dipole magnet which scattered particles pass through along the Hadron Arm of the experiment. Photos of the SBS dipole are shown in Figure 2-9. Figure 2-9a is a view looking *upstream* at the SBS magnet from

(a) **(b)**

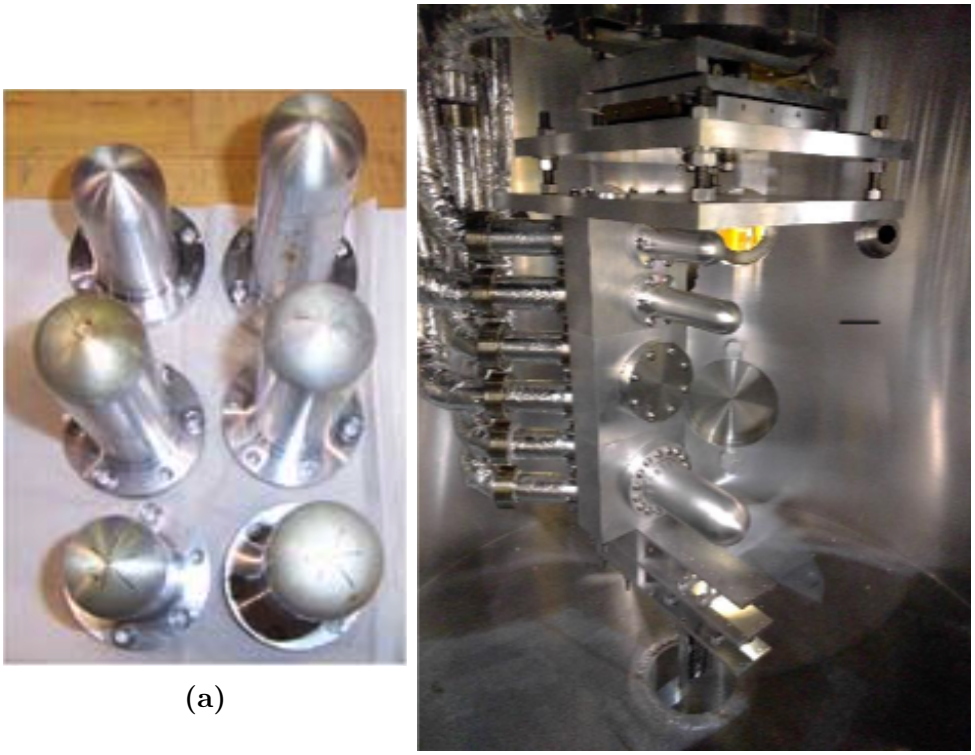
Figure 2-9: SBS Dipole Magnet. **(a)** Looking upstream at the SBS magnet. **(b)** Side-view of the SBS magnet.

HCal. Figure 2-9b is a side view of the magnet. The magnet has a large aperture with a vertical:horizontal aspect ratio of 2:1. The field of the magnet is oriented in the *horizontal* (B_y) direction, i.e., parallel to the ground. The maximum field is approximately 1.69 T with an integral field strength of 2.0 T-m (2.5 T-m with pole shims) [138]. The magnet has an opening through its return yoke which the scattered particles pass through. The adjustable field integral, $B \cdot dl$, will allow for variations in the separation between the two hadron peaks on HCal [47].

The variability of the SBS dipole is useful for multiple reasons. For one, the larger the separation between the proton and neutron peaks is, the easier it is to separate and un-correlate them from one another. The ultimate goal is to count the individual numbers of protons and neutrons which land on HCal. Thus, it helps to have two cleanly separated peaks without any overlap. This isn't always achievable at every kinematic setting. As the scattered nucleon momentum increases, the less the proton deflects due to the SBS dipole field. Therefore, as the proton momentum increases (decreases) with each kinematic setup, the SBS dipole field scale also needs to increase

(decrease) to result in a similar amount of proton-to-neutron separation.

2.5 Targets



(a) **(b)**

Figure 2-10: Hall A cryotargets. **(a)** Cryotarget 'cigar-tube' style cells [7]. **(b)** GMn and nTPE cryogenic target ladder. Each cylinder seen in the photo is a different target along the ladder [77].

All of the experiments in the SBS Program are elastic electromagnetic form factor experiments involving an electron beam incident on a nucleon. The GMn and nTPE experiments each used two separate production targets, a liquid hydrogen target (LH2 or simply hydrogen) and a liquid deuterium target (LD2 or simply deuterium). The LH2 target is used as a proton target primarily for calibration purposes whereas the LD2 target serves as a neutron target for experimental production data. LD2 is used since there are no free stable neutrons. LD2 provides stable target neutrons for our incoming electron beam. Interactions will also occur with the bound protons of LD2, which is precisely why we need the SBS dipole along the Hadron Arm for proton-

neutron identification. The targets are cryogenically cooled, and thus, referred to as *cryotargets*. The cryotargets for Hall A are shown in Figure 2-10. The individual target cells are shown in Fig. 2-10a while Fig. 2-10b shows a photo of the so-called *target ladder* for GMn and nTPE. Each *rung* on the ladder is a different target that can be moved into place in the beamline. The target ladder consists of the following targets: liquid hydrogen, liquid helium, liquid deuterium, carbon (hole), beryllium oxide, and titanium. The targets that are not LH2 or LD2 are used for various calibrations, setups, and test functions. The production targets are each in a cylindrical tube that is 15 cm long. The liquefaction for each target occurs due to heat exchange with helium vapor in the target loop/manifold (see Figure 2-11) [77, 71].

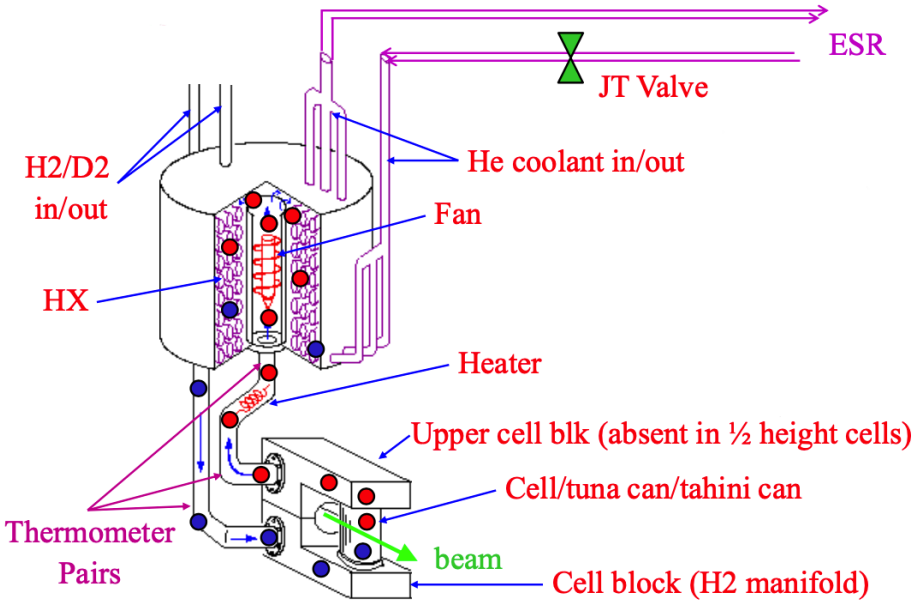
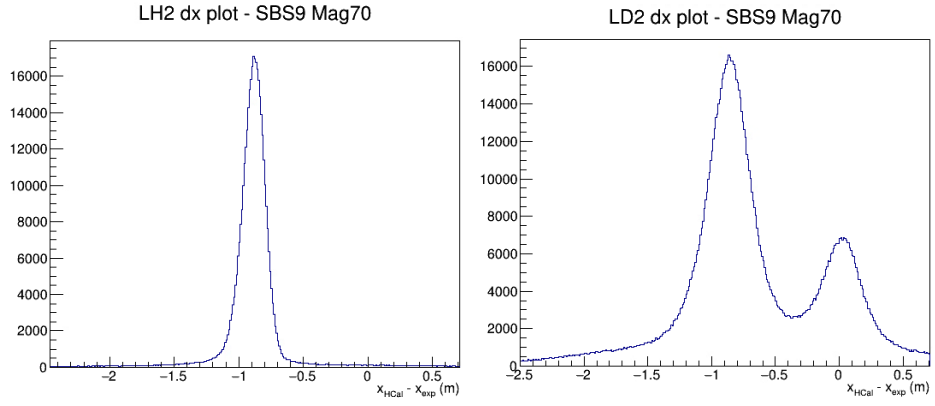


Figure 2-11: Anatomy of the GMn and nTPE target loop [77].

2.5.1 Liquid Hydrogen Target

The liquid hydrogen (LH2) target provides a stable, scattering proton that can be used to calibrate various parameters like momentum, optics, BigBite magnet settings, and SBS magnet field scales. LH2 data is used for calibrations either before taking production data on LD2 and/or during general, post-experiment data analysis. The

LH2 target had a nominal temperature of 19 K. Figure 2-12a shows the *delta x* plot for the scattered protons from the LH2 target. There is a single peak in this plot, because there are only protons present for scattering. This plot is from the SBS9 kinematic with an SBS dipole field setting of 70%. Because the field scale is non-zero, we expect the protons to be deflected. This is precisely why the peak is not centered around x = 0 in the figure. The SBS field deflects protons in the negative-x direction.



(a) **(b)**

Figure 2-12: LH2 and LD2 dx plots. These plots are from the SBS9 kinematic with an SBS field scale of 70%. The non-zero field scale is why the proton peak is not centered around x = 0. The neutron peak is centered around x = 0 since the neutrons are not deflected by the SBS dipole. **(a)** LH2 dx plot showing the single proton peak. **(b)** LD2 dx plot showing the separated proton and neutron peaks.

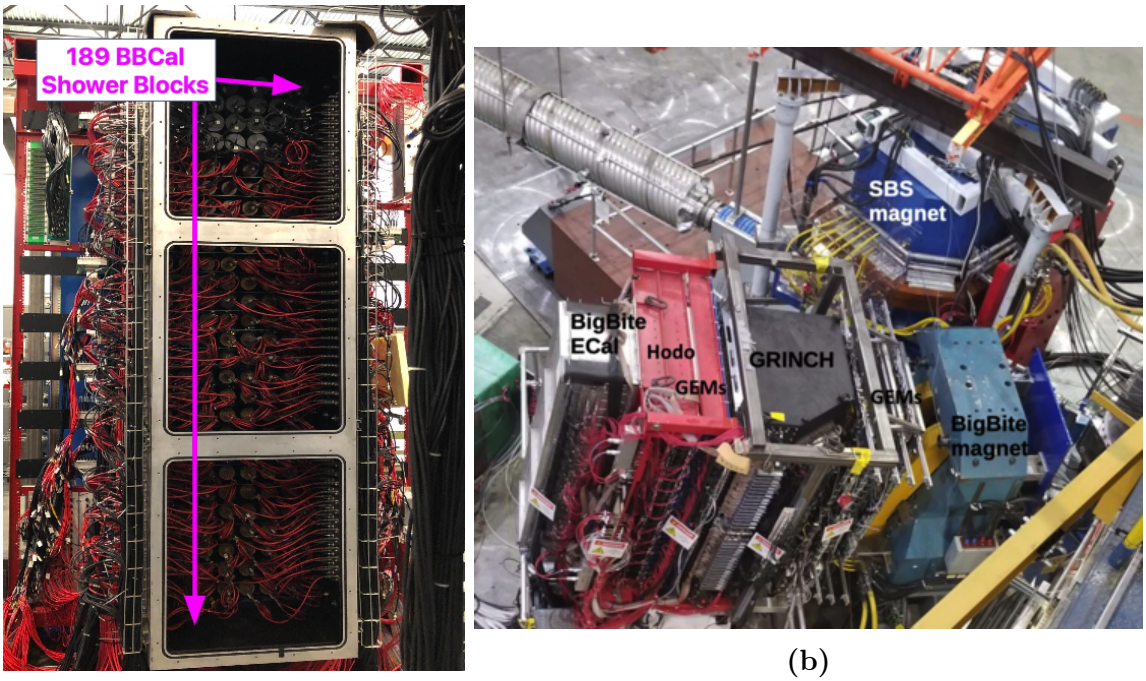
2.5.2 Liquid Deuterium Target

The liquid deuterium (LD2) target is used for the final production data that will be used to extract the quasi-elastic yields and subsequent electromagnetic form factor ratios. The LD2 target essentially provides us with a neutron target. Free neutrons decay with a lifetime of roughly fifteen minutes. Therefore, implementing a purely free neutron target for an experiment such as GMn or nTPE is unfeasible. Liquid deuterium targets are a very stable and proven means for providing a target neutron (and proton) in an elastic scattering experiment like that of the SBS Program [89]. Figure 2-12b shows the dx plot for the scattered hadrons (protons and neutrons) from the LD2 target. Two peaks are shown in the plot. The left peak corresponds to

scattered protons and the right peak corresponds to scattered neutrons. The peaks are separated from one another due to the presence of the SBS magnetic field along the scattered hadron arm. Like Fig. 2-12a, this plot is from the SBS9 kinematic with a SBS field scale of 70%. We therefore see similar proton deflection distances from zero in each plot. We can also see that the widths of the peaks for LD2 are wider than those for LH2. This widening is due to the effects of *Fermi motion*. Further discussion on these type of plots can be found in Section 4.3.2.

2.6 Detector Packages

2.6.1 BigBite Calorimeter



(a) **(b)**

Figure 2-13: BBCal on the BigBite Arm. **(a)** A photo of the Shower blocks of BigBite. **(b)** A photo of the BigBite stack with each sub-detector labelled. BBCal is at the back of the detector stack. Image source: [119].

The BigBite Calorimeter, or BBCal, measures the energy of the scattered electron on the e^- Arm. Figure 2-13a) shows the so-called *shower* blocks on the inside of

BigBite Calorimeter, and Figure 2-13b shows the detectors of the BigBite stack with BBCal and the end of the detector stack. The cells in the BigBite Calorimeter measure the electron's energy and provide the location of the high energy electron track with centimeter-level accuracy. This information is then used by the GEM trackers for high-precision track determination. BBCal is also responsible for sending the master trigger signals to the DAQ for event-taking [43].

The BigBite Calorimeter consists of two sets, or layers, of calorimeters: the *PreShower* (PS) and *Shower* (SH) Calorimeters. As the name implies, the PreShower Calorimeter is in front (upstream) of the Shower Calorimeters. Each of these calorimeters are made of rectangular lead-glass blocks that are each coupled with a photomultiplier tube (PMT). The PMT on a block can detect the effects of a charged particle (electron in this case) interacting with the lead-glass cell and measure the energy deposited therein. The aggregate sum of signals from a cluster of lead-glass blocks is compared to a threshold and used as the main trigger input for the experiment. A low-energy threshold can also be set on the PS Calorimeter data for applying a particle identification (PID) during the analysis. For instance, requiring a PS minimum energy of 0.150 MeV on all data to be analyzed results in an effective pion-rejection cut [128]. This is common practice for most of the analysis discussed within this thesis.

Figure 2-13a shows the general configuration of the PS and SH layers in BBCal. Since the PS is upstream of the SH, the charged particles first interact with the PreShower before moving through the Shower. The PreShower Calorimeter consists of 189 lead-glass blocks laid out in 27 rows that are two shower blocks deep per row. The cells are placed such that their long side is perpendicular to the direction of the scattered electron. Each of the PS blocks are 9cm x 9cm x 30cm. The Shower layer has 27 rows made up of 7 blocks each. This layer has its long dimension along the direction of the scattered electron. Each Shower block measures 8.5cm x 8.5cm x 37cm [127, 119].

The BigBite Calorimeter defines the most fundamental, yet utmost important, variables and parameters. It sets the basis from which many other detectors must

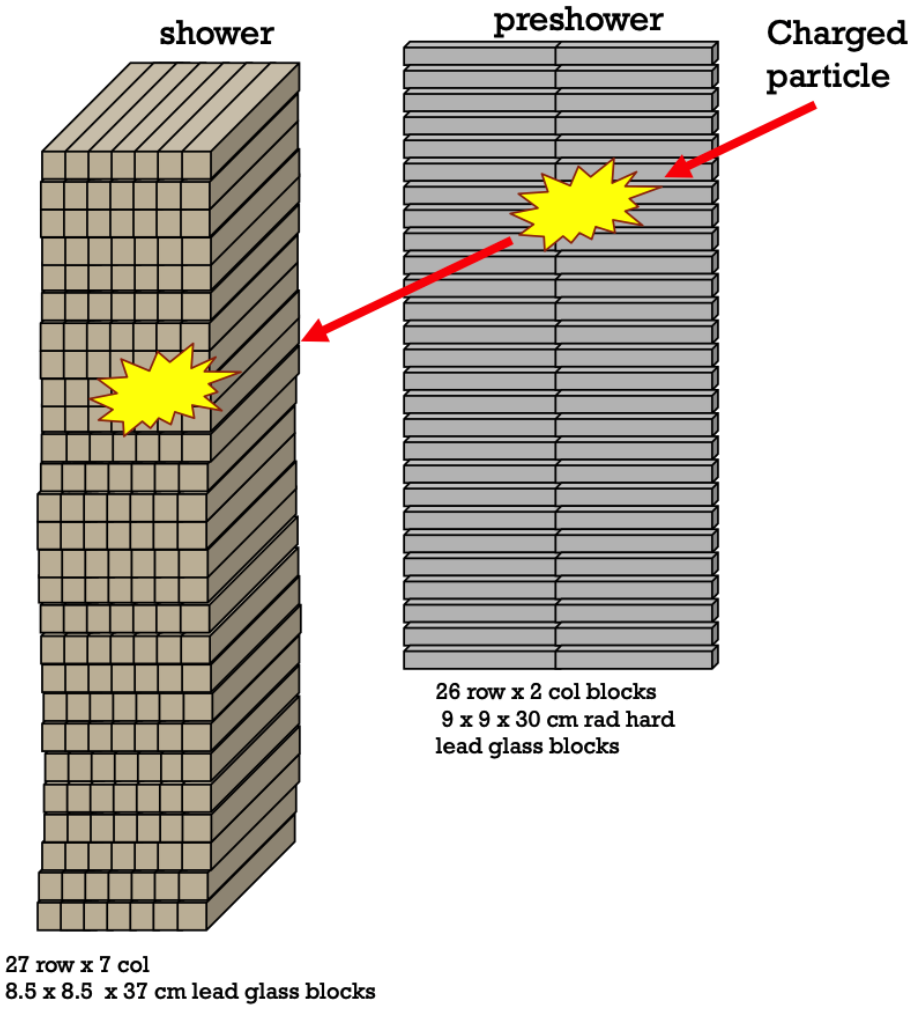


Figure 2-14: Drawing of the to sub-calorimeters that make up the BigBite Calorimeter: the PreShower and the Shower calorimeters. The PreShower is in front of the Shower and interacts with the scattered electron before the Shower. Image source: [127].

operate or calibrate from. BBCal undergoes a slew of calibrations, but the primary one is that of the scattered electron's total energy divided by its momentum ($E_{e'}/p_{e'}$). For the momentum transfer, Q^2, regimes of SBS4, SBS8, and SBS9 the mass of the electron is negligible and therefore, the $E_{e'}/p_{e'}$ distribution should be centered around 1. Figure 2-15 shows the energy over momentum plot before and after BBCal calibrations. The strong fields produced by the 48D48 SBS dipole magnet affect the performance of the PMTs on BigBite. These fringe field effects can be alleviated through this calibration of BBCal. This not only calibrates the energy determination

from the calorimeter response, but also helps to remove position bias from the main DAQ triggers provided by BBCal [119].

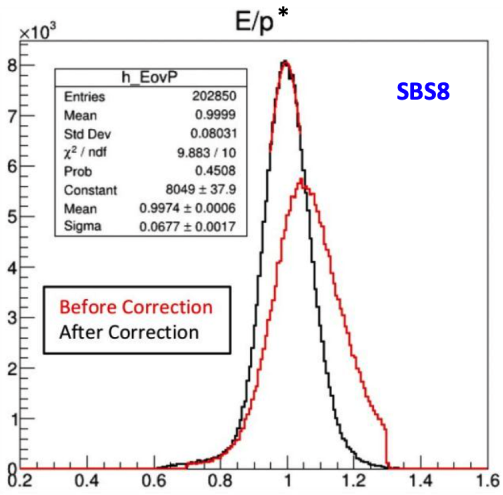
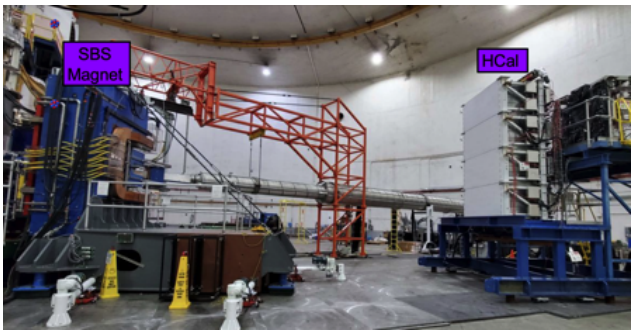


Figure 2-15: Calibration of BBCal for the *energy over momentum* plot of the scattered electron. The plot shows the distribution before and after calibration. For this Q^2, $E_{e'}/p_{e'}$ should be centered around 1.0. Plot source: [119].

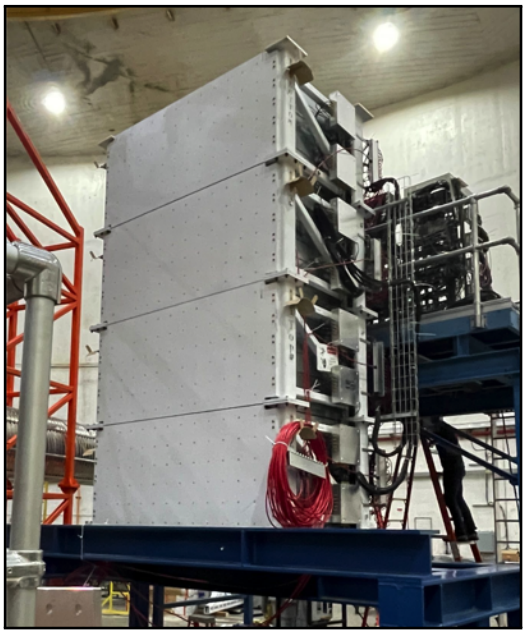
2.6.2 HCal

The Hadron Calorimeter, or HCal, measures the scattered proton and/or neutron on the Hadron Arm. Figure 2-16a shows the Hadron Arm of the GMn and nTPE experiments. The SBS dipole magnet can be seen on the left, and the Hadron Calorimeter is on the right. Figure 2-16b shows a close-up photo of the HCal detector. As the scattered hadron interacts with the HCal, it deposits its energy into the cell with which it interacts. From there, the measurements from HCal are used to determine the hadron type, energies, position, and time-of-flight [119, 97].

HCal is a modular sampling calorimeter made from a total of 288 modules. The modules have their long dimensions along the direction of the scattered hadron and are arranged in 24 rows of 12 modules each (see Fig 2-17). A single HCal block is made up of alternating stacks of scintillators and iron absorbers, a Wavelength Shifter, and a single photo-multiplier tube. All of the PMTs are on the backside of HCal, so the hadrons interact with the end of the block opposite the PMT. As the hadrons interact



(a)



(b)

Figure 2-16: The Hadron Arm is comprised of the SBS dipole magnet and the Hadron Calorimeter. **(a)** The Hadron Arm with the two primary components labelled. **(b)** The Hadron Calorimeter (HCal).

with an HCal cell they shower at each iron layer. These showers generate photons in the scintillator layers which are transmitted through a wavelength shifter to the block's PMT via lightguides [16]. Figure 2-18 shows the physical design of a typical HCal cell. Figures 2-18a and 2-18b show the individual components in a 3D render and schematic, respectively. Figure 2-18c shows a photograph of an HCal cell lit via an LED. The LED simulates a photon that generates an event in the scintillators. The photo shows how the light travels through the wavelength shifter via lightguides until it reaches the output at the PMT.

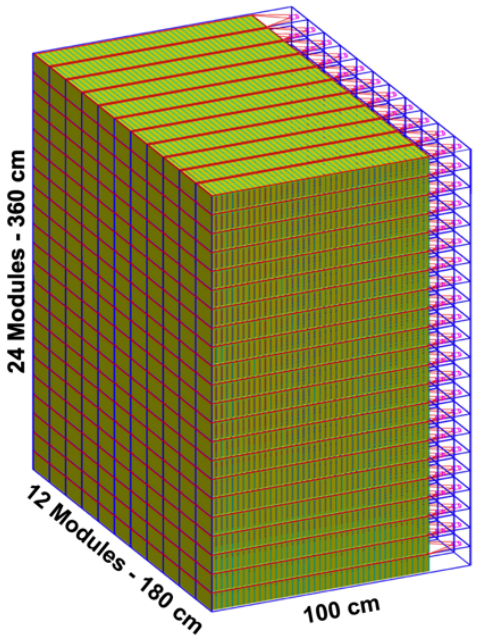


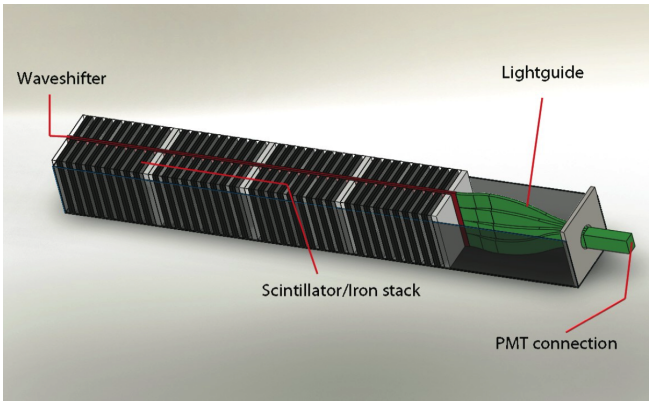
Figure 2-17: The module layout and configuration of the Hadron Calorimeter. There are 288 modules in total. These are arranged in 24 rows of 12 modules each.. Image source: [39].

The Hadron Calorimeter is the only detector on the scattering hadron arm and so it must be held to high standards of performance and operation. The detection efficiency of protons and neutrons for HCal plays a critically important role in the experimental extraction of GMn and nTPE. For our particular methods, we require detection efficiencies of at least 95% [47]. Preliminary analysis shows that the experimental detection efficiencies of HCal are within the required bounds necessary for a proper extraction.

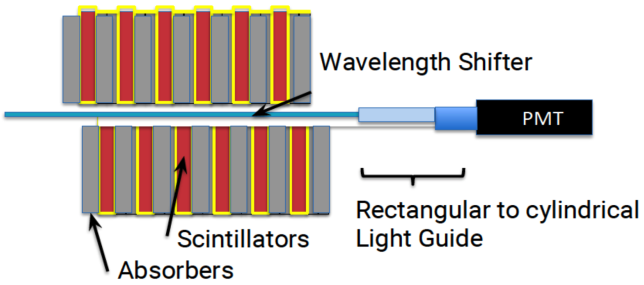
The design of HCal ensures that it matches the acceptance of the SBS magnet. It has a well-defined linear energy response, good energy resolution, the necessary 95% efficiency with trigger threshold at 25% peak signal, a spatial resolution of $\sim$7 cm RMS, a time resolution of 0.5 ns RMS, and an angular resolution of 5 mrad [97].

2.6.3 Momentum and Optics From BB and SBS

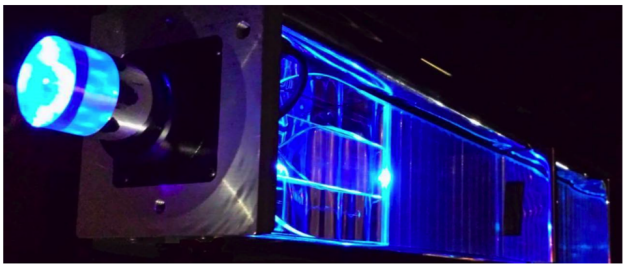
The momentum of the scattered electrons and hadrons are determined via the energy of clusters detected at each calorimeter. Before these measurements can be



(a)



(b)



(c)

Figure 2-18: The design and configuration of the individual blocks of HCal. **(a)** 3D render of the internal design of an HCal cell. Image source: [97]. **(b)** A schematic for the setup of the Wavelength Shifter, scintillators, absorbers, and PMTs within an HCal cell. Image source: [16]. **(c)** Photo of an HCal cell under test with an LED. Image source: [16].

extracted, a very important calibration of the momentum and energy must first take place. These calibrations use imaging (optics) and tracking in order to reconstruct scattering angles and correlate them with the scattered particles' momentum. This calibration of the so-called *optics* relies on the BigBite Magnet, the GEMs, a *sieve plate*, and known kinematics relations for a charged particle through a dipole (Eq.

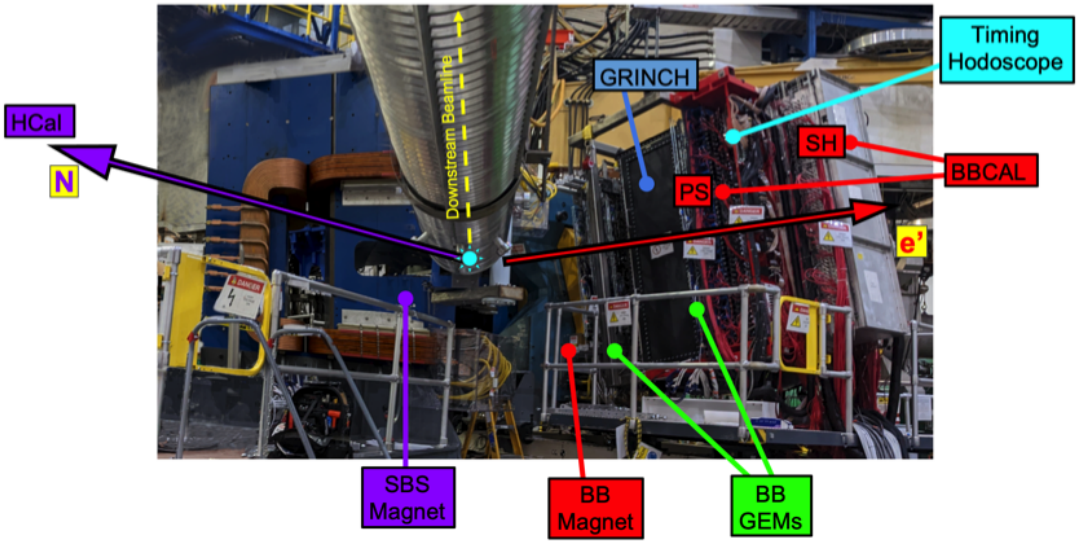


Figure 2-19: Upstream view of the SBS magnet, exit beamline, scattering chamber, and BigBite Spectrometer. HCal is not shown as it 11 m away from the SBS magnet for this configuration.

2.1).

Figure 2-19 shows the two spectrometers (minus the Hadron Calorimeter) of the nTPE and GMn experiments. The Hadron Calorimeter is not shown in the Fig. 2-19 because in this configuration the HCal was 11 meters away from the SBS magnet and could not be contained in such a close-up photograph. The sub-detectors, target location, and exit beam line are labelled in the picture. As a reminder, the Electron Arm defines the q-vector, momentum p_e, trajectory, reaction vertex, and trigger time correlation. The Hadron Arm measures the proton (neutron) momentum p_p (p_n), hadron trajectory, and time-of-flight.

Figure 2-20 shows a schematic of the BigBite magnet and the deflection effect it has on a charged particle passing through it. As a charged particle passes through the yoke of the BigBite magnet, it undergoes a deflection due to the field of the dipole. This deflection is a function of the field integral and the momentum of the scattered electron. The relation for this interaction is:

$$p_e\theta_e \approx 0.3 \int B \cdot dl \tag{2.1}$$

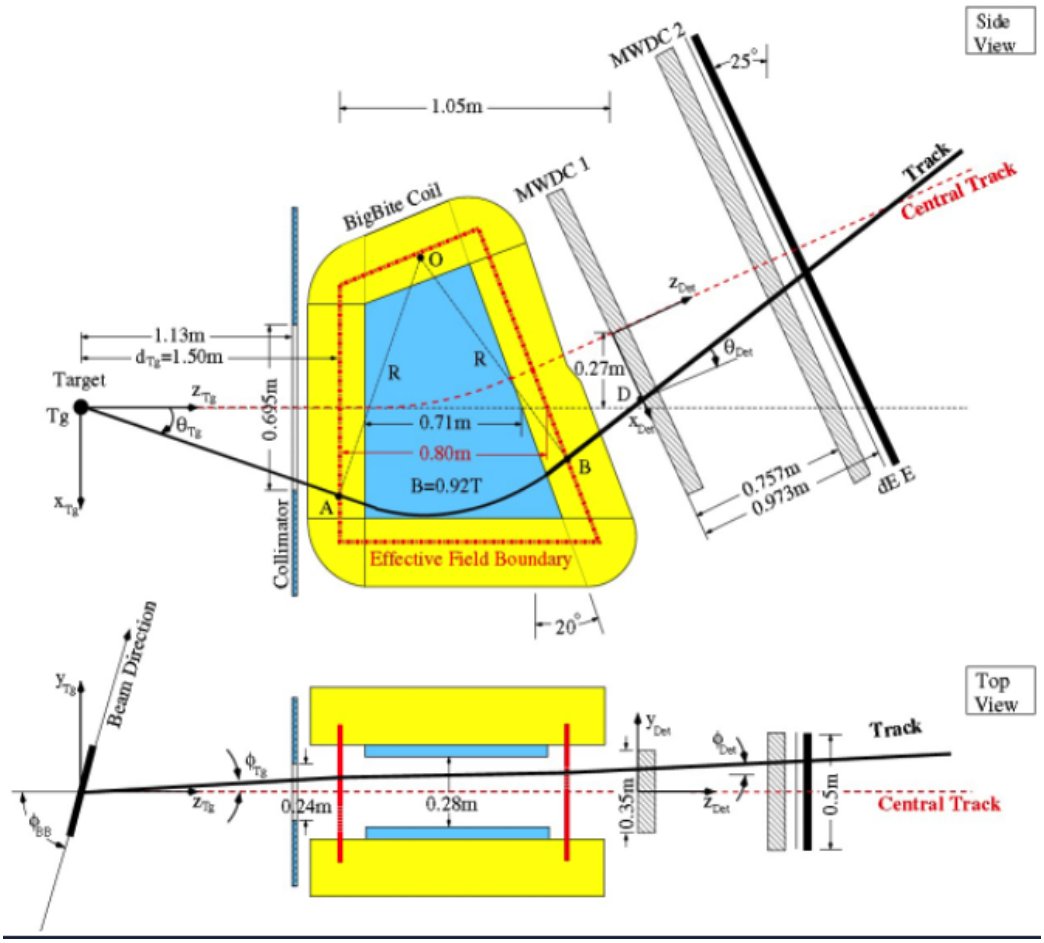
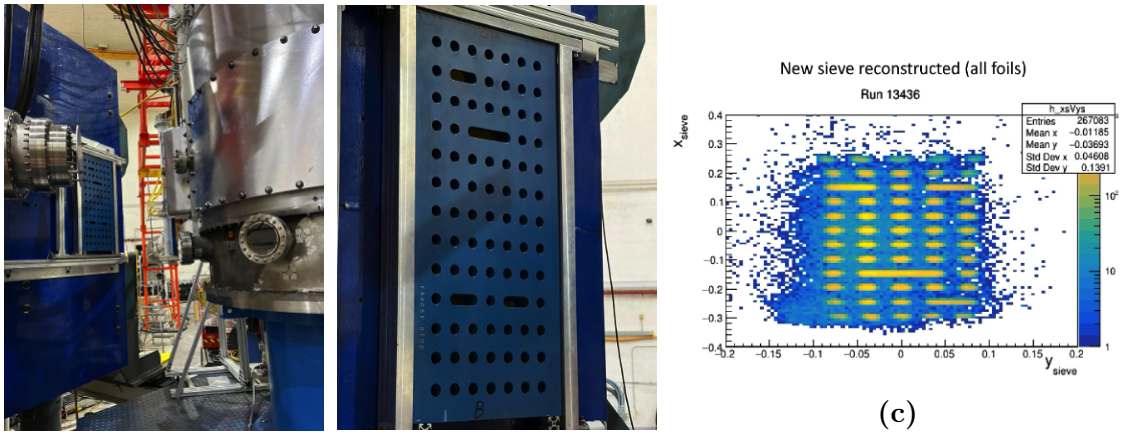


Figure 2-20: The BigBite magnet deflects a charged particle (scattered electron) as it passes through the yoke. The deflected trajectory and related parameters are shown. Image source: [94].

where p_e is the momentum of the scattered electron, θ_e is the scattered electron's angle of deflection through the magnetic field, and the RHS of the equation is the scaled field integral for the BigBite dipole. The BigBite magnet is adjustable and can be set to very precise settings via control of its supply current. From this, we can extract the right-hand-side of Eq. 2.1. Then, we just need scattered electron's deflected angle through the magnet.

The use of the term "optics" makes perfect sense when we consider how the momentum is calibrated using GEM tracking reconstruction and a sieve plate. In order to determine an accurate deflection angle, a precise positioning and trajectory of the scattered electron needs to be determined. A large steel plate known as a *sieve*



(a) **(b)** **(c)**

Figure 2-21: Sieve plate for optics and momentum calibrations. **(a)** The sieve plate is just beyond the exit port of the target scattering chamber on the Electron Arm. **(b)** Close-up photo of the sieve plate. **(c)** Hitmap on the sieve plate showing where scattered electrons have landed. Image source: [102].

plate is used for this. The plate is shown in Figures 2-21a and 2-21b. The sieve plate has a well-defined geometry of holes and slits across its surface. The quality of reconstructed GEM hitmaps is compared against the known hole and slit configuration of the sieve plate. Figure 2-21c shows a reconstructed hitmap plot of the sieve plate during optics calibration. Since the precise location and sizes of the holes and slots of the sieve plate are known, this hitmap can be used to determine an absolute position and orientation of the GEM detectors and their relative positions within the Hall. Once this has been established, the GEM and experimental coordinate systems can be calibrated to the ideal optics coordinate system [102].

Figure 2-22 shows a technical drawing of the sieve plate on the left with the corresponding reconstructed hitmap on the right. The orange circles on the figures show the location of the same hole on the sieve plate. This indicates a proper reconstruction and calibration of the GEM/internal coordinate systems. Once this is established, the momentum calibration can be determined by scanning various scattering angles with known BigBite fields and applying the relation of Eq. 2.1 in order to extract the scattered electron momentum, p_e.

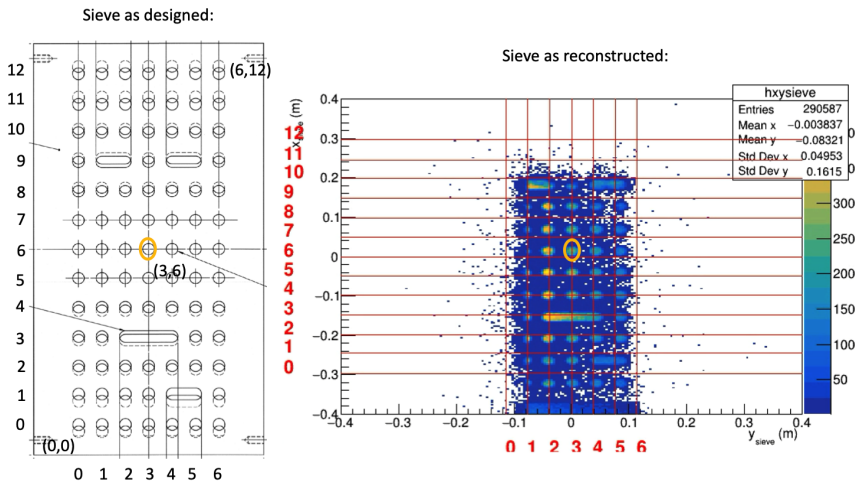


Figure 2-22: Sieve plate dimensional drawing and corresponding GEM hitmap. The orange circle indicates the same hole in the sieve slit in both figures. Image source: [100]

2.6.4 GEM Detectors

The GEM detectors are the particle trajectory and tracking detectors of the experiment. All of the GEMs for nTPE and GMn were situated on the BigBite arm. In total, the experiment contained five layers of GEM detectors. Four layers were at the very front of the BB detector stack (immediately downstream of the BigBite magnet) and a single layer was positioned after GRINCH and just before the BBCal PreShower. As the scattered electron passes through the GEM detectors its position is measured. The measured particle positions on each layer can be aggregated and correlated in order to reconstruct the particle's trajectory.

Figure 2-23 provides a visualization of how tracks can be reconstructed through the GEM layers on the BB Spectrometer. Figure 2-23a shows the five separate GEM layers of the BigBite Arm. These are histograms, or *hitmaps*, where the red spots correspond to the regions with the largest populations of hits. If we consider that each of these layers are at a different position along the scattering axis, ($\hat{z}$), then we can reconstruct the total trajectory by projecting a straight line through clusters across all planes. Figure 2-23b shows the same GEM plane hitmaps as they are positioned and separated along the scattering axis. A linear regression can be formed

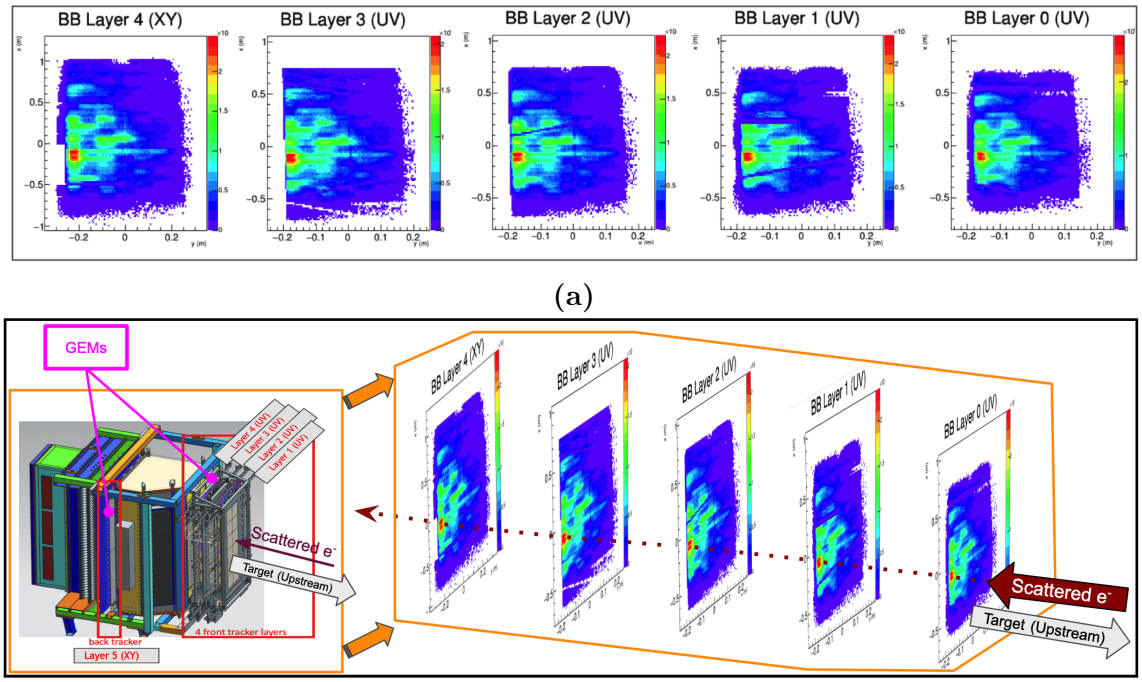


Figure 2-23: Trajectory reconstruction using layers of GEM detectors. **(a)** Hitmaps on the five GEM layers of the BB Stack. The largest population of hits correspond to the red spots. **(b)** A single track reconstructed through hits across multiple GEM layers.

from the coordinates on each plane and correlated between the detector and Hall coordinate systems to define the particle's absolute path in the scattering event. A more detailed discussion of the design and working principles of these GEMs can be found in Chapter 3.

2.6.5 GRINCH

As mentioned in Sec. 2.6.1, it is common practice to apply a BBCal PreShower cut of 0.150 MeV in order to reject pions. This can be effective but is slightly crude. The primary means for particle identification between pions and electrons lies with the GRINCH detector. The **G**as **R**ing **I**magi**N**g **CH**erenkov, or GRINCH for short, is a heavy-gas Cherenkov threshold detector. Figure 2-24 shows the position of the GRINCH within the BigBite detector stack. It sits between the four Front Tracker GEMs and the Rear Tracker GEM layer. In the right photo of Fig. 2-24, the GRINCH

is identifiable as the large black rectangle in the center of the picture.

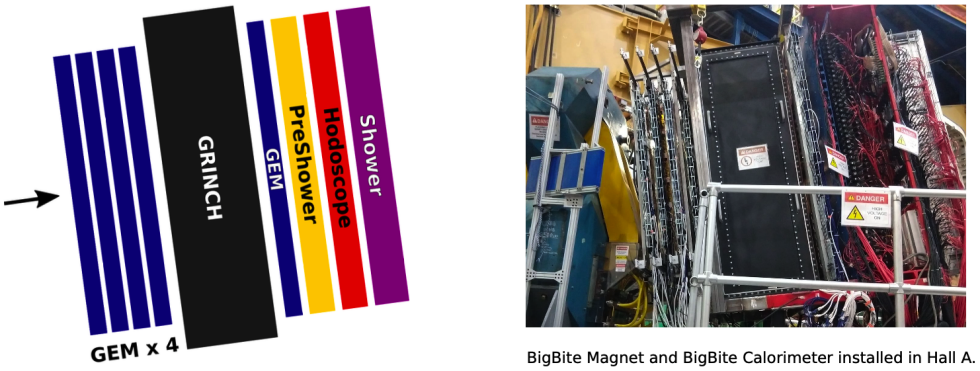


Figure 2-24: The GRINCH detector on the BigBite detector stack. It sits between the front and rear GEM trackers. Image source: [112]

When a particle travels through a transparent medium with a velocity greater than the speed of light in that medium, it emits photons through what is known as *Cherenkov Radiation*. The medium in which this happens may be any solid or liquid as long as it is transparent. Given their smaller masses, electrons have greater velocity than a pion at the same momentum. Recall that the index of refraction is related to the speed of light in a vacuum and the speed of light through a particular medium. Knowing this, a medium can be selected such that its index of refraction allows for particles to achieve speeds greater than light in that medium. When a particle exceeds the speed of light it emits Cherenkov radiation. This is similar to a sonic boom produced by objects as they cross the speed-of-sound barrier. Cherenkov radiation is, in a sense, the "optic boom" for photons and the speed-of-light barrier [41]. The internal heavy gas medium of the GRINCH is optimized such that Cherenkov radiation occurs only for electrons and not for pions. Therefore, electrons would emit photons and pions would not. This then allows for a very effective way to cut pion events out of the data [81, 112].

The GRINCH detector contains an array of photo-multiplier tubes (510 in total) arranged in a honeycomb array and 4 high reflectivity cylindrical mirrors. The entire internal volume of the GRINCH is filled with C_4F_8 heavy gas. A so-called pion threshold can be applied such that particles (pions) below a certain momentum can be filtered out. Figure 2-25 shows the basic working principle of the GRINCH detector.

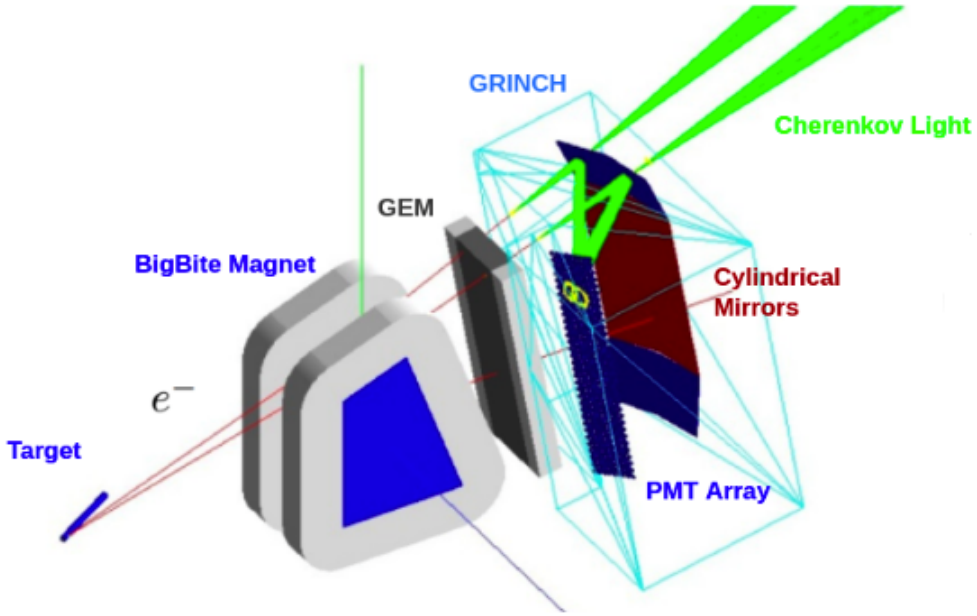


Figure 2-25: The GRINCH is filled with a heavy gas which allows for Cherenkov radiation for electrons but not for pions. The radiation occurs without affecting the scattered electron. Image source: [112]

As the scattered electrons pass through the GRINCH, it creates a ring of light due to the Cherenkov radiation light cone generated through the heavy gas. This ring of light is detected as clusters across the PMT array. These clusters correspond to electrons and provide a means to reject pions by actively selecting electron-correlated events [112, 12].

2.6.6 Timing Hodoscope

The BigBite Timing Hodoscope resides on the BigBite detector stack, between the PreShower and Shower layers of BBCal. It is an array of stacked plastic scintillators. The scintillator bars measure 600 x 25 x 25 mm and have their long dimension placed perpendicular to the scattering electron track. The total stack of scintillators contains 90 bars that each have two PMTs — one on each end [5]. A diagram of the hodoscope and picture of it on the BigBite Stack is shown in Figure 2-26.

The primary purpose of the Timing Hodoscope is to provide very precise trigger timing information for each event. As the electron traverses through a scintillator on

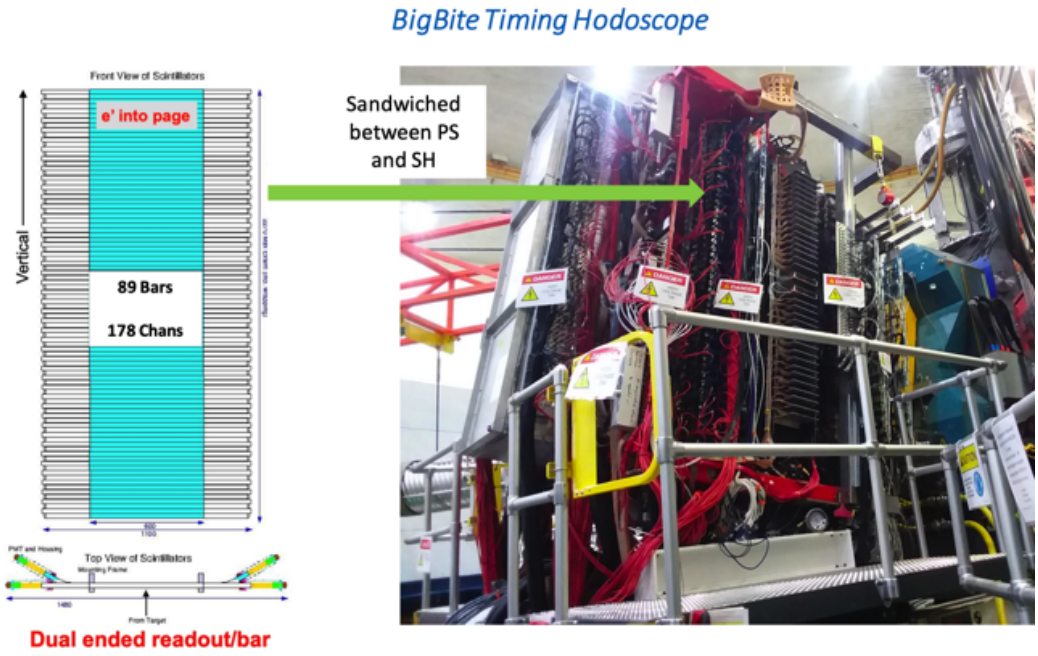


Figure 2-26: The BigBite Timing Hodoscope is on the BigBite Spectrometer between the PreShower and Shower layers of BBCal. The Timing Hodoscope provides high-precision trigger and timing information. Image source: [65]

the hodoscope, it will generate a photon that will be detected by the PMTs on either end. This measurement can be correlated with BBCal to calibrate the event trigger timing and be used to cut out random background. The timing of the electron hit on the hodoscope can also be referenced for time-of-flight measurements of the nucleons on the Hadron Arm. The fired scintillator(s) can also be used to reduce tracking and coordination combinatorics by providing an additional vertical hit position for trajectory reconstruction. A horizontal hit position measurement is possible as well. The horizontal position can be determined by analyzing the time difference between the two PMTs on a single scintillator. Given the known length of the scintillator bar, the speed of light in the plastic scintillating material, and the timing resolution of the PMTs ($\sim$100 ps), the horizontal position of the interacting electron can be calculated [54].

2.7 GEM Data-Acquisition System

A GEM detector only has the ability to "detect" if it is coupled with a *data-acquisition system* (DAQ system or simply DAQ). A DAQ system is a collection of hardware and software components which work together to allow for the measurement and/or control of a physical device, or set of devices. In brief, it converts physical stimulus (analog signal) into digital form (digital signal). The fundamental components and basic flow of our typical GEM DAQ setup is shown in Figure 2-27. The first component to receive the analog signals from the GEMs are the *APV25 Readout Cards* (APVs). The cards are *triggered* to collect the raw signals due to input from a Trigger Logic configuration. The signals from the APVs then interface through a series of digitizers, multiplexers, and relays (MPDs, VTPs, TIs, Network Interface, etc.) until they are collected in a storage device (hard drive, server, tape, farm, etc.) for decoding and analysis. The primary components of this system are discussed in the subsequent sub-sections.

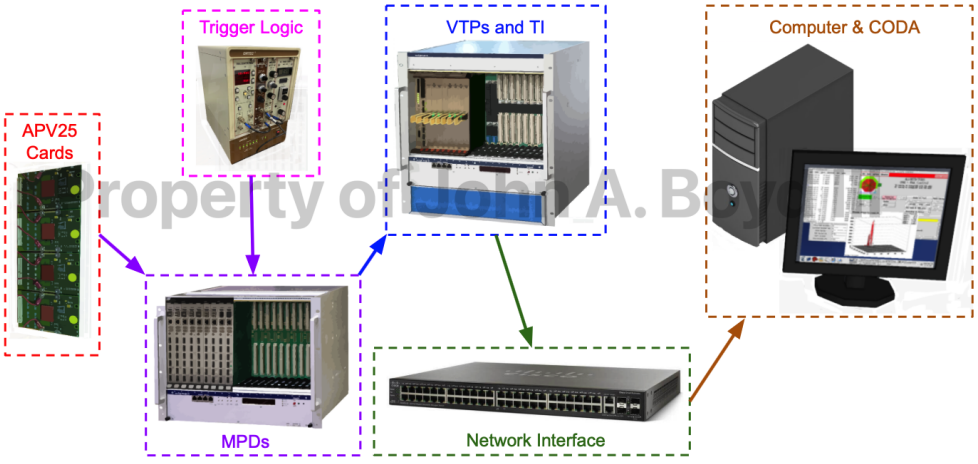


Figure 2-27: A typical GEM Data-acquisition system. These are the fundamental components needed for a basic GEM DAQ. A simplified signal flow stream is shown as well.

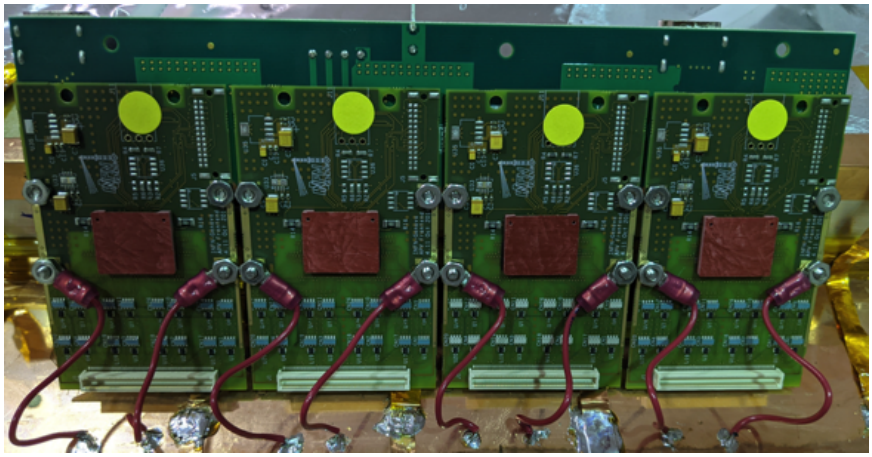


Figure 2-28: A bank of four APVs as used on the UVa GEMs. The Panasonic connectors which interface the APV25 cards to the Readout Board can be identified as the white rectangular pieces near the bottom of the picture.

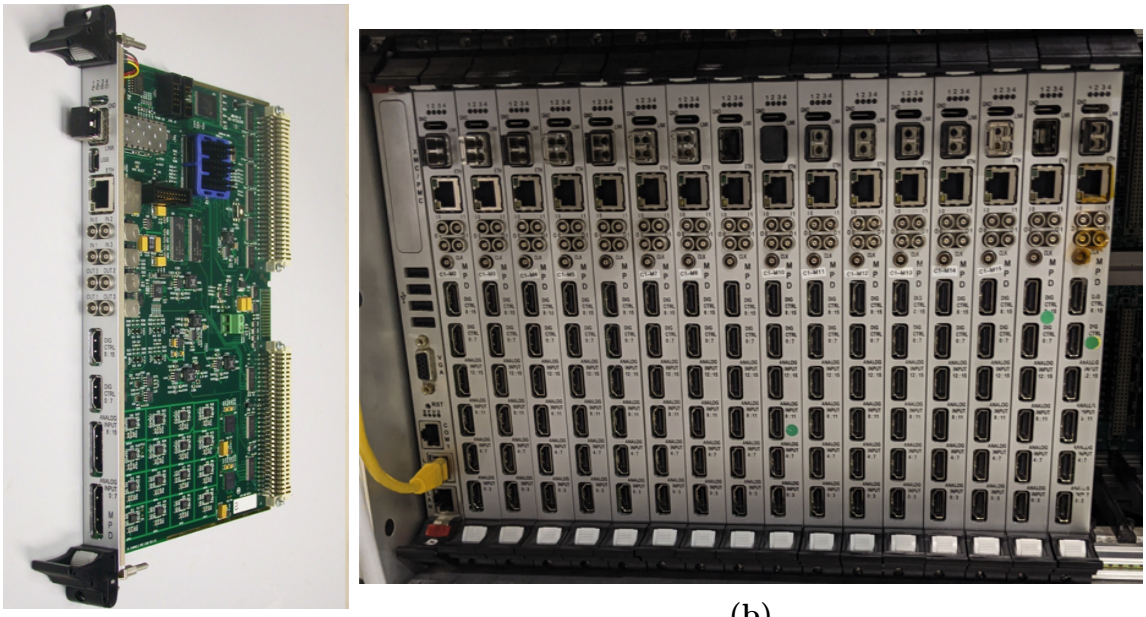
2.7.1 APV25 Readout Card

The APV25 Readout Cards are the primary analog signal readout hardware used on our current line of UVa GEMs. A photo of four APVs is shown in 2-28. To reduce signal noise, the cards were grounded to the Readout Board via the red wires shown in the picture. Each APV card has a 128-pin Panasonic connector which allows it to plug in directly to the GEM Readout Board and interface with 128 individual readout strips. These can be seen on the bottom of the picture and look like a horizontal rectangle on each card. The APV25 card was originally designed as an analogue pipeline ASIC for the CMS tracker read-out system at CERN, and performs analogue pre-processing of data from the detector prior to transmitting analog signal further downstream in the DAQ [72].

The UV GEMs have a total of 7680 readout strips, 3840 for the U-strips and 3840 for the V-strips. Therefore, a total of 60 APV25 cards are required to fully read out all strips of a UVa UV GEM detector. The analogue samples are written at the LHC frequency of 40 MHz. This creates some potential issues in the case of sampling frequency mismatch. For instance, a phenomenon known as *crosstalk* may arise from a misalignment of sampling frequencies [20]. These cards are also capable of many useful auxiliary functions such as marking and queuing memory locations

for selective output, FIR filter processing, preamplifier shape function deconvolution, and analogue multiplexing [72].

2.7.2 MPDs



(a) **(b)**

Figure 2-29: JLab's Multi-Purpose Digitizers (MPDs). **(a)** A single MPD. **(b)** An array of MPDs.

The next hardware component in the DAQ chain after the APVs are the *Multi-Purpose Digitizers*, or *MPDs*. See Figure 2-29 for pictures of some typical JLab MPDs. MPDs receive the analog signals from APV cards, digitize them, and then transmit control and configuration signals downstream in the DAQ system. Each MPD can interface with a crate via a *VME* interface; however, each card also contains an optical port for high-speed data transmission. The data-transmission speeds of the optical fibers exceed that of the VME interface, so this is how we prefer to transmit data through the MPDs [95]. Each MPD receives sequencing signals from a trigger and clock system. This aids in proper alignment of events in the data stream. The connection between the APV cards (more specifically, the APV card interface boards) are completed using high-speed HDMI cables.

2.7.3 VTPs

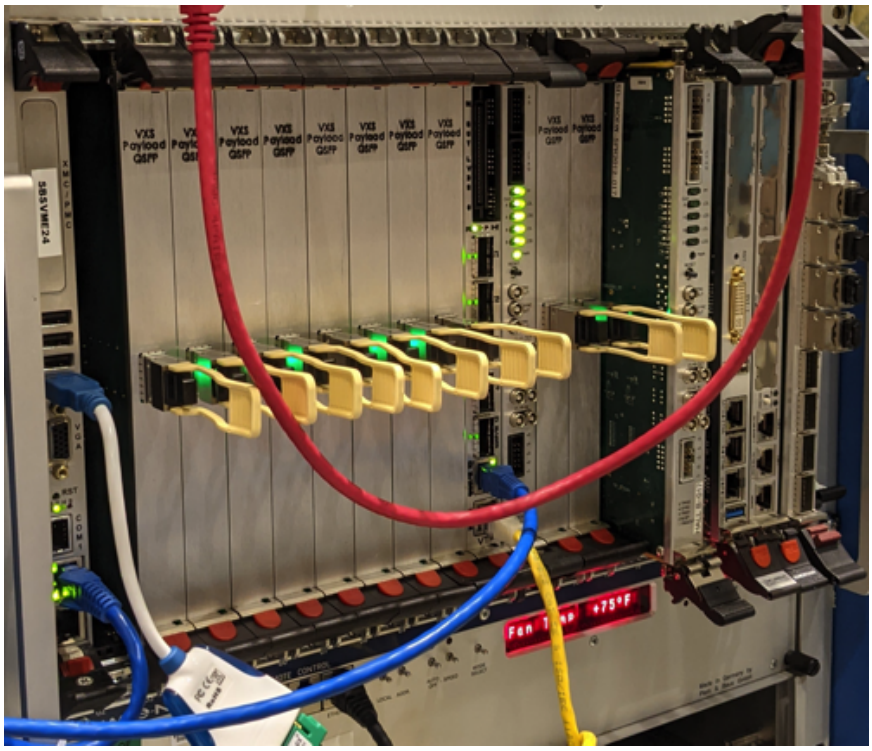


Figure 2-30: An array of VXS Trigger Processors (VTPs).

The digitized signal from the MPDs are sent to switch card modules called *VTPs*, short for *VXS Trigger Processors*. These are VXS switch cards which take on the leading role of central trigger processing as seen in typical DAQ systems. Figure 2-30 shows an array of VTPs used for the SBS Experiments at JLab. These modules are integrated downstream of the MPDs in order to further process and communicate their data more effectively. As it currently stands, MPDs are required as the front-end interface to the APV cards. The MPDs that are used rely on older, and slower data rates, therefore they are coupled with the newer and faster VTPs in order to increase their capabilities. Not only do the VTPs have much faster onboard processing and data-transfer rates, but they also provide increased FPGA resources for data processing logic as well as higher event building rates, triggers, and processing diagnostics. The MPDs more or less act as the ADC conduit between the APVs and VTPs.

2.7.4 CODA Software Platform

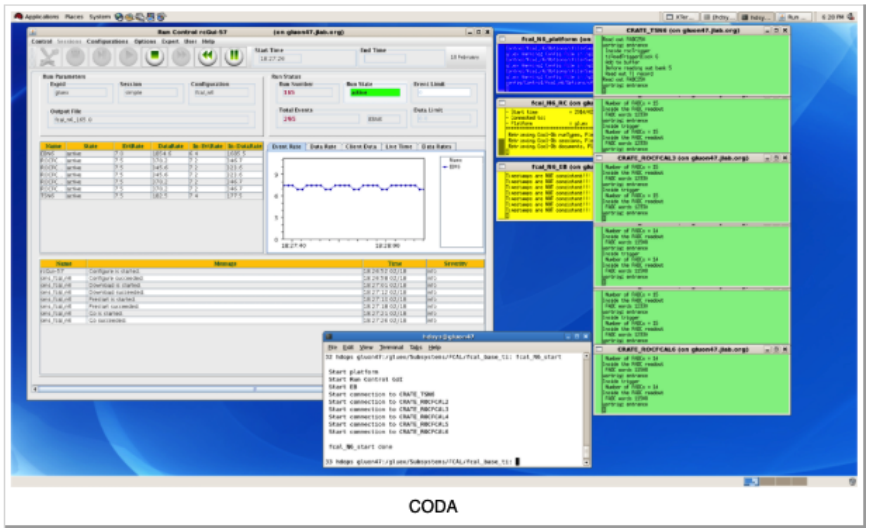


Figure 2-31: Screenshot of the CODA computer interface for data-taking.

At Jefferson Lab, all data *runs* are taken through a computer software system known as *CODA* (short for *CEBAF Online Data Acquisition*). The CODA software platform allows for the implementation of acquiring experimentally measured data from all connected detector systems and through all data-acquisition components. CODA allows for the starting and stopping of event/data runs, and stores everything neatly in a ROOT file correlated to a particular run number and/or configuration [79]. A screenshot of the CODA software package is shown in Figure 2-31.

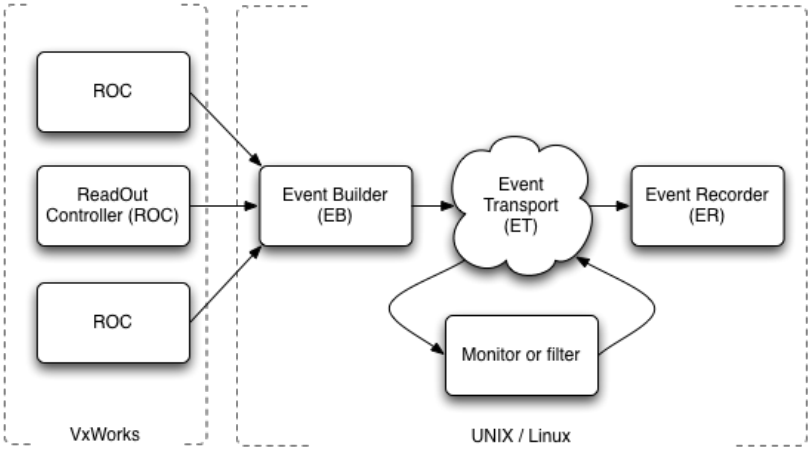


Figure 2-32: Schematic of the basic flow scheme for a simple CODA implementation.

Figure 2-32 shows a typical flow through a basic implementation of CODA. Data from a detector is digitized and pre-processed via the front-end hardware. Then, CODA manages the collection, tagging, ordering, hierarchy, and structure as it compiles the data into single ROOT files for subsequent parsing and analysis. CODA is used on both very small scale and extremely large scale systems alike. It has been implemented for small projects such as reading out a single GEM chamber on a table top as well as full, large-scale high-energy in-beam experiments (this experiment for instance). It is maintained and operated by faculty and staff at Jefferson Lab and is the most powerful tool in the data-acquisition system.

Chapter 3

Gaseous Electron Multipliers

The gas electron multiplier (GEM) is a type of gaseous ionization detector for charged particles. GEMs were introduced in 1997 by Fabio Sauli, who had originally developed them for particle physics experiments [114]. GEM detectors are in the fast radiation family of detectors and offer multiple advantages when compared to other similar-use devices such as photo multipliers (PMTs), silicon detectors (SiDs), micromegas detectors, and others. One major advantage is that GEM detector technology is relatively less expensive than the comparable silicon-type trackers [40]. When compared to similar detectors, GEMs also have the ability to sustain relatively higher rates with a better capability to effectively cover large areas while providing greater spatial resolution [66].

Figure 3-1 shows the basic operation of a GEM detector and a single ionizing particle that avalanches until it can be collected on a readout plane. A GEM detector is filled with an ionizing gas and consists of a drift cathode foil (entrance plane), at least one GEM foil, and a readout board (exiting plane). Large electric potentials are applied across each of the internal GEM foil layers. Each GEM foil within the detector provides an amplification gain of about 20. When a charged particle enters the detector, it ionizes in the gas and emits further charged particles which propagate and accelerate through the electric fields in the drift and hole regions of the GEM, creating a successive chain of ionizations (an avalanche) through the layers of the detector, until the amplified grouping of charged particles from the single initial ionization is

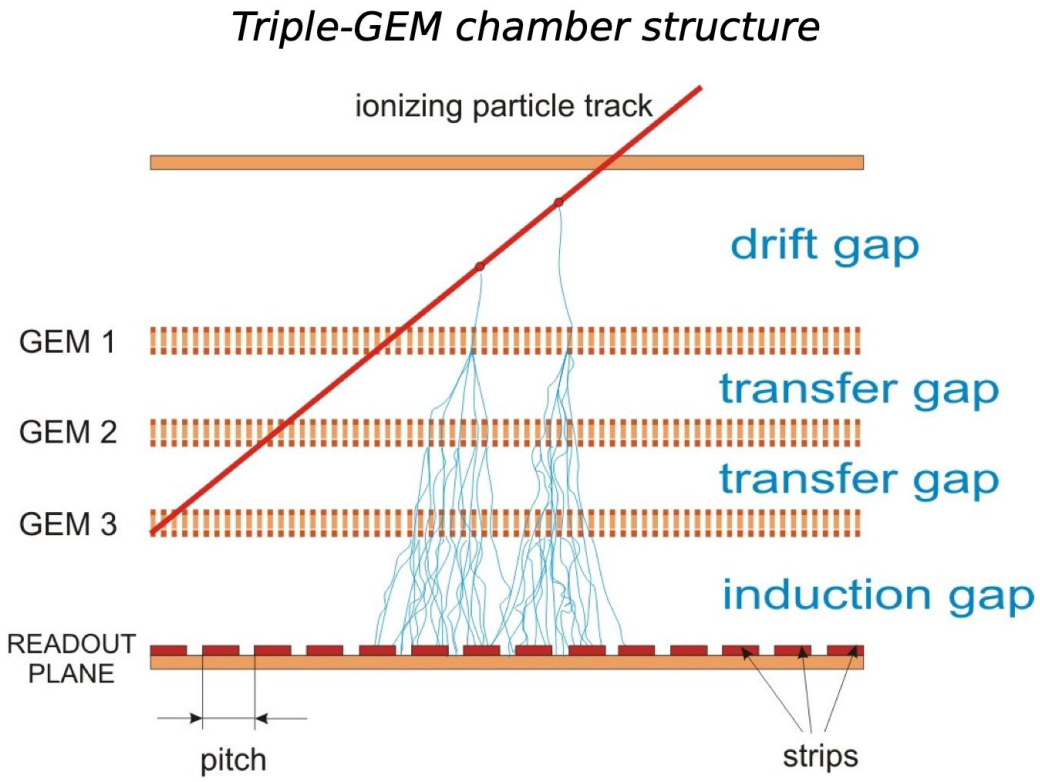


Figure 3-1: A single ionizing particle develops an avalanche through the layers of a Triple-GEM and the amplified signal reaches the readout plane. Image source: [15].

detected on an electronic readout board within the detector.

3.1 An Overview of GEM Detectors

3.1.1 Anatomy of a GEM Foil

A GEM foil is a thin polyimide foil (50 μm thick in our case) that is coated with a few ($\sim$5) microns of copper on both sides [115]. Figure 3-2 shows the physical layout and geometry of a typical GEM foil. It is important that the polyimide layer be of a material that has a low outgassing rate, because in the detector environment, conditions need to maintain stability, and outgassing of the polyimide during operation can negatively affect performance. Additionally, the polyimide layer also provides advantages from its high temperature resistance [74].

The polyimide layer insulates the two metal sides which then act as anode and

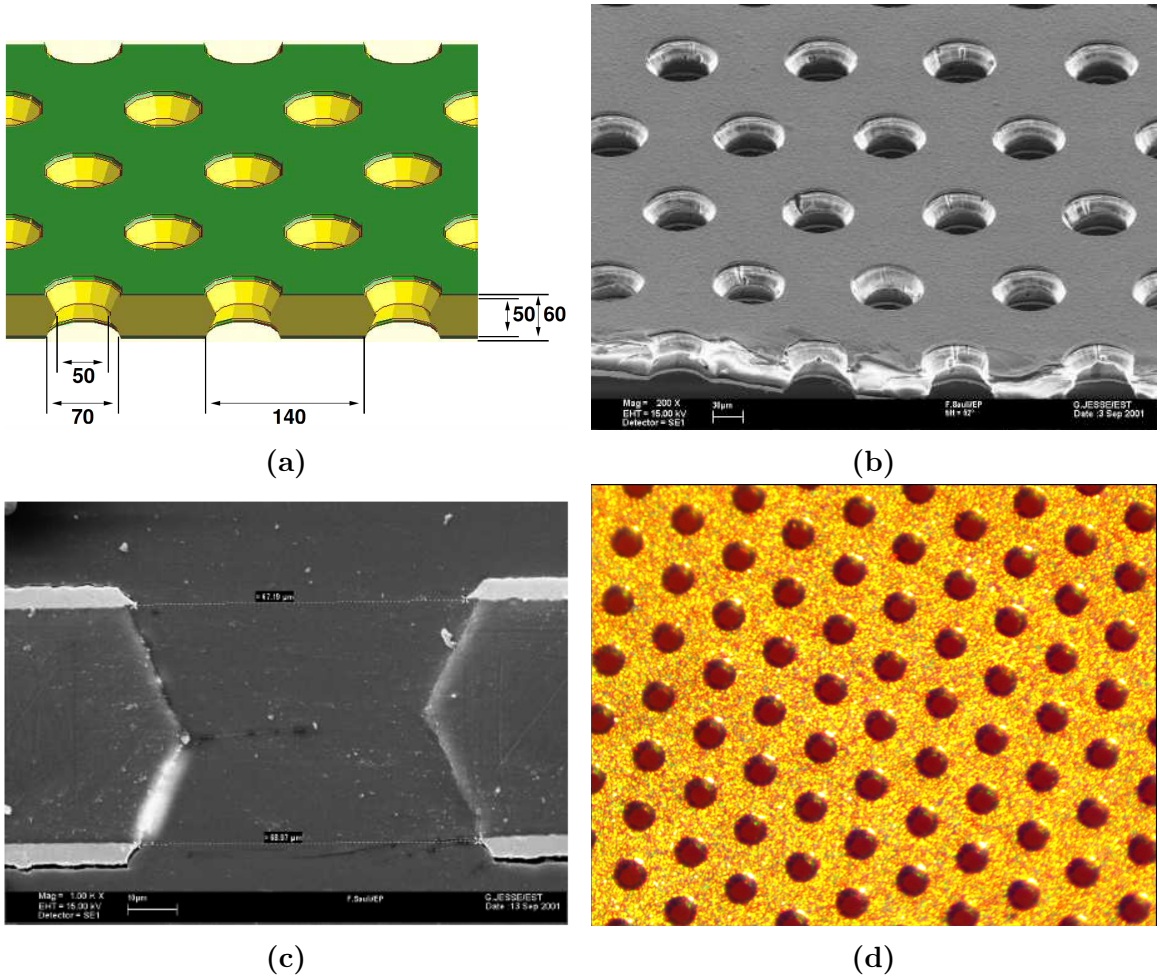


Figure 3-2: GEM foil geometry. **(a)** Computer-generated drawing showing typical foil dimensions. Image source: [136]. **(b)** Electron microscope image of perforation matrix. Image source: [115]. **(c)** Cross-section of single perforation. Image source: [110]. **(d)** Optical microscope image of a UVa GEM foil at 4.5x magnification.

cathode when potential is applied across them. The total composite of the polyimide foil sandwiched between the two conductors is what will be referenced by the terms "GEM foil" or "foil".

The foil is perforated with a high density of holes which are typically 70 μm in major diameter on a 140 μm pitch [115]. Due to the (photo)lithography manufacturing process, the perforations on the GEM foils are not purely cylindrical in profile, but rather have an hourglass-like shape (double conical), with a typical minor diameter of 50 μm and major diameter of 70 μm (see Figure 3-2c) [53].

3.1.2 Anatomy of a GEM Chamber

Single-foil GEM detectors

In its most basic form, a GEM detector, or chamber, consists of a (drift) cathode at the entry plane of the detector, a GEM foil, and an electronics readout on the exit plane (Figure 3-3). This configuration consisting of just one foil is often aptly referred to as a "single foil GEM detector", or even simply a "single GEM". The volume within the active area of the GEM detector is occupied by an ionizing gas. The gas is required for the avalanche and gain mechanisms which occur during the operating process of the GEM detector. The specifics of these details are explained in Section 3.2.4.

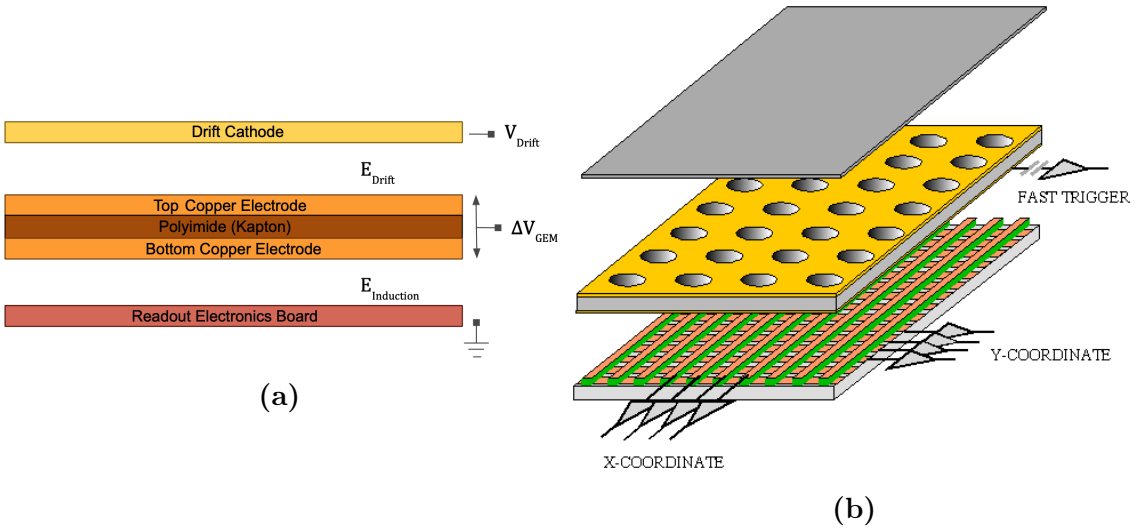


Figure 3-3: "Single GEM" detector. **(a)** Drawing of the various layers in a Single GEM configuration. **(b)** 3D drawing of a Single GEM with drift cathode, GEM foil, and an X-Y readout plane. Image source: [115].

Voltage is applied to each of the layers in the detector – the drift cathode and each electrode of the GEM foil – with the readout board referenced as ground. With this applied voltage, electric fields arise in the holes of the GEM foils as well as the surrounding regions of the foil. In a Single GEM, these surrounding regions are the Drift region between the drift cathode and the top GEM electrode, and the Induction region between the bottom GEM electrode and the readout board (See Figure 3-3a).

These regions are important to the operational interactions and physics of the detector for multiple reasons. These are discussed in further detail in Section 3.3.2. For now, it is important to note that the ions in the fields of the upper (drift) regions will drift towards the holes, where they will gain large amounts of energy and cause further ionization of the gas molecules in the GEM holes. These released electrons will also acquire sufficient amounts of energy in their release from the holes into the induction region below. Another important factor to note is that each GEM foil effectively increases the gain of an initial charge from an ionizing particle by approximately 20 [40]. This is critical, because the charge from an initial ionization alone is not enough for signal generation in the readout electronics, so the gain from the GEM layer is required for the particle's detection.

Double-foil, Triple-foil, and "Higher Order" GEMs

A single foil GEM detector acquires an approximate twenty-fold gain from the only electrode layer within. In order to increase the gain further, additional GEM foil layers can be added in series within a detector to result in a multi-layer GEM (see Figure 3-4). Each cascading layer of GEM foils provides a gain of about 20 and thus, a double-foil GEM detector has a total gain of around 400, a triple-foil GEM has gain around 8000, and so forth for each additional GEM layer in the chamber. As is shown in Figure 3-1, the avalanche of electrons produced through ionization at the top GEM foil layer (GEM 1) are now the generating particles for an avalanche on the second foil (GEM 2), and so forth as the process continues through each additional foil layer of the detector. The majority of GEMs that our group utilizes are triple-foil detectors and therefore, unless otherwise stated, discussion and reference to a GEM detector or chamber is assumed to be that of the triple-foil type.

3.1.3 The GEM Readout Board

The final internal layer of a GEM detector is the Readout Board or Readout Plane (Figure 3-4). This is where the final cascade of charge is collected and read

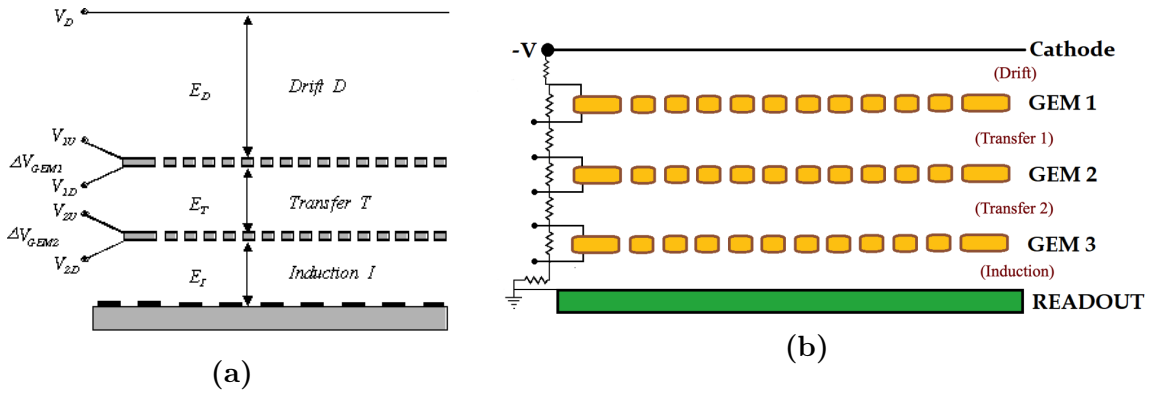
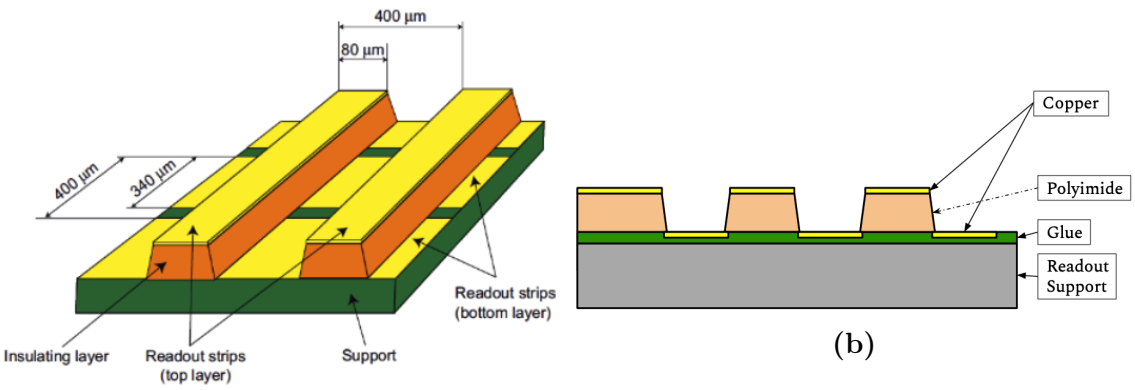


Figure 3-4: Configuration of layers for multi-layer GEMS. The regions between each foil are labelled. Each GEM detector has a Drift Region at the top and an Induction Region at the bottom. The layers between each multi-GEM layer are the Transfer Regions. **(a)** Double GEM. Image source: [115]. **(b)** Triple GEM.

out. The primary use of the readout plane is to reconstruct a coordinate for the incoming particle's incidence on the GEM detector. Each readout plane can provide two readout coordinates. These coordinates can vary based on the particular layout and construction of the readout board. Some examples include Cartesian (orthogonal), planar non-orthogonal, polar, pads, and zig-zag [115]. All of the GEMs used during the first run of the SBS Program utilized two-dimensional readouts that were oriented in either XY-Cartesian or stereo UV.

The location of an incident particle is determined by the location where the cascade of charge is collected on the readout plane. In the following discussion, refer to Figures 3-5 and 3-6.

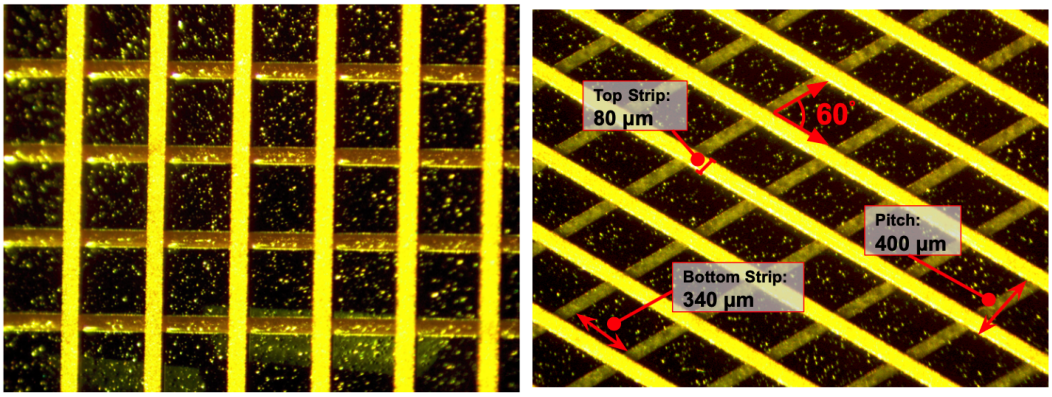
The readout plane consists of two sets of thin parallel copper strips separated and insulated by a thin layer of polyimide. This creates a top layer of strips that sits slightly higher than the bottom layer of strips, which are upon the support frame. The angle and orientation between the two planes of wires determines the readout coordinate system. In an XY Cartesian configuration, the top layer of wires lie orthogonal to the bottom plane of wires (Figure 3-6a). In a stereo angle, or UV, configuration, the angle between the two layers is non-orthogonal. The UV GEMs that we constructed had a stereo angle of 60^o between the two sets of copper strips (Figure 3-6b). In any configuration of readout board, the selection of any strip on



(a) (b)

Figure 3-5: GEM readout planes. **(a)** Isometric drawing of a 2D readout plane. The two orthogonal strips are shown atop the readout support structure below them. Image source: [85]. **(b)** 2D side-view drawing of a readout plane with its two separate sets of strips on the support frame.

the top layer and any strip on the bottom layer will provide a coordinate pair that specifies a location on the plane of the board.



(a) (b)

Figure 3-6: Two different readout plane configurations. **(a)** XY Cartesian: the top layer of strips is orthogonal to the bottom layer. A strip pair creates an XY-coordinate. **(b)** Stereo angle UV: the two sets of strip planes are separated by a stereo angle of 60^o. A strip pair corresponds to a single UV-coordinate.

When an avalanche of electrons lands on the readout board it will deposit its charge, distributed across a set of strips in the two layers. This collected charge can be measured and used to coordinate the kinematic point along the plane of the GEM detector. Figure 3-7 shows the measured ADC readout of consecutive U strips for a single coincidence event. The large peak located between APV Physical Strips

99 and 105 corresponds to the electrons transferred from the avalanche through the GEM layers. The central avalanche coordinate of this peak can be determined using a center-of-gravity calculation [114].

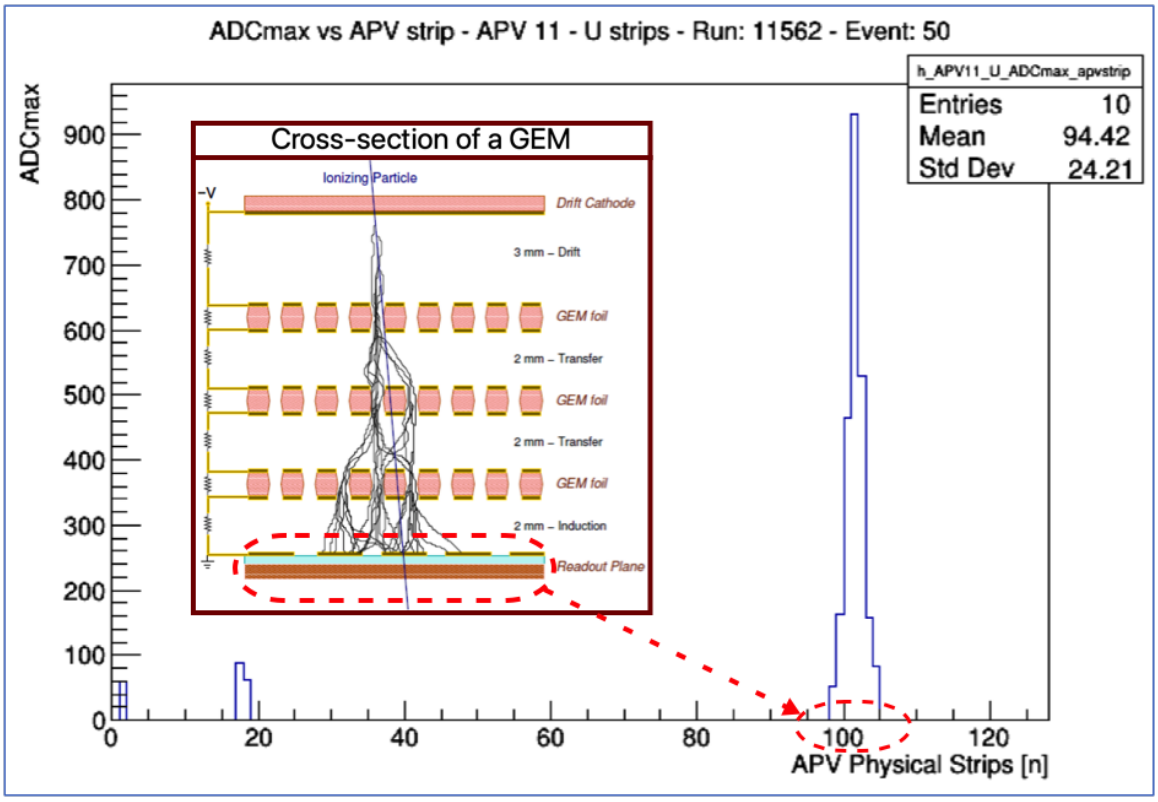


Figure 3-7: Readout from an APV showing the ADC signals read out from various strips on a GEM detector. The event of interest here is around strips 99 through 105. The center of the peak is on strip 102 with and ADCmax reading of approximately 920, in arbitrary ADC units. Sub-image source: [48].

A similar plot can be developed for the same event along the other coordinate axis, "V". The two central avalanche locations from both U and V dimensions then define the coordinate for the primary particle interaction, (U, V), on that GEM plane.

Due to the fact that the top strip layer lies above the bottom layer (they are separated by an insulating layer of polyimide, with gaps to expose the bottom layer), the two sets of strips would not get equal charge sharing if they were built to equal dimensions. If the strips on the top and bottom layers are the same size, the charge will be about 50% larger on the top layer than on the bottom [35]. This is due to more electric field lines converging on the top layer strips than the conducting strips

at the bottom of the gaps. To compensate for this, the strips on the top layer can be made narrower than the bottom strips. Equal charge sharing is a requirement for our GEMs. Figure 3-6 shows that the strips in our GEMs are configured such that the two sets of strips collect a proportionally equal amount of charge. In addition to particle coordination, the signals from the lower layer strips can also be used either as a trigger, or for energy discrimination [115].

3.2 Interactions of Charged Particles With Matter

GEM detectors are sensitive to a multitude of charged particles. This includes electrons, heavy charged particles (muons, protons, ions, etc.), and to a certain extent, photons. In order to better understand how GEMs respond to various types of particles, it is helpful to discuss the different ways that charged particles can interact with matter in general. This section will provide a short description of those interactions.

When a charged particle encounters, or passes through a medium, it can interact with that matter. This interaction can be through the strong, weak, or electromagnetic forces. In comparison, the electromagnetic force is several orders of magnitude stronger than the other forces, at the length scales relevant for particle detectors, and therefore is the most probable interaction to occur [33, 37]. Neutral particles, such as photons and neutrons, can also interact with matter. However, those interactions are quite different from that of charged particles with matter.

3.2.1 Heavy charged Particles

A charged particle is one that is surrounded by its Coulomb field and interacts with the nucleus of all atoms and orbital electrons it encounters as it passes through matter [98]. Charged particles are classified as either heavy ($m_{particle} \gg m_e$) or light ($m_{particle} \approx m_e$) [33]. The dominant processes for energy loss for heavy charged particles interacting with matter are ionization and excitation [92]. For a particle P with charge state Z, these are described by the following [37]:

$$
\begin{aligned}
Ionization &: P^Z + atom \longrightarrow atom^+ + e^- + P^Z, \\
Excitation &: P^Z + atom \longrightarrow atom^* + P^Z \\
&\quad where\ atom^* \longrightarrow\ atom + \gamma, \\
Ionization + Excitation &: P^Z + atom \longrightarrow atom^{*+} + e^- + P^Z.
\end{aligned}
\tag{3.1}
$$

When a charged particle hits the elements of a tracking system, such as a detector's gas volume, a copper layer in a GEM foil, or a drift cathode foil, some energy is lost in the process of ionization. When a charged particle passes near the electron cloud of an atom in a material, the electric field of the charged particle interacts with the electron cloud and can remove an electron, thereby ionizing the atom. Through this process, the charged particle will lose energy. The average energy loss along the flight path, $-\frac{dE}{dx}$, of a relativistic heavy charged particle, P, of mass $m_{particle} \gg m_e$, and energy E, as it passes through and interacts with matter is described by the Bethe-Bloch formula [123, 83, 55]:

$$
-\frac{dE}{dx} = \frac{K}{A}\frac{z^2 Z}{\beta^2}\left[\frac{1}{2}\ln\left(\frac{2m_e\beta^2\gamma^2 W_e^{max}}{I^2}\right) - \beta^2 - \frac{\delta(\gamma\beta)}{2} - \frac{C_e}{Z}\right]. \tag{3.2}
$$

In Equation 3.2 the following relations and terms are used:

$$
K = 4\pi N_A r_e^2 m_e c^2 \approx 0.307\ \frac{MeV}{gc^2}
$$

$$
W_e^{max} = \frac{2m_e c^2\beta^2\gamma^2}{(1 + 2\frac{\gamma m_e}{M} + (m_e)^2)} \qquad \text{Maximum energy transfer in a single collision}
$$

$$
r_e = \frac{e^2}{4\pi\epsilon_0 m_e c^2} \approx 2.8\ \text{fm} \qquad \text{Classical radius of the electron}
$$

$$
N_A \approx 6.022 \cdot 10^{23} \qquad \text{Avogadro's Number}
$$

$$
m_e \approx 511\ \text{keV} \qquad \text{Mass of the electron}
$$

$$
\beta = {}^{v}/_{c} \qquad \text{Particle velocity}
$$

$$
\gamma = {}^{1}/\sqrt{1-\beta^2} \qquad \text{Lorentz Factor}
$$

$$
z: \qquad \text{Charge of the incoming particle}
$$

M:	Mass of the incoming particle
Z:	Charge number of the medium
A:	Atomic mass of the medium
I:	Mean excitation energy of the medium
$\delta(\beta\gamma)$:	Density correction due to polarization at high energy
C_e:	Atomic shell correction at low energy

For a charged particle through a gaseous ionization detector, derivation of the Bethe-Bloch formula relies on three major assumptions [64]. The first is that when energy is transferred between the incoming particle and the medium, the direction of flight of the ionizing particle is not changed. A heavy, relativistic, charged particle only transfers a small fraction of its energy in any single collision, and so, as it travels through matter, it does so in straight paths with negligible deflection as it continuously loses energy through a large number of collisions with atomic electrons [92]. Secondly, the gas molecules of the volume are considered to be at rest. Lastly, the ionizing particle is assumed to be much heavier than an electron ($m_{particle} \gg m_e$). Figure 3-8 shows the relationship between the stopping power of copper and positive muons, plotted as a function of $\beta\gamma$ (muon momentum). The range of validity is $0.05 < \beta\gamma < 500$ for particles with $m_{particle} > m_\mu$ [49]. This range is shown in Figure 3-8 as the region between the second and third vertical grey bars. The Bethe-Bloch equation does not account for the fact that at lower velocities, heavy charged particles may capture free electrons as they traverse through the medium, thus reducing the particle's net charge and therefore the effective stopping power of the medium. The plot also shows the point for "Minimum ionization" (the minimum in the ionization energy loss curve) as $\beta\gamma \approx 3$. Particles with $\beta\gamma \approx 3$ are referred to as *minimum ionizing particles*, or *MIPs* [133]. Typically, relativistic particles have energy loss rates near this ionization minimum until radiative losses set in at higher momenta [96].

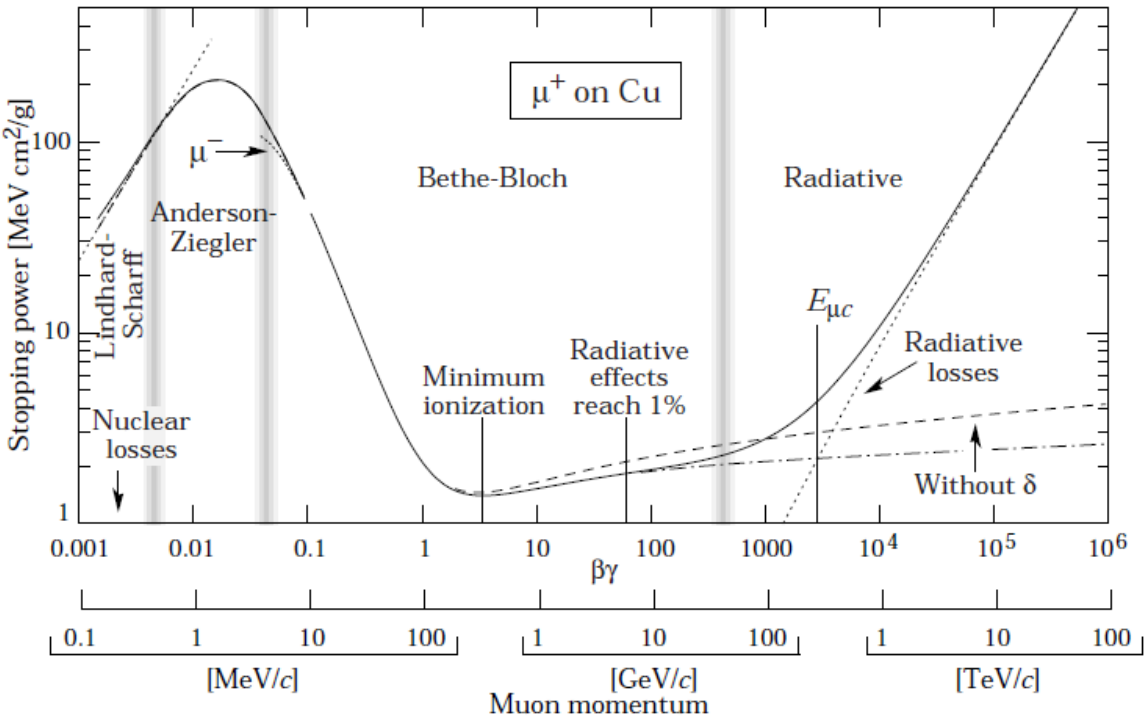


Figure 3-8: Plot of Bethe-Bloch formula for positive muons in copper as a function of muon momentum, represented here by $\beta\gamma$. Image source: [96].

3.2.2 Light Charged Particles

The two mechanisms by which light charged particles ($m_{particle} \sim m_e$) deposit energy are collision and radiative losses. As with heavy charged particles, collisional losses are through interactions with orbital electrons or atoms in other matter. Similarly, this also leads to ionization or excitation of the atom. In addition to collisional losses, these particles can emit electromagnetic radiation called *bremsstrahlung*, or "braking radiation". Charged particles which undergo a change in acceleration will always emit bremsstrahlung, and so, this process occurs as light charged particles interact with the electrostatic fields of nuclei they encounter as they pass through the detector medium [34, 133]. Heavy charged particles also experience bremsstrahlung, but this effect is much less important than in the case of light charged particles. The energy of the bremsstrahlung photon is directly related to the change in acceleration [82]. The larger the change, the more energetic the photon.

Due to the equivalence of masses of a light charged particle and an electron,

particle collisions of this class are like those of identical particles. This results in the possibility for large scattering angles. Therefore, the tracks of light charged particles can be twisted and winding, as opposed to a straight path that is expected for the heavier charged particles. This also results in acceleration changes for each drastic path deviation and therefore couples bremsstrahlung with the collisional losses [42].

3.2.3 Neutral Particles in Matter

Unlike a charged particle, a neutral particle is not surrounded by its own Coulomb electric field. However, it can still interact with orbital electrons and atoms as it penetrates matter. Examples of neutral particles which may interact include photons, neutrons, neutral kaons (K^0, $\overline{\mathrm{K}}^0$, ...), and neutrinos.

Photons are electrically neutral particles, but they directly couple with charged particles through electromagnetic interaction and can transfer their energy to charged particles. They can interact with matter in three different ways: the photoelectric effect, Compton scattering, and electron-ion pair production. In the photoelectric effect, a photoelectron is emitted upon the absorption of a photon by an electron.

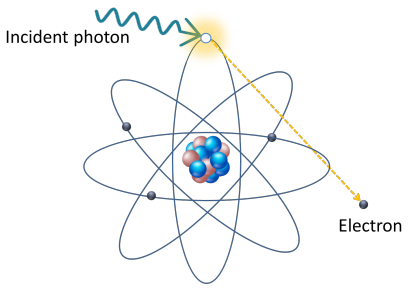


Image source: [57].

The released energy is fixed, and is $E_e = E_\gamma - E_{Binding}$. The electrons in the deepest shells (K, L, or M) are favored for this mechanism due to their proximity to the nucleus [37]. Compton scattering occurs when a photon with $E_\gamma > E_{Binding}$ elastically scatters from a free atomic electron (quasi-free electron). This scattering occurs through large deflection angles. Also, the photon only transfers some, not all,

of its energy to the electron, and therefore does not disappear. Rather, it is redshifted upon scattering [70]. Pair production can occur in the field of a nucleus or in the field of an electron. For near-nucleus pair production this requires photons with energies greater than $2m_e$ ($E_\gamma \geq 2m_e$), and for near-electron production, the condition is $E_\gamma \geq 4m_e$ [37, 70]. This is why at high energies, pair production is the leading effect of photon interactions.

3.2.4 Interactions and Gas Mixtures

The operating gas that fills a GEM chamber is a crucial ingredient for the proper function and performance of the detector. The gas mixture inside the detector provides the primary ionizations as it interacts with high energy electrons, as well as further ionization and avalanches. It is therefore important to select a gas that can provide optimal transport capabilities, ample signal production, and has sufficient gain capabilities. Noble gases require relatively low ionization energies, are chemically inert, and have only a few degrees of freedom. These attributes make them the ideal choice of gas for GEM detectors [25]. Of the noble gases, argon has become the preferred choice as the fill gas. In addition to possessing all of the positive features above, argon is a low cost choice that also has a relatively small electron attachment coefficient. This minimizes the amount of interaction interference to free electrons, thereby guaranteeing that they will reach the gain region of the detector [38, 64].

Even though pure argon gas is the theoretical preference, there are some fundamental drawbacks to using a pure gas in a GEM [25]. As stated previously, possible energy losses for charged particles include ionization and excitation that arise from the collisions between the gas molecules and electrons. Upon excitation, it is possible that a secondary electron will not be produced to further the ionization and avalanche processes. In these cases, the excited atom emits a photon instead (either visible or ultraviolet) and decays to its ground state [38]. This photon may propagate on to create additional ionization elsewhere in the detector and create spurious events outside of the actual event hit location [63]. In order to suppress these unwanted decays/emissions, the argon can be mixed with a polyatomic gas such as

carbon dioxide (CO_2) or Eethane (CH_4). The added polyatomic gas, commonly referred to as the *quencher* or *stabilizer*, aids by absorbing the photons energy through rotational and vibrational modes, effectively suppressing the spurious photon-induced effects [64, 125]. In our case, it is CO_2 that is added to the argon.

Stabilizing additives can be mixed in at various ratios. The most common ratio is 70% argon and 30% CO_2 by volume, but any concentrations can be used (80/20 Ar/CO_2 is also common). The energy loss rate of the interacting particle through the gas can now be re-defined in terms of a gas mixture. If we consider the corresponding weights by volume, w_i, of the mixed gases we find the average energy loss with a mixture of gasses to be [130]:

$$\left\langle \frac{dE}{dx} \right\rangle = \sum w_i \left\langle \frac{dE}{dx} \right\rangle_i \tag{3.3}$$

Increasing the concentration of CO_2 increases the operating voltage of the detector and lowers the measurable count rate [63]. Increasing the amount of CO_2 lowers the concentration of argon, and thus lowers the probability of primary ionization and decreases the amplification gain. In our GEMs we use 70/30 Ar/CO_2. This ratio provides great protection from photon discharge while also not diminishing the gain below a usable threshold [125]. This relatively higher concentration of CO_2 does however increase the operating voltage ($\sim$3800V), but not beyond safe and readily available levels.

As ionization and excitation occur, the argon is depleted of its active ability to further ionize or excite. Also, during standard GEM operation, any impurities that form within the gas volume need to be flushed out[1]. Therefore, it is important for the gas to circulate through the detector as opposed to remaining static. This is achieved by having a designated gas input and output that allows for through-flow. The gas could either travel once-through the detector or be recirculated in order to lower gas consumption. In the once-through configuration, the gas typically exits

[1]Even under standard operating conditions, discharges may occur throughout the internal gas volume of a GEM detector. These discharges can manifest physical debris inside the detector which could be catastrophic. Also, gases or vapors could be released (outgassed) from various detector components. These outgassed impurities affect the gas mixture and need to be removed.

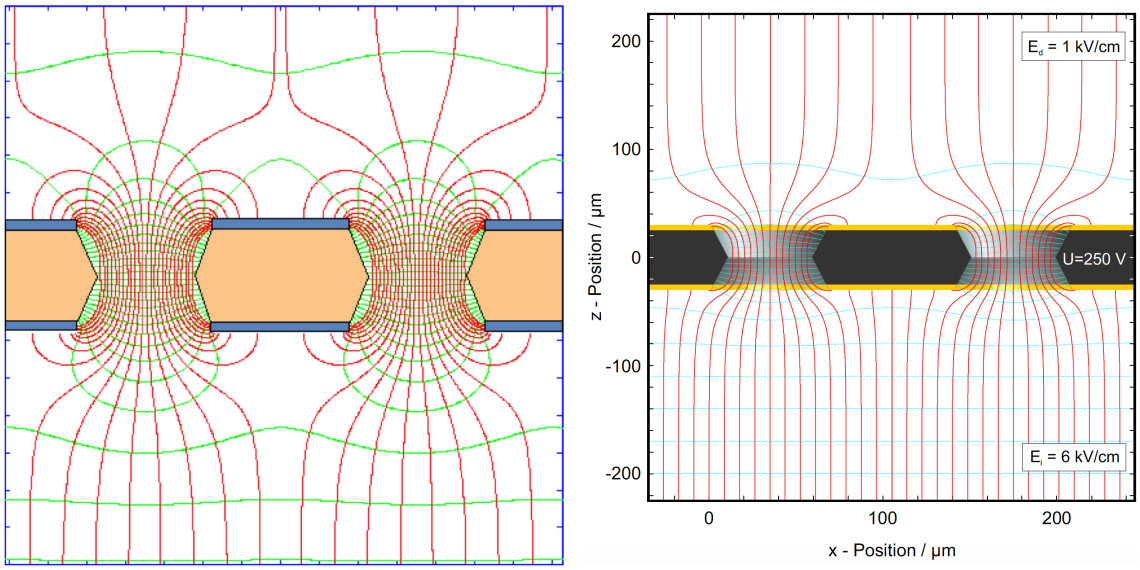
to atmosphere whereas with the recirculation method, the exiting gas is re-purified and then injected back into the pre-input mixer (if used). Regardless of whether the gas system is once-through or recirculation, a typical operation flow rate is around a few detector volumes per hour. GEM gain increases as the gas flow rate increases; however, higher flow rates increase the internal gas pressure of the detector and can cause serious damage to the GEM components or detector structure [63]. Therefore, there is a careful balance that needs to be maintained between gas flow rate, required GEM gain, and the structural integrity of the detector at the given internal gas pressure.

3.3 Ionization, Gain, and Charge Transfer

3.3.1 Electric Fields in GEM Foil Holes

When high voltage (∼400 V) is applied across a foil's anode and cathode electrodes (the two copper coating layers which sandwich the polyimide insulator between them), a very large electric field is generated within the small hole regions across the foil. The electric field lines are shown in Figure 3-9. Due to the strong electrostatic fields inside these hole regions, any electrons and ions that enter this area will accelerate and gain energy. The detector volume is populated with a mixture of argon and carbon dioxide and so, the electrons will collide with the gas atoms and create an avalanche of electrons through successive ionization collisions.

The typical diameter of a single perforation in a GEM foil is approximately 70 microns (see Figure 3-2). Thus, given standard techniques and methods, direct measurement of the generated fields in these regions is not feasible. Therefore, we must rely on calculations and simulations. The calculations for the field of the hole regions depends on the particular geometry of the holes, the voltage across the foil electrodes, and the external electric fields of the foil (in the drift regions above and below the GEM foil). If we consider an integral of the electric field over the area of the GEM hole center, we can acquire the mean electric field inside the hole region of the foil as



(a) **(b)**

Figure 3-9: Computed electric fields within the perforated regions of a GEM foil. **(a)** General high voltage application. Image source: [115]. **(b)** Fields for a voltage of U = 250V across electrodes. Image sources: [126, 117].

[76]:

$$E_{\text{hole}} = \frac{\int_A \vec{E} \cdot \vec{n} \quad d^2r}{\int_A d^2r}, \tag{3.4}$$

where A is the area over the center of the hole, and $\vec{n}$ is the corresponding normal vector. Extensive work into and beyond these calculations and simulations have been carried out and provided important reference. Calculations arising from 3.4 result in a linear parameterization of the electric field for the hole region of a GEM as [126]:

$$E_{\text{hole}} = a \cdot U_{\text{GEM}} + b \cdot (E_D + E_I), \tag{3.5}$$

where a and b are two geometry-dependent parameters, and the subscripts D and I denote the electric fields in the drift and induction regions, respectively.

Figures 3-9 and 3-10 show computer-generated fields for the hole regions of GEM foils. Figure 3-10 shows the calculation of the electric field along a perpendicular axis to the GEM hole orientation for various diameters. The computed electric fields correspond to GEM foils with 5 μm thick copper electrodes, 50 μm thick polyimide

between electrodes, and 500 V applied across the electrodes [13]. The figure shows hole diameters ranging from 10 μm to 100 μm. For our GEMs, with a diameter of approximately 70 μm, we see that the electric field at the center of the hole is approximately 65 kV/cm.

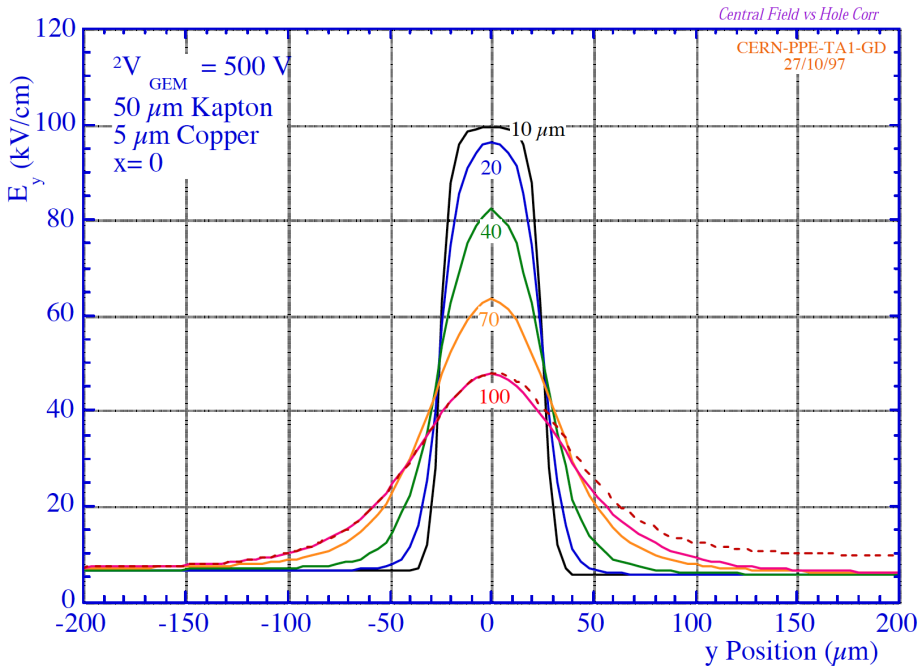


Figure 3-10: Electric field computed along a line through the center of the holes, for different hole diameters. The position y = 0 corresponds to the middle of the GEM foil hole. Image source: [13].

3.3.2 Ionization in the Gas Detector

When a charged particle passes through a medium, ionization occurs and charge is released [124]. This ionization occurs because the charged particle has deposited more energy than the average ionization energy of the gas mixture. In gas volumes, such as the interior of a GEM chamber, this results in the release of electron-ion pairs. Once released in a gas volume where an electric field is present, these electrons and positive gas atoms of the electron-ion pair are free to accelerate towards the anode and the cathode, respectively. The free electrons accelerate in the applied field and gain kinetic energy between collisions with the neutral gas atoms. As this energy becomes greater than the ionization energy of the atoms in the gas, secondary

ionizations occur between the neutral gas atoms and the electrons. This process repeats as the ionization chain, or avalanche, grows as it moves through the GEM foils and the drift regions. Now, the initial ionization from the incident charged particle on the detector – which alone is not enough to be detected by the readout electronics – avalanches through the detector through successive ionization chains, until is has gained sufficient amplification such that it can deposit measurable charge on the readout board [38]. This measurable charge translates as an induced voltage pulse in the readout electronics that can then be correlated to the incident particle event. This process can be visualized in Figure 3-1.

The total number of primary ionization electrons, n_{tot}, is given by [131]:

$$n_{\text{tot}} = \frac{\Delta E}{W_I}, \tag{3.6}$$

where ΔE is the total energy deposited by the detected particle, and W_I is the ionization energy. The ionization energies for argon and carbon dioxide are $W_{I,\,\text{Argon}} = 26\ eV$ and $W_{I,\,CO_2} = 33\ eV$, respectively [64].

If we consider an electron traveling in the direction of an applied field in an ionizing gas, along with the average number of ionizing collisions it makes, α_T, for some distance travelled, x, we can develop a relation for the number of these electrons, n_x, which would make it to the anode as:

$$n_x = n_0 e^{\alpha x}, \tag{3.7}$$

where n_x denotes the number of electrons created along some distance x, n_0 is the initial number of primary electrons, and α is *Townsend's first ionization coefficient* [99].

3.3.3 Optical and Electrical Transparency

GEM foils are populated with a high density of perforations which provide the working mechanisms for amplification, but also affect the overall capabilities of the detector for better or for worse. The various physical parameters defined by the

GEM foil hole design in conjunction with the fields applied therein are paramount when considering the efficiencies of the GEM. For instance, we can consider the ratio of perforations to the entire surface area (or rather, the non-perforated surface area remaining) and how that relation contributes to the effectiveness of the foil's gain. Also, the two metal sides of the GEM foil that are separated by the insulating polyimide are held at two different potentials. The ratio of these two differing applied electric fields will play an important part in the effectiveness of ionization, avalanche propagation, and overall performance.

The optical transparency of a GEM foil relates to the ratio of perforations to total GEM foil surface area. This parameter is a function of the perforation diameters and the pitch (linear center-to-center distance between neighboring perforations). The GEM foil holes are arranged in an equilateral triangular fashion so that the pitch defines the distance, on-center, between any of the three neighboring perforations [136]. With these relations we can define the optical transparency, τ_{opt} as:

$$\tau_{\text{optical}} = \frac{\pi d^2}{2\sqrt{3}p^2}, \tag{3.8}$$

where d is the major (larger) diameter of the hole and p is the pitch [135]. The GEM foils developed by our group at UVa have $d = 70\ \mu m$ and $p = 140\ \mu m$, and thus have an approximate optical transparency of $\tau_{opt} \approx 0.227$. This represents the maximum available surface area for the charge carriers to enter and possibly result in cascading ionization.

The probability that one of these particles would enter a GEM hole, as opposed to landing on the copper surface, is related to the collection efficiency (or electrical transparency) of the GEM foil. This parameter depends on the ratio between the applied electrical fields on the two metal sides of a GEM foil. It is the ratio of the number of electrons which are available for entry into a hole region, $N_{available}$, and the number of electrons actually collected into those regions, $N_{collected}$:

$$C = \frac{N_{collected}}{N_{available}}. \tag{3.9}$$

We can form a similar relation for the electrical fields outside and just above the hole regions, and the fields within the holes. The ratio of the two fields (external and within the hole), denoted as $\mathscr{F}$ becomes [64]:

$$\mathscr{F} = \frac{E_{\text{external}}}{E_{\text{hole}}}. \tag{3.10}$$

Given the optical transparency, τ_{opt}, and the potentials across the GEM foils (effectively $E_{external}$ and E_{hole}), the collection efficiency can be tweaked in order to match experimental requirements. It may be possible to reach a collection efficiency that is unity. However, there is plenty of give and take to consider in this optimization. This is discussed further in Section 3.3.4.

3.3.4 Extraction Efficiency

In addition to the probability of a charge carrier entering a GEM hole, there is also the probability that one will be able to "escape" it as it exits the downstream side of the foil as well. This is known as the extraction efficiency. The extraction efficiency is a ratio of the number of electrons extracted downstream, $N_{extracted}$, to the number of electrons produced in the hole region, $N_{produced}$, and can be written as [27]:

$$X = \frac{N_{\text{extracted}}}{N_{\text{produced}}}. \tag{3.11}$$

Using a model for the electric flux through a GEM hole [126] we can relate the extraction and collection efficiencies to their relative fluxes, $\Phi_{extracted}$ and Φ_{hole}, and find the following relation:

$$\frac{X}{C} = \frac{\Phi_{\text{extracted}}}{\Phi_{\text{hole}}}. \tag{3.12}$$

This can be re-written in terms of the corresponding electric fields and the ratio between the external and hole regions of the GEM, $A_{\text{external}}/A_{\text{hole}}$, and the result is [126]:

$$\begin{aligned}\frac{X}{C} &= \frac{\Phi_{\text{extracted}}}{\Phi_{\text{hole}}} \\ &= \frac{E_{\text{external}} \cdot A_{\text{external}}}{E_{\text{hole}} \cdot A_{\text{hole}}} \\ &= \mathscr{F} \cdot \frac{1}{\tau_{\text{optical}}}\end{aligned} \tag{3.13}$$

where we have interpreted the ratio of areas, $\mathrm{A}_{external}/\mathrm{A}_{hole}$, as the inverse of the optical transparency $1/\tau_{optical}$ [64]. This relation can be used to help optimize the GEM's efficiency by way of its physical design and operational voltages. From Equation 3.8 we have that $\tau_{optical} \approx 0.227$. With the parameter $\tau_{optical}$ set, we can then manipulate the field ratio parameter, $\mathscr{F}$, to get as close as we can to our desired efficiency. In our case, we would strive for 100% efficiency, but we can never get that in practice.

Efficiency optimization is not as straight-forward as it may appear from these simple relations. The process is iterative and requires both some give and take in balancing between practical and safe GEM operating voltages, as well as realistic efficiency expectations. In comparisons of GEM foils and varying applied voltages and pitch geometry, it was shown that collection efficiency increases while extraction efficiency decreases as the perforation pitch decreases [18]. Recall that the pitch is the center-to-center distance between the GEM foil perforations. Therefore, for a fixed hole diameter, as the pitch decreases, the number of holes on the foil surface increases and likewise, the total area of holes increases. This directly increases the optical transparency as the pitch gets smaller. This then diminishes the area of metal on both sides of the foil and therefore the area between the holes, and leads to less electrons being lost across the hole regions [116]. Consequently, this requires a much higher voltage across the induction gap in order to maintain a nominal extraction efficiency. However, this increase in the voltage across the foil decreases the chance of an ion escaping the hole regions and therefore the overall extraction efficiency [18, 116].

3.3.5 Gain

The electric fields in the hole regions of a GEM are the primary generators of the cascading avalanche of particles, or *gain*, through the foil and each of the hole regions. The *effective gain* of a GEM foil can be formulated in terms of its charge collection, gain, and extraction efficiency[134]:

$$G_{eff} = C \cdot G \cdot X \tag{3.14}$$

Here, C is the electrical transparency (or collection efficiency), G is the gain (the amplification factor within the hole region of the GEM), and X is the extraction efficiency. The gain of a single GEM is reliant on the available ionizing gas and the field within the hole regions. More on this can be found in Sections 3.2.4 and 3.3.2, respectively. In the most common GEM designs, amplification across the GEM foil through the GEM holes typically sets in around 250V or higher [40]. Beyond that, we combat an exponential plateau for both the gain and efficiency (see Figure 3-11).

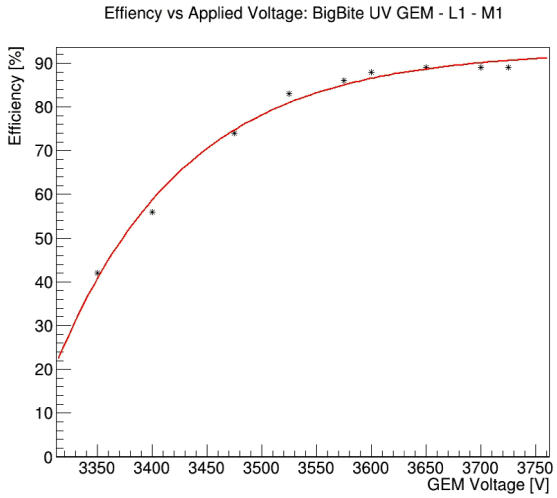


Figure 3-11: Efficiency vs. Applied GEM voltage plot. In order to develop the characteristic efficiency curve, the efficiency of the GEM chamber is measured at various set voltages. This plot can be used to determine the plateau for maximum efficiency and to select the nominal operating voltage.

This voltage-dependence of the efficiency and gain is a fundamental property that is always carefully managed during GEM operation. Once the relationship between

voltage, gain, and efficiency is plotted from measurements at various voltage steps, it is used to set operational limits on the GEM foil/detector.

Figure 3-12 illustrates the basic principles of collection efficiency, gain, and extraction efficiency. Each sub-figure is a cross-sectional view of a GEM foil hole region with arrows drawn in to show the behavior of particles in different scenarios. The figures are not to scale and are for illustrative representation only.

Figure 3-12a shows four particles which could possibly enter the hole region. Only three of these particles actually enter the hole and therefore the collection efficiency here is $C = 3/4 = 75\%$. Figure 3-12b shows a single particle resulting in an avalanche of three exiting particles and therefore, the gain here is $G = 3$. The third figure, 3-12c illustrates the extraction efficiency. Here, five particles are exiting the hole region. However, only three are able to escape. Therefore, here we have that $X = 3/5 = 60\%$.

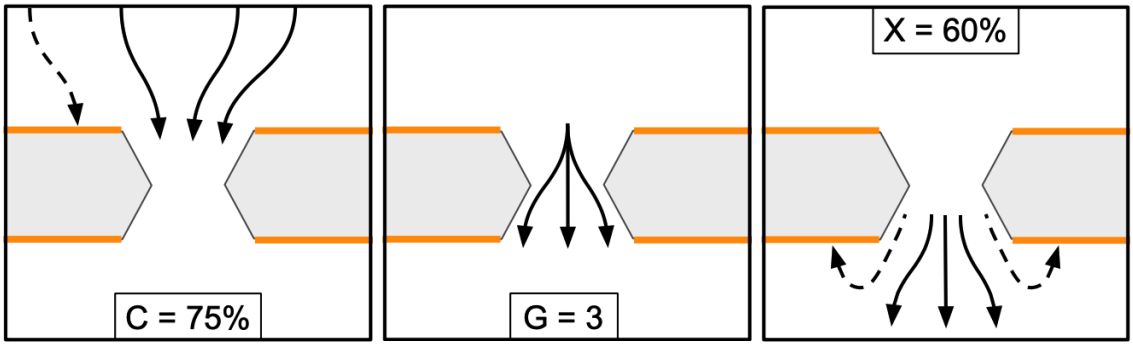


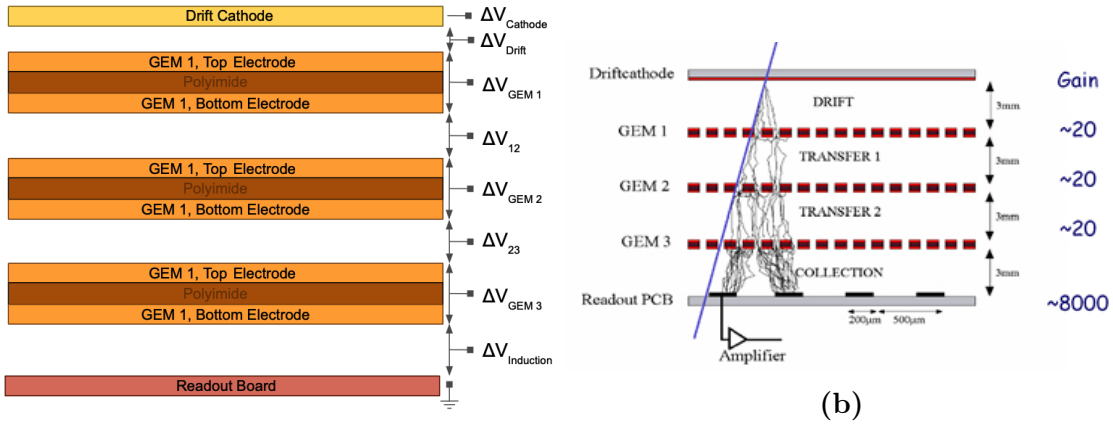
Figure 3-12: Drawings of the three GEM gain factors. Figures show cross-sections of a GEM foil hole with the orange line representing the copper sides and the grey representing the insulating polyimide. **(a)** Collection efficiency, C. **(b)** Hole region gain, G. **(c)** Extraction efficiency, X. Each of these represent the propagation for electrons. Corresponding figures can be made in the opposite direction to represent the drifting ions [135].

3.3.6 Multiple GEM Foils in Succession

In most of the previous sections of this Chapter we have primarily focused on single GEM foils — the single layer of polyimide coated on both sides with metal. As discussed earlier, additional layers of GEM foils can be added incrementally to

promote various desirable characteristics from the detector. Additional GEM foil layers provide higher gain while reducing ion back flow and the discharge rate by orders of magnitude [113]. This is because adding in more GEM foils inherently reduces the strain on each foil. Another advantage of using multiple GEM foils is their intrinsic ability for ion gating at particular voltage settings [64].

Some challenges to adding more foil layers are increased power requirements, additional electrical hardware, a possible need for more control electronics, and added construction complexity. The GEM detectors built by our group at UVa for the SBS Program were all triple GEMs (Figure 3-13). The primary foil and layer structure of our GEMs were (from top to bottom) a Drift Cathode, GEM Foil 1, GEM Foil 2, GEM Foil 3, and the Readout Electronics plane on the bottom (Figure 3-4b).



(a) (b)

Figure 3-13: Triple GEMs. **(a)** A schematic of a triple GEM detector. The layers for this type of detector are (from top to bottom): Drift Cathode foil, GEM Foil 1, GEM Foil 2, GEM Foil3, and the Readout Plane. **(b)** A drawing of an ionizing particle cascading, with gain, through the layers of a triple GEM. Image source: [30].

The foils used in our GEM detectors each provide a multiplication factor of about 20 and so, the resulting charge amplification is approximately $20^3 \approx 8000$ [24]. We can extend this to a general formulation for any number of stacked GEM foils, N_{foils}, to find the total effective gain for the combination as:

$$G_{\text{eff, tot}} = \prod_{i=1}^{N_{\text{foils}}} C_i \cdot G_i \cdot X_i. \tag{3.15}$$

Recall that the gain on a single GEM foil is the ratio between the number of electrons collected on the anode (downstream) foil and the number of electrons available just before the cathode (upstream) foil. Similarly, the total gain for a stack of multiple GEM foils is the ratio of the number electrons on the anode of the final, bottom most GEM foil and the number of electrons available just before the topmost GEM's cathode foil.

After undergoing this much gain, a single ionizing particle entering the GEM chamber would have enough charge to be sufficiently measured by our readout hardware and electronics. Figure 3-13b provides a schematic for a cascading avalanche that can occur when an ionizing charge interacts with the top cathode layer.

3.4 Physical Design and Principles

The first run of the SBS Program, at maximum, had twelve separate GEM detectors. These GEMs were contributed by our research group and by a group at INFN Roma. The initial experimental setup had twelve GEMs consisting of six INFN Roma XY GEMs and six UVa GEMs. Our six GEMs consisted of two UV GEMs and four XY GEMs. By the end of the experiment, all of the GEMs in the experimental setup were those provided by our group — four UV and four XY.

3.4.1 GEM Chamber General Design Layout

Though there are some noticeable differences between the construction, utility, and appearances between our XY and UV GEMs, their general construction and operation as a GEM chamber remain the same. The primary difference lies in the readout foil coordinate configuration as either UV or XY. The UV GEMs will be the model for explaining the physical design and structure of these detectors.

Figure 3-14 shows an exploded view of an SBS UV GEM chamber. The structural frames and support pieces are labelled on the left side of the image and the GEM foils and readout are labelled on the right. Though the GEMs may be oriented and operated in any direction once assembled and fastened down, our typical nomenclature

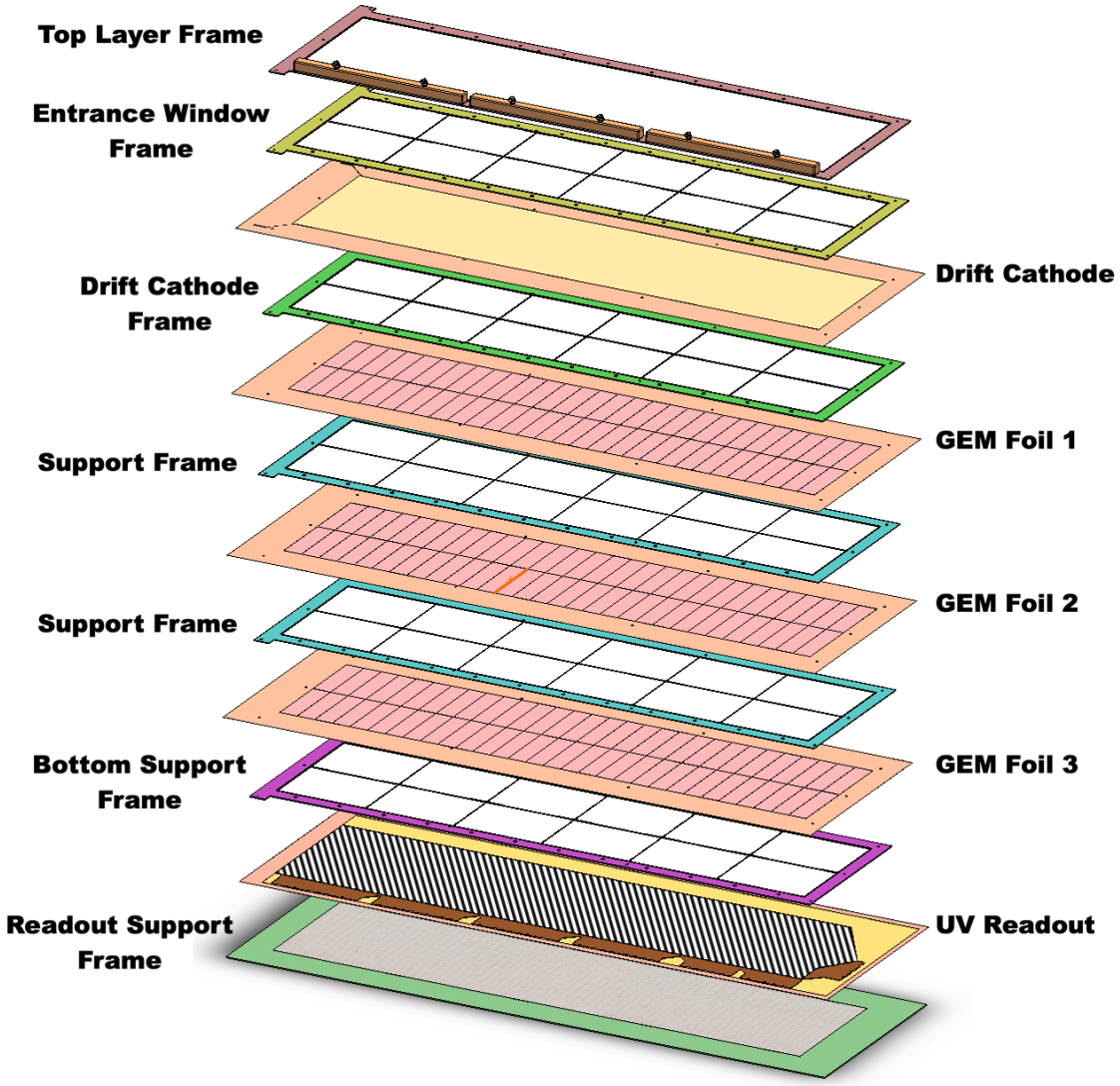


Figure 3-14: Exploded view of an SBS UV GEM chamber. Frame and support structures are labelled on the left, and GEM foils and readout board are labelled on the right.

follows this figure with the topmost piece being the Top Layer Frame and the bottom most piece being the Readout Support Frame. The naming convention for the copper foils also follows a top-down pattern as: Drift Cathode, GEM Foil 1, GEM Foil 2, GEM Foil 3, and the Readout Board at the bottom. The entire series of frames and foils are glued, clamped, and bolted together during the construction of the chamber (Sec. 3.5). The layout of the frames and foils are designed in such a way that the gas flow volume and ports are fully integrated into the assembly (Sec. 3.4.4).

3.4.2 Foil Layers Within a GEM Detector

This section aims to focus primarily on the macro-level form and function of the Drift Cathode, GEM foils, and Readout Board. Most of the previous discussion has focused on the finer structures of these components — down to the level of microns or less. Understanding these elements at small scales is fundamentally important. However, these are large-scale detectors that sit on tabletops and mount to weldments; they require electrical cabling, power, and data-acquisition systems in order for any real measurements to occur. Therefore, it is also important to understand them as we can touch them in the lab and see them with our naked eyes.

The Drift Cathode Foil

The Drift Cathode foil is the most simple of the detector's primary foils. There are some auxiliary layers of the detector which are very primitive in comparison to the Drift Cathode, but in terms of principle foils, the Drift Cathode is the most basic. As the name implies, the Drift Cathode foil is the cathode for the first drift region (above GEM 1) and the leading cathode electrode for the entire detector. Figure 3-13a shows each layer of the GEM chamber as an electrode with its applied voltage. In terms of potential differences, the Drift Cathode is at the greatest voltage as the cathode, and the Readout Board is referenced at ground as the anode. The cathode foil consequently becomes the greater potential for GEM Foil 1 and therefore defines the first ionization and drift region. Figure 3-15 shows how voltages can be

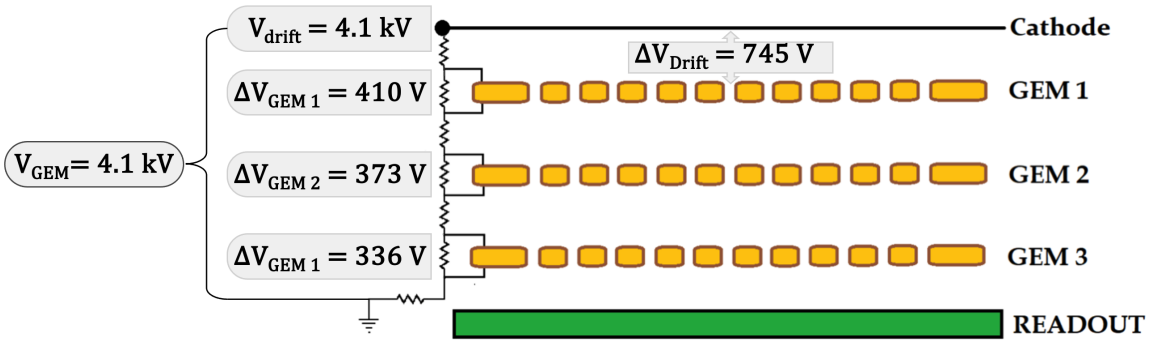
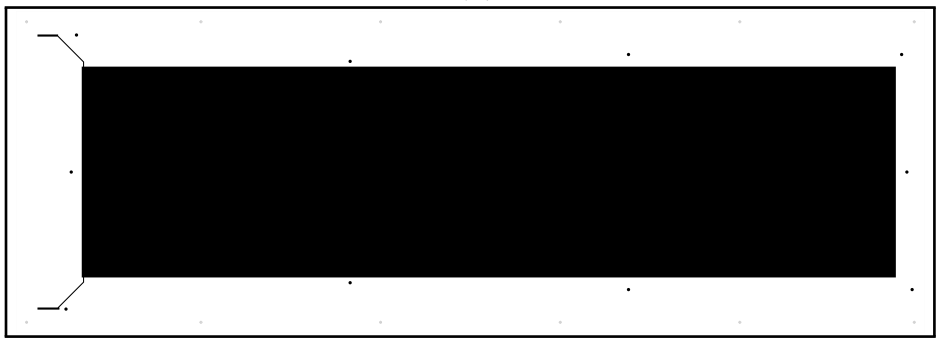


Figure 3-15: A typical voltage distribution across the foils of a GEM detector with an overall voltage setting of 4.1 kV.

distributed across the foils of a GEM detector. The configuration shown has an overall GEM supply voltage of 4.1 kV. The voltage drops across each region (Drift, GEM1, GEM2, etc.) are defined by the resistance values between each foil or drift region.

(a)



(b)

Figure 3-16: UV Drift Cathode foil photo and schematic. **(a)** Raw UV Drift Cathode Foil. **(b)** Electrical (Gerber) schematic for the Drift Cathode Foil. Black regions indicate areas of copper.

Figure 3-16 shows a picture of a "raw", unframed UV Drift Cathode foil (Fig. 3-

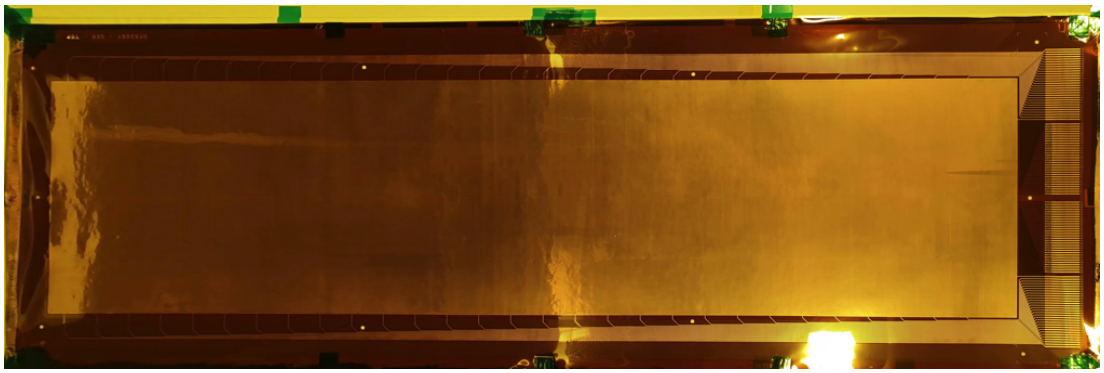
16a) and its Gerber printed circuit board (PCB) schematic (Fig. 3-16b). Apart from some gas flow permeability (see Sec. 3.4.4), the Drift Cathode is just a simple copper foil layer. There are two electrodes on the copper foil which serve as application points for the applied voltage. These can be seen on the left side of the two figures (one electrode is coming off of each left corner at an angle). These electrodes are simple pads and traces which allow the current to flow onto the large copper surface area occupying most of the foil. Note that the black regions inside the rectangular bordered region of the Gerber schematic (Fig. 3-16b) indicate copper surfaces. Therefore, as a Drift Cathode, it is basically a single large-area copper foil — UV foil active area dimensions are 1500 mm x 400 mm.

GEM Foils: General Design and Layout

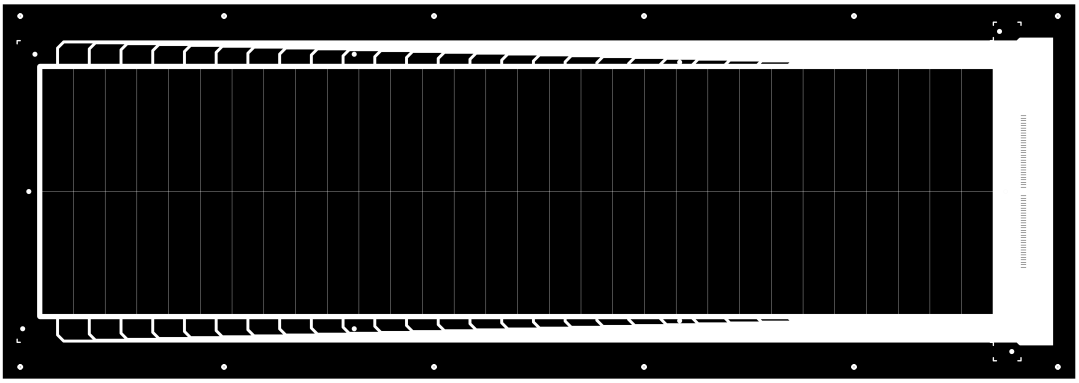
We know that each GEM foil is a thin layer of polyimide coated on both sides with a few microns of copper, and that the entire active area is covered in a dense matrix of perforations. The two metal sides have voltage applied across them. This, along with the ionizing gas inside of the chamber, provides the foil and hole regions with the mechanisms to generate cascading ionization. On the macro-level, the foils are non-rigid rectangular sheets of copper with polyimide traces which segment off sectors within the active area of the foil and electrode traces on one end of the sheet. Figure 3-17a shows a top-view of a "raw" unframed UV GEM foil and Figure 3-17b shows the Gerber electrical schematic.

In Figure 3-17b you can see that most of the foil appears to be copper, as indicated by the black areas. All of the white areas are regions where the copper has been etched away down to the central polyimide layer. This essentially leaves a (reverse) polyimide trace where the copper has been removed. These traces become isolating boundaries which then separate and electrically insulate various parts of the copper sheet from the other areas. For example, the copper is removed in such a way that the active area of the GEM foil is segmented into sixty separate and electrically independent sectors.

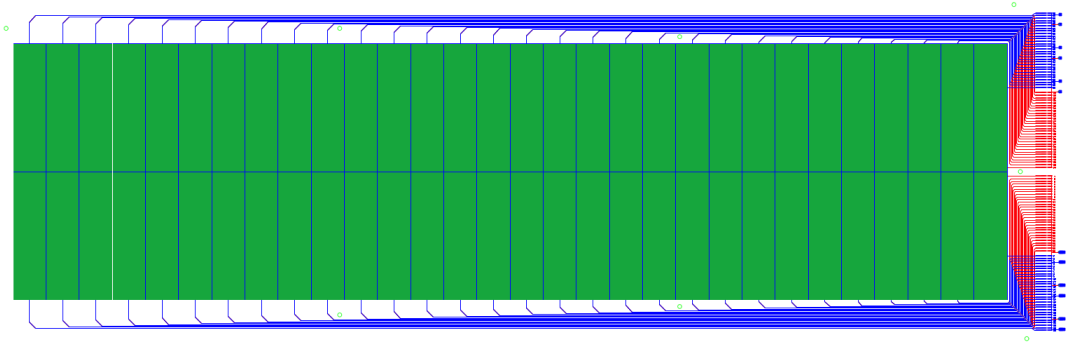
The UV GEMs have thirty segments on each of the top and bottom foil electrodes



(a)



(b)



(c)

Figure 3-17: UV GEM foil photo and schematic. **(a)** Unframed UV GEM foil laid out on a table in our lab's clean room prior to assembly. **(b)** Electrical (Gerber) schematic of the active area of a UV GEM foil. Black regions indicate areas of copper. **(c)** Layered Gerber Schematic of a UV GEM. Green areas are the copper masking, blue indicates polyimide tracing for top sectors, and red indicates polyimide tracing for the bottom sectors.

resulting in a total of sixty sectors. The XY GEMs only have their top side segmented while the bottom is left unsegmented. XY GEMs have fifteen sectors on the top electrode and one single electrode on the bottom. From the top view you can only see the top sectors. However, on UV GEMs, the bottom sectors are symmetrically placed on the bottom electrode. Figure 3-17c shows a layered Gerber schematic for the UV GEM foil. The green areas of the figure are the copper active area (or masking in Gerber-speak), the blue lines indicate the polyimide (reverse) tracing for the sectors on the top side, and the red lines indicate the polyimide traces for bottom sectors. Each of the sixty sectors has a copper trace that leads to the electrical connection side of the GEM foil (the right edges of the figures in Figure 3-17). A close-up of three of these independent sectors and their traces is shown in Figure 3-18.

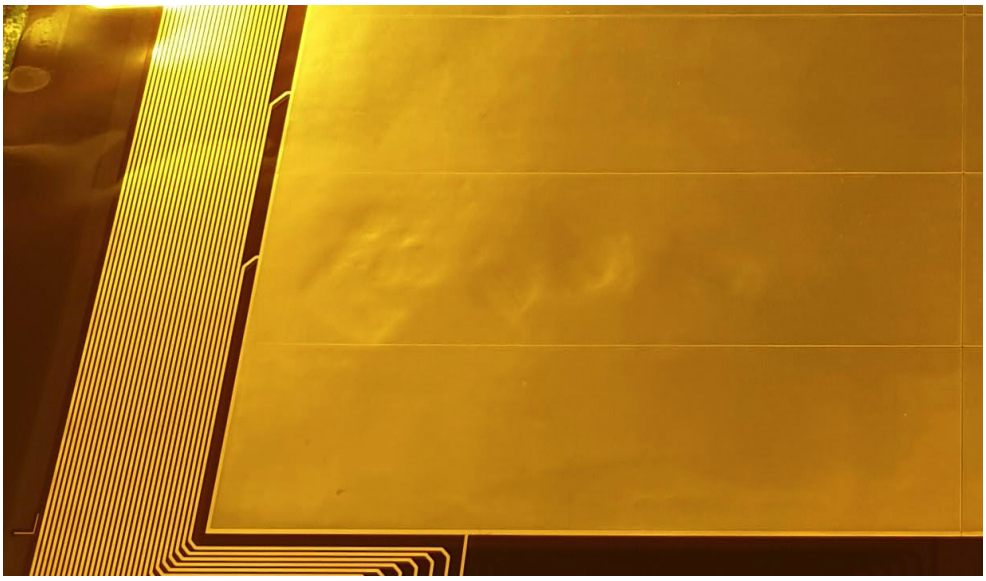


Figure 3-18: Close-up of three independent UV GEM foil sectors with their copper traces.

All of the traces from each of the independent sectors lead to a common end of the foil in order to facilitate easier connections. On the UV GEMs, both top and bottom electrode traces lead to the same end. The XY GEMs have connectors for half of the sectors on one end and the other half on a different end. Figure 3-19 shows the first top sector of a UV GEM, along with the first sixty electrical pads – thirty for the top electrodes (bottom right side of figure) and thirty for the bottom electrodes (bottom

left side of figure). The UV GEMs are unique in that you can access the bottom and top electrode pads from the top side of the foil. This allows for solder connections to all sectors, top or bottom, to be made from just the top side of the foil. This makes assembly, testing, repair, and modifications much easier due to simpler and more direct access.

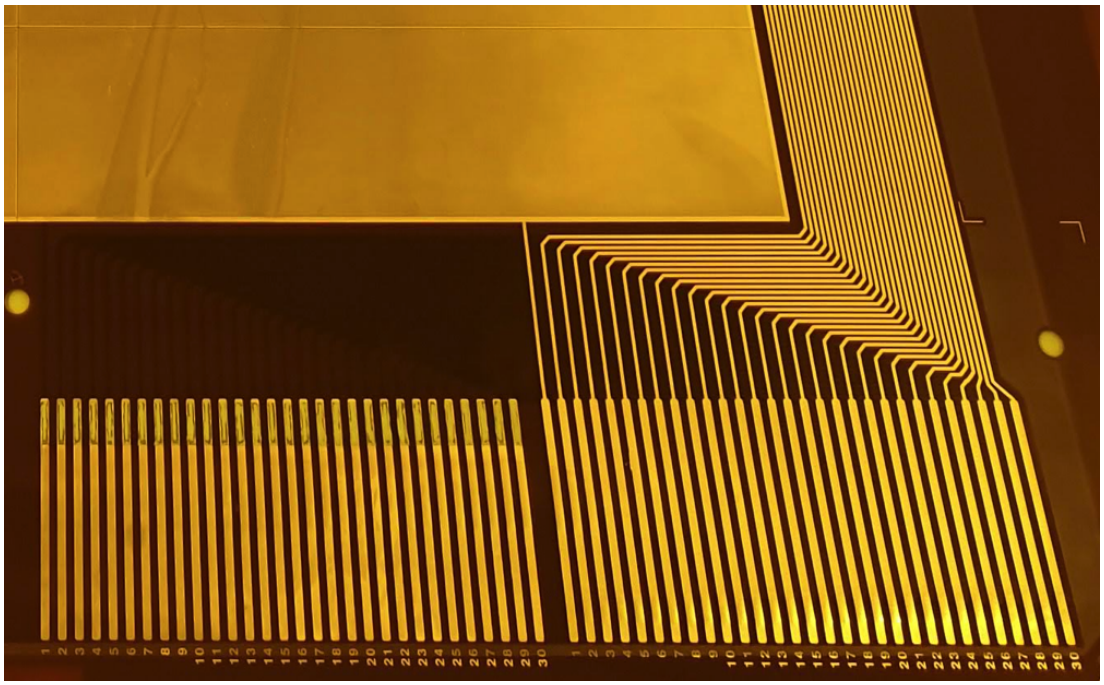


Figure 3-19: Close-up of the first UV GEM foil sector with its trace to the connection array. The lines are the copper traces and pads for the first top and bottom sectors.

Figure 3-20 shows all 120 (60 top and 60 bottom) of the electrical connection pads for the UV GEM. The strips are separated into four sections:

- Right-most grouping: top sectors 1 through 30,
- Second from right: bottom sectors 1 through 30,
- Third from right: bottom sectors 31 through 60,
- Left-most grouping: top sectors 31 through 60.

GEM Foils: Sectors

The separation of the GEM electrodes into sections on the top and bottom electrodes serves multiple critical functions. Figure 3-21a illustrates a GEM foil that only has sectors on one electrode. The segmentation of the top electrode reduces

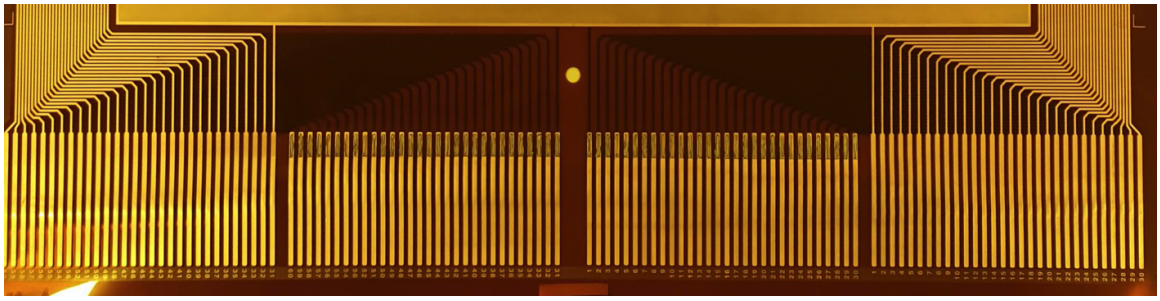


Figure 3-20: All 120 connector pads and final traces for the top and bottom sectors of the UV GEM foil.

the probability of unwanted discharge propagation and excess discharge energy [115]. During normal operation, GEM detectors can exhibit discharging within the chamber, between the top and bottom electrodes of a foil. This is typically induced at higher beam intensities and particle rates. In order to minimize the amount of energy that is discharged, it was found that a GEM sector should not exceed 100 cm^2. This restricts the capacitance of the sectors to below 5 nF [114]. Each of our GEM sectors have a nominal capacitance of about 4.7 nF. Also, in order to impose that a discharge is limited to the capacitance of only a single sector, each sector has its own very large inline protective resistor (10 MΩ in our case). This allows all other sectors to remain unaffected by a discharge in a single sector. Creating smaller area sectors also helps with the recovery time of the foil sector after a discharge. If we consider the sectors as individual capacitors with areas of A = 100 cm^2 we have:

$$\mathrm{R} = 10\ M\Omega \qquad \mathrm{C} = 4.7\ nF$$

$$\text{Recovery time, } \tau = RC = (1 \cdot 10^6) \cdot (4.7 \cdot 10^9) = 0.047s \text{ or } 47\ ms$$

The capacitance is a function of the area, and thus any increase in the size of the sector would only further increase this recovery time. Therefore, splitting an electrode into multiple sectors with these smaller areas allows for faster recovery times after a discharge. Reducing the size of the sectors minimizes the discharge energy and recovery time, but there is a balance between cost, manufacturability, component and hardware needs, and overall complexity, that always needs to be considered.

Separation of the electrodes into sectors also allows for localized repair or in-

dividual sector deactivation without the loss of the entire foil in cases when local malfunctions or defects occur. A local malfunction can be caused by multiple factors and can take on many forms. An example is a physical short between a top and bottom electrode caused by some bridging or closing of the gap between them. An errant foreign particle inside a GEM chamber could possibly become lodged within a GEM hole or between two foils, for instance. If the object is electrically conductive, it can close a loop between electrodes or foils, resulting in a short circuit. If, however, the object is not conductive, it can still cause a serious issue. Upon application of very high voltage across GEM components, the foreign object may burn up and carbonize between electrodes or foils, thereby creating a short or break, ultimately disabling that layer. It may also be possible to encounter local defects. This could be in the form of manufacturing-level defects whereby the perforations may be missing, irregular in shape, or clogged in the procurement process.

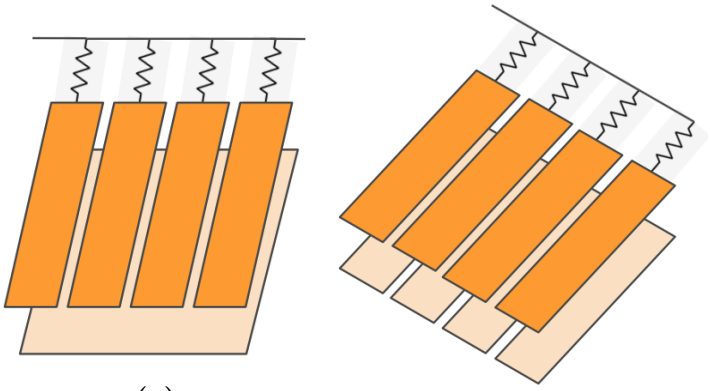


Figure 3-21: Two different GEM foil electrode sector layouts. The insulating polyimide layer between the two electrodes is not shown for display purposes. **(a)** A GEM foil that has multiple independent sectors on only one electrode while the other has no sectors (non-segmented). **(b)** GEM foil that has sections on both electrodes.

The two combinations of sector types used in the XY and UV GEMs are shown in Figure 3-21. The difference between the two is whether or not only one side of the GEM foil has sectors (Fig. 3-21a), or if both sides are sectored off (Fig. 3-21b). Each configuration has its own advantages and disadvantages. The more sectors on a GEM foil the harder it is to make and the more it costs to produce. Each sector requires

additional masking and tracing in the photo lithography process [50]. Also, for every sector added, an additional protective resistor is required, along with an added set of connections all the way back to the voltage divider boards. This increases complexity in manufacturing, assembly, and test.

The XY GEMs that we built were made with foils only segmented on the top side. A problem that can occur with this design is that if a single sector shorts out with the single non-segmented bottom electrode, this can disable the entire detector [62]. That short circuit becomes the shortest path in the electrical chain and consequently makes the other areas of the GEM non-functioning. The UV GEMs however, are built with top and bottom segmented electrodes. In this design, if a sector shorts out during operation, it does so only with its matching pair sector on the opposite electrode. As mentioned above, the protective resistors fully isolate each sector. Therefore, a short is only between the pair sectors and does not disable the entire chamber. Having sectors on both top and bottom electrodes also provides the advantage of limiting the voltage drop in the detector's voltage divider board during operation at high particle rates [62].

If a sector is discovered to be shorted out or inoperable in some way, then if possible, that sector is disconnected from the circuitry in order to eschew any ill-effects which that problematic sector may cause. Figure 3-22 shows the hit maps of two independent XY GEM detectors. On the left side of the hit maps there are two rectangular areas with low populations of hit events. These areas are two sectors, one on each GEM chamber, which are "dead" and have been disconnected from the detectors circuitry. These are triple GEMs, and when a faulty sector is identified, that sector alone is disabled. This means the other two GEM foils within that chamber still have active sectors inline with the disabled sector. With two active sectors providing a cumulative gain of about 400, some hits may still get measured in that region on these hit maps.

When a bad sector is identified, it is always best to disconnect that sector if possible. It may not always be possible due to the restricted access to the detector or other constraints. If the sector is bad due to a short circuit, then this could

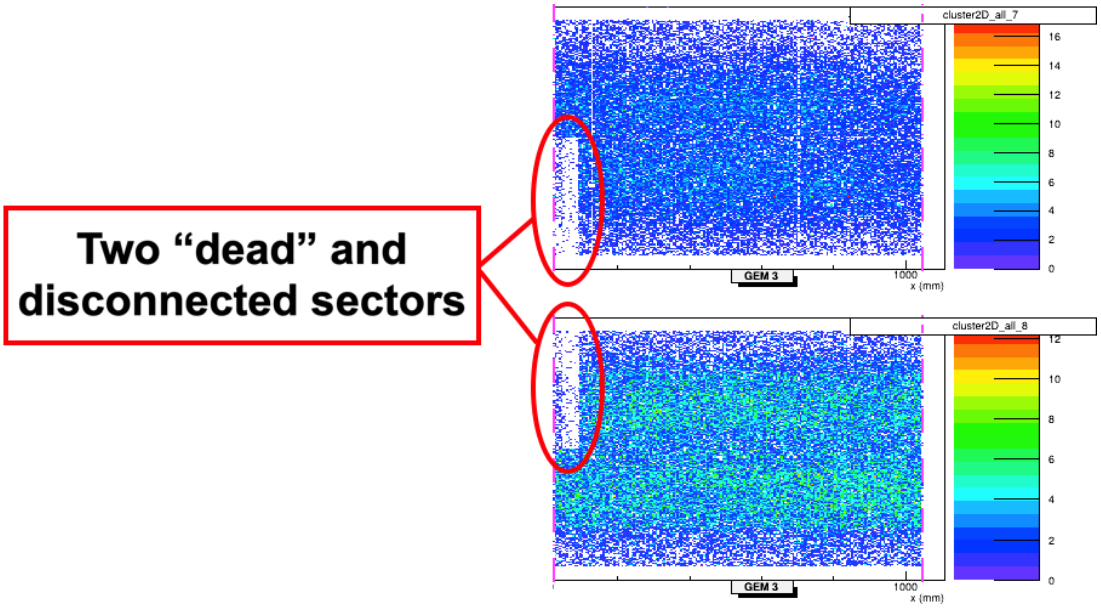
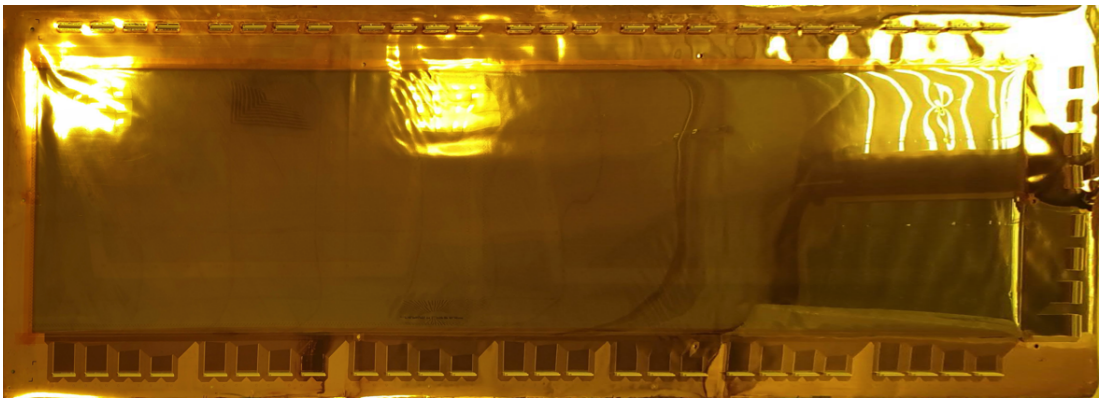


Figure 3-22: GEM hit map showing two chambers that each have a single non-functioning sector. Some hits can still be seen in this region because these are triple GEMs and only one sector on one out of three of the GEM foils was disconnected. The other two foil layers have active sectors which may still provide enough gain to read out some minimal signals.

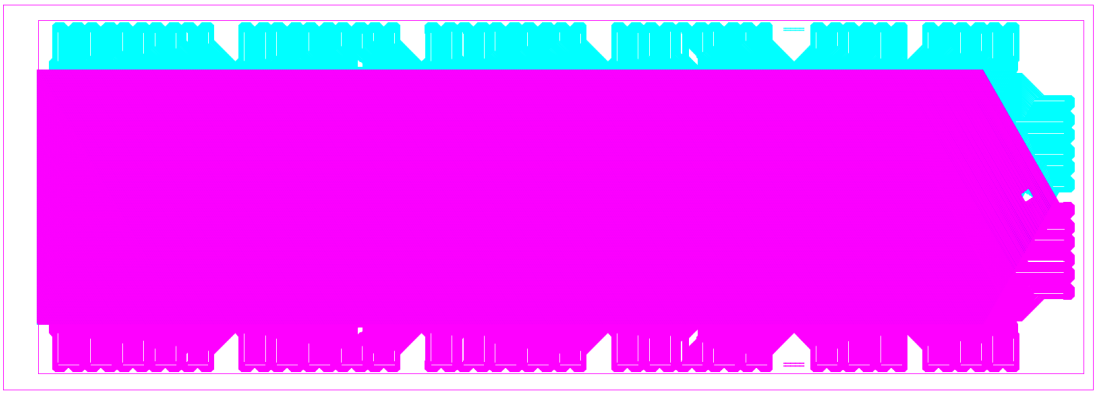
significantly reduce the voltage across that entire GEM foil, resulting in a less efficient and lower gain detector performance. This is especially true for the foil just before the Readout Board (GEM 3). If a shorted sector is not disconnected from the GEM circuit then the current flowing through that circuit will be noticeably higher. At full operating voltage, a single dead sector can raise the current draw by about 7 μA.

The Readout Foil

The readout foil is the most complex of the primary foils of our GEM detectors. Refer to Figure 3-23 for an image of an unframed UV Readout Foil (Fig. 3-23a) and for a Gerber electrical schematic (Fig. 3-23b). In Figure 3-23a you can see that most of the foil appears to be copper. This is mostly true. The sub-layers (the two sets of coordinate wire planes) of the Readout Foil are all mounted upon a semi-rigid copper sheet (approx. 500 μm thick). This relatively thicker sheet of copper provides much needed support and structure to the more flimsy layers which make up the readout strip planes. This copper backing is also used as the primary ground



(a)



(b)

Figure 3-23: UV Readout Foil photo and schematic. **(a)** Raw UV Readout Foil. **(b)** Electrical (Gerber) schematic for the UV Readout Foil. The magenta region corresponds to the top strips — connected to the lower seven connector sections. The cyan region corresponds to the bottom strips – connected to the upper seven APV connector segments.

reference. Therefore, in the manufacturing process, all ground terminations required from this layer are connected directly to this copper backing. In addition to the two sets of copper strips that each define one of the readout coordinates, the Readout Foil also incorporates the primary panasonic connectors that allow for direct measurement and readout of each strip. These connectors are the point at which the APV cards connect to the board for individual readout of each strip. You can see these on three of the edges of the readout board. The UV GEMs have 60 APV panasonic connectors, 30 APVs for each coordinate "U" and "V".

Figure 3-23b shows the layered electrical scheme for the UV Readout Foil. The two regions indicate the two sets of thin copper wires for each coordinate axis. The

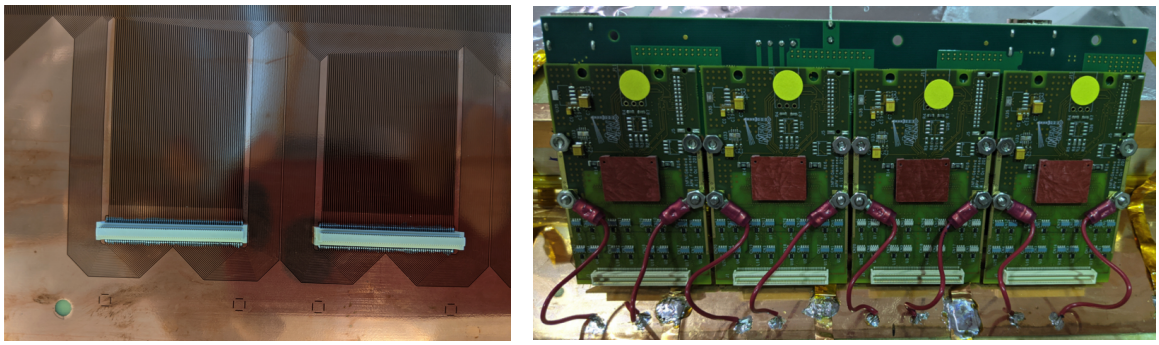
magenta region corresponds to the top layer of copper strips on the "U" axis, and the cyan area (which is mostly blocked by magenta region being on top of it) corresponds to the bottom strips, or "V" axis. In this schematic the copper backing sheet is not shown. This should be noted, as this indicates that the coordinate strips are therefore not connected or grounded to that copper sheet. There are some ground connections made to copper backing from various electrical components of the detector, but none of those are a part of these coordinate strip circuitry.

Readout-level Electronics Hardware

GEM detectors require a variety of auxiliary electronics in order to properly function and perform. Most of these are not fitted directly onto the detectors or even on a composite layer of detectors – such as MPDs, VTPs, HV and LV power supplies, etc. There are several primary components that mount directly onto the GEM detector: the APV25 cards and their associated backplanes, the voltage distribution board(s), and the protective resistor banks.

The APV25 cards that translate and map the collected charge from the Readout Board strips mount directly onto the Readout Plane. In Figure 3-23a, close inspection reveals 60 electrical (panasonic) connectors on the top, right, and bottom sides of the image. These are the plugs where each of the APV25 cards connect. A close-up image of these plugs can be see in Figure 3-24a. Each one of these panasonic connectors inserts into a single APV25 card. For various power and communication routing purposes, banks of APV25 cards are connected via a single backplane. More specific details on the design and operation of these cards and backplanes can be found in Section 2.7. Figure 3-24b shows four APV25 cards mounted to their 4-slot backplane, ready for connection directly to the readout board.

In addition to the APV cards, the GEM detectors are fitted with a PCB that controls the voltage distribution across the layers and electrodes. Each GEM layer's electrode (top and bottom) get their own protective resistor board. The UV and XY GEMs each have their own versions of these components, but they perform exactly the same functions. The high-voltage distribution boards take the incoming HV from



(a) **(b)**

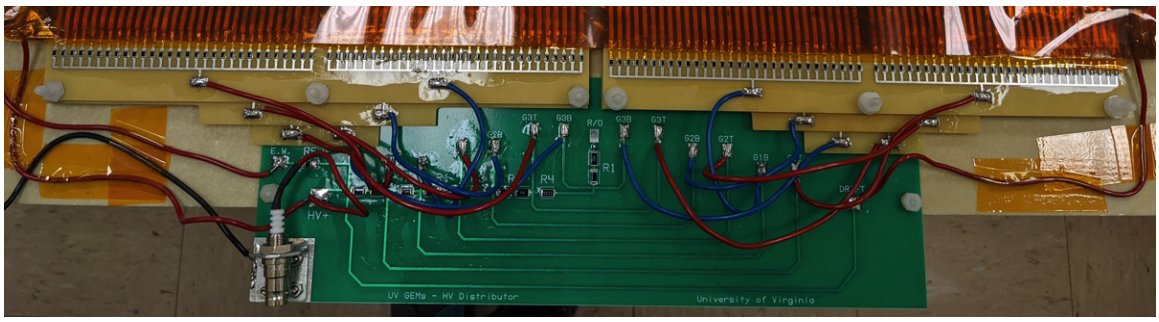
Figure 3-24: UV APV25 connections and cards. **(a)** Panasonic connectors that are directly on the Readout Board. Each connection gets inserted directly into a single APV25 card. **(b)** Four APV25 cards mounted to their backplane, ready for connection on the Readout Board.

an SHV connector and step it down to the voltages required by each foil of the GEM detector (Fig. 3-13a). Between the main HV input and the global ground reference, there are a series of resistors. Each resistor is sized such that the voltage drop across it matches the delta V needed between the electrodes/foils at that given point. Between each resistor, an electrical trace leads to an open pad. These pads provide solder points for wires that connect between the divider boards and the protective resistor boards.

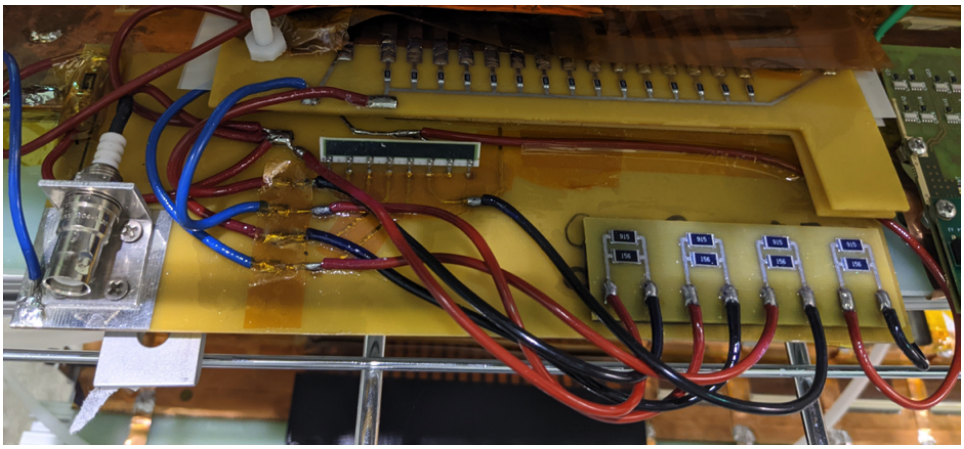
The protective resistor boards are simple PCBs that insert large resistors (10 $M\Omega$) inline with each voltage trace. Refer to the sub-section on GEM foil sectors in Section 3.4.2 for more details on the purpose of these protective resistors. Each GEM foil sector gets its own protective resistor. Therefore, the UV GEMs have 120 per foil (60 for the top and 60 for the bottom) and the XY GEMs get 31 (30 for the top segmented electrode and 1 for the non-sectioned bottom electrode). Though these are relatively simple printed circuit boards, their function is principle to the detector's successful operation.

3.4.3 GEM Frames

The Drift Cathode, GEM foils, and readout are all non-rigid sheets that require support and some external forces to hold their fitted shape. The frames of these GEMs



(a)



(b)

Figure 3-25: High-voltage dividers and protective resistor boards. **(a)** The HV distributor board (green piece in center of picture) and the individual protective resistor boards (top left and top right portions of the image) for the UV GEMs. **(b)** XY GEM HV distribution board (center of view) and protective resistors (top portion of image).

are not only the skeletal components that provide structural form and integrity, they also serve other fundamental purposes such as setting the gap distances between each individual foil layer and providing support in the middle of the active area to prevent the GEM foils from collapsing on one another. The frames also create the passages which allow for proper gas flow throughout the chamber.

Each of the GEM foils in our detectors get mounted directly to their support frame. Therefore, for the UV GEM, which has five foils there are a total of seven frames per detector: Top Layer, Entrance Window, Drift Cathode Support, GEM Support 1, GEM Support 2, Bottom Support, and the Readout Support frames. There is a frame for each foil and two additional frames on the top of the chamber

that provide gas input and mounting for an aluminized polyimide entrance window (Sec. 3.4.5).

Each of the GEM frames is made from G-10 (or garolite) laminated fiberglass. This material is perfect for the GEM frames, because it is a great electrical insulator which is required in order to properly isolate the foils and other components from one another. The frame material also has extremely high strength and dimensional stability over a wide range of temperatures [90]. It is also a material that can be procured relatively easily and be machined to high precision (±0.10 mm). This is very important since the drift, transfer, and induction regions of the GEMs are defined by the thicknesses of the frames. Also, the gas flow routing requires tight-area machining given the detectors size constraints and therefore an ability to machine to high-precision is a requirement of the frame material.

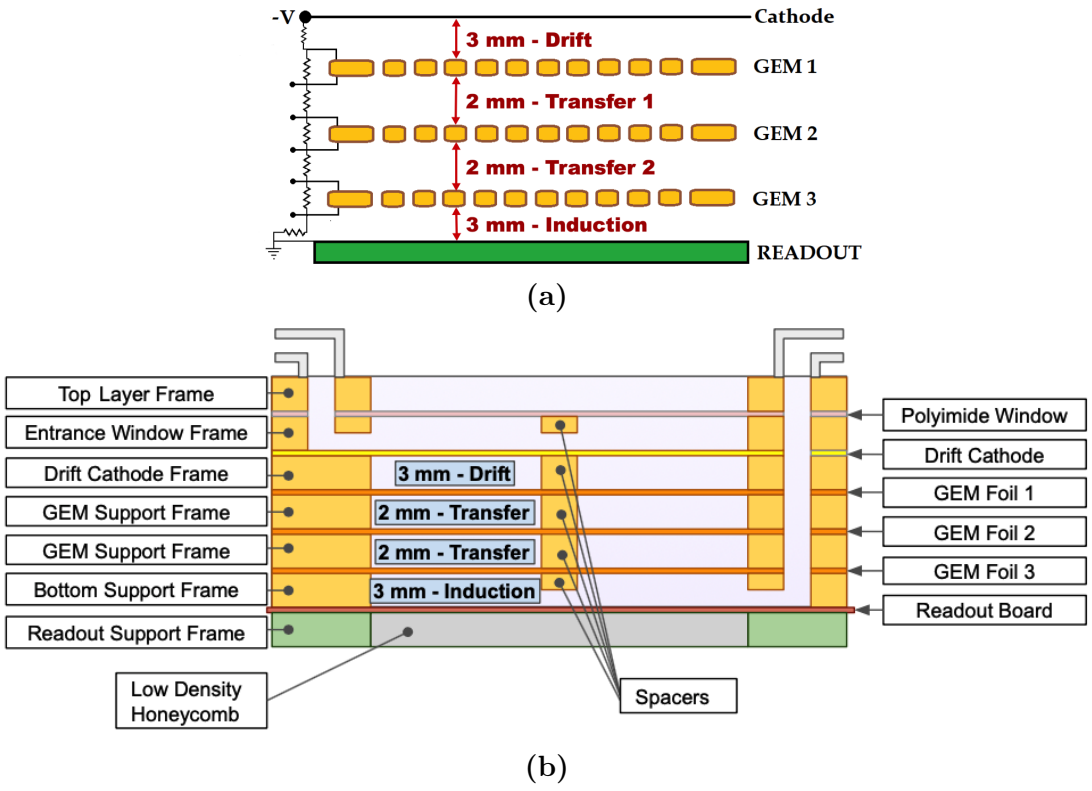


Figure 3-26: Triple GEM internal structure and layout. **(a)** The potential gradient is defined by the distances defined by the width of each frame. **(b)** Frames, foils, and gap regions. Each foil is fastened between a support frame.

Figure 3-26 shows cut-away drawings of the internal layer structure of our GEM

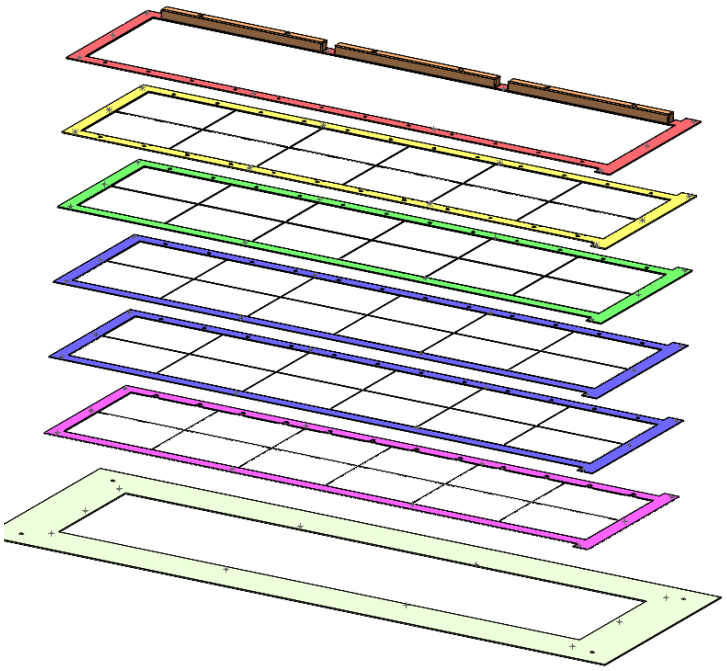
chambers. In Figure 3-26 we see each of the foil layers with the gap distances shown, in millimeters. We have that the distance for the Drift region is 3 mm, the two transfer regions are both 2 mm, and the induction gap is 3 mm. These distances are maintained by the thickness of the frames that occupy that region. Figure 3-26b shows the same regions but also includes the GEM frames. As we can see from these two figures, the frame thicknesses match that of the gap distances:

- Drift Gap = 3 mm = Drift Cathode frame thickness
- Transfer Gap 1 = 2 mm = GEM Support Frame 1 thickness
- Transfer Gap 2 = 2 mm = GEM Support Frame 2 thickness
- Induction Gap = 3 mm = Bottom Support Frame thickness

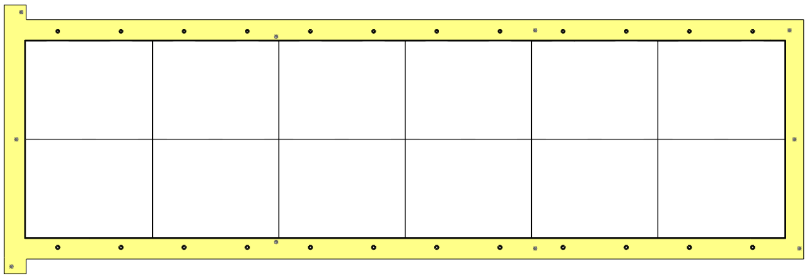
In the center of Figure 3-26b there are frame sections labelled as "Spacers". These are narrow and thin mullions of frame that extend into the open active area window portion of the frames. These can also be seen in the open areas of the frame drawings in Figure 3-27. Figure 3-27a shows an exploded view of a UV GEM as frames only (no foils) and Figure 3-27b shows only a single frame (Entrance Window) for a more close-up reference.

The "active area" window regions of the UV frames are 40 cm by 1500 cm (50 cm by 60 cm for the XY GEMs). Even with proper tensioning of the foils, this window gap is large enough that we must consider sagging due to gravity and electrostatic attraction forces. During detector assembly, the foils are stretched and glued to their associated frames. This helps to prevent sagging and wrinkles that would become sources of shorting out.

It is also important to consider the attractive electrostatic force on neighboring foils when high voltages are applied. This would only further exacerbate any sagging of the foils already in that direction. The stretching of the foils onto their frames re-directs the tension directly onto the frames. The frames are thin and long and therefore susceptible to buckling and flexion. This has the effect of negating the tension applied in the foil assembly however, if all of the layers are properly secured with adhesive and fasteners in the assembly process this is all mitigated. In the case



(a)



(b)

Figure 3-27: GEM frames. **(a)** Exploded view of the frames of a UV GEM. Th open active areas have mullions which support the foils and prevent them from sagging or collapsing in on one another. **(b)** A close-up view of the UV Entrance Window frame. The mullion spacers segment the large active area window into 12 smaller sections.

of sagging or frame flexion that leads to a decrease in distance between neighboring foils, the frame designs incorporate window spacers, or muntins, to act as physical support and barrier between foil layers.

The frame spacers span the open window regions and help prevent foil collapse and subsequent short circuiting. There are five of these spacers on the UV GEMs. This reduces the 40 mm window span down to two 20 mm sections and the 1500 mm open area segments into six 250 mm gaps. These reduced window sizes make foil collapse less of a problem. These spacers are directly in the active area of the GEMs and so, we try to minimize the number that are included. There is a fine balance between the number of spacers implemented and the loss of active area due to their presence. The more spacers, the more protection from foil collapse; however, each spacer creates a region of null readout due to the multiple scattering they cause.

The frames of our GEM chambers are an integral component for gas flow throughout the volume of the detectors. The inlet/outlet ports, flow passageways, and volume uniformity are all defined by the GEM frames. The inlet and outlet ports are directly a part of the outermost GEM frames. Also, the path for gas flow from the input ports, through the chamber, and out the exit ports is defined by grooves, channels, and holes machined directly into the GEM frames.

3.4.4 GEM Chamber Gas Distribution

The gas flow scheme for the UV and XY GEMs is shown in Figure 3-28. The black arrows indicate the general direction of macro gas flow throughout the chamber. The inlet of the chamber is on the top left of the figure and the outlet is on the top right. On the inlet side, the gas flows in through ports on the top of the frames. All frames have ports on the outlet side which allow for gas to flow up from the bottom (readout surface) of the GEM.

The gas comes into the chamber via the ports on the top left of the diagram and enters into the chamber volume between the Polyimide Entrance Window and the Drift Cathode foil. The entrance window is non-porous while the cathode and GEM foils each contain the dense matrix of perforations described in previous sections. In

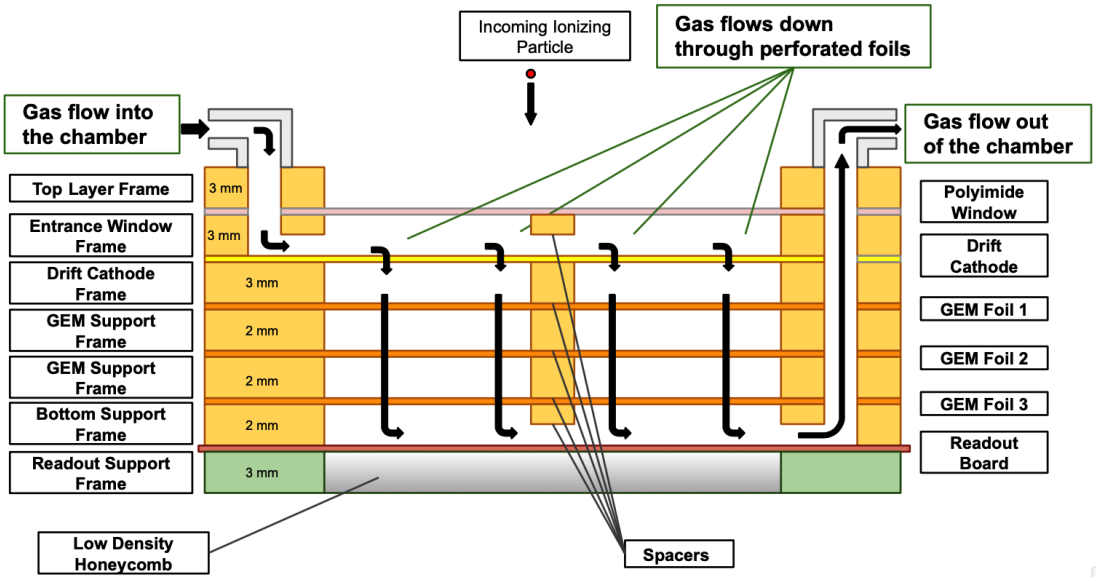


Figure 3-28: Cut-away drawing of gas flow through the UV GEMs. The black arrows show the macro flow path of the gas. Note that the gas flows through each of the foils via the perforation matrix of the foils.

addition to creating the regions for ionization and avalanching, these holes allow the gas to permeate across the layers and therefore flow through and past each of them. This starts at the very top of the Drift Cathode (and below the Entrance Window). The gas then flows through each of the foil layers until it reaches the non-porous readout board plane. From here the gas exits the chamber on the bottom right side of the figure. The spacers in the frames each have notches which allow gas to flow freely across their boundaries. The spacers are there to support the foils but need not restrict gas flow throughout the chamber. During the design phase of the frames, proper gas flow and circulation throughout the chamber was always an important consideration.

Under normal operating conditions the GEMs ionizing gas gets depleted and needs to be replenished. Recall that the electrons will collide with the gas atoms and create an avalanche of electrons through successive ionization collisions. Also, during standard operation, impurities may form within the gas volume of the detector[2].

[2]Under standard operating conditions, discharges may occur throughout a GEM detector. These discharges can produce potentially catastrophic physical debris inside the detector. Additionally, outgassing may occur from various detector components. These outgassed impurities affect the gas mixture and need to be removed.

Therefore, the gas needs to be flushed and replenished continuously throughout the entire volume of the chamber that is subject to the active area. Steady gas circulation at a reasonable rate and operating pressure was integrated into the design of these detectors. The placement of the gas inlets and outlets creates a cross-flow through the chamber. The gas enters in the top left of the figure and immediately fills volume above the Drift Cathode while also permeating down through the subsequent layers. As the gas moves through the layers, it will also fill the volumes at each level until it reaches the bottom. The exiting port is on the opposite side from the inlet port, and therefore the gas must flow across and down (cross-flow) through the chamber.

Gas circulation through a GEM chamber is a critical aspect of the detector's operation and performance. Therefore, in the design phase of the UV frames a rigorous gas flow study was performed in order to ensure that under the expected conditions, there would be uniform and refreshing gas circulation throughout the entire volume of these GEMs (Figs. 3-29 and 3-30). The computer-aided design (CAD) software used for all UV part design and flow simulation was SOLIDWORKS and SOLIDWORKS Flow Simulation, respectively. SOLIDWORKS Flow Simulation allows you to take the 3D parts and assemblies, and subject them to various simulated constraints. These included such things as gas properties (mixture ratios, densities, temperature, etc.), material properties, and environment properties (atmospheric pressure, temperature, humidity, etc.). Specific gas flow and pressure parameters are then set to match what is expected during operation. For instance, the expected total flow rate into these detectors was approximately 20 SCFM. The UV GEMs have a total of 12 inlet ports and so the flow into each port was set to $^{20}/_{12} \approx 1.667$ SCFM at STP.

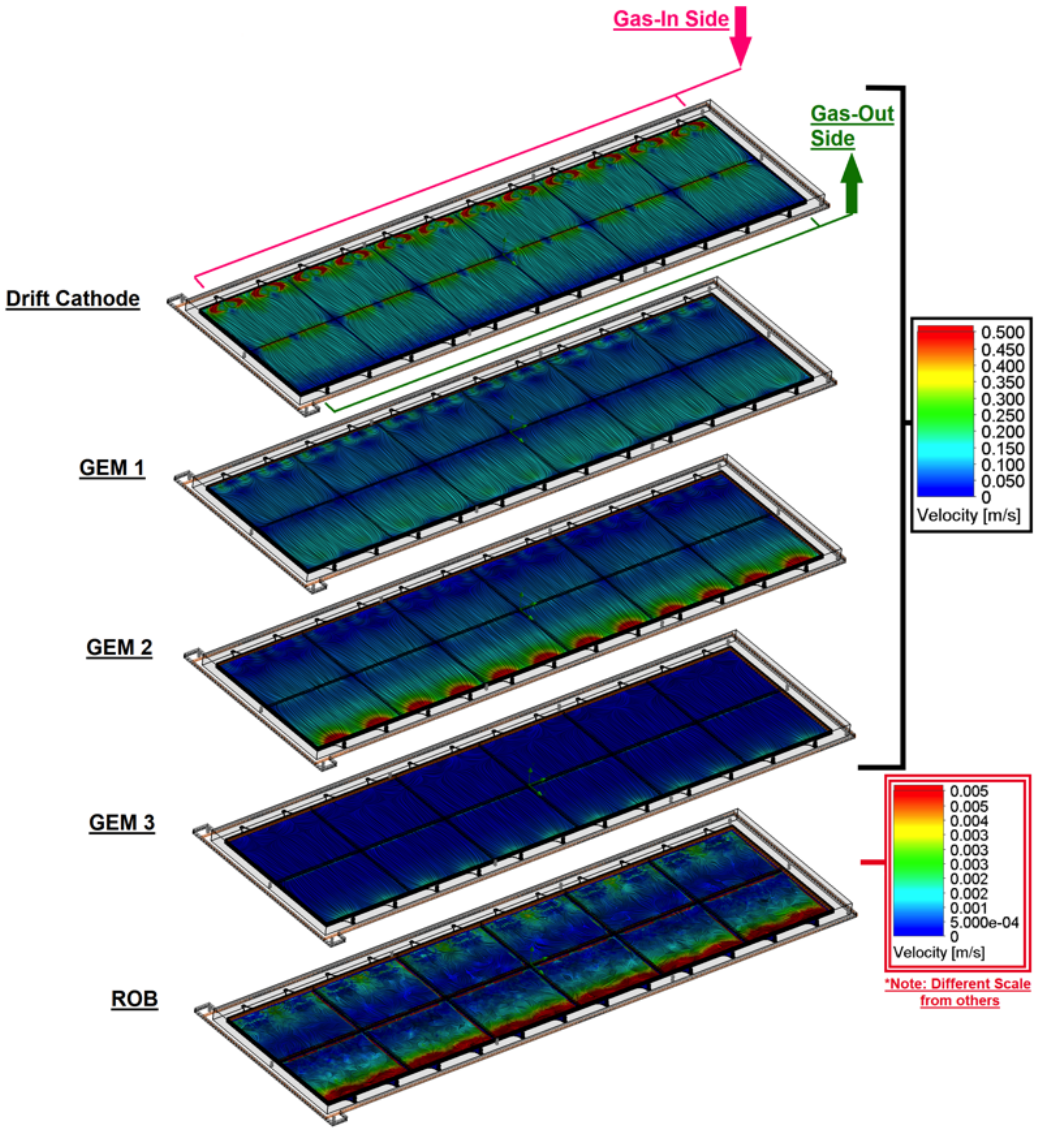


Figure 3-29: Exploded layer view of UV gas flow simulation.

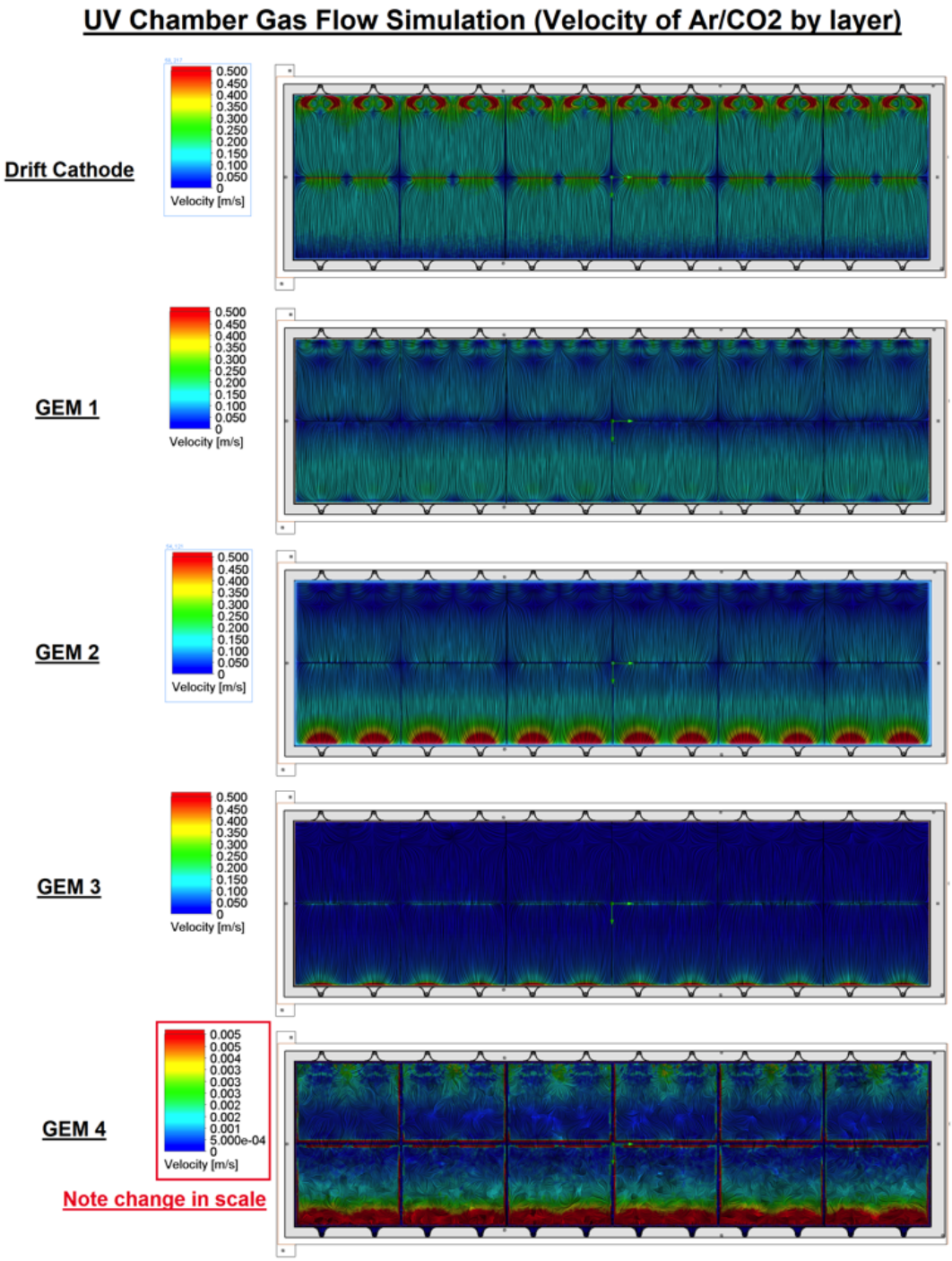


Figure 3-30: Top view of the gas flow simulation on each layer.

Figure 3-29 shows the final flow simulation across all layers of a UV GEM. The gas input into the chamber is on the top portion of the figure. From there, the gas flows across and down the layers until it reach the bottom where it then flows back

up and out the "Gas-Out Side". Figure 3-30 shows a top view of each layer's gas flow simulation. Most of the stagnant regions were eliminated. However, some are still visible in the figures. There can never be full elimination of these stagnation zones as there will always be regions, such as corners or other confined areas, where surface friction mitigates transport of the gas molecules in the volume. Figures 3-29 and 3-30 show the latest revisions of the SBS UV chambers' designs. These show vast improvements from the previous designs.

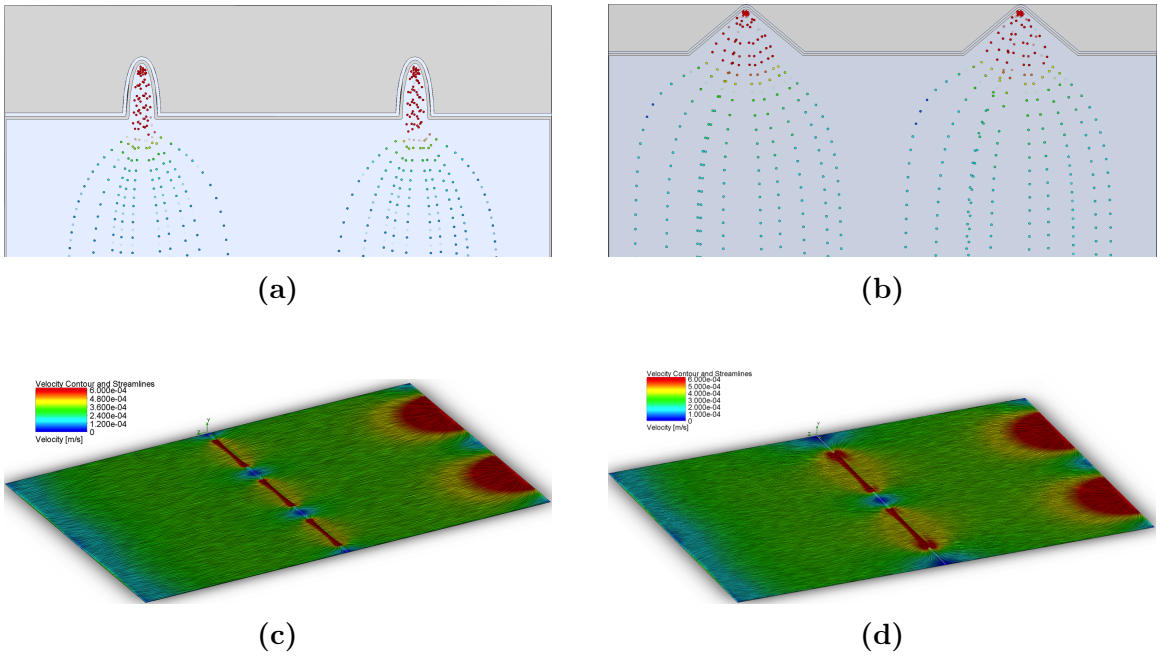


Figure 3-31: Effects of different frame designs on gas flow and circulation. Left: previous frame designs. Right: updated frame design with added features. **(a)** Previous inlet port grooves had a narrower interface. **(b)** Modified inlet port expansion groove allows for faster and more uniform flow distribution. **(c)** Spacers with 3 notches have a larger annular flow area but also have more stress points — these points are highly susceptible to breakage. **(d)** A frame window spacer with two notches allows for better flow uniformity with fewer stress points.

Figure 3-31 shows two iterations of gas flow architecture redesign. Figures 3-31a and 3-31c show an older design that did not provide the desired amount of gas flow, uniformity, and spacer support. The gas inlet flanges were much narrower, and this created higher velocities at the interface to the central chamber volume (See Fig. 3-31a). This increase in velocity distributed the gas to the center more quickly but did

not provide as much uniformity on the inlet side wall as was needed. Figure 3-31b shows the modified revision to the design. The expansion at the gas inlet is noticeably wider and subsequently has a wider gas distribution pattern near the interface wall. In earlier revisions, the mullion spacers had 3 notches to allow for cross-distribution about the chamber. Following analysis, it was found that two longer notches provided much for flow and uniform distribution through the volume. The two-notch design also had fewer stress points in its design and was therefore ultimately stronger and less prone to breakage. These simple changes provided drastic improvements in reliability and performance of our GEMs and have forged their way into our new typical designs.

3.4.5 Aluminized Polyimide Entrance Window

The top layer on each of our GEM detectors is called the Entrance Window. It is a layer of aluminized polyimide (polyimide with a few microns of aluminum on one side). This foil is integral to the general operation and stability of the detector. The layer itself is the "cap" on the gas inside the volume. If it were not for the Entrance Window, the internal volume of the chamber would be exposed to atmosphere above the Drift Cathode — recall that the Drift Cathode has the GEM perforation matrix across it. In previous designs, a simple non-aluminized polyimide layer was used. It was subsequently found that under high levels of charge deposition across surface of the Drift Cathode foil, the polyimide Entrance Window would attract and "stick" to the Drift Cathode. Figure 3-32 shows an older iteration GEM chamber that has a collapsed Entrance Window. The top window is semi-clear and therefore not aluminized — our aluminized polyimide foils are metallic and shiny on the coated side. You can see that the top polyimide layer, or window, is "sucked" to down to the Drift Cathode layer below it.

A drawing of a non-aluminized polyimide Entrance Window is shown in Figure 3-33. The top half of Figure 3-33 shows accumulated charge between the Entrance Window and the Drift Cathode foil. This occurs because of the electric field at the Drift Cathode (-4100 V). Dielectric properties of the polyimide develop an attraction between the two surfaces. This attractive force is so strong that it can "suck" the

Figure 3-32: GEM chamber with non-aluminized polyimide window that has collapsed.

polyimide layer down onto the top of the cathode foil. When this happens, the polyimide physically covers and closes off the gas permeation perforations. This restricts the normal gas flow that moves from above the cathode foil, through the perforations, and then throughout the subsequent volume of the detector. This effectively kills the chamber and makes it inoperable.

The solution is to incorporate an aluminized polyimide in place of the simple pure polyimide sheet. The aluminized variety is a polyimide sheet that has a few microns of aluminum on one side. Figure 3-34 is a drawing of how the aluminized polyimide layer is incorporated as the Entrance Window. The aluminized side of the window is placed on the exterior-facing surface of the detector and a 1 GΩ resistor is connected in parallel between the aluminum and the -4100 V connection to the Drift Cathode

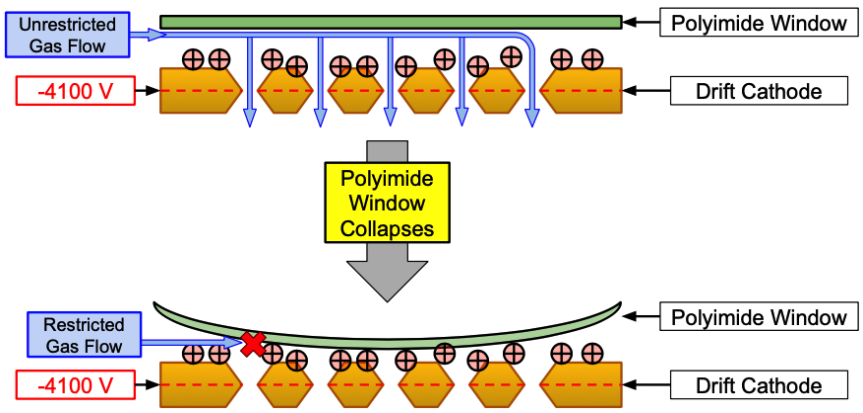


Figure 3-33: Drawing of a non-aluminized polyimide Entrance Window and how it can collapse onto the Drift Cathode foil below it.

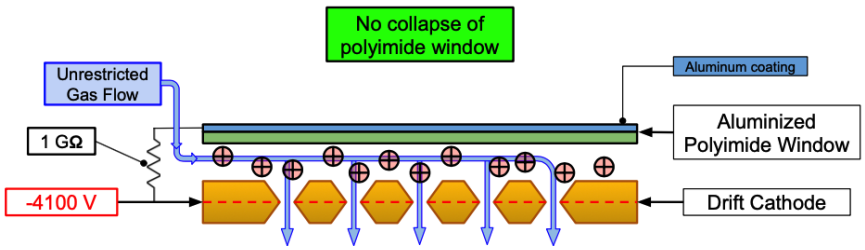


Figure 3-34: Drawing of the aluminized polyimide Entrance Window and how it mitigates collapse.

foil. This creates a small field on the outer surface of the polyimide window which now balances the electrostatic forces which previously caused it to attract to and stick on the cathode foil. The addition of the aluminized polyimide Entrance Window is now a standard feature on all of the GEMs we design.

3.4.6 Two Types of UVa GEMs: XY and UV

Two types of UVa GEM detectors were designed and implemented for GMn and nTPE. These are the so-called XY and UV GEMs. The names of these GEMs are in reference to their readout configuration. The strips of the XY GEM are oriented in a typical Cartesian coordinate setup: X and Y axes that are orthogonal to each other. The UV GEMs have a non-orthogonal and rotated readout configuration: U and V axes with a stereo angle of 60^o between them. Beyond the orientation of their readout strips, the two GEM types have some additional differences.

XY GEMs

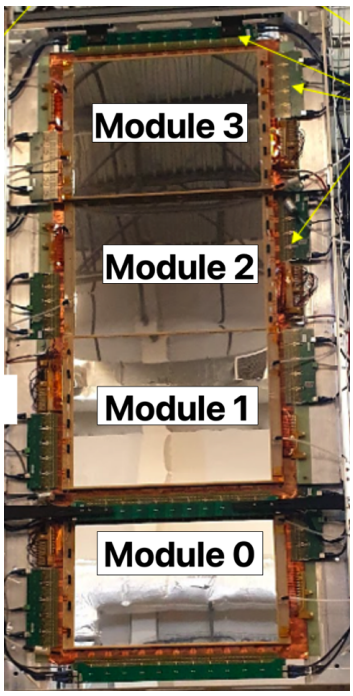


Figure 3-35: An XY GEM layer. This is a composite layer made up of 4 individual XY GEMs labelled Module 0 thru 3, as shown.

The XY GEMs have an active area of 50x60 cm^2 with the longer dimensions inline with the vertical. These GEMs are fitted with a total of 20 APVs: 8 APVs along the horizontal (shorter) length and 12 APVs along the vertical (longer) length. The GEM foils on the XY GEMs incorporate sectorization on their top GEM electrodes only (see Figure 3-21a). There are 10 sectors on each GEM foil layer in an XY GEM. In order to maximize their effectiveness, the XY GEMs were constructed in "composite" layers where a composite single layer was made up of 4 individual XY GEMs. Figure 3-35 shows a picture of a single XY layer. There are 4 modules to an XY layer and they are numbered from 0 through 3, starting with the lowest most module.

The layout and positioning of the 4 XY GEMs in a layer was designed to minimize potential dead active area regions between the overlapping GEM detectors. Figure 3-36 shows a hitmap on an XY GEM layer. The regions for each module are labelled. This particular hitmap was chosen because it contains some problematic sections which highlight the boundaries between individual detectors and makes it easier to

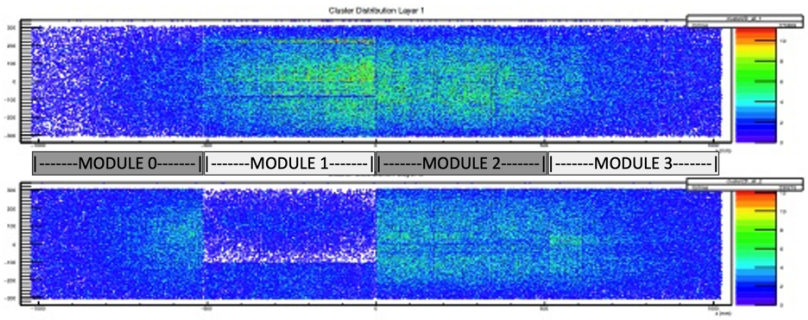


Figure 3-36: Hitmaps on an XY GEM layer. Each layer as 4 GEM modules. The regions for individual detectors are shown and in some regions you can see the boundaries along the hitmap.

see their separations. In nominal operation, these boundaries are not as noticeable and the layer appears uniform and continuous (apart from differences in relative efficiencies per detector).

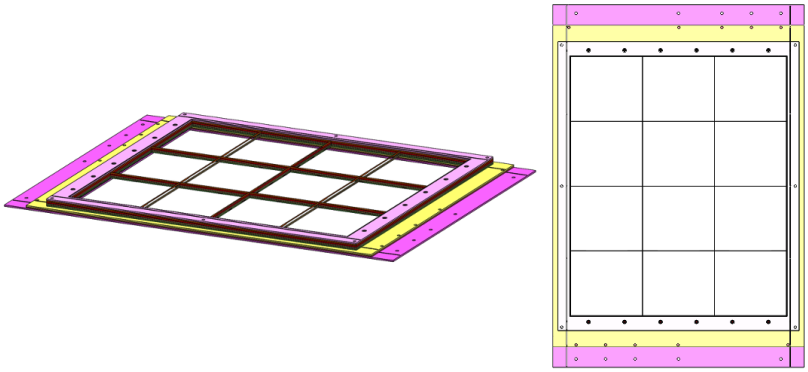


Figure 3-37: Frame layout and design of a typical XY GEM detector.

A CAD drawing of a typical XY GEM frame assembly is shown in Figure 3-37. The frames of this detector were designed with more internal support strips than the UV GEMs. This was possible because the active area portion of these frames were only 50x60 cm^2 as opposed to 40x150 cm^2 for the UV GEMs. The shorter length meant that the XY GEM frames were more rigid and could handle more support strips without as much deflection or warping occurring.

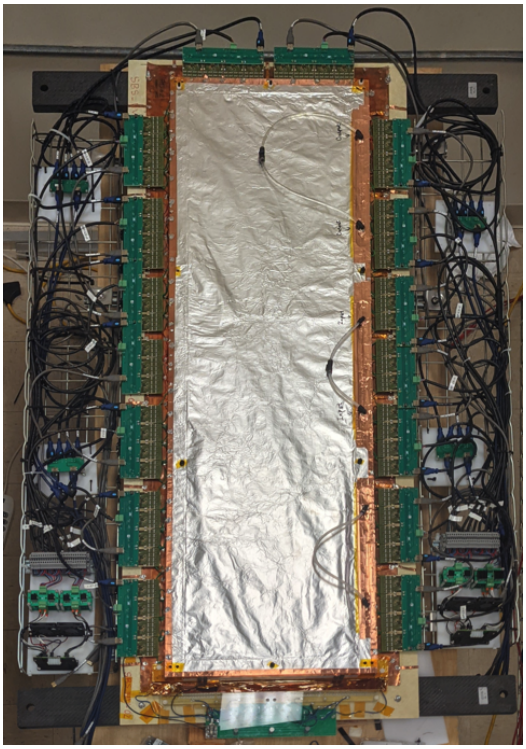


Figure 3-38: A UV GEM with shielding installed, ready for installation in Experimental Hall A.

UV GEMs

The UV GEMs are large detectors that cover a similar active area as compared to that of a composite XY GEM layer. The active area of a single UV GEM is $40cm$ by $150cm$. An overhead photo of an assembled, ready-to-be-installed, UV GEM is shown in Figure 3-38. There are many advantages to using a GEM with a single large active area versus a composite GEM layer. There are no chances of any dead regions due to overlapping GEM detectors or other framing components. The amount of hardware and support material is much less when you are only installing a single GEM. Each GEM detector needs to be rigidly mounted to some form of support frame or structure. Therefore, for each GEM you add to a composite, the more hardware you incorporate per detector. Another advantage of a single large area GEM comes from maintenance and operation. It is much simpler to deal with in-situ maintenance and adjustment when you are dealing with just a single GEM detector as opposed to a tandem set of them, each with their own hookup and hardware.

The primary difficulty in using a single, large GEM detector is the difficulty that arises in its construction, assembly, and testing. Due to flexion and uniformity issues, it is much more difficult to build large GEMs. The smaller GEM frames are more rigid and handle tension of the stretched foils better than the larger UV GEM frames. Extra care must be given when handling and transporting a single large GEM also. Any bending or torsion of the detector through movement can be destructive to these detectors.

For the most part the internal construction, design, and layouts between the XY and UV are the same; however, there are some noticeable differences. The UV GEM was the first of its kind to incorporate sectorization on both top and bottom electrodes of each GEM foil (see Fig. 3-21b). This allowed for loss of sectors without shorting out of the entire GEM foil. An exploded view of a UV GEM (along with the GEM foils within) is shown in Figure 3-14. The streamlined designed of this detector proved very effective, as none of these GEMs have been compromised during in-beam operation at the time of this writing.

3.5 GEM Detector Fabrication and Testing

Each of our GEMs are designed with specific needs and purposes in mind. The XY SBS GEMs are limited in length because of the orthogonal coordinate axes which run parallel to the frame structure of the GEM. If these GEMs were any longer then it would require the readout strips in that dimension to be longer as well. Since we want to match the X and Y strip lengths as much as possible, the length of the GEM is limited (size is 40 cm by 50 cm). The UV GEMs, on the other hand, can be longer since the readout strips do not run parallel to the frame lines but rather at a 30 degrees (60-degree stereo). Also, as we improve the designs and functionality of the chambers through changes in gas flow performance, entrance window material, and frame design, the general form factors of the GEMs evolve as well. We develop all of our GEMs in-house at our lab at the University of Virginia and perform an array of tests on them before giving them the "OK" for production.

3.5.1 Frame Preparation

Figure 3-39: GEM frames hanging to dry in the clean room. The frames have all of their edges and corners thoroughly touched up with sand paper in order to remove any burrs, points, or sharp edges.

The GEM frames are made from a glass cloth reinforced epoxy commonly referred to as G10 fiberglass. The frames are made at custom fabrication shops and arrive at our lab mostly ready for assembly and installation. In order to help protect the extremely sensitive foils and components of the detector, the frames are sanded along all edges to remove any burrs or sharp points. They are then thoroughly cleaned in a hot hydro-sonic bath and hung to dry for at least several days. Figure 3-39 shows sanded and cleaned frames hanging to dry in our clean room. They are hung on a drying rack that has a forced convection system above the hanging frames.

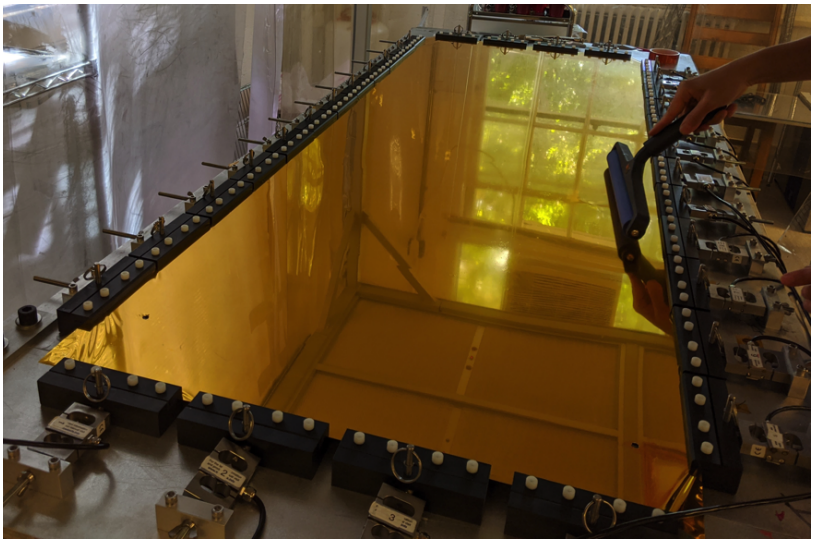


Figure 3-40: Using a static-cling type duster on a stretched Entrance Window.

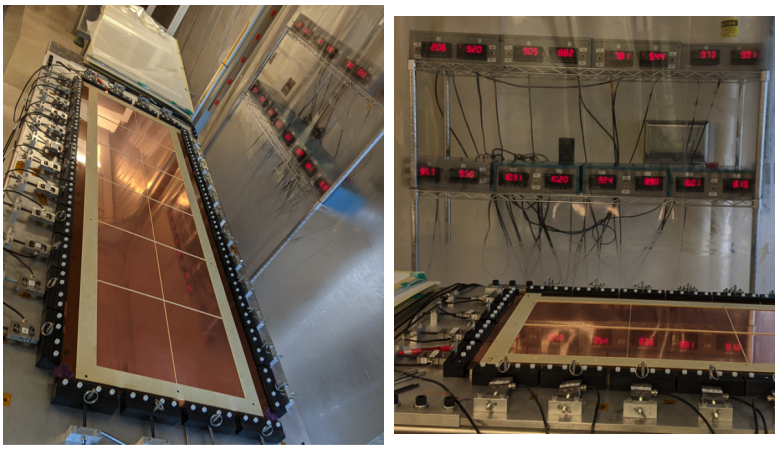
3.5.2 Foil Preparation

The GEM foils arrive at our lab directly from CERN. Even though they are made to fit the constrained form factor of our GEMs, they arrive on sheets that have excess material and thus need to be trimmed to size before assembly. Handling the GEMs is a very sensitive process that requires a delicate touch and a keen eye on cleanliness. All activities involving raw foils, in any manner, are performed in our clean room. Throughout the handling and assembly, all surfaces are continually cleaned and wiped down in order to reduce the possibilities of foreign particles finding their way inside the GEM detector. Even the smallest dust particle can be catastrophic inside of an assembled GEM. Our clean room is equipped with many GEM-safe dusting rollers which use static electricity to pull dust particles off of the foils without applying any force. This allows for easy dust removal without the fear of any physical damage occurring to the sensitive foils. A picture of an Entrance Window being cleaned with a static-cling dusting roller can be seen in Figure 3-40.

3.5.3 Foil Stretching

The subject of stretching foils before and while they are fastened to the GEM frames was introduced in Section 3.4.3. Before being glued to its associated frame,

each foil is mechanically stretched in order to reduce wrinkles and sagging which would reduce the gap between neighboring foils and subsequently increase the risk of short circuit and/or collapse between foils. Though stretching the foils can mitigate wrinkles and sagging, it can also create unwanted deformation in the holes of the perforation matrix of the GEM electrodes [88]. Therefore it is important to apply just enough tension to smooth out the wrinkles in the foil surface without causing any deformation in the holes of the GEM foils.



(a) **(b)**

Figure 3-41: GEM foil stretching table that was custom-built for the UV foils. **(a)** A stretched GEM foil and a frame on top of it. **(b)** The bank of force gauges showing the force on the 16 load sensor-equipped stretching clamps.

A custom-built stretching apparatus was built to apply a measurable force evenly around the perimeter of a foil before and while it is glued to a frame. This stretching table can be seen in Figure 3-41. Figure 3-41a shows a GEM foil stretched in the apparatus. A frame has been glued on top of the foil. There are a total of 32 clamps around the perimeter of the stretching area (Fig. 3-42). You first evenly attach the jaws of the stretching clamps around the edge of the GEM foil. Each of the clamps has a threaded road and fastener which when tightened, pulls the clamp away from the center of the table, thereby applying tension upon the foil held in the clamps. Half of the clamps (16) are fitted with load cells, which measure the tension applied to the GEM foil. Each of the tension adjustment screws are tightened or loosed until the proper tension is set across the surface of the foil. The foils are left in tension

through the frame gluing and application process.

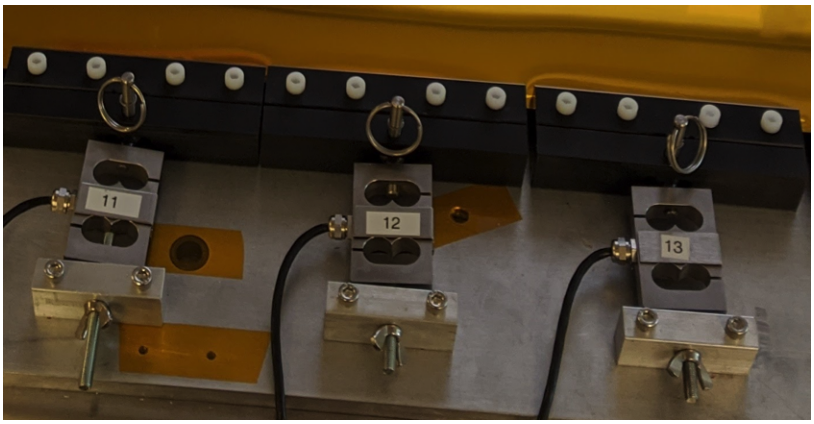


Figure 3-42: Clamps of the stretching table that are fitted with force measuring load cells.

3.5.4 Assembling the Detector

The layout for the stack of frames for a UV GEM chamber is shown in Figure 3-14. Each of these frames is glued on top of the foil just below it. The Readout Board is glued to the Readout Support Frame, GEM Foil 3 is glued to the Bottom Support Frame, and so forth. The Readout Board is not stretched, but it is weighted down during application to its support frame. Each of the GEM foils, cathode, and windows are held in tension on the foil stretcher as a frame is glued to it.

The frame and foil shown in Figure 3-41 had already been glued together when those photos were taken. A specially-formulated two-part epoxy is applied to each of the frames before it is laid atop its respective GEM foil (which is held in tension until the adhesive has completely dried). The stretching table is fitted with guide pins, which ensure the proper alignment and orientation of the foils and frames as they are glued together. Once the frame is in place on the foil, it receives a smooth and protective plate on top of it. This plate allows weights to be distributed across the components so that there is uniform adhesion between the foil and frame.

The layers of frames and foils are built from the bottom up, starting with the readout board on the main support frame. The next layer would be stretched GEM Foil 3, which has already been glued to the Bottom Support Frame. Once the epoxy

is applied to the bottom surface of GEM Foil 3 and the sub-assembly is laid atop the Readout Foil, a weight-distribution plate is again used to ensure uniform adhesion. This process is repeated one foil-framed sub-assembly at a time. The final layer is the framed aluminized polyimide entrance window. A stack of glued foils and frames on a readout board is shown in Figure 3-43.

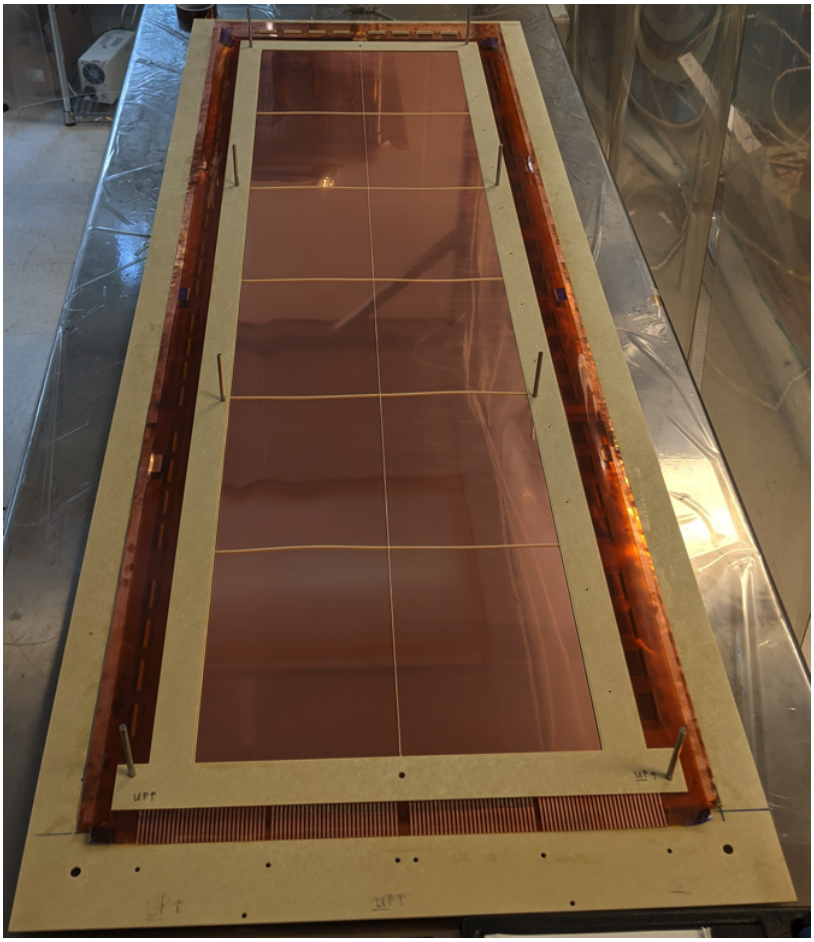


Figure 3-43: Multiple framed foils stack on a UV GEM readout assembly.

3.5.5 High-Voltage Testing

Throughout the entire gluing and assembly process, the primary GEM foils undergo strict high-voltage testing. The readout, GEM, and cathode foils are very sensitive items which can be compromised and/or damaged easily. Though we worked diligently to handle everything with the utmost cleanliness and care, it is inevitable

that damage and/or malfunction of a foil will occur during construction of the detector. Conducting these tests helps to track the vitality of the components through each step.

The test involves applying 550V across a set of sectors on a GEM foil as the current is measured and stored with a pico-ammeter. The sectors should be isolated from one another and have a typical capacitance of about 4.2 μF. When high voltage is applied, a very sharp spike in the current should appear and then dissipate off as the charge has accumulated. A typical malfunction would have the sectors appear as a short-circuit. Instead of the typical capacitive spike and discharge, the current would jump sharply and remain steady at that elevated voltage. High voltage application in this manner also provides the possibility to "burn off" any foreign particles which may be present on or in the foil. The fast application of the voltage and current can effectively remove anything that may become problematic down the road. These tests are performed on every GEM foil sector before and after they are glued to a frame. Once glued, the framed foil gets installed into the GEM detector. Final sector tests are performed once the entire chamber has been assembled and sealed.

The GEM foils are sensitive to humidity in the air and require this testing to be performed in a dry nitrogen-rich environment. Without this, excess sparking and discharging would occur and could damage the foils and components. A custom high voltage test compartment that could be purged and filled with nitrogen gas was constructed for all of these tests. A GEM foil under test in this box can be seen in Figure 3-44.

The GEM foil is placed within the testing box and connected to a specially-designed electrical connector which matches the sector traces on the GEM foil. The connector consists of pins, which make contact with each of the traces on the foil (Figure 3-45b). The pins are attached to a mounting plate fitted with guide pins which aligns it properly in place (Figure 3-45a). Once the connection pin array is engaged onto the GEM foil, it creates a connection between the GEM sectors and a panel of connectors on the outside of the test box. This allows for testing of the sectors without opening the box. This is necessary, because during testing the box is

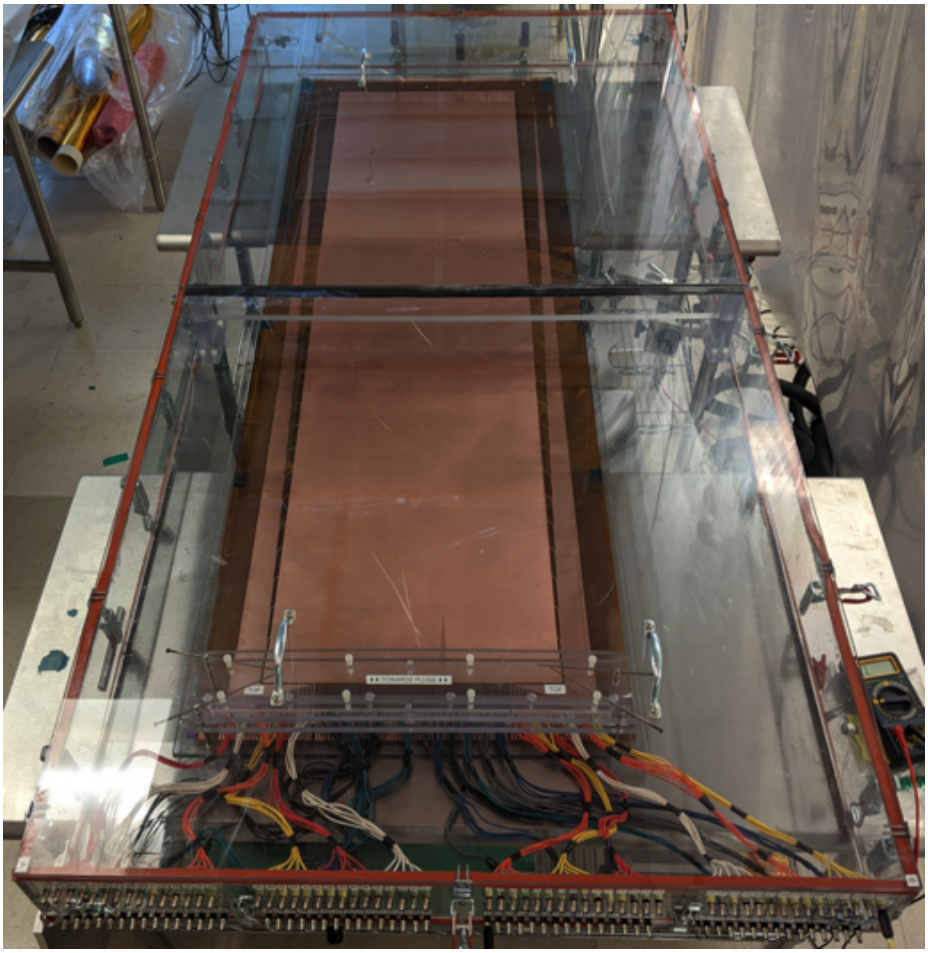
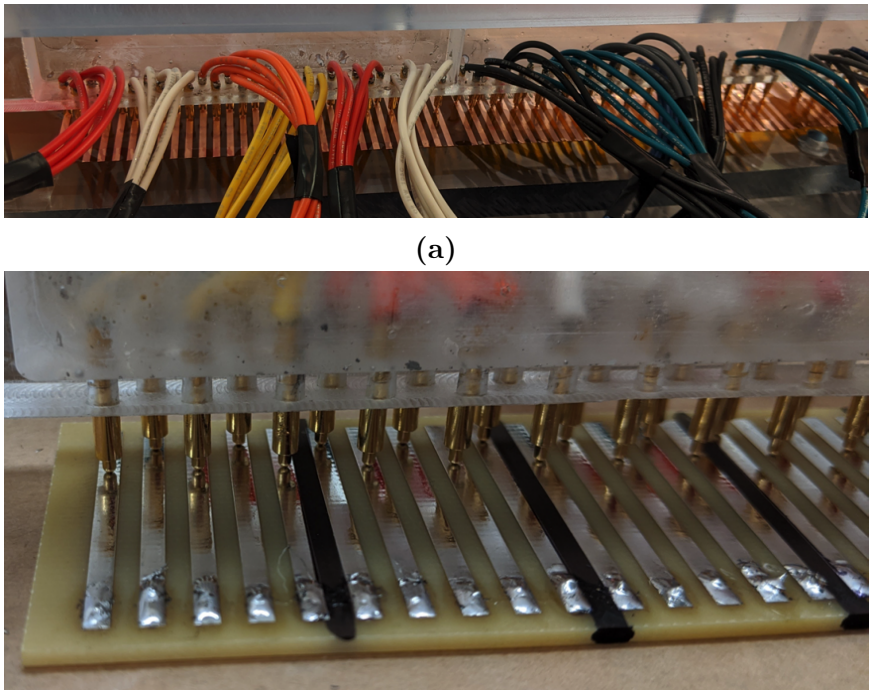


Figure 3-44: Testing a raw UV foil in the nitrogen-filled high-voltage box.

sealed and filled with nitrogen gas. The test box if fitted with inlet and outlet hose connectors which allow nitrogen gas to be fed into the box to create the required dry oxygen-free environment.

The test box was only designed to fit at maximum, a stretched foil glued to a frame. As the framed foils are glued to the detector stack they are no longer tested inside this box. Once the entire set of layers has been built up and the detector assembly is complete, each sector on all GEM foil layers is tested again. At this point, the detector is a sealed unit with its own input and output gas flow connections. The internal volume of the GEM detector is then filled with nitrogen for subsequent testing. This means that these tests can also be performed outside the testing box and even on a bench top. The same tests are performed and logged. At this point it



(a)

(b)

Figure 3-45: Connections to the GEM foil in the dry nitrogen HV test box. **(a)** The pin mounting plate which aligns the pins with each of the traces on the GEM foil. **(b)** Close-up of the pins engaged to a mock-up set of GEM traces.

is important to note the location of any "dead" sectors. However, this will be more self-evident using hitmaps from subsequent cosmic tests.

3.5.6 Cosmic Testing

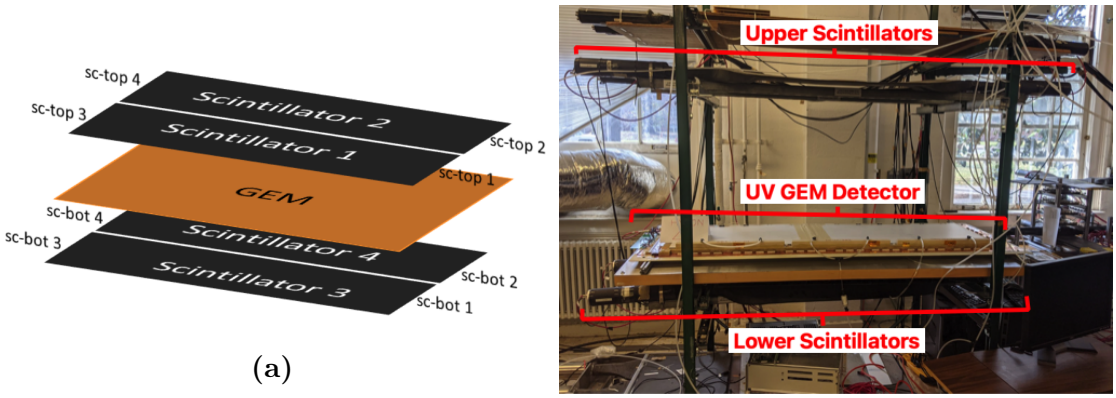
Once a GEM chamber is fully assembled and tested for stability, it can then undergo what are called "cosmic tests" or simply "cosmics". This fundamental test of the detector is the first opportunity to extract two-dimensional plots of particle hits upon the GEM chamber. These tests are called "cosmics" because they are not performed in-beam but rather utilize the abundant cosmic-ray muons (or atmospheric muons) that shower down from the atmosphere. These cosmic-ray muons are produced when a high energy galactic particle collides with nuclei in Earth's upper atmosphere. These interactions produce showers of particles which in turn produce more showers, and so on. Some of these post-collision particles are secondary mesons

(positive or negative pi mesons) which subsequently decay into positive or negative muons. These muons are short-lived particles with relatively high-energy and can thus penetrate quite well and easily reach the Earth's surface. Typical cosmic-ray flux at the Earth's surface at sea level is approximately $^{10{,}000}/_{(minute \cdot m^2)}$ [69]. These high energy particles are what allow us to perform table-top particle detection and coordination.

Though this test can be performed with the GEM detectors in any orientation, the highest cosmic muon fluxes occur for a horizontally (parallel to the ground) oriented GEM chamber [14]. Typical GEM cosmic tests are a relatively slow process with data-acquisition trigger rates typically on the order of 2 Hz for a detector that is approximately 50 cm x 50 cm. A useful cosmic measurement on a GEM will typically have several thousand triggers. Therefore, most of our in-lab tests have occurred with the GEM detectors laying flat. However, the SBS "XY" GEMs for the GEn-RP and related experiments were tested in a vertical position with positive results.

Any coordinate-based GEM measurement (tests not including HV sweeps, current draw measurements, etc.) involves a power supply system, a data-acquisition system, and the detector under test. The power supply system contains the voltage supply and control units for most of the data-acquisition system and any other incorporated electronics hardware. The data-acquisition system is comprised of a triggering system, a hardware interface setup, and a computer control and storage system. A more in-depth look at a GEM DAQ system is given in Section 2.7. For our initial tabletop cosmic tests there was a set of triggering scintillators situated above and below our GEM detector (See Figure 3-46a). Each pair of scintillators were configured in a logic OR. Then, the OR output from each pair was placed in a logic AND. Therefore, if a cosmic muon were to pass through either of the top pair of scintillators and then through either of the bottom pair of scintillators, an analog Boolean TRUE would be sent out as a trigger to our DAQ system telling it to record an event. A photo of a UV GEM on the UVa-based cosmic test stand is shown in Figure 3-46b.

Due to a limiting number of APV25 Readout Cards at the time, our initial cosmics tests on the UV GEMs required three separate measurements in order to get a



(a) (b)

Figure 3-46: UVa UV GEM cosmic test setup. **(a)** Schematic of the cosmic setup for scintillators and GEM chamber. Two scintillators were placed above the GEM detector and two scintillators were placed below the GEM. **(b)** Photo of the GEM detector on the lab test bench with the pairs of scintillators above and below it.

full scan across the entire active area of the detector. We only had enough APV25 readout cards to cover a third of the detector at a single time. Therefore, for the first test, the APV25 readout cards were placed on one end of the detector for a measurement. Then, the readout cards were physically moved to the center of the detector for a second measurement. The cards were then shifted down the detector in order to make the final measurement. These three measurements were then overlayed with each other in order to produce the first set of cosmic readout measurements on our UVa-built UV GEMs. Figure 3-47 shows the very first set of cosmic measurements performed on the then brand new UV GEMs. The figure shows how three measurements were taken and then spliced in order to show the measurements across the entire active area of the detector. These measurements gave us the first signs of life from these detectors. Not only could they successfully hold high-voltage without shorting out, but they could also coordinate charged particles that entered the chamber and cascaded down to the readout board. Cosmic tests are a very useful tool for determining the vitality of a GEM detector for particle tracking. These tests are implemented often because they are fundamental and straight-forward.

In the rightmost plot of Figure 3-47 we can easily identify dead GEM foil sectors and non-functioning Readout Board strips. These plots are histograms for detected

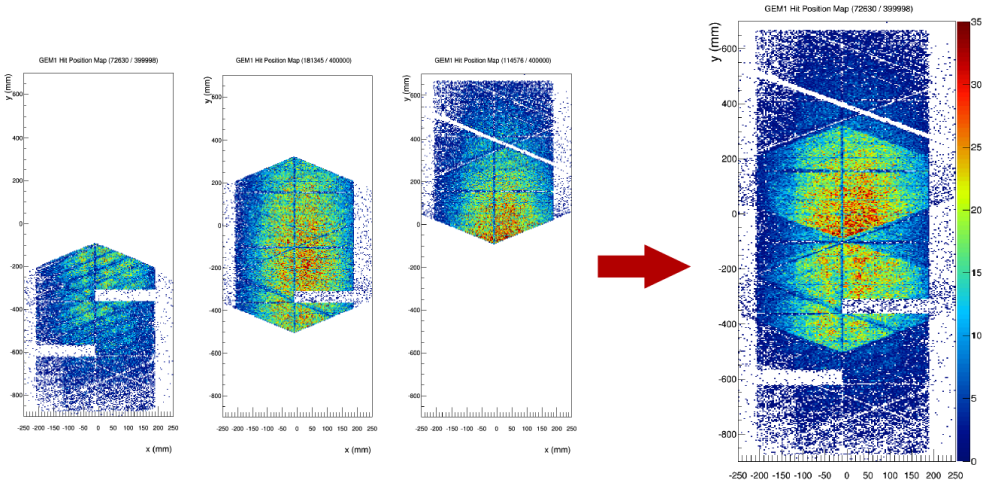


Figure 3-47: The first UV GEM cosmic measurements. Due to limited hardware the cosmic tests were taken with readout cards placed at three positions along the GEM detector. Afterwards, the three plots were then spliced together to display the measurements across the entire chamber.

particles on the detector and so, the colors indicate the number of events recorded across the detector surface. The dead sectors show up as mostly white rectangles in the plot(s). Recall that on the UV GEM the foil sectors have this rectangular shape and orientation (see Figures 3-17 and 3-18). Disconnected, dead, or shorted Readout Board strips appear as the diagonal white lines. Recall that the UV GEM Readout has U and V strips at a 60^o stereo angle.

3.6 Experimental GEM Setup in Hall A

The experimental setup for the GMn and nTPE experiments consisted of three primary arms: the incoming electron beam, the Electron Arm, and the Hadron Arm. See Figure 3-48 for reference. The electron and nucleon arms are both downstream of the target. The incoming electron beam impinges on the deuteron target, after which the electron will scatter down the Electron Arm and the nucleons will traverse the Hadron Arm. Each of the scattering arms contains a magnet used to deflect charged particles. Because of this, each arm is also referenced by each magnet's name. Along the Electron Arm is the BigBite magnet, and so this arm is often called the BigBite (or BB) arm. The detectors along this line may also get referenced

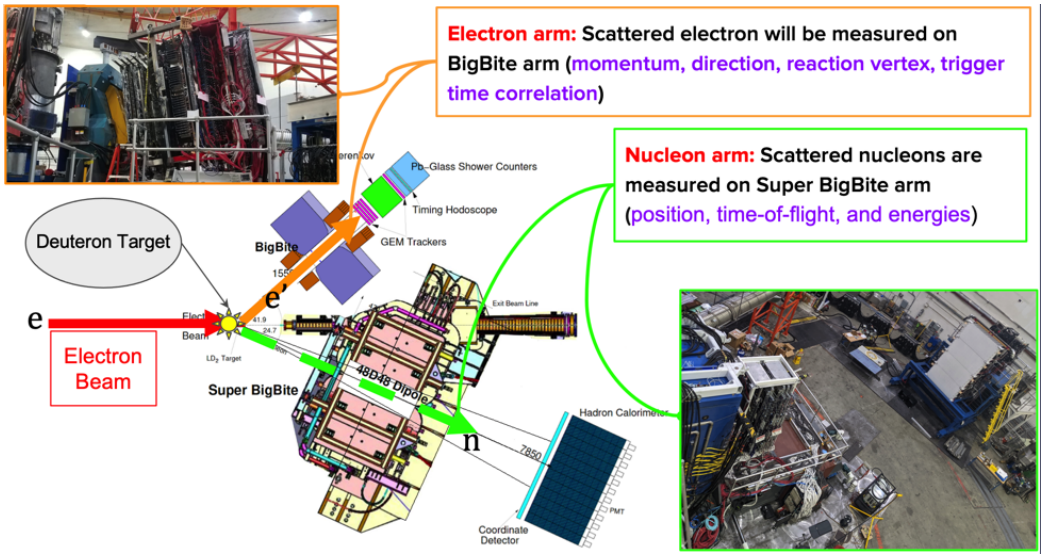


Figure 3-48: Experimental setup for GMn and nTPE with the three primary arms of the experimental highlighted.

in terms of their position along the BigBite arm. For instance, the entire set of detectors on the Electron Arm may be referenced as the BigBite (or BB) detector stack, the BB detectors, or even simply BigBite. Similarly, the nucleon arm contains the SuperBigBite magnet and so it is often referenced with this title. The nucleon (or hadron) arm is much simpler in that it only contains two primary components (the SuperBigBite magnet and the Hadron Calorimeter). However, the aforementioned naming convention may still be used.

The Electron Arm of the experimental setup contains most of the sub-detector packages for GMn and nTPE. The BigBite arm is used to extract multiple variables from the scattered electron: momentum, direction, trajectory, reaction vertex, and trigger time correlation to name a few. Information from the scattered electron arm can be used to make projections of scattered nucleons on the Hadron Arm. These projections help to isolate search regions along HCal. This minimizes the amount of data required since we only need to maintain HCal data within the projected search region. It also makes finding the actual hits on HCal easier since we only need to search within that small region, as opposed to the entire face of HCal.

3.6.1 GEMs on the BigBite Stack

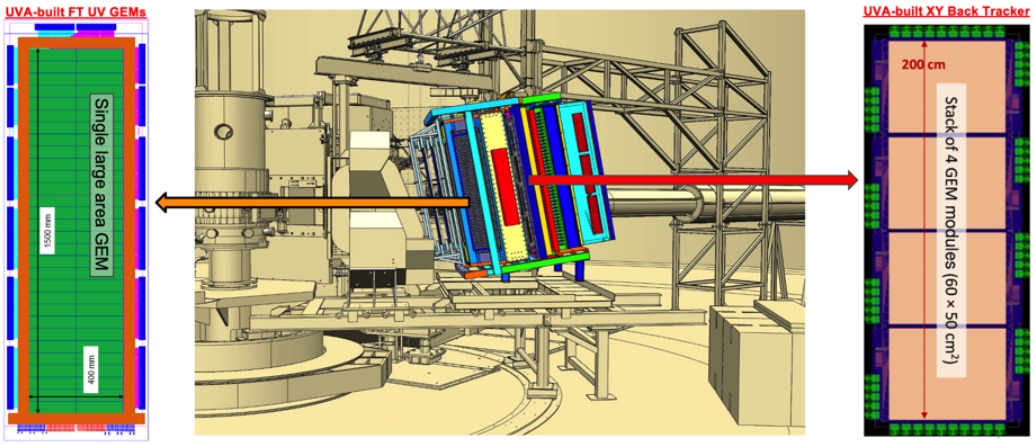


Figure 3-49: The BigBite detector stack and the location of the Front and Rear GEM trackers. Sub-image source: [78].

All of the GEMs for the GMn and nTPE experiments were installed on the BigBite stack along the scattered electron arm (See Fig. 3-49). The BigBite consisted of two sections of GEM detectors. The GEM Front Tracker were GEM layers at the very front of the BigBite detector stack. These were the first components to interact with the scattered electrons after they passed through the aperture of the BigBite magnet. Immediately downstream of the Front Tracker GEMs was the GRINCH detector and just downstream of that were the Back Tracker GEMs. See Section 2.4.1 for more detail.

In total, there were five independent layers of GEM detectors along the BigBite detector package. Initially, the BigBite GEMs were those provided by our research group at UVa and some that were provided by the GEM group at INFN-Roma. Due to certain experimental circumstances, only UVa-developed GEMs were present during the nTPE experiment. In total, there were two GEM layer types: single large area GEM layers and composite GEM layers. The single large area GEM layers consisted of a single detector that comprised the entire layer. These GEMs are the so-called UV GEMs. The composite GEM layers each consisted of either three (INFN-Roma) or four (UVa) GEM detectors constructed onto a single holding frame and configured to provide a unified active area across each detector. Figure 3-49 shows the two UVa

GEM types. On the left is the single, large area UV type GEMs which were used in the Front Tracker. On the right side of the figure is shown a UVa XY type GEM layer which consists of four independent GEM detectors situated together to form a single layer.

The integration of XY and UV type readout configurations is an important attribute of this detector stack layout. The combination of UV and XY layers allows for a decrease in combinatorics when isolating search regions for reconstructing particle tracks. Added ambiguity arises when only a single readout coordinate configuration is used. Introducing a second coordinate system provides an independent set of position variables which can be used when searching through hit regions along multiple readout strips across multiple GEM active areas.

Figure 3-50 shows a more detailed drawing of the GEM layers and their locations amid the BigBite detector package. In this figure, the scattered electron enters the detector stack from the right after having scattered from the deuteron target upstream. As the charged electron passes through the layers of GEMs, it will be detected along each subsequent layer. The GEM layers are configured parallel to each other and separated along the path of the scattered electron. Figure 3-51 shows each of the five BB GEM layers as single planes. These are hitmaps which show the positions of highest concentration for detection of a charged particle. The red regions near (-0.18, -0.1) mark the areas of highest concentration. If we take the measured charged particle positions on each GEM layer and correlate them along the axis of the scattered electron, we can then reconstruct the particle's trajectory.

We can visualize this by viewing these hitmaps as planes in three dimensional space as shown in Figure 3-52. The individual hit positions on each layer can be correlated along the path of the scattered electron. A straight-line projection can then be formed for the approximate track through the detectors. This can furthermore be correlated with measurements from the BigBite Calorimeter at the back of the detector stack. This calorimeter will measure the energy deposited by an electron in one of its detector cells. The position of this fired cell along the face of BBCal can then be correlated with the particle track through the GEM layers in order to reconstruct a refined and

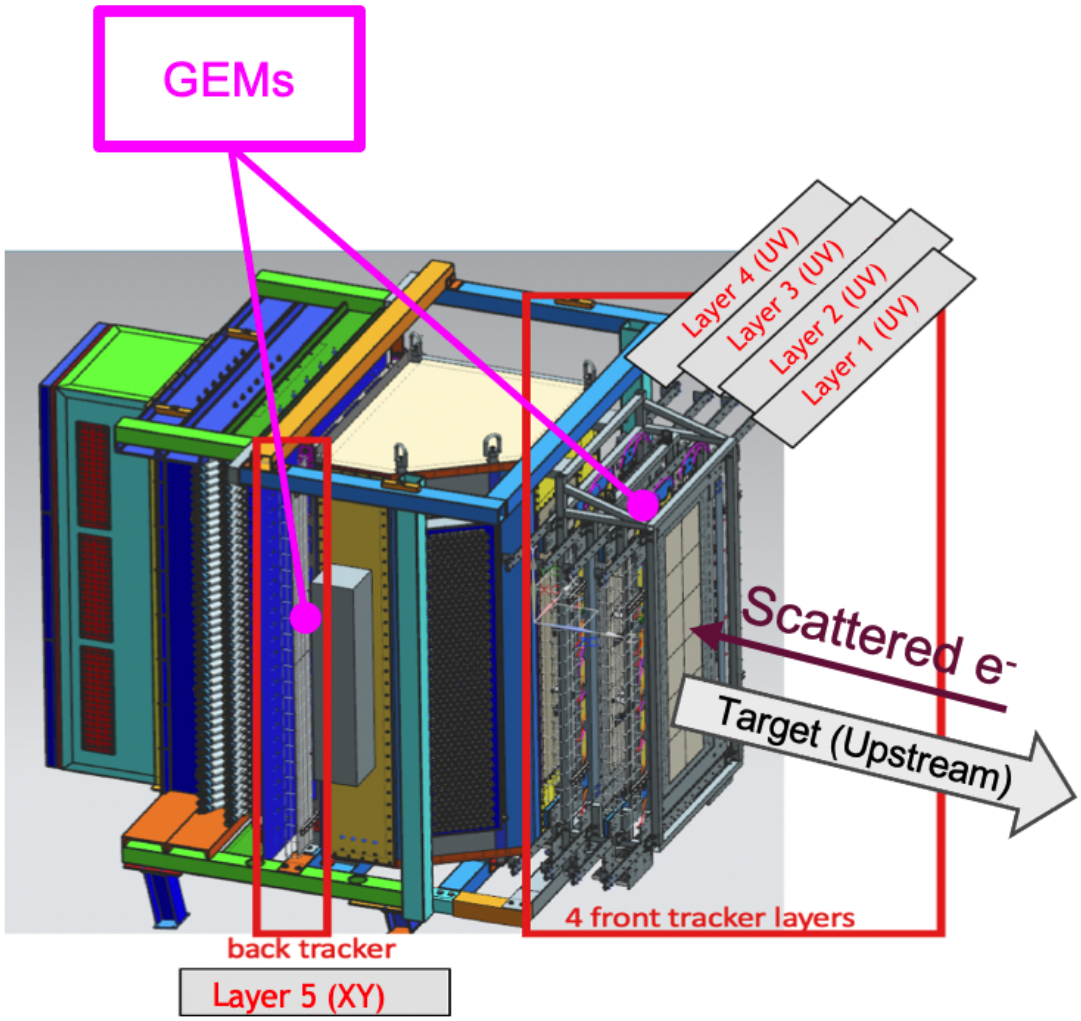


Figure 3-50: GEM layer positions on the BigBite stack. Direction of incoming scattered electron is shown. Location of upstream target is noted. Image source: [78].

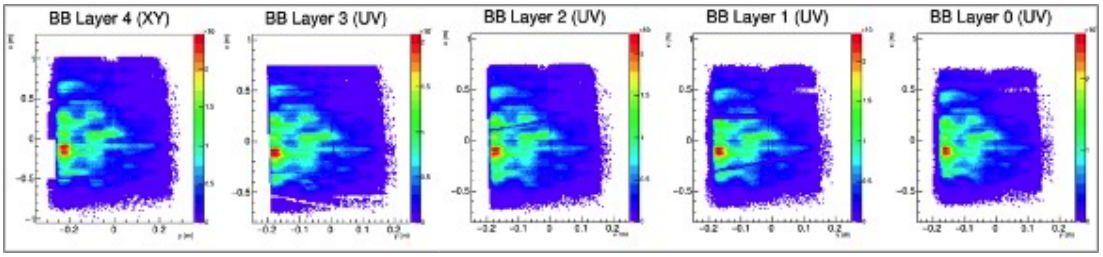


Figure 3-51: The five GEM layers of the BigBite stack. These hitmaps show the positions where the highest concentration of charged particle were detected on each layer in red.

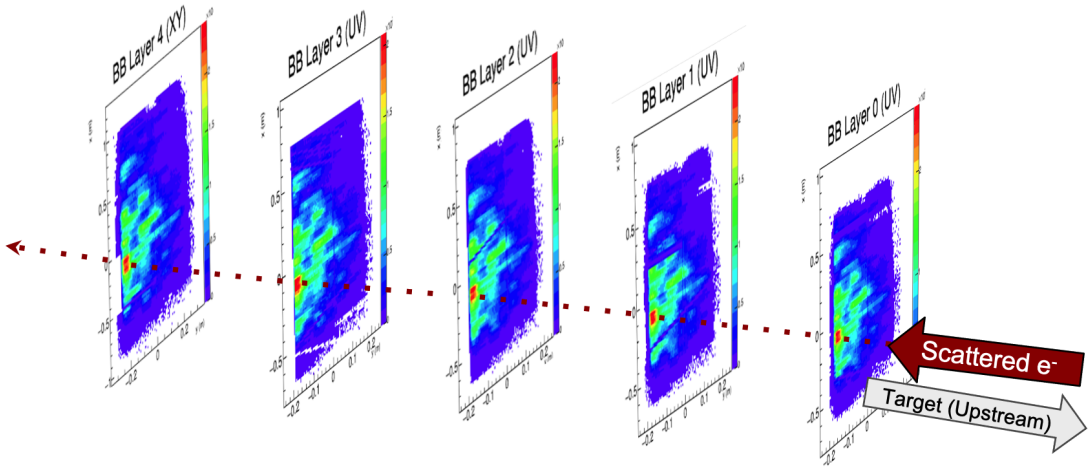


Figure 3-52: Visualization of track reconstruction across the multiple GEM layers of the BigBite detector stack.

accurate scattered electron trajectory.

3.6.2 Gas Distribution System for the GEMs

Each of the GEM detectors in this setup has its own internal gas volume which flows a mixture of argon and carbon dioxide through the detector at a specific rate. The replenishing rate of the chamber is critical to its performance and lifespan (see Section 3.2.4). Therefore, a gas distribution system was specifically designed for the GEM detector setup in Hall A. The primary components of the system were the gas sources (argon and carbon dioxide cylinders/bottles), a flow manifold and pressure regulator system, an online monitoring system, and the connections to the GEM detectors. A schematic of the GEM gas distribution system is shown in Figure 3-53.

In order to allow for bottle change-out of the argon and/or CO2 during normal beam operations, the gas cylinders were placed in a gas shed on the premises of JLab (a location near the Hall A Counting House). To avoid using costly pre-mixed cylinders of Ar/CO2, two separated gas bottles were used. In order to get the desired ratio of argon to CO2, a line from each gas source was fed into a mixing system. Downstream of this, the mixture ratio could be monitored online to ensure that a proper ratio was provided to the detectors. The precise Ar/CO2 gas mixture was then routed to a regulator panel (see Figure 3-54). This regular panel consisted of a primary

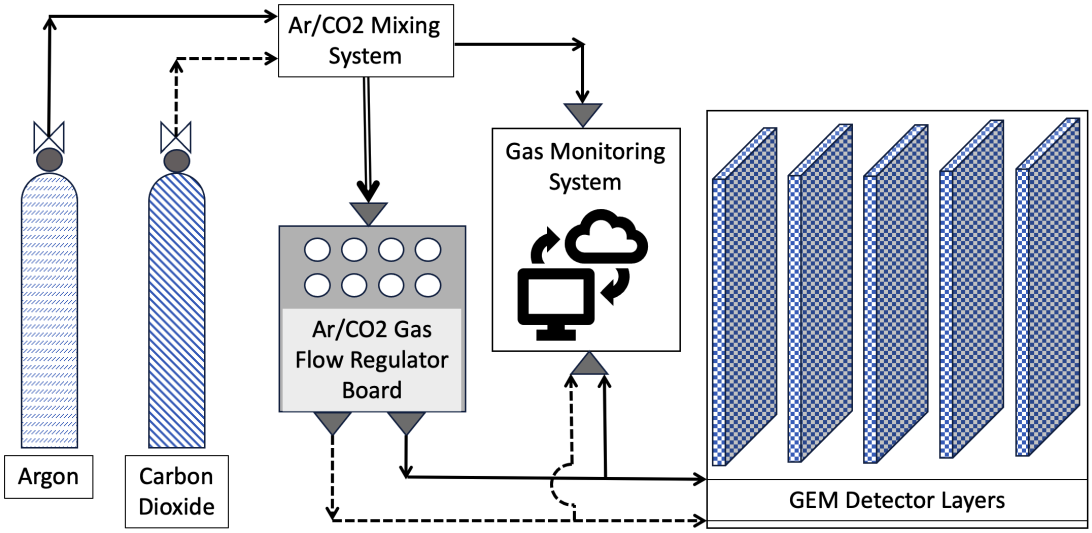


Figure 3-53: Schematic of the gas distribution system for GEM detectors on the BigBite stack.

pressure regulator to reduce the initial tank pressure from the gas cylinders down to a safe operating pressure for the detectors downstream. From there, a manifold was used to send a separate gas line to each GEM detector. Each gas line had its own independent flow regulator which allowed for precise control of the gas flow rate to each individual GEM detector. Each of these measurements were configured for live readout via an applet or an online interface. This allowed for easy detection of unwanted variations to flow rates or mixing ratios. Filters were installed at the inlet point of each GEM layer. This prevented any particulates from entering the GEMs and causing catastrophic failure.

3.6.3 GEM High Voltage Power Supply

The operating principle of our GEM detectors relies on large electric potentials across each GEM foil. For a single detector, this sums to approximately 3.7 kV at a nominal 745 μA. In order to achieve and maintain this, each GEM was powered with a dedicated high-voltage module rated at 6 kV and 1 μA per channel. The power supplies were W-IE-NE-R MPOD EHS 8060n installed into a W-IE-NE-R controllable crate. Each MPOD module had 8 channels which each met the necessary

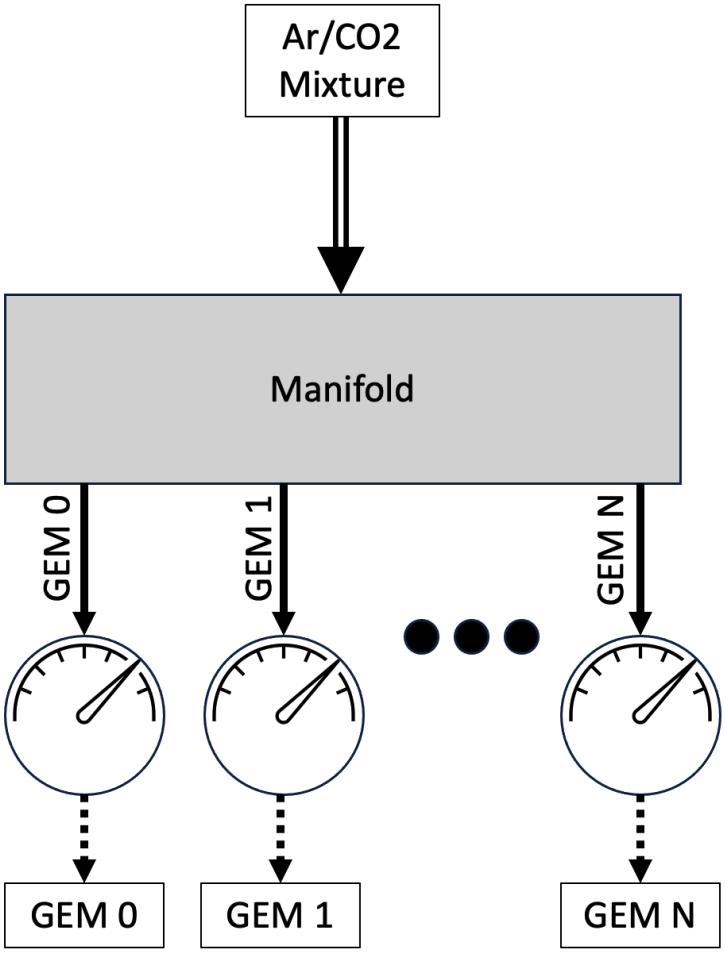


Figure 3-54: Schematic of the gas manifold for routing the Ar/CO2 mixture to each GEM at its required flow rate.

power requirements. A single SHV supply cable is connected from a channel at the power supply to the high-voltage divider board of each GEM (Fig. 3-25). W-IE-NE-R crates allowed for live monitoring and control of system power, voltage settings, current limits, voltage/current ramp rates, and various other parameters. This was a critical function due to limited access to the experimental hall during beam operations. It was often necessary to adjust various parameters for certain beam conditions. As the beam current would increase so too would the current draw for each high-voltage power supply. Therefore, it was necessary to monitor and adjust the over-current trip limits with any changes in beam condition. This equipment also allowed us to log voltage and current in a table or strip chart during operations.

3.6.4 GEM Low Voltage Power Supply

The readout hardware for our GEM detectors are the APV25 cards. These cards require their own power source in order to function properly. Under nominal conditions each APV requires three supply voltages: VDD = $+1.25V$ ($+90\ mA$, GND = $0V$ ($+64\ mA$), and VSS = $-1.25V$ ($-155\ mA$). Each channel of the APV consumes approximately 2.31 mW. Therefore, for all 128 channels on a single APV the total power consumption equates to approximately 296 mW [72].

In order to provide the proper voltage to the APV readout cards, a low voltage regulator board, originally designed by CERN, was implemented. The board takes in a $5V$ DC supply (which in our case was supplied by a 120V voltage power supply) and converts it to $+2.5V$ and $+1.25V$ against a common ground. The board was designed in such a way to reference these voltages against the reference such that it matched the requirements stated above for VDD and VSS. The low-voltage regulator boards were originally designed for use at CERN in radiation-sensitive environments. Therefore, they were radiation tolerant and could operate smoothly even while installed immediately next to the GEM layers.

Chapter 4

Experimental Technique

This chapter discusses the experimental methods and techniques for extracting exploratory results from data-tuned simulations. To achieve this, we need to combine our physical observables from experimental data with those from Monte Carlo simulations. This involves some slight re-formulation of the purely theoretical models presented in earlier chapters. It is imperative to approach the analysis and interpretations presented here with care and to recognize the exploratory nature of this thesis.

4.1 Experimental Kinematics and Configurations

The GMn and nTPE experiments ran *in series* with one another and used the exact same experimental setup in terms of detectors, hardware, and electronics. In the final run plan, there were a total of six kinematic configurations: SBS4, SBS7, SBS8, SBS9, SBS11, and SBS14. The principle difference between each of these kinematics are their Q^2 values. The nominal range for momentum transfers was 3.0 $\leq Q^2 \leq$ 13.6 $(\text{GeV/c})^2$. Each configuration used the same experimental components, but were physically reconfigured[1] for various beam energies, scattering arm angles, target-to-calorimeter distances, etc. Unfortunately, the number in a configuration's name — the "4" in SBS4, for instance — does not necessarily correspond to its Q^2 value or beam energy. The kinematics are also not numbered chronologically in terms

of, say, ascending (or descending) energy, momentum, or even position in the final run plan. In fact, they were named based on their positions in the original run plan. Due to various reasons, the original run plan needed to be modified and the configuration lineup changed. The original naming convention was maintained even though the kinematic ordering had been altered.

Technically speaking, the GMn kinematics are 4, 7, 9, 11, and 14, while the kinematics for nTPE are 8 and 9. SBS8 and SBS9 have the same Q^2 value of 4.5 $(\text{GeV/c})^2$. However, they have different electron scattering angles corresponding to two different epsilon values: $\epsilon_1 \approx 0.838$ (SBS8) and $\epsilon_2 \approx 0.599$ (SBS9). In order to arrive at my result for nTPE, I rely on an extraction of G_M^n at each nTPE point. Therefore, SBS8 will be included as part of my GMn kinematic set[2].

The nominal range of beam energies was $3.70 \leq E_{Beam} \leq 9.91$ (GeV). The solid angular acceptance for the BigBite arm is between 30 and 60 msr. The nominal experimental range of electron scattering angles was $26.5^o \leq \theta_e \leq 49.0^o$. The Super BigBite arm has a solid angle range between 15 and 45 msr. The nominal scattered hadron angles had a range of 13.3^o to 31.9^o [139]. In order to achieve a given Q^2, we simply need to define the electron beam energy and the electron scattering angle. From there, we can solve for the subsequent kinematic variables (hadron scattering angle(s), nucleon momentum(s), electron momentum, etc.). Once the scattering angles have been determined, the scattering arms can be moved to those positions. Then, the detectors are moved to positions along their respective scattering arm which will maximize their acceptance regions. The nominal and experimental kinematic configuration parameters are listed in Tables 4.1 and 4.2, respectively.

[1]The setup in Hall A was designed so that each scattering arm could be moved and (re)configured for multiple physical configurations. Figure 2-3 shows a photo of Hall A with grey semi-circles in the lower portion of the photo. These correspond to *tracks* which allow the detector packages to move and change their positions in the hall.

[2]The extraction of G_M^n via the ratio method does not differentiate between varying epsilon values at a similar Q^2 point. Therefore, it is a bit redundant for SBS8 to be included in the nominal GMn kinematic set. However, I think it serves as a great cross-check between methods and results since we can directly compare it with SBS9. For this reason, I will include G_M^n for SBS8 with my results for GMn.

Name	Program	E_{Beam}(GeV)	Q^2 (GeV/c)^2	θ_{BB} (Deg)	θ_{SBS} (Deg)	HCal Dist (m)	p_N (GeV/c)	p_e, (GeV/c)
SBS-4	GMn	3.70	3.0	36.0	31.9	11	2.35	2.11
SBS-7	GMn	8.00	10.0	40.0	16.1	14	6.20	2.67
SBS-8	nTPE	6.00	4.5	26.5	29.9	11	3.22	3.59
SBS-9	GMn, nTPE	4.03	4.5	49.0	22.5	11	3.21	1.63
SBS-11	GMn	9.91	13.6	42.0	13.3	14.5	8.13	2.67
SBS-14	GMn	6.00	7.5	46.5	17.3	14	4.84	2.0

Table 4.1: Nominal experimental configurations for the GMn and nTPE experiments.

Name	Program	E_{Beam}(GeV)	Q^2 (GeV/c)^2	θ_{BB} (Deg)	θ_{SBS} (Deg)	HCal Dist (m)	p_N (GeV/c)	p_e, (GeV/c)
SBS-4	GMn	3.7393	3.0326	36.0	31.9	11	2.373	2.122
SBS-7	GMn	7.9308	9.8842	40.0	16.1	14	6.20	2.6636
SBS-8	nTPE	5.9826	4.6299	26.5	29.9	11	3.22	3.5826
SBS-9	GMn, nTPE	4.0268	4.5047	49.0	22.5	11	3.21	1.6263
SBS-11	GMn	9.889	13.5514	42.0	13.3	14.5	8.10	2.6675
SBS-14	GMn	5.9828	7.4686	46.5	17.3	14	4.80	2.0028

Table 4.2: Experimental measurements for each kinematic configuration of the GMn and nTPE experiments.

4.2 Particle Deflection on the Scattering Arms

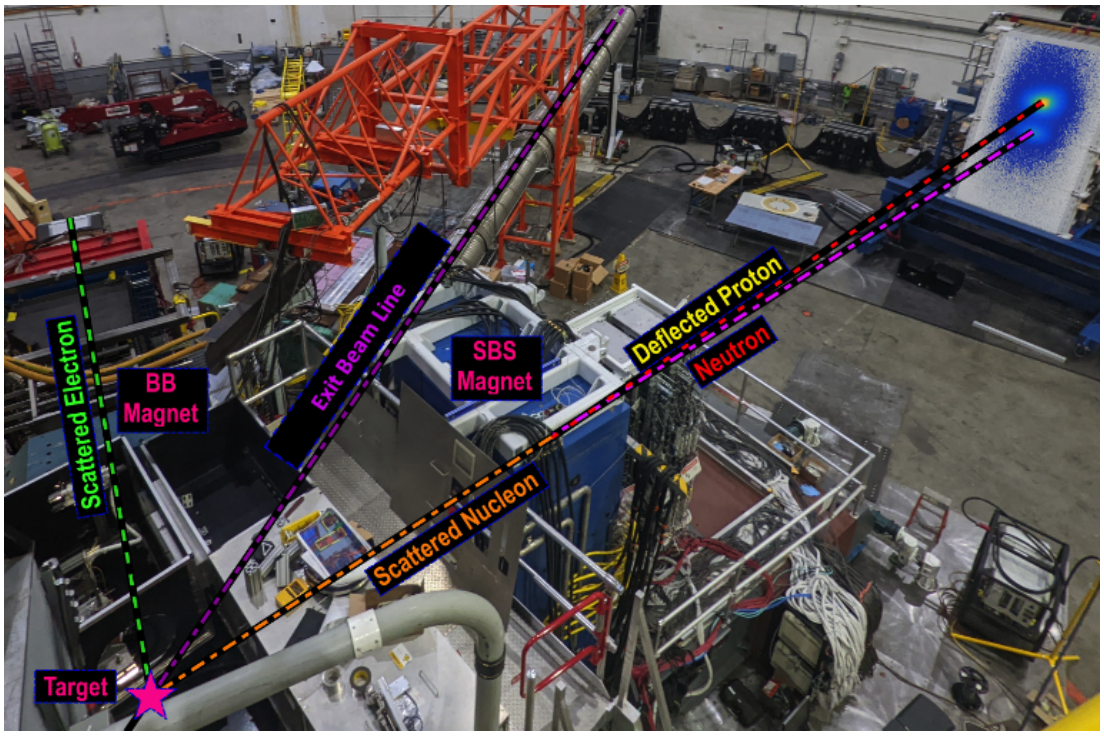


Figure 4-1: Downstream regions of the experimental setup. From the target, the electron scatters along the BB arm and the nucleons scatter along the SBS arm.

Figure 4-1 is a picture of the downstream region of the experimental setup for GMn and nTPE. There are three primary lines shown in the picture. They are, from left-to-right: the scattered electron arm, the exit beam line, and the scattered nucleon arm. The incoming electron beam would enter from the bottom left of the picture and impinge on the target (marked by the star in the bottom left corner). The scattered electrons then move down the Electron Arm where they pass through the BigBite magnet, the BB detector stack, and subsequently land on BBCal. The scattered nucleons traverse the Hadron Arm where they pass through the SBS magnet and then land on HCal.

The magnets along each arm play key roles in our experimental measurements. Both magnets deflect charged particles that pass through their yokes. On each of their respective scattering arms, the electrons and protons are deflected *upwards* towards the ceiling of Hall A by either the BigBite or SBS magnet, respectively. To deflect

both protons and electrons in the same upward direction, the magnets are set at opposite polarities. This exploits the fact that protons and electrons have opposite charges and effectively filters out hadrons on the Electron Arm and vice versa. This is a very crucial feature that needed to be accounted for.

The electron beam and target are both unpolarized. Therefore, we have azimuthal symmetry for the scattered leptons and hadrons. BBCal and HCal have no inherent way of distinguishing electrons from nucleons. They are sampling calorimeters which simply measure charged particles. Therefore, we need a way to screen hadrons from the Electron Arm and electrons from the Hadron Arm. The BigBite and SBS magnets provide that screening. The BigBite magnet directs electrons through the BB detector stack while deflecting protons down towards the floor and out of the acceptance of any detectors. Conversely, the SBS magnet directs protons up towards the ceiling and electrons down towards the floor such that they miss HCal.

4.3 HCal Observables: dx, dy, and dxdy

4.3.1 HCal Coordinates

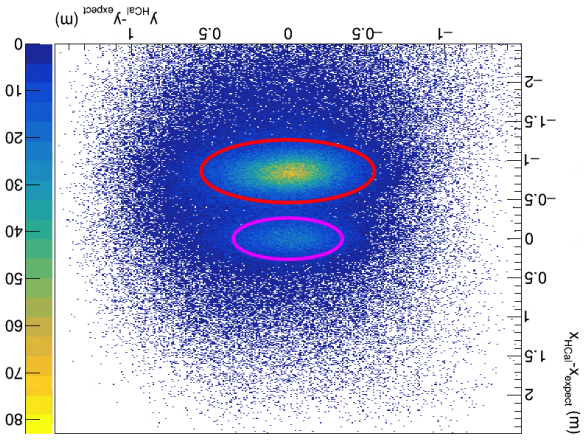


Figure 4-2: A basic HCal *dxdy* plot, or *hitmap*. The red ellipse indicates protons and the magenta ellipse indicates neutrons. The plot is upside down because the physical coordinates in the experimental hall are flipped from the coordinate system in the analysis database.

When nucleons land on HCal, their position, time of flight, and energy are mea-

sured. The position is measured with a resolution of ∼7 cm RMS, a time of flight accuracy of ∼700 ps, and an energy resolution of approximately 30% (see Section 2.6.2) [28]. Figure 4-2 shows a basic *dx* vs. *dy*, or simply *dxdy*, plot of particle hits across the surface of HCal[3]. The dxdy histogram (and plot) is an extremely important variable that will be discussed in greater detail in Section 4.3.2. For now, I will assume a basic interpretation of the dxdy plot as *a plot that shows the hit positions of neutrons and protons measured on HCal.*

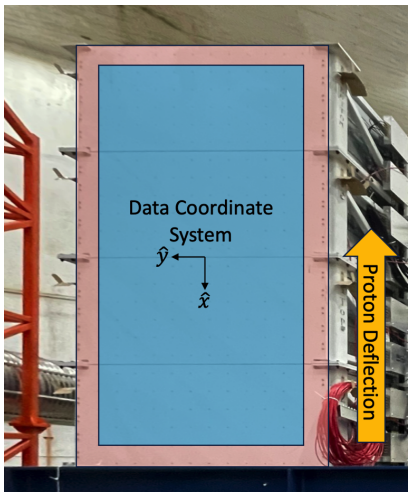


Figure 4-3: Photo of the SBS HCal in Experimental Hall A. Super-imposed onto the face of the detector are indicators of the physical area (pink), the acceptance (blue), and the coordinate system used in the analysis database. Note that this coordinate orientation is upside down and flipped from the physical setup.

Two distinct regions can be distinguished in Fig. 4-2. These are the *spots* for protons and neutrons. The red ellipse highlights the region for protons and the magenta ellipse indicates the neutrons. The plot is shown upside down and horizontally flipped to indicate the differences between the physical coordinate system in the experimental hall where this detector sits, and the coordinate system in the analysis database. The data stream is oriented using the *Beam Coordinate System* (BCS). There are many intricacies involved with understanding the BCS, but the most important detail to remember is that it is typically upside down and flipped from our standard, physical real-world orientation. Therefore, most of the two-dimensional hitmaps shown

[3]These are actually clusters of hits where each cluster corresponds to a single particle and its shower.

throughout this paper are typically upside down and flipped. The differences between the physical and data coordinate systems for HCal can be seen in Figure 4-3. The z-axes are along the scattering axes of each respective arm: incoming beamline, BB arm, and SBS arm.

In Figure 4-2, the neutrons (magenta ellipse) are centered around (0, 0). This indicates that they are along the straight-line projection from the target vertex. Note that in this figure, the y-axis is the horizontal axis, and that the x-axis is the vertical axis. The protons (red ellipse) are shifted from the origin towards the negative x direction. This is as expected. The protons are deflected up, towards the ceiling, from the Experimental Hall A perspective, and down in the data coordinate orientation. From here on, unless otherwise noted, the HCal dxdy plots will not be rotated and flipped as you see in Figure 4-2. They will be in their standard BCS frame.

4.3.2 The dxdy Plot

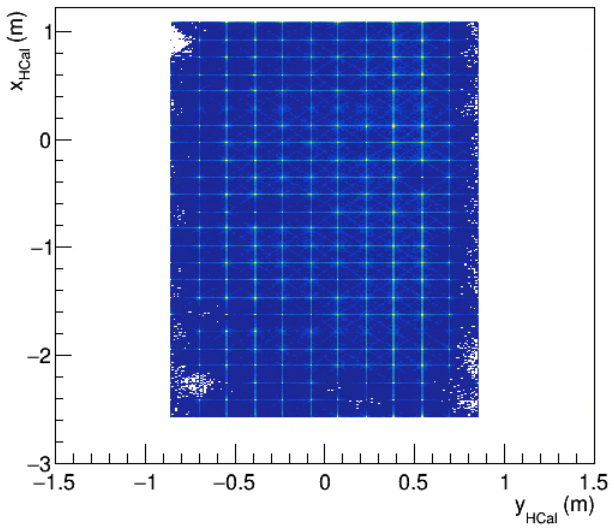


Figure 4-4: Measured hits on HCal. This shows all events in a given set of runs. The lines on the plot designate the centroids of each HCal block, not their edges.

On an event-by-event basis, the scattered protons and neutrons are spread across the entire face of HCal. The range and deviations in particle momenta and trajectories allow for this. Figure 4-4 shows the aggregate of hits on HCal for a given set of runs. The highly concentrated regions indicate the centers, of each HCal block. These are

where most of the energy is detected in each block [17]. Though this plot helps us to understand the general operation and effectiveness of measurements on HCal, it is hard to deduce anything about individual nucleons from here. So, in order to do that we use a dxdy plot (short for dx vs. dy).

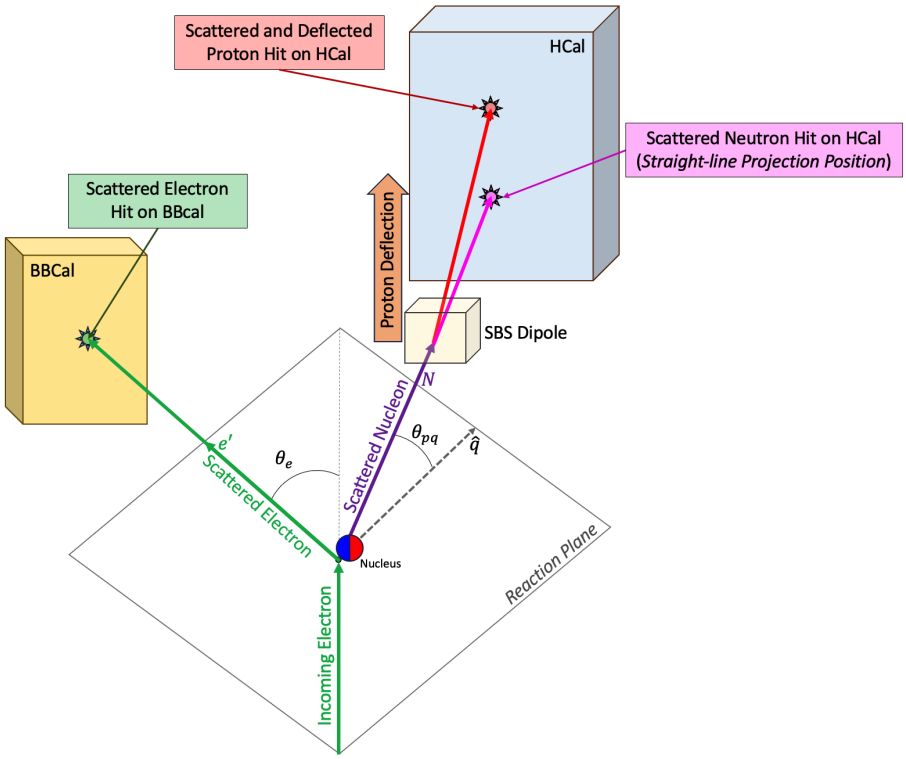


Figure 4-5: Basic e-N scattering kinematics and projections of expected hadron positions on HCal. A straight-line projection can be calculated purely from BigBite measurements to determine an expected hit position for the hadron on HCal.

Figure 4-5 shows the reaction of an incident electron (green) on a target nucleus (red and blue circle). The electron scatters and lands on BBCal while the nucleons scatter to HCal. Applying kinematic relations and various measurements strictly from the Electron Arm, we can calculate the expected position of a scattered nucleon on HCal. The goal is to find a straight-line projection for the nucleon, from the target to the surface of HCal. This calculation does not account for any deflection of the proton. Therefore, when we measure a nucleon at a different position than what we expect, it is assumed to be, or *grouped* as, a proton. This deviation from the expected

point is precisely the "dx" or "dy" we are referring to in the plot's name:

dx = (x position of Measured Hit on HCal) − (x position of Expected hit on HCal)

dy = (y position of Measured Hit on HCal) − (y position of Expected hit on HCal)

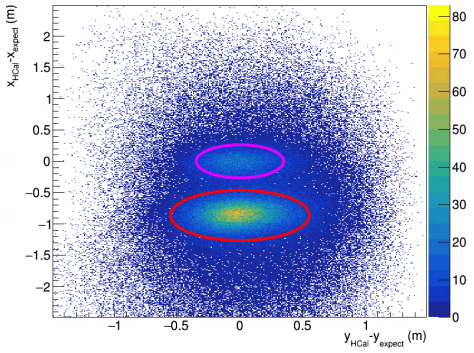


Figure 4-6: A typical dx vs. dy, or *dxdy*, HCal plot. This plot shows the displacements between expected and measured hits on HCal. The expected positions are the straight-line projections from HCal based on BBCal momentum measurements. The proton spot appears off-center because they are deflected by the SBS magnet.

When these are plotted against each other, the clustered regions, or spots, for each nucleon become visible (see Fig 4-6).

The 2D dxdy plot is very useful as is, but we can also utilize each of its variables independently as well. We can do this directly by accessing the dx and dy variables in their respective histograms, or we can extract them from y or x axes projections of the 2D dxdy histogram. The wording can get tricky here. If we project the 2D dxdy plot onto the y-axis, we are getting the dx variable. Similarly, a projection onto the x-axis gives us dy. Remember, the coordinates are upside down and flipped, so the x-dimension runs on the vertical (ordinate) axis of the plot, and the y-dimension is the horizontal (abscissas) axis. This can be tricky at first, but with some practice you will get used to the idea that down on the plot is the top of HCal, and that projecting onto the y-axis gives us dx. The general process for arriving at a dx plot from the projection of a dxdy plot is shown in Figure 4-7.

The right-most plot in Figure 4-7 shows a well-resolved dx plot. In this plot we can clearly see a separation of two peaks. The peak centered at $x = 0$ is the neutron peak,

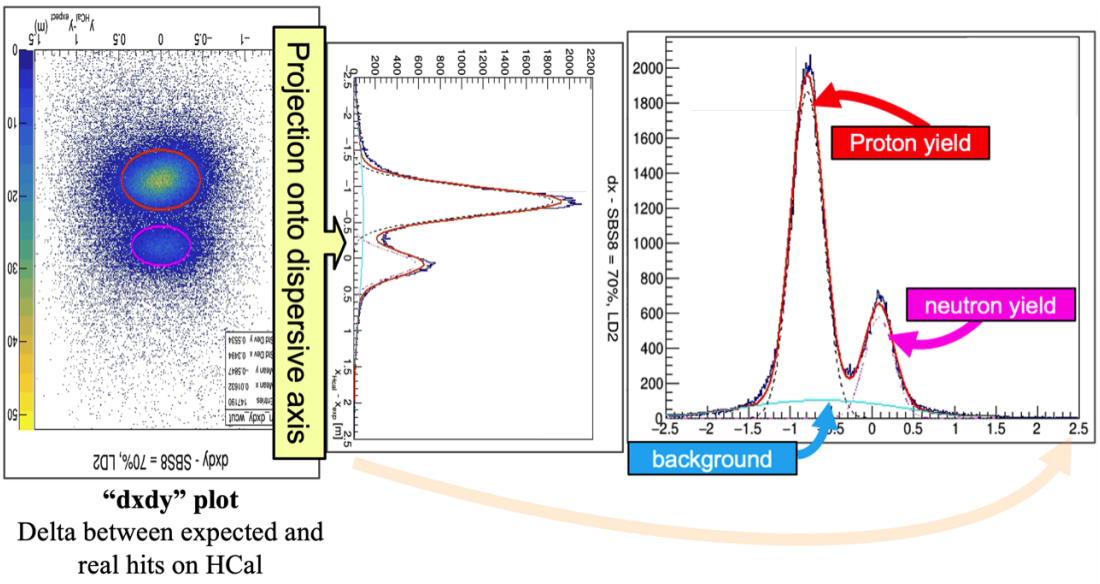
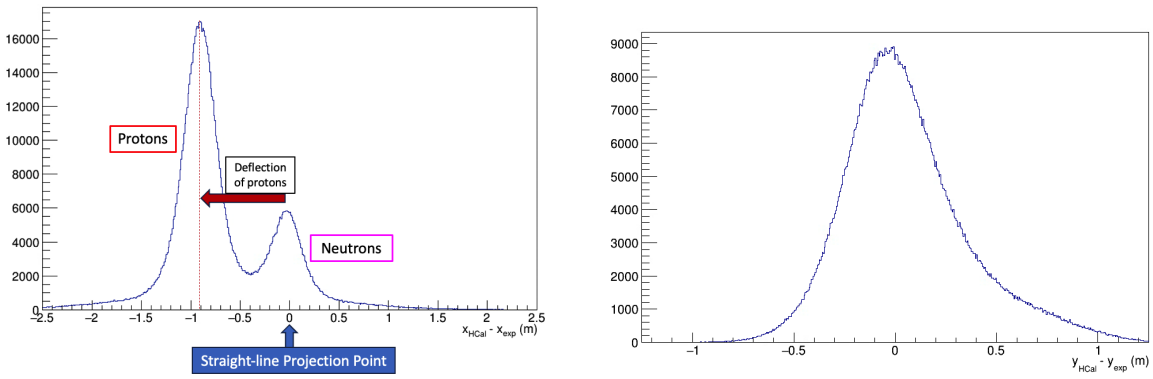


Figure 4-7: Projecting a 2D dxdy plot onto the y-axis and retrieving a dx plot. The various signal regions are shown.

and the other peak, centered at approximately $x = -0.8$, is the proton peak. The distance between the centers of these peaks is the average amount of deflection put on the proton as it passed through the SBS dipole magnet. Therefore, at various SBS field scales, the separation (proton deflection) distance will also vary. An increased SBS magnet field will cause greater deflection, while a smaller field causes less. This particular plot had an SBS field scale setting of 70% of maximum field.

Figure 4-8 shows dx and dy plots. The dx plot in Figure 4-8a highlights the regions corresponding to protons and neutrons. It also shows the zero-point corresponding to the straight-line projection coordinates on HCal. In Fig. 4-8b, we see the dy plot. This plot is useful as it provides us with criteria for making multiple *cuts* on the data during analysis. A *cut* here refers to a means by which we can filter out results which do not meet a certain criteria or threshold. For instance, a *dy cut* will be applied in the quasi-elastic selection process. In this cut, all data that exists beyond a certain upper and lower limit of dy will be rejected. The dy plot is also useful, because it is a good representation of the momentum spread for the events at hand. The experimental x-axis is the axis on which we observe the dispersion between protons and neutrons due to the SBS magnet. The experimental y-axis however, is considered



(a) dx: from Y-projection of dxdy **(b)** dy: from X-projection of dxdy

Figure 4-8: X and Y axes projections of an HCal dxdy plot. **(a)** A dx plot with labels for the straight-line projection point, the neutron peak, the direction of proton deflection, and the proton peak. **(b)** A dy plot. This demonstrates the momentum spread of the reactions. X is the dispersive direction so, any spread in Y comes from the momentum distribution.

the *non-dispersive* axis, since it has no inherent source of dispersion — there is no magnet-induced deflection in this direction. The distribution that we see in the dy plot arises from the momentum distribution of the scattering particles. The dx and dy plots also show the effects of *Fermi smearing* and *radiative tails*.

4.4 GMn and nTPE Simulations

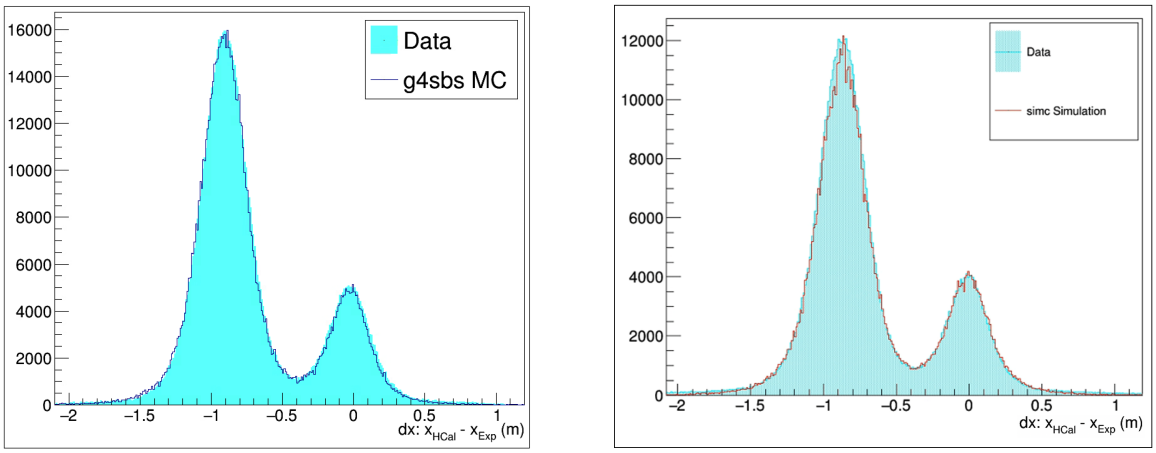
This section will introduce the basic underlying ideas behind the simulations used in this analysis. This includes an introduction to the simulation framework, why it is used, and how it is used. Some basic functionality and calculations are introduced and used to explain the motivation in applying simulations to extract form factors. However, a more detailed presentation can be found in Chapter 5.

It is typical in these types of experiments to extract a cross-section by simulating the experiment in a *Monte Carlo* (*MC*), calibrating or tuning that simulation to experimental data, and then performing the final extraction using the simulated yields [11]. The selection cuts applied to the simulation will be the same as those applied to the data. Some simulation parameters (e.g. the SBS magnetic field which determines the proton peak position in dx) need to be tuned with experimental data

parameters. Once tuned, we can determine our experimental cross-section by taking the ratio between the yields from data and the yields from simulation multiplied by the simulation's built-in cross-section [11].

The analysis described in this thesis employs a Geant4-based simulation package. Geant4 is physics software specifically designed for simulating particles passing through matter [36]. It is a very well-supported analysis package with decades of implementation. Geant4 allows you to re-create the entire experimental setup by building and implementing each detector (or sub-detector) package. Therefore, the physical dimensions, positions, and material properties all need to be programmed into the Geant4 project for the particular experiment. The Geant4 project for the entire SBS Program is called *g4sbs*. A second software package called *SIMC* is also used to simulate events for g4sbs.

4.4.1 g4sbs and SIMC



(a) Comparison of *pure g4sbs* and data. **(b)** Comparison of SIMC and data.

Figure 4-9: Comparisons of data and simulations. **(a)** Comparison of *pure g4sbs* simulated and experimental dx plots. The primary areas for disagreement are the tails of each peak, the peak widths, and the trough region between peaks. **(b)** Comparison of SIMC-simulated and experimental dx plots. The MC-simulated data matches much better because it incorporates additional radiative corrections.

g4sbs

g4sbs is a Monte Carlo simulation framework developed for the SBS Program. It is a self-contained Geant4 application which couples with the ROOT Analysis Framework to allow for streamlined additions of SBS detectors and hardware using already existing ROOT structures of the SBS analysis framework (Analyzer, PODD, SBS Offline, SBS Replay, etc.). The g4sbs package has access to all standard Geant4 libraries, in addition to specialized libraries for the SBS and related systems. g4sbs handles almost everything needed to properly simulate quasi-elastic electron-nucleon scattering. A comparison of g4sbs simulation and experimental data is shown in Figure 4-9a. The g4sbs machinery generates a very nice simulated dx plot that lines up well with the experimental data. However, there are some consistently dissimilar regions between the two datasets. In particular, the tails of each simulated g4sbs peak are not as broad as those from experimental data. We can also see that the simulated proton and neutron peaks are not quite as wide as the peaks in the experimental data. Though, not as prominent in Figure 4-9a, the so-called *trough* region between the peaks is not usually in such nice agreement. g4sbs has the ability to incorporate extended corrections onto Geant4 in order to account for neutron and proton detection efficiencies at HCal and first order nuclear corrections, but it does not provide any calculation support for various radiative corrections. To handle these additional contributions and corrections, an additional software kit called *SIMC* was added in parallel to g4sbs.

SIMC

In order to extend the capabilities of g4sbs, an MC called *SIMC* was implemented. SIMC was originally developed for Experimental Hall C, and has been adapted for use in Hall A and the SBS Program. Though SIMC is a Monte Carlo capable of fully simulating coincidence reactions and spectrometer optics/apertures, it is not a completely standalone add-on to Geant4 like g4sbs. Therefore, it was possible to incorporate SIMC on top of our g4sbs framework and make use of its added capabili-

ties for radiative effects, quasi-elastic scattering for H(e,e′p), Final State interactions, Coulomb Corrections, etc. These missing functions are what prevented us from fully tuning the g4sbs simulation to our experimental data. Figure 4-9b shows an overlay of a SIMC-generated dx plot on a plot from experimental data. In comparison to Fig. 4-9a, SIMC fits the tails, widths, and trough regions of the peaks much better.

As previously mentioned, the aspects of SIMC which provide improvement over g4sbs are the inclusion of radiative effects. A short discussion on these concepts is important to understanding their role in the simulation and subsequent analysis.

Electrons in the presence of a nuclear field will radiate due to changes in their velocity brought on by Coulomb interactions. As an electron accelerates or decelerates, it radiates through the process of bremsstrahlung. This affects both incoming and outgoing electrons due to either its proximity with the coulomb field of the nucleus involved in the scattering process, or the field of a nearby nucleus. During these interactions, the electrons can emit and reabsorb virtual photons, or emit real, soft, photons. When the interaction occurs with the nucleus involved in the scattering, the process is called *internal bremsstrahlung*. When it involves nearby nuclei not directly involved in the scattering, it is called *external bremsstrahlung*. When either of these occur, they have the effect of modifying the interaction cross-section and electron kinematic properties (momentum, energy, angle, etc.) [51]. These radiative effects are expected in experiments such as GMn or nTPE due to the primary interaction being that of an electron beam incident on a deuterium target. The dx plots extracted from data exhibit broader tails on the neutron and proton peaks when compared to g4sbs simulations, which do not incorporate these radiative effects. However, when we compare SIMC outputs with the same experimental data, we see that accounting for these effects produces a nicer comparison.

In addition to introducing radiative corrections to g4sbs, SIMC also accounts for the effects of *Fermi motion*, or *Fermi smearing*. In typical quasi-elastic electron-nucleon scattering formalism, the target nucleon is considered to be at rest. In reality, we know that the proton and neutron are not at rest. Nucleons within the nucleus move collectively with an average momentum. Their motion inside the nucleus can be

described by the Fermi statistic and their independent energy levels near the Fermi surface can be described by the Fermi momentum [84]. This Fermi momentum is closely related to the average motion just mentioned. If we now consider the nucleon as a non-stationary target, the incoming electron now interacts with an individual nucleon moving in the Fermi sea [58]. This has the effect of broadening the quasi-elastic peak and skewing its position [84, 58, 59]. This broadening, or lack thereof, is a source of disagreement between simulated and experimental data. If we look at SIMC data (Fig. 4-9b), we see that the simulated peaks, as a whole, are in better agreement. There are some incongruities between the tips of the peaks, but this is only due to limited statistics available for that particular plot. As the analysis matures and more data, i.e. statistics, are brought in, the tips of the SIMC peaks will smooth out.

From here on, unless specifically noted otherwise, any reference to g4sbs will also be a reference to Geant4 and all-inclusive packages there-in. Similarly, unless otherwise noted, the terms Monte Carlo (or MC), g4sbs, and SIMC, will be used synonymously for simulated data.

Geant4 allows for the complete re-creation of our experimental setup in its software environment. Within the simulation package, we are able to build and implement each detector package just as it sits in the experimental hall. This incorporates everything from basic detector dimensions and positions, down to how the materials themselves interact with charged particles. A three-dimensional wireframe representation of the GMn and nTPE experiments as they are built into g4sbs is shown in Figure 4-10. This figure can be compared to other figures which show the real experimental setup in the hall: 2-19, 4-1, and 4-5. The g4sbs configuration was modelled to accurately match the various experimental setups.

The g4sbs framework allows us to simulate a given electron beam incident upon a choice of targets. It then determines the subsequent downstream scattering of the electrons and nucleons from simulated beam-target interactions. The simulation contains the BigBite and SuperBigBite magnets as well, so any deflection brought on by their fields would be induced upon any scattered charged particles which pass through their apertures. Keep in mind that we are not defining the kinematics we

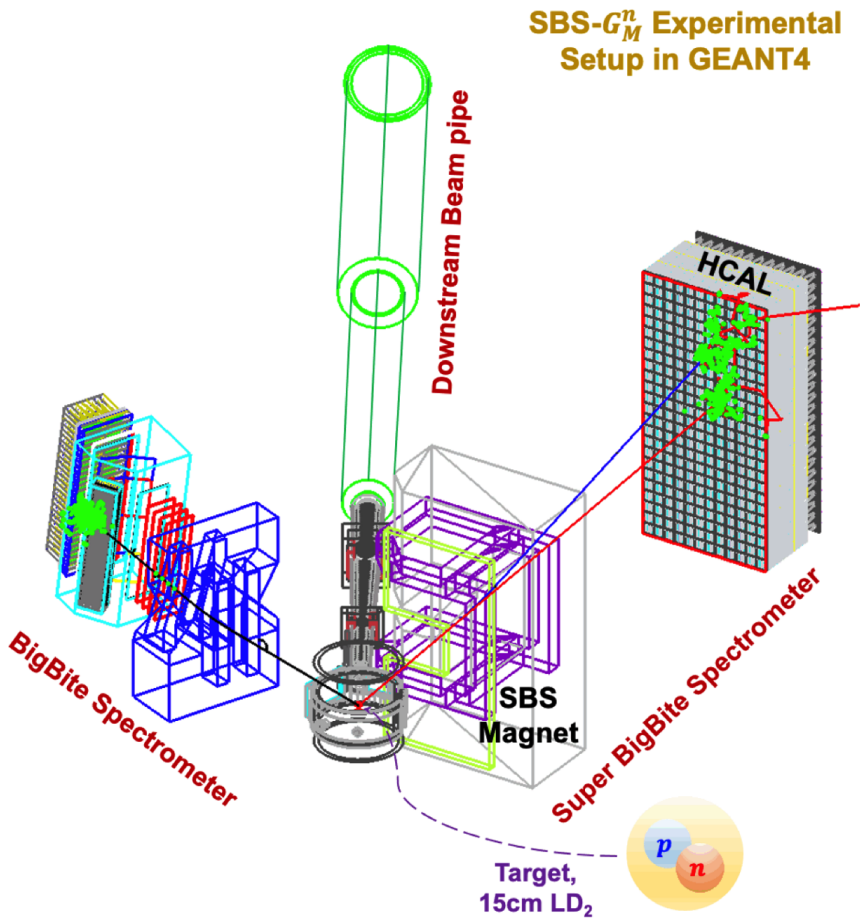


Figure 4-10: The GMn and nTPE experimental layout in the g4sbs software platform. Image source: [46]

expect to see, but rather providing the simulation with upstream experimental parameters (beam energy, target properties, magnetic field settings, etc.) and looking at what it produces from those specifications.

4.4.2 Simulation dx histograms and plots

The data we collect from the generated simulations has a data structure that is nearly identical to that of our experimental data. Figure 4-11 shows a dx plot generated by the g4sbs simulation package. The simulation stores events for protons and neutrons separately. Therefore, after running the simulation, our output file will contain a histogram containing the dx values for the proton (MC_dx_p) and a separate histogram containing dx values for the neutron (MC_dx_n). The total

dx histogram containing both protons and neutrons is obtained by adding these two histograms together:

$$\text{MC\_dx histogram} = \text{MC\_dx\_p histogram} + \text{MC\_dx\_n histogram}$$

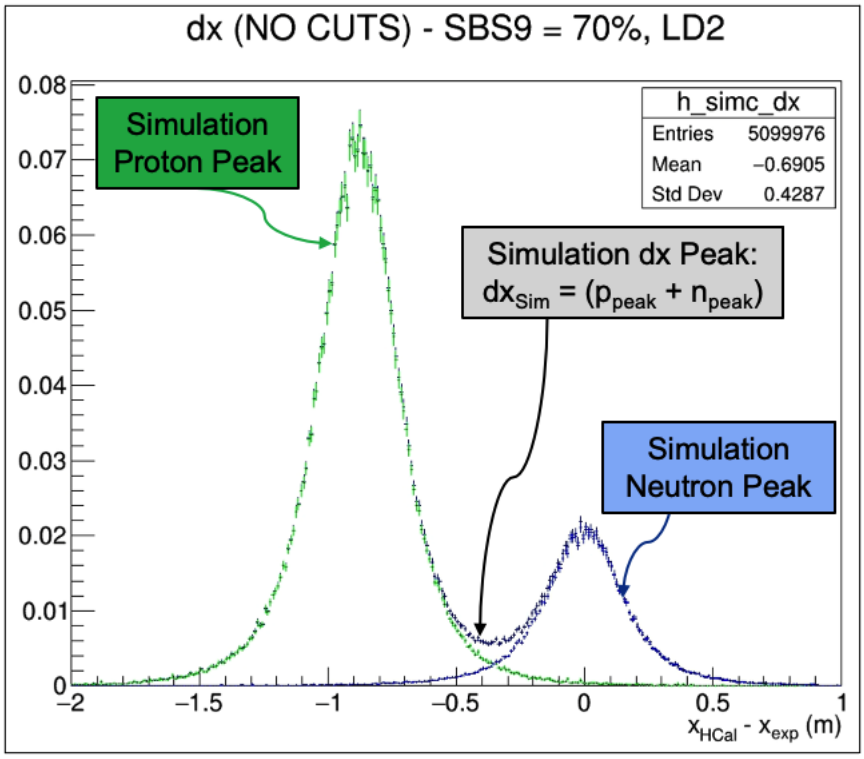


Figure 4-11: A dx plot for simulated data. The g4sbs simulation framework generates proton and neutron dx histograms separately. They are added to a single histogram which subsequently becomes the total dx plot.

4.4.3 Simulation Event Weighting

The events generated in g4sbs all spawn from within the target volume. Though the simulation package requires information about the incoming beam, it does not re-create or simulate the beam upstream of the target and then simulate its interactions with a simulated target. Instead, it uses a *generator* to initialize the events. The g4sbs and SIMC simulation frameworks contain various generators which produce reaction particles at the vertex of the beam and target. These simulated particles travel outward, and downstream, from the beam-target-coincidence point as if an

incoming electron beam of known parameters had been incident upon the target. The g4sbs package contains multiple types of generators which consider the probability of a given reaction and produce its downstream events. For a given type of event, a certain generator must be selected. In this analysis, the following generators were implemented [101]:

- Elastic generator: elastic electron-nucleon scattering. Electron polar and azimuthal angles are generated *flat* within user-defined limits. All other variables are calculated from energy-momentum conservation. Cross-sections follow from internal application of the Kelly fit for GMn, GMp, and GMn. GEn cross-section follows from a fit to a previous Hall A GEn experiment.

- Inelastic generator: Inclusive inelastic electron-proton or electron-deuteron scattering. Electron polar and azimuthal angles are generated flat within user-defined limits. Electron energies are generated flat within a user-defined range. Cross-sections follow from Christy-Bosted parameterizations of inclusive p(e, e′) and d(e, e′).

- Particle gun generator: Scattered particles only (called a particle gun). Generates the outgoing particle angles and energies given user-defined inputs and limits.

When creating a set of simulation data, you specify a multitude of values ranging from the number of events to generate, beam energy, target material/element, target length, detector positions, magnet positions, etc. The software then sets out to generate the requested number of events. The simulation only generates events on the physical part of the user-defined phase-space. Because of this, the simulation may need to iterate more times than the number of requested events. The comparison between the *number of tries* and *total number of events generated* is called the *generation efficiency* or *reconstruction efficiency*. This information will be used in order to help properly *weight* the histograms produced by the simulation. g4sbs and SIMC both incorporate form factor parameterizations within their machinery for calculating nucleon cross-sections. However, the histogram data in the output files comes out

flat. This means that the data comes out only accounting for phase space and is not cross-section weighted. Therefore, we must add in a weight function for this.

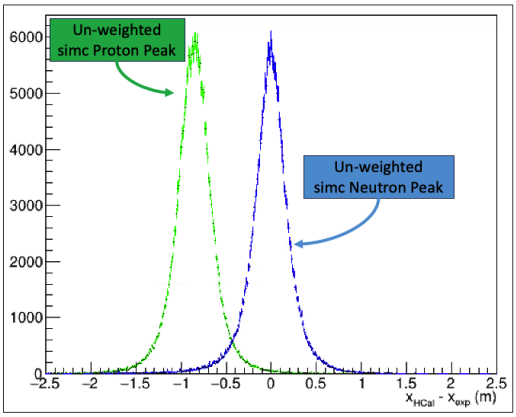


Figure 4-12: Unweighted MC dx plot. Without any weights, the proton and neutron dx histograms are roughly the same size. Adding an appropriate weight function will bring the ratio of histograms close to the expected ratio for neutron-to-proton cross-sections.

An un-weighted dx histogram from MC is shown in Figure 4-12. In the figure, both histograms are roughly the same size. This indicates that the histograms have been filled without a weighting function. Once we apply a proper weight function, Figure 4-12 will look more like Figure 4-11 in terms of the relative amplitudes of each peak. In this analysis, the following weight function is used:

$$\text{Final Weight} = \frac{(\text{MC Weight Value}) \cdot (\text{Luminosity}) \cdot (\text{Generation Volume})}{(\text{Number of attempts to generate events})} \tag{4.1}$$

The parameters in Eq. 4.1 are provided as outputs when generating the simulations. Each simulation file or event has a unique set of values corresponding to it. The *MC Weight value* is an event-based value:

$$\text{MC Weight Value} = (\text{cross-section}) \cdot (\text{Spectral Function}) \quad \left(\frac{\mu\text{b}}{\text{MeV sr}^2}\right)$$

The *Luminosity*, phase space *Generation Volume*, and *Number of attempts to generate events* are all values printed out to a text-based file per simulation run. When

analyzing MC data, these values are pulled from the output files and used to calculate the *Final Weight*. The Final Weight is then used directly as the weight value when *filling*[4] any cross-section-dependent histograms during the analysis and extraction processes.

The nucleon data produced by the Monte Carlo simulation allows for direct extraction of the nucleon yields by simply taking the integral of each corresponding histogram. When the Monte Carlo generates an event, it assigns a truth value for the scattered nucleon. For instance, if the event generated a scattering proton then the simulation output variable *fnucl* is set to 1. Conversely, for neutrons, fnucl would return a value of 0. Therefore, we know exactly which dx and dy events correspond to a neutron or proton. In order to get the yields for either nucleon we simply take the integral of the corresponding histogram.

4.4.4 Scaling MC to Data

The MC and data comparison plots presented up until this point have all incorporated some underlying amplitude normalization in order to get the simulated histograms on the same scale as the experimental data. This normalization was applied to make the plots easier to introduce and interpret. As a means to simply check the MC shape against that of the experimental data, we can use a basic scaling between the maximum values in each histogram[5]. This is fine for preliminary checks, but is not robust enough for a final, or even exploratory, physics extraction. For a proper analysis, we will need to precisely tune the amplitudes of the simulated nucleon peaks so that they match *background-subtracted* experimental data[6], and then assess their agreement using a statistical hypothesis (chi-squared, or χ^2) test. We start by in-

[4]When *filling* a histogram in ROOT c++, you use a command called Fill(). The Fill() commands takes as its input, the value to fill a histogram with, and its associated event-based weight value.

[5]The maximum value in any of our dx plots will be the maximum value of the proton peak. A basic way to scale MC to match data is to apply a linear scaling to the entire MC dx histogram of $^{(\text{Maximum Value of Data})}/_{(\text{Maximum Value of MC})}$. This will match the proton peak heights fairly well. If the proper weighting was applied while filling the MC dx_n histogram, the height of the neutron peaks should also be pretty close in amplitudes, as well.

[6]It is important to note that my analysis uses *background-subtracted* experimental data. All of my scale factor tuning and χ^2 minimization utilizes experimental data that has had some form of a background subtracted from it. This is discussed further in Section 5.3.8.

troducing *scale factors* which linearly scale each simulated nucleon's histogram. The composite simulation dx peak is then formed by combining the simulated proton and neutron histograms. The quality of the fit between the data and composite simulation dx plots after scaling is then assessed via a χ^2 test between the two histograms.

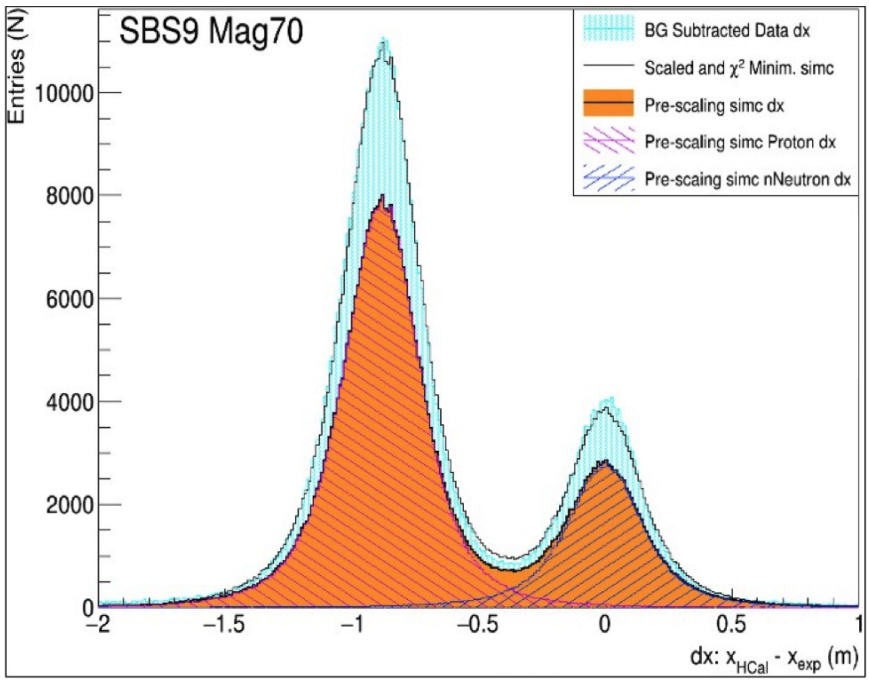


Figure 4-13: Comparisons between pre-scaled and post-scale MC dx peaks with experimental data. The cyan peak is experimental data. Everything else is from simulation. The black outline around the cyan peak is the *tuned* simulation dx peak (post-scaling). The orange peaks correspond to the raw, un-scaled simulation data (pre-scaling).

Figure 4-13 is a dx plot showing various stages of the simulated data compared against experimental data. In Fig. 4-13, the cyan peak is our experimental data. Everything else comes from the simulation. The black line along the cyan edge is the tuned simulation dx peak after all is said and done. Essentially, it is the final output from precisely tuning MC data to match the cyan region. The Monte Carlo simulation creates histograms for scattered proton and neutron events separately. The black line corresponds to a histogram that is made by adding the two individual proton and neutron histograms together. The orange region of the plot shows the raw output from the Monte Carlo simulation; this data has no scale factors applied to it. Within the orange region you can see two striped regions. The pink stripes on

the left correspond to the MC proton peak, and the blue stripes on the right are the MC neutron peak. In order to properly tune the black line, each of the individual peaks (pink and blue stripes) are independently scaled such that the χ^2 between the background-subtracted data and the sum of the two nucleon peaks is minimized. This minimization process effectively tunes the simulation to experimental data and allows me to determine the scale factor for the neutron peak, $f_{(\text{scale, n})}$, and the scale factor for the proton peak, $f_{(\text{scale, p})}$, which are used to minimize the χ^2 test. We can define the scale factors as follows:

$$f_{(\text{scale, n})}: \quad \chi^2\text{-minimizing scale factor the neutron}$$
$$f_{(\text{scale, p})}: \quad \chi^2\text{-minimizing scale factor the proton}$$

The scale factors essentially tell us how much our simulation's built-in parameterization model needed to be corrected in order to match the real data. If the cross-sections and form factor parameterizations in our Monte Carlo machinery were exact, then the ratio of these scale factors would be unity[7]. This will be discussed further in the following sections. As it turns out, these scale factors are our *actual* extraction parameters. They are the primary variables that we need from this analysis in order to extract our G_M^n values.

4.5 Extracting GMn Experimentally

Extracting the value of G_M^n experimentally differs slightly from the theoretical foundation described in Chapter 1. Theoretically, we seek to extract the ratio of quasi-elastic yields via the differential cross-sections of protons and neutrons. However, we can't directly access the quasi-elastic yields from our experimental data. Figure 4-14 shows an LD2 dx plot with the three primary peaks labelled. This is a relatively *clean* plot without much background contamination. Also, the total fit (red line) matches

[7]The ratio of scale factors is an integral component of the final extraction as is defined as: $F_{n/p} = f_{(\text{scale, n})}/f_{(\text{scale, p})}$.

fairly well. We could simply take the integrals of each fit and from there, determine the ratio of yields. This would be okay as a rough estimate or basic approximation, but in general is not very accurate, nor does it rely on any model-dependent physics for the fit functions.

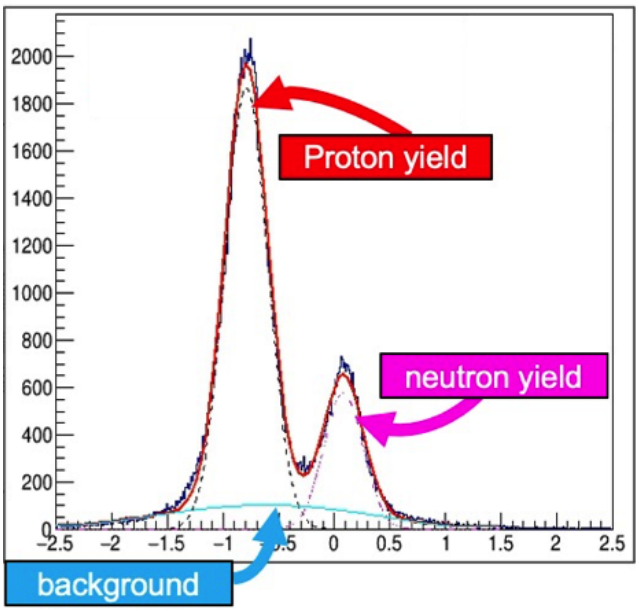


Figure 4-14: LD2 dx plot with labelled nucleon peaks and background approximation. We can approximate the shapes of each nucleon peak and the background in experimental data, but untangling them to get an accurate value for the quasi-elastic yields is un-realistic.

The plot of Fig. 4-14 shows various fits: a polynomial background fit (cyan), a proton Gaussian (blue dash), a neutron Gaussian (magenta dash) and the overall fit from all three (red). The implementation of Gaussian fits and a polynomial background are not driven by any physics models. They are functional forms which produce relatively well-matched fits that can be easily applied *ad hoc*. This is precisely why we implement our Monte Carlo simulation — the simulated MC peaks will replace the nucleon fits shown in Figure 4-14. The data from our simulations is generated directly from physics-driven models which will be more accurate, and more appropriately match the data when tuned via the scale factors previously mentioned. Therefore, I now need to slightly re-formulate the Ratio Method extraction technique (see Sec. 1.5) so that it accounts for how the quasi-elastic yields are actually ex-

tracted in this analysis: the quasi-elastic yields are determined from the simulation and incorporate the ratio of scale factors used to tune simulation to data.

There is a vast amount of world data that exists for the form factors of the proton. In general, the charged proton is much easier to measure than the neutral neutron. Also, there are *free proton targets* but no *free neutron targets*. This is precisely we why utilize the deuteron as our scattering target. It provides a relatively simple neutron target; deuterium consists of a proton and a neutron. To accent the fact there is not much world data for the neutron, recall that the nTPE experiment is the first measurement of the Rosenbluth Technique on the neutron.

The MC simulation framework that we implement uses a built-in set of world data form factor parameterizations when determining the nucleon cross-sections per generated event. Given that proton data is more established and therefore better-supported than the neutron, we have a bit more confidence in our proton model and can take it as *canon*. Therefore, if the proton model is essentially *correct*, we can then attribute any deviations between the simulation and data cross-sections as coming from the Monte Carlo's built-in parameterizations for the neutron. From this point forward, I will be considering the ratio of scale factors, $f_{(\text{scale, n})}/f_{(\text{scale, p})}$, as a type of correction to the MC's built-in neutron cross-sections. We can now re-formulate our derivations for extracting G_M^n from our experimental data and simulations, in terms of the neutron and proton scale factors.

Extraction of G_M^n begins with the ratio of neutron and proton yields (Eq. 1.16):

$$R_{n/p} = \frac{N_{e,e'n}}{N_{e,e'p}} = \frac{\frac{d\sigma}{d\Omega}|_{n(e,e')}}{\frac{d\sigma}{d\Omega}|_{p(e,e')}} \equiv \frac{\frac{\sigma_{Mott}}{1+\tau_n}\left((G_E^n)^2 + \frac{\tau_n}{\epsilon}(G_M^n)^2\right)}{\frac{\sigma_{Mott}}{1+\tau_p}\left((G_E^p)^2 + \frac{\tau_p}{\epsilon}(G_M^p)^2\right)} \approx \frac{\sigma_{R,n}}{\sigma_{R,p}} \tag{4.2}$$

where

$$\tau_N = \frac{Q^2}{4M_N^2}, \quad \epsilon = [1 + 2(1+\tau)tan(\theta/2)]^{-1}, \quad \text{and} \quad M_n \approx M_p \approx M_N. \tag{4.3}$$

Here, $\sigma_{R,n}$ and $\sigma_{R,p}$ are the reduced cross-sections of the neutron and proton, respectively. I have also assumed that $\sigma_{Mott,n} \approx \sigma_{Mott,p}$. In general,

$$\sigma_R = G_E^2 + \frac{\tau}{\epsilon} G_M^2. \tag{4.4}$$

In summary, the ratio of quasi-elastic yields can be related as:

$$\boxed{R_{n/p} = \frac{N_{e,e'n}}{N_{e,e'p}} \approx \frac{\sigma_{R,n}}{\sigma_{R,p}}} \tag{4.5}$$

The yields here are defined as quasi-elastic for the following reasons:

- SIMC simulation machinery internally accounts for radiative corrections, Fermi motion, and HCal detection efficiency.
- Quasi-elastic event selection and cuts have been applied during tuning and calibration.

Now, we can extract the quasi-elastic neutron and proton yields directly by taking the integrals of the scaled and calibrated MC neutron and proton histograms. These histograms have already been scaled for χ^2-minimization and therefore carry along with them the applied scale factors. Then, the yields, $N_{n,\mathrm{MC}}$ and $N_{p,\mathrm{MC}}$, that are determined by taking the integral of the respective histograms, are related to the cross-sections by the following relations:

$$N_{n,MC} = f_{(\text{scale, n})}(\sigma_{R,n,MC}) \quad \text{and} \quad N_{p,MC} = f_{(\text{scale, p})}(\sigma_{R,p,MC}). \tag{4.6}$$

Here, $\sigma_{R,n,MC}$ and $\sigma_{R,p,MC}$ come from the parameterizations built into the Monte Carlo for the reduced cross-section of each nucleon, n or p. From this we can define the ratio of quasi-elastic yields taken from the scaled MC histograms as:

$$R_{n/p} = \frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}} \left(\frac{\sigma_{R,n,MC}}{\sigma_{R,p,MC}} \right). \tag{4.7}$$

From equations 4.5 and 4.7 we find:

$$\left(\frac{\sigma_{R,n}}{\sigma_{R,p}} \right) = \frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}} \left(\frac{\sigma_{R,n,MC}}{\sigma_{R,p,MC}} \right) \tag{4.8}$$

We then move all proton-related variables over to the left hand side:

$$\left(\frac{\sigma_{R,p,MC}}{\sigma_{R,p}}\right) = \frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}}\left(\frac{\sigma_{R,n,MC}}{\sigma_{R,n}}\right) \tag{4.9}$$

At this point, we re-address the fact that we consider our built-in proton parameterizations to be *true*. This assumption means that our *parameterized values for the proton* should match our *experimental values for the proton*, i.e.,

$$\sigma_{R,p,MC} \approx \sigma_{R,p}$$

With this, we can simplify Eq. 4.9 in the following way:

$$\begin{aligned} \left(\frac{\sigma_{R,p,MC}}{\sigma_{R,p}}\right) &= \frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}}\left(\frac{\sigma_{R,n,MC}}{\sigma_{R,n}}\right) \\ 1 &= \frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}}\left(\frac{\sigma_{R,n,MC}}{\sigma_{R,n}}\right) \end{aligned} \tag{4.10}$$

Solving for the reduced cross-section of the neutron we get:

$$\sigma_{R,n} = \left(\frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}}\right)\sigma_{R,n,MC}. \tag{4.11}$$

Recalling that,

$$\sigma_{R,n} = (G_E^n)^2 + \frac{\tau_n}{\epsilon_n}(G_M^n)^2 \tag{4.12}$$

we can form a relation between Eqs. 4.11 and 4.12:

$$\sigma_{R,n} = \quad (G_E^n)^2 + \frac{\tau_n}{\epsilon_n}(G_M^n)^2 \quad = \quad \left(\frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}}\right)\sigma_{n,MC}. \tag{4.13}$$

This reduced cross-section, σ_R is the only observable that we can extract in a truly model-independent fashion. Now that we have it, we can then solve for G_M^n to find

$$\boxed{(G_M^n)^2 = \left(\frac{\epsilon_n}{\tau_n}\right)\left[\left(\frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}}\right)(\sigma_{R,n,MC}) - (G_E^n)^2\right]} \tag{4.14}$$

or,

$$\boxed{G_M^n = \sqrt{\left(\frac{\epsilon_n}{\tau_n}\right)\left[\left(\frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}}\right)(\sigma_{R,n,MC}) - (G_E^n)^2\right]}} \tag{4.15}$$

where we have used the following relations,

$\tau_n = \dfrac{Q^2}{4M_n^2}$	Neutron Bjorken scaling factor
$\epsilon = [1 + 2(1+\tau)\tan(\theta/2)]^{-1}$	Virtual photon polarization (neutron)
$f_{(\text{scale, n})}$	χ^2-minimized scale factor for MC, neutron
$f_{(\text{scale, p})}$	χ^2-minimized scale factor for MC, proton
$\sigma_{R,n,MC}$	MC reduced C.S. parameterization, neutron
G_E^n	Neutron electric FF from parameterization

4.6 Extracting nTPE Experimentally

In order to extract a value for nTPE via Rosenbluth Separation, we must again re-formulate the extraction technique introduced in Chapter 1. We need to relate the theoretical derivation to what we can actually access and extract from experimental data and simulations. Namely, we need to introduce the scale factors, $f_{(\text{scale, n})}$ and $f_{(\text{scale, p})}$, into the extraction process, just as we did for G_M^n.

In order to make it easier to differentiate between the two kinematics in the following derivation, I will introduce a color-coding scheme of the following form:

$$\epsilon_1 \longleftrightarrow \text{SBS8}$$

$$\epsilon_2 \longleftrightarrow \text{SBS9}$$

For the Rosenbluth Technique, we are ultimately looking to extract the Form Factor Ratio, G_E/G_M, via the Rosenbluth Slope, RS $= S = \sigma_L/\sigma_T$. The Rosenbluth Separation technique starts off with the implementation of a relation called the *super ratio*. The

super ratio is the *ratio of yield ratios*, and will be denoted by §. We can define the super ratio between SBS8 (ϵ_1) and SBS9 (ϵ_2) as:

$$\S = \frac{(R_{n/p})_{\epsilon_1}}{(R_{n/p})_{\epsilon_2}} = \frac{\left(\frac{\sigma_{R,n}}{\sigma_{R,p}}\right)_{\epsilon_1}}{\left(\frac{\sigma_{R,n}}{\sigma_{R,p}}\right)_{\epsilon_2}} \tag{4.16}$$

or

$$\S = \frac{(\sigma_{R,n})_{\epsilon_1}}{(\sigma_{R,n})_{\epsilon_2}} \frac{(\sigma_{R,p})_{\epsilon_2}}{(\sigma_{R,p})_{\epsilon_1}} \tag{4.17}$$

Recall that the reduced cross-section is:

$$\sigma_R = \sigma_T + \epsilon\sigma_L.$$

We also have the Rosenbluth Slope defined as:

$$\text{RS} = S = \frac{\sigma_L}{\sigma_T} \tag{4.18}$$

This lets us write the reduced cross-section as:

$$\sigma_R = \quad \sigma_T(1 + \epsilon\frac{\sigma_L}{\sigma_T}) \quad = \quad \sigma_T(1 + \epsilon S) \tag{4.19}$$

It should also be noted that we can have a Rosenbluth Slope for either the neutron or proton, i.e.,

$$\text{Neutron: } S^n \qquad\qquad \text{Proton: } S^p$$

We therefore have the following relation:

$$\frac{(\sigma_{R,n})_{\epsilon_1}}{(\sigma_{R,n})_{\epsilon_2}} = \frac{\sigma_T^n\,[1 + \epsilon S^n]_{\epsilon_1}}{\sigma_T^n\,[1 + \epsilon S^n]_{\epsilon_2}}$$

Recall that $\sigma_T = \tau G_M^2$ and that G_M is only a function of Q^2. Therefore, as long as the Q^2 values between SBS8 (ϵ_1) and SBS9 (ϵ_2) are equivalent to one another, we

have that

$$\frac{(\sigma_{R,n})_{\epsilon_1}}{(\sigma_{R,n})_{\epsilon_2}} = \frac{\cancel{\sigma_T^n}^{1}\,[1+\epsilon S^n]_{\epsilon_1}}{\cancel{\sigma_T^n}\,[1+\epsilon S^n]_{\epsilon_2}} \tag{4.20}$$

This holds true for the proton form of this relation as well. Equation 4.20 is a very useful simplification that will be used repeatedly. It is important to keep it in mind during the following derivation. We use this relation immediately and expand the right hand side of Equation 4.17.

$$\S = \frac{(\sigma_{R,p})_{\epsilon_2}}{(\sigma_{R,p})_{\epsilon_1}} \frac{(\sigma_{R,n})_{\epsilon_1}}{(\sigma_{R,n})_{\epsilon_2}} = \left(\frac{[1+\epsilon S^p]_{\epsilon_2}}{[1+\epsilon S^p]_{\epsilon_1}}\right)\left(\frac{[1+\epsilon S^n]_{\epsilon_1}}{[1+\epsilon S^n]_{\epsilon_2}}\right) \tag{4.21}$$

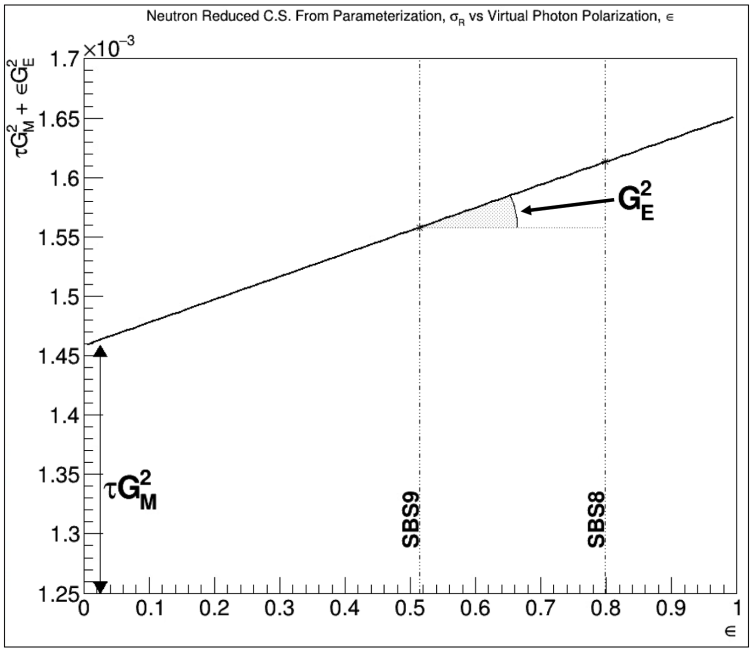


Figure 4-15: Neutron reduced cross-section from parameterization vs. virtual photon polarization, ϵ. The reduced cross-section of the neutron is calculated at the nominal epsilon points for SBS8 and SBS9 using the Ye parameterization ([140]). The linear-dependence in ϵ allows for a graphical interpretation that the slope of σ_R at ϵ_1 (SBS8) and ϵ_2 (SBS9) are the same.

Figure 4-15 is a plot of the reduced cross-section vs. epsilon for SBS8 (ϵ_1) and SBS9 (ϵ_2). The Rosenbluth Slope values of Eq. 4.21 each correspond to one of the

epsilon points (SBS8 $\leftrightarrow \epsilon_1$ and SBS9 $\leftrightarrow \epsilon_2$). We are forming a linear extrapolation between these two epsilon points and finding the slope. Therefore, it is fair to say that the slopes calculated at each epsilon point are equal to one another. They are along the same line after all. Therefore, we can set the Rosenbluth Slopes across the different kinematics equal to each other.

$$S^n_{\epsilon_1} \approx S^n_{\epsilon_2} \approx S^n \qquad \text{and} \qquad S^p_{\epsilon_1} \approx S^p_{\epsilon_2} \approx S^p$$

Now that we have equated the Rosenbluth Slopes between ϵ_1 and ϵ_2, Equation 4.21 can be re-written as:

$$\S = \left(\frac{1+\epsilon_2 S^p}{1+\epsilon_1 S^p}\right)\left(\frac{1+\epsilon_1 S^n}{1+\epsilon_2 S^n}\right) \tag{4.22}$$

We can now put this back into Equation 4.16 to get,

$$\frac{(R_{n/p})_{\epsilon_1}}{(R_{n/p})_{\epsilon_2}} = \S = \left(\frac{1+\epsilon_2 S^p}{1+\epsilon_1 S^p}\right)\left(\frac{1+\epsilon_1 S^n}{1+\epsilon_2 S^n}\right) \tag{4.23}$$

The left hand side of Eq. 4.23 can be expressed in terms of scale factor ratios and yield ratios to become:

$$\frac{(R_{n/p})_{\epsilon_1}}{(R_{n/p})_{\epsilon_2}} = \left(F_{np}\frac{N_n}{N_p}\right)_{\epsilon_1}\left(\frac{N_p}{F_{np}N_n}\right)_{\epsilon_2} \tag{4.24}$$

We can decompose the right hand side of Equation 4.23 further by applying a Taylor Series Expansion. If we consider a small range of epsilon values and small slope values (S^n), we can expand this expression via a Taylor Series as:

$$\S \sim \left(\frac{1+\epsilon_2 S^p}{1+\epsilon_1 S^p}\right)[1+(\epsilon_1-\epsilon_2)S^n + \cancelto{0}{\mathcal{O}(S^n)^2}] \tag{4.25}$$

where we have dropped the higher order terms. We can now combine Equations 4.24 and 4.25. This results in:

$$\left(F_{np}\frac{N_n}{N_p}\right)_{\epsilon_1}\left(\frac{N_p}{F_{np}N_n}\right)_{\epsilon_2} \approx \left(\frac{1+\epsilon_2 S^p}{1+\epsilon_1 S^p}\right)[1+(\epsilon_1-\epsilon_2)S^n]$$

Finally, we can solve for the Rosenbluth Slope of the neutron, S^n, in terms of our yields, scale factor ratios, Rosenbluth Slopes from proton parameterizations, and our epsilon values. The resulting equation is[8],

$$\boxed{S^n = \left[\left(F_{np}\frac{N_n}{N_p}\right)_{\epsilon_1}\left(\frac{N_p}{F_{np}N_n}\right)_{\epsilon_2} - \left(\frac{1+\epsilon_2 S^p}{1+\epsilon_1 S^p}\right)\right]\left[\left(\frac{1+\epsilon_2 S^p}{1+\epsilon_1 S^p}\right)(\epsilon_1-\epsilon_2)\right]^{-1}} \tag{4.26}$$

Calculating the contribution from neutron TPE

At this point in the calculation, we will have extracted the values of G_M^n for SBS8 and SBS9, used them to determine the reduced cross-section at each epsilon point, and then extracted the Rosenbluth Slope of the neutron, S^n, via Rosenbluth Separation. The Rosenbluth Separation Technique is sensitive to the effects of two-photon exchange whereas the Polarization Transfer method is not. Therefore, if we take the difference between our experimentally-extracted S^n and those from polarization transfer measurements, we can assess the difference between the two techniques:

$$\boxed{\text{nTPE} = S^n - (S^n)_{\text{OPE}}} \tag{4.27}$$

In Equation 4.27, S^n is our experimentally extracted Rosenbluth Slope for the neutron (between SBS8 and SBS9) and S^n_{OPE} is the Rosenbluth Slope calculated using parameterizations strictly from single-photon experiments. The choice of parameterization for S^n_{OPE} is critical. We are assessing the difference between measurements that are sensitive to the two-photon exchange and those limited to the single-photon interactions. Therefore, we need to be sure that S^n_{OPE} comes only from single-photon measurements.

[8] I am writing Eq. 4.26 as an equality and not as an approximation primarily for simplicity sake. At this point, it is assumed that we understand what goes into this formulation and that there are multiple underlying assumptions. However, given the scope of this analysis and my presumption that the approximations don't profoundly reduce the accuracy of the relation, I will write it as an equality from here on

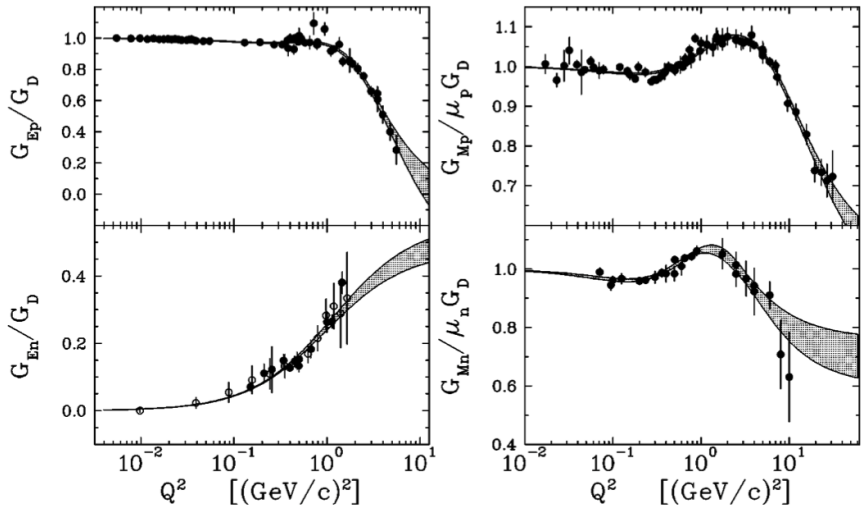


Figure 4-16: Global data and their parametric fits by Kelly for G_E^p, G_M^p, G_E^n, and G_M^n. g4sbs and SIMC use Kelly's parameterizations for G_E^p, G_M^p, and G_M^n, and Riordan's equation for G_E^n. Plots reproduced from [75].

4.7 A Note on World Data and Parameterizations

Experiments to measure the electromagnetic form factors of the nucleons can be traced back several decades (see Section 1.8). As experimental capabilities advance and Q^2 increases, more and more data points are added to the set of world data for these form factors. As these datasets expand, functional forms of the data emerge. Fits to these global data sets can be made which then provide a means to calculate form factor values at a particular kinematic point by way of parameterized functions.

A great example of how parameterizations are extracted from the global data set of nucleon electromagnetic form factor measurements is shown in Figure 4-16. These plots have been re-produced from [75]. The plots show the world data set for each associated nucleon form factor and the derived parameterization as the banded region tending to it. Essentially, the parameterization is the best fit across all given data points. There is much more behind the process than merely applying a fit, but for the scope of this paper that description should suffice. These experiments span multiple measurement techniques across various levels of precision, and each have their own unique constraints. The finer details of these considerations can be accessed from the referenced source materials directly.

The specific parameterizations that are important in this analysis are those built into the simulation machinery and those that will be used to calculate the counterpart form factors in the final steps of extraction. The second point here is discussed in more detail during its application in Chapter 6. On the first point, two parameterizations are implemented in g4sbs and SIMC for calculating the form factors and cross-sections: g4sbs implements Kelly's parameterization for G_E^p, G_M^p, and G_M^n ([75]), and Riordan's parameterization for G_E^n ([56]). An important point to note for an extraction of nTPE is that the Kelly data set used for parameterization emphasizes recoil or target polarization over other methods when possible [75]. Recall that recoil and target polarization methods are less sensitive to two-photon exchange contributions, and therefore, the calculations which follow from Kelly's fits should not be heavily influenced by TPE effects. This will become significant later when attempting to compare the extracted G_M^n values from this experiment with results from measurement techniques that are less affected by TPE effects.

Chapter 5

Exploratory Analysis

This chapter will build upon the experimental techniques discussed in previous chapters and discuss some of the intricacies of the analysis process. The goal is to realize the preceding discussions by transforming our experimental data into tangible physical results. We have laid the foundation and built our framework. Now comes the fun part. While I've tried my best to address numerous concerns, there are certainly aspects left untouched and further work remaining. It is essential to keep the exploratory nature of this thesis in mind throughout the following chapters.

5.1 Analysis Software

The software packages discussed thus far have all been vital to this analysis. However, they have dealt more with the so-called back-end of the data. For instance, the data that is generated by our Monte Carlo software needs to be further processed in order to be transformed into the plots presented throughout this thesis. This is true for the experimental data as well. The raw experimental data sits in a format which is directly inaccessible and not configured in any data structure that can be used for analysis. Therefore, we first need to implement a certain set of software tools to transform our raw data before we can start to play with it.

5.1.1 SBS Offline and SBS Replay

SBS Offline and SBS Replay are two coupled analysis packages that reconstruct, or *replay*, raw experimental data into a form that is properly structured, organized, and accessible. Technically speaking, SBS Replay is more of a library and repository that contains analyzer database files, scripts, and other macros. But more on this after a quick discussion on SBS Offline.

SBS Offline is the software which configures via the SBS Replay library in order to take in raw experimental data, and convert it to a standard SBS-format ROOT file. SBS offline contains SBS-specific libraries and source code for all of the detectors and sub-detectors. It houses all of the raw decoders for the DAQ interfaces such as MPDs, VTPs, and VETROC. Even though SBS Replay is called the "replay software", it only really initiates SBS Offline, which performs the actual replaying of the data. The term *replay* is chosen wisely here. In effect, SBS Offline replays the process of the data as it was collected during the experiment. In this way, it can take the raw data, and play it through all of the pathways and channels that the experimental data flowed through. This is why it must contain all of the detector decoders and libraries. It will replay the data through each associated detector and structure it as a ROOT tree with variables organized by detector, device, time, etc.

SBS Replay is more or less the access point for SBS Offline. It houses the databases used to configure a replay in SBS Offline. SBS Replay is the repository where the replay scripts are housed. These scripts are used to select specific configurations, databases, replay settings, and various other implementations before commencing a replay through the SBS Offline framework. Often the term *replay* will be used in reference to running data through the SBS Offline software via the SBS Replay interface.

Experimental data and simulated data both need to be replayed in order to be transformed into our standard format ROOT file. The simulation runs in the same way that a real experiment would. A simulated set of scattered particles is generated, and they interact digitally with everything downstream of them. This means that as

the simulated scattered electron traverses through the simulated GEM detectors, it creates simulated ions which have simulated avalanches, so on and so forth. Each of these processes, though simulated, are stored in the simulation output just as the raw data is stored during the taking of experimental data. Therefore, in its raw format, the simulated data is, like the raw experimental data, not wholly accessible until it is replayed[1]. The simulation data requires one extra step though, digitization.

In the live experiment, the real measurements spawn from raw analog signals in each detector. These are digitized pretty quickly after the analog signal is received in most detectors. For instance, the analog signal in a GEM is the charge deposited onto the readout strips. This charge is read and amplified by the APV card which is digitized by the MPD. The same process also needs to occur for the simulated dataset. The raw simulated data for the GEMs is in a sense sitting at the simulated APVs waiting to be digitized. Therefore, before replaying simulated data, it is transformed using a software package called *libsbsdig*. Libsbsdig converts g4sbs outputs such as position, energy, and time, into digitized detector signals that are in a data structure which can be directly used in the reconstruction and replay process [103]. Typically, I will refer to the full process of digitization and replay as simply a *replay*; a full replay is inherently assumed to incorporate the digitization process.

5.1.2 ROOT

The most integral analysis tool that we have in our SBS bag of tricks is ROOT. ROOT is a data analysis framework developed in C++ by CERN in the 1990's. It was designed to facilitate the challenges of high-energy physics data analysis. It is an incredibly powerful tool that is implemented in almost every step of this analysis. We access the ROOT framework via C++ scripts and ROOT's C++ interpreter which allows for direct access and manipulation of data. ROOT contains an enormous set of advanced statistical tools including multi-dimensional histograms, fitting tools, and

[1]Before it is replayed, the simulation data is accessible at a basic level. There is a foundation of so-called *cured* analysis performed during the generation of simulated data. This is stored in the pre-replay simulation files but is by no means in a *final state*.

various minimization functions. ROOT has a very robust set of visualization tools for plot design, creation, and export. Most importantly, ROOT allows for the handling of huge (on the order petabytes or more) datasets in a form that allows for advanced query and access to the data within (ROOT trees) [23]. Unless otherwise stated, the terms C++ and ROOT will be interchangeable throughout this thesis.

5.2 Initial Calibrations and Replay

Given the proper access rights and capabilities, anyone can take the raw experimental data, digitize them, and then replay them into ROOT files that have the standard SBS TTree structure. The process of this digitization and replay is straightforward. However, it can be very time consuming. There are also some preliminary cuts, thresholds, configurations, and many other definable parameters that can, and should be, set before the replay. In order to alleviate the process of everyone making their own ROOT files and to ensure that everyone is using a common set of baseline settings, a common set of replays were created for the entire set of runs for GMn and nTPE.

The first set of replays were created during the running of the experiment. This initial volume of replays was the zeroth *pass* through the data and is therefore referred to as *Pass 0*. While many of the pre-replay cuts and selections were determined during the commissioning of the experimental setup, many of them were also determined on the fly during the actual running of the experiment. During the live running of the experiment, there is constant monitoring and calibrating of the detectors and software. These are parameters such as *energy conversion factors* and *timing cuts*. As the system characteristics evolve over the runtime of the experiment, a familiarity with the various parameters develops and shapes the initial intuition of where to place the initial cuts, thresholds, etc. Therefore, though Pass 0 was replayed with the proper settings to get initial looks at the data, it had plenty of room for improvement. New rounds of improvements led to new, subsequent replay passes. These follow the previous name schema with a *Pass 1* on data shortly after the end of the experiment

and a *Pass 2* which was performed in January of 2024.

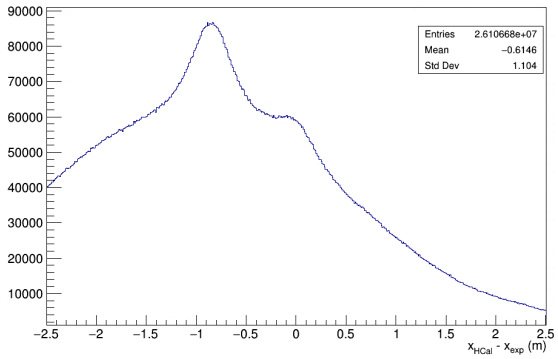
The major changes between each pass were primarily focused on calibrations to the detector packages (BBCal, HCal, GEMs, GRINCH, etc.). The first pass (Pass 1) included calibrations from all detector systems with the primary ones being for BBCal and HCal. As a GEM student, my calibration task was to assess and correct *digital crosstalk* in the GEM data [20]. Other GEM-related calibrations for Pass 1 included Gain Matching of the APVs ([137]) and Time Deconvolution ([108]). Though Pass 0 and Pass 1 both showed great looking data that gave us very promising outlooks, it still required additional calibration to get it to an *exploratory analysis* stage. Further passes beyond Pass 2 will most likely be required for the data to be considered for preliminary analysis.

To the uninitiated, spotting the differences between Passes 0, 1, and 2 may be a challenge of patience. Nonetheless, it should be noted that each subsequent pass has provided better and better data each time. Background and spurious signals are cleaned up through calibrations such as those mentioned above, as well as many others. Timing and energy calibrations on BBCal and HCal allow for finer and more accurate gradations when applying cuts and selections. Calibrations to the GRINCH allow for better particle ID (PID) of pions which make selecting them out a much less-crude process. The exhaustive list of calibrations and their effect is quite extensive, and will be compiled slowly over time. For now, two major sources of documentation exist for referencing BBCal ([45]) and HCal ([121]) calibrations.

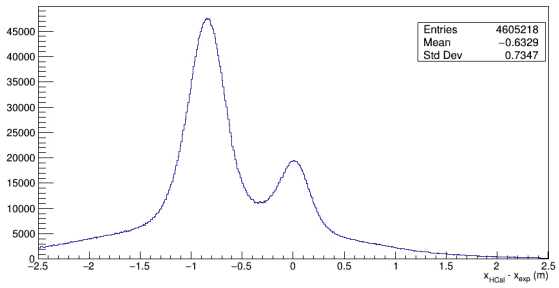
5.3 Elastic Event Selection and Cuts

The fully replayed files in each pass contain all events that will be used for this analysis. Though some preliminary calibrations and cuts were applied before the full *cooking of data*, we can rest assured that any events removed were either background, spurious, not quasi-elastic, or simply *bad*. This analysis relies on a *clean* extraction of quasi-elastically scattered electrons, protons, and neutrons. Therefore, we need to apply various physics-motivated cuts and thresholds to our data in order to help

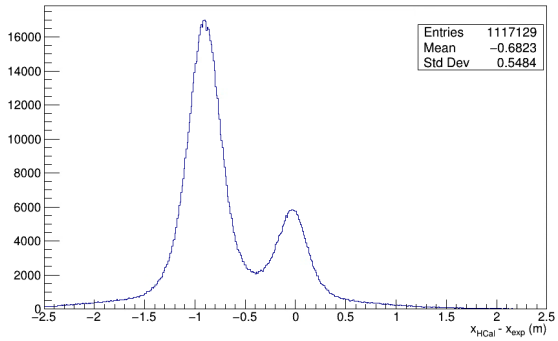
select out these *good* events.



(a) SBS8 dx, no cuts.



(b) SBS8 dx, basic cuts.



(c) SBS8 dx, all quasi-elastic cuts.

Figure 5-1: The evolution of the SBS8 dx plots as more cuts are applied. **(a)** No cuts applied. You can make out the tops of the nucleon peaks but beyond that, any further shape is difficult to resolve. **(b)** Energy and timing cuts have been applied. These include cuts on BBCal PreShower, the invariant mass W^2, HCal min. cluster energy, etc. Here, we can see the two nucleon peaks. **(c)** The HCal acceptance-matching fiducial cut is applied and further cleans up the plot.

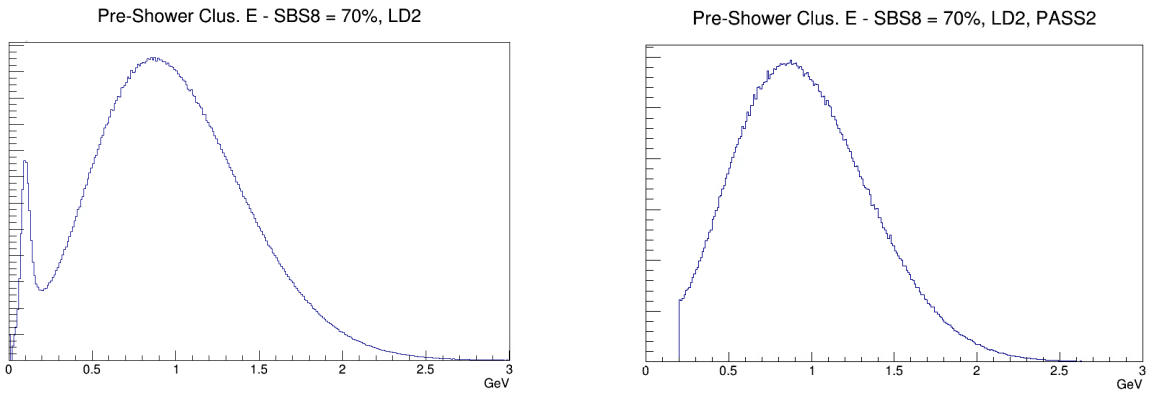
The effect of applying quasi-elastic cuts can be seen in Figure 5-1. This set of plots shows the effect of adding additional cuts onto a data set. The first figure (Fig. 5-1a) is the data as it arrives from raw replays of Pass 2. No cuts have been applied to this data while producing this particular dx plot. In Fig 5-1a we can make out the tops

of the two nucleon peaks; the neutron peak is just barely visible. From this plot, it would be nearly impossible to extract any reasonable yield from either nucleon. The next figure (Fig. 5-1b) incorporates a general set of top-level quasi-elastic selection cuts. These cuts include a BBCal PreShower minimum energy threshold, a minimum cluster energy for HCal blocks, a minimum number of hits on GEMs, a bandpass cut on the invariant mass (W^2), and a vertex cut which ensures that each event originates from within the target volume.

The differences between Figures 5-1a and 5-1b are striking. We can now easily resolve the two peaks independently and could probably make a decent extraction to this dataset as is. However, there is one important cut we still need to make. Though it may not be visible to the naked eye (at least we hope not at this point), there could be some discrepancies between the populations of the two peaks. We need to apply a so-called *fiducial cut*. This cut is applied in Figure 5-1c. The fiducial cut is an HCal acceptance-matching cut which helps correct for imbalances in proton and neutron yields on HCal due to the deflections caused by the SBS magnet (see Sec. 5.3.7). A proper simulation fit to the data and subsequent extraction can now be underway. The major quasi-elastic cuts and their effects are discussed further in the following sections.

5.3.1 BBCal PreShower (Pion) Cut

The BigBite Calorimeter has two independent layers of blocks which make it up. These are the PreShower (PS) blocks followed by the shower (SH) blocks. As the name implies, the PreShower layer is positioned in front (target side) of the shower blocks and is the layer in which the first particle *showers* are produced. An *electromagnetic* shower is produced when high energy electrons deposit their energy into the lead-glass blocks of BBCal. The layers are designed such that the thickness of a PreShower block is not enough to fully stop oncoming electrons, but will deplete a fraction of their energy. The rest of the energy is dumped entirely into the shower layer. Because of the energy profile, the PS energy deposit can be used as a handle for pion background rejection [44]. Pions are heavier than electrons, and are therefore harder to stop than



(a) **(b)**

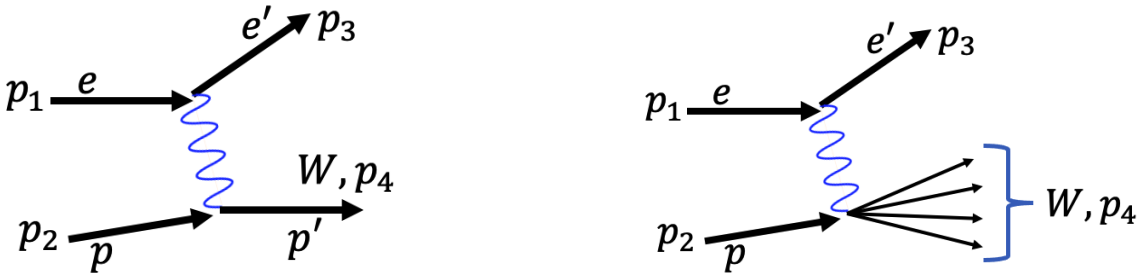
Figure 5-2: BigBite Calorimeter PreShower cluster energy without **(a)** and with **(b)** the pion-rejection cut. **(a)** The pion peak can be seen on the left side of the plot around 0.130 GeV. **(b)** A pion threshold cut has been applied at 200 MeV.

electrons. Because of this, the electrons deposit more energy than pions in the BBCal PreShower. This is the basis of the BBCal PreShower cut. The mass of the pion is approximately 130 MeV [129]. When we look at a plot of the BBCal PreShower energy (Fig. 5-2a) we can see an initial peak on the left side of the plot. This corresponds to the energy deposited by pions. In order to apply a safe rejection of pions, we can cut everything below 200 MeV. This provides a very effective pion cut while also not reducing our dataset too greatly. See Figure 5-2b.

5.3.2 Invariant Mass Cuts

A very important and powerful cut for elastic event selection is the *invariant mass cut*. In this analysis the invariant mass is defined by W. We can also make reference to the invariant mass squared, W^2. Let us work through a quick derivation of the invariant mass so that we can understand the cut and its applications a bit better.

Figure 5-3 shows the basic kinematic diagrams for elastic (Fig. 5-3a) and inelastic (Fig. 5-3b) scattering of an electron incident on a target proton. Figure 5-3b shows a proton begin to split into multiple pieces. This would be the case for *deep inelastic scattering*. In the case of only excitation of the proton, we have simply *inelastic scattering*. In Figure 5-3, p_1 and p_3 correspond to the electron's incoming and outgoing momentum, respectively. Similarly, the proton has p_2 and p_4 for its incoming and



(a) Elastic scattering **(b)** Inelastic scattering

Figure 5-3: Elastic and inelastic scattering of an electron and target proton.

outgoing momentums. From these diagrams we can form the following relations:

$$\begin{aligned} \text{4 Momentum Transfer to the proton: } q &= p_1 - p_3 \\ \text{Invariant Mass: } W^2 &= (q + p_2)^2 \\ &= q^2 + 2qp_2 + p_2^2 \\ &= q^2 + 2qp_2 + M_p^2 \end{aligned}$$

Here, M_p is the mass of the proton. Due to baryon number conservation, we have that $W^2 \geq M_p^2$ since the proton is the lightest baryon. From this we arrive at the following:

$$W^2 - M_p^2 = q^2 + 2qp_2 \left\{ \begin{array}{ll} = 0 & \text{Elastic Scattering} \\ > 0 & \text{Inelastic Scattering} \end{array} \right\} \tag{5.1}$$

Isolating just the elastic term and applying that the proton is initially at rest ($p_2 = (M_p, 0, 0, 0)$) and $W^2 \geq M_p^2$, we get:

$$W^2 - M_p^2 = \cancelto{0}{(q^2 + 2qp_2)} \text{ (Elastic)}$$

or

$$W^2 = M_p^2$$

Therefore, we can apply a bandpass cut on W^2 centralized around the average nucleon

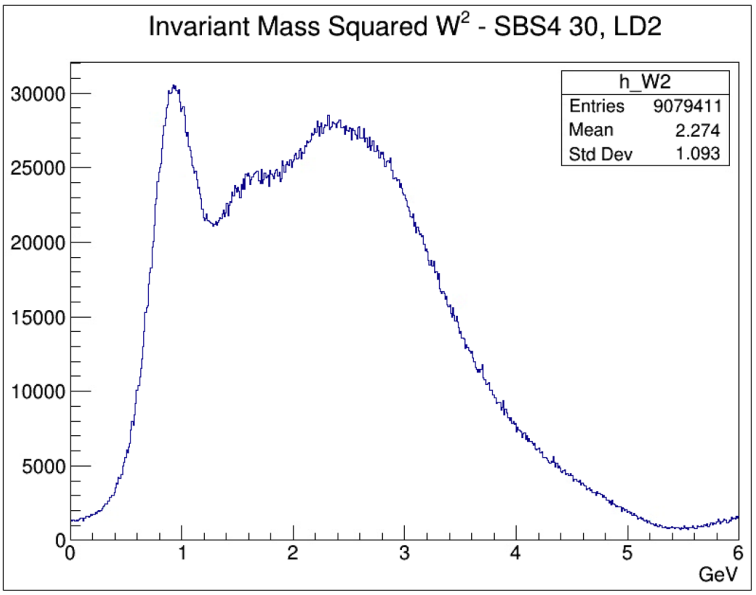


Figure 5-4: Invariant mass, W^2 for SBS4 with an SBS field scale of 30% and an LD2 target.

mass squared (≈ 0.880 GeV^2)[2]. W^2 for SBS4 is shown in Figure 5-4. On this plot you can see a peak jutting out at around 1 GeV. This corresponds to our elastic peak. This peak is not centralized around 0.880. This is primarily due to background contamination. It is also helpful to look at regions where Eq. 5.1 is greater than zero. This helps to characterize the various inelastic backgrounds present in our data.

To determine the W^2 cut, we first fit a Gaussian to the elastic peak region of W^2. The use of a Gaussian fit here has no physics-motivation behind it. It simply works well and is easy to apply. The Gaussian fit in ROOT provides three fit parameters: a Gaussian norm, mean, and sigma. The mean will provide us with the central value of the W^2 cut, and the sigma will give us a means to set upper and lower bounds. For instance, a W^2 cut may be applied in the form:

$$(W^2_{\text{(Gaus. mean)}} - W^2_{\text{(Gaus. sigma)}}) \leq W^2 \leq (W^2_{\text{(Gaus. mean)}} + W^2_{\text{(Gaus. sigma)}})$$

where $W^2_{\text{(Gaus. mean)}}$ and $W^2_{\text{(Gaus. sigma)}}$ are the Gaussian mean and sigma from the fit, respectively.

[2]The nucleon masses are nearly equal, i.e., $M_n \approx M_p$. Therefore, we can use the average of the nucleon masses, M_N, to approximate values across both nucleons. Therefore, we use $M_N = {}^{(M_n+M_p)}/_2$ in place of either M_n or M_p.

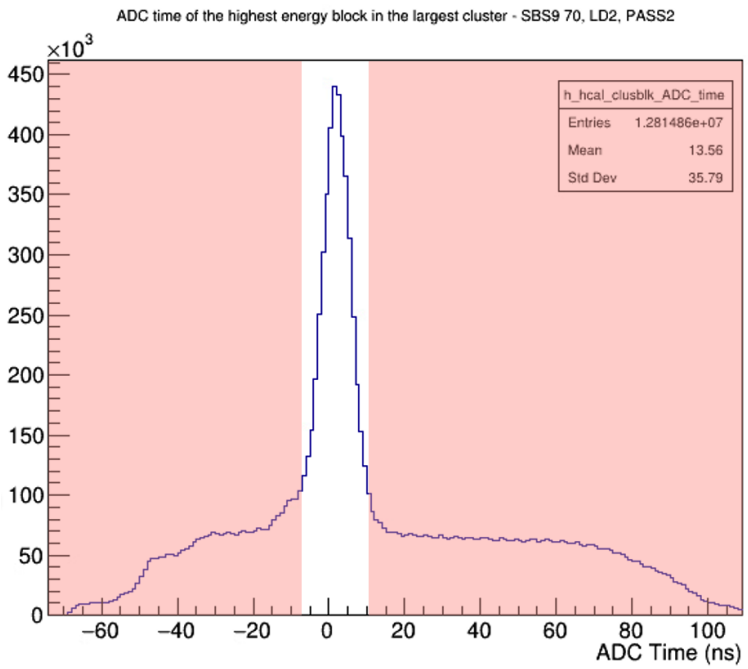


Figure 5-5: ADC time for highest energy block in largest BigBite cluster. The red shaded regions show approximately where a bandpass cut would be applied. This cut would reject events which do not coincide with the timing window of the highest energy measurements and thus events which are probably not quasi-elastic.

5.3.3 Timing Cut

We can make a quasi-elastic cut on the ADC time of the highest energy BigBite block in the largest cluster on BigBite. It is highly likely that this event would be associated with a primary quasi-elastic event. If we look at Figure 5-5, we can see a sharp peak centered around zero. On either side of the peak we can see a plateau that is associated with background. Therefore, we can place a *bandpass* cut around this central peak and only accept events which arrive within this time window. Any out-of-time events will be rejected by this cut. These rejected events are represented by the red shaded regions on the plot.

5.3.4 HCal Cluster Selection

As scattered particles arrive at HCal, they deposit their energies into the detector's blocks (also called *modules* or *cells*) in the form of hadronic and electromagnetic

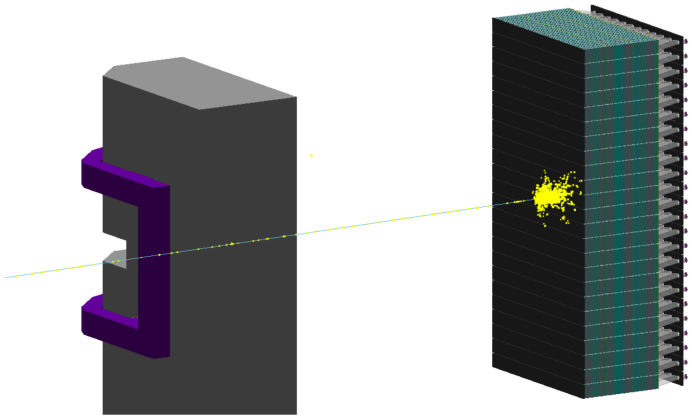


Figure 5-6: Simulated scattering event and cluster on HCal. Image source: [39].

showers (see Fig. 5-6). This results in multiple clusters of signals from multiple HCal modules. In these clusters, there is one block considered to be the *primary block*. This is the block with the highest energy cluster. The cluster with the highest energy primary block is the default selection in the analysis framework. Though this is often the right choice, it isn't guaranteed[3]. To improve this selection process, I implemented an additional *best cluster selection* algorithm which extends the functionality of the default selection method by introducing additional correlations with block/cluster coincidence times.

The basic principle of this new *best cluster selection* method is as follows. The clusters are sorted in decreasing order by their total energy. This removes any previous ambiguity in energy sorting which may have already been present[3]. Then, for each cluster we lookup the corresponding ADC time (similar to what was discussed in Section 5.3.3 but using HCal ADC time instead of BBCal timing). Then we take the mean value from the central peak of our ADC time plot (similar to that of Fig. 5-5 but again, using HCal timing information) and take the difference between the ADC time of each cluster and the overall mean ADC time, i.e., $\text{ADC}_{\text{Diff. time}} = \left|\text{ADC}_{\text{Mean time}} - \text{ADC}_{\text{Cluster time}}\right|$.

There are multiple ways in which we can implement this ADC time difference into

[3]It was discovered that some events weren't properly sorted by their energy. The zeroth element of the variable array should always be the highest energy; however, a few percent of total events were found to not have been properly sorted. A manual *pre-sorting* eschewed any instances of improper sorting.

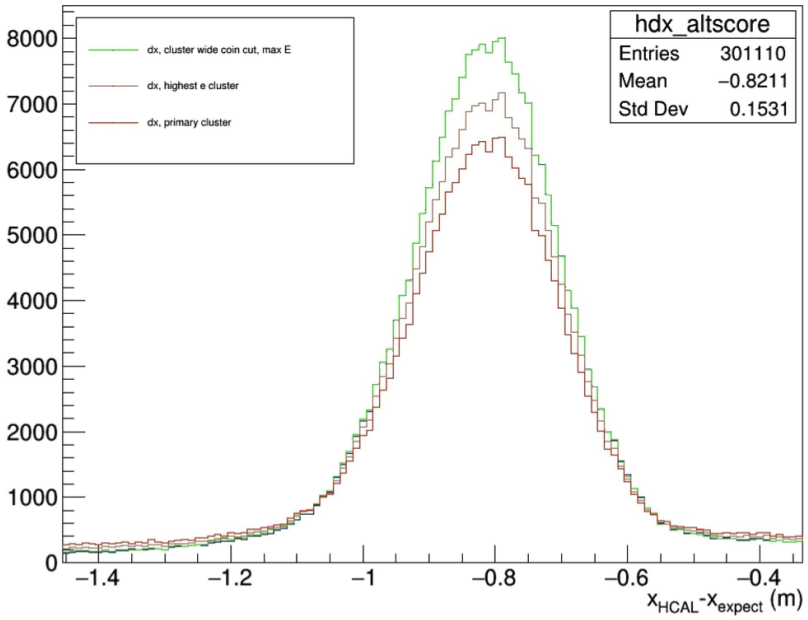


Figure 5-7: Comparison of cluster selection methods for LH2. Image source: [122].

our best cluster selection process[4]. My approach is to simply put an upper limit on accepted $\text{ADC}_{\text{Diff. time}}$ values and then select the cluster corresponding to the highest energy from that subset. This is basically applying an additional and tighter ADC time cut on just the subset of highest energy HCal clusters.

A comparison of various best cluster selection methods is shown in Figure 5-7. This plot shows the number of quasi-elastic events extracted from a dataset using different HCal best cluster selection methods. My selection algorithm is shown in green. The two brown lines correspond to methods which make selections based purely on highest energy. The light brown line is the selection method which uses only the highest energy cluster, and the dark brown line is the default primary cluster assigned to the zeroth array element discussed above. The light brown line shows that simply pre-sorting the HCal cluster array by energy and then selecting the highest energy element noticeably improves upon the default primary cluster. Above that, we see marked improvement by incorporating coincidence timing into the selection process.

[4]Several alternative methods and approaches have been tested and analyzed by others and myself. These include a score-based selection algorithm and selections based on a weighted probability distribution function for ADC time and energy. All of these methods show good potential and would need to be highly scrutinized in a final analysis.

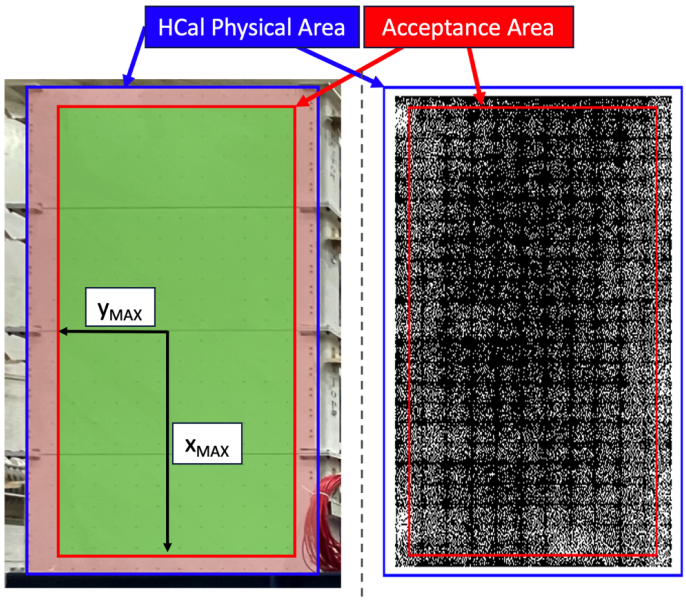


Figure 5-8: The HCal acceptance cut. **(a)** A picture of HCal which shows the boundary of its physical area in blue and the region defined as the acceptance area in red. **(b)** An HCal hitmap showing the regions of physical area in blue and the acceptance cut in red.

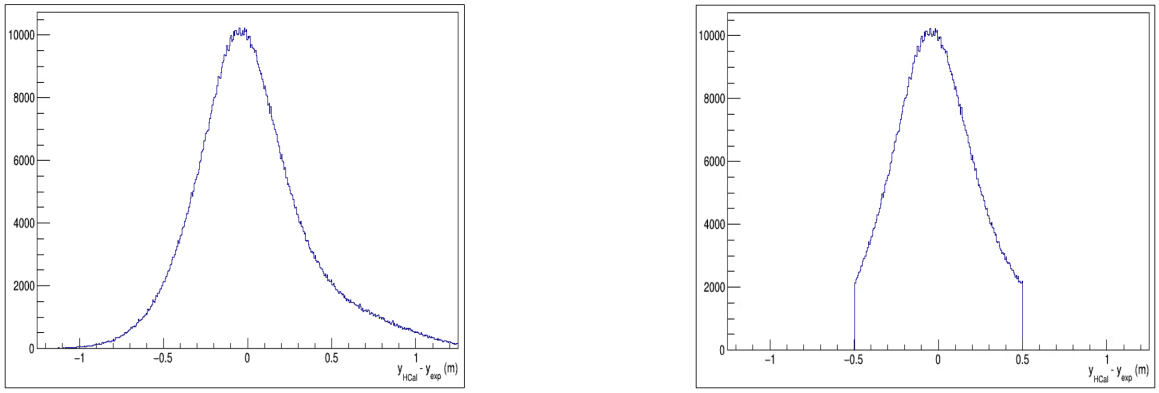
5.3.5 HCal Acceptance Cut

The Hadron Calorimeter is made up of 288 individual modules, or blocks. There are, in total, twenty four rows and twelve columns of blocks. Each block is fitted with a light guide and is capable of effectively measuring deposited energies. The acceptance area of HCal is very large and was designed to minimize the difference between the ratio of acceptance-to-efficiency for neutrons-to-protons. We want to ensure that we are measuring the total amount of energy for any scattered particle that lands on HCal. Therefore, we have to be careful about particles which land near the outer edge of the detector. If they land in such a way that parts of their shower can escape the detector without being fully sampled, it could result in erroneous energy and position measurements. Therefore, as a safety precaution, an acceptance cut is placed on the active area of HCal. Figure 5-8 shows the physical boundary of the grid of HCal modules and the smaller region used to define the acceptance area. In each figure, the blue line marks the physical surface area of the face of HCal while the red region marks the area of the acceptance cut.

The acceptance cut on HCal is performed by setting coordinate thresholds based on boundaries defined by an integer number of HCal blocks. For instance, the most basic acceptance cut would be to remove one block from the perimeter. This basic perimeter cut removes events with clusters that have landed on the edge of the calorimeter and are poorly reconstructed. This is precisely the region that is defined in red in Figure 5-8 and the basis for the HCal acceptance cut used in this analysis.

5.3.6 dy cuts

The utility of the dy plot is discussed in Section 4.3.2. Recall that in this analysis, "x" is considered the dispersive axis since protons and neutrons separate along this axis. The Y-axis on the other hand, is non-dispersive and is a representation of the momentum distribution, Fermi smearing, and radiative effects throughout the events. Figure 5-9a shows a standard dy plot without any cuts. During calibration, a Gaussian will be fit to the central portion of this distribution. From this fit, we get the Gaussian mean and sigma. The dy cut thresholds are defined using the central value defined by the Gaussian mean and some multiple of the Gaussian sigma on either side. For instance, a typical SBS8 dy cut will remove anything from beyond two sigmas of the mean. The cut shown in Fig. 5-9b is an overly aggressive and exaggerated cut for representative purposes only.



(a) **(b)**

Figure 5-9: The dy plot without **(a)** and with **(b)** cuts. Note, the dy cut shown in the **(b)** is an exaggerated and aggressive cut for display purposes only. The standard dy cut is typically around two Gaussian sigmas wide.

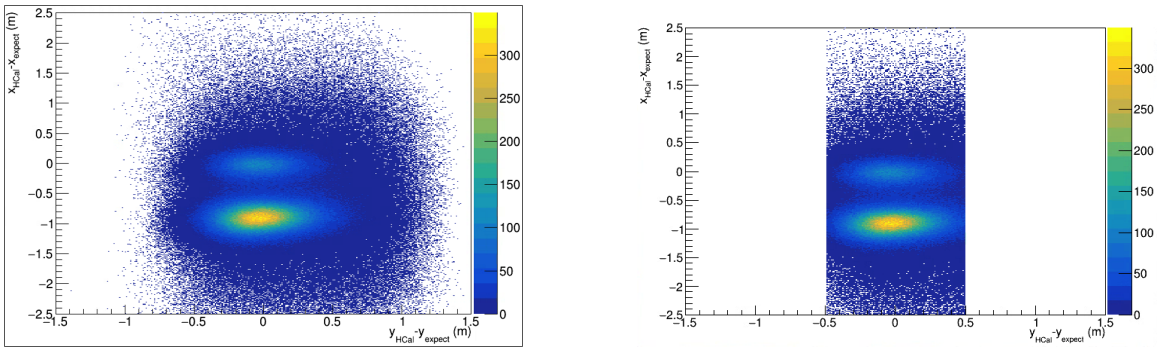


Figure 5-10: Comparison of a dxdy plot with and without dy cuts. **(a)** No dy cut. **(b)** With a dy cut. This cut allows us to remove a fair amount of background in the non-dispersive direction.

This cut allows us to isolate events which are correlated along the dispersive axis. We are inherently cutting events which may be quasi-elastic but are beyond our average momentum expectations. By tightening in on the dy region, we select events which contribute along the central axis of dispersion. This cut also removes a fair amount of background. A better way to visualize this cut is to see its effect on the two-dimensional dxdy plot. This is shown in Figure 5-10. Figure 5-10a shows a standard dxdy cut while in Figure 5-10b, we see that either side of the central peak areas have been removed. This is an easy and effective way of removing non-elastic background in the areas outside of our elastic peak regions. We have to be careful about getting too close to the regions of our elastic peaks, but even a conservative cut on dy at say, ± 0.75 on the plots shown, would have a noticeable and positive effect.

5.3.7 Fiducial Cut

In order to remove any spurious nucleon bias, a so-called *fiducial cut* is required. The fiducial cut is, in a sense, an *acceptance matching* cut which ensures an equal acceptance region for both neutrons and protons. An imbalance can arise when the deflection on the proton is such that only one nucleon can land on HCal while the other would miss the detector completely. One case is shown in Figure 5-11. In both of these figures, the direction of proton deflection is downwards. In Figure 5-

11a we have the case where a neutron lands near the lower portion of the plot (this corresponds to the top of HCal in Hall A). In this case, the proton is deflected so much that it would miss HCal and fly over it. Therefore, we would record a neutron but would never be able to record the proton that corresponds to the same kinematic event. To account for this, we can apply a boundary threshold for accepted hits.

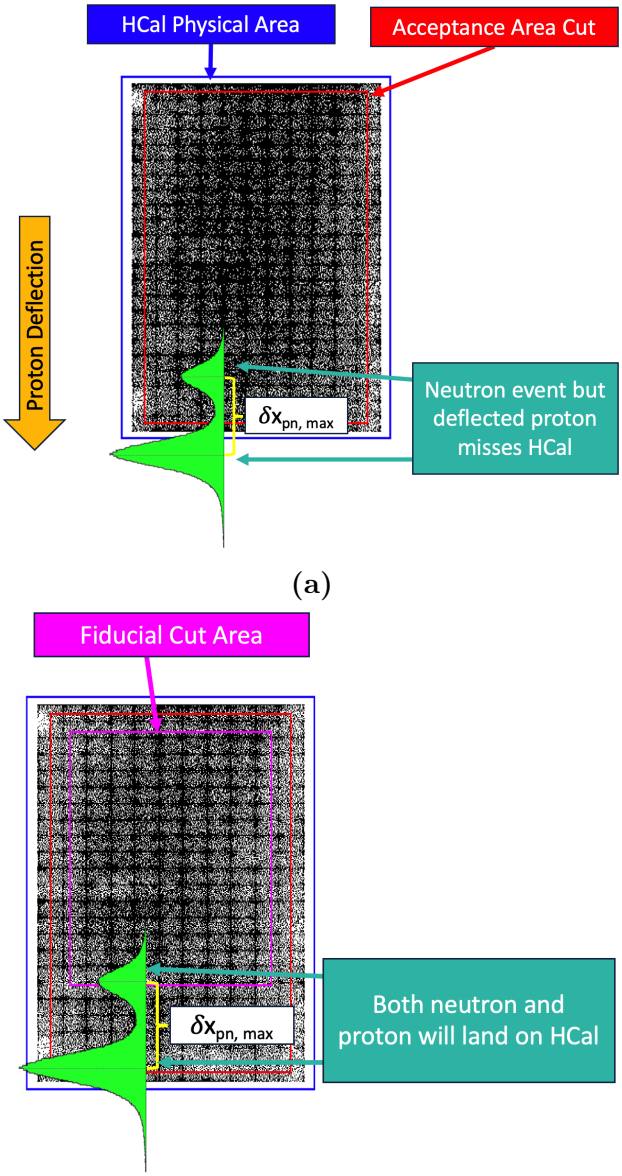


Figure 5-11: The fiducial cut. **(a)** The case when a neutron lands within the acceptance area of HCal but the deflected proton misses the detector. In this case, a neutron is detected but there is no proton. **(b)** A fiducial cut is applied such that we only keep neutrons if the deflected proton is able to land on HCal.

The fiducial boundary is shown in Figure 5-11b as the magenta region. In this case, if a neutron lands on this lower fiducial boundary, there is enough clearance to the top of HCal such that a deflected proton can still land on the detector. If a neutron passes below this boundary (in this figure's perspective), we reject that event. This effectively removes any bias of neutrons over protons in these cases. The same accounting is handled on the other side of HCal as well. It could be possible that a neutron would miss the bottom of HCal, but a deflected proton is bent up enough to land on the detector. The fiducial boundary accounts for the deflection distance between neutrons and protons on the top and bottom edges of HCal.

The fiducial cut is very important, as it directly affects the validity of our neutron-to-proton yield ratio, a quantity at the heart of this analysis. There are some critical considerations that must be maintained while devising and implementing the fiducial cut. These are:

1. The primary axis affected by the fiducial cut is the dispersive X-axis
   (a) Proton deflection is along the X-axis and is the leading reason to apply the fiducial cut.
   (b) The fiducial cut is also applied along the Y-axis. Even though we don't expect any imbalance or nucleon bias in this direction, the fiducial cut along this dimension does provide slight corrections to variances in momentum, Fermi smearing, and radiative effects.
2. The fiducial boundary relies on the average proton deflection distance per kinematic. This is shown in Figure 5-11 as $\delta x_{\text{pn, max}}$[5].
   (a) For all events in a given kinematic (SBS field scale setting, target, beam energy, Q^2, etc.) we form a dx plot like that of Fig. 5-1b. From there, we can extract the average distance between the central values of the nucleon peaks for the entire kinematic.

[5]The naming of the variable $\delta x_{\text{pn, max}}$ is a bit mis-leading here. This terminology derives from the implementation in software of the average observed separation distance between the protons and neutrons being used as a limit in the fiducial cut. A better name would be $\delta x_{\text{pn, ave}}$.

(b) Once determined, the average proton deflection value is stored in configuration files and becomes *fixed* for that entire kinematic. These values are retrieved at the onset of any analysis or extraction.

3. The application of the fiducial cut is based purely on expected nucleon positions as calculated by measurements from BigBite, i.e., only electron information.

   (a) From electron information alone, we can project an expected hit position for the scattered nucleon on HCal.

   (b) The criteria (or *logical truth value*) which determines whether or not the analysis applies the fiducial cut for a given event, is determined by the coordinates of the fiducial boundary, the projected/expected (x,y) hit positions on HCal, and the average proton deflection distance for that kinematic.

   (c) For each event, the average proton deflection value is added to the expected x hit position on HCal. If this value falls outside of the pre-determined fiducial boundary, the event is *cut*.

   (d) There is no dispersion along the Y-axis so, instead of the average proton deflection value, I use half the *width* of the dy peak (see Fig. 5-9a) as the cut margin.

4. A safety margin, or multiplier, called f_{mult} is introduced into the fiducial cut algorithm to allow for variable tuning.

   (a) The value for f_{mult} is determined per kinematic through a calibration process. Once determined, this value is then *fixed* and retrieved from configuration files at the onset of any analysis or extraction.

   (b) f_{mult} provides a means to account for the width of the proton and neutron peaks in the dx plots. The average proton deflection value is based on the distance between the centers of each nucleon peaks. This does not account for the *widths* of each nucleon peak. Introducing the multiplier, f_{mult}, essentially provides the ability to accommodate those features.

(c) Initially, f_{mult} was chosen *empirically* based on widths determined from fits to the proton and neutron peaks. Eventually, the value for f_{mult} was determined via studies that assessed the stability of the cut.

5. The pseudo-algorithm for the fiducial cut is as follows:

Fiducial cut along HCal X-axis:

Apply X fiducial cut:

$(f_{mult})(x_{(\text{HCal, Expected})} + \text{Ave. Proton Deflection}) > \text{X Fiducial Boundary}$

No X fiducial cut applied:

$(f_{mult})(x_{(\text{HCal, Expected})} + \text{Ave. Proton Deflection}) < \text{X Fiducial Boundary}$

Fiducial cut along HCal Y-axis:

Apply Y fiducial cut:

$(f_{mult})(y_{(\text{HCal, Expected})} + \left(\frac{1}{2}\right)(\text{dy peak width})) > \text{Y Fiducial Boundary}$

No Y fiducial cut applied:

$(f_{mult})(y_{(\text{HCal, Expected})} + \left(\frac{1}{2}\right)(\text{dy peak width})) < \text{Y Fiducial Boundary}$

6. The same fiducial cuts that are applied to the data are also applied to the Monte Carlo simulations.

   (a) In fact, ***all*** cuts applied to the experimental data are also applied to the simulation. The simulated data is held to the exact same cut and selection criteria as the experimental data.

   (b) The HCal acceptance and fiducial boundary areas are determined from experimental data and then matched in the simulation environment.

The importance of the fiducial cut cannot be over stressed. It ensures that we maintain a balanced and fair counting of the nucleons in our experiment. It may be possible for one nucleon to land on HCal while the other one doesn't. The proton could get deflected by the SBS magnet such that it flies over the top of HCal but the un-deflected neutron still hits HCal. Alternatively, a low-scattering neutron could fly

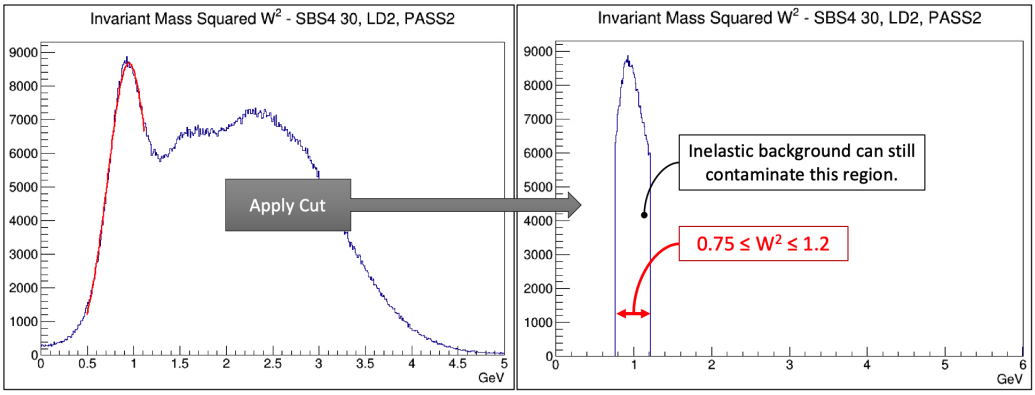


Figure 5-12: Comparisons between W^2 with and without a tight, sharp cut. Even with a tight cut around $0.75 \leq W^2 \leq 1.2$ we could expect the inelastic background to the right of $W^2 = 1.2$ to *bleed* into our cut region.

below the detector and miss, while the proton is deflected upwards and still lands on HCal. The fiducial cut allows us to account and correct for these imbalances.

5.3.8 Background Subtraction

In previous sections we have discussed various ways of applying cuts on energy thresholds, coordinate boundaries, timing coincidences, and invariant mass. Each of these processes provide great elastic event isolation and background rejection. However, even after applying them with the tightest of constraints, we will still have contamination from various inelastic sources. If we look at Figure 5-12, we can see that if we keep only events for which $0.75 \leq W^2 \leq 1.2$, this isn't a clean cut and will have sharp cut-off boundaries. Therefore, even though we reject the values beyond this selection range, we can expect to have some contamination and bleed through within our elastic cut. W^2 isn't the only source of contamination; there are multiple contributions to our background that will need to be handled. My approach is to handle them all at once using a background subtraction method. I first determine a fit function for the background contamination and then subtract that from the data. In order to visualize how the background can be accounted for, we can approximate the shapes of the nucleon peaks and the background, and implement a total fit onto our dx plot (see Fig. 5-13).

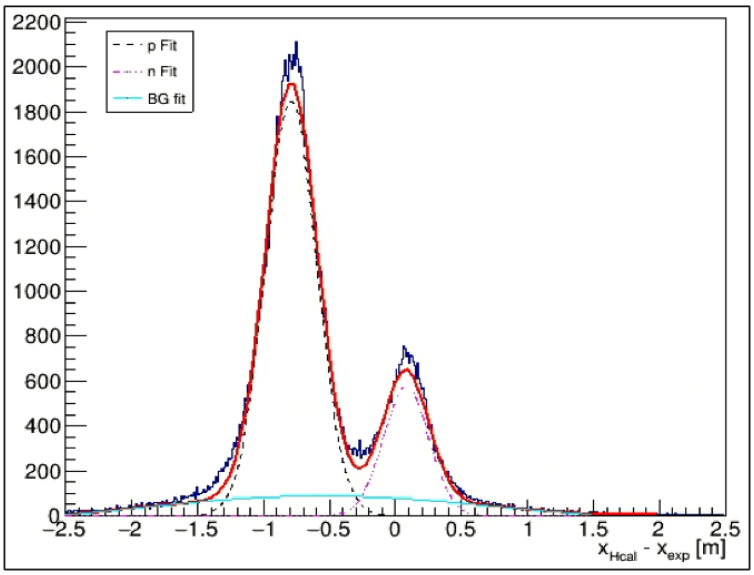


Figure 5-13: A *total dx fit* comprised of three functions: a Gaussian for the proton, a Gaussian for the neutron, and a third order polynomial for the background.

The total fit function in Figure 5-13 is made of three separate fits: a Gaussian fit for the proton peak (black dashed), a Gaussian fit for the neutron peak (magenta dashed), and a polynomial fit for the background (cyan)[6]. The total fit from these three ($p_{\text{Gaus}} + n_{\text{Gaus}} + \text{BG}_{\text{polyN}}$) is shown as the solid red line.

The use of Gaussian fits to the nucleon peaks is not motivated by any physics-driven insight for the shapes of those peaks. Gaussians do not fit our peaks exactly, but can come reasonably close. They also provide very useful fit parameters that go beyond the implications of the Gaussian function itself. From the Gaussian fit we obtain three fit parameters: *norm*, *mean*, and *sigma*. The norm is useful for performing quick data-scaling or amplitude-matching. The mean provides us with the central value of the peak and is a parameter that is used countless times in this analysis. The Gaussian sigma provides a value that can be related to the *width* of our peak — this parameter is also used time and time again throughout this entire analysis. The choice to fit with Gaussians mainly comes down to the fact that they are quick and easy to apply, they fit reasonably well, and they provide extremely

[6]The order of polynomial fit used for the background varies depending on the background one expects or has calibrated against. In order to discourage any over-fitting, we will always opt for the lowest order polynomial which provides a reasonable fit.

useful fit parameters.

Similar comments can be said for the background function. There is no *a priori*, or physics-motivated, reasoning for using a polynomial fit to the background. A polynomial fit is easy to apply and provides relatively nice fits. Also, the motivation for which order of polynomial to use comes down to the one that provides the fit with the lowest χ^2 value. Ultimately, the choice of what background function to use in the final analysis will need to be rigorously studied and critiqued before a final result can be established. Those studies are beyond the scope of this analysis.

The total fit (red) in Fig. 5-13 fits the dx plot well, with the exception of the *tail* regions and the *trough* between the peaks. We can also see a disagreement between the fit and data at the tips of each peak. Despite this, the overall fit provides a reasonable approximation of the background function (cyan).

The background function that I will subtract from data doesn't come this total fit function. The total fit is purely speculative and will not provide an accurate *shape*. However, it will be useful later when I need to normalize my eventual background function to the experimental data. The function that I employ will come from a separate fit to some form of background data. It is very important to consider what data will be used for this fit process. A couple of options were considered in this analysis. One option is to employ our Monte Carlo inelastic generator. The second is to perform a so-called *dy anti-cut projection*. This technique involves cutting the central elastic neutron peak regions out of the two-dimensional dxdy plot and projecting the remainder of the histogram onto the Y-axis. This projected data comes from points outside of the elastic peak regions and corresponds to *good background*. Both techniques were tested, but I opted to implement the dy anti-cut projection method.

The MC inelastic generator has yet to be fully characterized, or *vetted*, and is often referenced as a *toy model*. On the other hand, the dy anti-cut method uses real experimental data, but it may suffer from statistical limitations — we essentially have infinite statistics with any MC generator. Ultimately, one would need to repeat the analysis for each background subtraction method and compare the differences in

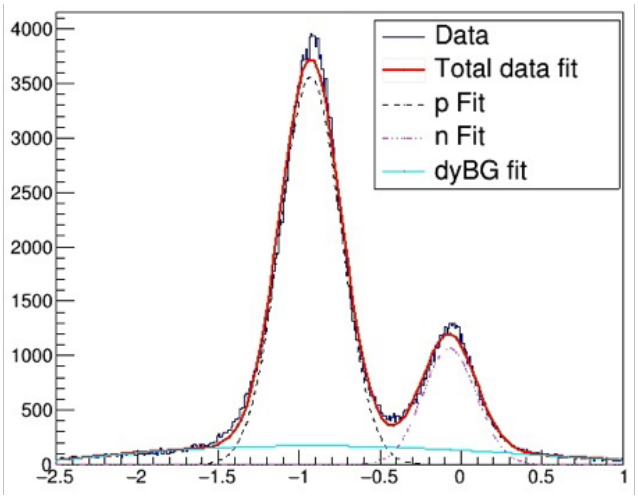


Figure 5-14: Initial *total dx fit* with a proton Gaussian, neutron Gaussian, and a third order polynomial for the background.

the results. This could then be used to properly estimate the systematics associated with each method.

Background Subtraction Method

1) Initial Fit: First, we perform an initial fit to our dx data (Fig. 5-14). The total fit is three-part: Gaussian proton, Gaussian neutron, and an N-order polynomial. The order of the polynomial doesn't need to be known exactly *a priori*. However, it should match the eventual order used in the background fit. If necessary, an iterative approach can be taken here for determining the order N of the background polynomial. We will use many of the parameters from this total fit later in our process.

2) dxdy Anti-cut: Next, we select a dy region which will be removed with an anti-cut. This dy region should be far enough away from the elastic region so as to not contain any *elastic contamination*, but not too far away that it is statistics-limited. A cut, fit, and minimization process can be used to select these bounds, or it can be achieved simply through trial and error. A 3-sigma dy anti-cut is shown in Figure 5-15.

3) Anti-cut dxdy Y-Projection and fit: Once we have the dxdy with the central region removed (dy anti-cut), we can project the remainder of the histogram onto the Y-axis. This is gives us a histogram in dx of inelastic data — our background

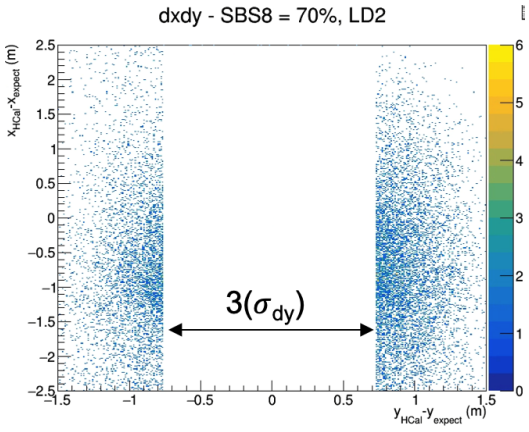


Figure 5-15: A dxdy plot with a 3-sigma dy anti-cut.

from the anti-cut. We can fit this dataset and extract a function for our background. Figure 5-16 shows the Y-projection and fit.

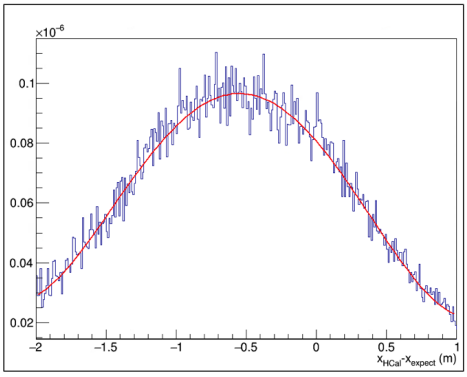


Figure 5-16: Inelastic background with polynomial fit.

4) Scale the Background Function: The Y-projection from the dxdy anti-cut provides a great data set for modelling the inelastic background. However, it will not be normalized to the background signal we would expect to find under our elastic peaks. Therefore, we need to scale it up accordingly. This is where the total dx fit from the first step comes into play. The initial total dx fit includes a background function normalized to the overall dx plot. Therefore, we can use the parameters from the initial background fit as scale factors for normalization.

Consider a 3rd order polynomial. Our background fits then have the form:

$$\text{Initial background :} \quad A_i + B_i x + C_i x^2 + D_i x^3$$

$$\text{Anti-cut background :} \quad A_{AC} + B_{AC} x + C_{AC} x^2 + D_{AC} x^3$$

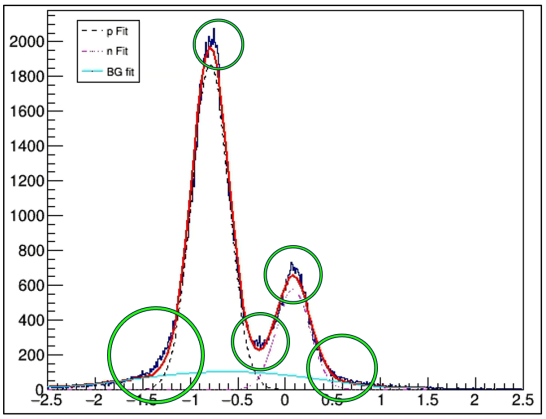
We can use the ratio of scalar coefficients, A_i/A_{AC}, to scale up the anti-cut background fit. We will take the value from this ratio and multiply each parameter of the anti-cut background fit by it[7]. Afterwards, it should be properly normalized.

$$\text{Scalar ratio :} \quad M = A_i/A_{AC}$$

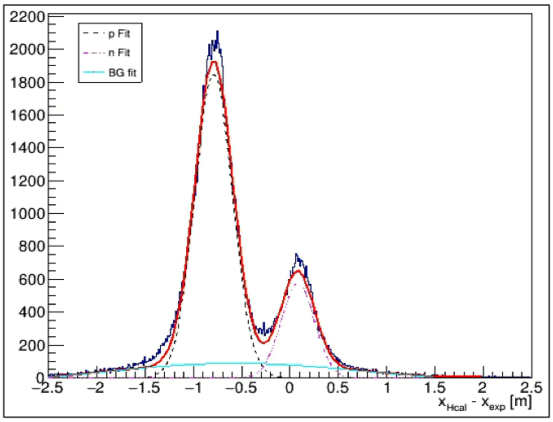
$$\text{Modified Anti-cut Background function :} \quad M(A_{AC} + B_{AC} x + C_{AC} x^2 + D_{ACi} x^3)$$

The final dx fit with a dy anti-cut background is shown in Figure 5-17a. The total fit, shown in red, appears to match the black experimental data points better than it did in the initial fit (Fig. 5-17b). The green circles in Figure 5-17a highlight the areas where we see marked improvement by using dy anti-cut projection background.

[7]These fits are ROOT TF1 functions. A TF1 function holds all of the coefficients of a fit function. These can be determined directly from fitting data, or be user-defined. The normalized background function here is simply a new TF1 function with coefficients defined by taking the original anti-cut background fit parameters and multiplying each of them by the ratio of scalar components.



(a) dy anti-cut projection background



(b) Initial total dx fit background

Figure 5-17: Comparison of total dx fits **(a)** with and **(b)** without background from a dy anti-cut projection

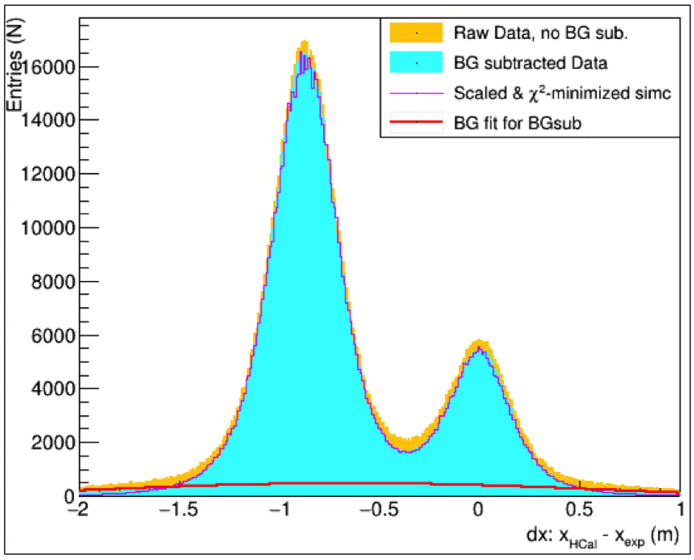


Figure 5-18: Comparison of data and simulation before and after background subtraction. The raw data without any background subtraction is in yellow. The dy anti-cut background fit is the red line. The data after background subtraction is cyan. The χ^2-minimized simulation peak is in magenta.

5) Subtract the Background: We now have a data-driven background function which we can use to create a new histogram. This new histogram can then be subtracted directly from our data dx histogram. Figures 5-18 and 5-19 show various comparisons between raw data and background-subtracted data. In Figure 5-18 we can see the raw data in yellow. The red line corresponds to the normalized background fit

created from the dy anti-cut projection. The cyan region is the background-subtracted experimental data. The magenta line corresponds to scaled and χ^2-minimized simulation data. We can see that it nicely outlines the cyan region.

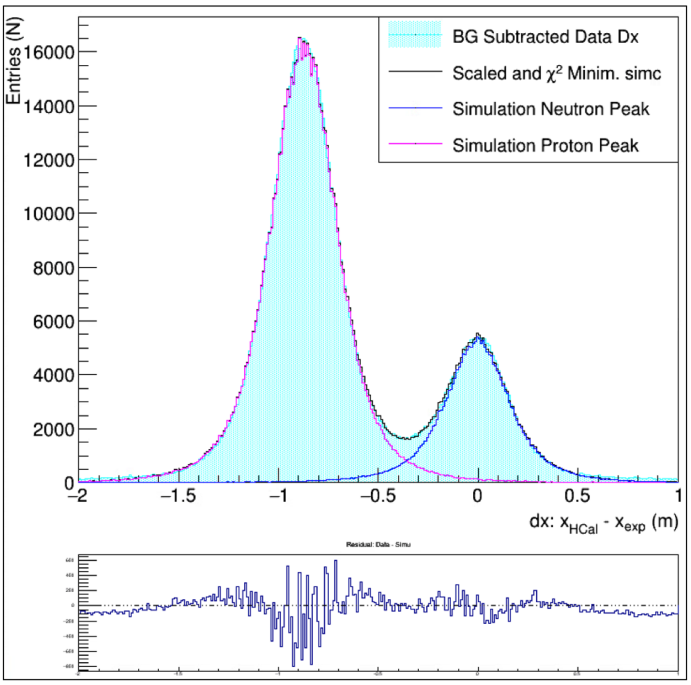


Figure 5-19: Comparisons of data and simulation after background subtraction. The background subtracted data is in cyan. The χ^2-minimized simulation peak is the black line that outlines the cyan region. A residual plot of data minus simulation is shown beneath the main plot.

To better visualize the agreement, it is easier to look at a plot of just the background-subtracted data and the tuned simulation. This is shown in Figure 5-19. The data with background subtraction is the cyan region. The black line along the cyan edge is the χ^2-minimized simulation data. Also shown are the simulated proton (magenta) and neutron (blue) peaks. Below this plot is a *residual plot* of the data minus the simulation. From this residual plot, we can see that the agreement between the two plots isn't exact. Most of this disagreement likely stems from any mis-calibration in the simulation magnetic field scales or statistical limitations. This will be discussed further in later sections.

5.4 Fitting Methods and Techniques

All of the fits discussed throughout this thesis are embedded in analysis scripts and run as part of the framework during the extraction process. These fits are all implemented using ROOT's built-in TF1 fit functions and/or they are derived from ROOT TMinuit minimization methods. When constructing these fit/minimization functions, one can define various initial parameter values, set variables as *fixed* or *flexible*, set thresholds, define end conditions, etc. Determining the setup of these functions can be a bit of an art, and can result in either a very restrictive or loose fit result, either of which may or may not be desirable. Certain fits can, and should, be held to tighter constraints to ensure that ROOT doesn't go off the rails during its optimization — if set conditions are too tight or too loose, ROOT can produce some very eccentric fits. This can be especially true when using polynomial functions of order N $>$ 2. Depending on the data being fit, ROOT may have a difficult time determining the correct sign for each of the coefficients, resulting in the *best fit* not actually being the best one.

These issues can be mitigated by either making the *correct* initial *guesses* for the fit functions or by implementing an adaptive and iterative approach when performing the fits. A mix of both were applied across my analysis efforts. The first approach is more appropriate for cases such as fitting the proton and neutron peaks. For each kinematic setting, we typically know within $\pm$5 cm where the center of each dx peak will be. From there, we can define a reasonable range for the width of each peak and allow the amplitude to vary freely. This provides great results, but it does rely on some *a priori* intuition or knowledge. When information about the peaks is not known beforehand, an iterative approach can be taken in order to determine reasonable initial fit values. This method involves recursively applying a fit, assessing the results, and adjusting the fit parameters until an acceptable fit agreement is found. This was the method used in finding the initial values for the dy anti-cut background. When possible, I will always try to implement an adaptive approach with as few constraints as possible.

5.5 HCal Detection Efficiency

HCal detection efficiency is a systematic uncertainty that does not cancel out in the Ratio Method. This extraction method relies on an accurate accounting of scattered neutrons and protons. Therefore, it is important that we understand the efficiencies at which we can measure the nucleons. My analysis will ultimately rely on the HCal detection efficiencies that are built into the Monte Carlo framework. However, in order to determine the efficacy of the MC-based efficiency values, I conducted a study comparing simulated and experimental HCal detection efficiencies. My study is constrained to a single kinematic point (SBS4) and is for protons only[8]. I must remind the reader that this analysis is purely exploratory in nature. There is still a vast amount of work that needs to be completed in order to reliably estimate the systematic errors associated the HCal detection efficiencies.

MC HCal Detection Efficiency

My analysis of the Monte Carlo HCal detection efficiency closely follows the procedure and practices laid out by J. Cornejo in [39]. In my study, I analyzed 80 nucleon momentum points ranging from approximately 1 to 9 (GeV/c). The nucleon momentum points for this thesis are 2.35 (SBS4), 3.22 (SBS8), and 3.21 (SBS9) (GeV/c). At each point, I store the total energy deposited in all of the scintillators of a 4x4 block of cells on the active face of HCal. Figure 5-20 shows the collected energy at four different momentum values. These momentum values are also called *p bins*, or *momentum bins*.

The plots in the top row of Fig. 5-20 each contain a shaded region near their left side. This indicates a region that is *below threshold*. In most of these plots, this shaded region also contains a *spike* in the data — this spike is only visible in the first, third, and fourth plots. This signal is not correlated with hadron energy deposition and therefore needs to be removed [120]. For each momentum bin, we can isolate this

[8] I was not able to conduct a study on experimental HCal detection efficiency of the neutron. Therefore, for the neutron, I am left with assuming that all of the detection efficiencies built into g4sbs are accurate and correct. There is no reason for me to doubt them, and I do not.

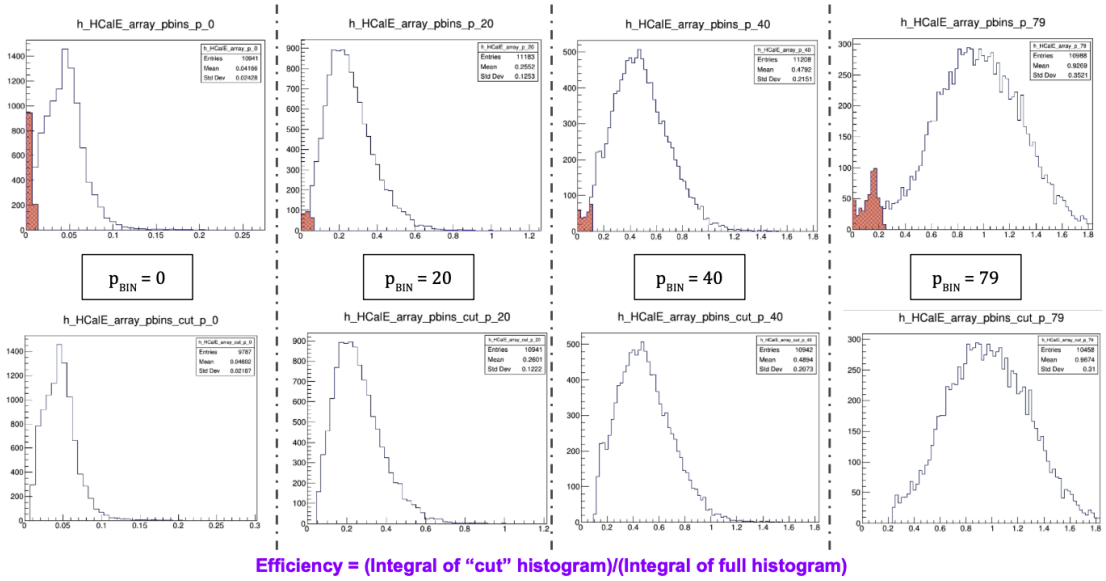


Figure 5-20: Four MC HCal *momentum bin* selections. In each plot on the top row, the energy below threshold is shown by the shaded region. The plots on the bottom row have that region removed. The efficiency is the ratio of these plots' integrals.

region given that it is always *"1/4 the energy of the peak of the energy distribution"* [39]. With that, we define a threshold energy $E_{thresh} = (1/4)E_{Peak}$, where E_{Peak} corresponds to the energy bin with the most entries. The hadron detection efficiency is then determined by dividing the total number of hadrons with energy $E_h > E_{thresh}$, by the total number of events at HCal.

The plots in the lower row of Fig. 5-20 show the histograms after the energies below threshold have been removed. We refer to these histograms as the *cut histograms*. The original, un-cut, histogram is simply referred to as the *full histogram*. Therefore, for each momentum bin, we can calculate the detection efficiency as:

$$\text{Monte Carlo HCal Detection Efficiency} = \frac{\text{Integral of } \textit{cut} \text{ histogram}}{\text{Integral of } \textit{full} \text{ histogram}}$$

This is calculated across all 80 momentum steps for the neutron and proton separately. The plot of the calculated MC HCal detection efficiencies is shown in Fig. 5-21. The proton efficiency curve is shown in red and the neutron in blue. The calculations for efficiency of each bin are plotted with a 3rd order polynomial fit across each

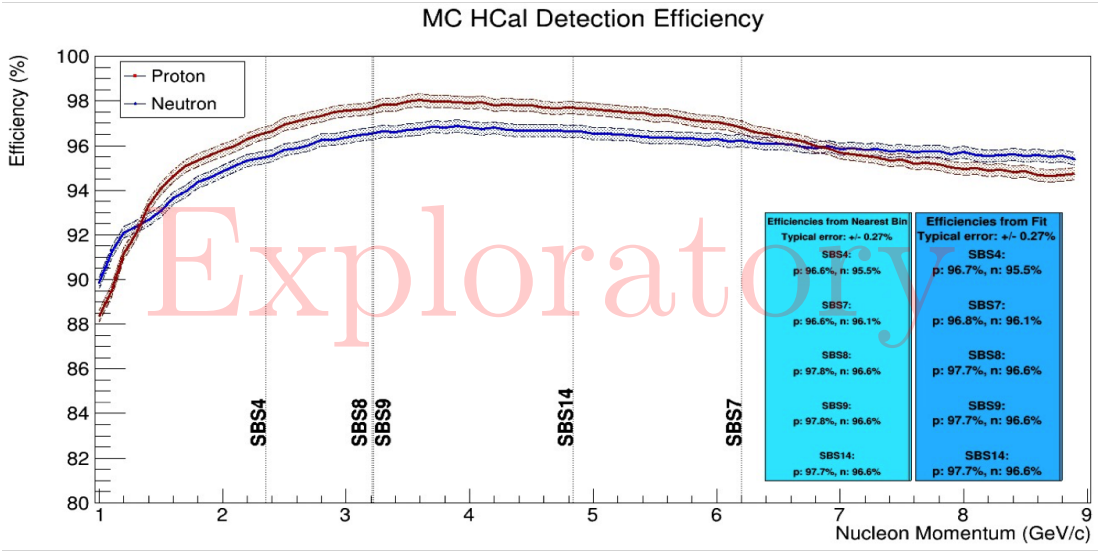


Figure 5-21: **[Exploratory]** Monte Carlo HCal proton and neutron detection efficiencies. Momentums corresponding to kinematic points are shown. A legend is provided which shows calculated efficiency values based on either an evaluation of the fit function or the value from the closest histogram bin.

data set[9]. The plot insets provide efficiency values at various kinematic points. The inset labelled "Efficiencies from Nearest Bin" contains values taken directly from the nearest histogram bin to the corresponding momentum point. The inset labelled "Efficiencies from Fit" contains values from evaluating each polynomial fit function at the given momentum value. The uncertainties are statistical only, and populations for MC simulation data sets were selected to be on-scale with those used in data-simc-comparison studies.

Remarks on MC HCal Detection Efficiency Plot

There is one region of Figure 5-21 that requires further explanation. Near the Nucleon Momentum point of 7 GeV/c we see an unexpected characteristic. The two efficiency curves meet at a *cross-over point*. After its initial rise, the proton peak begins to decrease near 3.6 GeV/c. From this point forward, it continues to decrease. The neutron curve also follows a similar trend. However, its rate of decrease is

[9]There is no *a priori* physics-motivated reason for the use of a 3rd order polynomial here. This was the lowest-order polynomial that provided reasonable agreement.

markedly less than that of the proton curve. Hence, their intersection near 7 GeV/c. This is an unexpected finding that has not been fully-explained and is still being studied. At this point I can only speculate possible causes for this and even then, I hesitate to proceed too far. Therefore, I encourage the reader to stay tuned to the JLab community for more updates.

HCal Detection Efficiency from Data

A data-based study of HCal detection efficiency was conducted using Pass 0 data. I make to attempts to correlate the data with the hardware settings during experimental running and data-taking. Additionally, I do not make any comparisons between the hardware and software triggers for either calorimeter. I am also not considering characteristics such as how well-balanced the gain and light collection is for each calorimeter signal. These are all considerations that will need to be included in a full-scale analysis of HCal detection efficiency. This study simply compares dx from HCal with W^2 from BBCal. The nature of this thesis is entirely exploratory, but the use of Pass 0 data with the added limited considerations, makes this particular study *pre-exploratory*. Therefore, the values presented here should be carefully taken with a grain of salt.

In order to calculate the hadron detection efficiency from data, we use a slightly different approach from what we used for the Monte Carlo study. In comparison to MC, we have a limited set of experimental data — there are six SBS kinematic points, seven if you include SBS8 — which exist at varying levels of calibration and background contamination. Therefore, with a data-based HCal detection efficiency study, we won't have the luxury of being able to analyze 80 data points like we did with the Monte Carlo study. For my study, I focused on the Pass 0 SBS4 LH2 dataset with no SBS magnetic field. Since we are using a hydrogen target, there is no need to distinguish protons from neutrons. Its all protons. Therefore, we do not necessarily need to consider any data with the SBS magnet turned on. Our largest dataset for SBS4 with LH2 is for the SBS dipole magnet at 0%. Therefore, to maximize my statistics, I have opted to use this particular kinematic configuration.

The data-based HCal nucleon detection efficiency is calculated as:

$$\text{Experimental HCal Detection Efficiency} = \frac{\text{Detected number of nucleons}}{\text{Expected number of nucleons}}$$

An additional limitation for determining the experimental HCal detection efficiency is that we don't have a means to analyze neutrons on their own; there are no free neutrons. Therefore, this study is limited to determining the *proton detection efficiency*. This is precisely why I am only using LH2 data. In this study I will compare the yields of protons from HCal measurements (detected number of protons) to those from BBCal W^2 measurements (expected number of protons). In other words:

$$\text{Experimental HCal Proton Detection Efficiency} = \frac{\text{Detected number of protons}}{\text{Expected number of protons}}$$

$$= \frac{\text{Yield from LH2 HCal dx}}{\text{Yield from LH2 BBCal } W^2}$$

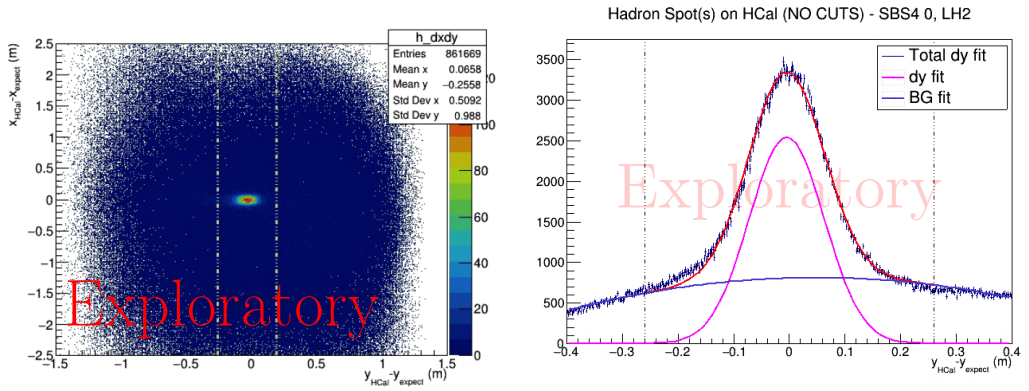
In this particular study, we need to make sure that the values we extract from HCal are not biased by any BBCal-dependent cuts or selections, and vice versa. For instance, if we are looking for proton counts on HCal, we need to be sure that we *turn off* all BBCal cuts. This includes the cuts on Shower and PreShower, for example. Similarly, when looking at data from the BigBite arm, we need to make sure that we don't include basic HCal cuts that are typically applied as a default. This includes a cut like the *minimum HCal cluster energy*. If we don't remove this cut, then our events on BBCal will be correlated with, and biased by events on HCal. We are seeking data that is isolated to each scattering arm without any input or bias from the other.

For the proton count on HCal, I am using the integral of the dx histogram. Recall that we are using an LH2 target. Therefore, this histogram is simply the counts of protons which land on HCal. For this dx histogram, I will apply an HCal acceptance

cut (the particle must land on HCal) but not the fiducial cut — the fiducial cut relies on expected nucleon hit positions based on BigBite measurements. We will also want to apply the best cluster selection method for this variable as well. This selection only depends on HCal cluster energy and HCal coincidence timing.

To determine the proton counts on BigBite, I will access the invariant mass, W^2, histogram. The invariant mass is a good candidate for determining protons at BigBite because W^2 is calculated purely from BigBite variables and elastic events are correlated with the invariant mass around $W^2 \sim M_p^2$, where M_p is the mass of the proton. When determining W^2, we still need to apply a basic HCal acceptance cut for projected proton hits on HCal. We need to ensure that the invariant mass events on BigBite are associated with scattered protons which would have actually landed on HCal. We are trying to determine whether or not protons are detected after they have arrived at HCal. For us to detect protons, they need to be able to hit the active area of HCal. Therefore, I require the expected proton hits on HCal to land within the physical acceptance area of the detector. Any discrepancies that I find between the number of entries in the dx and W^2 histograms is assumed to derive from the proton detection inefficiencies at HCal.

Figure 5-22 shows the four principle histograms used in this analysis. The starting point is the HCal dx measurement. Figure 5-22a shows a raw dxdy plot for LH2 data from the SBS4 kinematic. Our data-under-scrutiny is that little dot in the center. In order to make future plot-fitting easier, we can apply a very wide dy cut to this data. The two vertical lines in Fig. 5-22a mark possible positions for applying cuts which reject everything outside of these boundaries. These positions would be determined form the X-projection of this dxdy plot (the dy plot shown in Fig. 5-22b). A fit to dy provides us with a mean and sigma which we can use in defining our boundaries. Once the dy cut is applied, we can then extract our dx plot from the Y-projection of dxdy. This is shown in Figure 5-22c. The goal here is to extract the yield of the signal without the background. Therefore, I apply a fit to extract the shape of the background function, shown in blue. Once I have this background function, I can then subtract it from the raw data. This leaves me with the dx signal, shown in pink.



(a) dxdy plot, dy cut positions marked **(b)** dy plot for determining cut positions

(c) dx plot after a wide dy cut **(d)** W^2 with signal and BG fits

Figure 5-22: **[Exploratory]** The various plots used to calculate the experimental HCal proton detection efficiency for SBS4.

The integral of this pink dx peak provides my *detected number of protons on HCal* value.

Figure 5-22d shows the invariant mass plot for the same dataset. SBS4 was chosen because it is a relatively *clean* dataset in terms of background contamination. This is why there is not much background present in Fig. 5-22d. This makes fitting the signal and background a bit easier than it would be for a noisier kinematic setting. My goal is to extract the yield of the W^2 signal without the background. I follow a similar process as was just described for dx. In Figure 5-22d, we see a background fit in blue. This function is subtracted from the raw data. This leaves us with the isolated W^2 signal, shown in pink. The integral of this pink W^2 peak gives us the *expected number of protons*. We now have everything we need to calculate our experimental

HCal proton detection efficiency for SBS4. The extracted values resulted in a proton detection efficiency of 94.65% ± 0.41%.

Comparison of MC and Data HCal Detection Efficiencies

The proton detection efficiency from Monte Carlo was calculated to be 96.5% ± 0.27% whereas the experimental data provided 94.65% ± 0.41%. These are not within bounds of each other. This ultimately needs to be reconciled as best as possible. That is beyond the scope of this work, but will surely be addressed in the future. A major point to consider for my study is that it was performed with Pass 0 data — this is essentially a *pre-exploratory* study. It is possible that performing the exact same analysis with Pass 2 data may produce better results that are in closer agreement. In any case, this topic is sure to be a primary focus and a point of rigor for future analyses.

5.6 HCal Inefficiencies and Data Issues

5.6.1 Data Quality for SBS9

At this point in time, my estimates of G_M^n for SBS4, SBS8, and SBS9 are within reasonable expectations of predicted values. There are however, some issues which arose during data-taking for SBS9 which deserve a deeper discussion. The issues and topics discussed within this section are still under intense scrutiny. The methods I present here are my initial studies on the subject and are entirely pre-exploratory in nature. This is an extremely important issue that certainly deserves heavier focus in future analyses.

The discrepancies with the current SBS9 dataset are apparent when comparing the scale factor and yield ratios from SBS9 with those of SBS4 and SBS8. Also, extractions of G_M^n from an un-corrected SBS9 dataset do not fall within bounds of global parameterizations, or the value extracted from SBS8. Recall that G_M^n is a function of Q^2 only. The Q^2 values for SBS8 and SBS9 are nominally the same, 4.5

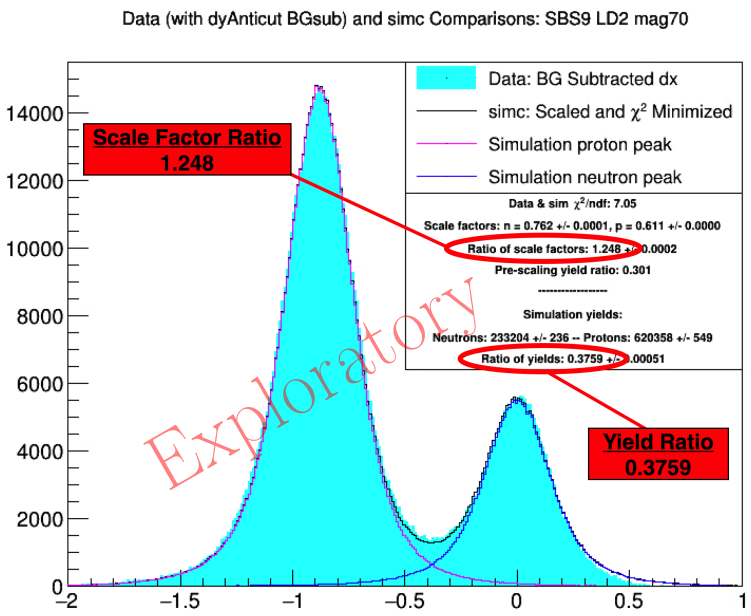


Figure 5-23: [Exploratory] SBS9 dx plot highlighting the effects of HCal inefficiencies on the data and resulting extractions. In this plot we see a Scale Factor Ratio, $F_{n/p}$, of 1.248 and a Yield Ratio, $R_{n/p}$, of 0.3759. These values should *theoretically* be closer in value to 1.0 and 0.35, respectively.

$(\mathrm{GeV/c})^2$. Therefore, their G_M^n values should also be equivalent. The experimental Q^2 values for SBS8 and SBS9 are of course not exactly equal to each other, nor are they exactly the same as the nominal value of 4.5: $Q^2_{SBS8} = 4.3686\ (\mathrm{GeV/c})^2$ and $Q^2_{SBS9} = 4.3927\ (\mathrm{GeV/c})^2$. Regardless of the slight disagreement between them, their Q^2 values are close enough that we should expect their G_M^n values to also be reasonably similar. This would be true if everything was based purely on the momentum-transfer values alone. Unfortunately, this is not the case due to some external factors.

The underlying issue with the SBS9 dataset has been traced back to performance issues with the Hadron Calorimeter during the running of SBS9 (and SBS8 to some extent) [104, 118]. Several of the blocks in HCal had malfunctioned and either stopped working altogether or, operated at abnormally low efficiency levels. In either case, this resulted in poor data quality for those regions of HCal. Unfortunately, these regions corresponded to a primary proton acceptance area on the face of HCal. The result is a skew in the population of measured protons. I hypothesize that we should have been able to detect more protons in SBS9, but were *losing* many of those events.

Figure 5-23 shows an SBS9 dx plot exhibiting this characteristic. We can see that the ratio of yields reported for that plot's analysis is 0.3759. The ratio of yields is the count of neutrons divided by the count of protons. We would nominally expect a value less than this[10].

The yield ratio alone is insufficient in proving that we are in fact measuring less protons than we expect, as opposed to say, measuring more neutrons. Figure 5-23 also highlights a scale factor ratio, $F_{n/p}$, of 1.248 extracted in that analysis. From this plot, we can also see the individual scale factors for each nucleon: $f_{(\text{scale, n})} = 0.762$ and $f_{(\text{scale, p})} = 0.611$. As a reminder, $F_{n/p}$, is the ratio of scale factors used in cross-matching the nucleon peaks between experimental and simulated data. This is an indicator of the relative mis-match between the simulation-based calculations for nucleon yields and what we observe in data. The world data set for the protons is better supported than that of the neutron. Thus, we take it with a bit more certainty (see Sections 4.5 and 4.4.4). For the individual scale factors and subsequent scale factor ratio just discussed, this implies that the neutron peaks required much *less* scaling than the proton peak. Therefore, in order to match the data, the simulated proton peak needed to be scaled down *more* than the neutron peak. Typically, we would expect the neutron peak to be the one that requires the majority of the scaling.

When the scale factors are less than 1.0, the simulated data is being scaled down[11]. The smaller the value, the more the simulation needs to be scaled down. This implies that the experimental proton data here is in disagreement with the simulated proton calculations. This opposes our fundamental declaration that we hold the simulated proton calculations to be well-established and *true*. This indicates to me that there is a deficiency in the proton measurements for this dataset.

Figure 5-24 shows a plot of proton detection efficiencies for each block of HCal. The central region of the plot should ideally not include the darker regions that are present near the coordinate position of (3, 12). In this region of the plot, we can

[10]The yield ratio can be approximated given your choice of parameterizations, a value of Q^2, and some algebra. For the kinematic explored in this analysis, we typically expect the yield ratios to be approximately $1/3$.

[11]Conversely, if the scale factors are greater than 1.0, the simulated peaks are being scaled up.

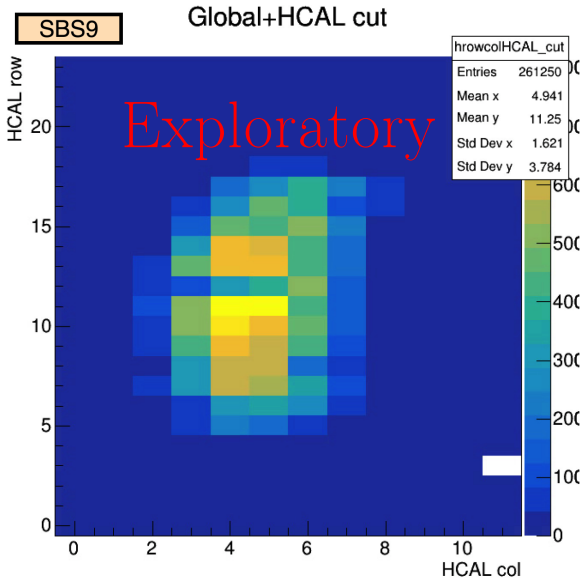


Figure 5-24: **[Exploratory]** HCal inefficiencies for SBS9. The darker regions in the central area of the plot correspond to banks of malfunctioning HCal blocks. Plot source: [104].

identify a single, low-efficiency *stripe* just above the bright yellow portion near the center of the figure. This is our culprit. We can correlate this low-efficiency region with a separate performance scan of the blocks in HCal (Figure 5-25). In this figure, red blocks indicate malfunctioning cells of HCal. We see a similar striping in Fig. 5-25 as we do in Fig. 5-24. Further studies have shown that these regions do in fact coincide with one another [104, 118].

As it turns out, this is a very unfortunate region of HCal for this problem to have occurred. As compared to SBS8, the envelope for elastic protons in SBS9 is quite small. The central region of this envelope just happens to overlap with this lower-performing region of HCal. At this location, we would expect the proton cross-section to be a maximum [104]. Any *dead* or low-efficiency cells in this region of HCal would result in a considerable loss of proton data.

5.6.2 Corrections for SBS9 Data

There are multiple ways in which we can correct for the low efficiency HCal blocks of SBS9. The simulation in its normal state does not account for any low-efficiency

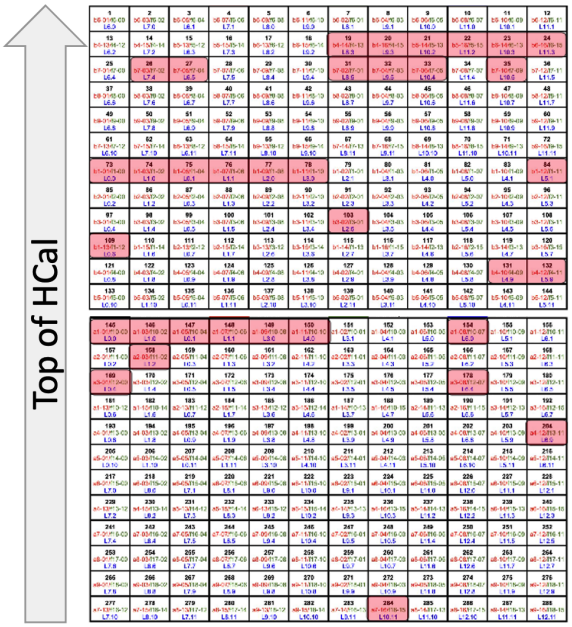


Figure 5-25: [Exploratory] Map of HCal blocks and their operational status. The red blocks indicate malfunctioning cells. Image source: [118].

blocks. Therefore, we need to modify our simulated setup so that it also contains the same HCal block efficiency distribution that we see in the experimental data.

A possible reason for the low efficiency issue is background contamination due to poor sealing and *light-leakage* into the HCal block PMTs. This results in a large amount of background and subsequent signal saturation [118]. One approach would be to modify the experimental setup in the simulation environment. It could be possible to replicate the physical gaps between the PMTs and blocks, and replicate our real-world problems. This would be an exhaustive process that may never be able to achieve a suitable outcome. A more realistic approach would be to apply a correction factor to the simulated block's data stream in an attempt to account for the contamination. This is a feasible solution which could be developed in such a way that it allows for continual calibration and updates as the physical setups and data structures evolve. I am not taking either of these approaches in my analysis since they both require timelines which are beyond the scope of this thesis.

The correction that I have implemented involves making an *HCal efficiency map* similar to that shown in Figure 5-26. I can use this efficiency map as a source for

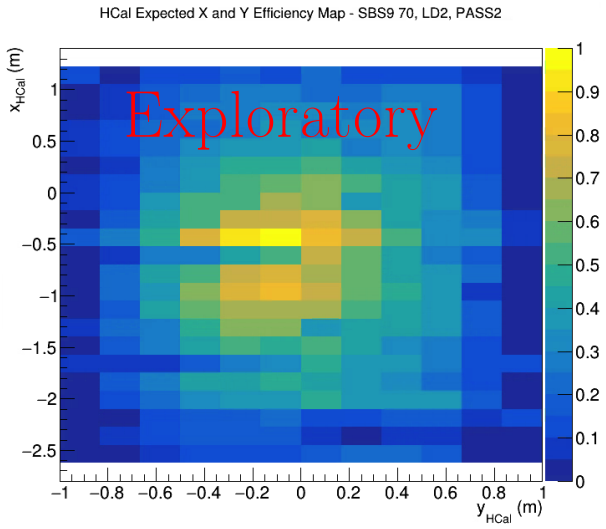
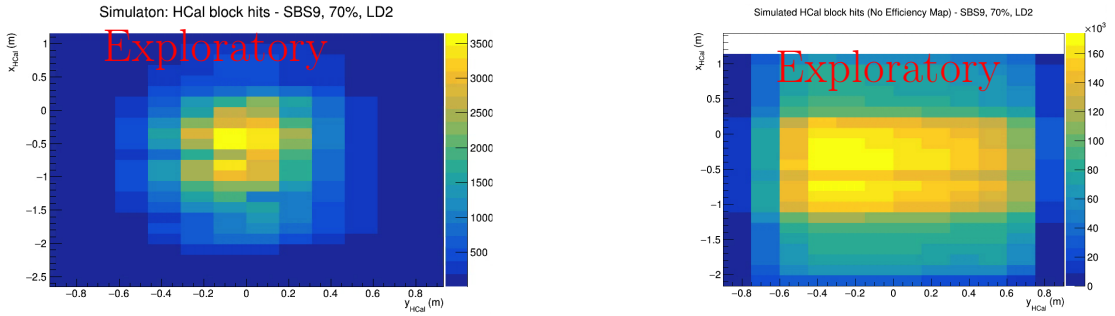


Figure 5-26: **[Exploratory]** SBS9 data-based HCal expected x and y efficiency map. This plot is made from experimental data and is used to weight events in the simulation.

weight coefficients that can be multiplied into the simulation's Final Weight value (see Sec. 4.4.3 and Eq. 4.1). The basic principle involves creating an 2D histogram of expected x and y hit positions on HCal. This histogram is defined so that each histogram bin corresponds to a single block in HCal. The histogram also has the same *BCS dimensions* as HCal in both x and y. The histogram is normalized such that the maximum value is 1.0.

For each event in the simulated data, we find the bin in our HCal efficiency map which corresponds to the x and y hit positions on HCal. The value of this bin will range between 0.0 and 1.0, and is an efficiency value associated with the corresponding block in HCal. This efficiency value is then multiplied into the simulation Final Weight (Eq. 4.1). Events corresponding to hits on low efficiency blocks will be weighted down accordingly. This is a pretty simple implementation, but proves quite effective in simulating the HCal efficiency distribution we see in SBS9 data. Figure 5-27a shows the simulated distribution of hits across HCal blocks after implementing this correction method. We can see that we have effectively reproduced the efficiency distribution that we find in data (Fig. 5-24). Figure 5-27b shows the un-corrected simulated HCal hit map without implementing the efficiency map. We can verify this



(a) Using HCal efficiency map

(b) Without HCal efficiency map

Figure 5-27: **[Exploratory]** SBS9 simulated HCal hits with and without data-based HCal efficiency map. **(a)** SBS9 simulated HCal hits using data-based HCal efficiency map. The simulation pulls efficiency values from the HCal efficiency map. This efficiency is used to weight events based on which HCal block they are interacting with. **(b)** Standard hit map without the HCal efficiency map correction applied.

method further by performing an analysis and extraction of SBS9 and comparing the results.

Figure 5-28 shows a dx plot for SBS9 after implementing an HCal efficiency map. We can see that our Scale Factor Ratio, $F_{n/p}$, is now 0.945. Recall that without the efficiency map $F_{n/p}$ was 1.248. In general, we are expecting the Scale Factor Ratio to be a value close to 1.0. Our new Scale Factor Ratio of 0.945 is therefore better aligned with our expectations and indicates that the simulated cross-sections are in better agreement with the experimental data.

We can also assess the effectiveness of this method by looking at the *pre-scaling yield ratio*. The pre-scaling yield ratio is the raw n:p yield ratio as it is generated in the simulation, without any scaling applied. If the simulation is well-modelled to the data, this value should be close to the n:p yield ratio we find experimentally. Due to the inefficiency-induced *loss* of protons, the SBS9 experimental data has a larger than expected n:p yield ratio of approximately 0.360. Before applying the efficiency map, the simulation had a pre-scaling yield ratio of 0.301. After implementing the correction, this value goes up to 0.384, which is much closer to our experimental value of 0.360.

The efficiency map applied in this analysis can be described as a *relative efficiency map*. In this type of map, the values are based on the ratio between the efficiency

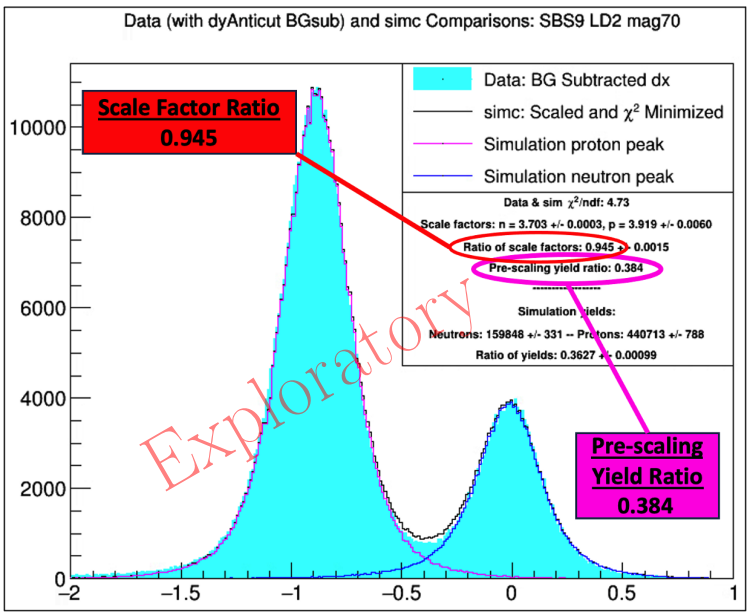


Figure 5-28: [Exploratory] SBS9 dx plot after applying HCal efficiency map corrections. The Pre-scaling yield ratio is now closer to the value that we see in data. This reflects the relative proton detection inefficiencies and *loss* of proton counts.

at a given point and the maximum efficiency value found within the useful range of expected x and y hit coordinates [60]. This particular type of map was used because it may provide better compensation against the built-in HCal detection efficiencies of the MC simulation. Further testing and additional studies will be needed in order to fully examine the characteristics and efficacy of these claims though.

In addition to the implementation of an efficiency map, applying a large cut on the HCal minimum cluster energy also provided decent corrections. The typical cut on this variable is a rejection of everything below 5 MeV. If we increase this threshold to 300 MeV, it also provides us with some level of correction. The affected HCal blocks have elevated amounts of background, most likely due to loose connections and/or light leaks. Raising the HCal minimum cluster energy effectively cuts out a lot of this background [122]. This method was tested and does provide comparable results to the efficiency map method. However, it results in a large loss of events and therefore, statistics. Luckily, SBS9 has the most raw events of the kinematics considered in this analysis (SBS4, SBS8, and SBS9). Even after I increase the threshold on HCal minimum cluster energy, the remaining SBS9 dataset still has more events than either

SBS4 or SBS8. Nevertheless, I will always prefer to maintain as many statistics as possible. Therefore, I do not apply this technique in my analysis.

The corrections applied here appear effective at simulating the effects of inefficient blocks across HCal during the SBS9 kinematic. Everything discussed here was only implemented in an *ad hoc* manner and from an exploratory standpoint. Additionally, any corrections implemented here are not considered in any of my systematic uncertainty analyses.

Chapter 6

Exploratory Experimental Results

The concluding chapters of this thesis will present and discuss projected values of G_M^n and the discrepancies between Rosenbluth Separation and single-photon exchange measurements. The full scope of the analyses presented here spans Pass 0, Pass 1, and Pass 2 data. My concluding analysis was conducted using only Pass 2 data. Everything presented in this chapter should be considered purely exploratory in nature.

6.1 Exploratory GMn Projections

The goal of this analysis is to extract the neutron form factor, G_M^n, at multiple Q^2 points. The GMn experimental kinematic points are SBS4 ($Q^2 = 3.0$ (GeV/c)^2) and SBS9 ($Q^2 = 4.5$ (GeV/c)^2). We also pick up an additional measurement point when we include the overlapping nTPE SBS8 kinematic ($Q^2 = 4.5$ (GeV/c)^2). This section will discuss the realization of the methods presented in Section 4.5.

6.1.1 GMn Observables

In order to extract G_M^n via the ratio method, we are looking for the ratio of quasi-elastic D(e,e′)n and D(e,e′)p yields. In this analysis, the primary extraction variables are the data-simulation scale factors for the neutron and proton ($f_{(\text{scale, n})}$

and $f_{(\text{scale, p})}$), and the yields of each nucleon (Sec. 4.4.4). Encoded within the scale factors are the built-in parameterization functions used to compute the aforementioned quasi-elastic yields. These scale factors effectively correct for differences between the simulation's built-in parameterizations and the experimental data. The nucleon yields are extracted by taking the integrals of the scaled and data-matched simulated proton and neutron histograms.

6.1.2 Exploratory GMn Results At A Glance

	SBS4	**SBS8**	**SBS9**
$\boldsymbol{Q^2}$ **(GeV/c)^2**	2.937	4.368	4.393
Scale Factor Ratio, $\boldsymbol{F_{n/p}}$	0.9353	0.9823	1.013
$G_M^n/(\mu_n G_D)$	0.9785	0.9323	0.9183

Table 6.1: **[Exploratory]** Summary of Q^2, $F_{n/p}$, and G_M^n for SBS4, SBS8, and SBS9.

Table 6.1 is a summary of my experimental findings for SBS4, SBS8, and SBS9. These values are extracted from the *exploratory extraction dx plots* shown in Figure 6-1. These represent the final stages from which my extraction of G_M^n is performed. In these plots, the experimental data has been processed through all elastic selections, and the simulation has been calibrated to match. The simulation files used in these comparisons are initialized using the exact same elastic and acceptance cuts as the experimental data. Furthermore, our Monte Carlo framework is designed to include features such as radiative effects, Fermi motion, and HCal nucleon detection efficiencies into the simulated data. The plots of Fig. 6-1 contain simulation data which has been scaled and χ^2-minimized against experimental data. It is through the data-simulation tuning process that I described earlier, where I retrieve the values of $f_{(\text{scale, n})}$ and $f_{(\text{scale, p})}$.

The lower portions of the plots in Figure 6-1 are residual plots between the data

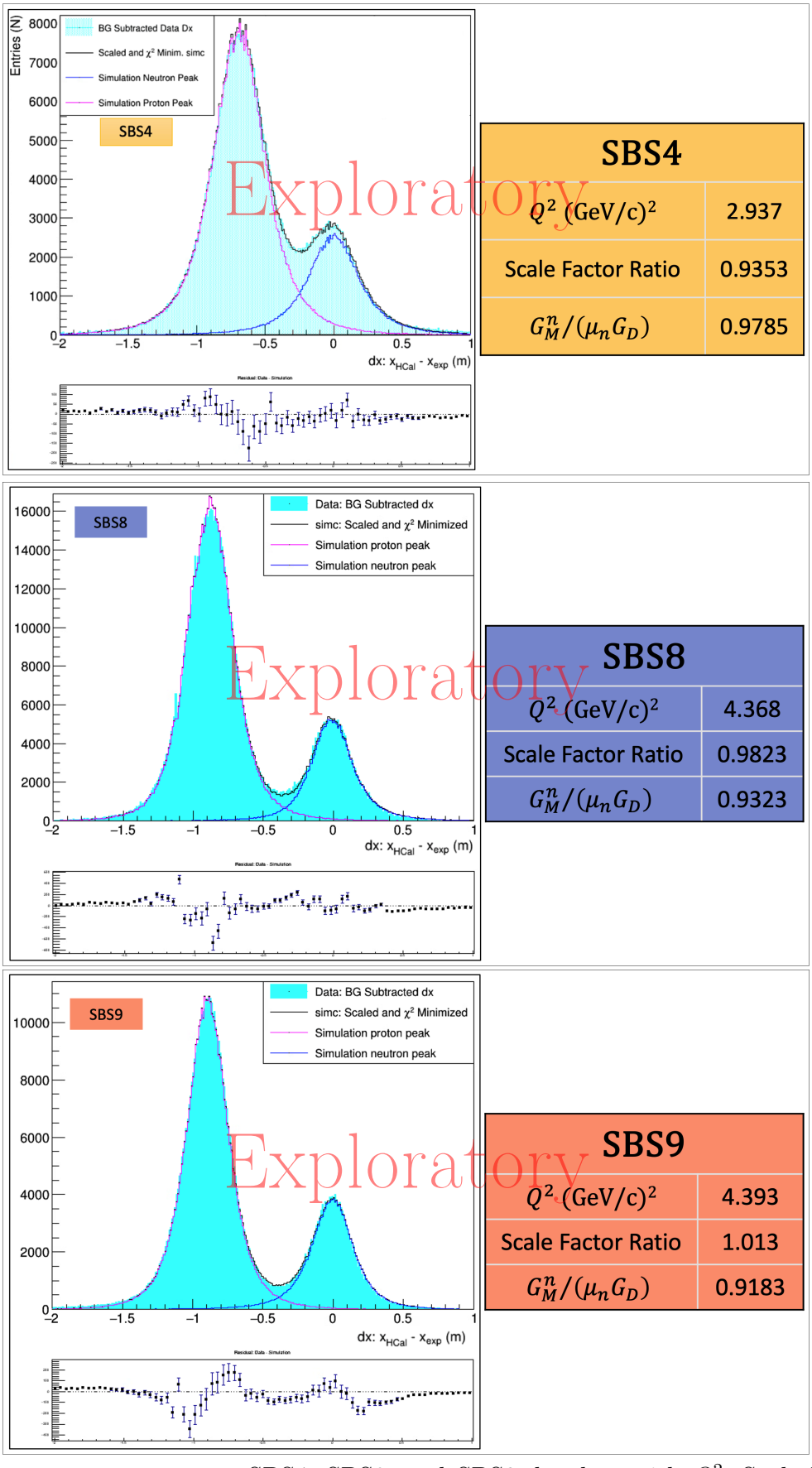


Figure 6-1: [Exploratory] SBS4, SBS8, and SBS9 dx plots with Q^2, Scale Factor Ratio, and G_M^n. These plots include all cuts, background subtraction, and MC scaling. These are the final plots of my analysis and extraction process.

and simulation, i.e., data minus simulation. These are good indicators of how well the simulation matches (or doesn't match) the data. In a perfect match-up, we would expect a residual plot with only random fluctuations consistent with statistical error around the horizontal axis. However, we can see areas which indicate a mis-match between the simulation and experimental data. The large residuals under the center of each nucleon peak correlate to either statistical limitations or differences in scaling and normalization[1]. The asymmetries along the residual plot highlight any mis-alignment between the peak centers caused by the field scale of the SBS dipole not being exactly dialed in when configuring and calibrating the simulation[2]. In future and further-developed analyses, the tuning between the two data sets would need to be properly established and critiqued.

6.1.3 A Note on World Data Parameterizations

In accordance with Sections 4.5 and 4.7, the extracted values of G_M^n rely on calculated values of the Sach's electromagnetic form factors from world data parameterizations. Our simulation is built with the Kelly parameterization for G_E^p, G_M^p, and G_M^n ([75]), and Riordan's function for G_E^n ([56]). When performing our extraction of G_M^n using Eq. 4.15, there is no requirement to use the exact parameterizations that are built into our simulation. Generally speaking, this formulation is model-independent; however, in this analysis, we need to be very careful about when and where we apply certain parameterizations.

Each parameterization is developed with a certain set of criteria, assumptions, choice of data, etc. These considerations need to be carefully assessed when determining which parameterization to use. For instance, the Kelly parameterization ([75]) prioritizes polarization transfer measurements over others. Therefore, it is in a

[1]The statistical limitation here is primarily driven by experimental data. In the simulation, we potentially have infinite statistics. However, we still need to fit those infinite statistics to experimental datasets with finite and limited statistics. Therein lies our limitation.

[2]The simulation is configured with an adjustable parameter for the field strength of the SBS magnet. In my analysis, I determine this value empirically by making small adjustments and then comparing the resulting plots/fits. Therefore, I was somewhat limited in how well I was able to properly match the simulated and experimental SBS magnetic field strengths.

sense a *single-photon parameterization*. On the other hand, the Ye parameterization ([140]) doesn't include such a bias, but it does incorporate a correction for two-photon exchange effects. These differences may seem subtle, but can have profound implications. This will be a central point for the nTPE extraction since we will need to compare our Rosenbluth Separation results against parameterizations which derive strictly from single-photon measurements.

6.2 Exploratory nTPE Projections

6.2.1 nTPE Observables

The method for determining the differences between Rosenbluth Separation and single-photon exchange measurements follows directly from the discussion of Chapter 4. This analysis will rely on values determined during the GMn analysis and extraction. These are the neutron and proton scale factors, the nucleon yields, and the two form factors $G^n_{M,\,\mathrm{SBS8}}$ and $G^n_{M,\,\mathrm{SBS9}}$. Recall that SBS8 corresponds to ϵ_1 and SBS9 with ϵ_2. My calculation will also require parameterizations for G^p_E, G^p_M, $G^n_{E,\,\mathrm{TPE}}$, and $(G^n_E/G^n_M)_{\mathrm{OPE}}$[3].

6.2.2 Exploratory nTPE Results At A Glance

Our goal is to re-create the plot of Figure 4-15 with values derived from experimental data. We can directly form the reduced cross-section for each epsilon point using $\sigma_{R,n} = \tau(G^n_M)^2 + \epsilon(G^n_E)^2$. For this analysis, we will use the nominal values of tau and epsilon as calculated using the data-based values for Q^2 at each kinematic point (see Table 6.2).

Figure 6-2 is a plot of our experimental neutron reduced cross-section vs. epsilon. Following the discussions that prelude and accompany Fig. 4-15, this plot includes a linear extrapolation between the two data points. The y-intercept of this line gives

[3]Here we just need the ratio of (G^n_E/G^n_M) in order to calculate the Rosenbluth Slope. We can use a parameterization for the form factor ratio, or calculate the ratio using individual parameterizations for G^n_E and G^n_M.

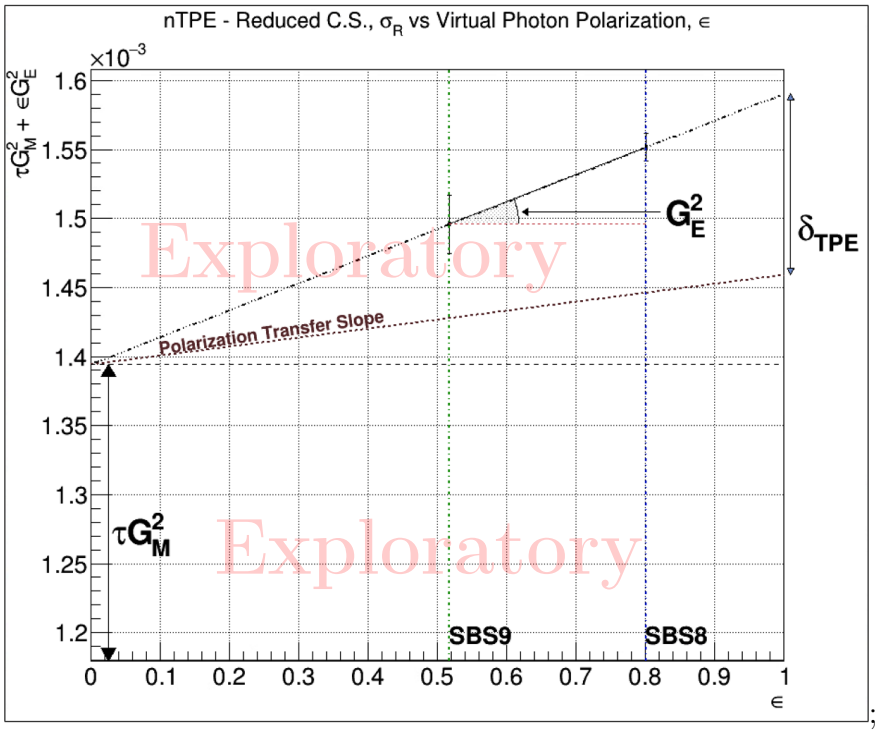


Figure 6-2: [Exploratory] Experimental neutron Reduced C.S. vs. epsilon for nTPE. *Polarization Transfer Slope* calculated using Kelly parameterizations [75].

	SBS4	SBS8	SBS9
Q^2	2.937 $(\text{GeV/c})^2$	4.368 $(\text{GeV/c})^2$	4.392 $(\text{GeV/c})^2$
τ	0.834	1.237	1.244
ϵ	—	0.8012	0.5176
		$\bar{\tau} \approx 1.2405$	

Table 6.2: Experimental nTPE values for Q^2, τ, and ϵ. The value of tau used in the concluding exploratory extraction is the average value of tau from SBS8 and SBS9.

us $\tau(G_M^n)^2$ and the slope provides $(G_E^n)^2$. We can then find the Rosenbluth Slope directly as: RS $= S^n = {}^{(G_E^n)^2}/_{\tau(G_M^n)^2}$. These values are summarized in Table 6.3.

Experimental Values		
	$\sigma_{R,n}$	$\sigma_{R,n}$ Error
SBS8	1.555e-3	9.9973e-6
SBS9	1.496e-3	2.1224e-5

y-intercept, $\tau(G_M^n)^2$	y-intercept Error	Slope, $(G_E^n)^2$	Slope Error
1.3944e-3	3.3627e-5	1.9603e-4	4.9854e-6

Calculated Values - Kelly Parameterization	
Normalized Rosenbluth Slope, RS $= S^n = (G_E^n)^2/\tau(G_M^n)^2$	Normalized Rosenbluth Slope Error
0.1406	0.0049

Table 6.3: [Exploratory] Normalized Rosenbluth Slope values extracted from Figure 6-2. The following relation is used for the Rosenbluth Slope, RS $= S^n = (G_E^n)^2/\tau(G_M^n)^2$.

We now need to compare our experimental Rosenbluth Slope values with those from single-photon measurements. Figure 6-2 includes a dashed line labelled "Polarization Transfer Slope". This line corresponds to a Rosenbluth Slope calculated using the Kelly parameterization. Recall that the Kelly parameterization prioritizes polarization transfer methods over others, and that polarization transfer measurements are mostly insensitive to two-photon exchange effects. Therefore, we can essentially consider Kelly to be a single-photon exchange parameterization [75]. We make our comparison using the relation

$$nTPE = S^n - S^n_{OPE} \tag{6.1}$$

with the following values:

Experimental S^n with statistical error only: 0.1406 ± 0.0049 (6.2)

S^n_{OPE} from parameterization: 0.0668 ± 0.0232 (6.3)

or,

Exploratory nTPE with only statistical error and *no systematic error*:

0.0738 ± 0.0237

6.2.3 A note on nTPE Error

A huge source of uncertainty in this nTPE calculation comes from the G_E^n parameterizations. G_E^n parameterizations may carry with them 18% or more uncertainty [75, 140, 56, 10]. This is a non-trivial contribution when we consider how G_E^n directly factors into our calculations for S^n and S^n_{OPE}, and therefore nTPE. This analysis considered two approaches:

1. Set G_E^n error to 1% to "simulate" high-precision values for G_E^n from the already completed GEN-II and currently-running GEn-RP experiments[4].

2. Use the G_E^n error as calculated via the parameterizations.

The GEN-II ([26]) and GEn-RP ([6]) experiments will yield high-precision values for G_E^n at the same Q^2 as SBS8 and SBS9 (4.5 (GeV/c)^2). Given the exploratory nature of this analysis and the confidence in upcoming high-precision G_E^n measurements, I am opting to look at the best case scenario. That is, an ***nTPE extraction that***

[4]The GEN-II and GEn-RP experiments are expected to achieve the highest level of precision yet for measurements of G_E^n [26, 6]. Though these experiments will most likely provide G_E^n values with errors greater than 1%, we can hope for the best case scenario and drive this value down in our approximations [86].

results from a G_E^n value with low ($\sim 1\%$) uncertainty. We can see the effects of this on the error bars in Figure 6-2. They are much smaller than what would be expected if I had incorporated the errors associated with G_E^n parameterizations. These small error bars correlate with the relatively small error that I report with my results. A standard error approximation would most likely include an uncertainty that is at least two, or three, times as much as what I have presented.

6.3 Error Analysis

The following discussion on error analyses corresponds to each of the three kinematics discussed in this thesis: SBS4, SBS8, and SBS9. A single analysis framework is used to analyze all three kinematics using the same method for each. The data structure and format for all kinematics is identical. The differences that arise for each form factor extraction derive from several factors:

1. Quality of raw data
2. Number of raw events for each kinematic and magnet setting:
   (a) SBS4, LD2, SBS Magnet 30%: 74.3 million
   (b) SBS4, LD2, SBS Magnet 50%: 13.3 million
   (c) SBS8, LD2, SBS Magnet 70%: 405.2 million
   (d) SBS9, LD2, SBS Magnet 70%: 743.8 million
3. Number of elastic events
   (a) Quality of elastic cut calibration
   (b) Effectiveness of elastic cuts
4. Background contamination
   (a) Determination of bounds for dy anti-cut
   (b) Quality of fit to dy anti-cut background

(c) Effectiveness of background subtraction method

5. Data-Simulation calibrations

   (a) χ^2-minimization technique

   (b) Nucleon peak scaling technique

   (c) Quality of fits for data-simulation tuning

Some of these factors are beyond our control, e.g., the quality of the raw data, the number of events available, and the amount of background contamination present. These are conditions of the experiment during run-time and operation, and will be the primary source of our uncertainty.

6.3.1 GMn Statistical Uncertainties

The ROOT analysis framework provides statistical error for most of its processes[5]. For instance, ROOT can easily, and directly, provide the error associated with performing a fit to a histogram, calculating an integral from a histogram, or the errors for the parameters used in a χ^2-minimization operation. Although ROOT provides the numerical values for various uncertainties, error propagation must still be handled *manually* for downstream calculations involving those extracted values and their errors.

For instance, ROOT will provide the errors, $\delta f_{(\text{scale, n})}$ and $\delta f_{(\text{scale, p})}$, associated with the values of $f_{(\text{scale, n})}$ and $f_{(\text{scale, p})}$ that we get from our data-simulation calibration:

$$f_{(\text{scale, n})} \pm \delta f_{(\text{scale, n})} \quad \text{and} \quad f_{(\text{scale, p})} \pm \delta f_{(\text{scale, p})}$$

[5]Special care must be taken when relying on errors provided by the ROOT data analysis framework. Error is only accurately propagated in ROOT if *ALL* uncertainties are properly defined when objects and variables are initialized/constructed in the environment. By default, ROOT weights each histogram entry with a value of 1, and the error for each bin is the square root of the bin content [29]. This could result in every downstream error calculation being inaccurate. Therefore, it is extremely important to properly initialize and define all errors in ROOT for proper propagation.

However, when we take the ratio of $(f_{(\text{scale, n})}/f_{(\text{scale, p})})$, we must handle the error propagation ourselves, i.e.,

$$\delta\left(\frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}}\right) = \left(\frac{f_{(\text{scale, n})}}{f_{(\text{scale, p})}}\right)\sqrt{\left(\frac{\delta f_{(\text{scale, n})}}{f_{(\text{scale, n})}}\right)^2 + \left(\frac{\delta f_{(\text{scale, p})}}{f_{(\text{scale, p})}}\right)^2} \tag{6.4}$$

Note that this error formulation assumes that variables such as the scale factors, yields, etc., are completely independent and uncorrelated. A full error analysis would need to re-evaluate these considerations and further assess the correlations between variables.

With this in mind, we need to assess our primary formulae and determine how to propagate the errors accordingly. In Eq. 4.15, we have errors associated with $R_{n/p}$, $F_{n/p}$, $\sigma_{R,n,\text{param}}$, and G_E^n. The values of ϵ and τ for each kinematic will be the nominal values listed in Table 6.2. The error for the yield ratio, $R_{n,p}$, comes from a similar formulation as that of Eq. 6.4. Again, I am considering the yields as independent and uncorrelated variables, and their uncertainties are the errors associated with taking the integral of the simulated nucleon histograms, i.e., the integral errors from ROOT[5]. We now have everything we need to estimate the error associated with $(G_M^n)^2$. We start with the relation,

$$(G_M^n)^2 = \left(\frac{\epsilon_n}{\tau_n}\right)\left[(\sigma_{R,n,\text{ param}})(R_{n/p})(F_{n/p})^2 - (G_{E,\text{ param}}^n)^2\right]. \tag{6.5}$$

Note that terms involving parameterizations are now generalized with the subscript "param" whereas previous notation used "MC". I have adopted this new naming convention to be more general and to allow for model-independence where applicable. From here, we can evaluate the two terms inside the brackets separately. Looking at the first term we get,

$$\frac{\delta\left((\sigma_{R,n,\text{param}})(R_{n/p})(F_{n/p})^2\right)}{(\sigma_{R,n,\text{param}})(R_{n/p})(F_{n/p})^2} = \sqrt{\left(\frac{\delta\sigma_{R,n,\text{ param}}}{\sigma_{R,n,\text{ param}}}\right)^2 + \left(\frac{\delta R_{n/p}}{R_{n/p}}\right)^2 + (2F_{n/p}\delta F_{n/p})} \tag{6.6}$$

The error associated with the second term in the brackets of Eq. 6.5, $(G_E^n)^2$, is simply,

$$\delta\big((G^n_{E,\text{ param}})^2\big) = 2G^n_{E,\text{ param}}(\delta G^n_{E,\text{ param}}). \tag{6.7}$$

Now, we carry over the constant terms and combine Equations 6.6 and 6.7 to find,

$$\delta\big((G^n_M)^2\big) = \sqrt{\left(\frac{\epsilon_n}{\tau_n}\right)\left[\left[\frac{\delta\big((\sigma_{R,n,\text{param}})(R_{n/p})(F_{n/p})^2\big)}{(\sigma_{R,n,\text{param}})(R_{n/p})(F_{n/p})^2}\right]^2 + \left[\delta\big((G^n_{E,\text{ param}})^2\big)\right]^2\right]} \tag{6.8}$$

In Eq. 6.8, the two bracketed terms under the radical are from Equations 6.6 and 6.7, and are treated as uncorrelated. The calculated statistical errors for $F_{n/p}$, $R_{n/p}$, and G^n_M are listed in Table 6.4.

Statistical Errors for $F_{n/p}$, $R_{n/p}$, and G^n_M			
	$\delta F_{n/p}$	$\delta R_{n/p}$	δG^n_M
SBS4	0.0056	0.0013	0.0254
SBS8	0.0046	0.0016	0.0240
SBS9	0.0035	0.0010	0.0149

Table 6.4: **[Exploratory]** Statistical errors associated with $F_{n/p}$, $R_{n/p}$, and G^n_M. Errors for the individual scale factors and nucleon yields are extracted from ROOT and then manually propagated. The errors for G^n_M are calculate using Eq. 6.8.

6.3.2 GMn Systematic Uncertainties

It cannot be over-stressed that the work presented throughout this thesis is exploratory in nature. This should be even further stressed in regards to my systematic uncertainty analysis. I have performed a limited study using only a select set of contributors of systematic uncertainty in hopes that it can provide some future insight. I must re-iterate and emphasize that as a whole, there is still a significant amount of systematic studies and benchmarking that remain to be done.

The main contributors to the systematic error in this analysis derive primarily from the elastic cuts, the various fits and parameterizations applied throughout the analysis process, and the multiple corrections that have been introduced. Given the

scope of this analysis I was only able to assess a limited selection of elastic cuts:

Sources considered in my limited systematic uncertainty study:

- **Invariant mass, W^2:** Symmetric cut around $\sim M_p^2$ for elastic events
- **HCal cluster minimum energy, $HCal\_clus\_e\_min$:** cuts events below the specified threshold.
- **Minimum BBCal PreShower energy, PS_{min}:** Pion cut. Typical value is 0.15 GeV. Everything below this value is cut, i.e., pions.
- **Total energy over momentum, E/p:** Wide, symmetric cut around 1.0. Selects elastic events, but is primarily applied as a means to cut background.
- **Fiducial cut multiplier, $Fcut\_mult$:** Acceptance-matching cut for HCal. Accounts for imbalance between nucleon pairs detected at HCal due to the deflection caused by the SBS magnet.
- **dy cut multiplier, dy_mult:** Symmetric cut around the central value of dy. Primarily used as a background cut, but inherently cuts on the momentum distribution of scattered nucleons.
- **HCal acceptance cut multiplier, $HCal\_XY\_mult$:** Quality-assurance cut. Removes events near the edge of HCal which may be poorly reconstructed.

In actuality, there are more contributing factors which have been omitted, rather than included, in this study. The following is a *non-exhaustive* list of cuts which should be included, or at least considered, for a full-scale analysis:

Sources *not* considered for systematic uncertainty:

- Background subtraction (dy limits, normalization, fit function, etc.)
- Timing cuts

- Corrections for HCal inefficiencies (type of efficiency map used, weighting, etc.)
- HCal acceptance asymmetries
- HCal detection efficiency
- ***and many more...***

In order to assess my set of systematic variables, I looked at the effects of varying a single cut while holding the others constant. The "varying" of the cut can take multiple forms depending on which particular cut is being studied. For threshold cuts, like PS_{min} or $HCal\_clus\_e\_min$, I iterated over a range of threshold values near the nominal cut value. For example, the typical PS_{min} cut is 0.15 GeV. In my study, I varied the PreShower cut between 0.14 GeV $\leq PS_{min} \leq$ 0.40 GeV with step sizes of 0.005 GeV. For each interval, I performed a full analysis and extracted the values of $R_{n/p}$, $F_{n/p}$, and G_M^n. I then calculated the standard deviation for each variable and used this as the value for systematic uncertainty. This same technique applies to elastic cuts which use *multipliers*; these include the fiducial cut multiplier, the dy cut multiplier, and the HCal acceptance cut multiplier.

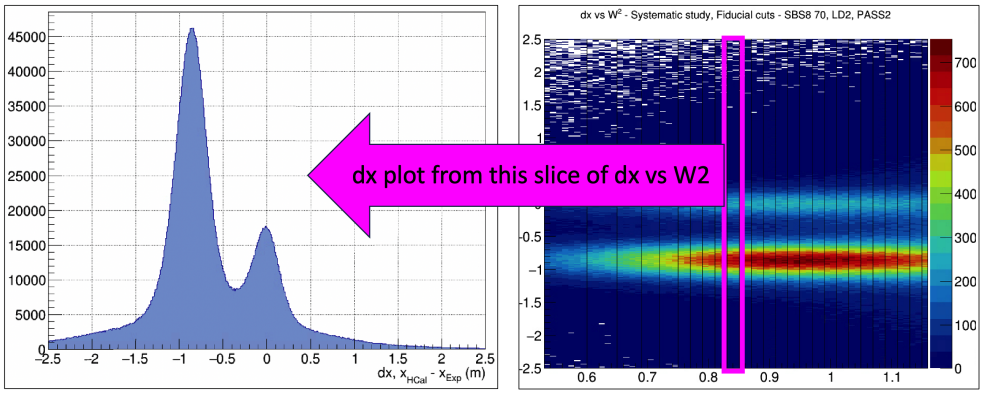


Figure 6-3: A dx plot made from the Y-Projection of a W^2 slice as part of a systematic study. For each W^2 slice we can perform a full analysis and extraction and then compare the results across all slices.

Some elastic cuts are defined with an upper and lower bound. For instance, a typical invariant mass cut is 0.75 GeV $\leq W^2 \leq$ 1.2. Another range-based elastic cut is E/p. For these types of cuts, I employed a slightly different method: the so-called

method of slices. Please use Figure 6-3 as a reference in the following discussion. To investigate the systematic effects of the W^2 cut, I first create a 2D histogram of dx versus W^2. I then partition the histogram into vertical *slices* along the W^2 axis. The bounds of each partition are determined so that all slices contain the same number of events. The Y-Projection of each slice provides a corresponding dx histogram which can then be fed into the analysis machinery for extracting $R_{n/p}$, $F_{n/p}$, and G_M^n. For each variable, I calculate the standard deviation across all of the slices, and use that value as the systematic uncertainty associated with that elastic cut. These individual errors are then added in quadrature to determine their total systematic contribution.

The accumulated findings for this systematic study are listed in Table 6.5. It goes without saying that this study is extremely limited. In total, I have only studied six elastic cuts when there are definitely many more that need to be considered. Also, the step sizes that I used were relatively coarse. Future studies should include a more exhaustive list of cuts with a much finer variation size.

	$F_{n/p}$			$R_{n/p}$			G_M^n		
Elastic Cut	**SBS4**	**SBS8**	**SBS9**	**SBS4**	**SBS8**	**SBS9**	**SBS4**	**SBS8**	**SBS9**
W^2	0.0013	0.0024	0.0019	0.0048	0.0080	0.0006	0.0068	0.0032	0.0007
$HCal\_clus\_e\_min$	0.0038	0.0080	0.0013	0.0028	0.0080	0.0004	0.0059	0.0066	0.0005
PS_{min}	0.0044	0.0050	0.0037	0.0017	0.0080	0.0015	0.0070	0.0070	0.0016
E/p	0.0032	0.0016	0.0009	0.0013	0.0060	0.0005	0.0052	0.0088	0.0005
$Fcut\_mult$	0.0025	0.0010	0.0005	0.0032	0.0030	0.0002	0.0040	0.0070	0.0002
dy_mult	0.0035	0.0022	0.0037	0.0065	0.0070	0.0013	0.0040	0.0067	0.0015
$HCal\_XY\_mult$	0.0044	0.0041	0.0261	0.0016	0.0013	0.0088	0.0069	0.0055	0.0115
Combined Systematic Uncertainty	**0.0085**	**0.0058**	**0.0033**	**0.0074**	**0.0056**	**0.0045**	**0.0100**	**0.0066**	**0.0032**

Table 6.5: [Exploratory] Systematic uncertainty contributions to $R_{n/p}$, $F_{n/p}$, and G_M^n from elastic cuts for SBS4, SBS8, and SBS9.

6.3.3 nTPE Statistical Uncertainties

In short, the statistical uncertainties for this nTPE study have been derived purely from ROOT-based error values[5]. The values reported in Table 6.3 and Equation 6.2 all derive from the plot of Figure 6-2. The linear extrapolation between the two reduced cross-section points for SBS8 and SBS9 is determined using ROOT's built-in functionality. After the linear fit is determined, I can directly extract the parameters for the y-intercept and slope. ROOT also provides the error associated with these values. These are what I report in Table 6.3 for Y-intercept Error and Slope Error. There are errors bars associated with the two reduced cross-section data points in Figure 6-2. These are determined from the experimentally-extracted G_M^n and parameterized G_E^n values for SBS8 and SBS9. Each G_M^n value is attached to an experimental error value, but the error on G_E^n is approximated according to the discussion in Sec. 6.2.3. A full-scale and detailed analysis on statistical uncertainties for nTPE is beyond the scope of this thesis.

6.3.4 nTPE Systematic Uncertainties

The analysis presented here does not include a study on systematic uncertainties for nTPE. All of the error that I quote in my exploratory results for nTPE are purely statistical. The topic of systematic uncertainty is an extremely important factor for the final nTPE analysis. The robustness and reliability of any Rosenbluth Separation depends heavily on the uncertainties therein. The nTPE proposal provides estimated values for systematic uncertainties which could be used as a reference guide [4]. My results will forego any representation of systematic uncertainty and will remain purely statistical (and exploratory).

Chapter 7

Conclusion

The GMn experiment was conducted to measure the magnetic form factor of the neutron, G_M^n, at multiple kinematic settings corresponding to different momentum-transfer, Q^2, values. These kinematic points were SBS4, SBS7, SBS8[1], SBS9, SBS11, and SBS14. This thesis presents an exploratory analysis and extraction of G_M^n for kinematics SBS4, SBS8, and SBS9. In addition to a G_M^n extraction, we have developed the necessary components needed to compare experimental results for Rosenbluth Separation with the single-photon polarization transfer measurements for the neutron. This chapter aims to summarize the findings herein, discuss any pitfalls, and to make an outlook towards future work on the subject.

7.0.1 Summary of GMn Analysis and Findings

The three exploratory extractions of G_M^n are summarized in Table 7.1. It is important to compare these experimentally-extracted values of G_M^n to those from previous experiments. In this discussion, I will graphically and numerically compare my extracted values to the Kelly G_M^n parameterization ([75]). Kelly's formulation is used in our Monte Carlo simulation framework for calculating G_M^n; Kelly also prioritizes polarization transfer measurements and is therefore essentially a single-photon exchange parameterization. Additionally, I will also plot my extracted values on top of

[1]SBS8 is technically an nTPE point, but it corresponds to an identical experimental setup as the rest of the GMn kinematics. Therefore, it can easily be included in the set of GMn points.

existing world data in order to help visualize how they all fit together.

Experimental GMn Extractions				
	Exper. Q^2	$G_M^n/\mu G_D$	**Syst. Error**	**Stat. Error**
SBS4	2.937	0.9772	0.0154	0.0100
SBS8	4.368	0.9323	0.0174	0.0066
SBS9	4.393	0.9183	0.0117	0.0032

Table 7.1: [Exploratory] Summary of experimental G_M^n extractions for SBS4, SBS8, and SBS9.

Kelly Parameterizations			
	Exper. Q^2	$G_M^n/\mu G_D$	**Param. Error**
SBS4	2.937	0.9890	0.0125
SBS8	4.368	0.9218	0.0130
SBS9	4.393	0.9209	0.0130

Table 7.2: [Exploratory] Summary of Kelly G_M^n parameterizations for SBS4, SBS8, and SBS9.

Table 7.2 lists the Kelly parameterizations and errors for G_M^n at the experimental Q^2 for each kinematic. When we compare the Kelly parameterizations with my extractions listed in Table 7.1, we find that each experimental result overlaps with its corresponding parameterization. We can better visualize the agreement by plotting these values together. Figure 7-1 shows my experimental G_M^n values overlaid with the Kelly parameterization. This shows the agreement between the two. Each of my experimental values are near the central Kelly parameterization curve, and my error bars all seem reasonable.

It is also useful to plot the experimental results with the world data for G_M^n. This is shown in Figure 7-2. The red dashed line corresponds to the same Kelly parameterization shown in Fig. 7-1. It can be a bit difficult to isolate my experimental points amidst the other data points in this figure. Therefore, I have provided a *zoomed in* version of the same plot (see Figure 7-3). It is no surprise that our G_M^n results

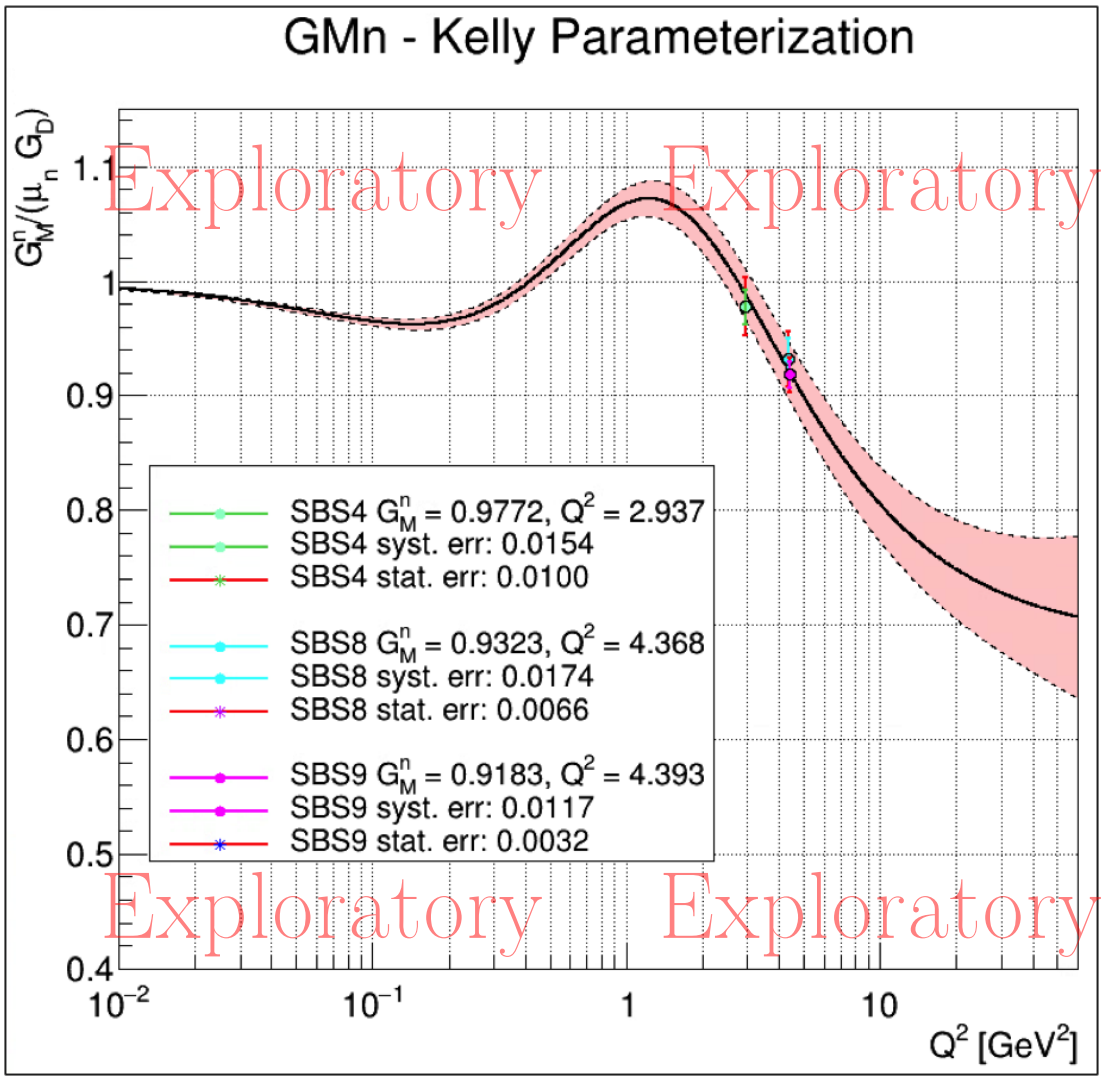


Figure 7-1: [Exploratory] Experimental G_M^n values overlaid with the Kelly parameterization.

for SBS4, SBS8, and SBS9 are in agreement with the world data, given that these three points are within the Q^2 range of previous measurements. At the same time, we can see that the two higher Q^2 points, SBS8 and SBS9, have significantly better statistical errors than the best world data available at that Q^2 range.

7.0.2 Summary of nTPE Analysis and Findings

From the values of G_M^n extracted for SBS8 and SBS9 I was able to perform an exploratory LT separation and extract a Rosenbluth Slope between those two points.

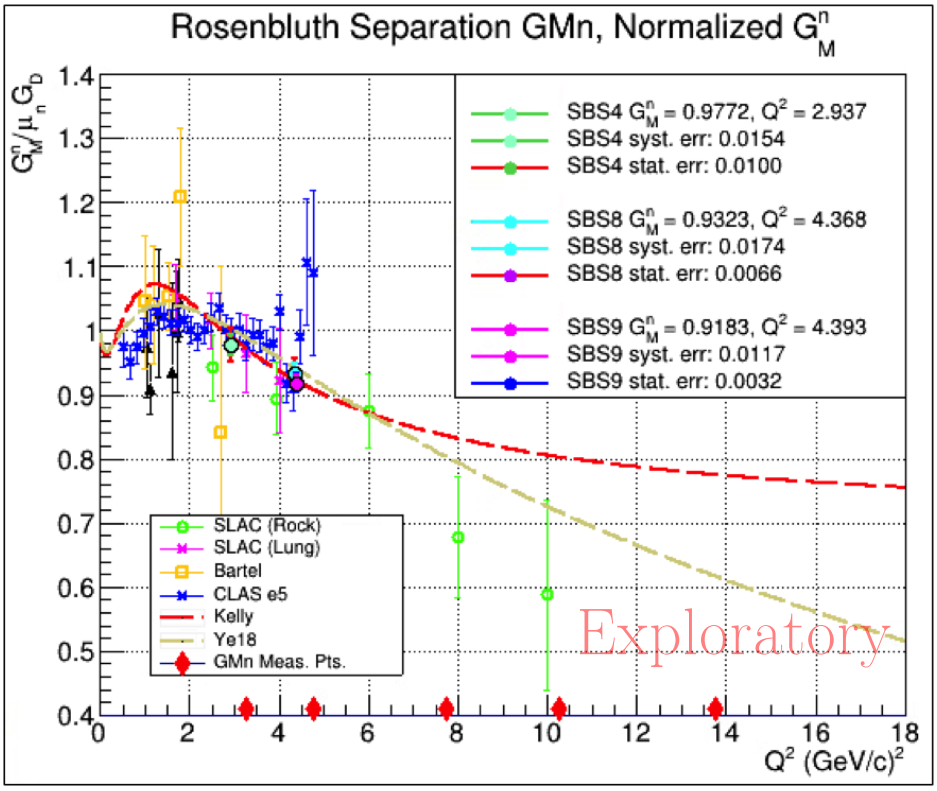


Figure 7-2: [Exploratory] Experimental G_M^n values overlaid with world G_M^n data.

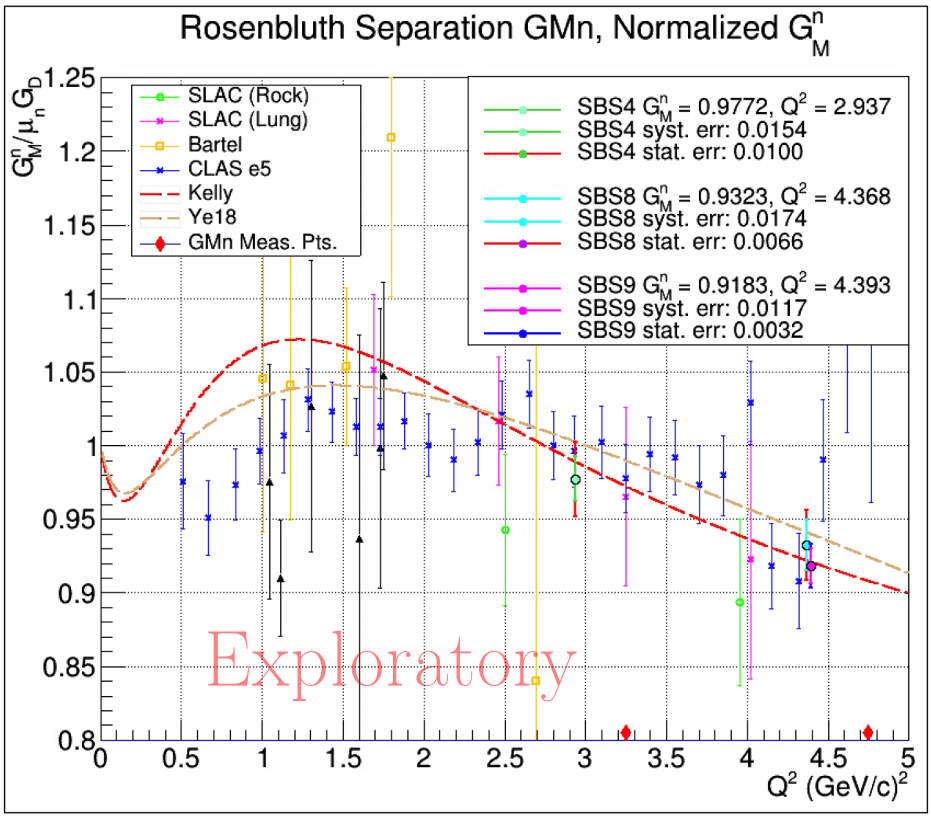


Figure 7-3: [Exploratory] Experimental G_M^n values overlaid with world G_M^n data in the lower Q^2 range (0 - 5 (GeV/c)^2). The Y-axis has also been limited to 0.8 - 1.25. This helps distinguish the relative agreement of our experimental results amidst the world data.

This is the first measurement of the neutron via the Rosenbluth Technique and therefore there are no previous measurements to compare against. However, we do have the projected value for nTPE that was provided in the proposal[2], as well as theoretical neutron calculations from parameterizations.

The nTPE proposal states the following projection for nTPE [4]:

$$\textbf{Estimated value for nTPE from proposal: } 0.063 \pm 0.012 \pm 0.01 \tag{7.1}$$

where the first uncertainty is due to the accuracy of G_E^n/G_M^n and the second uncertainty is derived from the projected precision of this experiment. My exploratory value for nTPE is:

$$\boxed{\textbf{Exploratory nTPE: } 0.0738 \pm 0.0237} \tag{7.2}$$

At this point, a direct comparison between Equations 7.1 and 7.2 is not really fair or warranted. My error analysis for nTPE is extremely limited: I only consider statistical error and I make some serious approximations along the way. The uncertainty on my experimental nTPE value (Eq. 7.2) could be under-estimated at this very early stage of analysis, even at $\sim 32\%$. The percent difference between my experimental nTPE value and the one listed in the proposal is approximately 16%. I can't draw many conclusions from these values given the underlying simplifications and approximations. However, this is the first measurement of its kind and there has been a substantial amount of work to arrive at this point. Further analysis and future studies will hopefully prove this out further.

An additional means to assess the nTPE results presented here is by comparing the form factor ratio from our Rosenbluth Separation measurement with those from polarization transfer data. This was discussed earlier in regards to the Form Factor Ratio Puzzle (Sec. 1.9). Figure 7-4 shows parameterization calculations of G_E^n/G_M^n for LT Separation (red dashed line) and Polarization Transfer (blue dashed line) methods [2, 93]. The single magenta data point corresponds to my experimental

[2]The projected values in the nTPE proposal are estimates inspired by the two-photon results for the proton [4].

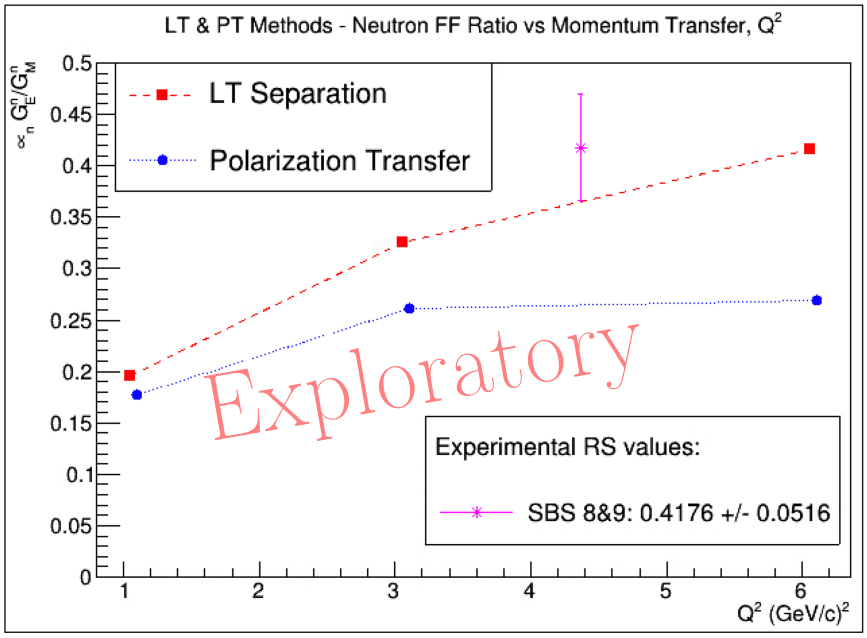


Figure 7-4: **[Exploratory]** Rosenbluth Slope vs. Q^2 for SBS8 and SBS9. The blue data points are from the form factor parameterization in [93]. The red data points are taken directly from Blunden, Figs. 12 and 13 of [2]. The single magenta point is the neutron form factor ratio from my experimental Rosenbluth Slope extraction. Plot reproduced from [2].

value for G_E^n/G_M^n. This ratio can be calculated from the values of Table 6.3. We have:

$$G_E^n \approx 0.0140 \pm 6.78\text{e-}6, \quad G_M^n \approx 0.0335 \pm 0.0041,$$

and so,

$$\frac{G_E^n}{G_M^n} \approx 0.4176 \pm 0.0516$$

These errors are based on previously discussed approximations and are statistical only.

This figure needs to be taken with multiple grains of salt. Everything discussed in regards to this is purely speculative, hypothetical, and based on my very exploratory findings. Figure 7-4 is intended to highlight the Form Factor Ratio Puzzle (FFRP). Recall that the FFRP is a discrepancy that has been identified between measurements using the Rosenbluth Technique and single-photon exchange measurements. As Q^2 increases in Fig. 7-4, we see a growing separation between the two datasets.

This is the unexpected behavior that the FFRP refers to. My experimental value of 0.4176 tends towards the region for LT Separation which indicates some promise. I will refrain from over-speculating on the subject, but it is interesting to see how my experimental data point orients in comparison with the two parameterizations. A leading thought is that higher order radiation effects, i.e., two-photon exchange, is the cause of the discrepancy between LT separation and polarization transfer measurements. Hopefully, future work will tighten up the experimental findings and provide a more definitive result.

7.0.3 Future Work

The nature of this thesis leaves much work left to be completed, improved, and redone. This includes repeating the entire analysis with the latest versions of data, performing a significantly broader and more critiqued systematic analysis, and improving any deficiencies in data or infrastructure. It would take too long to expound upon each of these in great detail and so, I will address a few points which I find most important.

Calibrations

We should always consider our experimental machine to be in a state of ongoing fine-tuning and calibrations. It has now been over two years since we completed the GMn experiment, and we are still learning more about each detector system with every passing day. Most of my analysis used data in the Pass 2 stage. At the time of this writing, future passes on the data are already planned and/or in the works. After each previous pass on the data, analysis results have looked progressively cleaner and the data responded to cuts with better stability. Therefore, any analysis that uses future passes is sure to be noticeably improved. Until the final result is determined — and perhaps even after — calibrations will forever be underway.

HCal inefficiency corrections

An ad hoc solution has been implemented to correct for low efficiency HCal blocks in SBS9. Moving forward, a broader set of solutions will need to be implemented and rigorously tested. The low efficiency most likely stems from bad connections and/or *light leaks* between the PMTs and conjoining surfaces. This can have multiple negative effects including over-saturated signals, overwhelming background contamination, and poor detection response. In any case, these issues manifest as poor (or no) signal from the affected blocks. These poorly-operating blocks can be replicated in the simulation machinery in such a way that the simulated outputs correlate with the measured signal responses we see in data. A *direct* way of implementing this may be to introduce a scalar multiplier or gain coefficient to the outputs of each simulated HCal block. These coefficients could then be optimized to match each signal to the experimental data. Implementing this should allow the simulation to account for lower proton detection efficiencies where necessary. For the time being, the implementation of an HCal efficiency map with subsequent event weighting has proven effective at resolving the issue. This technique seems formidable and would greatly benefit from further study and scrutiny. There are multiple types of efficiency maps that can be used. Each one should be tested and scrutinized thoroughly.

Systematics

This is perhaps the area that leaves the most work left to be done. In my analysis of GMn, I only evaluated six sources of systematic uncertainty and made no attempts for nTPE. Systematics play an important role in both experiments, but are a central component for nTPE. I have tried to describe each of the cuts that have been omitted from this analysis and the effects they may have, but this is a difficult task. The deeper I look, the longer the list grows. Nevertheless, this is sure to be a central focus for any related analyses efforts going forward.

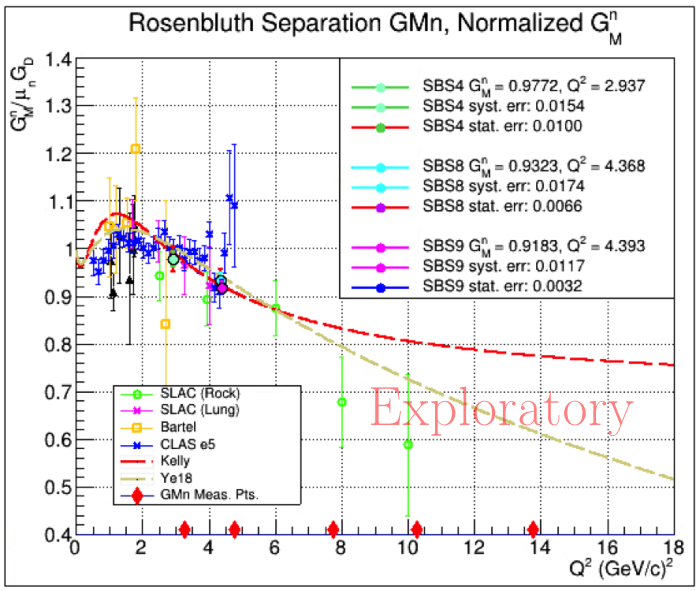


Figure 7-5: [Exploratory] World data and parameterizations for G_M^n with experimental values for SBS4, SBS8, and SBS9.

Further Exploration for GMn

There is so much left to explore between these two experiments. The only way to pin any hypotheses or theories down is to keep pushing forward and see what evolves. The GMn experiment pushes the Q^2 bounds of G_M^n measurements up to 13.7 (GeV/c)^2. Figure 7-5 shows the world data set for G_M^n. The plot includes red diamond markers which indicate the nominal measurement points for the GMn experiment. My experimental G_M^n extractions are in the lower range of these GMn points. The higher Q^2 measurement points (SBS7, SBS11, and SBS14) may prove to be more illuminating. For instance, Figure 7-5 includes two G_M^n parameterizations: Kelly (red dashed line, [75]) and Ye (tan dashed line, [140]). In the plot, we can see that the two parameterizations diverge near 6.5 (GeV/c)^2. The experimental extractions for SBS7, SBS11, and SBS14, will provide insight on which trend may hold precedence. My experimental values are in a Q^2 region where the two parameterizations agree reasonably well. Therefore, from SBS4, SBS8, and SBS9 alone, it is hard to make a conclusion towards one parameterization over another. Also, these aren't the only two parameterizations out there. There are several others that we can, and should, compare with. The true tests of these parameterizations comes once the analyses of

the higher Q^2 points are completed. These bound-pushing GMn measurements have truly brought us to the next frontier.

Further Exploration for nTPE

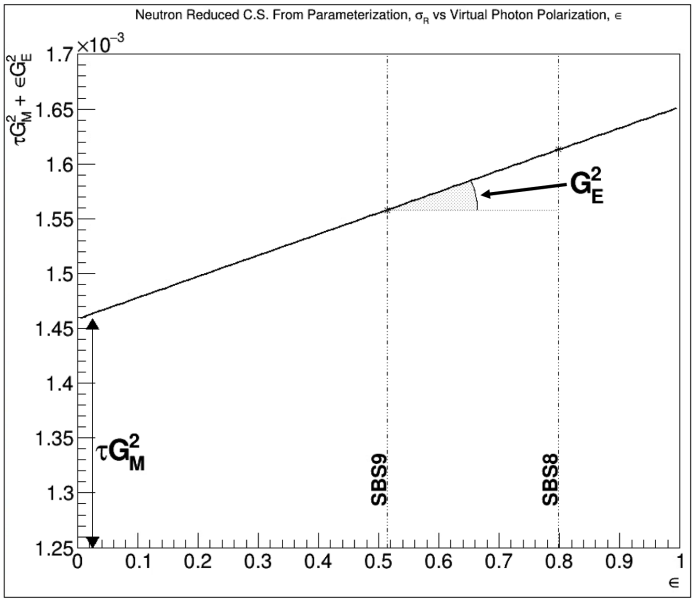


Figure 7-6: Neutron reduced cross-section from parameterization vs. epsilon with nominal values of SBS8 and SBS9. The neutron reduced cross-section is calculated at nominal epsilon points for SBS8 and SBS9 using Ye's parameterization ([140]).

I have left plenty of room for others to continue exploring nTPE. One hypothetical that comes to mind is to extend the nTPE measurements in epsilon. Figure 7-6 shows a nominal neutron reduced cross-section versus epsilon plot for our two kinematics SBS8 and SBS9. The two data points were calculated using Ye's parameterization at nominal values of epsilon for SBS8 and SBS9. If we explored more than two epsilon points then perhaps some new and interesting physics could emerge. The single-photon exchange approximation predicts a straight line between the two epsilon points. However, only having two epsilon points means me are restricted to a simple linear extrapolation between points. By expanding our data set in epsilon, we could find out if this relationship is truly linear all the way through. Who knows, there could be something much more interesting than higher order photon exchange lurking in there.

7.0.4 Final Remarks

Although there is still a huge amount of work left to be completed, I think we are off to a great start. We have successfully extracted the neutron magnetic form factor, G_M^n, at three separate kinematics using the ratio between quasi-elastic $e-n$ and $p-n$, for scattering from a deuterium target. The implementation of the ratio method resulted in high-precision measurements and the cancellation of many uncertainties. From the G_M^n values for SBS8 and SBS9 we were then able to implement Rosenbluth Separation and compare the results with those from single-photon exchange measurements.

The GMn and nTPE experiments took place in Hall A of Jefferson using the BigBite and Super BigBite spectrometers. The scattered electrons were detected along the BigBite arm while the scattered protons and neutrons were simultaneously measured on the Super BigBite arm. The scattered hadrons flight path took them through the yoke of the SBS dipole magnet. Traversing through this region caused deflections in the protons but not the neutrons. This was the process by which we were able to separate and distinguish protons from neutrons. The BigBite arm contained a large detector package which consisted of several GEM layers, the GRINCH, the timing hodoscope, and the BigBite Calorimeter.

Our research group provided the majority of the GEM detectors utilized throughout the entire SBS Program of experiments. Our GEMs are the primary tracking detectors in these experiments and from day one, never ceased to amaze us. I was lucky enough to participate in the designing, manufacturing, construction, testing, commissioning, running, and data analysis for the UV GEMs — a full-cycle experience that I believe few graduate students get to experience. The UV GEMs introduced some never-before-seen features, and pushed the bounds of performance with their high rate capabilities and sustainability.

Overall, the extracted values of G_M^n for SBS4, SBS8, and SBS9, are all satisfactory with reasonably-sized uncertainties. The calculated difference between our Rosenbluth Separation result with those from single-photon measurements is also

within the projected bounds listed in the nTPE proposal. It should be noted that my experimental result had an uncertainty of approximately 32% and so, finding overlap wasn't too incredibly difficult. Regardless, this marks the first measurement of Rosenbluth Separation on the neutron.

Bibliography

[1] D. Adikaram et al. Towards a resolution of the proton form factor problem: New electron and positron scattering data. *Phys. Rev. Lett.*, 114:062003, Feb 2015.

[2] A. Afanasev, P.G. Blunden, D. Hasell, and B.A. Raue. Two-photon exchange in elastic electron–proton scattering. *Progress in Particle and Nuclear Physics*, 95:245–278, 2017.

[3] W. M. Alberico, S. M. Bilenky, C. Giunti, and K. M. Graczyk. Electromagnetic form factors of the nucleon: New fit and analysis of uncertainties. *Physical Review C*, 79(6), June 2009.

[4] S. Alsalmi et al. Measurement of the two-photon exchange contribution to the electron-nuetron elastic scattering cross section. *A New Proposal to Jefferson Lab PAC48*, 2009.

[5] John R.M. Annand and Rachel Montgomery. Bigbite timing hodoscope update. https://sbs.jlab.org/DocDB/0000/000024/001/BBhodo.07.18.pdf, July 2018.

[6] J.R.M. Annand et al. E12-17-004: Measurement of the ratio gen/gmn by double-polarized 2h(e,e'n) reaction. *Jefferson Lab: PAC45*, May 2017.

[7] D.S. Armstrong, B. Moffit, and R. Suleiman. Target density fluctuations and bulk boiling in the hall a cryotarget. *Jefferson Lab Technical Note*, JLAB-TN-03-017, May 2003.

[8] J. Arrington. How well do we know the electromagnetic form factors of the proton? *Phys. Rev. C*, 68:034325, Sep 2003.

[9] J. Arrington. Implications of the discrepancy between proton form factor measurements. *Phys. Rev. C*, 69:022201, Feb 2004.

[10] J. Arrington, P.G. Blunden, and W. Melnitchouk. Review of two-photon exchange in electron scattering. *Progress in Particle and Nuclear Physics*, 66(4):782–833, 2011.

[11] John Arrington. A-b-simc. http://www.jlab.org/~johna/SIMC_documents/simc.ps, July 2001.

[12] Todd Averett. Update on teh bigbite gas ring imaging cherenkov (grinch) detector for sbs. https://sbs.jlab.org/DocDB/0000/000021/001/SBS_GRINCH_update_20-Jul-2018.pdf, July 2018.

[13] S. Bachmann, A. Bressan, L. Ropelewsky, F. Sauli, A. Sharma, and D. Mörmann. Charge amplification and transfer processes in the gas electron multiplier. *Nuclear Instruments and Methods in Physics Research*, 438:376–408, December 1999.

[14] M. Bahmanabadi. A method for determining the angular distribution of atmosphereic muons using a cosmic ray telescop. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 916:1–7, February 2019.

[15] Dmitriy Baranov and Oleg Rogachevsky. Simulation of the gem detector for bm@n experiment. *EPJ Web Conf.*, 138:1104, 2017.

[16] Scott Barcus. g_m^n hcal energy deposition study. https://hallaweb.jlab.org/wiki/images/archive/4/4f/20200518140632!HCal_Energy_Deposition.pdf, December 2020.

[17] Scott Barcus. Sbs hadron calorimeter commissioning results. https://sbs.jlab.org/DocDB/0001/000163/001/HCal_Presentation_Feb2022.pdf, February 2022.

[18] P. Bhattacharya, B. Mohanty, S. Mukhopadhyay, N. Majumdar, and H. Natal da Luz. 3d simulation of electron and ion transmission of gem-based detectors. *Nuclear Instruments and Methods A*, 870:64–72, 2017.

[19] D. Borisyuk and A. Kobushkin. Two-photon exchange in elastic electron scattering on hadronic systems. *Ukrainian Journal of Physics*, 66(1):3, January 2021.

[20] John Boyd. Crosstalk analysis. https://sbs.jlab.org/DocDB/0005/000527/001/Crosstalk%20Summary%20-%20Mar%2010%2C%202023.pdf, March 2023.

[21] R. Bradford, A. Bodek, H. Budd, and J. Arrington. A new parameterization of the nucleon elastic form factors. *Nuclear Physics B - Proceedings Supplements*, 159:127–132, September 2006.

[22] R. Bradford, A. Bodek, H. Budd, and J. Arrington. A new parameterization of the nucleon elastic form factors. *Nucl.Phys.Proc.Suppl*, 159:127–132, May 2006.

[23] René Brun, Fons Rademakers, and Suzanne Panacek. Root, an object oriented data analysis framework. https://cds.cern.ch/record/491486/files/p11.pdf.

[24] C. Büttner, M. Capeáns, W. Dominik, M.Hoch, J.C. Labbé, G. Manzin, G. Million, L. Ropelewski, F. Sauli, and A. Sharma. Progress with the gas electron multiplier. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 409:79–83, 1998.

[25] A. Buzulutskov, L. Shekhtman, A. Bressan, A.Di Mauro, L. Ropelewski, F. Sauli, and S. Biagi. Gem operation in pure noble gases and the avalanche confinement. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 433(1):471–475, 1999.

[26] A. Camsonne et al. Measurement of the neutron electromagnetic form factor ratio gen/gmn at high q-squared. *A New Proposal to Jefferson Lab PAC34*, December 2008.

[27] A. Cardini, G. Bencivenni, W. Bonivento, C. Deplano, P. de Simone, F. Murtas, D. Pinci, M. Poli Lener, and D. Raspino. Results from the measurement of gem-based detector parameters performed with x-rays. https://cds.cern.ch/record/1744770/files/LHCb-TALK-2002-057.pdf, November 2002.

[28] Gordon D. Cates. The jlab sbs program and nucleon form factors at high q^2. https://indico.ectstar.eu/event/58/contributions/1367/attachments/1098/1417/01_CatesG.pdf, September 2019.

[29] CERN. *ROOT Data Analysis Framework*. CERN, May 2018.

[30] CERN. Homeland security. https://cms.cern/content/homeland-security, 2021.

[31] E. E. Chambers and R. Hofstadter. Structure of the proton. *Phs. Rev.*, 103, September 1956.

[32] M. E. Christy et al. Form factors and two-photon exchange in high-energy elastic electron-proton scattering. *Phys. Rev. Lett.*, 128:102002, Mar 2022.

[33] Lucia Di Ciaccio, editor. *ESIPAP IPM*, Interactions of Particles/Radiation with Matter, CERN, January 2018.

[34] Jeffrey Coderre, editor. *Principles of Radiation Interactions*, Energy Deposition: electrons, MIT, 2004.

[35] ALICE Collaboration. Alice technical design report of the time projection chamber. Technical Report ALICE TDR 7, CERN/LHC, January 2000.

[36] Geant4 Collaboration. Geant4 a simulation toolkit: Introduction to geant4. https://geant4-userdoc.web.cern.ch/UsersGuides/IntroductionToGeant4/fo/IntroductionToGeant4.pdf, December 2023.

[37] Johann Collot, editor. *ESIPAP IPM*, Interaction of particles with matter (lecture 1), CERN, February 2019.

[38] Mara Corbetta. Studies on gas mixture and gas recirculation effects on gem detectors operation at the cern gif++ facility. Master's thesis, Universitá degli Studi Milano-Bicocca, March 2018.

[39] Juan Carlos Cornejo. Simulations of hcal efficiency and resolution. https://sbs.jlab.org/DocDB/0000/000086/002/HCAL_Simulations_report.pdf, May 2021.

[40] Gabriele Croci. *Study of relevant parameters of GEM-based detectors*. PhD thesis, Università Degli Studi Di Milano - Bicocca, September 2007.

[41] Michelangelo D'Agostino. Cherenkov light. *Symmetry, A joint Fermliab/SLAC publication*, August 2009.

[42] Iain Darby, editor. *Radiation Interactions*, An Introduction to the Interaction of Radiation with Matter, International Atomic Energy Agency, June 2012.

[43] Provakar Datta. Probing the neutron's current distribution with unprecedented q^2 reach. https://www.jlab.org/sites/default/files/user-liaison/jluo2022/poster_JLOU22_PDatta.pdf.

[44] Provakar Datta. Perforamance of the bigbite calorimeter during sbs-gmn experiment. https://indico.jlab.org/event/529/contributions/10270/attachments/8180/11693/F%26C_MIT_gmn%26bbcal_2022.pdf, August 2022.

[45] Provakar Datta. Bbcal energy calibration updates for gmn/ntpe pass 2 readiness. https://logbooks.jlab.org/entry/4135857, Febraury 2023.

[46] Provakar Datta. Sbs-gmn analysis update. https://indico.jlab.org/event/602/contributions/11902/attachments/8858/13065/Hall_A_win23.pdf, January 2023.

[47] C.W. de Jager et al. The super-bigbite spectrometer for jefferson lab hall a. Technical report, Jefferson Lab, January 2010.

[48] Valérie De Smet. *Study of a GEM tracker of charged particles for the Hall A high luminosity spectrometers at Jefferson Lab*. PhD thesis, 09 2011.

[49] Marco Delmastro, editor. *ESIPAP IPM*, Experimental Particle Physics: Particle Interactions in Particle Detectors, European Scientific Institute - Archamps, January 2014.

[50] Serge Duarte. *Gas Electron Multipliers: Development of large area GEMS and spherical GEMS*. PhD thesis, Universität Bonn, August 2011.

[51] Dipangkar Dutta. Radiative corrections - the simc way.

[52] I. Estermann, R. Frisch, and O. Stern. Magnetic moment of the proton. *Nature*, 132:169–170, July 1933.

[53] A. Karadzhinova et al. Impact of gem foil hole geometry on gem detector gain. *Journal of Instrumentation*, 10, December 2015.

[54] J. Annand et al. Bigbite timing hodoscope. https://hallaweb.jlab.org/wiki/images/f/fb/BB-Hodo-technical-specs.pdf.

[55] Richard Clinton Fernow. *Introduction to Experimental Particle Physics*. Cambridge University Press, 1986.

[56] Feuerbach et al. A measurement of g e n at high momentum transfer in hall a.

[57] Laboratory for Nuclear Technologies. Interaction of gamma-rays with matter. https://www.fe.infn.it/radioactivity/educational/detection.html, 2021.

[58] Sverker Fredriksson. An estimate of the effect of nuclear fermi motion on large transverse momentum inclusive spectra from high-energy proton-nucleus collisions. *Nuclear Physics B*, 94(2):337–343, 1975.

[59] Sverker Fredriksson. On large transverse momentum production in high-energy proton-nucleus collisions. *Nuclear Physics B*, 111(1):167–177, 1976.

[60] Eric Fuchey. Elasticeventselection_multicluster_inputeff. private communication/shared scripts, April 2024.

[61] T. R. Gentile and C. B. Crawford. Neutron charge radius and the neutron electric form factor. *Phys. Rev. C*, 83:055203, May 2011.

[62] K. Gnanvo. Gems chambers for sbs. https://hallaweb.jlab.org/12GeV/SuperBigBite/meetings/col_2019aug5/talks/KG_SBSCollMeeting20190806.pdf, August 2019.

[63] R Guida, B Mandelli, and M Corbetta. Effects of gas mixture quality on GEM detectors operation. *J. Phys.: Conf. Ser.*, 1498:012036. 7 p, 2020.

[64] Lea Hallermann. *Analysis of GEM Properties and Development of a GEM Support Structure for the ILD Time Projection Chamber*. PhD thesis, Universität Hamburg, 2010.

[65] David Hamilton. Bigbite timing hodoscope performance update, June 2023.

[66] Frank Hartmann. Advances in tracking detectors. *Annual Review of Nuclear and Particle Science*, 61:197–221, November 2011.

[67] Leigh Harwood. The jlab 12 gev energy upgrade of cebaf. *Proceedings of the PAC2013, Pasadena, CA, USA*, MOZAA1, 2013.

[68] Brian S. Hendesrson. Results from the olympus experiment on the contribution of hard two-photon exchange to elastic electron-proton scattering. *Proceedings of Science*, September 2017.

[69] Wu Hui-Yin, Zhao Sheng-Ying, Wang Xiao-Dong, Zhang Xian-Ming, Qi Hui-Rong, Zhang Wei, Wu Ke-Yan, Hu Bi-Tao, and Zhang Yi. "discriminating cosmic muon and x-ray based on rising tim using gem detector". *Chinese Physics*, C, 2015.

[70] Karl Jakobs, editor. *Particle Detectors*, Interaction of Neutral Particles with Matter, Albert-Ludwigs-Universität Freiburg, 2015.

[71] Jr. Joe D. Wilson, Mark Stapleton, and William Hunewell. Overview of recent jlab cryogenic operations. https://www.slac.stanford.edu/econf/C0605091/present/WILSON.PDF, 2006.

[72] Lawrence Jones. Apv25-s1 user guide version 2.2. https://cds.cern.ch/record/1069892/files/cer-002725643.pdf, September 2001.

[73] M. K. Jones et al. G_{E_p}/G_{M_p} ratio by polarization transfer in $\vec{e}\ p \rightarrow e\ \vec{p}$. *Phys. Rev. Lett.*, 84:1398–1402, Feb 2000.

[74] V. Hauer K. Battes, C. Day. Outgassing behavior of different high-temperature resistant polyers. *Journal of Vacuum Science & Technology*, 36, November 2017.

[75] J. J. Kelly. Simple parametrization of nucleon form factors. *Phys. Rev. C*, 70:068202, Dec 2004.

[76] M. Killenberg, S. Lotze, J. Mnich, S. Roth, R. Schulte, B. Sobloher, W. Struczinksi, and M. Tonutti. Modelling and measurement of charge transfer in multiple gem structuers. *Nuclear Instruments and Methods in Physics Research*, 498:369 – 383, 2003.

[77] Jefferson Lab. Cryotarget training. https://userweb.jlab.org/~smithg/target/Qweak/HallACTgt_Training.pdf.

[78] Jefferson Lab. Gmn experiment. https://hallaweb.jlab.org/experiment/SBS/GMn/GMn_Website.html#pictures, 2021.

[79] Thomas Jefferson National Accelerator Laboratory. Coda cebaf on-line data acquisition user's manual. https://coda.jlab.org/drupal/content/coda-documentation, March 1993.

[80] Jeffrey Lachniet. *A high precision measurement of the neutron magnetic form factor using the CLAS detector.* PhD thesis, Carnegie Mellon University, June 2005.

[81] Michael F. L'Annunziata. *Handbook of Radioactivity Analysis.* Elsevier, 2004.

[82] Andrew J. Larkoski. *Elementary Particle Physics: An Intuitive Introduction*. Cambridge University Press, New York, NY, 1st edition, 2019.

[83] Claude Leroy and Pier-Giorgio Rancoita. *Principles of Radiation Interaction in Matter and Detection*. WORLD SCIENTIFIC, 4th edition, 2016.

[84] Hui Liu, Na-Na Ma, and Rong Wang. Nuclear fermi momenta of 2h, 27al and 56fe from an analysis of clas data. *Nuclear Physics A*, 1018:122377, 2022.

[85] Nilanga Liyanage. Solid instrumentation and gem chamber. https://www.jlab.org/conferences/hadron-china/Beijing2010/files/slides/Talks/28pm/Solid_Instrumentation_and_GEM_chambers__Nilanga_Liyanage.pdf, July 2010.

[86] Nilanga Liyanage. email. private communication/email, April 2024.

[87] Symmetry Magazine. The standard model of particle phyics. https://www.symmetrymagazine.org/standard-model/.

[88] Shivali Malhotra, Othmane Bouhali, Abdulla Al-Tamimi, and Mohammad Al-Gammal. Effect of deformation due to stretching of the gas electron multiplier (gem) foil. https://inspirehep.net/files/f7d362cc017ab26eb3f23911e8a1ca12, 2020.

[89] M. L. Marshak, P. Schmueser, J. L. Day, P. Kalbaci, A. D. Krisch, J. K. Randolph, G. J. Marmer, and L. G. Ratner. Proton-neutron elastic scattering on a deuterium target. *Phys. Rev. D*, 2:1808–1819, Nov 1970.

[90] MatWeb. G-10 fiberglass epoxy laminate sheet. https://www.matweb.com/search/DataSheet.aspx?MatGUID=8337b2d050d44da1b8a9a5e61b0d5f85&ckck=1, 2023.

[91] R D McKeown. The jefferson lab 12 gev upgrade. *Journal of Physics: Conference Series*, 032014, 2011.

[92] Ling Jian Meng. Interaction of heavy charged particles. https://courses.grainger.illinois.edu/npre441/sp2020/lctr%20notes/lecture_chapter4_2020_interaction_heavy_charged_particle.pdf, 2020.

[93] P. Mergell, Ulf-G. Meißner, and D. Drechsel. Dispersion-theoretical analysis of the nucleon electromagnetic form factors. *Nuclear Physics A*, 596(3–4):367–396, January 1996.

[94] M. Mihovilovic et al. Methods for optical calibration of the bigbite hadron spectrometer. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors, and Associated Equipment*, 2012.

[95] Paolo Musico. Mpd specifications. https://coda.jlab.org/drupal/system/files/pdfs/HardwareManual/MPD/MPD4_UserManual_r2.5.pdf, October 2018.

[96] K. Nakamura and Particle Data Group. Review of particle physics. *Journal of Physics G: Nuclear and Particle Physics*, 37(7A), 2010.

[97] Catia Petta. Hcal-j: Hadron calorimeter for the study of nucleon form factors at jefferson lab. https://indico.cern.ch/event/696724/contributions/2994185/attachments/1859127/3054775/05_Petta_CHERNE18_Petta.pdf, May 2018.

[98] Ervin B. Podgoršak. *Interactions of Charged Particles with Matter*. Radiation Physics for Medical Physicists. Springer, Berlin, Heidelberg, 2009.

[99] Daniel Q. Posin. The townsend coefficients and spark discharge. *Phys. Rev.*, 50:650–658, Oct 1936.

[100] Andrew Puckett. Straight track analysis sbs8. https://sbs.jlab.org/DocDB/0001/000157/001/str_trk_sbs8.pdf.

[101] Andrew Puckett. Docmentation of g4sbs. https://hallaweb.jlab.org/wiki/index.php/Documentation_of_g4sbs, February 2019.

[102] Andrew Puckett. Sbs software and analysis. https://indico.jlab.org/event/503/contributions/9279/attachments/7518/10463/Puckett_SBS_software.pdf, February 2022.

[103] Andrew Puckett. Sbs software and analysis. https://indico.jlab.org/event/503/contributions/9279/attachments/7518/10463/Puckett_SBS_software.pdf, February 2022.

[104] Andrew Puckett. Assessing uniformity of hcal proton detection efficiency, sbs-4,8,9. https://sbs.jlab.org/DocDB/0004/000482/002/HCALprotonEfficiencyUniformity.pdf, December 2023.

[105] Andrew Puckett. Sbs-gmn analysis overview and status (note "results"). https://indico.jlab.org/event/721/contributions/13211/attachments/10053/14936/PuckettGMN_talk.pdf, July 2023.

[106] I. A. Qattan et al. Precision rosenbluth measurement of the proton elastic form factors. *Phys. Rev. Lett.*, 94:142301, Apr 2005.

[107] B. Quinn, B. Wojtsekhowski, R. Gilman, and E-09-019 collaboration. Update of e-09-019 to the jefferson lab program advisory committee (pac35); precision measurement of the neutron magnetic form factor up to $q^2 = 13.5$ (gev/c)^2 by the ratio method. https://hallaweb.jlab.org/12GeV/SuperBigBite/documents/ERR2017/update10.pdf.

[108] Anuruddha Rathnayake. Time deconvolution principle. https://sbs.jlab.org/DocDB/0002/000223/001/Time%20Deconvolution%20Principles.pdf, May 2022.

[109] Dipak Rimal. *Proton Form FActor Puzzle and the CEBAF Large Acceptance Spectrometer (CLAS) Two-Photon Exchange Experiment.* PhD thesis, Florida International University, 2014.

[110] Leszek Ropelewski. Gem at cern. http://leszek.web.cern.ch/GEM_at_CERN_seminar.pdf, October 2014.

[111] E. Rutherford. The scattering of α and β particles by matter and the structure of the atom. *Philos. Mag*, 6, 1911.

[112] Maria Satnik. Sbs grinch gas cherenkov detector. https://www.jlab.org/sites/default/files/user-liaison/jluo2022/poster_2022_JLUO_SATNIK.pdf, 2022.

[113] Fabio Sauli. "recent topics on gaseous detectors. *National Instruments and Methods in Physics Research A*, 2010.

[114] Fabio Sauli. The gas electron multiplier (gem): Operating principles and applications. *Nuclear Instruments and Methods in Physics Research*, 805:2 – 24, January 2016.

[115] Fabio Sauli. Gem: Gas electron multiplier development and applications. http://www.physics.smu.edu/yejb/teaching/7360_2018s/notes/Lecture_Notes/sauli_32002.pdf, 2018.

[116] Fabio Sauli, Steffen Kappler, and Leszek Ropelewski. Electron collection and ion feedback in gem-based detectors. *IEEE Transactions on Nuclear Science*, December 2002.

[117] O. Schäfer. *Ein Monitorsystem für Gasbasierte Detektoren am International Linear Collider (ILC) - in german.* PhD thesis, Universität Rostock, 2006.

[118] Sebastian Seeds. Hcal update - 10.05.21. private communication/shared slides, October 2021.

[119] Sebastian Seeds. Measurement of the contribution of two photon exchange to the electron-neutron elastic scattering cross section (ntpe) using the suepr bigbite spectrometer. https://www.jlab.org/sites/default/files/user-liaison/jluo2022/JLUO%20Seeds%20nTPE.pdf, April 2022.

[120] Sebastian Seeds. The sbs hadronic calorimeter. https://sbs.jlab.org/DocDB/0002/000298/001/Student%20Analysis%20HCal%20Talk.pdf, August 2022.

[121] Sebastian Seeds. Hcal energy calibration quality checks. https://sbs.jlab.org/DocDB/0004/000441/012/HCal_Energy_Calibration_SBS4_LH2_First_HV_Set_Pass2.pdf, September 2023.

[122] Sebastian Seeds. Sbs hcal clustering: Scoring methods. https://sbs.jlab.org/DocDB/0004/000445/001/SBS%20HCal%20Clustering.pdf, September 2023.

[123] Emilio G. Segrè. *Nuclei and Particles: An Introduction To Nuclear and Subnuclear Physics*. W.A. Benjamin, 2nd edition, 1977.

[124] M. R. Silva. Ionizing radiation detectors. https://pdfs.semanticscholar.org/9f03/2bd6c7d4d013173cfc1ec7ffa1d88450c53d.pdf?_ga=2.267328461.2138071203.1631042163-2100087839.1631042163, 2015.

[125] Frank Simon. *Commissioning of the GEM Detectors in the COMPASS Experiment*. PhD thesis, Technish Universität München, November 2001.

[126] B. Sobloher. *Simulationsstudien zu GEM-Folien Für die Auslese einer TPC - in german*. PhD thesis, RWTH Aachen, 2002.

[127] Arun Tadepalli. Bigbite calorimeter status. https://indico.jlab.org/event/430/contributions/7812/attachments/6473/8673/sbs_collaboration_meeting_feb17.pdf, February 2023.

[128] Arun Tadepalli, Bogdan Wojtsekhowski, Mark Jones, and Eric Fuchey. Bigbite shower and preshower current status. , January 2021.

[129] M. Tanabashi et al. π^0 mass. *Physical Review D*, 98, 2018.

[130] M. Tanabashi, K. Hagiwara, K. Hikasa, K. Nakamura, Y. Sumino, F. Takahashi, J. Tanaka, K. Agashe, G. Aielli, C. Amsler, M. Antonelli, D. M. Asner, H. Baer, Sw. Banerjee, R. M. Barnett, T. Basaglia, C. W. Bauer, J. J. Beatty, and etc Belousov. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.

[131] Werner Witzeling Thomas Lohse. The time projection chamber. *Advanced Series on Directions in High Energy Physics*, 9:81–155, 1992.

[132] Mark Thomson. Handout 5: Electron-proton elastic scattering. https://www.hep.phy.cam.ac.uk/~thomson/partIIIparticles/handouts/Handout_5_2011.pdf, 2011.

[133] Mark Thomson. *Modern Particle Physics*. Cambridge University Press, 5th edition, 2019.

[134] Dezső Varga. Gain fluctuations in gem detectors. https://indico.cern.ch/event/323839/contributions/751969/attachments/627137/862974/gainfluctuations.pdf, 2014.

[135] Adrian Vogel. *Beam-Induced Backgrounds in Detectors at the ILC*. PhD thesis, Universitä Hamburg, 2008.

[136] M. Weber. *Simulationen zur Verwendung von GEM-Folien als Auslesemodul einer TPC - in german*. PhD thesis, RWTH Aachen, 2003.

[137] Zeke Wertz. Gain match update & summary. https://sbs.jlab.org/DocDB/0004/000408/001/GainMatch_Update.pdf, July 2023.

[138] Robin Wines. Magnet and infrastructure. https://hallaweb.jlab.org/12GeV/SuperBigBite/Dry_Run/Slides/Magnet-and-Infrastructure.pdf, October 2013.

[139] Bogdan Wojtsekhowski. Super bigbite experiments. https://indico.jlab.org/event/721/contributions/13209/attachments/10036/14933/SBS%20experiments.pdf, July 2023.

[140] Zhihon Ye, John Arrington, Richard J. Hill, and Gabriel Lee. Proton and neutron electromagnetic form factors and uncertainties. *Physics Letters B*, 777:8–15, 2018.

[141] Zhihong Ye, John Arrington, Richard J. Hill, and Gabriel Lee. Proton and neutron electromagnetic form factors and uncertainties. *Physics Letters B*, 777:8–15, 2018.