

©Copyright 2019  
Aaron T. Fienberg

# Measuring the Precession Frequency in the E989 Muon $g - 2$ Experiment

Aaron T. Fienberg

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

Doctor of Philosophy

University of Washington

2019

Reading Committee:

David Hertzog, Chair

Blayne Heckel

Ann Nelson

Program Authorized to Offer Degree:  
Physics

University of Washington

## **Abstract**

Measuring the Precession Frequency in the E989 Muon  $g - 2$  Experiment

Aaron T. Fienberg

Chair of the Supervisory Committee:  
Director of CENPA, Professor David Hertzog  
Physics

The E989 Muon  $g - 2$  Experiment aims to measure the anomalous magnetic moment of the muon,  $a_\mu$ , to an unprecedented precision of 140 parts per billion (ppb). There currently stands a greater than  $3\sigma$  discrepancy between the best measurement of  $a_\mu$  and its theoretical value predicted using the Standard Model. The E989 experiment seeks to either resolve or confirm this discrepancy, which is suggestive of new physics interactions within reach of many contemporary experiments. To achieve the E989 target precision, the anomalous precession frequency,  $\omega_a$ , of muons in a magnetic storage ring must be determined with a systematic uncertainty below 70 ppb. This frequency is imprinted on the time-dependent energy distribution of decay positrons observed by 24 electromagnetic calorimeters. These calorimeters feature a novel design optimized expressly for the stringent demands of the  $\omega_a$  measurement. This dissertation outlines the motivation for and measurement principles behind E989, discusses the requirements, prototyping, testing, commissioning, and operations of the electromagnetic calorimeters, and presents a preliminary, blinded analysis of data collected in the spring of 2018.

# TABLE OF CONTENTS

	Page
List of Figures . . . . .	v
Chapter 1: Introduction . . . . .	1
1.1 Magnetic Dipoles . . . . .	2
1.2 Spin Precession . . . . .	3
1.3 Magnetic Dipole Moments of Spin-1/2 Particles . . . . .	4
1.4 Muon $g - 2$ in the Standard Model . . . . .	9
1.4.1 Quantum Electrodynamics . . . . .	9
1.4.2 Electroweak . . . . .	11
1.4.3 Hadronic Vacuum Polarization . . . . .	12
1.4.4 Hadronic Light-by-Light Scattering . . . . .	14
1.4.5 Combined Standard Model $a_\mu$ Prediction . . . . .	16
1.5 The Brookhaven Muon $g - 2$ Experiment . . . . .	17
1.6 New Physics Interpretations of the Brookhaven Result . . . . .	19
Chapter 2: The E989 Muon $g - 2$ Experiment . . . . .	23
2.1 Measurement Principle . . . . .	23
2.2 Storage Ring and Electrostatic Quadrupoles . . . . .	25
2.3 Muon Decay . . . . .	27
2.4 Beam Injection . . . . .	31
2.5 Magnetic Field Measurement . . . . .	36
2.6 Anomalous Precession Frequency Measurement . . . . .	39
2.7 Required Running Time . . . . .	47
Chapter 3: Systematic Uncertainties of the Precession Frequency Measurement . . .	50
3.1 Decay to Detection . . . . .	51
3.2 Precession Frequency Bias from a Changing Phase . . . . .	53



3.3	Systematic Effects from Beam Dynamics . . . . .	55
3.3.1	Muon Dynamics in the Storage Ring . . . . .	55
3.3.2	Differential Decay . . . . .	60
3.3.3	Muon Losses . . . . .	63
3.3.4	Beam Debunching and Fast Rotation . . . . .	63
3.3.5	Coherent Betatron Oscillations . . . . .	67
3.3.6	Electric Field and Pitch Corrections . . . . .	71
3.4	Systematic Effects from Detectors . . . . .	76
3.4.1	Pileup . . . . .	77
3.4.2	Gain Changes . . . . .	79
3.5	Estimation of Detector-Based Systematic Biases . . . . .	80
3.5.1	Estimated Size of the Bias from Gain Perturbations . . . . .	82
3.5.2	Estimated Size of the Bias from Pileup . . . . .	83
3.6	Summary of the Precession Frequency Error Budget . . . . .	87
Chapter 4:	The E989 Calorimeter . . . . .	89
4.1	Hardware . . . . .	89
4.1.1	Lead Fluoride . . . . .	90
4.1.2	Silicon Photomultipliers . . . . .	92
4.1.3	Laser Calibration System . . . . .	98
4.2	Reconstruction Software . . . . .	104
4.2.1	Input to Output . . . . .	105
4.2.2	Pedestal Correction . . . . .	107
4.2.3	Pulse Shapes and Templates . . . . .	108
4.2.4	Template Fitting . . . . .	113
4.2.5	Energy Calibration . . . . .	118
4.2.6	Timing Corrections . . . . .	121
4.2.7	Clustering . . . . .	126
4.3	Beam Tests . . . . .	130
4.3.1	November 2013 SLAC Test Beam Experiment . . . . .	131
4.3.2	July 2014 SLAC Test Beam Experiment . . . . .	132
4.3.3	March 2016 Frascati Test Beam Experiment . . . . .	139
4.3.4	June 2016 SLAC Test Beam Experiment . . . . .	142

Chapter 5:	Commissioning	146
5.1	Summer 2017 Commissioning Run	146
5.1.1	Storage Diagnostics	148
5.1.2	First $\omega_a$ Measurement	151
5.1.3	Vertical Beam Position and Radial Fields	152
5.1.4	Lost Proton CBO Studies	153
5.1.5	Final Storage Rates	157
5.2	2017-2018 Commissioning	157
Chapter 6:	E989 Run 1 Precession Frequency Analysis	161
6.1	The 60-Hour Dataset at a Glance	161
6.2	Data Quality Conditions	163
6.3	Blinding	165
6.4	Five-Parameter T-Method Fit to the 60-Hour Dataset	166
6.5	Corrections to the Five-Parameter Model	168
6.5.1	Beam Oscillations	168
6.5.2	Muon Losses	173
6.5.3	The Pileup Correction	175
6.5.4	The Gain Correction	181
6.6	Extraction of $\omega_a$	183
6.6.1	T-Method Fit to the 60-Hour Dataset	183
6.6.2	Energy-Binned Analysis of the 60-Hour Dataset	187
6.6.3	Asymmetry-Weighted Analysis of the 60-Hour Dataset	191
6.6.4	Per-Calorimeter Fits	196
6.7	Systematic Uncertainty Assessment	197
6.7.1	Systematic Uncertainty from Beam Oscillations	197
6.7.2	Systematic Uncertainty from Pileup	202
6.7.3	Variation of the Artificial Dead Time	206
6.7.4	Systematic Uncertainty from Changing Detector Gains	208
6.7.5	Systematic Uncertainty from Muon Losses	216
6.7.6	Systematic Uncertainty from Differential Decay	220
6.7.7	Systematic Uncertainty from the Electric Field and Pitch Corrections	220
6.7.8	Summary of the Systematic Uncertainties	221

6.8 Outlook . . . . .	222
Bibliography . . . . .	224
Appendix A: Pileup Corrected Bin Uncertainties . . . . .	231
A.1 Pileup Correction Overview . . . . .	231
A.2 Uncertainty of a Single Energy Bin . . . . .	233
A.3 T-Method Uncertainty Enhancements . . . . .	234
A.4 A-Weighted Uncertainty Enhancements . . . . .	238
A.5 Effect on The Extracted $R$ Value . . . . .	238
A.6 Algebraic Derivation of Equation A.11 . . . . .	239
Appendix B: Correlation Matrices from the T-Method and A-Weighted Fits . . . . .	241

# LIST OF FIGURES

Figure Number	Page
1.1 General interaction between a fermion and an off-shell photon . . . . .	6
1.2 Leading-order and one-loop QED contributions to $g$ . . . . .	8
1.3 Standard Model Feynman diagrams contributing to muon $g - 2$ . . . . .	10
1.4 Leading-order hadronic vacuum polarization . . . . .	12
1.5 Hadronic $R$ -ratio data used to calculate $a_\mu^{\text{Had,VP}}$ . . . . .	14
1.6 Hadronic light-by-light scattering . . . . .	15
1.7 Visual decomposition of $a_\mu^{\text{SM}}$ . . . . .	16
1.8 Possible new physics contribution to $a_\mu$ . . . . .	20
1.9 Parameter space of the dark photon model relevant to muon $g - 2$ . . . . .	22
2.1 The E989 storage ring . . . . .	27
2.2 Equipotential contours from the electrostatic quadrupoles . . . . .	28
2.3 Muon decay distribution in the rest frame . . . . .	29
2.4 Muon decay distribution in the lab frame . . . . .	31
2.5 The muon campus beamline . . . . .	33
2.6 Diagrammatic illustration of injection into the storage ring . . . . .	35
2.7 Azimuthally averaged magnetic field contours . . . . .	37
2.8 Illustration of muon decay in the storage ring . . . . .	40
2.9 Illustration of the T-Method for $\omega_a$ analysis . . . . .	42
2.10 $\omega_a$ precision versus threshold . . . . .	44
3.1 Momentum dependent drift time and acceptance . . . . .	52
3.2 Beam debunching in the $g - 2$ storage ring . . . . .	66
3.3 Beam oscillations observed by detectors at fixed positions . . . . .	70
3.4 The electric field correction with an off-center beam . . . . .	76
3.5 A calorimeter pileup pulse pair . . . . .	78
3.6 Precession frequency bias with different gain perturbations . . . . .	84
3.7 Precession frequency bias with different detector dead times . . . . .	87

4.1	Model of the E989 calorimeter . . . . .	90
4.2	Locations of the E989 calorimeters . . . . .	91
4.3	Spectrum of detected photons in the E989 calorimeter . . . . .	93
4.4	SiPM PDE, pulse-integral, and pulse-integral per pe versus overvoltage . . . .	96
4.5	The E989 SiPM readout board . . . . .	98
4.6	Schematic of the laser calibration system . . . . .	99
4.7	Example photostatistics calibration . . . . .	101
4.8	In-fill laser pulses for measuring calorimeter gain variations . . . . .	103
4.9	Example digitizer island . . . . .	107
4.10	Example pulse templates . . . . .	109
4.11	Illustration of the template building procedure . . . . .	112
4.12	Tests of SiPM pulse shape stability . . . . .	114
4.13	Template fits to island waveforms . . . . .	118
4.14	Time shift versus pulse time separation in double pulse template fits . . . . .	119
4.15	A possible lost muon trajectory . . . . .	120
4.16	Lost muon energy peak before and after energy calibration . . . . .	122
4.17	Characteristic pulse from the T0 detector . . . . .	123
4.18	Hit time differences between neighboring calorimeter channels . . . . .	125
4.19	An island with three identified clusters . . . . .	128
4.20	November 2013 calorimeter prototype . . . . .	131
4.21	July 2014 calorimeter prototype . . . . .	133
4.22	Setup of the July 2014 test beam experiment . . . . .	135
4.23	Energy response from the July 2014 Test Beam . . . . .	136
4.24	Calorimeter energy resolution versus energy . . . . .	137
4.25	Pulse shape dependence on impact position for different crystal wrappings . .	139
4.26	March 2016 calorimeter prototype . . . . .	140
4.27	June 2016 production calorimeter at SLAC . . . . .	143
4.28	Crystal by crystal response to a 3 GeV electron . . . . .	144
4.29	Gain corrections at the SLAC test beam . . . . .	145
5.1	Online display of calorimeter hit times with early quadrupole ramp down . . .	150
5.2	The first E989 kicker timing and inflector current scans . . . . .	151
5.3	The first E989 wiggle plot . . . . .	152

5.4	Surface coil scan current scan to center the stored beam . . . . .	154
5.5	Lost proton time spectrum from the 2017 commissioning run . . . . .	155
5.6	Accumulated statistics versus time throughout E989 Physics Run 1 . . . . .	158
6.1	The equilibrium radius distribution of stored muons in the 60-Hour Dataset .	163
6.2	Calorimeter energy and time spectra from the 60-Hour Dataset . . . . .	164
6.3	$\omega_a$ precision versus energy threshold . . . . .	167
6.4	Five-parameter T-Method fit to the 60-Hour Dataset . . . . .	169
6.5	Radial distribution of the stored beam as seen by the trackers . . . . .	171
6.6	CBO envelope and frequency shift in the 60-Hour Dataset . . . . .	172
6.7	Muon losses measured in the 60-Hour Dataset . . . . .	175
6.8	Energy dependence of the pileup correction in the 60-Hour Dataset . . . . .	177
6.9	Energy dependence of the pileup correction in the 60-Hour Dataset . . . . .	178
6.10	Representative in-fill gain measurement from 60-Hour Dataset . . . . .	181
6.11	Full T-Method fit to the 60-Hour Dataset . . . . .	183
6.12	Asymmetry versus start time with and without the pileup correction . . . . .	186
6.13	$R$ and $N_0$ versus fit start time . . . . .	187
6.14	$\chi^2/ndf$ , $N_0$ , and $A$ versus positron energy . . . . .	189
6.15	Blinded $\omega_a$ and $\phi$ versus positron energy . . . . .	190
6.16	$\tau$ and $K_{loss}$ versus positron energy . . . . .	191
6.17	Weighting of positron energies in the T-Method and A-Weighted analyses . . .	192
6.18	A-Weighted fit to the 60-Hour Dataset . . . . .	193
6.19	$\chi^2/ndf$ , $R$ , $\tau$ , and $K_{loss}$ by calorimeter . . . . .	198
6.20	Results of the pileup multiplier scan . . . . .	202
6.21	Pileup contaminations of the T-Method and A-Weighted histograms . . . . .	206
6.22	Time separation of consecutive clusters with different artificial dead times . .	208
6.23	$\omega_a$ analysis with different artificial dead times . . . . .	209
6.24	The effect of a small gain perturbation on the calorimeter energy spectrum . .	212
6.25	$K_{loss}$ versus positron energy with and without the artificial gain correction . .	213
6.26	$g - 2$ asymmetry versus start time with differing artificial gain corrections . .	214
6.27	Fractional muon losses during the 60-Hour Dataset . . . . .	217
6.28	Truncated normal distribution . . . . .	218
A.1	Pileup uncertainty enhancement factors by energy . . . . .	235

A.2	Energy pairs relevant to pileup in the T-Method analysis . . . . .	236
A.3	Pileup uncertainty factors in the T-Method . . . . .	237
A.4	Pileup uncertainty factors in the A-Weighted analysis . . . . .	239
B.1	Correlation matrix from the T-Method fit . . . . .	242
B.2	Correlation matrix from the A-Weighted fit . . . . .	243

## **DEDICATION**

to Hillary and Freja



## ACKNOWLEDGMENTS

First and foremost, I wish to recognize the support of my mentors and peers from the muon  $g - 2$  group at the University of Washington: David Hertzog, Pete Alonzi, Jarek Kaspar, Kim Siang Khaw, Peter Kammel, Erik Swanson, Nathan Froemming, Matthias Smith, Rachel Osofsky, Jason Hempstead, Brynn MacCoy, Hannah Binney, and Josh LaBounty. The opportunity to work alongside these individuals was surely the most valuable component of my graduate education; our discussions were always thoroughly edifying.

Also invaluable were the relationships I formed with scientists from the E989 collaboration. I am particularly grateful to Chris Polly for his mentorship during my year-long residence at FNAL, to Graziano Venanzoni for hosting me at INFN Frascati and for his continual encouragement, and to Lawrence Gibbons, David Sweigart, Antoine Chapelain, Antonio Anastasi, Anna Driutti, James Mott, Nick Kinnaird, and Joe Price for their perennial availability to collaborate on difficult problems and to provide thoughtful feedback.

This dissertation could not have been written without the intellectual foundation and confidence I received from my undergraduate physics professors at Earlham College: Ray Hively, John Howell, and Michael Lerner.

Finally, I would like to thank my mom, Susan, my dad, Rick, my brothers, John and Luke, and my partner, Hillary, for the parts they played in my becoming who I am.

## Chapter 1

### INTRODUCTION

These pages describe many years worth of effort toward the measurement of a single quantity: muon  $g - 2$ . Over one hundred scientists—experimentalists and theorists both—have contributed to this effort, along with dozens of engineers and technicians. The aspects of the  $g - 2$  measurement emphasized here are those to which the author has directly and significantly contributed.

Measuring muon  $g - 2$ , or equivalently the muon’s anomalous magnetic moment,  $a_\mu \equiv \frac{g_\mu - 2}{2}$ , requires the precise determination of both a magnetic field strength and a spin precession frequency. Each is equally important, though the majority of this document pertains to the precession frequency measurement. Chapter 1, this chapter, introduces the concepts of magnetic dipoles and spin precession, discusses the current state of experiment and theory relevant to muon  $g - 2$ , and motivates a new measurement with improved precision. Chapter 2 then introduces the E989 Muon  $g - 2$  Experiment at Fermilab, outlining the physical principles that enable the experiment. The major systematic effects that complicate the E989 precession frequency measurement are discussed in Chapter 3. Next, Chapter 4 presents the E989 calorimeter and characterizes its performance. The E989 calorimeter is designed expressly to limit many of the systematic effects identified in Chapter 3. The E989 commissioning phase is briefly discussed in Chapter 5, and finally Chapter 6 presents preliminary, blinded analysis of data from E989 Run 1, collected in the spring of 2018.

### 1.1 Magnetic Dipoles

Magnetic dipoles are objects that—when viewed from a sufficient distance—produce electromagnetic vector potentials of the form [1]

$$\vec{A}(\vec{x}) = \frac{\mu_0}{4\pi} \frac{\vec{\mu} \times \vec{x}}{|\vec{x}|^3}. \quad (1.1)$$

$\vec{\mu}$ , the magnetic dipole moment, is a quantity that conveys the orientation and strength of a magnetic dipole. Taking the curl of the above vector potential yields a characteristic magnetic field  $\vec{B}$  with position dependence identical to that of the electric field produced by an electric dipole. Magnetic dipoles can be thought of as objects that create magnetic fields with this characteristic form. Stronger fields come from objects with larger magnetic dipole moments.

Localized electric current distributions can generate magnetic dipole moments. The magnetic dipole moment of a classical current distribution with current density  $\vec{J}$  is [1]

$$\vec{\mu} = \frac{1}{2} \int \vec{x}' \times \vec{J}(\vec{x}') d^3x'. \quad (1.2)$$

Examples of classical objects with nonvanishing magnetic dipole moments include current carrying loops of wire and rotating charged spheres. In cases where the current distribution comprises a population of charged particles with identical charge and mass densities, Equation 1.2 becomes [1]

$$\vec{\mu} = \frac{Q}{2M} \vec{L}, \quad (1.3)$$

where  $M$  is the total mass of the population,  $Q$  is the total charge, and  $\vec{L}$  is the net angular momentum. Thus, magnetic dipole moments have the dimensions of a charge-to-mass ratio multiplied by an angular momentum.

What is of most importance to the measurement described in this dissertation is not the magnetic field created by a magnetic dipole, but rather the effect an external magnetic field has *on* a magnetic dipole. The interaction between a magnetic dipole and an external magnetic field  $\vec{B}$  is encoded in the potential energy expression [1]

$$U = -\vec{\mu} \cdot \vec{B}. \quad (1.4)$$

The configuration's potential energy depends on the magnetic dipole's orientation relative to the external magnetic field. Alignment of the magnetic dipole moment with the field yields the lowest energy state. Despite its simplicity, Equation 1.4 has a number of important implications. For example, magnetic dipoles tend to align with applied magnetic fields, and they experience forces when subjected to magnetic field gradients. Equation 1.4 applies to both classical systems and quantum mechanical systems.

## 1.2 Spin Precession

Consider a particle at rest with spin  $\vec{S}$  and magnetic dipole moment  $\vec{\mu}$ . With no other directions defined,  $\vec{\mu}$  must be parallel or antiparallel to  $\vec{S}$ , expressible as

$$\vec{\mu} = \gamma \vec{S}, \quad (1.5)$$

where  $\gamma$ , the ratio of the particle's magnetic moment to its angular momentum, is called the gyromagnetic ratio [2]. The classical systems described by Equation 1.3 have gyromagnetic ratios equal to  $Q/(2M)$ .

The Hamiltonian describing the interaction of this particle with an external magnetic field can be determined from Equation 1.4. It is

$$H = -\gamma \vec{S} \cdot \vec{B}. \quad (1.6)$$

The dynamics of this system will be considered from a quantum mechanical perspective. Adopting the notation from J.J. Sakurai's *Modern Quantum Mechanics* [3], the time-evolution operator of a system with a time-independent Hamiltonian such as the one above is

$$\mathcal{U}(t, t_0) = \exp \left[ \frac{-iH(t - t_0)}{\hbar} \right]. \quad (1.7)$$

Setting  $t_0$  to 0 and considering the specific Hamiltonian from Equation 1.6,

$$\mathcal{U}(t, 0) = \exp \left[ \frac{i\gamma \vec{S} \cdot \vec{B}t}{\hbar} \right]. \quad (1.8)$$

In the coordinate system where  $\vec{B}$  points along the z-axis, this becomes

$$\mathcal{U}(t, 0) = \exp\left[\frac{i\gamma S_z B t}{\hbar}\right]. \quad (1.9)$$

The above equation can be compared to the operator that generates rotations of angle  $\phi$  about the z-axis,  $\mathcal{D}_z(\phi)$  [3],

$$\mathcal{D}_z(\phi) = \exp\left[\frac{-iS_z \phi}{\hbar}\right]. \quad (1.10)$$

Equation 1.9 is equivalent to Equation 1.10, except  $\phi$  has been replaced with  $-\gamma B t$ . Thus, time evolution of the system in question is equivalent to spin rotation about the axis of the external magnetic field with angular frequency

$$\vec{\omega} = -\gamma \vec{B}. \quad (1.11)$$

The phenomenon of spin rotation in external magnetic fields is called spin precession. Spin precession is a direct consequence of the potential energy expression given in Equation 1.4. The precession frequency is directly proportional to both the magnitude of the external magnetic field,  $B$ , and the gyromagnetic ratio,  $\gamma$ , of the precessing particle. If  $B$  is known, then measuring the precession frequency provides  $\gamma$ . Alternatively, if  $\gamma$  is known, then measuring the precession frequency provides  $B$ . Chapter 2 describes how the E989 experiment leverages these principles to measure muon  $g - 2$ .

### 1.3 Magnetic Dipole Moments of Spin-1/2 Particles

The observed gyromagnetic ratios of subatomic particles with intrinsic angular momentum, such as electrons, protons, and muons, are not consistent with Equation 1.3. Neither classical electrodynamics nor non-relativistic quantum mechanics is capable of providing accurate predictions of these particles' magnetic moments [1, 4]. The deviation of a subatomic particle's magnetic moment from the classical formula can be quantified with a dimensionless g-factor as follows:

$$\vec{\mu} = \pm g \cdot \frac{e}{2m} \vec{S}. \quad (1.12)$$

In the above equation,  $\vec{\mu}$  is the particle's magnetic moment,  $g$  is the aforementioned g-factor,  $e$  is the elementary charge,  $\vec{S}$  is the particle's spin, and  $m$  is the particle's mass. The choice of positive or negative sign depends on whether the subatomic particle in question is positively or negatively charged.

In general, different particles have different g-factors. The electron and muon g-factors,  $g_e$  and  $g_\mu$ , are each about 2, whereas  $g_p$ , the proton g-factor, is about 5.6 [5]. The focus here will be on structureless spin-1/2 particles, particularly electrons and muons.

Paul Dirac's *The Quantum Theory of the Electron*, published in 1928, presented a relativistic wave equation describing the dynamics of electrons subjected to external electromagnetic fields [6]. Dirac's equation famously explained the contemporaneous experimental observations suggesting that  $g_e = 2$ . For a bit of historical perspective, see Lee Roberts's *Historical Introduction to Electric and Magnetic Moments* [7] or the PhD dissertation of Charles C. Polly [8]. Only theories respecting the principles of both quantum mechanics and special relativity have made accurate predictions regarding the magnetic dipole moments of elementary particles.

With modern notation and treatment taken from Matthew Schwartz's *Quantum Field Theory and the Standard Model* [4], the Dirac equation in the presence of external electromagnetic fields reads

$$(i\cancel{\partial} - e\cancel{A} - m)\psi = 0, \quad (1.13)$$

implying that

$$\left[ (i\partial_\mu - eA_\mu)^2 - \frac{e}{2}F_{\mu\nu}\sigma^{\mu\nu} - m^2 \right] \psi = 0. \quad (1.14)$$

The term  $-\frac{e}{2}F_{\mu\nu}\sigma^{\mu\nu}$  corresponds to a magnetic dipole interaction with the external field. In the non-relativistic limit, the above equation yields the Hamiltonian

$$H = \frac{\vec{p}^2}{2m} + V(r) + \frac{e}{2m}\vec{B} \cdot (\vec{L} + 2\vec{S}). \quad (1.15)$$

The  $g_e = 2$  prediction—evident in the  $\frac{2e}{2m}\vec{B} \cdot \vec{S}$  interaction—comes from the  $-\frac{e}{2}F_{\mu\nu}\sigma^{\mu\nu}$  contribution to Equation 1.14.

At present,  $g_e$  has been measured to 0.28 parts per trillion [9]. It deviates from the prediction of the Dirac equation by approximately 0.1%. This deviation is not a new discovery; indeed, by the late 1940s there was evidence that  $g_e$  is not exactly equal to 2 [4]. Accordingly, the anomalous magnetic moment of an elementary particle,  $a$ , is defined as

$$a \equiv \frac{g - 2}{2}. \quad (1.16)$$

It quantifies the deviation of  $g$  from the Dirac equation's prediction of 2.

The existence of nonvanishing anomalous magnetic moments can be explained using the framework of quantum field theory, in which a fermion's magnetic moment can be understood in terms of the interaction between the fermion field and the photon field, with the leading-order  $g = 2$  contribution coming from the tree-level diagram and the anomalous part coming from higher-order loop corrections.

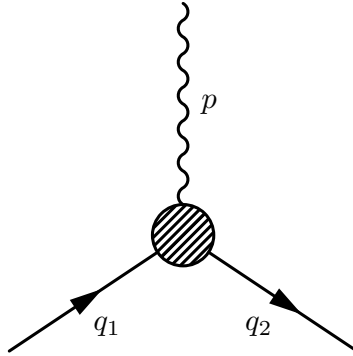


Figure 1.1: A Feynman diagram depicting the general interaction between a fermion and an off-shell photon. The off-shell photon represents an external electromagnetic field.

A Feynman diagram depicting the general interaction between a fermion and an external electromagnetic field is shown in Figure 1.1. Neglecting terms that violate parity symmetry, the matrix element describing this interaction can be expressed in terms of two independent form factors as follows [4]

$$i\mathcal{M}^\mu = (-ie) \bar{u}(q_2) \left[ F_1 \left( \frac{p^2}{m^2} \right) \gamma^\mu + \frac{i\sigma^{\mu\nu}}{2m} p_\nu F_2 \left( \frac{p^2}{m^2} \right) \right] u(q_1). \quad (1.17)$$

The second term is spin-dependent, and it has a form reminiscent of the one in Equation 1.14 that gave rise to the  $\vec{S} \cdot \vec{B}$  interaction. It is this second term that generates the anomalous part of the magnetic moment. Specifically,  $a$  is directly related to  $F_2$  evaluated at  $p = 0$ ,

$$a = F_2(0). \quad (1.18)$$

Rigorous derivations of this relation are available in most introductory quantum field theory textbooks, for example [4, 10, 11]. It is important to note that the  $F_2$  interaction, with two fermion fields, a photon field, and a factor of the momentum transfer, has the structure of a dimension-5 operator. Thus, it cannot appear in the Lagrangian density of a renormalizable field theory [4]. The anomalous part of the magnetic dipole moment is therefore generated purely by loop corrections to the diagram in Figure 1.1 [12]. In a renormalizable field theory, for instance QED,  $a$  is *predictable* based on parameters such as coupling constants and particle masses.

Any loop correction to the QED vertex can correct  $a$ . The dominant contributions to the electron and muon anomalous moments,  $a_e$  and  $a_\mu$ , come from the one-loop diagram shown in Figure 1.2. It is the only order  $\alpha$  correction to  $a$  and provides a positive contribution of

$$a = \frac{\alpha}{2\pi}. \quad (1.19)$$

This contribution was presented first by Julian Schwinger in 1948 [4]. Derivations of the leading-order QED contribution to  $a$  are available in the same introductory quantum field theory textbooks cited for Equation 1.17.

Equation 1.19 depends only on  $\alpha$ , the fine-structure constant. Thus, the leading-order contributions to  $a_\mu$  and  $a_e$  are the same. This is not true for higher order QED contributions or for contributions coming from electroweak and hadronic interactions [5, 13]. Beyond the one-loop QED diagram, the differing masses of muons and electrons lead to different sized corrections at order  $\alpha^2$  and beyond. Electroweak contributions come with leading factors of  $m_e^2/m_W^2$  in the electron case and  $m_\mu^2/m_W^2$  in the muon case. Thus, the electroweak contributions to  $a_\mu$  are approximately  $m_\mu^2/m_e^2 \approx 43,000$  times larger than they are to  $a_e$ . In



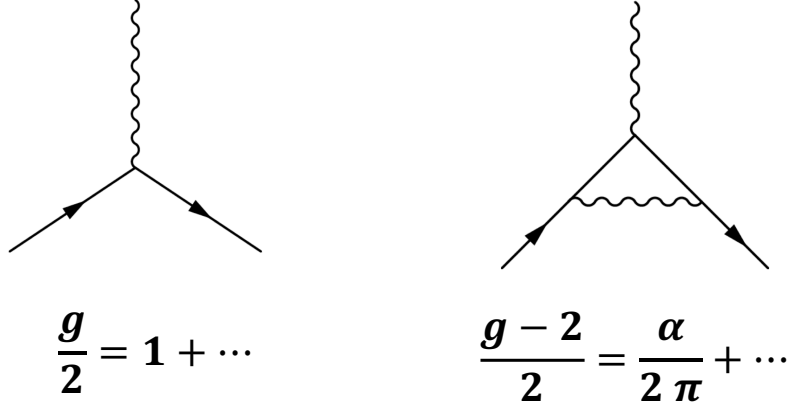


Figure 1.2: Leading-order and one-loop QED contributions to  $g$ . The one-loop diagram on the right provides an  $\alpha/\pi$  positive contribution to  $g$ , which is an approximately 0.1% correction to the leading-order value of 2.

general, the effects of interactions at high energy scales on the anomalous magnetic moment of a lepton are proportional to the lepton mass squared, [12]

$$\delta a_l \propto \frac{m_l^2}{\Lambda^2}. \quad (1.20)$$

In the above equation,  $\delta a_l$  represents a contribution to the anomalous magnetic moment of a lepton from an interaction at energy scale  $\Lambda$ , and  $m_l$  is the mass of the lepton.

Although  $a_e$  can be measured more precisely than  $a_\mu$ , the general pattern described by Equation 1.20 suggests that  $a_\mu$  is 43,000 times more sensitive to new physics interactions at high energy scales. This gain in sensitivity is approximately 20 times larger than the uncertainty of the best current  $a_\mu$  measurement over the uncertainty of the best current  $a_e$  measurement [13].

The electron and muon anomalous magnetic moments,  $a_e$  and  $a_\mu$ , are quantities that can be precisely predicted using the Standard Model. Measurements of  $a_e$  allow for precise tests of QED, or, alternatively, precise measurements of  $\alpha$ , while measurements of  $a_\mu$  are more sensitive to electroweak and hadronic interactions and place more stringent constraints on new physics models [12]. The remainder of this chapter will focus solely on  $a_\mu$ , first out-

lining the Standard Model prediction and then discussing the current state of experimental measurement.

#### 1.4 Muon $g - 2$ in the Standard Model

The Standard Model prediction for the anomalous magnetic moment of the muon is traditionally expressed as the sum of three components [5]

$$a_\mu^{\text{SM}} = a_\mu^{\text{QED}} + a_\mu^{\text{EW}} + a_\mu^{\text{Had}}. \quad (1.21)$$

These correspond to contributions from quantum electrodynamics, electroweak, and hadronic interactions. The hadronic part is further divided into two subcomponents:

$$a_\mu^{\text{Had}} = a_\mu^{\text{Had,VP}} + a_\mu^{\text{Had,LbL}}, \quad (1.22)$$

which are the so-called hadronic vacuum polarization and hadronic light-by-light scattering contributions. Each component of  $a_\mu^{\text{SM}}$  has both a value and an uncertainty; the uncertainty's source—and correspondingly its size—varies between the different components. Figure 1.3 shows representative Feynman diagrams belonging to each of the above categories. In the following subsections, the identified contributions will be defined and their values will be presented. Full evaluations of  $a_\mu^{\text{SM}}$  have recently been completed by Keshavarzi et al. and Davier et al. [14, 15].

##### 1.4.1 Quantum Electrodynamics

$a_\mu^{\text{QED}}$  comes from Feynman diagrams consisting solely of photons and leptons [5]. It can be expressed as a perturbative series in  $\alpha$ , the fine structure constant: [16]

$$a_\mu^{\text{QED}} = \sum_{n=1}^{\infty} \left( \frac{\alpha}{\pi} \right)^n a_\mu^{(2n)}. \quad (1.23)$$

The coefficient  $a_\mu^{(2)}$  is exactly 0.5, as determined from the one-loop diagram in Figure 1.2. Coefficients up to  $a_\mu^{(10)}$ , five loops, have been computed by Aoyama et al. [16, 17]. The

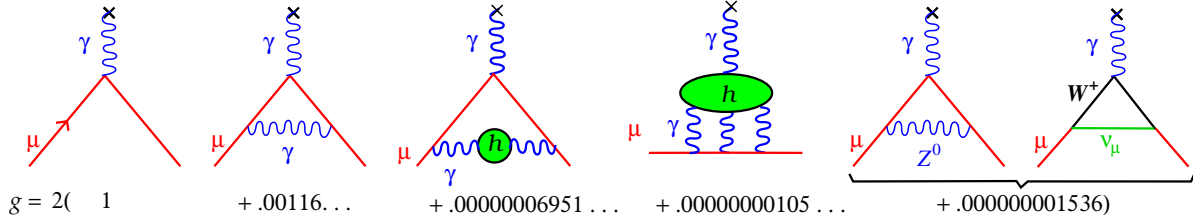


Figure 1.3: Standard Model Feynman diagrams contributing to muon  $g - 2$ , courtesy of Lee Roberts. The contributions to  $g_\mu$  from the interaction categories to which each diagram belongs are displayed along the bottom. From left to right, the diagrams are the leading-order tree-level QED diagram, and then the leading contributions to  $a_\mu^{\text{QED}}$ ,  $a_\mu^{\text{Had,VP}}$ ,  $a_\mu^{\text{Had,LbL}}$ , and  $a_\mu^{\text{EW}}$ .

coefficients beyond  $a_\mu^{(2)}$  have errors arising from lepton mass measurements and statistical uncertainties introduced by the employed Monte Carlo integration techniques.

The value of  $a_\mu^{\text{QED}}$  depends both on the coefficients in Equation 1.23 and on the value of  $\alpha$ . Aoyama et al. present two values of  $a_\mu^{\text{QED}}$ , one derived using  $\alpha(\text{Rb})$ , measured through the recoil velocity of rubidium atoms [18, 19], and one using  $\alpha(a_e)$ , measured through electron  $g - 2$  [9, 19]. The uncertainty of  $\alpha(a_e)$  is smaller than that of  $\alpha(\text{Rb})$  by more than a factor of two [19]. However, Aoyama et al. caution that because the extraction of  $\alpha(a_e)$  depends on a QED calculation that is highly correlated with the one used for  $a_\mu^{\text{QED}}$ , it may be preferable to use the more independent  $\alpha(\text{Rb})$  despite its larger uncertainty [16]. A new value of  $\alpha$  extracted from the recoil of cesium atoms,  $\alpha(\text{Cs})$ , was published in 2018 by Parker et al. [20]. The new value  $\alpha(\text{Cs})$  has a slightly smaller uncertainty than does  $\alpha(a_e)$ , and thus it may be an attractive choice for future evaluations of  $a_\mu^{\text{QED}}$ .

In any case, the uncertainty of the standard model prediction for  $a_\mu$  is not driven by the QED contribution. Using  $\alpha(\text{Rb})$ , Aoyama et al. report [17]

$$a_\mu^{\text{QED}} = (11658471.8971 \pm 0.007) \times 10^{-10}. \quad (1.24)$$

$a_\mu^{\text{QED}}$  has a sub part per billion relative uncertainty, which can be compared to the order 100 parts per billion (ppb) precision with which  $a_\mu$  can currently be measured. As will be seen,  $a_\mu^{\text{QED}}$  is by far the largest component of  $a_\mu^{\text{SM}}$ , but—compared to the other Standard Model

contributions and to the precision reasonably achievable by contemporary experiments—its uncertainty is negligible.

#### 1.4.2 *Electroweak*

Electroweak diagrams are those including weak bosons or the Higgs particle [5]. This categorization holds regardless of whatever other particles or fields are present.  $a_\mu^{\text{EW}}$  has currently been calculated to two loops. For the details of this calculation, the reader is directed to the references [21,22]. The relative contribution of electroweak effects to  $a_\mu$  is very small because of the large masses of the  $W$ ,  $Z$ , and Higgs relative to the muon. Before the discovery of the Higgs at the LHC, the uncertainty of  $a_\mu^{\text{EW}}$  had a significant component from the unknown Higgs mass. Measurements of the Higgs mass subsequently reduced the overall uncertainty of  $a_\mu^{\text{EW}}$  by more than a factor of two. Incorporating the Higgs mass and its experimental uncertainty, Gnendiger et al. report that [22]

$$a_\mu^{\text{EW}} = (15.36 \pm 0.10) \times 10^{-10}, \quad (1.25)$$

with errors dominated by three-loop contributions and hadronic effects that enter from quarks in two-loop diagrams.  $a_\mu^{\text{EW}}$  is a one part per million perturbation to  $a_\mu$ . The uncertainty of the electroweak contribution is approximately 15 times larger than that of the QED contribution, but it is still negligible compared to the precision of contemporary experiments.

In 2018, Ishikawa et al. published an independent, numerical calculation of the two-loop electroweak contribution to  $a_\mu$  [23]. Their result is different from the one just shown by  $0.26 \times 10^{-10}$ , which is 2.6 times the uncertainty quoted above. Even if one were to inflate the uncertainty of  $a_\mu^{\text{EW}}$  to account for this slight discrepancy, it would not change the conclusion that the uncertainty from electroweak contributions to muon  $g - 2$  is negligible compared to the uncertainty from hadronic contributions, which will be discussed next.

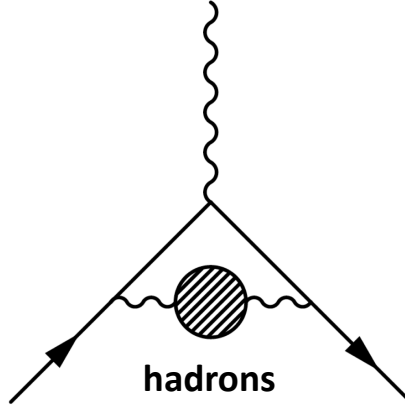


Figure 1.4: A Feynman diagram depicting the leading-order hadronic vacuum polarization contribution to muon  $g - 2$ .

#### 1.4.3 Hadronic Vacuum Polarization

Hadronic contributions to  $a_\mu$  are those that involve quarks and are absent weak bosons or the Higgs particle. Quark confinement and the large QCD coupling constant at low energies preclude the use of perturbation theory in calculating these contributions [5, 24]. Consequently, the  $a_\mu^{\text{SM}}$  uncertainty arising from hadronic effects is an order of magnitude larger than the uncertainty arising from QED or electroweak effects. The lowest order—and largest—contribution to  $a_\mu^{\text{Had}}$  is hadronic vacuum polarization,  $a_\mu^{\text{Had,VP}}$ . The leading-order  $a_\mu^{\text{Had,VP}}$  diagram is shown in Figure 1.4, wherein the loop appearing along the bottom photon line consists of hadrons, for example  $\pi^+\pi^-$ .

Although  $a_\mu^{\text{Had,VP}}$  cannot be calculated perturbatively, it can be related to experimental data by employing dispersion relations and the optical theorem [4, 24]. The only assumptions required for this approach are unitarity and analyticity. A commonly used relation to extract the leading-order (LO) contribution to  $a_\mu^{\text{Had,VP}}$  reads [14, 24, 25]

$$a_\mu^{\text{HVP,LO}} = \frac{\alpha^2}{3\pi^2} \int_{m_{\pi_0}^2}^{\infty} \frac{ds}{s} R(s) K(s) \quad (1.26)$$

$$R(s) = \sigma^0(e^+e^- \rightarrow \text{hadrons}) \bigg/ \frac{4\pi\alpha^2}{3s}, \quad (1.27)$$

where  $\sigma^0$  denotes a bare cross section from which the effects of vacuum polarization and initial state radiation have been subtracted,  $m_{\pi_0}$  is the pion mass,  $R(s)$  is the hadronic  $R$ -ratio, and  $K(s)$  is a kernel function that behaves as  $m_\mu^2/s$  at low energies [14]. The next-to-leading-order (NLO) hadronic vacuum polarization contribution can be extracted in a similar fashion [14, 24]. As the  $R$ -ratio is taken from measurement, the cross section uncertainties reported by various experiments dominate the uncertainty of  $a_\mu^{\text{HVP,LO}}$ . Accordingly, the  $a_\mu^{\text{HVP,LO}}$  prediction has become more precise over time as more cross section data has become available.

Extracting  $a_\mu^{\text{HVP,LO}}$  in the manner just described involves combining exclusive cross section measurements from different channels and also combining results from different experiments at different colliders, see Figure 1.5. This is a challenging process. See references [14, 15] for two recently completed analyses. Despite their independent approaches and slightly discrepant results in certain channels, their final values agree within  $0.2 \times 10^{-10}$ , which is less than 10% of the smaller of the two uncertainties. The LO and NLO results calculated by Keshavarzi et al. are

$$a_\mu^{\text{HVP,LO}} = (693.26 \pm 2.46) \times 10^{-10} \quad (1.28)$$

$$a_\mu^{\text{HVP,NLO}} = (-9.82 \pm 0.04) \times 10^{-10}. \quad (1.29)$$

While hadronic vacuum polarization is responsible for only 60 parts per million (ppm) of the Standard Model  $a_\mu$  prediction, it accounts for approximately half of the combined uncertainty.

Although it currently provides the most precise value, the dispersion relation presented in Equation 1.26 is not the only way to extract  $a_\mu^{\text{Had,VP}}$ . And, as the hadronic contributions to muon  $g - 2$  are responsible for such a large part of the Standard Model prediction's uncertainty, alternative approaches are desirable for independent verification. A calculation from first principles is possible using lattice QCD and QED. In 2018, Blum et al. released a full lattice calculation of  $a_\mu^{\text{HVP,LO}}$  that is in agreement with the  $R$ -ratio result, although with a significantly larger uncertainty [26]. By augmenting their lattice data at very short

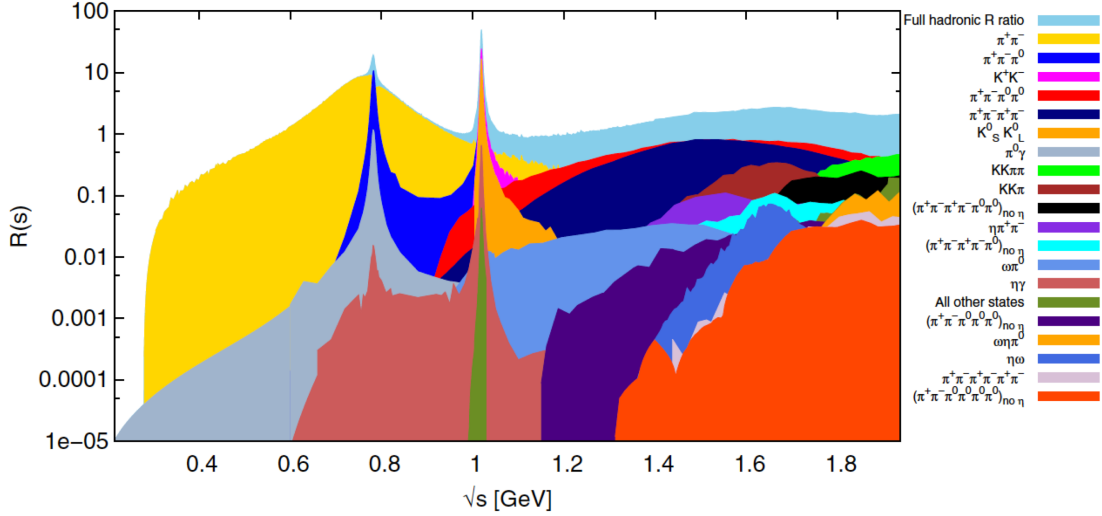


Figure 1.5: Hadronic  $R$ -ratio data used to calculate  $a_\mu^{\text{Had,VP}}$ , reproduced from Keshavarzi et al. [14]. Each color represents a different hadronic final state. Integrating the  $R$ -ratio along with a kernel function provides  $a_\mu^{\text{Had,VP}}$ , as shown in Equation 1.26.

and long distances with  $R$ -ratio data, they arrived at a result with an uncertainty similar to the one reported by Keshavarzi et al. This is encouraging, and in the coming years lattice calculations are likely to provide completely independent measurements of  $a_\mu^{\text{HVP,LO}}$  at precisions competitive with those from traditional dispersive analyses. Another novel approach, which has not yet been attempted, is to extract  $a_\mu^{\text{HVP,LO}}$  from the running of  $\alpha$  in the spacelike domain, measured through  $\mu e \rightarrow \mu e$  scattering. This would require a dedicated experiment, such as the one proposed by Abiendi et al. [27].

#### 1.4.4 Hadronic Light-by-Light Scattering

Figure 1.6 shows the hadronic light-by-light scattering process responsible for  $a_\mu^{\text{Had,LbL}}$ . It appears at the same order as  $a_\mu^{\text{HVP,NLO}}$ , and its effect has historically been very difficult to calculate. In *The Muon  $g - 2$* , Jegerlehner et al. describe the hadronic light-by-light contributions to  $a_\mu$  as “The most problematic set of hadronic corrections...” [24]. Consequently, many theorists are actively working to improve the determination of  $a_\mu^{\text{Had,LbL}}$ . It is beyond

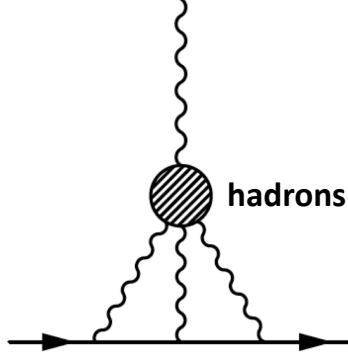


Figure 1.6: A Feynman diagram depicting the hadronic light-by-light scattering contribution to muon  $g - 2$ .

the scope of this dissertation to describe their work in detail. For a summary of the current status as of November 2018, see reference [28].

Before 2014, all  $a_\mu^{\text{Had,LbL}}$  estimates were based on model calculations [28]. Since then, there has been a vigorous effort to produce independent values using dispersive approaches and lattice calculations [29, 30]. While Equation 1.26 described  $a_\mu^{\text{HVP,LO}}$  in terms of a single integral, no such relation exists for  $a_\mu^{\text{Had,LbL}}$ . A separate integral is required for each intermediate state of the hadronic light-by-light scattering process [29]. See the 2018 paper by Hoferichter et al. for a complete dispersive analysis of the leading contribution to  $a_\mu^{\text{Had,LbL}}$  [31].

New dispersive analyses and lattice calculations are likely to provide improved determinations of  $a_\mu^{\text{Had,LbL}}$  in the near future. For now, the  $a_\mu^{\text{Had,LbL}}$  values employed by recent combined estimates of  $a_\mu^{\text{SM}}$  come from model calculations [14, 15]. Keshavarzi et al. use the value suggested by Nyffeler, which is [32]

$$a_\mu^{\text{Had,LbL}} = (9.8 \pm 2.6) \times 10^{-10}. \quad (1.30)$$

Notice that, in the Keshavarzi et al. analysis, the hadronic light-by-light process is the largest source of uncertainty in  $a_\mu^{\text{SM}}$ . This is despite a central value that is smaller even than the



electroweak contribution.

#### 1.4.5 Combined Standard Model $a_\mu$ Prediction

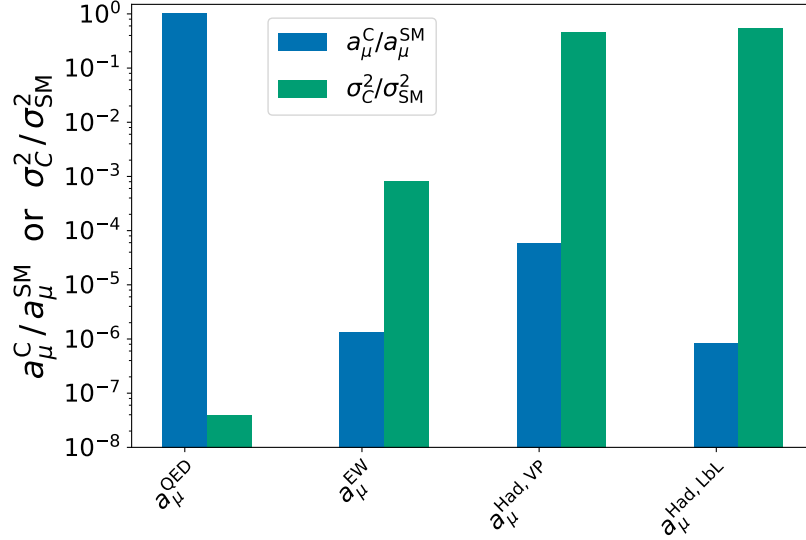


Figure 1.7: Visual decomposition of  $a_\mu^{\text{SM}}$  and its variance,  $\sigma_{\text{SM}}^2$ .  $a_\mu^C$  and  $\sigma_C^2$  refer to the contribution of an individual component to its respective total. The Standard Model prediction for  $a_\mu$  is dominated by the QED component, but its variance is dominated by the hadronic components.

With all its components calculated, one arrives at a combined Standard Model prediction of  $a_\mu$ . Keshavarzi et al. report the combined value [14]

$$a_\mu^{\text{SM}} = (11659182.04 \pm 3.56) \times 10^{-10}, \quad (1.31)$$

corresponding to a relative uncertainty of 300 ppb. This value includes a small NNLO hadronic component that was not mentioned above. Additionally, its uncertainty accounts for correlations between the different hadronic orders. Other analyses arrive at slightly different results, but agreement is generally quite good and certainly well within the stated uncertainty of any given analysis. See Figure 1.7 for a visual decomposition of  $a_\mu^{\text{SM}}$ .

### 1.5 The Brookhaven Muon $g - 2$ Experiment

As mentioned earlier, muon  $g - 2$  measurements are potentially sensitive to unknown physics interactions. The more precise the measurement, the greater the sensitivity. Furthermore, testing the precise Standard Model prediction presented in the previous section is interesting in and of itself, even if disagreement were merely to indicate an error in the extremely difficult calculation. These two motivations are not independent, as experimental precision sufficiently beyond that of the theoretical prediction is not particularly valuable. For these reasons, experimentalists have found it worthwhile to expend great effort on increasingly precise measurements of muon  $g - 2$ , and theorists have been no less busy constructing their myriad interpretations of the results.

The first muon magnetic moment measurement occurred in 1957. Since then, there have been a number of muon  $g - 2$  experiments, each improving on the last. Lee Roberts succinctly summarized the fascinating history of these experiments in his article *The History of Muon ( $g - 2$ ) Experiments* [33]. Here the focus will be on the most recently completed experiment, The Brookhaven Muon  $g - 2$  Experiment, or E821 [34]. The measurement techniques employed by the E821 experiment are nearly identical to those of the E989 Muon  $g - 2$  Experiment; they will be discussed in the next chapter.

In short, E821 measured  $a_\mu$  by comparing the spin precession and cyclotron frequencies of muons in a known magnetic field. The difference between the spin precession frequency and the cyclotron frequency is called the anomalous precession frequency—see the next chapter for details. Technically speaking, E821 did not measure  $a_\mu$ , but rather the ratio of the aforementioned anomalous precession frequency to the spin precession frequency of free protons in the same field. The E821 collaboration then related this quantity to  $a_\mu$  using the ratio of the muon magnetic moment to the proton magnetic moment, which was taken from a completely separate experiment that had determined the required magnetic moment ratio by measuring the hyperfine interval of muonium [34, 35].

It is worth appreciating how many experiments have now been identified as providing

parameters necessary either for the  $a_\mu$  measurement or for the calculation of its Standard Model value. The Standard Model predicts that the results of all these experiments will be related in a very particular way. Testing for consistency between the  $a_\mu$  measurement,  $a_\mu^{\text{exp}}$ , and  $a_\mu^{\text{SM}}$  amounts to verifying that this predicted relationship holds. As a reminder, the involved measurements include but are not limited to the rubidium recoil velocity, the Higgs mass, the hadronic  $R$ -ratio, the muonium hyperfine interval, and the frequency ratio from E821.

The current value of  $a_\mu^{\text{exp}}$  is [5, 36]

$$a_\mu^{\text{exp}} = (11659208.9 \pm 6.3) \times 10^{-10}. \quad (1.32)$$

The relative uncertainty, dominated by the E821 measurement, is 540 ppb, quite similar to that of the theoretical prediction. Additionally, the experimental uncertainty is smaller than the electroweak and hadronic light-by-light contributions to  $a_\mu^{\text{SM}}$ , meaning that—despite their smallness—they must be included and understood to interpret the E821 result.

Calculating the difference between theory and experiment requires choosing a particular Standard Model analysis. With the value determined by Keshavarzi et al. [14]

$$a_\mu^{\text{exp}} - a_\mu^{\text{SM}} = (26.86 \pm 7.24) \times 10^{-10}. \quad (1.33)$$

Note that the value reported here is not quite the same as the one reported in Keshavarzi et al. because of differing choices for the exact value of  $a_\mu^{\text{exp}}$ . Here,  $a_\mu^{\text{exp}}$  comes from *CODATA recommended values of the fundamental physical constants: 2014* [5, 36].

The difference between E821's  $a_\mu$  measurement and the Standard Model prediction is nonzero with a significance of  $3.7\sigma$ . Or, if one instead chooses the recent Standard Model analysis by Davier et al., the significance is  $3.5\sigma$  [15]. The absolute size of the discrepancy is about twice that of the electroweak contribution to  $a_\mu^{\text{SM}}$ , and with a sign such that failing to include the electroweak contribution would increase the discrepancy.

The E821 collaboration's final report, published over a decade before this writing, declared a  $2.2$ - $2.7\sigma$  discrepancy between their result and the best Standard Model analyses

available at the time [34]. In the intervening years, increased scrutiny of the theory prediction, additional collider data, and improved understanding of the hadronic light-by-light contribution have yielded no resolution. On the contrary, the difference between theory and experiment has only become more statistically significant. Besides the theory, the E821 measurement—which was quite challenging and required careful treatment of many subtle systematic effects—has also been subjected to intense scrutiny. However, no error that could resolve the discrepancy has been discovered.

Before a declaration can be made with any degree of certainty, a new measurement—or, ideally, multiple new measurements—is necessary. The difference between the E821 result and the Standard Model prediction does not meet the  $5\sigma$  discovery threshold, and, even (or perhaps especially) if it did, an independent measurement would certainly be warranted. Two new  $a_\mu$  experiments are expected to release results in the relatively near future. The first, The E989 Muon  $g - 2$  Experiment, is a direct successor to E821 and the subject of the remainder of this document [37]. The second, The Muon  $g - 2$  Experiment at J-PARC, utilizes a very different experimental setup and thus will provide an extremely valuable crosscheck should the E989 Experiment confirm the muon  $g - 2$  discrepancy [38].

## 1.6 New Physics Interpretations of the Brookhaven Result

This section addresses the following question: assuming the theorists have made no errors in their calculation, and assuming the experimentalists have made no errors in their measurement, what could explain the observed muon  $g - 2$  discrepancy? Although these assumptions will be challenged by the ongoing work to provide independent estimates of the hadronic contributions to  $a_\mu^{\text{SM}}$  and by the upcoming new measurements of  $a_\mu^{\text{exp}}$ , there is currently no compelling reason to believe that the discrepancy will not endure.

For an answer to the question just posed, one need not read beyond the title of Czarnecki and Marciano’s article *Muon anomalous magnetic moment: A harbinger for “New Physics”*, published in 2001 [12]. New physics, interactions and particles not included in the Standard Model, correct the muon-photon coupling through loops in the same way that the electroweak

and strong interactions do, see Figure 1.8 for a possible example. Thus, the muon  $g - 2$  discrepancy signals the existence of unaccounted for effects. Furthermore, the size of the discrepancy—even if it were zero—severely constrains the parameters of new physics models.

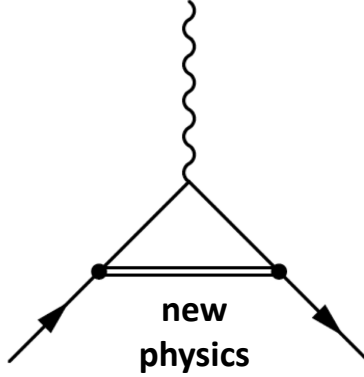


Figure 1.8: A Feynman diagram depicting a possible new physics contribution to muon  $g - 2$  entering at one loop. The double line connecting the incoming and outgoing muons represents a new physics particle. Some possibilities of what such a particle could be are discussed in the text.

The impact that a new physics process will have on  $a_\mu$  of course depends on the details of the proposed interaction. Recall that the muon  $g - 2$  discrepancy is about twice the size of the electroweak contribution. Looking in a little more detail, one finds that the leading diagram involving the  $Z$  yields a correction with the opposite sign from the leading diagram involving the  $W$ ; the  $W$  diagram considered in isolation shifts  $a_\mu$  by about  $39 \times 10^{-10}$ , which is just a bit larger than the E821 discrepancy [13]. This is presented to illustrate that particles with masses of approximately 100 GeV (like the  $W$ ) can correct  $a_\mu$  by an amount of the same order as the discrepancy, provided they couple to muons in a fashion similar to the  $W$  and enter at one loop. The exact size of the correction would depend on the mass of the exchanged particle and on the involved coupling constants.

Czarnecki and Marciano discuss a number of new physics possibilities, among them supersymmetry and radiative mass mechanisms [12]. The masses of new physics particles that could explain the  $g - 2$  discrepancy vary significantly between the considered models. For

some, the new physics scale is above one TeV. Additionally, as with the electroweak example,  $a_\mu$  contributions from new physics interactions can come with either a positive or a negative sign. Thus, for many new physics models, the definite sign of the observed  $g - 2$  discrepancy eliminates large swathes of parameter space that might otherwise be allowed [12].

Possible resolutions of the muon  $g - 2$  discrepancy cannot be considered in isolation. New particles must be observed in multiple ways and by multiple experiments for their discovery to be credible. Indeed,  $a_\mu$  experiments alone cannot distinguish between models that explain their measurements equally well. They can, however, draw attention to regions of parameter space that may then be explored with direct searches. For example, the LHC can address much of the supersymmetry parameter space relevant to the  $g - 2$  discrepancy. A 2017 article by Hagiwara et al. states that while LHC searches have highly constrained much of the relevant parameter space, some regions that could resolve the muon  $g - 2$  discrepancy have not yet been excluded [39].

Another model that has been proposed to explain the muon  $g - 2$  discrepancy is that of the dark photon. In this model, an additional U(1) group with a massive gauge boson is introduced [41]. The massive gauge boson, called the dark photon,  $\gamma'$ , then mixes with the Standard Model photon,  $\gamma$ . A parameter  $\epsilon$  controls the mixing strength. This model has attracted interest because of its potential to simultaneously explain both muon  $g - 2$  and the positron excess observed in cosmic ray spectra [42]. Additionally, the simplest version of the model has a two dimensional parameter space—spanned by the dark photon mass,  $m_{\gamma'}$ , and the mixing parameter,  $\epsilon$ —that can be addressed with existing experiments, see, for example, [40, 43, 44]. As of the NA48/2 Collaboration’s 2015 publication, the region of this parameter space that resolves the muon  $g - 2$  discrepancy has been entirely excluded, see Figure 1.9. While it appears that dark photons cannot account for the E821 result, their story serves as an example of how  $a_\mu$  measurements can guide the search for new physics in general.

Yet more possibilities remain. For example, light ( $<1$  GeV) scalar and axion-like pseudoscalar particles, which may not yet have been otherwise seen, are capable of resolving

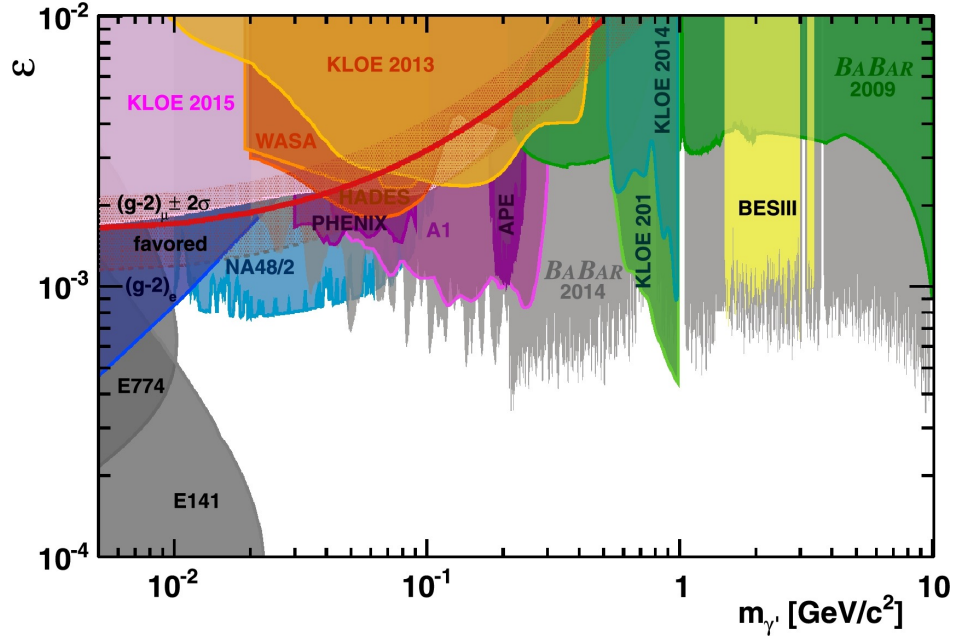


Figure 1.9: Parameter space of the dark photon model relevant to muon  $g - 2$ , reproduced from [40]. On the x-axis is the dark photon mass and on the y-axis is the parameter governing the mixing between the dark photon and the Standard Model photon. The red band is the region of parameter space that resolves the muon  $g - 2$  discrepancy. Each exclusion band comes from a separate experiment. Combining all experiments, the simple dark photon model is entirely ruled out as an explanation for the  $g - 2$  discrepancy.

the  $g - 2$  discrepancy [45, 46]. Each of these scenarios can be further addressed by existing experiments, as discussed in the references.

If the Brookhaven muon  $g - 2$  result is not an experimental artifact or a very unlikely statistical fluctuation, then the Standard Model prediction for  $a_\mu$  is incorrect. If the prediction is incorrect, and no errors were made in its determination, then physics beyond the Standard Model has been observed. That particle physicists consider this an exciting prospect is evident in the over 2000 citations garnered by the Brookhaven experiment's final report [34]. To confirm their result, an independent measurement of  $a_\mu$  is absolutely essential.

## Chapter 2

### THE E989 MUON $g - 2$ EXPERIMENT

The New Muon  $g - 2$  Experiment at Fermi National Accelerator Laboratory (FNAL), designated E989, aims to improve on the current best measurement of  $a_\mu$  by a factor of four in precision, with a target combined uncertainty of 140 parts per billion (ppb) [37]. Coupled with expected reductions in the Standard Model prediction's uncertainty, E989 is designed to either resolve the discrepancy between theory and experiment or confirm it with a significance of  $5\sigma$  or greater. To achieve this unprecedented precision, E989 uses an intense, pure muon beam provided by FNAL and will ultimately accumulate a dataset sufficient to reach 100 ppb statistical uncertainty. Systematic uncertainties are controlled both with detector and calibration systems constructed specifically to meet the unique challenges of the experiment and with computer models and simulations. With 100 ppb statistical uncertainty, the experiment's total  $a_\mu$  systematic uncertainty must be held below 100 ppb to reach the target combined uncertainty. This chapter describes the principles behind the  $a_\mu$  measurement and the design of the E989 experiment, with an emphasis on how the design enables the aforementioned levels of precision.

#### 2.1 *Measurement Principle*

A point particle with spin, mass  $m$ , and charge  $e$  placed in an external magnetic field will follow a circular path in the plane defined by the magnetic field, and its spin direction will rotate in the same plane. The particle's circular motion is called cyclotron motion and its spin rotation is called spin precession. With a uniform magnetic field, the absence of an electric field, and particle motion entirely perpendicular to the magnetic field, the equations



governing this motion are [1]

$$\vec{\omega}_s = -\frac{ge\vec{B}}{2m} - (1 - \gamma)\frac{e\vec{B}}{m\gamma} \quad (2.1)$$

$$\vec{\omega}_c = -\frac{e\vec{B}}{m\gamma}. \quad (2.2)$$

In the above equations,  $\omega_s$  is the spin precession angular frequency,  $\omega_c$  is the orbital angular frequency,  $\gamma$  is the relativistic Lorentz factor, and  $g$  is the particle's  $g$ -factor. The second term in Equation 2.1, which did not appear in Equation 1.11, is a relativistic correction called Thomas precession [1]. It accounts for the continuous rotation of the orbiting particle's frame of reference.

The rate of change of the angle between the particle's momentum and its spin is called the anomalous precession frequency,  $\omega_a$ , and is the difference between  $\omega_s$  and  $\omega_c$ :

$$\vec{\omega}_a = \vec{\omega}_s - \vec{\omega}_c = -\left(\frac{g-2}{2}\right)\frac{e\vec{B}}{m} = -a\frac{e\vec{B}}{m}, \quad (2.3)$$

where  $a$  is the anomalous magnetic moment of the particle in question. Note in the above equation that  $\omega_a$  is directly proportional to  $a$ : if  $g$  were exactly equal to 2, the spin would precess at the orbital frequency and the angle between the spin and momentum would never change. By observing anomalous precession in a magnetic storage ring, one can measure  $g - 2$  directly rather than  $g$ . This is the principle on which the E989 experiment is based. The mechanism for measuring the anomalous precession frequency,  $\omega_a$ , will be discussed later in this chapter.

A precise measurement of  $\omega_a$  alone is not sufficient to determine  $a_\mu$ . One must also know the other quantities in Equation 2.3 with comparable or better precision. Introducing  $\omega_p$  and  $\mu_p$ , the proton Larmor precession frequency in the storage ring field and the proton's magnetic dipole moment, and  $g_e$ ,  $\mu_e$ , and  $m_e$ , the  $g$ -factor, magnetic dipole moment, and

mass of the electron, one can use the following relations to rewrite Equation 2.3:

$$B = \frac{\hbar \omega_p}{2\mu_p}$$

$$e = \frac{4m_e \mu_e}{\hbar g_e}.$$

The first of these equations is an expression of the magnetic field in terms of the precession frequency of free, non-relativistic protons and the second comes simply from the definition of  $g_e$ . Combining these equations with Equation 2.3 yields the following relation:

$$a_\mu = \frac{g_e}{2} \frac{\omega_a}{\omega_p} \frac{m_\mu}{m_e} \frac{\mu_p}{\mu_e}. \quad (2.4)$$

Allowing for slight nonuniformities in the magnetic field, the quantity  $\omega_p$  should be replaced with  $\tilde{\omega}_p$ , where  $\tilde{\omega}_p$  is the particle distribution weighted spatial average of the proton Larmor precession frequency. The average must be weighted by the particle distribution because the magnetic field at a given location only contributes to  $\omega_a$  if particles are present there. Making this replacement, the equation reads

$$a_\mu = \frac{g_e}{2} \frac{\omega_a}{\tilde{\omega}_p} \frac{m_\mu}{m_e} \frac{\mu_p}{\mu_e}. \quad (2.5)$$

This is the identity E989 will use in practice to determine  $a_\mu$ ;  $\omega_a$  and  $\tilde{\omega}_p$  are measured in the E989 storage ring, whereas  $g_e$ ,  $m_\mu/m_e$ , and  $\mu_p/\mu_e$  are known from prior experiments precisely enough not to contribute significantly to E989's experimental uncertainty (see Table 2.1).

## 2.2 Storage Ring and Electrostatic Quadrupoles

The E989 magnetic storage ring, shown in Figure 2.1 is the same one previously used in the E821 Muon  $g - 2$  Experiment at Brookhaven National Lab. It is an iron electromagnetic with superconducting coils operating at 5200 A [37]. When energized, the magnet provides a highly uniform 1.451 T field in a toroidal storage region with a 90 cm diameter and a 7.112 m orbit radius. Inside the storage region, the azimuthal peak-to-peak variations of

Table 2.1: Shown above are the quantities required to determine  $a_\mu$  and their associated uncertainties.  $\omega_a/\tilde{\omega}_p$  will be measured in the E989 experiment; the quoted 140 ppb uncertainty is expected at the experiment’s completion. The values and uncertainties for  $g_e$ ,  $m_\mu/m_e$ , and  $\mu_p/\mu_e$  are CODATA’s recommended values [36]

Quantity	Source	Uncertainty or target uncertainty [ppb]
$g_e$	Quantum cyclotron spectroscopy [9]	0.00026
$m_\mu/m_e$	Muonium spectroscopy [35]	22
$\mu_p/\mu_e$	Hydrogen spectroscopy [36]	3
$\omega_a/\tilde{\omega}_p$	E989 experiment [37]	140
Total $a_\mu$ uncertainty		$\sim 140$

the magnetic field are on the order of 100 ppm. In the 9 cm diameter storage region, the azimuthally averaged field uniformity is on the order of 1 ppm.

The force exerted by a magnetic field is always perpendicular to the field, and thus a uniform magnetic field is not sufficient to stably store particles in three dimensions. Any momentum component parallel to the magnetic field would lead to rapid particle loss. To prevent such losses, electrostatic quadrupole plates are placed inside of the vacuum chambers. These plates are charged to voltages between  $\pm 20$  kV and  $\pm 30$  kV, and they provide vertical focusing at the expense of radial defocusing. Figure 2.2 shows the equipotential contours of the electrostatic quadrupole fields. Together, the storage ring’s magnetic field and the electrostatic quadrupole field provide restoring forces in both the vertical and radial directions [37, 47].

An electric field perpendicular to a particle’s momentum affects the particle’s spin precession frequency. The expression for  $\omega_a$  including the effect of a transverse electric field

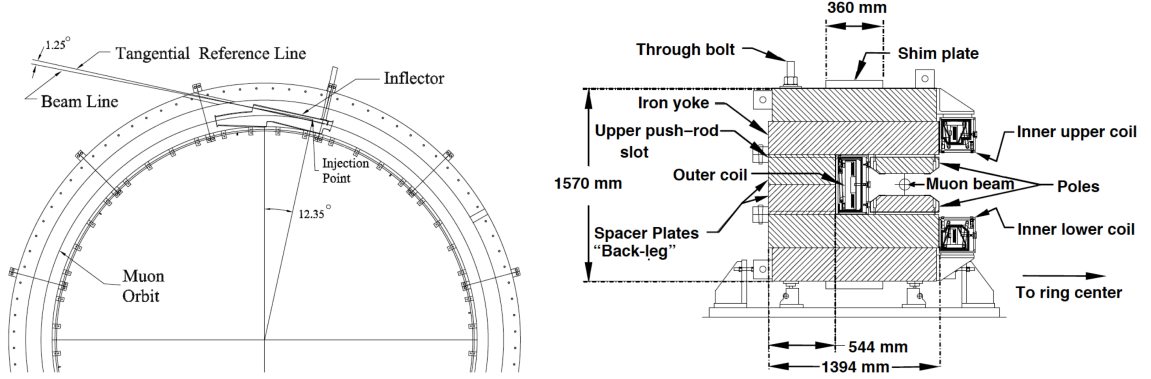


Figure 2.1: Top down and cross sectional view of the E989 storage ring, reproduced from [37]. The injection process and the inflector will be discussed more later in this chapter.

is [37]

$$\vec{\omega}_a = -\frac{e}{m} \left[ a_\mu \vec{B} - \left( a_\mu - \frac{1}{\gamma^2 - 1} \right) \frac{\vec{\beta} \times \vec{E}}{c} \right]. \quad (2.6)$$

For a proper choice of  $\gamma$ , or equivalently stored particle energy, the effect of the electric field on the anomalous precession frequency will become very small. With a  $\gamma$  of 29.3,  $1/(\gamma^2 - 1)$  is nearly exactly equal to  $a_\mu$  and the electric field effect becomes a small, sub-ppm correction to  $\omega_a$ . The so-called magic  $\gamma$  of 29.3 corresponds to a muon momentum of 3.094 GeV/c, which is, by design and for this reason, the momentum stored by the E989 storage ring.

### 2.3 Muon Decay

Muons are unstable particles and decay with a lifetime of  $2.2 \mu\text{s}$  in the frame of reference where they are at rest. Positive muons predominantly decay into a positron and two neutrinos,

$$\mu^+ \longrightarrow e^+ \nu_e \bar{\nu}_\mu, \quad (2.7)$$

with a branching fraction of nearly 100% [5]. Muon decay proceeds through the weak interaction and therefore exhibits parity violation. The parity violation manifests as a correlation

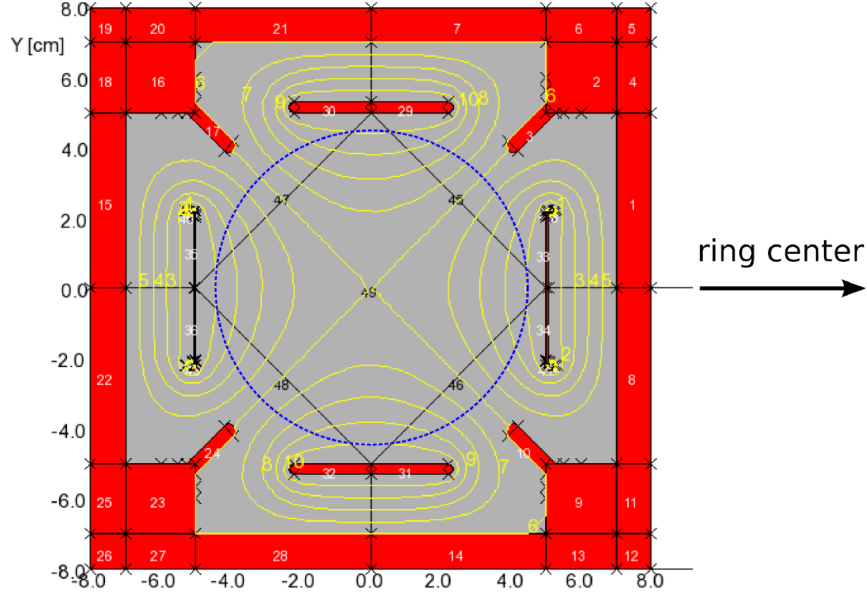


Figure 2.2: Equipotential contours from the electrostatic quadrupoles, reproduced from [37]. The blue curve outlines the muon storage region. The electrostatic quadrupole fields are vertically focusing and radially defocusing.

between the decay positron momentum and the muon spin, i.e. [5, 37]

$$\langle \vec{S}_\mu \cdot \vec{p}_e \rangle \neq 0. \quad (2.8)$$

This correlation could not exist in a system obeying parity symmetry because under a parity transformation the vector  $\vec{p}_e$  changes sign but the axial vector  $\vec{S}_\mu$  does not. As will be explained, the correlation between the muon spin and positron momentum is the physical mechanism through which  $\omega_a$  is measured in the E989 experiment.

The differential decay distribution in the muon's rest frame can be expressed in terms of the positron energy  $E$  and the angle between the positron momentum and muon spin  $\theta$  as [37]

$$\frac{d^2 P}{dE d\cos(\theta)} = N_r(E) [1 \pm A_r(E) \cos(\theta)], \quad (2.9)$$

where the positive sign is taken for positive muons and the negative sign for negative muons. The  $r$  subscripts indicate that these functions refer to quantities in the muon rest frame. The

function  $A_r(E)$  is called the *asymmetry* and encodes how strong the correlation between the positron momentum and muon spin is for decay positrons of a given energy. It is bounded between  $-1$  and  $1$ , where  $A_r(E) = 1$  means that positrons of energy  $E$  are never emitted opposite to the muon spin and  $A_r(E) = -1$  means that positrons of energy  $E$  are never emitted along the muon spin. The asymmetry,  $A_r(E)$ , increases toward  $1$  as  $E$  approaches  $m_\mu c^2/2$ , the maximum possible energy an emitted positron can have in the muon rest frame. This means that, for positive muons, higher energy decay positrons are more likely to be emitted along the muon spin, and this correlation becomes stronger at higher positron energies.  $N_r(E)$  also increases with  $E$ , so decay positrons in general are preferentially emitted in the direction of the muon spin. See Figure 2.9 for a graphical representation of this behavior.

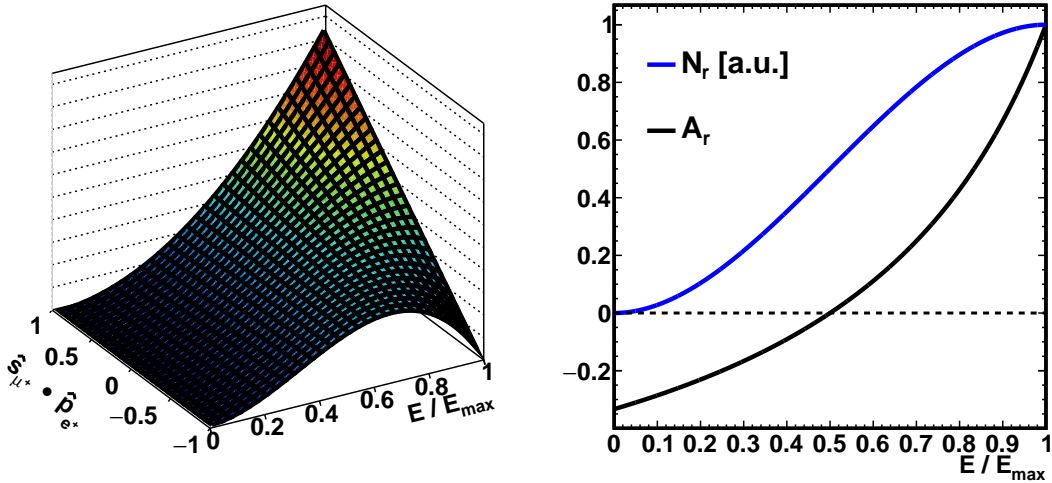


Figure 2.3: Muon decay distribution in the rest frame. The plot on the left shows the differential decay distribution as a function of  $\hat{S}_\mu \cdot \hat{p}_{e^+}$ , equivalent to  $\cos(\theta)$  in Equation 2.9. The highest energy positrons are most likely to be emitted in the direction of the muon spin. The plot on the right shows the functions  $N_r$  and  $A_r$ , defined in the text.  $E_{max}$  is the maximum possible energy a decay positron can have, which is very nearly  $m_\mu c^2/2$ .

To determine what will be observed in the E989 experiment, the muon decay distribution must be Lorentz transformed into lab frame coordinates. In the lab, the muons are highly

relativistic with a  $\gamma$  of 29.3. To complete the transformation, another angle must be introduced: the angle between the muon spin and the boost direction, or, equivalently, the angle between the muon spin and the muon momentum in the lab. Neglecting the mass of the positron, the energy in the lab frame is given by the expression

$$E_{lab} = \gamma E_{rest} (1 - \beta \vec{p}_e \cdot \hat{b}), \quad (2.10)$$

where  $\hat{b}$  points along the boost axis, opposite to the muon momentum. The lab frame energy has a strong dependence on the emission angle of the positron. As the emission angle is correlated with the muon spin direction, the energy distribution in the lab frame depends on the angle between the muon spin and the muon momentum. For example, when the spin and momentum are aligned, positrons are more likely to be emitted parallel to the muon momentum and are therefore more likely to be observed with higher energies in the lab frame.

Calling the angle between the muon spin and its momentum  $\alpha$ , the energy distribution in the lab frame is

$$\frac{dP}{dE} = N(E) [1 + A(E) \cos(\alpha)]. \quad (2.11)$$

The functions  $N(E)$  and  $A(E)$  are distinct from  $N_r(E)$  and  $A_r(E)$ .  $N(E)$  and  $A(E)$  describe the lab frame energy distribution.  $A(E)$  encodes how the probability of observing a positron with energy  $E$  in the lab frame depends on the angle between the muon spin and momentum, with positive values indicating that energy  $E$  positrons are more likely to be observed when the spin and momentum are aligned and negative values indicating that they are less likely to be observed when the spin and momentum are aligned. Figure 2.4 shows the decay positron energy distribution in the lab frame for three different values of  $\alpha$ .

The  $\cos(\alpha)$  modulation of the positron energy distribution in the lab frame enables the  $\omega_a$  measurement. The rate of change of  $\alpha$  is exactly  $\omega_a$ , and thus the energy distribution in the lab frame changes with a period of  $2\pi/\omega_a$ . Replacing  $\alpha$  with  $\omega_a t - \phi$ , the positron decay distribution as observed in the lab frame is

$$\frac{d^2P}{dE dt} \propto e^{-t/(\gamma\tau_\mu)} N(E) [1 + A(E) \cos(\omega_a t - \phi)]. \quad (2.12)$$

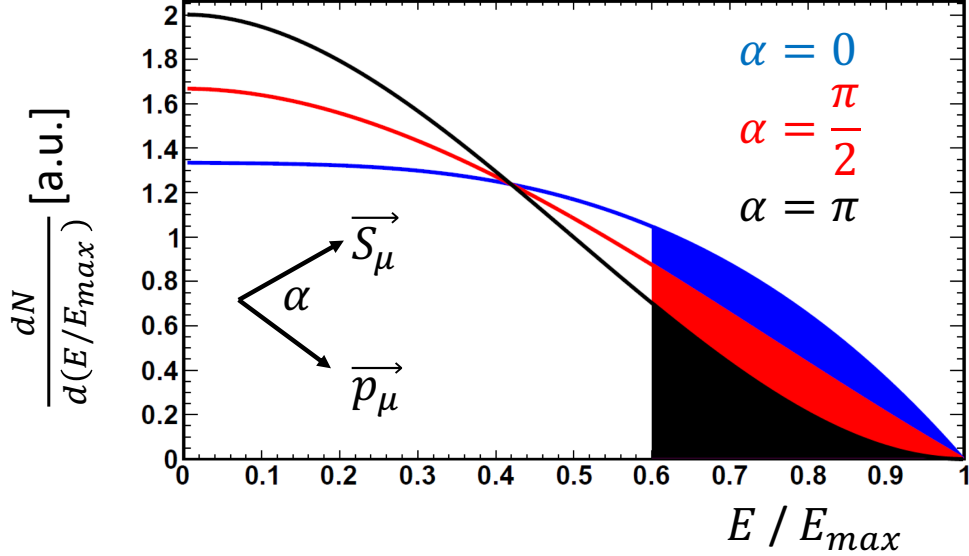


Figure 2.4: Muon decay energy distribution in the lab frame for three different values of the angle between the muon spin and its momentum,  $\alpha$ . When the spin and momentum are aligned, the decay positron energy distribution shifts in favor of higher energies. In the lab frame,  $E_{max}$  is approximately 3.1 GeV. The fraction of observed decay positrons above a fixed energy threshold depends on  $\alpha$  and thus oscillates at  $\omega_a$ . This is illustrated by the solid regions above  $E/E_{max} = 0.6$ ; the area of the solid blue region, representing the probability of observing a high-energy positron when the muon spin and momentum are parallel, is more than twice the area of the solid black region, representing the same when the muon spin and momentum are antiparallel. Techniques for extracting  $\omega_a$  from the observed decay positron spectrum will be discussed more in Section 2.6.

The  $\omega_a$  extraction in the E989 experiment is achieved by measuring the modulation of the detected positron energy distribution and fitting its periodicity.

## 2.4 Beam Injection

The FNAL muon campus beamline is constructed to deliver pure pulses of highly polarized muons to the E989 storage ring. The beam is delivered in bursts, called fills, separated by at least 10 ms and with an average rate of 11.4 Hz. The distribution of delivered muons has a momentum RMS of approximately 2% centered around 3.094 GeV/ $c$  and a temporal length



of 120 ns. Of these injected muons, only 1% to 2% can be stored. With a 4.5 cm radius storage region and a 7.112 m orbit radius, the E989 ring can at best store muons within approximately 0.5% of the design momentum.

The polarized muon beam originates from decaying pions, which are in turn produced by focusing an 8 GeV proton beam from the FNAL accelerator on a pion production target. Positive particles within  $\pm 10\%$  of 3.11 GeV/ $c$  are selected and steered into a beamline containing focusing/defocusing optics (FODO) optimized to capture 3.094 GeV/ $c$  muons produced as pions decay in flight. This section of beamline is called the M2 and M3 lines and is approximately 280 m in length. 3.11 GeV/ $c$  pions have a decay length of 170 m, so only about 20% of the pions in the M2 and M3 lines will traverse the full 280 m without decaying. Pion decays in the M2 and M3 lines are the primary source of muons in the particle beam delivered to the E989 experiment [37, 48].

Charged pions, like muons, decay through the weak interaction. The parity violating nature of the weak interaction is responsible both for pions predominantly decaying to muons rather than electrons (with a branching fraction of greater than 99.98%) and for the polarization of the produced muon beam [5, 37]. An explanation for why this is so goes as follows. Charged pions decay to a charged lepton and a neutrino

$$\pi^+ \rightarrow l^+ \nu_l. \quad (2.13)$$

Neutrinos produced in the weak interaction have a definite helicity state, left-handed for neutrinos and right-handed for antineutrinos. In the rest frame of the pion, the charged lepton and the neutrino must be produced back to back to conserve momentum. The pion is a spin zero particle, so the decay products must also be in a zero angular momentum state. Therefore, the charged lepton in the rest frame of the pion will also be in a definite helicity state with the same handedness as the neutrino. For a massless charged lepton, this decay would be forbidden because the weak interaction does not couple to massless left-handed antiparticles or massless right-handed particles. The amplitude for pions to decay to a given charged lepton channel must be suppressed by the mass of the charged lepton to regain

the proper behavior in the limit where the mass goes to zero. The partial decay rate is proportional to the square of the amplitude, so the rate of charged pion decay to electrons is suppressed by a factor of  $(m_e/m_\mu)^2$  relative to the rate of decay to muons [2, 5].

The definite helicity state of decay muons in the pion rest frame becomes a correlation between muon energy and polarization along the pion momentum direction in the lab frame. Only the highest energy decay muons are captured in the M2 and M3 lines. These muons are produced with their momenta and spin along the boost axis in the rest frame and thus they are highly polarized along the pion momentum direction in the lab frame. High-energy positive muons are produced in a left-handed helicity state; their spin is opposite their momentum. Selecting 3.094 GeV/ $c$  muons from 3.11 GeV/ $c$  pion decays yields a muon beam with a net polarization of about 95% [37]. See Figure 2.5 for a schematic of the FNAL muon campus beamline.

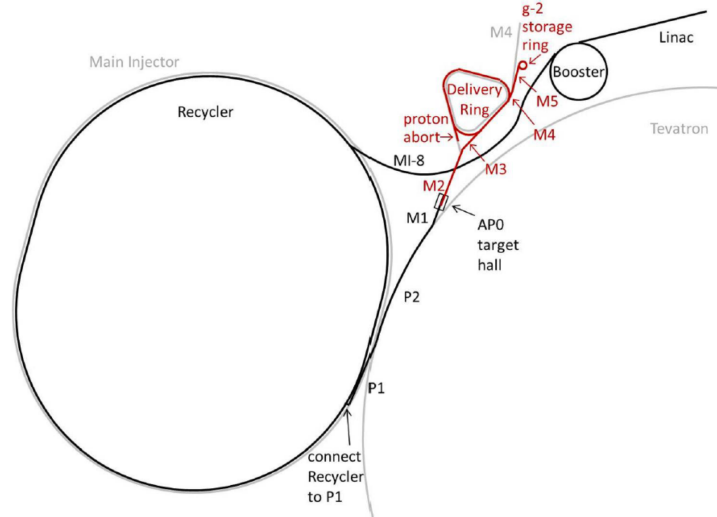


Figure 2.5: The muon campus beamline, in red, reproduced from [37]. Pions are produced from 8 GeV protons in the AP0 target hall. The beamline components are described in the text.

At the end of the M2 and M3 lines, the beam is transferred into a facility called the Delivery Ring, or DR. The DR is a 505 m long ring of beamline around which the beam

circulates. Upon entering the DR, the beam contains a mixture of muons, protons, and the 20% of pions that did not decay in the M2 and M3 lines.  $3.094 \text{ GeV}/c$  protons have a  $\gamma$  of 3.3 and travel at about 95% the speed of light. Muons with a  $\gamma$  of 29.3 travel within 0.1% of the speed of light. Therefore, after each trip around the DR the protons fall 25 m behind the muons. After four circulations, the protons are 100 m behind the muons. Accounting for the approximately 120 ns temporal width of the beam pulse, the trailing edge of the muon beam and the leading edge of the proton beam will be separated by 200 ns after four DR orbits. This separation is sufficient to kick the protons into an abort channel without affecting the muon beam. Additionally, four times around the DR is approximately 12 pion decay lengths, which allows sufficient time for virtually all pions remaining in the beam to decay.

After proton removal in the DR, the remaining beam is mostly muons, but with a 30% to 40% positron contamination. The muon-positron beam is extracted from the DR into a 130 m long final stretch of beamline, called the M4 and M5 lines, which leads to the E989 storage ring. It is then focused through a small aperture bored through the storage ring magnet iron and into an apparatus called the Superconducting Inflector Magnet, or the inflector [49]. The magnetic field inside of the inflector is constructed to cancel the fringe field of the storage magnet. Without it, muon injection into the storage ring would not be possible because passage through the fringe field would deflect the beam into the magnet iron.

The beam exits the downstream end of the inflector inside of the E989 storage ring vacuum. The exit of the inflector is displaced 77 mm radially outward from the center of the storage region. Particles this far displaced from the center of the storage region cannot be stored: the storage ring is designed to store particles in a 45 mm radius around the center of the storage region. To store the particles that enter the storage ring vacuum, a pulsed kicker magnet fires as the muon beam crosses the ideal orbit  $90^\circ$  around the ring from the injection site. The kicker pulse length is ideally less than the muon orbital period, 149 ns, so that each particle is only kicked once. The ideal kick is exactly strong enough to deflect design-momentum muons onto the ideal orbit. Synchrotron radiation, though negligible for

muons, causes the loss of all beam positrons within  $4\mu\text{s}$  of injection [1, 50].

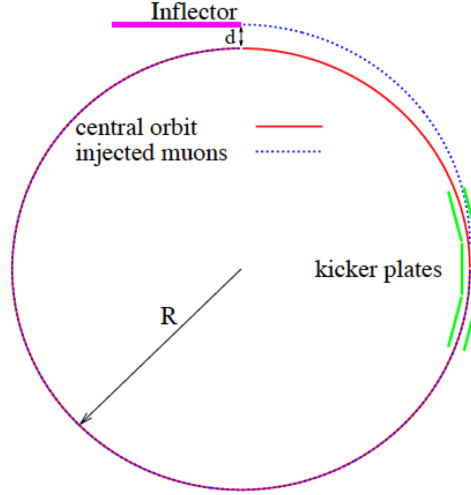


Figure 2.6: Diagrammatic illustration of injection into the storage ring, reproduced from [37]. During the first turn after injection, the kickers provide the deflection necessary to place muons onto the central storage orbit.

Particles exit the inflector with a range of radial angles,  $x'$ , defined as the angle between the radial momentum and the longitudinal momentum. To sufficient approximation,  $x' = p_x/p_z$  where  $x$  is the radial direction and  $z$  is the longitudinal direction. This approximation is valid because  $p_z$  is much larger than  $p_x$ . One can define the vertical angle,  $y'$ , in the same way. Particles exiting the inflector with  $x'$  between  $-5\text{ mrad}$  and  $5\text{ mrad}$  will cross the ideal orbit with  $x'$  between  $-10.8\text{ mrad}$  and  $-12\text{ mrad}$ . The minimum of  $-10.8\text{ mrad}$  occurs for particles exiting the inflector with  $x' = 0$ , or exactly tangent to the ideal orbit. The ideal kick would provide an  $x'$  impulse between  $10.8\text{ mrad}$  and  $12\text{ mrad}$  and place the injected muons onto the design orbit with no radial deflection. As the muons exit the inflector with a range of  $x'$ , not all can be kicked ideally. The required magnetic field for a given deflection angle,  $\Delta x'$ , is approximately

$$B_{kick} = -B_0 \frac{\Delta x'}{\theta_{kick}}, \quad (2.14)$$

where  $\theta_{kick}$  is the azimuthal angle in the storage ring over which the kick is delivered.

The E989 kicker spans approximately  $45^\circ$  in the storage ring, so to deflect muons between 10.8 mrad and 12 mrad the kicker must provide a magnetic field in the range from 200 G to 250 G. A diagrammatic illustration of the injection process is shown in Figure 2.6.

## 2.5 *Magnetic Field Measurement*

The magnetic field measurement consists of all the experimental components, techniques, and equipment used to measure  $\tilde{\omega}_p$ , the muon distribution weighted average of the free proton Larmor precession frequency in the storage ring. It is related to  $a_\mu$  through Equation 2.5; the precision of E989's final  $a_\mu$  measurement can be no better than the precision of the  $\tilde{\omega}_p$  measurement. Thus, the magnetic field measurement and the precession frequency measurement are equally important.

The pulsed nuclear magnetic resonance (NMR) technique is used throughout the magnetic field measurement. In pulsed NMR, the net magnetization of a sample of protons in a material such as water or petroleum jelly is rotated  $90^\circ$  relative to an external magnetic field using a radio frequency (RF) pulse called a  $\pi/2$  pulse. Subsequently, the proton sample's magnetization will precess about the external magnetic field until the magnetization relaxes back into its equilibrium orientation, which is aligned with the external magnetic field. Pickup coils oriented perpendicular to the external field are connected to waveform digitizers that record the current induced in the coils by the precessing protons. The resulting signal is called a free induction decay. NMR probes contain the proton samples, the pickup coils, and the necessary circuitry to read out the free induction decay signal. The same RF coil is used to deliver the  $\pi/2$  pulse and to read out the free induction decay. Each free induction decay signal yields a 10 ppb precision measurement of the proton precession frequency in the probe. The statistical uncertainty of the  $\tilde{\omega}_p$  measurement is entirely negligible [37, 51].

The magnetic field in the muon storage region is intermittently measured by sending a trolley holding 17 NMR probes around the inside of the storage region. Rails to guide the trolley are built into the vacuum chambers for this purpose. The trolley measurements produce a three dimensional map of the magnetic field inside the storage region. Knowledge

of the spatial variations of the field are necessary to correctly calculate  $\tilde{\omega}_p$ . The trolley, by necessity, cannot be present in the storage region during muon storage. If it were, all the muons would be rapidly lost from scattering. During E989 data taking, storage ring usage will alternate between trolley runs and muon storage runs. Figure 2.7 shows an azimuthally averaged map of the magnetic field in the muon storage region as measured by the trolley.

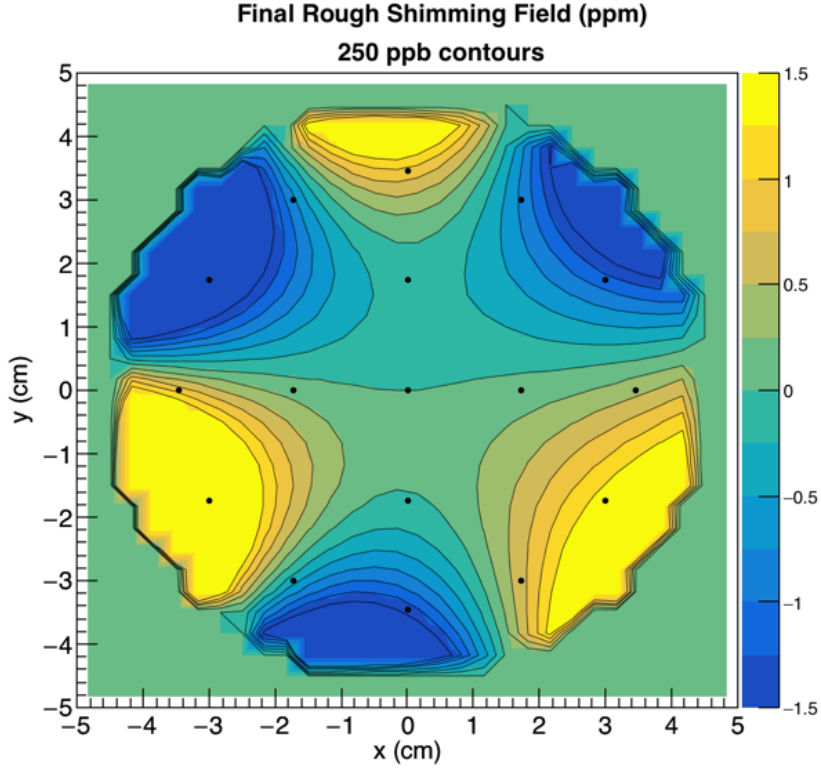


Figure 2.7: Azimuthally averaged magnetic field contours from late summer 2016, courtesy of Rachel Osofsky. The azimuthally averaged field nonuniformities are on the order of 1 ppm.

Another suite of NMR probes is used to interpolate between trolley runs. These are the 378 fixed NMR probes placed around the storage ring. They are on the outside of the vacuum chambers but inside the high magnetic field region. The fixed probes are operated concurrently with muon storage and measure the storage ring field at least once per accelerator super cycle during E989 data taking runs. They allow for any time variations in the mag-

netic field to be detected and accounted for in the  $\tilde{\omega}_p$  analysis. Additionally—being placed symmetrically around the ring—they provide spatial information regarding field variations, though not the spatial variations inside of the storage region itself.

The precession frequency measured by the NMR probes is not quite equal to the free proton precession frequency needed to calculate  $a_\mu$ . The frequency measured in the probes receives corrections from the molecular properties of the proton sample, the magnetization of the probe materials, and the shape of probe sample. The general expression for an NMR probe with a water sample is [51]

$$\omega_{probe} = [1 - \sigma(H_2O, T) + \delta_b + \delta_p + \delta_s] \omega_p, \quad (2.15)$$

where  $\sigma$  is the temperature dependent diamagnetic shielding, and the  $\delta$ 's are corrections from the bulk susceptibility, paramagnetic impurities, and magnetization of the probe, respectively. A calibration is required to relate the probe precession frequency to the free proton precession frequency. Much of the experimental uncertainty of the magnetic field measurement comes from the systematic uncertainty introduced by the calibration procedure.

A special NMR probe is used for absolute calibration of the field measurement. The total correction needed to relate the frequency measured by this absolute calibration probe to the free proton precession frequency is known to 35 ppb. By inserting the absolute calibration probe into the storage region and measuring the field in the same location as the central trolley probe, the absolute calibration is transferred to the central trolley probe. Yet another special probe, called the plunging probe, transfers the calibration from the central trolley probe to the 16 other trolley probes. The plunging probe is connected to a 3D motion system that can position it in the same location as any of the 17 trolley probes. The offsets between the trolley probes and any given fixed probe are determined when the trolley passes the location of a fixed probe.

Other than the 35 ppb uncertainty of the absolute calibration, additional significant sources of systematic uncertainty in the  $\tilde{\omega}_p$  measurement include the trolley probe calibration

Table 2.2: Known sources of systematic uncertainty for the E989 magnetic field measurement. The target combined systematic uncertainty is 70 ppb.

Source of uncertainty	Target value [ppb]
Trolley measurements	30
Trolley probe calibration	30
Fixed probe interpolation	30
Muon distribution weighted average	10
Time dependent external fields, others	10
Absolute calibration probe correction	10
Total quadrature sum	70

procedure, the fixed probe time interpolation, and uncertainties in the trolley’s field measurement originating from field gradients and trolley position uncertainties. See Table 2.2 for a full accounting of the expected  $\tilde{\omega}_p$  systematic uncertainties. The total uncertainty budget for the magnetic field measurement is 70 ppb systematic uncertainty and negligible statistical uncertainty.

## 2.6 Anomalous Precession Frequency Measurement

The lab frame decay positron energy spectrum depends on the angle between the muon spin and the muon momentum. For a population of muons stored in the E989 storage ring, the overall positron energy spectrum depends on the average of this angle over the whole population. With a beam of muons polarized at injection time, the anomalous precession—the changing angle between the muon spin and momentum—will be coherent and the observed positron energy spectrum will oscillate. The anomalous precession frequency,  $\omega_a$ , is extracted by measuring the period of this oscillation.

When a muon decays, the produced decay positron has less energy than its parent



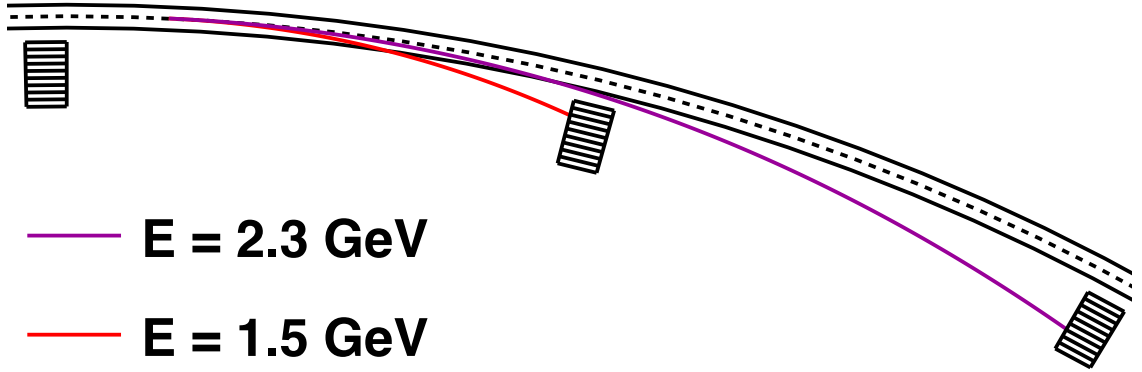


Figure 2.8: Illustration of muon decay in the storage ring. The black curves delineate the storage region, and the dashed line is its center. The rectangular objects represent the electromagnetic calorimeters. When a muon decays in the storage region, its decay positron will curl inward and may strike a calorimeter. The red and magenta curves show the paths that decay positrons of two different energies would take if they were born at the same location with momenta tangent to the ideal orbit. In the legend,  $E$  refers to the decay positron energy. Higher energy decay positrons travel farther before encountering a calorimeter.

muon and correspondingly a smaller orbit radius in the storage ring's magnetic field; decay positrons curl toward the center of the ring. There are 24 electromagnetic calorimeters stationed around the inner radius of the storage ring to intercept these decay positrons and measure their energies and hit times. This process is illustrated in Figure 2.8. With 24 calorimeters equally spaced azimuthally, the probability that a decay positron will hit a calorimeter—called the acceptance—becomes very high, about 80%, for the decay positrons nearest the endpoint of the energy spectrum [37]. These high-energy decay positrons carry the most signal for the precession frequency measurement because their emission probability is the most highly modulated at  $\omega_a$ . After each muon injection, the calorimeters record data for  $700\ \mu\text{s}$ . The boosted muon lifetime is  $64.4\ \mu\text{s}$ ; by the end of the  $700\ \mu\text{s}$  measurement period, virtually all of the stored muons have decayed. The calorimeters will be discussed in detail in Chapter 4.

There are a number of ways to extract  $\omega_a$  from the information collected by the calorime-

ters, but a straightforward and effective way, the way that will be discussed here, is to histogram decay positron hit times and to apply energy dependent weights in such a way as to enhance the  $\omega_a$  oscillations. For example, an analysis technique called the T-Method, or threshold method, uses a weight of 1 for positrons above a certain energy threshold and 0 for positrons below that threshold. Because the probability a muon will decay to a high-energy positron depends on the angle between its spin and momentum, the number of detected positrons over a non-zero energy threshold will oscillate at  $\omega_a$ . With energy dependent weights  $w(E)$ , the functional form describing the histogram produced using such a technique is

$$\begin{aligned}
 f(t) &= N \int_0^{E_{max}} w(E) \frac{d^2 P}{dE dt} dE \\
 f(t) &= \int_0^{E_{max}} N_0 e^{-t/(\gamma\tau_\mu)} [w(E)N(E) + w(E)N(E)A(E) \cos(\omega_a t - \phi)] dE \\
 f(t) &\propto N_0 e^{-t/(\gamma\tau_\mu)} [\langle w \rangle + \langle wA \rangle \cos(\omega_a t - \phi)].
 \end{aligned} \tag{2.16}$$

Figure 2.9 is an illustration of how one constructs a T-Method histogram.

Absent any additional effects, histograms produced using the weighting technique described above can be fit with a five parameter function of the form

$$f(t) = N e^{-t/\tau} [1 + A \cos(\omega_a t - \phi)]. \tag{2.17}$$

The five parameters are the overall normalization,  $N$ , the boosted muon lifetime,  $\tau$ , the overall asymmetry,  $A$ , the initial phase,  $\phi$ , and the anomalous precession frequency,  $\omega_a$ . Parameters are extracted from the time histogram with a  $\chi^2$  minimization fit. Such a parameter extraction will have a statistical uncertainty; using the functional form above, it is possible to calculate what the uncertainty will be for a data set of a given size.

Provided the fluctuations of the  $\omega_a$  time histogram's bin contents are Gaussian, the covariance matrix  $\sigma_{ij}$  pertaining to the parameters  $p_i$  and  $p_j$  extracted from the  $\chi^2$  minimization

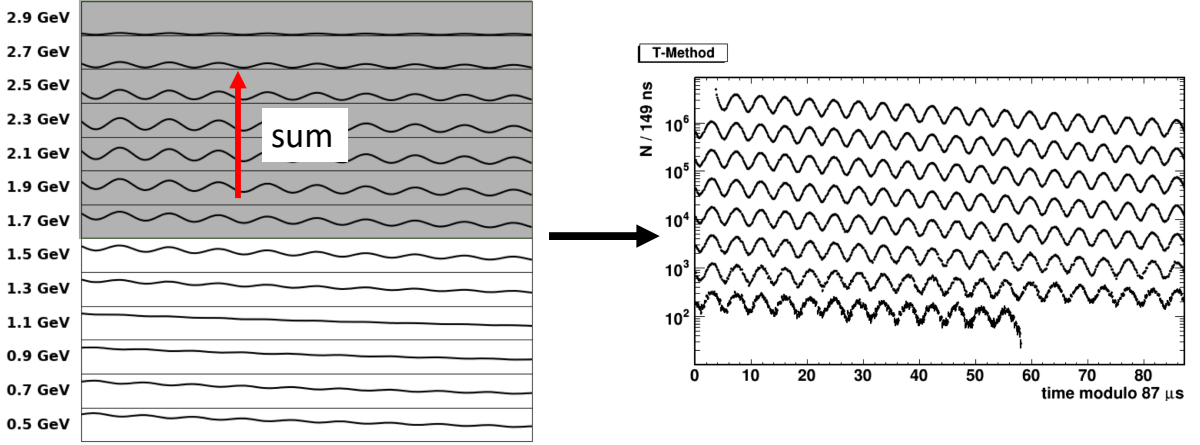


Figure 2.9: Illustration of the T-Method for  $\omega_a$  analysis. On the left are the calorimeter signals in a number of different energy windows. Summing the signals from all energies above a certain threshold yields the one dimensional T-Method histogram shown on the right. The T-Method signal is wrapped on itself every  $87\mu\text{s}$  to show the entire  $700\mu\text{s}$  measurement period. The period of the oscillations visible in the T-Method histogram is  $2\pi/\omega_a$ . Thus, by fitting the T-Method histogram with the appropriate model, one obtains  $\omega_a$ .

procedure is given by [5, 37, 52]

$$\begin{aligned}\sigma_{ij}^{-1} &= \frac{1}{2} \frac{\partial^2 \chi^2}{\partial p_i \partial p_j} = \sum_{bins} \frac{1}{\sigma_{bin}^2} \frac{\partial f(t_{bin}, \vec{p})}{\partial p_i} \frac{\partial f(t_{bin}, \vec{p})}{\partial p_j} \\ &\approx \frac{\langle w \rangle N_{hist}}{\int f(t) dt} \int \frac{1}{\sigma^2(t)} \frac{\partial f(t, \vec{p})}{\partial p_i} \frac{\partial f(t, \vec{p})}{\partial p_j} dt,\end{aligned}\quad (2.18)$$

where

$$\sigma^2(t) = N_0 e^{-t/(\gamma\tau_\mu)} \left[ \langle w^2 \rangle + \langle w^2 A \rangle \cos(\omega_a t - \phi) \right]. \quad (2.19)$$

In the above, it is used that the variance of a bin in a weighted histogram is given by the sum of the weights squared. Additionally, in approximating the sum over histogram bins as an integral, the sum of the bin contents of the histogram is replaced with the integral of the histogram divided by the bin width. Finally, the total sum of the histogram contents is replaced with  $\langle w \rangle N_{hist}$ , where  $\langle w \rangle$  is the average weight value and  $N_{hist}$  is the number of discrete events in the histogram. The above substitutions are based on the approach of Bennett et al. [52]. The necessary integrals can be completed analytically to leading order in

small parameters  $(\gamma\tau_\mu\omega_a)^{-1}$  and  $\langle wA \rangle^2/\langle w \rangle^2$ , and from there the statistical precision  $\sigma_{\omega_a}/\omega_a$  can be determined. The result is [37, 52]

$$\frac{\sigma_{\omega_a}}{\omega_a} = \sqrt{\frac{2\langle w^2 \rangle}{N_{hist}\langle wA \rangle^2\gamma^2\tau_\mu^2\omega_a^2}}. \quad (2.20)$$

Any weighting function can be used in tandem with an energy threshold; positrons reconstructed below that threshold are not included in the histogram. Potential weighting functions include  $w(E) = E$ , or energy weighted,  $w(E) = A(E)$ , or asymmetry-weighted, and  $w(E) = 1$ , the pure threshold method with no weighting. All these analyses will be conducted on the E989 data set. Neglecting detector effects entirely, the asymmetry-weighted technique provides the smallest uncertainty in the  $\omega_a$  extraction with a 30% lower statistical uncertainty than a fully optimized non-weighted analysis, see Figure 2.10. In fact, with the approximations employed here, the asymmetry-weighted method can be shown to be statistically optimal [8, 52]. When detector effects such as acceptance are included, the statistical gain of the asymmetry-weighted technique is more modest, only about 10% over the threshold technique. As the threshold method provides nearly as good statistical precision as the optimal method while also having the advantage of being simple and relatively straightforward to characterize, it is considered the baseline analysis technique for the  $\omega_a$  extraction in the E989 experiment.

For a pure threshold method with  $w(E) = 1$ , Equation 2.20 reduces to

$$\frac{\sigma_{\omega_a}}{\omega_a} = \frac{\sqrt{2}}{\sqrt{N_{hist}A\gamma\tau_\mu\omega_a}}. \quad (2.21)$$

Here  $A$  refers to its ultimate value in the final  $\omega_a$  histogram fit and includes the effects of averaging the asymmetries of all particles over the chosen threshold and the less than 100% polarization of the injected muon beam. Solving for  $N_{hist}$ , one obtains an equation for the required number of particles in the histogram as a function of the target statistical  $\omega_a$  precision:

$$N_{hist} = \frac{2}{\epsilon_{\omega_a}^2 A^2 \gamma^2 \tau_\mu^2 \omega_a^2}, \quad (2.22)$$

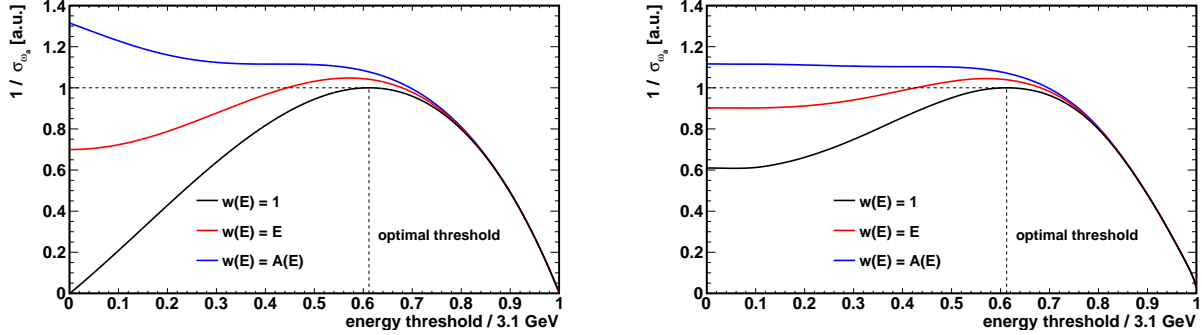


Figure 2.10: Relative  $\omega_a$  precisions obtained with different weighting methods and energy thresholds, calculated using Equation 2.20. A model for detector acceptance is included in the right graph but not the left. The optimum energy threshold for a given weighting function is the energy threshold maximizing these curves. Dashed lines show the relative precision obtained by a pure threshold method with the optimal threshold of 1.9 GeV. While the asymmetry-weighted technique at 0 threshold always provides the best precision, the inclusion of detector acceptance effects brings the pure threshold method precision within 15% of the asymmetry-weighted method. The pure threshold method has the advantage of being simpler to characterize and control.

where  $\epsilon_{\omega_a}$  is  $\sigma_{\omega_a}/\omega_a$ .

The best threshold to optimize statistical precision is 1.9 GeV. The average asymmetry for particles above 1.9 GeV is approximately 0.4, but this scales linearly with the polarization. At the expected 95% polarization of the E989 injected beam, the asymmetry will be 0.38. The other values required to carry out this calculation are shown in Table 2.3. The calculation yields a required 160 billion, or  $1.6 \times 10^{11}$  over threshold detected particles. In other words, assuming a pure threshold analysis and an optimal energy threshold of 1.9 GeV, there must be at least 160 billion entries in the decay positron hit-time histogram to fit  $\omega_a$  with a statistical precision of 100 ppb. The next section will outline how these 160 billion high-energy decay positrons will be obtained.

Rather than forming a single weighted positron hit-time histogram, an alternative approach to the  $\omega_a$  analysis forms a family of non-overlapping and non-weighted hit-time histograms each containing positrons in the energy range  $E$  to  $E + \Delta E$ . Such an analysis is

Table 2.3: The quantities used to determine the number of particles needed in the full E989 dataset to meet the experiment’s statistical precision target of 100 ppb. The asymmetry value takes into account the 95% polarization of the E989 muon beam. With these values as input, Equation 2.22 yields a value of 160 billion required particles.

Quantity	Value
Target statistical precision, $\epsilon_{\omega_a}$	100 ppb
Asymmetry, $A$	0.38
Muon lifetime, $\tau_\mu$	2.2 $\mu s$
Stored muon Lorentz factor, $\gamma$	29.3
Anomalous precession frequency, $\omega_a$	1.44 MHz
Number of particles in $\omega_a$ histogram	160 billion

referred to as an energy-binned analysis. An example of an energy-binned analysis applied to the E821 data can be found in the PhD dissertation of Frederick Gray [53]. In an energy-binned analysis, each of the sub histograms is fit separately using the five parameter function given in Equation 2.17. The result is a separate  $\omega_a$  value for each positron energy range,  $\omega_{a,E}$ . As the sub histograms contain completely non-overlapping sets of positron hits, the statistical fluctuations contributing to each  $\omega_{a,E}$  are independent and the  $\omega_{a,E}$ ’s can be averaged to obtain a single combined value,  $\langle\omega_a\rangle$ , for the whole dataset. To obtain the smallest  $\langle\omega_a\rangle$  uncertainty, the proper weight for a given  $\omega_{a,E}$  in the average is  $\sigma_{\omega_{a,E}}^{-2}$ . Thus, the more precise  $\omega_{a,E}$  values are given greater weight. With these weights, the statistical uncertainty of  $\langle\omega_a\rangle$  is

$$\frac{1}{\sigma_{\langle\omega_a\rangle}^2} = \sum \frac{1}{\sigma_{\omega_{a,E}}^2}. \quad (2.23)$$

Using Equation 2.21 and assuming the energy bin width  $\Delta E$  is small enough that  $N(E)$  and  $A(E)$  can be treated as constant,

$$\frac{1}{\sigma_{\omega_{a,E}}^2} = \frac{1}{2} \cdot N_{hist} N(E) A(E)^2 \gamma^2 \tau_\mu^2 \Delta E. \quad (2.24)$$

Table 2.4: The major categories of experimental uncertainty in the E989 experiment.

Uncertainty source	Target value
$\omega_a$ statistical uncertainty	100 ppb
$\omega_a$ systematic uncertainty	70 ppb
$\tilde{\omega}_p$ systematic uncertainty	70 ppb
Quadrature sum	$\sim 140$ ppb

This equation assumes  $N(E)$  is normalized such that  $\int N(E) dE = 1$ . With small  $\Delta E$ ,  $\sigma_{\langle\omega_a\rangle}^2$  can be calculated as follows:

$$\begin{aligned}
\frac{1}{\sigma_{\langle\omega_a\rangle}^2} &= \sum \frac{1}{\sigma_{\omega_a,E}^2} \\
&\approx \frac{1}{2} \cdot N_{hist} \gamma^2 \tau_\mu^2 \int N(E) A(E)^2 dE \\
&\approx \frac{1}{2} N_{hist} \langle A^2 \rangle \gamma^2 \tau_\mu^2.
\end{aligned} \tag{2.25}$$

Comparing this to Equation 2.20, it is apparent that—in the small  $A$  limit—the precision of an energy-binned analysis with small  $\Delta E$  is equivalent to the precision of an asymmetry-weighted analysis. The energy-binned analysis has the advantage that the asymmetry function  $A(E)$  need not be known in advance. Additionally, an energy-binned analysis maintains the same precision even if the  $\omega_a$  oscillations in different positron energy ranges have different phases. Energy dependent phases will be discussed in the next section. An energy-binned analysis requires better understanding and control of the calorimeter energy response than a pure threshold based analysis.

In addition to the 100 ppb projected statistical uncertainty of the  $\omega_a$  measurement, there is a 70 ppb systematic uncertainty budget. The systematic uncertainties of the  $\omega_a$  measurement will be discussed in detail in the next chapter. Table 2.4 enumerates the three major uncertainty categories in the E989 experiment.

## 2.7 *Required Running Time*

The time required to record 160 billion decay positrons depends on a number of factors: the number of protons impacting the pion production target per shot, the fraction of these protons that yield positive pions in the correct energy range, the fraction of these pions that decay to muons in the correct energy range, the fraction of these muons that are captured and transmitted to the E989 storage ring, the fraction of delivered muons that are stored, the fraction of stored muons that yield detected over threshold positrons, and the number of beam pulses per second. Some of these factors are determined by the capabilities of the FNAL accelerator and others by the properties of the E989 storage ring and detectors. Before they can be measured, these factors are determined with numerical calculations and simulations.

The fraction of delivered muons that are stored in the ring, called the storage fraction, is affected both by the characteristics of the delivered beam—for example its physical width—and by the performance of the kicker and the inflector. Injection parameters that must be optimized include kicker voltages and timings, and the current setting (i.e. magnetic field strength) of the inflector. As will be discussed later, it is a significant effort to optimize these parameters. The injection parameter optimization involves coordination with FNAL beam operators who, with feedback from the E989 detectors, change beamline settings to yield the highest number of stored muons. It is expected that with a fully optimized injection, 2% of the injected muons will be stored. The reason this fraction is so small is that a large number of the injected muons are outside of the momentum range that can possibly be accepted by the E989 storage ring [37].

Of the stored muons, about 11% will yield a high-energy detected decay positron. This is the product of the probability that the muon will decay to an over threshold positron and the probability that the positron will hit a calorimeter. The number of muons that will be transmitted to the E989 ring per proton incident on the pion production target has been studied carefully with multiple beamline simulation software packages: it is expected that



$7.3 \times 10^{-7}$  muons will arrive at the storage ring per proton on target. When E989 is running at full intensity,  $10^{12}$  protons on target are expected per pulse. For practical reasons that will be discussed later, the  $\omega_a$  fit does not begin until  $30 \mu\text{s}$  after the beam is injected, so the effective number of stored muons when considering the length of data taking required is a factor of about 1.6 lower than the number stored after injection. Combining these factors, 1000 detected high-energy positrons are expected per beam pulse. Therefore, 160 million beam pulses are required to obtain the necessary 160 billion high-energy positrons.

The average rate of beam injections, or fills, during full intensity data taking is 11.4 Hz. It would take just over 160 days running nonstop at 11.4 Hz to reach the 160 million fills necessary for 100 ppb statistical uncertainty. Allowing for systematic tests, down time, and activities such as magnetic field scans with the NMR trolley, it was originally expected that the E989 experiment would accumulate its full dataset with one to two years of running. The numbers and factors discussed in this section are summarized in Table 2.5. As of the end of Run 1, which occurred in the spring of 2018, the E989 experiment had recorded approximately 12 billion high-energy decay positron hits suitable for use in the  $\omega_a$  analysis. Much of the 2018 running time was devoted to commissioning, which is discussed briefly in Chapter 5.

Table 2.5: Factors contributing to the estimated running time required to accumulate the complete E989 dataset. The final estimate of 420 days is approximate, but it is expected that E989 will accumulate its full dataset with two to three years of full-rate operations.

Factor contributing to calculation	Value
Protons on target per pulse	$10^{12}$
Muons transmitted to E989 per proton	$7.3 \times 10^{-7}$
Storage probability per injected muon	0.02
Fraction of muons remaining after $30 \mu s$	0.63
Detected high-energy positrons per muon	0.11
Over threshold positrons per fill	1000
Fills needed for 160 billion positrons	160 million
Fill rate	11.4 Hz
Required data accumulation time	3900 hours
Run time per day	17 hours
Time required for data accumulation	230 days
Time required for commissioning	150 days
Time for systematic studies	50 days
Estimated total E989 running time	430 days

## Chapter 3

### SYSTEMATIC UNCERTAINTIES OF THE PRECESSION FREQUENCY MEASUREMENT

The precession frequency measurement is robust: it relies only on a time oscillation in the number of detected high-energy positrons. It is not required to have a model for the muon decay distribution, that the calorimeter acceptance as a function of energy be known, that the calorimeters have a linear energy response or be particularly well calibrated, or that the positron hit time reconstruction be absent energy-dependent biases. Differing decay distributions and calorimeter responses modify the asymmetry, phase, and normalization of the observed oscillation, but they cannot affect the frequency. Despite the robustness of the precession frequency measurement, there are effects that are not accounted for by the basic five parameter model given in Equation 2.17. These effects have the potential to bias the  $\omega_a$  extraction. For all such effects, the size of the bias is estimated. If the bias is small enough, it is considered a contribution to the  $\omega_a$  systematic uncertainty. Otherwise, a correction to either the final  $\omega_a$  value or to the positron hit time histogram is applied, and the associated systematic uncertainty is the uncertainty of the applied correction. In general,  $\omega_a$  systematic uncertainties can be classified either as beam dynamics effects or detector effects. Beam dynamics effects are caused by behaviors of the stored muon population that are not accounted for by the fit model, whereas detector effects are systematic biases imparted by the non-ideal properties of the calorimeter. This chapter outlines the known systematic effects present in the  $\omega_a$  measurement and emphasizes the performance requirements these effects place on the calorimeters.

### 3.1 *Decay to Detection*

A general description of the process by which a decay positron becomes an entry in the final hit time histogram is useful to understand the source of many  $\omega_a$  systematic uncertainties. When a muon decays in the storage region, the produced decay positron will have a smaller momentum than its parent muon. The orbit radius in the uniform magnetic field is proportional to the particle's momentum, so the lower momentum positrons will curl toward the center of the ring. The range of decay positron momenta varies from essentially 0 to a maximum value of the parent muon momentum. Accordingly, there is a great range of orbit radii that a decay positron may have. Figure 2.8 illustrated the trajectory of two decay positrons with different energies. A high-momentum decay positron has an orbit radius only slightly smaller than that of its parent muon's and thus takes a long time to exit the storage region and potentially hit a calorimeter. Low-momentum positrons curve much more sharply and exit the storage region more quickly. The drift time, or the time elapsed between the muon decay and the positron hitting the calorimeter, then varies as a function of decay positron momentum. The average drift time ranges from 1 ns at 100 MeV/ $c$  to 25 ns at 3 GeV/ $c$ . Additionally, high-energy decay positrons leave the storage region with a larger azimuthal momentum component so they are more likely to continue on to hit the face of a calorimeter: the acceptance depends on the positron momentum and increases with increasing positron momentum. The energy dependent acceptance itself serves to weight high-energy positrons more than low-energy positrons in the hit time histogram. As it has already been shown in Section 2.6 that the general form of the fit function does not depend on the weighting scheme, the momentum dependence of the detector acceptance is not harmful. The decay positron momentum dependences of the drift time and acceptance are shown in Figure 3.1.

In addition to their decay positron momentum dependence, the acceptance and drift time also depend on the parent muon's decay position and polarization. The drift time depends on the path length from the decay position to the calorimeter, and the acceptance depends on the azimuthal angle traversed by a decay positron while it is in the radial region

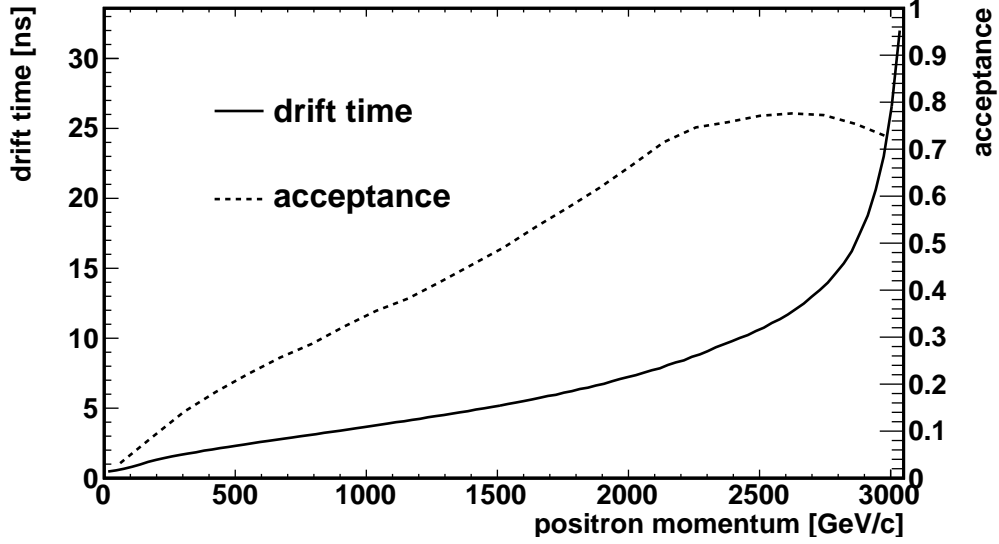


Figure 3.1: Momentum dependence of positron drift time and acceptance. Acceptance is defined as the probability that an emitted positron will hit a calorimeter. As described in the text, both drift time and acceptance increase with increasing positron momentum.

containing the calorimeters. For positrons born farther from the calorimeters, the drift time increases because the positrons have to travel farther, and the acceptance decreases because the positrons enter the calorimeter region with a smaller azimuthal momentum component. Thus, the drift time increases for positrons born farther from the calorimeters and the acceptance decreases. As the polarization direction at decay time modulates the angular distribution of emitted positrons, so does it affect the acceptance and drift time. Positrons emitted outwardly relative to the calorimeters travel longer before detection.

Decay positron drift times directly correspond to phases in the  $\omega_a$  oscillation. The probability of a certain energy positron being emitted varies with the angle between the muon spin and momentum at decay time, but the positron is measured when it hits the calorimeter at a slightly later time. For a given energy positron with drift time  $t_d(E)$ , the detected number as a function of hit time rather than decay time is

$$N(E, t) = N(E) e^{-(t-t_d(E))/\tau} [1 + A(E) \cos(\omega_a(t - t_d(E)) - \phi)]. \quad (3.1)$$

A drift time of  $t_d$  corresponds to a phase offset of  $\omega_a t_d$ . For drift times in the range of 1 ns to 25 ns, the range of phase offsets is 1.4 mrad to 35 mrad. In summing over all energies, the net effect is a small change in the initial phase and a very slight dilution of the asymmetry. In an energy-binned analysis, there is no sum over energies and thus the dilution of the asymmetry is much smaller. Modifications to the observed initial phase and average asymmetry are not problems because the form of the fit function is unchanged and, regardless of the varying phase offsets, the oscillation frequency is the same for all energies. It will be shown that biases can arise, however, if drift times change systematically over the 700  $\mu$ s measurement period.

When a decay positron hits a calorimeter, the calorimeter reports an energy and a time. The reported energy and time will be the output of reconstruction software applied to a recorded electrical signal and will not be exactly equal to the true positron hit time and energy. For example, there will be a detector energy resolution and time resolution, and the energy response of the calorimeter is not be completely linear. Energy nonlinearities cannot bias the  $\omega_a$  measurement as they are equivalent to changing the energy dependent weighting scheme. Energy dependent time offsets are also not a problem: they are equivalent to the drift time effect discussed above. Again, however,  $\omega_a$  biases can arise if the detector response changes in a systematic way over the 700  $\mu$ s fill.

### **3.2 Precession Frequency Bias from a Changing Phase**

Recall that the precession frequency measurement relies on fitting the measured decay positron time distribution with the function

$$f(t) = N e^{-t/\tau} [1 + A \cos(\omega_a t - \phi)], \quad (3.2)$$

where  $\phi$  is the stored muon population's average anomalous precession phase at  $t = 0$ , accounting for the effects discussed above. Consider now replacing  $\phi$  with a time dependent  $\phi(t)$ . How the stored muon population's average phase could come to change over time will be discussed in more detail later. Expressing  $\phi(t)$  as a power series  $\phi_0 + \frac{d\phi}{dt}t + \mathcal{O}(\frac{d^2\phi}{dt^2})$ , with

the derivatives evaluated at  $t = 0$ ,

$$\begin{aligned}
 \cos(\omega_a t - \phi) &= \cos \left[ \omega_a t - \phi_0 - \frac{d\phi}{dt} t + \mathcal{O}\left(\frac{d^2\phi}{dt^2}\right) \right] \\
 &= \cos \left[ \left( \omega_a - \frac{d\phi}{dt} \right) t - \phi_0 + \mathcal{O}\left(\frac{d^2\phi}{dt^2}\right) \right] \\
 &= \cos \left[ \omega'_a t - \phi_0 + \mathcal{O}\left(\frac{d^2\phi}{dt^2}\right) \right]
 \end{aligned} \tag{3.3}$$

One finds an observed precession frequency of  $\omega'_a = \omega_a - d\phi/dt$ . The linear component of any phase change over a muon fill is indistinguishable from a precession frequency shift of  $d\phi/dt$ . Many of the systematic effects present in the precession analysis can be understood through their effect on the observed initial phase, and any effect that changes the observed phase is a potential systematic error. With  $\omega_a$  approximately equal to  $1.4 \text{ rad}/\mu\text{s}$ , a phase change of  $0.07 \text{ mrad}$  over a  $700 \mu\text{s}$  fill would bias  $\omega_a$  by more than the  $70 \text{ ppb}$  systematic error budget. Accordingly, an uncorrected shift in the average positron drift time or detector response of  $50 \text{ ps}$  over the  $700 \mu\text{s}$  muon fill would exceed the systematic error budget. Therefore, the muon population's average initial phase and the detector time response must be understood and controlled to better than this level to meet the goals of the E989 experiment. Phase changes that are not linear in time cannot be correctly fit with the simple five parameter model and thus necessitate adjustments to the fit function.

It is worth emphasizing that the sorts of phase changes leading to  $\omega_a$  biases are those that reproducibly occur over the time scale of a muon fill, in the same way fill after fill. If somehow the injected muon beam's average phase were drifting slowly over the lifetime of the experiment, the drift—if uncorrected—would somewhat reduce the statistical precision of the measurement, but it would not directly bias the extracted frequency. On the other hand, if the reconstructed positron hit times had rate dependent offsets, every single fill the positrons detected early on would be offset in time differently from those detected later, and thus the observed phase would shift in the same way each and every fill. Such effects are referred to as early-to-late effects and are particularly concerning because of their potential to bias the extracted precession frequency. Limiting early-to-late effects was among the highest

E989 design priorities. The following are some potential early-to-late effects:

- beam decoherence
- muon losses from the storage ring
- longer lifetimes for higher energy muons
- rate dependent energy and time reconstruction
- detector dead time.

### ***3.3 Systematic Effects from Beam Dynamics***

The five parameter model discussed to now assumes that stored muons are all at the design momentum, are all on perfectly circular orbits in the exact center of the storage ring, and are uniformly distributed in azimuth. Unsurprisingly, these assumptions must be relaxed in a real experiment. In practice, the ring stores a range of momenta. The particle distribution inside the storage region is complex and affected by the kicker strength and the electrostatic quadrupole voltages. Individual muons oscillate about their equilibrium orbits. These realities combine to produce rich phenomena that one must understand to correctly model the experiment. Effects rooted in these phenomena are categorized as beam dynamics. Some examples follow.

#### *3.3.1 Muon Dynamics in the Storage Ring*

A discussion of muon motion in the storage ring is required before proceeding with examples of systematic effects from beam dynamics. Stored muons are subject to the magnetic field from the storage ring and the electric field from the electrostatic quadrupoles. This section will work under the approximation of ideal fields, where the magnetic field is uniform and the electric field is independent of azimuth and linear in both the radial and vertical displacement from the center of the storage region. There are a number of known deviations from



these ideal conditions, including discrete quadrupole plates, higher order multipoles, and azimuthal magnetic field variations. Nevertheless, studying the ideal conditions is sufficient for understanding the origin of many beam dynamics effects.

In cylindrical coordinates, where  $\hat{y}$  points upwards and  $\hat{\theta}$  points counterclockwise, a muon position  $\mathbf{x}$  is expressed as  $\rho\hat{\rho} + y\hat{y}$ . With  $R_0$  as the center of the electrostatic quadrupoles and  $k$  a scaling constant for the quadrupole field strength, the ideal fields for positive muon storage are

$$\mathbf{B} = B_0\hat{y}, \quad (3.4)$$

$$\mathbf{E} = k(\rho - R_0)\hat{\rho} - ky\hat{y}. \quad (3.5)$$

Also considering the Lorentz force and Newton's second law,

$$\begin{aligned} \dot{\mathbf{p}} &= e(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B}) \\ \ddot{\mathbf{x}} &= \frac{e}{m\gamma}(\mathbf{E} + \dot{\mathbf{x}} \times \mathbf{B}) \end{aligned} \quad (3.6)$$

In the above equation,  $\gamma$  is treated as unchanging. This is justified as the changes in energy as a muon moves through the approximately 20 kV quadrupole potential are very small compared to stored muon energy, which is about 3 GeV. Combining Newton's second law with the expression for the ideal fields gives three differential equations of motion, one for each vector component in cylindrical coordinates:

$$\ddot{\rho} - \rho\dot{\theta}^2 = \frac{e\rho\dot{\theta}B_0}{m\gamma} + \frac{ek(\rho - R_0)}{m\gamma}, \quad (3.7)$$

$$2\dot{\rho}\dot{\theta} + \rho\ddot{\theta} = -\frac{e\dot{\rho}B_0}{m\gamma}, \quad (3.8)$$

$$\ddot{y} = -\frac{eky}{m\gamma}. \quad (3.9)$$

The first of these equations relates the radial acceleration to the magnetic field, tangential velocity, and the radial quadrupole electric field. The second relates the tangential acceleration to the radial velocity and the magnetic field. The third expresses a linear restoring force

in the vertical direction, coming from the vertically focusing quadrupole field. One simple solution is

$$\begin{aligned}\rho &= R_0, \\ \dot{\theta} &= \omega_c = -\frac{eB_0}{m\gamma}, \\ y &= 0\end{aligned}$$

This solution describes a muon stored at the center of the quadrupole potential moving in a perfect circle around the ring. As this ideal muon has only a tangential (often called longitudinal) component to its velocity, the relation  $\rho\dot{\theta} = v$  holds. This defines a specific momentum,  $p_0 = eR_0B_0$ , for which the ideal storage condition is possible. The E989 storage ring is designed such that  $p_0$  is  $3.094\text{ GeV}/c$ , the momentum at which the electric field correction to the anomalous precession frequency vanishes.

The differential equation governing vertical motion has the form of a simple harmonic oscillator. Rewritten in terms of the cyclotron frequency,  $\omega_c$ ,

$$\begin{aligned}\ddot{y} &= -\left(\frac{kR_0}{vB_0}\right)\omega_c^2 y \\ n &\equiv \frac{kR_0}{vB_0},\end{aligned}\tag{3.10}$$

$$\ddot{y} = -n\omega_c^2 y.\tag{3.11}$$

The quantity  $n$  is called the field index and, as will be seen, is a useful metric for the strength of the quadrupole fields [37]. The above equation indicates that vertical perturbations from the center of the quadrupole potential will result in simple harmonic oscillation in  $y$  with angular frequency  $\omega_y = \sqrt{n}\omega_c$ . This is called vertical betatron oscillation.

To understand the effect of horizontal perturbations from the ideal orbit, consider solutions of the form

$$\begin{aligned}\dot{\theta} &= -\omega_c + \dot{\delta\theta}, \\ \rho &= R_0 + x.\end{aligned}$$

The quantities  $\delta\theta$  and  $x$  are functions of time and represent small deviations from the ideal trajectory described above. As the storage ring aperture is only about 45 mm in radius and  $R_0$  is 7.112 m,  $x$  represents a less than 1% perturbation to  $R_0$ . Accordingly, higher powers of the perturbations  $x$  and  $\delta\theta$  can be neglected. Using the above solutions in the equations of motion yields

$$\ddot{x} = -\left(\omega_c^2 - \frac{ek}{m\gamma}\right)x + \mathcal{O}(x^2)$$

Dropping the higher order terms and rewriting in terms of  $n$  gives

$$\ddot{x} = -(1-n)\omega_c^2 x. \quad (3.12)$$

This equation indicates that small radial perturbations from the ideal orbit result in simple harmonic oscillation in the radial direction with angular frequency  $\omega_x = \sqrt{1-n}\omega_c$ . This is called horizontal betatron oscillation. Additionally, it is apparent that for  $n$  values larger than 1, there is no restoring force in the radial direction. In this case, particles cannot be stored. Thus, there is a limit on the quadrupole strength beyond which the horizontally defocussing effect becomes too large. Typical  $n$  values in E989 are approximately 0.1, well away from this limit.

As mentioned above, the ideal orbit at the center of the quadrupole potential is only possible for muons at the specific momentum  $p_0$ . Muons at different momenta will have equilibrium orbits at radii other than  $R_0$ . The equilibrium radius for an off-momentum muon ( $p \neq p_0$ ) can be determined by finding the radius at which the centripetal acceleration will result in uniform circular motion. As the muon will not be at the center of the quadrupole potential, this calculation must include both the magnetic field and the radial electric field. Defining  $x_{eq}$  such that  $R_{eq} = R_0 + x_{eq}$ ,

$$\frac{v^2}{R_0(1+x_{eq})} = \frac{e}{m\gamma} (vB_0 - kx_{eq}).$$

Defining the momentum deviation  $\Delta p$  as  $p - p_0$ , this can be rearranged into

$$1 + \frac{\Delta p}{p_0} = \left(1 - n \frac{x_{eq}}{R_0}\right) \left(1 + \frac{x_{eq}}{R_0}\right)$$

Thus, the leading order expression for the equilibrium radius as a function of the momentum deviation is

$$x_{eq} = \frac{R_0}{1-n} \left( \frac{\Delta p}{p_0} \right). \quad (3.13)$$

Higher momentum muons are stored at higher radii, and lower momentum muons at lower radii. For a given momentum, the quadrupole field index determines the equilibrium radius. At higher quadrupole settings, i.e. larger  $n$  values, off-momentum muons are pushed farther from the center of the storage region. This is a manifestation of horizontal defocussing.

The cyclotron frequency also depends on the momentum deviation. As the muons in the E989 experiment are highly relativistic, with  $\gamma = 29.3$ , their velocities are very close to  $c$  and do not change much with changing momentum. The stored muon momentum spread is on the order of 0.1%; the velocity spread is on the order of 1 ppm. Thus, for all muons the expression  $R_{eq}\omega_c = v$ , with constant  $v$ , is true to the part-per-million level. Using the equation for equilibrium radius as a function of momentum deviation, the above relation gives

$$\omega_c = \omega_{c,0} \left[ 1 - \frac{(\Delta p/p_0)}{1-n} \right], \quad (3.14)$$

where  $\omega_{c,0}$  is the cyclotron frequency for on-momentum ( $p = p_0$ ) muons.

Off-momentum muons also oscillate about their equilibrium orbits. Their horizontal and vertical oscillation frequencies are close to horizontal and vertical betatron frequencies derived in this section. These oscillation frequencies do have a momentum dependence though. The horizontal and vertical betatron frequencies can be written in the form

$$\begin{aligned} \omega_x &= Q_x \omega_c \\ \omega_y &= Q_y \omega_c. \end{aligned}$$

The quantities  $Q_x$  and  $Q_y$  are called the horizontal and vertical tunes. Recall they are equal to  $\sqrt{1-n}$  and  $\sqrt{n}$ , respectively, for on-momentum muons. The betatron frequencies of off-momentum muons differ from those of on-momentum muons both because  $\omega_c$  is momentum dependent and because the tunes are momentum dependent. The linear dependence of the

tunes on  $(\Delta p/p_0)$  is called the chromaticity, defined such that [54]

$$Q_{x/y}\left(\frac{\Delta p}{p_0}\right) = Q_{x/y} + Q'_{x/y} \cdot \left(\frac{\Delta p}{p_0}\right), \quad (3.15)$$

where  $Q'_{x/y}$  is the  $x$  or  $y$  chromaticity. Correctly calculating the  $x$  and  $y$  chromaticities requires moving beyond the idealized assumptions used in this section. The relative spread in betatron frequencies is the same order of magnitude as the momentum spread, i.e. a 0.1% momentum spread leads to a betatron frequency spread on the order of 0.1%.

To summarize, muons in the E989 storage ring will undergo simple harmonic motion about their equilibrium orbits with oscillation frequencies determined by the quadrupole field index,  $n$ . Muon equilibrium positions are all centered vertically in the quadrupole potential, but their radii are momentum dependent. High-momentum muons are stored at larger radii and have smaller cyclotron frequencies. The spread in equilibrium radius also depends on the quadrupole field index. Stronger quadrupole fields defocus the beam and push off-momentum muons farther from the center of the storage region. Individual stored muons undergo oscillations of the form

$$y(t) = A_y \cos(Q_y \omega_c t - \phi_y), \quad (3.16)$$

$$x(t) = x_{eq} + A_x \cos(Q_x \omega_c t - \phi_x), \quad (3.17)$$

$$Q_y = \sqrt{n} + \mathcal{O}\left(\frac{\Delta p}{p_0}\right), \quad (3.18)$$

$$Q_x = \sqrt{1-n} + \mathcal{O}\left(\frac{\Delta p}{p_0}\right). \quad (3.19)$$

More information regarding muon dynamics in the storage ring can be found in the references [37, 54–57].

### 3.3.2 Differential Decay

The stored muon momentum spread results in a range of decay lifetimes in the laboratory reference frame. Time dilation dictates that the observed lifetime will be  $\gamma\tau_\mu$ , where  $\tau_\mu$  is the muon lifetime of about  $2.2\mu\text{s}$  [5]. Neglecting the ppm level changes in velocity,  $\gamma$  is directly

proportional to the momentum,  $p$ . As low-momentum muons are likely to decay earlier, the average stored muon momentum increases over time. The equilibrium storage radius is proportional to momentum, so the average storage radius also increases over time. Positron drift times correspond to phase offsets in the observed precession signal, and drift times increase with increasing distance between the calorimeter and the muon decay vertex. Thus, as low-energy muons decay and the average beam position moves outward, the observed precession phase changes. This phenomenon is referred to as differential decay and is exactly the sort of early-to-late effect discussed earlier that can bias the  $\omega_a$  measurement.

To obtain an estimate for the size of this bias, consider an initial momentum distribution of the form  $\rho_0(p)$  with mean  $p_0$  normalized such that  $\int \rho_0(p) dp = 1$  and  $\int (p - p_0) \rho_0(p) dp = 0$ . Over time, the distribution will distort according to  $\rho(p, t) = e^{-t/(\gamma\tau_\mu)} \cdot \rho_0(p)$ . Or, written entirely in terms of momentum,  $p$ ,  $\rho(p, t) = e^{-tm_\mu v/(p\tau_\mu)} \cdot \rho_0(p)$ . This distortion comes from low-momentum muons decaying more quickly than high-momentum muons, and it will cause the stored population's average momentum to change over time. This change is given by

$$\langle \Delta p \rangle(t) = \frac{\int (p - p_0) \cdot \rho(p, t) dp}{\int \rho(p, t) dp}. \quad (3.20)$$

The storage ring can only accept muons within a narrow momentum range, so  $\rho_0(p)$  must be nearly 0 when  $(\Delta p/p_0)$  is larger than a few tenths of a percent. Thus, in the following calculation,  $(\Delta p/p_0)$  can be treated as small. Expanding the exponential about  $p_0$ ,

$$\begin{aligned} \rho(p, t) &= e^{-tm_\mu v/(p\tau_\mu)} \rho_0(p) \\ &= e^{-tm_\mu v/(p_0\tau_\mu)} \left[ 1 + \frac{tm_\mu v}{p_0} \left( \frac{\Delta p}{p_0} \right) + \mathcal{O} \left( \frac{\Delta p}{p_0} \right)^2 \right] \rho_0(p) \end{aligned}$$

Then, calculating the average momentum shift versus time,

$$\begin{aligned}\langle \Delta p \rangle(t) &= \frac{\int (p - p_0) \left[ 1 + t m_\mu v / (\tau_\mu p_0^2) (p - p_0) + \mathcal{O} \left( \frac{\Delta p}{p_0} \right)^2 \right] \rho_0(p) dp}{\int \left[ 1 + t m_\mu v / (\tau_\mu p_0^2) (p - p_0) + \mathcal{O} \left( \frac{\Delta p}{p_0} \right)^2 \right] \rho_0(p) dp} \\ \langle \Delta p \rangle(t) &\approx \frac{t m_\mu v / \tau_\mu (\sigma_p / p_0)^2}{1 + \mathcal{O} (\sigma_p / p_0)^2} \\ \langle \Delta p \rangle(t) &\approx t \cdot \frac{m_\mu v}{\tau_\mu} \left( \frac{\sigma_p}{p_0} \right)^2.\end{aligned}$$

Expressed as the derivative of average momentum with respect to time,

$$\frac{d\langle p \rangle}{dt} = \frac{m_\mu v}{\tau_\mu} \left( \frac{\sigma_p}{p_0} \right)^2. \quad (3.21)$$

So, to leading order, the stored beam's average momentum increases linearly in time. The rate of change scales quadratically with the relative momentum width.

Assuming a stored momentum width of 0.15%, the differential decay effect results in a 25 ppm average momentum shift over a 700  $\mu$ s muon fill. This is approximately 100 keV/c, or in terms of equilibrium radius, about 200  $\mu$ m. These are small numbers compared to the average momentum and radius of 3.09 GeV/c and 7.112 m, but they must be considered in the context of the precession measurement's 70 ppb total systematic uncertainty budget. Any momentum dependence of the precession phase will directly bias the observed frequency:

$$\begin{aligned}\frac{\Delta \omega_a}{\omega_a} &= -\frac{1}{\omega_a} \frac{d\langle p \rangle}{dt} \cdot \frac{d\langle \phi \rangle}{d\langle p \rangle} \\ &= -\frac{m_\mu v}{\omega_a \tau_\mu} \left( \frac{\sigma_p}{p_0} \right)^2 \cdot \frac{d\langle \phi \rangle}{d\langle p \rangle}.\end{aligned} \quad (3.22)$$

As mentioned, the observed phase will have some momentum dependence from the positron drift times. Tracking simulations report that the average drift time changes by less than 1 ns per 10 MeV change in the average stored momentum. Using the above equation, and again assuming a 0.15% momentum width, the frequency bias from changing drift times alone will be less than 10 ppb, a small and likely negligible contribution. However, a nonzero  $d\langle \phi \rangle / d\langle p \rangle$  may also arise from the phase evolution that occurs along the injection beamline.

A 1 mrad/MeV shift in the average phase per average momentum would cause a 70 ppb frequency bias. Detailed beamline simulations are necessary to evaluate the size of the effect, which may not be negligible for the final E989 analysis. An alternative discussion of the differential decay systematic error including the tracking simulation referenced above can be found in the internal E989 note by Metodiev et al. [58].

### 3.3.3 *Muon Losses*

Not all muons will remain stored until they decay. Some will inevitably be lost from the ring during the precession measurement. This could occur, for example, if a muon with a large betatron amplitude oscillates out of the storage region and scatters off an object such as a beam collimator. Betatron oscillations can be driven to larger amplitudes by field nonuniformities. Muon losses do not intrinsically bias the precession frequency measurement, although they do complicate the fitting procedure by distorting the observed hit time spectrum. Biases can occur, however, if the loss probability is correlated with the precession phase at injection time. Consider, for example, muons produced by pion decays in the target station before pion momentum selection. These muons will have a different average spin direction than those produced after pion momentum selection and also may have different average injection angles and position offsets. If these muons are more or less likely to be lost than the muons produced along the M2 and M3 lines, the stored muon population's average phase will change over time. This reasoning applies generally to muons produced at different locations along the injection beamline, and, again, careful beamline simulations are conducted to determine the degree to which muon losses result in a time dependence of the stored muon population's phase [59].

### 3.3.4 *Beam Debunching and Fast Rotation*

The five-parameter fit model implicitly assumes that stored muons are uniformly distributed in azimuth. This is in fact not the case. Muons are injected through the inflector channel, which is at a fixed location. The temporal extent of the injection is less than one cyclotron



period, and the particle intensity varies throughout the injection pulse. The result is a spatially localized beam bunch that travels around the storage ring. Only detectors near the bunch at a given time will observe positron decays; event rates at each detector vary with the cyclotron period. Within the E989 collaboration, this effect is called fast rotation. It is the highest physical frequency present in the detector signals, where physical means not an artifact of the readout electronics.

To develop some intuition, consider an injection pulse narrow enough in time to be treatable as a delta function. In this case, at the injection time  $t = 0$ , the beam is localized to a single azimuthal location. The pulse will be allowed a finite momentum width and, correspondingly, will have a range of cyclotron frequencies. This range of frequencies will cause the beam to spread out over time, and eventually the highest frequency muons at the head of the bunch will catch up to the lowest frequency muons at the tail. After many orbits, the beam will become more or less uniformly distributed in azimuth. This process is called debunching. Although the cyclotron and spin precession frequencies are momentum dependent, as discussed in Chapter 2 the anomalous precession frequency,  $\omega_a$ , is *not* momentum dependent. Thus, the  $\omega_a$  signal remains coherent even as the muons spread uniformly throughout the storage ring.

Beam debunching can be described mathematically in the following way. The cyclotron frequency distribution will be represented by a density function,  $\rho(\omega)$ . The density of muons at an azimuthal location  $\phi$  at time  $t$ ,  $N(\phi, t)$ , is the sum of the density of muons that arrive at azimuth  $\phi$  at time  $t$  on their first turn around the ring and the density that arrive there on their second turn and the density that arrive there on their third and so on. Labeling the turn number with  $n$ , and assuming all muons are localized at  $\phi = 0$  at  $t = 0$ :

$$N(\phi, t) = \sum_{n=-\infty}^{\infty} \int \rho(\omega) \delta(\omega t - 2\pi n - \phi) d\omega.$$

This equation can be rewritten with the delta function expanded about its 0 and in terms of  $\omega' = \omega - \phi/t$  to give

$$N(\phi, t) = \frac{1}{t} \int \rho\left(\omega' + \frac{\phi}{t}\right) \sum_{n=-\infty}^{\infty} \delta\left(\omega' - n\frac{2\pi}{t}\right) d\omega'. \quad (3.23)$$

Writing the equation in this way reveals the structure  $\sum_{n=-\infty}^{\infty} \delta(\omega' - n2\pi/t)$ , a Dirac comb with period  $2\pi/t$ . Expressing the Dirac comb as a Fourier series and changing the integration variable back to  $\omega$  yields the equation

$$N(\phi, t) = \frac{1}{\pi} \sum_{n=1}^{\infty} \int \rho(\omega) \cos[n(\omega t - \phi)] d\omega + \frac{1}{2\pi} \int \rho(\omega) d\omega. \quad (3.24)$$

The above form reveals the debunching behavior. Consider, for example, the  $\phi = 0$  position. At  $t = 0$ , all the cosine terms will equal one and the sum will diverge, in accordance with the input to this model that the beam begins with a delta function density in azimuth. At later times, the cosine terms begin to go out of phase for different values of  $\omega$ . At very late times, the value of the cosine argument varies rapidly with changing  $\omega$ , and the first integral will tend to cancel for all values of  $n$ . This leaves only the second term, which is uniform in  $\phi$ . See Figure 3.3 for an example of how this might look at three different azimuthal locations. This debunching signal,  $N(\phi, t)$ , is called the fast rotation signal.

The frequency components of the injected beam determine the fast rotation signal's structure. The process can be inverted to extract the frequency distribution—and thus the momentum and equilibrium radius distributions—from the measured fast rotation signal. As will be discussed later, this fast rotation analysis is a critical piece of the precession frequency measurement. There are two independent fast rotation analysis methods currently being pursued within the E989 collaboration. One is to discretize the frequency distribution and, with the value of each frequency component as a free parameter, fit the observed fast rotation with a  $\chi^2$  minimization routine. The second is to directly project out the frequency components of the fast rotation signal with a Fourier transform. Each of these techniques has its own systematic issues and challenges; agreement between them is an important control. The details of the fast rotation analysis will not be described here, but they can be found in the references [60–62].

It must be noted that the beam delivered by FNAL to the E989 experiment is far from the delta function limit—it extends over approximately 120 ns. The above treatment can be extended to a more general case by replacing the frequency density function with a joint

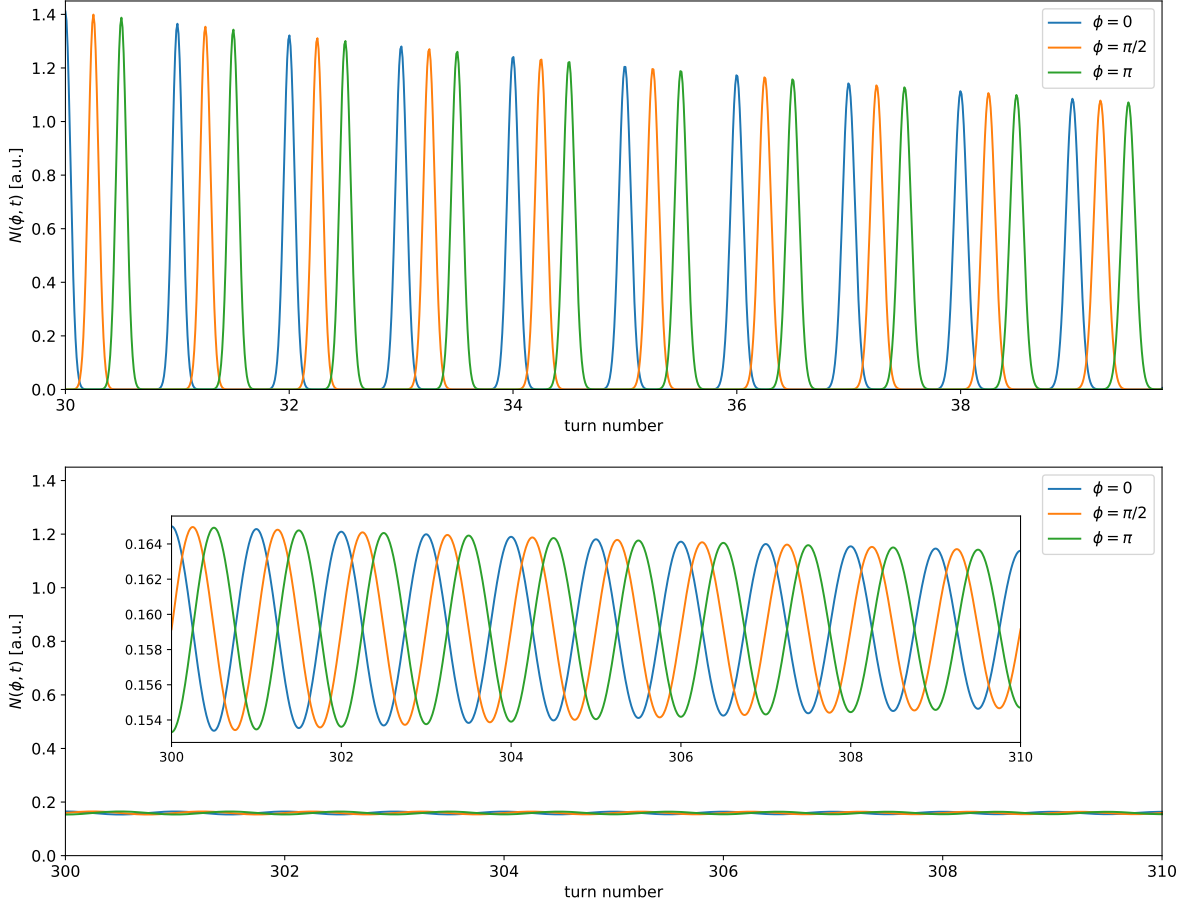


Figure 3.2:  $N(\phi, t)$ , as described in the text, for three different values of  $\phi$ . At early times, the beam is tightly grouped and appears only once per cyclotron period. The slow spreading of the bunch is evident in the decreasing peak density on each subsequent turn. As one would expect, when the beam is tightly bunched it passes different azimuthal locations at different times. At later times, the bunch has overlapped itself and the distribution is mostly uniform. A full  $g - 2$  fill is thousands of turns; the debunching effect is predominant only early on. In this illustration,  $\rho(\omega)$  is a Gaussian with a relative width of 0.15%.

density function  $\rho(\omega, t_0)$ , where  $t_0$  is the muon injection time. The integrals would then be taken over both  $\omega$  and  $t_0$ . The qualitative behavior, however, is largely unchanged. It is still true that the frequency components of the injected beam determine the debunching structure. It is an active effort in the E989 collaboration to determine how to robustly extract the frequency components of the injected beam from the measured fast rotation signal.

### 3.3.5 *Coherent Betatron Oscillations*

Calorimeters are sensitive to betatron oscillations. The probability that a decay positron will hit a calorimeter, the calorimeter acceptance, depends on the muon decay position. The average acceptance at a given time, then, depends on the distribution of muon positions at that time. In general, the overall acceptance depends on the mean, width, and higher moments of the muon beam. If all muons had random and uniformly distributed betatron phases, each muon oscillating one way would have a corresponding muon oscillating the other way. The moments of the spatial beam distribution, and thus the calorimeter acceptance, would then be constant. However, the injection process—particularly having the inflector exit at a large radial offset and the kick that occurs on the first turn around the ring—does not leave the muons with random, uniformly distributed betatron phases. For example, an imperfect kick will leave the beam as a whole offset from the center of the storage ring, causing its mean position to oscillate. This coherent motion of the beam in the radial direction, specifically its observation at a fixed azimuthal position, is called coherent betatron oscillation, or CBO.

The observed CBO frequency is not equal to the horizontal betatron frequency. To see why, consider, for simplicity, a group of mono-energetic muons that are tightly bunched longitudinally. The simplest version of this would be a single muon. Let  $x$  be some quantity of the muon beam that oscillates in time sinusoidally such that  $x(t) = \cos(\omega_x t)$ . This quantity could be the mean horizontal or vertical position of the distribution, the width, or anything that oscillates and can be measured by a detector. If there is a detector placed at a well defined azimuthal position, such as a calorimeter, it will sample the muon distribution once

per cyclotron period and collect a set of measurements  $m_i = \cos(2\pi i\omega_x/\omega_c)$ . (In reality, each calorimeter samples a range of azimuth, but the key is that a given calorimeter only looks at a fixed slice of the ring and does not see the whole beam.)

The set of measurements  $m_i$  is not sufficient to determine  $\omega_x$  because the same set of measurements would be produced by oscillations of the form  $x(t) = \cos[(k\omega_c - \omega_x)t]$  where  $k$  is any integer. The smallest of these frequencies, the one where  $k$  is closest to  $\omega_x/\omega_c$ , is the observed frequency. This effect is called aliasing, and it applies to any discretely sampled signal. The different frequencies  $(k\omega_c - \omega_x)$  are aliases of  $\omega_x$ .

The observed frequency will only be the single muon oscillation frequency,  $\omega_x$ , if  $\omega_x < \omega_c/2$ . In the specific case of horizontal betatron motion,  $\omega_x$  is about 0.9 to 0.95 times  $\omega_c$ , so  $k = 1$  and the observed frequency,  $\omega_{CBO}$ , is equal to  $(1 - \sqrt{1 - n})\omega_c$ . The beam width's oscillation also has components at  $2\omega_x$ . In that case,  $k = 2$  and the observed frequency is  $2\omega_{CBO}$ . The vertical oscillation frequency,  $\omega_y$ , is less than  $\omega_c/2$  and thus is observed in fixed-position detectors as  $\omega_y$ , unchanged. The components of the vertical width oscillation at  $2\omega_y$ , which dominate, are greater than  $\omega_c/2$  and are observed as  $\omega_{vw} = \omega_c - 2\omega_y$ . The  $vw$  subscript refers to the common name of this frequency, the vertical waist. The vertical waist is where the beam's vertical width is at a minimum; this minimum travels around the ring periodically at the vertical waist frequency.

Consider now another detector separated from the first one in azimuth by an angle  $\Delta\phi$ . It samples the tightly bunched muons at times  $t_i = 2\pi i/\omega_c + \Delta\phi/\omega_c$ ,  $\Delta\phi/\omega_c$  being the time it takes for the muon bunch to travel from the first detector to the second. The oscillation amplitude and frequency seen by the second detector must be the same as that seen by the first, but the phase could be different. So, one can write the oscillation seen by the first detector as  $D_1(t) = \cos(\omega_{CBO}t)$  and the oscillation seen by the second as  $D_2(t) = \cos(\omega_{CBO}t + \delta)$ . To solve for  $\delta$ , one can use that the oscillation seen by the detector and the true oscillation must match at the times when the detector samples the muons. Enforcing this condition at  $t = 0$

gives:

$$\begin{aligned}\cos(\omega_{CBO}\Delta\phi/\omega_c + \delta) &= \cos(\omega_x\Delta\phi/\omega_c) \\ \omega_{CBO}\Delta\phi/\omega_c + \delta &= \pm\omega_x\Delta\phi/\omega_c.\end{aligned}$$

The  $-$  sign in the previous equality must be chosen in order to match at all  $t_i$ . Continuing,

$$\begin{aligned}\delta &= -\Delta\phi/\omega_c(\omega_x + \omega_{CBO}), \quad \omega_{CBO} = k\omega_c - \omega_x \\ \delta &= -k\Delta\phi,\end{aligned}$$

where  $k$  is the alias label described previously.

The phase difference in the oscillation seen by two detectors separated by  $\Delta\phi$  depends on both the azimuthal separation and  $k$ , while  $k$  in turn depends on the ratio  $\omega_x/\omega_c$ . Figure 3.3 illustrates this effect for four different combinations of  $\omega_x/\omega_c$  and  $\Delta\phi$  of relevance in E989. The CBO oscillations, for which  $k$  is one, will be out of phase for detectors on opposite sides of the ring. One might expect this would lead to a cancellation when the signals from all detectors are combined. While this is partially true, slight differences in overall acceptances between calorimeters, caused in part by the differing intervening objects between the detectors and the storage region, prevents this cancellation from being perfect. The CBO frequency is present in the observed positron hit time spectrum even after summing the signals from all calorimeters.

Coherent betatron oscillations do not directly bias the precession frequency. They do, however, necessitate adjustment of the five parameter fit model. The CBO oscillations modulate the overall calorimeter acceptance at the percent level. As there is a range of frequencies originating from the beam's momentum spread, the coherent oscillations decrease over time. Additionally, the dependence of the acceptance on the muon decay position is not the same for positrons of different energies. This leads to a distortion of the observed energy spectrum at  $\omega_{CBO}$ , which in turn modulates the observed anomalous precession asymmetry and phase. There is a similar, though smaller, effect from the vertical waist. To trust the ultimate precession frequency fit to data, these effects must all be well understood and

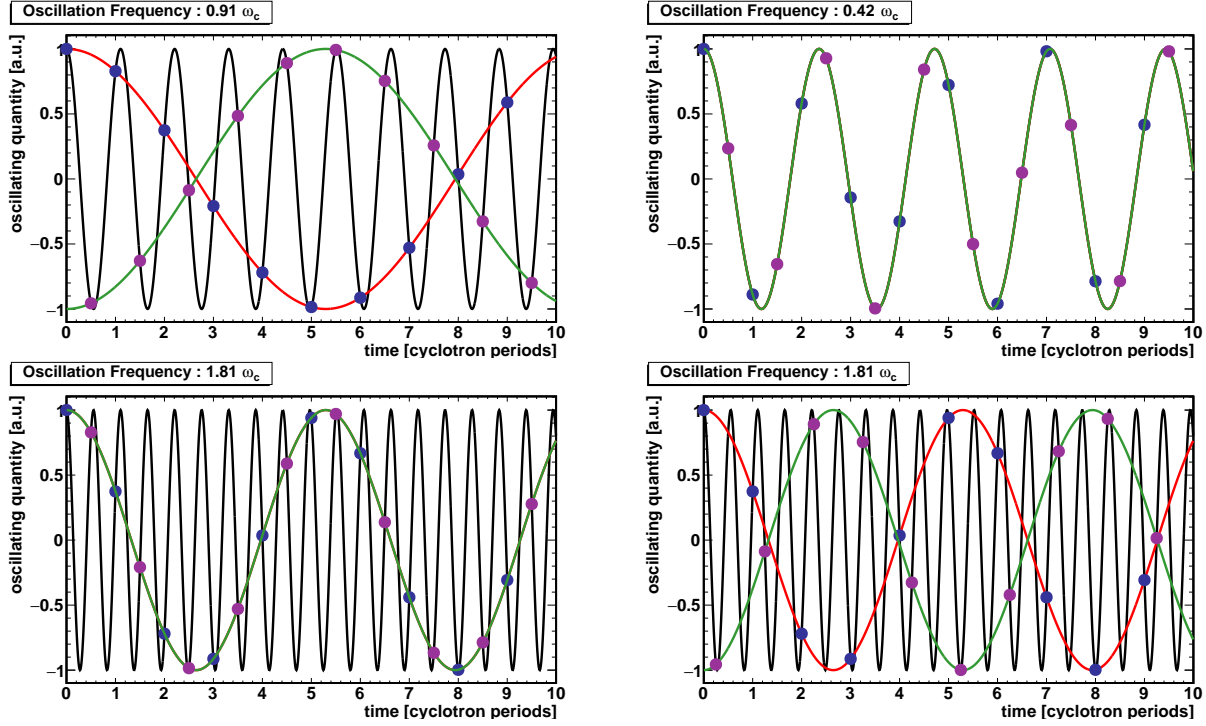


Figure 3.3: In all plots, the black curve shows the single muon oscillation frequency, the blue dots and red curve show the observations made by one detector, and the purple dots and green curve show the same for another detector at a different position around the ring. *Top Left:* Horizontal betatron frequency with two detectors separated by  $\Delta\phi = \pi$ . Here  $k$  is 1 and the two signals are exactly out of phase. *Top Right:* Vertical betatron frequency with two detectors separated by  $\Delta\phi = \pi$ . Here  $k = 0$  and  $\omega_x/\omega_c < 1/2$ , so the observed frequency is equal to the true frequency and both signals are perfectly in phase. *Bottom Left:* Twice the horizontal betatron frequency with two detectors separated by  $\Delta\phi = \pi$ . This frequency is noteworthy because the higher moments of the muon beam distribution have an oscillation at twice the betatron frequency.  $k = 2$  here, so two detectors separated by  $\Delta\phi = \pi$  have signals exactly in phase. *Bottom Right:* Same as *Bottom Left* except the two detectors are separated by  $\Delta\phi = \pi/2$ . Now the signals are exactly out of phase.

Table 3.1: Frequencies expected in the positron hit time histogram. Values are taken from E989 Run 1 and correspond to a quadrupole field index,  $n$ , of 0.107.

Name	Symbol	Typical value
Cyclotron frequency	$f_c$	6.71 MHz
Anomalous precession frequency	$f_a$	0.23 MHz
Coherent betatron frequency	$f_{CBO}$	0.37 MHz
Vertical betatron frequency	$f_y$	2.19 MHz
Vertical waist frequency	$f_{VW}$	2.32 MHz

included in the fit model. See Table 3.1 for typical values of these frequencies in the E989 experiment.

### 3.3.6 Electric Field and Pitch Corrections

The expressions used to now for  $a_\mu$  in terms of the anomalous precession frequency and the magnetic field require modification in the presence of betatron motion and quadrupole electric fields. A more complete expression—assuming a perfectly uniform magnetic field and no muon electric dipole moment (EDM)—is [1, 37]

$$\vec{\omega}_s - \vec{\omega}_c \approx \omega_a = \frac{e}{m} \left[ a_\mu \vec{B} - a_\mu \left( \frac{\gamma}{\gamma + 1} \right) (\vec{\beta} \cdot \vec{B}) \vec{\beta} - \left( a_\mu - \frac{1}{\gamma^2 - 1} \right) \frac{\vec{\beta} \times \vec{E}}{c} \right]. \quad (3.25)$$

The first term in the above equation provides the simple expression for  $a_\mu$  in terms of the magnetic field; this first term was used to derive Equation 2.5. The second term is a correction arising from muon motion parallel to the magnetic field. As individual muons undergo betatron motion in the vertical direction, this correction will not in general be 0. The third term, mentioned earlier, is the electric field correction that motivates the choice of 3.094 GeV/ $c$  for the ring momentum. Recall that for this momentum  $1/(\gamma^2 - 1) = a_\mu$ , and the third term vanishes. The beam’s finite momentum width, however, leads to a nonzero bias despite this momentum choice. Neither of these effects is negligible in E989; they are



each a few 100 ppb. The measured value of  $\omega_a$  must be corrected for these biases before it can be used to determine  $a_\mu$ .

First, consider only the  $(\vec{\beta} \cdot \vec{B})\vec{\beta}$  term. Recall that the calorimeters are sensitive to the rate of change of the angle between the muon spin and the muon momentum. Thus, the relevant component of the precession frequency is the one perpendicular to  $\vec{\beta}$ .  $(\vec{\beta} \cdot \vec{B})\vec{\beta}$  provides corrections to  $\omega_a$  that are parallel to  $\vec{\beta}$  and are proportional to  $(\vec{\beta} \cdot \vec{B})$ , which changes sign as the muon undergoes vertical betatron oscillations. The vertical betatron frequency is almost ten times faster than the anomalous precession frequency, so the effect of these  $\vec{\omega}_a$  components parallel to  $\vec{B}$  tends to average to 0 over any given precession period. The vertical oscillations do, however, reduce the size of the perpendicular component of  $\omega_a$  according to

$$\omega_{a,\perp} = \frac{e}{m} a_\mu B \cos(\psi), \quad (3.26)$$

where  $\psi$ , the pitch angle, is the angle between the muon momentum and the plane of the storage ring.  $\psi$  is generally small, on the order of milliradians. Expanding to leading order in  $\psi$  and averaging over the muon population,

$$\left\langle \frac{\Delta\omega_a}{\omega_a} \right\rangle_{pitch} = -\frac{\langle\psi^2\rangle}{2}. \quad (3.27)$$

The observed precession frequency is lowered by the vertical betatron motion of the beam. This effect is commonly called the pitch correction. The pitch angles in the ring cannot be directly measured, but they can be related to the vertical betatron amplitudes. For small angles,  $\psi = v_y/v_z = v_y/(\omega_c R_0)$ . Using this relation and the solutions to the equations of motion found in Section 3.3.1,

$$\begin{aligned} \psi &= \frac{1}{\omega_c R_0} \frac{dy}{dt}, \\ \langle\psi^2\rangle &= \frac{n}{R_0^2} \langle y^2 \rangle. \end{aligned} \quad (3.28)$$

The pitch correction is then related to the vertical beam distribution, which can be measured by tracking detectors and cross checked with simulations. The overall pitch correction in

terms of the vertical beam distribution is

$$\left\langle \frac{\Delta\omega_a}{\omega_a} \right\rangle_{pitch} = -\frac{n\langle y^2 \rangle}{2R_0^2}. \quad (3.29)$$

The systematic uncertainty from the pitch correction is not the pitch correction itself, which can be determined using the above equation, but the uncertainty of that correction, which is driven by the imperfect knowledge of the stored beam's vertical distribution.

Next, consider the  $\vec{\beta} \times \vec{E}$  term. The dominant component here is the  $\beta_z E_r$  term, which is in the vertical direction, parallel to the leading term in the expression for  $\omega_a$ . Isolating this term, and using the relation  $\gamma^2 - 1 = \beta^2 \gamma^2$ , the shift in the observed precession frequency is

$$\left( \frac{\Delta\omega_a}{\omega_a} \right)_E = -\frac{\beta_z E_r}{B_y c} \left( 1 - \frac{1}{a_\mu \beta^2 \gamma^2} \right).$$

To further simplify this, remember that the design momentum of the ring is such that  $1/a_\mu^2 = \gamma_0^2 - 1$ . Additionally, without much loss of accuracy, one can expand  $\gamma$  about the design value and treat  $\beta$  as unchanging over this range of  $\gamma$ . Doing so yields

$$\left( \frac{\Delta\omega_a}{\omega_a} \right)_E = -\frac{2\beta E_r}{B_y c} \left( \frac{\Delta p}{p_0} \right). \quad (3.30)$$

This shift in the observed precession frequency is called the electric field correction.

Examining this equation, one finds the only terms that vary much from one muon to another are  $E_r$  and  $(\Delta p/p_0)$ . The product of these determines the sign of the correction. In the idealized version of E989, the ring is constructed such that muons with momentum  $p_0$  sit exactly in the center of the quadrupoles, where  $E_r = 0$ . Remember also that a muon's equilibrium radius is proportional to its momentum. Muons with momenta higher than  $p_0$  will sit at slightly higher radii, where  $E_r$  is positive. For them, the product  $E_r(\Delta p/p_0)$  is positive. Muons with momenta lower than  $p_0$  will sit at lower radii, where  $E_r$  is negative. For them, the product in question is also positive. So, studying the above equation, it appears that for an ideally constructed E989 experiment, the sign of the electric field correction is decidedly negative for all muons other than those with momentum exactly equal to  $p_0$ , which receive no correction.

Consider now the case where the quadrupole plates are slightly misaligned such that the center of their potential, where  $E_r = 0$ , is not at the equilibrium radius of design-momentum muons. Now there are two important radii to consider: the one where  $E_r = 0$ , and the equilibrium radius of design-momentum muons. In the region between these two radii, the electric field correction is positive. Everywhere else, it is negative. The average electric field correction over the entire muon population is sensitive to quadrupole plate misalignments.

The radial electric field is a function of  $x$ , the displacement from the center of the storage region. The average radial electric field felt by a single muon is related to that muon's average radial coordinate—namely its equilibrium radius. Equation 3.13 shows that  $(\Delta p/p_0)$  is also related to the equilibrium radius. Equation 3.30 can be rewritten entirely in terms of muon equilibrium radii. To do so, however, one must keep in mind that the equilibrium radius is related to the deviation from the momentum stored at the center of the quadrupole potential, not the deviation from the design momentum. Equation 3.30, however, is in terms of deviation from the design momentum. To make all necessary substitutions, one needs an additional variable,  $p_r$ , which is the momentum stored at the center of the quadrupole potential. Doing the requisite algebra, and using the definition of the field index,  $n$ , from Equation 3.10,

$$\left(\frac{\Delta\omega_a}{\omega_a}\right)_E = -\frac{2\beta^2 n(1-n)}{R_0^2} \left(x_{eq}^2 - x_{eq} \frac{R_0}{1-n} \frac{p_0 - p_r}{p_r}\right). \quad (3.31)$$

In the above equation, multiplicative factors of  $p_r/p_0$  have been dropped. Any quadrupole misalignment could not be larger than a few mm at most, making these terms per mille perturbations to an order ppm correction. In Equation 3.31,  $x_{eq}$  is relative to the center of the quadrupole potential, not the center of the ring. Starting with the ideal case where  $p_0 - p_r = 0$ , after averaging over the muon population the total electric field correction is

$$\left\langle \frac{\Delta\omega_a}{\omega_a} \right\rangle_E = -\frac{2\beta^2 n(1-n)}{R_0^2} \langle x_{eq}^2 \rangle, \quad (3.32)$$

$$\left\langle \frac{\Delta\omega_a}{\omega_a} \right\rangle_E = -\frac{2\beta^2 n(1-n)}{R_0^2} (\langle x_{eq} \rangle^2 + \sigma_{x_{eq}}^2) \quad (3.33)$$

The size of the electric field correction is determined by the distribution of equilibrium radii, which is extracted from the fast rotation analysis discussed earlier. As with the pitch

correction, the contribution to the E989 error budget is not the electric field correction itself, but the correction's uncertainty. The correction's uncertainty has components from the uncertainty of the equilibrium radius distribution and also from the uncertainty of the quadrupole alignment.

Moving beyond the ideal case,  $p_0 - p_r$  could be nonzero for at least two reasons. One reason is quadrupole misalignment. If the center of the quadrupole potential were moved by a distance  $h$ , the expression for the electric field correction would become

$$\left\langle \frac{\Delta\omega_a}{\omega_a} \right\rangle_E = -\frac{2\beta^2 n(1-n)}{R_0^2} \left( \langle x_{eq}^2 \rangle + \langle x_{eq} \rangle \frac{h}{1-n} \right). \quad (3.34)$$

Uncertainty in the quadrupole alignment couples linearly with an offset in the equilibrium radius distribution.

Another reason could be that the quadrupoles are perfectly aligned, but the ring's magnetic field is changed. The magnetic field is controlled precisely enough that this could not happen accidentally at any appreciable level, but it may be changed deliberately for systematic tests. In that case,

$$\left\langle \frac{\Delta\omega_a}{\omega_a} \right\rangle_E = -\frac{2\beta^2 n(1-n)}{R_0^2} \left( \langle x_{eq}^2 \rangle + \langle x_{eq} \rangle \frac{R_0}{1-n} \frac{\Delta B}{B_0} \right). \quad (3.35)$$

Fast rotation analysis of 2018 data revealed that the equilibrium radius distribution is not centered at  $x_{eq} = 0$ . Without any adjustments, a non-centered equilibrium radius distribution can increase the size of the electric field correction considerably. The above equation shows, however, that this increase can be countered by reducing the storage ring's magnetic field. See Figure 3.4. In general, the smallest uncertainty in the electric field correction can be achieved when the beam is centered in the quadrupole potential and the quadrupole potential is aligned to design-momentum muons. In this case, uncertainties in the beam distribution have minimal impact on the correction.

The electric field and pitch corrections are required adjustments to the measured precession frequency. The corrections themselves combine to approximately 700 ppb [37]. The size of these corrections is related to the transverse distribution of the beam, specifically the

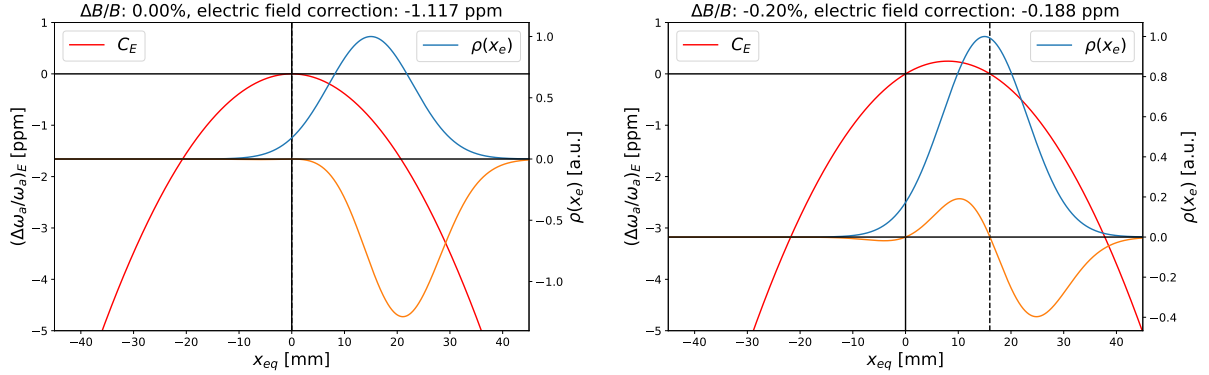


Figure 3.4: The electric field correction for an off-center beam. In each graph, the red curve is the electric field correction as a function of equilibrium radius, the blue curve is the stored beam’s equilibrium radius distribution, and the yellow curve is the product of the two. The integral of the yellow curve is proportional to the average electric field correction. The left plot shows the situation where the magnetic field is tuned such that design-momentum muons orbit at the center of the quadrupoles. The right plot shows the situation where the magnetic field has been reduced—design-momentum muons now orbit at the dashed line. In the case of an off-center equilibrium radius distribution, the size of the correction can be reduced by changing the storage ring’s magnetic field.

radial distribution for the electric field correction and the vertical distribution for the pitch correction. Measuring these transverse distributions is a critical component of the precession frequency measurement. The uncertainty of these corrections comes from how well the beam distributions can be known, as well as from uncertainty in the alignment of the quadrupole plates. The combined E989 error budget for the electric field and pitch corrections is 30 ppb.

### 3.4 Systematic Effects from Detectors

In the laboratory reference frame, positrons from muon decay are produced according to the probability density function given in Equation 2.12. The functions  $N(E)$  and  $A(E)$  can be calculated theoretically. The quantities  $E$  and  $t$ , however, can never be observed. Rather, the calorimeters report detected energies and hit times that necessarily differ somewhat from the pure, decay-time quantities. If the detection process transforming the pure quantities into measured quantities varies over a muon fill, systematic biases can arise. Such effects

could be rooted in detector hardware, software, or both. The primary known detector effects of concern are pileup, also called detector dead time, and gain instability, which is a time-dependent shift of the conversion factor from positron energy to detected energy. This section describes these effects and how they can bias the measured precession frequency.

#### 3.4.1 Pileup

Pileup is the term for when two or more particles impact a detector so close in time that only one hit is recorded. This hit is generally recorded with an energy close to the sum of the energies of the unresolved hits. See Figure 3.5 for an example. The maximum time separation below which distinct hits cannot be resolved is called the detector dead time. A nonzero detector dead time distorts the measured positron spectrum and modifies the observed asymmetry, phase, and normalization. For example, in the presence of pileup, events reconstructed as 2 GeV positrons are really a combination of 2 GeV positrons and pairs of smaller energy pulses that happen to add to 2 GeV. These smaller energy pulses have different phases and asymmetries than true 2 GeV positrons.

If pileup did nothing other than change the normalization, asymmetry, and phase versus detected energy, it would not be a concern in the precession frequency analysis. The distortions, however, are time dependent. To see why, consider a time window of width  $\Delta t$  immediately following a detected pulse. If  $\Delta t$  is small compared to  $1/r$ , where  $r$  is the average hit rate over the time window, then the probability another hit will fall inside the time window is approximately equal to  $r\Delta t$ . If  $\Delta t$  is the detector dead time, which in E989 is always small compared to the inverse rate, then another pulse falling inside the time window is a pileup pulse. It will not be distinguished from the first pulse, and a single combined pulse will be recorded. Therefore, the probability a given pulse will be a pileup pulse is approximately equal to  $r\Delta t$ , where  $r$  is the instantaneous hit rate and  $\Delta t$  is the detector dead time. As  $r(t)$  is the overall pulse rate as a function of time, the rate of pileup pulses over time is

$$r_{pu}(t) \approx r^2(t)\Delta t. \quad (3.36)$$

The overall hit rate in the absence of pileup is  $r(t)$ , and the rate of pileup pulses is  $r^2(t)\Delta t$ .  $r(t)$  decreases over a fill as the stored muons decay, so the size of the pileup perturbation decreases as well. This time dependent perturbation of the observed positron distribution parameters, especially the phase, will bias the precession measurement if uncorrected. The size of the perturbation can be decreased by either reducing the detector dead time or reducing the rate of positron hits. For this reason, the E989 calorimeter was engineered to have as small a dead time as possible; its dead time is less than 5 ns. On the other hand, a great effort was expended to optimize the beam injection into the E989 storage ring and increase the number of stored muons per fill. Doing so raises the positron hit rate and makes the pileup perturbation larger. The size of the pileup perturbation's precession frequency bias will be estimated in section 3.5. The E989 pileup error budget is 40 ppb.

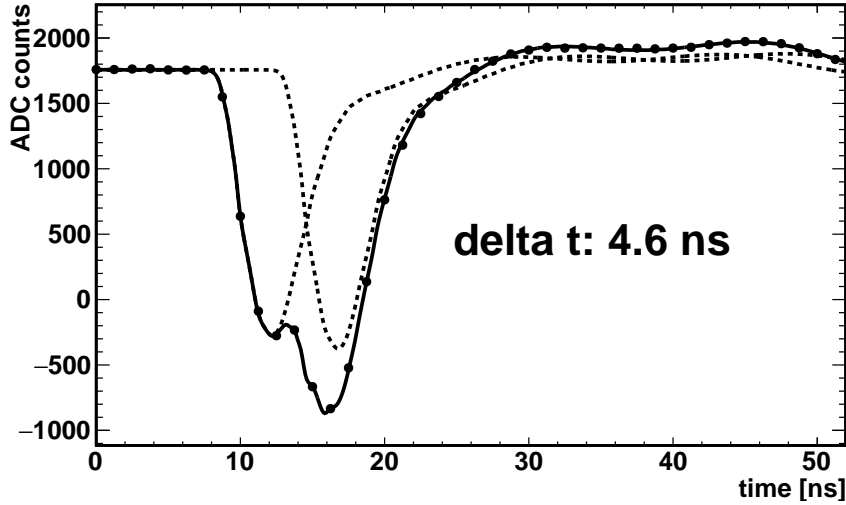


Figure 3.5: A digitized trace recorded as two electron pulses impact a prototype E989 calorimeter during a test beam experiment, reproduced from [63]. Dashed lines show the traces that would have been recorded had either pulse arrived individually. As the pulse waveforms move closer together, they merge into a single, ill-defined pulse. Depending on the reconstruction routine, this event may be reconstructed as two pulses or a single, larger energy pulse. The latter case would be pileup.

### 3.4.2 Gain Changes

Gain is a conversion factor between the physical energy deposition in a calorimeter and the quantifiable detected signal. A quantifiable detected signal could be an amount of charge collected, an ADC value, or whatever physical value is recorded following a calorimeter pulse. Typically, through a calibration procedure, the gain is measured so the detected signal can be converted back into physical energy units, such as MeV. At first approximation—for a given configuration of settings—the gain is a fixed property of a detector. In reality, however, the gain can vary with external factors such as temperature and hit rate. Temperature variations occur over time scales much longer than a single muon fill, so it is gain variation with hit rate that is liable to cause a precession frequency systematic bias.

The muon decay lifetime in the E989 storage ring is about  $64\ \mu\text{s}$ , and decays are observed over  $700\ \mu\text{s}$  fills. A fill lasts about 11 decay lifetimes; throughout a fill, the rate of muon decays drops by five orders of magnitude. For example, if there were 10,000 muons present at the beginning of a fill, the instantaneous rate of decays in the ring at  $t = 0$  would be about 160 MHz.  $700\ \mu\text{s}$  later, the instantaneous rate of decays would be about 3 kHz. Care must be taken to account for any changes in detector gain that occur as a result of these large rate changes.

The pernicious effect of time-in-fill dependent gain changes is most easily understood in the context of a T-Method analysis. As a reminder, a T-Method analysis histograms the hit times of positrons above a certain energy threshold. Neglecting systematic effects, the resulting histogram will on average follow the form given in Equation 2.17. The parameters  $N$ ,  $A$ , and  $\phi$  depend on the chosen energy threshold. If there are uncorrected, repeatable detector gain shifts over the course of a fill, the energy threshold will effectively change. A decreased gain is equivalent to a higher energy threshold. Thus, the parameters  $N$ ,  $A$ , and  $\phi$  become time dependent. As has been shown, phase shifts bias the observed precession frequency. Additionally, the assumed fit function will be generally incorrect in the case of time dependent  $N$ ,  $A$ , and  $\phi$ .



Other analysis techniques, such as energy-binned and asymmetry-weighted analyses, are also susceptible to changing gains. The size and sign of the precession frequency bias can be different between these different procedures. These differences will be explored and quantified in section 3.5. The differing effects of gain changes pileup in different analysis procedures is a strong point in favor of applying numerous analysis techniques to the E989 data and demanding agreement. The E989 uncertainty budget for changing detector gains is 20 ppb. To achieve this target, systematic gain changes over the  $700\ \mu\text{s}$  long muon fills must be controlled to the sub-per-mille level [37].

### **3.5 Estimation of Detector-Based Systematic Biases**

The sizes of the precession frequency biases from pileup and gain changes can be estimated using a toy Monte Carlo study. The starting point for this study is a joint probability density function  $\rho_d(E, t)$ , where  $E$  and  $t$  are the detected energy and time rather than pure energy and time. Under certain approximations, the form of this function is identical to 2.12 except that  $\phi$  becomes a function of  $E$ . The functions  $A(E)$  and  $N(E)$  in reference to detected energy and time are different from the theoretical distributions, which are in terms of the pure decay quantities. To transform from the theoretical distributions to what would be observed at a detector, detector resolution, energy dependent acceptance, and drift time must be considered. In this study,  $\rho_d$  was calculated under the following assumptions:

- Calorimeter resolution is piecewise defined to be a Gaussian distribution with mean  $E$  and a relative width of  $1.5\% + 4.5\%/\sqrt{E/\text{GeV}}$ , summed in quadrature, for  $E_d > E - \sigma(E)$  and an exponential tail for  $E_d \leq E - \sigma(E)$ . The parameters of the exponential tail are chosen to enforce that the resolution function is continuous and has a continuous first derivative  $\sigma(E)$  is the standard deviation of the Gaussian distribution described above. This model is consistent with the test beam measurements described in section 4.3.
- Acceptance increases linearly from 0 to 75% over the range from 0 MeV to 2350 MeV and then remains flat. No time dependence is considered. This model qualitatively

agrees with the acceptance extracted from **Geant4** simulations [64].

- The drift time distribution is approximated as a delta function at the mean drift time for a given energy. The mean drift time for a given energy decay positron is taken from **Geant4** simulations.
- Leading order theoretical equations for  $N(E)$  and  $A(E)$  are used.

Using the models described above, one obtains a decay density function in terms of detected quantities:

$$\rho_d(E, t) \propto e^{-t/(\gamma\tau_\mu)} N_d(E) [1 + A_d(E) \cos(\omega_a t - \phi_d(E))]. \quad (3.37)$$

The  $d$  subscripts indicate that these functions apply to the detected energy and time and are not the theoretical  $A$  and  $N$  functions appearing in Equation 2.17.

For a dataset of  $N$  muons,  $N \cdot \rho_d(E, t) \Delta E \Delta t$  gives the expected number of events in a two-dimensional energy, time histogram bin centered at  $E, t$  with energy bin width  $\Delta E$  and time bin width  $\Delta t$ . Datasets of any size can be emulated using this distribution. Fluctuations in bin contents will obey Poisson statistics. The toy Monte Carlo procedure employed here is:

1. Create a two-dimensional energy and time histogram.
2. Set each bin content to  $N_{muons} \cdot \rho_d(E, t) \Delta E \Delta t$ , where  $\Delta E$  and  $\Delta t$  are the energy and time bin widths and  $E$  and  $t$  are the central energy and time of the bin.
3. If desired, include statistical fluctuations by changing each bin content by a random number drawn from a Gaussian distribution with mean of 0 and width of the square root of the bin content.
4. Replace each bin content with the closest integer.

5. Generate T-Method, asymmetry-weighted, energy-weighted, and other  $\omega_a$  analysis oriented histograms through operations on this two-dimensional histogram.

This procedure is accurate provided the bin widths are small compared to the rate of change of the asymmetry, number, or phase with respect to time or energy. Additionally, statistical fluctuations can only be modeled with Gaussian distributions if bin contents are not too small. The dataset emulated here had  $2.3 \times 10^{12}$  stored muons at  $t = 0$ , which corresponds to  $1.6 \times 10^{11}$  detected positrons over 1.86 GeV after  $30 \mu\text{s}$ . Bin widths were 149 ns in time 16.5 MeV in energy. All fits began at  $30 \mu\text{s}$ .

### 3.5.1 *Estimated Size of the Bias from Gain Perturbations*

This toy Monte Carlo procedure can produce an estimated size of the precession frequency bias from changing gains. It is convenient to define a gain function  $g(t)$  in the following way: if a positron would be measured with energy  $E$  absent gain perturbations, the measured energy with gain perturbations,  $E_p$ , is  $g(t) \cdot E$ . So,  $g(t) = 1$  corresponds to an unperturbed gain and  $g(t) < 1$  denotes that measured energies are lower than they should be. Typically, gain perturbations will be small—on the order of 1 per mille—so  $g(t)$  can be expressed as  $1 + \delta g(t)$  where  $\delta g(t)$  is a small parameter.

In an energy-binned approach, gain perturbations shift positron hits from one energy bin into another. As each bin has its own phase, asymmetry, and normalization, gain changes over the course of a fill cause phase, asymmetry, and normalization shifts that bias  $\omega_a$ , potentially in a different way for each energy bin. In a T-Method analysis, gain changes are equivalent to a time dependent energy threshold and thus the derivative of the T-Method  $N$ ,  $A$ , and  $\phi$  with respect to energy threshold determine the T-Method sensitivity to gain shifts. In an asymmetry-weighted analysis, changing gains change the way positrons of a given energy are weighted in the histogram. To determine the effect of gain perturbations on these different  $\omega_a$  analyses, gain perturbations of varying sizes are included in the emulated datasets. So that the systematic effect caused by the gain perturbations can be isolated,

statistical fluctuations are not included. After generating a gain perturbed detected positron distribution, T-Method, asymmetry-weighted, and energy-binned histograms are generated and fit to extract  $\omega_a$ . The resulting  $\omega_a$  values are different from the true value; the difference is the bias caused by the gain perturbations.

Knowing  $\rho_d(E, t)$ , the unperturbed detected positron distribution, one can obtain the gain perturbed positron distribution,  $\rho_{gp}(E_p, t)$ . With  $\rho_d(E, p)$  normalized to 1, we have

$$\begin{aligned} 1 &= \iint \rho_d(E, t) dE dt, \quad E = \frac{E_p}{g(t)}, \quad dE = \frac{dE_p}{g(t)} \\ 1 &= \iint \frac{1}{g(t)} \rho_d\left(\frac{E_p}{g(t)}, t\right) dE_p dt. \end{aligned}$$

Thus, the gain perturbed positron distribution function is

$$\rho_{gp} = \frac{1}{g(t)} \rho_d\left(\frac{E_p}{g(t)}, t\right). \quad (3.38)$$

The factor of  $1/g(t)$  in front accounts for how shifting gains change not only which energy bin a positron will be assigned to, they also change the width of each energy bin. For example, with a 1% gain reduction, a 100 MeV wide bin becomes a 101 MeV wide bin. The gain perturbed positron distribution is used to generate emulated datasets in the same way described earlier. The  $g(t)$  function used in this study is of a double-exponential form based on lab and test beam measurements [63, 65]. The results are shown in Figure 3.6.

This study suggests that a 0.1% gain perturbation will create a systematic bias on the order of 100 ppb, which is a result consistent with other similar studies [37, 66]. The gain must then be corrected to  $10^{-4}$  to reach the E989 systematic error target. The exact value of the  $\omega_a$  bias depends on the fit start time in the  $g - 2$  cycle, as well as the model used for  $\phi(E)$ . Data-driven estimates of the gain perturbations and associated biases present in the E989 Run 1 data will be presented in Chapter 6.

### 3.5.2 Estimated Size of the Bias from Pileup

The detected positron distribution will be distorted by unresolved pulse pairs that arrive within the calorimeter dead time of one another. A realistic dead time to consider for

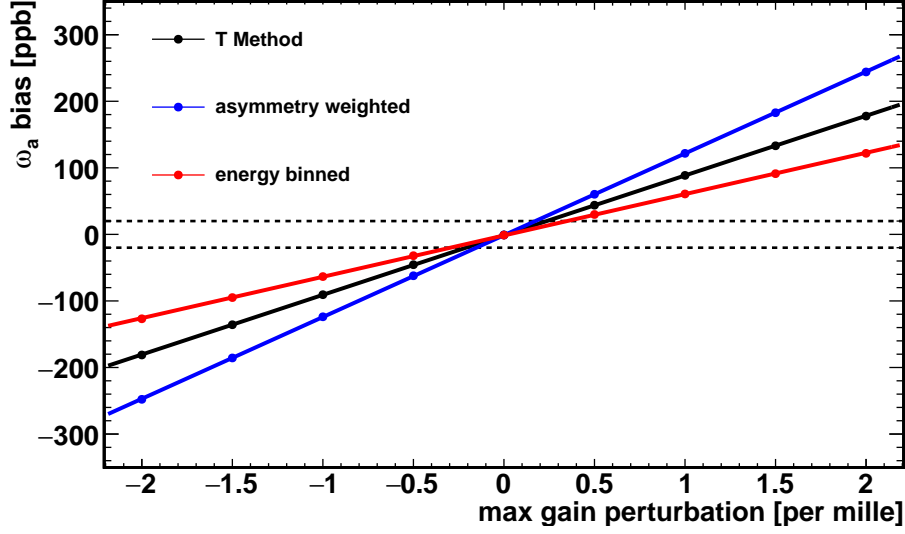


Figure 3.6:  $\omega_a$  bias versus maximum size of gain perturbation for asymmetry-weighted, T-Method, and energy-binned  $\omega_a$  analyses. Dashed lines are at  $\pm 20$  ppb, the E989 systematic error budget allotted for gain perturbations.

the E989 calorimeters is about 5 ns. It is useful to describe the pileup distorted positron distribution,  $\rho_{pu}$ , as a sum of two components:

$$\rho_{pu}(E, t) = \rho_d(E, t) + \delta\rho_{pu}(E, t). \quad (3.39)$$

In the above equation,  $\rho(E, t)$  is the positron distribution without any pileup perturbations and  $\delta\rho_{pu}(E, t)$  is the dead-time-dependent distortion to the distribution. In addition,  $\delta\rho_{pu}$  is the correction you would have to subtract from a pileup contaminated positron distribution to recover the unperturbed distribution. Uncorrected pileup effects can cause large  $\omega_a$  biases, and different analysis procedures display different sensitivities to them. In this section, a model for  $\delta\rho_{pu}$  is presented and the relative pileup sensitivities of the T-Method, asymmetry-weighted, and energy-binned  $\omega_a$  analyses are evaluated using that model. As double pileup is the leading order effect, for simplicity the model described here considers only double pileup effects, i.e. the effect of more than two pulses arriving within a single dead time window will not be considered. However, the procedure that will be described can be adapted to include

triple pileup effects, as discussed in Chapter 6.

The size of the pileup perturbation depends on the per-calorimeter per-fill hit rate and the detector dead time. Let  $\mathcal{N}_{fill}(E, t)$  be a rescaled version of  $\rho_d(E, t)$  that integrates to the average number of positron hits in a single calorimeter in a single fill. Then,  $\mathcal{N}_{fill}(E, t) dE$  is the single calorimeter rate of positron hits at time  $t$  in an energy slice between  $E$  and  $E + dE$  and  $\int \mathcal{N}_{fill}(E, t) dE$  is the time dependent total hit rate for a single calorimeter. In this study, the E989 design value of 16,000 stored muons at  $t = 0$  is used, which corresponds to a hit rate of about 3.5 MHz per calorimeter at  $t = 0$ .  $\mathcal{N}_{fill}(E, t)$  can be used to generate the time and energy dependent pileup perturbation.

For a dead time  $\Delta t$  that is very small compared to  $\tau$  and  $\omega_a$ , and assuming individual decays are entirely independent from one another, the number density of pileup pulses in a fill as a function of the energy of the first pulse,  $E_1$ , the energy of the second pulse,  $E_2$ , and the time of the first pulse,  $t$ , is

$$\mathcal{N}_{pu}(E_1, E_2, t) \equiv \frac{d^3 N_{pu}}{dE_1 dE_2 dt} = \Delta t \cdot \mathcal{N}_{fill}(E_1, t) \mathcal{N}_{fill}(E_2, t). \quad (3.40)$$

In the above,  $N_{pu}$  is a number of pileup pulse pairs and  $\mathcal{N}_{pu}$  is a number density with units of one per energy squared per time.

To determine exactly how a pileup pulse pair will perturb the measured positron distribution, one needs to know how the reconstruction procedure responds to two pulses within the dead time window. A simple case considered here is that when faced with two pulses of parameters  $E_1$ ,  $E_2$ , and  $t$  as defined above, the reconstruction will report a single pulse at time  $t$  with energy  $E_1 + E_2$ . Thus, the perturbation relative to the true spectrum will include the loss of a count at  $(E_1, t)$ , the loss of a count at  $(E_2, t)$ , and the gain of a count at  $(E_1 + E_2, t)$ . Thus, the energy and time dependent double pileup perturbation is

$$\delta \mathcal{N}_{pu}(E, t) = - \int \mathcal{N}_{pu}(E, E_2, t) dE_2 - \int \mathcal{N}_{pu}(E_1, E, t) dE_1 + \int \mathcal{N}_{pu}(E - E_2, E_2, t) dE_2. \quad (3.41)$$

$\delta \mathcal{N}_{pu}(E, t)$  is the energy and time dependent per-calorimeter per-fill pileup perturbation. It can be rescaled by the number of hits per calorimeter per fill to obtain the desired overall

positron distribution pileup perturbation,  $\delta\rho_{pu}(E, t)$ . In the end, one finds that the size of the double pileup perturbation scales linearly with the detector dead time and with the number of hits per calorimeter per fill.

With  $\rho_{pu}(E, t)$  in hand, one can emulate pileup perturbed  $g-2$  experiments and determine the relative susceptibilities of the different analysis procedures. As before, the T-Method, asymmetry-weighted, and energy-binned analyses are considered. Positron hits reconstructed above 3300 MeV are not included in any analysis. Additionally, two different minimum energy cuts for the asymmetry-weighted and energy-binned analyses were used. Figure 3.7 shows how the  $\omega_a$  bias changes with the detector dead time. For the model used in this study, a change in the detector dead time is equivalent to a change in the per calorimeter per fill rate. As expected, all analysis procedures are unbiased in the limit where the dead time goes to 0. The T-Method bias and asymmetry-weighted bias change linearly as the dead time increases. The low-threshold energy-binned bias, however, appears to remain quite flat until the dead time is increased beyond 4 ns. The sign of the  $\omega_a$  bias in the asymmetry-weighted analysis changes depending on the low-end energy cut. As with the gain study, the precise value of the  $\omega_a$  bias depends highly on the fit start time within a  $g-2$  cycle, the model used for  $\phi(E)$ , and other parameters of the detector effect model.

The results shown in Figure 3.7 indicate that detector dead times must be significantly below 1 ns to keep the precession frequency bias below 40 ppb. Detector dead times of a few ns cause precession frequency biases that are hundreds of ppb. It is not feasible to reach the target E989 pileup error simply by reducing detector dead times to sub-ns levels. Rather, the dead time is brought as low as possible and then the remaining pileup contamination is removed on average from final histogram before fitting. This procedure is called pileup subtraction. There are multiple pileup subtraction routines currently being pursued in the E989 collaboration. Pileup subtraction will be discussed more in Chapter 6.

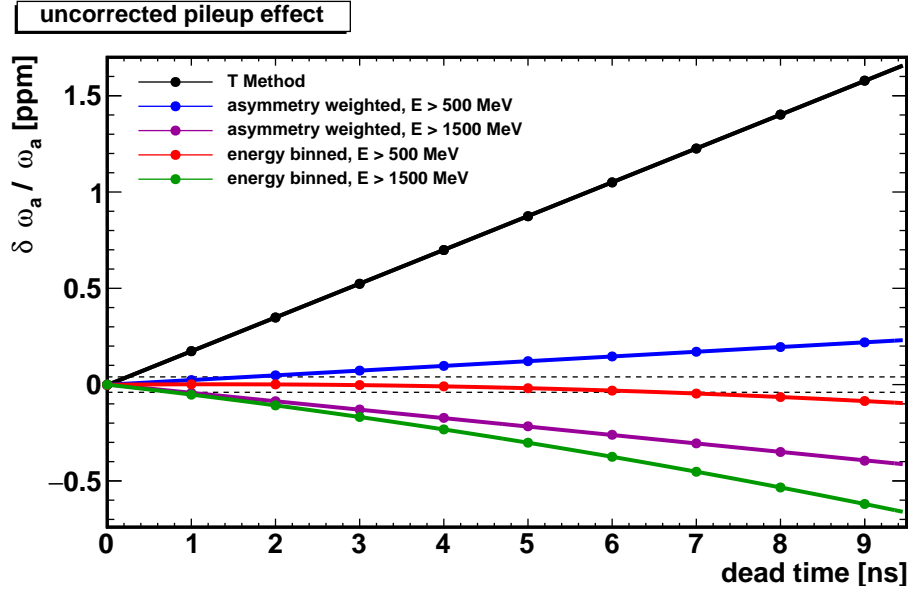


Figure 3.7:  $\omega_a$  bias versus dead time for different  $\omega_a$  analysis procedures and minimum energy cuts. Note that the sign of the bias in an asymmetry-weighted analysis depends on the minimum energy cut. Dashed lines show the  $\pm 40$  ppb level that is ultimately targeted in E989.

### 3.6 Summary of the Precession Frequency Error Budget

This chapter has outlined the known sources of systematic uncertainty in the E989 anomalous precession frequency measurement. They and their target values are summarized in Table 3.2. These targets are motivated by considering realistic improvements—achievable through upgraded instrumentation and simulation tools—to the best systematic uncertainties reached in the E821 experiment. The E989 Technical Design Report (TDR) contains detailed arguments justifying the values quoted in the table [37]. The systematic error budget leaves no room for the unanticipated and undiscovered. It is the challenging responsibility of the E989 data analyzers and simulators to demonstrate that they have reached these targets and that there are no additional, unaccounted-for effects.



Table 3.2: Known sources of systematic uncertainty in the anomalous precession frequency measurement and their E989 target values. The target values quoted for beam dynamics and detector effects are quadrature sums of the individual effects within those respective categories.

Uncertainty source	Target value
Beam dynamics effects	< 50 ppb
Lost muons	20 ppb
CBO	30 ppb
E-field and pitch corrections	30 ppb
Others?	? ppb
Detector effects	< 50 ppb
Gain changes	20 ppb
Pileup	40 ppb
Others?	? ppb
Total quadrature sum	< 70 ppb

## Chapter 4

### THE E989 CALORIMETER

The calorimeters are the primary instruments in the E989  $\omega_a$  measurement. They are responsible for reporting the energies and times of incident positrons. For an unbiased frequency measurement, any systematic differences between these reported energies and times and the true quantities must be stable across the  $700\,\mu\text{s}$  of a muon fill. The calorimeter hardware and reconstruction software should limit the effects of gain changes and pileup as much as possible, and the electronics must operate in magnetic fields ranging from 0.8 T to 1.5 T. Reaching the E989 target precision is possible only with a high-performance calorimeter optimized for the specific needs of the  $\omega_a$  measurement. This chapter discusses the design, testing, and characterization of the E989 calorimeter.

#### 4.1 *Hardware*

The E989 calorimeters feature 6 high by 9 wide arrays of lead fluoride ( $\text{PbF}_2$ ) crystals. Each crystal is coupled directly to a silicon photomultiplier (SiPM), which responds to incident Cherenkov photons with electrical current. Produced electrical current passes through a custom electronic readout board where it is converted into a voltage signal. These voltage signals are recorded by 12-bit, 800 MS/s waveform digitizers. Waveforms judged as candidate positron pulses by an online data reduction system are then stored for offline analysis. Additionally, a laser fiber is coupled to the front face of each crystal. The laser system provides periodic laser pulses for gain calibration and time synchronization. A model of the E989 calorimeter and its housing is shown in Figure 4.1 [37].

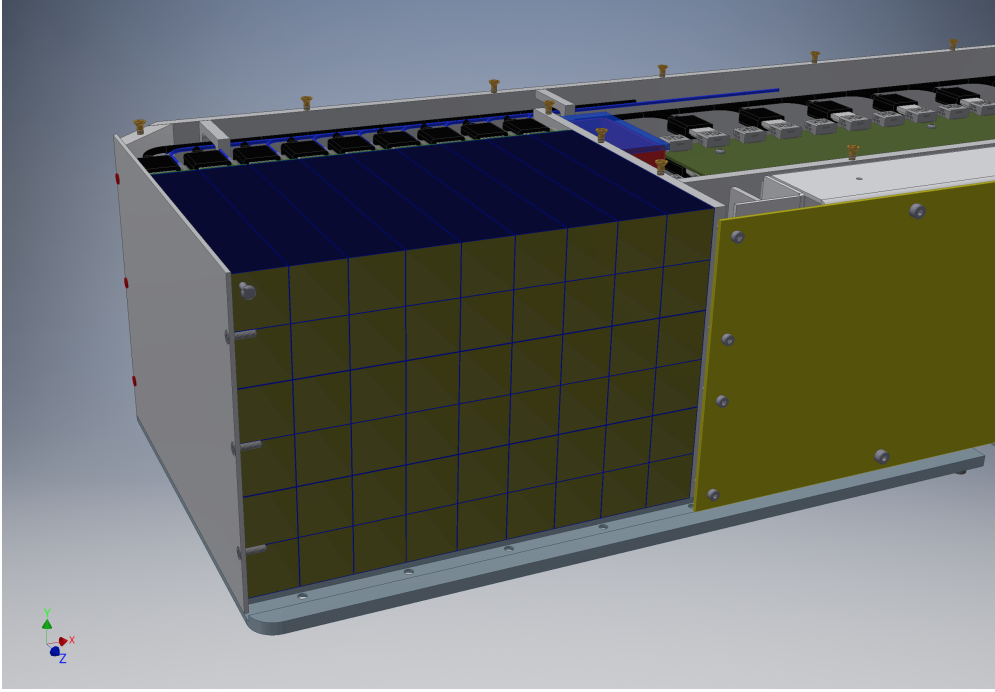


Figure 4.1: Model of the E989 calorimeter, courtesy of Jarek Kaspar. The active volume comprises a  $6 \times 9$  array of  $\text{PbF}_2$  crystals. There are 24 calorimeters equally spaced around the E989 storage ring, illustrated in Figure 4.2

#### 4.1.1 Lead Fluoride

Lead fluoride is a dense material, has a high index of refraction, and is transparent to optical light (see Table 4.1). High density allows for decay positrons to deposit virtually all of their energy in a relatively compact calorimeter. The crystals used in E989 are  $2.5\text{ cm} \times 2.5\text{ cm} \times 14\text{ cm}$  rectangular prisms; 14 cm is approximately 15 radiation lengths. Additionally, the crystals have transverse dimensions larger than the  $\text{PbF}_2$  Molière radius, leading to energy depositions from typical decay positrons that are contained almost entirely within one or two crystals. (The Molière radius is the radius of a cylinder transverse to the electromagnetic shower in which 90% of the shower's energy is contained [5].) Calorimeter segmentation into  $6 \times 9$  arrays enables pileup resolution through recognition of spatially separated energy depositions. With 24 calorimeters, there are a total of 1296  $\text{PbF}_2$  crystals

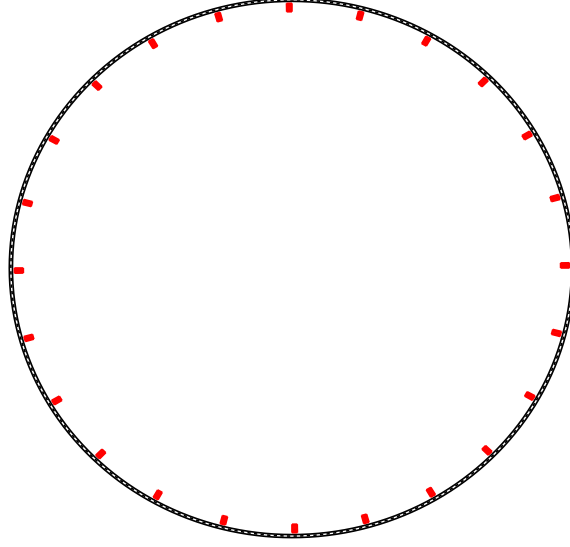


Figure 4.2: Illustration of the locations of the 24 E989 calorimeters relative to the muon storage region. The calorimeters are shown in red. The radius of the E989 storage ring is 7.112 m.

deployed in the E989 experiment.

As a positron deposits its energy in a  $\text{PbF}_2$  crystal, optical and ultraviolet photons are produced through the Cherenkov process. Cherenkov radiation occurs when a charged particle passes through a material faster than the speed of light in that material [1]. Given  $\text{PbF}_2$ 's index of refraction ( $n$ ) of 1.8, Cherenkov radiation will only occur for positrons traveling faster than  $c/1.8$ . This corresponds to a minimum positron kinetic energy of approximately 100 keV; positrons below this energy will not produce Cherenkov radiation. The positron energies of interest in the  $\omega_a$  measurement are greater than 1 GeV, significantly above the Cherenkov threshold.

Cherenkov radiation is also by nature very fast, meaning it is produced only while charged particles are traversing the crystals. There is no delay as one would expect from a scintillator. Considering the 14 cm crystal lengths, the relevant time scale for Cherenkov light production is on the order of 1 ns. Even allowing photons several reflections within a crystal before being

Table 4.1: Relevant properties of lead fluoride ( $\text{PbF}_2$ ) [67].

Property	Value
Density	7.77 g/cm <sup>3</sup>
Molière radius	2.1 cm
Index of refraction	1.8

observed, one should expect all the associated Cherenkov light to be collected within 10 ns at most of a positron impact. This naturally short time scale is very advantageous in reducing the number of unresolved pileup pulses. E989 uses commercially produced  $\text{PbF}_2$  crystals, grown by SICCAS [67].

#### 4.1.2 Silicon Photomultipliers

At the rear face of each  $\text{PbF}_2$  crystal sits a silicon photomultiplier, or SiPM. SiPMs fulfill a purpose similar to that of photomultiplier tubes: they output a current that increases with the intensity of incident light. They are able to operate at very low light intensities and detect optical pulses with sizes ranging from a single photon to thousands of photons. SiPMs are solid state devices able to operate in regions with high magnetic fields [68]. They are also compact: a SiPM and its readout board fit easily within the 25 mm×25 mm boundaries of a  $\text{PbF}_2$  crystal face. SiPMs maintain high photon detection efficiency (PDE) at near ultraviolet wavelengths, an important region for Cherenkov photons (Figure 4.3). These properties make SiPMs an attractive choice for Cherenkov photon counters in E989. SiPMs exhibit a high temperature sensitivity and are intrinsically nonlinear.

SiPMs are arrays of discrete avalanche photodiodes (APD) operating in Geiger mode. Each APD is called a pixel. Operation in Geiger mode means that each pixel is reverse biased above the APD breakdown voltage,  $V_{bd}$ . Following a photon impact on a pixel's photosensitive surface, the photoelectric effect can lead to production of charge carriers

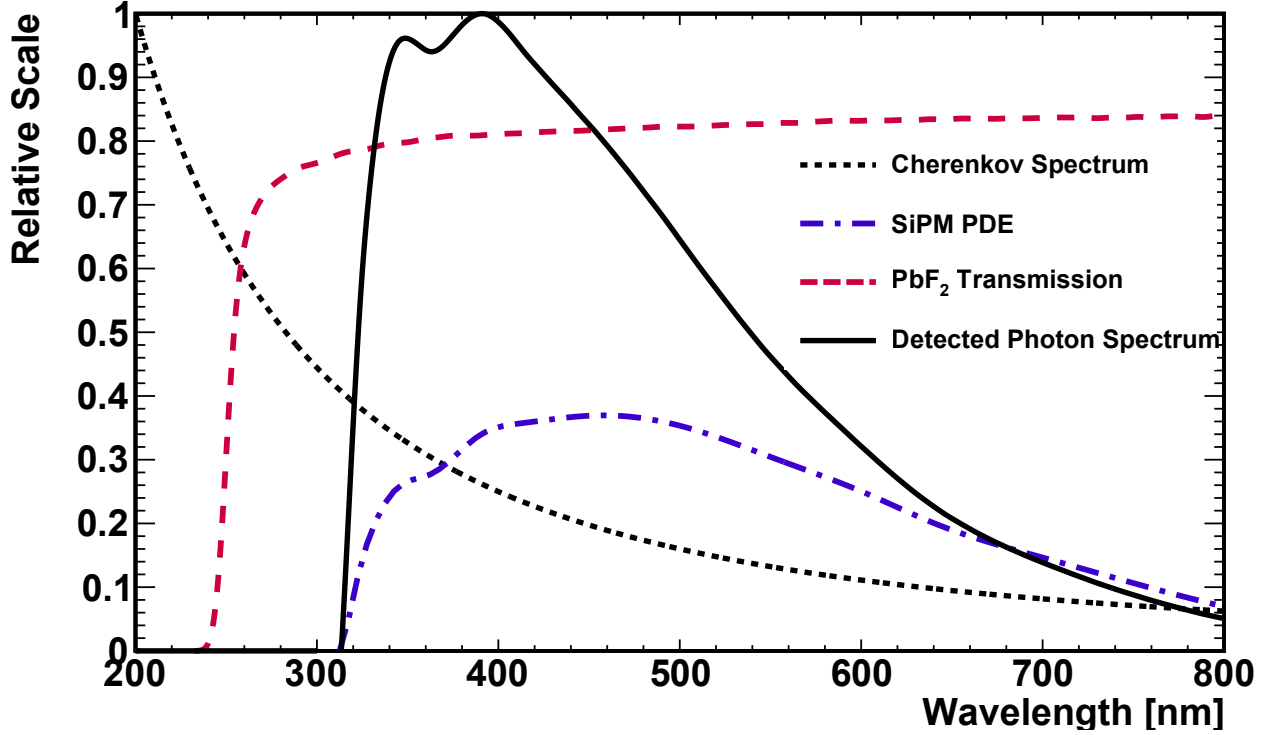


Figure 4.3: Cherenkov photon spectrum, reproduced from [69]. SiPM photon detection frequency, and the detected photon spectrum (product of the others) in the E989 calorimeter. SiPMs are sensitive to photons with wavelengths down to 300 nm. PbF<sub>2</sub> is mostly transparent to photons with wavelengths down to 240 nm.

(electron-hole pairs). As these charge carriers travel through the electric field produced by the applied reverse bias, they collide with atoms in the crystal lattice and release more charge carriers, which in turn travel through the electric field and collide with more atoms. This process is called an avalanche, and it drives a pixel into conduction. Each pixel contains a resistor in series with the APD, called the quenching resistor. As current begins flowing through a pixel following an avalanche, there is a voltage drop across the quenching resistor. Once the voltage drop across the quenching resistor brings the bias across the APD down to the breakdown voltage, the APD stops conducting and the avalanche ceases. At this point, the APD acts like a capacitor and the current draw slowly decays as the capacitor charges back to the applied bias voltage, a process called pixel recovery. Avalanches themselves are

very fast, yielding pulse rise times of approximately 3 ns [63, 68]. Most of the measured current flows during the recovery process. The charge collected during the recovery is that required to charge the pixel capacitor from the breakdown voltage,  $V_{bd}$ , to the bias voltage,  $V_{bias}$ . Therefore, the charge collected from a single pixel following an avalanche is  $Q = (V_{bias} - V_{bd}) \cdot C_{pixel}$ . The quantity  $V_{bias} - V_{bd}$  is called the overvoltage [68].

While the above is a simplified explanation of SiPM operating principles, it provides sufficient information for understanding some critical SiPM properties. For one, the equation for the collected charge following an avalanche makes no reference to the number of incident photons. This is what the term Geiger mode refers to: a given pixel either fires (has an avalanche) or it does not. A pixel is not sensitive to the number of incident photons responsible for the avalanche. Once a pixel fires, it must recover before it can fire again. A SiPM's ability to count photons comes from its large number of pixels. This design leads to the nonlinearity intrinsic to SiPMs. A SiPM can never count more photons than the number of pixels it has.

To further understand the nonlinearity, consider the differential equation

$$dN_{fired} = \epsilon \cdot \left(1 - \frac{N_{fired}}{N_{tot}}\right) \cdot dN_{\gamma}, \quad (4.1)$$

where  $N_{fired}$  is the number of fired pixels,  $\epsilon$  is the PDE (the probability an incident photon will induce an avalanche),  $N_{tot}$  is the total number of SiPM pixels, and  $N_{\gamma}$  is the number of incident photons. This differential equation is a continuous approximation of the statement that for each additional incident photon, the probability that photon will cause a pixel to fire is the PDE multiplied by the probability it will hit a pixel that has not already fired. The probability that a photon will hit a pixel that has not already fired is  $(1 - N_{fired}/N_{tot})$ , assuming uniform illumination of the SiPM's active surface. The solution to the above differential equation is

$$N_{fired} = N_{tot} \left(1 - e^{-\epsilon N_{\gamma}/N_{tot}}\right). \quad (4.2)$$

This equation has the expected limits. As the number of incident photons becomes very large compared to the number of pixels, the number of fired pixels approaches the total number of

pixels. When the number of incident photons is very small, a leading order Taylor expansion will be sufficient. The expression in that case is

$$N_{fired} \approx \epsilon N_{\gamma}. \quad (4.3)$$

In the limit of a small number of incident photons relative to the total number of pixels, the SiPM behaves linearly. The SiPMs used in E989 have 57,344 pixels, and the typical number of fired pixels from a high-energy decay positron is about 2000. Using the above approximation, with 2000 incident photons there is a 2% reduction in the number of fired pixels compared to the perfectly linear case.

The charge collected following a SiPM avalanche is related not solely to the applied bias voltage, but to the overvoltage, which is the difference between the applied voltage and the APD breakdown voltage. This is critical because the breakdown voltage is not a static quantity: it is temperature dependent [68]. Lab measurements have yielded a breakdown voltage temperature dependence of 70 mV/°C. Overvoltages typically range from 1 V to 2 V. Thus, assuming a linear dependence of detector gain on overvoltage, a 1°C temperature drift could change the detector gain by 5% or more. This is a huge effect considering the stringent gain stability requirements of the E989 precession frequency measurement. In E989, temperature-dependent gain drifts are removed with an offline software correction determined from laser pulse data, as will be described later in this chapter.

The SiPM PDE is also not a static quantity. It depends both on photon wavelength and overvoltage. The former effect can be mostly neglected after averaging over the relevant photon wavelength spectrum, but the latter effect has measurable consequences. Gain is often defined as the charge collected per fired pixel divided by the elementary charge [68]. For the purposes of E989, the relevant gain is the size of the recorded pulse divided by the number of incident photons, or, alternatively, divided by the total energy of the positron impacting the calorimeter. For gain defined in this way, it is not sufficient to consider only the charge collected per fired pixel because the PDE affects the number of fired pixels per incident photon. While gain defined as charge collected per fired pixel is linear in overvoltage,



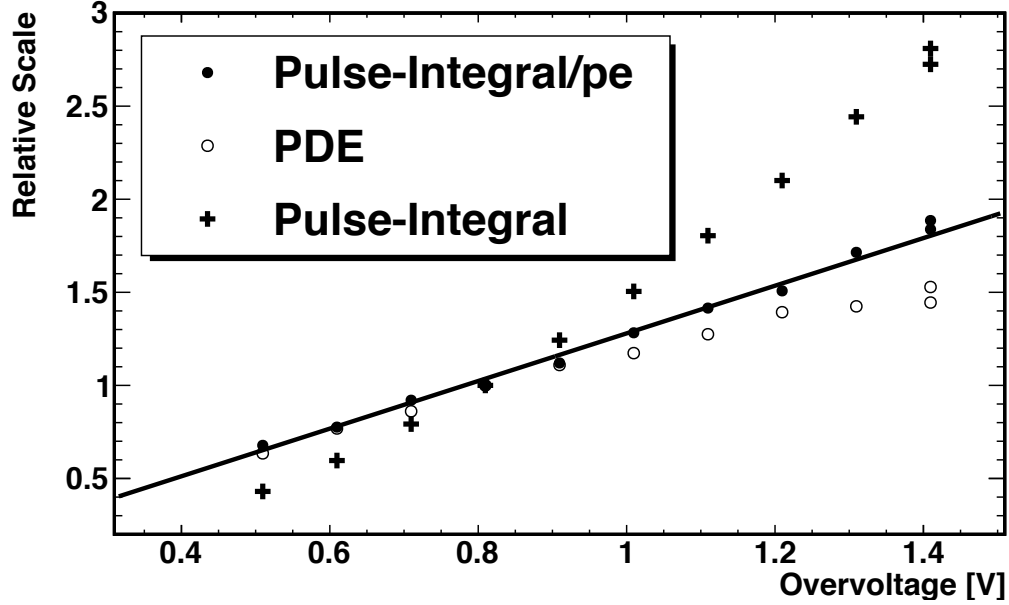


Figure 4.4: SiPM PDE, pulse-integral, and pulse-integral per pe as a function of overvoltage for a fixed intensity of incident light provided by a laser, reproduced from [69]. Photoelectrons or photon equivalents, pe, is the number of fired pixels estimated by looking at the statistical distribution of detected pulse sizes over many laser shots. The pulse-integral represents the size of the pulse. As expected, the pulse-integral per pe increases linearly with overvoltage. The pulse-integral itself, however, increases faster than linearly with overvoltage. This is explained by an increase in PDE with overvoltage. Overvoltage of 0 is defined as the point where the linear fit to pulse-integral / pe vs overvoltage intercepts the x-axis.

the dependence of PDE on overvoltage means that the size of the pulse detected for a fixed amount of incident light varies quadratically, or even faster, with overvoltage. Figure 4.4 shows a measurement of this effect. SiPM pulse size is typically characterized in units of photoelectrons or photon equivalents (pe). The number of pe is related to the integral of the SiPM pulse's current signal through a calibration process. In a linear approximation, the number of pe equals the number of fired pixels. The calibration procedure that yields pe per pulse-integral will be described in Section 4.1.3. While pulse-integral per fired pixel increases linearly with overvoltage, the pulse-integral itself increases faster-than-linearly with overvoltage. The dependence of PDE on overvoltage increases the gain sensitivity to temperature

Table 4.2: E989 SiPM properties and other relevant quantities.

Quantity	Value
Manufacturer	Hamamatsu
Model number	S12642-4040PA-50
Typical bias voltage	66–69 V
Characteristic pulse width	10 ns
Active area	12 mm×12 mm
Number of pixels	57,344
Typical number of fired pixels from a positron shower	2000

beyond what it would be otherwise. Additional effects not described here include afterpulsing and crosstalk, which both increase with overvoltage and contribute to the faster-than-linear change of observed signal size with overvoltage [68].

A custom electronic readout board was designed for the purposes of E989. The current pulses from each SiPM pass through multiple stages of amplification and are ultimately output as differential voltage signals. The final stage of amplification uses a programmable, variable gain differential amplifier. The gains of each of the 1296 SiPMs are tuned individually using this amplifier to equalize responses across all channels. A board mounted temperature sensor is also included with each SiPM to aid in understanding and controlling temperature variations. The output signal passes through a pole-zero network that cancels the exponential tail of the current pulse resulting from pixel recovery. The ultimate output pulses have a 5 ns full width at half maximum. See Figure 4.5 for photographs of the SiPM and its readout board. The circuit schematic is available in the article by Kaspar et al. [63]. Important features of the SiPMs used in E989 are summarized in Table 4.2.

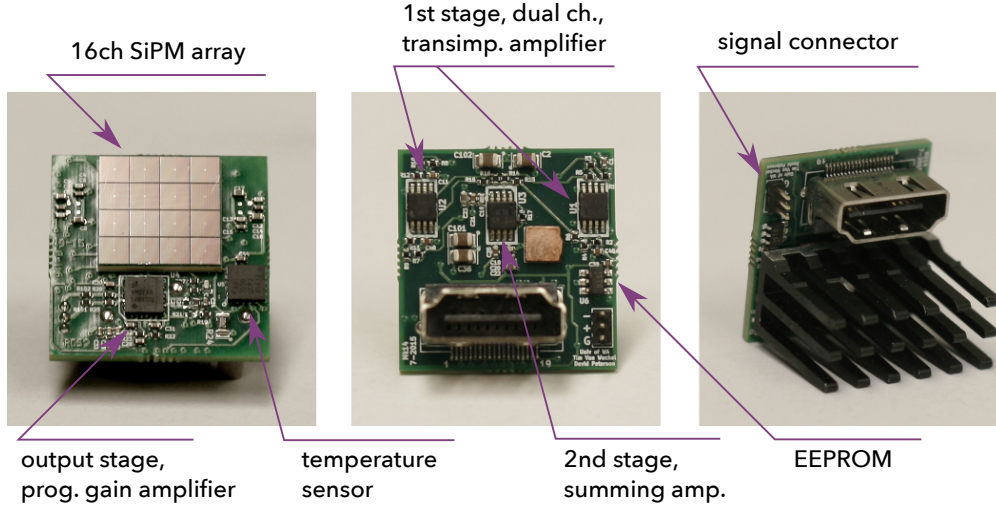


Figure 4.5: The E989 SiPM readout board, reproduced from [63]. A circuit schematic is available in the original article.

#### 4.1.3 Laser Calibration System

To achieve the experiment's target systematic uncertainty, the E989 calorimeter gains after all corrections must be stable to better than 0.1% over the time scale of a  $700\text{ }\mu\text{s}$  muon fill. Additionally, the temperature dependent SiPM gains must be understood and corrected over long time periods, hours to days, to maintain consistent calorimeter energy measurements. A laser calibration system has been constructed and installed to address these requirements. The laser calibration system features six independent laser heads that, through a complex system of mirrors, beam splitters, diffusers, and laser fibers, provide on demand laser pulses to each SiPM. A collection of neutral density filters can be remotely placed in the paths of the laser beams. This allows the laser pulse intensity seen by the SiPMs to be varied through a known set of values. Furthermore, the laser system provides a pulse at the beginning of

every muon fill, called the sync pulse, that is used for timing alignment between channels. In addition to allowing for gain and timing measurements and corrections, the laser system is an invaluable tool for calorimeter testing and diagnostics [66,70]. A schematic of the laser calibration system is shown in Figure 4.6.

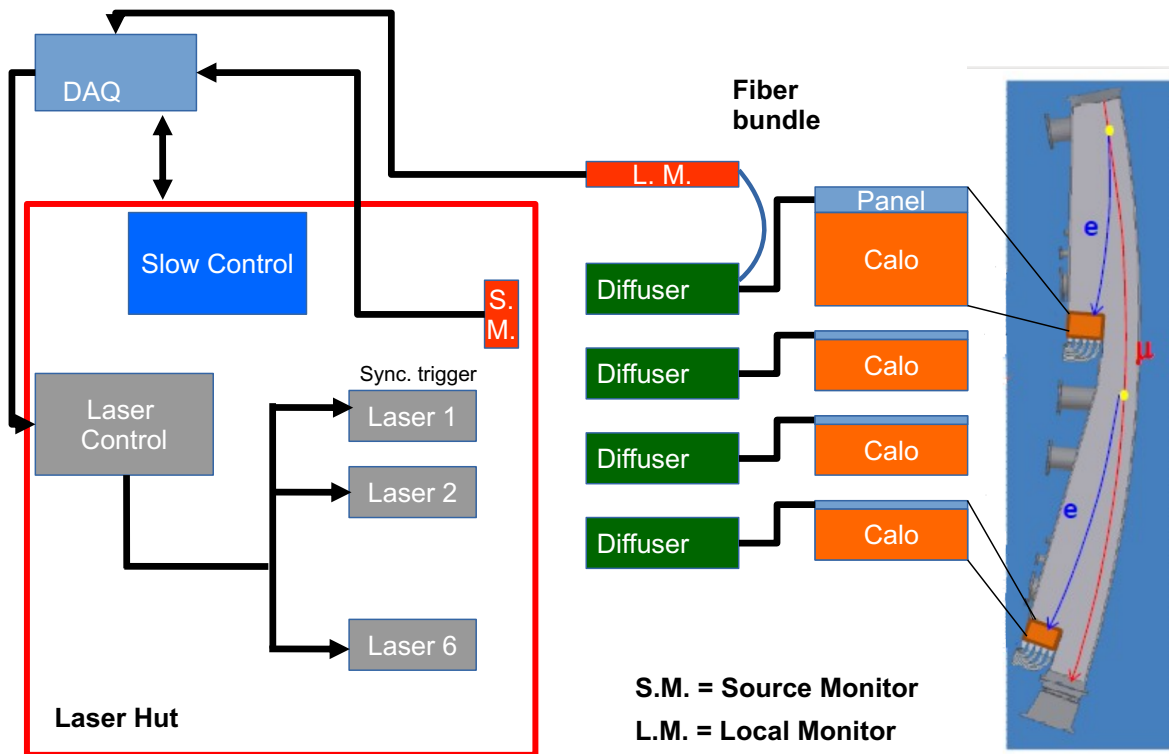


Figure 4.6: Schematic of the laser calibration system, courtesy of Andrea Fioretti. Each of the six laser heads feeds four calorimeters. There are separate source monitors for each laser head and separate local monitors at the end of the light distribution chain for each calorimeter. The front panel of the calorimeter housing is engineered to provide laser pulses to all 54 channels simultaneously.

In a leading order approximation, one can assume that the average response of a calorimeter channel to a laser pulse is proportional to that channel's gain and that a change in gain

will affect the channel's response to laser pulses in the same way it will affect the channel's response to positron pulses. While the latter assumption has been confirmed at test beam experiments, as will be discussed in Section 4.3, the former is violated in the case where the laser pulse intensity itself changes systematically from one shot to the next. To reach the required accuracy, variations in laser pulse intensity must be accounted for. Laser pulse intensity is monitored from shot to shot using PIN diode source monitors that receive a large fraction of the laser light before it is distributed to the calorimeter channels. Additionally, to ensure the distribution system itself is stable, local monitor photomultiplier tubes receive light from the end of the distribution system at each calorimeter. There are 12 PIN diode source monitors, two per laser, and 24 photomultiplier tube local monitors, one per calorimeter. After accounting for variations in the laser pulse intensity and distribution chain using these monitors, changes in SiPM gain can be measured by comparing the size of a laser pulse at some time of interest to the size of a laser pulse at a reference time, typically the time of the last absolute calibration.

The laser system can be used for absolute calibration. Consider the statistical distribution governing the number of fired pixels from one laser pulse to another. It is sufficient to neglect SiPM nonlinearities in this consideration because absolute calibrations are conducted at low light levels. There are three contributions to the observed statistical fluctuations of the number of fired pixels:

1. electrical noise,
2. Poisson fluctuations of the number of incident photons triggering avalanches,
3. laser intensity fluctuations from shot to shot.

As long as the laser intensity fluctuations are not too large, these three components can be treated as uncorrelated and the expression for the variance of the number of fired pixels is

$$\sigma_p^2 = n^2 + \langle p \rangle + \epsilon^2 \langle p \rangle^2, \quad (4.4)$$

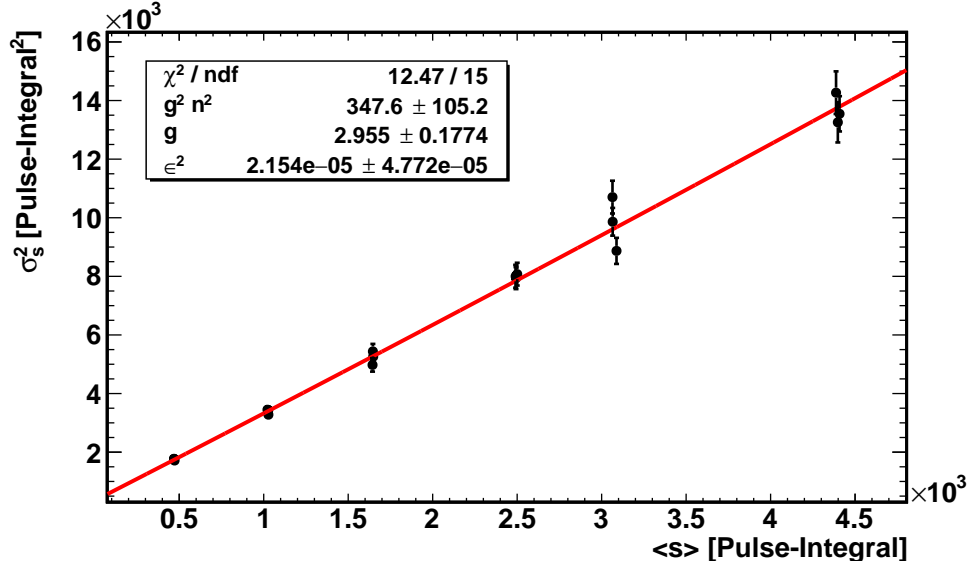


Figure 4.7: Example of a photostatistics calibration, described in the text, from E989 Run 1. For this detector, the extracted gain,  $g$ , is approximately 3 pe per pulse-integral. Considering this gain in conjunction with the range of  $\langle s \rangle$  values shown on the x-axis, one determines that the laser pulses employed in this calibration run produced SiPM signals ranging from 170 pe to 1500 pe.

where  $n$  is the electrical noise level in units of fired pixels,  $\langle p \rangle$  is the average number of fired pixels over many laser shots, and  $\epsilon$  is a unitless measure of the relative size of the laser fluctuation, which should be on the order of a percent or less for this treatment to be valid. The three terms in the above equation correspond to the three contributions mentioned before, with the key feature that the contribution to the variance from discrete photon statistics is equal to the mean number of fired pixels.

What is measured in the calorimeters is not directly the number of fired pixels, but an electrical signal that is proportional to the number of fired pixels; the measured signal  $s$  is equal to  $g \cdot p$ , where  $g$  is the channel gain and  $p$  is the number of fired pixels. Noting that  $g \cdot \sigma_p = \sigma_s$ , rewriting the above equation in terms of  $s$  gives

$$\sigma_s^2 = g^2 n^2 + g \langle s \rangle + \epsilon^2 s^2. \quad (4.5)$$

The coefficient of the linear term in the above expression is  $g$ , the gain, in units of signal per fired pixel. Signal could be collected charge, ADC, or any proxy for pulse size. This relation can be used in practice by observing a large number of laser pulses, measuring the variance and mean of the distribution of measured pulse sizes, and repeating for a number of different laser settings by varying the neutral density filters. The variances of the resulting statistical distributions, one for each neutral density filter setting, can be plotted against the means and fit with a parabola. The linear term extracted from such a fit will be the gain,  $g$ . Using the laser calibration system, all 1296 SiPMs can be subjected to this procedure simultaneously and the gains determined for all of them. Through combination with an external measurement of average pixels fired per MeV of incident particle energy, as can be extracted from test beam experiments (see Section 4.3), one obtains an absolute calibration of each channel in terms of signal per MeV. An example calibration curve is shown in Figure 4.7. Absolute calibrations conducted in this way are called photostatistics calibrations. Photostatistics calibrations are accurate to approximately 10% and are limited by variations in the optical properties of different PbF<sub>2</sub> crystals and their SiPM couplings. In E989 Run 1, the SiPM amplifier settings were chosen to equalize gains as measured by this procedure.

Long term gain tracking is achieved through out-of-fill laser pulses, i.e. laser pulses fired when there are no muons present in the ring. Reference signal sizes in the calorimeter channels and in the laser monitors are recorded at absolute calibration time and the out-of-fill pulse sizes are compared to these references. The comparison provides a measurement of how much the gains have drifted since the last calibration. Neglecting the local monitor correction, the gain at time  $t$  in terms of the SiPM signals  $s$  and the monitor signals  $m$  is

$$g(t) = g(0) \cdot \frac{s(t)}{s(0)} \cdot \frac{m(0)}{m(t)}. \quad (4.6)$$

The first ratio,  $s(t)/s(0)$ , is based only on the SiPM signals and corrects the gain under the assumption that the laser intensity has not changed. The second ratio,  $m(0)/m(t)$ , is based on the source monitor signals and accounts for any drifts in average laser intensity.

The SiPM and monitor signals are averaged over many pulses to remove the effects of shot-to-shot statistical fluctuations. This procedure is effective because the response of the PIN diode source monitors is intrinsically stable over long periods of time [71].

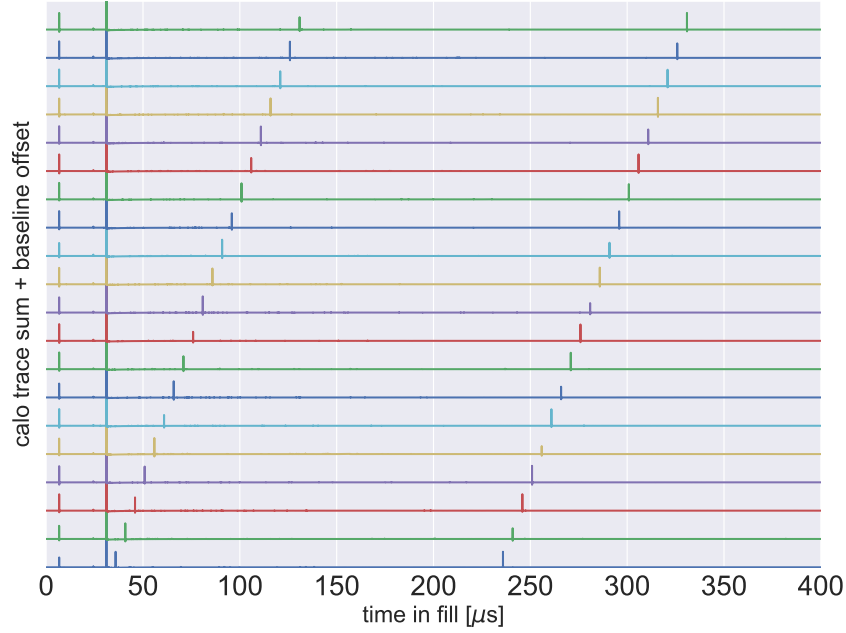


Figure 4.8: Each curve is the sum of all SiPM voltage signals, or traces, within a given calorimeter across the first 400  $\mu\text{s}$  of a muon fill. Within each trace, there are four clear pulses: the first is the sync pulse used for timing synchronization on all fills, the second and largest pulse is the beam injection, following which most of the injected particles are lost and produce large signals in nearby calorimeters, and the last two pulses are the in-fill laser pulses described in the text. Each subsequent vertically offset curve is the next fill in which in-fill laser pulses appear. For the data shown here, taken from early in E989 Run 1, the in-fill laser pulses were shifted in 4  $\mu\text{s}$  steps. Later on, the time shift was reduced to 2  $\mu\text{s}$ .

The laser calibration system also provides a measurement of consistent, early-to-late gain variations during muon fills. The measured early-to-late gain perturbation is used to correct reconstructed pulse energies during analysis. The goal is for residual in-fill gain changes after this correction to be small enough that the ultimate gain systematic uncertainty is below



20 ppb. The in-fill gain measurement is achieved as follows: during approximately 10% of muon fills, two to three in-fill laser pulses are fired, each separated by  $200\,\mu\text{s}$  from the last. The next time in-fill laser pulses are fired, the pulse times are shifted by a configurable amount of time, typically  $2\text{--}4\,\mu\text{s}$ . In this way, over time, the entire time domain of a muon fill is sampled with laser pulses. The SiPM gain at the time of each pulse is determined from the reconstructed SiPM pulse sizes corrected for laser intensity fluctuations using the source monitors and normalized to the sync pulse, which is fired when the SiPMs are in an unperturbed state. See Figure 4.8 for a visualization of the in-fill laser pulse pattern.

Further capabilities of the laser calibration system that will not be discussed in detail here include a programmable control board that can trigger the laser at an exponentially decreasing rate to simulate conditions produced by muon decay [72] and remote controllable mirrors that can redirect light from two lasers into a single calorimeter to test its response to separate pulses with a small but well known time separation. These features are valuable tools for verifying calorimeter performance and for measuring systematic effects. The laser calibration system is a vital feature of the E989 experiment, and it is essential for achieving the target systematic uncertainties.

## 4.2 *Reconstruction Software*

The term reconstruction refers to the process by which raw recorded data is transformed into meaningful quantities such as positron hit energies and times. These quantities are called reconstructed objects or physics objects and are the basic ingredients for physics analyses. Raw data can be understood as the data recorded from a detector system before any processing has occurred. Reconstruction software is the computer code responsible for the transformation from raw data to physics objects. There is no single way to do reconstruction; the E989 collaboration already has two independent methods for calorimeter reconstruction, and there may be more in the future. The resulting physics objects will necessarily differ somewhat based on diverging decisions made during the respective algorithm and software development processes. These differences between reconstruction procedures aid in char-

acterizing and understanding each given approach. Applying multiple reconstructions to the same raw data helps verify correctness and is an important check on systematic effects. This section will describe one approach to calorimeter reconstruction. All following analysis discussion is based on the reconstruction procedure described here.

#### *4.2.1 Input to Output*

Raw calorimeter data takes the form of digitized waveforms. These waveforms are voltage versus time traces output from each SiPM readout board. The employed custom digitizers record continuously throughout each  $700\ \mu\text{s}$  muon fill at  $800\ \text{MS/s}$  and have 12-bit resolution [37, 73, 74]. Given the  $11.4\ \text{Hz}$  muon fill rate, recording all the raw traces would consume close to  $20\ \text{GB/s}$  of storage space, an intractably high number. To overcome this limitation, the  $700\ \mu\text{s}$  digitized waveforms pass through GPUs for online preprocessing before anything is saved to disk [37, 75]. The GPUs comb through each trace to search for voltage spikes above some threshold. Such spikes indicate activity in the calorimeter. Only over threshold samples plus a configurable number of consecutive pre- and post-samples are saved to disk. These short sub traces are called islands. Island lengths are dynamic, longer when there is more calorimeter activity, but the typical length of an island triggered by an isolated positron pulse is  $40\ \text{ns}$ . During E989 Run 1, the data acquisition was configured to save islands from every SiPM waveform in a given calorimeter whenever any one of them went over threshold, with all the islands taken from the same range of sample numbers within the original digitized waveform. All the data described in this document was collected under that configuration. Later, in the summer accelerator shutdown following Run 1, the data acquisition was reconfigured to save data only from the triggering SiPM and its nearest neighbors; this was to further reduce the data rate. The Run 1 data acquisition configuration was sufficient to reduce the raw data rate by a factor of 100 compared to saving the full,  $700\ \mu\text{s}$  digitized waveforms. Digitizer islands are the raw input to the calorimeter reconstruction chain. See Figure 4.9 for an example island.

The primary objective of any calorimeter reconstruction process is to take these raw

islands and produce collections of positron hit times and energies. This is the information needed to conduct an  $\omega_a$  analysis. Additionally, but less critically, the reconstruction can provide estimates of impact positions and angles on the calorimeter face. There are a number of common steps required by any reconstruction procedure, although the order and boundaries can vary. These steps are as follows.

- Trace corrections: correcting the island waveforms for known and undesired digitizer effect.
- Pulse finding: converting digitized waveforms into one or more pulse size and pulse time measurements.
- Energy calibration: converting pulse sizes into meaningful, absolute energy units that are common between all channels.
- Gain corrections: adjusting these energies for any gain drifts since the last absolute calibration.
- Timing corrections: adjusting pulse times for fixed inter-channel timing offsets, fill-to-fill relative timing jitter, and variations of the beam injection time.
- Clustering: the aggregation of calibrated crystal hit times and energies into best estimates for positron hit times and energies. The reconstructed hit times and energies are called clusters.

One can imagine walking through these steps for the island shown in Figure 4.9. First, the pulses in each channel would be summarized with a pulse-integral and a peak time. Pulse-integrals are a convenient proxy for pulse energy, as the integral of a voltage versus time trace in a given circuit is proportional to the charge associated with the pulse. The pulse-integrals then must be converted into units of deposited MeV through a known conversion factor taken from a calibration and incorporating gain correction factors provided by the laser

system. Then, the pulses must be aligned to a common reference to remove any differences in digitizer trigger times or pulse rise times. Finally, the collection of pulses must be combined, or clustered, into a single energy and time, which is the best guess of the energy and time of the impacting decay positron responsible for the constituent pulses. By following these steps, the reconstruction software takes calorimeter islands and produces calibrated clusters.

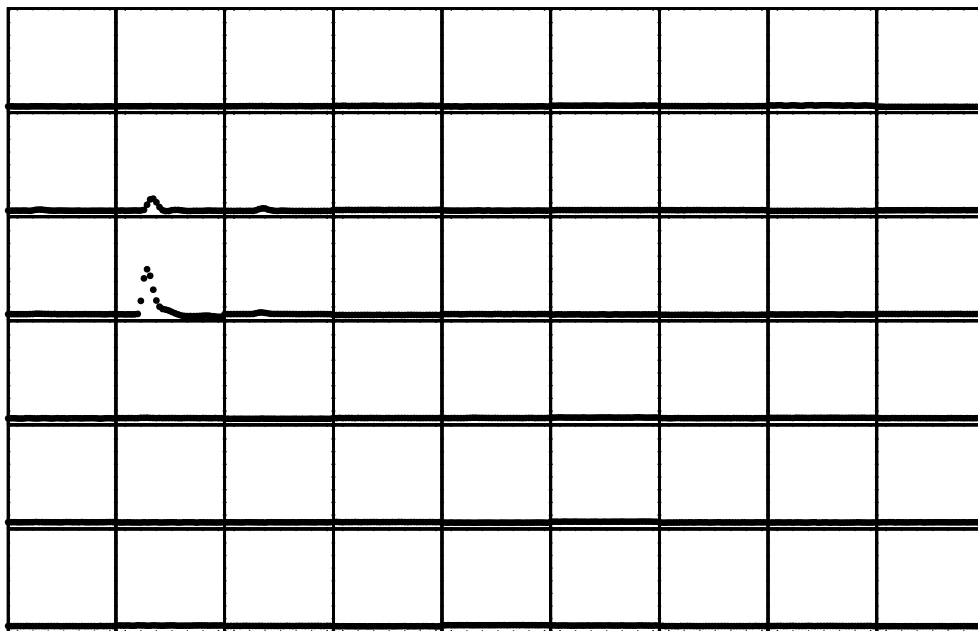


Figure 4.9: An example digitizer island resulting from an isolated positron impact in a calorimeter. Each square contains a voltage versus time signal for one of the 54 SiPMs in this calorimeter. These waveforms are 35 samples—about 44 ns—long. They are arranged following the physical layout of the calorimeter. The left side is closest to the storage ring. There is clear activity in the upper left, where there is a large pulse in the second column from the left and fourth row from the bottom. There are smaller pulses in the surrounding channels. Activity in any crystal causes the data acquisition system to store islands from all 54 channels.

#### 4.2.2 Pedestal Correction

Each digitizer waveform sits atop a baseline, or pedestal. The pedestal is the average sample value when there is no calorimeter activity. The pedestal value is not generally 0, and can be

adjusted through the digitizer configuration. The custom digitizers used in E989 feature two interleaved 400 MS/s ADC chips that together achieve the overall 800 MS/s digitization rate. These two chips have slightly differing pedestal values, typically by 1 to 5 ADC counts out of the 4096 possible ADC values. This behavior imparts an even-odd effect to the digitizer waveforms whereby every other sample sits slightly higher or lower. The size of this effect has been measured for each channel, and the first step in the reconstruction process is to correct all island waveforms for the even-odd pedestal differences. The correction process is, for each channel, to subtract from every other sample the measured average difference between odd and even samples.

#### *4.2.3 Pulse Shapes and Templates*

The pulse finding procedure is based on templates, empirical functions describing the voltage versus time signals resulting from SiPM pulses. For the E989 Run 1 analysis, 250 ns-long templates were constructed for each detector channel. See Figure 4.10 for example templates from a typical SiPM. Ideally, template functions have only two parameters: the pulse-integral (pulse size) and the peak time. Describing waveforms with functions of only these two parameters assumes that pulse shapes are independent of size and time and that any pulse can be described by translating and scaling the proper SiPM template. The correctness of a template based procedure is tied to the validity of these assumptions. Using templates has the advantage of removing the need to model the complex behavior of the SiPM readout circuits, and as long as the assumption of a fixed pulse shape is reasonably satisfied, templates are well suited to resolving pileup pulses. Before discussing how templates are used to fit pulses and resolve pileup, the template construction procedure will be discussed and the fixed pulse shape assumption examined.

A template is a function of time,  $T(t)$ , that describes a SiPM's pulse shape. A leading order approximation of  $T$  could come from collecting a number of digitized SiPM pulses, normalizing them by pulse-integral, aligning them by their peak sample, averaging them, and then interpolating between the samples of the resulting average waveform. While this is

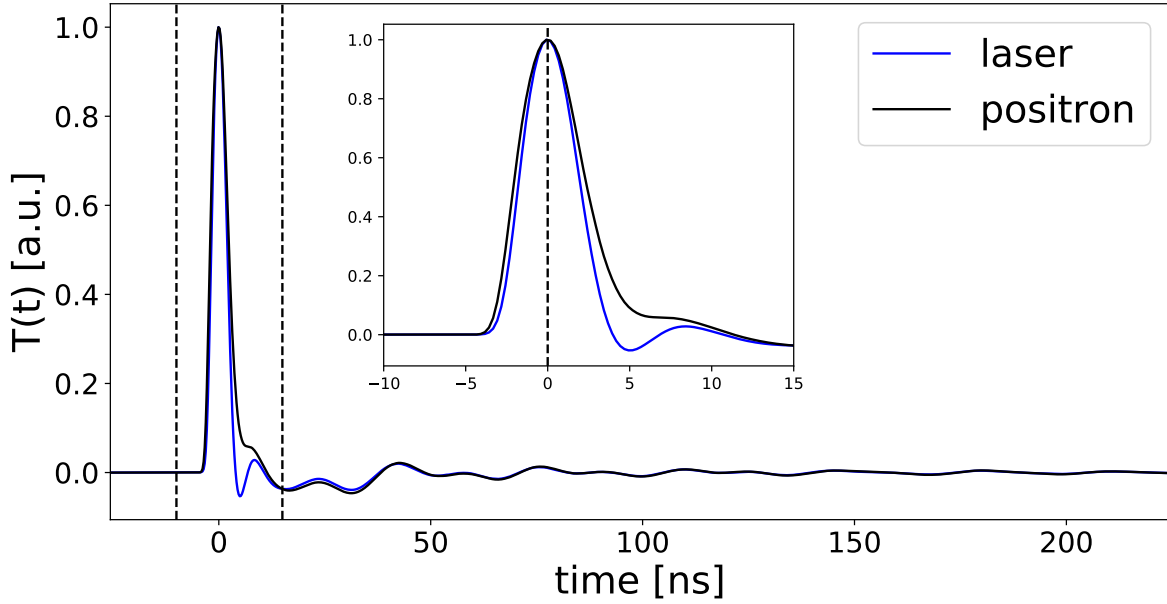


Figure 4.10: Example SiPM templates modeling a single detector channel's response to pulses from differing sources. The blue and black curves show templates built using laser pulses and positron pulses, respectively. Laser pulses are narrower than positron pulses and have a sharp undershoot following the initial peak. The templates extend over 250 ns.

quite close to how templates are generated, it does not yield acceptable results because the SiPM pulse is fast compared to the digitizer sampling period of 1.25 ns. The peaks of these fast pulses are defined by only three or four digitizer samples. To get a better estimate of the average pulse shape, one must find a way to sample more finely. This can be achieved by aligning individual pulses not by their peak sample time, but by their estimated peak time on a continuous domain. With a noise level that is small relative to the difference between the peak sample and its neighbors, the true pulse peak will be within  $\pm 1/2$  of a digitizer clock tick of the peak sample. Assuming the source of recorded pulses has no relation to the digitizer clock, pulses should be sampled at random phases and pulse peaks should be uniformly distributed between digitizer samples.

There are multiple ways to estimate the pulse peak time from a digitizer trace. One way would be to match a parabola to the peak sample and its two neighbors and then take

the time corresponding to the parabola's maximum. This procedure is not ideal when the pulse is sampled so coarsely that the pulse shape is not well described by a parabola in the three clock tick region surrounding its peak. Another approach is based on a quantity called pseudotime,  $t_p$ . Calling the peak sample of a trace  $s_0$ , the sample to the left of the peak  $s_-$ , and the sample to the right  $s_+$ , pseudotime is defined as [34]

$$t_p \equiv \frac{2}{\pi} \arctan\left(\frac{s_0 - s_-}{s_0 - s_+}\right). \quad (4.7)$$

This quantity is useful as it categorizes each pulse with a number between 0 and 1. When the pulse peak is as far to the left of the maximum sample as possible,  $s_0 = s_-$ . In this case, according to the above equation,  $t_p = 0$ . Similarly, when the peak is as far to the right of the maximum sample as possible,  $t_p = 1$ . Based on the form of the above function, as the peak moves across a digitizer bin,  $t_p$  varies smoothly from 0 to 1. It is not obvious how exactly the pseudotime varies between these limits. For instance, the pseudotime will not in general equal 0.5 when the pulse peak sits exactly on a digitizer sample.

A more useful quantity for aligning pulses would increase linearly over a range of 1 unit as a pulse peak moves from the left edge of a digitizer bin to the right edge. To contrast with pseudotime, this quantity is called realtime,  $t_r$ . Realtime is defined to have the property that shifting a waveform's peak sample to the time  $0.5 - t_r$  places its continuously interpolated peak at  $t = 0$ . Given that both pseudotime and realtime increase smoothly over a range of 1 unit as a pulse peak moves across a digitizer bin, there must exist a unique map from pseudotime to realtime. This map can be constructed from a distribution of pseudotimes extracted from pulses with random peak times.

First consider a pulse that is symmetric within the  $\pm 1/2$  clock tick region surrounding its peak. The E989 SiPM pulses are close to this limit. For such a pulse,  $t_p = 0$  implies  $t_r = 0$  and  $t_p = 1$  implies  $t_r = 1$ . Additionally, with random peak times relative to the digitizer boundaries, a quarter of all pulses will have realtimes less than 0.25, half will have realtimes less than 0.5, three quarters less than 0.75, etc. Therefore, the pseudotime that corresponds to a realtime of 0.25 will be the one below which sits 25% of the pseudotime distribution. In

other words, the map from pseudotime to realtime for a pulse that is symmetric near its peak is the cumulative distribution function, or CDF, of the pseudotime distribution. Assuming a normalized pseudotime probability distribution  $\rho(t_p)$ ,

$$t_r(t_p) = \int_0^{t_p} \rho(t'_p) dt'_p. \quad (4.8)$$

Given a distribution of pseudotimes taken from data, one can numerically calculate the needed map to realtime. The map from pseudotime to realtime depends on the pulse shape of the detector in question.

For a pulse that is asymmetric in the  $\pm 1/2$  clock tick region surrounding its peak,  $t_p = 0$  does not imply  $t_r = 0$ . However, it is still the case that  $t_p$  increases from 0 to 1 as the pulse peak moves across a digitizer bin. Therefore, for an asymmetric pulse, Equation 4.8 need only be adjusted by a constant offset that encodes the value of  $t_r$  when  $t_p = 0$ . Ignoring the offset will lead to a pulse template whose peak is not perfectly centered at  $t = 0$ . This offset is accounted for by the time correction procedure described in Section 4.2.6 and can thus be discarded in the template building process.

After generating a map from pseudotime to realtime, pulses can be aligned on a very fine grid and the average pulse shape measured with high time resolution. A cubic spline is used to interpolate between these finely binned points and provides a continuous template function  $T(t)$ , see Figure 4.11. Slight pulse shape differences necessitate building custom templates for each of the 1296 SiPMs. Additionally, to achieve the best performance, separate templates are needed for laser pulses and positron pulses, as was shown in Figure 4.10. This is because the distribution of photon detection times at the SiPM differs significantly between laser pulses and beam pulses.

The template building procedure described here can be used to check whether pulse shapes vary significantly with pulse energy, rate, or time. To check whether pulse shapes vary with size, the template building procedure was applied to laser pulses collected at seven different neutral density filter settings. These settings span a factor of ten in pulse-integral and range in equivalent positron energy from 200 MeV to 2000 MeV, well matched



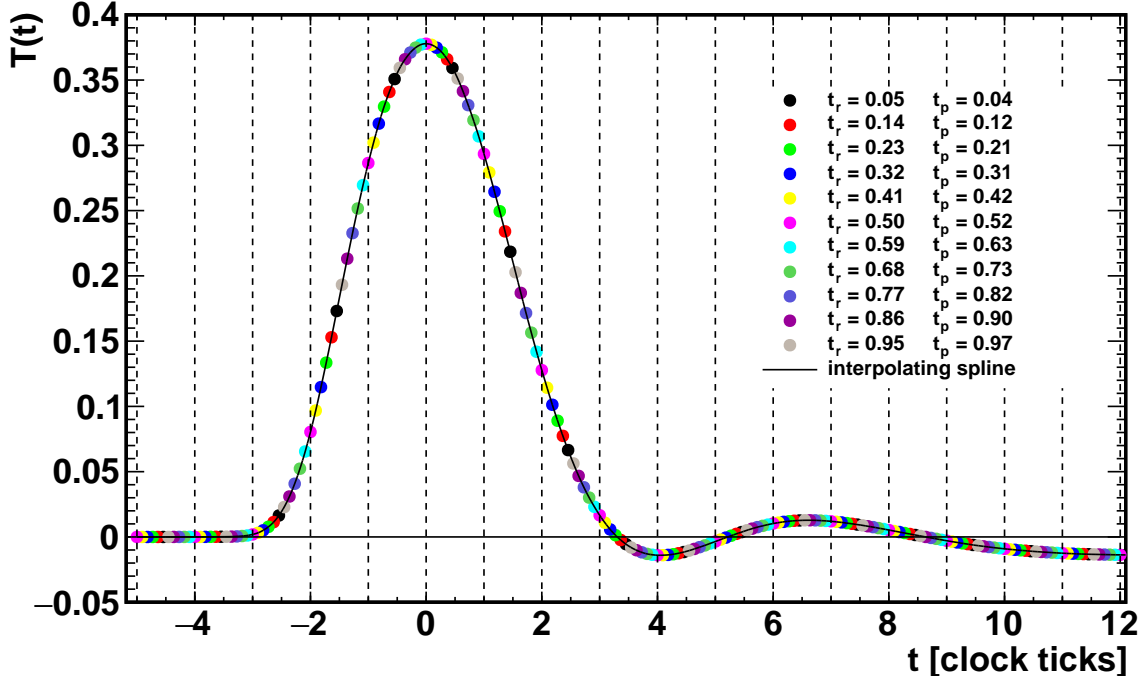


Figure 4.11: A template built from laser pulses. Different colors signify different realtimes as discussed in the text. Points of a given color follow the samples that would be recorded for a single digitized pulse with the specified realtime. Aligning pulses by their realtimes allows for a very fine sampling of the average pulse shape. The legend specifies the realtime and pseudotime for each of the differently colored traces. There are 1.25 ns per clock tick.

to the range of positron energies expected during production running. After normalizing, no discernible difference was observed between the templates built from the differently sized pulses.

Another concern was whether the SiPM pulse shape would be affected by rate. This was also investigated using the laser system. The laser was configured to fire an isolated train of eight pulses with an instantaneous rate of 5 MHz. Templates were built for the first and last pulses in the train. If the SiPM exhibited a rate dependent pulse shape, one would expect these two templates to differ. No significant difference was observed [63].

Furthermore, one may be concerned that aging effects could distort the SiPM pulse shapes

over long periods of time. Comparing templates built from data collected at different times can allay this worry. Laser pulse templates built in the spring of 2017 and spring of 2018 have been compared and show no significant differences. Templates are rebuilt periodically to maintain confidence that pulse shapes have remained stable. Comparisons of positron templates built at the end of E989 Run 1 with those built at the beginning showed that the pulse RMS widths had on average decreased by 1% over a four month period. The experiment hall temperature had increased by 4°C during over this same period, and thus the change in average pulse shape is most likely attributable to temperature and not to aging. The hall temperature stability is expected to be improved for Run 2.

While there is strong evidence that SiPM pulse shapes are stable with size, rate, and time, they do depend on the distribution of photon arrival times at the photosensitive surface. The photon arrival time distribution does not change significantly with differing pulse rates or neutral density filters, although inserting a particularly dense filter will increase the prevalence of statistical fluctuations. However, simulations have indicated that when a positron impacts a given crystal, the distribution of photon arrival times at that crystal's SiPM differs from the distributions at the neighboring SiPMs. This effect has been confirmed with test beam data [69]. As the majority of the electromagnetic shower is contained in the central crystal, neglecting to account for impact position dependent pulse shapes does not create a significant error in the ultimate time and energy extraction. Figure 4.12 depicts the results of the template stability tests that have just been described.

#### 4.2.4 *Template Fitting*

A process for generating empirical functions to model SiPM pulse shapes, called templates, has been described. Templates  $T(t)$  are normalized by the condition

$$\int_{-\infty}^{\infty} T(t) dt = 1, \quad (4.9)$$

where  $t$  is in units of digitizer clock ticks, and templates are time aligned such that their peaks occur at  $t = 0$ . Under the assumption explored above that all pulses are describable

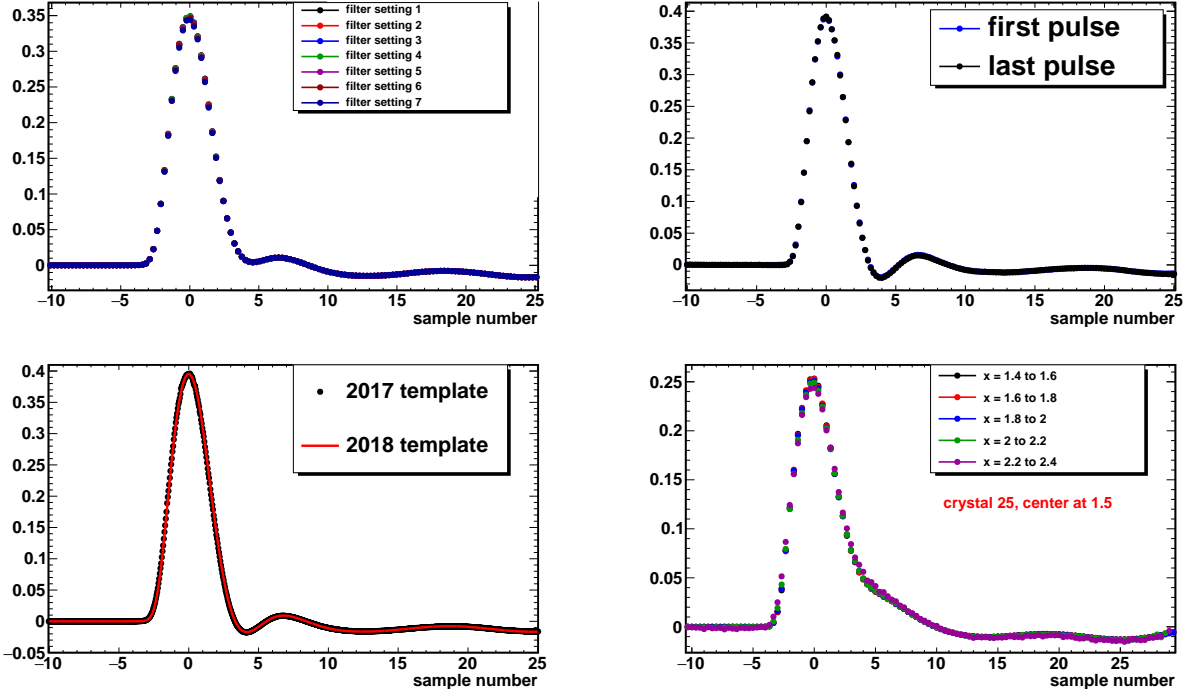


Figure 4.12: Results of the pulse shape stability tests described in the text. The top left shows the pulse shape’s insensitivity to pulse size. The top right shows the pulse shape’s insensitivity to rate: first and last refer to the first and last pulses in a train of eight, each separated by 200 ns. The bottom left is a comparison between laser templates built in spring 2017 and spring 2018. The bottom right depicts the pulse shape dependence on electron impact position,  $x$ . The pulse is narrower when the electron impacts a crystal directly than when the electron impacts a neighbor.

with a translated and scaled template, individual pulse waveforms are modeled by the family of functions

$$f(t|s, t_0, P) = s \cdot T(t - t_0) + P. \quad (4.10)$$

Following the template normalization and alignment conditions, the parameters  $s$  and  $t_0$  correspond respectively to the pulse-integral and the pulse peak time. The parameter  $P$  accounts for the digitizer pedestal. This description generalizes to the case of multiple pulses. If there are  $n$  pulses present, the resulting waveform is described with a superposition

of template functions

$$f(t|s_i, t_i, P) = \sum_{i=0}^n s_i \cdot T(t - t_i) + P, \quad (4.11)$$

where the  $s_i$  are the integrals of each individual contributing pulse, the  $t_i$  are pulse peak times, and  $P$  is the overall pedestal. The pulse finding technique used to extract pulse energies and times from digitized waveforms is called template fitting, whereby digitized waveforms are fit to functions of the above form.

The  $s$ ,  $t$ , and  $P$  parameters are extracted with a  $\chi^2$  minimization procedure. With  $n$  pulses and  $m$  digitizer samples  $\{D_0 \dots D_m\}$  taken at times  $\{t_0 \dots t_m\}$ , the function to minimize takes the following form:

$$\chi^2 = \sum_{i=0}^m \sigma_i^{-2} \left( D_i - \sum_{j=0}^n s_j T(t_i - t_{0,j}) - P \right)^2. \quad (4.12)$$

There are  $2n + 1$  fit parameters: the  $n$  times, the  $n$  pulse-integrals, and the pedestal. An important observation about this function is that it is quadratic in all parameters other than the times, which is to say  $f(t|s_i, t_i, P)$  is linear in all parameters other than the  $t_i$ . For a fixed set of time parameters, the optimal integral and pedestal parameters can be determined exactly by solving a system of linear equations [76]. This reduces the dimensionality of the parameter space that must be numerically searched from  $2n + 1$  to  $n$ . In practice, the  $\chi^2$  optimization was implemented as a Newton's method search over the time parameter subspace [76]. At each step in the time subspace search, the optimal pulse-integral and pedestal parameters are calculated. Linear algebra operations are handled with the EIGEN3 software library [77]. This process capable of handling on the order of  $10^5$  pulses per second per CPU.

In Equation 4.12,  $\sigma_i$  is the statistical uncertainty assigned to each digitizer sample. It was found through simulations that assigning an equal uncertainty to all samples yields the most robust behavior and provides the best pileup resolution capabilities [78]. Thus, in the E989 reconstruction code described here, all digitizer samples are assigned an equal uncertainty of 1.

In the  $\chi^2$  minimization procedure, the number of pulses,  $n$ , is not a fit parameter. Rather, it is externally fixed before the fit occurs. As the number of pulses in an island is an output of the reconstruction procedure, the determination of how many pulses to include in the fit model is a distinct reconstruction task, related to but not the same as the  $\chi^2$  minimization. There is no single way to determine the number of pulses to hypothesize, and there is ambiguity in the case of small pulses that sit near the digitizer noise level and in the case of pulse pairs with very small time separations. The procedure used in this analysis is as follows.

1. Hypothesize a single pulse and fit.
2. Calculate the residuals, the differences between the fit function and each digitizer sample.
3. If any of these residuals are over a configurable threshold, repeat with an additional hypothesized pulse. If not, stop.

This procedure is capable of finding an arbitrary number of pulses on a given island.

In addition to deciding how many pulses to hypothesize, the reconstruction software must determine which templates to use; each SiPM has a separate template for positron and laser pulses. During typical muon fills, there are no laser pulses and so all pulses should be fit with positron templates. Similarly, laser shots are usually fired before or between muon fills, and in these cases there exists no ambiguity regarding which template to use. During the fills used for gain measurements, however, there are both laser pulses and positron pulses. Laser shots have the signature feature that they create pulses in all 54 channels of a calorimeter simultaneously. Positron pulses are localized to only a few channels. Islands with a very large number of over threshold crystals reliably contain laser pulses, but they may also contain positron pulses. To correctly fit multiple pulses on islands containing laser pulses, both template types must be employed in a single fit. In-fill laser pulses are separated by  $200\ \mu\text{s}$ , so there will never be more than one laser pulse on a given island. Multi-pulse fits on islands

containing laser pulses are handled by trying every possible combination of fits where one pulse takes the laser template and the others the beam template. Of these trials, the one with the smallest resulting  $\chi^2$  is recorded as the correct interpretation.

The fitting described to now treats each island independently. Such an approach is not ideal because each SiPM pulse causes a subsequent baseline perturbation that persists for hundreds of nanoseconds. This perturbation is clear in Figure 4.10 and should be accounted for when fitting later pulses. After all islands in a muon fill have been processed by the template fitting procedure, pulses that have been found to be within the 250 ns template length of one another are fit again simultaneously using the independent island fit results as starting parameter guesses. These multi-island simultaneous fits are called chain fits. As the baseline perturbations following SiPM pulses are included in the templates, the chain fits intrinsically account for them. See Figure 4.13 for four example template fits, including a chain fit.

Time-in-fill dependent shifts in pulse time reconstruction directly bias the  $\omega_a$  measurement. The template fitter—being the process that first extracts times from digitized waveforms—could be responsible for generating such an effect. In each fill, the probability that an island trace will contain multiple pulses decreases with time. If the time extracted from a given pulse is different when taken from a multi-pulse fit compared to a single-pulse fit, a time-in-fill dependent time reconstruction shift would result. A GEANT4 simulation study was conducted to test for this effect. Single positrons and pairs of positrons with varying time and spatial separations were fired into a calorimeter and SiPM pulses were simulated based on the energy deposition in each crystal. The pulses were reconstructed using the template fitting procedure described here. Time offsets between the positron impact times and the resulting pulse peak times were calibrated using the single pulse events. Double pulse fit times were then analyzed using this calibration. As shown in Figure 4.14, no timing shifts were detected.

To summarize, the template fitter takes digitized islands and produces uncalibrated pulse energies and times. It is responsible for resolving temporal pileup as much as possible and

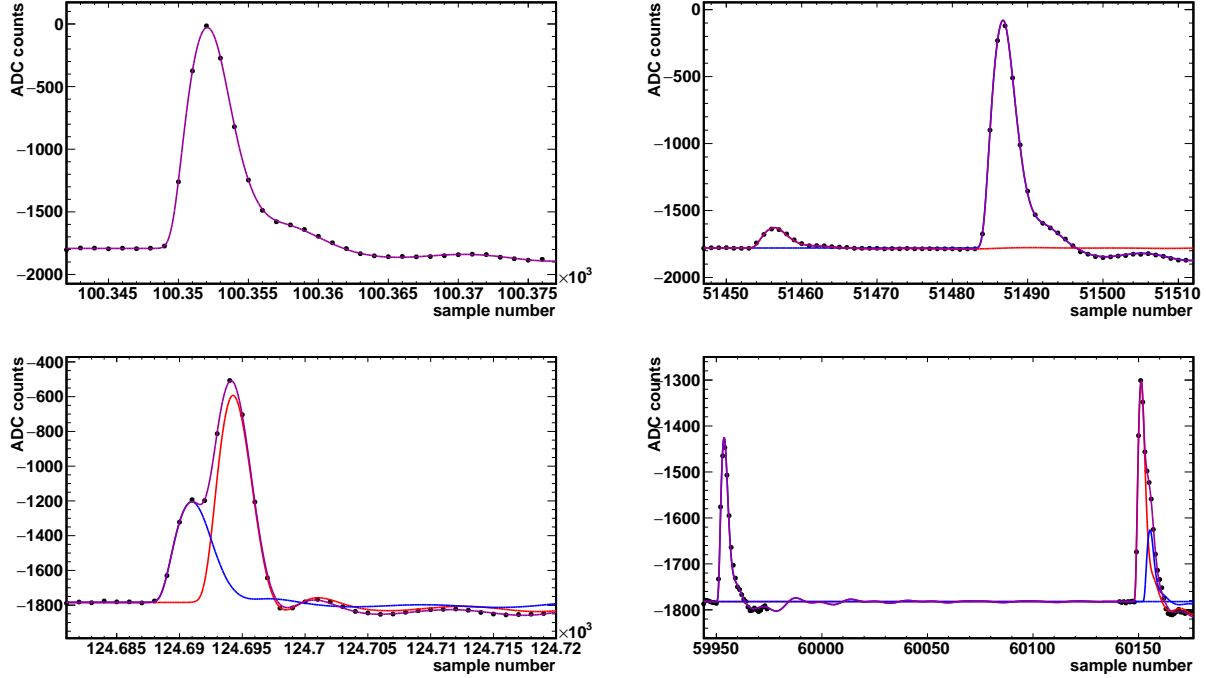


Figure 4.13: Template fits of four differing varieties, selected from E989 Run 1 data. In all graphs, the black points are digitizer sample values and the smooth curves are fit results. In the top left, a single pulse was identified. In the top right, two pulses were identified: a small one followed by a large one. The bottom left depicts a multi pulse fit containing both a laser pulse and positron pulse. The two components are shown in red and blue, and the difference in pulse shape between the two is evident. The wider first pulse is a positron pulse and the second narrower pulse is from the laser. The bottom right shows a chain fit containing two separate islands. The baseline perturbations from the first pulse persist into the second island, where two pulses separated by 5 ns were identified.

for providing accurate, unbiased time and energy measurements. Template fitting is not the only way to accomplish this task, but as of this writing no other approach has emerged that can match its performance or reliability.

#### 4.2.5 Energy Calibration

After pulses have been identified and their pulse-integrals recorded, the next reconstruction step is to convert these pulse-integrals into meaningful energy units. To do so, a calibration

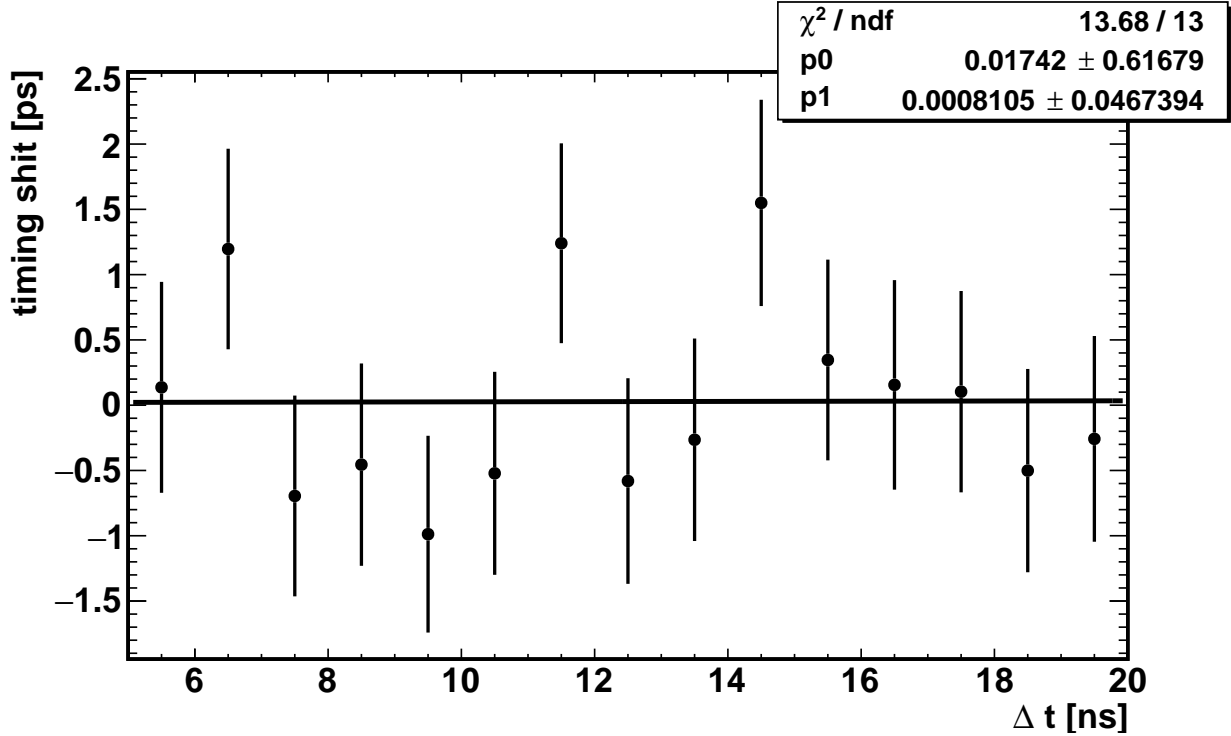


Figure 4.14: Timing shift, defined as reconstructed time minus calibrated hit time, versus pulse pair time separation. This plot was generated from a simulation described in the text.  $p_0$  and  $p_1$  refer to zeroth and first order terms in a polynomial fit to the data.

constant in units of MeV/pulse-integral is required for each calorimeter channel. A technique for obtaining these constants using discrete photon statistics has already been described. As photostatistics calibrations do not account for differing optical properties between crystals, other techniques may provide more accurate results. The energy calibration constants used in the reconstruction approach described here were determined by observing the energy deposition signature of lost muons in each crystal.

As mentioned in Section 3.3.3, some number of muons are inevitably lost from the storage ring before they decay, especially at early times in a muon fill. As these muons travel out of the storage region, they are likely to strike a calorimeter. Muons are 200 times more massive than positrons and generally travel straight through a  $\text{PbF}_2$  crystal without stopping. These lost muon impacts leave a precise energy signature [59]. Furthermore, as a lost muon does



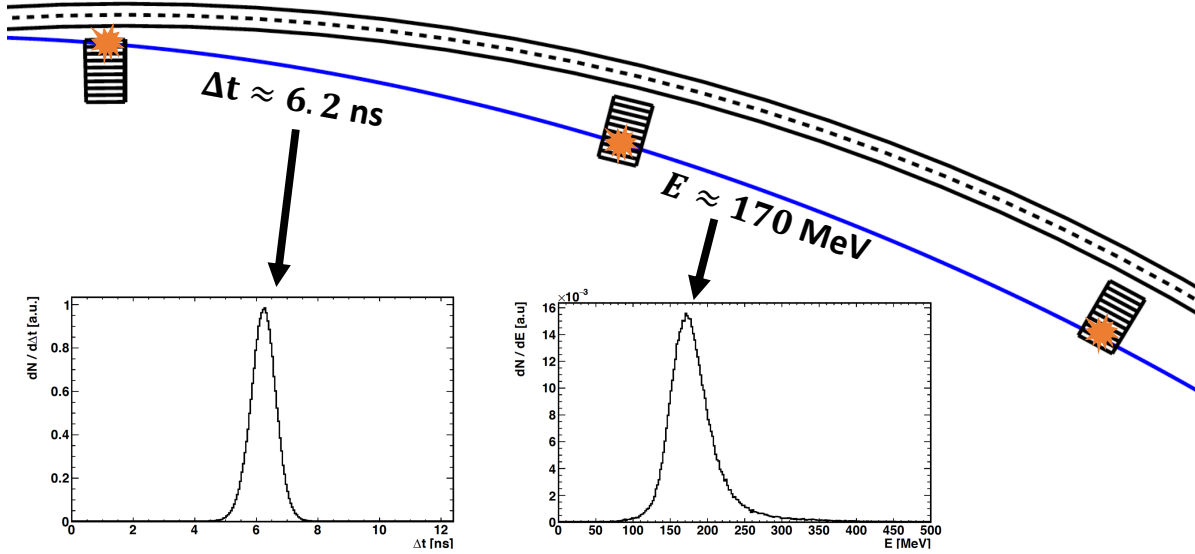


Figure 4.15: A possible lost muon trajectory. The solid black curves delineate the boundaries of the muon storage region. The dashed curve depicts the equilibrium orbit for on-momentum muons. The blue curve depicts the path of a muon lost from the storage region. This muon passes through three calorimeters, leaving a distinct energy signature that is used for energy calibration and a distinct calorimeter-to-calorimeter time difference that is used for time synchronization.

not stop after encountering a calorimeter, it often continues to hit a second and a third (see Figure 4.15). The time separation between calorimeter hits in such a case is always very close to 6.2 ns, calculable as the distance between calorimeters divided by  $c$ . Lost muon calorimeter pulses are easily identified by looking for coincidence events in which two or three consecutive calorimeters see pulses separated by 6.2 ns. As lost muons typically do not induce electromagnetic showers within the  $\text{PbF}_2$  crystals, discriminating power is further increased by selecting for islands containing only a single SiPM pulse.

Selecting isolated, single crystal pulses coincident between two or three calorimeters provides a signal usable for energy calibration [79]. The peak of the pulse-integral distribution from each SiPM should correspond to the same amount of deposited energy, as it is assumed that these selection criteria leave nothing remaining except for lost muon signals. It is assumed that a lost muon will deposit the same amount of energy regardless of which crystal

it traverses. Differences in SiPM gains and in  $\text{PbF}_2$  optical properties will result in differing pulse-integral distributions between channels. The correct gain constants, accounting for both SiPM and  $\text{PbF}_2$  differences, are the ones that align the peaks of all the distributions. Additionally, a common factor is needed to correct for differences between the  $\text{PbF}_2$  response to lost muons and to positron showers. Multiple approaches exist for finding this common factor. Examples include calculating it directly from simulations, matching the end point of the final reconstructed calorimeter energy distributions to the expected 3.1 GeV, and placing the optimal T-Method threshold at a chosen value. For the analysis described in this document, the common factor was chosen to place the optimal T-Method threshold at 1.7 GeV. Figure 4.16 shows the lost muon energy peak before and after the lost muon energy calibration.

While the earliest E989 calorimeter calibrations were based on photostatistics, lost muon calibrations have since become the default. To maintain consistent energy measurements over time scales of hours to days, the gain corrections provided by the laser system are required. In short, lost muon calibrations provide the baseline equalization factors between calorimeter channels, and the laser provided gain correction factors control for temperature dependent drifts from these baseline values.

#### 4.2.6 *Timing Corrections*

The  $\omega_a$  measurement is fundamentally time-based. Correct time reconstruction is critical. The ideal time measurement would be that of a muon decay relative to the beam injection. As positron energy dependent phases do not cause systematic errors, positron hit time measurements are sufficient proxies for muon decay times. What the template fitter produces, however, is a SiPM pulse's peak time relative to the first sample of a digitizer trace. Digitizer traces begin  $32\ \mu\text{s}$  before the beam injection. There is no guarantee that each of the 1296 digitizer traces begins at the exact same time, nor is there any guarantee that there is an unchanging offset between a given digitizer's start time and the beam injection time. Moreover, in the presence of differing SiPM pulse shapes, one cannot assume that the offset

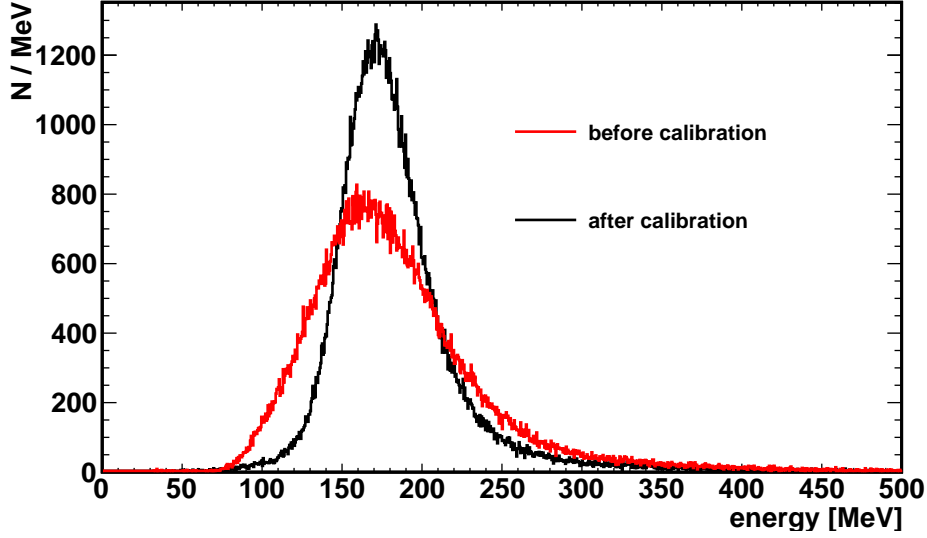


Figure 4.16: The lost muon energy peak before and after energy calibration and long-term gain correction. The above histograms were constructed from a small subset of the E989 Run 1 data and include all calorimeters. Lost muon events were selected from triple calorimeter coincidences. Lost muons leave a distinct energy deposition of approximately 170 MeV when they pass through a calorimeter. In E989 Run 1, the SiPM amplifier gain settings were chosen to equalize channel energy responses based on absolute photostatistics calibrations conducted with the laser system. The lost muon calibration and long-term gain corrections further improved channel equalization, as evidenced by the decreased width of the black curve relative to the red.

between a SiPM pulse’s peak time and the responsible positron’s hit time is the same from one channel to another. Therefore, per-fill and per-channel corrections are needed to align all timing measurements to the same reference.

The time correction procedure amounts to offsetting each pulse time by the combination of two factors: a fill-dependent factor accounting for the time difference between the beam injection and the trace start, and a fill-independent factor accounting for the pulse shape. All factors are channel-dependent. The total correction is

$$t = t' - t_{bd} - t_c, \quad (4.13)$$

where  $t$  is the corrected time,  $t'$  is the time reported by the fitter,  $t_{bd}$  is the fill-dependent

time difference between the beam injection and the trace start, and  $t_c$  is the fill-independent pulse shape factor.

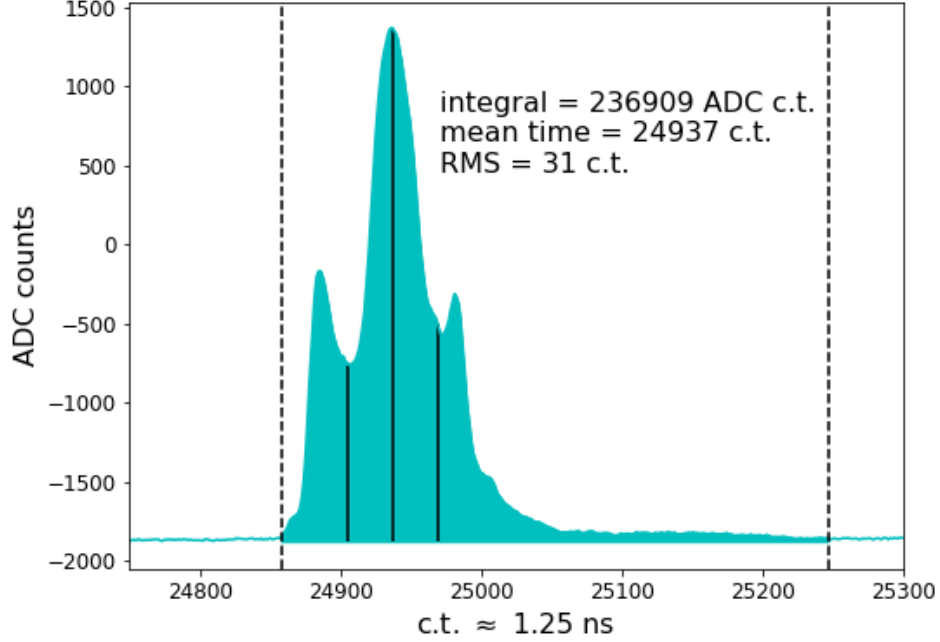


Figure 4.17: Characteristic pulse from the T0 detector, courtesy of Hannah Binney. The black lines denote the pulse’s mean time, or centroid, and the mean time  $\pm 1$  RMS. The mean and RMS of the pulse were calculated between the dashed lines. The reconstructed T0 pulse time is used to synchronize calorimeter hits to the beam injection on a fill-by-fill basis. This pulse is from E989 Run 1.

The sync pulse delivered by the laser system before each muon injection provides the necessary per-fill correction. If there were a known, fixed offset between the sync pulse time and the beam injection time, the sync pulse along with the fixed offset would be all one needed to determine  $t_{bd}$  for each channel. This is not the case, as there is an additional per-fill offset between the sync pulse time and the beam injection time originating both from the non-fixed time difference between the FNAL provided accelerator trigger and the beam injection and from the independence of the clocks used by the calorimeter digitizers and

the laser control board. This offset is measured by feeding an additional laser fiber into a scintillating entrance counter, called the T0 (pronounced T-zero) counter. The T0 pulses are recorded with the same type of digitizer used for the calorimeter SiPMs. An example digitized T0 pulse from E989 Run 1 is shown in Figure 4.17. For each muon fill, the T0 counter provides an independent measurement of both the sync pulse time and the beam injection time; the difference between these measurements is independent of any trace's start time. With a sync pulse measurement in each calorimeter channel,  $t_{sd}$ , and a sync pulse to beam injection time measured by the T0 counter,  $t_{bs}$ , the quantity  $t_{bd}$  is determined as

$$t_{bd} = t_{sd} + t_{bs}. \quad (4.14)$$

The above reasoning assumes that the sync pulse time relative to the beam injection is the same in all channels. This assumption is violated by variations in laser fiber lengths and slight differences in firing time among the six laser heads. These effects are not fill dependent, but are physical differences between channels that remain unchanging in time. Therefore, they can be included with the channel-dependent time corrections,  $t_c$ , which then must account for the combined effects of the pulse shape differences between channels, the fiber length differences between channels, and the firing time differences between laser heads.

The channel-dependent time correction factors  $t_c$  can be decomposed into per-crystal and per-calorimeter constants. The per-crystal offsets, originating from differences in fiber length and pulse shape, are determined from islands with large pulses in multiple neighboring crystals. Following the sync pulse alignment, the distribution of pulse-time differences between each pair of neighboring crystals is measured. A subset of these distributions is shown in Figure 4.18. A least squares problem is then solved to find the constants that bring the average time differences between neighbors as close to 0 as possible. This procedure is successful at bringing most of these mean time differences to within 5 ps of 0 [79].

The per-calorimeter constants accounting for differences between the six laser heads are found from lost muon coincidence events, such as the one illustrated in Figure 4.15. Again, a least squares problem provides the constants needed to equalize the time differences between

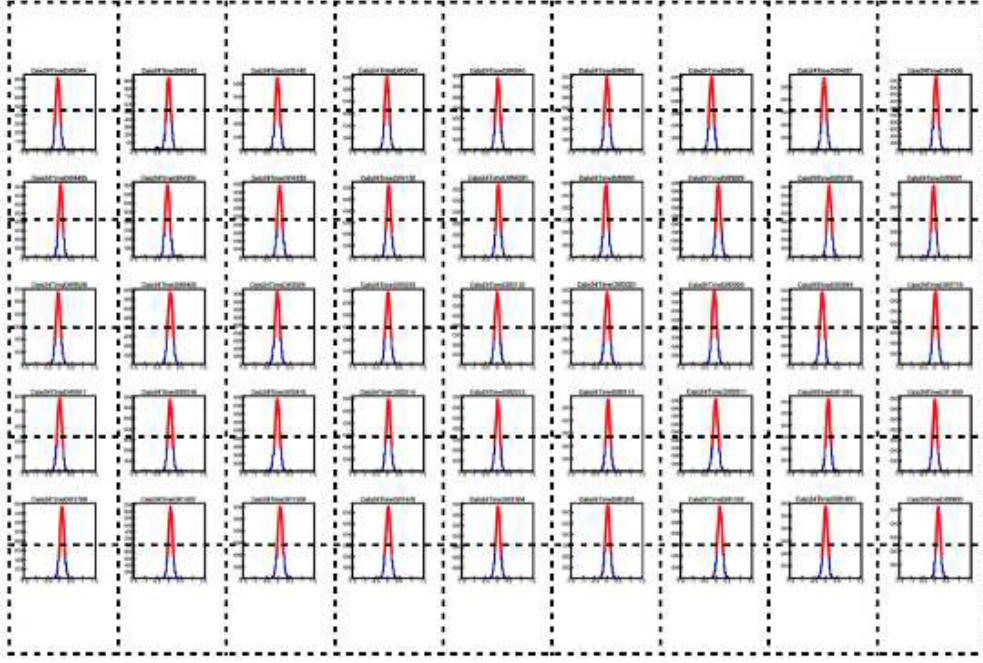


Figure 4.18: Hit time differences between neighboring calorimeter channels measured from positron showers, courtesy of Kim Siang Khaw. The channel-dependent time correction factors are determined as the values that bring the means of these distributions to 0.

each pair of calorimeters [79]. Finally, the fill-independent, channel-dependent constants  $t_c$  are the sums of the per-crystal offsets found from pulse time differences between neighboring crystals and the per-calorimeter offsets found from lost muon coincidences. The final time correction applied to each pulse is

$$t = t' - t_{sd} - t_{bs} - t_c. \quad (4.15)$$

Not all effects have been accounted for here. For example, the laser fiber that feeds the T0 counter is longer than the one that feeds the calorimeters. This time difference does not change on a per-fill basis, and—after applying the per-channel and per-calorimeter corrections—it is the same for all channels. Global, unchanging time offsets that apply to all 1296 channels in the same way are not a concern as they do nothing more than shift the measurement's overall time origin. The corrections described here provide the needed per-fill

and per-channel timing alignments.

#### 4.2.7 Clustering

Following the energy and time corrections, template fit results from all channels are expressed in common, physical units. The term crystal hit refers to fully calibrated template fit results. Combining crystal hits into best guesses for decay positron impact parameters is the final reconstruction step. Reinterpreted decay positron objects are called clusters, and the process that produces them is called clustering.

Clustering is straightforward in most cases. The island shown in Figure 4.9 contains a single positron impact. A reasonable choice for the interpreted decay positron energy is the sum of the crystal hit energies found on that island. Regarding the decay positron hit time, one might take the energy weighted time of all crystal hits in the island or the time of the largest crystal hit. The two approaches produce very similar results—within 100 ps generally—because of both the channel timing alignments and the tendency of positron showers to be mostly contained within a single crystal. Although it provides slightly worse precision, the time of the crystal hit with the maximum energy was chosen for the analysis described here. The reason is that using the energy weighted time couples calibration errors to timing reconstruction errors, whereas using the time of the largest contributing hit does not.

Although not strictly necessary for the  $\omega_a$  measurement, determining decay positron hit positions is useful for understanding the recorded data and checking for unexpected systematic effects. A logarithmic weighting approach provides the best compromise found so far between optimizing precision and minimizing the extracted position bias. The reconstructed horizontal position is determined using the equation [80]

$$x_{reco} = \frac{\sum_i w_i \cdot x_i}{\sum_i w_i}, \quad (4.16)$$

$$w_i = \max \left( 0, W_0 + \log \frac{E_i}{\sum_j E_j} \right). \quad (4.17)$$

The sums are taken over crystal hits, the  $x_i$  are the central positions of the crystals,  $E_i$  is the crystal hit energy for crystal  $i$ , and  $W_0$  is a free parameter found from simulations and test beam studies. Larger values of  $W_0$  place higher importance on shower tails in determining the reconstructed position. The optimal  $W_0$  value for the E989 calorimeter is approximately 4, although the optimal value varies with positron energy [80]. Substituting  $y$  for  $x$  in the above equation allows for reconstruction of the vertical position. For a 2 GeV positron impact, the logarithmic weighting approach provides 2 mm position resolution in both the horizontal and vertical directions.

When multiple positron hits occur in the same island, clustering becomes more complicated. Crystal hits must be correctly assigned to their respective clusters. In most cases, considering only the crystal hit times suffices. Crystal hits originating from the same positron impact are expected to be reconstructed very close in time, much closer than the typical island length of 40 ns. In the first step of the clustering algorithm, crystal hits are time sorted. A crystal hit separated by more than a configurable number of nanoseconds, typically 2.5 ns, from the previous is considered to be part of a separate positron shower. This sorting process, called time partitioning, groups crystal hits by time proximity. Subsequent processing treats each of these time partitioned groups independently. Figure 4.19 shows an island with three identified clusters that were separated through time partitioning.

It is possible that two or more positrons will hit the same calorimeter near enough in time to one another that they are unresolved by the time partitioning. If these close-in-time positrons hit the same crystal, the template fitter may report a single pulse. There is nothing the clustering algorithm can do in that case. However, if the positron impacts are spatially separated, one can leverage the calorimeter segmentation to resolve them. Spatial separation occurs after time partitioning; its purpose is to use spatial information to resolve any pileup events remaining after the template fitting and time partitioning.

Spatial separation algorithm development is an ongoing project within the E989 collaboration. As pileup is projected to be the single largest source of systematic uncertainty in the E989 experiment [37], there is great value in applying varying pileup resolution techniques



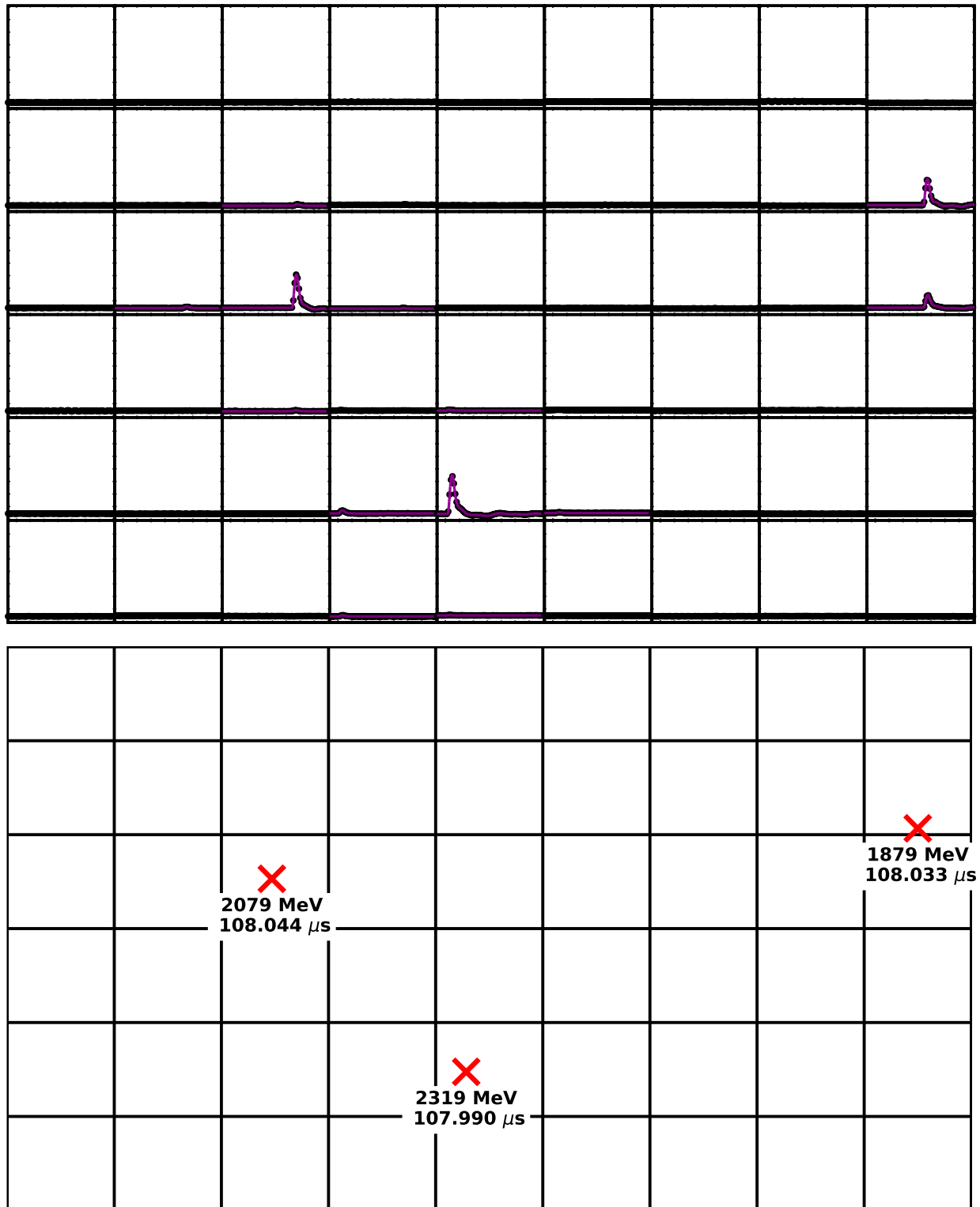


Figure 4.19: An island with three identified clusters. On top are the original SiPM pulses with their template fits, and on the bottom are the reconstructed cluster positions, energies, and times.

and seeking agreement between them. One approach is to not use spatial separation at all. As discussed in 3.5.2, pileup resolution in the reconstruction software is likely not sufficient to reach the experiment's final systematic uncertainty target. Analyzers must apply pileup subtraction procedures to correct all hit time histograms for residual pileup contamination. Not using spatial separation has the advantage of simplicity: one need not model correlations between positron energy and hit position when constructing a pileup correction. The disadvantage is that the size of the required correction will be larger. Different analyzers have made different decisions regarding this issue.

A simple spatial separation algorithm that performs reasonably well goes as follows:

1. Locate the crystal hit with the maximum energy. Label it a cluster center and record its energy,  $E_{max}$ .
  - (a) Looking only at crystals not within a  $3 \times 3$  grid surrounding a cluster center, find the crystal hit with the maximum energy. If this crystal hit's energy is greater than  $r_{cut} \cdot E_{max}$ , label it a cluster center.
  - (b) Repeat if a new cluster center was found.
2. Assign each cluster an energy equal to the sum of the crystal hit energies in the surrounding  $3 \times 3$  grid.
3. Scale these energies, maintaining their relative sizes, so that the sum of all cluster energies is equal to the sum of all crystal hit energies input to this procedure.
4. Assign each cluster a time equal to the time of its cluster center.

This algorithm searches for large energy depositions far from the crystal with the largest pulse. The  $r_{cut}$  algorithm parameter determines how large of an energy deposition far from the maximum will be considered a separate cluster. Simulation studies suggest that 12% is a reasonable value for  $r_{cut}$  [81]. This technique works quite well above the 1.8 GeV T-Method energy threshold. According to the aforementioned simulation study, it resolves

75-80% of pileup events remaining in the T-Method energy range after template fitting and time partitioning. Unfortunately, the same simulation revealed that false positives become a problem at lower energies. Low-energy positrons are more likely to enter calorimeters at oblique angles and spread their energy across many columns. This algorithm would have to be adjusted for use in an energy binned or asymmetry weighted analysis including lower energy positrons. Nevertheless, it is expected that spatial separation will reduce the number of pileup events by 70-80%.

The reconstruction process described here is sequential and modular. It moves unidirectionally from template fitting to calibrations and corrections to clustering. The output from each step is recorded so that intermediate steps can be reconfigured and run again without the need to repeat the entire reconstruction process. Given that template fitting is the most computationally intensive calorimeter reconstruction task, it is convenient to be able to apply improved calibration constants without having to rerun the template fitting. Additionally, the reconstruction path can fork. Current E989 production jobs have two independently developed clustering modules that run on the same crystal hits. The respective output clusters are consumed by different analysis groups.

### **4.3 Beam Tests**

Prototype calorimeters were characterized through four test beam experiments: three at SLAC's End Station Test Beam Facility (ESTB) and one at INFN Frascati's Beam Test Facility (BTF). These tests guided calorimeter design decisions, provided key measurements of the calorimeter response not easily obtainable in situ, and built confidence in the calorimeter's overall performance. Each prototype of incrementally increasing complexity served to test or measure a different aspect of the calorimeter. The final test beam experiment featured a production calorimeter that was then transported to FNAL and installed. The author of this dissertation played a leading role in all four test beam experiments. Their setups and key findings will now be described.

#### 4.3.1 November 2013 SLAC Test Beam Experiment

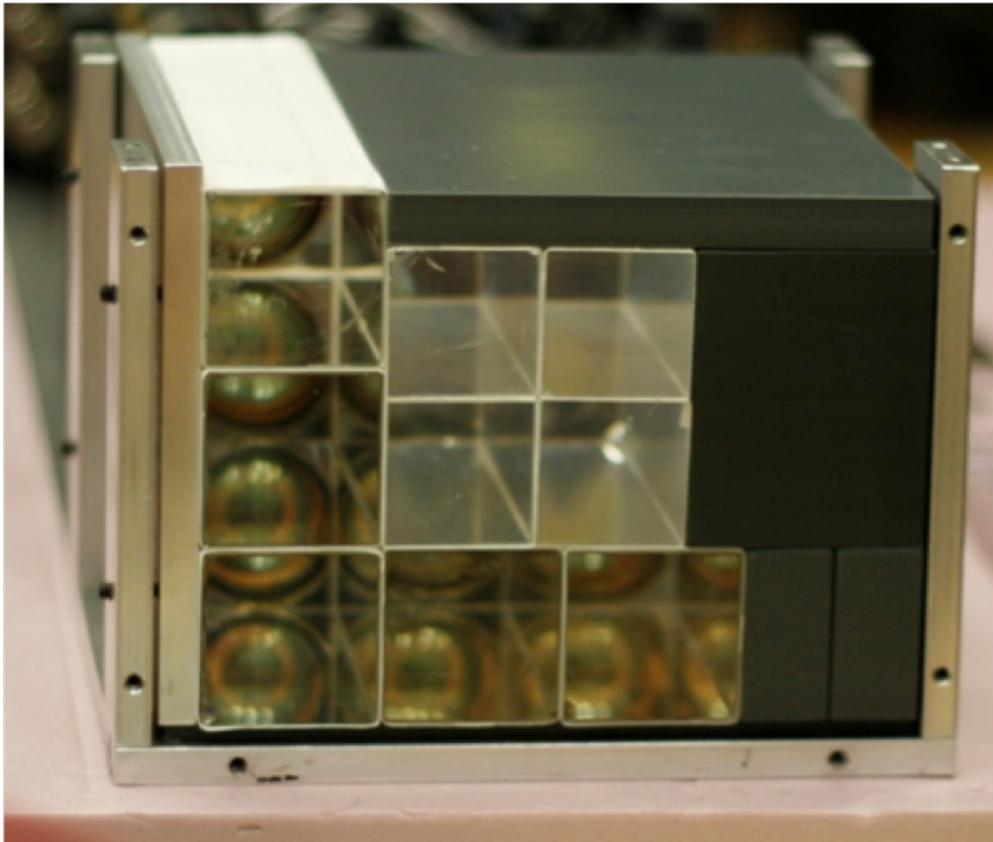


Figure 4.20: Prototype calorimeter from the November 2013 SLAC ESTB test beam experiment. The smaller crystals were coupled to SiPMs and the larger crystals to PMTs.

The first SLAC test beam experiment occurred in November 2013. The prototype calorimeter consisted of four  $2.5\text{ cm} \times 2.5\text{ cm} \times 14\text{ cm}$   $\text{PbF}_2$  crystals, the same dimensions used in the production calorimeters, and five larger  $3\text{ cm} \times 3\text{ cm} \times 14\text{ cm}$   $\text{PbF}_2$  crystals. The smaller crystals were coupled to SiPMs with early versions of the custom built electronic readout board described previously. The larger crystals were coupled to commercial photomultiplier tubes. The SiPMs were optically coupled to their respective crystals with NuSil LS-5257 optical grease. Bicon BC-630 optical grease was used for the PMTs. This early prototype is shown in Figure 4.20.

ESTB provides a narrow electron beam with tunable energy and rate. The objective of this first test beam experiment was to verify that  $\text{PbF}_2$  exhibits a light yield and energy resolution sufficient for the needs of the E989  $\omega_a$  measurement. The ESTB beamline was configured to deliver on each pulse a Poisson distribution of electrons with a mean value of 1. The beam's transverse size was between 1 mm and 2 mm in  $x$  and  $y$ , with a slight energy dependence. Additionally, the facility provided a remotely movable table that enabled the adjustment of the position of the detector under test in both the horizontal and vertical directions. This facilitated tests of the position dependence of the prototype detector's response to the electron beam. The beam pulse rate varied between 5 Hz and 10 Hz. In the November 2013 test, the accelerator provided electron energies ranging from 2.5 GeV to 4 GeV. This energy range is similar to that used in the  $\omega_a$  analysis, although it extends beyond the 3.1 GeV endpoint. Single electron events were selected for analysis. The details of the measurements and their results will not be described here as they were repeated more carefully and reliably at the following test beam experiment. Nevertheless, the November 2013 test beam was a significant landmark in the E989 calorimeter's development history as it provided the first confirmation that the  $\text{PbF}_2$  calorimeter would meet the requirements of the  $\omega_a$  measurement [37].

#### *4.3.2 July 2014 SLAC Test Beam Experiment*

The second SLAC calorimeter test occurred over two weeks in July 2014. This section is adapted from the resulting publication [69]. The tested prototype comprised 28 production geometry  $2.5\text{ cm} \times 2.5\text{ cm} \times 14\text{ cm}$   $\text{PbF}_2$  crystals arranged in a four high by seven wide array, see Figure 4.21. Every crystal was coupled to a SiPM with the same dimensions and pixel count as those used in the production experiment. An early version of the SiPM electronic readout board was used. Objectives of this test included measuring the energy dependence of the calorimeter energy resolution, the impact position dependence of the calorimeter response, and the effect of different crystal wrappings. It was expected that wrapping crystals in a white, reflective paper would improve the light yield but increase the pulse width relative to

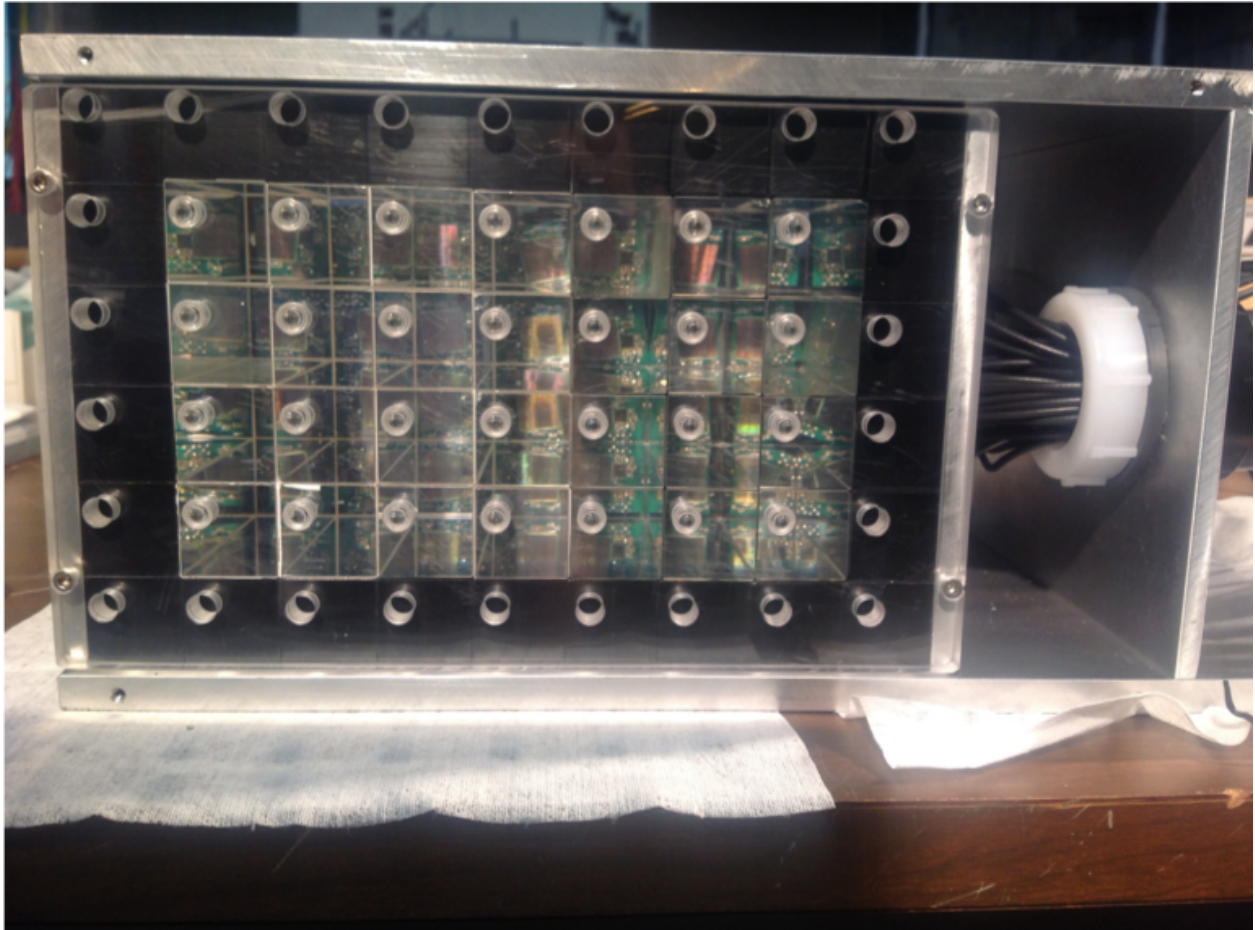


Figure 4.21: Prototype calorimeter from the July 2014 SLAC ESTB test beam experiment. All crystals were coupled to SiPMs. The crystals in the  $4 \times 4$  group on the left were wrapped in reflective white Millipore paper, whereas the remaining crystals in the  $4 \times 3$  group on the right were wrapped in absorptive black Tedlar.

a dark, absorptive wrapping. To enable a measurement of the wrapping effect, each crystal in the first four consecutive columns was wrapped in a single, non-overlapping layer of reflective white Millipore® paper, whereas each crystal in the remaining three columns was wrapped in matte black absorbing Tedlar™. The Millipore Immobilon-P is a polyvinylidene fluoride membrane with  $0.45\text{ }\mu\text{m}$  pores, and it is a Lambertian (diffusive) mirror. The upstream face for all crystals was left unwrapped to permit the injection of light from a prototype laser calibration system. Millipore wrapped crystals will be referred to as white-wrapped and Tedlar wrapped crystals as black-wrapped.

SiPM pulses were recorded with high-speed waveform digitizers. For most studies, the SiPMs coupled to the 16 white-wrapped crystals were digitized using a 12-bit CAEN DT5742 switched capacitor desktop digitizer sampling at 1 GS/s, while eight of the black-wrapped crystals were digitized using a 12-bit Struck SIS3350 digitizer sampling at 0.5 GS/s. These digitization rates are both similar to the 800 MS/s digitization rate used by the custom E989 digitizers and enabled the application of the template building and template fitting procedures described in Section 4.2.

A key component of the July 2014 test was the prototype laser calibration system, driven by a PicoQuant LDH-P-C-405M pulsed diode laser. Laser light passed through a diffuser system—both an integrating sphere and an engineered diffuser were tested—and into an optical fiber bundle consisting of 30 fibers. A remote controllable neutral density filter wheel was placed between the laser beam and the diffuser. The emitting ends of the optical fibers were coupled via SMA connectors to a port in front of each crystal, allowing for calibration and gain tracking of each individual calorimeter channel. See Figure 4.22 for an illustration of this setup and a crystal numbering scheme that will facilitate discussion of the subsequent tests.

Each SiPM was calibrated using the photostatistics technique described in Section 4.1.3. After pulse-integrals were converted to numbers of pixels fired (pe), a correction was applied for pixel saturation effects following Equation 4.2. This correction improved the calorimeter linearity. Light yield and energy resolution were then measured for both the white-wrapped

1	2	3	4	17	18	19
5	6	7	8	20	21	22
9	10	11	12	23	24	25
13	14	15	16	26	27	28

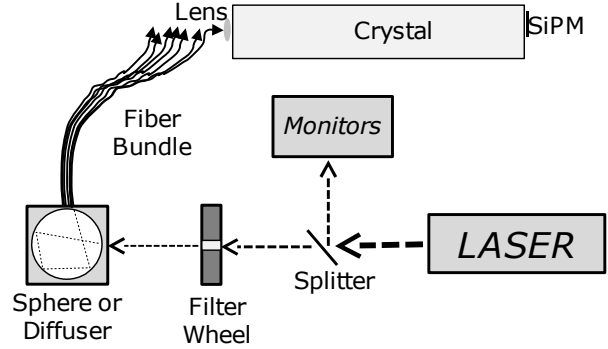


Figure 4.22: Diagrammatic representation of the July 2014 prototype calorimeter. Crystals 1-16 were wrapped in white Millipore paper. Crystals 17-28 were wrapped in black Tedlar.

and the black-wrapped crystals. The measurement was conducted with beam energies ranging from 2 to 4.5 GeV. With the numbering defined in Figure 4.22, characteristics of the white and black-wrapped crystals were measured by aiming the electron beam into crystals 11 and 24, respectively.

Offline analysis after the fact revealed that one quarter of the pixels in SiPM 24—the central SiPM in the black array—were not functioning and therefore were not being included in the final stage of amplification. This amounted to an effective decrease in light yield to 75% of what it would have been had the SiPM been operating correctly. All pe values calculated for SiPM 24 were increased by a factor of 4/3 to correct for this effect.

For every beam event, the pulses in each of the nine crystals within a cluster were converted into pe and summed together. As shown in Figure 4.23, the measured pe distribution was not symmetric; it displayed a low-energy tail from incomplete longitudinal shower containment. The effective mean values were extracted from Gaussian fits over an asymmetric fit region.

A linear relationship between pe extracted from the fit and the beam energy in the range from 3 to 4.5 GeV was found. The resulting slope was  $(1.45 \pm 0.05)$  pe/MeV with an offset of  $(-80 \pm 200)$  pe for the white array and  $(0.76 \pm 0.04)$  pe/MeV with an offset of  $(-150 \pm 160)$  pe for the black array. These values were consistent with expectation and previous measurements.



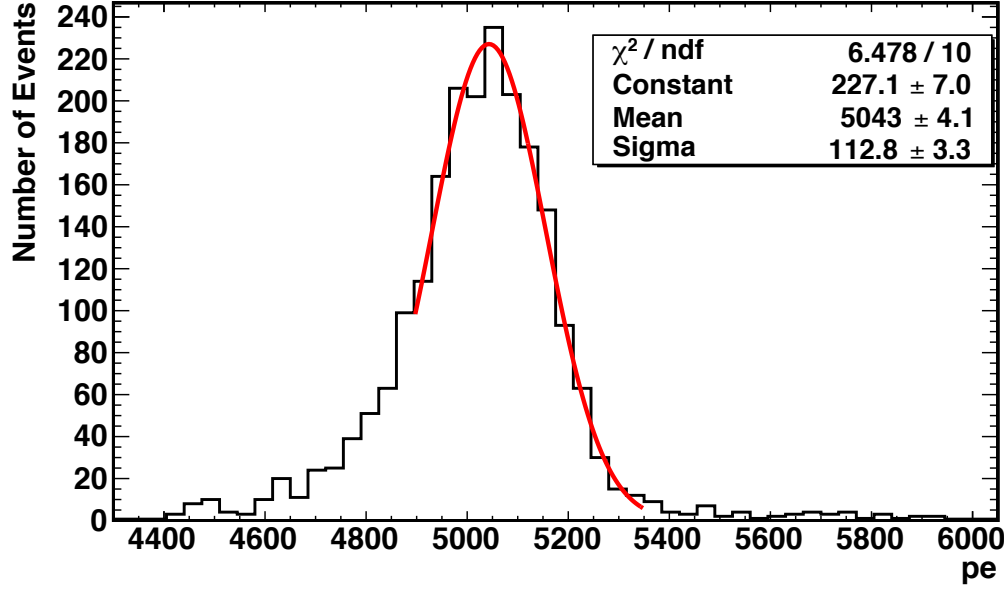


Figure 4.23: The total number of fired pixels measured among nine white-wrapped crystals arranged in a  $3 \times 3$  grid centered on crystal 11 when a 3.5 GeV electron impacts the center.

Energy resolution, defined as  $\sigma_E/\sqrt{E}$ , is expected to improve at higher energies as the contribution from discrete photon statistics decrease with  $1/\sqrt{E}$ . Fluctuations in the electromagnetic shower containment within a finite sized detector also contribute to energy resolution. According to GEANT4 calculations, containment fluctuations introduce an additional 1.5% contribution to the prototype calorimeter's energy resolution. This contribution does not change significantly over the range of beam energies provided by the SLAC accelerator. Assuming discrete photon statistics and containment fluctuations are the dominant contributions to the energy resolution, energy resolution should change with energy according to

$$\frac{\sigma_E}{E} = \sqrt{(1.5\%)^2 + \frac{a^2}{E/\text{GeV}}} . \quad (4.18)$$

Energy resolution was measured at each beam energy. The resulting energy resolution versus energy curves were fit with the above function. The fits yielded  $a = (3.4 \pm 0.1)\%$  with a reduced  $\chi^2$  of 10.6/8 for the white array and  $a = (5.0 \pm 0.3)\%$  with a reduced  $\chi^2$  of 3.5/4 for

the black array, see Figure 4.24. The data points were extracted using the aforementioned asymmetric Gaussian fits to the pe distributions measured at each beam energy.

The obtained  $a$  term values are dominated by the photostatistics contributions predicted from the light yield values. Any remaining contributions to the energy width can be assigned to uncertainties in the calibration of various SiPMs together with position fluctuations of the beam. In principle, the momentum width of the beam can contribute to the constant term, which was held fixed in the fit. SLAC asserted that the beam's momentum width was less than 1%.

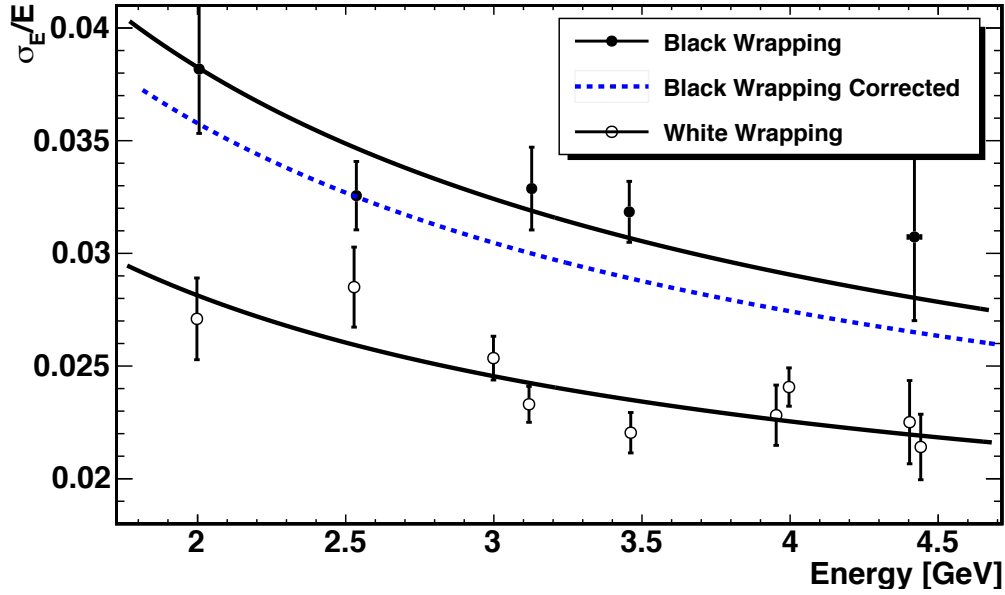


Figure 4.24: Energy resolutions of  $3 \times 3$  arrays of  $\text{PbF}_2$  crystals with black and white wrappings as a function of energy. Fit functions are of the form described in the text. The blue dashed line is the result of correcting the black-wrapped curve for dead SiPM channels discovered after the fact.

The energy resolution measurement for the black-wrapped array suffered because of the dead quadrant in SiPM 24; lower crystal coverage results in an increased variance of the photostatistics term. While one can correct the light yield by increasing the measured pe in SiPM 24 by  $4/3$ , the increased weighting does not account for the increased photostatistics

Table 4.3: Summary of results for  $3 \times 3$  arrays. Note the Tedlar-wrapped array's energy resolution has been corrected following the procedure described in text.

	Millipore	Tedlar
Light yield [pe/MeV]	$1.45 \pm 0.05$	$0.76 \pm 0.04$
Resolution $[\%/\sqrt{E/\text{GeV}}]$	$3.4 \pm 0.1$	$4.6 \pm 0.3$

contribution to the relative variance, which behaves as  $1/N_{\text{pe}}$ , where  $N_{\text{pe}}$  is the mean number of photoelectrons detected by the array.

The correction process to account for the dead quadrant can be separated into two conceptual steps. First, the obtained photostatistics contribution is subtracted quadratically from the energy resolution. Next, the expected photostatistics contribution is added back in. Let  $L_{\text{tot}}$  be the total light yield of the  $3 \times 3$  array in pe/GeV and  $f_c$  the fraction of  $L_{\text{tot}}$  collected in the central crystal with a fully working SiPM. The value for  $f_c$  in the black-wrapped array is 0.85, and  $L_{\text{tot}}$  is 760 pe/GeV. The obtained photostatistics contribution to the relative variance when 3/4 of the central SiPM is in operation is  $[EL_{\text{tot}}(\frac{3}{4}f_c + 1 - f_c)]^{-1}$ . The expected contribution if all channels are operating is  $(EL_{\text{tot}})^{-1}$ . Because the photostatistics contribution to the relative variance is of the form  $E^{-1}$ , the described procedure corrects  $a$  in the energy resolution expression. Therefore, the result  $a_{\text{corrected}} = (4.6 \pm 0.3)\%$  is obtained. Table 4.3 contains a summary of the light yield and energy resolution results.

The last July 2014 test beam finding that will be discussed here pertains to measured pulse shape differences between black-wrapped and white-wrapped crystals. Recall from Section 4.2.3 that the template fitting procedure is predicated on the assumption that SiPM pulse shapes do not change, at least not with a fixed light source. Recall also that pulse shapes do in fact vary slightly with positron or electron impact position. An important test beam finding was that pulse shapes vary far more with impact position for white-wrapped crystals than they do for black-wrapped crystals. The prototype SiPM boards used in the July 2014 test provided much wider pulses than those ultimately achieved with the final

readout board. With these wide pulses, no pulse shape dependence on impact position was detected for black-wrapped crystals. Conversely, for white-wrapped crystals the pulse full width at half maximum increased by 50% as the electron beam moved from the center of one crystal to the center of that crystal's neighbor. This result is shown qualitatively in Figure 4.25 and quantitatively in Table 4.4. As the template fitter's pileup resolving capabilities are diminished by such large pulse shape variations, black wrapping was selected for the production calorimeters despite its lower light yield. Additional studies from the July 2014 test beam experiment, such as those pertaining to impact position and impact angle dependence of the calorimeter response, are discussed in the article by Fienberg et al. [69].

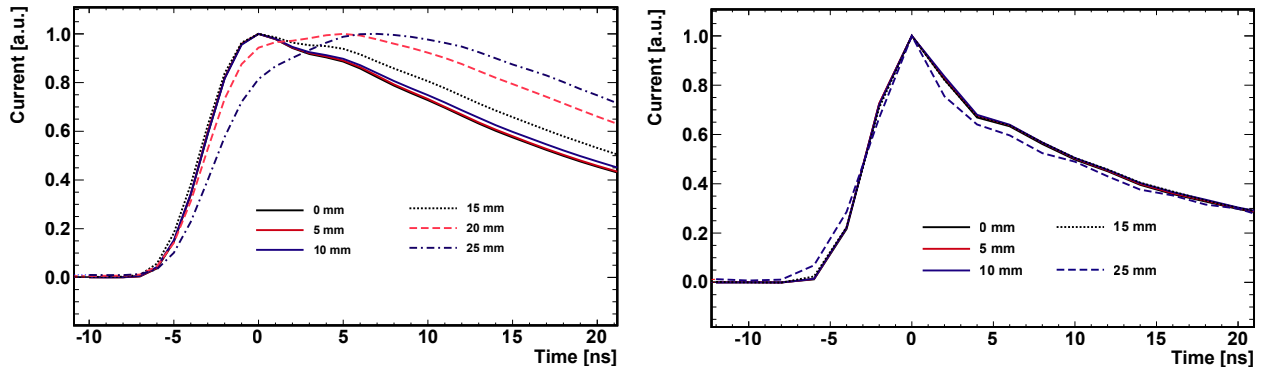


Figure 4.25: Pulse shape dependence on electron impact position for different crystal wrappings. The left is a white-wrapped crystal, and the right is a black-wrapped crystal. There is a large effect in the white-wrapped case. Numbers refer to the horizontal distance between the electron beam impact position and the center of the crystal in question.

#### 4.3.3 March 2016 Frascati Test Beam Experiment

In March 2016, a test beam experiment was conducted at Laboratori Nazionali di Frascati's Beam Test Facility (BTF). Unlike the previous SLAC efforts, which were intended to characterize the  $\text{PbF}_2$  and SiPM responses, this experiment was designed primarily as a test of the laser calibration system. Specific objectives were to test the final version of the complete calibration system chain and to measure the equivalent energy, in units of MeV, delivered

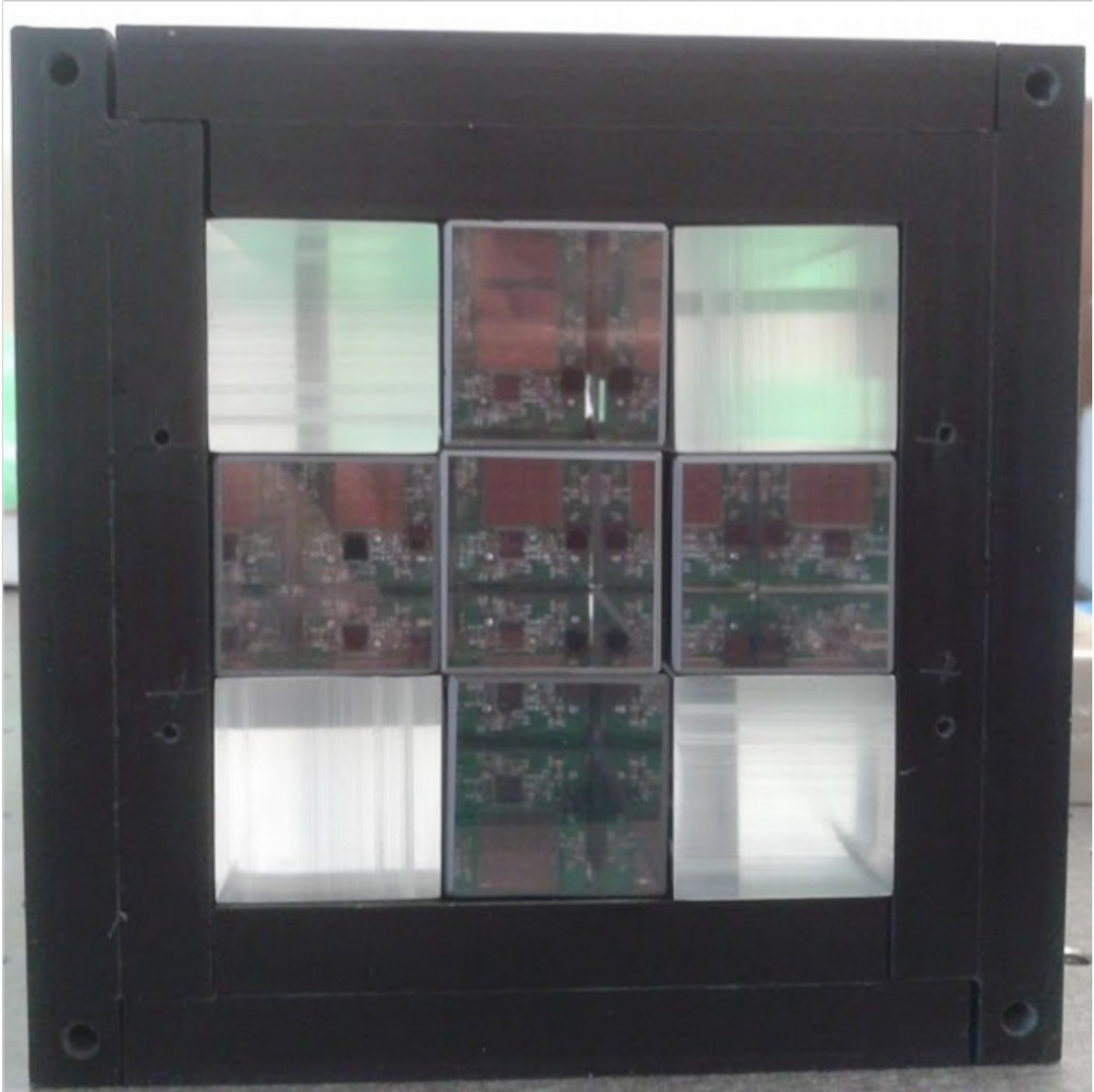


Figure 4.26: Prototype calorimeter tested at Laboratori Nazionali di Frascati's Beam Test Facility (BTF) in March 2016. There are five  $\text{PbF}_2$  crystals; the outer corners of the array are Plexiglass dummies.

Table 4.4: Pulse full width at half maximum evolution during the position scan shown in Figure 4.25. “Distance from Beam” of 0 indicates that the beam is hitting the center of a crystal. The scan is performed horizontally from that point.

Distance from Beam [mm]	Relative FWHM	
	Millipore	Tedlar
0	1	1
5	$1.01 \pm .01$	$1.01 \pm .02$
10	$1.05 \pm .02$	$1.02 \pm .02$
15	$1.17 \pm .01$	$1.02 \pm .03$
20	$1.40 \pm .01$	
25	$1.52 \pm .02$	$0.96 \pm .05$

by the laser system to the calorimeter SiPMs. Specific details regarding the laser system’s components and configuration are available in the technical publication resulting from this test [71]. The general configuration is as described before: a pulsed diode laser beam is split between a set of source monitors and a fiber bundle, which proceeds to deliver light to each SiPM and to a set of local monitors.

The test calorimeter consisted of five production geometry  $\text{PbF}_2$  crystals in a cross orientation, see Figure 4.26. In addition, four Plexiglass dummy crystals completed the  $3 \times 3$  grid. Four of the five crystals, including the center of the cross, were wrapped in black paper. The remaining crystal was wrapped in white. In contrast to the previous tests that have been described, SiPMs were glued to the rear crystal faces rather than coupled with optical grease. Glue provides better long term stability in optical coupling compared to grease and was used in production calorimeters as well. The final version of the SiPM readout boards were available and used in this test.

The BTF electron beam delivers pulse energies between 100 and 500 MeV, significantly lower than the energies provided by SLAC. These energies are common among E989 decay

positrons, so the opportunity to measure the calorimeter response to them was quite valuable. As with the previous tests, SiPM waveforms were digitized and reconstructed with template fits, and absolute calibrations in terms of pe were conducted using the discrete photon statistics technique. A light yield of 0.9 pe/MeV was measured for the black-wrapped crystals using single 450 MeV electrons. This value is about 10% higher than what was measured in the 2014 SLAC test. As different crystals were used in each test, a small difference is not surprising given the varying  $\text{PbF}_2$  optical properties introduced by the manufacturing process. 7.1% energy resolution was achieved, consistent with the curve measured at SLAC extrapolated down to 450 MeV.

Knowing the calorimeter's light yield, one can express the laser pulse energy in terms of equivalent electron or positron MeV. The result obtained at BTF was about 800 MeV, meaning that a laser pulse on average results in the same number of fired pixels as an electromagnetic shower from an 800 MeV electron or positron. The amount of light split to the source monitors relative to the calorimeter was not the same at the test beam as in the production system. Correcting for this known difference, the maximum equivalent energy providable to each calorimeter channel by the production laser system was projected to be approximately 10 GeV. This result was favorable as it confirmed the laser system's capability to deliver to each and every calorimeter channel pulses with equivalent energies spanning the entire range expected from decay positrons in the E989 experiment. This capability was later confirmed in the production system.

#### *4.3.4 June 2016 SLAC Test Beam Experiment*

The final SLAC test beam experiment was conducted in June 2016 and featured a full production calorimeter, see Figure 4.27. Final versions of the laser light distribution, electronic readout, and data acquisition as described in Section 4.1 were all included in the setup. Besides serving as a necessary test of the full E989 calorimeter system, SLAC provided a unique opportunity to practice the difficult task of equalizing and calibrating the 54 calorimeter channels. At SLAC, calibration constants can be obtained by aiming the beam into the



Figure 4.27: The inside of the production calorimeter tested at SLAC in June 2016. There are 54  $\text{PbF}_2$  crystals, each black-wrapped, in a  $6 \times 9$  array. The crystals are all to the right side of the photo; most of the space inside the calorimeter housing is dedicated to electronics.

center of each crystal and measuring its energy response. These constants can then be compared to results from a photostatistics calibration. Additionally, as the longest of all the test beam experiments, the June 2016 SLAC run was ideal for testing the efficacy of long term gain corrections.

Calibrations were conducted in multiple stages. First, calibration constants in units of  $\text{pe/pulse-integral}$  were measured through a photostatistics calibration. The programmable gain amplifiers on each SiPM readout board were then adjusted in an attempt to equalize the responses of all channels. This process was repeated a number of times. After the final iteration, all but one of the 54  $\text{pe/pulse-integral}$  constants were within  $\pm 5\%$  of the mean. A 3 GeV electron beam was then scanned across the face of all 54 crystals. With a perfect calibration, the observed beam energy should be independent of the crystal used to measure it. Some crystals were not usable in this analysis because of issues with the data taking process. Of those that were included, most calibrated responses were within 5% of the mean. A few crystals, however, deviated by more than 10% from the mean, see Figure 4.28. This result represents the best achieved calibration using the photostatistics technique, which is unable to account for differing optical properties and SiPM efficiencies



between channels. The lost muon technique outlined in Section 4.2.5 was later developed to further improve the energy calibration procedure. Many hours elapsed over the course of this scan, and thus the application of long term gain corrections was necessary.

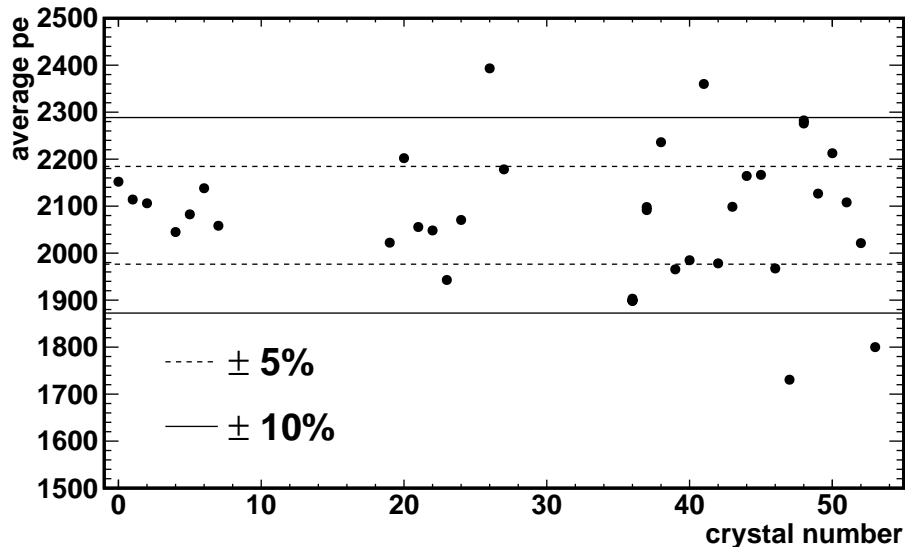


Figure 4.28: Crystal by crystal response to a 3 GeV electron. Most are within 5% of the mean, but a few deviate by more than 10%. Shown here is the number of pe observed within a single channel, not a pe sum over all channels.

Long term gain corrections were derived from laser pulses in the calorimeter SiPMs and source monitors following the procedure described in Section 4.1.3. During stable periods of running, these corrections were in general quite effective. Figure 4.29 depicts the calorimeter response over a continuous 7-hour running period before and after applying long term gain corrections. The electron beam energy was 3 GeV during this measurement. It was typical during the SLAC run for SiPM temperatures to vary by up to 1°C over a period of 7 hours. Mean temperatures varied from day to day, following the weather. Gain drifts on the order of 1% to 10% are expected to result from such temperature changes. Figure 4.29 shows the uncorrected calorimeter response drifting by 3% over the first five hours of the measurement. The overall offset between the uncorrected response and the corrected response indicates that

the temperature averaged across the 7-hour measurement was lower than the temperature at the time of the previous absolute calibration. After gain corrections, the calorimeter response was stable to better than 0.1% per hour.

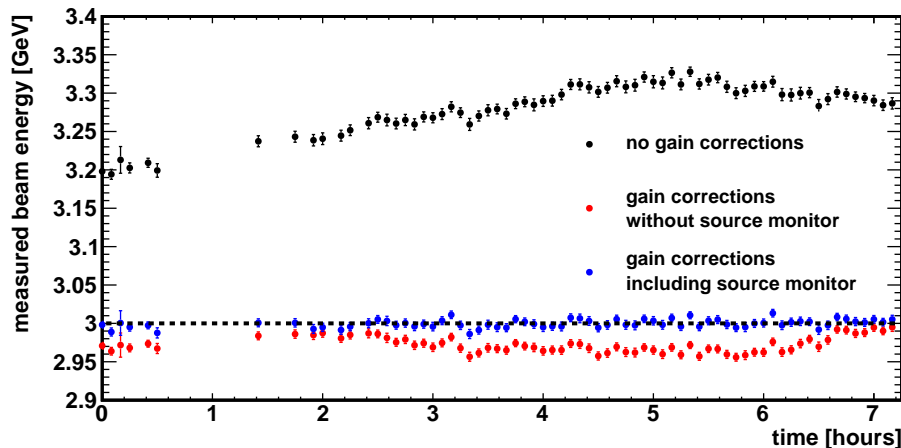


Figure 4.29: The June 2016 test beam calorimeter’s energy response to a 3 GeV electron beam over 7 hours, with and without gain corrections. The source monitor corrections referenced in the legend account for drifts in the energy of the laser pulses used to measure calorimeter channel gains. Including the source monitors improves the result. With all corrections in place, the calorimeter response is stable to better than 0.1% per hour.

The June 2016 SLAC test beam experiment was an invaluable test of the complete calorimeter system. It provided an opportunity to vet near-final versions of the calorimeter hardware, temperature and gain monitoring and control software, reconstruction techniques, the data acquisition system, and the laser calibration system. Most of these tests have not been described here. At the end of the run, the detector team was confident that the calorimeter as designed would meet—and in many cases exceed—the demands of the  $\omega_a$  measurement. The months following the June 2016 test beam were consumed by the construction, quality control, and installation of the 24 production calorimeters.

## Chapter 5

# COMMISSIONING

E989 is a precision experiment; success depends on accumulating a dataset large enough to reach the target statistical uncertainty and varied enough in its conditions to explore the major systematic effects. To measure the anomalous precession frequency with 100 ppb statistical precision, the ultimate dataset must contain over  $10^{11}$ , one hundred billion, high-energy decay positrons. Table 2.5 presents an estimate of how long it might take to collect such a dataset. The number of detected high-energy decay positrons per fill, proportional to the number of stored muons per fill, is a key ingredient in this calculation. The experiment's running time scales directly with this number. At the design rate of 1000 high-energy decay positrons per fill, one to two years of running is sufficient. At 500 high-energy decay positrons per fill, the experiment would take two to four years. With rates much lower than that, the experiment's target precision begins to slip entirely out of reach. Therefore, as soon as the FNAL particle beam was available, the E989 collaboration directed all effort to the vital project of assessing and optimizing the experiment's muon injection and storage efficiency.

### ***5.1 Summer 2017 Commissioning Run***

FNAL's accelerator division first delivered beam to the E989 experiment on May 23, 2017. On May 31, particles were successfully injected through the inflector magnet and traveled a full rotation around the storage ring, producing signals in all 24 calorimeters. The accelerator shut down for the summer on July 7, 2017. During the short intervening five week commissioning run, E989 reached many milestones. Included among these are the first stored muons, the first  $\omega_a$ , the first momentum distribution measurement, and the first observations and measurements of beam dynamics in the storage ring. Time was divided between tuning

the injection beamline and tuning the storage ring. Tuning in both cases refers to varying operational parameters of components such as power supplies and pulser timings to optimize the number of particles delivered to the experiment and the fraction of incoming particles that are successfully stored. As storage ring tuning depends on a stable injected beam for reliable diagnostic feedback, tuning the ring and tuning the beamline are generally mutually exclusive tasks. Furthermore, the accelerator division often uses signals from the storage ring to guide their efforts and thus prefers that the storage ring be kept as stable as possible during their tuning sessions.

At the time of the summer 2017 commissioning run, transmission through the delivery ring had not yet been optimized. Recall from Chapter 2 that the delivery ring allows for proton removal using time-of-flight separation. To avoid beam losses, excepting a few dedicated tests, the delivery ring was bypassed entirely throughout the run. Consequently, with a shorter beamline and no proton removal, the injected beam contained a large number of protons and pions. The ratio of protons to muons was approximately 60 to 1. For the most part, when the storage ring is tuned to store 3.1 GeV muons it will also store 3.1 GeV protons. One need only adjust the kicker pulse time to account for the time-of-flight difference between muons and protons, the latter being more massive and therefore slower. To leverage the higher statistics provided by the large number of injected protons, storage studies were conducted with the aim to maximize the number of stored protons. For studies requiring muons, the kicker pulse time was adjusted earlier from the proton storage optimum by a known offset of about 60 ns. While other detectors, such as beam profile monitors (fiber harps) and trackers, provided invaluable information during the commissioning run, the discussion here will focus on results obtained using the calorimeters. The efforts of the summer 2017 commissioning run will be described in some detail as the author of this dissertation was on-site throughout.

### 5.1.1 Storage Diagnostics

Online diagnostic tools provide real time feedback regarding beam storage. As soon as an event is available from the data acquisition system, it is shipped through a local network to a dedicated data quality monitoring (DQM) computer and subjected to a simplified version of the reconstruction procedure described in Chapter 4. Interactive monitoring web pages containing plots and tables update as new data becomes available. The DQM event processing time is usually less than 100 ms. Such immediate feedback is quite useful in tuning exercises where optimal beam storage settings must be found empirically.

The calorimeter reconstruction produces a collection of calorimeter hit times, energies, and positions. Varying derived plots and variables are all different manipulations of these hit times and energies. For example, a number of scalar variables found useful throughout commissioning are constructed simply by counting hits within a fill that pass certain time and energy cuts. One such variable was named *decay positrons*, and it counts high-energy decay positrons relatively late in time. Typically it is calculated as the number of hits found later than  $30\,\mu\text{s}$  after beam injection and with energies greater than  $1.8\,\text{GeV}$ , though the exact values of these cuts changed throughout the commissioning run. The decay positrons variable is a very clean signal for measuring muon storage efficiency as there are no other sources of high-energy hits besides muon decays, and almost all muons in the ring after  $30\,\mu\text{s}$  will remain stored until they decay. Its exact value is also meaningful as it is designed to use the same cuts as a T-Method analysis; it should be approximately 1000 per fill after achieving the experiment's design storage rate. Accounting for calorimeter acceptance, the muon decay energy distribution, and the boosted muon decay lifetime, 6% of stored muons will result in calorimeter hits satisfying the *decay positrons* time and energy cuts.

Storage studies during the summer 2017 commissioning run were conducted primarily with protons. Protons do not produce high-energy calorimeter hits. When a proton hits a calorimeter, it acts as a minimum ionizing particle and typically travels straight through without depositing much energy. Proton hits are usually reconstructed with energies near

200 MeV, although there is a high-energy tail. The signal for muon storage efficiency comes from positrons produced by muon decays in flight. Protons, being stable particles, orbit in the storage ring “forever” and produce no signal in the calorimeters. To obtain a suitable signal for measuring proton storage efficiency, protons must be deliberately ejected from the ring. This was accomplished by ramping down the electrostatic quadrupoles, which provide vertical focusing. As the quadrupoles are ramped down, each particle’s vertical betatron oscillation amplitude grows until the particle leaves the storage region and is lost. Many of these particles scatter off of some material, lose energy, curl inwards in the magnetic field, and strike a calorimeter. See Figure 5.1 for two examples of calorimeter hit time spectra from runs where the quadrupoles were ramped down during the data acquisition window. The procedure for storage studies with protons was to ramp down the quadrupoles a few hundred microseconds after beam injection and count the number of resulting calorimeter hits. Any proton still in the ring when the quadrupoles are turned off, hundreds of microseconds after beam injection, can safely be considered as stored. While it is not clear exactly what fraction of stored protons will result in a calorimeter hit when the quadrupoles are turned off, it is surely the case that the more stored protons there are the more such hits there will be, which is all one needs from a diagnostic tuning variable. The primary variable used for proton storage optimization was named *proton launch*; it is the number of calorimeter hits after the quadrupole ramp down.

Two controllable parameters that have a large effect on storage rates are the inflector magnet current and the kicker pulse timing. The inflector current affects the size of the residual magnetic field inside the inflector channel through which all injected particles pass. Different inflector magnetic field values result in different transverse injection angles relative to the ideal orbit. The injection angle changes the efficacy of the kick and thus the storage efficiency. The inflector current resulting in optimal storage depends on the properties of the beam delivered by the accelerator. For example, if the beam arrives with some nonzero average transverse angle, a different inflector current may be needed to compensate. Kicker pulse timing variation is needed primarily to ensure that the kicker fires when the injected

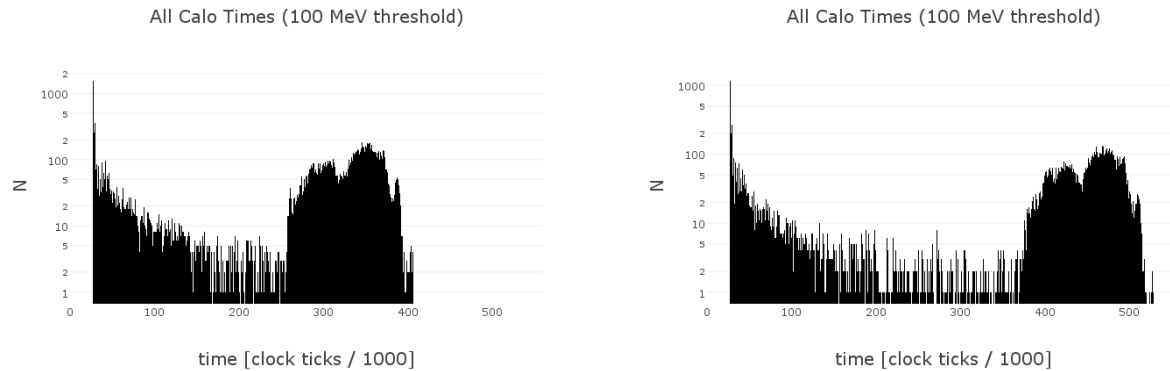


Figure 5.1: Online displays of calorimeter hit times with different quadrupole ramp down times. Clock ticks are 1.25 ns. The spike at early times is the beam injection. The following long tail is a combination of muon decays and particle losses. The large enhancement at late times occurs when the quadrupoles are slowly turned off and all stored particles are ejected from the ring. On the left, the quadrupoles begin to turn off at about 270,000 clock ticks and, on the right, they begin to turn off at about 380,000 clock ticks. This late time signal was used for storage tuning in the 2017 E989 commissioning run.

beam is physically passing the kicker plates. The E989 kicker pulse is not an ideal square shape, so fine time adjustments are also needed to empirically determine the best alignment between the kicker pulse shape and the injected beam's time profile. As with the inflector current, the optimal kicker pulse timing depends on properties of the injected beam, which change each time the accelerator division tunes the beamline. As beamline tuning was a routine activity during the commissioning run, so were kicker timing scans and inflector current scans. The objective of these scans was to find the inflector current and kicker timing settings that yielded the largest measured *proton launch* and *decay positrons*.

For an interesting piece of history, see Figure 5.2. This figure shows the results of the first ever E989 inflector current and kicker timing optimization scans. The scans took place in the early hours of June 9, 2017. Clear storage optima were found for both parameters, with the expected offset between optimal kicker timing for muon storage and for proton storage.

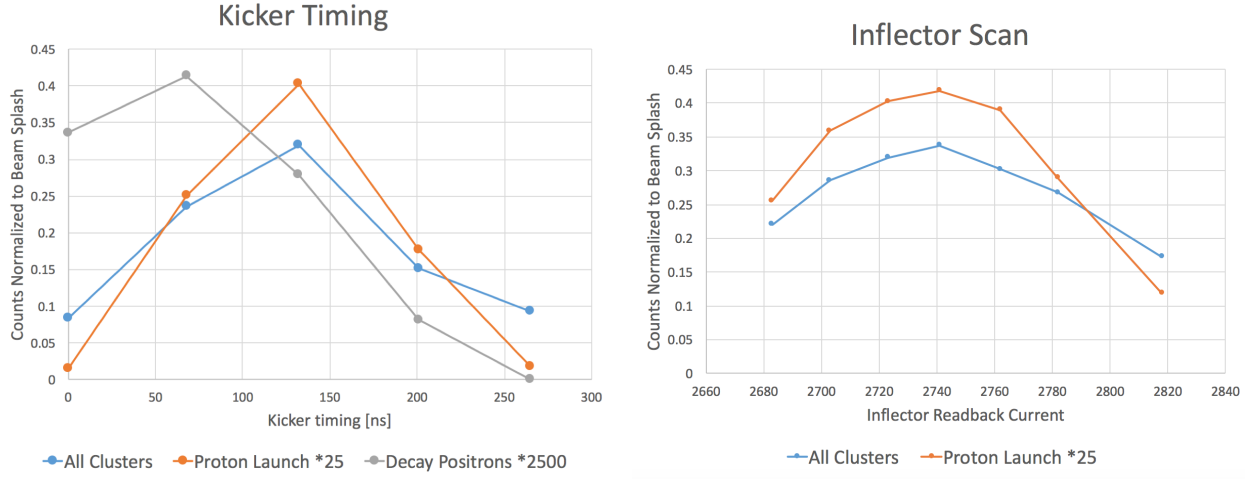


Figure 5.2: The first E989 kicker timing and inflector current scans, conducted in the early morning of June 9, 2017. Inflector current is in units of amperes. The diagnostic variables *proton launch* and *decay positrons* are described in the text. *All clusters* counts all calorimeter hits later than one microsecond after beam injection. Beam splash refers to a measure of the initial pulse seen in all calorimeters at injection time, and it is used as a normalization to remove the effect of varying injected beam intensity. *Decay positrons* measures muon storage and *proton launch* measures proton storage. As expected, the kicker pulses timing for optimal muon storage is approximately 60 ns earlier than it is for optimal proton storage.

### 5.1.2 First $\omega_a$ Measurement

On the evening of June 10, 2017, the E989 kicker timing was configured for muon storage. Muon data collection occurred throughout the night, and the resulting dataset was analyzed the following morning. T-Method analysis of the calorimeter hits yielded an  $\omega_a$  measurement with 0.1% precision; approximately 10,000 hits passed the 1.8 GeV energy cut, see Figure 5.3. This was a celebrated achievement, coming less than two weeks after the first beam injection through the inflector magnet. The storage rates, however, were nowhere near where they needed to be. At the E989 design storage rate, it would take less than one second to collect 10,000 high-energy positrons. On June 10, 2017, it took an entire night. This was not a complete mystery, as neither the beamline nor the kickers nor the quadrupoles were operating at full power. Nevertheless, it served as a reminder of how much work was left to be done.



### 5.1.3 Vertical Beam Position and Radial Fields

The storage ring's magnetic field is not purely in the vertical direction. At the level of 10 ppm relative to the 1.5 T storage field, there are also radial and longitudinal field components with azimuthally varying strengths. Subject to the Lorentz force, charged particles orbiting the storage ring will be shifted vertically in the presence of radial magnetic fields. Equating the force exerted by the average radial field to that exerted by the vertically focusing quadrupoles, one finds that a nonzero average radial magnetic field will shift the equilibrium orbit vertically

according to [82]

$$\frac{\Delta y}{R_0} = \frac{1}{n} \cdot \frac{B_r}{B_0}, \quad (5.1)$$

where  $R_0$  is the storage ring radius,  $n$  is the field index defined in 3.3.1, and  $B_r$  is the azimuthally averaged radial field. Prior to the 2017 commissioning run, the average radial field was measured to be 70 ppm [83]. With typical  $n$  values on the order of 0.1, a 70 ppm average radial field will shift the equilibrium orbit by about 5 mm. This is not desirable because it effectively reduces the available storage volume, and because it complicates the treatment of systematic effects such as the pitch correction and the electric field correction.

The E989 storage ring features a correction mechanism for handling nonzero average radial fields. Concentric current coils above and below the vacuum chambers, called surface coils, are capable of producing an azimuthally uniform radial field by passing current in opposite directions above and below [37]. The surface coils are another important tool in the tuning and commissioning endeavor, and a number of surface coil current scans were conducted during the commissioning run.

The results from a typical surface coil current scan are shown in Figure 5.4. In this scan, completed in June 2017, data was collected at varying surface coil current settings. At each setting, the calorimeter data was analyzed and the vertical reconstructed position, averaged over all calorimeters, was recorded. The expectation was that the average radial field, and accordingly the beam's average vertical position, would change linearly with the surface coil current. That is exactly what was observed. The same measurement was conducted using retractable beam profile monitors and consistent results were obtained. The surface coils were capable of compensating for the nonzero average radial magnetic field and centering the stored particle beam.

#### 5.1.4 *Lost Proton CBO Studies*

In addition to maximizing the number of stored particles, another tuning objective was to reduce the observed coherent betatron oscillation (CBO) amplitude. As discussed in

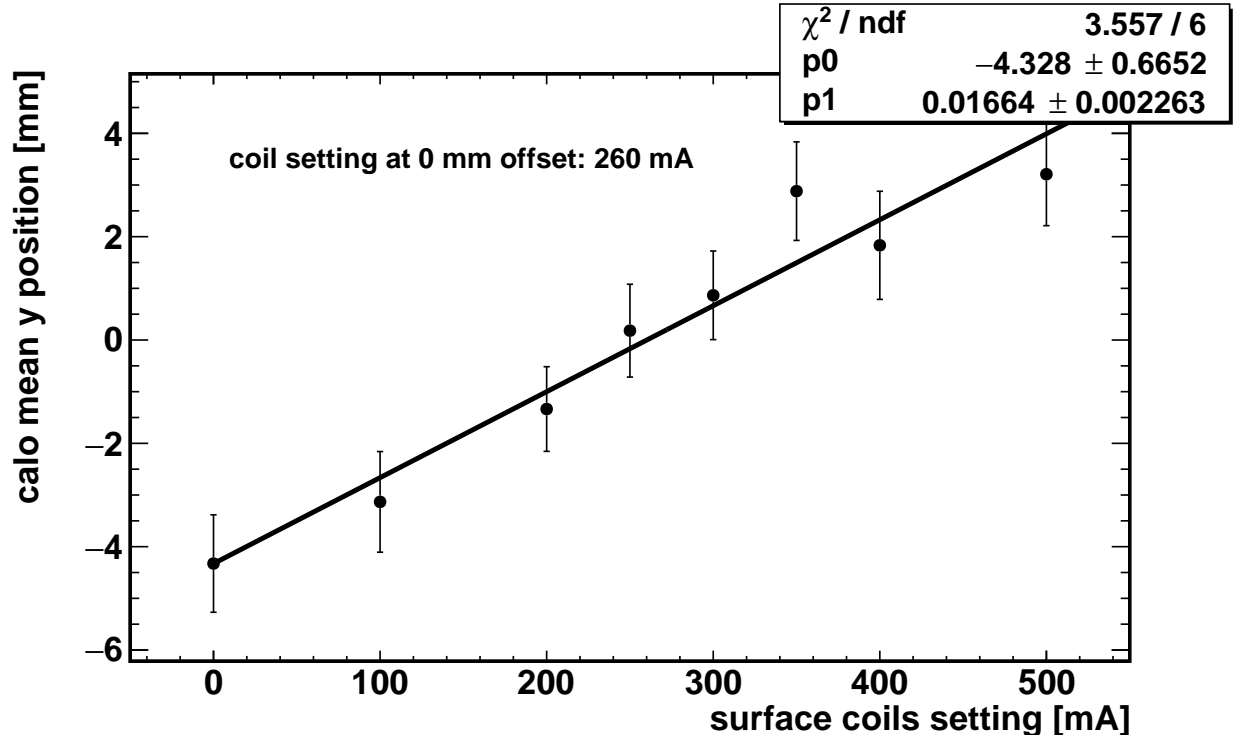


Figure 5.4: Surface coil current scan conducted in June 2017. The objective was to find the current setting that brought the stored beam as close to the vertical center of the storage volume as possible. Vertical positions shown in this graph are reconstructed from calorimeter hits. The optimal setting was found to be 260 mA, in agreement with the result obtained using beam profile monitors.

3.3.5, the CBO frequency appears in calorimeter hit spectra and complicates the precession frequency fitting. An optimal injection followed by an optimal kick will both maximize storage efficiency and minimize the CBO amplitude. Conversely, a sub-optimal kick will leave the beam centroid offset from the center of the storage region and create a large coherent oscillation. The quadrupole high voltage setting determines the CBO frequency.

The CBO frequency is easily observed in the hit time spectrum of protons lost from the storage ring, obtained by selecting calorimeter hits in the energy range from 100 MeV to 500 MeV. Losses are generally induced when particles scatter off of material positioned just outside of the storage region. This is most likely to occur when a particle reaches an extrema

of its oscillation about its equilibrium orbit. Furthermore, unlike with decay positrons, the lost proton time spectrum does not contain any other oscillations with comparably large amplitudes. The lost proton time spectrum,  $L_p(t)$ , fit well to a function of the following form

$$L_p(t) = N_0 e^{-t/\tau_L} [1 + A_{CBO} \cos(\omega_{CBO} t - \phi_0)]. \quad (5.2)$$

During the 2017 commissioning run, the lost proton CBO amplitude,  $A_{CBO}$ , was between 20% and 50%, depending on the particular data taking configuration. The loss lifetime,  $\tau_L$ , also varied quite a bit throughout the run but was generally between  $50 \mu\text{s}$  and  $80 \mu\text{s}$ . See Figure 5.5 for an example lost proton spectrum fit to the above function.

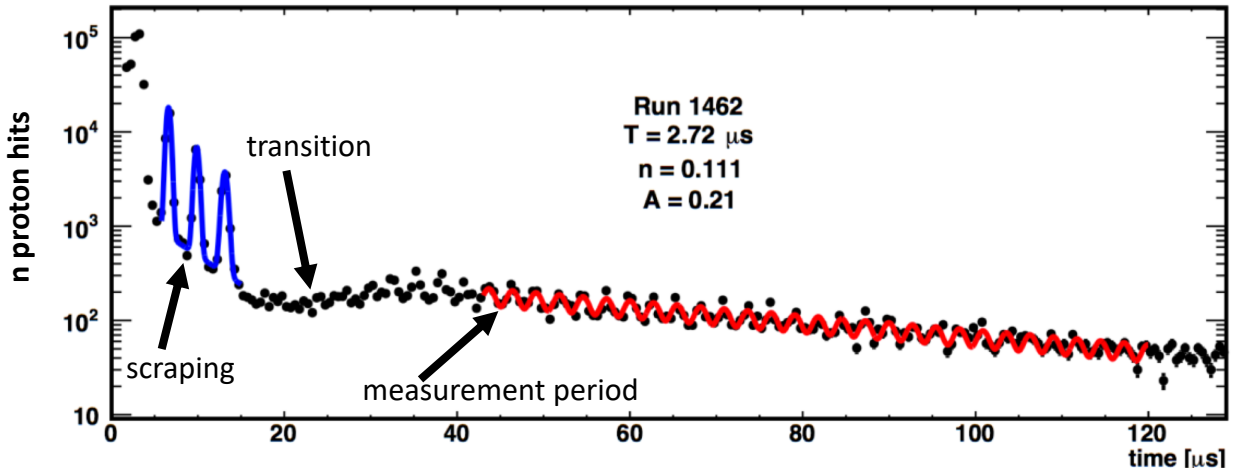


Figure 5.5: A lost proton time spectrum from the summer 2017 commissioning run. The large peaks at early times are caused by scraping, which is when the quadrupoles are charged asymmetrically to force early losses of particles at the edges of the storable phase space. The transition is when the quadrupoles are ramped from the scraping potential to the measurement potential. The last period is the measurement period, in which the time spectrum is fit for the CBO frequency and amplitude. The listed measured parameters are the CBO period, the quadrupole field index, and the CBO amplitude.

Lost proton measurements enabled convenient, calorimeter based extractions of CBO amplitudes and frequencies for all tested kicker and quadrupole high voltage configurations. Expectations were that the CBO frequency would be primarily determined by the quadrupole

high voltage, which sets the field index, and the CBO amplitude would be primarily determined by the kicker strength, which sets the initial phase space conditions of the stored beam. Following the derivations and definitions in Section 3.3, the quadrupole field index,  $n$ , should be directly proportional to the quadrupole high voltage. The CBO frequency is related to  $n$  by

$$\omega_{CBO} = \left(1 - \sqrt{1 - n}\right) \omega_c, \quad (5.3)$$

where  $\omega_c$  is the cyclotron frequency for the same momentum. It should be noted that, because of their larger mass and accordingly smaller velocity, protons take longer than muons to orbit the ring and thus  $\omega_c$  for protons is about 5% lower than it is for muons. Similarly, as the definition of  $n$  given in Equation 3.10 is inversely proportional to particle velocity,  $n$  for protons is about 5% larger than it is for muons.

The expected relationships between quadrupole high voltage and  $n$  and between  $n$  and  $\omega_{CBO}$  were confirmed experimentally, both with the calorimeters and with the beam profile monitors. The CBO amplitude, however, appeared to be independent of the kicker high voltage setting. This unexpected observation was confirmed with multiple detector systems. As all models and common sense had predicted that increasing the kicker strength would reduce the CBO amplitude, finding the CBO amplitude to be independent of kicker strength—even when the strength was reduced by a factor of 4—was quite curious. At the time of this test, kicker high voltage calibration was still a work in progress, and the kick strength in absolute units of kV or G was not well known. The only way to reproduce the kicker strength scan’s result with simulations and models was to posit a kick strength more than a factor of 2 below the design value. With kicks that weak, the CBO amplitude plateaus at approximately 20 mm, the apparent maximum value that can be sustained while still storing beam [84]. Thus, an important finding of the kicker strength scan was that the applied kick was drastically weaker than intended. It was unclear exactly how far there was to go, but the evidence suggested that at least a factor of 2 increase in strength was needed to reach the design value and achieve an optimal kick.

### 5.1.5 *Final Storage Rates*

E989's first commissioning run ended on July 7, 2017. Despite the many achievements and successes of the run, the best achieved storage rates painted a grim picture. The experiment had achieved 30 high-energy decay positrons per fill, 33 times fewer than the 1000 high-energy decay positrons per fill that were designed for. A factor of 4 improvement was expected from increased beam intensity on the pion production target, but that still left close to a factor of 10 that had to be recovered from improved transmission efficiency along the injection beamline and improved performance of the storage ring itself. Regarding the storage ring, increasing the kicker strength was a clear priority toward which significant effort was expended. Beyond that, during the interim between the summer 2017 accelerator shutdown and the resumption of operations the following November, lessons learned from the first commissioning run were applied to the improvement of virtually all aspects of the E989 experiment.

## 5.2 **2017-2018 Commissioning**

Beam returned to the E989 experiment in November 2017. The collaboration quickly re-established conditions from the end of the previous run and resumed its efforts to increase storage efficiency. Unlike the summer 2017 commissioning run, which lasted only five weeks, the 2017-2018 run extended from November 2017 through July 2018, about eight full months of running. Delivery ring proton removal was successfully demonstrated early on and used throughout the run. The number of high-energy decay positrons per fill was 30 in November, 200 in February, 400 in March, and passed 500 in May. By the end of the run, the experiment was consistently measuring 600 high-energy decay positrons per fill. In March, the experiment's leadership declared that storage rates were sufficient to obtain a meaningful physics result and declared the start of Run 1 physics data collection.

The 2017-2018 commissioning will not be described in detail. While much of the collaboration worked to bring muon storage rates to level required for the target  $a_\mu$  precision,

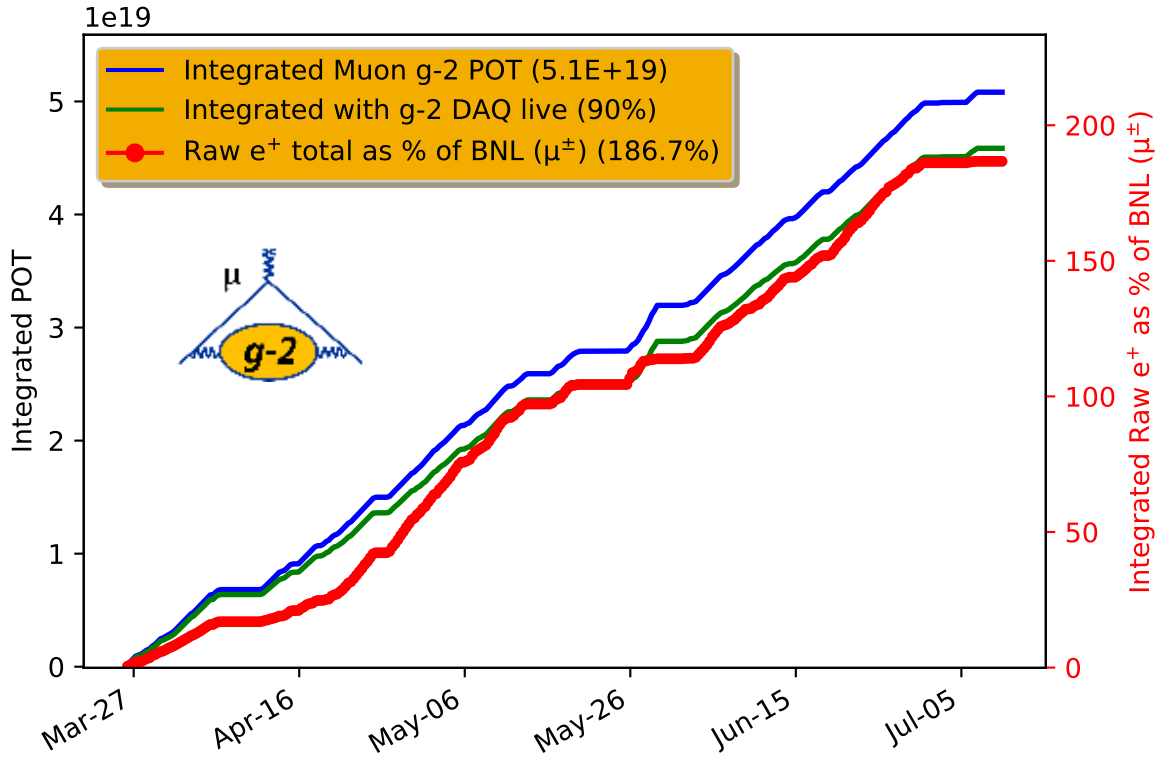


Figure 5.6: Accumulated statistics versus time throughout E989 Physics Run 1, courtesy of Mark Lancaster. The left vertical axis is the total number of accumulated accelerator protons incident on the pion production target. The right vertical axis shows the number of detected high-energy positrons relative to the final dataset of the Brookhaven  $g-2$  Experiment, E821. By the end of Run 1, E989 had accumulated nearly twice the number of high-energy positrons that were in E821’s final combined dataset.

the author of this dissertation was developing an  $\omega_a$  analysis framework in anticipation of the upcoming physics dataset. The 2017-2018 commissioning efforts utilized to great effect the online diagnostic tools described earlier in this chapter. Among the most important hardware improvements during the 2017-2018 commissioning period were the following:

- The FNAL accelerator division achieved delivery ring proton removal while maintaining muon injection intensity sufficient for the E989 experiment.
- Beamline parameters along the M4 and M5 lines were optimized for maximum muon

storage. This optimization effort is described in detail in the University of Washington PhD dissertation of Nathan Froemming [85].

- Storage vacuum conditions were improved through the efforts of the ring engineering staff.
- An error in the kicker voltage readings was discovered. Following the correction of this error, the kicker high voltage was increased substantially, and the *decay positrons* storage diagnostic variable increased from 225 per fill to 530 per fill.
- In-vacuum trolley runs provided maps of the magnetic field in the muon storage region.
- Uniformity of the storage ring's magnetic field was improved by configuring the surface coils to cancel large multipole components measured by the NMR trolley. The surface coil correction was a component of the University of Washington PhD research conducted by Rachel Osofsky.
- The absolute magnetic field calibration system was installed.

Despite these successes, a number of suboptimal conditions persisted into the physics run. Some examples of these suboptimal conditions are the following:

- A large CBO amplitude, measured by the tracking detectors to be 10 mm at  $30\ \mu\text{s}$  [86].
- An average stored muon momentum significantly higher than the design momentum.
- A large muon loss rate.
- Regular quadrupole and kicker sparks and other similar failures that interrupted data taking and resulted in a physics dataset with a number of different storage ring configurations.



Each of the above abnormalities affected the  $\omega_a$  analysis that will be described in the next chapter. The large CBO amplitude and the high momentum bias suggested the need for a further increase of the kicker high voltage, which is a planned improvement for the 2019 physics run.

Thanks to the hardware improvements made during the 2017-2018 commissioning, by the end of Run 1, E989 had accumulated a dataset approximately twice as large as that of the Brookhaven  $g - 2$  experiment, as shown in Figure 5.6. The remainder of this document presents an  $\omega_a$  analysis applied to data from E989 Run 1.

## Chapter 6

### E989 RUN 1 PRECESSION FREQUENCY ANALYSIS

E989 Run 1 occurred between late March and early July of 2018. Data collection conditions—for example quadrupole and kicker voltage set points and injected beam intensities—varied significantly throughout the run. Run 1 was thus divided into a number of subdatasets with stable conditions. Each subdataset must be fit separately to properly account for systematic effects such as beam oscillations.

One subdataset, named the 60-Hour Dataset, was defined and processed well in advance of the others. It was therefore the first physics dataset analyzed for  $\omega_a$ . This chapter describes an anomalous precession frequency analysis of the E989 Run 1 60-Hour Dataset.

#### **6.1 The 60-Hour Dataset at a Glance**

The 60-Hour Dataset was collected during a 60 hour period with stable running conditions that occurred between April 22 and April 25, 2018. Over this period, approximately one billion high-energy decay positrons—that is, positrons passing the T-Method energy cut described in Section 2.6—were collected. Table 6.1 lists the 60-Hour Dataset running conditions relevant to the  $\omega_a$  analysis.

Practical hardware concerns prevented the E989 collaboration from operating the electrostatic quadrupoles and the pulsed kicker magnet at their design voltages during the 60-Hour Dataset. In fact, this was true throughout Run 1. Storage conditions were consequently suboptimal. Specifically, the ring’s storage efficiency was lower than desired, and the distribution of stored muon equilibrium radii was not centered at the design value of 7.112 m. Fast rotation analyses—described in Section 3.3.4—found an asymmetric cyclotron frequency distribution skewed toward lower frequencies and therefore higher momenta and higher equi-

Table 6.1: Running conditions of the E989 Run 1 60-Hour Dataset, which occurred between April 22 and April 25 of 2018.

Quantity	Value
Number of fills	$\sim 1.4 \times 10^6$
Number of high-energy $e^+$	$\sim 1 \times 10^9$
Kicker voltage setting	130 kV
Quadrupole voltage setting	18.3 kV
Field index ( $n$ )	0.107
Coherent betatron frequency	0.37 MHz
Vertical betatron frequency	2.19 MHz
Vertical waist frequency	2.32 MHz

librium radii, see Figure 6.1. This is exactly what one would expect from a suboptimal kick: muons are injected with a positive radial offset, so high-momentum muons—having higher equilibrium radii—are injected closer to their ideal orbits and thus require smaller kicks.

Figure 6.2 shows the raw calorimeter hit spectrum measured in the 60-Hour Dataset. The expected features are present. All energy bins display oscillations at the anomalous precession frequency except the one where the  $g - 2$  asymmetry is 0. The large peak in the energy spectrum at 170 MeV is indicative of muon losses. As will be presented shortly, the expected beam oscillation frequencies are also detectable in the calorimeter hit spectra.

Considering statistics alone, the 60-Hour Dataset is sufficient for a 1.3 ppm  $\omega_a$  measurement. Following appropriate corrections, the systematic effects discussed in Chapter 3 are expected to contribute less than 100 ppb of uncertainty. An  $\omega_a$  measurement from the 60-Hour Dataset in isolation is therefore expected to be entirely dominated by statistics. However, it is incumbent upon the analyzer to demonstrate that this is the case. A complete analysis of the 60-Hour Dataset must include an evaluation of all known systematic effects

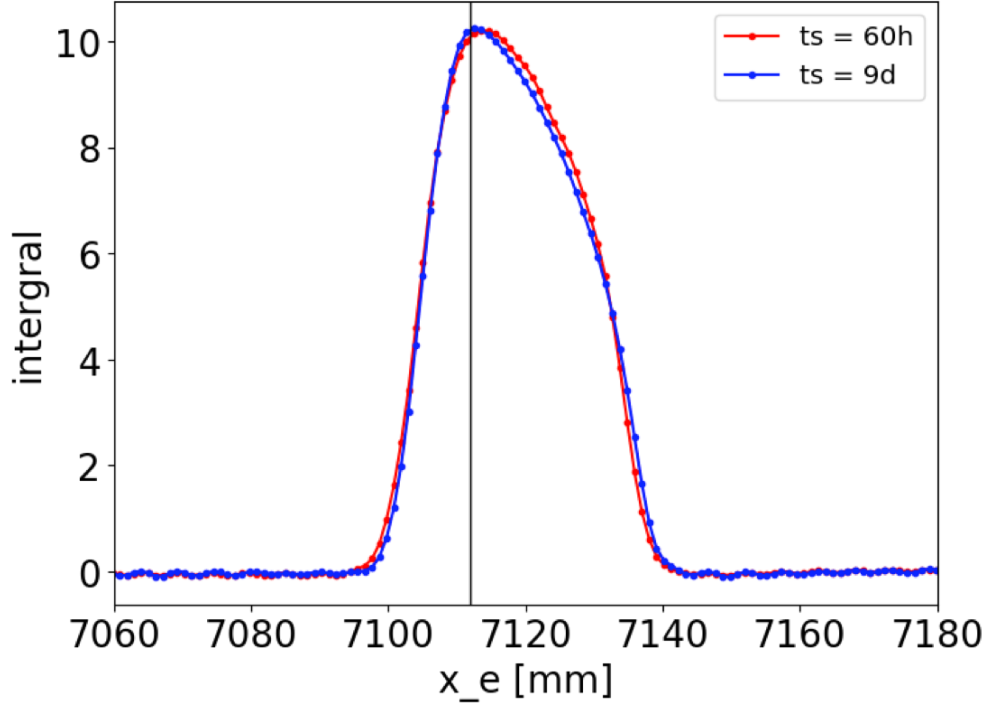


Figure 6.1: The equilibrium radius distribution of stored muons in the 60-Hour Dataset, in red, and from another Run 1 dataset, called the 9-Day, in blue, courtesy of Kim Siang Khaw. The black line depicts the design radius at the center of the storage region. A suboptimal kick resulted in preferential storage of high-momentum muons, which have large equilibrium radii.

and must be statistically well-behaved.

## 6.2 Data Quality Conditions

Although the 60-Hour Dataset featured generally stable running conditions, there were brief and intermittent failures of the accelerator beamline and of some experiment subsystems. For example, the injected beam pulse occasionally had an anomalously low intensity or an unusual time profile, beam delivery entirely ceased for brief periods, and quadrupole or kicker sparks temporarily disturbed the ring's ability to store particles. As these sorts of data taking interruptions are unavoidable, data quality cuts are necessary when analyzing any dataset

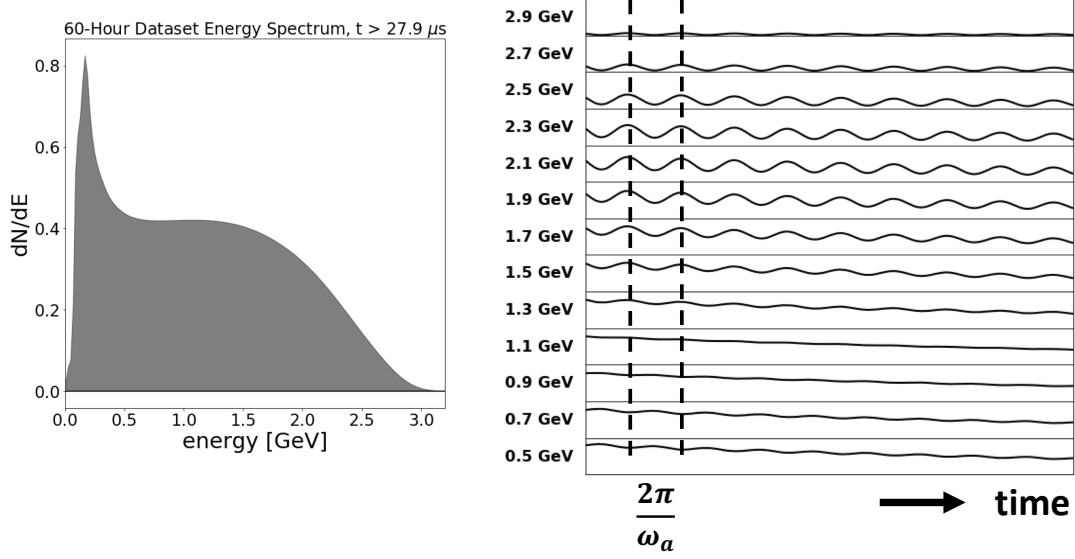


Figure 6.2: Calorimeter energy and time spectra from the 60-Hour Dataset. On the left is the energy spectrum measured over all times greater than  $30.2 \mu\text{s}$ . The peak at 170 MeV is from lost muons. On the right is the calorimeter hit time spectrum measured in 200 MeV energy bins from 0.5 to 3.0 GeV. The  $\omega_a$  oscillations are visible in each energy bin, as is the phase inversion that occurs at 1 GeV. Each energy bin's time spectrum is drawn with the same y-axis range.

to veto muon fills taken under conditions not suitable for physics data analysis.

A data quality group was convened within the collaboration and charged with defining a set of official conditions that muon fills used for data analysis should satisfy. These conditions are essentially a list of cuts, or allowable ranges, that are applied to signals from measurement systems including (among others) the T0 detector, the calorimeters, and voltage readouts from the kickers and quadrupoles. The data quality cuts remove fills in which the storage ring, the injection beamline, or any calorimeter is not operating as desired. The values chosen for data quality cuts are specific to each dataset. Unless otherwise stated, only muon fills passing the official data quality cuts were considered for the analysis presented in this chapter. Data quality cuts removed approximately 9% of the muon fills in the 60-Hour

Dataset.

### 6.3 *Blinding*

There are two levels of blinding in the  $\omega_a$  measurement. The first is implemented in the clock system and is called hardware blinding. Hardware blinding amounts to adjusting the 40 MHz master clock signal that drives the waveform digitizers: the nominal 40 MHz frequency is shifted by some value in the range of  $\pm 25$  ppm. The exact value of the shift is known only by two FNAL personnel who are not otherwise involved with the  $g - 2$  experiment. When units involving time ( $\mu\text{s}$ , MHz, etc.) are presented in this document, those units should be understood to be correct only up to the  $\pm 25$  ppm clock blinding. The secret clock frequency is only revealed once all  $\omega_a$  analyses have been completed to the satisfaction of the collaboration at large.

The second level of blinding occurs in software. In practice, analyzers do not fit for  $\omega_a$  directly. To ensure that analyzers are not influenced by one another's results and that the internal consistency of a given analysis is evaluated independently from any prior bias regarding what the result should be, fits are conducted in terms of a blinded parameter  $R$  that is related to  $\omega_a$  by

$$\omega_a(R) = 2\pi \cdot 0.2291 \text{ MHz} \cdot [1 + (R - \Delta R) \times 10^{-6}]. \quad (6.1)$$

In words,  $R$  is a unitless parameter, a 1 unit shift of  $R$  corresponds to a 1 ppm shift of  $\omega_a$  relative to 0.2291 MHz, and  $\Delta R$  is a fixed, secret offset. Each analyzer has a separate  $\Delta R$  value taken randomly from a range of approximately  $\pm 24$ . Software blinding allows for analyzers to first conduct independently blinded analyses and later, after they have completed all their consistency checks, directly compare results without ever having to know the secret clock frequency.

Even after all  $\omega_a$  analyzers have set their  $\Delta R$  values to zero and the secret clock frequency has been revealed,  $a_\mu$  cannot be determined from  $\omega_a$  without the complementary magnetic field measurement,  $\omega_p$ . The  $\omega_p$  analysis also employs software blinding and is conducted in

parallel by multiple teams. Only after both the  $\omega_a$  and  $\omega_p$  analyses have been demonstrated to be internally consistent will all blinding be removed, the numbers combined, and a value for  $a_\mu$  determined.

#### **6.4 Five-Parameter T-Method Fit to the 60-Hour Dataset**

Section 2.6 described an  $\omega_a$  analysis technique called the T-Method, wherein one histograms the hit times of positrons above a chosen energy threshold. The simplicity of the T-Method makes it an ideal first analysis approach for building confidence in one's understanding of the data.

Equation 2.17 gives the expected functional form of a T-Method histogram in the absence of beam dynamics and detector effects. Including the software blinding, this form suggests the following fit model:

$$N(t) = \mathbf{N_0} e^{-t/\tau} [1 + \mathbf{A} \cos(\omega_a(\mathbf{R}) - \phi)]. \quad (6.2)$$

The five variables in bold are parameters to be extracted from the data. The five-parameter model is not expected to fully describe the T-Method time spectrum obtained from data. It does not contain any of the known beam oscillations, losses, pileup, or time-dependent gain effects. Nevertheless, it is interesting to attempt to fit with the five-parameter model and then to observe the structures present in the residuals.

The optimal T-Method threshold is determined empirically from the data. As explored in Section 2.6, there is a single threshold that maximizes the statistical figure of merit,  $NA^2$ , and therefore yields the smallest  $\omega_a$  uncertainty. In the absence of detector acceptance effects, the optimal threshold is approximately 1.9 GeV. As shown in Figure 6.3, the optimal threshold found in the 60-Hour Dataset was about 1.7 GeV. The difference between the theoretical optimum and that obtained from data depends on detector acceptance and energy calibration.

After determining the optimal threshold, the T-Method histogram is constructed by counting calorimeter hits with reconstructed energies above that threshold as a function of

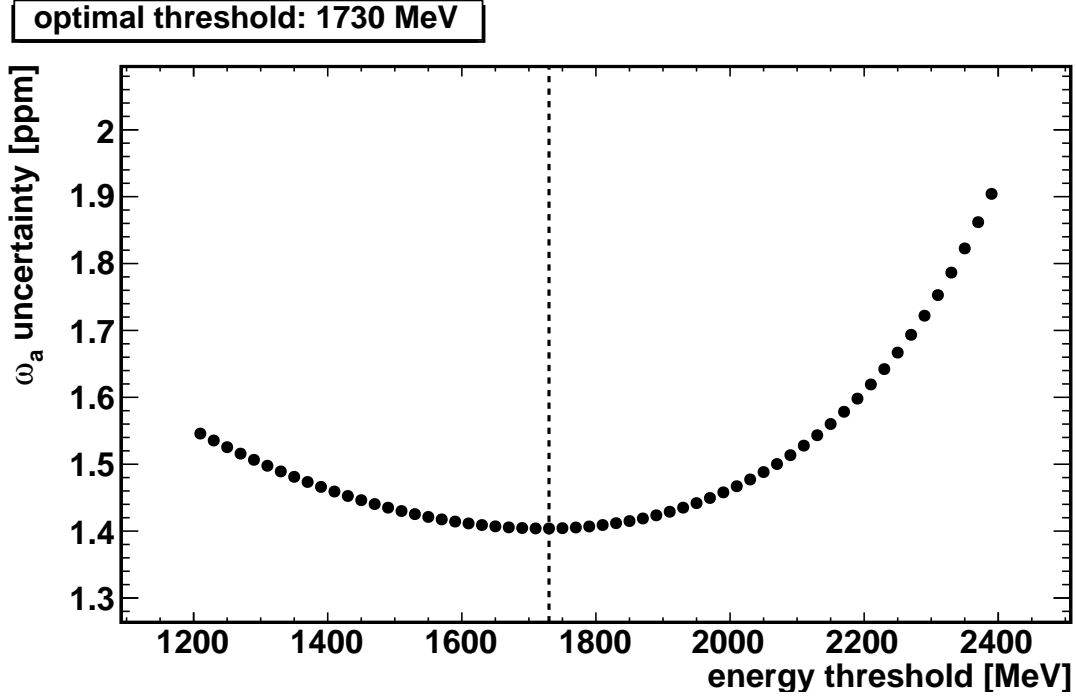


Figure 6.3:  $\omega_a$  precision versus threshold in the 60-Hour Dataset. For each threshold, a T-Method histogram is constructed and fit with the five-parameter model (Equation 6.2).  $\omega_a$  precisions are determined from the fit results. The best precision is obtained with an energy threshold of about 1.7 GeV.

time-in-fill, or time relative to the muon beam injection. The time bin width is set to the average cyclotron period determined from the fast rotation analysis. Doing so reduce the influence of the beam debunching on the analysis histogram. Hits from all calorimeters are included in a single T-Method histogram.

The time domain of the  $\omega_a$  fit is called the measurement period. Fits should begin late enough in the fill that the quadrupole plates have reached their nominal values following beam scraping. Additionally, the majority of muon losses occur early in the fill, so starting the fit later reduces the number of losses during the  $\omega_a$  measurement. The relative contributions of pileup and gain perturbations are also largest early in the fill. Thus, later measurement start times reduce the impact of many systematic effects that can bias  $\omega_a$ . This benefit comes at the expense of reduced statistical power. A  $30.2\,\mu\text{s}$  start time was chosen for this analysis.



Model parameters are extracted using a  $\chi^2$  minimization. All  $\omega_a$  fits presented in this document were conducted using the ROOT software package [87]. The task of the  $\chi^2$  minimization is to find the set of parameters  $\vec{p}$  that minimize the quantity

$$\chi^2 = \sum_{bins} \frac{[N_{bin} - f(t_{bin}, \vec{p})]^2}{\sigma_{bin}^2}. \quad (6.3)$$

In the T-Method case, bin contents are large—over 100 counts in all bins—and constructed by counting discrete events. Thus, the bin variances,  $\sigma_{bin}^2$ , are taken to be the bin contents,  $N_{bin}$ . If the underlying data is indeed distributed according to the fit model, and the bin variances are assigned correctly,  $\chi^2$  values obtained with the above procedure should follow a  $\chi^2$  distribution with  $N_{bins} - N_{params}$  degrees of freedom [5]. The best-fit  $\chi^2$  value, therefore, meaningfully characterizes the goodness-of-fit: one can calculate the probability that a fit with the correct model would yield a  $\chi^2$  as high or higher than the best-fit  $\chi^2$ . If this probability is sufficiently low, one can safely conclude that the employed fit model did not fully describe the data.

Figure 6.4 shows the result of a five-parameter T-Method fit to the 60-Hour Dataset. As expected, the five-parameter fit alone is not adequate, and the fit residuals display many structures. The predominant oscillation frequency present in the residuals is the CBO frequency. Oscillations are also detectable at the vertical waist frequency, the anomalous precession frequency, and at the sum and difference frequencies of the anomalous precession frequency and the CBO frequency. The frequency spectrum of the residuals displays no significant peaks at any unexpected locations. A falling peak at 0 MHz is evidence of slow effects not captured by the five-parameter model. Some known effects with slow components are pileup, gain perturbations, and muon losses.

## 6.5 Corrections to the Five-Parameter Model

### 6.5.1 Beam Oscillations

Physical oscillations of the stored beam affect the calorimeter hit spectra. The primary mechanism by which this occurs is the dependence of the calorimeter acceptance on the

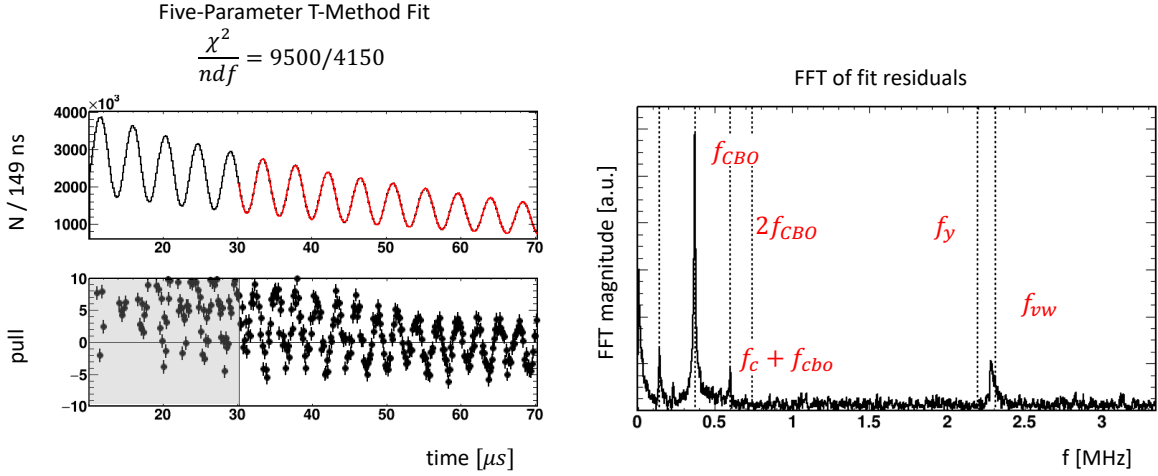


Figure 6.4: Five-parameter T-Method fit to the 60-Hour Dataset. On the left is the fit at early times and the pulls (pulls are the fit residuals divided by the bin errors). The full fit range is not shown. The grayed out region shows the fit extrapolated beyond the fit start time of  $30.2 \mu\text{s}$ . The probability that the correct fit model would yield a  $\chi^2$  value as high as that shown here is virtually 0. On the right is a Fourier transform of the fit residuals. Peaks appear at expected beam oscillation frequencies. There is also a slowly falling peak at 0 MHz.

muon decay position. In other words, the probability that a decay positron will strike a calorimeter varies with the position and momentum direction of its parent muon at the time of decay. To account for this effect, the fit model must allow for a time-dependent acceptance,  $Ac(t)$ :

$$N(t) = Ac(t) \cdot N_0 e^{-t/\tau} [1 + A \cos(\omega_a(R) - \phi)]. \quad (6.4)$$

In a simplified case where there is only CBO with a time-independent amplitude, the above equation would become

$$N(t) = [1 + A_{CBO} \cos(\omega_{CBO}t - \phi_{CBO})] \cdot N_0 e^{-t/\tau} [1 + A \cos(\omega_a(R) - \phi)], \quad (6.5)$$

where  $A_{CBO}$  is the relative size of the acceptance modulation at the CBO frequency. If one were to expand this equation, there would be three oscillatory terms: one with  $\cos(\omega_a)$ , one with  $\cos(\omega_{CBO})$ , and one with the product  $\cos(\omega_{CBO}t) \cdot \cos(\omega_a t)$ . The product term gives rise to the peaks at  $\omega_{CBO} + \omega_a$  and  $\omega_{CBO} - \omega_a$  that were visible in Figure 6.4.

The beam oscillation amplitudes are not time-independent. As shown in Section 3.3, the characteristic radial and vertical oscillation frequencies depend on muon momentum. As the storage ring accepts a range of momenta, there will be a commensurate range of beam oscillation frequencies. Furthermore, the realistic storage potential is not perfectly harmonic, and thus the beam oscillation frequencies are amplitude dependent. The spread in oscillation frequencies causes any coherent beam motion to decohere over time, with a decoherence time roughly equal to the inverse width of the angular frequency distribution. For a 0.15% RMS momentum width and a field index of 0.107, back-of-the-envelope estimates yield approximate decoherence times of  $250\ \mu\text{s}$  for the CBO and  $40\ \mu\text{s}$  for the vertical waist. The CBO persists the furthest into the  $\omega_a$  measurement period, so it is the most important to model correctly.

In addition to the overall acceptance, beam oscillations also affect the  $g - 2$  phase and asymmetry. The  $g - 2$  phase is directly related to the positron drift time. As the beam moves closer and farther from the calorimeters, the average drift time changes, and so too does the average  $g - 2$  phase. Moreover, as the position dependence of the acceptance is different for different positron energies, the observed energy distribution depends on the position of the muon when it decays. This energy-dependent acceptance variation leads to a periodic distortion of the observed energy spectrum at the CBO frequency. The  $g - 2$  asymmetry and phase both depend on positron energy, so the T-Method asymmetry and phase must oscillate at the CBO frequency as well. This is more subtle than the overall acceptance modulation, but it is detectable and must be taken into account.

The tracker system is capable of extrapolating positron trajectories to their associated muon decay vertices. By doing so, the tracking detectors provide the transverse distribution of the stored beam as a function of time-in-fill. Figure 6.5 shows the radial distribution as a function of time as measured by the tracking detectors. The CBO and its decoherence are clearly visible. The precise measurements of the CBO frequency and decoherence time provided by the trackers inform the calorimeter analysis. The decoherence envelope can be described by an exponential decay law (see Figure 6.6).

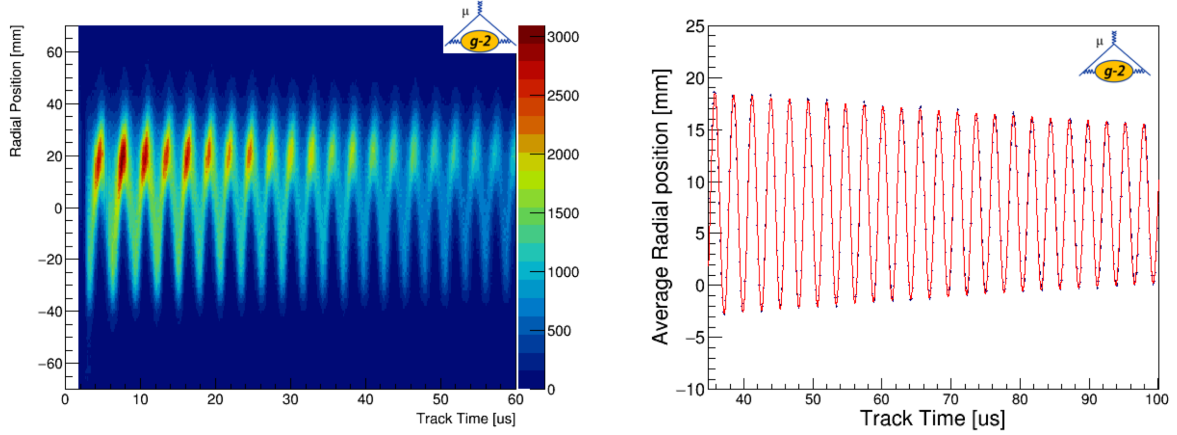


Figure 6.5: Radial distribution of the stored muon beam as measured by the tracking detectors. The left plots shows the full radial distribution and the right plot shows the average radial position as a function of time. The observed oscillations are the CBO. The expected decoherence—manifesting as a reduction in the oscillation amplitude over time—is evident.

As shown in Figure 6.6, careful analysis of the tracker data revealed that the CBO frequency slowly increased throughout each muon fill in the 60-Hour Dataset [88]. This was later attributed to three broken resistors in the quadrupole system, which caused a much longer than expected recovery from the scraping configuration. While the changing frequency was detectable in the calorimeter data, the trackers were able to provide the most precise measurement of its functional form. This form was taken from the tracker-based analysis and used in the calorimeter fits. In principle, the slow quadrupole recovery should also create a changing vertical waist frequency. However, the vertical waist decoheres so quickly relative to the time scale of the frequency change that the assumption of a fixed frequency was sufficient to model the observed oscillations.

To account for all the beam motion effects described in this section, the fit model is adjusted in the following way:

$$N(t) = N_0 \cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot e^{-t/\tau} \cdot [1 + A(t) \cos(\omega_a(R) - \phi(t))], \quad (6.6)$$

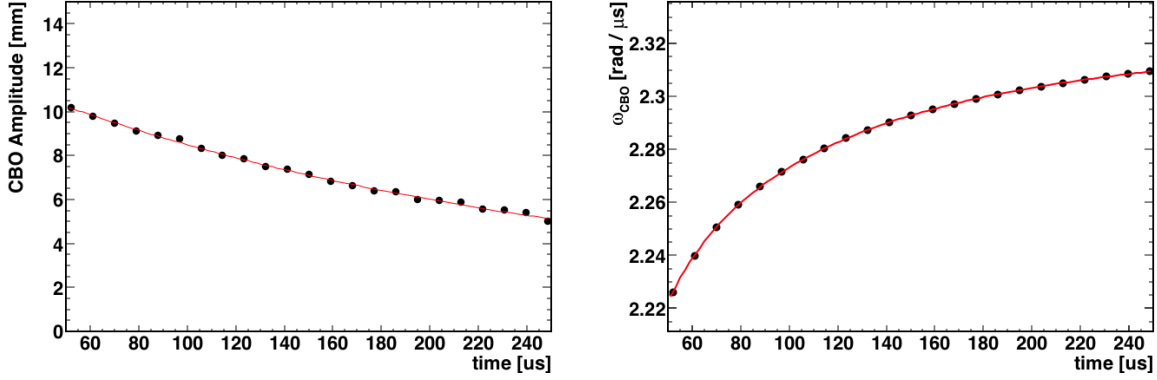


Figure 6.6: CBO envelope and frequency shift in the 60-Hour Dataset as measured by the tracking detectors, courtesy of James Mott. Each graph contains one point per CBO oscillation period.

where

$$\begin{aligned}
 N_{CBO}(t) &= 1 + A_{CBO,N} \cdot e^{-t/\tau_{CBO}} \cos(\omega_{CBO} \cdot t - \phi_{CBO,N}) \\
 N_{VW}(t) &= 1 + A_{VW,N} \cdot e^{-t/\tau_{VW}} \cos(\omega_{VW} \cdot t - \phi_{VW,N}) \\
 \phi(t) &= \phi_0 + A_{CBO,\phi} \cdot e^{-t/\tau_{CBO}} \cos(\omega_{CBO} \cdot t - \phi_{CBO,\phi}) \\
 A(t) &= A_0 \left[ 1 + A_{CBO,A} \cdot e^{-t/\tau_{CBO}} \cos(\omega_{CBO} \cdot t - \phi_{CBO,A}) \right] \\
 \omega_{CBO}(t) &= \omega_{CBO,0} [1 + \delta_{CBO}(t)].
 \end{aligned}$$

This treatment introduces 12 new fit parameters: the CBO and vertical waist frequencies ( $\omega_0$  and  $\omega_{VW}$ ), their decoherence times ( $\tau_{CBO}$  and  $\tau_{VW}$ ), the amplitude and phase of the overall acceptance modulation from the CBO ( $A_{CBO,N}$  and  $\phi_{CBO,N}$ ), the amplitude and phase of the overall acceptance modulation from the vertical waist ( $A_{VW,N}$  and  $\phi_{VW,N}$ ), and the amplitudes and phases of the CBO modulations of the T-Method  $g-2$  asymmetry and phase ( $A_{CBO,\phi}$ ,  $A_{CBO,A}$ ,  $\phi_{CBO,A}$ , and  $\phi_{CBO,\phi}$ ). The changing CBO frequency is described by  $\delta_{CBO}(t)$ , which is fixed from the tracker analysis and introduces no new fit parameters. As will be later seen, none of these parameters is significantly correlated with  $\omega_a$ . Because the trackers and calorimeters have generally different acceptances, it is unclear whether it is

appropriate to fix parameters such as the decoherence times based on tracker analyses. In this analysis, all the above parameters are left free in the  $\chi^2$  minimization.

### 6.5.2 Muon Losses

Equation 6.6 contains the implicit assumption that all muons will remain stored until they decay. This is not the case: the fit model needs further adjustment to account for muon losses during the measurement period. As discussed in Section 4.2.5, lost muons are detectable via their characteristic energy deposition and their tendency to strike multiple calorimeters in sequence, each approximately 6.2 ns after the last. Lost muon candidates are identified as collections of hits coincident between three calorimeters wherein each calorimeter hit is between 50 MeV and 200 MeV and the consecutive time differences are each between 3.75 ns and 8.75 ns. Within a proportionality constant related to the detection efficiency, these selection criteria provide a measurement of the muon loss rate as a function of time that is minimally contaminated by background. Selecting for quadruple coincidences rather than triple coincidences yields a consistent result, although with a factor of 5 reduction in statistics.

Let  $L(t)$  be the number of triple coincidences passing the cuts described above as a function of time. Provided that the lost muon detection efficiency is constant throughout the measurement period and that any background contamination is negligible,  $L(t)$  will be proportional to the muon loss rate as a function of time. The assumption of constant detection efficiency is supported by the observation that  $L(t)$  has the same shape whether it is measured using triple coincidences or quadruple coincidences. An alternative measurement of  $L(t)$  is possible by correlating calorimeter hits with events in the tracking detectors. Although that approach has not been pursued in this analysis, preliminary results from other analyzers show consistency between  $L(t)$  obtained with a coincidence-based measurement and  $L(t)$  obtained with a tracker-based measurement [89]. Figure 6.7 shows  $L(t)$  as measured in the 60-Hour Dataset.

With knowledge of  $L(t)$ , one can write a differential equation governing the number of muons present in the storage ring at any time following the beam injection. This requires

introducing a new variable,  $\epsilon$ , which is the loss detection efficiency. The true loss rate is then  $L(t)/\epsilon$ . A muon can leave the storage ring either by decaying or by escaping. Thus, the change in the number of stored muons in an infinitesimal time interval  $dt$  is

$$dN_{stored} = -\frac{N_{stored}}{\tau} dt - \frac{1}{\epsilon} L(t) dt. \quad (6.7)$$

By inspection, the solution to this differential equation is

$$N_{stored}(t) = N_0 \cdot \left( 1 - \frac{1}{\epsilon N_0} \int_0^t e^{t'/\tau} L(t') dt' \right) \cdot e^{-t/\tau}. \quad (6.8)$$

$N_0$  is the number of muons stored in the ring at  $t = 0$ . Defining the variable  $K_{loss}$  as  $1/(\epsilon N_0)$ , the above equation becomes

$$N_{stored}(t) = N_0 \cdot \left( 1 - K_{loss} \int_0^t e^{t'/\tau} L(t') dt' \right) \cdot e^{-t/\tau}. \quad (6.9)$$

Muon losses cause an effective reduction in the  $N_0$  parameter as a function of time.

Considering muon losses, the fit model becomes

$$N(t) = N_0 \cdot \left( 1 - K_{loss} \int_0^t e^{t'/\tau} L(t') dt' \right) \cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot e^{-t/\tau} \cdot [1 + A(t) \cos(\omega_a(R) - \phi(t))]. \quad (6.10)$$

As the loss detection efficiency is not known a priori,  $K_{loss}$  is a free parameter that must be extracted from the fit. This brings the total count to 18. The integral  $\int_0^t e^{t'/\tau} L(t') dt'$  is calculated numerically for each bin of the T-Method histogram. Though one may worry that the integral depends on  $\tau$ , which is itself a fit parameter, the ultimate results of the  $\omega_a$  analysis are quite insensitive to the exact value of  $\tau$  used when calculating this integral. In the analysis presented here, an iterative procedure is employed whereby the value of  $\tau$  extracted from the previous iteration is used to calculate the lost muon correction for the next. Only two iterations are required for this procedure to stabilize.

It is important to note that the  $K_{loss}$  parameter depends solely on two quantities: the number of muons stored at  $t = 0$  and the loss detection efficiency. Its value should therefore be independent of calorimeter number, energy bin, or analysis technique. The  $K_{loss}$  parameter

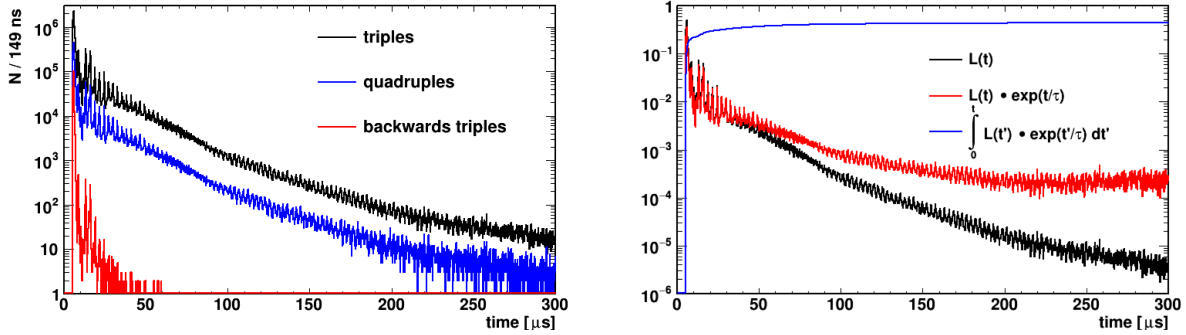


Figure 6.7: Muon losses measured in the 60-Hour Dataset. On the left are the number of triple coincidences and quadruple coincidences measured using the selection criteria described in the text. The shapes are identical. Also shown is the number of nonphysical backwards triples, which are defined using the same selection criteria as for triples except that the direction of increasing hit times is reversed. The very small number of backwards triples indicates a negligible background from uncorrelated coincidences. On the right are the ingredients for creating the muon loss correction in Equation 6.9.

is also critical for determining the relative number of muon losses to muon decays during the measurement period. Knowledge of this quantity is necessary to evaluate the experiment's systematic error from muon losses.

### 6.5.3 The Pileup Correction

The calorimeter data must be corrected for unresolved pileup. Section 3.5.2 estimated the  $\omega_a$  bias resulting from uncorrected pileup, and it can be quite large. Additionally, the fit model given in Equation 6.10 is not valid when there is significant pileup contamination. This section presents an approach for removing pileup contamination present in the calorimeter hit spectra.

First the expected form of the pileup contamination will be discussed. Following the notation from Section 3.5.2, let  $\rho(E, t)$  be the calorimeter hit distribution in the absence of any pileup perturbation, i.e. in the limit of zero detector dead time. Section 3.5.2 discusses the expected pileup perturbation to  $\rho(E, t)$  from unresolved pulse pairs (double pileup).



Repeated here in a slightly different form,

$$\delta\rho_{pu,d}(E,t) \propto \Delta t \left[ \int \rho(E - E_2, t) \cdot \rho(E_2, t) dE_2 - 2\rho(E, t) \int \rho(E_2, t) dE_2 \right], \quad (6.11)$$

where  $\rho_{pu,d}$  is the double pileup perturbation,  $\Delta t$  is the detector dead time, and  $r(t)$  is the overall hit rate a function of time. The first term above corresponds to the false counts that are gained as a result of reconstructing two independent pulses as a single, and the second term corresponds to the two true counts that are lost. The former will in principle be affected by nonlinearities in the reconstruction's treatment of unresolved pulse pairs. These nonlinearities are not included in the pileup correction approach described here. Nevertheless, as will be shown, the generated correction appears to sufficiently remove all pileup in the 60-Hour Dataset. It will be convenient to define the double pulse sum term

$$\rho_{d+}(E, t) = \int \rho(E - E_2, t) \cdot \rho(E_2, t) dE_2, \quad (6.12)$$

so that

$$\delta\rho_{pu,d}(E, t) \propto \Delta t \left[ \rho_{d+}(E, t) - 2\rho(E, t) \int \rho(E_2, t) dE_2 \right] \quad (6.13)$$

Equation 6.13 describes the contamination of the measured energy spectrum from double pileup in terms of the uncontaminated spectrum  $\rho(E, t)$ . It can also be used to generate a pileup correction from the measured hit spectrum, which is itself contaminated by pileup. Recall that the relative double pileup contamination appears at the order  $r(t) \cdot \Delta t$ . Even with a conservative detector dead time and no spatial cluster separation employed in the reconstruction,  $r(t) \cdot \Delta t$  will be at most 1% to 2% and thus can be treated as small. Observe that if in Equation 6.11 one were to replace  $\rho$  with  $[\rho + r(t)\Delta t \cdot f(E, t)]$ —where  $f$  is some function of the same order as  $\rho$ —the difference between the obtained expression and the original would appear at order  $r^2(t)\Delta t^2$ , order  $10^{-4}$ . Thus, using the pileup contaminated hit spectrum to generate the expected double pileup contamination is only wrong by order  $r^2(t)\Delta t^2$ . This is the key observation motivating the pileup correction described in this section. More details surrounding how the pileup correction works in practice will be described shortly.

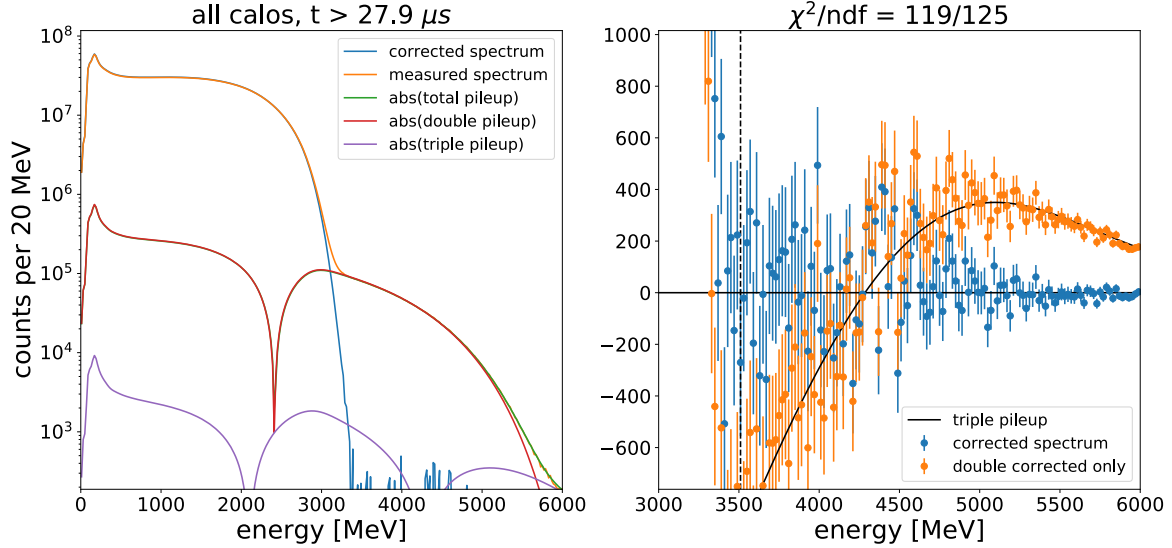


Figure 6.8: The measured energy spectrum, summed over all calorimeters, along with the total pileup correction and its double and triple pileup components. On the left, the absolute values of the pileup spectra are shown. The double pileup spectrum changes sign at 2.4 GeV. The triple pileup spectrum changes sign at 4.3 GeV and 2.1 GeV. On the right is the total pileup corrected energy spectrum and a spectrum that is corrected only for double pileup. The triple pileup correction accounts for the structure remaining after the double pileup correction. The generated pileup spectrum correctly models the detected energy spectrum in the non-physical energy region above 3.5 GeV. The black dashed line on the right denotes the lowest energy used to calculate the displayed  $\chi^2$ , obtained by comparing the corrected spectrum with 0 in each bin. The measured energy spectrum shown here was generated using a 7 ns artificial dead time.

One can also determine the expected contamination from triple pileup. As just mentioned, double pileup appears at order  $r(t) \cdot \Delta t$ . The detector dead time  $\Delta t$  can be artificially increased during reconstruction, and doing so is useful for validating the pileup correction. For sufficiently large rates or dead times, triple pileup contamination—which appears at order  $r(t)^2 \Delta t^2$ —becomes important. The treatment of double pileup shown above assumes that all pulse pairs within the detector dead time of one another will yield a false count at the sum energy and the loss of a count at each of the two constituent pulse energies. This assumption is not valid in the case where there are three pulses all within the detector dead

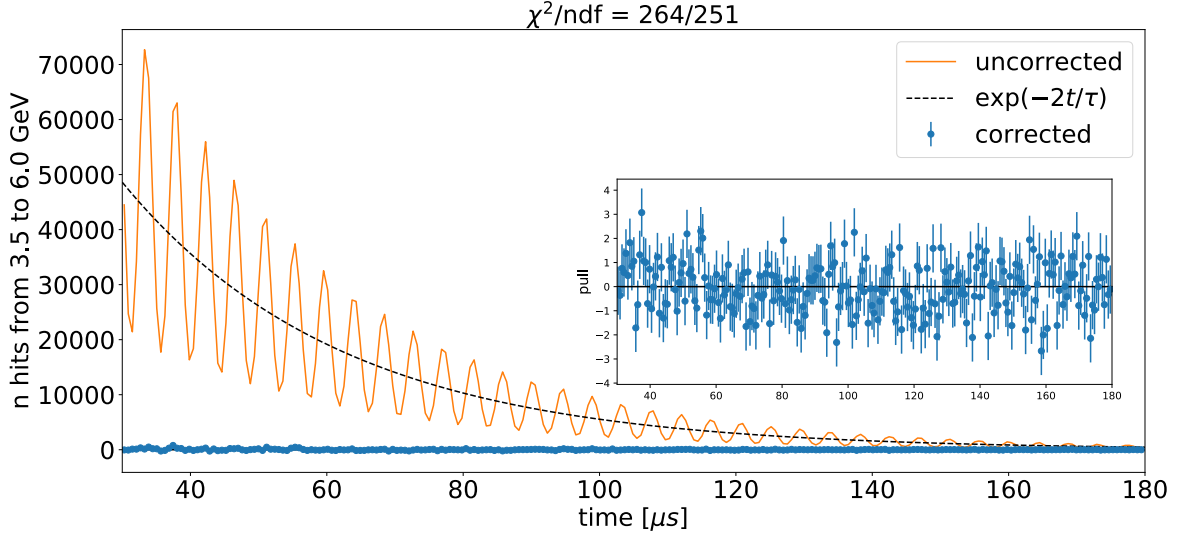


Figure 6.9: The sum of all counts between 3.5 GeV and 6 GeV in both the measured and pileup corrected energy spectra as a function of time. As with Figure 6.8, the associated reconstruction used a 7 ns artificial dead time. The pileup correction removes the non-physical tail of the energy spectrum at all times, indicating that it has the correct time structure. The dashed line is a guide to the eye. In general, the pileup spectrum decays at half the muon lifetime.

time. In the case of three pulses, one expects a loss of three true counts and a gain of one false count. The double pileup treatment, however, would count three pulse pairs and thus erroneously remove six true counts and add three false counts. A triple pileup correction must then account both for the reconstruction's treatment of groups of three unresolved pulses and for the error in the double pileup correction that occurs at the order of triple pileup.

For a more concrete example, consider the case in which three hits with respective energies  $E_1$ ,  $E_2$ , and  $E_3$  occur at once. There are three pairs:  $(E_1, E_2)$ ,  $(E_2, E_3)$ , and  $(E_1, E_3)$ . The double pileup correction would therefore subtract counts at each pulse energy twice and add counts at the sum energies  $E_1 + E_2$ ,  $E_2 + E_3$ , and  $E_1 + E_3$ . What one would reasonably expect to occur in the reconstruction, however, is the loss of a single count at each of the energies  $E_1$ ,  $E_2$ , and  $E_3$ , and the gain of a false count at the energy  $E_1 + E_2 + E_3$ . Considering that

the triple pileup correction and the double pileup correction are applied in concert, the triple pileup correction must add a count at  $E_1 + E_2 + E_3$ , subtract the counts at  $E_1 + E_2$ ,  $E_2 + E_3$ , and  $E_1 + E_3$  that were mistakenly added by the double pileup correction, and finally add a count at each of  $E_1$ ,  $E_2$ , and  $E_3$ . Applying both the double and triple pileup corrections to this trio of pulses then leads to the desired behavior of a single subtraction at each pulse energy and a single addition at the sum of the three.

Defining

$$\rho_t(E_1, E_2, E_3, t) = \rho(E_1, t) \cdot \rho(E_2, t) \cdot \rho(E_3, t),$$

the above reasoning suggests a triple pileup correction,  $\delta\rho_{pu,t}$ , of the form

$$\begin{aligned} \delta\rho_{pu,t}(E, t) \propto \Delta t^2 \bigg[ & \iint \rho_t(E - E_2 - E_3, E_2, E_3, t) dE_2 dE_3 \\ & - 3 \iint \rho_t(E - E_2, E_2, E_3, t) dE_2 dE_3 \\ & + 3 \iint \rho_t(E, E_2, E_3, t) dE_2 dE_3 \bigg]. \end{aligned} \quad (6.14)$$

This can be rewritten in terms of the double pulse sum term,  $\rho_{d+}$ , as follows

$$\begin{aligned} \delta\rho_{pu,t}(E, t) \propto \Delta t^2 \bigg[ & \int \rho(E - E_d) \cdot \rho_{d+}(E_d, t) dE_d \\ & - 3\rho_{d+}(E, t) \cdot \int \rho(E_3, t) dE_3 \\ & + 3\rho(E, t) \cdot \left( \int \rho(E_2, t) dE_2 \right)^2 \bigg]. \end{aligned} \quad (6.15)$$

The same argument given for the double pileup contamination equation applies to the triple: if in Equation 6.14 one were to replace  $\rho$  with  $[\rho + r(t)\Delta t \cdot f(t, E)]$ , the result would only change by order  $r(t)^3 \cdot \Delta t^3$ . For an order one percent pileup contamination, the error in the triple pileup correction stemming from using the pileup contaminated spectrum rather than the true spectrum is of order  $10^{-6}$ . It is not reasonable to go beyond triple pileup as the approximations implicit in this procedure are expected to yield relative errors larger than  $10^{-6}$ .

The groundwork necessary to present the pileup correction is now complete. The steps for correcting an individual calorimeter's hit spectrum are as follows:

1. Take the measured calorimeter hit distribution,  $N(E, t)$ , as an estimate of  $\rho(E, t)$ .
2. Use Equations 6.13 and 6.15 to generate quantities proportional to the expected double and triple pileup contaminations,  $\delta\rho_{pu,d}(E, t)$  and  $\delta\rho_{pu,t}(E, t)$ . As  $N(E, t)$  is itself pileup contaminated, based on the arguments above these two quantities will only be correct to order  $r(t) \cdot \Delta t$  and  $r(t)^2 \cdot \Delta t^2$ , respectively.
3. Fit the nonphysical tail of the observed hit spectrum above 3.5 GeV to a linear combination of the double and triple pileup contaminations just obtained. Doing so fixes their normalizations. This step is justified because the endpoint of the true energy spectrum is approximately 3.1 GeV, so any hits above 3.5 GeV must come purely from pileup contamination.
4. Generate a corrected hit spectrum  $N_c(E, t) = N(E, t) - \delta\rho_{pu,d}(E, t) - \delta\rho_{pu,t}(E, t)$ . This is expected to be correct to order  $r(t) \cdot \Delta t$ .
5. Go back to Step 2, except this time use the first-pass corrected spectrum,  $N_c(E, t)$ , to generate the expected pileup contaminations. This yields a second-pass spectrum correct to order  $r(t)^2 \cdot \Delta t^2$ .

$N(E, t)$  for each calorimeter is in practice a finely binned two-dimensional histogram. In this analysis, time bins were approximately 25 ns wide—precisely  $2\pi/(6\omega_c)$ —and energy bins were 20 MeV wide. Integrals were calculated numerically according to the bin contents of  $N(E, t)$ . The pileup correction described above adjusts the contents of each bin. Following the pileup correction, the second-pass corrected histograms are rebinned, projected, and summed over calorimeters to produce T-Method and other analysis histograms. The procedure described here captures the time and energy dependence of the pileup spectrum, see Figure 6.8 and Figure 6.9.

Following the pileup correction, analysis histogram bins no longer obey Poisson statistics. Each bin uncertainty must be increased by an amount of the same order as the relative pileup

contamination in that bin. Appendix A contains a complete calculation of the bin uncertainty enhancement factors arising from the pileup correction described in this section. Including them does not affect the  $\omega_a$  value extracted from the 60-Hour Dataset.

#### 6.5.4 The Gain Correction

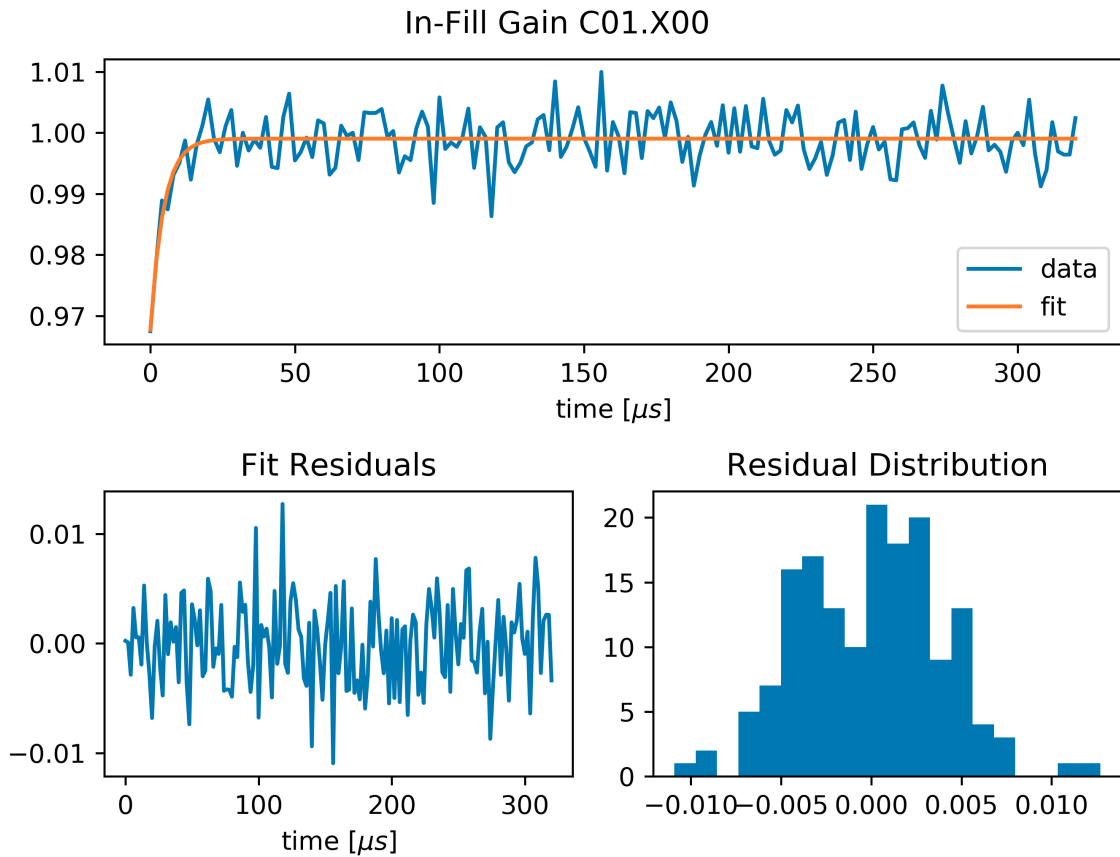


Figure 6.10: Representative in-fill gain measurement from 60-Hour Dataset for a single crystal, courtesy of Matthias Smith. The large flash from the beam injection is responsible for the gain reduction at early times. The flash is largest at calorimeters near the injection site and crystals closest to the storage ring, so an individualized gain recovery curve is necessary for each of the 1296 calorimeter channels. The recovery time constants vary but are all less than  $10\ \mu\text{s}$ . The gain reduction from the flash is therefore very small—a few per mille for the most perturbed crystals—at the beginning of the  $\omega_a$  measurement period, which is  $30.2\ \mu\text{s}$  after the beam injection.

Long-term and in-fill gain corrections are implemented following the procedures described in Section 4.1.3. Details of the in-fill gain correction are available in a series of presentations and notes; a recent example by Matthias Smith is cited here [90]. Figure 6.10 shows a representative in-fill gain curve for a single crystal, as measured by the laser calibration system.

In the 60-Hour Dataset, the in-fill laser pulses were sufficient to measure any given calorimeter channel's gain to a relative precision of  $6 \times 10^{-4}$  at each sampled time-in-fill. In one out of every 11 fills, the laser system delivered two pulses separated by  $200 \mu\text{s}$ . Each time the laser fired, the time of these two pulses relative to the beam injection was adjusted. Through this procedure, the laser pulses evenly sampled the time between the beam injection and  $400 \mu\text{s}$  after the injection with a grid spacing of  $2.5 \mu\text{s}$ . Fills with laser pulses in the measurement window were not included in the  $\omega_a$  analysis.

The in-fill laser pulses revealed the clear signature of a hardware gain reduction following the beam injection. This reduction was 1% to 3% for the crystals closest to the beam injection site. The gain recovery time constant was less than  $10 \mu\text{s}$  for all crystals, and thus the perturbation was almost entirely gone by the beginning of the  $\omega_a$  measurement period,  $30.2 \mu\text{s}$ . Nevertheless, correcting for the gain perturbation following the beam injection is important to achieve the target precision of the E989 experiment. Gain recovery curves like the one shown in Figure 6.10 were constructed for each crystal and fit with an exponential recovery function. These recovery functions were then used during reconstruction to apply per-hit, time-dependent gain adjustments, effectively removing the influence of the hardware gain reduction induced by the beam injection flash.

Work toward an additional laser-based gain correction is currently in progress. While the in-fill laser pulses enable a correction for flash-induced gain perturbations, in E989 Run 1 they were not sensitive to rate-dependent gain perturbations. Rate-dependent gain perturbations will be measured through dedicated short-time double-pulse (STDP) studies, wherein two laser pulses are provided to the same calorimeter channel with a variable time separation on the order of  $10 \text{ ns}$  [91]. The STDP data will be analyzed to determine the characteristic gain

reduction of each calorimeter channel following a pulse as a function of the energy of the pulse and the time since the pulse. As with the in-fill correction, the information obtained from the STDP studies will be used to correct the reconstructed data in software. The STDP correction was not available for the data analysis presented here. Once the STDP analysis is complete, the 60-Hour Dataset—along with all other E989 Run 1 subdatasets—will be reprocessed to apply the resulting correction. Systematic uncertainties from the neglected STDP correction will be discussed in Section 6.7.4.

## 6.6 Extraction of $\omega_a$

### 6.6.1 T-Method Fit to the 60-Hour Dataset

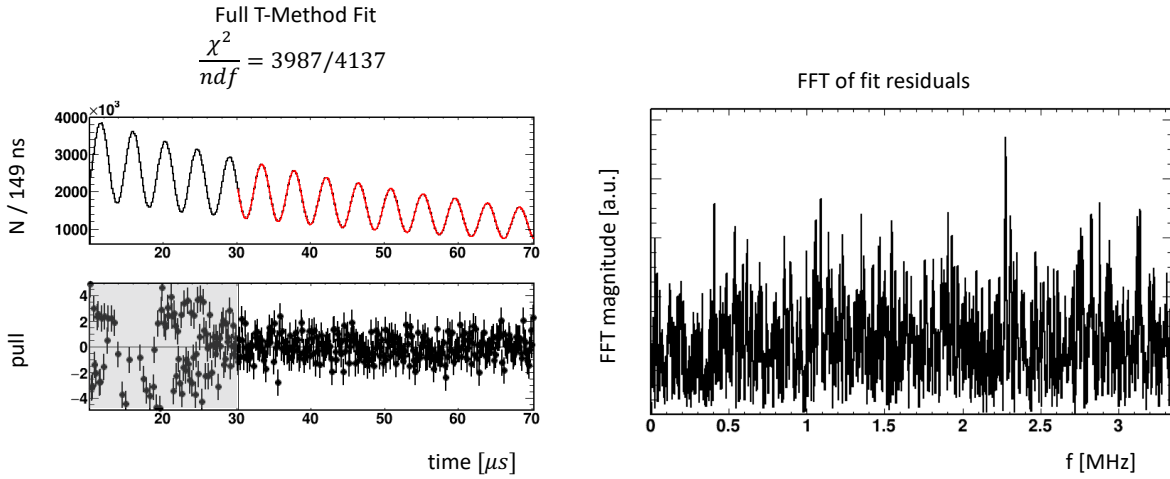


Figure 6.11: Full T-Method fit (Equation 6.10) to the 60-Hour Dataset corrected for pileup and gain. The vertical scale of the Fourier spectrum shown on the right is a factor of 12 smaller than the one shown in Figure 6.4. Within the measurement window, the pulls follow a standard normal distribution, and the Fourier spectrum of the residuals no longer displays any large peaks. The pulls begin to quickly diverge from 0 in the grayed out region before the measurement window, during which the quadrupoles are relaxing from their scraping configuration to their storage configuration.

Following the pileup and gain corrections, a T-Method histogram with a 1.73 GeV threshold was constructed and fit with the model given in Equation 6.10. A 2.5 ns deadtime was



applied in the reconstruction. Figure 6.11 shows the fit residuals in the time and frequency domain as Figure 6.4 did for the uncorrected five-parameter fit. No clear structures are evident. The best-fit parameters, their uncertainties, and their correlations with  $R$  as determined from the fit are presented in Table 6.2.

The T-Method analysis yielded an  $\omega_a$  value with 1.3 ppm statistical uncertainty. This is in good agreement with the expected T-Method precision for the 60-Hour Dataset calculated using Equation 2.20. The fit reveals that, when all calorimeters are combined, the coherent betatron oscillations impart a 0.4% oscillation to the acceptance at  $t = 0$ , and the vertical waist imparts a 0.8% oscillation to the acceptance at  $t = 0$ . The decoherence lifetimes are of the same order—though a bit smaller—than the back-of-the-envelope predictions in Section 6.5.1. With an 82.7% correlation,  $\phi$  is the only parameter strongly coupled to  $R$  in the  $\chi^2$  minimization. The next largest  $R$  correlation coefficient is 3.5% with the  $A_{CBO,N}$  parameter.

As discussed in Chapter 3, effects that cause the  $g - 2$  phase to change with time-in-fill bias the extracted  $\omega_a$  value. Pileup and gain perturbations are two possible sources of a time-dependent phase. Effects such as these are also liable to create a time dependence of the  $g - 2$  asymmetry. The asymmetry varies largely with positron energy, and thus anything that effectively creates a time-dependent threshold or causes the erroneous inclusion of low-energy positrons will affect the asymmetry. Varying the start time of the  $\omega_a$  fit allows one to search for unexpected time dependence of any fit parameters that were assumed to be constant.

The idea behind a start time scan is the following: as the fit model assumes that each parameter is constant, any observed parameter drift with fit start time is indicative of some effect that has not been accounted for. This process is slightly complicated by the allowed statistical drifts of fit parameters as data is removed. Later start times will have larger statistical parameter uncertainties. For values  $p_1$  and  $p_2$  representing a parameter  $p$  measured at an early start time  $t_1$  and a later start time  $t_2$ , the one standard deviation allowed drift

Table 6.2: Results of the T-Method fit to the 60-Hour Dataset.

Parameter	Meaning	Value	Correlation with R [%]
$R$	Blinded $\omega_a$	$-43.3 \pm 1.3$	100
$N_0$	Normalization	$(3.406 \pm 0.003) \times 10^6$	-0.4
$\tau$	Muon lifetime	$(64.438 \pm 0.004) \mu\text{s}$	-0.3
$A$	$g - 2$ asymmetry	$(37.587 \pm 0.004)\%$	0.5
$\phi$	$g - 2$ phase	$(-2.0909 \pm 0.0002) \text{ rad}$	82.7
$\omega_{CBO,0}$	CBO frequency	$(2.3048 \pm 0.0004) \text{ rad}/\mu\text{s}$	-1.2
$\tau_{CBO}$	CBO decoherence time	$(170 \pm 10) \mu\text{s}$	-2.6
$A_{CBO,N}$	CBO $N_0$ modulation	$(0.45 \pm 0.02)\%$	3.5
$\phi_{CBO,N}$	Phase of CBO $N_0$ modulation	$(1.80 \pm 0.03) \text{ rad}$	-1.7
$A_{CBO,A}$	CBO $A$ modulation	$(0.04 \pm 0.03)\%$	-2.3
$\phi_{CBO,A}$	Phase of CBO $A$ modulation	$(0.5 \pm 0.7) \text{ rad}$	1.3
$A_{CBO,\phi}$	CBO $\phi$ modulation	$(0.4 \pm 0.3) \text{ mrad}$	1.0
$\phi_{CBO,\phi}$	Phase of CBO $\phi$ modulation	$(-2.0 \pm 0.6) \text{ rad}$	0.5
$\omega_{VW}$	VW frequency	$(14.39 \pm 0.01) \text{ rad}/\mu\text{s}$	-1.1
$\tau_{VW}$	VW decoherence lifetime	$(20 \pm 4) \mu\text{s}$	0.9
$A_{VW}$	VW $N_0$ modulation	$(0.8 \pm 0.3)\%$	-0.9
$\phi_{VW}$	Phase of VW $N_0$ modulation	$(0.4 \pm 0.4) \text{ rad}$	-1.2
$K_{loss}$	Muon loss correction amplitude	$6.5 \pm 0.3$	-0.4

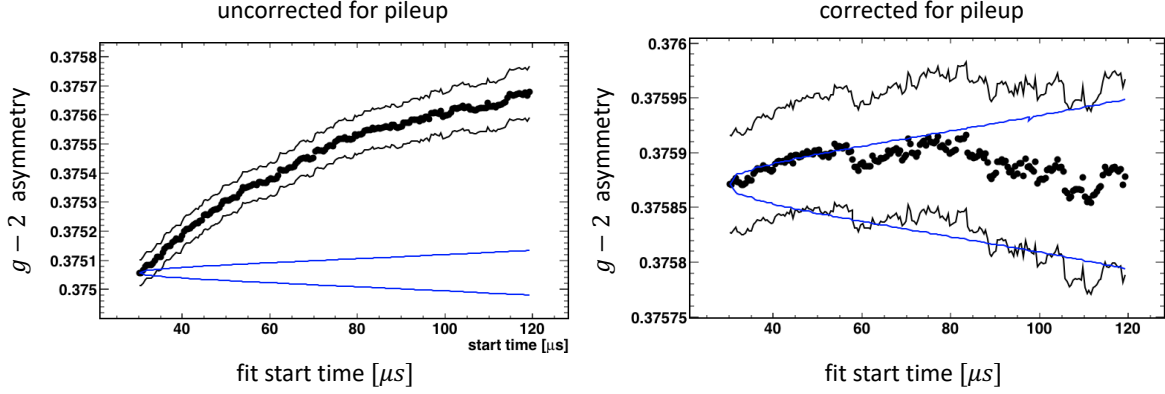


Figure 6.12: Asymmetry versus fit start time with and without the pileup correction. The blue curves indicate the statistical  $1\sigma$  allowed drifts from the starting value. The black curves delimit the  $1\sigma$  statistical parameter uncertainties obtained from each fit start time. Without the pileup correction, the asymmetry starts low and then climbs toward the value from the pileup corrected fit. It quickly departs from the allowed band and ends up quite far from its starting value. This is consistent with the expectation from pileup, specifically that low-energy positrons with small  $g - 2$  asymmetries contaminate the T-Method histogram at early times. In the case with the pileup correction, the asymmetry remains mostly within the  $\pm 1\sigma$  band, making a brief excursion and then returning.

from  $p_1$  to  $p_2$  is approximately given by [34, 92]

$$\sigma_{p_2-p_1}^2 = \sigma_{p_2}^2 - \sigma_{p_1}^2. \quad (6.16)$$

In a start time scan,  $\sigma_{p_2}^2$  will be larger than  $\sigma_{p_1}^2$  because  $p_2$  is derived from a smaller dataset than is  $p_1$ . Note that all data used to extract  $p_2$  is contained in the dataset used to extract  $p_1$ .

In the start time scans presented here, all parameters except  $N_0$ ,  $\tau$ ,  $A$ ,  $\phi$ , and  $R$  are fixed to their best-fit values from the  $30.2\mu\text{s}$  start time fit. See Figure 6.12 for an illustration of the usefulness of this technique. Even though the pileup bias is much less than the statistical uncertainty in the 60-Hour Dataset (this will be discussed more in Section 6.7), the asymmetry versus start time scan clearly indicates a failure of the fit model when there is no pileup correction. Were pileup unanticipated, this start time scan would have clearly revealed its presence. Similarly, in the absence of the muon loss correction,  $N_0$  vs start time

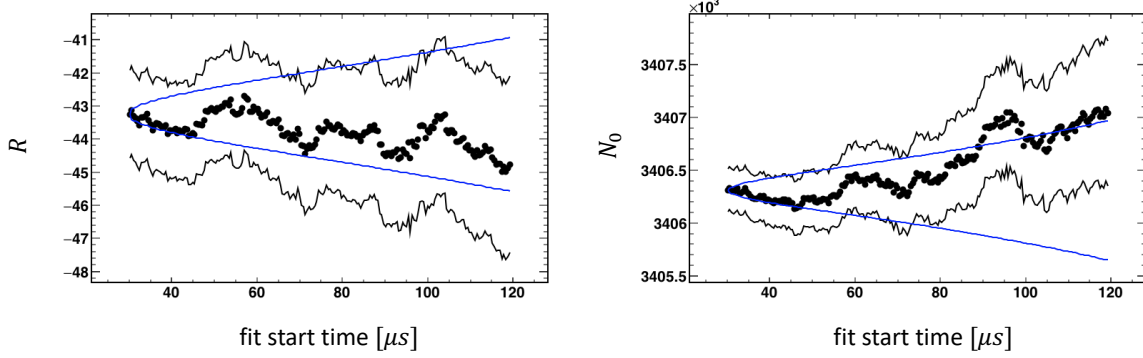


Figure 6.13:  $R$  and  $N_0$  versus fit start time.  $R$  remains within or at the border of the  $1\sigma$  band for all sampled start times.  $N_0$  ends barely outside of the  $1\sigma$  band.

quickly diverges from its starting value. The power of start time scans to uncover systematic effects increases with the statistical precision of the dataset. Figure 6.13 shows the behavior of  $R$  and  $N_0$  versus start time. Both appear consistent with purely statistical variations. As  $\phi$  and  $R$  are highly correlated,  $\phi$  versus start time looks very similar to  $R$  versus start time, and  $\tau$  versus start time looks very similar to  $R$  versus start time. They are therefore not presented here.

### 6.6.2 Energy-Binned Analysis of the 60-Hour Dataset

Section 2.6 described the energy-binned  $\omega_a$  analysis technique, in which a family of hit time histograms partitioned by positron energy are constructed and fit individually. Averaging the resulting  $\omega_a$  values yields a combined measurement with improved precision compared to the T-Method analysis. The full theoretical statistical gain of an energy-binned analysis cannot be realized in the 60-Hour dataset because, after positrons hits have been separated by energy, there are not enough counts in late time bins to fit over the full measurement window used in the T-Method analysis. In the energy-binned analysis presented here, fits extended from  $30.2 \mu\text{s}$  to  $400 \mu\text{s}$ .

Positrons were sorted into 40 energy bins from 500 MeV to 2900 MeV. Each energy bin

was thus 60 MeV wide. As in the T-Method analysis, the time bin width was matched to the cyclotron period extracted from the fast rotation analysis. The parameters  $\omega_{CBO}$ ,  $\tau_{CBO}$ ,  $\omega_{VW}$ , and  $\tau_{VW}$  were fixed to their best-fit values from the T-Method analysis. Each of the 40 histograms was then fit with Equation 6.10. Figure 6.14 shows that the energy-binned fits are well behaved and reproduce the expected qualitative shapes for  $N(E)$  and  $A(E)$ . Additionally, it shows the  $\omega_a$  statistical analyzing power carried by positrons of different energies. Information regarding the muon spin direction is carried predominantly by high-energy positrons above 2 GeV, as expected. These are also the energies where detector effects such as gain and pileup have the largest impact.

It is important to verify that consistent  $\omega_a$  values are obtained at each positron energy. The detector systematic effects that have been discussed at length cause energy and time-dependent phase shifts that bias  $\omega_a$  differently at different energies. Uncorrected detector-based systematic effects are expected to create biases on the order of 100 ppb, which is much smaller than the statistical uncertainty of an  $\omega_a$  measurement in any particular energy bin—even with the expected statistics of the ultimate combined E989 dataset. Thus, consistency of  $\omega_a$  versus energy is necessary but not sufficient for verifying the correct treatment of systematic effects. Beam dynamics effects can also cause energy dependent biases via the energy dependence of the calorimeter acceptance. Figure 6.15 shows that the per-energy  $\omega_a$  values extracted from the 60-Hour Dataset are consistent. There is a hint that there may be an effect increasing the measured  $R$  values near 2.6 GeV. However, there is no known effect that could bias  $\omega_a$  in only a narrow energy range. Analysis of further datasets will reveal whether the 2.6 GeV increase is real or a statistical artifact.

Figure 6.15 also displays the  $g - 2$  phase as a function of positron energy. This is an important input to systematic error studies like the one presented in Section 3.5. Chapter 3 discussed the implications of the energy dependence of the  $g - 2$  phase. One source of the phase's energy dependence is the positron drift time (see Figure 3.1). The shape of the phase versus energy extracted from the data does not match that of the drift time versus energy, and thus one must conclude there are other equally or even more important effects.

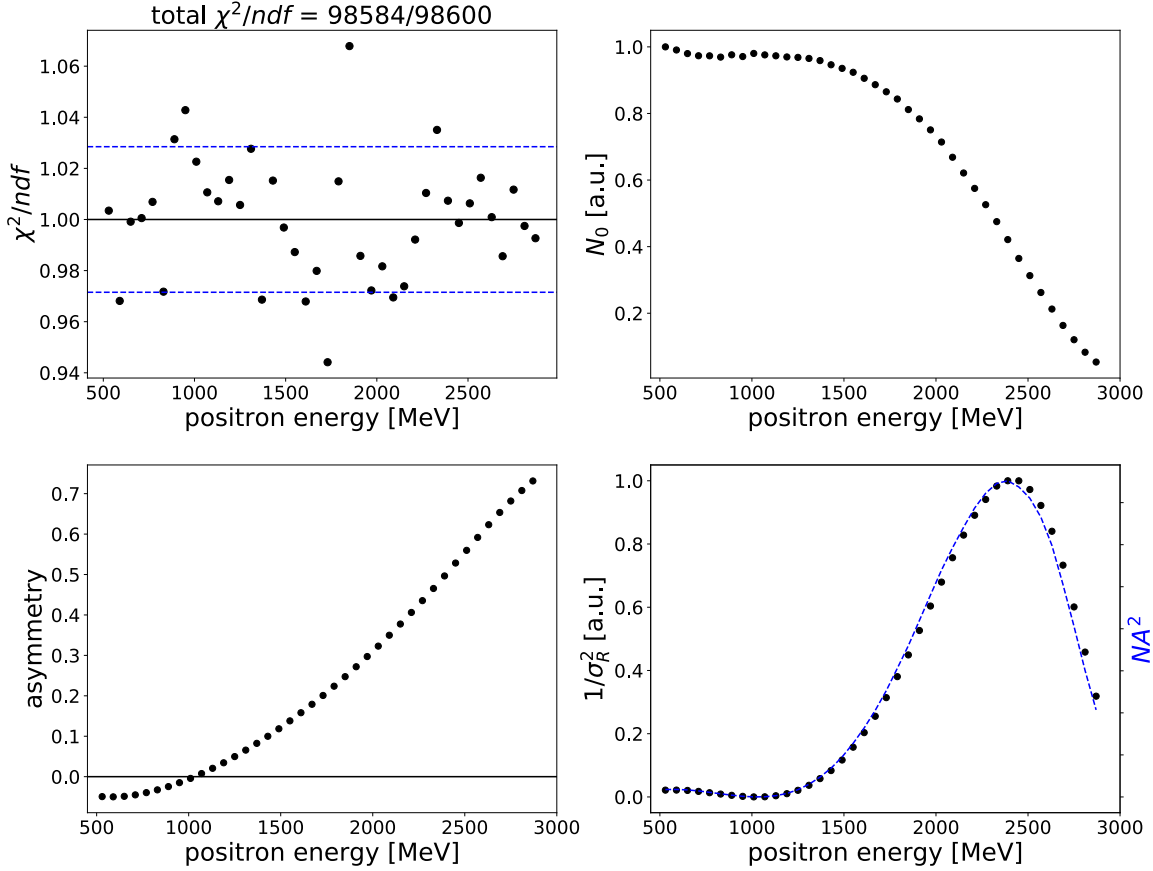


Figure 6.14:  $\chi^2$ ,  $N$ , and  $A$  versus energy as extracted from the energy-binned fits. Good fits were achieved in all energy bins. The total number of degrees of freedom and total  $\chi^2$  displayed on the top left are the sums of the respective quantities from each energy bin. The total  $\chi^2$  is within one standard deviation of the expected mean, assuming a  $\chi^2$  distribution with a number degrees of freedom equal to the total over all the energy-binned fits. The normalization and asymmetry follow the expected qualitative patterns, specifically  $N_0$  decreases with energy while the asymmetry increases at high energies and crosses zero at approximately 1 GeV. The bottom right plot shows the statistical power of the  $\omega_a$  extraction in each energy bin. It closely follows  $NA^2$ , shown by the blue dashed curve. The actual statistical power tends toward higher asymmetries than predicted by the leading order  $NA^2$  approximation.

One such effect is that the calorimeter acceptance as a function of energy (also shown in Figure 3.1) is not time independent: it, like the positron energy distribution, oscillates at

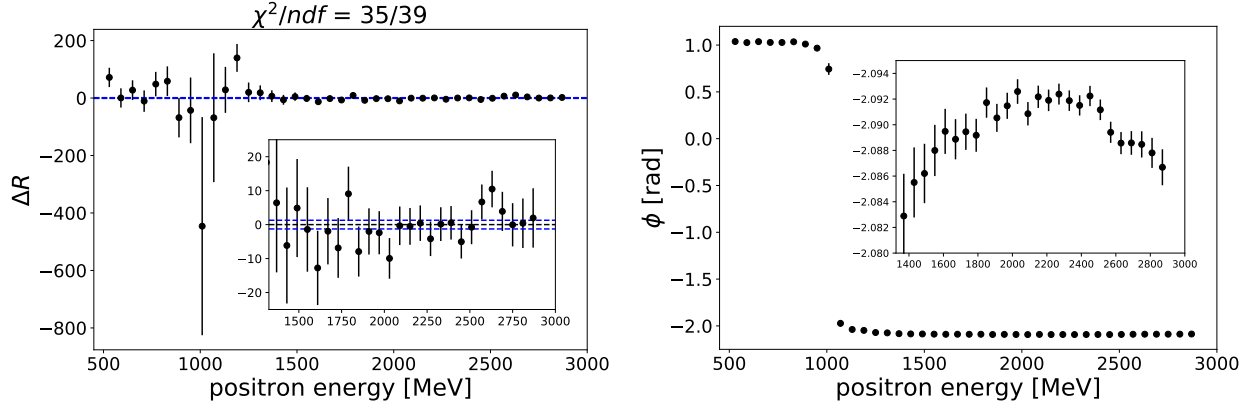


Figure 6.15: Blinded  $\omega_a$  and  $\phi$  versus positron energy.  $R$  values are shown relative to the average. The blue dashed lines are drawn at plus and minus one standard deviation of the mean. The  $\chi^2$  indicates that the deviations from the mean are consistent with a purely statistical variation. The uncertainties become very large at the energies where the  $g - 2$  asymmetry is close to zero. The plot on the right indicates that the  $g - 2$  phase differs by less than 10 mrad over the range from 1.7 GeV to 3 GeV.

$\omega_a$ . This acceptance oscillation is created by the changing angular distribution of decay positrons following the rotation of the stored muon population's average spin direction. The energy-binned analysis of the 60-Hour Dataset revealed that the  $g - 2$  phase changes less with energy than one would expect solely from considering the drift time: the phase changes by less than 10 mrad from 1.7 GeV to 3 GeV, whereas a drift-time-based calculation would suggest a change of 20 mrad to 30 mrad. Therefore, acceptance-based effects must partially counteract the drift time effect.

Energies near the spectrum's endpoint are exceptionally sensitive to gain and pileup effects. This is simply because the fractional impact of such effects is greatly enhanced by the naturally small number of counts in very high-energy bins. For example, a gain perturbation of  $10^{-3}$  will impart a 2% distortion to the energy spectrum at 2.7 GeV. This increased sensitivity is beneficial in that it provides a powerful handle on systematic effects: it is easy to see a 2% shift in the energy spectrum, but difficult to directly measure a  $10^{-3}$  gain reduction. Figure 6.16 shows that there does appear to be an unaccounted-for effect

perturbing the high-energy bins. Both the  $\tau$  and  $K_{loss}$  parameters decrease at the highest energies, while in the absence of any systematic effects these parameters should be consistent across energy bins. This is indicative of an uncorrected gain-like effect reducing the number of counts in the high-energy bins at early times. The implications of such an effect will be discussed in Section 6.7.

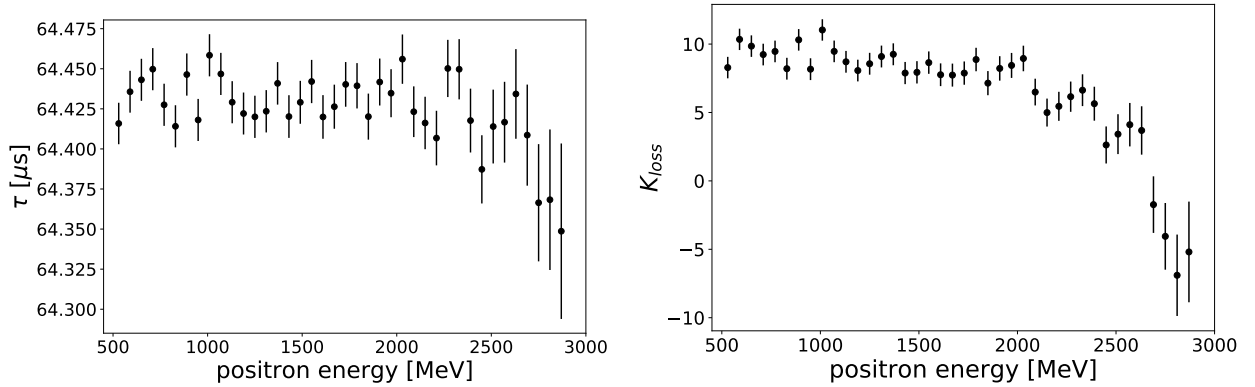


Figure 6.16:  $\tau$  and  $K_{loss}$  versus positron energy. The reduction of both at high energies is an indication of an imperfect gain correction or a similar energy-dependent effect.

While the energy-binned analysis provides invaluable information, it is not the ideal choice for analyzing the 60-Hour Dataset. For a given measurement window, an asymmetry-weighted analysis will provide virtually the same statistical power as an energy-binned analysis. Furthermore, in an asymmetry-weighted analysis one can combine hits from all desired energy bins and calorimeters into a single histogram, allowing for a comfortable extension of the measurement window into times where any single energy-binned histogram would have very few counts. Such an analysis is presented next.

### 6.6.3 Asymmetry-Weighted Analysis of the 60-Hour Dataset

The asymmetry-weighted (A-Weighted) analysis technique is similar to the T-Method except that instead of incrementing histogram bin counts by one for each over threshold positron, bin values are increased by the energy-dependent weight  $A(E)$ , where  $A$  is the  $g - 2$  asymmetry



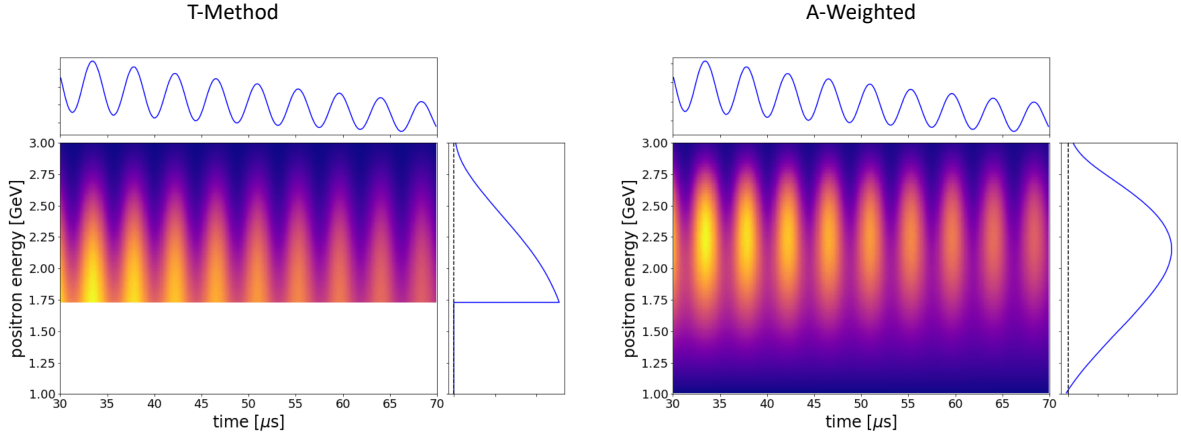


Figure 6.17: Relative weighting of positron energies in the T-Method and A-Weighted analyses. Each central plot shows the all-calorimeter energy and time distribution from the 60-Hour dataset scaled by the energy dependent weights used in the T-Method and A-Weighted analyses. In the T-Method, the weights are 0 for any positron under 1.73 GeV and 1 otherwise. In the A-Weighted analysis, the weights are equal to  $A(E)$  for positrons between 1 GeV and 3 GeV and 0 otherwise. On top of each distribution is the sum over energies in each time bin, the histogram that is ultimately fit for  $\omega_a$ . On the right of each is the sum over time bins for each energy, which illustrates the relative contributions of different positron energies in each analysis technique. The A-Weighted analysis gives additional weight to high-energy positrons.

and  $E$  is the energy of the detected positron. As discussed in Section 2.6, such an analysis provides the best precision achievable from a single histogram constructed with energy-dependent weights.

An A-Weighted histogram was constructed from the finely-binned energy and time distribution provided by the pileup correction procedure (Section 6.5.3). The asymmetry at the center of each 20 MeV energy bin was evaluated through interpolation of the empirical  $A(E)$  curve obtained from the energy-binned analysis (Figure 6.14). After scaling each bin content by the appropriate asymmetry value, the distribution was summed across all calorimeters and over energies between 1 GeV and 3 GeV. The low end of this energy range was selected to avoid potential contamination from noise pulses or the tail of the lost muon energy peak, and the high end was selected because no meaningful asymmetry can be defined beyond

the energy spectrum end point of 3 GeV. See Appendix A for a discussion of A-Weighted histogram bin uncertainty assignments.

As the  $g - 2$  asymmetry increases with energy, the A-Weighted histogram preferentially weights high-energy positrons. This is the source of the A-Weighted procedure's improved precision relative to the T-Method. The preferential weighting of high-energy positrons also increases the impact of gain perturbations and pileup, which are both most prevalent near the endpoint of the energy spectrum. Therefore, as with the high-energy bins discussed in the previous section, fit stability in the A-Weighted analysis is a good sign that detector effects have been correctly handled. See Figure 6.17 for an illustration of the relative contributions of positrons with different energies to the T-Method and the A-Weighted analyses.

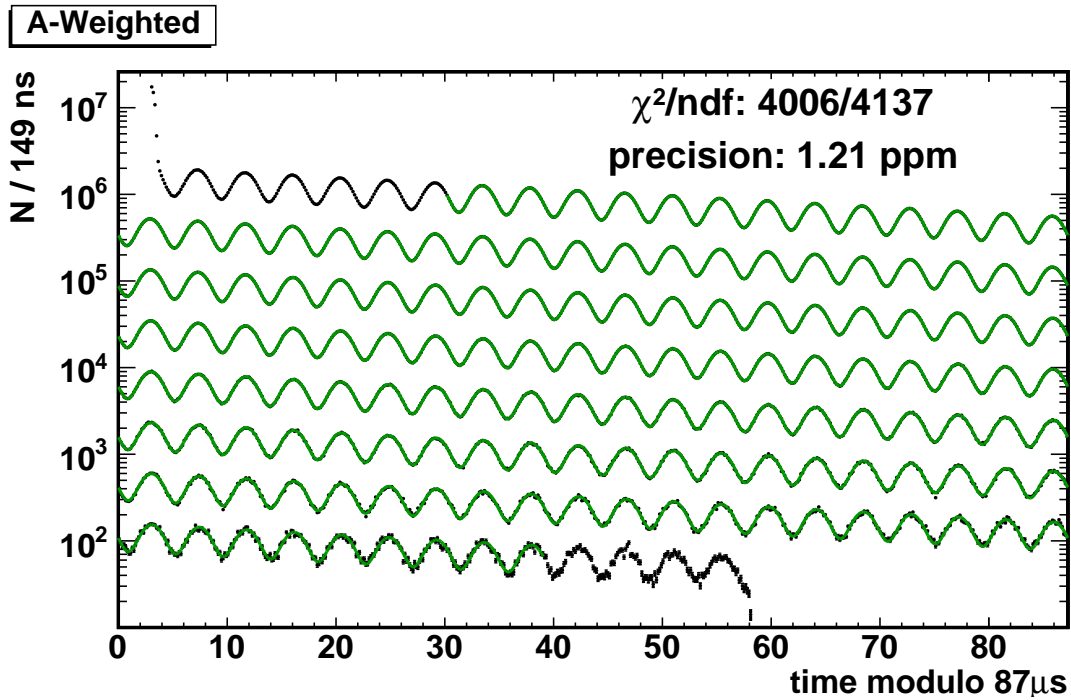


Figure 6.18: A-Weighted fit to the 60-Hour Dataset. The black points are the histogram bin values and the green curve is the fit function. Each of the stacked rows displays  $87 \mu\text{s}$  of data. Time increases with descending rows. The fit extends from  $30.2 \mu\text{s}$  to  $650 \mu\text{s}$ .

The A-Weighted histogram was fit with the same model as the T-Method histogram

(Figure 6.18). Best-fit values of energy dependent parameters—such as the  $g - 2$  phase—are expected to differ between the two approaches. Table 6.3 shows the obtained fit parameters and their correlations with  $R$ , as Table 6.2 did for the T-Method. The results are very similar. There are small changes in the amplitude and phase parameters, as expected from the different energy weighting. The statistical precision of the extracted  $\omega_a$  value improved by approximately 10% compared to the T-Method; it is the best achieved by any analysis technique applied to the 60-Hour Dataset. Start time scans and fit residuals in the time and frequency domains look by eye identical to those obtained from the T-Method analysis.

Analyses that employ different weighting schemes and data subsets produce  $R$  values that are permitted to differ somewhat from one another. Exactly how much they are allowed to differ is an active question within the E989 collaboration. A reasonable lower bound on the allowed difference comes from treating the analysis result with the larger uncertainty as having been derived from a subset of the data used for the result with the smaller uncertainty. In this case, Equation 6.16 would apply. This can only be a lower bound, because one can imagine a case in which two analysis techniques achieved the same precision but through the use of different data or different weighting schemes. In such a case, a full analysis of the correlations between the two techniques would be required to determine the allowed statistical difference between the obtained results. Furthermore, each analysis technique has different susceptibilities to detector-based systematic effects.

The T-Method analysis yielded a blinded  $R$  value of  $R_T = -43.25 \pm 1.33$ . The A-Weighted analysis gave  $R_A = -42.53 \pm 1.21$ . Using the simple approximation for a lower bound on the allowed difference,

$$\begin{aligned}\sigma_{R_T-R_A} &= \sqrt{1.33^2 + 1.21^2} \\ &= 0.55.\end{aligned}$$

The difference  $R_T - R_A$  is  $-0.72$ , approximately  $-1.3 \cdot \sigma_{R_T-R_A}$ . This is another manifestation of the feature already identified in Figure 6.14: in the 60-Hour Dataset,  $R$  appears higher when measured with positrons near 2.6 GeV than it does when measured with positrons at

Table 6.3: Results of the A-Weighted fit to the 60-Hour Dataset.

Parameter	Meaning	Value	Correlation with R [%]
$R$	Blinded $\omega_a$	$-42.5 \pm 1.2$	100
$N_0$	Normalization	$(1.596 \pm 0.001) \times 10^6$	-0.5
$\tau$	Muon lifetime	$(64.437 \pm 0.003) \mu s$	-0.3
$A$	$g - 2$ asymmetry	$(35.659 \pm 0.004)\%$	0.5
$\phi$	$g - 2$ phase	$(-2.0905 \pm 0.0002) \text{ rad}$	82.7
$\omega_{CBO,0}$	CBO frequency	$(2.3049 \pm 0.0003) \text{ rad}/\mu s$	-1.3
$\tau_{CBO}$	CBO decoherence time	$(180 \pm 10) \mu s$	-2.5
$A_{CBO,N}$	CBO $N_0$ modulation	$(0.47 \pm 0.01)\%$	3.4
$\phi_{CBO,N}$	Phase of CBO $N_0$ modulation	$(1.78 \pm 0.03) \text{ rad}$	-1.7
$A_{CBO,A}$	CBO $A$ modulation	$(0.09 \pm 0.02)\%$	-1.3
$\phi_{CBO,A}$	Phase of CBO $A$ modulation	$(1.0 \pm 0.3) \text{ rad}$	2.1
$A_{CBO,\phi}$	CBO $\phi$ modulation	$(0.3 \pm 0.2) \text{ mrad}$	0.9
$\phi_{CBO,\phi}$	Phase of CBO $\phi$ modulation	$(-1.9 \pm 0.7) \text{ rad}$	0.5
$\omega_{VW}$	VW frequency	$(14.39 \pm 0.01) \text{ rad}/\mu s$	-1
$\tau_{VW}$	VW decoherence lifetime	$(20 \pm 3) \mu s$	0.8
$A_{VW}$	VW $N_0$ modulation	$(1.0 \pm 0.3)\%$	-0.8
$\phi_{VW}$	Phase of VW $N_0$ modulation	$(0.5 \pm 0.4) \text{ rad}$	-1.1
$K_{loss}$	Muon loss correction amplitude	$6.4 \pm 0.2$	-0.5

lower energies. High-energy hits have a larger influence on the A-Weighted analysis than they do on the T-Method, and thus the A-Weighted analysis of the 60-Hour Dataset yields a higher  $R$  than does the T-Method analysis. This deviation is not particularly significant, and thus it is no cause for alarm unless it appears in future datasets as well.

#### 6.6.4 Per-Calorimeter Fits

The analyses presented thus far were conducted on histograms containing events from all calorimeters. It is also possible to conduct analyses on each calorimeter individually. Doing so allows for additional consistency checks. Certain parameters, such as  $R$ ,  $\tau$ , and  $K_{loss}$  are expected to be the same across all calorimeters. Conversely,  $A$ ,  $\phi$ ,  $N_0$ , and the beam oscillation parameters may differ based on imperfect energy calibrations and differing materials present between the calorimeters and the storage region, which can distort the observed energy spectra in a calorimeter-dependent way. This section presents the consistency of T-Method fit parameters between calorimeters.

Per calorimeter fits were conducted between  $30.2\ \mu\text{s}$  and  $400\ \mu\text{s}$ . Truncating the fit window was necessary to avoid low-count histogram bins at late times. An all-calorimeter analysis over the same truncated fit window was also conducted for the purposes of comparison, and the obtained  $R$  value agreed with the per-calorimeter average. The all-calorimeter fit over the truncated time region yielded an  $R$  value statistically consistent with the one obtained over the full measurement window extending to  $650\ \mu\text{s}$ . In the single-calorimeter fits, all parameters were allowed to freely vary except for  $\omega_{VW}$  and  $\tau_{VW}$ , which were fixed to the values obtained from the all-calorimeter fit.

The range of momenta stored by the  $g - 2$  ring corresponds to a range of cyclotron frequencies. Some residual effects from orbital motion of the beam remain even after binning positron hits by the average cyclotron period,  $T_c$ . These residual effects are almost entirely eliminated when summing hits across all calorimeters. However, each individual calorimeter is only able to observe a small slice of the storage ring, and thus the beam's orbital motion is more important when considering single calorimeter spectra. To reduce the impact of this

effect, when constructing per-calorimeter histograms a random, per-fill time offset sampled uniformly from the range  $(-T_c/2, T_c/2)$  was applied to all clusters at the histogramming stage. This random offset scrambles the phase of the cyclotron motion, eliminating its effect on the calorimeter fits.

Figure 6.19 shows the results of the per-calorimeter fits. Consistent  $R$ ,  $\tau$ , and  $K_{loss}$  values were obtained for all calorimeters. These values were also consistent with those obtained from the all-calorimeter fit described earlier. The parameter consistency between calorimeters validates the treatment of beam oscillations, particularly the CBO and the VW. The CBO acceptance modulation is a factor of 6 to 10 larger in single calorimeters than it is in the all-calorimeter sum. This is because the CBO phase difference between calorimeters—described in Section 3.3.5—creates a large cancellation when all calorimeter signals are combined.

## 6.7 Systematic Uncertainty Assessment

The E989  $g - 2$  Experiment was designed for a 100 ppb measurement of  $\omega_a$ . In this section, data-driven estimates of the major E989 systematic uncertainties will be presented. It is difficult to constrain these systematic uncertainties to the expected tens of parts per billion given the relatively large statistical uncertainty of the 60-Hour Dataset measurement. Thus, the estimates in this section should be taken as preliminary and subject to change following the analysis of larger future datasets. A number of systematic uncertainties are driven by the imperfect knowledge of beam properties, for instance the exact form of the beam oscillation decoherence envelopes. As more data is collected and more simulations are completed, knowledge of these properties will improve and the associated systematic uncertainties will naturally reduce in size.

### 6.7.1 Systematic Uncertainty from Beam Oscillations

Assuming that the models employed to fit the CBO and VW are correct, beam oscillations do not impart a systematic bias to  $\omega_a$ . Appendix B contains the full parameter correlation matrices for the T-Method and A-Weighted fits. The parameter of interest,  $R$ , does not have

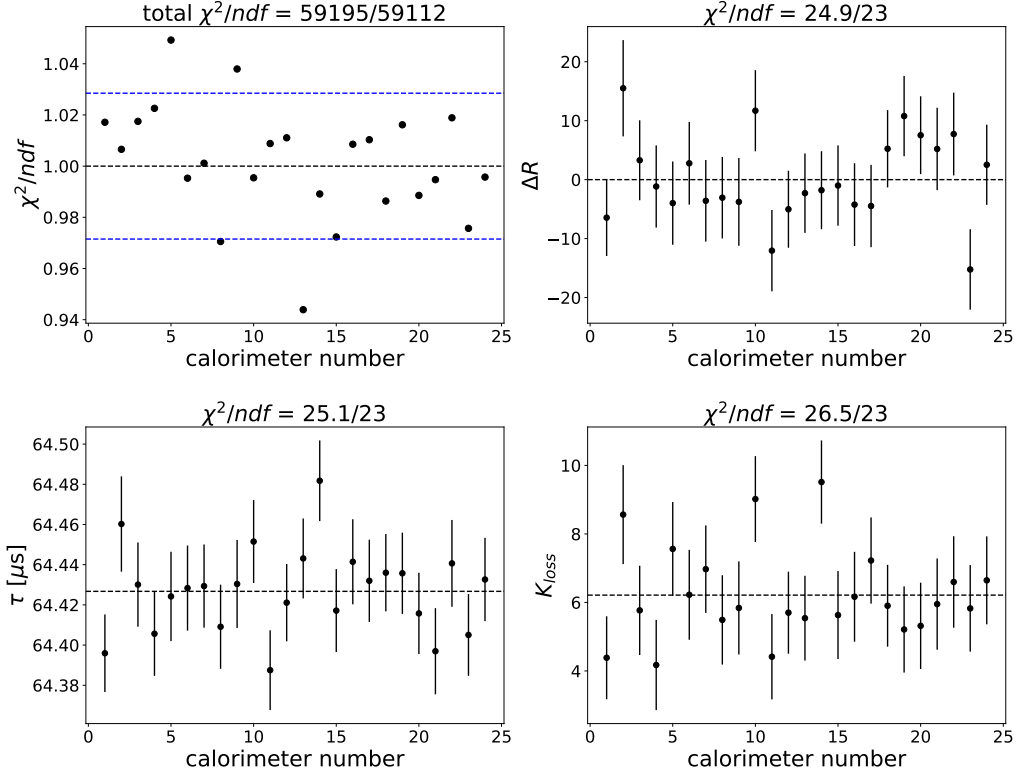


Figure 6.19:  $\chi^2/ndf$ ,  $R$ ,  $\tau$ , and  $K_{loss}$  by calorimeter. Good fits were achieved for all calorimeters. The total number of degrees of freedom and total  $\chi^2$  displayed on the top left are the sums of the respective quantities from each calorimeter. In the top right, the average  $R$  across all calorimeters has been subtracted from each point.  $\chi^2$  values displayed for each of the  $R$ ,  $\tau$ , and  $K_{loss}$  plots are calculated from a fit to a constant value. The dashed lines are the resulting best-fit constants.

a significant correlation with any of the beam oscillation parameters. Thus, the necessity to include beam oscillations in the fit does not even dilute the statistical power of the experiment. If one had perfect knowledge of all the beam oscillation parameters and needed only to fit for the original five parameters, the uncertainty of the extracted  $R$  would improve by less than one percent. This is truly a negligible contribution. The systematic uncertainty

arising from beam oscillations is determined by the degree of confidence in the applied fit model.

The CBO amplitude was assumed to reduce in size over time according to an exponential decay law. Figure 6.6 showed that this assumption was well motivated by tracker-based measurements of the stored beam oscillations. However, it is possible that the envelope differs from an exponential form in ways to which the tracker is insensitive. The exact form of the decoherence envelope depends on the stored muon momentum distribution and the nonlinearities of the quadrupole fields. In the future, the form may become highly constrained by further tracker analysis and beam dynamics simulations. Here, the sensitivity of the  $R$  to the CBO envelope is extracted from the calorimeter data, and the associated systematic uncertainty is determined by considering that sensitivity along with reasonable alternative models for the envelope.

The T-Method and A-Weighted analyses were conducted with two alternative models for the envelope:

$$A_{CBO}(t) = A \left( e^{-t/\tau_{CBO}} + C \right), \quad (6.17)$$

and

$$A_{CBO}(t) = A e^{-t/\tau_{CBO}} [1 + B \cos(2\pi t/T - \phi)]. \quad (6.18)$$

Each of these models describes the tracker data more or less equally well, and the second is motivated by an analytical beam dynamics evaluation [57, 86]. Attempts to use a Gaussian envelope yielded a considerably worse  $\chi^2$  for both the T-Method and A-Weighted fits, and thus a Gaussian envelope will not be considered as a viable alternative to the exponential envelope. Parameters of the above equations were taken from fits to the tracker data, except for  $A$  which was replaced by the  $A_{CBO}$  fit parameters. See Table 6.4 for a summary of these results.

For both the T-Method and A-Weighted analyses,  $R$  varied by tens of parts per billion when the assumed form of the CBO envelope was varied. Equation 6.17, an exponential recovery plus a constant baseline, yielded fits with the smallest  $\chi^2$ . The  $R$  values extracted



Table 6.4: Sensitivity of  $R$  to different CBO envelope functions.

CBO Envelope Function	T-Method		A-Weighted	
	$\chi^2$	$R$	$\chi^2$	$R$
$Ae^{-t/\tau_{CBO}}$	3987	-43.25	4006	-42.53
$A(e^{-t/\tau_{CBO}} + C)$	3982	-43.21	4002	-42.49
$Ae^{-t/\tau_{CBO}} [1 + B \cos(2\pi t/T - \phi)]$	3989	-43.27	4010	-42.55

with the different envelope forms spanned a range of 0.06. A change in  $R$  of 0.06 corresponds to a 60 ppb shift of  $\omega_a$ . The two alternative envelope functions moved  $R$  in opposite directions. As the data fit more-or-less equally well with all three functions, a systematic uncertainty of  $\pm 30$  ppb is assigned based on the imperfect knowledge of the CBO envelope function.

The parameters of the tracker-provided model of the changing CBO frequency varied slightly between the two tracking stations. To assess the sensitivity of  $R$  to the exact form of the changing CBO frequency, the T-Method and A-Weighted fits were repeated using the parameters from the two different stations.  $R$  was minimally affected, changing by 0.01 in both the T-Method and the A-Weighted analyses. The sensitivity to the CBO envelope function was the same regardless of the frequency model used. This is as expected from the covariance matrices shown in Appendix B: the CBO frequency is not highly correlated with the CBO amplitude or decoherence time constant.

The sensitivity of  $R$  to the VW oscillations must also be assessed. Given that the time-dependent CBO frequency has been attributed to the broken resistors in the quadrupole system, the VW frequency must also be time-dependent. This is because the VW frequency, like the CBO frequency, depends on the quadrupole field index. This has been neglected given the VW perturbations to the calorimeter hit spectra are already much smaller than the CBO perturbations, and the VW decoheres much more quickly. If the VW is entirely left out of the fit, the extracted  $R$  value changes by only 0.05, though the  $\chi^2$  values are significantly

increased. As an adjustment of either  $\omega_{VW}$  or  $A_{VW}$  from its best-fit value all the way to zero effects such a small change in  $R$ , there is no way that percent-level changes or other small perturbations could be significant. Thus, the systematic uncertainty of  $\omega_a$  originating from incorrect modeling of the VW will be considered negligible.

Finally, potential acceptance effects at  $2\omega_{CBO}$  will be discussed. Such effects were entirely left out of the fit model but are expected to be present at some level, originating from oscillations in the width of the stored beam. No component at  $2\omega_{CBO}$  is visible in the Fourier spectrum of the all calorimeter fit residuals. However, a peak is visible in the Fourier spectrum of the residuals of some, but not all, single calorimeters fits. Cancellation in the all-calorimeter spectrum is expected owing to the phase difference of the beam oscillations seen by detectors at different azimuthal locations. Generally, when fitting all calorimeters individually and then averaging the resulting  $R$  values, a result consistent with that of the all-calorimeter analysis is obtained—provided the fits are all over the same measurement window. The per-calorimeter average  $R$  varies by .01 when  $2\omega_{CBO}$  terms are included compared to when they are not. Thus, the effects of  $2\omega_{CBO}$  oscillations on the all-calorimeter fit are expected to appear at the 10 ppb level. As they were not included in the all-calorimeter fit, a systematic uncertainty of 10 ppb is assigned. Table 6.5 summarizes the  $\omega_a$  systematic uncertainties arising from imperfect knowledge of the beam oscillations. They are the same for the A-Weighted and T-Method analyses.

Table 6.5:  $\omega_a$  systematic errors from beam oscillations.

Source	Value
Form of the CBO envelope	30 ppb
Form of the changing CBO frequency	10 ppb
Oscillations at $2\omega_{CBO}$	10 ppb
Quadrature sum	33 ppb

### 6.7.2 Systematic Uncertainty from Pileup

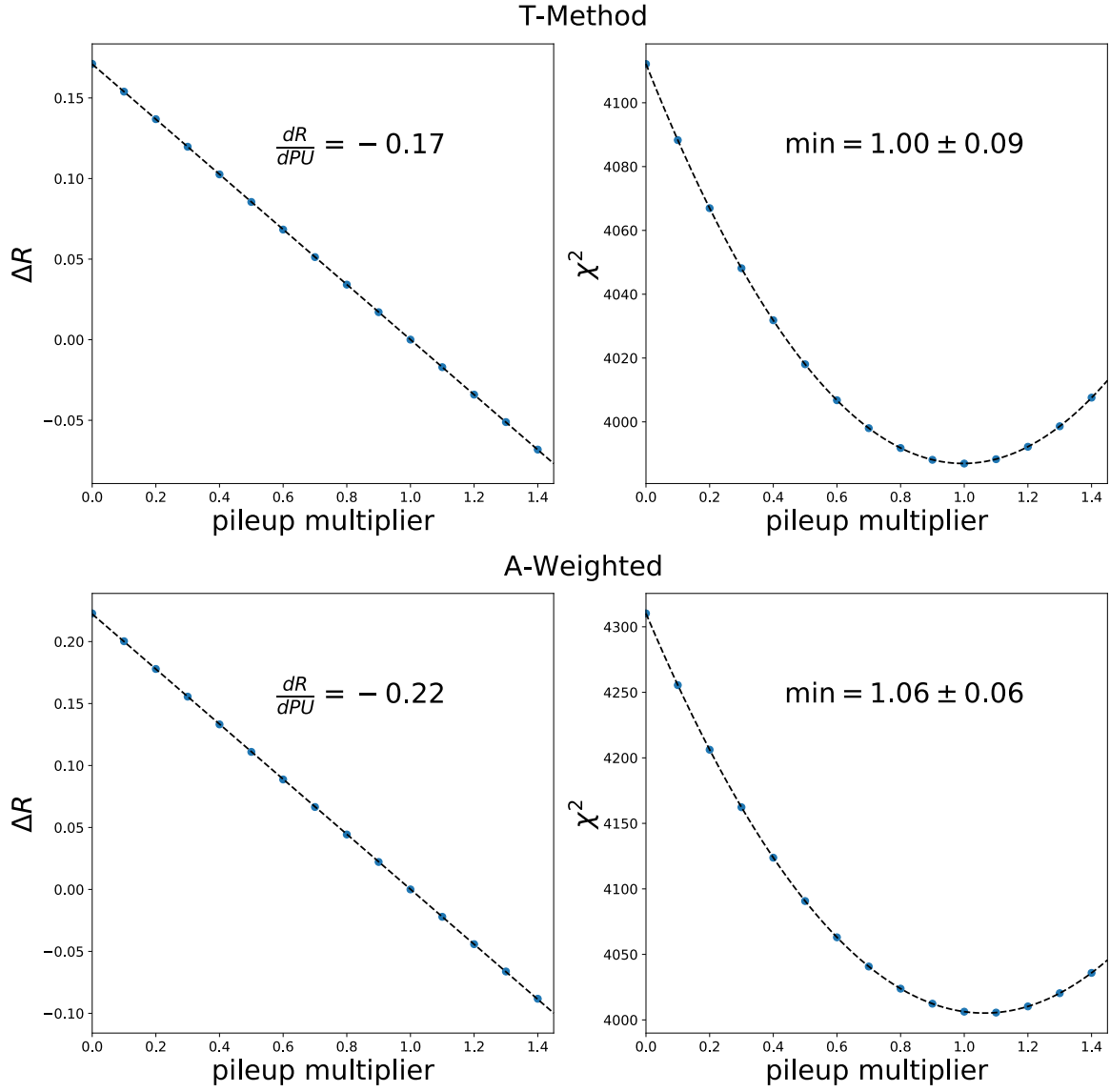


Figure 6.20: Results of the pileup multiplier scan described in the text.

The pileup correction increases the statistical uncertainty of the  $\omega_a$  measurement. The increase is very small—less than 1% (Appendix A)—and is not a significant contributor to the uncertainty budget of the E989 experiment. The pileup systematic uncertainty, however,

could be significant and must be assessed. Systematic uncertainties arise from imperfections of the pileup correction. Recall that the E989 experiment's target systematic uncertainty from pileup is 40 ppb.

As was done in the Brookhaven E821 experiment, the pileup systematic uncertainty will be assessed on two axes: uncertainty from imperfect knowledge of the pileup correction's amplitude and uncertainty from imperfect knowledge of the pileup correction's phase [8]. The phase uncertainty is rooted in the ambiguity of the pileup spectrum's time at the level of the reconstruction dead time. The optimal correction amplitude depends on the time shift, and thus the uncertainties arising from these two components are correlated. The E821 collaboration chose to add uncertainties from the pileup phase and the pileup amplitude linearly, and the same conservative treatment will be applied here. The E821 collaboration also assessed an uncertainty arising from unseen pileup, or pileup pulses below the reconstruction threshold [53]. As the reconstruction threshold in E989 is much lower, no uncertainty from unseen pileup will be assigned.

Table 6.6: Sensitivity of  $R$  to the pileup multiplier;  $dR/dPU$  is the slope of  $R$  versus the pileup multiplier,  $\sigma_{PU}$  is the uncertainty of the optimal pileup multiplier, taken from the curvature of  $\chi^2$  versus pileup multiplier, and  $\sigma_R$  is the systematic uncertainty from imperfect knowledge of the pileup correction's amplitude.

	$dR/dPU$	$\sigma_{PU}$	$\sigma_R$
T-Method	-0.17	0.09	0.015
A-Weighted	-0.22	0.06	0.013

Sensitivity to the pileup correction's amplitude was determined by repeating the  $\omega_a$  analysis with different artificial scalings of the pileup spectrum. Full 18 parameter fits were conducted at each step. Scale factors ranged from 0 to 1.4, where 1 corresponds to the scale obtained from fitting the pileup spectrum to the high-energy tail of the uncorrected energy spectrum. The results of this study are shown in Figure 6.20 and summarized in Ta-

ble 6.6. Uncertainty of the optimal pileup multiplier is taken as the inverse square root of the quadratic coefficient of the  $\chi^2$  versus pileup multiplier curve, or, equivalently, the distance from the minimum that increases the  $\chi^2$  by 1. The pileup multipliers yielding the minimum  $\chi^2$  values were consistent with one within their uncertainties for both analysis techniques, which is a valuable consistency check of the pileup correction procedure. The A-Weighted analysis was slightly more sensitive to the pileup correction scale than the T-Method, leading to a larger change of  $R$  with pileup scale, but also a smaller uncertainty of the optimal scale. Obtained  $\omega_a$  systematic uncertainties from the pileup amplitude were 15 ppb for the T-Method and 13 ppb for the A-Weighted analysis. Interestingly, the systematic error from the pileup amplitude does not change much with the artificial dead time used in reconstruction. Increasing the amount of pileup increases the slope of  $R$  versus pileup scale, but it also reduces the uncertainty of the optimal scale. These two effects tend to compensate for each other.

The fit  $\chi^2$  is extremely insensitive to the pileup correction's phase. Repeating the  $\omega_a$  analysis with the pileup spectrum shifted by an entire cyclotron period, 149 ns, changed the fit  $\chi^2$  by 0 to 2, depending on the analysis technique and the direction of the shift. The change in  $R$  was linear with the pileup time shift over this region, and the slope of  $R$  versus time shift was used to calculate the  $\omega_a$  systematic uncertainty. As the  $\chi^2$  is very insensitive to time shifts of the pileup spectrum, a scan similar to the one used for the pileup amplitude is not effective in determining the uncertainty of the pileup phase. For this analysis, a 2.5 ns dead time was employed during the reconstruction. It is therefore not reasonable to shift the pileup spectrum by more than 2.5 ns. Assuming no knowledge other than that the pileup spectrum should be shifted by some value between -2.5 and 2.5 ns, the uncertainty of the pileup correction's time,  $\sigma_{t,pu}$  can be calculated as

$$\sigma_{t,pu} = \sqrt{\int_{-2.5 \text{ ns}}^{2.5 \text{ ns}} t^2 \frac{1}{5 \text{ ns}} dt}$$

$$\sigma_{t,pu} \approx 1.4 \text{ ns.} \quad (6.19)$$

Thus, 1.4 ns is taken as the uncertainty of the pileup correction's time. This can be combined

with the sensitivity of  $R$  to the pileup phase to determine an  $\omega_a$  systematic uncertainty. Values resulting from such a determination are summarized in Table 6.7.

Table 6.7: Sensitivity of  $R$  to time shifts of the pileup spectrum, i.e. pileup phase;  $dR/dt_{pu}$  is the slope of  $R$  versus the pileup time shift,  $\sigma_{t,pu}$  represents the range of reasonable pileup time shifts as described in the text, and  $\sigma_R$  is the systematic uncertainty from imperfect knowledge of the pileup correction’s phase.

	$dR/dt_{pu}$	$\sigma_{t,pu}$	$\sigma_R$
T-Method	$0.003 \text{ ns}^{-1}$	1.4 ns	0.004
A-Weighted	$0.027 \text{ ns}^{-1}$	1.4 ns	0.038

The A-Weighted analysis is nearly ten times more sensitive to time shifts of the pileup spectrum than is the T-Method. Figure 6.21 shows why. The A-Weighted pileup spectrum displays much sharper oscillations than the T-Method pileup spectrum. As the uncertainty of the pileup phase cannot easily be extracted from the fit itself—at least not with the statistics of the 60-Hour Dataset—the systematic uncertainty from the pileup phase increases linearly with the reconstruction dead time through Equation 6.19. It is therefore beneficial to reconstruct the data with the smallest possible dead time. The large pileup phase contribution to the A-Weighted systematic uncertainty could be reduced by improving the overall pileup separation efficiency in the reconstruction. That this is possible has already been demonstrated by another analysis group within the collaboration using an alternative reconstruction procedure [93].

The systematic uncertainties from pileup are summarized in Table 6.11. The total is significantly larger for the A-Weighted analysis than it is for the T-Method. While in the case of the 60-Hour Dataset the reduced statistical uncertainty of the A-Weighted analysis greatly exceeds the increased pileup systematic uncertainty, this will not necessarily be the case with a larger dataset. Conversely, with a larger dataset the pileup phase will almost certainly be better constrained. Assuming no improvement to the calorimeter reconstruction,

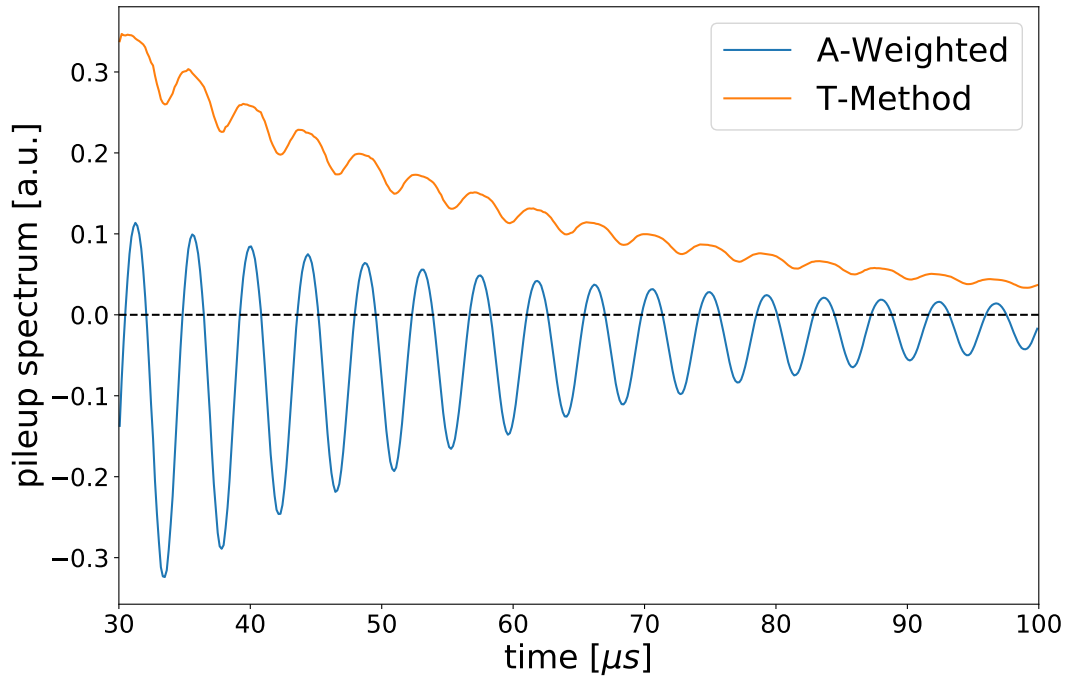


Figure 6.21: Pileup contaminations of the T-Method and A-Weighted histograms, defined as the difference between an uncorrected histogram and a pileup corrected histogram. The oscillations of the A-Weighted pileup spectrum are much sharper than those of the T-Method, which causes the A-Weighted analysis to be more sensitive to small time shifts of the pileup correction. The chosen T-Method threshold is near the threshold at which the oscillatory component of the pileup spectrum changes sign. This reduces the sensitivity of the T-Method analysis to small time shifts of the pileup correction.

the muon storage rate improvements expected in Run 2 and Run 3 will increase the relative pileup contamination. In both the T-Method and A-Weighted analyses, it appears that the ultimate target of 40 ppb is achievable.

### 6.7.3 Variation of the Artificial Dead Time

An alternative validation of the pileup correction will now be presented. The reconstruction software allows for an artificial dead time (ADT) to be imposed at the clustering stage.

Table 6.8:  $\omega_a$  systematic uncertainties from imperfections of the pileup correction, separated into contributions from uncertainty of the pileup amplitude and uncertainty of the pileup phase. As the pileup phase and the pileup amplitude are correlated, the systematic uncertainties are added linearly.

	Amplitude contribution	Phase contribution	Linear sum
T-Method	15 ppb	4 ppb	20 ppb
A-Weighted	13 ppb	38 ppb	51 ppb

Clusters closer in time than the ADT are combined. Figure 6.22 shows the time difference of consecutive clusters following reconstruction passes with different ADT configurations. With a properly functioning pileup correction,  $\omega_a$  values obtained with different ADT's should agree within an allowed drift from the increasing statistical uncertainty with increasing ADT (Appendix A). The allowed drift can be estimated using Equation 6.16. Increasing the ADT from 2.5 ns to 15 ns will enhance the  $R$  uncertainty by approximately 1%, corresponding to an allowed drift of approximately  $0.14\sigma_R$ . In the case of the 60-Hour Dataset, this allowed drift is larger than the estimated systematic uncertainties listed in Table 6.11. For the planned 100 ppb dataset, however, extraction of consistent  $\omega_a$  values with varying ADT's will be a powerful demonstration that systematic effects from pileup are understood.

The entire 60-Hour Dataset was processed with seven different ADT's between 2.5 ns and 15 ns. With each ADT, the dataset was analyzed for  $\omega_a$  with and without the pileup correction. The results of this study are shown in Figure 6.23. With the statistics of the 60-Hour Dataset, the  $\pm 1\sigma$  bands at 15 ns of ADT are approximately 200 ppb wide. They are larger in the A-Weighted case because of the larger overall pileup contamination. In both cases, the result at 15 ns ADT agrees with the one at 2.5 ns ADT. Additionally, the corrected points do not move off the baseline unidirectionally, but wander randomly within the bands as would be expected from a purely statistical drift. With a 100 ppb dataset, the bands at 15 ns would be approximately 20 ppb wide. Thus, the ability of an ADT sweep to constrain



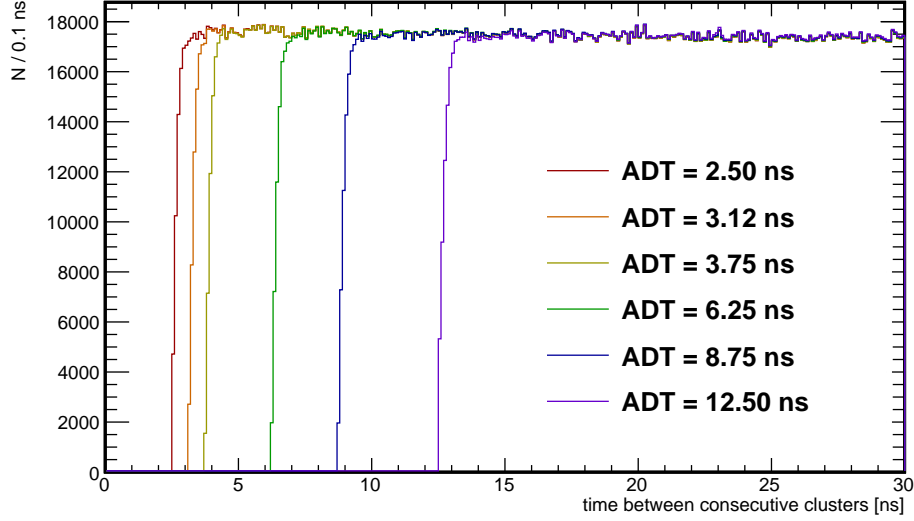


Figure 6.22: Time separation of consecutive clusters over 1 GeV with different artificial dead times, courtesy of Jason Hempstead. The same subset of the 60-Hour Dataset was processed in each case.

the pileup systematic uncertainty improves with the statistical precision of the dataset to which it is applied.

#### 6.7.4 Systematic Uncertainty from Changing Detector Gains

The systematic uncertainty from imperfect gain corrections must be assessed. In the future, dedicated studies with the laser calibration system and greater numbers of in-fill laser pulses will provide very precise measurements of the predictable calorimeter gain perturbations. In the 60-Hour Dataset, there is an ambiguity in the gain correction function at the level of  $0.5 \times 10^{-4}$  at the fit start time of  $30.2 \mu\text{s}$ . Figure 6.16 displayed evidence that, in the 60-Hour Dataset, the detector gains were not completely stabilized by the in-fill gain correction.

The first component of the gain-based systematic uncertainty comes from the extraction of the flash recovery function parameters (Section 6.5.4). The curves obtained from the in-fill laser pulses were fit with exponential recovery functions, and the obtained function

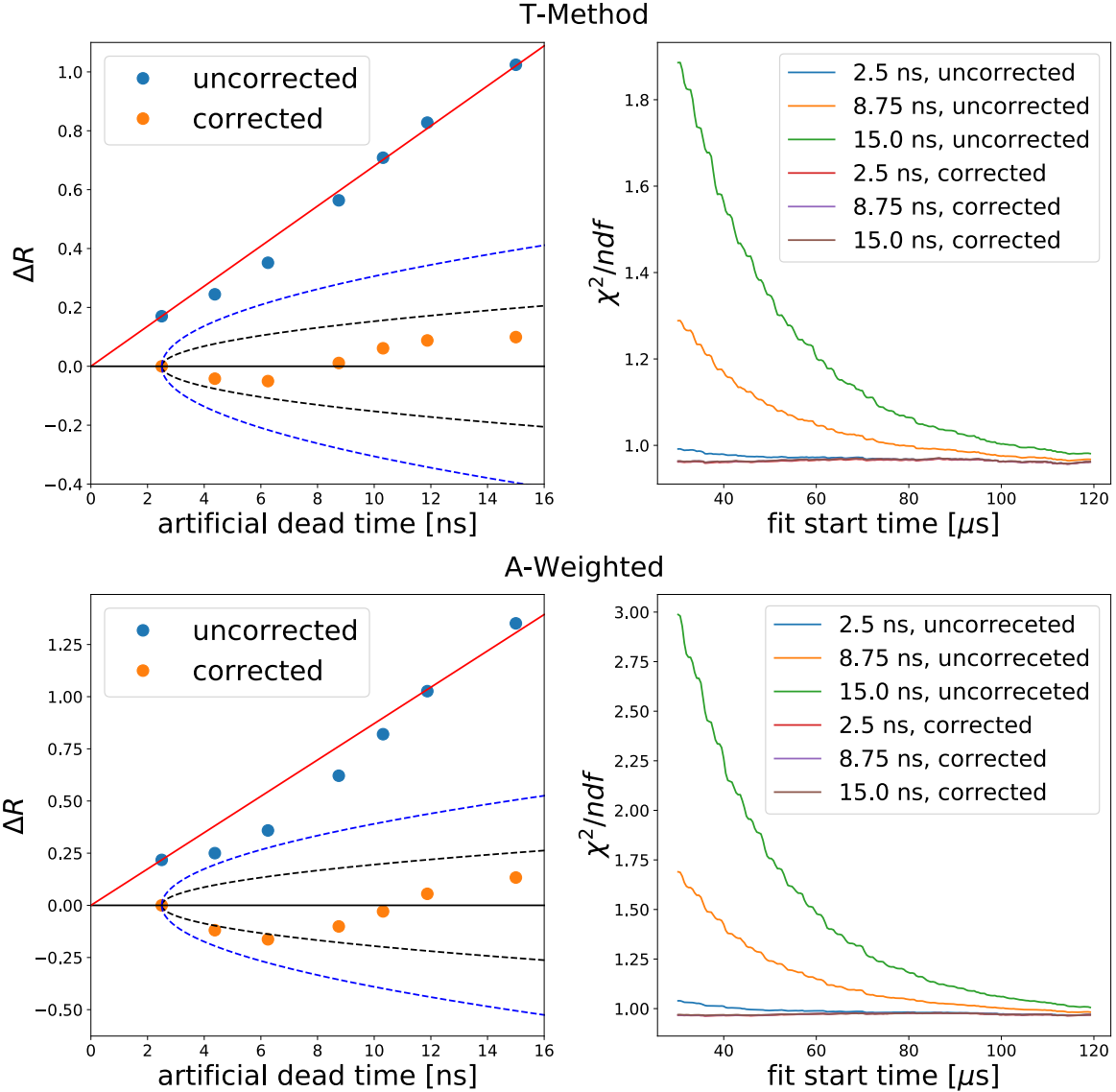


Figure 6.23:  $\omega_a$  analysis of the 60-Hour Dataset with different ADT's. The black and blue dashed lines on the left respectively correspond to the  $1\sigma$  and  $2\sigma$  allowed statistical drift from the starting point at 2.5 ns ADT. The  $R$  values are plotted relative to the corrected point at 2.5 ns. The red lines are guides to the eye, extrapolations of the pileup bias extracted from the point at 2.5 ns to larger ADT's. In both the T-Method and A-Weighted analyses, the corrected point at 15 ns lands well within the allowed bands despite a pileup correction of 1 ppm—close to the measurement's entire statistical uncertainty. On the right is  $\chi^2/\text{ndf}$  versus fit start time for three representative ADT's. After the pileup correction, the obtained  $\chi^2/\text{ndf}$  is stable versus start time, even with a 15 ns ADT.

parameters had uncertainties on the order of 10%. Additionally, the effect of the flash recovery function on the extracted  $R$  value was determined to be linear in the correction amplitude. Here, the systematic uncertainty of  $R$  owing to the flash recovery correction is estimated as 10% of the shift in  $R$  between when no correction is applied and when the full correction is applied. This is a very conservative estimate, as it implicitly assumes complete correlation of the recovery function parameters across all of the calorimeter channels. Even with this conservative treatment, the systematic uncertainty of  $R$  from the flash recovery is very small, approximately 10 ppb for both the T-Method and A-Weighted analyses. The exact values are given in Table 6.9. The smallness of this effect is not surprising given that the flash recovery is almost entirely complete by the fit start time.

Table 6.9: Estimated  $\omega_a$  systematic uncertainties from the flash recovery model, also called the in-fill gain correction. In the table, the referenced correction is the in-fill gain correction.

	Uncorrected $R$	Corrected $R$	Correction uncertainty	$\sigma_R$
T-Method	-43.255	-43.370	10%	0.011
A-Weighted	-42.530	-42.627	10%	0.010

The second component of the gain-based systematic uncertainty comes from consistent gain perturbations that are not accounted for by the in-fill gain correction. Such perturbations certainly exist because the in-fill gain correction models only the flash recovery, but there must be—at some level—a rate-dependent gain effect from the positron decays themselves. This effect cannot be much larger than  $0.5 \times 10^{-4}$ , or else it would have been directly detected through the in-fill laser pulses. The implications of a small yet uncorrected gain perturbation will now be explored.

First consider the effect a small gain perturbation would have on the calorimeter energy spectrum. This was discussed in Chapter 3. Equation 3.38 encodes how a change in calorimeter gain affects the energy spectrum. Let  $\rho_p(E)$  be the instantaneous, gain-perturbed energy

spectrum at some unspecified time, let  $1 - \delta g$  be the relative gain perturbation at that time, and let  $\rho(E)$  be the unperturbed energy spectrum. The gain perturbation,  $\delta g$ , should be understood as a function of time that smoothly approaches zero by the end of a muon fill, after the detectors have presumably recovered to their unperturbed states. The perturbed energy spectrum,  $\rho_p(E)$ , smoothly transitions into the unperturbed energy spectrum,  $\rho(E)$ , as the gain perturbation approaches zero. With these quantities defined, Equation 3.38 becomes

$$\rho_p(E) = \frac{1}{1 - \delta g} \rho\left(\frac{E}{1 - \delta g}\right). \quad (6.20)$$

Any  $\delta g$  of interest must be very small, less than one per-mille, so the above can be expanded to leading order in  $\delta g$  to obtain

$$\rho_p(E) = \rho(E) + \delta g \left[ \rho(E) + E \frac{d\rho}{dE} \right]. \quad (6.21)$$

Finally, this can be expressed as a fractional perturbation:

$$\frac{\rho_p(E) - \rho(E)}{\rho(E)} = \delta g \left( 1 + \frac{E}{\rho(E)} \frac{d\rho}{dE} \right). \quad (6.22)$$

Figure 6.24 shows this fractional perturbation by energy assuming a  $\delta g$  of  $5 \times 10^{-4}$ . It is quite large at high energies, where  $\rho(E)$  is small.

Gain perturbations that may be below the sensitivity of the laser system can inflict percent-level early-to-late baseline distortions on high-energy bins. As will be discussed in the next section, the muon loss effect is at most a percent-level baseline distortion. Thus, very small gain effects can match or exceed the size of the muon loss correction. Here it is argued that such gain effects are a likely cause of the  $K_{loss}$  parameter's instability at high energies, which was identified in Section 6.6.2.

The 60-Hour Dataset was reanalyzed assuming three different gain perturbations remaining after the in-fill gain correction. Investigations into the possible form of rate-dependent gain perturbations suggest that, after a rapid initial decrease, the gain will recover as [65]

$$\delta g(t) \approx K e^{-t/\tau} [1 + A_g \cos(\omega_a t - \phi)]. \quad (6.23)$$

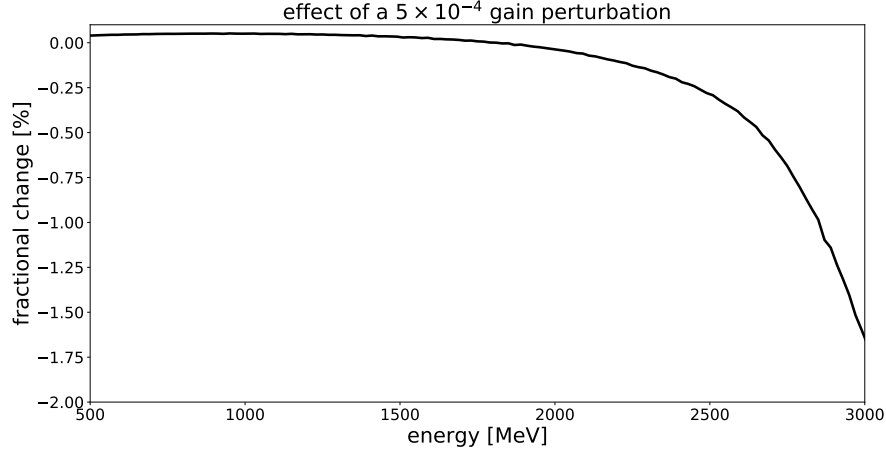


Figure 6.24: The effect of  $5 \times 10^{-4}$  gain perturbation on the calorimeter energy spectrum, obtained by replacing  $\rho(E)$  in Equation 6.22 with the measured energy spectrum from the 60-Hour Dataset.

This neglects effects from the stored beam oscillations. The gain oscillation asymmetry,  $A_g$ , can be as large as 0.2, and the phase of the oscillation,  $\phi$ , should match that of the positron hit time spectrum. Setting  $K$  to  $10^{-3}$  yields a perturbation of  $0.6 \times 10^{-4}$  at the fit start time, which is about the maximum level that could remain undetected by the laser calibration measurements. The three trial gain perturbation functions all followed Equation 6.23 and had  $K = 10^{-3}$ . They differed in the value chosen for  $A_g$ , the gain oscillation asymmetry. The values -0.2, 0, and 0.2 were used. The value 0.2 is justified as it is the overall asymmetry of the energy-weighted hit rate [65].

The pileup-subtracted hit spectrum was corrected for residual gain perturbations assuming each of the three trial gain functions. The finely-binned, pileup-corrected calorimeter spectrum was taken as  $\rho_p(E)$  and transformed into  $\rho(E)$  following Equation 6.22, which can be inverted to order  $\delta g$  to obtain  $\rho(E)$  as a function of  $\rho_p(E)$ . This procedure will be referred to as an artificial gain correction, as it is applied to the measured hit spectrum on average rather than during the reconstruction. Following the artificial gain correction, the full  $\omega_a$  analysis was repeated. Figure 6.25 shows  $K_{loss}$  versus energy with and without the

artificial gain correction. A gain perturbation of  $6 \times 10^{-4}$  can indeed explain the instability of  $K_{loss}$  with positron energy.

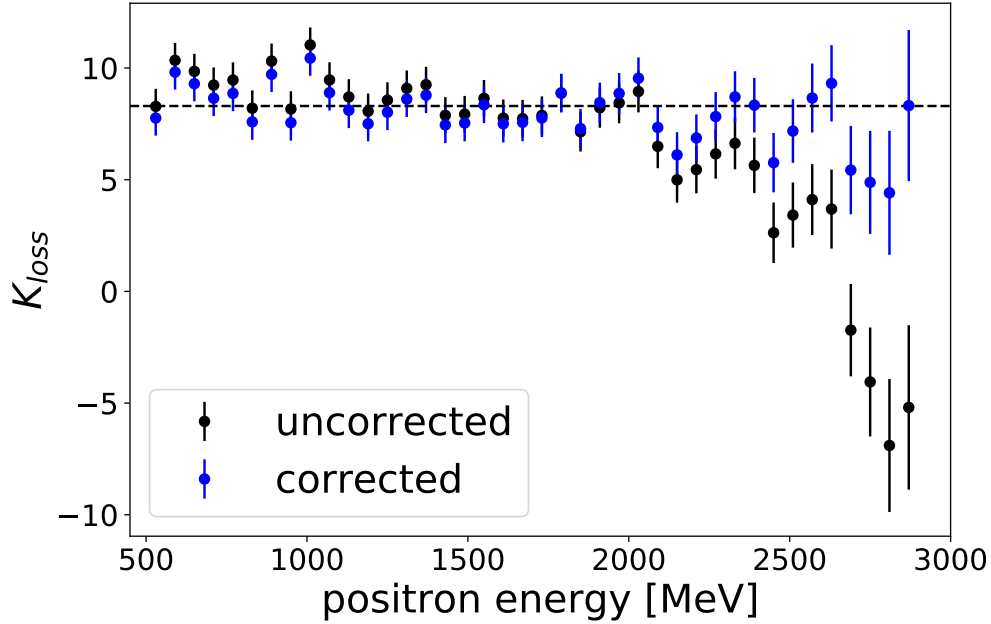


Figure 6.25:  $K_{loss}$  versus positron energy with and without the artificial gain correction. An uncorrected early-to-late gain perturbation of  $6 \times 10^{-4}$  at the fit start time can explain the instability of  $K_{loss}$  versus positron energy.

The three trial values for the gain oscillation asymmetry all stabilize  $K_{loss}$  versus positron energy equally well. To determine which is most reasonable, one can appeal to  $g - 2$  asymmetry start time scans. A gain oscillation at  $\omega_a$  would affect the  $g - 2$  asymmetry, and—as the gain perturbation reduces in size with time—one would expect to observe a shifting  $g - 2$  asymmetry with fit start time. As shown in Figure 6.26, this effect is visible even with the statistics of the 60-Hour Dataset. One expects  $A_g$  in Equation 6.23 to be positive so that the gain perturbation is largest when the hit rate is largest, and, indeed, the asymmetry is most stable versus start time when  $A_g$  is positive. Additionally, the possibility of negative  $A_g$  can be eliminated based on the results of the start time scan.

The main purpose of this exercise was to determine a reasonable systematic uncertainty

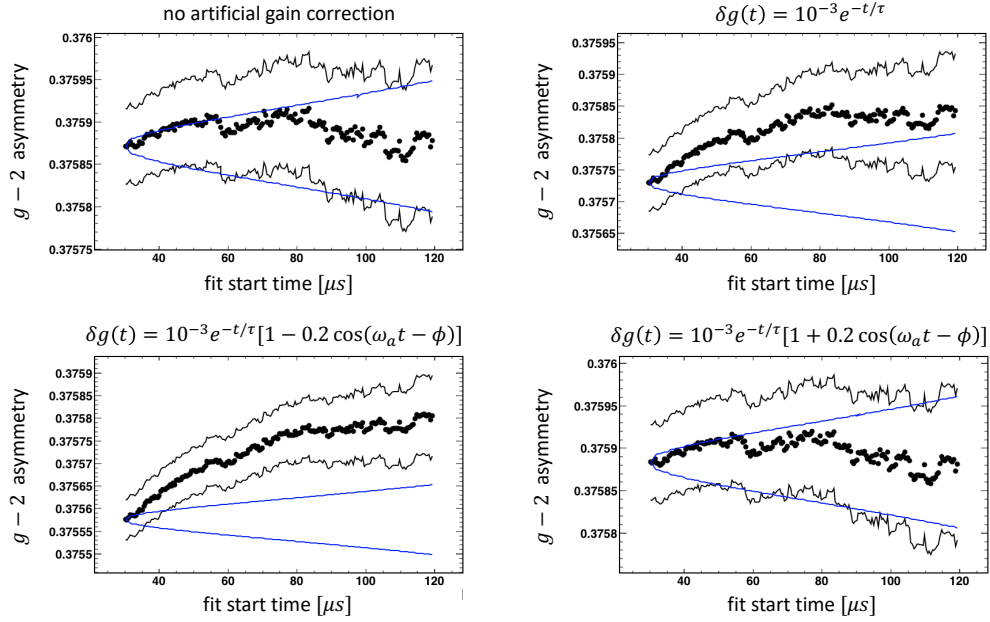


Figure 6.26:  $g - 2$  asymmetry versus start time with differing artificial gain corrections. The displayed functions are the hypothesized gain perturbations that were then removed using the procedure described in the text. A gain correction without an oscillatory component destabilizes the asymmetry versus start time. Choosing the wrong oscillation phase worsens the instability. When the phase is chosen such that the gain perturbation is largest when the hit rate is largest, the asymmetry becomes stable versus fit start time.

from rate-dependent gain perturbations not removed by the in-fill gain correction. Evidence has been presented suggesting that these perturbations could be as large as  $6 \times 10^{-4}$  at the start of the  $\omega_a$  fit. Of the trial gain functions considered, only the choice of  $A_g = 0.2$  yielded stable start time scans, and thus the others can be discounted. The changes in  $R$  and  $\chi^2$  when correcting for the trial gain function with  $A_g = 0.2$  are summarized in Table 6.10. In both cases,  $\omega_a$  is affected at the level of 60 ppb. While the trial gain function is well motivated, there is, as of this writing, no conclusive, direct evidence that it is correct. Nevertheless, the study just described provides reasonable estimates of the sensitivity of  $R$  to uncorrected, rate-dependent gain perturbations. The improvements of the fit  $\chi^2$  and of  $K_{loss}$  versus positron energy do suggest that some additional gain correction is warranted. For the purposes of the

systematic uncertainty estimates in this document, the full  $R$  shifts given in Table 6.10 will be taken as upper bounds on the systematic uncertainty contribution from rate-dependent gain changes. Through ongoing analysis of the laser calibration data and upcoming analysis of Run 1 datasets besides the 60-Hour Dataset, these upper bounds will be converted into proper systematic uncertainties before the first E989  $a_\mu$  measurement is published.

Table 6.10: Effect of the trial gain function described in the text on  $R$  and  $\chi^2$ . Artificially correcting for the trial gain function shifted  $\omega_a$  by 55 ppb in the T-Method and 70 ppb in the A-Weighted analysis. For both analyses, the fit  $\chi^2$  became smaller after the correction. The differences  $\Delta R$  and  $\Delta\chi^2$  are defined as the corrected quantity minus the uncorrected quantity.

	$\Delta R$	$\Delta\chi^2$
T-Method	-0.055	-3
A-Weighted	-0.070	-7

Table 6.11:  $\omega_a$  systematic uncertainties from changing detector gains. Only an upper bound on the uncertainty stemming from rate-dependent detector gains is presented here.

	Flash recovery model	Rate dependent gain changes	Quadrature sum
T-Method	11 ppb	< 55 ppb	< 56 ppb
A-Weighted	10 ppb	< 70 ppb	< 71 ppb

The gain systematic uncertainty assessments from this section are summarized in Table 6.11. The ultimate Run 1 systematic uncertainty from changing gains is expected to be significantly smaller than the upper bounds shown in the table. Dedicated laser measurements are planned to constrain the possible size of the rate-dependent gain perturbations. Assuming linear scaling of the systematic bias, the rate-dependent gains need only be known



to within 25% of the range explored in this section to bring the total gain systematic uncertainty below the target of 20 ppb.

#### 6.7.5 *Systematic Uncertainty from Muon Losses*

The remainder of this section is concerned with systematic uncertainties from beam dynamics effects—specifically muon losses, differential decay, and the electric field and pitch corrections—that cannot be fully constrained by the  $\omega_a$  fit itself. These effects were introduced in Chapter 3. Each of them requires its own in depth analysis, supported with beam dynamics simulations, to fully quantify, and each has its own dedicated working group within the E989 collaboration. As their analyses are not yet complete, only upper bounds for the associated systematic uncertainties will be estimated here.

Muons losses during the measurement period directly bias  $\omega_a$  if the loss probability for any given muon is not independent from its initial  $g - 2$  phase. If there does exist a dependence of the loss probability on the initial phase, muon losses will cause a time evolution of the stored muon population’s average phase and bias the observed  $\omega_a$  according to Equation 3.3. This effect exists regardless of whether the muon loss correction described in Section 6.5.2 is included in the  $\omega_a$  fit. To exactly determine the magnitude and sign of this bias, one must fully understand all the correlations between stored muon phase space variables, and one must understand how these variables relate to muon losses. To this end, detailed beam injection and storage simulations are in progress. Without this knowledge, however, one can still estimate an upper bound on the  $\omega_a$  bias from muon losses.

In the presence of losses, the number of muons stored in the ring at any given time obeys Equation 6.9. The term  $(1 - K_{loss} \int_0^t e^{t'/\tau} L(t') dt')$  encodes the correction to pure exponential decay caused by losses. The absolute scale is set by the  $K_{loss}$  parameter, which is extracted from the previously discussed fits to the calorimeter hit spectra. The  $K_{loss}$  value used here comes from the T-Method fit following the artificial gain correction described in the previous section. As the gain correction improved the stability of  $K_{loss}$  with positron energy, it is likely that the gain-corrected hit spectra provide a more reliable measurement of  $K_{loss}$  than do the

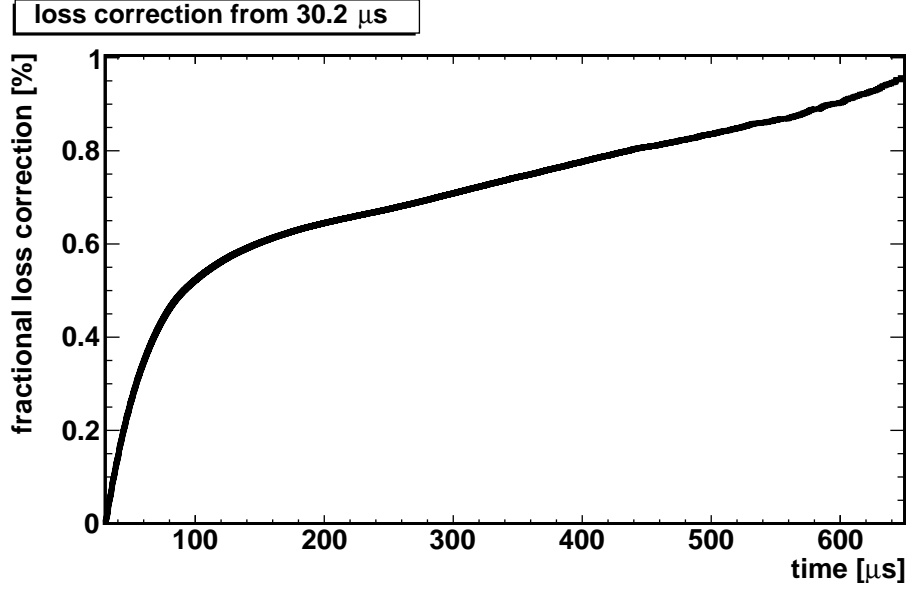


Figure 6.27: Muon loss fraction,  $f_{loss}$ , from the 60-Hour Dataset. This quantity is relative to the fit start time of  $30.2 \mu\text{s}$ . The absolute scale is set by the  $K_{loss}$  parameter extracted from the T-Method analysis. A fractional loss correction of, for example, 0.5% at  $100 \mu\text{s}$  means that  $100 \mu\text{s}$  into the fill there are 0.5% fewer muons stored in the ring than there would be in the absence of losses. This metric requires no knowledge of anything before the start of the fit at  $30.2 \mu\text{s}$ .

uncorrected ones. The artificial gain correction increased the measured value of  $K_{loss}$  by approximately 20%.

Losses that occur before the fit start time are not relevant to the  $\omega_a$  measurement. Thus, useful metrics regarding the number of muon losses must be relative to the start of the measurement window,  $t_s$ . One such metric is the loss fraction,  $f_{loss}(t)$ , defined as follows:

$$f_{loss}(t) = \frac{K_{loss}}{1 - K_{loss} \int_0^{t_s} e^{t'/\tau} L(t') dt'} \int_{t_s}^t e^{t'/\tau} L(t') dt', \quad (6.24)$$

whereby  $f_{loss}$  measures the cumulative number of losses relative to the number of muons present in the ring at the start of the  $\omega_a$  measurement window. An  $f_{loss}(t)$  of 1% means that at time  $t$  there are 1% fewer stored muons remaining than there would be from exponential decay alone. Figure 6.27 shows this quantity as a function of time in the 60-Hour Dataset. It approaches 1% by the end of the fill and is thus comparable in size to the gain and pileup

effects in high-energy bins.

As a worst case scenario, imagine that all lost muons come from the extreme low (or high) end of the momentum distribution. In this situation, losses would cause the average stored muon momentum to change over time. This alone would not be an issue unless there were a correlation between muon momentum and initial phase. Such a correlation is created by the injection beamline, particularly the delivery ring. High-momentum muons, having a higher cyclotron period, spend more time in the delivery ring than low-momentum muons do, and thus their spins precess more before they are injected into the  $g - 2$  storage ring. Simulations have suggested that the change in phase per momentum could be as large as 0.3 mrad per MeV/ $c$ . Assuming the average stored muon momentum changes at a constant rate, the  $\omega_a$  bias is

$$\frac{\Delta\omega_a}{\omega_a} = \frac{1}{\omega_a} \cdot \frac{d\langle p \rangle}{dt} \cdot \frac{d\langle \phi \rangle}{d\langle p \rangle}. \quad (6.25)$$

To obtain a number,  $\frac{d\langle p \rangle}{dt}$  must be estimated.

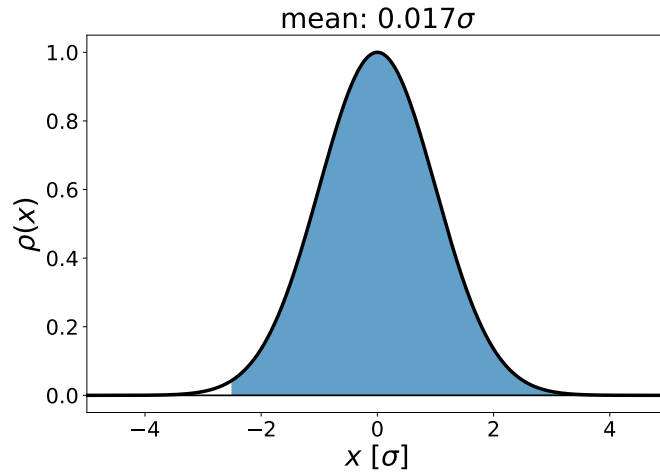


Figure 6.28: Illustration of a normal distribution with 0.6% of its total probability removed from the low side. The black curve shows the full normal distribution, and the blue the truncated distribution. Removing the bottom 0.6% of the probability distribution shifts the mean from 0 to  $0.017\sigma$ . This number is used for the muon loss systematic error estimate described in the text.

The 60-Hour Dataset loss fraction  $100\ \mu\text{s}$  after the start of the fit is approximately 0.6%. In the worst case scenario described above, the bottom 0.6% of the momentum distribution would be lost by  $100\ \mu\text{s}$ . The change in average momentum that this would cause depends on the exact shape of the momentum distribution. In the case of a normal distribution, removing the bottom 0.6% of the total probability shifts the mean by  $0.017\sigma$  (Figure 6.28). While it is known that the stored momentum distribution is not normal, this number allows for a concrete estimate of the  $\omega_a$  bias from muon losses. Figure 6.27 shows that the loss fraction increases steeply at early times, but its rate of change decreases at  $100\ \mu\text{s}$  and then remains more-or-less constant. Most of the data used for the  $\omega_a$  analysis comes in the first  $100\ \mu\text{s}$ . For this estimate, we will assume a constant rate of change of  $0.017\sigma$  per  $100\ \mu\text{s}$ .

The last ingredient needed to complete this estimate is the standard deviation of the stored momentum distribution, the momentum width. The fast rotation analysis provides a distribution of cyclotron frequencies, which can be converted to a momentum distribution. In the 60-Hour Dataset, the momentum width was found to be  $3.4\ \text{MeV}/c$ . Thus, the change of  $0.017\sigma$  corresponds to an average momentum shift of  $0.06\ \text{MeV}/c$  over  $100\ \mu\text{s}$ . This and the estimate of  $\frac{d\langle\phi\rangle}{d\langle p\rangle}$  listed above can be combined to provide an estimated upper bound for the lost muon systematic bias. Doing so,

$$\begin{aligned}\frac{\Delta\omega_a}{\omega_a} &= \frac{1}{\omega_a} \cdot \frac{d\langle p\rangle}{dt} \cdot \frac{d\langle\phi\rangle}{d\langle p\rangle} \\ &= \frac{1}{1.44\text{rad}/\mu\text{s}} \cdot \frac{0.06\ \text{MeV}/c}{100\ \mu\text{s}} \cdot \frac{0.3\ \text{mrad}}{\text{MeV}/c} \\ &= 125\ \text{ppb}.\end{aligned}$$

In the worst case scenario where all lost muons come from the extreme end of the momentum distribution, the  $\omega_a$  bias is estimated to be 125 ppb. Many imperfect assumptions and approximations contributed to this estimate, but it does provide a plausible scale for the size of the lost muon effect. This result does not mean the E989 Run 1 systematic uncertainty from lost muons will be 125 ppb, as simulations and tracker measurements will reveal where in stored muon phase space the losses are actually occurring. If losses do tend toward one

side or the other of the momentum distribution, a correction will be applied and the resulting systematic uncertainty will be that correction's uncertainty, not the bias estimated here.

#### 6.7.6 *Systematic Uncertainty from Differential Decay*

Section 3.3.2 presented the differential decay effect. The boosted decay lifetime depends on muon momentum, and this alone is enough to create a time-dependent distortion of the stored muon momentum distribution. Equation 3.22 gave the size of the resulting bias as a function of the stored momentum width and  $\frac{d\langle\phi\rangle}{d\langle p\rangle}$ . Repeated here,

$$\frac{\Delta\omega_a}{\omega_a} = -\frac{m_\mu v}{\omega_a \tau_\mu} \left(\frac{\sigma_p}{p_0}\right)^2 \cdot \frac{d\langle\phi\rangle}{d\langle p\rangle}.$$

Estimates for  $\frac{d\langle\phi\rangle}{d\langle p\rangle}$  and the momentum width were presented in the previous section. With these numbers, the differential decay bias is 12 ppb. This is very small in the context of E989 Run 1, but may become important in later runs. With  $\frac{d\langle\phi\rangle}{d\langle p\rangle}$  known, the sign of the differential decay bias is determined. Unlike muon losses, differential decay can only ever cause the average momentum to increase with time. Therefore, a correction can easily be applied if the bias is deemed too large to tolerate.

#### 6.7.7 *Systematic Uncertainty from the Electric Field and Pitch Corrections*

The electric field and pitch corrections (Section 3.3.6) are not determined from analyses like the one described in this chapter. The electric field correction depends on the stored muon equilibrium radius distribution (Figure 6.1), which is extracted from the beam debunching envelope through the fast rotation analysis. The pitch correction is derived from the transverse distribution of stored muons, specifically the marginal distribution in  $y$ . This is measured with the tracking detectors. Both of these analyses are still in progress. Those involved are confident that, for Run 1, each correction will be determined to better than 50 ppb. Assuming 50 ppb in both cases, the limit on the quadrature sum is 70 ppb. The final uncertainties cannot be known before the analyses are complete.

### 6.7.8 Summary of the Systematic Uncertainties

Data-driven estimates for the known Run 1  $\omega_a$  systematic uncertainties have been presented. In some cases, only conservative upper bounds could be determined. The estimates are summarized in Table 6.12. It is of course possible that more effects will be discovered in the Run 1 data, but initial results are promising. Known avenues for improvement have been identified in all cases where the Run 1 estimates differ significantly from their targets.

Table 6.12: Estimated E989 Run 1 systematic uncertainties and the target values presented in Chapter 3. Run 1 values are for the T-Method.

Uncertainty source	Run 1 estimate	Target value
Beam dynamics effects	< 150 ppb	< 50 ppb
Lost muons	< 125 ppb	20 ppb
CBO	30 ppb	30 ppb
E-field and pitch corrections	< 70 ppb	30 ppb
Detector effects	< 60 ppb	< 50 ppb
Gain changes	< 60 ppb	20 ppb
Pileup	20 ppb	40 ppb
Total quadrature sum	< 160 ppb	< 70 ppb

## 6.8 Outlook

The 60-Hour Dataset is one of seven E989 Run 1 subdatasets with conditions suitable for  $\omega_a$  analysis. Relative unblinding of the  $\omega_a$  values extracted from the 60-Hour Dataset by up to seven independent analysis teams is planned for early spring of 2019. Following thorough data quality filtering and production processing, the remaining six subdatasets will be analyzed for  $\omega_a$  using the same techniques described in this chapter. Relative unblinding of these numbers is planned for mid-2019. If there is agreement between all analyzers and the internal review committees are satisfied, the Run 1 result will be published shortly thereafter.

The combined Run 1 statistical uncertainty of  $\omega_a$  is projected to be 350 ppb. In the worst-case scenario where no improvements beyond the upper bounds shown in Table 6.12 are realized, the systematic uncertainty of the Run 1  $\omega_a$  measurement would be 160 ppb. Thus, the Run 1  $\omega_a$  result will be statistics-limited. Regarding the magnetic field measurement, the projected Run 1 uncertainty of  $\tilde{\omega}_p$  is 140 ppb. The uncertainty of the E989  $a_\mu$  measurement is entirely dominated by the uncertainty of the ratio  $\omega_a/\tilde{\omega}_p$ , which can be decomposed into the statistical uncertainty of the  $\omega_a$  measurement and the systematic uncertainties of the  $\omega_a$  measurement and the  $\tilde{\omega}_p$  measurement. Using the worst-case value for the  $\omega_a$  systematic uncertainty,

$$\begin{aligned}\sigma_{a_\mu, \text{Run 1}} &\approx \sqrt{(350 \text{ ppb})^2 + (160 \text{ ppb})^2 + (140 \text{ ppb})^2}, \\ \sigma_{a_\mu, \text{Run 1}} &\approx 410 \text{ ppb}.\end{aligned}$$

If the  $\omega_a$  systematic uncertainty were improved to 70 ppb, the design target, the above result would change to 380 ppb. These numbers can be compared to the uncertainty of the combined  $a_\mu$  result from Brookhaven, 540 ppb.

The E989 Run 1 measurement will be the first experimental value of  $a_\mu$  with sub-ppm precision since the final Brookhaven result was published, nearly two decades before this writing. For the E989 Run 1 result and the Brookhaven result to agree within  $1\sigma$ , the difference between the E989 value and the Standard Model prediction must be at least  $3\sigma$ .

Independent confirmation would effectively eliminate the possibility that the long-standing muon  $g-2$  discrepancy was caused by a statistical fluctuation. By the time the Run 1 result is published, collection of the  $\sim 200$  ppb Run 2 dataset will be nearly finished and its analysis will be well on its way. Run 3 will occur in 2020, completing the E989 combined dataset. With the expected improvements to muon storage rates, the E989 experiment will achieve its target precision of 140 ppb.



## BIBLIOGRAPHY

- [1] John David Jackson. *Classical electrodynamics*. Wiley India, 2011.
- [2] David Griffiths. *Elementary Particles*. WILEY-VCH, 2008.
- [3] J. J. Sakurai and San Fu Tuan. *Modern quantum mechanics revised edition*. Addison-Wesley Pub. Co., 1994.
- [4] Matthew Dean Schwartz. *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.
- [5] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98(3):030001, 2018.
- [6] P. A. M. Dirac. The Quantum Theory of the Electron. *Proceedings of the Royal Society of London Series A*, 117:610–624, February 1928.
- [7] B. Lee Roberts. Historical introduction to electric and magnetic moments. *Lepton Dipole Moments Advanced Series on Directions in High Energy Physics*, page 1–9, 2009.
- [8] Charles C. Polly. *A Measurement of the Anomalous Magnetic Moment of the Negative Muon to 0.7 ppm*. PhD thesis, University of Illinois, Urbana IL, 2005.
- [9] D. Hanneke, S. Fogwell, and G. Gabrielse. New Measurement of the Electron Magnetic Moment and the Fine Structure Constant. *Phys. Rev. Lett.*, 100:120801, 2008.
- [10] Mark Allen. Srednicki. *Quantum field theory*. Cambridge University Press, 2018.
- [11] Michael Edward Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Westview Press, 2016.
- [12] Andrzej Czarnecki and William J. Marciano. The Muon anomalous magnetic moment: A Harbinger for 'new physics'. *Phys. Rev.*, D64:013014, 2001.
- [13] Fred Jegerlehner. Leading-order hadronic contribution to the electron and muon  $g - 2$ . *EPJ Web Conf.*, 118:01016, 2016.
- [14] Alexander Keshavarzi, Daisuke Nomura, and Thomas Teubner. Muon  $g - 2$  and  $\alpha(M_Z^2)$ : a new data-based analysis. *Phys. Rev.*, D97(11):114025, 2018.

- [15] Michel Davier, Andreas Hoecker, Bogdan Malaescu, and Zhiqing Zhang. Reevaluation of the hadronic vacuum polarisation contributions to the Standard Model predictions of the muon  $g - 2$  and  $\alpha(m_Z^2)$  using newest hadronic cross-section data. *Eur. Phys. J.*, C77(12):827, 2017.
- [16] Tatsumi Aoyama, Masashi Hayakawa, Toichiro Kinoshita, and Makiko Nio. Complete Tenth-Order QED Contribution to the Muon  $g - 2$ . *Phys. Rev. Lett.*, 109:111808, 2012.
- [17] Tatsumi Aoyama, Toichiro Kinoshita, and Makiko Nio. Revised and Improved Value of the QED Tenth-Order Electron Anomalous Magnetic Moment. *Phys. Rev.*, D97(3):036001, 2018.
- [18] Rym Bouchendira, Pierre Clade, Saida Guellati-Khelifa, Francois Nez, and Francois Biraben. New determination of the fine structure constant and test of the quantum electrodynamics. *Phys. Rev. Lett.*, 106:080801, 2011.
- [19] Peter J Mohr, David B Newell, Barry N Taylor, and Eite Tiesinga. Data and analysis for the CODATA 2017 special fundamental constants adjustment. *Metrologia*, 55(1):125, 2018.
- [20] Richard H. Parker, Chenghui Yu, Weicheng Zhong, Brian Estey, and Holger Müller. Measurement of the fine-structure constant as a test of the Standard Model. *Science*, 360(6385):191–195, 2018.
- [21] Andrzej Czarnecki, William J. Marciano, and Arkady Vainshtein. Refinements in electroweak contributions to the muon anomalous magnetic moment. *Phys. Rev. D*, 67:073006, Apr 2003.
- [22] C. Gnendiger, D. Stöckinger, and H. Stöckinger-Kim. The electroweak contributions to  $(g - 2)_\mu$  after the Higgs boson mass measurement. *Phys. Rev.*, D88:053005, 2013.
- [23] Tadashi Ishikawa, Nobuya Nakazawa, and Yoshiaki Yasui. Numerical calculation of the full two-loop electroweak corrections to muon  $(g-2)$ , 2018.
- [24] Fred Jegerlehner and Andreas Nyffeler. The muon  $g - 2$ . *Physics Reports*, 477(1):1 – 110, 2009.
- [25] Gilberto Colangelo, Martin Hoferichter, and Peter Stoffer. Two-pion contribution to hadronic vacuum polarization, 2018.

- [26] T. Blum, P. A. Boyle, V. Gülpers, T. Izubuchi, L. Jin, C. Jung, A. Jüttner, C. Lehner, A. Portelli, and J. T. Tsang. Calculation of the hadronic vacuum polarization contribution to the muon anomalous magnetic moment. *Phys. Rev. Lett.*, 121(2):022003, 2018.
- [27] G. Abbiendi et al. Measuring the leading hadronic contribution to the muon  $g-2$  via  $\mu e$  scattering. *Eur. Phys. J.*, C77(3):139, 2017.
- [28] Nils Asmussen, Antoine Gérardin, Andreas Nyffeler, and Harvey B. Meyer. Hadronic light-by-light scattering in the anomalous magnetic moment of the muon. In *15th International Workshop on Tau Lepton Physics (TAU2018) Amsterdam, Netherlands, September 24-28, 2018*, 2018.
- [29] Gilberto Colangelo, Martin Hoferichter, Massimiliano Procura, and Peter Stoffer. Hadronic light-by-light contribution to  $(g-2)_\mu$ : a dispersive approach. *EPJ Web Conf.*, 175:01025, 2018.
- [30] Harvey B. Meyer and Hartmut Wittig. Lattice QCD and the anomalous magnetic moment of the muon. *Prog. Part. Nucl. Phys.*, 104:46–96, 2019.
- [31] Martin Hoferichter, Bai-Long Hoid, Bastian Kubis, Stefan Leupold, and Sebastian P. Schneider. Pion-pole contribution to hadronic light-by-light scattering in the anomalous magnetic moment of the muon. *Phys. Rev. Lett.*, 121(11):112002, 2018.
- [32] Andreas Nyffeler. Precision of a data-driven estimate of hadronic light-by-light scattering in the muon  $g-2$ : Pseudoscalar-pole contribution. *Phys. Rev.*, D94(5):053006, 2016.
- [33] B. Lee Roberts. The History of the Muon  $(g-2)$  Experiments. In *15th International Workshop on Tau Lepton Physics (TAU2018) Amsterdam, Netherlands, September 24-28, 2018*, 2018.
- [34] G. W. Bennett et al. Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL. *Phys. Rev.*, D73:072003, 2006.
- [35] Weiwen Liu et al. High precision measurements of the ground state hyperfine structure interval of muonium and of the muon magnetic moment. *Phys. Rev. Lett.*, 82:711–714, 1999.
- [36] Peter J. Mohr, David B. Newell, and Barry N. Taylor. CODATA recommended values of the fundamental physical constants: 2014. *Rev. Mod. Phys.*, 88:035009, Sep 2016.

- [37] J. Grange et al. Muon (g-2) Technical Design Report, 2015.
- [38] Yutaro Sato. Muon g-2/EDM experiment at J-PARC. *PoS*, KMI2017:006, 2017.
- [39] Kaoru Hagiwara, Kai Ma, and Satyanarayan Mukhopadhyay. Closing in on the chargino contribution to the muon  $g - 2$  in the MSSM: current LHC constraints. *Phys. Rev.*, D97(5):055035, 2018.
- [40] M. Ablikim et al. Dark Photon Search in the Mass Range Between 1.5 and 3.4 GeV/ $c^2$ . *Phys. Lett.*, B774:252–257, 2017.
- [41] Maxim Pospelov. Secluded U(1) below the weak scale. *Phys. Rev.*, D80:095002, 2009.
- [42] Mauro Raggi and Venelin Kozhuharov. Results and perspectives in dark photon physics. *Riv. Nuovo Cim.*, 38(10):449–505, 2015.
- [43] J. P. Lees et al. Search for Invisible Decays of a Dark Photon Produced in  $e^+e^-$  Collisions at BaBar. *Phys. Rev. Lett.*, 119(13):131804, 2017.
- [44] J. R. Batley et al. Search for the dark photon in  $\pi^0$  decays. *Phys. Lett.*, B746:178–185, 2015.
- [45] Chien-Yi Chen, Hooman Davoudiasl, William J. Marciano, and Cen Zhang. Implications of a light “dark Higgs” solution to the  $g_\mu$ -2 discrepancy. *Phys. Rev.*, D93(3):035006, 2016.
- [46] W. J. Marciano, A. Masiero, P. Paradisi, and M. Passera. Contributions of axionlike particles to lepton dipole moments. *Phys. Rev.*, D94(11):115033, 2016.
- [47] Y. K. Semertzidis et al. The Brookhaven muon (g-2) storage ring high voltage quadrupoles. *Nucl. Instrum. Meth.*, A503:458–484, 2003.
- [48] Diktys Stratakis et al. Accelerator performance analysis of the Fermilab Muon Campus. *Phys. Rev. Accel. Beams*, 20(11):111003, 2017.
- [49] A. Yamamoto et al. The superconducting inflector for the BNL g-2 experiment. *Nucl. Instrum. Meth.*, A491:23–40, 2002.
- [50] A. Chapelain. Attempt at Estimating the Beam-Line Positron Contamination from Fiber Harps Data and Simulation. E989 DocDB 10758, 2018.
- [51] X. Fei, V. W. Hughes, and R. Prigl. Precision measurement of the magnetic field in terms of the free-proton NMR frequency. *Nucl. Instrum. Meth.*, A394:349–356, 1997.

- [52] G. W. Bennett et al. Statistical equations and methods applied to the precision muon (g-2) experiment at BNL. *Nucl. Instrum. Meth.*, A579:1096–1116, 2007.
- [53] Frederick Earl Gray. *A Measurement of the Anomalous Magnetic Moment of the Positive Muon to 0.7 Parts Per Million*. PhD thesis, University of Illinois, Urbana IL, 2003.
- [54] D. L. Rubin. E989 Note 86: Evolution of the Muon Distribution in the g-2 Ring, 2016.
- [55] J. Grange. E989 Note 57: Shift in  $\omega_p$  due to couplings between longitudinal and transverse magnetic field inhomogeneities (distortion of the closed orbit), 2015.
- [56] M. J. Syphers. E989 Note 127: Ring Tunes, 2017.
- [57] M. J. Syphers. E989 Note 142: Nonlinear Detuning, 2018.
- [58] E. M. Metodiev et al. Farley/Metodiev Differential Decay Systematic Error, 2016.
- [59] Jason Crnkovic, Sudeshna Ganguly, William Morse, and Diktys Stratakis. Lost Muon Study for the Muon g-2 Experiment at Fermilab. In *Proceedings, 8th International Particle Accelerator Conference (IPAC 2017): Copenhagen, Denmark, May 14-19, 2017*, page WEPIK119, 2017.
- [60] J. Bailey et al. Final Report on the CERN Muon Storage Ring Including the Anomalous Magnetic Moment and the Electric Dipole Moment of the Muon, and a Direct Test of Relativistic Time Dilation. *Nucl. Phys.*, B150:1–75, 1979.
- [61] Y. Orlov, C. S. Ozben, and Y. K. Semertzidis. Muon revolution frequency distribution from a partial-time Fourier transform of the g-2 signal in the muon g-2 experiment. *Nucl. Instrum. Meth.*, A482:767–775, 2002.
- [62] D. Seleznev et al. E989 Note 130: Extraction of the Muon Beam Frequency Distribution via the Fourier Analysis of the Fast Rotation Signal, 2018.
- [63] J. Kaspar et al. Design and performance of SiPM-based readout of PbF<sub>2</sub> crystals for high-rate, precision timing applications. *JINST*, 12(01):P01009, 2017.
- [64] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.
- [65] A. T. Fienberg. E989 Note 102: Gain Perturbation Calculation, 2017.
- [66] Antonio Anastasi. *The Calibration System of the E989 Experiment at Fermilab*. PhD thesis, Messina U., 2017.

- [67] Shanghai SICCAS High Technology Corporation. PbF<sub>2</sub> Crystal. <http://www.siccas.com/PbF2LeadfluorideCrystal.htm>. Accessed: December 21, 2018.
- [68] A. Ghassemi et al. MPPC. [https://www.hamamatsu.com/resources/pdf/ssd/mppc\\_kapd9005e.pdf](https://www.hamamatsu.com/resources/pdf/ssd/mppc_kapd9005e.pdf). Accessed: December 21, 2018.
- [69] A. T. Fienberg et al. Studies of an array of PbF<sub>2</sub> Cherenkov crystals with large-area SiPM readout. *Nucl. Instrum. Meth.*, A783:12–21, 2015.
- [70] A. Anastasi et al. Test of candidate light distributors for the muon ( $g-2$ ) laser calibration system. *Nucl. Instrum. Meth.*, A788:43–48, 2015.
- [71] A. Anastasi et al. Electron beam test of key elements of the laser-based calibration system for the muon  $g-2$  experiment. *Nucl. Instrum. Meth.*, A842:86–91, 2017.
- [72] A. Anastasi et al. The laser control of the muon  $g-2$  experiment at Fermilab. *JINST*, 13(02):T02009, 2018.
- [73] Antoine Chapelain. Development of the electromagnetic calorimeter waveform digitizers for the Fermilab Muon  $g-2$  experiment. *PoS*, EPS-HEP2015:280, 2015.
- [74] David A. Sweigart. A new MicroTCA-based waveform digitizer for the Muon  $g-2$  experiment. *PoS*, ICHEP2016:845, 2016.
- [75] W. Gohn. Data Acquisition with GPUs: The DAQ for the Muon  $g-2$  Experiment at Fermilab. *PoS*, ICHEP2016:174, 2016.
- [76] A. T. Fienberg. Template Fitting with Newton’s Method. E989 DocDB 4597, 2016.
- [77] G. Guennebaud, B. Jacob, et al. Eigen v3. <http://eigen.tuxfamily.org>, 2010.
- [78] F. Gray and G. Onderwater. E821 Note 416: Pulse Fitting in G2Too, 2002.
- [79] K. Khaw. Recon West Update: A Tale of Time and Energy. E989 DocDB 11055, 2018.
- [80] M. Mazouz, L. Ghedira, and E. Voutier. Determination of shower central position in laterally segmented lead-fluoride electromagnetic calorimeters. *JINST*, 11(07):P07001, 2016.
- [81] A. T. Fienberg. Calorimeter Event Reconstruction. E989 DocDB 2974, 2015.
- [82] W. M. Morse. E821 Note 268: Effects of Skew Multipoles in the Storage Ring Magnet, 1996.

- [83] R. Osofsky. Radial and Longitudinal Field Measurements. E989 DocDB 4671, 2016.
- [84] K. Khaw. E989 Note 120: CBO Amplitude Versus Kicker Strength, 2017.
- [85] Nathan S. Froemming. *Optimization of Muon Injection and Storage in the Fermilab  $g - 2$  Experiment: From Simulation to Reality*. PhD thesis, University of Washington, Seattle WA, 2018.
- [86] J. Mott. CBO systematics from tracker frequency model. E989 DocDB 15376, 2018.
- [87] R. Brun and F. Rademakers. ROOT: An object oriented data analysis framework. *Nucl. Instrum. Meth.*, A389:81–86, 1997.
- [88] J. Mott. Changing CBO Frequency. E989 DocDB 14208, 2018.
- [89] J. Price. Omega Europa Summary. E989 DocDB 15355, 2018.
- [90] M. W. Smith. In Fill Corrections. E989 DocDB 15022, 2018.
- [91] M. W. Smith and P. Girotti. E989 Note 157: Short-Time Double Pulse: Initial Tests, 2018.
- [92] D. Kawall. E821 Note 322: Statistical and Systematic Errors in Fitting  $\omega_a$ , 1998.
- [93] D. A. Sweigart. T/E-binned Analysis on 60 h. with Recon East. E989 DocDB 15340, 2018.

## Appendix A

### PILEUP CORRECTED BIN UNCERTAINTIES

This appendix presents the bin uncertainties for T-Method and A-Weighted histograms following the pileup correction procedure described in Section 6.5.3. The effect of the pileup imparted uncertainty enhancements on A-Weighted and T-Method fits from the 60-Hour Dataset are also discussed.

#### **A.1 Pileup Correction Overview**

The pileup correction presented in Section 6.5.3 is based on an empirical per-calorimeter per-fill positron hit distribution  $\rho(E, t)$ . Note that when spatial information is not used for pileup discrimination during reconstruction, the spatial hit distribution of positrons need not be considered. The measured, pileup perturbed hit distribution is taken to be  $\rho(E, t)$ . The error in using a pileup perturbed hit distribution to generate the double pileup correction appears at the order of triple pileup. The procedure can be improved further by modeling triple pileup and running multiple iterations of the correction procedure. However, the effect of additional bin uncertainties created by the pileup correction is already so small that treatment up to the level of double pileup should be sufficient here. Thus, bin uncertainty effects from triple pileup will not be considered.

Repeating Equation 6.11, the perturbation to the measured hit spectrum from double pileup is:

$$\delta\rho_{pu,d}(E, t) \propto \Delta t \left[ \int \rho(E - E_2, t) \cdot \rho(E_2, t) dE_2 - 2\rho(E, t) \int \rho(E_2, t) dE_2 \right],$$

where  $\Delta t$  is the detector dead time. One important observation from this is that the pileup correction at a given time is independent from the pileup correction at any other time.



Thus, while the pileup correction can create correlations between energy bins, it cannot create correlations between time bins. This equation also assumes that the reconstruction reports the energy of unresolved double pileup as the sum of the energies of the constituent pulses, with no gain reduction. While this assumption is surely violated at some level, it appears sufficient for generating a pileup correction whose shape matches the high-energy tail observed in energy distributions constructed from data.

In practice, the pileup correction begins with a finely binned histogram representing  $\rho(E, t)$ , except aggregated over an entire dataset rather than just a single fill. I will call this object  $N_{ij}$ , with  $i$  as an energy bin and  $j$  a time bin. Furthermore—as the pileup correction in each time bin is independent—I will work with the object  $N_i$ , which is the number of counts in energy bin  $i$  at some unspecified time. Whatever equations are derived will apply to any time bin. When considering a single calorimeter,  $N_{ij}$  is related to  $\rho(E, t)$  by:

$$\begin{aligned} N_{ij} &= N_{fills} \iint_{bin\ ij} \rho(E, t) dE dt \\ &\approx N_{fills} \cdot w_E \cdot w_t \cdot \rho(E, t), \end{aligned} \quad (\text{A.1})$$

where  $w_E$  and  $w_t$  refer to the bin widths. Following this, the pileup perturbation in a given energy bin calculated from the discrete bin counts  $N$  is:

$$\delta N_i = \epsilon \left( \sum_{j < i} N_j N_{i-j-1} - 2N_i \sum_j N_j \right), \quad (\text{A.2})$$

$$\epsilon = \frac{\Delta t}{w_t} \cdot \frac{1}{N_{fills}}. \quad (\text{A.3})$$

The proportionality constant  $\epsilon$  can be extracted from the data by fitting the pileup correction to the observed high-energy tail and then related to the dead time using the above equation. It is a small parameter. The pileup corrected histogram,  $N_i^c$ , is  $N_i - \delta N_i$ :

$$N_i^c = N_i + \epsilon \left( 2N_i \sum_j N_j - \sum_{j < i} N_j N_{i-j-1} \right). \quad (\text{A.4})$$

The corrected spectrum,  $N_i^c$ , is then used to generate T-Method, A-Weighted, or energy-binned histograms. In general, variables with a  $c$  superscript will refer to pileup corrected objects whereas those without a  $c$  superscript will refer to uncorrected objects.

## A.2 Uncertainty of a Single Energy Bin

In this section, the variance of a pileup corrected energy bin is calculated. Through Equation A.4, the content of an energy bin in the pileup corrected histogram depends on the contents of all energy bins in the pileup perturbed histogram. The pileup perturbed histogram, being what is originally measured in the experiment, obeys counting statistics, and the statistical fluctuations in each energy bin are independent. Consequently, the statistical fluctuations of the energy bin contents in the pileup corrected histogram are correlated with each other. Calculating the variance of each pileup corrected energy bin alone is not sufficient to determine the variance of T-Method or A-Weighted histogram bins, which are built from sums over a number of the correlated energy bins.

The variance of a pileup corrected energy bin can be determined by considering its dependence on each of the uncorrected energy bins, which have independent statistical fluctuations:

$$\begin{aligned}\sigma_{N_i^c}^2 &= \sum_j \sigma_{N_j}^2 \left( \frac{\partial N_i^c}{\partial N_j} \right)^2 \\ \sigma_{N_i^c}^2 &= \sum_j N_j \left( \frac{\partial N_i^c}{\partial N_j} \right)^2.\end{aligned}\tag{A.5}$$

That last step comes from assuming the uncorrected bin contents obey Poisson statistics. To continue, we need to calculate the derivative that appears in the above equation. This can be done directly from Equation A.4:

$$\begin{aligned}\frac{\partial N_i^c}{\partial N_j} &= \delta_{ij} + 2\epsilon \left( \delta_{ij} \sum_k N_k + N_i - N_{i-j-1} \right) \\ \frac{\partial N_i^c}{\partial N_j} &= \delta_{ij} + 2\epsilon (\delta_{ij} I + N_i - N_{i-j-1}),\end{aligned}\tag{A.6}$$

where  $\delta_{ij}$  is the Kronecker delta and  $I$  is  $\sum_j N_j$ , the total number of hits in the time bin in question.

Equation A.5 and Equation A.6 together provide the uncertainty of a pileup corrected

energy bin. The expressions simplify greatly when dropping order  $\epsilon^2$  terms:

$$\begin{aligned} \left( \frac{\partial N_i^c}{\partial N_j} \right)^2 &= \delta_{ij} [1 + 4\epsilon (I + N_i)] + \mathcal{O}(\epsilon^2) \\ &\approx \delta_{ij} [1 + 4\epsilon I] \end{aligned}$$

The  $N_i$  term is dropped because  $I$ , the integral over all energy bins, is much larger than  $N_i$ , the number of hits in any given single energy bin. Completing the sum in Equation A.5 gives

$$\sigma_{N_i^c}^2 = N_i (1 + 4\epsilon I) + \mathcal{O}(\epsilon^2). \quad (\text{A.7})$$

So, to leading order in  $\epsilon$ , the energy bin variances are uncorrelated even after the pileup correction. It is interesting to see how the pileup corrected energy bin variances compare to  $N_i^c$ , what they would be in the case of Poisson statistics. Using Equation A.4 and dropping terms of order  $\epsilon^2$ , the result is

$$\sigma_{N_i^c}^2 = N_i^c \left[ 1 + \epsilon \left( 2I + \frac{1}{N_i^c} \sum_{j < i} N_j^c N_{i-j-1}^c \right) \right]. \quad (\text{A.8})$$

The bin uncertainty is increased relative to what it would be in the case of Poisson statistics. With a 2.5 ns deadtime and no spatial separation, the uncertainty enhancement—defined as  $\sigma_{N_i^c} / \sqrt{N_i^c}$ —is about 1.005 for most energies, but it increases dramatically for the energies with large pileup contamination factors, see Figure A.1.

### A.3 T-Method Uncertainty Enhancements

The T-Method counts all hits above a defined threshold. In other words, the number of pileup-corrected T-Method counts in a given time bin,  $N_T^c$ , is

$$N_T^c = \sum_{i \geq i_T} N_i^c, \quad (\text{A.9})$$

where  $i_T$  is the first over threshold bin. Though the previous section derived the variances of each bin  $N_i^c$ , they cannot simply be summed to get the T-Method bin variance  $\sigma_{N_T^c}^2$  because

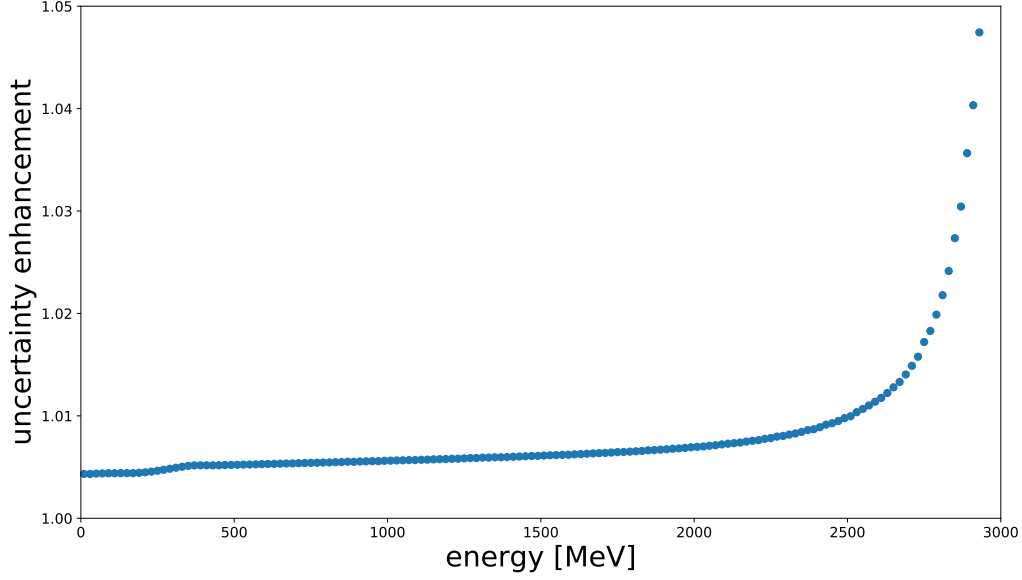


Figure A.1: Pileup uncertainty enhancements at  $30 \mu\text{s}$  by energy, calculated from the 60-Hour Dataset using a 2.5 ns reconstruction deadline and no spatial separation.

of the statistical correlations imparted by the pileup correction procedure. One approach to get the variance would be to calculate, by brute force,

$$\begin{aligned}\sigma_{N_T^c}^2 &= \sum_j N_j \left( \frac{\partial N_T^c}{\partial N_j} \right)^2 \\ \sigma_{N_T^c}^2 &= \sum_j N_j \left[ \sum_{i>i_T} (\delta_{ij} + 2\epsilon\delta_{ij}I + 2\epsilon N_i - 2\epsilon N_{i-j-1}) \right]^2.\end{aligned}\tag{A.10}$$

This works and can be computed relatively easily, but the nature of the T-Method allows for a simpler treatment.

As shown in Figure A.2, there are only two types of pileup pairs that are relevant to the T-Method: those in which both pulses are over threshold and those in which both pulses are under threshold but have energies summing to be over threshold. In any other case, the total number of hits will be recorded correctly. Therefore, the number of pileup corrected

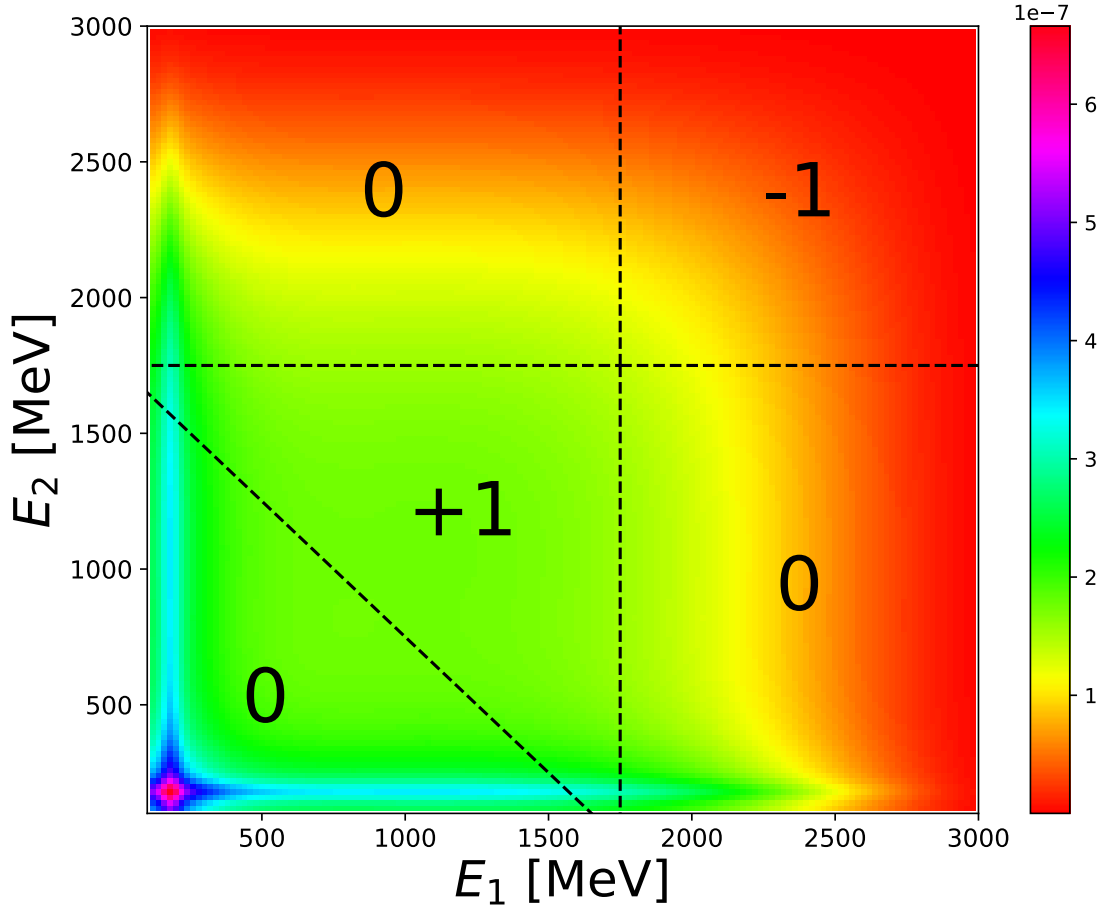


Figure A.2: Regions in  $E_1$  versus  $E_2$  space relevant to the T-Method pileup correction. The z-axis encodes the probability a pileup pulse pair will have energies  $(E_1, E_2)$ . This graphic assumes a 1.75 GeV threshold. If both pulses are over threshold, a count will be lost from the T-Method spectrum. If both pulses are under threshold but their energies add to be over threshold, a false count will be added. In all other cases, the correct number of counts will be recorded. Because the integral over the +1 region is greater than the integral over the -1 region, in general the T-Method pileup spectrum is positive.

T-Method counts can be written as

$$N_T^c = N_T + \epsilon [N_T^2 - N_{pair}], \quad (\text{A.11})$$

where  $N_T$  is the number of T-Method counts before the pileup correction, and  $N_{pair}$  is the number of under threshold pairs whose energies add to be over threshold. The quantities  $N_T^2$

and  $N_{pair}$  respectively correspond to the integrals over the  $-1$  and  $+1$  regions in Figure A.2. As these two quantities are derived from completely separate pulses, they have independent statistical fluctuations. Equation A.11 can also be derived algebraically from Equation A.4, as shown at the end of this appendix. To order  $\epsilon$ , then, the result is

$$\sigma_{N_T^c}^2 = N_T (1 + 4\epsilon N_T). \quad (\text{A.12})$$

Or, to order  $\epsilon$  and in terms of  $N_T^c$ ,

$$\sigma_{N_T^c}^2 = N_T^c \left[ 1 + \epsilon \left( 3N_T^c + \frac{N_{pair}^c}{N_T^c} \right) \right]. \quad (\text{A.13})$$

Figure A.3 shows the T-Method pileup uncertainty enhancements,  $\sigma_{N_T^c}/\sqrt{N_T^c}$  in the 60-Hour Dataset, calculated with both Equation A.10 and Equation A.13. The enhancement factor is about 1.004 at  $30 \mu s$  and displays complicated oscillatory behavior.

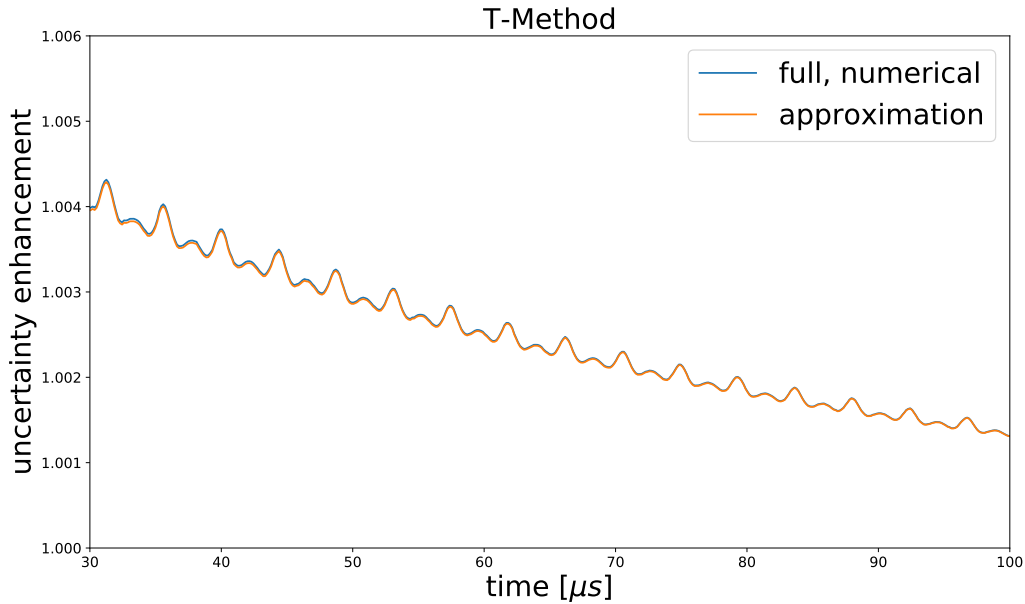


Figure A.3: 60-Hour Dataset T-Method Pileup uncertainty enhancement factors calculated with Equation A.10 and Equation A.13.

#### A.4 A-Weighted Uncertainty Enhancements

The A-Weighted histogram combines hits within a chosen energy range. Each hit is weighted with its asymmetry, which is a function of the hit energy. For sufficiently narrow energy bins, the pileup corrected asymmetry weighted histogram  $A_H^c$  is

$$A_H^c = \sum_{i=i_{min}}^{i_{max}} A_i N_i^c, \quad (\text{A.14})$$

where  $A_i$  is the asymmetry evaluated at the center of energy bin  $i$ . The variance is

$$\begin{aligned} \sigma_{A_H^c}^2 &= \sum_j N_j \left[ \sum_{i=i_{min}}^{i_{max}} A_i \left( \frac{\partial N_i^c}{\partial N_j} \right) \right]^2 \\ \sigma_{A_H^c}^2 &= \sum_j N_j \left[ \sum_{i=i_{min}}^{i_{max}} A_i (\delta_{ij} + 2\epsilon \delta_{ij} I + 2\epsilon N_i - 2\epsilon N_{i-j-1}) \right]^2. \end{aligned} \quad (\text{A.15})$$

The asymmetry's variation with energy prevents any convenient cancellation from occurring as it did in the T-Method case. Nevertheless, numerically computing the sum in Equation A.15 is relatively straightforward. To obtain an enhancement factor analogous to that shown in Figure A.3, the adjusted uncertainty should be compared to what the uncertainty would be in the case with no pileup contamination,  $\sqrt{\sum A_i^2 N_i^c}$ . Figure A.4 shows the A-Weighted and T-Method uncertainty enhancements. The A-Weighted enhancement is larger. This is attributable to the larger overall pileup contamination in the A-Weighted histogram, which is in turn rooted in its preferential weighting of high-energy positrons.

#### A.5 Effect on The Extracted $R$ Value

The pileup uncertainty enhancement factors should be accounted for when fitting a pileup-corrected histogram. Although an order one percent increased statistical uncertainty is completely negligible in the overall error budget, one may worry that the time-dependent enhancement factors could change the  $\chi^2$  optimum and thus shift  $R$  by an appreciable amount. However, this appears not to be the case. In the 60-Hour Dataset, fits including the pileup uncertainty enhancements and those assuming that pileup corrected histograms obey Pois-

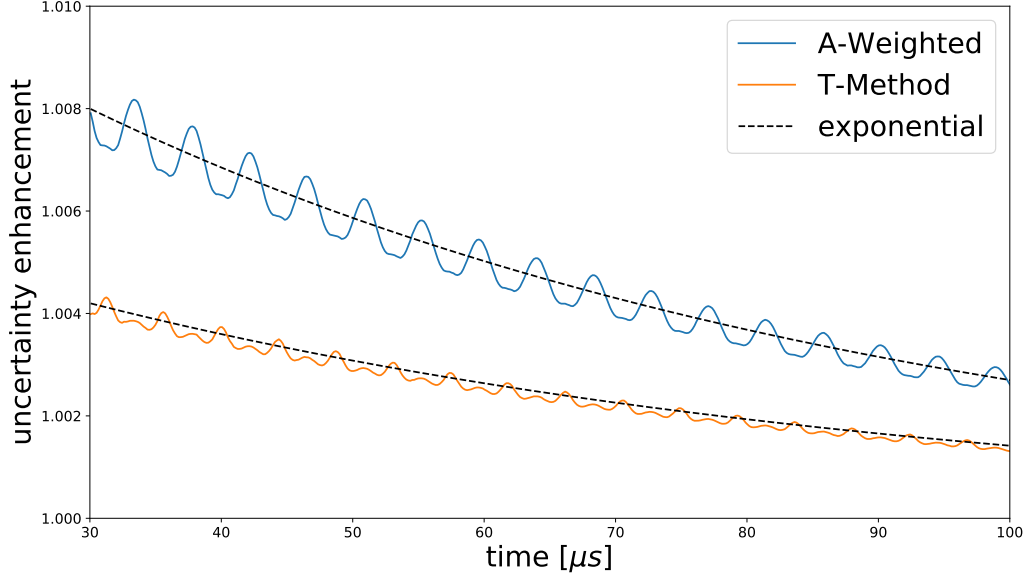


Figure A.4: 60-Hour Dataset T-Method and A-Weighted pileup uncertainty enhancement factors. The exponentials are drawn to guide the eye and fall with 64.44  $\mu s$  time constants.

son statistics yielded identical R values and nearly identical precisions, i.e. changes were less than 10 ppb in both cases.

### A.6 Algebraic Derivation of Equation A.11

Here Equation A.11 is derived directly from Equation A.4. Isolating only the perturbation to  $N_T$  caused by pileup,  $\delta N_T$ ,

$$\begin{aligned}
 \frac{\delta N_T}{\epsilon} &= \sum_{i>i_T} \left[ 2N_i \sum_j N_j - \sum_{j<i} N_j N_{i-j-1} \right] \\
 &= \sum_{i>i_T} \left[ 2N_i \left( \sum_{j>i_T} N_j + \sum_{j\leq i_T} N_j \right) - \sum_{j<i} N_j N_{i-j-1} \right] \\
 &= 2N_T^2 + 2 \sum_{i>i_T} \sum_{j\leq i_T} N_i N_j - \sum_{i>i_T} \sum_{j<i} N_j N_{i-j-1}.
 \end{aligned}$$



The last term above is the sum over all pulse pairs that combine to an energy greater than the T-Method threshold. It can be rewritten in the following way:

$$\sum_{i>i_T} \sum_{j<i} N_j N_{i-j-1} = \sum_i \sum_{j>i_T-i-1} N_i N_j.$$

Using this identity,

$$\begin{aligned} \frac{\delta_{N_T}}{\epsilon} &= 2N_T^2 + 2 \sum_{i>i_T} \sum_{j<=i_T} N_i N_j - \sum_i \sum_{j>i_T-i-1} N_i N_j \\ &= 2N_T^2 + 2 \sum_{i>i_T} \sum_{j<=i_T} N_i N_j - \sum_{i>i_T} \sum_{j>0} N_i N_j - \sum_{i<=i_T} \sum_{j>i_T-i-1} N_i N_j \\ &= 2N_T^2 + 2 \sum_{i>i_T} \sum_{j<=i_T} N_i N_j - \sum_{i>i_T} \sum_{j>i_T} N_i N_j - \sum_{i>i_T} \sum_{j<=i_T} N_i N_j - \sum_{i<=i_T} \sum_{j>i_T-i-1} N_i N_j \\ &= N_T^2 + \sum_{i>i_T} \sum_{j<=i_T} N_i N_j - \sum_{i<=i_T} \sum_{j>i_T-i-1} N_i N_j. \end{aligned}$$

The sum over  $j$  in the last term above can be divided into two parts: one in which  $j > i_T$  and one in which  $i_T - i - 1 < j \leq i_T$ . Doing so,

$$\frac{\delta_{N_T}}{\epsilon} = N_T^2 + \sum_{i>i_T} \sum_{j<=i_T} N_i N_j - \sum_{i<=i_T} \sum_{j>i_T} N_i N_j - \sum_{i<=i_T} \sum_{i_T-i-1 < j \leq i_T} N_i N_j.$$

The second and third terms above are equal and opposite, except with the index labels swapped. The fourth term is the sum over all pulse pairs with an energy sum greater than the T-Method threshold but in which both pulses individually are under the T-Method threshold. This is the quantity from Section A.3 called  $N_{pair}$ . So, finally,

$$\frac{\delta_{N_T}}{\epsilon} = N_T^2 - N_{pair}.$$

This matches Equation A.11.

## Appendix B

### CORRELATION MATRICES FROM THE T-METHOD AND A-WEIGHTED FITS

This appendix contains the full parameter correlation matrices for the A-Weighted and T-Method fits to the E989 60-Hour Dataset, described in Chapter 6. Both tell the same qualitative story:  $R$ , the blinded  $\omega_a$  value, displays no strong correlation with any parameter other than  $\phi$ , the  $g - 2$  phase. Other strong correlations are between  $N_0$ ,  $\tau$ , and  $K_{loss}$ , and internally among the  $CBO$  parameters and among the  $VW$  parameters. There is no strong correlation between the  $VW$  parameters and the  $CBO$  parameters. See Chapter 6 for a description of all the fit parameters.

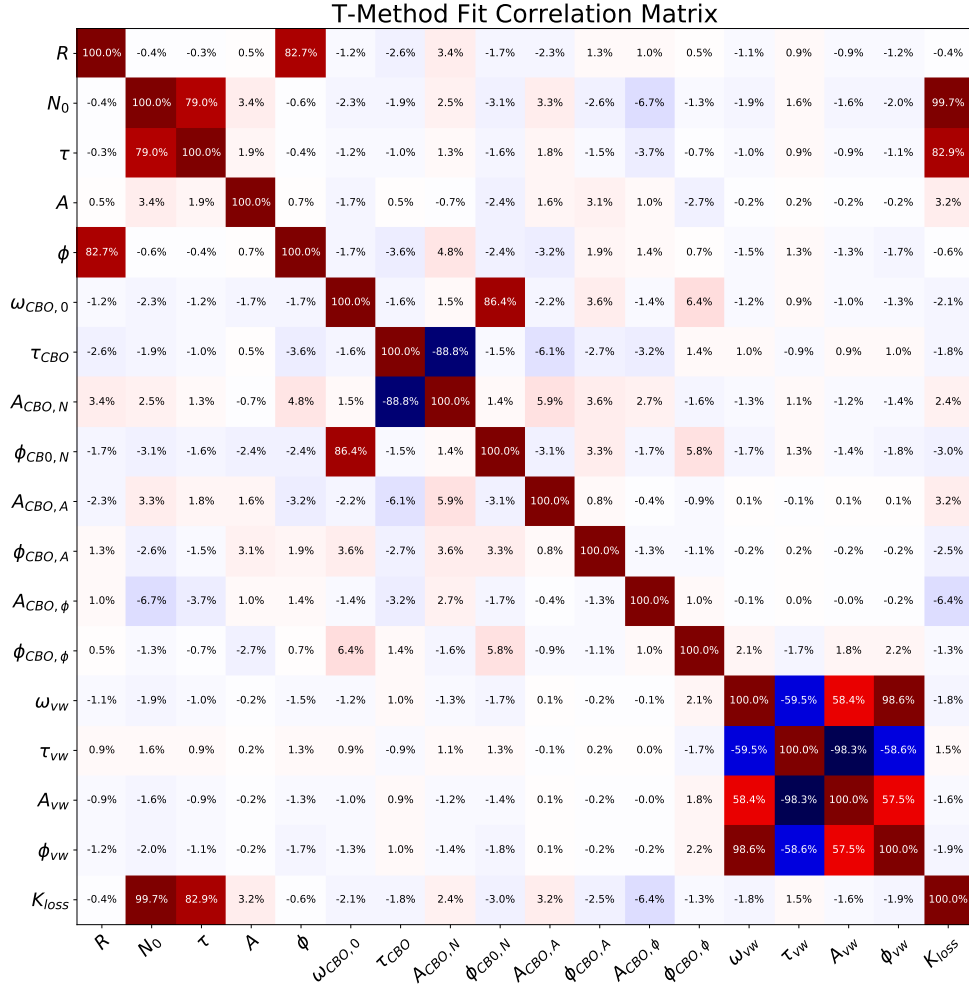


Figure B.1: Full parameter correlation matrix from the T-Method fit to the E989 60-Hour Dataset, described in Section 6.6.1.

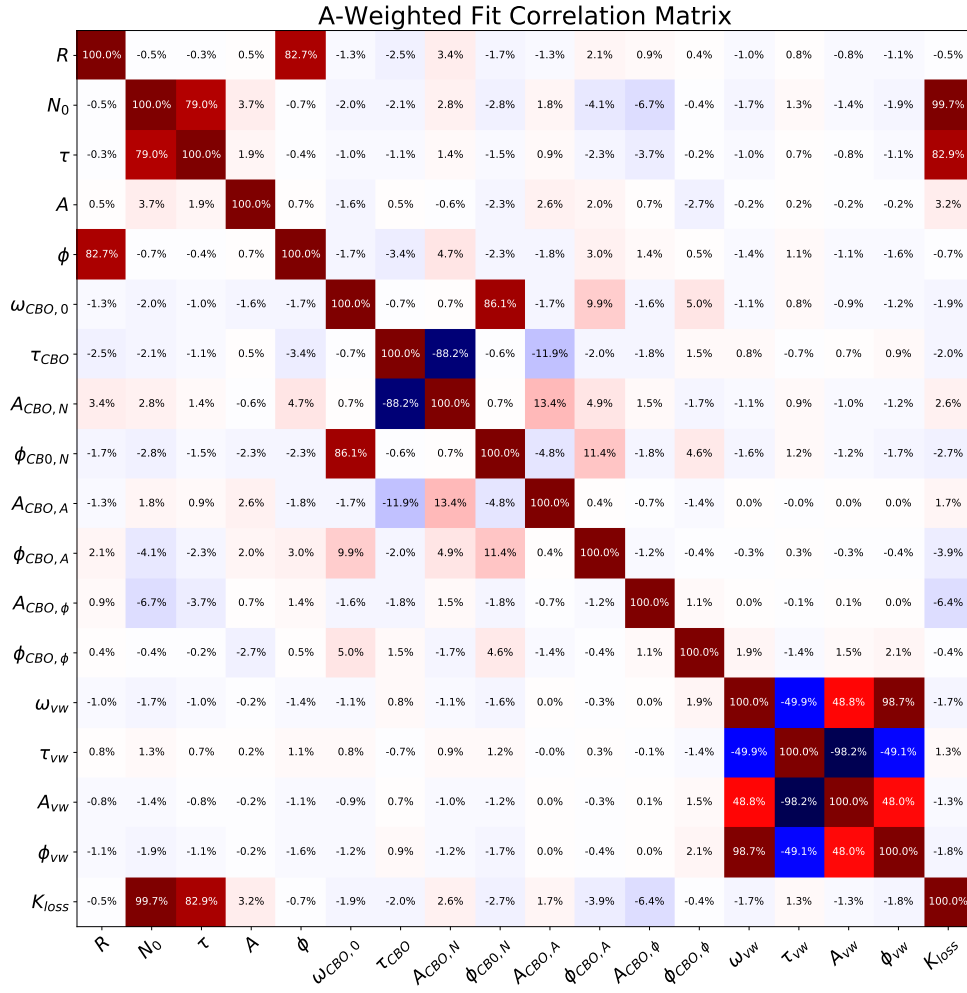


Figure B.2: Full parameter correlation matrix from the A-Weighted fit to the E989 60-Hour Dataset, described in Section 6.6.3.