# UNIVERSITÀ DI PISA

**Università degli Studi di Pisa**

Dipartimento di Fisica E. Fermi
Corso di Dottorato di Ricerca in Fisica - XXXV Ciclo

# Measurement of the muon anomalous precession frequency at the Muon $g - 2$ Experiment at Fermilab

Candidato:
**Paolo Girotti**
**Matricola 595929**

Supervisor:
**Dott. Marco Incagli**

**Anno Accademico 2021–2022**

# Summary

The anomalous magnetic moment of the muon, $a_\mu = \frac{g-2}{2}$ is the fractional deviation of the muon $g$-factor from the value of 2. It emerges as the cumulative effect of the virtual particles participating in the muon interaction with a magnetic field via quantum loop corrections. Its value encodes all the possible interactions between the virtual particles and, for this reason, represents an important test of the Standard Model (SM). In particular, any deviation from the SM theoretical evaluation could be due to new physics contributions.

The new Muon $g - 2$ (E989) Experiment at Fermilab is currently operating to repeat and improve the previous E821 experiment at Brookhaven National Laboratory (BNL), aiming to reduce the experimental error by a factor of 4 to the final accuracy of 140 parts per billion (ppb). On April 7th, 2021, the E989 collaboration published the first result based on the first year of data taking (Run-1), measuring $a_\mu = 0.001\ 165\ 920\ 40(54)$ with a precision of 460 ppb. The measured value is consistent with the BNL measurement and strengthens the long-standing tension with the data-driven SM prediction to a combined discrepancy of $4.2\sigma$. On the theory side, however, new efforts involving lattice-QCD techniques are starting to question the current consensus on the theoretical prediction, demanding new improvements on both the experimental and theoretical sides. The E989 collaboration is now finalizing the analysis of Run-2 and Run-3 data and a new publication is expected in the first half of 2023 with a combined statistical uncertainty of 200 ppb.

The anomalous magnetic moment $a_\mu$ is measured as the ratio between the muon spin anomalous precession frequency, $\omega_a$, and the average magnetic field experienced by the muons as they circulate in the storage ring. This thesis presents a precession frequency analysis of the Run-1 data and an evaluation of the related systematic uncertainties. A new positron reconstruction developed for the analysis of the subsequent data-taking periods, aiming to reduce some of the major systematic uncertainties of the $\omega_a$ measurement, is presented. The author's involvement in the production of the Run-2/5 data and in the precise calibration of the detectors is discussed. Finally, the complete Run-1 $a_\mu$ result is presented.

# Contents

# Chapter 1

# Introduction

The story of our understanding of the subatomic world begins with the discovery of the electron by J. J. Thomson in 1897. More than a century of fascinating experiments and revolutionizing theories have led to the current Standard Model (SM) of particle physics. Nearly all the constituent parts of this theory and its implications have been validated experimentally over time, with the discovery of the top quark (1995), the tau neutrino (2000), and the Higgs boson (2012) being among the most recent ones.

Despite the astonishing number of physical phenomena that can be described with the SM, however, some questions about the subatomic world remain unanswered. As an example, the model cannot describe gravitation and the presence of dark matter and dark energy postulated by cosmological observations.

Many experiments around the world are now trying to perform tests of the SM which require high volumes of data. There are two main strategies: the search for interactions that should be suppressed within the SM but can be enhanced by extensions of the model, and high-precision measurements of fundamental quantities in search of discrepancies with the values predicted by the SM.

In the latter category, one of the quantities that have received continuous improvements from both the experimental and theoretical standpoint over the last decades is the magnetic moment of the muon. The great interest in this physical property comes from the fact that its value encodes all of the existing interactions that can happen between subatomic particles. Its precise measurement, when compared to the theoretical prediction, serves as a test of the completeness of the SM. Any observed discrepancy, otherwise, would suggest the presence of *new physics*. In any case, the measurement would constrain the formulation of any extension to the current model.

In this dissertation, I will present the experimental principles of the measurement of the muon anomaly and my contribution to the first results delivered by the Muon $g - 2$ Experiment in operation at Fermilab.

This chapter will present the theoretical background and a brief historical digression of the muon measurements, and Chapter 2 will describe the experimental principles that enable a precise measurement of the muon anomaly. Chapter 3 will present the new Muon $g - 2$ Experiment at Fermilab. Chapter 4 will discuss the production workflow of the E989 data and my involvement as *production manager*. Chapter 5 will present a new reconstruction of the positrons, of which I led the development. Chapter 6 will describe the data analysis of the muon anomalous precession frequency and Chapter 7 will present the results and the evaluation of the systematic uncertainties. Finally, Chapter 8 will present the complete muon anomaly measurement based on Run-1 data that was published in April 2021.

## 1.1   The magnetic moment

In classical physics, the magnetic dipole moment $\vec{\mu}$ of a charged object is a measure of how much torque it experiences when placed in a magnetic field:

$$\vec{\tau} = \vec{\mu} \times \vec{B} \,, \tag{1.1}$$

and can be expressed as a function of the object charge $q$, mass $m$, and angular momentum $\vec{L}$:

$$\vec{\mu} = \frac{q}{2m} \vec{L} \,. \tag{1.2}$$

For subatomic particles, the classical Equation 1.2 becomes

$$\vec{\mu} = g \frac{q}{2m} \vec{S} \,, \tag{1.3}$$

where $\vec{S}$ is the spin of the particle and $g$ is the so called *Landé g-factor*. The $g$ factor is an observable that can be measured by placing the charged particle in a magnetic field and observing the rate at which it precesses, according to Equation 1.1.

**Anomalous magnetic moments**

According to the Dirac theory, $g$ is equal to 2 for an elementary particle with spin $\frac{1}{2}$ [1, 2]. This would be true if the particle only interacted with the magnetic field. However, the magnetic moment cannot be measured without the influence from virtual exchanges made explicit in quantum field theories. All particles are surrounded by a cloud of virtual particles that are temporarily created from the vacuum. The *screening* effect given by this cloud leads to a slight change in the observed magnetic moment from what would be expected for a bare particle. The difference between the observed $g$ value and 2 is called magnetic anomaly and is conventionally written as a fractional deviation:

$$a = \frac{g - 2}{2} \,.\tag{1.4}$$

The first evidence of this anomaly is due to Kusch and Foley who measured in 1948 a value of $a_e = 0.00119(5)$ for the electron [3], giving no space to doubt in the result given such a small error. At about the same time, quantum electrodynamics (QED) was reaching a turning point. For almost two decades, QED calculations were impeded by the problem of non-convergent infinite series. The solution of renormalization was proposed by Feynman, Schwinger, and Tomonaga, and by the end of 1948 Schwinger had calculated [4] the first-order correction to the magnetic moment to yield $a_e = \frac{\alpha}{2\pi} = 0.00116$, which is within the quoted error of the Kusch and Foley experiment.

The true value of $g$ depends on all the possible interactions that can happen within the virtual cloud influencing the interaction between the particle and the magnetic field. The main contribution is given by QED interactions, whose terms scale as powers of the fine structure constant $\alpha$. The more particles are involved in the virtual interaction, the smaller the contribution to the value of $g$. The current precision of the experimental values of the electron magnetic anomaly requires the calculation to the fifth order in $\alpha$ for a comparison between experiment and theory. The most recent calculation is [5, 6] :

$$
\begin{aligned}
a_e^{th}(2017) = {} & 0.5 \left(\frac{\alpha}{\pi}\right) - 0.328478965579 \left(\frac{\alpha}{\pi}\right)^2 \\
& + 1.181241456 \left(\frac{\alpha}{\pi}\right)^3 - 1.912245764 \left(\frac{\alpha}{\pi}\right)^4 \\
& + 6.599(223) \left(\frac{\alpha}{\pi}\right)^5 + 1.74(2) \cdot 10^{-12} \\
= {} & 0.001\,159\,652\,182\,031\,(720) \,.
\end{aligned}\tag{1.5}
$$

The last term arises from the contributions of Electro-Weak (EW) and Quantum Chromo Dynamics (QCD) interactions. The theoretical value of $a_e$ depends on the precise measurement of $\alpha$, which is the major source of uncertainty in Equation 1.5. Recent measurements based on the recoil of Cesium and Rubidium atoms give statistically different values [7, 8]:

$$\alpha^{-1}(\text{Cs; 2018}) = 137.035\,999\,046\,(27)\,, \tag{1.6a}$$

$$\alpha^{-1}(\text{Rb; 2020}) = 137.035\,999\,206\,(11)\,, \tag{1.6b}$$

and the tension of $5.4\sigma$ is currently unresolved.

The experimental measurement of $a_e$ has recently reached the incredible precision of 0.1 ppb [9]:

$$a_e(\text{Exp; 2023}) = 0.001\,159\,652\,180\,59\,(13) \tag{1.7}$$

which in turn can be used to predict [9]:

$$\alpha^{-1}(a_e;\,2023) = 137.035\,999\,166\,(15)\,. \tag{1.8}$$

Table 1.1 lists the measured *g-factor* values for different particles, as well as the theoretical predictions as of 2022.

| Particle | Experimental value | Precision | Ref. | Theoretical prediction | Ref. |
|---|---|---|---|---|---|
| Electron | 2.00231930436118(26) | $1 \cdot 10^{-13}$ | [10] | 2.002319304364(1) | [6] |
| Muon | 2.00233184122(82) | $4 \cdot 10^{-10}$ | [11] | 2.00233183620(86) | [19] |
| Tau | 2.036(34) | $1.7 \cdot 10^{-2}$ | [10] | 2.00235442(10) | [13] |
| Proton | 5.5856946893(16) | $3 \cdot 10^{-10}$ | [14] | 5.58 | [15] |
| Neutron | -3.82608545(90) | $2 \cdot 10^{-7}$ | [14] | -3.72 | [15] |

Table 1.1: Experimental and theoretical values of $g$ for various particles.

## 1.2   The muon anomaly

Among the particles of the Standard Model, recent years have been characterized by an increasing interest in the measurement of the magnetic anomaly of the muon.

The muon is usually considered as a heavier version of the electron. Both particles are charged leptons: the lepton universality dictates that they are identical except for mass and the fact that muons are intrinsically unstable. Together with the tauon, they complete the charged lepton family. Table 1.2 enumerates lepton properties compiled by

the Particle Data Group [10].

| Property | e | $\mu$ | $\tau$ |
|---|---|---|---|
| Charge | -1 | -1 | -1 |
| Spin | 1/2 | 1/2 | 1/2 |
| Mass | 0.511 MeV/$c^2$ | 105.7 MeV/$c^2$ | 1776.9 MeV/$c^2$ |
| Lifetime | $> 2.0 \cdot 10^{36}$ s | $2.2 \cdot 10^{-6}$ s | $290.3 \cdot 10^{-15}$ s |
| Main decay mode | - | $\mu \to e\bar{\nu}_e\nu_\mu$ | $\tau \to \mu\bar{\nu}_\mu\nu_\tau$ |
| Branching ratio | - | $\sim 100\%$ | $17.39\%$ |

Table 1.2: Lepton properties

A precise measurement of the muon anomaly $a_\mu$ is of great interest because it probes all of the interactions between the lepton and virtual particles. The whole Standard Model is tested, since all possible particles can contribute either via direct virtual interaction with the muon or with higher-order loop corrections, adding to the value of $a_\mu$. Figure 1.1 depicts the lower-order interactions for the electromagnetic, weak, and strong forces.
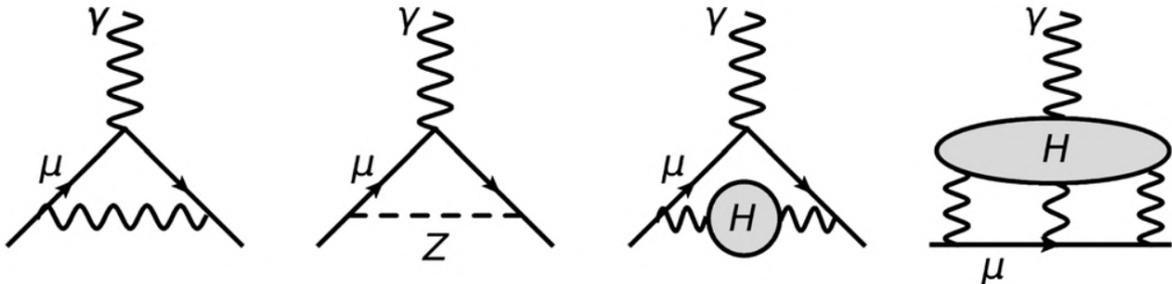


Figure 1.1: Lower-order examples for EM, weak, and QCD interactions. Figure from [11].

Moreover, eventual measurements that differ from the SM theoretical prediction would indicate hints of *new physics*. Such results would help set new constraints on the multitude of currently proposed Beyond Standard Model (BSM) theories.
Muons are particularly important in such a measurement because BSM interactions with massive particles contribute with mass suppression terms, $\propto (\frac{m_{lepton}}{M})^2$. The electron $g-2$ has already been measured with around $\mathcal{O}(10^3)$ times more precision than the muon [16], but the relative mass ratio between the electron and the muon enhances the sensitivity to these terms by a factor of $(\frac{105.66}{0.511})^2 \approx 43000$ [17]. Following the same reasoning, $\tau$ leptons would be even better for detecting new heavy particles, but their very short

lifetime ($2.9 \times 10^{-13}$ s), multiple decay branches, and other practical reasons make such an experiment not feasible with current technology.

## 1.3   Standard Model contributions to $a_\mu$

As anticipated in the previous section, the value of $a_\mu$ arises from the exchange of virtual particles. The contributions that affect this value come from many sectors of theory, that are Quantum ElectroDynamics (QED), ElectroWeak theory (EW), and Quantum ChromoDynamics (QCD), with the latter further divided into three classes. The lowest-order contribution arises from Hadronic Vacuum Polarization (HVP). Higher-Order contributions (HOHVP) contain an HVP insertion along with an additional loop. The last class is Hadronic Light-By-Light (HLBL) scattering.

$$a_\mu^{SM} = a_\mu^{QED} + a_\mu^{EW} + a_\mu^{HVP} + a_\mu^{HOHVP} + a_\mu^{HLBL} \tag{1.9}$$

As seen in Table 1.4 the dominant contribution comes by far by QED, being nearly two million times the experimental error. On the other hand, the error on the Standard Model comes almost entirely from the hadronic terms.

### 1.3.1   QED terms

The QED contribution to the value of $a_\mu$ originates from loops containing virtual leptons and photons, and can be evaluated as a perturbative expansion in $\alpha/\pi$:

$$a_\mu^{QED} = \sum_{n=1}^{\infty} C_n \left(\frac{\alpha}{\pi}\right)^n \tag{1.10}$$

where the $C_n$ coefficients are finite thanks to the renormalizability of QED. The first coefficient $C_1 = \frac{1}{2}$ is the Schwinger term and represents the contribution of a single virtual photon exchanged between the two muon legs as depicted in the leftmost diagram of Figure 1.1.

These loops are well understood and have been calculated up to the fifth order from over than 13'000 Feynman diagrams. Some examples of fourth and fifth-order QED contributions are shown in Fig. 1.2. The current best estimate for the QED contribution

to $a_\mu$ is [18, 19]:

$$a_\mu^{QED} = 0.5 \left(\frac{\alpha}{\pi}\right) + 0.765857425(17) \left(\frac{\alpha}{\pi}\right)^2 + 24.05050996(32) \left(\frac{\alpha}{\pi}\right)^3$$
$$+ 130.8796(63) \left(\frac{\alpha}{\pi}\right)^4 + 753.29(1.04) \left(\frac{\alpha}{\pi}\right)^5$$
$$= 116\,584\,718\,931(104) \cdot 10^{-14} , \tag{1.11}$$

where the error mainly comes from the uncertainty in $\alpha$, which is:

$$\alpha^{-1} = 137.035\,999\,046(27) \rightarrow 0.2 \text{ ppb} \tag{1.12}$$

obtained from the precise measurement of the recoil velocity of Cesium-133 $h/M$, the Rydberg constant and $M/m_e$ [20]:

$$\alpha^2 = \frac{2R_\infty}{c} \frac{M}{m_e} \frac{h}{M} . \tag{1.13}$$

While over 99% of the value of $a_\mu$ comes from the QED sector, the precision at which it is possible to calculate the QED loops results in an error that is much smaller than the EW and hadronic contributions (Table 1.4).



(a)                                        (b)
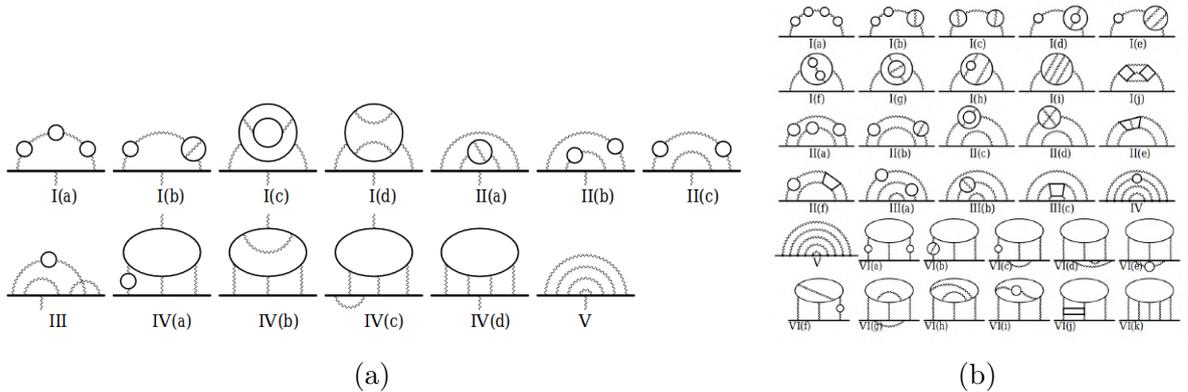
Figure 1.2: Vertex diagrams representing 13 (a) and 32 (b) gauge-invariant subsets contributing to the lepton $g - 2$ for QED at the eighth (a) and tenth (b) order. Solid and wavy lines represent lepton and photon lines, respectively. Figures from [18].

### 1.3.2 EW term

The corrections due to the weak force are very small compared to the QED contribution. In the EW diagrams the heavy masses of the gauge bosons will produce contributions

with characteristic scales of $\left(\frac{m_\mu}{m_{Z^0,H,W^\pm}}\right)^2$. Because the masses of $Z^0$, $H$, and $W^\pm$ are much larger than the muon $m_\mu$, these processes are suppressed. The lowest order and largest contributions to the weak corrections are represented by the two diagrams pictured in Fig. 1.3. The diagram on the left is similar to the Schwinger diagram, but with



Figure 1.3: The largest contributing diagrams from the weak interaction.

a Z (or Higgs) boson propagating instead of a photon. The single loop contributions first calculated in 1972 [21] yield

$$a_\mu^{EW}(1-loop) = 194.8 \cdot 10^{-11} \tag{1.14}$$

The term containing the W boson contribution is suppressed by a term $(m_\mu/M_W)^2$ and it introduces a perturbation of approximately 3.3 ppm. On the other way, the single loop with an exchange of a Z boson reduces the overall EW contribution with a negative perturbation of -1.6 ppm, while the remaining loop with a Higgs boson ($M_H \sim 125$ GeV) doesn't contribute significantly due to the mass suppression. The total contributions from the second order weak corrections sum up to $-40 \cdot 10^{-11}$ [17], so that the overall EW term is [22]

$$a_\mu^{EW} = 153.6(1.0) \cdot 10^{-11} . \tag{1.15}$$

### 1.3.3   QCD terms

The hadronic contributions to $a_\mu$ arise from loop diagrams with virtual hadrons. The strong coupling is large at low energies, and the contributions from these processes cannot be calculated perturbatively as for QED and EW. Moreover, the low energy processes dominate the hadronic contribution. For these reasons the calculations are difficult to be carried out and a semi-phenomenological approach is required. As a result, the current uncertainty on $a_\mu^{QCD}$ dominates the Standard Model value of $a_\mu$.

The hadronic contribution to $a_\mu$ can be separated into two parts:

$$a_\mu^{QCD} = a_\mu^{HVP} + a_\mu^{HLbL} \,, \tag{1.16}$$

where the first term is related to the Hadronic Vacuum Polarization and the second to the Hadronic Light-by-Light scattering; both processes are depicted in the third and fourth diagrams of Figure 1.1.

**Hadronic Vacuum Polarization**
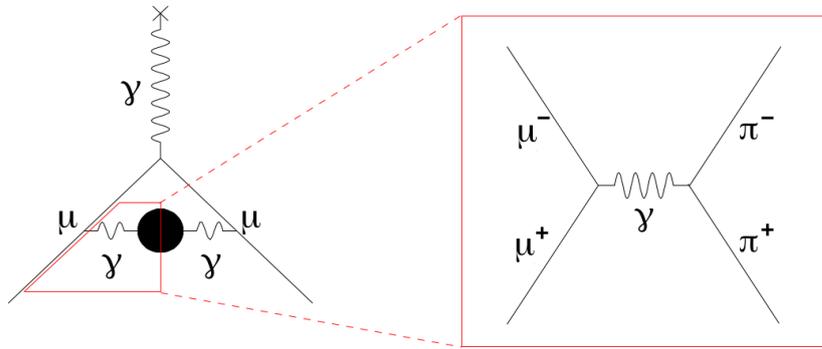


Figure 1.4: The Feynman diagram on the left shows the lowest-order Hadronic Vacuum Polarization (HVP), where the black circle represents any possible combination of quark matter.

The general form of Hadronic Vacuum Polarization (HVP) is quite similar to the QED one. The muon radiates a photon or another boson, which creates a particle pair that annihilates before being recaptured with the muon, as seen in Fig. 1.4. The difference with QED is that the particle pair is composed of hadronic matter, such as $\pi^0$, $\pi^\pm$, $\rho^0$, etc. These loop terms can be calculated relating to the cross section of hadron production from the annihilation of $e^+e^-$, using the dispersion relation [23, 24, 25, 26]:

$$a_\mu^{HVP} = \frac{1}{3}\left(\frac{\alpha}{\pi}\right)^2 \int_{4m_\pi^2}^{\infty} \frac{\mathrm{d}s}{s} \frac{\sigma_{e^+e^- \to hadrons}(s)}{\sigma_{e^+e^- \to \mu^+\mu^-}(s)} K(s) \,, \tag{1.17}$$

where $K(s)$ is the kinematic factor which can be expressed as:

$$K(s) = \int_0^1 dx \frac{x^2(x-1)}{x^2 + (1-x)\frac{s}{m_\mu^2}} \,. \tag{1.18}$$

The current recommended values for the leading order and higher orders contributions

of the hadronic vacuum polarization calculated using the data-driven $\sigma_{e^+e^-\to hadrons}$ values are summarized in Table 1.3. The overall HVP contribution is listed as well.

| Term | Value $[\times 10^{-11}]$ |
| --- | --- |
| $a_\mu^{HVP,LO}$ | $6931 \pm 40$ |
| $a_\mu^{HVP,NLO}$ | $-98.3 \pm 0.7$ |
| $a_\mu^{HVP,NNLO}$ | $12.4 \pm 0.1$ |
| $a_\mu^{HVP}$ | $6845 \pm 40$ |

Table 1.3: Hadronic Vacuum Polarization contributions to the value of $a_\mu$ as calculated with the dispersion relation of Equation 1.17. Values from [19].

**Lattice calculation**

An independent way to compute the QCD contributions to the muon anomaly is based on the Lattice technique. This approach describes space-time as a discrete grid where each point is separated by a spacing $D$. The computation is performed in this simulated lattice and, by extrapolating the results for $D \to 0$ (and for volume $\to \infty$), the value corresponding to the real universe is extracted. This method does not rely on any other experimental measurement, such as the $\sigma_{e^+e^-\to hadrons}$ cross sections for the dispersive approach described above. In recent years, also thanks to the increase of available computational resources, this *ab-initio* approach has reached uncertainties comparable with the other methods and the experimental values.

The first group to publish a result with competitive precision was the Budapest-Marseille-Wuppertal (BMW) collaboration on April 7th, 2021, the same day as the first publication from the Fermilab $g - 2$ collaboration, with a value of $a_\mu^{HVP,LO} = 7075(55) \cdot 10^{-11}$ [27]. The value presents a $2.2\sigma$ tension with the data-driven approach listed in Table 1.3 and, when added to the overall calculation, brings $a_\mu$ closer to the experimental value.

Very recently, three other groups provided preliminary results on the same quantity measured in a reduced region of energies which accounts for $\sim 30\%$ of the total value [30, 31, 29], all in agreement with the BMW results. Figure 1.5 shows the current predictions for this value. The tension that is now consolidating between the two theoretical approaches for the estimation of $a_\mu^{HVP}$ is being referred as *the new g − 2 puzzle* and remains unexplained as of today.
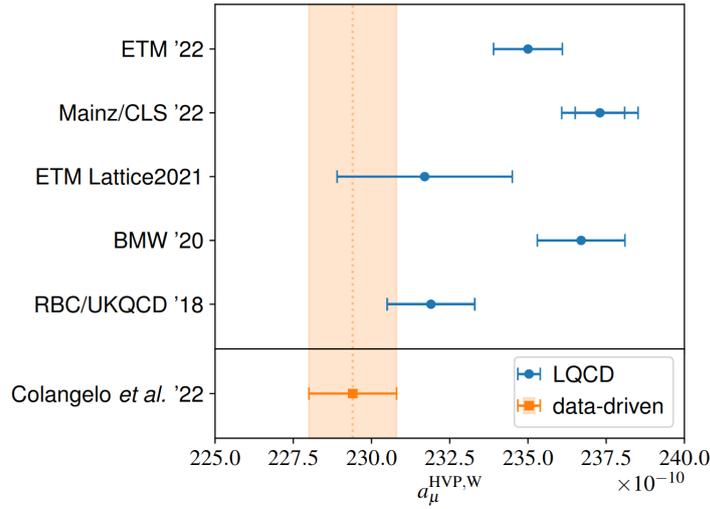
Figure 1.5: Current lattice-QCD predictions for $a_\mu^{HVP,LO}$ in a reduced region of energies. The estimate based on the data-driven approach [28] is shown in orange. Figure from [29].

**Hadronic Light-by-Light**

The general form for hadronic light-by-light (HLbL) contains more interaction vertices than HVP and therefore contributes to a lesser extent to the total muon anomaly. The propagating muon interacts with three photons and those photons interact with a QCD loop which interacts with the external field (Figure 1.6). The HLbL scattering differs from other hadronic contributions in that it cannot be related to experiments, so it must be estimated from a model. Several evaluations have taken place, both with the data-driven and lattice-QCD approaches, with the "best estimate" value currently accepted of [19]:

$$a_\mu^{HLbL} = 92(18) \cdot 10^{-11} \, . \tag{1.19}$$
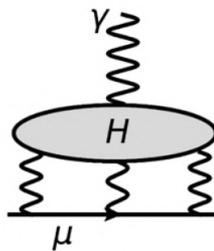


Figure 1.6: Feynman diagram for the general case of hadronic light-by-light scattering. Figure from [11].

### 1.3.4   The Standard Model prediction for $a_\mu$

The history of the theoretical prediction of the muon anomaly has now spanned for more than 70 years, and it is far from coming to an end. As the new computing capabilities now enable previously impossible calculations, and experiments continue to gather new quality data, the precision of the Standard Model prediction is constantly improving. However, as mentioned in Section 1.3.3, new techniques based on lattice simulations of QCD are bringing new tensions within the field.

In 2017 a collaboration was formed, the Muon $g-2$ Theory Initiative[1], whose goal is to combine all the calculations relative to $a_\mu$ and provide a consensus on a combined value. The latest publication from the group is from 2020 [19], and the values for the QED, EW, and QCD contributions to $a_\mu$, as well as the total Standard Model prediction are listed in Table 1.4. The quoted values do not currently include the lattice-QCD results.

| Term | Value ($\cdot 10^{-11}$) |
|---|---|
| $a_\mu^{QED}$ | $116\,584\,718.931 \pm 0.104$ |
| $a_\mu^{EW}$ | $153.6 \pm 1.0$ |
| $a_\mu^{HVP}$ | $6\,845 \pm 40$ |
| $a_\mu^{HLBL}$ | $92 \pm 18$ |
| Total SM | $116\,591\,810 \pm 43$ |

Table 1.4: Standard model contributions to $a_\mu$ [19].

## 1.4   Experimental value of $a_\mu$

The history of the experimental attempts to measure the muon magnetic moment went hand-in-hand with the theoretical development since the middle of 20th century. At each step, both the theoretical predictions and the experimental confirmations added digits to the expected and measured values of $a_\mu$. The muon anomaly is now one of the most precisely measured quantities in physics. This section will briefly describe the various experiments that marked the path up to the current Muon $g-2$ Experiment at Fermilab, which is the subject of this dissertation.

---

[1]https://muon-gm2-theory.illinois.edu/

### 1.4.1   Early experiments

The first muon experiment was performed in 1957 by Garwin and collaborators at the Nevis cyclotron of Columbia University [32]. The experiment observed the Larmor precession of single muons stopped in a carbon target submerged in a magnetic field. The magnitude of the field could be tuned to increase or decrease the rate of precession. This experiment was able to determine a value for $g_\mu = 2.00 \pm 0.10$ for the muon by correlating the counts measured from a fixed counter with the strength of the magnetic field following Equation 1.20 as shown in Figure 1.7:

$$\omega_s = g_\mu \frac{eB}{2m_\mu c} \tag{1.20}$$

.
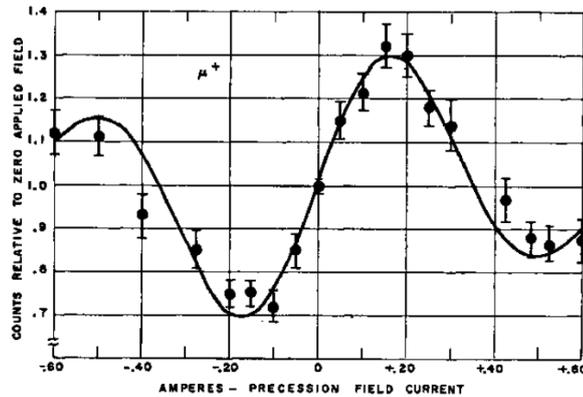


Figure 1.7: Larmor precession data from the Garwin experiment used to determine the muon *g-factor*. Figure from [32].

Similar experiments continued during the next years with the aim to improve the precision of this measurement. An experiment at the cyclotron of the University of Liverpool managed to measure the value $g_\mu = 2.004 \pm 0.014$ by selecting stopped muons with a coincidence of three counters and observing the number of decay positrons hitting the forward detector through time [33]. Figure 1.8 shows the apparatus and the positron plot.

So far, the precision of these experiments was not sufficient to measure anomalous contributions to $g_\mu$.

In 1963 another experiment was performed at Columbia University, with a technique similar to the Garwin experiment. In this experiment, the magnetic field was also measured via nuclear magnetic resonance (NMR) in terms of the Larmor precession frequency of protons in a polarized water sample. The result was presented as the ratio $\lambda$ of the two
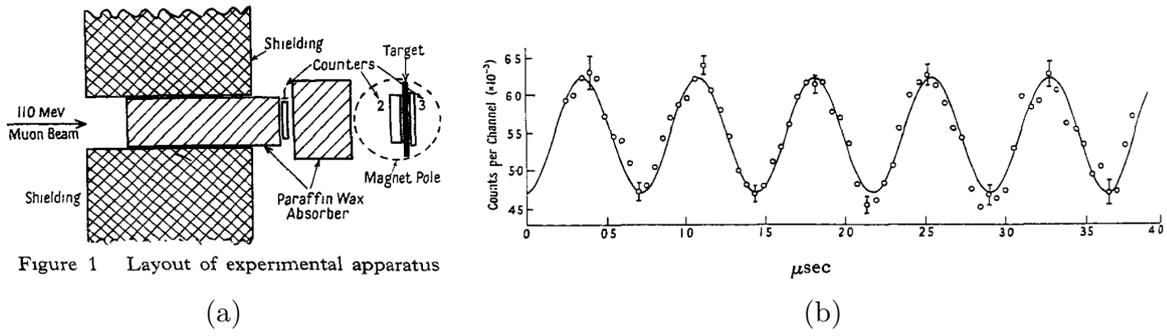
Figure 1.8: (a) Experimental apparatus of the Liverpool experiment. (b) Time distribution of emitted positrons from stopped positive muons in a ∼100 Gauss magnetic field. The exponential decay factor has been removed. Figures from [33].

frequencies or magnetic moments obtaining:

$$\lambda = \frac{\omega_\mu}{\omega_p} = \frac{\mu_\mu}{\mu_p} = 3.18338(4) \, . \tag{1.21}$$

The result corresponds to a measured *g-factor* of:

$$g_\mu = 2.002\,353(25) \rightarrow a_\mu = 0.001\,177(13) \tag{1.22}$$

which is the first measurement of $a_\mu$ and is in agreement with the Schwinger prediction of the first order QED contribution.
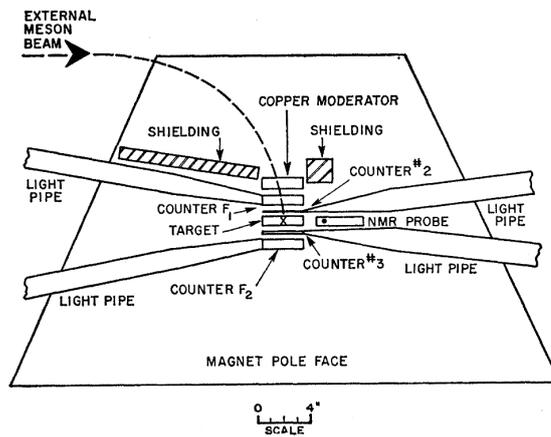


Figure 1.9: Experimental apparatus of the Columbia University experiment by Hutchinson et al.. Figure from [34].

### 1.4.2   Experiments at CERN

In the same years, Lederman and other collaborators were launching a Muon $g - 2$ measurement campaign at CERN. Overall, three experiments from 1961 to 1979 were carried out, each one with increasing precision.

**CERN I**

The first, CERN I, was set to reach an experimental precision of 1% on $a_\mu$. The experiment consisted of a muon beam injected into a dipole magnet that was 6 m long and wide enough to contain the entire muon orbit. The magnet contained a small gradient to slowly drift the muon circular orbits toward the far end of the magnet, as depicted in Figure 1.10. A higher gradient toward the end facilitated the final ejection of the muons into a methylene-iodide stopping target. A set of counters placed before and after the target would then provide the trigger for an incoming muon and measure the decay positron. The direction of the emitted positron is correlated to the direction of the muon spin, which depends on the time spent preceding inside the magnet. The result was:

$$a_\mu(\text{CERN-I}) = 0.001162(5) , \qquad (1.23)$$

with a new record precision of 0.43% [35, 36, 37].
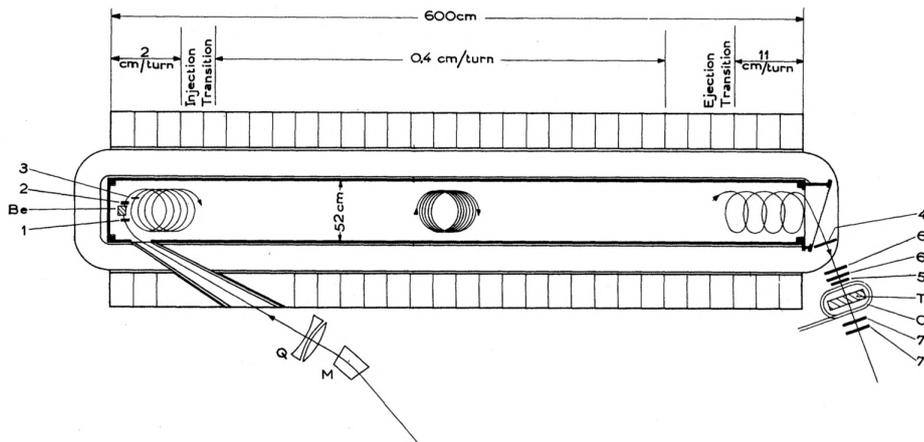


Figure 1.10: Experimental apparatus of the CERN I experiment. Figure from [35].

**CERN II**

For the next experiment, CERN II, a completely new technique involving a storage ring was developed. The main benefit from a storage ring is that high energy muons can be stored, and their lifetime is greatly increased by the relativistic dilation of time. The spin precession can then last longer and be measured with much higher precision.

The storage ring was composed by a series of 40 independent bending magnets producing a magnetic field of 1.7 T. A beam of 10.5 GeV proton accelerated by the new Proton Synchrotron (PS) hits a target placed in the ring producing 1.27 GeV/c muons that in turn could circulate for almost 200 $\mu$s thanks to a Lorentz boost factor of $\gamma = 12.06$. The vertical focusing of the muon beam was provided by a small gradient in the radial direction of the magnetic field. The positrons emitted by the muon decay were then detected by a series of six counters placed on the inner side of the ring (Figure 1.11). The time distribution of the observed positrons can be described by:

$$N(t) = Ne^{-t/\gamma\tau_\mu} \left[1 + A\cos(\omega_a t + \phi)\right] . \tag{1.24}$$

The CERN II experiment provided a new measurement of $a_\mu$ with an uncertainty almost 20 times smaller than the CERN I experiment:

$$a_\mu(\text{CERN-II}) = 0.001\,166\,16(31) , \tag{1.25}$$

with a precision of 266 parts per million [38].

**CERN III**

The third and last of the CERN-based experiments, `CERN-III`, took the storage ring idea and improved on it even further, setting the benchmark for the approach which is still in use today.

The collaboration found a way to increase both the lifetime of the muons inside the ring and the stability of the muon beam. The magnetic field is uniform, and a series of electrostatic quadrupoles provides the weak vertical focusing. In principle, charged quadrupoles would affect the precession of the moving muons, but if the machine operates at the *magic* momentum of 3.1 GeV/c the effect is perfectly canceled out. This phenomenon will be explained in more detail in Section 2.1. Luckily, the desired momentum implies a relativistic factor $\gamma = 29.3$, which is more than double of the CERN II one, extending the storage time from $\sim 200$ $\mu$s to $\sim 500$ $\mu$s. The storage ring had a diameter of 14 meters with a magnetic field of 1.5 T.
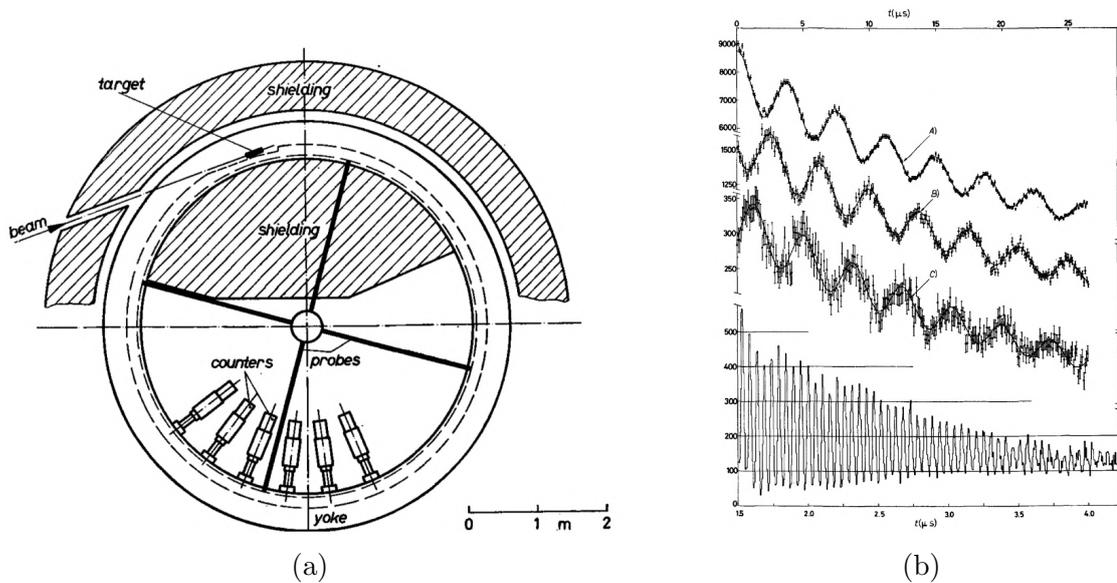
Figure 1.11: Experimental apparatus (a) and the *wiggle* plot (b) of the CERN II experiment. Figures from [38].

The target was moved outside of the storage region, thus greatly reducing the impact from the hadronic flash at beam injection. The decay positrons were detected with a set of 24 electromagnetic calorimeters equally distributed along the ring. Figure 1.12 shows the apparatus scheme and the observed positron data.

Many other technical improvements contributed in achieving the final result of [39, 40]:

$$a_\mu(\text{CERN-III}) = 0.001\,165\,924(8.5)\,, \tag{1.26}$$

with an incredible precision of 7.3 parts per million and an improvement of 36 times with respect to CERN II. The result is the combination of the measurement of both positive and negative muons.

### 1.4.3 The Muon $g-2$ Experiment at BNL (E821)

More than 10 years after the end of the CERN campaign, a group that included many of the CERN collaborators started developing a new Muon $g-2$ experiment at the Brookhaven National Laboratory (BNL) using beams from the Alternate Gradient Synchrotron (AGS). In the following 10 years of development, they built a new superconducting storage ring magnet with a much higher uniformity in the magnetic field. A new passive magnetic inflector and a set of pulsed magnetic kickers would inject the muon beam in the correct orbit while keeping the inflector away from the storage region. The

(a)                                           (b)
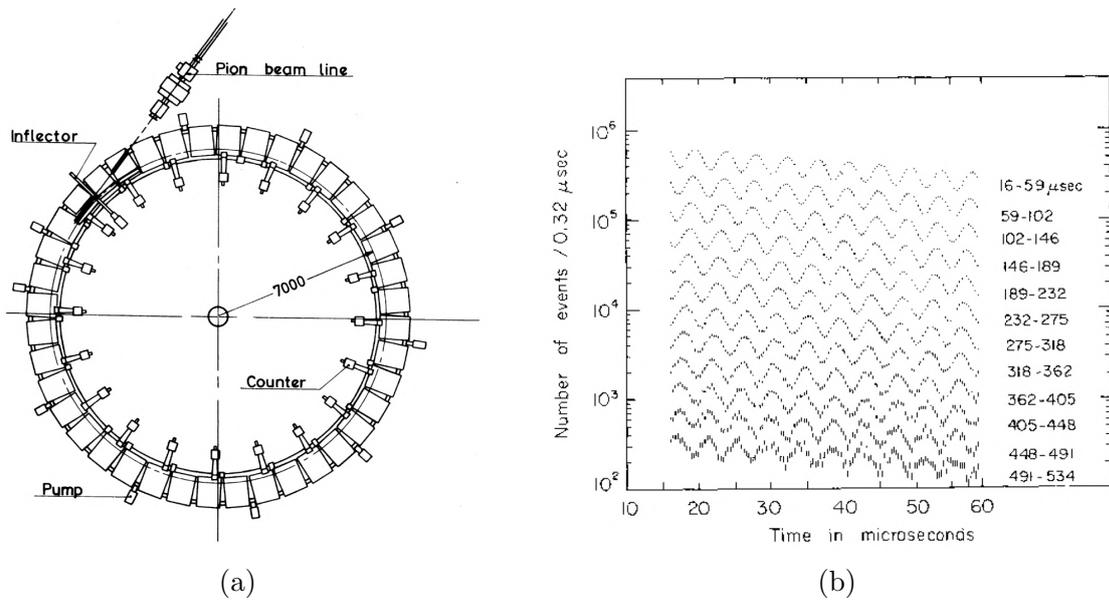
Figure 1.12: Experimental apparatus (a) and the *wiggle* plot (b) of the CERN III experiment. Figures from [39].

same successful idea of utilizing electrostatic quadrupoles and the *magic* muon momentum was applied.

The AGS accelerator complex was capable of delivering beams which were almost 20 times more intense than the PS at CERN, and the proton target station was moved even further away from the injection point, so that most of the pions would decay into muons before entering the ring, producing a cleaner beam.

Together with these, some more improvements of the BNL experiment were necessary to keep systematic errors at a controlled level:

- The storage ring was constructed with three continuously wound superconductors, as opposed to the series of 40 independent conventional bending magnets used in CERN III.

- The inflector incorporates a superconducting shield to minimize the disruption of the field in the storage region, and unlike the CERN inflector, allows it to operate in a static DC mode.

- An NMR system capable of making an in-situ measurement of the field in the storage ring was designed, which, unlike CERN III, does not require cycling the magnet power.

- In the BNL experiments, the decay electron signals from the calorimeters are

recorded by waveform digitizers and stored for later analysis instead of relying on a hardware trigger.

After the data taking concluded in 2001, the BNL muon $g-2$ collaboration delivered the final result for $a_\mu$ by combining positive and negative muons data [41]:

$$a_\mu(\text{BNL}) = 0.001\,165\,920\,80(63)\,, \tag{1.27}$$

with a precision of 0.54 ppm and an improvement of almost 14 times with respect to CERN III. The precision of this measurement was then enough to be sensitive to EW effects and higher order QCD contributions. At that time, one of the best estimations from the Standard Model prediction was[2] [42]:

$$a_\mu(\text{SM, 2006}) = 0.001\,165\,918\,20(73)\,, \tag{1.28}$$

resulting in a discrepancy of:

$$\Delta a_\mu(\text{BNL - SM, 2006}) = (26.1 \pm 9.4) \cdot 10^{-10}\,, \tag{1.29}$$

with a significance of $2.7\sigma$.

### 1.4.4   The Muon $g-2$ Experiment at Fermilab (E989)

While the discrepancy between experiment and theory might be attributed to miscalculations in the theory or systematic errors in the BNL experiment, or a very rare statistical fluctuation, no such errors have been found despite repeated attempts to resolve it. Moreover, after the BNL measurement, improvements from the theoretical prediction increased the precision of the Standard Model value to the current:

$$a_\mu(\text{SM, 2020}) = 0.001\,165\,918\,10(43)\,, \tag{1.30}$$

bringing the discrepancy with the BNL experimental value to $3.7\sigma$ [19].
As the discrepancy could not be solved, even more interest in Muon $g-2$ experiments started to grow. The most intriguing and exciting explanation for such a discrepancy would be the effects from physics Beyond the Standard Model (BSM). As discussed in Section 1.2, the value of $a_\mu$ receives contributions from all particles that couple to the

---

[2]At the time of the BNL publication, there were two main HVP evaluations. One of them, which included data from hadronic tauon decays, was not considered because of inconsistencies with the $e^+e^-$ annihilation data [41].

muon through virtual loops, so it is possible that new undiscovered particles are the source of this discrepancy.

In order to either confirm or disprove the discrepancy, a new experiment at Fermilab was proposed in the late 2000s. The Fermilab accelerator complex is capable of delivering an even higher luminosity than the Brookhaven's AGS. The BNL magnetic ring was already a sufficiently accurate device for an improved measurement, and, together with economical reasons, the decision was made to move the superconducting coils as a single piece all the way from Brookhaven (NY) to Fermilab (IL). This operation, nicknamed *the Big Move* lasted 35 days and successfully brought the 15-meter-wide magnet to its destination after more than 5000 km on land and sea during the summer of 2013.

The goal of the new Muon $g-2$ Experiment at Fermilab is to measure the muon anomaly with a precision of 140 ppb by observing $\sim 20$ times the muons accumulated at BNL. Many improvements from the beamline to the detectors are in place to reduce the systematics at the necessary level. A detailed description of the apparatus will be provided in Chapter 3. The experimental technique, which will be discussed in Chapter 2, is similar to the one of the BNL experiment and involves a pure beam of polarized muons at the *magic* momentum of 3.094 MeV/c.

The Muon $g-2$ collaboration, of which I am part since the summer of 2017, released in April 2021 the first measurement of $a_\mu$, relative to the first year of data taking (Run-1). The detailed description of this measurement will be presented in Chapter 8. The measured value, with a precision of 460 parts per billion (ppb), is [11]:

$$a_\mu(\text{FNAL, 2021}) = 0.001\,165\,920\,40(54)\,, \tag{1.31}$$

which is consistent with the previous BNL experiment.

## 1.5   Future perspectives

The number of muons analyzed in the first E989 result corresponds to roughly $\sim 5\%$ of the total collected statistics. Right now the analysis of Runs 2 and 3 is being finalized and a new publication is expected in 2023. The experiment is currently collecting more positive muons in the sixth year of running, which started in November 2022. A final publication containing the entire statistics is currently projected for 2025. In addition,

the E989 experiment will aim to set a new experimental limit to the measurement of the Muon Electric Dipole Moment (EDM), which was also performed by the BNL collaboration, to the upper value of $|d_\mu| < 10^{-21}$ e·cm [43].

From the theoretical point of view, many efforts are coming from the international $g - 2$ Theory Initiative to improve the actual estimation on the Standard Model prediction. In particular, the lattice QCD value described in Section 1.3.3 has to be independently verified and eventually included in the final average.

### 1.5.1    The J-PARC E34 Experiment

A new Muon $g - 2$ experiment is now being developed at J-PARC in Japan, aiming to provide an independent measurement of $a_\mu$ with a completely new approach. The experiment will utilize a low-emittance 300 MeV/c muon beam, which is produced by re-acceleration of thermal muons regenerated by the laser resonant ionization of muonium atoms ($\mu^+ e^-$, or Mu) emitted from a silica aerogel. Figure 1.13 shows the schematic drawing.

The use of a low emittance beam eliminates the need for strong focusing by an electric field, which introduces a large correction to the $\omega_a$ measurement and a sizable systematic error in the BNL and Fermilab experiments. Moreover, the absence of electric fields allows to select a momentum different from the *magic* one of 3.1 GeV. The muon beam will be stored in a highly uniform (1 ppm local uniformity) 3.0 T magnetic storage ring that is 20 times smaller than the Fermilab storage ring. The positrons from the muon decay will be measured by a tracker detector composed of 40 silicon strip sensors arranged radially.

The first phase of the experiment is expected to begin in 2027 with a target precision of about 0.5 ppm, similar to that of BNL or the Run-1 result from Fermilab [44]. In addition, the J-PARC experiment will measure the electric dipole moment (EDM) of the muon, if any, as well.

### 1.5.2    The MUonE Experiment

A new measurement technique has been proposed to measure just the hadronic component $a_\mu^{HVP,LO}$ of $a_\mu$. It is based on a high precision measurement of the differential cross
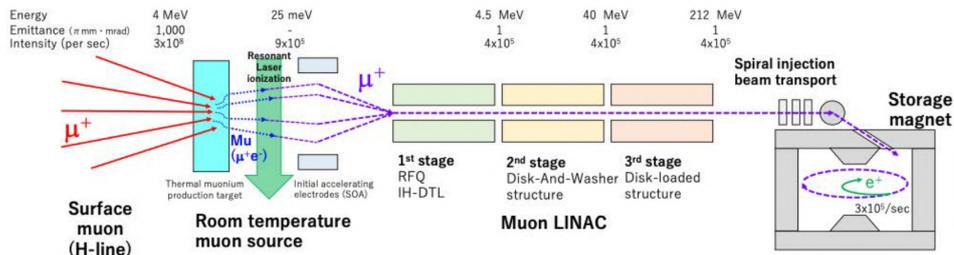
Figure 1.13: Schematics of the muon beamline for the J-PARC Muon $g-2$ experiment. Figure from [44].

section of the muon-electron scattering $\mu e \to \mu e$ as a function of the space-like squared momentum transfer. The new MUonE experiment, currently being developed at CERN, will attempt this by scattering a 160 GeV muon beam, currently available at CERN's muon M2 beamline, on the atomic electrons of a low-Z target consisting of 40 beryllium layers, each 1.5 cm thick. Each layer will be separated by $\sim 1$ m, and three tracking modules will be placed after each beryllium target. A final calorimeter placed at the end of the apparatus will measure the energies of the muons and the scattered electrons.
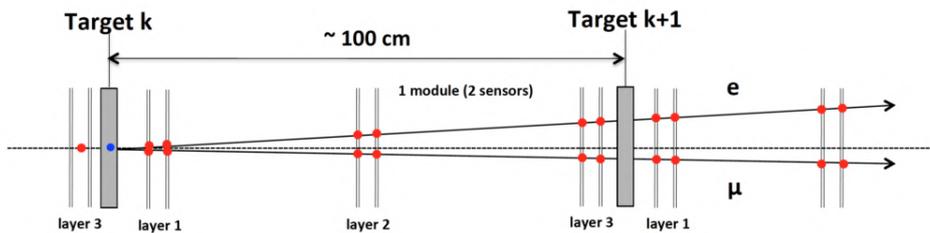


(a)
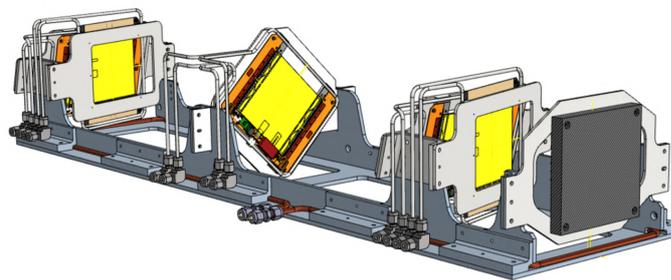


(b)

Figure 1.14: (a) Schematics of the MUonE experiment detector. The target is a thin Beryllium layer while the tracking layers are made of two silicon tracker detectors. (b) CAD drawing of one of the 40 tracking stations. Figures from [45, 46].

The MUonE experiment will directly measure the hadronic contribution to the running of the electromagnetic coupling constant ($\Delta\alpha_{Had}$) in the space-like region [47]. The $a_\mu^{HVP}$ term can be extracted with the following equation [45]:

$$a_\mu^{HVP} = \frac{\alpha}{\pi} \int_0^1 dx (1-x) \Delta\alpha_{Had}[t(x)] \, , \qquad (1.32)$$

where $\alpha$ is the fine structure constant, and the integration variable $x$ is related to the space-like momentum transfer $t$ through the formula:

$$t(x) = \frac{x^2 m_\mu^2}{x-1} < 0 \, . \qquad (1.33)$$

The total target thickness will be $\sim$60 cm, and the CERN M2 beamline can provide a luminosity of $\sim 1.3 \cdot 10^7$ $\mu$/s. The expected integrated luminosity of 3 years of data taking should accumulate $\sim 4 \cdot 10^{12}$ elastic events with electron energy $> 1$ GeV, sufficient to achieve a final statistical error of $\sim 0.3\%$ on $a_\mu^{HVP}$.

With a systematic uncertainty of the same order of magnitude, the MUonE experiment could bring a result comparable with both the Standard Model predictions obtained with data-driven and lattice-QCD techniques. It could therefore clarify the current tension between the two predictions described in Section 1.3.3, allowing for a firmer interpretation of the upcoming measurements at Fermilab and J-PARC.

# Chapter 2

# The experimental technique

## 2.1 Measuring $a_\mu$

A charged particle with mass $m$, placed in a uniform external magnetic field will follow a circular path because of the Lorentz force, and this motion is called cyclotron motion. The spin of the particle, if any, will also rotate (precess) in the same plane of the circular orbit. In the absence of electrical fields, and with the particle velocity perpendicular to the magnetic field, the equations governing this motion are:

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

$$\vec{\omega}_c = -\frac{e\vec{B}}{m\gamma} \, . \tag{2.2}$$

$\omega_s$ is the spin precession frequency and $\omega_c$ is the cyclotron frequency. $\gamma$ is the relativistic Lorentz factor, and $g$ is the particle's g-factor. The second term of Eq. 2.1 is a relativistic correction to the Lorentz force, called Thomas precession [48]. It accounts for the rotation of the particle's frame of reference. The rotation of the particle spin with respect to its momentum is called anomalous precession frequency, $\omega_a$, and is equal to:

$$\vec{\omega}_a = \vec{\omega}_s - \vec{\omega}_c = -\left(\frac{g-2}{2}\right)\frac{e\vec{B}}{m} \equiv -a_\mu \frac{e\vec{B}}{m} \, . \tag{2.3}$$

Interestingly, if the particle's g-factor is equal to 2 as predicted by Dirac equation then $a_\mu = 0$ and there would be no anomalous precession. Figure 2.1 depicts the two scenarios of a particle beam with $g = 2$ and $g > 2$.

More generally, in the presence of electric fields $\vec{E}$ and for non-perpendicular momen-

(a) $g = 2$                                         (b) $g > 2$
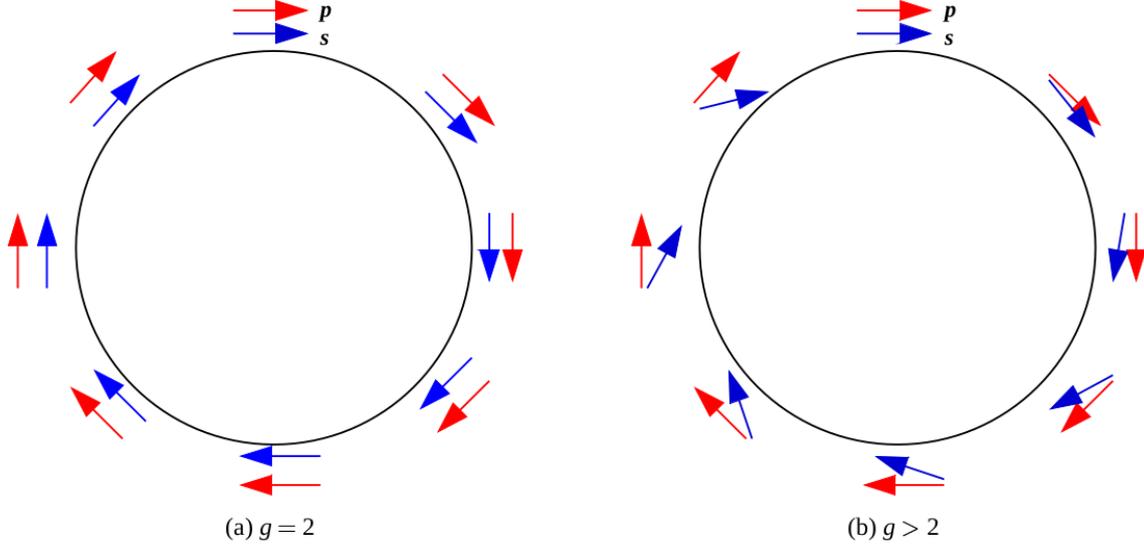
Figure 2.1: Illustration of the muon spin (blue) and momentum (red) vectors for a particle orbiting in a magnetic field if $g = 2$ (a) and $g > 2$ (b).

tum and magnetic field vectors, Equation 2.3 becomes:

$$\vec{\omega}_a = -\frac{e}{m}\left[a_\mu\vec{B} - \left(a_\mu - \frac{1}{\gamma^2 - 1}\right)\vec{\beta} \times \vec{E} - a_\mu\frac{\gamma}{\gamma + 1}(\vec{\beta} \cdot \vec{B})\vec{\beta}\right]\,, \qquad (2.4)$$

where $\vec{\beta}$ is the particle velocity in units of $c$.

Electric fields are typically necessary in $g - 2$ experiments, as is the case for the Fermilab E989 one, in order to provide vertical focusing of the beam. The second term $\vec{\beta} \times \vec{E}$ would introduce a perturbation to the particle precession. Interestingly, the multiplicative factor in front of it can be canceled out for a specific *magic* relativistic gamma:

$$\gamma \approx 29.3 \ \rightarrow \ \left(a_\mu - \frac{1}{\gamma^2 - 1}\right) = 0\,, \qquad (2.5)$$

which corresponds, for the muon, to a *magic* momentum $p_\mu = 3.094$ GeV/c. The experiments at CERN III, BNL, and the current Fermilab E989 exploited this feature by using muons with magic momentum to reduce the influence of the E-field on the measured precession.

Assuming that the motion of the particles is, on average, perpendicular to the magnetic field vector, the last term of Equation 2.4 can be neglected as well, and the anomalous precession frequency can be approximated with the expression of Equation 2.3. It

is therefore possible to calculate the anomalous magnetic moment of the muon, $a_\mu$, from the measurement of $\omega_a$ and $B$:

$$a_\mu = \frac{\omega_a}{B}\frac{m}{e} \, . \tag{2.6}$$

A precise determination of the anomalous precession frequency and of the magnetic field magnitude leads to a direct and precise measurement of the anomaly value.

The value of $\omega_a$ can be measured with the observation of the decay positrons as a function of time. The Fermilab E989 Experiment achieves this with a set of electromagnetic calorimeters. The value of the magnetic field can be measured in terms of the shielded Larmor precession frequency of the proton ($\omega_p'$) when submerged in the same magnetic region where the muons are stored:

$$B = \frac{\hbar\tilde{\omega}_p'}{2\mu_p'} \, , \tag{2.7}$$

where $\mu_p'$ is the proton's magnetic dipole moment and $\tilde{\omega}_p'$ is the spatial average of $\omega_p'$ weighted by the beam distribution. In the Fermilab E989 Experiment the measurement of $\omega_p'$ is done using Nuclear Magnetic Resonance (NMR) probes and the beam distribution is measured with two tracking stations. Finally, to reduce the uncertainty on the $m/e$ ratio, the electric charge can be written as:

$$e = \frac{4m_e\mu_e}{\hbar g_e} \, , \tag{2.8}$$

where $g_e$, $\mu_e$, and $m_e$ are the g-factor, the magnetic dipole moment, and the mass of the electron. The equation for the muon magnetic anomaly becomes:

$$a_\mu = \frac{\omega_a}{\tilde{\omega}_p'}\frac{g_e}{2}\frac{m_\mu}{m_e}\frac{\mu_p'}{\mu_e} \tag{2.9}$$

The first fraction $\omega_a/\tilde{\omega}_p'$ is what is measured at the Muon $g-2$ Experiment at Fermilab. The other quantities are known with high precision from other experiments.

## 2.2   Muon production

An experiment aiming to measure the precession of the muons in a magnetic field requires that the beam of muons is polarized when injected into the storage magnet. After the discovery of parity violation in weak interactions, the decay of charged pions became

an easy way to produce polarized muons.

The pion is the lightest of all hadronic matter and is produced in copious numbers when an energetic proton beam impacts a high Z target. When a charged pion decays, the two constituent quarks annihilate, and the $W$ boson is the only possible propagator because of charge conservation. The final state must also have a net charge, and the light mass of the pion ($\sim 140$ MeV) implies that the only possibilities are to have a muon or an electron paired with the corresponding neutrino.

Typically, the phase space of the possible outcomes would favor the lighter electron, but the parity violation of the weak force greatly enhances the decay into a muon and a neutrino [10]:

$$R_\pi = \frac{\Gamma(\pi \to \mu + \nu_\mu)}{\Gamma(\pi \to e + \nu_e)} = 8.1 \cdot 10^3 \,, \tag{2.10}$$

so that the probability of having a muon in the final state is $> 99.9\%$

To better understand this phenomenon, Figure 2.2 shows the $\pi^+$ decay in the rest frame of the pion. In this frame, the neutrino and the anti-lepton are emitted back-to-back because of momentum conservation. Neutrinos have left-handed chirality in the Standard Model, and, in the limit of massless neutrinos, the neutrino is always emitted with a left-handed helicity state, as the chirality and helicity operators coincide for $\beta \to 1$.

Since the pion has zero spin, to conserve total angular momentum the opposite anti-



Figure 2.2: Diagram depicting the pion decay in the rest frame, with helicity constraints.

lepton must also have left-handed helicity. In the weak decay, both the positron and the anti-muon prefer to have right-handed helicity due to the $(1 - \gamma^5)$ dependence of the decay probability. As this is not possible in the pion decay, the more massive muon is heavily favored.

The decay muon momentum direction is spatially isotropic in the rest frame of the pion, as the latter has zero spin. When a beam of pions is boosted into the lab frame, however, both the highest energy (forward decay) or lowest energy (backward decay) muons exhibit strong spin polarization, with higher energy pions having their spin anti-parallel to their momentum vector.

The same conclusions can be drawn for a beam of negative pions, although the results are reversed. Both the outgoing anti-neutrino and the negative muon are right-handed. Selecting the highest momentum decay muons now results in a beam with its polarization aligned to the momentum.

## 2.3   $\omega_a$ measurement

The first quantity to measure in Equation 2.9 is the anomalous precession frequency $\omega_a$, which is the rate at which the muon spin rotates with respect to its momentum. The measurement of this quantity is possible by observing the positrons emitted in the muon decay as a function of time.

### 2.3.1   Muon decay

Muons decay into an electron and two neutrinos with nearly 100% probability:

$$\mu^- \to e^- \bar{\nu}_e \nu_\mu \tag{2.11}$$

$$\mu^+ \to e^+ \nu_e \bar{\nu}_\mu \tag{2.12}$$

As in the case of pion decay, parity violation plays an important role here. In the rest frame of the muon, the highest energy decay positrons come from decays in which the neutrinos are emitted collaterally, as depicted in Fig. 2.3. In this scenario, half of the initial rest mass of the muon is carried away by the decay positron ($E_{max} \approx 53$ MeV), while the other half is shared by the two neutrinos. Since the neutrino and anti-neutrino are traveling in the same direction, and the weak decay dictates they must have left and right helicity respectively, their spins must have opposite direction. With the neutrino's spins canceling, the conservation of angular momentum forces the decay positron to carry the spin of the parent muon.

The parity-violating nature of the weak interactions prefers to generate right-handed positrons in the muon decay. This is achieved when the positron is emitted in the same direction as the muon spin. For relativistic positrons, the differential decay distribution in the muon mass frame is [49]:

$$\frac{d^2 P}{dy\, d\cos\theta} \propto n(y)(1 + a(y)\cos\theta)\,, \tag{2.13}$$
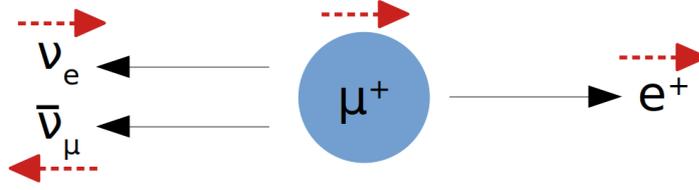
Figure 2.3: Helicity constraints in the muon decay when the neutrinos are emitted in the same direction.

where $y = E/E_{max}$ is the fraction of energy carried by the positron, and $\theta$ is the angle between the muon spin and the positron direction. The quantity $n(y)$, also called the Michel Spectrum, is a proxy for the overall probability of a decay positron with the energy $y$. The decay asymmetry $a(y)$ indicates the degree of correlation between the muon spin and positron momentum. Both functions are normalized to have a maximum value of one:

$$n(y) = y^2(3 - 2y)\,, \tag{2.14}$$

$$a(y) = \frac{2y - 1}{3 - 2y}\,. \tag{2.15}$$

The scenario depicted in Figure 2.3 where the positron has maximum energy ($y = 1$) corresponds to the most probable outcome and the one with maximum asymmetry ($a = 1$). The $n(y)$ and $a(y)$ functions are shown in Figure 2.4a.

The decay distribution in the rest frame of the muons is then boosted into the lab frame. The total number of detected positrons in a fixed direction as a function of time can be described by:

$$N(t, y) = N(y)[1 + A(y)\cos(\phi(t))]\,, \tag{2.16}$$

where $\phi$ is the angle between the muon spin and the muon momentum. $N(y)$ and $A(y)$ are the lab frame probability and asymmetry values, which can be derived to be:

$$A(y) = \frac{-8y^2 + y + 1}{4y^2 - 5y - 5}\,, \tag{2.17}$$

$$N(y) = (y - 1)(4y^2 - 5y - 5)/5\,. \tag{2.18}$$

The two functions are normalized to have a maximum value of one and are shown in Figure 2.4b. The angle $\phi(t)$ changes over time according to the muon precession frequency:

$$\phi(t) = \omega_a t + \phi_0\,, \tag{2.19}$$

where $\phi_0$ is the initial phase. Finally, the count of positrons provided by detectors placed in a fixed position around the muon beam is:

$$N(t, E) = N(E)e^{-t/\gamma\tau_\mu}[1 + A(E)\cos(\omega_a t + \phi_0)] , \qquad (2.20)$$

where the exponential term represents the decay rate of the muons, with a boosted lifetime of $\gamma\tau_\mu \approx 64.4\ \mu$s for the *magic* momentum as in the Fermilab E989 Experiment. The number of positrons above an energy threshold $E_{th}$ oscillates at the precession frequency $\omega_a$, which in turn can be measured by fitting Equation 2.20 to the time distribution histogram produced by the detectors. Figure 2.5 shows an example of such measurement, and the entire process will be described in detail in Chapter 6.



(a)                                                    (b)
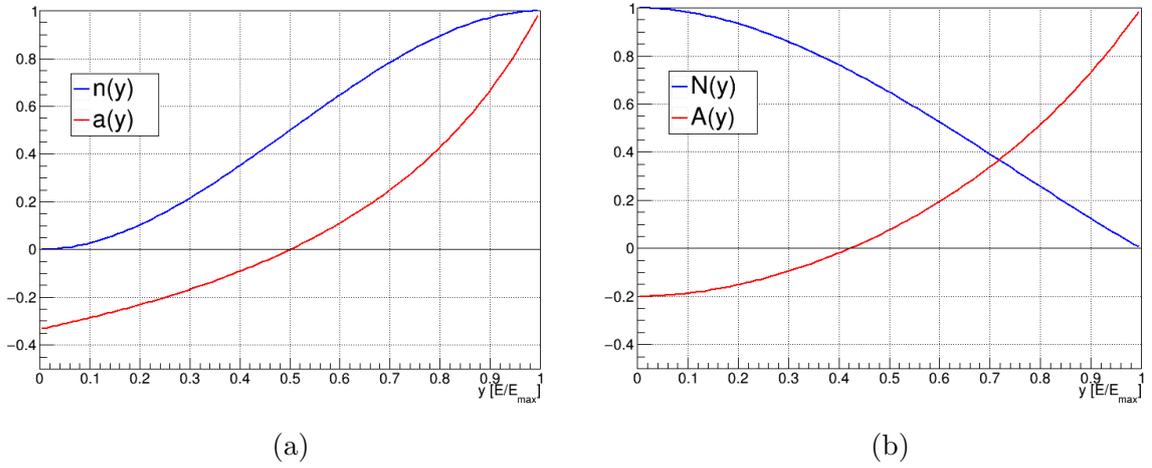
Figure 2.4: Decay rate, $n(y)$, and decay asymmetry, $a(y)$, of the muon decay in the rest frame (a) and the laboratory frame (b). $E_{max} = 53$ MeV in the rest frame and $E_{max} = 3094$ MeV in the laboratory frame.

## 2.4   $\tilde{\omega}'_p$ measurement

The second quantity to be measured in Equation 2.9 is the magnetic field in terms of the Larmor precession frequency of the proton. A Nuclear Magnetic Resonance probe operates by pulsing a radio frequency pulse into a sample of petroleum jelly, rotating the spin of the protons inside the sample by 90° relative to the magnetic field direction. As the RF pulse ends, the sample magnetization relaxes back to the equilibrium orientation, parallel to the magnetic field of the storage ring. During the relaxation period, the proton will precess at the frequency $\omega_p$, inducing a Free Induction Decay (FID) signal onto a pickup coil perpendicular to the magnetic field. The amplitude of the signal decays as
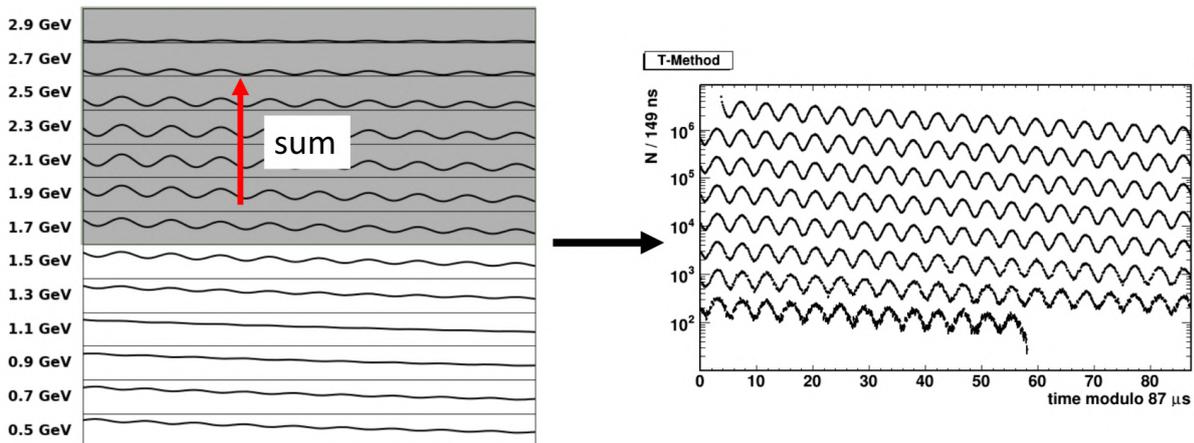
Figure 2.5: The observed number of positron oscillates in time according to the muon precession frequency $\omega_a$. The oscillation amplitude $A(E)$ depends on the positron energy, as shown on the left. Summing the signals from all energies above a certain threshold yields the one-dimensional "wiggle" histogram shown on the right. The signal is wrapped on itself every 87 $\mu$s to show the entire 700 $\mu$s measurement period of the E989 experiment. Figure from [50].

the spin gets back to its original position. An example of such a signal is shown in Figure 2.6.
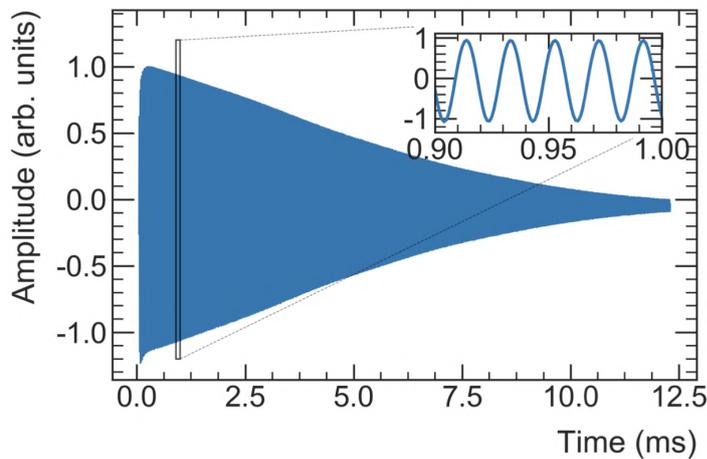


Figure 2.6: Example of the Free Induction Decay signal from one of the NMR probes of the Fermilab E989 Experiment. The zoomed portion shows the oscillation corresponding to the Larmor precession frequency of the proton $\omega_p$. Figure from [51].

The NMR probes need to measure the field in the same region where the muon beam travels. However, placing the probes while the muons are circulating would destroy the beam. The solution adopted by the Fermilab E989 experiment is to measure the field by

equipping a movable trolley with 17 NMR probes and making it scan the entire storage region once every three days, with no muon beam. The field stability is then monitored between two consecutive trolley measurements with a set of NMR probes positioned directly above and below the storage region. The trolley probes are then calibrated in terms of the free proton precession frequency using a high precision probe containing a sample of pure water.

The field measurements provide a three-dimensional map $\omega_p(x, y, \phi)$, where $x$, $y$, and $\phi$ are the transverse and azimuthal coordinates. This map must be averaged with the position distribution of the muons in the storage region $M(x, y, \phi)$ with a convolution:

$$\tilde{\omega}_p = \langle \omega_p(x, y, \phi) \times M(x, y, \phi) \rangle . \tag{2.21}$$

More details on the field measurement techniques are reported in Section 8.4. The muon distribution can be measured with a set of tracking stations, described in the next section.

## 2.5   Beam dynamics

In a fictional experiment where the magnetic field is perfectly uniform, and the beam is monochromatic and perfectly collimated, $\omega_a$ and $\omega_p$ would be the only quantities to be measured. However, a spread in the momentum distribution of the muons, the presence of electric fields, and the finite transversal distribution of the beam generate a concert of oscillations and motions of the beam that must be characterized and considered for the final $a_\mu$ computation.

In the Fermilab E989 Experiment, a magnetic inflector allows the entrance of the beam into the storage ring, a set of magnetic kickers provide a kick to move the muons into the right orbit, and a set of electrostatic quadrupoles provide vertical focusing. More details on these systems will be discussed in Chapter 3.

A quadrupole is characterized by a field index, $n$, that is related to the electric field gradient in the storage region by the following expression [49]:

$$n = -\frac{R_0}{\beta B} \frac{\partial E_y}{\partial y} , \tag{2.22}$$

where $R_0$ is the orbit radius, $\beta$ is the particle speed, $B$ is the magnetic field, and $E_y$ is the

vertical component of the electric field. For a quadrupole field that provides continuous vertical focusing over the full azimuth, the muon beam is affected by harmonic oscillations called Betatron Oscillations (BO) with frequencies:

$$\omega_x^{BO} = \omega_c \nu_x = \omega_c \sqrt{1 - n} \,, \tag{2.23}$$

$$\omega_y^{BO} = \omega_c \nu_y = \omega_c \sqrt{n} \,, \tag{2.24}$$

where $\omega_c$ is the cyclotron frequency and $\nu_x$, $\nu_y$ are the *tuning* factors. As these frequencies are governed by the electric field strength, the voltage applied to the quadrupole plates has to be carefully chosen in order to avoid beam resonances that would lead to large beam oscillations and possible beam loss. In fact, resonances occur when the betatron oscillations are multiples of the cyclotron frequency, or, in general:

$$a\nu_x + b\nu_y = c \,, \tag{2.25}$$

with $a, b, c \in \mathbb{Z}$. The E989 experiment ensures that the quadrupole voltages do not correspond to resonant values, as shown in Figure 2.7. A typical field index of $n = 0.108$ results in the tuning factors $\nu_x = 0.944$ and $\nu_y = 0.329$.

The beam oscillations introduce a time-dependent modulation in the number of detected positrons due to the radial and vertical acceptance of the detectors. The horizontal betatron oscillation frequency is higher than half of the cyclotron one, since $\nu_x > 0.5$. The modulation observed by the calorimeters is therefore at the aliased value of

$$\omega_{CBO} = \omega_c - \omega_x^{BO} \,, \tag{2.26}$$

called Coherent Betatron Oscillation (CBO).

In general, in an ideal storage ring, betatron oscillations are stable in time, however, due to the momentum spread in the stored beam and to the incomplete coverage in the quadrupoles field, the betatron motions decohere with time. The bundle of slightly different oscillation frequencies gets out of sync after some time and, overall, the effect is reduced. A simple model for the decoherence of both horizontal and vertical oscillations is an exponential function of the form:

$$D(t) = Ne^{-t/\tau} \,, \tag{2.27}$$

where $N$ is the betatron modulation as observed by the detectors. A more detailed description of these beam dynamic effects will be discussed in Chapter 6.
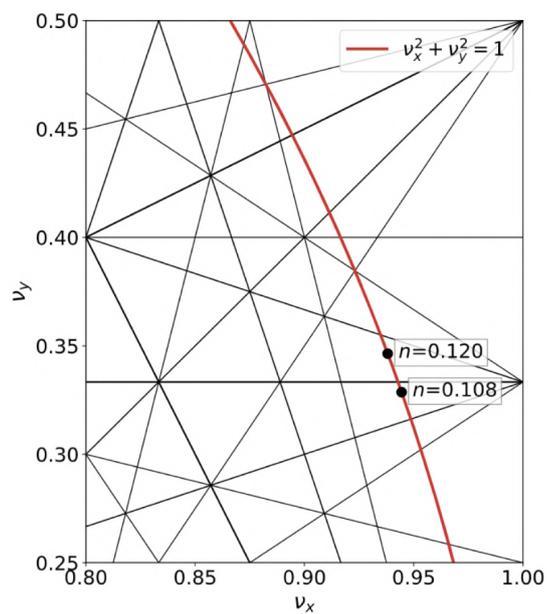
Figure 2.7: Tune diagram showing the resonances in the E989 ring. The black lines represent the distribution of points corresponding to resonant frequencies. The red line corresponds to the possible tuning values of the quadrupoles. The two black points are the Run-1 settings.

# Chapter 3

# The Muon $g - 2$ Experiment at Fermilab
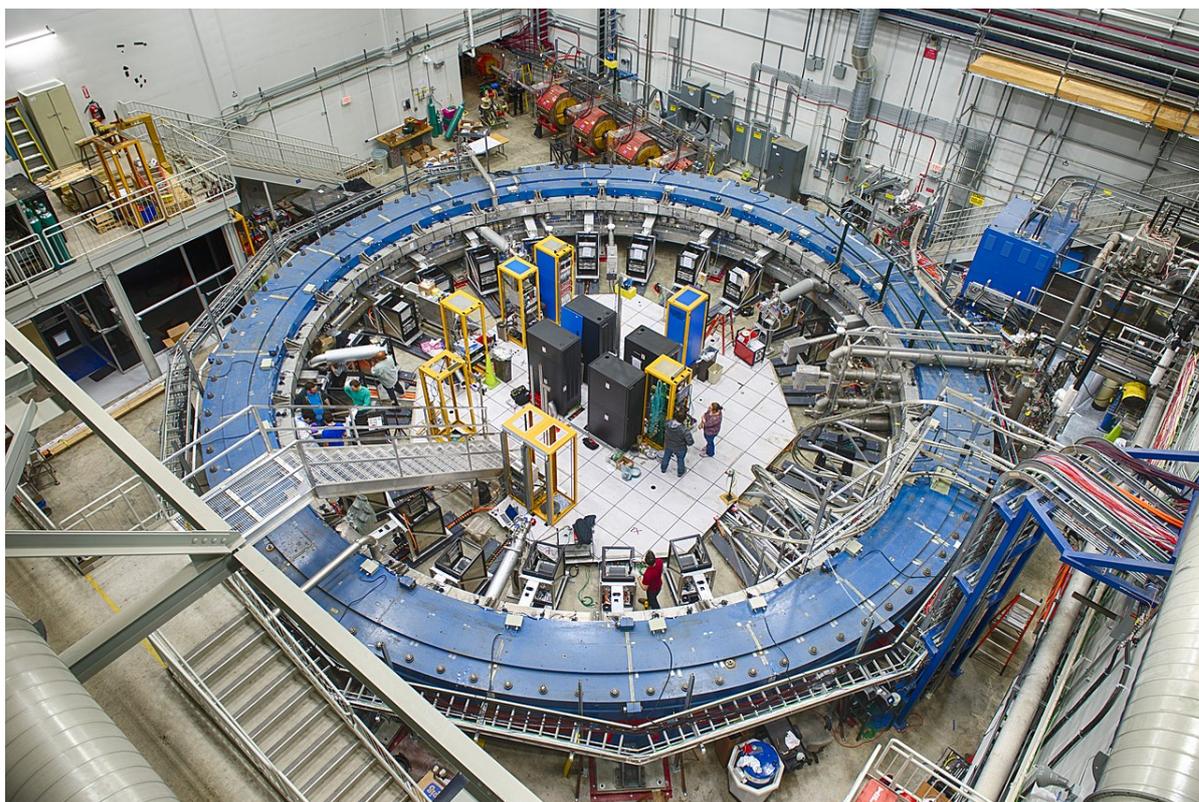


Figure 3.1: The Muon $g - 2$ Experiment at Fermilab.

The new Muon $g-2$ Experiment (E989) in operation at Fermi National Accelerator Laboratory (Fermilab) aims to measure the anomalous magnetic moment of the muon with a precision of 140 parts per billion, a factor 4 better than the previous BNL E821 experiment. The BNL magnetic ring was moved from Brookhaven to Fermilab and installed in the muon campus, at the end of the FNAL accelerator chain. Fermilab accelerators can provide a much more intense beam allowing the experiment to collect 21 times the statistics of the BNL experiment in few years of data collection. While BNL E821 improved on the CERN III experiment in a revolutionary manner, primarily by the invention of direct muon injection into the storage ring, the FNAL E989 experiment introduced a broad suite of refinements focused on optimizing the beam purity and rate, the muon storage efficiency, and modernizing the instrumentation used to measure both $\omega_a$ and $\omega_p$ [49].

The E989 experiment has currently collected $\sim 20$ times the statistics of E821 at the time of this writing and is running its sixth year of data taking. On April 7th 2021, the first $a_\mu$ measurement base on the first year, Run-1, was published in Physical Review Letters [11] together with three accompanying papers. This result, corresponding to $\sim 6\%$ of the total statistics, confirmed the BNL measurement with a similar precision and increased the experimental discrepancy with the data-driven Standard Model prediction to $4.2\sigma$, making a significant impact on the scientific press. The second and third years of data taking, Run-2 and Run-3, are now being analyzed and a new publication is expected in the first half of 2023, with a precision improved by a factor 2 with respect to the Run-1 measurement. The latest datasets, Run-4 to Run-6, are currently being produced, and analysis will start soon. They should increase the final precision, together with the previous datasets, by another factor of two, allowing to reach the final goal of 140 ppb. The total final uncertainty expected for the E989 experiment is subdivided into 100 ppb of statistical and 100 ppb of systematic uncertainties. The systematic uncertainty is further divided into $\sim 70$ ppb for the $\omega_a$ measurement and $\sim 70$ ppb for the $\omega_p$ measurement. The BNL budget on the $\omega_a$ systematic uncertainty was 180 ppb, and many improvements on the hardware side have been developed for the E989 experiment to reduce this value by a factor of three or better.

Some of the various improvements needed to reach these goal uncertainties are [49]:

- Higher proton rate with less protons per bunch. Since the detected positron number is directly proportional to the protons on target, the Fermilab experiment will have to deliver $4 \cdot 10^{20}$ total protons. These high numbers are within reach thanks to the

Fermilab beam complex which is expected to annually deliver $\sim 2 \cdot 10^{20}$ protons with an energy of 8 GeV on an Inconel core target;

- A very long pion decay line. A limiting factor at BNL was the 120 m beamline between the pion production target and the storage ring; since the decay length of a 3.1 GeV/c pion is $\approx 173$ m, the beam injected into the storage ring contained both muons and a significant number of undecayed pions. Those pions create an enormous burst of neutrons when they intercept materials. The new decay line is more than 2000 m long, also thanks to the four orbits around a 500 m long Delivery Ring: virtually no pion will reach the muon storage ring without decaying;

- 6-12 times larger muon yield per proton and 3 times the muon beam rate: the muon storage ring is filled at a repetition rate of 12 Hz, which is the average rate of muon spills that consists of sequences of successive 700 $\mu$s spills with 11 ms spill-separations, compared to 4.4 Hz at BNL;

- Improved detectors and new electronics: the detectors and electronics are newly constructed to meet the demands of measuring the anomalous spin precession frequency to the 70 ppb level. Better gain stability and corrections for overlapping events in the calorimeters are crucial improvements addressed in the current design. A new tracking system allows for better monitoring of the stored muon population, and to establish corrections to $\omega_a$ that arise from the electric field and vertical oscillations;

- Better monitoring of B-field variations: the storage ring magnetic field, and thus $\omega_p$, are measured with an uncertainty that is approximately 2.5 times smaller than E821. This is done by placing critical Nuclear Magnetic Resonance (NMR) probes at strategic locations around the ring and shimming the magnetic field by placing wedges and small steel foils to achieve high uniformity;

- A continuous monitoring and re-calibration of the detectors, whose response varies on several timescales from nanoseconds to days: a high-precision laser calibration system monitors the gain fluctuations of the calorimeter SiPMs at 0.04 % accuracy [52].

This chapter presents the E989 instrumentation to inject and store the muon beams and the detectors to measure the beam, the decay positrons, and the magnetic field strength. A schematic drawing of the E989 experiment is shown in Figure 3.2.
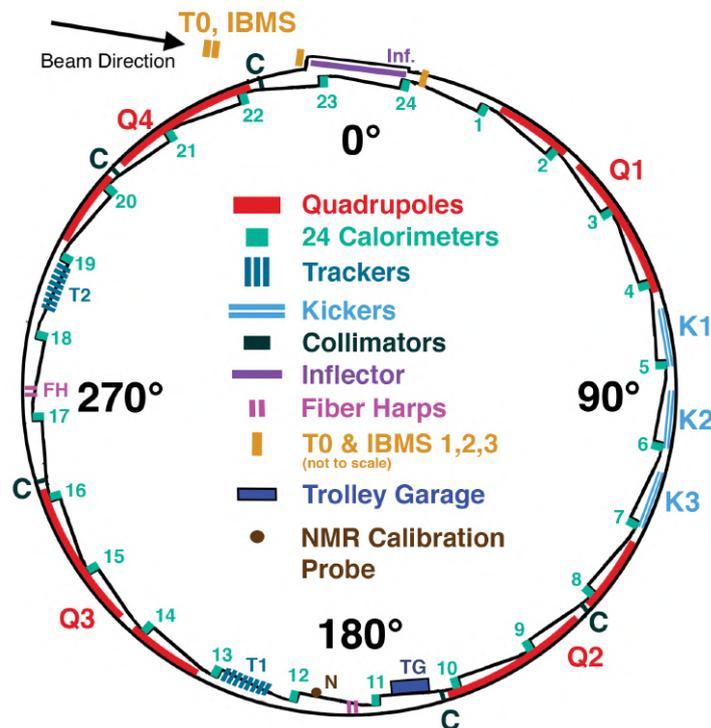
Figure 3.2: Schematic drawing of the storage ring with all the detectors and instrumentation. Figure from [53].

## 3.1    Production and injection of the muon beam

The FNAL muon campus beamline is constructed to deliver pure pulses of highly polarized muons to the E989 storage ring. The muons originate from decaying pions, which are in turn produced by focusing a proton beam on a target. The overall structure of the accelerator complex is schematized in Figure 3.3. The protons begin their journey in the linear accelerator Linac and accelerate through the Booster. From there, the protons continue into the Recycler Ring where they are grouped into high intensity bunches with a short temporal width of $\sim$120 ns. Each proton bunch contains $\mathcal{O}(10^{12})$ protons with 8 GeV kinetic energy. The protons then propagate from the Recycler to the AP0 target hall through (the antiproton production hall used by the Tevatron) where they collide with an *Inconel* target. The collision produces $\mathcal{O}(10^{9})$ positive secondary particles of which many are pions. The secondary particles are focused via an electrostatic lithium lens into a secondary beam which goes through a momentum filter shortly after focusing. Momentum selection yields a beam of 3.1 GeV/c with a momentum spread of $\pm 0.10\frac{\Delta p}{p}$. The secondary beam then proceeds through M2 and M3 beamlines into the Delivery Ring.
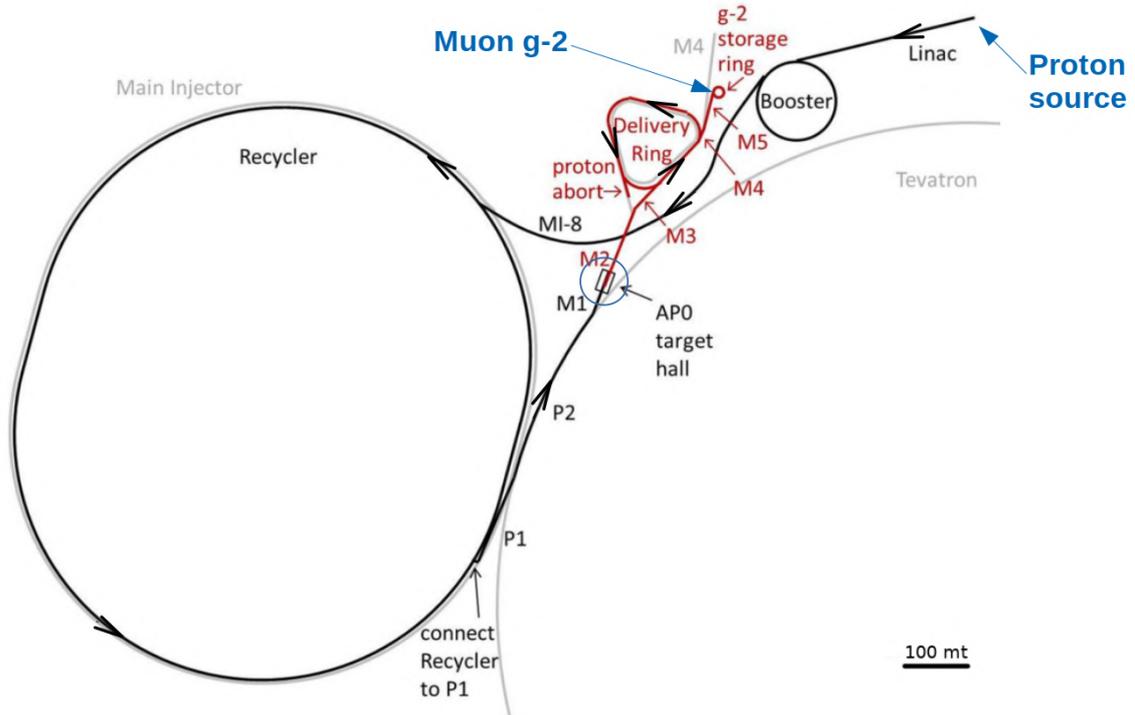
Figure 3.3: Diagram of the accelerator complex used to produce muons for $g-2$. Figure from [49].

The goals in the Delivery Ring are twofold. First, the beam cycles around the Delivery Ring to create a spatial separation between the pions/muons and the more massive protons (slightly lower velocity for the same momentum), so that the protons can be removed. Protons with momentum 3.094 GeV/c have a relativistic $\gamma$ 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 Delivery Ring, the protons fall 25 m behind the muons. Secondly, essentially all pions decay in flight into muons, so that the outgoing beam is a very pure muon beam. The pion decay line is ~2 km long while the one of Brookhaven was only 120 m. Four orbits around the Delivery Ring are enough to achieve both goals.

After the Delivery Ring, the now muon beam is extracted onto the path toward the Storage Ring. Through the pion decay process the high and low energy muons have a net spin polarization (as discussed in Section 2.2), and the beamline design acceptance is narrow around the filtered secondary energy of 3.1 GeV/c. The muons produced at 3.094 GeV/c by the pion beam are forward decays and thereby achieve a net spin polarization of around 95%. 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. A bunch of muons produced in the beamline is referred to as a "fill". These fills deliver $\mathcal{O}(10^4)$ muons to the storage ring at an average rate of 11.4 Hz [49]. 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.

### 3.1.1 Beam structure

The muon beam entering the E989 ring is composed of a cycle of 16 bunches repeating every 1.4 seconds. The cycle structure, shown in Figure 3.4, consists of two group of 8 bunches. The bunches within a group are separated by 10 ms, while the two groups are separated by $\sim$ 200 ms. Each bunch circulates inside the E989 storage ring for $\sim$ 750 $\mu$s, making a total of $\sim$ 5000 turns before being dumped. The bunch length is shorter than 120 ns, as the cyclotron period is 149.2 ns.
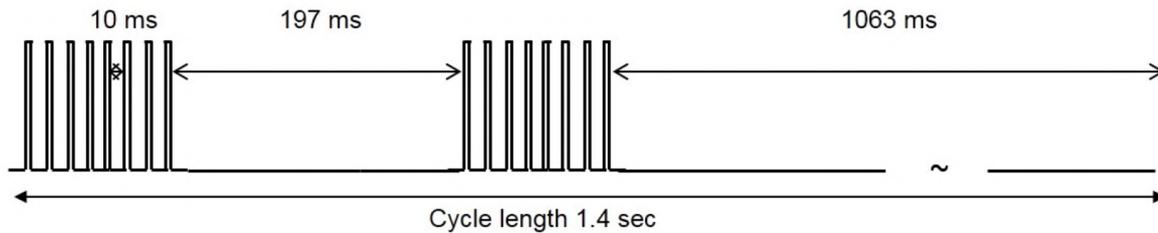


Figure 3.4: Muon beam repetition cycle. Figure from [49].

### 3.1.2 Beam injection

The highly purified muon beam is extracted from the Delivery ring and brought to the E989 ring through the M4/M5 beamline. At the end of the M5 beamline, the beam is focused by four magnetic quadrupoles and injected into the $g-2$ storage ring. The E989 magnet is made of a continuous iron yoke, and the beam enters through a hole in the outer side of the magnet. The beam then passes through a superconducting *inflector* magnet, which provides an almost field free region, until arriving to the edge of the storage region. Without the inflector, the muon injection into the storage ring would not be possible because passage through the fringe field would deflect the beam into the magnet iron.

### 3.1.3    Injection detectors

Before entering the magnet yoke, the muon beam passes through a scintillator detector called *T0*. It is composed of a 1 mm thick piece of plastic scintillator coupled via light guides to two photomultiplier tubes (PMTs). This detector provides the time reference $t = 0$ for the measurement period of the muon bunch. It also measures the beam time profile and the integrated beam intensity. These quantities are used for determining the beam storage efficiency and for data quality monitoring across the data taking period. Figure 3.5 shows a picture and a schematics of the detector.



(a)                                                                       (b)
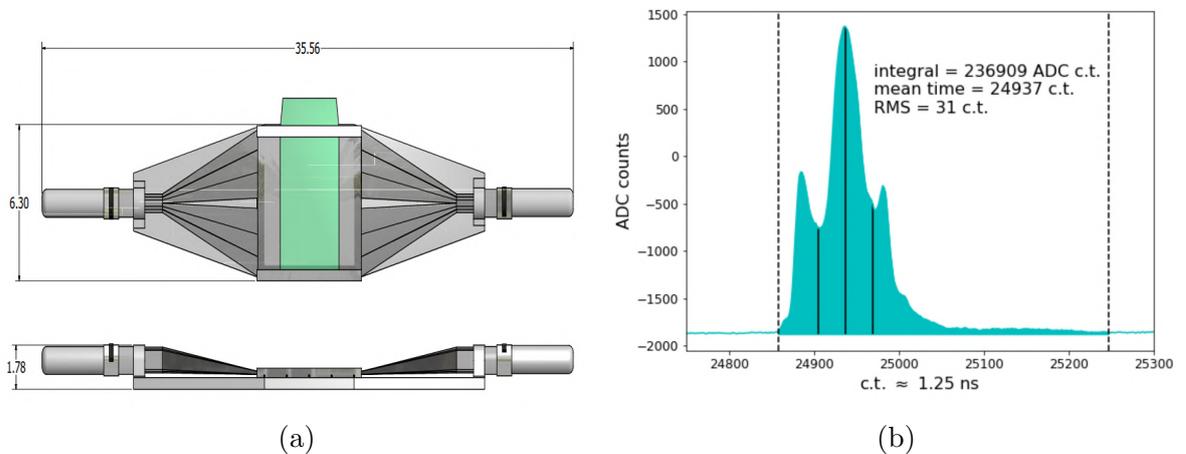
Figure 3.5: (a) CAD drawing of the T0 detector. The muon beam passes through the plastic scintillator (green). (b) Typical beam pulse as measured by T0. The dashed vertical lines represent the integration period over which the beam intensity is measured.

After the T0 detector, three scintillating fiber detectors measure the horizontal and vertical beam profile before and after the inflector magnet. They are called Inflector Beam Monitoring System (IBMS) and are made of a $16 \times 16$ grid of scintillating fibers read by a 1 mm$^2$ Hamamatsu SiPMs. The third IBMS detector (IBMS3) only has the vertical fibers measuring the X-plane profile, and it is in a retracted position during normal data taking as, otherwise, it would destroy the stored beam over time. Figure 3.6 shows a picture of the third IBMS detector and a typical measured beam profile.

## 3.2    The Storage Ring

The E989 magnetic Storage Ring, shown in Figures 3.1 and 3.7 is the same one that was previously used in the E821 Muon $g-2$ Experiment at Brookhaven National Lab.

(a)                         (b)
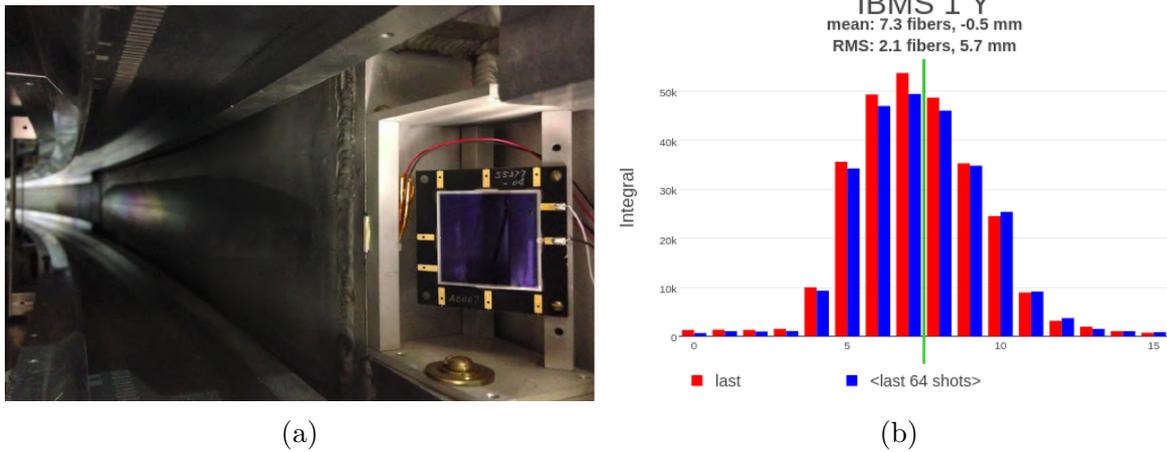
Figure 3.6: (a) Picture of the third IBMS detector placed after the inflector magnet. (b) Example of the vertical beam profile as measured by IBMS1.
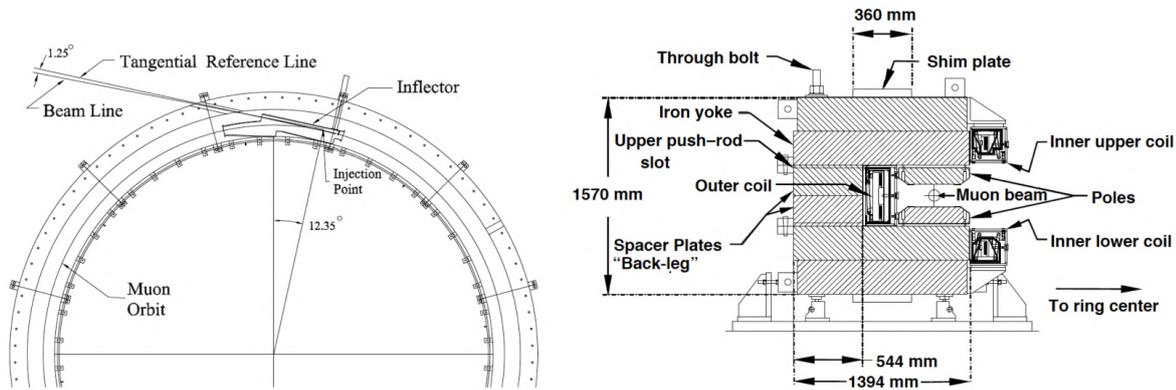


Figure 3.7: Top down and cross sectional view of the Storage Ring. Figures from [49].

The magnetic field is generated by three superconducting Niobium-Titanium coils in a Copper matrix. They operate with a current of 5176 A, providing an uniform vertical 1.45 T dipole field [49]. The C-shape of the magnet cross section faces the interior of the ring so that positrons from the muon decay, which spiral inward, can travel unobstructed by the magnet yoke to detectors placed around the interior of the storage ring.

The muon anomalous precession frequency is proportional to the magnetic field strength, so a very uniform field is essential for the success of the experiment. During the assembly of the ring, edge shims and iron foils were used to provide passive shimming and fine tune the magnetic field over the entire azimuthal angle and to control the transverse gradients. In addition to that, a set of active surface correction coils is used to achieve field uniformity in the storage region to better than one part per million [51]. Moreover, the magnet power supply is adjusted continuously to correct for temperature variations

that cause thermal expansion of the ring. Between Run-1 and Run-2, an insulating blanket was installed around the magnet to provide better thermal stability. After Run-2, a new air conditioning system was installed in the experimental hall for the same reason. The magnetic field is measured by 378 Nuclear Magnetic Resonance (NMR) probes located in grooves carved out from the aluminum vacuum chamber above and below the storage volume.

### 3.2.1   Kickers

The exit of the inflector is displaced 77 mm radially outward from the center of the storage region, and the beam is injected in an orbit slightly shifted from the designed one. A set of three magnetic kickers adjust the orbit into the nominal one by decreasing the magnetic field slightly during the first turn of the beam. The kicker system is composed of three pairs of aluminum plates placed at 1/4 turn after the injection point, as shown in Figure 3.8a. The aluminum plates are $\sim 1.27$ m in length and are pulsed by a $\sim 4000$ A current for deflecting the muons by $\sim 10$ mrad outwards. The aluminum material has been chosen to avoid any perturbation to the main magnetic field during the measurement window. The pulse shape, shown in Figure 3.8b, has a FWHM of $\sim 150$ ns, so that the muons are deflected only during the first turn [54]. However, small ringing fluctuations in the kicker current are visible in the few following cyclotron periods, and additional eddy currents are induced for many milliseconds after the kick. These oscillations are measured with a dedicated magnetometer system, and their effect on the muon precession is evaluated. This will be discussed in more detail in Chapter 8.

### 3.2.2   Electrostatic quadrupoles

While the magnetic field provides natural horizontal focusing, the vertical focusing is provided by a set of electrostatic quadrupoles (ESQs). The system is composed of four sets of quadrupoles, each made of a long (spanning 26°) and a short (spanning 13°) section. A view from the inside of the vacuum chamber is shown in Figure 3.9a, where the four quadrupole plates are seen around the storage region. The quadrupoles provide a linear restoring force in the vertical direction at the expense of a slight horizontal defocusing.

The quadrupole plates are charged before each beam injection with a charging scheme shown in Figure 3.9b. Some quadrupoles plates are connected to a two-step power supply that provides a $\sim 13$ kV field before reaching the final nominal $\sim 18$ kV field. This

|  |  |
|:---:|:---:|
| (a) | (b) |
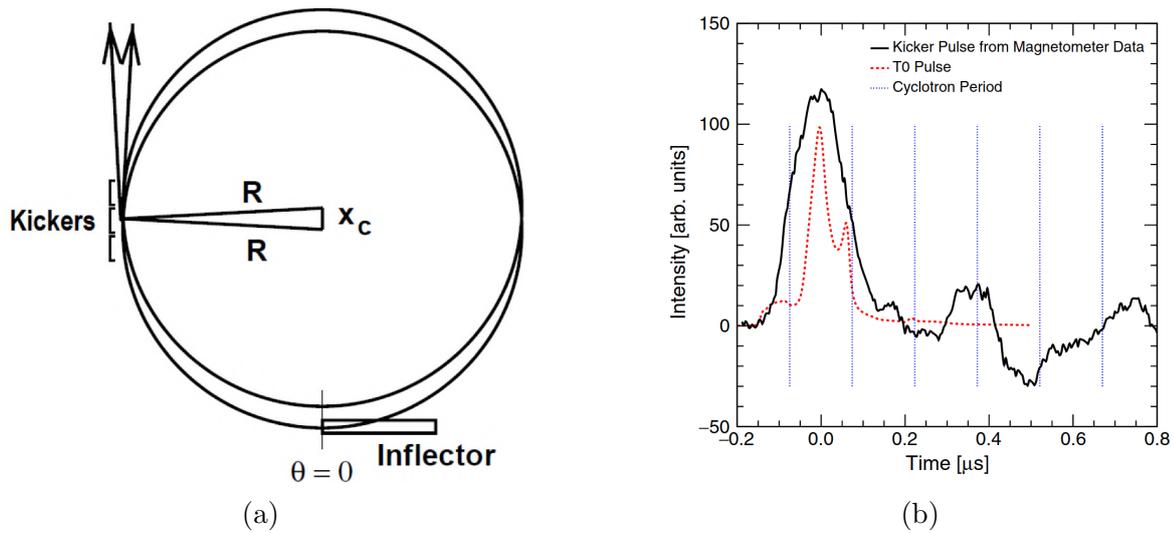
Figure 3.8: (a) Schematic illustration of the muon orbits before and after the kick. During the first turn after injection, the kickers provide the deflection necessary to place muons onto the central storage orbit. (b) Kicker pulse shape as measured with a laser magnetometer. The T0 injection pulse is also overlaid for time comparison. Figure from [53].

middle step is introduced to perform a *scraping* of the muon beam to remove the off-momentum particles. The scraping process moves the beam horizontally and vertically from the nominal orbit so that the particles at the edge of the storage region eventually hit one of the collimators (whose position is marked in Figure 3.2) and exit the storage region.

Both the nominal and scraping charging patterns are designed so that the final voltage is achieved before the start of the measurement period, 30 $\mu$s after injection. In Run-1, two ESQ resistors got damaged and affected the charging characteristic times as shown in Figure 3.9b, resulting in off-nominal voltages up to 300 $\mu$s after injection. The effect of these broken resistors on the measurement of $\omega_a$ will be discussed in Section 6.5.3.

### 3.2.3    Beam monitoring detectors

As the beam circulates inside the storage ring, it oscillates and breathes in the cross section around the nominal orbit radius. The movements of the beam are mainly caused by the weak focusing fields generated by the magnet and the quadrupole system. These oscillations are observed as modulations in the count of the decay positrons, and a precise modeling of the beam movements is required to keep the systematic uncertainties under

<div align="center">(a)                                               (b)</div>
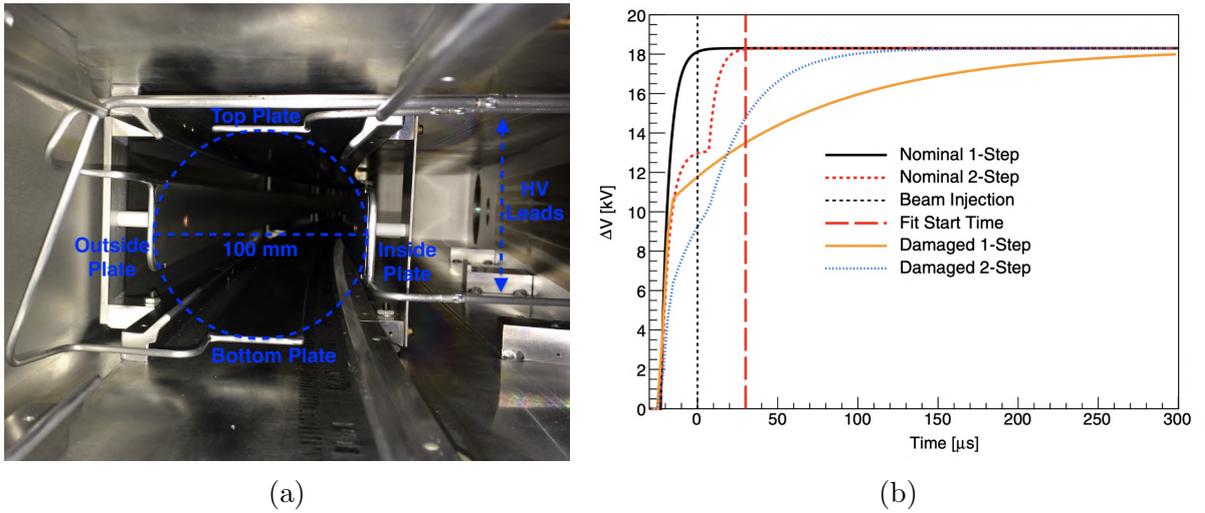
Figure 3.9: (a) Picture of a quadrupole section inside the storage ring. (b) Quadrupole voltages as a function of time. The black line represents a one-step charging mode with 5 $\mu$s time constant, while the red dotted line represents the two-step process. The yellow and blue lines represent the quadrupole containing the two damaged resistors. The voltage varies well beyond the $\omega_a$ fit start time of 30 $\mu$s. Figure from [53].

control.

**Trackers**

Two tracking stations are located at approximately 180° and 270° after the injection point, as shown in Figure 3.2. Each station is composed of 32 planes of straw tubes detectors assembled into 8 modules. A picture of a module is shown in Figure 3.10b. When a muon decays into a positron, the latter curls inward and can pass through the various planes of a tracker station before hitting a calorimeter placed afterward. Figure 3.10a shows an example of a high energy decay positron traversing a full tracking station. The straw tubes are filled with Argon-Ethane gas, and a thin tungsten wire positioned in the axis of each straw collects the drift electrons produced by the ionization of the passing positron. The reconstruction of the hits in multiple planes into a track allows to measure the positron momentum and to extrapolate the muon decay vertex. From the reconstructed decay vertex it is possible to monitor the beam distribution in the storage ring in the proximity of the two tracking stations. In addition, by matching the track with a calorimeter hit, it is possible to perform particle identification using the relation between momentum and energy. This is used to identify muons that exit the storage region, e.g. after hitting a collimator, before decaying into positrons.

(a)



(b)
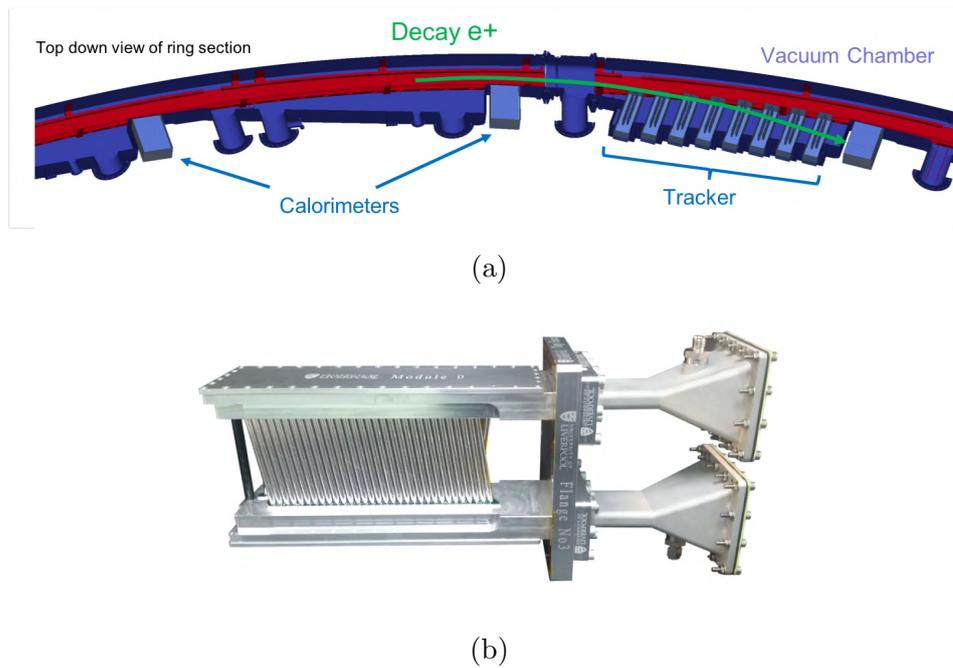
Figure 3.10: (a) Diagram of the tracker measurement of a decay positron. (b) Picture of a tracker module composed of four planes of straw tubes arranged in U-V configuration.

**Fiber harps**

A set of two fiber harps detectors are positioned slightly before each tracking station. They are composed of seven 0.5 mm diameter scintillating fibers read out by SiPMs. For each pair of harps, one has horizontal fibers and the other one has vertical fibers, so that together they can directly and destructively measure the beam $x-y$ profile. They are stored in a retracted position during normal data taking, but can be inserted to measure the beam momentum distribution, the cyclotron frequency, and the de-bunching of the muon beam across time. They are also used to increase the number of lost muons for calorimeter calibration purposes. Figure 3.11 shows a picture of an inserted y-plane fiber harp detector.

## 3.3  Calorimeters

The E989 experiment is equipped with 24 electromagnetic calorimeters for detecting the positrons emitted by the decay of the muons. When a muon decays, the produced decay positron has less energy than its parent muon and correspondingly a smaller orbit radius in the storage ring's magnetic field, and for this reason it curls toward the center of the ring. The 24 calorimeters are stationed around the inner radius of the storage ring
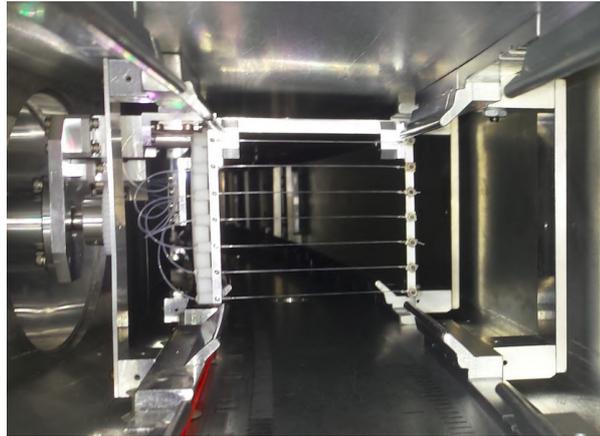
Figure 3.11: Picture of a fiber harp module inserted into the storage region.
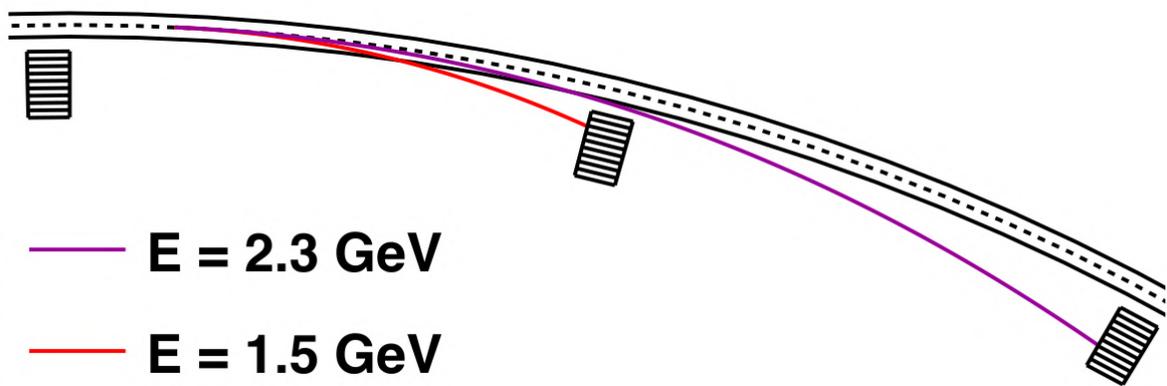


Figure 3.12: Example of two positrons of different energy hitting the calorimeters. The black curves delineate the storage region, and the dashed line is the nominal orbit. Figure from [50].

to intercept these decay positrons and measure their energies and arrival times. This process is illustrated in Figure 3.12. With 24 calorimeters equally spaced azimuthally, the probability that a decay positron will hit a calorimeter (called the acceptance) is very high, about 80%, for the highest energy positrons [49].

The calorimeters are the primary instruments for the $\omega_a$ measurement. The main purpose of the calorimeters is to establish a time and energy for each detected positron. Each one consists of a 6 high by 9 wide array of lead fluoride ($PbF_2$) crystals, each one is a 14 cm long block with a square cross section of side 2.5 cm (Fig. 3.13). Every crystal is coupled with a monolithic 16-channel Hamamatsu MPPC (S12642-0404PA-50) Silicon PhotoMultiplier (SiPM) detector (Fig. 3.14), with a sensitive area of 144 mm$^2$. The properties of these calorimeters are particularly suited to the needs of the Muon $g-2$

Experiment: PbF$_2$ has very high density (7.77 g/cm$^3$), a 9.3 mm radiation length and a Molière radius of $R_M = 22$ mm for energy deposition. High density allows for decay positrons to deposit virtually all of their energy in a relatively compact calorimeter. The crystals are 14 cm long, which is approximately 15 radiation lengths. The width of 2.5 cm is higher than the Molière radius, thus the energy depositions of a typical decay positron is contained almost entirely within a matrix of $3 \times 3$ crystals. The segmentation of the detector decreases the likelihood of two decay events occurring simultaneously in the same SiPMs. These double events are referred to as *pileup* and were a major source of uncertainty in E821 [49].
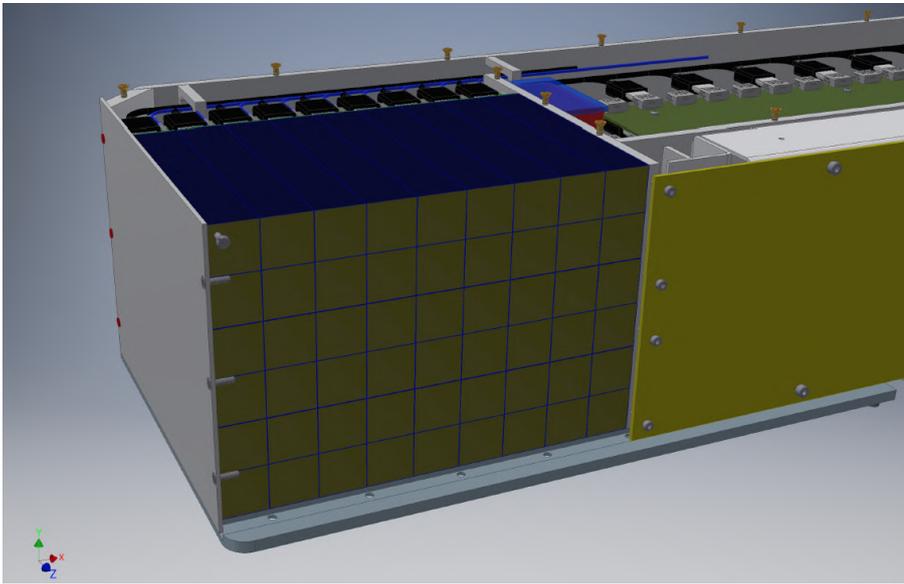


Figure 3.13: Model of the E989 calorimeter. The active volume comprises a 6 by 9 array of PbF$_2$ crystals. There are 24 calorimeters equally spaced around the E989 storage ring. Figure from [49].

Incoming positrons deposit energy in each crystal by producing an electromagnetic shower, and the charged particles participating to the shower emit Čerenkov radiation. The refraction index of PbF$_2$ is 1.8: Čerenkov radiation is produced only by positrons that travel faster than $c/1.8$, that is with a kinetic energy of $\sim$100 keV. All the positrons emitted by the decaying muons have an energy significantly higher than 100 keV. The choice of a pure Čerenkov material is driven by the almost instantaneous signal produced when a positron strikes a crystal, and the light signal is produced within few nanoseconds. In addition, the PbF$_2$ crystals have a very low magnetic susceptibility, which is optimal for working in a high magnetic environment without perturbing the magnetic field itself. Each crystal is wrapped in a black Tedlar® material which, while

absorbing a portion of the emitted photons, ensures a faster response of the crystal [55].

The SiPM is composed of 57344 pixels arranged in 16 sub-sensors on a $1.2 \times 1.2$ cm$^2$ device with 50 $\mu$m pitch. Each pixel is an avalanche photodiode coupled with a quenching resistor serving to arrest the avalanche and allow the device to recover with a time constant of $\sim$10 ns. The produced electrical current is summed for all the pixels and converted into a voltage signal by a custom electronic readout board. These signals are then digitized into 12-bit waveforms with a sampling rate of 800 MHz [56].
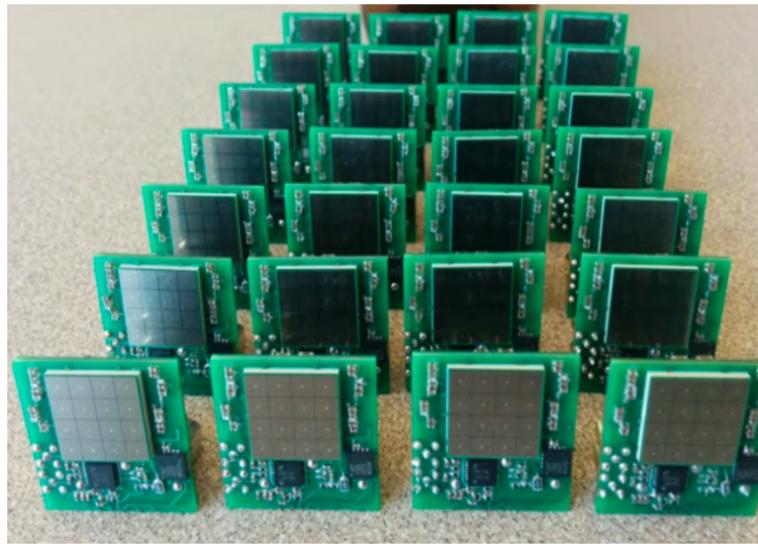


Figure 3.14: The silicon photomultiplier (SiPM) detectors and their readout boards. Figure from [55].

Choosing SiPMs over PMTs is advantageous, as they can be placed inside the storage ring field without perturbation, avoiding the long light guides that would be needed for remote PMTs as in E821. On the other hand, SiPMs are very sensitive to temperature and bias voltage, thus making their gain calibration a challenging and important task. The gain of the 1296 SiPMs is constantly monitored during the data taking period using a Laser Calibration System.

## 3.4   Laser calibration system

The E989 experiment aims to measure the muon anomalous precession frequency $\omega_a$ with a precision of 100 ppb, consisting of 70 ppb statistical and 70 ppb systematic uncertainties. To reach this level of accuracy, the systematic uncertainty related to the gain
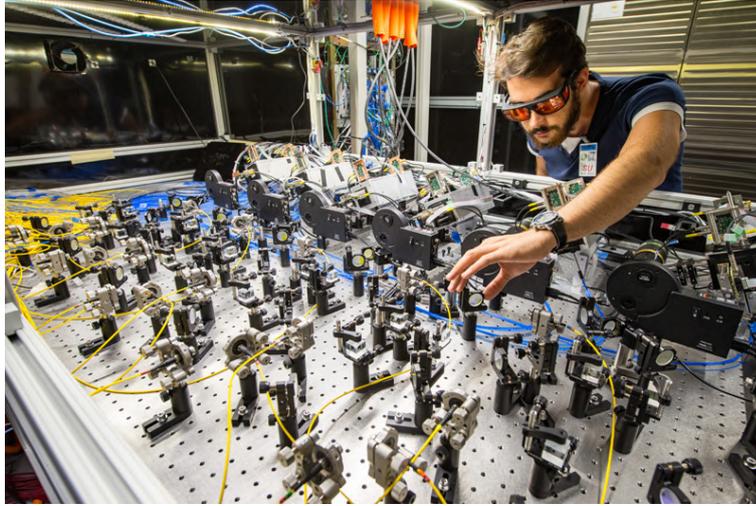
Figure 3.15: The optical table of the Laser Calibration System. On the right: me.

fluctuation has to be lowered from the 120 ppb final value of E821 to a final goal of 20 ppb [49]. The Laser Calibration System (LCS) (Figure 3.15) was designed and built by the INFN group in collaboration with the Istituto Nazionale di Ottica (INO) to precisely monitor the gain fluctuations of the SiPM detectors (SiPM) at the 0.04% level on the short timescale of a single beam fill (700 $\mu$s) and at the percent level on the timescale of an entire run (days).

### 3.4.1    Light distribution system

To achieve this goal, the LCS sends almost simultaneous light pulses to all the 1296 calorimeter crystals. The pulses are also sent to the T0 detector for providing time synchronization. The laser system is assembled in a light-tight and thermally controlled *Laser Hut* located right next to the E989 storage ring. The reason why this room is located outside the ring is to avoid electromagnetic perturbations of the local field induced by the current flow used to excite the lasers.

On the optical table, a set of 6 laser heads (Picoquant LDH-P-C-405M) fires short (0.6 ns FWHM) pulses of light into a distribution system which splits the 6 beams into 24 and sends the laser light into 25 meter long silica *launching* fibers that bring the light toward the calorimeters. Before reaching the distribution system, the laser light passes through a beam splitter which redirects 30% of it to a Source Monitor (SM) which measures the laser source stability. The non-diverted 70% passes through a motorized filter wheel

containing a set of neutral density (ND) filters before being split into four beams. The schematics of the optical table is shown in Figure 3.16. The laser wavelength is $405 \pm 10$ nm and it was chosen according to the spectrum of light of the Čherenkov signal produced by the positrons, the spectral transmission of the crystals, and the quantum efficiency of the SiPMs.
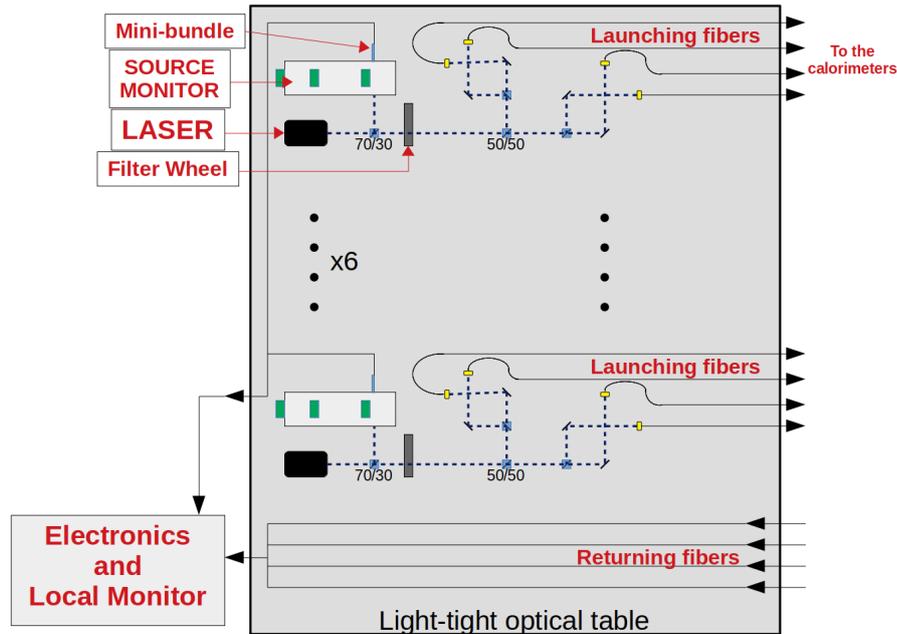


Figure 3.16: Schematic of the optical table inside the Laser Hut. Most of the major components are shown and labeled.

### 3.4.2    Connection to calorimeters

The far end of each launching fiber enters in a light distribution box, where it illuminates a bundle of 54 short fibers via an optical diffuser. These fibers bring the light to each crystal of the calorimeter via a light distribution plate made of 1 cm thick Delrin material placed in front of the calorimeter crystals (Figure 3.17). The panel holds 54 optical prisms that steer the light at 90° into each calorimeter crystals. In addition to the fiber bundle, the diffused light illuminates two additional 25 m long fibers, one made of quartz and the other made of PMMA. These two fibers are sent back to the laser hut and their light is measured by a Local Monitor (LM).
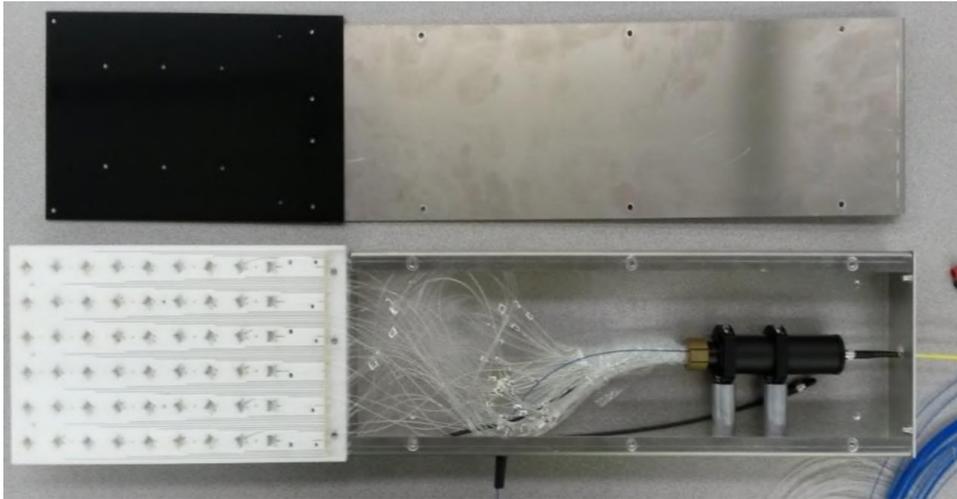
Figure 3.17: Light distribution box of one calorimeter. The launching fiber (yellow) on the right brings the light from the Laser Hut. The white panel on the left distributes the light to the 54 crystals. Figure from [52].

### 3.4.3   Source Monitor

The SM consists of an aluminum box containing a commercial integrating sphere (*Thorlabs IS200*) equipped with two large area (1 cm$^2$) photodiodes (*Hamamatsu S3590-18*) and one PhotoMultiplier Tube (PMT) (*Hamamatsu H5783*). In addition, a low activity Am source coupled to a NaI crystal illuminates the PMT with an average rate of 6 Hz, providing an absolute calibration reference. The pin diodes are used for calibrating the laser intensity over the long timescales of the data taking period, and are sampled at 100 MHz rate. A typical pulse is shown in Figure 3.19a. A schematic of the source monitor is shown in figure 3.18.

### 3.4.4   Local Monitor

As mentioned earlier, two fibers return to the Laser Hut from each calorimeter and their light is measured by the Local Monitors. Each of the 24 LMs is a *Photonics XP2982* PMT sampled at 800 MHz using the same digitizer electronic which is also used for the calorimeters. The PMTs are contained in light-tight plastic cases that hold 10 tubes each, installed in an electronic rack next to the optical table. Before reaching the PMTs, the light passes through a band-pass filter centered at 405 nm, with 10 nm half width, which filters out possible background light. The LM receives two signals separated by about 250 ns. The first is brought by a fiber attached to the corresponding SM integrating sphere (∼3 meters), whereas the second comes from the returning fiber from the calorimeters
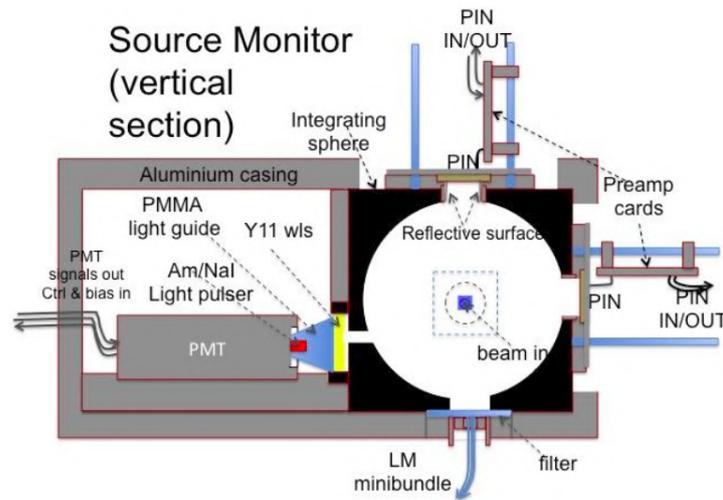
Figure 3.18: Schematic of the vertical section of the Source Monitor. Figure from [52].

($\sim$50 meters). Figure 3.19b shows a typical signal measured by the LM. The two pulses can be directly compared since the expected gain fluctuations of a PMT at this timescale are negligible.



(a)                                                                              (b)
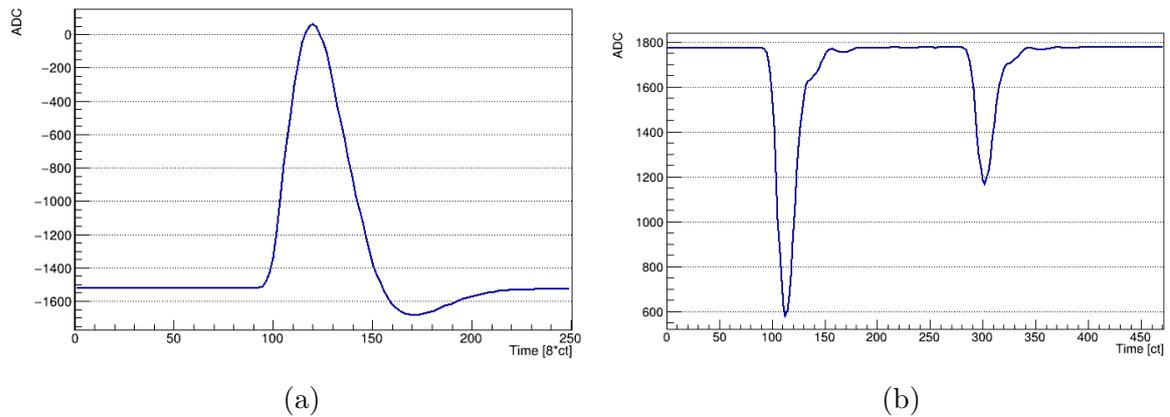
Figure 3.19: (a) Laser pulse as measured by the Source Monitor pin diodes. The sampling rate is 100 MHz. (b) Laser pulses as measured by the Local Monitor. The first pulse comes from the SM fiber bundle, while the second has traveled to and from the calorimeters.

## 3.4.5 Operating modes

The laser is operated in two distinct modes. The first is enabled during physics runs when the muon beam circulates in the storage ring, and used to synchronize the SiPMs to the T0 detector and to measure the gain variations of the SiPMs within the muon fill.

The second is devoted to dedicated calibration runs, with no beam, in order to study the SiPM response to two pulses very close in time. The latter is referred as the *double-pulse* mode.

## Standard mode

When the laser is operated in the standard mode during physics runs, three sets of pulses are generated, which are represented in Figure 3.20:

- SYNC: a Begin-of-Fill (BoF, also known as SYNC) and an End-of-Fill (EoF) pulse are sent to all 1296 calorimeter channels a few tens of microseconds before muon injection and after beam dump, respectively. These signals are used to synchronize the response of the crystals in such a way that the final accuracy of the time reconstruction is at the $\sim 30$ ps level. This will be discussed in more detail in Section 5.3.

- IN FILL: During a prescaled subset of muon fills, the laser system fires a fixed number of pulses. The pulses are shifted in time, for each subsequent fill, in order to sample all times from the injection time up to several hundred microseconds later. These signals are used to measure the gain fluctuations within a fill. This will be discussed in more detail in Section 5.4.2.

- OUT OF FILL: muon injections are interleaved by a time gap of $\sim 10$ ms which allows the laser system to send a set of 4 pulses when no muons are present. These pulses are used as a long term calibration of the gain. This will be discussed in more detail in Section 5.4.3.

## Double-pulse mode

In the double-pulse mode, two consecutive laser pulses are sent to all crystals with a delay that can vary from 1 ns up to several hundreds of $\mu$s. The goal of this mode is to test the calorimeter response to two or more consecutive particles. In particular, the SiPM gain is reduced when two particles enter a crystal within a short time interval. There are two distinct time dependencies: one at very short time separations ($< 100$ ns), due to the recovery time constant of the quenching resistors of the SiPMs, and one at longer times ($\sim 10$ $\mu$s), due to the recovery time of the power supply. The double-pulse mode provides the possibility of checking periodically (each three days of data taking,
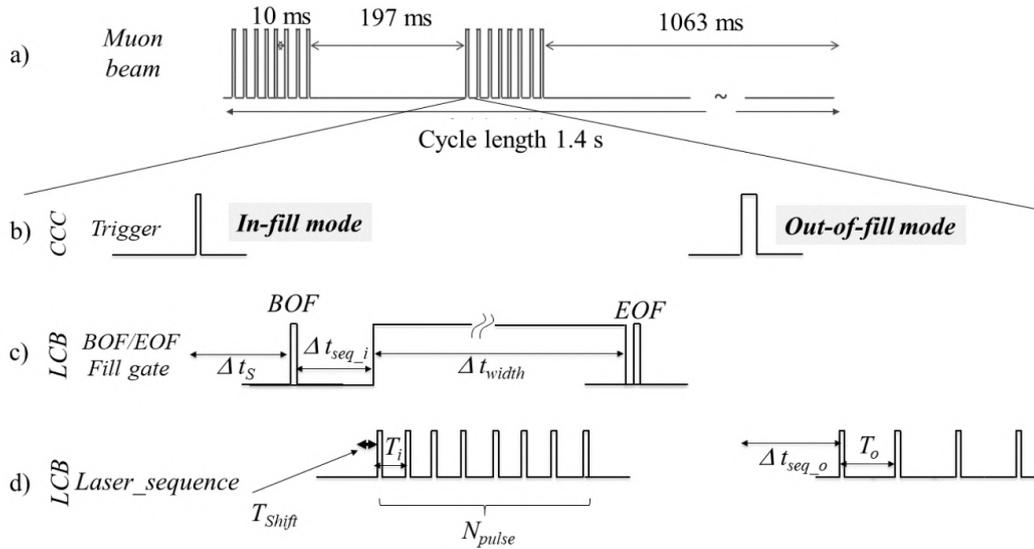
Figure 3.20: Time sequence of the different trigger pulses generated for each muon injection. (a) The sequence of the muon bunches as provided by the Fermilab accelerator complex. (b) For each muon fill, the central clock generates two signals before and after the presence of the muon beam in the storage ring. (c) The Laser Control Board generates two pulses right before beam injection (sync pulse) and right after beam dump (End-Of-Fill). (d) Series of In-Fill and Out-Of-Fill laser pulses. Figure from [52].

typically) the gain function for each of the 1296 crystals during data taking which allows for correction of these effects and for keeping the systematic error under control.

There are several different reasons why it is better to send pulses from two different lasers, rather than fire the same laser repeatedly:

- two lasers allow for choosing different light intensities for each pulse of the pair;

- the laser maximum repetition rate of 40 MHz does not allow the testing of nanosecond time scale;

- in case of two consecutive pulses, the laser output light for the second one can be systematically different from the first one, while light fluctuations for different lasers are uncorrelated;

- as the laser light output can fluctuate up to the percent level from pulse to pulse and this fluctuation is monitored by the SM (response time tens of microseconds), it cannot be corrected when pulses are too close in time.
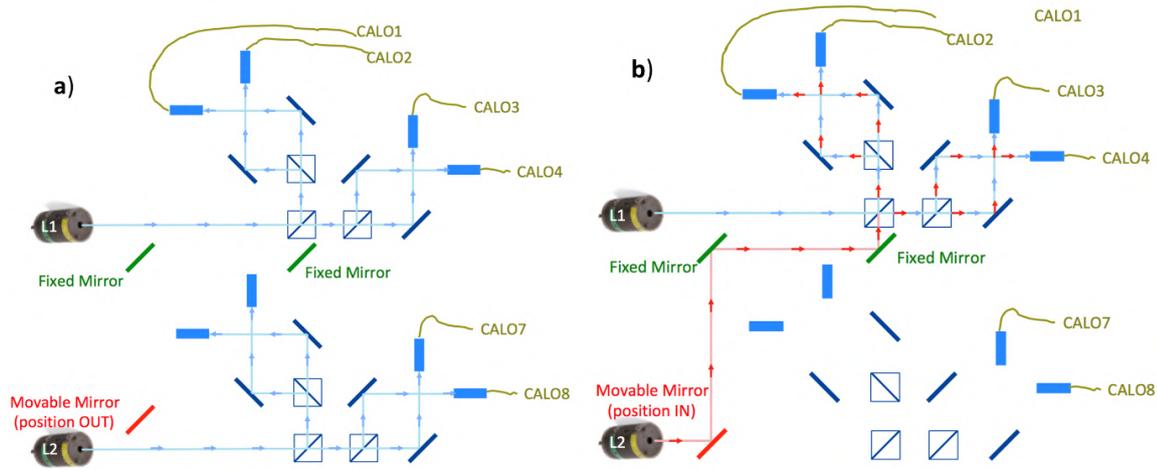
Figure 3.21: (a) Standard laser configuration where the light of each laser head is split in four and injected into four fibers to the assigned calorimeters. (b) Modified setup for double pulse operation in which one laser beam is superimposed into the path of its paired one by a movable mirror (position IN). A similar set of moving+fixed mirrors is present also on the upper laser setup to allow the symmetrical operation. Figure from [52].

For all these reasons, the laser system is capable of sending two different laser pulses to the same calorimeter by inserting a movable mirror in front of a laser head in order to redirect it to the optical path of the next laser head. This scheme is shown in Figure 3.21. An external delay generator (DG) (SRS DG645) is used to send prompt and delayed signals to the two laser heads respectively.

The Short-Term (20 ns) data structure can be measured by inserting the movable mirrors and by operating the delay generator so that the second pulse is delayed in the 0-80 ns range. An example of two double-pulse events are shown in Figure 3.22.

The longer time constant is more complex to measure. The gain drop is in fact due to the overlap of several pulses which overload the HV power supply. Because of this, the prompt signal is not provided by a single pulse, but by a burst of pulses. The test is performed as follows:

- The Laser Control Board sends a burst of $N$ laser pulses separated by an interval $\Delta t = 120$ ns, which is the minimum delay allowed by the laser driver. The filter wheel located after the laser head is set to the maximum transmittance value in order to saturate the SiPMs response. The number of pulses $N$ is what defines the

overall gain sag.

- After a delay in the range of [0,40] $\mu$s, a test pulse is sent from the second laser. An example of such a pulse structure is shown in Figure 3.23.
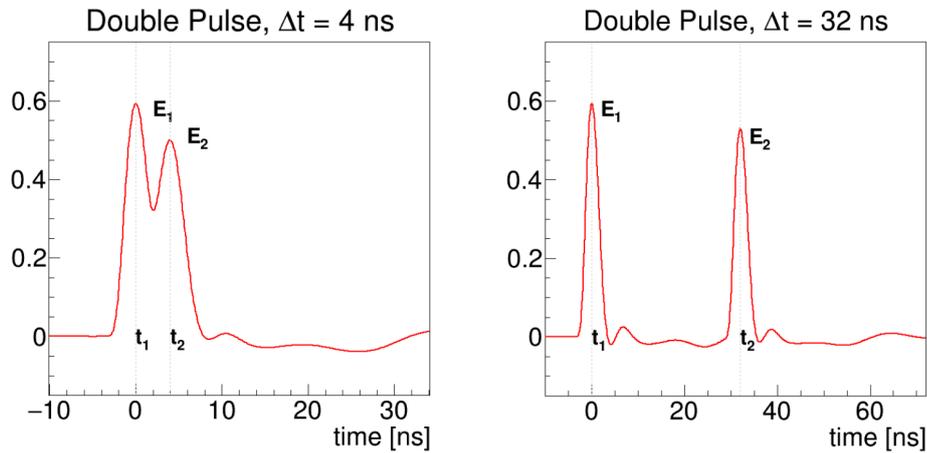


Figure 3.22: Example of two STDP signals produced with a delay of 4 ns (left) and 32 ns (right).
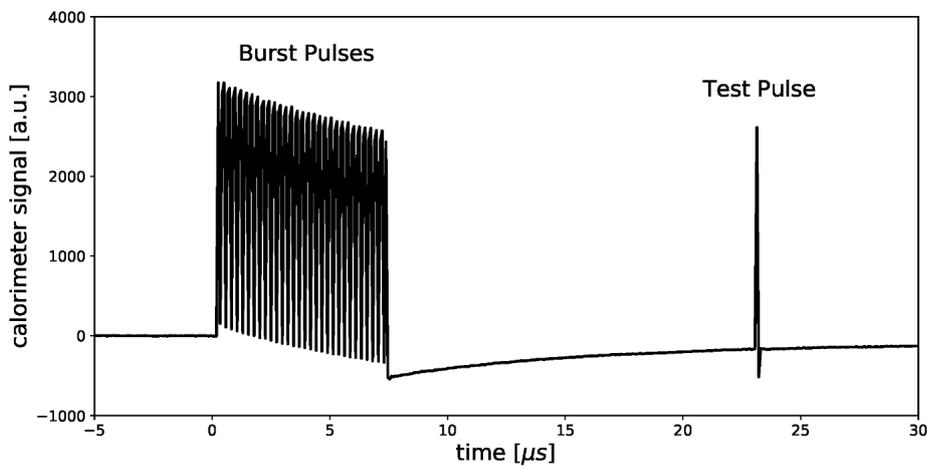


Figure 3.23: Example of a LTDP pulsing pattern. First, a burst of several, equally spaced, laser triggers is sent to simulate the arrival of multiple particles, and then a final test laser pulse is sent with a large and variable delay with respect to the burst.

## 3.5    Simulation

The E989 Experiment makes use of several simulation programs to precisely predict the beam behavior and the particle detection and to compare with real measurements. There are three main simulation programs currently in use and will be described briefly.

### 3.5.1    gm2ringsim

The *gm2ringsim* software is a Geant4 [57] simulation capable of simulating the beam motion from the injection into the storage ring to the detection in calorimeters and trackers [58]. The complete calorimeter and tracker systems are fully and realistically represented in the simulation. The geometry for the straw-tracker modules includes coordinates as determined in alignment surveys. The entire magnet, the vacuum chamber, the kicker and quadrupoles plates are constructed too, and four realistic electric and magnetic fields are implemented, including the radially dependent fringe field that extends toward the center of the ring. The time dependence of the kicker magnetic field is taken from direct magnetometer measurements made at the center of the plates. The fields associated with the quadrupole plates are implemented as a multipole expansion. They are dynamically evolved according to the scraping process and include the perturbations given by the damaged resistors of Run-1.

### 3.5.2    COSY

An independent simulation of the storage ring has been developed using a computational environment called COSY INFINITY [59]. This simulation focuses on the precise modeling of the electric and magnetic fields in the storage region, quantifying nonlinear effects given by the quadrupole system, as well as the higher order magnetic multipoles.
It is based on Differential-Algebraic (DA) methods which allow the preparation of transfer maps that describe the solutions of storage ring beam optics as ordinary differential equations (ODEs). These methods allow to quickly reach an high statistics particle tracking in contrast to conventional orbit-integration beam dynamics programs like gm2ringsim.

### 3.5.3  BMAD

BMAD is a library for simulating the dynamics of relativistic beams of charged particles into beam line elements [60]. The BMAD simulation of the E989 experiment is composed of the M5 beam line, the inflector, and the storage ring. The latter includes a static magnetic field, time-dependent quadrupole electric fields and the kicker magnetic fields. The fields in each element are defined by field maps, multipole expansions, or analytic expressions. Time dependence for pulsed kickers and quadrupoles is provided by custom code. This software has been used for the determination of systematic uncertainties due to the E-field and the pitch correction related to quadrupoles misalignment effect.

## 3.6  Current status

The E989 Experiment is currently active and collecting data for the sixth consecutive year of running, Run-6, which will be the last one. Run-1 has been completely analyzed and published on April 7th 2021, and the complete measurement is described in Chapter 8. Analysis of Run-2 and Run-3 is in progress, while Run-4 and Run-5 are being produced. Figure 3.24 shows the number of collected positrons from Run-1 to Run-5 in terms of the final statistics of the BNL experiment, with a total number of $\sim$19 $\times$BNL collected so far. With the current Run-6, the E989 experiment is expected to reach and exceed the design goal of 21 $\times$BNL of statistics with positive muons, which will ultimately allow for a final statistical uncertainty on the $\omega_a$ measurement of 100 ppb.

After Run-1, many upgrades have been implemented in order to improve the beam storage capability and the stability of the experimental conditions. Some of them are:

- The kicker system has been upgraded to provide higher voltages starting from Run-2 and finally reaching the design values by the end of Run-3. The nominal design voltage of 165 kV resulted in a stored beam centered around the nominal *magic* orbit radius of 7.112 m.

- The quadrupoles conditions have been stabilized and the broken resistors have been replaced.

- Starting from the end of Run-4, a new quadrupole operating scheme has been implemented with a radio-frequency (RF) system which reduces the beam betatron oscillations.

- New specialized detectors have been constructed to measure transient fields generated by the mechanical vibration of the quadrupole plates and by the eddy currents
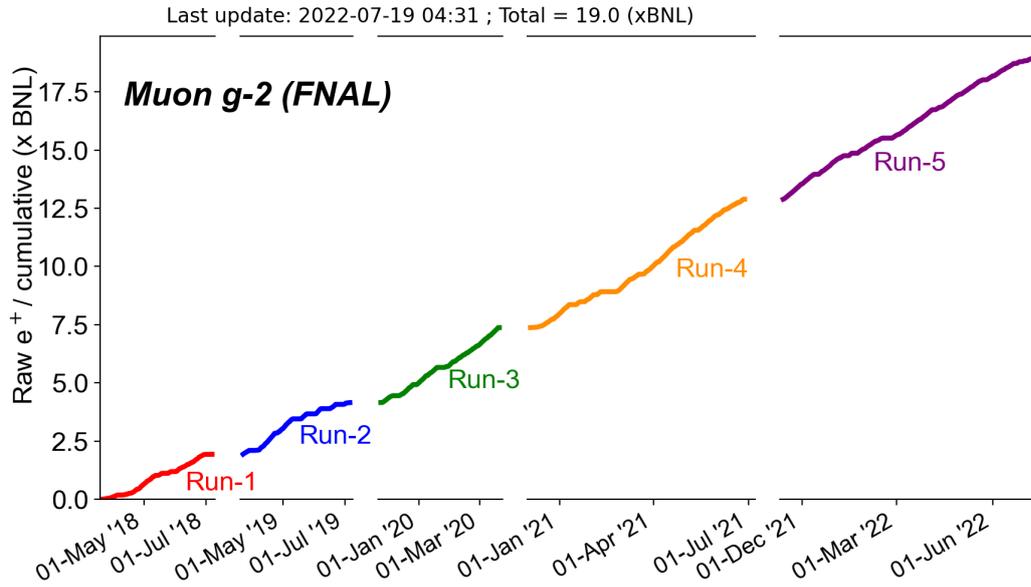
Figure 3.24: Number of accumulated positrons in the first five runs of the Muon $g - 2$ Experiment at Fermilab in terms of the total statistics of the BNL experiment.

induced by the fast kicker pulse. These measurement will be discussed in more detail in Chapter 8.

- A thermal blanket has been installed around the magnet yoke after Run-1 to reduce the daily fluctuations of the magnet temperature which affects its physical shape and magnetic field.

- An air conditioning system has been installed in both the experimental hall and in the laser hut to drastically reduce the daily and seasonal temperature fluctuations of the entire equipment from $\pm 2$ °C to $\pm 0.5$ °C or better. Figure 3.25 shows the temperature of the magnet yokes, the calorimeter SiPMs and the laser optical table from Run-1 to Run-5.

Overall, many systematic uncertainties on both the $\omega_a$ and $\omega_p$ measurements are expected to be significantly reduced in Run-2 and beyond with respect to Run-1 thanks to the mentioned hardware improvements.
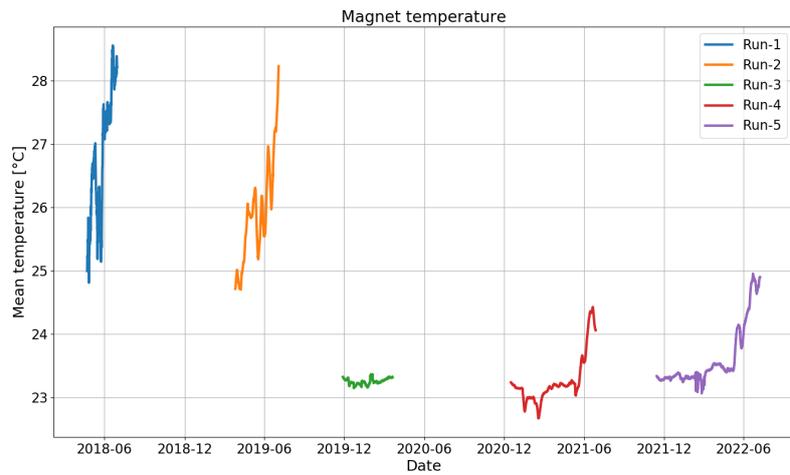
(a) Magnet yoke temperature
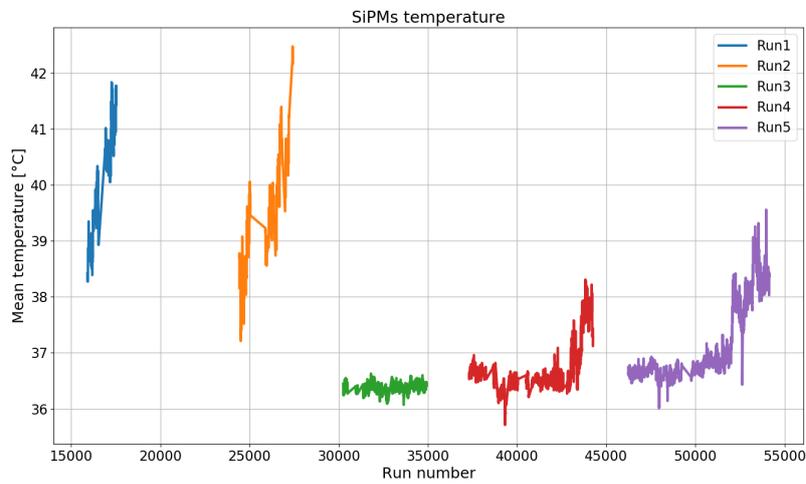


(b) Calorimeters temperature
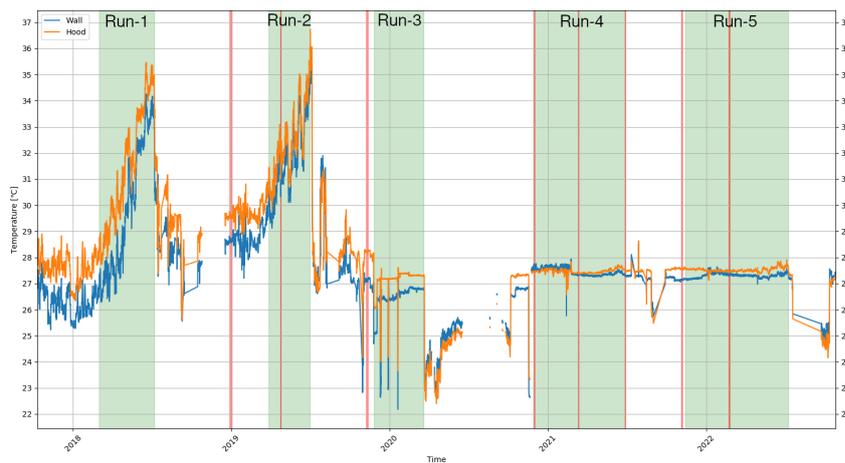


(c) Laser hut temperature

Figure 3.25: Temperature measurements from Run-1 to Run-5 of the magnet yoke (a), the calorimeter SiPMs (b) and the laser hut (c).

# Chapter 4

# Data acquisition and reconstruction

The Fermilab E989 is a high precision experiment, and in spite of the relatively limited number of channels compared to other particle physics experiments, the amount of data that is written, reconstructed, and analyzed is quite sizable.

The 24 homogeneous calorimeters are segmented as matrices of 54 crystals, each one coupled with a SiPM, totaling 1296 channels each sampled at 800 MHz. The two tracker stations are composed of 8 modules each, which in turn consist of 64+64 straw tubes in a U-V plane configuration, totaling 2048 channels read out at 400 MHz. Each $g - 2$ event is defined as a stored bunch, which lasts for 700 $\mu$s. In this time interval, muons circulate roughly 4500 times while decaying, producing $\sim$2000 hits on the calorimeters. Raw files are called subruns, defined to be $\sim$2 GB in size, each one containing $\sim$8 seconds of data and roughly 100 muon events (fills). Out-Of-Fill laser pulses fired to calorimeters between one bunch and the next are saved as well. When the experiment is running (typically from autumn to spring), data is collected continuously, producing $\sim$12000 subruns and 24 TB of data every day. Subruns are grouped into runs every $\sim$500 files, summing to 1 TB size. Finally, the runs are assembled into datasets for production and analysis. Typically, a dataset contains $\sim$ 100'000 subruns.

The current data collected in the first 5 years sums up to $\sim$7 PB. While some data-quality selection is performed after production, most of the data is actually used for the determination of $a_\mu$, as the goal of the experiment is to collect more than $10^{11}$ high energy positrons ($E > 1700$ MeV and $t > 30$ $\mu$s) for a final statistical uncertainty of 100 ppb. In order to achieve this, roughly 10 PB of raw data (including simulation) must be handled, produced, and analyzed entirely, granting the experiment a place in the *Big*

*Data* category.

The vast amount of data to be processed is mainly due to the fact that for each positron detected by the calorimeters, 9 SiPM waveforms are saved on average. Each waveform is sampled at 800 MHz, and there are ∼40 12-bit samples per waveform [56]. All the detected positrons must be saved and reconstructed: even if the low-energy ones do not contribute to the measurement of $\omega_a$, they participate in the pileup correction and in the study of the systematics.

The raw data is processed in three ways with different purposes:

- **online processing**: a small fraction ($\sim 10 - 50\%$) of data is analyzed on the fly from the detector back-ends for generating Data-Quality-Monitoring (DQM) purposes. Diagnostic and raw plots are generated for all the experiment detectors and can be consulted on a live website using the *plotly* suite.

- **nearline processing**: all the acquired data is analyzed in a simplified reconstruction ∼30 minutes after acquisition and the main quantities of interest are plotted on a website with the help of *Bokeh* [61] and a *ROOT* [62] visualizer. This is useful for quick physics analysis during the fine-tuning of the experiment.

- **offline processing**: the complete and fully calibrated reconstruction of the data, submitted on the distributed Open Science Grid with the management tool developed for Fermilab, Production Operations Management System (POMS) [63], with a version-controlled *art* software [64]. This is what generates the reconstructed output files used for the final analysis of the data.

## 4.1   Offline production workflow

The reconstruction of a dataset is a process that requires multiple sequential steps. A single uninterrupted procedure is not possible as the calorimeters need to be calibrated per subrun (a single raw file) and per dataset. For this reason, the production is split into a pre-production phase and a full-production phase. Both phases start from the raw files, but the pre-production only contains the minimum reconstruction needed for the extraction of the calorimeter calibrations. For convenience, a series of per-fill Data Quality Checks (DQC) are also performed at this stage.

The full reconstruction can be broken down into 8 steps as illustrated in Figure 4.1: *Pre-staging, Pre-production, Calibration and In-Fill DQC, Database upload and migration, Test production, Test DQC and validation, Full production*, and *Final DQC*. A typical dataset contains ∼50k files, and there are 70 datasets from Run-1 to Run-5. Starting from Run-2 the production was performed in a rolling scheme: each of the 8 steps happening simultaneously on different datasets, like a typical manufacturing production chain.
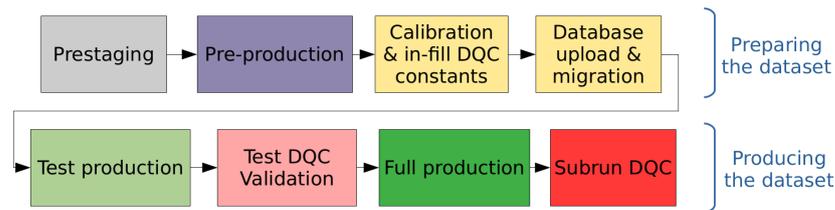


Figure 4.1: The 8 steps of the Muon $g - 2$ production.

### 4.1.1   Pre-production

The first step of pre-production involves copying the data from the tape system to the disk cache. This *pre-staging* of the data is necessary as the amount of raw data is too big to be always stored in disks. Moreover, "on-the-fly" direct copying from tape to the grid nodes would be extremely inefficient. As the amount of data to be reconstructed substantially increased from Run-1 to Run-5, this step became increasingly important. Careful planning is needed so that, while some datasets are being processed, the next ones are in the prestaging phase and ready as soon as the previous ones are completed. In addition, the tape system is shared among all the Fermilab experiments, and the resources must be carefully distributed to avoid periods of competition.

After a dataset has been fully pre-staged, pre-production is performed. It consists of an unpacking step followed by a light reconstruction of the positron data. At this stage, only pulse fitting and minimal calibration is executed. Consequently to the pre-production, analyzer jobs automatically run to extract gain calibration data and quality checks of individual muon fills. The calibration constants and quality cuts are then manually verified by calibration responsibles, before proceeding in uploading them to a *condition* database. A fill passes the quality checks if it registered the laser sync pulse, if the measured kicker strength and timing were nominal, and if the T0 detector measured

a nominal beam profile.

The final step in the pre-production phase consists in uploading the extracted constants and conditions to a *development* database and then migrating them to a *production* database. These constants are stored as *data* and *status* tables, each marked with an Interval of Validity (IoV) which can be a range of runs or subruns. The database is accessed via a PostgreSQL interface, and when a subrun is reconstructed a `https` query extracts the corresponding table from the production database.

### 4.1.2   Full-production

When pre-production and calibration is done, the dataset is ready to be fully produced. Full production includes unpacking of raw data and full reconstruction of all detectors, applying the calibration constants derived from the previous phase. Three independent reconstructions for the calorimeters are executed.

Before starting the full production, a portion of approximately 10% of the files, uniformly distributed across the dataset, is reconstructed to detect potential problems. Examples are miscalibrations, detector faults, software bugs, or changes in the experiment configuration. A dedicated validation step is performed on the produced files and many plots are generated to check for physical quantities, such as calorimeter energy spectra, and anomalies in the calibration constants applied to the data. The different reconstructions are compared with each other and long-term trends on the beam storage efficiency are checked. Once the data quality is assessed by shifters and experts, full production resumes on the remainder of the dataset.

**Data quality checks**

Finally, when all the files of a dataset have been produced, a second round of validation is performed and cuts are applied to select high-quality files only. In particular, quality tests include:

- number of positrons observed in a fill:

- number of muons interacting with collimators, or other obstacles, and exiting from the storage region before decaying;

- intensity of beam at injection;

- average fill-quality pass rate in a subrun.

Once this process is complete, typically $\sim 10\%$ of the files are discarded while $\sim 99\%$ of the positrons are actually kept[1]. The selection of good files is finally delivered to the analyzer teams.

## 4.2   Computing resources

The production of many petabytes of raw data requires a non-trivial amount of resources and careful management of the whole process. This section will summarize the resources needed for the production of the $g - 2$ data, the challenges faced and the lesson learned. Most of the tools for managing the production processes are provided by the *FabrIc for Frontier Experiment* (FIFE) project maintained by Fermilab [65]. These tools interface with the tape libraries, the disk servers, the distributed computing, and the $g - 2$ data acquisition back-ends.

### 4.2.1   Tapes and disks

As already mentioned in section 4.1.1, an important step is the handling of the data itself. Both the raw and produced files are stored in tape libraries containing thousands of 12 TB tape cassettes. Each file pertains to a *file family* meaning that different families are not mixed in the same tapes. This ensures that, when a file family can be discarded (i.e. obsolete data), the relative tapes can be recycled.
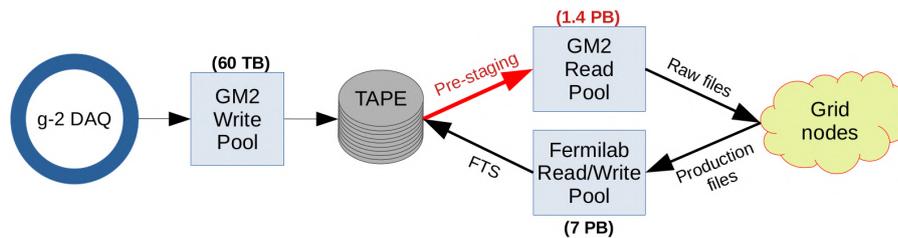


Figure 4.2: Data movement from DAQ to production. The pre-staging of raw files from tape to the dedicated GM2 read pool is highlighted in red.

The data can be copied from tape to disks with *samweb* commands interfacing the tape *ENSTORE* [66] system. The Muon $g - 2$ Experiment has a dedicated $\sim 1400$ TB disk pool for temporary storage of the raw data to be produced (Figure 4.2). All the

---

[1]Most of the discarded files are actually empty or have a very limited statistics.

output from production goes to a general $\sim 7$ PB pool and is copied to tape. A new *migration* mode involving the dump of an entire tape at once recently helped us increase the efficiency of the pre-staging step.

### 4.2.2   Computing nodes

The offline production of the data runs both on the Fermilab computing nodes and on the Open Science Grid (OSG) distributed around the world. The Muon $g-2$ Experiment has currently 5700 reserved slots in the Fermilab grid, where each slot consists of a CPU core and 2 GB of memory.

### 4.2.3   Shifters

While the multi-step *rolling* production scheme mentioned in section 4.1 greatly improved the production efficiency, the increased number of datasets required the organization of production shifts. Setting up such shifts and providing the full documentation and the correct instruction required a non-negligible effort by the production managers that I contributed to coordinate.

### 4.2.4   Simulation

The simulation of the E989 experiment is an important piece of the physics analysis and it is developed for studying the beam motion inside the storage ring. Many simulation packages are used to simulate various parts of both the beamline starting from the proton target and the storage ring and are described in Section 3.5. The simulation has to consider electric and magnetic fields, the beam dynamics, the muon decays, and the interaction with the detectors. This means to precisely track the particles over 220 km of flight distance, which corresponds to $\sim 5000$ turns inside the ring, traveled during the 700 $\mu$s storage duration. The most CPU-intensive parts of the simulation are performed with the help of High-Performance-Computing (HPC) jobs at the National Energy Research Scientific Computing Center (NERSC).

## 4.3   Outlook

The data production of the Muon $g-2$ Experiment is a challenging but required task for achieving a new world-best measurement of the muon anomalous magnetic moment. The data grew almost 20-fold from the Run-1 publication to the current amount, which

will allow a final statistical uncertainty of 100 ppb on $a_\mu$. During the last year, great efforts successfully increased the resources, the production speed, and improved the management and the efficiency of the data production workflow.

The full-production step, which is the most resource intensive one, improved by a factor of two between Run-2 and Run-3, and by another factor of two between Run-3 and Run-4, as shown in Table 4.1. Figure 4.3 shows the timeline of the full-production iterations for the various datasets. Some of them have been produced multiple times for testing or debugging the software. Dataset Run-4U was produced five times when major upgrades were introduced in the $g-2$ reconstruction code, one of which is the new ReconITA reconstruction that will be presented in the next Chapter.

The E989 production team is now starting to produce the Run-5 data, and a new production scheme is now being tested to perform the pre-production of Run-6 data directly as soon as it is generated from the experiment.

| Dataset | N° of files | Prod. time [days] | Rate [files/day] |
|---------|-------------|-------------------|------------------|
| Run-2 | 267152 | 129 | 2071 |
| Run-3 | 538622 | 122 | 4415 |
| Run-4 | 992005 | 111 | 8937 |

Table 4.1: Full-production performance per aggregated dataset.
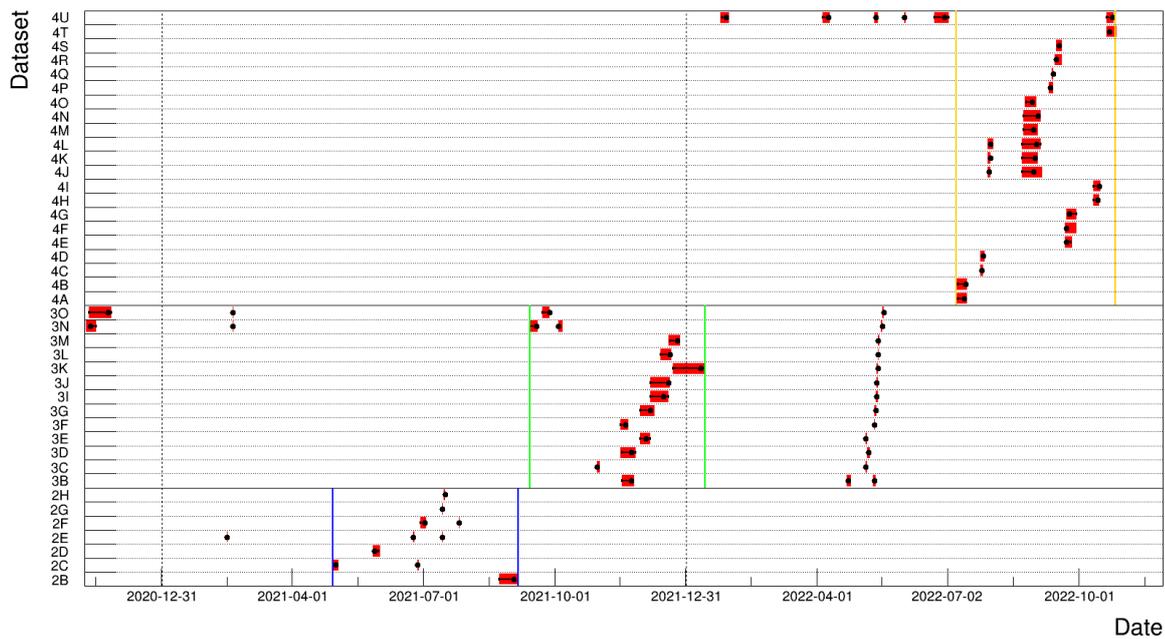
Figure 4.3: Timeline of the production of all the datasets from Run-2 to Run-4. For each data point, the red bar marks the elapsed time between the beginning and the completion of the dataset. The dot and the error bars mark the median (50%), the 5%, and the 95% percentiles completion points. Multiple data points on the same horizontal line correspond to multiple productions of the same dataset. The blue, green, and orange lines mark the beginning and the end of the Run-2, Run-3, and Run-4 final production campaigns respectively. Only the full-production step is displayed here.

# Chapter 5

# The reconstruction of calorimeter data

The measurement of the muon precession frequency is performed by observing the positrons emitted by the muons when they decay. The set of 24 calorimeters described in Chapter 3 has the purpose of measuring the arrival time and energy of the outgoing positrons. A precise time measurement is necessary for the extraction of the precession frequency $\omega_a$, and a very stable energy measurement is required for determining the selection of positrons according to the decay asymmetry. The positrons generate electromagnetic showers inside the $PbF_2$ crystals and the Čherenkov light produced by the charged particles is collected by each crystal's SiPM. The readout of each SiPM is sampled and digitized at 800 MHz. The reconstruction of such signals is an important step, having the goal of separating unique positrons and measuring their hit time and energy precisely while maintaining a calibrated and stable detector across time.

This chapter will describe the entire process that transforms raw calorimeter pulses into reconstructed positrons, as well as my contribution to several steps of this reconstruction chain. A new procedure for fitting the raw pulses and assembling the positron hits has been developed under the name of *ReconITA*.

## 5.1   Motivation

The ambitious goal of the Muon $g-2$ Experiment is to measure the muon anomaly $a_\mu$ with a precision of 140 ppb. This requires having both statistical and systematic uncertainties on $\omega_a$ to be below the 70 ppb level. Such a level of precision and confidence in

the measurement requires meticulous scrutiny of all the possible effects that could bias the measured value of $\omega_a$.

The analysis of the first year of data (Run-1) saw 6 different teams independently calculating the value of $\omega_a$, each utilizing a unique mix of reconstructions, analyses, and data-driven corrections. These approaches, in spite of being strongly correlated, have varying sensitivities to potential systematic effects, as well as varying statistical sensitivities. Among the reconstruction techniques of Run-1, there are three main approaches. Two of them are based on the reconstruction of the individual positrons hitting the calorimeters, whereas the third accumulates the total energy sum deposited in the calorimeters.

Before discussing the motivation and the role of the new *ReconITA* approach, I will briefly describe the three mentioned reconstruction techniques first.

### 5.1.1   The *local* method

As mentioned in Section 3.3, the calorimeters provide positron signals in the form of digitized waveforms with a typical length of 50 ns ($\sim$40 samples at 800 MHz). The waveforms of each SiPM that received a sufficient energy deposit and the surrounding ones are saved in the raw files.

The first reconstruction technique treats the signal from different SiPMs separately and reassembles them into positron clusters after being calibrated in energy and time. For this reason, it is called the *local* fitting method and will be referred to as *ReconWest* (RW) because it was developed mainly by A. Fienberg at the University of Washington [50]. The reconstruction chain involves the following sequential steps:

1. Pulse fitting
2. Time calibration
3. Energy calibration
4. Clustering

Step 1, *pulse fitting*, involves the preparation and the template fit of the waveforms. A typical waveform is shown in Figure 5.1. First, a *pedestal* correction is applied to the individual trace samples. The ADC converters (TI ADS5401) mounted on the custom Waveform Digitizers consist of two interleaved ADCs each operating at half of the ADC

sampling rate but 180° out of phase from each other. Therefore, half of the waveform samples are digitized by one ADC and half by the other [56]. A small offset between the two modules introduces an even-odd oscillation of a few ADC units (Fig. 5.1). These offsets are measured and corrected once per Run and for each SiPM. Then, a template fit procedure is applied to the pedestal-corrected traces. The pulse templates are extracted with a dedicated analysis at the beginning of each data taking period, by using data from dedicated calibration acquisitions. Signals from high energy positrons are accumulated and the average determines the template shape [50]. Each trace can contain one or more positron signals, and the fitting procedure tries to individuate and fit all of them.
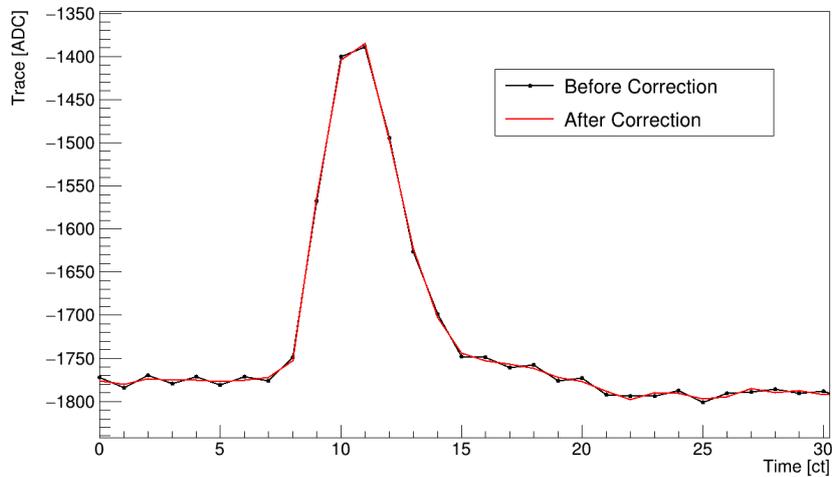


Figure 5.1: Example of a digitized trace before and after the pedestal correction. The small even-odd sample oscillation is introduced by the ADC converters.

Step 2, *time calibration*, takes care of the synchronization of the 1296 SiPMs and the time alignment with respect to the beam injection. The Laser Calibration System described in Section 3.4 is fundamental in this process, since a *sync* pulse, fired to all crystals a few microseconds before injection, serves as a time reference.

Step 3, *gain calibration*, converts the raw energies extracted from the template fit into physical MeV units. This step is composed of four sub-steps. The first three make use of the laser signals to precisely stabilize the gain of the 1296 SiPMs across time. The three corrections calibrate the channels at the $\mathcal{O}(10)$ ns, $\mathcal{O}(10)$ $\mu$s, and $\mathcal{O}(10)$ s timescales respectively as described in Section 5.4. Finally, the energy is converted into the absolute MeV scale, equalizing the response of all the different SiPMs.

Step 4, *Clustering*, is the final stage, where hits coming from different SiPMs are grouped together in a cluster representing a single positron. The positron time and energy measurements are extracted from the cluster information. If hits from different positrons are mistakenly grouped together in the same cluster, a *pileup* event is formed. Such events are then treated at the analysis stage.

## 5.1.2   The *global* method

An independent team developed a new reconstruction algorithm based on the strategy of looking at an entire calorimeter simultaneously instead of fitting individual crystals and clustering them afterward. For this reason, it is defined as the *global* fitting approach and will be referred to as *ReconEast* (RE) because it was mainly developed by D. Sweigart at Cornell University [56].

In this approach, the algorithm simultaneously fits all the crystals, within a calorimeter, whose waveforms are generated roughly in the same time frame. Each positron with an energy over a threshold of 60 analog-to-digital counts (ADC) above noise, corresponding to approximately 50 MeV, is identified with a 3×3 cluster of crystals. The pulse magnitude for each crystal pulse floats independently in the fit, but the peak time is constrained for each crystal to be at the same value. Clusters that share one or more crystals must be separated by at least 1.25 ns; otherwise, they will be merged into one larger cluster and treated later during analysis. This approach also inherently imposes spatial separation between positrons that hit a calorimeter very close in time, thus reducing the amount of pileup contamination.

## 5.1.3   The *integration* method

A third independent reconstruction aims to measure $\omega_a$ by not identifying individual decay positrons but combining the raw crystal waveform samples into contiguous 75 ns windows [67]. This *integration* approach will be referred to as the *QMethod* since the final product is a measure of the deposited charge in the calorimeters versus time. The time-rebinned waveforms span a time period of -6 $\mu$s $< t <$ 231 $\mu$s relative to the beam arrival time. The reduced time range and increased time binning were chosen to limit the rate and volume of the integrated energy data.

Ideally, a simple sum of the waveforms over the 54 crystals from a calorimeter would yield the integrated energy waveform for that calorimeter. As Figure 5.2b illustrates, while positron pulses appear clearly in a single waveform, the slow recovery of the pedestal dominates the signal. A threshold is therefore applied to separate the positron signals from the baseline as shown in Figure 5.2a, after the same gain corrections described for the event-based approaches are applied to the data.

The QMethod has the benefit of being immune to the systematic effects given by the pileup contamination, but has a lower statistical power because of the reduced measurement time window and the inclusion of low energy pulses with negative decay asymmetry.



(a)                                           (b)
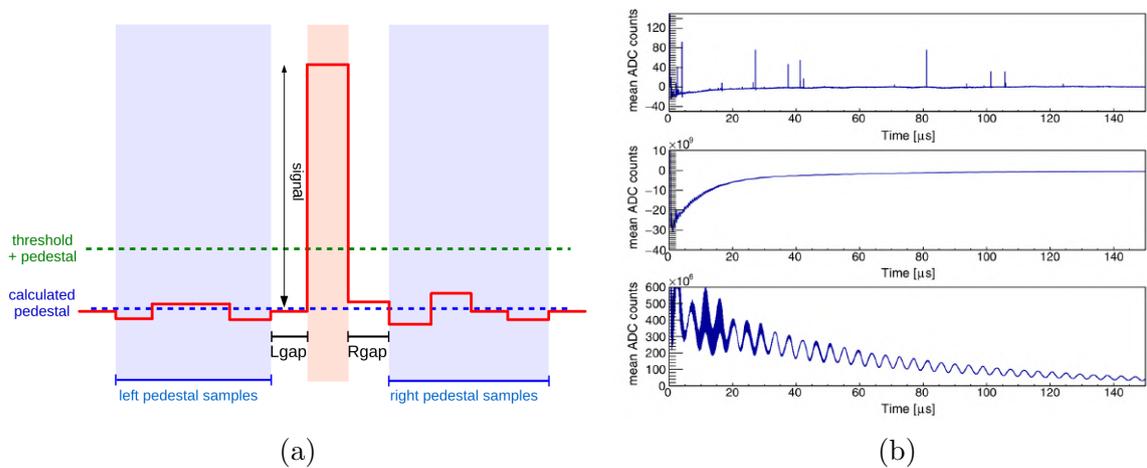
Figure 5.2: (a) Diagram illustrating the pedestal calculation algorithm and the application of the threshold for the threshold integrated energy waveforms. (b) Integrated signal for one crystal (top), the corresponding signal sum over all crystals for an entire dataset (middle), and the final pedestal-removed signal (bottom). Figures from [68].

### 5.1.4   ReconITA

The Run-1 $\omega_a$ measurement has an error vastly dominated by the statistical uncertainty, as will be discussed in Chapter 8. Some considerations and decisions on the techniques used for the Run-1 analysis were based on that fact. However, Run-1 is only ~5% of the currently acquired data, and with the much higher statistics of Run-2, Run-3, and beyond, the systematic errors will contribute increasingly more. It is then important to improve the reconstruction and analysis algorithms and study with great detail effects which were negligible in Run-1.

Prompted by this motivation, the Italian group started the **ReconITA** project, of which I led the software development, with the goal of reducing some of the systematic errors, in particular those who fall under the category of *Pileup contamination* and *Early-to-late effect* in Table 5.1.

*Pileup* is one of the major components, and is due to the contamination of double or triple positrons in the same reconstructed event. The *ReconITA* approach aims at reducing the amount of pileup contamination in the data by increasing the positron separation efficiency, while also improving the pileup subtraction technique, thus reducing drastically this systematic error.

Then, the *Early-to-late effect* is a systematic uncertainty that was introduced because of residual effects seen in the data whose nature was not completely understood. A reconstruction effect is among the possible explanations, and the *ReconITA* approach also aims at understanding and possibly removing this source of systematic error. More details on this matter will be discussed in Section 7.6.

Finally, the underlying purpose of the *ReconITA* project is to have a different approach on some of the key parts of the positron reconstruction chain, developing an independent reconstruction and analysis procedure, and building more confidence in the final E989 $\omega_a$ measurement.

| Uncertainty | Run-1A | Run-1B | Run-1C | Run-1D |
|---|---|---|---|---|
| Time synchronization | 4 | 1 | 1 | 1 |
| Gain corrections | 12 | 9 | 9 | 5 |
| Pileup contamination | 39 | 42 | 35 | 31 |
| Beam dynamics | 42 | 49 | 32 | 35 |
| Early-to-late effect | 21 | 21 | 22 | 10 |
| Hit randomization | 15 | 12 | 9 | 7 |
| Total systematic uncertainty | 64 | 70 | 54 | 49 |

Table 5.1: Systematic uncertainties of the $\omega_a$ measurement for the four subsets of Run-1. All values are in parts per billion.

**Outline**

The main contributions of the *ReconITA* reconstruction regard the first and last steps of the *local* fitting method, listed in Section 5.1.1. The pileup separation efficiency for very

close positrons is limited by the pulse fitting performance in distinguishing two signals at very low $\delta t$. If two almost simultaneous positrons hit a calorimeter in the same region, the two signals affecting the same crystal might be too close to be properly characterized by the template fitting routine. For this reason, the *local* fitting algorithm implements a 2 ct artificial dead time after each fitted pulse [50]. This means that the template fitter cannot separate two pulses closer than 2.5 ns.

The ReconITA solution to this problem is to redo the pulse fitting after the first reconstruction round has been completed, taking advantage of the reconstructed clusters to constrain the time of the fitted pulses. In this way, signals closer than 2.5 ns that could not be correctly distinguished in the first reconstruction, can then be separated thanks to the time information coming from the other surrounding crystals. Then, when the clustering algorithm is applied again at the end, it can further separate positrons at very small $\delta t$ while being robust in accurately reconstructing their energies.

The ReconITA reconstruction software starts from the output of step 3 described in Section 5.1.1 and perform its own clustering algorithm. Then, the raw traces are passed through the new pulse fitting algorithm with constrained peak times obtained from the clusters, in search of hidden pulses. These new hits are then calibrated in time and energy and a final round of clustering is applied. Figure 5.3 sketches this process. The details of each step will be discussed in the following sections.
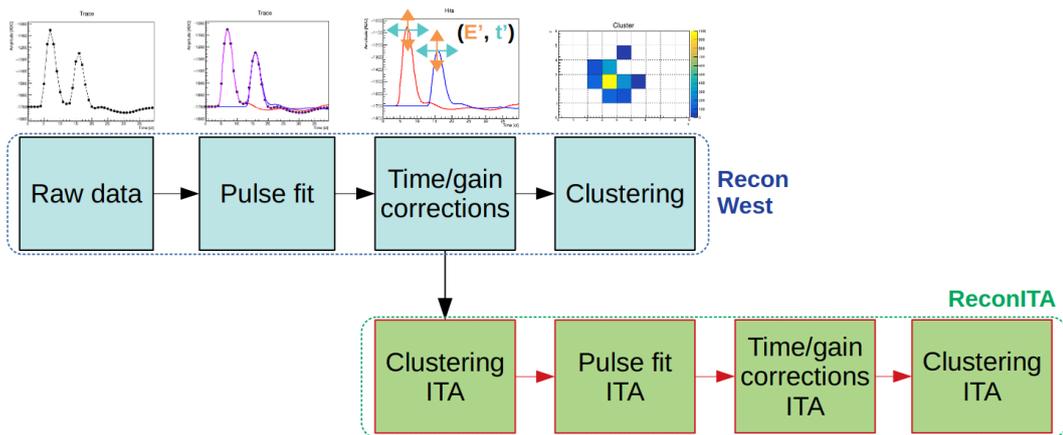


Figure 5.3: ReconITA flow diagram for reconstruction (green) and analysis (red). Our algorithm is appended at the end of the ReconWest reconstruction.

## 5.2   Pulse Fitting

The local-fitting approach only looks at one SiPM at a time, searching for over-threshold signals and fitting the traces with one or more templates depending on the observed number of pulses. The threshold is fixed at 75 ADC*ct, which corresponds to roughly 25 MeV, depending on the crystal[1]. The algorithm also performs noise suppression by removing the periodic pulses in the SiPMs which are sometimes produced by a 160 MHz noise oscillation of the baseline [56]. In addition, if two subsequent waveforms belonging to the same SiPM are sufficiently close in time, the template fits are repeated by propagating the first pulse into the second waveform (Figure 5.4).
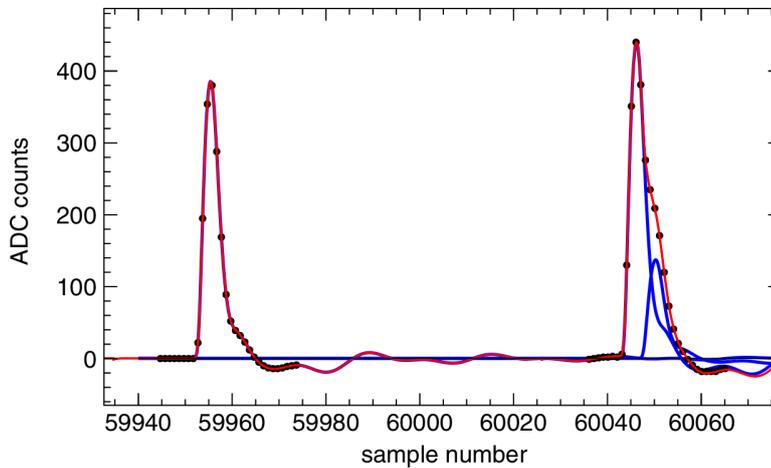


Figure 5.4: Example of a chain fit containing waveforms from two separate time windows and three pulses. The baseline perturbations from the first pulse persist into the second time window, in which two pulses separated by 5 ns are identified. Figure from [68].

One of the fitter parameters that can be tuned is an Artificial Dead Time (ADT), set to be 2 ct (2.5 ns), which essentially makes two pulses closer than this value indistinguishable by construction as anticipated in Section 5.1.4. This was introduced to limit the cases where small distortions of the waveform could lead to a double-pulse fit when only one signal is actually present.

In order to increase the pileup separation efficiency, which also heavily depends on the clustering algorithm, the ReconITA pulse fitter tries to decrease the dead time. The

---

[1]At the time of writing this thesis, the local approach has been upgraded to use thresholds which vary according to the SiPM gain. The new data reconstructed with this upgrade has not been fully analyzed yet.

time information given by the cluster reconstructed during the first round is used to constrain the time of one of the possible pulses present in the waveform. The template fitter then tries to find another pulse, with floating time, in the same waveform. However, a small but non-zero dead time has to be applied in any case, as a fit of two perfectly overlapping pulses would result in an indeterminacy of the two energies since the same template shape is used for both. The same technique is also applied when two nearby clusters were already successfully separated by the first reconstruction, in search of pulses belonging to the second cluster that were hidden by the tails of the first cluster pulses.

The most energetic hit of a cluster is used as a time reference for constraining the fit of the other hits pertaining to the same cluster and will be referred to as the *seed*. This choice is motivated by the distribution of the time distance of the cluster hits with respect to the most energetic one, shown in Figure 5.5, which is strongly peaked at $\Delta t = 0$. In addition, the most energetic hit is the one that provides the most precise fit parameters because of the higher signal-to-noise ratio.
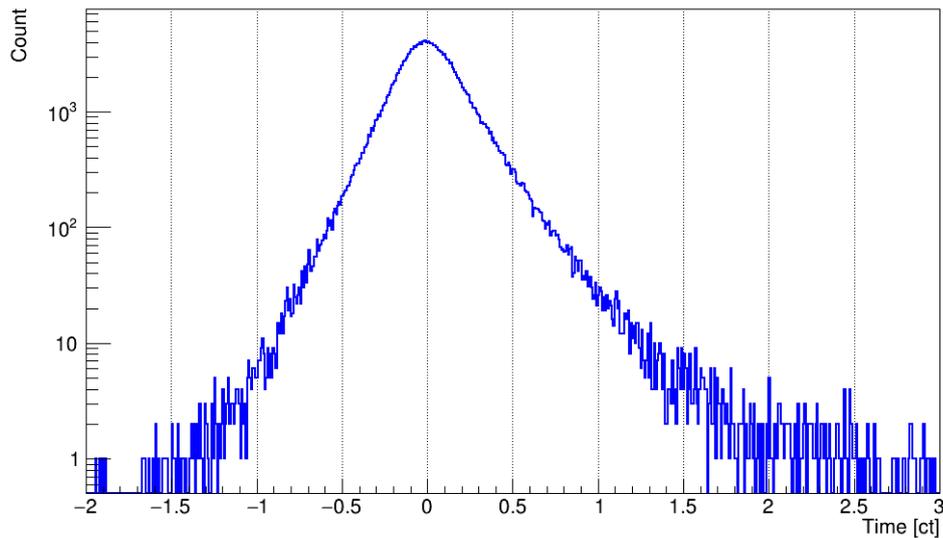


Figure 5.5: Time distribution of the individual hits with respect to the most energetic pulse in a cluster. Clusters at the end of the fills are selected to minimize pileup contamination. The most energetic pulse at $\Delta t = 0$ is omitted.

The algorithm then proceeds in performing a template fit with two pulses, one constrained to the seed, the other free to float. The constraint is not rigid, as is it applied

as a parabolic term added to the $\chi^2$ calculation of the fit:

$$\chi^2 = \left[ \sum_{i=0}^{N} \left( \frac{y(i) - f(t_1, t_2, E_1, E_2)}{y_{err}(i)} \right)^2 \right] + \left( \frac{t_1 - t_{constraint}}{\sigma_{constraint}} \right)^2 \tag{5.1}$$

The width of the quadratic term is determined by the parameter $\sigma_{constraint}$ which is set to 0.1 ct according to the cluster hit time distribution of Figure 5.5. Figure 5.6a shows a comparison between a standard and weighted $\chi^2$ scan, while Figure 5.6b shows an event in which the two pulses are correctly separated by ReconITA, whereas they are merged by the standard algorithm.
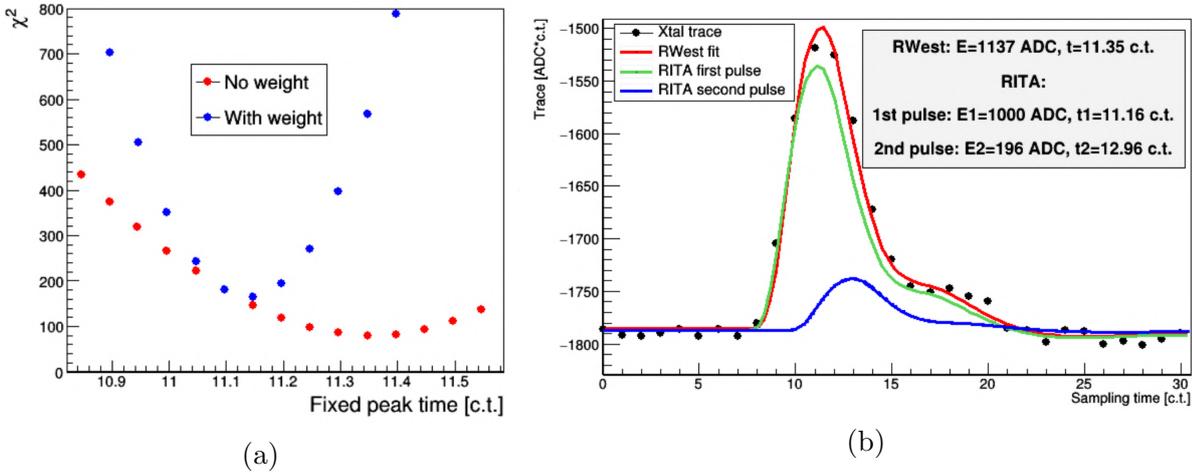


Figure 5.6: (a) Scan of the $\chi^2$ function versus $t_1$, the time of the constrained peak. (b) Example of an artificial island with two pulses separated by 1.8 clock ticks. The ReconITA fitter manages to correctly identify both peaks.

When the ReconITA algorithm fits two pulses, they are not always accepted as physically valid. Four criteria determine whether a trace contains two peaks:

- Time separation of at least 0.5 clock ticks
- Absolute energy of at least 75 ADC*ct ($\sim$25 MeV)
- Relative energy of each pulse above a varying threshold $T(E)$
- Ratio of energies above a time-varying threshold $T(\Delta t)$

The first one is an Artificial Dead Time, introduced because of the previously mentioned indeterminacy of the energies for overlapping pulses, but lowered to the value of 0.5 ct (0.625 ns) with respect to the standard algorithm. The second one is the minimum energy that can be extracted for each pulse, and has been chosen to be the same as the *local*

approach for consistency. The third and fourth criteria determine if the non-constrained pulse is physical or not. For example, waveform distortions introduced by noise could generate fake pulses. The next two paragraphs will discuss the last two criteria.

### 5.2.1 Simulation studies

A dedicated study was performed on artificially constructed waveforms to determine whether the result of a double-pulse constrained fit generates fake pulses and to construct a criterion to suppress such cases.

First, artificial traces are constructed by adding a single pulse template to a baseline (pedestal). The trace is sampled at 800 MHz frequency (1.25 ns per sample) and then each trace point is smeared according to a Gaussian distribution with mean $\mu = 0$ and variance $\sigma(t, E)$, which depends on the time distance from the peak and energy of the pulse. The variance $\sigma(t, E)$ has been extracted by analyzing real positron pulses from a sample near to the end of the beam storage time of 700 $\mu$s where pileup contamination is at minimum. From these events, the difference between the observed trace and the fitted pulse is analyzed. Figure 5.7a shows the fit residual distribution from the sample, aligned at the pulse peak for $t = 0$. The time variance of the distribution is extracted for each time slice of width 0.1 ct and for each energy value of the pulse. Figure 5.7b shows the trace variance distribution as a function of energy from 0 to 6000 ADC*ct ($\approx 2000$ MeV).



<center>(a)                              (b)</center>
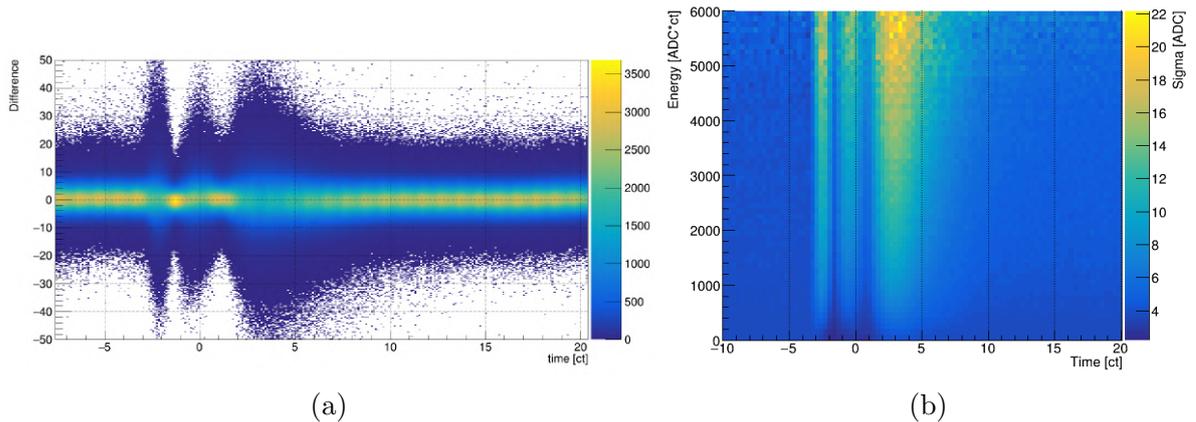
Figure 5.7: (a) Fit residual of a template fit of a single pulse waveform. The pulse peak is centered at $t = 0$, and the values on the y-axis are in ADC counts. (b) Fit residual variance $\sigma(t, E)$ as a function of pulse energy and trace sample time.

The artificial traces are then fitted with the ReconITA algorithm by fitting for two

pulses according to Equation 5.1, with one pulse free to float starting at $t_1 = 0$ and one constrained to be at an arbitrary $t_2 \neq 0$. As the artificial traces contain only one real pulse, the energy of the fitted pulse at $t_2$ can be used to construct a criterion to individuate fake pulses. Figure 5.8 shows the energy distribution of the second (fake) pulse for different energy ranges of the first (true) pulse, when $|t_2 - t_1| = 2$ ct. The energy is observed to depend on the energy of the true pulse.
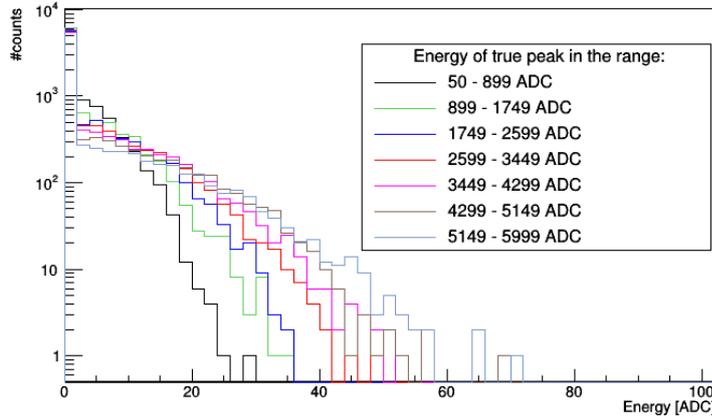


Figure 5.8: Energy distribution of the fake pulse found by a constrained fit at $t_2 = 2$ ct, for different energy ranges of the real pulse at $t_1 = 0$. Courtesy of L. Cotrozzi.

By defining the endpoint $E_n$ as the energy corresponding to the $n$-th percentile of the distribution, a cut such that only energies $E_2 > E_{99}(t, E_1)$ are considered ensures that 99% of the fake pulses are discarded from the output of the ReconITA fitter. Applying an energy threshold defined as:

$$\frac{min(E_1, E_2)}{E_1 + E_2} > \lambda, \tag{5.2}$$

the parameter $\lambda$ can be tuned to minimize the amount of fake pulses being created by the ReconITA fitter while maximizing the amount of true pileup separation.

Figure 5.9a shows the fraction $\epsilon_1$ of traces in which a fake pulse passes the energy cut of Equation 5.2 as a function of $\lambda$. The traces are divided into groups according to the total fitted energy $E_1 + E_2$, which is a proxy for the trace integral. The lowest value of $\lambda$ which keeps the $\epsilon_1$ value below the percent level strongly depends on the total energy of the trace.

A second study is performed by constructing a sample of artificial traces with two pulses, one at $t_1 = 0$ and one, $t_2$, distributed uniformly in the $[-3, +3]$ ct range. For both pulses, the energies are assigned randomly according to the observed crystal energy spectrum from real isolated positrons. The trace is smeared according to Figure 5.7b. The ReconITA fitter is then applied with one pulse free to float starting at $t_1 = 0$ and one constrained to be at $t_2$. The fraction of traces in which the fitter is able to separate the two pulses is defined as $\epsilon_2$. This efficiency depends on the $\lambda$ cut described above and is shown in Figure 5.9b. For the same value of $\lambda$, an higher pulse separation efficiency is achieved for lower energy values.

The two parameters $\epsilon_1$ (false positive rate) and $\epsilon_2$ (true positive rate) are compared as a function of $\lambda$ to form the ROC curves shown in Figure 5.10a. The chosen values for $\lambda$ in the pulse fitter criterion are those who maximize the pulse splitting efficiency $\epsilon_2$ while keeping $\epsilon_1 < 10^{-3}$, which is to allow a fake pulse for every 1000 fits. The $\lambda(E)$ value as a function of energy is shown in Figure 5.10b and is fitted with an exponential function. The result is the energy threshold $T(E) = \lambda(E)$ mentioned in the third criterion.



(a)                                                              (b)
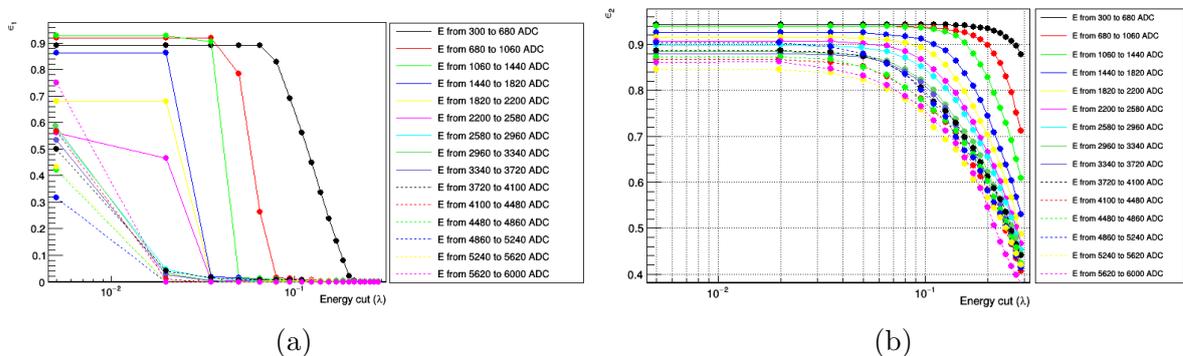
Figure 5.9: Fraction $\epsilon_1$ of traces where a fake pulse is extracted (a) and pulse separation efficiency $\epsilon_2$ of two real pulses (b) as a function of the threshold parameter $\lambda$. The total energy of the two fitted pulses $E_1 + E_2$ is color-coded. Courtesy of L. Cotrozzi.

## 5.2.2   Data-driven tuning

A first version of the fitter is tested on the positron data by applying the first three criteria listed in Section 5.2. To test whether the number of fake pulses matches the expected value of $\epsilon_1 = 10^{-3}$, a sample of clusters at the end of the muon fill is selected. The sample does not have pileup contamination and each positron is sufficiently separated
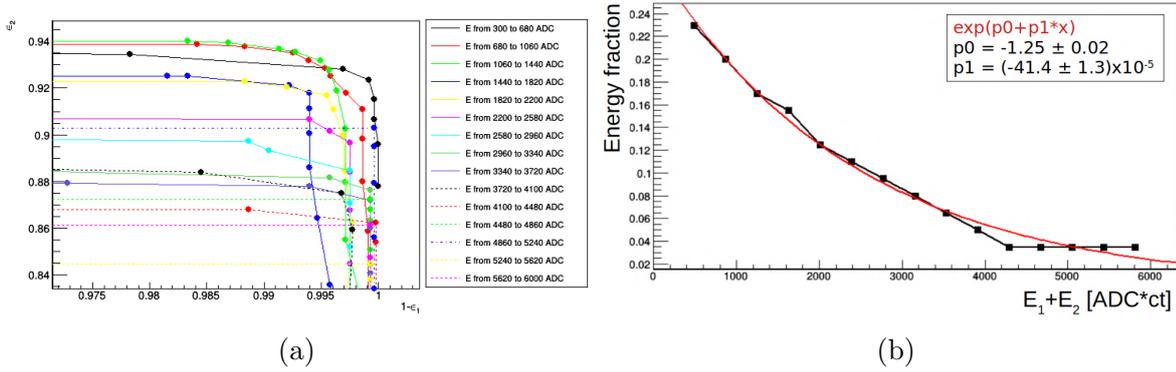
Figure 5.10: (a) ROC curve obtained by comparing $1 - \epsilon_1$ and $\epsilon_2$ as a function of $\lambda$. (b) Best value of $\lambda$ which keeps $\epsilon_1 < 10^{-4}$ as a function of the sum of the two fitted energies $E_1 + E_2$. The exponential fit defines the selection criterion for constrained fits. Courtesy of L. Cotrozzi.

from the next one in the same calorimeter. The number of pulses which get split by the fitter is found to be $\sim 3\%$, of which $\sim 1.2\%$ were already introduced by the ReconWest fitter. All of these are fake pulses, with a typically low energy ($E < 400$ MeV), and an additional cut is set to remove them.

For each time distance $|t_2 - t_1|$ between the two pulses in the same crystal, the 99th percentile of the $E_2/E_1$ distribution is found and shown in Figure 5.11, where $E_1$ and $E_2$ are the energies of the first and second pulses respectively. An exponential fit is then performed to define the new and fourth criterion listed in section 5.2, so that the ReconITA fit is rejected in the cases where the energy ratio is below the fitted exponential function. The same study on real data is repeated with the full set of criteria and the fraction of fake pulses introduced by the ReconITA fitter is reduced to $\sim 0.07\%$.

### 5.2.3   Multi-cluster fitting

The procedure and selection criteria described in the previous sections help separate positrons which are very close in time and produce overlapping peaks in the same SiPM trace. The same procedure is used to find additional pulses in two well separated clusters within the same waveform. The ReconWest algorithm uses different pulse thresholds for the highest peak in a trace and the other peaks found in the residual of the first [50]. This helps reducing fake hits coming from fluctuations in the pulse tail and from periodic noise. Thanks to the clusters information, a double pulse fit can be performed on every trace, constraining both pulse times to be at the two cluster times respectively. An example is given in Figure 5.12. In such cases, only the first two criteria ($\delta t > 0.5$ ct,
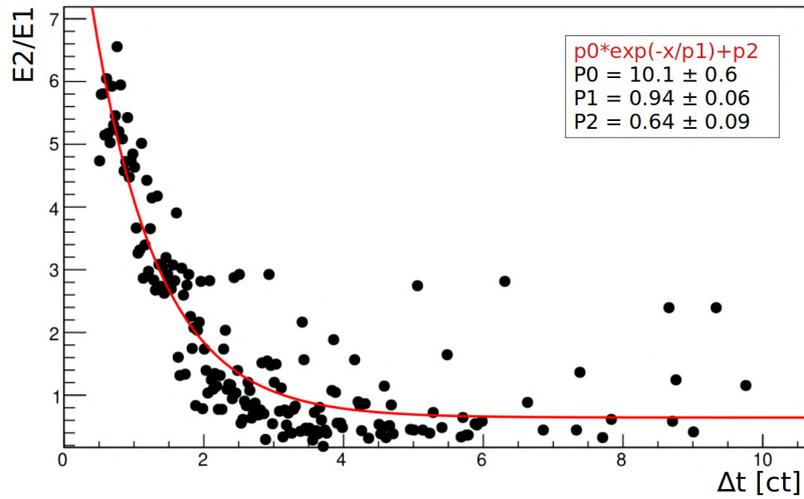
Figure 5.11: Distribution of the 99th percentile of the $E_2/E_1$ ratio versus time separation for the fake pulses. The fitted exponential is used to define the fourth pulse rejection criterion. Courtesy of L. Cotrozzi.

$E > 75$ ADC*ct) are used to select the physically-plausible fit results.
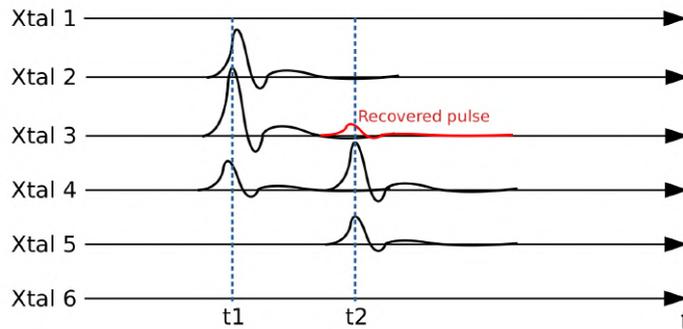


Figure 5.12: Two separated clusters populating the waveforms of some crystals at times $t_1$ and $t_2$. The two times are used to constrain the double pulse fits. The red pulse was too small to be fitted by the ReconWest method (the amplitude is exaggerated in this sketch), but recovered with the ReconITA constrained fit.

The simultaneous fitting of two clusters has proven very effective in recovering hits that would otherwise be hidden by the tails of a preceding pulse. Figure 5.13 shows the distribution of $\delta t$ between consecutive hits in the same crystal. Since two consecutive positrons are independent with respect to each other, the distribution should be flat at short timescales. The ReconWest distribution shows a decrease in the observed hits in the [0,40] ct region corresponding to the length of a typical trace. A sharp peak at $\Delta t \approx 3$ ct is also visible, and is due to a small fraction of fake hits being fitted as part

of the tail of the first cluster. This will be discussed in more detail in Section 5.5. On the other hand, the ReconITA distribution is flatter and proves to recover the missing hits in the [0,40] ct range. A sharp peak at $\Delta t \approx 1$ ct is still visible, and is due to the small residual number of fake pulses generated by the new fitter. It is to be noted that the height of the peak is proportional to the number of total positrons, while the flat baseline is proportional to the much smaller number of double coincidences, so a direct comparison between the two features is not significant.
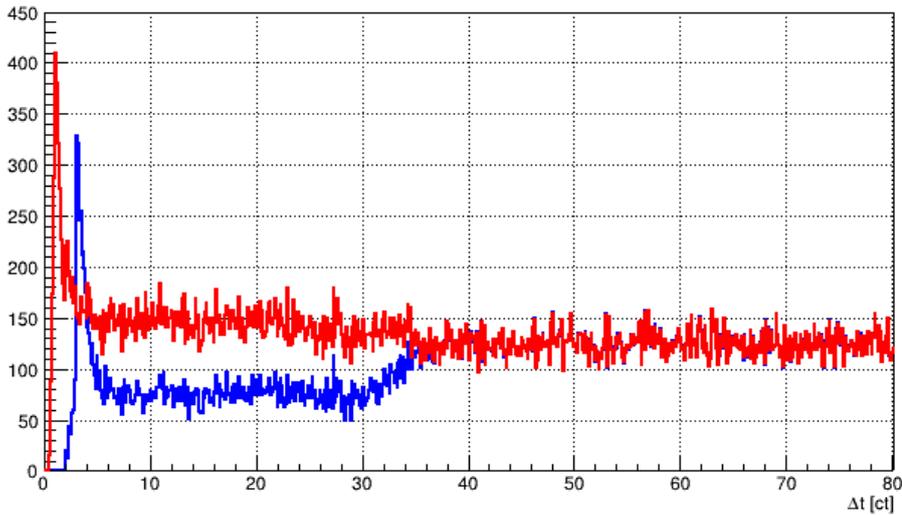


Figure 5.13: Distribution of double hits in the same SiPM, where the first hit is placed at t=0 ct. ReconWest (blue) and ReconITA (red) are compared. The missing hits in the [0-40] ct region, which is the average waveform length, are well recovered by ReconITA.

The hits that are recovered by this technique have low energy and are mostly $E < 100$ MeV, as expected. However, the probability of having a missed pulse because of a preceding positron decreases as half of the muon lifetime, since the probability of having two positrons in the same waveform is proportional to the square of the positron rate $\rho$:

$$\rho(t) \propto e^{-\frac{t}{\tau_\mu}} \qquad \rightarrow \qquad \rho^2(t) \propto e^{-\frac{t}{\tau_\mu/2}} \qquad (5.3)$$

Given the shorter lifetime, the missing energies could produce an early-to-late effect leading to a bias in the measured $\omega_a$. Such effect will be described in more detail in Section 7.3.4, and it is partially fixed by the ReconITA approach as demonstrated in Section 7.6. It has to be noted that a newer version of the ReconWest algorithm addresses this issue as well, but it was not yet analyzed.

### 5.2.4   Pulse fitting software

The pulse fitting software is composed of two main pieces:

- a C++/ROOT standalone library named `ConstrainedFitter` which is responsible of fitting one or more pulses given a trace and the initial guesses and constraints;

- a C++/art producer module named `ClusterFitter` which interfaces with the *art* data and orchestrates the pulse fitting of all the clusters in a single muon fill.

The *art* producer module sorts the clusters in time and by calorimeter, extracts the raw traces, and performs the pulse fits.
Since the ReconITA pulse fitter performs a second fitting round after the first Recon-West estimates, it is not necessary to re-fit every single pulse. For example, at the end of the 700 $\mu$s long fill the rate of positron decays is so low that near-pileup events - which ReconITA targets - are extremely rare. Therefore, most of the clusters are expected to remain untouched. Because of this, the ReconITA module actually recycles most of the ReconWest hits, and replaces the smaller percentage ($\sim 0.1\%$) that actually gets modified.

The crystal hit islands are chopped with a typical length of $\sim$40 ct, and the pulse fit is calculated over this length. However, the pulses produced by the SiPMs last much longer, with an oscillating tail visible, although small, up to 250 ct after the peak [50]. The ReconWest fitter, analyzing one crystal at a time, performs contiguous fits over separated islands when two pulses are found closer than 250 ct. The ReconITA fitter operates over single (or pairs of) clusters at a time, ignoring what happens before and after the event during the fit routine. To remove the tails belonging to preceding hits, the pulse traces of the original ReconWest present before the event that is being currently fitted are subtracted. Since the long tail is present only after the peak, there is no need to subtract the hits present after the fitted event.

The use of a constrained time in the pulse fitter requires that the crystals belonging to the same calorimeter are perfectly synchronized. The time synchronization step is normally applied to the hits after the pulse fitting in the *local approach*. In the Re-conITA software, the time correction is temporarily applied to the traces before the fit is performed.

Finally, after performing the constrained fit, either the original hits or the newly constructed hits get saved in the output according to the selection criteria described in the previous sections.

**Software performance**

The ReconITA software has been included in the official production workflow of the experiment. Being an independent reconstruction program, it inevitably requires additional time, memory, and disk resources. The code has been optimized to be efficient and to run the pulse fitting over the minimum number of traces necessary.

The same reconstruction chain is applied independently over each of the 24 calorimeters. This is a typical process that can easily be optimized with multi-threading. The code has been written with that parallelization, but unfortunately it cannot be fully exploited when running over the FermiGrid or OpenScienceGrid because a single core is allocated to each job. This is because grid *slots* are composed of one core and 2 GB of memory. Running with 24 cores would dramatically decrease the experiment efficiency as 48 GB of memory would be allocated while being unused, and, since most of the workflow cannot be parallelized anyway, 23 of the allocated cores would be rendered useless.

To reduce the processing time, a pre-selection is applied to run the fitter over the bare minimum number of clusters. All the clusters before 25 $\mu$s and after 750 $\mu$s are ignored, as the $\omega_a$ analysis is typically performed in the [30,700] $\mu$s range. Since the decay rate is exponential, the first [0,25] $\mu$s range would contribute significantly to the number of reconstructed positrons.
Then, clusters containing one single hit that are well separated from the others are ignored too, as there is not enough information to perform a significant constrained fit.
Finally, groups of four or more clusters are not processed, as the algorithm is not yet well suited for such complicated configurations. However, the probability of having four or more clusters within a few clock ticks in the same calorimeter is extremely rare and decays with a fast lifetime of $\frac{\gamma \tau_\mu}{4} = 16.1$ $\mu$s.

Figure 5.14 shows the impact of the ReconITA reconstruction chain on the time and memory resources. The slight increase of both has been approved by the offline production team of the Muon $g-2$ Experiment, and the ReconITA reconstruction is part of the official output starting from Run-4 data. Newer improvements in the production

flow actually reduced significantly the memory usage, shifting the peak far from the 2 GB mark, which is the limit of individual grid slots.



(a) Processing time                              (b) Memory usage
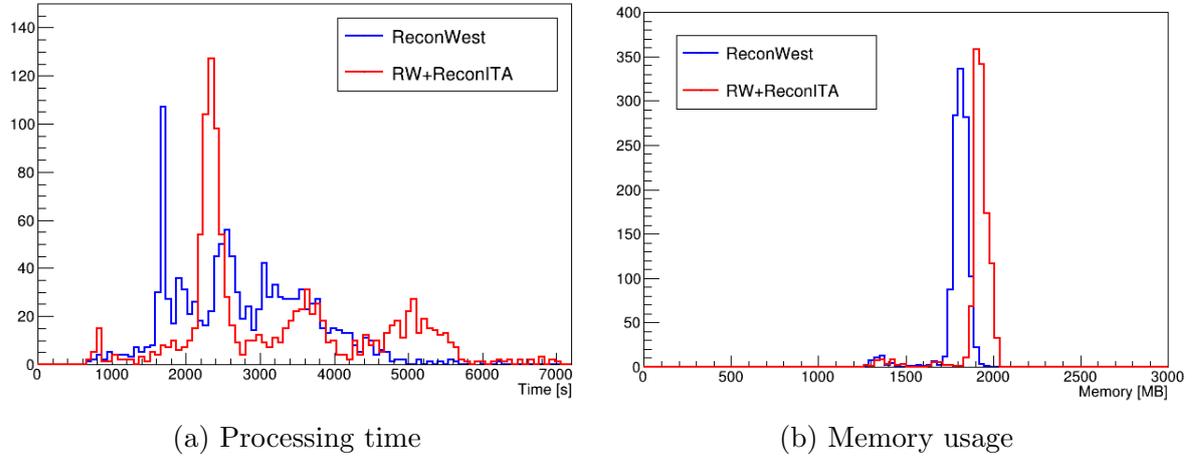
Figure 5.14: Time and memory usage for the calorimeter reconstruction with (red) and without (blue) ReconITA. The large time variability depends on the specific CPU models used by the grid jobs.

The last performance parameter that must be taken care of is the size of the output on disk. Adding a new reconstruction means saving bigger files too. Even though the experiment saves both raw and reconstructed data on tapes with high capacity, the number of files of the Muon $g - 2$ Experiment is in the order of millions (from Run-1 to Run-5 there are $\sim$3.5 million raw files). Therefore, a small increase of few megabytes per file can easily result in hundreds of additional terabytes of needed storage.

Without going too much into the details of how *art* works, it is sufficient to know that every time a calibration or fitting module (*producer*) creates or applies some modification to the crystal hits, a new *dataproduct* containing all the crystal hits is saved in the output file. The various ReconWest fitting and calibration steps that convert the raw traces into final clusters sum up to 315 MB. This is $\sim$20% of the size from calorimeter data. The similar steps for ReconEast have a size of 845 MB ($\sim$54%) and ReconITA finally adds 109 MB ($\sim$7%). The three reconstructions constitute $\sim$57% of the total output file size. The small footprint of ReconITA is achieved by dropping the intermediate calibration dataproducts from the output. Since most of the crystal hits are actually identical to the ReconWest fit results, it would be useless to keep all the redundant data. For this reason, only the fit results and the final hits are saved.

Without the intermediate steps, however, the *art* dataproducts lose the provenance information, which is the link between a dataproduct and its previous version. To provide

for it, I developed a *linker* module which saves the connections between the first and last
dataproducts, as shown in Figure 5.15.
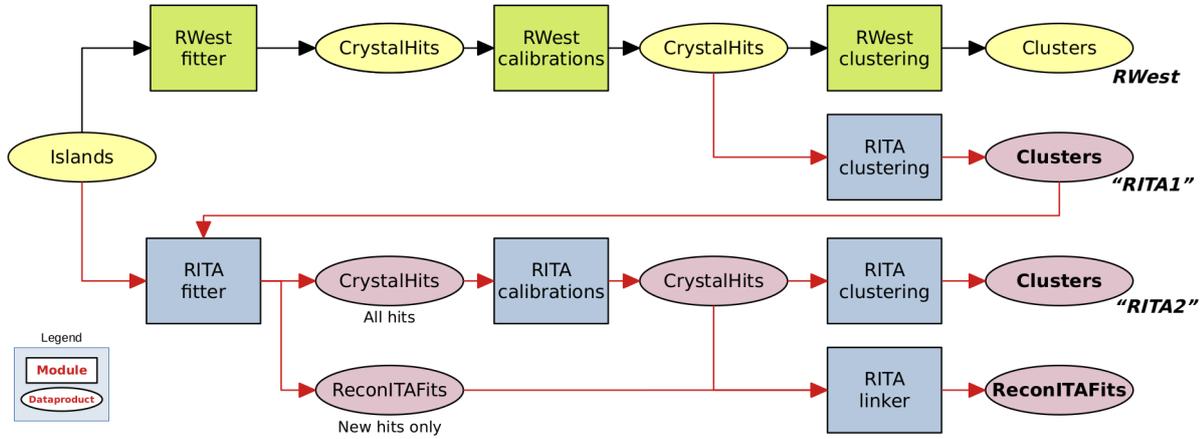


Figure 5.15: Workflow diagram of the calorimeter reconstruction. The top row is the
ReconWest chain, starting with the *islands* containing the raw traces and ending with
the cluster dataproducts. All the other blocks are part of the ReconITA reconstruction,
which produces two versions of clusters (*RITA1*, *RITA2*), both before and after the
complete refitting of the traces.

## 5.3   Time calibrations

A precise measurement of the muon precession frequency, $\omega_a$, requires that all the 1296
SiPMs that compose the calorimeters, and in particular those within the same calorime-
ter, must be precisely synchronized in time. As pulse fitting and clustering algorithms
approach the sub-nanosecond scale, it is necessary to have the hits aligned with greater
precision.

The Laser Calibration System described in Section 3.4 fires a synchronization pulse
to all calorimeters and the T0 detector few microseconds before the injection of every
beam bunch. This pulse is used to correct for the exact moment the individual digitizers
start their clocks with respect to the beam injection, while also setting the absolute time
reference for the beginning of each fill.

Dedicated studies using both muon and positron calorimeter data are performed once
per Run period to extract the relative synchronization between different SiPMs. Syn-
chronization among calorimeters is achieved by observing the signal deposited by lost

muons in two consecutive calorimeters. The synchronization of the crystals in the same calorimeter is achieved using the positron events, assuming that the energy deposition from a single positron is simultaneous in all the crystals.

With these techniques a precision of ∼100 ps can be achieved, being limited by the time spread of electromagnetic showers within a calorimeter. This value is much smaller than the sampling frequency, which is one sample every 1.25 ns, and it is highly reproducible at the level of ∼10 ps among different Runs throughout the data taking period.

The time correction for an individual pulse is then applied as:

$$t_{corrected} = t_{raw} - \delta t \,, \tag{5.4}$$

$$\delta t = t_{sync} + t_{T0} + \delta t_{SiPM} + \delta t_{Calo} \,, \tag{5.5}$$

where $t_{sync}$ is the laser sync pulse, $t_{T0}$ is the beam arrival time as measured by the T0 detector, and $\delta t_{SiPM}$ and $\delta t_{Calo}$ are the time corrections for the SiPMs within a calorimeter and between different calorimeters respectively. Figure 5.16 shows the distribution of $\delta t_{SiPM}$ for a calorimeter.
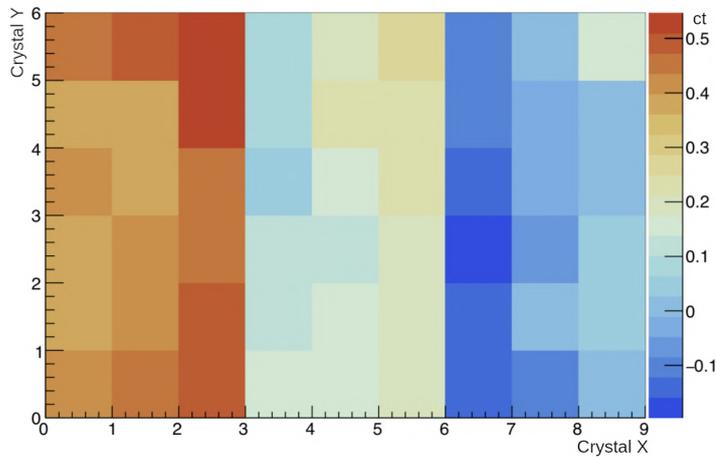


Figure 5.16: Time corrections for the synchronization of the SiPMs belonging to the same calorimeter. The three different groups of columns correspond to the different fiber lengths bringing the laser light to the front face of the calorimeter [69, 70].

## 5.4   Gain calibrations

The energy of the fitted hits is, at this point, uncalibrated. A constant gain response of the detector is crucial for the determination of the $\omega_a$ precession value. Any time-dependent variation in the energy response causes an effective change in the detector acceptance, thus changing the number of positrons entering the histogram across time.

The phase term $\phi$ of Equation 2.20 as measured by the calorimeters depends on the positron energy:

$$N(t, E) = N(E)e^{-t/\tau_\mu} \left[1 + A(E)\cos(\omega_a t + \phi(E))\right] . \tag{5.6}$$

This dependency will be explained in more detail in Section 7.2.2. Therefore, any variation in the energy response of the calorimeters induces a variation in the measured phase:

$$\frac{d\phi}{dt} = \frac{d\phi}{dE}\frac{dE}{dt} \tag{5.7}$$

By writing the phase as a Taylor expansion in time:

$$\phi(t) = \phi_0 + \frac{d\phi}{dt}t + \mathcal{O}(t^2) , \tag{5.8}$$

then any energy variation $\frac{dE}{dt} \neq 0$ leads to a direct bias on the measured $\omega_a^m$:

$$\omega_a t + \phi(t) = \left(\omega_a + \frac{d\phi}{dE}\frac{dE}{dt}\right)t + \phi_0 = \omega_a^m t + \phi_0 . \tag{5.9}$$

The fluctuation of the calorimeter gain was the leading systematic uncertainty in the BNL E821 experiment, with a value of 120 ppb [41]. The new Laser Calibration System built by the INFN-INO group for the new E989 experiment at Fermilab was specifically designed to reduce this systematic uncertainty. I contributed building and operating the system, as well as extracting the gain corrections that will be described in this section for all the data taking periods.

The laser system allows to precisely monitor and correct the calorimeters gain fluctuations. In order to keep the gain-related uncertainty below the design value of 20 ppb[49], a gain equalized at the $\mathcal{O}(10^{-4})$ level in the 700 $\mu$s timescale of a *fill*, and at the $\mathcal{O}(10^{-2})$ level over the entire data taking period must be achieved [52].
Three types of gain correction are extracted for each SiPM using the laser calibration system, and then applied to the reconstructed crystal hits. Each one covers the fluctuations

at different timescales:

- Short-Term Correction: this correction is applied when two or more positrons hit the same crystal in the nanosecond timescale;

- In-Fill Correction: used to equalize the calorimeter response within a fill. It mostly corrects for the effects due to the splash of particles caused by the beam injection;

- Long-Term Correction: temperature variations and environmental changes in general modify the SiPM gain. Although in principle these slow-changing fluctuations do not bias the measured value of $\omega_a$, it is preferable to have a uniform response throughout the entire data taking.

Finally, an absolute calibration of the detectors, while not directly affecting the measured value of $\omega_a$, evens out the response of one calorimeter and enables an equal treatment of all calorimeters. This calibration is extracted from the signals produced by the muons that exit the nominal orbit prematurely and hit the calorimeters before decaying. These are called *lost muons* and produce a MIP-like signal in the calorimeter crystals.

## 5.4.1   Short-Term correction

The gain of a SiPM is a function of the bias over-voltage, $V_b$. When a positron hits a calorimeter and a Čherenkov photon hits a SiPM, a charge avalanche is initiated. A quenching resistor arrests the flow of current by dropping the effective bias below the Geiger threshold. The drop in $V_b$ is proportional to the size of the charge pulse and thus proportional to the SiPM gain [55]. The SiPM behaves like a charged capacitor and shortly after the avalanche has ended the charges are replenished by the frontend electronics. This happens in the scale of $\mathcal{O}(10)$ ns, which is the same timescale at which two positrons must be carefully separated in order to reduce pile-up. As a consequence, a second positron hitting the same crystals within this time range will be reconstructed with a reduced gain, thus biasing the energy measurement (Figure 5.17). The effect is related to the average of number of pixels that fire per positron, which is proportional to the particle energy.

This effect can be precisely measured with the Laser Calibration System by using the so-called Short Term Double Pulse (STDP) technique. The laser is operated in the double-pulse mode (Section 3.4) with two lasers firing toward the same SiPMs with adjustable intensity and time delay. Before the beginning of each long data acquisition
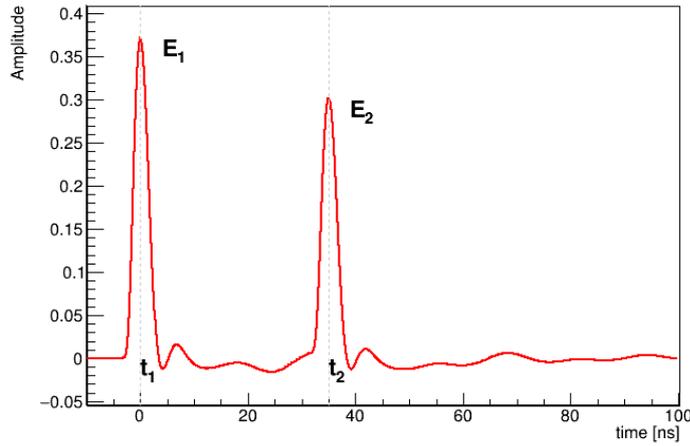
Figure 5.17: Example of the effect of one signal to the next one. The two pulses have the same energy, but the second is measured to be lower because of the reduced SiPM gain. The effect has been artificially amplified in this example for easier visualization.

period, right after the calorimeters are turned on and calibrated, a dedicated laser campaign is performed to extract this STDP correction. The campaign is then repeated in the middle and at the end of the run for consistency checks. Each campaign consists of a series of measurements taken by scanning time delays from 0 to 100 ns and varying the energy of the first pulse by using four different neutral density filters (Filter Wheels). The fully automated sequence is repeated twice: the first time the odd (1,3,5) lasers are pointed, by using movable mirrors, to the calorimeters calibrated by the even (2,4,6) lasers. The odd lasers are used to provide the second (test) pulse. This allows to test half of the calorimeters. Then, the role of odd and even lasers is inverted.

In these dedicated campaigns, double-pulses and single pulses are fired in an alternating sequence (Figure 5.19). When a double-pulse event is present, the first pulse induces the gain drop on the second one, whereas in the single-pulse events the first pulse is absent and the second unaffected pulse is then used as a reference.

The STDP effect on the second pulse is observed as a gain sag with exponential shape (Figure 5.18). The amplitude is directly proportional to the energy of the preceding pulse, while the lifetime remains constant and characteristic of each individual SiPM. The STDP correction is therefore:

$$E_2' = \frac{E_2}{G(\Delta t, E_1)}; \qquad G(\Delta t, E_1; P_1, \tau) = 1 - E_1 \cdot P_1 \cdot e^{-\Delta t/\tau}, \qquad (5.10)$$
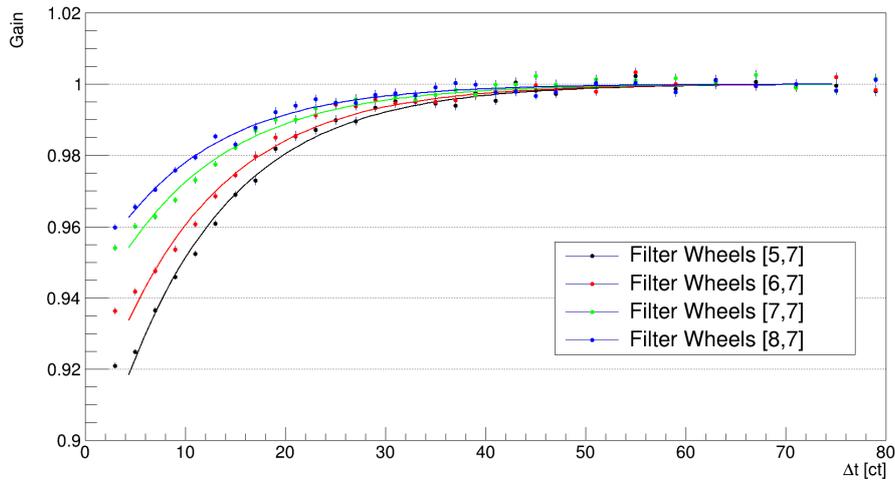
Figure 5.18: Example for one SiPM of the gain sag as a function of time separation $\Delta t$. The four different Filter Wheel combinations are plotted together. A lower value for the first filter wheel (black points) means a higher transmittance, hence a more energetic pulse on the SiPM and a bigger gain sag.
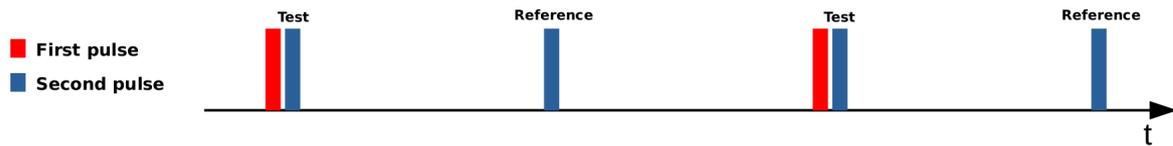


Figure 5.19: Pulsing scheme of the Short Term Double Pulse study.

where $E_1$ is the energy of the first pulse, $P_1$ is a proportionality constant, and $\tau$ is the recovery lifetime. Each energy setting is fitted with an exponential function and the $(P_1, \tau)$ parameters are obtained from linear fits to the four energy configurations (Figure 5.20).

**Temperature dependence**

The STDP effect is related to the bias voltage, which in turn is sensitive to temperature variations. It is then possible that the STDP effect can sensibly change in magnitude between the start and the end of each 8-months acquisition periods. For this reason, a shorter STDP campaign is performed approximately every three days, in coincidence with the trolley field measurements described in Section 2.4. The *Trolley-STDP* campaigns are shorter than the complete STDP campaigns performed to extract the nominal $(P_1, \tau)$ parameters and consist of a single energy configuration. It is, however, sufficient for monitoring long term trends.
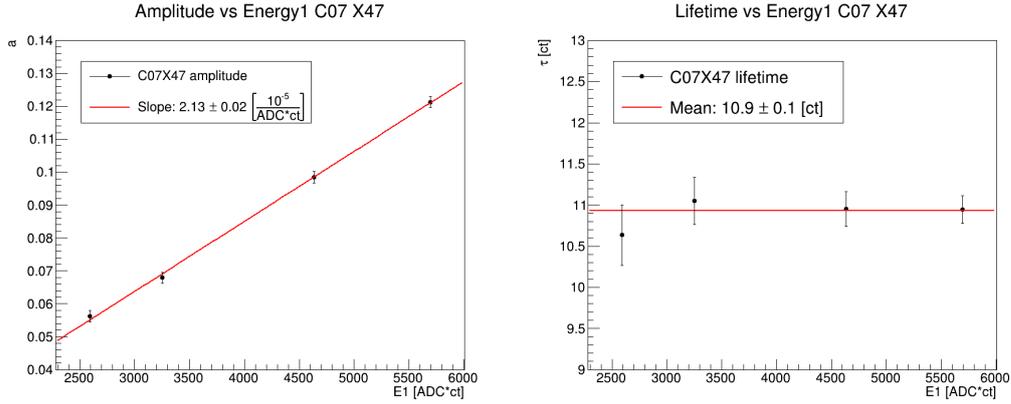
Figure 5.20: Extraction of the $P_1$ (left) and $\tau$ (right) parameters from the four Filter Wheel combinations. The amplitude and lifetime parameters are plotted against the energy $E_1$ of the first pulse in the double-pulse pair.
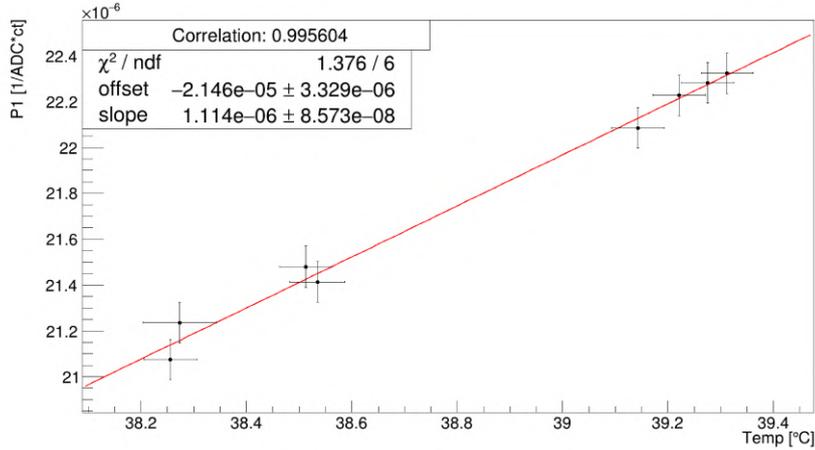


Figure 5.21: Averages of the $P_1$ parameter versus the SiPMs temperature, as measured in 8 *Trolley-STDP* campaigns taken during Run-2. The fitted slope is $\sim 5\ \%/^\circ$C.

The analysis of these periodic campaigns shows that the STDP amplitude scales linearly with temperature. The $P_1$ parameter increases, on average, by $\sim 5\ \%/^\circ$C (Figure 5.21). As a consequence, Equation 5.10 is modified as follows:

$$G(\Delta t, E_1, T; P_1, \tau, \alpha, T_{ref}) = 1 - E_1 \cdot P_1 \cdot (1 + \alpha \left(T - T_{ref}\right)) \cdot e^{-\Delta t/\tau}, \qquad (5.11)$$

where $\alpha$ is the temperature scaling factor, $T$ is the temperature of the calorimeter in a given run, and $T_{ref}$ is the temperature reference value taken during the initial STDP campaign.

The temperature effect has been measured and corrected for the Run-2 period only.
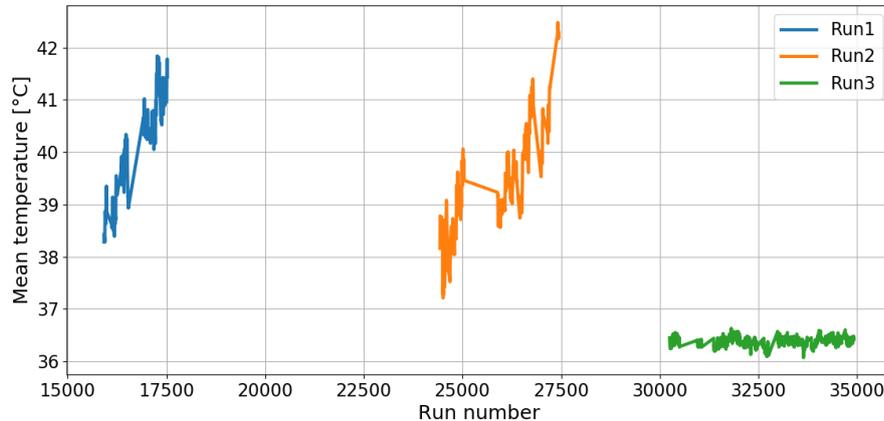
Figure 5.22: Average temperature of the SiPMs in Run-1, Run-2 and Run-3. The high stability from Run-3 is given by a new hall cooling system.

The effect has been discovered after the Run-1 analysis was completed and a systematic error was assigned to the latter. During the summer between Run-2 and Run-3 a new hall cooling system has been installed. From Run-3 onward the calorimeter temperature is stable at the $\pm 0.3$ °C level, so that a dedicated correction is no longer needed (Figure 5.22).

### 5.4.2   In-Fill correction

In the first few microseconds after beam injection, the calorimeters are temporarily "blinded" by the *splash* of particles that accompanies the muon beam. These particles quickly exit the storage region and mostly deposit their energy in the first calorimeters after the injection point. Then, between the start of the measurement period (30 $\mu$s) and the end (700 $\mu$s), the muon decay rate drops by more than four orders of magnitude because of the 64.4 $\mu$s lifetime. These transient effects contribute in generating an *in-fill* time variation of the SiPMs gain.

Correcting for this gain fluctuation is of crucial importance, since any in-fill effect directly biases the $\omega_a$ measurement as explained in section 5.4 and equation 5.9.

To measure the in-fill gain correction, laser pulses are fired during a fill at different times from the injection. In order to minimize the effect of the laser on normal data taking, only four pulses, 185 $\mu$s apart, are shot within the fill. In addition, the laser is fired once every 11 fills. These four pulses are then shifted by 2.5 $\mu$s each following fill

in order to cover the full range from 0 to 800 $\mu$s in 93 steps. The ranges covered by two adjacent pulses slightly overlap. This pattern has been slightly changed during Run-3, decreasing the shift to 1.5 $\mu$s, the distance to 162 $\mu$s and increasing the number of step to 117.

For the analysis of Run-1, Run-2 and Run-3 data, the fills containing laser pulses are not considered for the $\omega_a$ analysis. Although the laser pulses can be easily identified by looking for more than 50 simultaneous hits above 1 GeV in a calorimeter, they affect the detector response to the upcoming positrons. A careful analysis will be carried out for Run-4, Run-5, and Run-6 datasets so that these fills will contribute to the total statistics. The fills containing the laser pulses are therefore now used to correct the ones which do not. The prescale factor of 11 was chosen to be co-prime to the number of bunches in a booster cycle, 16, so that the fills used for the in-fill gain extraction are not systematically related to the same bunch profiles. Starting from Run-2, in-fill pulses are fired for each of the 16 bunches but only once every 11 booster cycles. There is no longer a need for co-prime numbers but the prescale factor remained the same. Starting from the end of Run-3, the prescale factor was increased to 22 as the average dataset was larger than in Run-1 and Run-2 thanks to the increasing stability of the E989 machine. A lower fraction of laser fills was then needed to extract precise In-Fill gain parameters. Finally, starting from Run-5 the laser pattern was changed by increasing the number of laser pulses in the first 50 $\mu$s of the fill, while decreasing the number of pulses later in the fill. This change was motivated by the fact that the gain deviation from 1 is fully contained in the first 50 $\mu$s.

The in-fill gain correction is extracted by looking at the laser energy measured by the SiPMs, equalized using the Local Monitor, in order to correct for shot-by-shot fluctuations of the laser driver. This energy ratio is then compared to the ratio of out-of-fill pulses, in order to correct for long term gain fluctuations. The out-of-fill values are averaged over all the pulses of a subrun, which typically contains 120 fills and $\mathcal{O}(500)$ pulses. The gain function is therefore:

$$g(t) = \left\langle \frac{E_{IF}}{LM_{IF}} \middle/ \left\langle \frac{E_{OOF}}{LM_{OOF}} \right\rangle_{subrun} \right\rangle_t \tag{5.12}$$

where $t$ is the time within a fill. A typical in-fill gain function is shown in Figure 5.23 where the time axis is cut at 300 kct (375 $\mu$s) for aesthetic reasons. The gain function
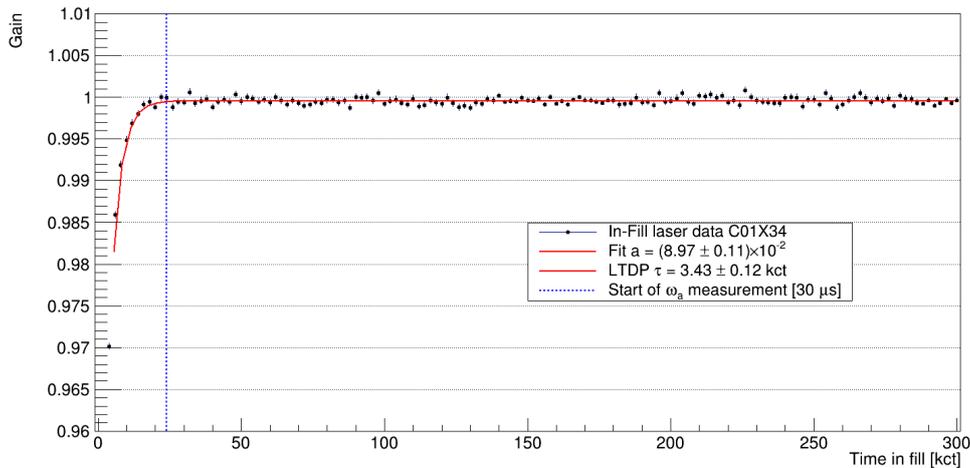
Figure 5.23: Example of an in-fill gain function extracted for Calorimeter 1 SiPM 34. The exponential curve (red) has fixed lifetime parameter from LTDP studies, and fitted for the amplitude parameter. The dashed line represents the start of the $\omega_a$ measurement period, at 30 $\mu$s = 24 kct.

can be fitted with a simple exponential of the form:

$$g(t) = N(1 - ae^{-t/\tau})$$

where $a$ is the amplitude of the sag and $\tau$ is the recovery time of the electronics. The asymptote $N$ is fitted as well since the complexity of the measurements contributing to Eq. 5.12 can bring small deviations from 1. Figure 5.24 shows the distribution of $a$ and $\tau$ values for all the 1296 SiPMs for a Run-3 dataset. The amplitude distribution reflects the different amount of splash particles at injection seen by different calorimeters around the ring. The lifetime distribution depends instead on the electronics of the SiPMs and shows two bands of values centered at $\sim$3.2 kct and 6.5 kct respectively. These bi-modal values can be seen as a specific pattern of the SiPMs in a calorimeter as shown in Figure 5.25 depending on their electronics [69]. The very short average recovery time of the SiPMs drastically reduces the effect of the splash on the measurement range starting at 30 $\mu$s. The average sag goes from $\sim$2.6% at injection to $\sim$0.04% at $t$ = 30 $\mu$s.

**LTDP studies**

The relatively small lifetime value means that few points deviate from 1 in the in-fill gain functions, as visible in Figure 5.23. This problem was mitigated starting from the end of Run-3 with decreased spacing between points and then in Run-5 with a much denser region from 0 to 50 $\mu$s with one point every 0.5 $\mu$s. This, together with the high correlation
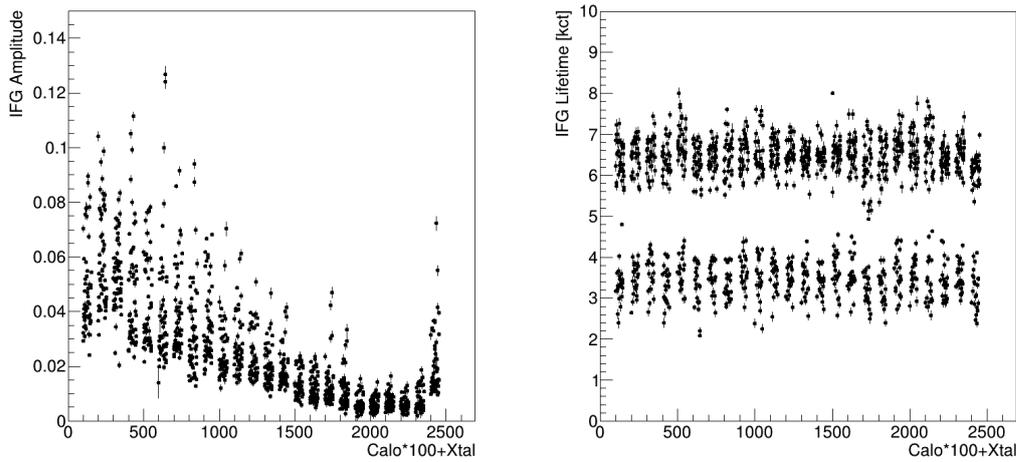
Figure 5.24: Distribution of IFG amplitude and lifetime values for dataset Run-3D. The amplitude is higher for the calorimeters close to the injection point which receive a bigger splash of particles. The lifetime values show a bimodal distribution, due to the different calorimeter electronics for a population of SiPMs.
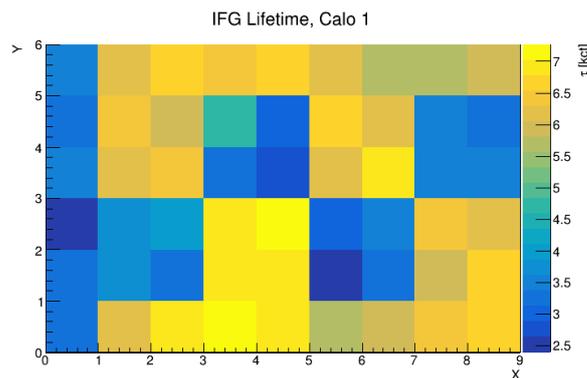


Figure 5.25: View of the bimodal pattern of the IFG recovery time in calorimeter 1.

between the amplitude and lifetime parameters, results in a bias in the measurement of the lifetime parameters which depends on the amount of splash seen by the SiPMs. This effect is visible as "V-shapes" in the lifetime distribution of the SiPMs when shown in sequential order from bottom to top (Fig.5.26). In addition to that, the crystals that receive a small splash have a sag so small that the lifetime cannot be measured with precision.

To solve this problem, starting from Run-2 the in-fill lifetime parameter is determined with dedicated laser studies instead of being fitted freely from the in-fill data. Run-1B and Run-1D datasets are also treated in the same way. These long-term double-pulse (LTDP) laser studies are performed once before each Run period, together with STDP

studies, and are described in Section 3.4. The lifetime values are measured with these studies with a precision of ~2% for all SiPMs as shown in Figure 5.27.
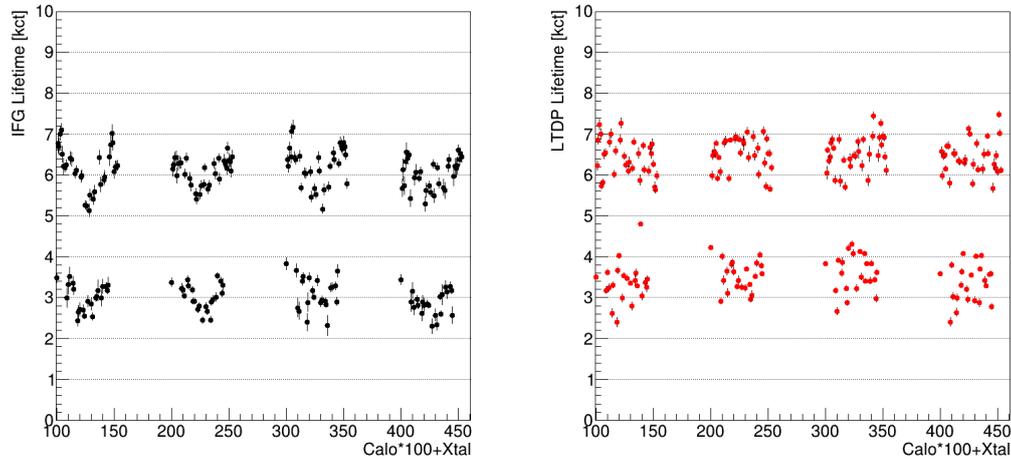


Figure 5.26: Comparison between the in-fill lifetime parameters extracted from a free exponential fit (left) and from the dedicated LTDP studies (right) for the first four calorimeters. The black points are distributed in *V-shapes* according to the different amount of splash that SiPMs receive in a calorimeter.
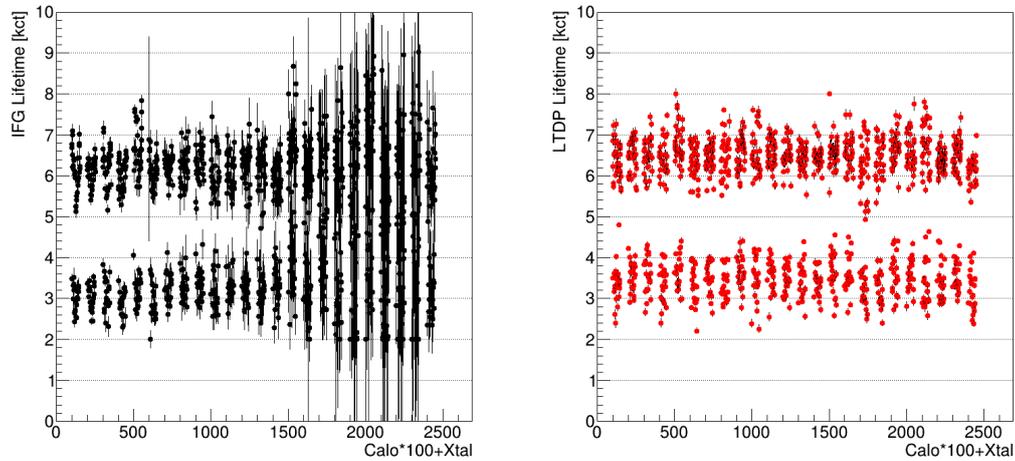


Figure 5.27: Comparison between the in-fill lifetime parameters extracted from a free exponential fit (left) and from the dedicated LTDP studies (right). Calorimeters 15 to 23 receive little to no splash, resulting in a high uncertainty on the free-fit lifetime parameter.

### 5.4.3   Long-Term correction

As described in Section 3.3, the Čherenkov light emitted by positron showers is collected by Silicon PhotoMultipliers (SiPM) because their response is very stable in the presence of strong electric and magnetic fields. However, they are very sensitive to temperature variations. In particular, the gain response changes by roughly $\sim$8 %/$^{\circ}$C as shown in Figure 5.28. While this is not an early-to-late effect and thus does not directly affect the $\omega_a$ measurement, a varying gain can change the effective spatial acceptance of the detectors, other than terribly complicating the combination of results.
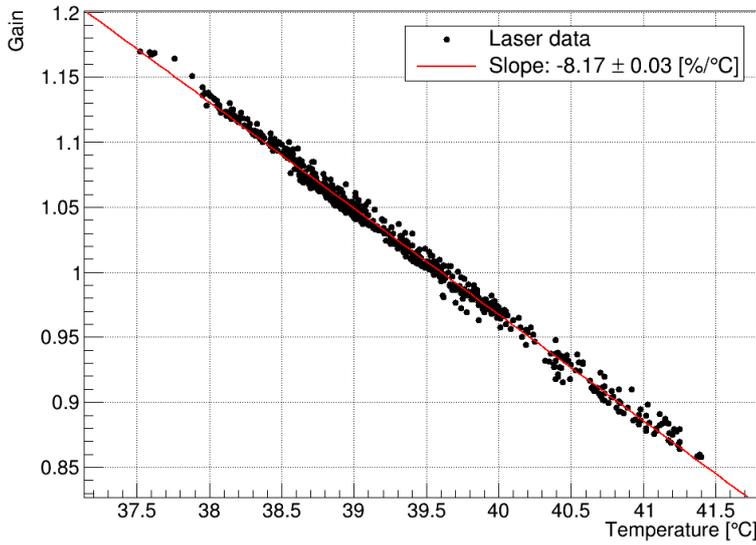


Figure 5.28: SiPM gain versus temperature as measured with the laser pulses.

For these reasons, the laser system fires a series of *out-of-fill* pulses, between each bunch fill and the next. This provides a continuous monitoring of the SiPM performance and allows to build a long-term gain correction. The correction is extracted for each subrun by averaging over the $\mathcal{O}(500)$ pulses as we already saw for the in-fill gain correction:

$$g(subrun) = \left\langle \frac{E_{OOF}}{SM_{OOF}} \right\rangle \tag{5.13}$$

where $E$ is the energy measured by the SiPMs and $SM$ is the laser source monitor which calibrate the shot-by-shot fluctuations of the laser drivers. Figure 5.29 shows the effect of the correction on positron data in Run-2. As the temperature varied by as much as 4 $^{\circ}$C (Fig. 5.22), the average measured energy fluctuated by $\pm$15 %. The long-term correction reduced the gain variation to a stability better than $\pm$0.5 %.
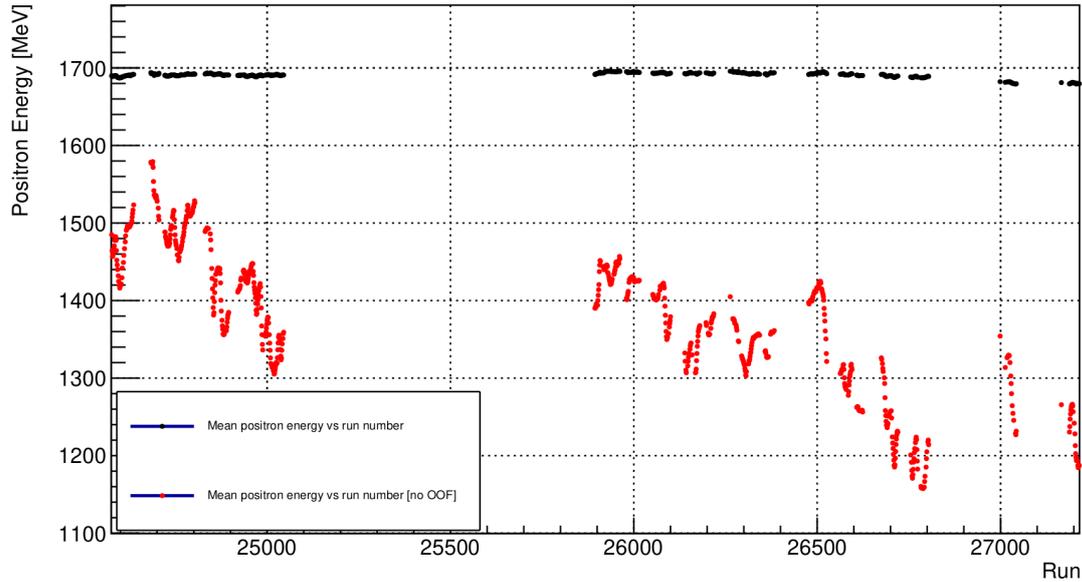
Figure 5.29: Mean energy of the positrons as measured with (black) and without (red) the long-term gain correction. Positrons in the 1000-3100 MeV range are considered for calculating the average. The graph contains the entirety of Run-2 data, which lasted 86 days from March to June 2019.

### 5.4.4   Absolute calibration

The final correction to the energy of the calorimeter hits is the absolute calibration of the energies. While the previous corrections only accounted for variations in gain, the energies are now still in ADC*ct units. Each SiPMs is configured with a different gain setting during the preparation phase of each Run period, according to the individual response to photons. The absolute calibration is needed for both equalizing the response of the SiPMs and for setting an absolute scale.

The equalization is achieved by normalizing the SiPM response to the same physical signal. A feature of the Muon $g-2$ Experiment is the signal from muons that escape the storage region and hit the calorimeters before decaying. While a high storage efficiency and small muon momentum spread is desired, the muons which are lost can provide an excellent calibration tool for the calorimeters. The *lost muons* are MIP particles which typically deposit energy in a single crystal of a calorimeter. Because of the small energy loss, they can even hit multiple calorimeters in their trajectory before being lost definitively. The coincidence of single-crystal hits on adjacent calorimeters with a proper Time-Of-Flight constraint ensure easy selection of such events (Figure 5.30).

The energy deposit of a lost muon in a crystal can be approximated with a Gaussian distribution. By collecting and fitting the distributions for each SiPM, a calibration constant can be obtained for each channel by equalizing the peak position on the energy axis to an arbitrary number.
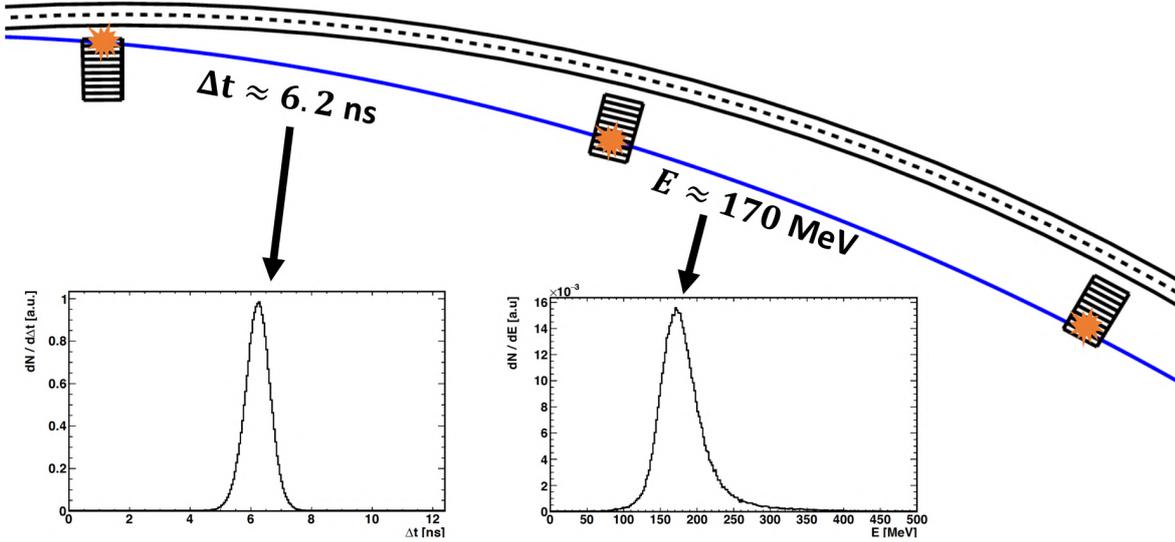


Figure 5.30: A lost muon depositing signals over three consecutive calorimeters. The dashed curve depicts the nominal muon beam orbit. The distributions of the measured Time-Of-Flight and the energy deposit is shown in the insets. Figure from [50].

Now that all channels are equalized in energy response, the absolute scale can be extracted using the positron signals. The positron spectrum endpoint is defined to be 3100 MeV by kinematic considerations. However, a precise measure of the endpoint is made difficult by the pileup contamination and the calorimeter resolution affecting the tail of the distribution. Instead, a simple $\omega_a$ analysis is conducted with a 5-parameter fit over the positron data:

$$N(t) = N \cdot e^{-t/\tau_\mu} \cdot (1 + A\cos(\omega_a t + \phi))  \qquad (5.14)$$

where only the positrons above a certain threshold are considered. By performing a scan over this energy threshold, the absolute scaling can be extracted by finding the threshold value that minimizes the statistical uncertainty of $\omega_a$. The optimal threshold is the one which maximizes the figure of merit $N \cdot A^2$ [68]. The absolute scaling is then set so that the optimal threshold coincides with 1700 MeV. This corresponds to having the MIP energy distribution peaking at 170 MeV. The result of the scan can be seen in Figure 5.31.

The energy deposited by a MIP particle in a 14 cm PbF2 crystal can be estimated using the information from the Particle Data Group [10]:

$$E = \int \langle \frac{dE}{dx} \rangle dx \simeq 1.551 \text{ MeV} \cdot \text{cm}^2 \text{ / g} \times 7.77 \text{ g / cm}^3 \times 14 \text{ cm} = 168.7 \text{ MeV}. \quad (5.15)$$

However, this calculation does not take into account the production of Čherenkov photons, which are what is actually measured by the SiPMs. In addition, the amount of light produced by a positron shower of a given energy might differ from the light produced by a MIP particle depositing the same energy. The close match between the PDG prediction and the actual calibration value is just coincidental.
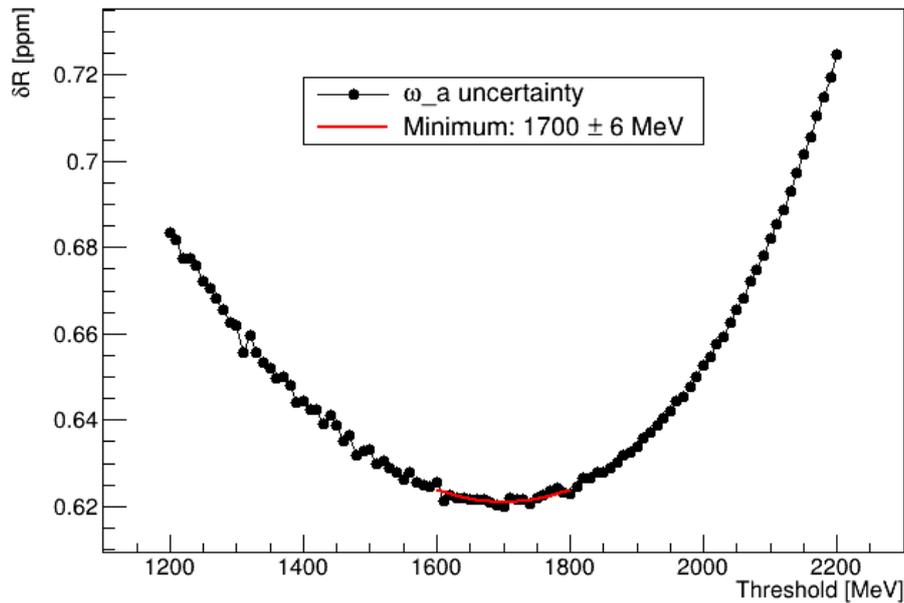


Figure 5.31: Scan of the energy threshold for a five-parameter fit of the positron data. The absolute calibration of the SiPMs is defined such that the threshold which minimizes the error on $\omega_a$ is 1700 MeV.

## 5.5 Clustering

The last step in the reconstruction chain (Fig. 5.3) is the clustering of individual hits into positron candidates. This step collects the fully calibrated crystal hits and assembles them into clusters, each one representing a positron. Ideally, the algorithm is always able to distinguish between different positrons and to assign the correct hits. However, when two positrons hit the same calorimeter in the same region at the same time, separation

is not always possible and the clustering algorithm creates the so-called pileup events.

The clustering algorithm developed for the ReconWest reconstruction, and used for the $\omega_a$-*Europe* analysis of Run-1 data, is based on the time partitioning of the hits. First, the hits in a calorimeter are listed in a time-ordered manner. Then, starting from the beginning of this list, the hits are accumulated into clusters. For each group of hits, a *running* time is computed as the energy-weighted average:

$$t_{running} = \frac{\sum_i E_i t_i}{\sum_i E_i} \, , \tag{5.16}$$

where the sum is computed over the hits belonging to the cluster being currently formed. Each new hit is then either added to the current cluster, or it initiates a new cluster as follows. By calculating the time difference between the new hit and the running time, $\Delta t = t_{i+1} - t_{running}$, the following criteria based on two parameters $t_{low}$ and $t_{high}$ are applied:

- $\Delta t < t_{low}$: the new hit is always included in the current cluster and the new running time is computed;

- $\Delta t > t_{high}$: the current cluster is complete and a new cluster is created starting with the new hit;

- $t_{low} < \Delta t < t_{high}$: the time difference between the new hit and the next one is computed $\Delta t_{next} = t_{i+2} - t_{i+1}$, and the following criteria are applied:

  a) $\Delta t_{next} > \Delta t$: the new hit is included in the current cluster;

  b) $\Delta t_{next} < \Delta t$: a new cluster is created.

The parameters values $(t_{low}, t_{high}) = (3, 5)$ ct were chosen for the reconstruction of Run-1 data. Since Run-2, they were reduced to $(t_{low}, t_{high}) = (2, 3)$ ct for increasing the pileup separation efficiency. This algorithm will be referred to as Time Partitioning (TP).

As anticipated in Section 5.1.4, one of the largest sources of systematic uncertainties in the measurement of $\omega_a$ is the presence of pileup. To increase the efficiency of pileup separation and consequently reduce the related systematics, I have developed a new clustering algorithm for the ReconITA reconstruction. The same algorithm is applied both at the end of the ReconWest reconstruction chain, and at the end of the ReconITA chain, as shown schematically in Figure 5.3. The output positrons from the

first iteration will be referred to as ReconITA1 (RITA1), whereas the final output will be called ReconITA2 (RITA2). The analysis of Run-2 and Run-3 conducted by the $\omega_a$-*Europe* group is based on the RITA1 clusters. The entire ReconITA reconstruction chain was included in the official production of the E989 experiment starting from Run-4, and the analysis of Run-4, Run-5, and Run-6 data will be based on the RITA2 clusters.

A new clustering approach was also developed by the University of Washington group as an improved variation of the ReconWest time partitioning clustering described above [71]. This was developed roughly in the same period of the development of the work discussed in this dissertation. For this reason, the next sections will compare the ReconITA approach to the standard ReconWest clustering.

## 5.5.1 Space partitioning

The E989 electromagnetic calorimeters are composed of 54 $PbF_2$ crystals arranged in a $6 \times 9$ matrix. Each crystal has a square cross section of side $\sim$2.5 cm. The $PbF_2$ material has a Molière radius of $\sim$2.1 cm [10], so the energy deposition of an electromagnetic shower generated by a positron should be 95% contained within a matrix of $3 \times 3$ adjacent crystals, with the central one having the most measured energy. The segmentation of the calorimeters should enable spatial separation of few positrons hitting different sections of the calorimeter simultaneously.

To exploit this feature, a space partitioning algorithm was developed. First, the calorimeter hits are subdivided into groups which are sufficiently separated from each other. To do this, the standard ReconWest time partitioning algorithm is applied with $(t_{low}, t_{high}) = (3, 5)$ ct. Then, for each group of hits belonging to a temporary cluster, the following seed-and-propagation algorithm is applied:

1. Find the most energetic hit in the group, the *seed*, and start creating a cluster;

2. Include all the hits in the adjacent crystals with energy $E > E_{th}$;

3. Continue including all the hits that are adjacent to one or more of the previously selected ones;

4. When no more adjacent hits can be found, find the most energetic hit in the remaining hits (new seed);

5. Repeat steps 1-4 until all the hits with $E > E_{th}$ have been included in clusters;

6. Finally, assign all the remaining hits with $E < E_{th}$ to the clusters having the closest seed.

The energy threshold $E_{th}$ is introduces so that low energy hits, which can be further away from the seed because of the geometry of the electromagnetic shower, do not act like "bridges" by connecting two independent positrons together. The value of $E_{th} = 100$ MeV has been chosen. An example of the algorithm is sketched in Figure 5.32.
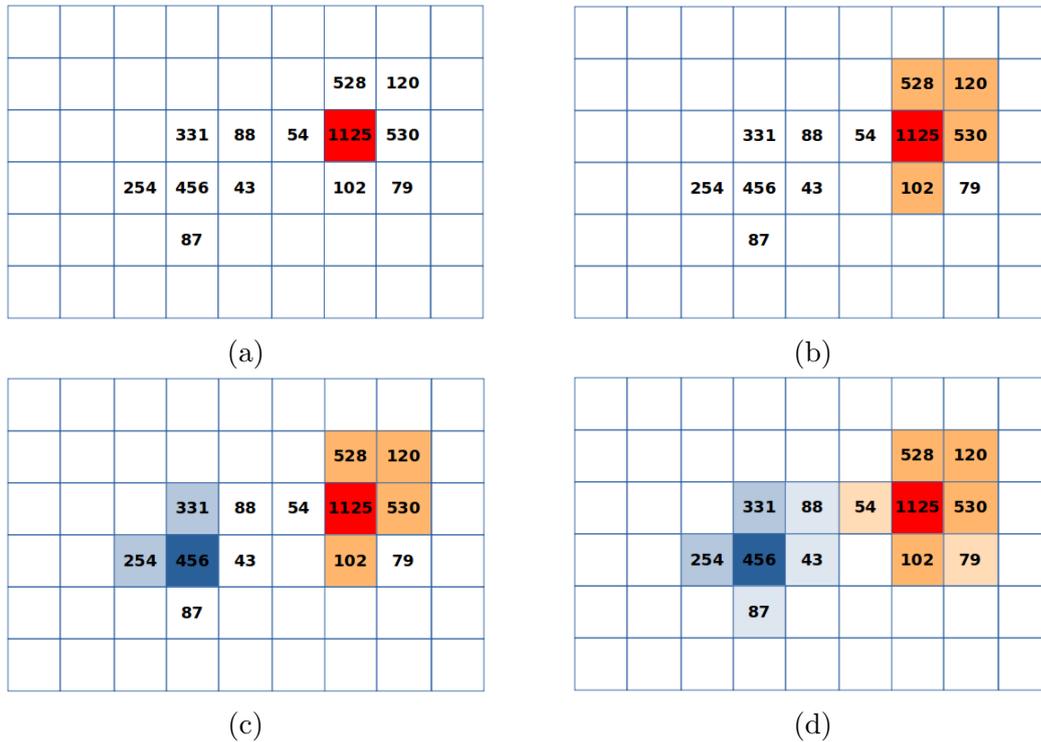


Figure 5.32: Example of the space partitioning algorithm applied to successfully separate two positron events. The $6 \times 9$ grids represent the calorimeters and the numbers are the deposited energies in MeV for each crystal. (a) The most energetic hit (seed) is located. (b) The adjacent hits are associated with the seed. (c) A new seed is located and propagated. (d) Finally, all the hits with $E < 100$ MeV are associated with the closest seeds.

## False pileup

A way to measure the algorithm efficiency of separating pileup events is to look at the distribution of the time distance between two consecutive clusters in a calorimeter. In fact, for very short timescales, smaller than the cyclotron period, the rate of muon decays is constant. By setting $t_0$ the time of a cluster, and $t_1$ the time of the next cluster in the same calorimeter, the distribution $\Delta t = t_1 - t_0$ is expected to be roughly flat up to

$\mathcal{O}(100)$ ns. Any clustering inefficiency in separating pileup events would be visible as a decrease in the distribution for small $\Delta t$. On the other hand, if the clustering algorithm ends up splitting a single positron into two or more clusters, an increase for small $\Delta t$ is expected.

Figure 5.33 shows a comparison between the Run-1 ReconWest time partitioning and the space partitioning algorithm described above. The ReconWest clustering shows zero separation efficiency between 0 and 3 ct as expected. The ReconITA space partitioning clustering shows a discrete improvement from 1 to 3 ct. However, a significant peak at $\Delta t \to 0$ and roughly 15 times higher than the baseline is introduced, indicating that individual positrons are split by mistake.

The source of this peak can be explained by a rare population of positrons that deposit fractions of their energy in multiple parts of the calorimeters. This is motivated by looking at the distribution of the sum of the energies, $E_0 + E_1$, of the two consecutive clusters. Figure 5.34 shows the distributions extracted for six slices of $\Delta t$ from $[0, 1]$ ct to $[5, 6]$ ct. The slices from $\Delta t > 1$ ct show the characteristic energy spectrum of pileup events, going from 0 to 6000 MeV. The first slice, corresponding to the peak of Figure 5.33, shows a distribution typical of single positrons, on top of a double positron spectrum, indicating that it is dominated by individual positrons. Moreover, these events are distributed throughout the entire beam storage time with the typical muon lifetime $\gamma \tau_\mu \approx 64.4$ $\mu$s. These events are referred to as *false pileup* and account for $\sim 1\%$ of all positrons. The physical explanation for this effect is not completely understood, but a possible reason is that sometimes the electromagnetic shower is initiated in the material between the storage region and the calorimeter crystals. This is motivated by the distribution of the false pileup events, selected with $\Delta t < 1$ ct and $E_1 + E_2 < 3100$ MeV, among the calorimeters, shown in Figure 5.35a. The number of false pileup events is significantly lower for calorimeters 1, 17, 18, 23, and 24, all associated by the absence of any quadrupole or kicker plate in front of them (Figure 3.2). The spatial distribution of the two energy depositions composing a false pileup is also different for the two calorimeters placed after the tracking stations (Figure 5.35b), indicating further interaction with the tracker material.

Since the majority of events separated by the space partitioning are actually false pileup events, this algorithm has been discarded.
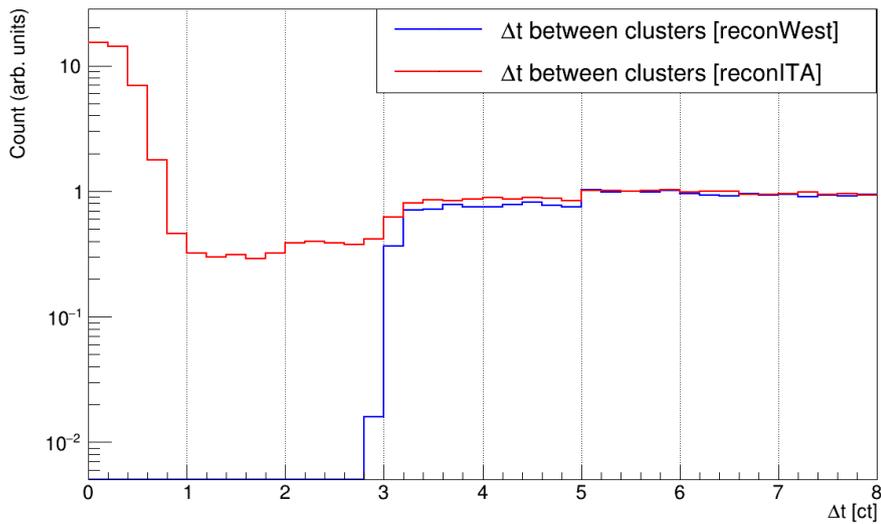
Figure 5.33: Distribution of time separation $\Delta t$ between two consecutive clusters in the same calorimeter. The ReconWest time partitioning (blue) and the ReconITA space partitioning (red) are compared. Both distributions are normalized to the value at $\Delta t = 8$ ct.
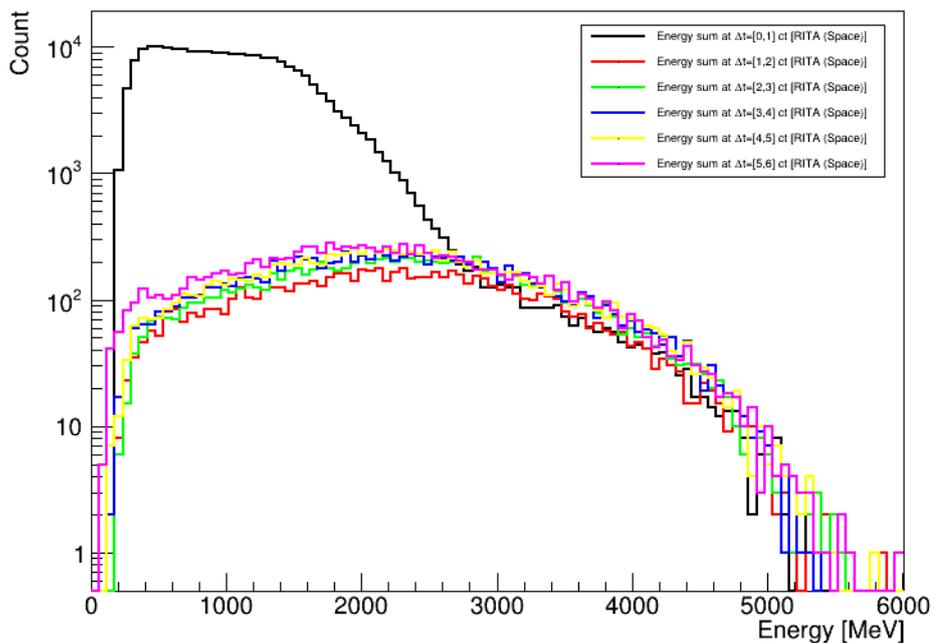


Figure 5.34: Distribution of the sum of energies $E_1 + E_2$ of the two clusters split by the space partitioning algorithm. Each time difference slice is assigned to a different color. The black line represents clusters with $\Delta t < 1$ ct which are dominated by the false pileup events.
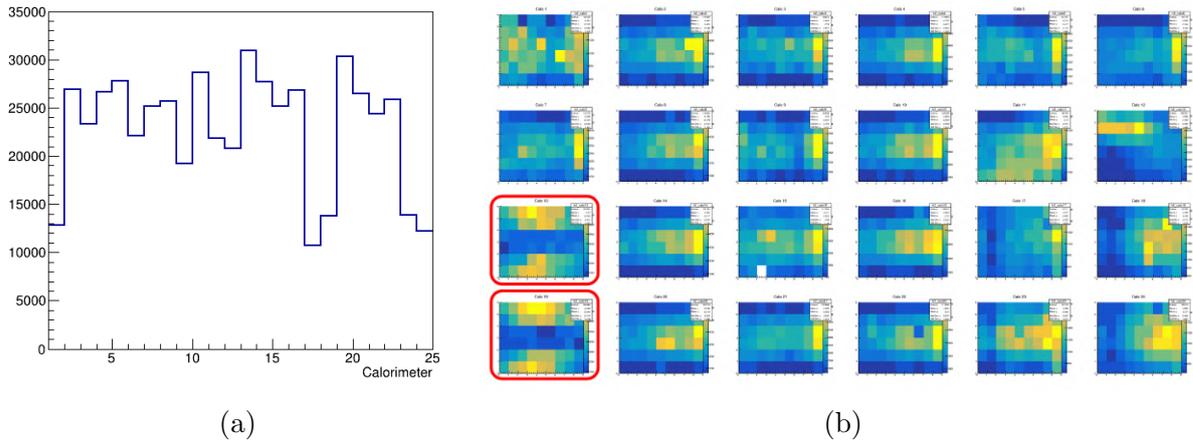
Figure 5.35: (a) Number of false pileup events versus calorimeter. Calorimeters 1, 17, 18, 23, and 24 have a significantly smaller contamination. (b) Distribution of false pileup events on the calorimeter crystals. The calorimeters 13 and 19, which are placed after a tracking station, show a different pattern and are circled in red.

## 5.5.2 Time propagation algorithm

While the spatial information cannot be used to increase the pileup separation efficiency, improvements can still be made by using the time information.

The seed-and-propagation algorithm has been modified to act using the time information of the hits, with slight modifications:

- The hits are included in the cluster initiated by the seed according to a time-based criterion;

- The low energy hits ($E < 100$ MeV) are associated at the end, but only if they meet the same time-based criterion, otherwise they create new independent clusters;

- If the seed crystal has a second hit which did not get associated to any other cluster, then it gets added to the seed's cluster.

The reason behind the third criterion will be discussed later.

**Time-based criterion**

The request of propagating the seed and include the "nearby" hits is based on the analysis of a clean sample of single positron clusters, extracted from the end of the fill when the probability of pileup is very small due to the exponential form of muon decay. In particular, the time distribution of the hits with respect to the cluster seed (the

(a)                                                    (b)
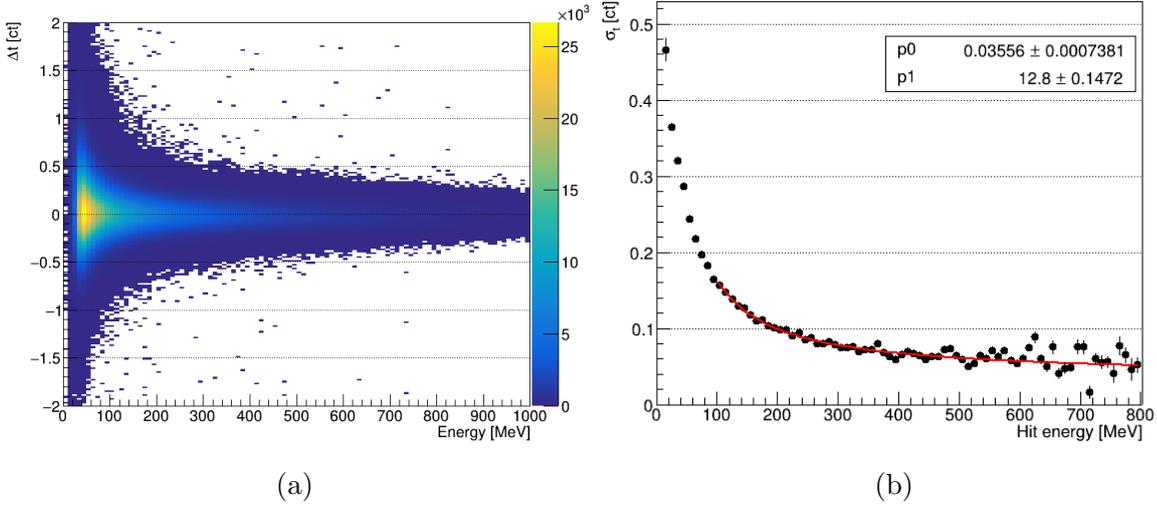
Figure 5.36: (a) Time distribution of hits with respect to the seed versus hit energy. (b) Variance of the distribution in (a) versus energy. The graph is fitted with $p_0 + p_1/x$ (red line).

most energetic one) is found to depend on their energy, as shown in Figure 5.36a. For each energy bin, the variance of the time distribution is extracted as shown in Figure 5.36b. An empirical function of the form $p_0 + p_1/x$ is fitted to the data, and the best fit parameter values are found to be:

$$\sigma(E) = 0.036 + \frac{12.8}{E} \, . \tag{5.17}$$

This information is used to build the criterion for the propagation algorithm. For two generic hits of energies $E_i$, $E_j$, their variances are added in quadrature in order to determine the expected deviation between them:

$$\sigma(E_i, E_j) = \sqrt{\sigma^2(E_i) + \sigma^2(E_j)} \tag{5.18}$$

To determine whether two hits $i$ and $j$ should be connected in the same cluster, the following criterion is used:

$$\Delta t_{ij} < k\sigma(E_i, E_j) \tag{5.19}$$

where $k$ is a parameter that can be tuned according to the algorithm performance. Several values for $k$ has been tested, and the best value is the one that maximizes the pileup separation while minimizing the splitting of single positrons. Figure 5.37 shows the positron spectrum for the ReconWest time partitioning algorithm compared to ReconITA with the parameter $k$ ranging between 4.0 and 7.0. The pileup tail is small

for all the values of $k$, with the value of $k = 4.0$ being the most performing one. To test whether the algorithm splits single positrons, the time distribution of the ReconWest clusters that get split by the ReconITA algorithm is extracted and fitted with a double exponential function of the form:

$$N(t) = A_1 e^{-t/\tau_\mu} + A_2 e^{-2t/\tau_\mu} \, , \tag{5.20}$$

where $A_1$ measures the contribution due to single positrons (distributed as $\tau_\mu = 64.4 \ \mu\text{s}$), and $A_2$ measures the contribution due to pileup events that get successfully separated (distributed as $\tau_\mu/2 = 32.2 \ \mu\text{s}$). The time distributions are shown in Figure 5.38a, while the fitted $A_1$ and $A_2$ amplitudes are plotted in Figure 5.38b as a function of $k$. A reasonable value that maximizes $A_2$ while keeping $A_1$ small is found for $k = 5.5$.
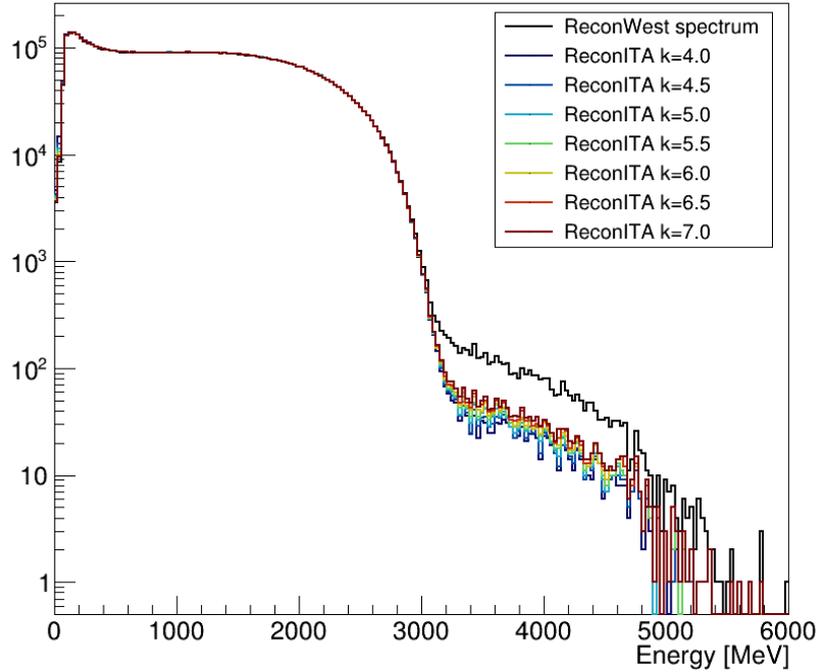


Figure 5.37: Energy distribution of the clusters for the ReconWest time partitioning clustering (black line) and the ReconITA time propagation algorithm for some values of $k$.

## Distance-based parameters

With the choice of $k = 5.5$, the time criterion becomes:

$$\Delta t_{ij} < 5.5\sigma(E_i, E_j) \, . \tag{5.21}$$

(a)                                                          (b)
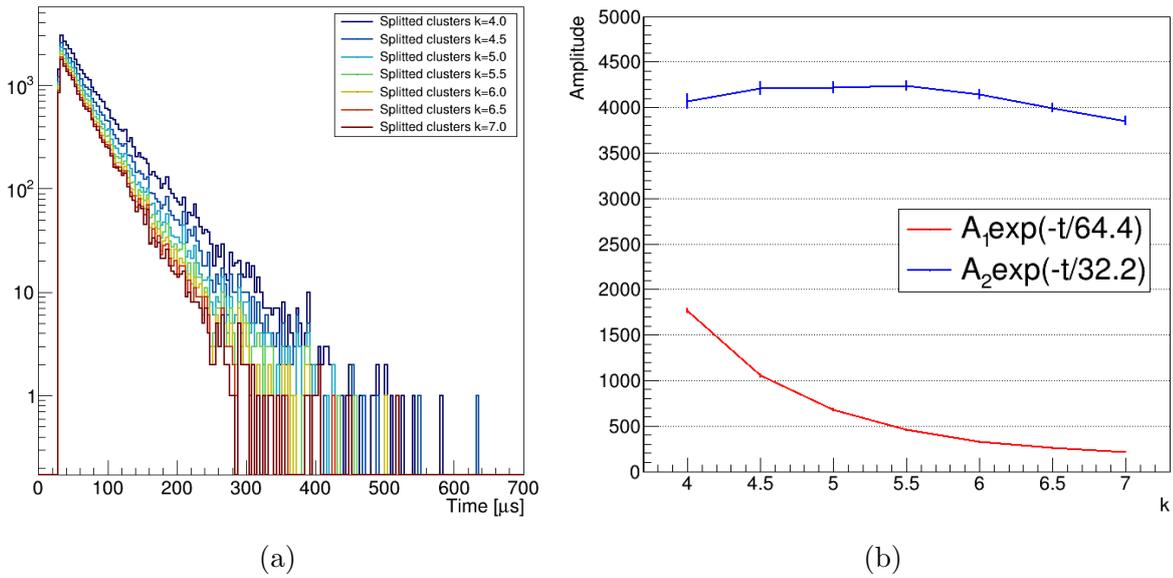
Figure 5.38: (a) Time distribution of the ReconWest clusters that get split into two by the ReconITA algorithm for some values of $k$. (b) Each distribution of Figure (a) is fitted with a double exponential (Equation 5.20). The two amplitudes $A_1$ (red) and $A_2$ (blue) are plotted as a function of $k$.

This equation, however, is not sufficient to make sure that the false pileup events described in Section 5.5.1 are always assembled into unique clusters instead of being split into two parts. Figure 5.39 shows the distribution of hits in a cluster as a function of time separation and space distance from the cluster seed. A selection of clusters late in the fill and with energy $E < 3100$ MeV ensures that no pileup contamination is present in the distribution. Two interesting features are visible:

- A small population of hits ($\sim 1\%$) is located far from the seed, from 3 to 8 crystals away, with the higher distance being the width of the entire calorimeter. The distribution of these hits is also projected in Figure 5.40a, and the time variance is shown in Figure 5.40b.

- The crystal of the cluster seed contains some additional hits at $\Delta t > 2$ ct (distance zero from seed), and fewer at $\Delta t < -2ct$.

The first feature is related to the presence of false pileup positrons, and an additional factor based on distance is needed to include them in the same cluster while applying the propagation algorithm. A term proportional to the distance is added to the time criterion and Equation 5.19 becomes:

$$\Delta t_{ij} < 5.5\sigma(E_i, E_j) + 0.03\Delta r_{ij}, \tag{5.22}$$

where the 0.03 factor has been chosen such that the inclusion factor is higher than the time variance for the distance $\Delta r = 8$ crystals.

The second feature indicates that sometimes the ReconWest pulse fitter finds an additional pulse located shortly after the main peak. The minimum time distance $|\Delta t| = 2$ ct at which these extra pulses are found corresponds to the Artificial Dead Time of the fitter. The distribution of the time separation between two consecutive hits in the same crystal, shown in Figure 5.41a, confirms that, while no pulses are found within 2 ct from a pulse, there is an excess of pulses from $\Delta t > 2$ ct up to $\Delta t \sim 8$ ct. The energy deposited by a positron in a crystal is measured as the integral of the pulse shape registered by the SiPM. In principle, the number of pulses found in the same template fit does not bias the energy measurement of the event, provided that the sum of the pulses equals the waveform integral. However, if these pulses are sufficiently distant to be separated into two different clusters by the time clustering algorithm, then the energy measurement is affected. However, it has to be noted that this is not a rate-dependent effect and should have little impact on the measurement of $\omega_a$.

To make sure that the clustering algorithm does not separate these rare extra pulses, any hit that is positioned in the same crystal of the seed within 8 ct from the latter is re-absorbed in the cluster unless it was already part of another cluster. This is the last criterion listed at the beginning of this section. Figure 5.41b shows the time separation of two consecutive hits in the same crystal when they belong to two different clusters. The ReconWest clustering shows the excess peak from $\sim 3$ ct as the extra pulses are not re-absorbed in the original cluster but create a new cluster with only one hit. The ReconITA clustering effectively treats these hits and the distribution appears flat as expected.

**Summary**

The goal of the new ReconITA clustering algorithm is to increase the pileup separation efficiency while minimizing the probability of splitting individual positrons into multiple clusters. Figure 5.42a shows the time separation between two consecutive clusters in the same calorimeter, comparing the performance of the ReconWest time partitioning and the ReconITA time propagation algorithms. The ReconITA approach manages to separate clusters at $\sim 100\%$ efficiency up to $\Delta t \approx 1.5$ ct, and then with a lower efficiency up to $\Delta t \approx 0.5$ ct. The distribution is also flatter above $\Delta t > 2$ ct indicating that the
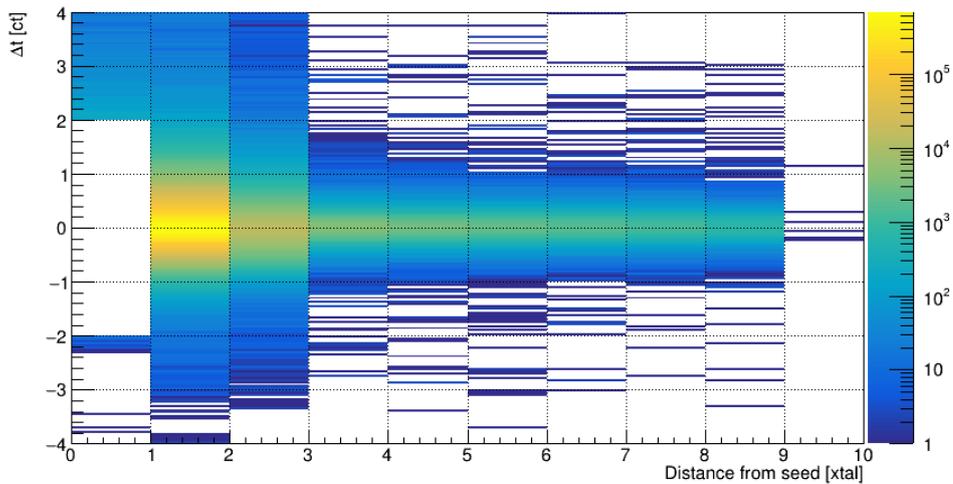
Figure 5.39: Distribution of hits in a cluster with respect to the cluster seed. On the x-axis the spatial distance in units of crystals ($\approx 2.5$ cm) is given, while on the y-axis the time difference is calculated. The first bin on the left is populated by the hits which are registered in the same crystal of the seed.

unwanted splitting of single positrons is minimized. This results in an overall reduction of pileup contamination of $\approx 50\%$. This can also be measured in the tail $E > 3100$ MeV of the cluster energy spectrum shown in Figure 5.42b as the tail above that value consists of pileup events solely.

This clustering algorithm is applied both at the end of the ReconWest reconstruction chain and at the end of the ReconITA chain. The performance between the two steps is quite similar, as shown in Figure 5.43a. Figure 5.43b shows the same time separation between two consecutive clusters but with the constraint for the second cluster to have two or more hits. The ReconWest peak after 3 ct is no longer visible, confirming that it was dominated by individual hits separated from the original cluster. The second ReconITA clustering (ReconITA2) shows a little excess between $\Delta t \approx 1.5$ ct and $\Delta t \approx 3$, whose effect is considered to be negligible on the $\omega_a$ measurement.

(a)                                                          (b)
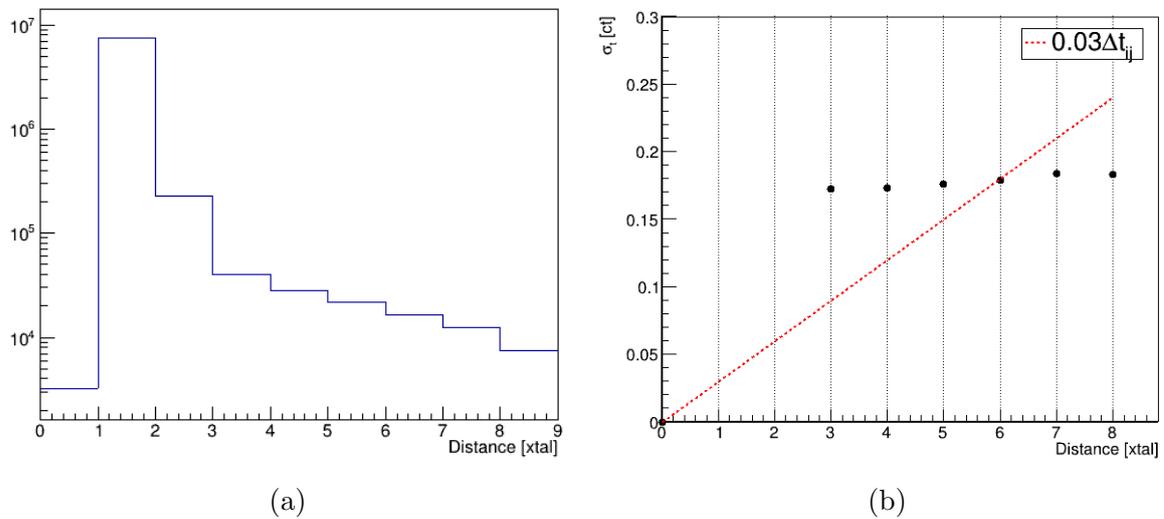
Figure 5.40: (a) Spatial distance distribution of the cluster hits from the most energetic one (seed). (b) Time separation variance as a function of the distance from the seed. The red dotted line is the additional term included in the criterion of Equation 5.22.



(a)                                                          (b)
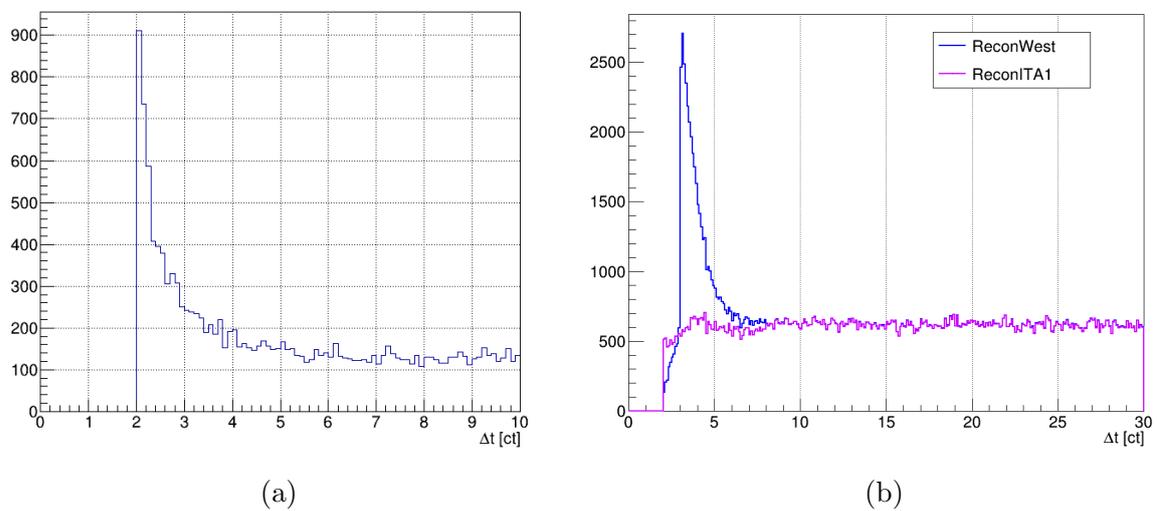
Figure 5.41: (a) Time separation between two consecutive hits in the same crystal. No clustering algorithm has been applied. The baseline of $\sim 100$ events can be attributed to the uniform distribution given by a second positron hitting the same crystal. No pulses are found between 0 and 2 ct because of the fitter dead time. (b) Time separation between two consecutive hits in the same crystal pertaining to two consecutive clusters. The ReconITA clustering algorithm shows a flatter distribution as expected.
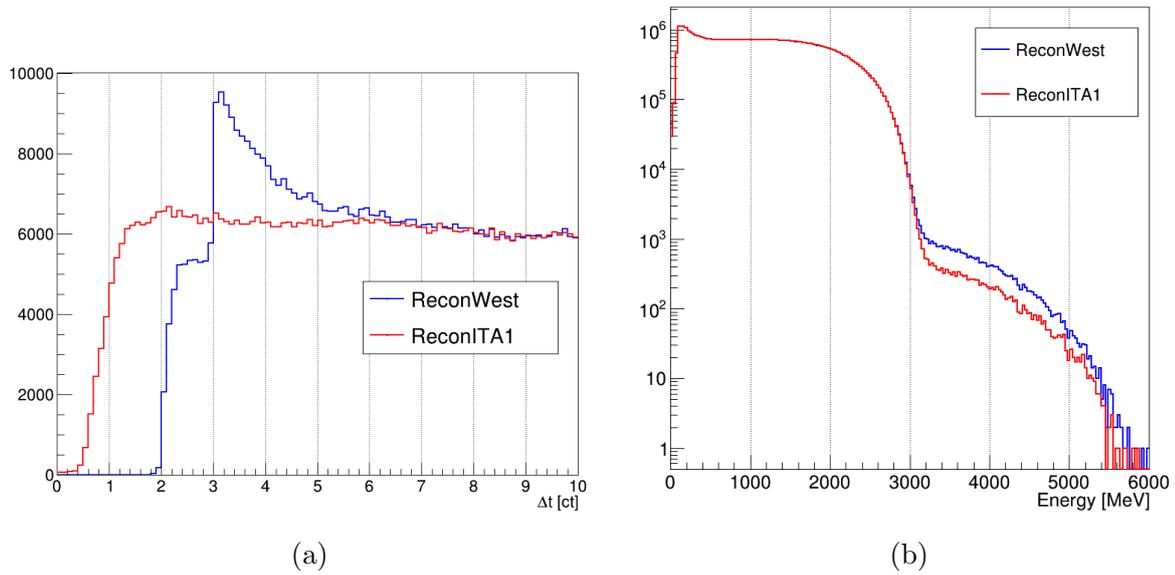
Figure 5.42: Comparison between clusters from ReconWest (blue) and ReconITA1 (red). (a) Time distance between two consecutive clusters. (b) Cluster energy spectrum before pileup correction.
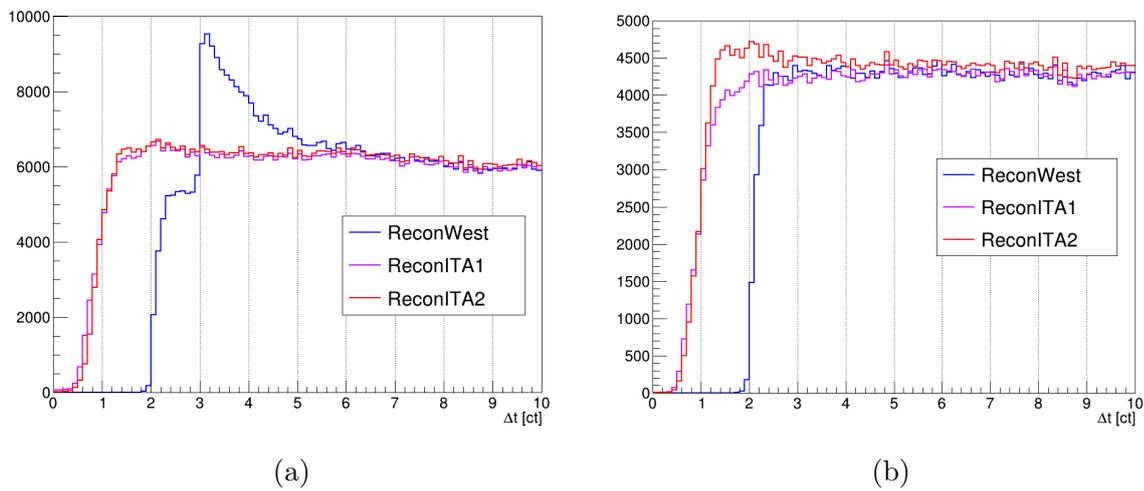


Figure 5.43: Comparison between clusters from ReconWest (blue), ReconITA1 (violet), ReconITA2 (red). ReconITA1 is the clustering applied at the end of the ReconWest reconstruction chain, while ReconITA2 is the clustering applied at the end of the ReconITA chain. (a) Time distance between two consecutive clusters. (b) The same distribution of (a) but selecting the second clusters to have more than one hit.

# Chapter 6

# The anomalous precession frequency measurement

The measurement of the anomalous precession frequency is one of the three key ingredients for obtaining the anomalous magnetic moment $a_\mu$. As explained in section 2.3.1, the parity-violation of the weak force in the muon decay allows the emitted positron to carry information of the original muon spin. The decay asymmetry is measured by the calorimeters by selecting positrons above a certain energy threshold. The positron spectrum in the laboratory frame of reference depends on the direction in which positrons are emitted during the decay, and, since the muon spin rotates with respect to the momentum direction, the positron spectrum as seen by calorimeters changes through time. The number of positrons of a particular energy bin will therefore oscillate with a frequency corresponding to the anomalous muon precession one: $\omega_a$.

As discussed in Chapter 5, the output of the reconstruction of the calorimeter data is the energy and detection time of the positrons produced by the muon decay. The distribution of positrons as a function of energy and time is shown in Figure 6.1a. By integrating over the energies above a certain threshold we obtain the iconic *wiggle* plot of Figure 6.1b. Before fitting for $\omega_a$ however, the data has to be corrected for detector effects such as pileup. In addition, the dynamics of the beam inside the storage ring introduce effects that add oscillations and distort the exponential shape. A measurement aiming to a precision of order 100 ppb requires a scrupulous control of all these beam-related effects, as well as an extremely accurate knowledge of the detector response to impinging particles.

This chapter presents the techniques used to measure the muon anomalous precession frequency for the data taken in Run-1, Run-2, and Run-3. The work presented here refers to the results obtained by the $\omega_a$-*Europe* group, of which I am part.



|  (a)  |  (b)  |
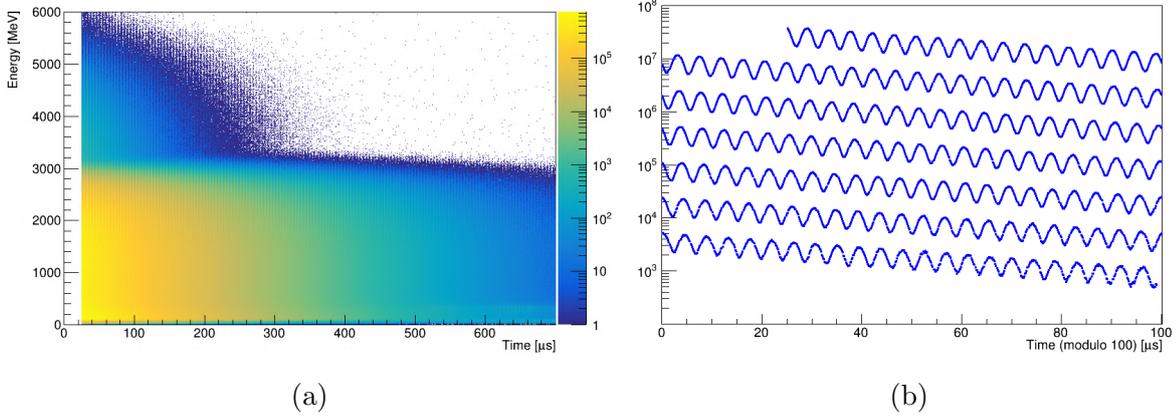
Figure 6.1: (a) Uncorrected energy-time histogram as produced by the reconstruction. (b) Wiggle plot obtained by integrating over the energies $E > 1700$ MeV of the histogram in (a). The time axis is wrapped onto itself every 100 $\mu$s.

## 6.1   Data sets

The reconstruction algorithm used for Run-1 is *ReconWest*, while Run-2 and Run-3 have been reconstructed using also the *ReconITA* clustering algorithm. The complete *ReconITA* procedure described in Chapter 5 has been used starting from Run-4.

Run-1, Run-2, and Run-3 are composed of 4, 7, and 11 datasets, respectively. The increasing number of subsets is due to the increased efficiency of the machine and the longer data collection times. Each dataset is characterized by having the same beam storage settings, as field index and kicker strength, and bounded by trolley field acquisitions. Datasets are typically limited to $\sim 10^5$ maximum files for convenience of production and data handling. Some datasets were discarded after the first round of production, such as 2A, 2I, 3A, and 3H, either because they were too small or were used for systematic studies, such as kicker strength tuning, and calibration purposes. While the analysis of the magnetic field and the data quality checks are carried out at the dataset level, the $\omega_a$ analysis combines multiple datasets in order to minimize the effects of systematic uncertainties which scale with statistics. For this reason, the combined Run-2 = [2B-2H], Run-3a = [3B-3M], and Run-3b = [3N-3O] are analyzed in

their entirety for the final estimation of $\omega_a$. During the first year of data taking, Run-1, the storage conditions varied significantly as both the E989 collaboration and the Fermilab Accelerator Division were trying to optimize the muon beam and its storage capabilities. In particular, for what regards the experiment, both the kicker currents and the kick timing, as well as the electrostatic quadrupole high voltages were modified several times during the data taking period. This forced us to analyze the datasets separately. Table 6.1 lists the statistics and storage conditions for all the mentioned datasets.

As explained in Chapter 4, the production process generates one output for each raw input file, and the reconstructed data is stored in *art dataproducts*. The positron data is then extracted from these files and stored in fewer `ROOT` files containing hundreds of subruns each. In particular, each file contains information on the time and energy of each positron, as well as the individual hits composing the positron's cluster. For convenience, positrons are stored on a *fill by fill* basis, which helps find consecutive positrons within the same 700 $\mu$s of beam storage, a feature useful to extract the pileup correction which will be discussed in section 6.3.

## 6.2    Hit time randomization

The muon beam enters the storage ring in bunches of $\sim$120 ns in length. The cyclotron period, i.e. the time of flight in the storage circumference, is $T_c = 149.2$ ns. Because of the bunched structure, the number of positrons detected by a calorimeter during the first turns is modulated by the cyclotron frequency. As the muons gradually de-bunch, the modulation flattens out, with a time constant of several microseconds, and muons become uniformly distributed within the ring, appearing as a continuous signal in the calorimeters.

During the first $\sim 50$ $\mu$s, however, the bunched structure of the muon beam is observed as an additional frequency in the time distribution of the positrons and it can affect the final result of the $\omega_a$ fit if not properly taken into account. In order to mitigate its effect, known as *fast rotation* signal, the wiggle histograms (Figure 6.1b) are binned using the cyclotron period as the bin width. This way, the modulation is averaged out within the bin. However, this method is not sufficient to remove the effect completely. In fact, when performing individual $\omega_a$ fits for each calo (Figure 6.2a), a sinusoidal oscillation of the $\omega_a$ value around the azimuth angle is present with an amplitude of $\sim$10 ppm. The frequency of the sinusoidal fit is fixed to the value of $2\pi/24$ corresponding to

| Dataset | Number of files | CTAGs [$\times 10^9$] | Field index | Kicker voltage |
|---------|-----------------|-----------------------|-------------|----------------|
| 1A | 23560 | 0.92 | 0.108 | 130 kV |
| 1B | 27556 | 1.28 | 0.120 | 137 kV |
| 1C | 39236 | 1.98 | 0.120 | 130 kV |
| 1D | 75512 | 4.00 | 0.108 | 125 kV |
| 2B | 13715 | 0.84 | 0.108 | 142 kV |
| 2C | 69303 | 4.11 | ” | ” |
| 2D | 60658 | 3.25 | ” | ” |
| 2E | 25178 | 1.25 | ” | ” |
| 2F | 23336 | 1.27 | ” | ” |
| 2G | 4929 | 0.26 | ” | ” |
| 2H | 6729 | 0.36 | ” | ” |
| 3B | 43213 | 1.57 | 0.107 | 142 kV |
| 3C | 12767 | 0.51 | ” | ” |
| 3D | 73347 | 3.52 | ” | ” |
| 3E | 29392 | 1.45 | ” | ” |
| 3F | 14746 | 0.72 | ” | ” |
| 3G | 33485 | 1.66 | ” | ” |
| 3I | 32619 | 1.58 | ” | ” |
| 3J | 25138 | 1.21 | ” | ” |
| 3K | 15656 | 0.72 | ” | ” |
| 3L | 11262 | 0.52 | ” | ” |
| 3M | 42315 | 1.80 | ” | ” |
| 3N | 67307 | 3.45 | 0.107 | 165 kV |
| 3O | 50292 | 2.52 | ” | ” |
| Run-2 | 203848 | 11.35 | 0.108 | 142 kV |
| Run-3a | 333940 | 15.27 | 0.107 | 142 kV |
| Run-3b | 117599 | 5.96 | 0.107 | 165 kV |
| Run-4 | 914766 | 40.19 | 0.107 | 165 kV |
| Run-5 | 1300000* | 50* | 0.107 | 165 kV |

Table 6.1: List of the datasets in Run-1, Run-2, and Run-3. CTAGs (Calorimeter TAGs) are the number of positrons reconstructed by the calorimeters with $E > 1700$ MeV and $t > 30$ $\mu$s. *Estimate, production not finalized yet.

one turn around the ring.

To further suppress this effect, a randomization of the hit times is performed. For each

positron, a random shift in the $[-\frac{T_c}{2}, \frac{T_c}{2}]$ ns range is applied to its detection time. Figure 6.2b shows the per-calorimeter fits after this randomization, where the oscillation amplitude is now compatible with zero.

The pseudo-random number generator (RNG) chosen to extract the time shifts is `RANLUX64` which is part of the `CLHEP` package. A unique random seed for each analyzed file is provided to the RNG in order to have repeatability when the data is reprocessed. The final $\omega_a$ value is obtained by averaging over 200 different random seeds.
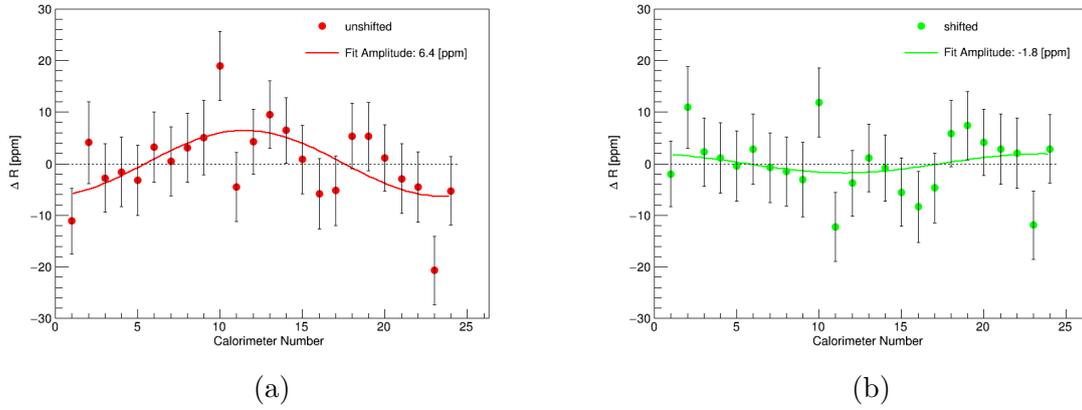


Figure 6.2: Per-calorimeter $\omega_a$ fits before (a) and after (b) applying a random time shift to the positron times.

## 6.3 Pile-up treatment

When two or more positrons hit the same calorimeter very close in time, the reconstruction procedure does not always manage to separate them. In such cases, a pileup event occurs and the affected positrons are reconstructed as a single particle. The probability of having two positrons in a time window $t + \delta t$ is proportional to the square of the positron rate $\rho$ and therefore the distribution of pileup events has an exponential lifetime of $\tau_\mu/2 \simeq 32.2$ $\mu$s. In fact, the positron "time density" $\rho(t)$ can be written as:

$$\rho(t) \propto e^{-t/\tau_\mu} \cdot \left(1 + A\cos(\omega_a t - \phi)\right). \tag{6.1}$$

Therefore, the probability of having a positron at time $t$ and one at time $t + \delta t$ hitting the same calorimeter can be written, for small values of $\delta t$ as:

$$\lim_{\delta t \to 0} \rho(t)\rho(t + \delta t) \propto e^{\frac{-t}{\tau_\mu/2}} \cdot \left(1 + A\cos(\omega_a t - \phi)\right)^2. \tag{6.2}$$

As a consequence, the pileup contamination, if not properly corrected, modifies the simple exponential decay introducing an early-to-late effect that can bias the measurement of $\omega_a$. It has to be noticed that even after being corrected, the pileup introduces a systematic uncertainty on $\omega_a$ which in Run-1 was among the largest ones, as shown in Table 5.1. This is one of the main reasons for which the *ReconITA* clustering, described in Section 5.5, has been developed. This new clustering, applied from Run-2 onward, improves the separation of pileup positrons achieving $\sim 100\%$ efficiency at $\Delta t \geq 1.5$ ns. This reduces the pileup contamination in the data, decreasing the associated systematic uncertainty proportionally.

### 6.3.1   Shadow method

The pileup events can be estimated and subtracted on a statistical basis. The probability of having two positrons at the same time is approximately the same as observing two individual positrons separated by a time $t_{gap}$. This time has to be small compared to the $\omega_a$ and $\tau_\mu$ timescales, in order for the muon density to remain constant. Because of the fast rotation effect, $t_{gap}$ should be either very close to 0 or a multiple of $T_c$ to ensure that the positron rate is not affected by the cyclotron frequency. In order to ensure that the two individual positrons are well separated and individually reconstructed, the value of $t_{gap} = T_c = 149.2$ ns has been chosen. The number of overlapping positrons in a window $[t, t + \delta t]$ is the same as the number of two single positrons in the windows $[t, t + \delta t]$ and $[t + t_{gap}, t + t_{gap} + \delta t]$. It is then possible to statistically simulate the effect of pileup events by artificially shifting the time of the second positron by $-t_{gap}$, combining it with the first positron, and obtaining an artificial pileup event to be subtracted from the data. This technique is called "shadow window method".
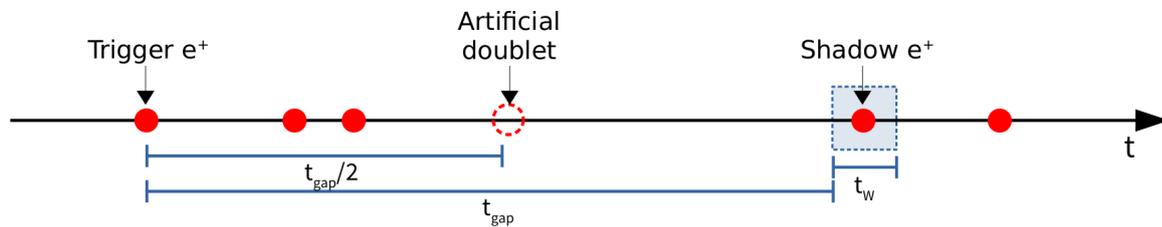


Figure 6.3: Example of a shadow method coincidence. The combined artificial doublet will be shifted by $t_{gap}/2$ after the trigger positron.

The shadow method algorithm searches for pairs of positrons hitting the same calorime-

ter separated by a time $t_{gap}$. More specifically, it searches for the second positron in a small *shadow* window of width $t_W$ starting $t_{gap}$ after the first *trigger* positron, as schematized in figure 6.3. If a coincidence is found, the two positrons are combined to form an artificial *doublet* with time and energy:

$$t_D = \frac{E_1 t_1 + E_2(t_2 - t_{gap})}{E_1 + E_2} + \frac{t_{gap}}{2} \,, \tag{6.3a}$$

$$E_D = E_1 + E_2 \,. \tag{6.3b}$$

The additional factor $\frac{t_{gap}}{2}$ accounts for the small but non-negligible variation of the positron rate between $t$ and $t + t_{gap}$. Assuming that the positron rate is purely exponential, $\rho(t) = e^{-t/\tau_\mu}$, the pileup doublet should be placed at the time $t'$ such that:

$$\rho^2(t') = \rho(t) \cdot \rho(t + t_{gap}) \,, \tag{6.4}$$

which is true for

$$t' = t + \frac{t_{gap}}{2} \,. \tag{6.5}$$

The real positron distribution, however, is not a pure exponential and contains, at first order, the oscillation corresponding to $\omega_a$. This remaining effect is treated as a systematic uncertainty and not considered in the correction.

For each coincidence, both the artificial doublet and the original single positrons are stored into individual time-energy histograms $D(t, E)$ and $S_D(t, E)$. The pileup correction is then applied by subtracting the doublets and adding the singlets to the data as follows:

$$N(t, E) = N'(t, E) + \alpha(S_D(t, E) - D(t, E)) \,, \tag{6.6}$$

where $N(t, E)$ is the true positron distribution and $N'(t, E)$ is the observed positron distribution. $\alpha$ is a scaling parameter used to tune the pileup correction amplitude. This accounts for inaccuracies of the shadow window method and for higher order effects due, for example, to "triple pileup". The optimal value for $\alpha$ is determined for Run-1 by a scan of the parameter over the range $[0, 2]$ and by searching for the value which minimizes the $\chi^2$ of the $\omega_a$ fit.

## 6.3.2    Reconstruction effects

If every positron present in the shadow window is combined to the trigger positron to form a doublet, this is equivalent to assuming that any two clusters within the time window $[t, t + t_W]$ are always merged into one by the positron reconstruction algorithms. This is not generally true, as the probability of separating the two components of a pileup event is not exactly a step function as shown in Figure 5.43a. This problem was solved in Run-1 by applying an artificial dead time (ADT) to the calorimeters at the analysis level. When two clusters, $(t_1, E_1)$ and $(t_2, E_2)$, are closer in time than $t_{ADT}$, they are merged in a single cluster with time and energy calculated as:

$$t = \frac{t_1 E_1 + t_2 E_2}{E_1 + E_2}\,, \qquad E = E_1 + E_2\,. \tag{6.7}$$

This way, the probability of merging two consecutive clusters is a box function equal to 1 in the $[0, t_{ADT}]$ range. The acceptance of the shadow window is a box function as well, equal to 1 in the $[t_{gap}, t_{gap} + t_W]$ range. By setting $t_{ADT} = t_W$, the number of artificial doublets built from shadow coincidences correctly represents the amount of pileup in the data.

The optimal value of $t_{ADT}$ is the minimum distance such that the pileup separation efficiency becomes a step function, so that the additional dead time does not significantly increase the pileup contamination in the data. The Run-1 clustering algorithm described in section 5.5 achieves 100% pileup separation at $\Delta t \geq 5$ ct. The chosen value for Run-1 is, therefore, $t_{ADT} = t_W = 5$ ct.


**Re-clustering**

As described in the previous section, the pileup probability, as well as the energy of the merged cluster, depends upon the details of the reconstruction algorithm. Mostly for this reason, I have developed the ReconITA clustering which is being applied for the analyses of Run-2 and Run-3, to be published in the first half of 2023, with the goal of improving the precision of the pileup correction in order to decrease the associated systematic uncertainties.

When two positrons are selected by the shadow method, the hits of the second one are shifted by $-t_{gap}$ and combined with the ones of the first cluster. The ReconITA clustering is then applied. If this second clustering is able to separate the two positrons, then the coincidence event is discarded and not considered for the pileup correction. This

strategy removes the need for an artificial dead time, and makes the pileup correction to not depend on the width of the window $t_W$. In any case, the shadow window width needs to be larger than the minimum distance such that the clustering algorithm manages to separate two positrons 100% of the time. For the ReconITA clustering this is true for $\Delta t \geq 1.5$ ct. A window size of $t_W = 4$ ct is used for Run-2 and Run-3.

**Pulse fit correction**

When a pileup event occurs, the showers of the two positrons might overlap spatially in the calorimeter. In such cases, the SiPMs will receive two signals very close in time that the reconstruction algorithm might not be able to distinguish, fitting a single pulse instead. Unless the two pulses are exactly simultaneous, the energy determined by fitting a single pulse would differ from the sum of the two pulses. Not accounting for this results in a systematic bias on the observed energy used in Equation 6.3b. Figure 6.4 shows an example of two overlapping pulses of equal energy which are 1 ct apart. The fit using a single pulse results in a slightly underestimated energy.
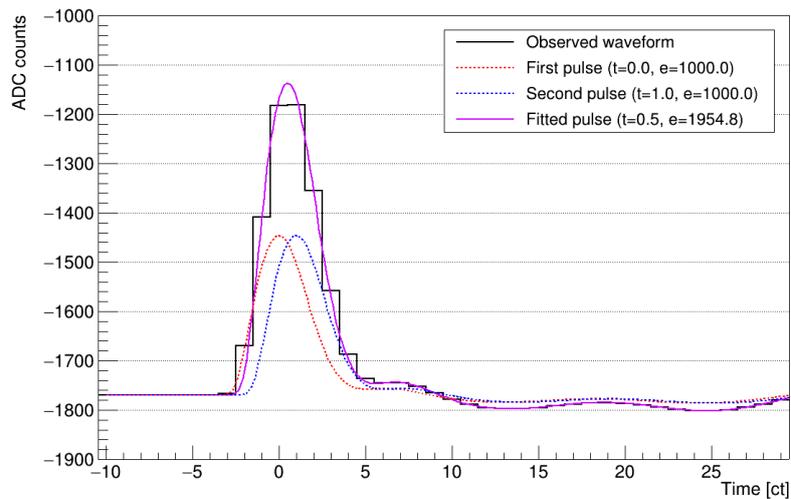


Figure 6.4: Template fit of two pulses when they are $\Delta t = 1$ ct (1.25 ns) apart. A single template fit is performed as the reconstruction is not able to distinguish between the two underlying pulses. The measured energy ($E_{fit} = 1954.8$ ADC*ct) is smaller than the sum of the two underlying pulses.

To study this effect, real SiPM pulses have been used to create artificial traces by summing two pulses with arbitrary time separation and energies. The traces are then sampled at 800 MHz (1.25 ns sampling time) and the points are multiplied by a smearing factor, extracted randomly from a Gaussian distribution of variance $\sigma = 0.003$. The smearing factor, representing the SiPM read noise, has been extracted from the positron

data by evaluating the variance of the trace fit residuals. Finally, the trace points are digitized into integer ADC values and the result is stored in a format which can be processed by the reconstruction algorithm as if it was real raw data. By analyzing the fitter response, the following effects are extracted as a function of the time separation $\Delta t$, $E_1$, and $E_2$:

- the efficiency of separating the two pulses,

- the bias in energy of the single fitted pulse,

- the bias in time of the single fitted pulse,

Figure 6.5 shows these corrections on average. When there are two or more pulses on the same SiPM in a shadow window coincidence, they are merged into one pulse according to the probability of Figure 6.5a. If merged, the following corrections are applied:

$$E' = f_{corr} \cdot (E_1 + E_2), \qquad t' = \frac{t_1 E_1 + t_2 E_2}{E_1 + E_2} + t_{corr}, \tag{6.8}$$

where both $f_{corr}$ and $t_{corr}$ depend on $\Delta t$, $E_1$, and $E_2$, according to the corrections of Figg. 6.5b and 6.5c respectively.



(a)

(b)

(c)
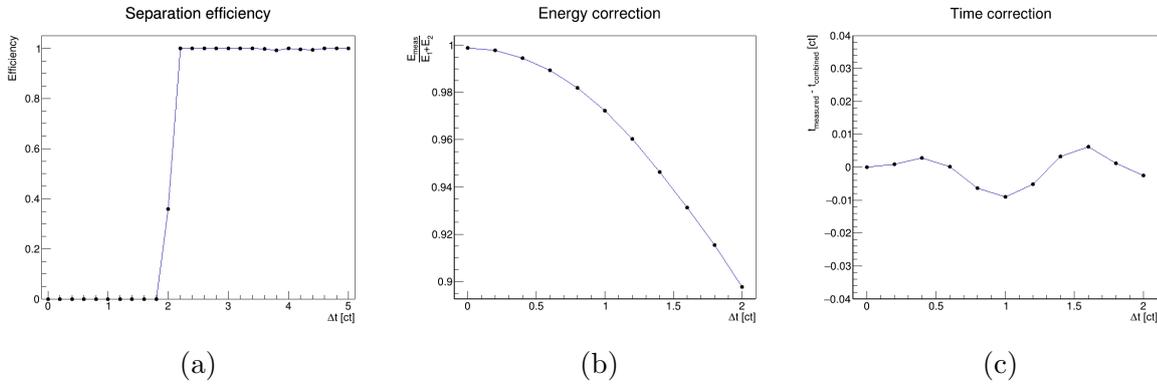
Figure 6.5: Pulse fit average corrections for multi-hit crystals in a pileup coincidence. (a) is the efficiency of separating the two individual pulses, (b) is the ratio between the fitted energy and the sum of the two pulses, and (c) is the difference between the fitted time and the energy-weighted average of the two pulses.

### 6.3.3  Triple coincidences

The probability of pileup events containing three simultaneous positrons goes like $\rho^3(t)$, so if the statistics is limited their effect can be ignored. For this reason, in Run-1

the positron spectrum is corrected only for the effect of "double" pileup (pileup of two positrons). The contribution of the "triple" pileup has been treated as a systematic uncertainty, which contributes on the final $\omega_a$ measure to less than 5 ppb.

From Run-2 and beyond, the larger datasets and higher statistics result in a contamination of triple pileup events which cannot be ignored. This is true even if the pileup effect is reduced by the ReconITA clustering (Figure 6.6).
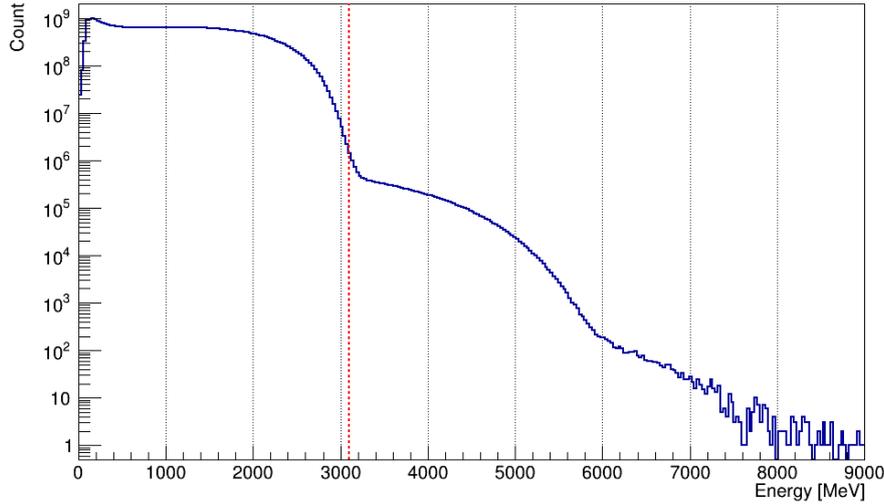


Figure 6.6: Run-2 positron energy spectrum before pileup correction is applied. The red dashed line corresponds to the positron endpoint of $E = 3100$ MeV. The tails visible at $E > 3100$ MeV and $E > 6200$ MeV are due to double and triple pileup contamination.

Triple pileup is corrected using the same shadow method described in section 6.3.1. Two shadow windows are used in this case at $t_{gap}$ and $2t_{gap}$ after the trigger positron. The three positrons are combined to form an artificial *triplet* with time and energy:

$$t_T = \frac{E_1 t_1 + E_2(t_2 - t_{gap}) + E_3(t_3 - 2t_{gap})}{E_1 + E_2 + E_3} + t_{gap} \,, \tag{6.9a}$$

$$E_T = E_1 + E_2 + E_3 \,. \tag{6.9b}$$

In this case, the triplet is shifted as $t' = t + t_{gap}$ in order to satisfy:

$$\rho^3(t') = \rho(t) \cdot \rho(t + t_{gap}) \cdot \rho(t + 2t_{gap}) \,. \tag{6.10}$$

While the technique is similar, double and triple pileup corrections cannot be extracted and applied independently. First of all, to avoid double counting, the two double coincidences $(t_1, t_2)$, $(t_2, t_3)$ that are part of a triple coincidence $(t_1, t_2, t_3)$ must not be

considered.

Then, we must consider the possibility that one of the positrons used to build a double coincidence can be a pileup event itself. This contamination is of the same order as the triple correction. The same effect is present in triplets as well, but ignored since of higher order: $\mathcal{O}(\rho^4)$. There are six possible combinations of events in a double coincidence, listed in Table 6.2.

| Trigger | Shadow | Combination | | Probability |
|:---:|:---:|:---:|:---|:---:|
| $e^+$ | $e^+$ | $2e^+$ | Doublet | $\rho^2(t)$ |
| $e^+$ | $2e^+$ | $3e^+$ | Triplet | $\rho^3(t)$ |
| $2e^+$ | $e^+$ | $3e^+$ | Triplet | $\rho^3(t)$ |
| $e^+$ | $3e^+$ | $4e^+$ | Quadruplet | $\rho^4(t)$ |
| $2e^+$ | $2e^+$ | $4e^+$ | Quadruplet | $\rho^4(t)$ |
| $3e^+$ | $e^+$ | $4e^+$ | Quadruplet | $\rho^4(t)$ |

Table 6.2: Contamination of true pileup events in double coincidences. The trigger and shadow columns show the number of real positrons contributing to the single clusters observed.

The first row in Table 6.2 is the desired coincidence of two individual positrons used to form an artificial doublet. The following rows are included in the "double pileup" correction as a contamination and must be subtracted. The last three combinations are ignored since they contribute at higher order $\mathcal{O}(\rho^4)$.

By calling $D(t, E)$ and $T(t, E)$ the distributions of double and triple pileup respectively, a pileup correction on the positron data takes the form:

$$N = N' + (S_D - D) + (S_T - T), \tag{6.11}$$

where $N(t, E)$ is the true positron distribution, $N'(t, E)$ is the observed cluster distribution, and $S_D$ and $S_T$ are the two and three true positrons contributing to the double and triple pileup respectively. The variables $(t, E)$ have been omitted for clarity.

Because of the contamination cases listed in Table 6.2, the observed pileup distribution $D'$ extracted with the shadow method differs from the true distribution $D$. The first row contributes as $D$, while the second and third rows, having $1 + 2$ positrons, contribute as $T$ each:

$$D' = D + 2T \tag{6.12}$$

Similarly, the single clusters selected by the shadow method in the double coincidences are contaminated too. The first row of Table 6.2 has two individual positrons, and therefore represents the real singlet distribution $S_D$. However, the second and third rows contribute each with a single positron and a pair of positrons, with probability $\rho^3(t)$. The single one contributes as much as a positron from a triple pileup coincidence, and it is distributed as a third of $S_T$. The pair of positrons has the energy of a double pileup $D$ but is distributed in time as a triple pileup event ($\propto \rho^3(t)$). Its contribution is therefore the same as $D$ but scaled by $\rho(t) = T(t)/D(t)$.

The sum of the three rows is:

$$
\begin{aligned}
S'_D &= S_D + \left( \frac{1}{3} S_T + \frac{T(t)}{D(t)} D \right) + \left( \frac{T(t)}{D(t)} D + \frac{1}{3} S_T \right) \\
&= S_D + \frac{2}{3} S_T + 2\rho(t) D ,
\end{aligned}
\tag{6.13}
$$

since

$$
\frac{T(t)}{D(t)} = \frac{\rho^3(t)}{\rho^2(t)} = \rho(t) .
\tag{6.14}
$$

As previously mentioned, by ignoring higher order terms $\mathcal{O}(\rho^4)$ the observed triple coincidences match with the true distributions:

$$
T' = T , \qquad S'_T = S_T .
\tag{6.15}
$$

By combining all the observed distributions from double and triple shadow method coincidences we have:

$$
\begin{cases}
D' = D + 2T \\
S'_D = S_D + \frac{2}{3} S_T + 2\rho(t) D \\
T' = T \\
S'_T = S_T
\end{cases}
.
\tag{6.16}
$$

Solving the system for the true distributions $D$, $S_D$, $T$, $S_T$:

$$
\begin{cases}
D = D' - 2T' \\
S_D = S'_D - \frac{2}{3} S'_T - 2\rho(t)(D' - 2T') \\
T = T' \\
S_T = S'_T
\end{cases}
.
\tag{6.17}
$$

Finally, Equation 6.11 becomes:

$$N = N' + \left( S'_D - D' + 2T' - \frac{2}{3}S'_T - 2\rho(t)(D' - 2T') \right) + (S'_t - T') \, , \qquad (6.18)$$

where $\rho(t)$ can be empirically extracted from the observed distributions:

$$\rho(t) = \frac{T(t)}{D(t)} = \frac{T'(t)}{D'(t) - 2T'(t)} \, . \qquad (6.19)$$

In analogy with Equation 6.6, to account for any inaccuracies of the shadow method procedure, the formula applied for each time-energy bin is the following:

$$N = N' + f_{scale} \left[ f_D \left( S'_D - D' + 2T' - \frac{2}{3}S'_T - 2\rho(t)(D' - 2T') \right) + f_T \left( S'_T - T' \right) \right] \, , $$
$$(6.20)$$

where $f_D$ and $f_T$ are fine-tuning scaling factors for double and triple pileup respectively, and $f_{scale}$ is a global scaling factor (equal to 1 by default) used for evaluating the pileup systematic uncertainty.

The optimal values for $f_D$ and $f_T$ are determined by looking at the ratio between the positron energy spectrum and the pileup correction in the energy ranges of $[3500, 4500]$ MeV and $[6500, 7500]$ MeV respectively. The results of the horizontal fits are plugged into $f_D$ and $f_T$, as shown in figure 6.7. The fitted value of $f_D = 1.08$ is close to 1 as expected, whereas the value of $f_T$ has been fixed to 1 because of the high uncertainty given by the low statistics above $E > 6000$ MeV. Figure 6.8 shows the time and energy distributions of the double and triple pileup corrections for the ReconITA clustering for dataset Run-2C. The double pileup contamination at $t = 30$ $\mu$s is $\sim 2 \times 10^{-3}$ and decays as $\tau_D = 32.2$ $\mu$s, while the triple pileup contamination is $\sim 4 \times 10^{-6}$ and decays as $\tau_T = 21.5$ $\mu$s.

## 6.4   Positron weighting techniques

The result of time randomization and pileup correction (Sections 6.2 & 6.3) is a fully corrected 2D time-energy histogram of the positrons detected by the calorimeters, shown in Figure 6.9.

The amplitude of the oscillation due to the anomalous precession frequency $\omega_a$ de-
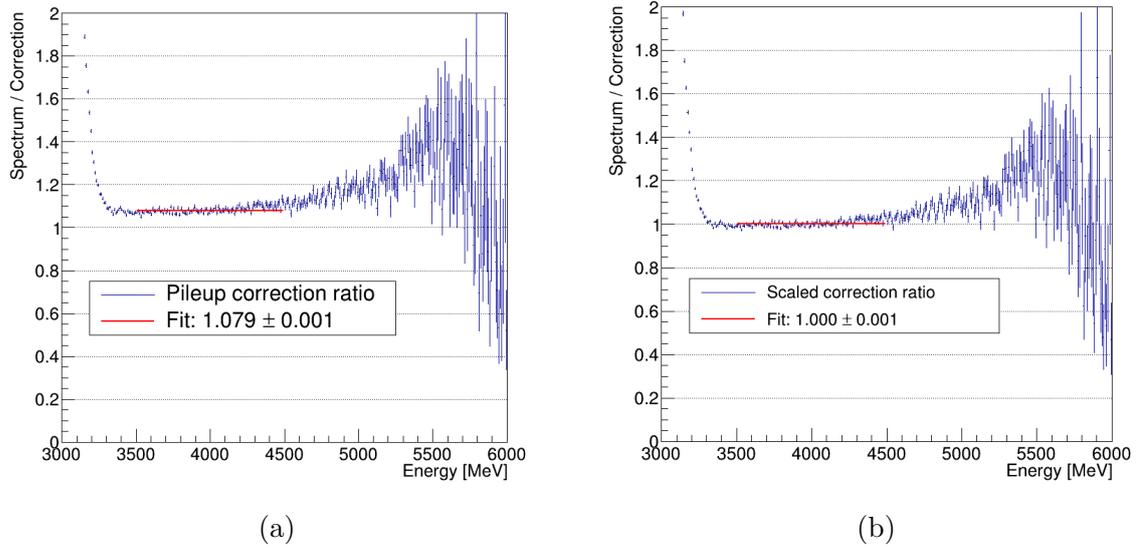
Figure 6.7: Ratio between the positron spectrum and the pileup correction (a) for dataset Run-2C. A horizontal fit in the $[3500, 4500]$ MeV range is performed to finely tune the correction amplitude. Figure (b) shows the same ratio with the applied scaling $f_D = 1.079$.
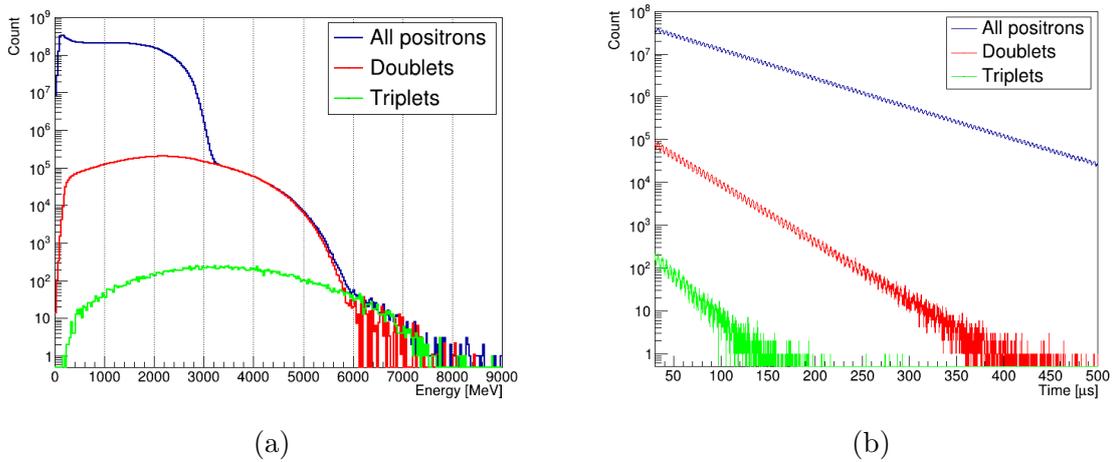


Figure 6.8: Pileup energy (a) and time (b) distributions.

pends on the positron energy as anticipated in Section 2.3. This asymmetry amplitude is maximum for 3.1 GeV positrons, negative for low energy positrons, and zero for $E \approx 1.0$ GeV positrons (Figure 6.10).

By integrating over the entire positron energy range, and assuming perfect detector efficiency, the observed rate would follow the standard muon exponential decay and no asymmetry would be visible. In reality, due to the calorimeter geometrical acceptance
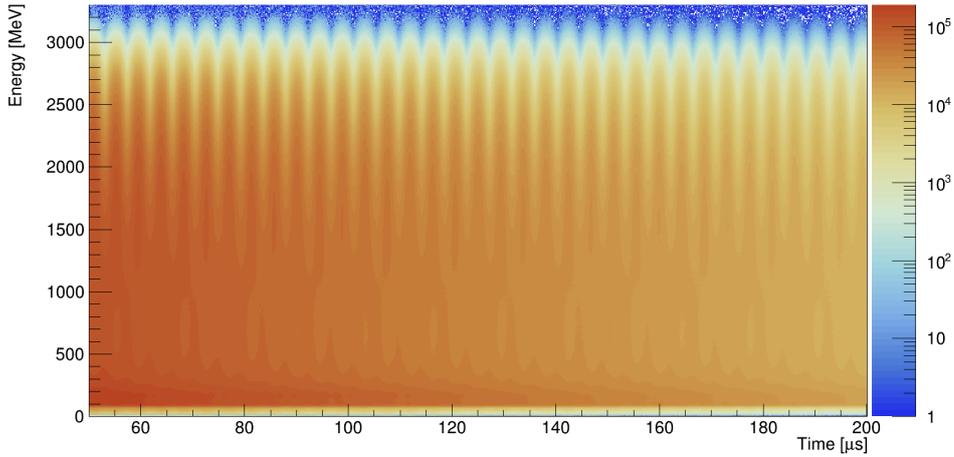
Figure 6.9: Time-energy histogram of positrons after pileup collection.
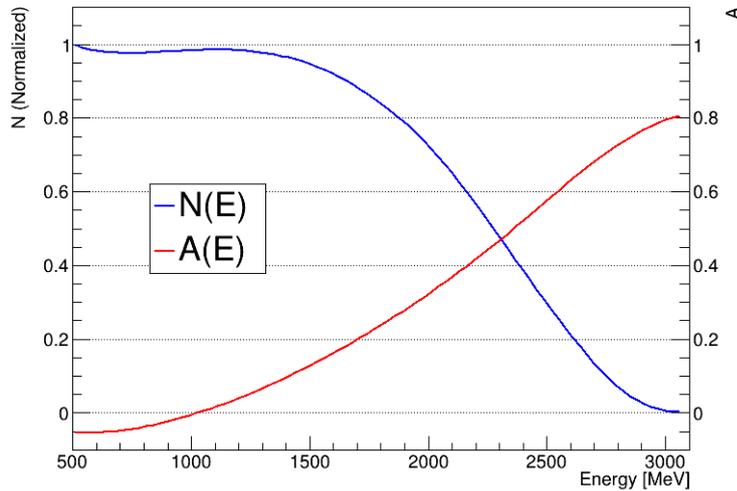


Figure 6.10: Number of positrons $N$ and decay asymmetry $A$ versus energy as measured by the calorimeters. For each 30 MeV bin both parameters are extracted with a $\omega_a$ fit of the form $N(t) = N \cdot e^{-t/\tau_\mu} \cdot (1 + A\cos(\omega_a t + \phi))$.

and to its efficiency as a function of energy, a residual asymmetry would still be observed. In order to maximize this asymmetry, and thus optimize the $\omega_a$ fit, a proper positron selection can be applied. More generally, a weighting function $w(E)$ is applied to the positrons depending on their energy, in order to build the *wiggle* plot. A simple way to do this is to apply a *box* function (Figure 6.11) which is 1 above a given energy threshold and 0 below. In this way, only high energy positrons, which have the highest asymmetry, are included in the fit. However, other choices are possible and will be described in the next sections. In general, the figure-of-merit (FOM) that has to be maximized to reach
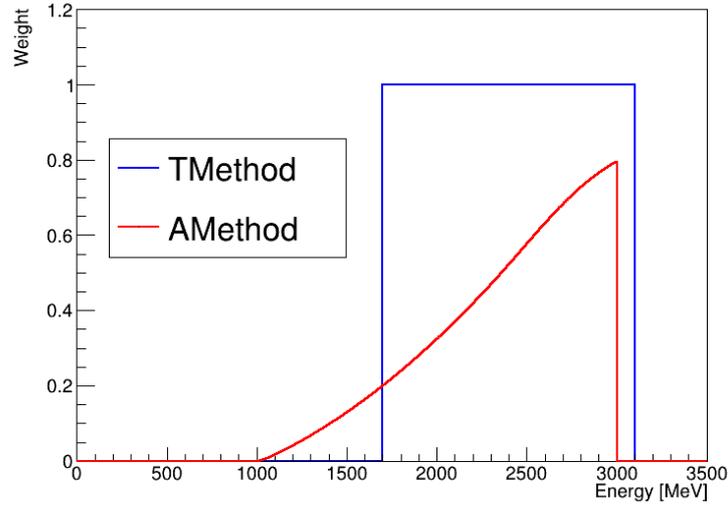
Figure 6.11: Weighting functions for the TMethod and AMethod.

the optimal statistical uncertainty on the $\omega_a$ fit is [72]:

$$\frac{1}{\sigma^2_{\omega_a}} \approx \frac{N_{th} \left\langle w(E) A(E) \right\rangle^2_E \tau^2}{2 \left\langle w^2(E) \right\rangle_E} \, , \tag{6.21}$$

where $N_{th}$ is the number of positrons with $E > E_{th}$ and $w(E)$ is the weighting function.

### 6.4.1   Threshold Method

As described above, a simple selection consists in using only the positrons above a given energy threshold $E_{th}$ in the fit. This is called the Threshold Method, or *TMethod*, and it is the reference method used also in previous experiments. The FOM of Equation 6.21 becomes:

$$\frac{1}{\sigma^2_{\omega_a}} \approx \frac{1}{2} N_{th} \left\langle A(E) \right\rangle^2_E \tau^2 \, . \tag{6.22}$$

By reducing the energy threshold, the number of events, and thus the statistical significance, tends to increase; at the same time, the asymmetry decreases until all the information on the muon spin direction is lost. To find the optimal threshold, a scan is performed by building many wiggle plots varying the threshold itself and then fitting the resulting histograms with the function

$$N(t) = N \cdot e^{-t/\tau_\mu} \cdot \left(1 + A\cos(\omega_a t + \phi)\right) \, . \tag{6.23}$$

Figure 6.12 shows the uncertainty $\sigma_{\omega_a}$ as a function of $E_{th}$ for the dataset Run-2C. The minimum is found for $E_{th} = 1700$ MeV.
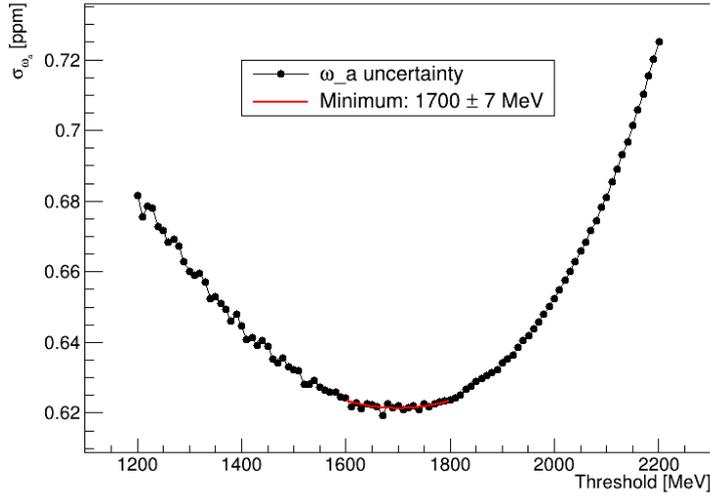


Figure 6.12: Statistical uncertainty of $\omega_a$ as a function of the energy threshold $E_{th}$ for the TMethod fits. The value of $E_{th} = 1700$ MeV minimizes the uncertainty.

## 6.4.2　Asymmetry Method

It is possible, however, to improve the statistical significance by weighting each event by its own asymmetry. In this way, high energy positrons, which are strongly correlated to the muon spin vector, receive a larger weight than intermediate energy ones when building the wiggle plot. It corresponds to setting $w(E) = A(E)$ and, for this reason, it is called Asymmetry Method or *AMethod*. This method is also the one which maximizes the theoretical statistical power of the positron data as it allows, in principle, to use all the positron events, even those with negative asymmetry. By setting $w(E) = A(E)$, Equation 6.21 becomes:

$$\frac{1}{\sigma_{\omega_a}^2} \approx \frac{N_{th} \langle A^2(E) \rangle_E^2 \tau^2}{2 \langle A^2(E) \rangle_E} = \frac{1}{2} N_{th} \langle A^2(E) \rangle_E \tau^2 . \tag{6.24}$$

The theoretical lower limit for the precision of $\omega_a$ is set by the Cramér-Rao limit, which in our case is [72]:

$$\frac{1}{\sigma_{\omega_a}^2} = \frac{1}{2} N_{th} \langle A(E) \rangle_E^2 \tau^2 . \tag{6.25}$$

Since $\langle A \rangle^2 \leq \langle A^2 \rangle$, the *AMethod* reaches this limit giving the best possible weighting function, shown in Figure 6.11. In principle, the best choice for the lower threshold

would be $E_{th} = 0$, thus integrating over all the detected positrons, including the ones with negative asymmetry. However, because of energy resolution, electronic noise, and contamination from muons exiting the storage area after interacting with the collimators (lost muons), the threshold $E_{th} = 1000$ MeV is used instead. As visible in Figure 6.13, the gain in precision from a lower threshold is negligible. Since the asymmetry is not well defined in the last few energy bins, an upper energy threshold is also applied at $E_{th}^{up} = 3000$ MeV.
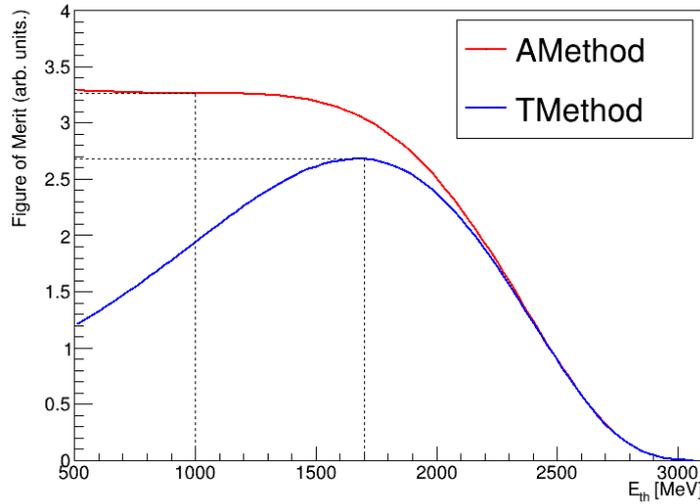


Figure 6.13: $\omega_a$ figure of merit as a function of the energy threshold $E_{th}$ for the TMethod and AMethod fits. The value of $E_{th} = 1700$ MeV maximizes the statistical power of the data for the TMethod.

### 6.4.3   Ratio Method

A third technique aiming at reducing the effect of slow and early-to-late effects has been developed. Before integrating over the energies, the data is treated in such a way that the exponential decay term in the wiggle plot is completely removed.

The positron data is randomly divided into four separate groups. For two of these groups, $u_+$ and $u_-$, the positron times are shifted by $+\frac{1}{2}T_a$ and $-\frac{1}{2}T_a$ respectively, where $T_a = \frac{2\pi}{\omega_a}$ is the $\omega_a$ period. The two remaining groups, $v_1$ and $v_2$, are left untouched. The following histograms are then built from these groups:

$$U(t, E) = u_+ + u_-, \qquad V(t, E) = v_1 + v_2. \tag{6.26}$$

The energy dependence is integrated out by using a weighting function, be it the one used for the *TMethod* or the *AMethod* described previously. Finally, the following ratio is built:

$$R(t) = \frac{V(t) - U(t)}{V(t) + U(t)}.$$

(6.27)

The result is, at the first order, a flat distribution modulated by the $\omega_a$ oscillation which can be fitted with the function:

$$R(t) = A\cos(\omega t - \phi) + R_0 + \mathcal{O}(A^2 e^{T/2\tau_\mu}).$$

(6.28)

The resulting function does not show the exponential decay, at first order, and it can be further improved by weighting the four groups $u_+$, $u_-$, $v_1$, $v_2$ proportionally to $e^{T/2\tau_\mu}$ : $e^{-T/2\tau_\mu} : 1 : 1$. Figure 6.14 shows the Ratio distribution after this reweighing. Only the $\omega_a$ oscillations are visible.
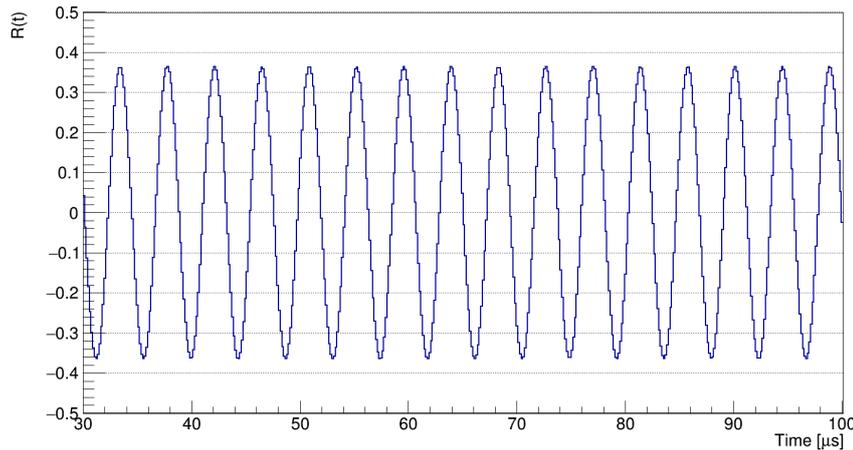


Figure 6.14: The wiggle plot $R(t)$ obtained with the Ratio Method.

## 6.5   The anomalous precession frequency fit

The result of integrating over the positron energies with one of the techniques described in the previous section is the final *wiggle* plot ready to be fitted. The precise extraction of the value of $\omega_a$ requires fitting with a functional form that describes the time distribution of the detected positrons in the calorimeters. The simple function described in Equation 6.23 is not sufficient; the fitting function has to incorporate all the effects from the anomalous precession, the muon exponential decay, the beam dynamics, the lost muons, and the beating frequencies.

This section describes the fitting model applied to the Run-1, Run-2, and Run-3 data.

## 6.5.1   Coherent Betatron Oscillation

The equation describing the physics of the muon decay and the anomalous spin precession is the five parameter function described in section 2.3:

$$N(t) = N \cdot e^{-t/\tau_\mu} \cdot [1 + A\cos(\omega_a t - \phi)] \ . \tag{6.29}$$

This function, however, does not perfectly match the observation from the calorimeters. Figure 6.15 shows the 5-parameter fit performed to the *wiggle* plot and the fit residuals. The oscillations visible in the residuals can be attributed to the movements of the beam while circulating around the storage ring.

A Fast Fourier Transform (FFT) of the residuals reveals several frequency peaks. The biggest one is relative to the radial Coherent Betatron Oscillation (CBO), with frequency $f_{CBO} \approx 0.37$ MHz. The two peaks adjacent to $f_{CBO}$ are related to the interference between the CBO itself and $\omega_a$, at $f_{CBO} \pm f_a$, where $f_a = \frac{\omega_a}{2\pi} \approx 0.23$ MHz. Further to the right, another peak corresponding to the Vertical Waist $f_{VW} \approx 2.28$ MHz is visible. A smaller peak next to the latter at $f_y \approx 2.22$ MHz is related to the vertical $(y)$ oscillation of the average muon beam position. Finally, the peak at very low frequency is an indication of slowly varying effects, and it is dominated by the muons which fail to stay in the storage region and exit before they decay into positrons. These slow effects are discussed in the next section.

In the Muon $g-2$ Experiment, both the radial betatron oscillation and the vertical waist are measured at the frequencies:

$$\omega_{CBO} = \omega_c - \omega_{CBO}^{true} \,, \tag{6.30a}$$

$$\omega_{VW} = \omega_c - \omega_{VW}^{true} \,, \tag{6.30b}$$

where $\omega_c$ is the cyclotron frequency ($f_c = \frac{\omega_c}{2\pi} \approx 6.7$ MHz). The reason is that the actual oscillations are faster than half of the cyclotron frequency, i.e. $\omega_{CBO}^{true} > \omega_c/2$. Since the time binning is set to $T_c$ as described in section 6.2, the Nyquist-Shannon sampling theorem implies that the *aliased* frequencies of Equations 6.30a-6.30b are measured instead.

The four oscillations of the beam with respect to the horizontal and vertical move-
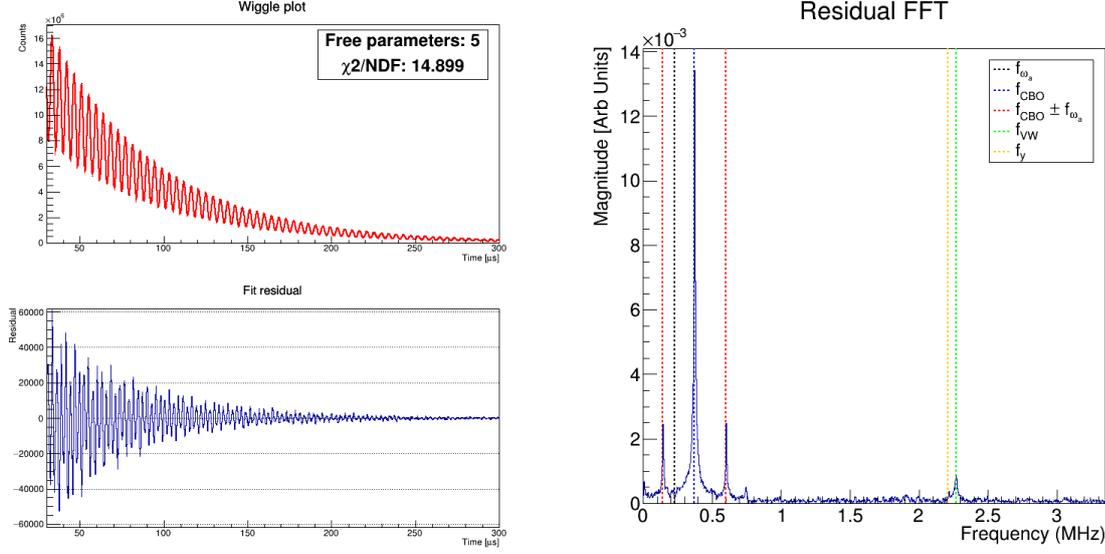
Figure 6.15: Run-2 wiggle plot fitted with the five parameter function of Equation 6.29. On the left, the fitted wiggle plot (top) and the fit residual (bottom). On the right, the Fast-Fourier-Transform analysis of the fit residual.

ments can be incorporated into Equation 6.29 as modulations of the following form:

$$N_{CBO}(t) = 1 + A_{CBO}\cos(\omega_{CBO}t - \phi_{CBO})e^{-\frac{t}{\tau_{CBO}}} , \tag{6.31a}$$

$$N_{2CBO}(t) = 1 + A_{2CBO}\cos(\omega_{2CBO}t - \phi_{2CBO})e^{-\frac{t}{\tau_{2CBO}}} , \tag{6.31b}$$

$$N_y(t) = 1 + A_y\cos(\omega_y t - \phi_y)e^{-\frac{t}{\tau_y}} , \tag{6.31c}$$

$$N_{VW}(t) = 1 + A_{VW}\cos(\omega_{VW}t - \phi_{VW})e^{-\frac{t}{\tau_{VW}}} , \tag{6.31d}$$

where the exponential terms model the decoherence lifetime. As shown in Section 2.5, the characteristic radial and vertical oscillation frequencies depend on the muon momentum. The storage ring accepts a range of momenta, so the muon beam is affected by a range of oscillation frequencies. This spread causes any coherent beam motion to decohere over time, with a decoherence time roughly equal to the inverse width of the angular frequency distribution.

These four modulations are added to Equation 6.29 bringing the number of parameters from five to 21:

$$N(t) = Ne^{-t/\tau_\mu}\left[1 + A\cos(\omega_a t - \phi)\right] \cdot N_{CBO}(t) \cdot N_{2CBO}(t) \cdot N_y(t) \cdot N_{VW}(t) . \tag{6.32}$$

The number of free parameters can be, however, reduced.

First of all, the width oscillations $\omega_{2CBO}$ and $\omega_{VW}$ have frequencies which are twice the ones of the oscillation of the mean:

$$\omega_{2CBO} = 2\omega_{CBO}\,, \qquad \omega_{VW} = 2\omega_y\,. \tag{6.33}$$

Similarly, the decoherence lifetimes are halved as well:

$$\tau_{2CBO} = \frac{\tau_{CBO}}{2}\,, \qquad \tau_{VW} = \frac{\tau_y}{2}\,. \tag{6.34}$$

Then, both the radial and vertical oscillation frequencies are related to the quadrupoles field index $n$ as described in section 2.5:

$$\omega_{CBO} = \omega_c\sqrt{n}\,, \qquad \omega_y = \omega_c\sqrt{1-n}\,. \tag{6.35}$$

The vertical oscillation frequency can therefore be expressed in terms of the radial CBO frequency:

$$\omega_y = \omega_{CBO}\frac{\sqrt{1-n}}{\sqrt{n}} = F_y \cdot \omega_{CBO}\sqrt{\frac{2\omega_c}{F_y \cdot \omega_{CBO}} - 1}\,, \tag{6.36}$$

where $F_y$ is a correction factor that takes into account the fact that quadrupoles do not cover the full azimuth and any non-linearities inside the ring. $F_y$ is free to float in the $\omega_a$ fit.

This reduces the number of free parameters from 21 to 17.

In addition to the modulation of the number of observed positrons, beam oscillations also affect the $g-2$ phase and asymmetry. The $g-2$ phase is directly related to the positron drift time, which in turn depends on the positron energy. As the beam oscillates closer and farther from the calorimeters, the average positron drift time change, and so does the average measured $g-2$ phase $\phi$. The energy dependence of the measured phase leads also to a periodic distortion of the observed energy spectrum. These effects require higher order corrections to $A$ and $\phi$, which vary over time:

$$A_{BO}(t) = 1 + A_A\cos(\omega_{CBO}t - \phi_A)e^{-\frac{t}{\tau_{CBO}}}\,, \tag{6.37a}$$

$$\phi_{BO}(t) = 1 + A_\phi\cos(\omega_{CBO}t - \phi_\phi)e^{-\frac{t}{\tau_{CBO}}}\,. \tag{6.37b}$$

The fitting function of Equation 6.32 then becomes:

$$
\begin{aligned}
N(t) =& Ne^{-t/\tau_\mu}\left[1 + A \cdot A_{BO}(t) \cdot \cos(\omega_a t - \phi + \phi_{BO}(t))\right] \cdot \\
& \cdot N_{CBO}(t) \cdot N_{2CBO}(t) \cdot N_y(t) \cdot N_{VW}(t) \,,
\end{aligned}
\tag{6.38}
$$

with 21 floating parameters.

## 6.5.2　Muon loss correction

A small fraction of the muons are lost from the stored beam before they decay into positrons. If their momentum is smaller (or higher) than the storage ring acceptance of 3.1 GeV ±0.15%, they will eventually hit one of the collimators and, by losing energy, exit the storage region. While slightly decreasing the number of muons participating in the $\omega_a$ signal, the collimators ensure a very sharp muon momentum distribution. The consequence of this is that the observed decay rate deviates from a pure exponential by the number of lost muons $L(t)$:

$$
\frac{dN}{dt} = -\frac{N}{\tau_\mu} - L(t) \,.
\tag{6.39}
$$

Solving this differential equation, we obtain:

$$
N(t) = Ne^{-t/\tau_\mu} \cdot \left(1 - k_{LM} \int_{t_0}^{t} L(t')e^{t'/\tau_\mu}dt'\right) \equiv Ne^{-t/\tau_\mu} \cdot \Lambda(t) \,,
\tag{6.40}
$$

where $k_{LM}$ is a floating parameter in the $\omega_a$ fit. The distortion caused by the lost muons is visible both as a non-flat component in the wiggle fit residual and as a low-frequency peak in the FFT of the same residual, as visible in Figure 6.22.

The function $L(t)$ can be extracted empirically by identifying the muons that hit the calorimeters before decaying into positrons. Therefore, $k_{LM}$ represents the fraction of the lost muons that are detected by the calorimeters.

### Muon selection

After a muon hits a collimator, it curls inwards as it does not have a momentum sufficient to stay in orbit. In its path, the muon can cross two or more calorimeters, as shown in Figure 6.16, before decaying or exiting the calorimeters acceptance. Such muon produces some characteristic signatures that can be used for its identification. For instance, it
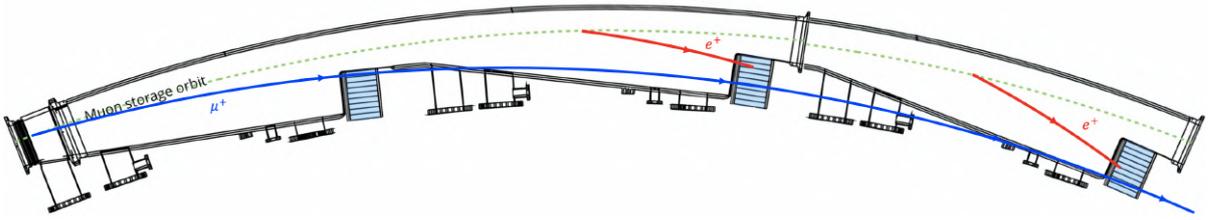
Figure 6.16: Diagram of a muon lost from the storage region. The particle hits three consecutive calorimeters before decaying or exiting the calorimeters acceptance region. Courtesy of M. Sorbara.

generally produces Čherenkov photons in only one crystal as its trajectory is mostly perpendicular to the calorimeter face. The deposited energy is the one of a minimum ionizing particle (MIP), which is described by the Bethe-Bloch relation [10] and it has a value of $E \approx 170$ MeV for a 3.1 GeV muon in a $PbF_2$ crystal of length 14 cm. The center of the distribution as measured in the calorimeters is actually defined to be 170 MeV by the absolute calibration as described in section 5.4.4. Finally, the lost muons can produce hits in consecutive calorimeters with a typical time of flight (TOF) of $t_{TOF} \approx T_c/24 = 6.22$ ns.
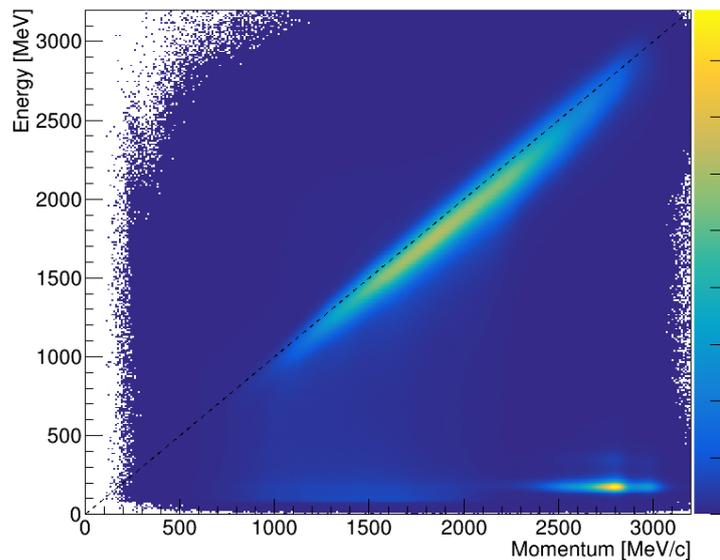


Figure 6.17: Cluster energy as measured by the calorimeter versus track momentum as measured by the tracker. The tracker detection efficiency decreases at low momentum due to its geometrical acceptance. In addition, the momentum of $\sim 3$ GeV/c particles is not precisely measured because of the large radius of curvature.

For the two calorimeters located after a tracking station, muons can be easily identified by comparing the momentum and the energy measured by the two detectors, as

shown in figure 6.17. The distribution along the diagonal is populated by positrons, for which $E \approx p$, while the muons are localized in the region at high momentum and low energy deposit. The cluster observed on the calorimeter is matched to the track measured by the tracker according to the following time and position cuts:

$$|t_{calo} - t_{tracker} - 2| < 8\text{ns} \,, \qquad |\Delta r| < 30\text{mm} \,, \tag{6.41}$$

where the shift of 2 ns in the time coincidence is the time needed by the muon to travel between the end of the straw tubes and the calorimeter back face.

By defining a cut on the $E/p$ ratio it is possible to extract a very pure sample of lost muons. Then, for the calorimeters hit by the selected particles, likelihood functions are built based on the deposited energy, the measured time of flight $t_{TOF}$, and the position distribution.

The lost muons used to build the empirical $L(t)$ are extracted from all 24 calorimeters by selecting the particles which satisfy the likelihood cuts in 3, 4, or 5 consecutive calorimeters. Double coincidences are not included, as they contain some residual positron background. However, they are used to estimate the related systematic effect. These triple, quadruple, and quintuple coincidences are then combined together to provide an estimation of $L(t)$ as shown in fig.6.18. The function is then integrated according to Equation 6.40 and stored as

$$J(t) = \int_0^t L(t')e^{t'/\tau_\mu}dt' \,. \tag{6.42}$$

Finally, the integral function is inserted into the $\omega_a$ fit function as

$$\Lambda(t) = 1 - k_{LM}J(t) \,. \tag{6.43}$$

### 6.5.3 Variable CBO frequencies

The beam oscillation frequencies implemented in section 6.5.1 depend on the electrostatic quadrupole field index as described in section 2.5. The field strength depends on the charge accumulated in the quadrupole plates and stabilize after the quadrupoles reach their nominal voltage. The charging process is designed to have a time constant of $\sim 5$ $\mu$s [49], so that the CBO terms are stable in the measurement window which starts at $t = 30$ $\mu$s.
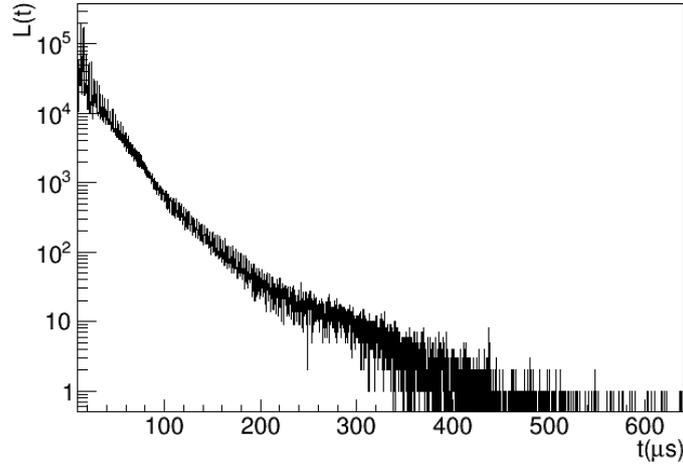
Figure 6.18: $L(t)$ function from triple, quadrupole, and quintuple coincidences.

During the Run-1 data-taking, two resistors connected to one of the quadrupoles were damaged and changed their resistance during time. This resulted in the quadrupole not reaching the nominal voltage before the $\omega_a$ fit start time of 30 $\mu$s, as shown in Figure 3.9b. As a consequence, this lead to a higher particle loss rate, a time dependent phase shift due to the correlation between the beam motion and the detector acceptance, and time dependent betatron frequencies.

To account for this, the CBO frequencies are allowed to vary as a function of time in the $\omega_a$ fit. The tracker system is capable of measuring the beam distribution as a function of time, and has the possibility to measure these effects directly. Figure 6.19 shows the varying CBO frequency as measured by the two tracker station for the Run-1D dataset, the one where the resistors were damaged the most. The frequency variation is modeled as a double exponential function:

$$\omega_{CBO}(t) = \omega_{CBO}^0 + \frac{A}{t}e^{-\frac{t}{\tau_A}} + \frac{B}{t}e^{-\frac{t}{\tau_B}} \tag{6.44}$$

where $\tau_B$ represents the nominal quadrupole charging lifetime and $\tau_A$ the slower lifetime introduced by the damaged resistors. In the $\omega_a$ fitting function (Equation 6.38) each instance of $\omega_{CBO}$ is substituted with the $\omega_{CBO}(t)$ of Equation 6.44. The four parameters $A$, $B$, $\tau_A$, $\tau_B$ are fixed to the values measured by the trackers, while $\omega_{CBO}^0$ remains free to float in the $\omega_a$ fit. Because of the relations of Equations 6.33 and 6.36, the oscillation frequencies $\omega_y(t)$, $\omega_{VW}(t)$, $\omega_{2CBO}(t)$ will vary with time too.
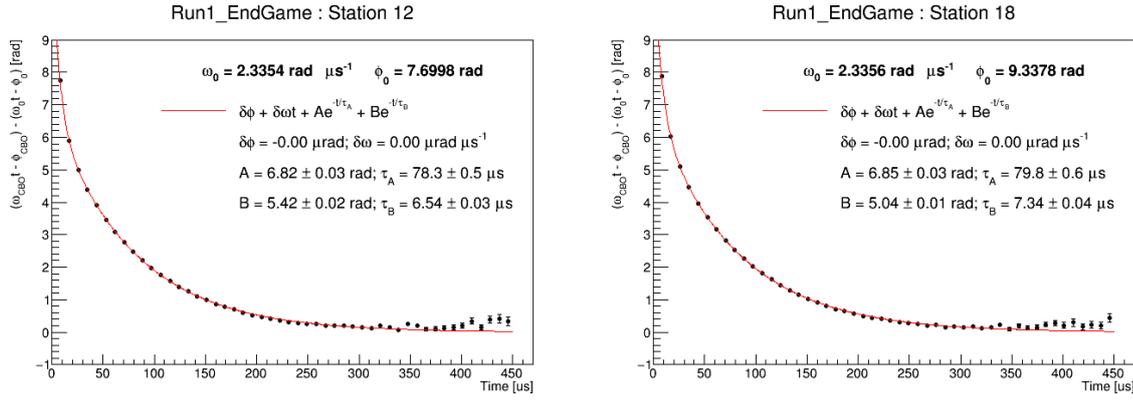
Figure 6.19: Variable CBO frequency as measured by the tracker detectors for the Run-1D dataset [53].

The damaged resistors have been replaced before the start of Run-2, so the effect of the quadrupoles on the variation of the CBO terms has been greatly suppressed. However, the second exponential term of Equation 6.44 remains and, while the time constant is small ($\tau_B \approx 7$ $\mu$s), the variable CBO term is included in the fit anyway.

The $\omega_a$ fitting function is, at this point, the following:

$$N(t) = N e^{-t/\tau_\mu} \left[ 1 + A \cdot A_{BO}(t) \cos(\omega_a t - \phi + \phi_{BO}(t)) \right] \cdot \tag{6.45}$$
$$\cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot N_{2CBO}(t) \cdot N_y(t) \cdot \Lambda(t) \,,$$

with 22 floating parameters.

### 6.5.4   CBO-VW beat frequency

Starting from Run-2, the running conditions of the experiment are stable enough that the $\omega_a$ analysis can be performed to the entire run combined. The number of positrons collected in all the Run-2 datasets is more than twice of the biggest dataset in Run-1 (Run-1D). Run-3a is even bigger at $\sim$1.35 times the statistics of Run-2.

One of the benefit of analyzing bigger datasets is the ability to be sensitive to small effects previously non-detectable. An example is the appearance of a small peak in the FFT of the residuals of the $\omega_a$ fit performed with the function of Equation 6.45, shown in Figure 6.26.

The new peak is one of the beating frequencies between the CBO and Vertical Waist

terms:

$$\omega_{CBO\pm VW}(t) = \omega_{CBO}(t) \pm \omega_{VW}(t)\,. \tag{6.46}$$

A new term is multiplied to the $\omega_a$ function of Equation 6.45:

$$N_{CBOVW}(t) = 1 + [A_+ \cos(\omega_+(t)t - \phi_+) + A_- \cos(\omega_-(t)t - \phi_-)]\, e^{-\frac{t}{\tau_{CBOVW}}}\,, \tag{6.47}$$

where the $+$ and $-$ subscripts refer to the $CBO+VW$ and $CBO-VW$ terms respectively. While only the $CBO-VW$ peak is visible in the FFT plot, both beating frequencies are included in the $\omega_a$ fit. This increases the number of floating parameters to 27.

## 6.5.5 Complete fitting function

The complete model used to fit the positron wiggle plot to extract the anomalous precession frequency includes all the modulating terms described in the previous sections:

$$N(t) = N e^{-t/\tau_\mu} \left[1 + A \cdot A_{BO}(t) \cos(\omega_a t - \phi + \phi_{BO}(t))\right] \cdot \tag{6.48}$$
$$\cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot N_{2CBO}(t) \cdot N_y(t) \cdot \Lambda(t) \cdot N_{CBOVW}(t)\,,$$

where the last term $N_{CBOVW}(t)$ is omitted for the Run-1 analysis because of the lack of sensitivity due to the lower statistics. In addition, the $N_y(t)$ term is very small in the Run-1A and Run-1B datasets for the same reason and set equal to 1.

The determination of the best-fit parameters is achieved by minimizing the Neyman $\chi^2$ using the `TMinuit` minimizer included in the `ROOT` software [62].
Because of the high number of parameters and the very small contribution of some terms, the fit routine does not converge if the fitting function is not adequately prepared. The algorithm developed for this analysis proceeds by fitting an increasingly higher number of parameters, while omitting or fixing the remaining ones. Table 6.3 shows the ordered list of functions used to reach the final 27-parameter model of Equation 6.48. For each of the functions, new terms are added according to the highest peak in the FFT of the residuals. At every step, the fit routine runs twice: the first time, all the parameters determined in the previous step are kept fixed in order to find an appropriate starting point for the new terms; the second time, all the parameters are released and the results are collected.
Uncertainties on the parameters at each step are computed using the `MnHesse` algorithm,

which is less time-consuming, while the final fit is performed using the `MnMinos` algorithm, which is more precise. The difference in the fitted $\omega_a$ value using the two different algorithms is $< 1$ ppb.

Figures 6.15 and 6.21$-$6.27 show the result at each step of the algorithm.

| Label | Equation | Notes |
|---|---|---|
| f5par | $N(t) = Ne^{-t/\tau_\mu}\left[1 + A\cos(\omega_a t - \phi)\right]$ | |
| f9par | $N(t) = Ne^{-t/\tau_\mu}\left[1 + A\cos(\omega_a t - \phi)\right] \cdot N_{CBO}(t)$ | |
| f12par | $N(t) = Ne^{-t/\tau_\mu}\left[1 + A\cos(\omega_a t - \phi)\right] \cdot N_{CBO}(t) \cdot N_{VW}(t)$ | |
| f13par | $N(t) = Ne^{-t/\tau_\mu}\left[1 + A\cos(\omega_a t - \phi)\right] \cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot \Lambda(t)$ | |
| f19par | $N(t) = Ne^{-t/\tau_\mu}\left[1 + A \cdot A_{BO}(t)\cos(\omega_a t - \phi + \phi_{BO}(t))\right] \cdot$ $\cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot \Lambda(t) \cdot N_{2CBO}(t)$ | |
| f20par | $N(t) = Ne^{-t/\tau_\mu}\left[1 + A \cdot A_{BO}(t)\cos(\omega_a t - \phi + \phi_{BO}(t))\right] \cdot$ $\cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot \Lambda(t)$ | Variable $\omega_{CBO}$ and $\omega_{VW}$ |
| f22par | $N(t) = Ne^{-t/\tau_\mu}\left[1 + A \cdot A_{BO}(t)\cos(\omega_a t - \phi + \phi_{BO}(t))\right] \cdot$ $\cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot \Lambda(t) \cdot N_{2CBO}(t) \cdot N_y(t)$ | Variable $\omega_{CBO}$, $\omega_{VW}$, and $\omega_y$ |
| f27par | $N(t) = Ne^{-t/\tau_\mu}\left[1 + A \cdot A_{BO}(t)\cos(\omega_a t - \phi + \phi_{BO}(t))\right] \cdot$ $\cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot \Lambda(t) \cdot N_{2CBO}(t) \cdot N_y(t) \cdot N_{CBOVW}(t)$ | Variable $\omega_{CBO}$, $\omega_{VW}$, and $\omega_y$ |

Table 6.3: List of labeled functions as used in the $\omega_a$ fit routine. The number in the labels represents the number of free parameters.

The choice of the fitting window is driven by achieving a balance between the gain in statistical power, and the reduction of systematic effects. An earlier fit start time would introduce more statistics, thanks to the exponential decay shape, but the beam takes time to stabilize after injection and the quadrupole scraping. For this reason, a fit start time is chosen to be around 30 $\mu$s after injection. A more precise value is determined such that the systematic effect due to the calorimeter gain fluctuations is minimized. As demonstrated in Figure 6.20 of reference [56], for any uncorrected slow effect, the sensitivity of $\omega_a$ is minimized when the fit starts on a node of the $\omega_a$ oscillation. For this reason, the fit start time for each dataset is chosen to be the time bin of the wiggle plot closest to the $\omega_a$ oscillation node near 30 $\mu$s. If the fit starts one time bin earlier, for example, the gain in statistical power is outweighed by the systematic precision lost.

On the other hand, both the statistical and systematic uncertainties are not greatly affected by the fit stop time. Table 6.4 lists the chosen start and stop times for the Run-1, Run-2, and Run-3 datasets.
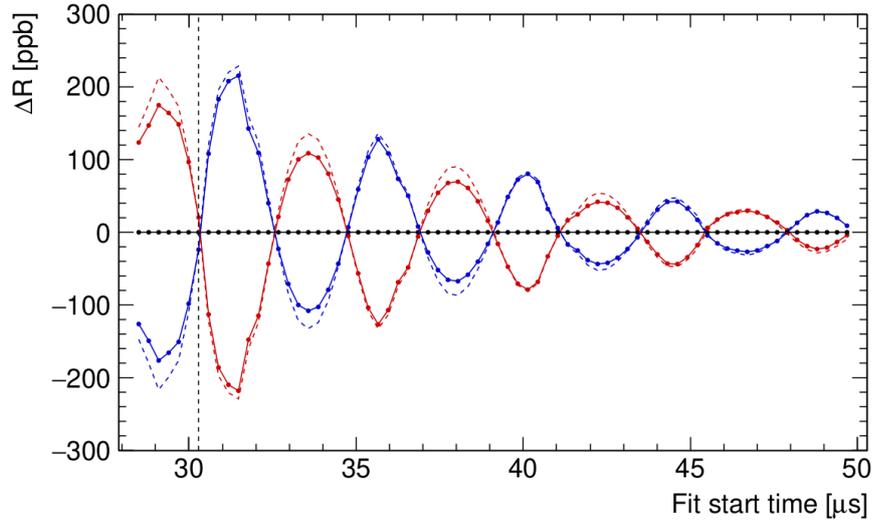
Figure 6.20: A scan over the fit start time when artificially scaling the In-Fill Gain correction amplitudes by a factor 0 (blue), 1 (black), and 2 (red), for the TMethod (solid) and AMethod (dashed). The change in R is with respect to that when using the nominal gain amplitude values. The chosen fit start time is marked by a vertical, dashed line. Figure from [56].

As mentioned in Section 6.5.3, the damaged quadrupoles resistors caused the beam to move even after 30 $\mu$s, during the measurement window. This effect got worse with time as the resistors deteriorated during the data acquisition. In order to reduce the related systematic effects on $\omega_a$ for the last Run-1 dataset, Run-1D, the start time has been shifted to the $\omega_a$ node closest to $\sim$50 $\mu$s.

| Dataset | Fit start time [$\mu$s ] | Fit stop time [$\mu$s ] |
|---|---|---|
| Run-1A | 30.1364 | 650.0210 |
| Run-1B | 30.1364 | 650.0210 |
| Run-1C | 30.1364 | 650.0210 |
| Run-1D | 49.8295 | 650.0210 |
| Run-2 | 30.1384 | 650.0644 |
| Run-3a | 30.1384 | 650.0644 |
| Run-3b | 30.1384 | 650.0644 |

Table 6.4: List of the $\omega_a$ fit start and stop times for the Run-1, Run-2, and Run-3 datasets.
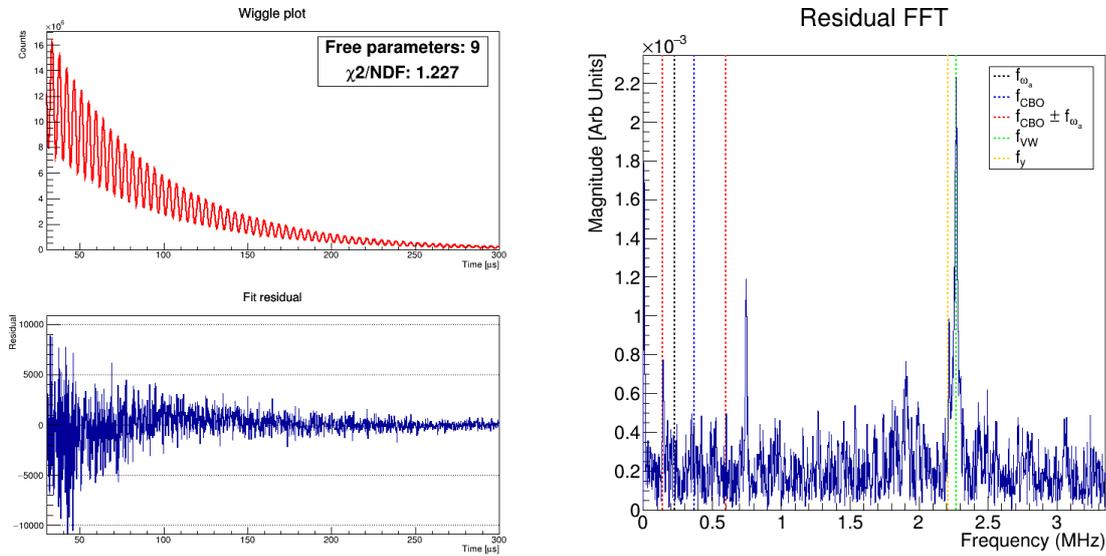
Figure 6.21: Run-2 wiggle plot fitted with the 9 parameter function *f9par* of Table 6.3.
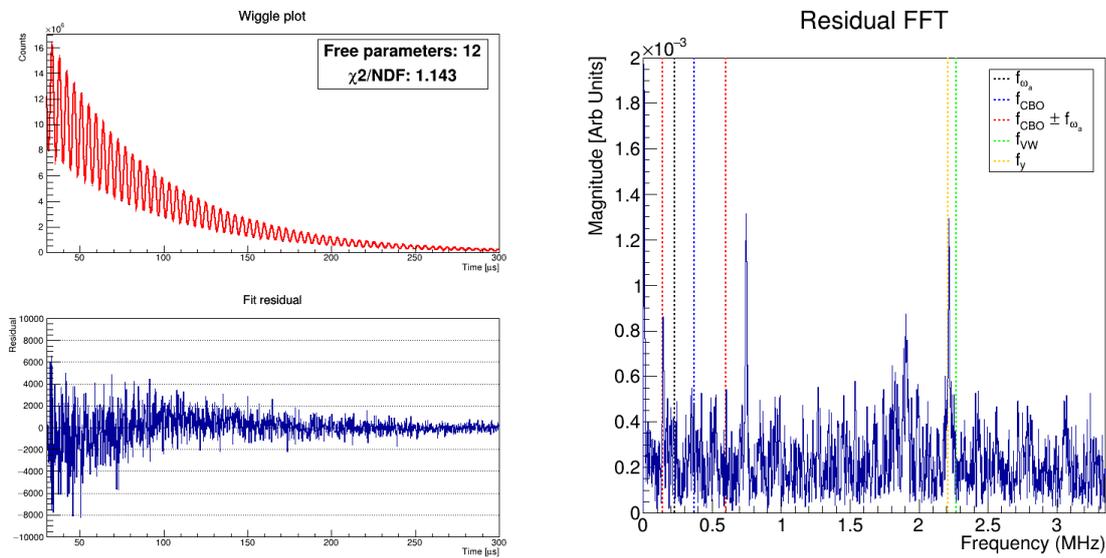


Figure 6.22: Run-2 wiggle plot fitted with the 12 parameter function *f12par* of Table 6.3.
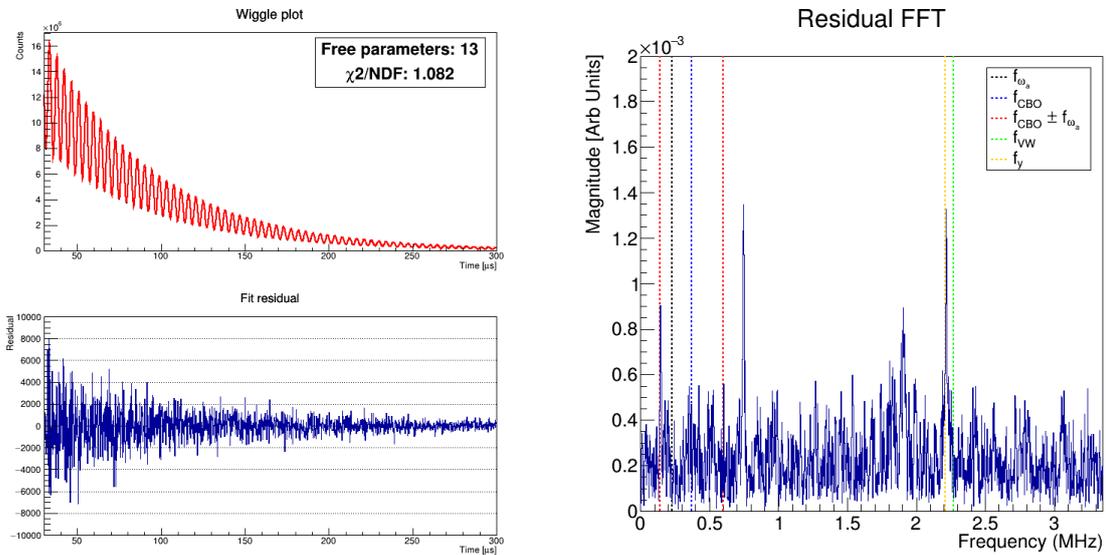
Figure 6.23: Run-2 wiggle plot fitted with the 13 parameter function *f13par* of Table 6.3.
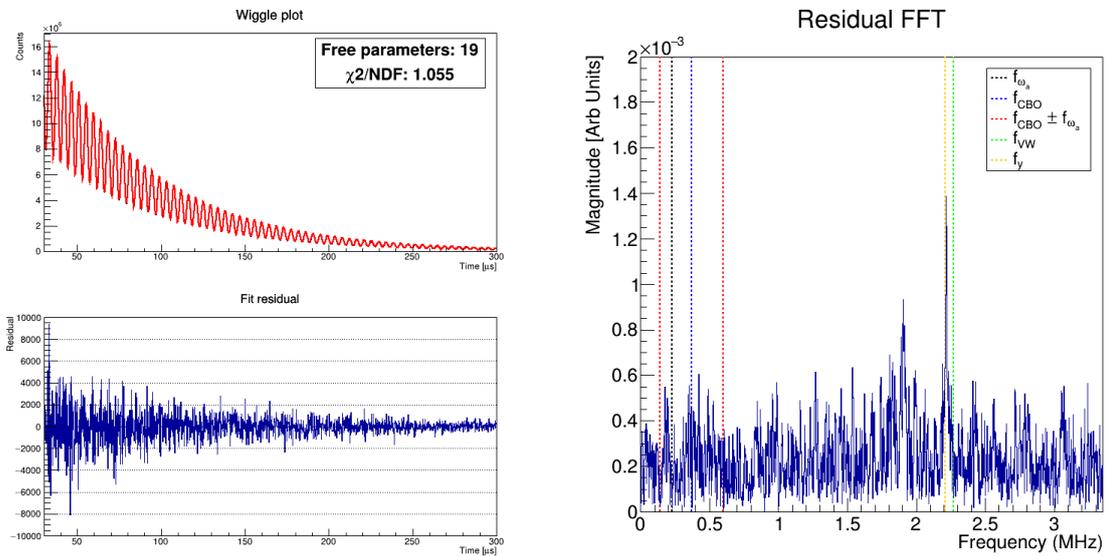


Figure 6.24: Run-2 wiggle plot fitted with the 19 parameter function *f19par* of Table 6.3.
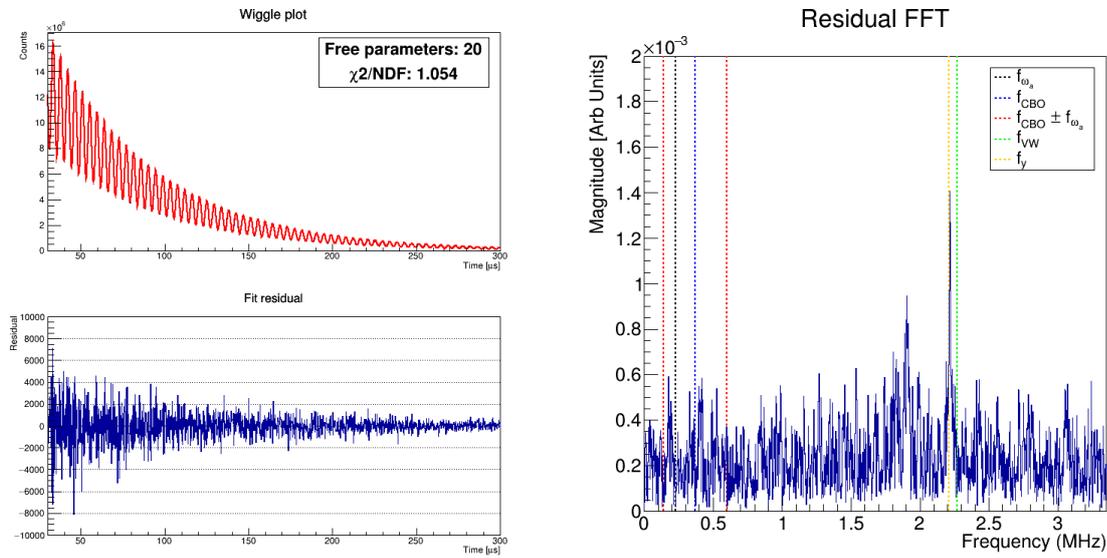
Figure 6.25: Run-2 wiggle plot fitted with the 20 parameter function *f20par* of Table 6.3.
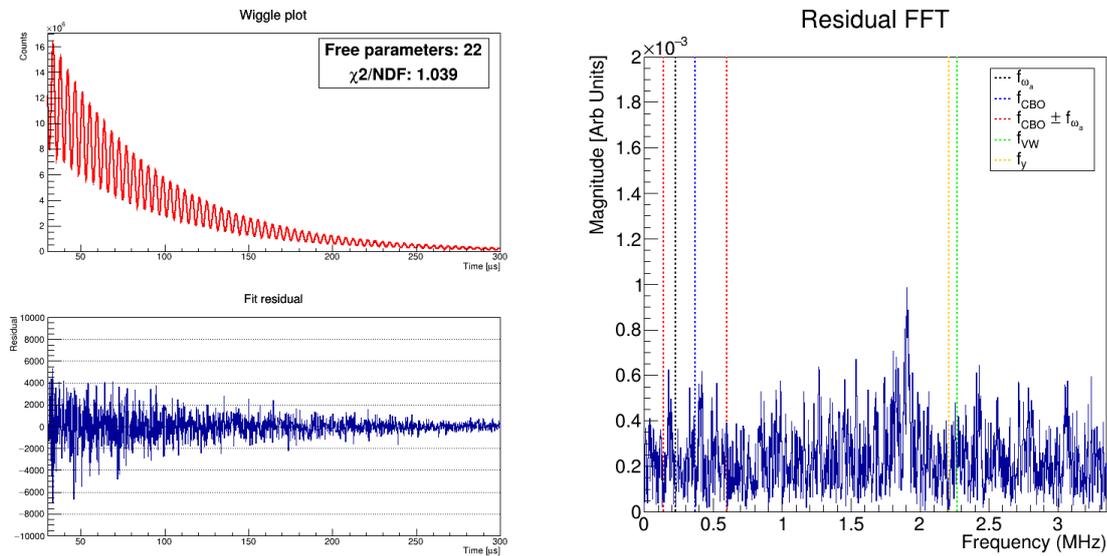


Figure 6.26: Run-2 wiggle plot fitted with the 22 parameter function *f22par* of Table 6.3.
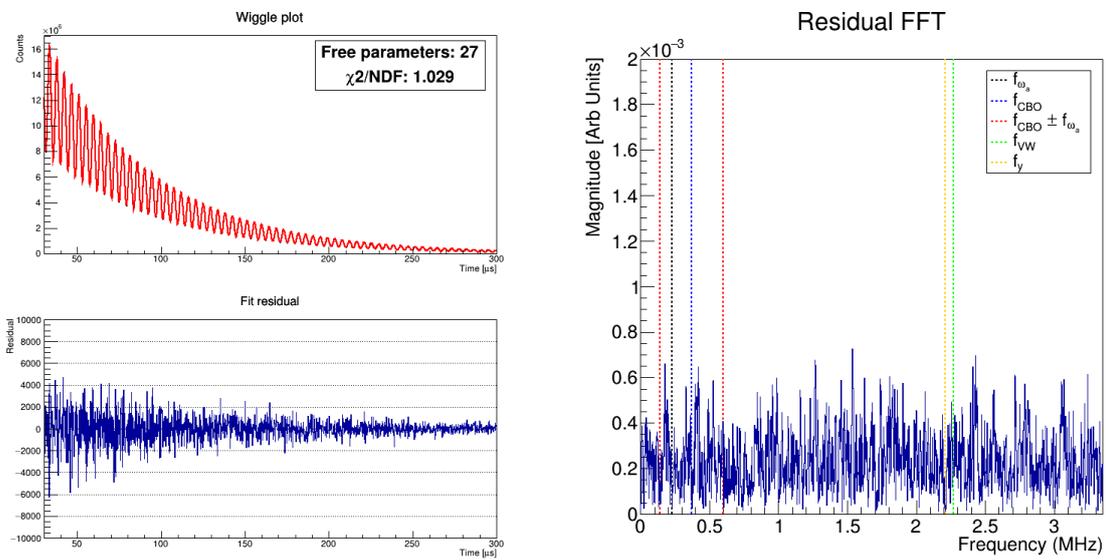
Figure 6.27: Run-2 AMethod wiggle plot fitted with the 27 parameter function *f27par* of Table 6.3. This is the most complete fitting function for Run-2 and Run-3. The $\chi^2/ndf$ value of $\sim$1, the flat residual, and the lack of any pronounced peak in the FFT plot are indications of a good fitting model.

# Chapter 7

# Results

This chapter presents the results of the anomalous precession frequency analysis performed on the data of Run-1, Run-2, and Run-3. The analysis presented here is conducted by the $\omega_a$-*Europe* group, a collaboration between the Italian (INFN) and UK groups. While the results of Run-1 have been released to the public in April 2021 [11, 68], the Run-2 and Run-3 analyses are still ongoing, with a planned publication in 2023. For this reason, the results are still blinded. Where possible, I will show the preliminary values, and I will focus on the general improvements on the systematic uncertainties with respect to Run-1 introduced by the new *ReconITA* reconstruction, mostly developed by me and described in Chapter 5.

## 7.1 Blinded analysis

The $\omega_a$ analysis of the Muon $g-2$ Experiment is performed as blinded, meaning that the measured value of $\omega_a$ is artificially shifted with respect to the true value. This is done to prevent unconscious biases of the analyzers when processing and analyzing the data. The blinding is applied in two different ways, one at the hardware level, and one at the software level.

### 7.1.1 Hardware blinding

The clock timings of the detectors are centrally provided by a 10 MHz Rubidium *master* clock. The Rubidium source provides the short-period oscillation stability, while a GPS receiver provides the long term stability. The *master* clock generates the signal for two precision synthesizers (SRS SG-382) which generate the clock signals for the field and positron measurements respectively. The synthesizer providing the clock to the

calorimeters generates a 40 MHz signal which then gets up-scaled to the sampling rate of 800 MHz in the calorimeter front-ends. The 40 MHz frequency is purposely detuned in the range [39.997,39.999] MHz by two Fermilab staff members not part of the Muon $g - 2$ collaboration. The actual "secret" frequencies are noted in two sealed envelopes and the synthesizers are locked behind a security door. Upon completion of the $\omega_a$ and $\omega_p$ analyses, and after a unanimous decision to unblind the results, the secret frequency is revealed.

### 7.1.2   Software blinding

A second level of blinding is applied to the $\omega_a$ fitting procedure, in order to avoid biases arising from the comparison of the results between different analysis teams. The fit routine replaces the $\omega_a$ term with a dimensionless parameter $R$ defined as:

$$\omega_a = \omega_a^{ref} \cdot \left[ 1 + (R + \delta R) \cdot 10^{-6} \right] \,, \tag{7.1}$$

where $\omega_a^{ref} = 2\pi \cdot 0.2291$ MHz and $\delta R$ is a secret offset that differs for all the analysis groups. The offset is generated randomly in the $[-24, +24]$ ppm range by transforming a secret passphrase into a random seed with a `MD5` hash function. In the following sections, I will always refer to the blinded $R$ values when talking about the measured values of $\omega_a$.

## 7.2   Run-1 results

Run-1 corresponds to the first period of acquisition of physics data for the Muon $g - 2$ Experiment. As mentioned in Section 6.1, Run-1 is composed of four separate datasets, collected between March and July 2018, distinguished by the different storage ring conditions. In this section, the results of the $\omega_a$ analysis described in Chapter 6 for each of the four datasets are reported.

The fit functions used for the final fit are listed in table 7.1, while the fit start and stop times were listed in Table 6.4 of Section 6.5.5. Figure 7.1b shows the Fast Fourier Transform of the fit residual with the complete fitting model for dataset Run-1D. The absence of peaks corresponding to the beam characteristic frequencies indicates that the fitting model described in Chapter 6 is correct. Moreover, the absence of a peak at low frequencies indicates that the treatment of lost muons and pileup contamination is also

correct.

| Dataset | Fitting function | Parameters |
|---------|------------------|------------|
| Run-1A | $N(t) = Ne^{-t/\tau_\mu} \left[1 + A \cdot A_{BO}(t)\cos(\omega_a t - \phi + \phi_{BO}(t))\right] \cdot$ $\cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot \Lambda(t) \cdot N_{2CBO}(t)$ | 20 |
| Run-1B | $N(t) = Ne^{-t/\tau_\mu} \left[1 + A \cdot A_{BO}(t)\cos(\omega_a t - \phi + \phi_{BO}(t))\right] \cdot$ $\cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot \Lambda(t) \cdot N_{2CBO}(t)$ | 20 |
| Run-1C | $N(t) = Ne^{-t/\tau_\mu} \left[1 + A \cdot A_{BO}(t)\cos(\omega_a t - \phi + \phi_{BO}(t))\right] \cdot$ $\cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot \Lambda(t) \cdot N_{2CBO}(t) \cdot N_y(t)$ | 23 |
| Run-1D | $N(t) = Ne^{-t/\tau_\mu} \left[1 + A \cdot A_{BO}(t)\cos(\omega_a t - \phi + \phi_{BO}(t))\right] \cdot$ $\cdot N_{CBO}(t) \cdot N_{VW}(t) \cdot \Lambda(t) \cdot N_{2CBO}(t) \cdot N_y(t)$ | 23 |

Table 7.1: List of fitting functions for the Run-1 datasets.
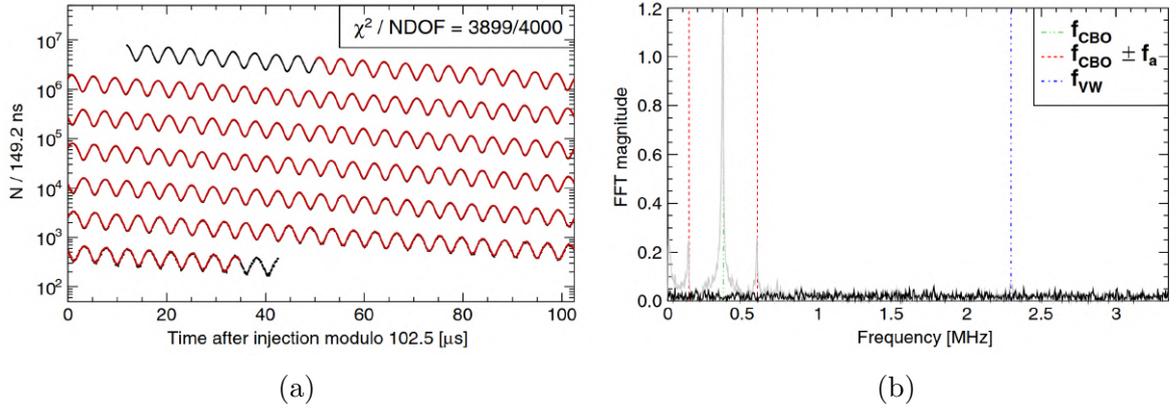


Figure 7.1: Wiggle fit (a) and Fast Fourier Transform of the fit residual (b) for the Run-1D dataset. The FFT spectrum shows no remaining peaks with respect to the 5-parameter fit (light gray), indicating an excellent performance of the fit model [68].

The final $\omega_a$-*Europe* central values of the main fit parameters for the four Run-1 datasets are listed in Table 7.2. The fit results are reported for one randomization seed only and for both the TMethod and AMethod techniques. The complete tables, which include all the fit parameters, and the correlation matrices, can be found in Appendix A.

The $\omega_a$ values are reported in their blinded form $R(\omega_a)$, but the same blinding offset is applied to all the four datasets. Note that the $R$ values need not to be the same, as the magnetic field $B$ may vary across the run, e.g. for seasonal temperature differences (see Section 3.6). The variation in the $\omega_{CBO}$ values is due to the different quadrupoles field indexes of $n = 0.108$ and $n = 0.120$ used throughout the run and listed in Table 6.1.

The TMethod and AMethod techniques use different positron energy thresholds as mentioned in Section 6.4 and therefore their results are related to slightly different, although highly correlated, samples. Using the simple approximation for a lower bound on the allowed difference for the Run-1D dataset:

$$\sigma_{\omega_a^T - \omega_a^A} = \sqrt{\sigma_{\omega_a^T}^2 + \sigma_{\omega_a^A}^2} = \sqrt{0.74^2 - 0.67^2} = 0.31 \text{ ppm}, \tag{7.2}$$

where the difference $|\omega_a^T - \omega_a^A|$ is 0.2 ppm, approximately $0.6 \cdot \sigma_{\omega_a^T - \omega_a^A}$, compatible with the hypothesis of a statistical fluctuation. The Neyman reduced chi-square, $\chi_{red}^2 = \chi^2/NDF = 1.009$ for the Run-1D AMethod, is consistent with unity within one standard deviation for all the four datasets and for both the TMethod and AMethod techniques:

$$\sigma(\chi_{red}^2) = \sqrt{\frac{2}{NDF}} \approx 0.02, \tag{7.3}$$

As mentioned in Section 6.4, the AMethod provides the highest possible statistical power for the determination of $\omega_a$. For this reason, the AMethod values have been chosen to extract the final value of $\omega_a$. Table 7.3 summarizes the blinded $\omega_a$ values for the $\omega_a$-*Europe* analysis obtained by averaging the AMethod values extracted with several randomization seeds. These values are the ones used in the final average for the computation of the $a_\mu$ result published in [11, 68]. The final, unblinded, results are presented in Section 8.2.

The next sections will present several consistency checks performed to confirm the stability and reliability of the extracted $\omega_a$ values. The anomalous precession frequency of the muons does not depend, within the allowed statistical deviations, on the specific detectors used for the analysis, nor on the time and energy selection over the $\omega_a$ fit is performed.

### 7.2.1   Start time scans

As discussed in Section 5.4, effects that cause the measured $g - 2$ phase $\phi$ to change within a fill introduce a direct bias to the extracted $\omega_a$ value. For example, pileup contamination and calorimeter gain fluctuations are two sources of time-dependent effects. In addition, anything that changes the measured positron energies will result in a time-dependent change in the effective energy threshold. A non corrected In-Fill gain, for example, would raise the effective energy threshold in the first microseconds of the fill. Since the $g - 2$ decay asymmetry varies with positron energy, as previously shown in

| Dataset | Parameter | Unit | T-Method | | A-Method | |
|---|---|---|---|---|---|---|
| | | | Value | Uncertainty | Value | Uncertainty |
| | $N$ | $-$ | 3594224. | 930. | 1617963. | 369. |
| | $\tau$ | $\mu$s | 64.4411 | 0.0037 | 64.4388 | 0.0033 |
| | $A$ | $-$ | 0.365025 | 0.000043 | 0.360934 | 0.000038 |
| Run-1A | $R(\omega_a)$ | ppm | $-24.21$ | 1.33 | $-23.74$ | 1.19 |
| | $\phi$ | rad | 2.09093 | 0.00022 | 2.09052 | 0.00020 |
| | $\omega_{CBO}$ | rad/$\mu$s | 2.33830 | 0.00044 | 2.33816 | 0.00035 |
| | $\chi^2/NDF$ | - | $4041.69/4135 = 0.977$ | | $4079.35/4135 = 0.987$ | |
| | $N$ | $-$ | 5002739. | 496. | 2251357. | 1038. |
| | $\tau$ | $\mu$s | 64.4275 | 0.0018 | 64.4275 | 0.0027 |
| | $A$ | $-$ | 0.365243 | 0.000036 | 0.360798 | 0.000032 |
| Run-1B | $R(\omega_a)$ | ppm | $-22.53$ | 1.12 | $-22.39$ | 1.01 |
| | $\phi$ | rad | 2.08329 | 0.00017 | 2.08272 | 0.00017 |
| | $\omega_{CBO}$ | rad/$\mu$s | 2.61524 | 0.00055 | 2.61497 | 0.00052 |
| | $\chi^2/NDF$ | - | $4072.97/4135 = 0.985$ | | $4199.43/4135 = 1.016$ | |
| | $N$ | $-$ | 7636215. | 2359. | 3433721. | 1009. |
| | $\tau$ | $\mu$s | 64.4390 | 0.0022 | 64.4378 | 0.0021 |
| | $A$ | $-$ | 0.366721 | 0.000030 | 0.361882 | 0.000026 |
| Run-1C | $R(\omega_a)$ | ppm | $-23.38$ | 0.91 | $-22.99$ | 0.81 |
| | $\phi$ | rad | 2.08073 | 0.00015 | 2.08020 | 0.00013 |
| | $\omega_{CBO}$ | rad/$\mu$s | 2.61026 | 0.00029 | 2.60999 | 0.00023 |
| | $\chi^2/NDF$ | - | $4182.25/4133 = 1.012$ | | $4235.98/4133 = 1.025$ | |
| | $N$ | $-$ | 15346879. | 9618. | 6878558. | 3322. |
| | $\tau$ | $\mu$s | 64.4419 | 0.0026 | 64.4415 | 0.0021 |
| | $A$ | $-$ | 0.370068 | 0.000024 | 0.364085 | 0.000021 |
| Run-1D | $R(\omega_a)$ | ppm | $-23.28$ | 0.74 | $-23.08$ | 0.67 |
| | $\phi$ | rad | 2.07614 | 0.00014 | 2.07557 | 0.00013 |
| | $\omega_{CBO}$ | rad/$\mu$s | 2.33594 | 0.00032 | 2.33591 | 0.00025 |
| | $\chi^2/NDF$ | - | $4048.39/4001 = 1.012$ | | $4036.15/4001 = 1.009$ | |

Table 7.2: Best fit results for the main parameters for the Run-1 datasets. The fitting functions are the ones listed in Table 7.1. The results are from a single random seed.

Figure 6.10, the resulting effect would be a time-dependent asymmetry term $A(t)$.

A powerful way to check for any time dependence of the fit parameters is to perform a scan of the $\omega_a$ fit start time. As the fit model assumes that each parameter is constant,

| Dataset | $R(\omega_a)$ [ppm] | $\sigma_{\omega_a}$ [ppm] |
|---------|---------------------|---------------------------|
| Run-1A  | -23.9467            | 1.1925                    |
| Run-1B  | -22.4867            | 1.0107                    |
| Run-1C  | -23.0652            | 0.8136                    |
| Run-1D  | -23.2811            | 0.6672                    |

Table 7.3: Blinded $\omega_a$ best fit values for the AMethod $\omega_a$-*Europe* analysis. These results are the average of several hit randomization seeds. The reported uncertainty $\omega_{\omega_a}$ is the statistical one.

any observed drift with respect to the fit start time is indicative of some effect that has not been accounted for. While doing this check, it has to be noted that shifting the start time of the fit changes the number of positrons participating to the fit, due to the exponential muon decay time, thus modifying the statistical effects. For a parameter measured at different start times $t_1$ and $t_2$, with $t_2 > t_1$, the one-sigma allowed statistical drift is approximately given by [41]:

$$\sigma_{diff} = \sqrt{\sigma_2^2 - \sigma_1^2 \left( 2\frac{A_1}{A_2} \cos(\phi_1 - \phi_2) - 1 \right)}. \tag{7.4}$$

In a start time scan, $\sigma_2$ will be larger than $\sigma_1$ because the former is derived from a smaller dataset than the latter. Equation 7.4 defines a region in the fit start time scans where the parameters can vary due to statistical fluctuations. The stability of the fitted parameters implies that all the systematic effects with time constants $\ll \tau_\mu$ are handled correctly in the analysis and they do not introduce any bias to the results.

Some parameters, such as the lifetime, amplitude and phase of the vertical oscillation term $N_y(t)$, as well as the higher order terms $N_{2CBO}(t)$, $A_{BO}(t)$, and $\phi_{BO}(t)$ are difficult to fit at later start times due to their small amplitudes and fast decoherence times. The same goes for the muon loss normalization term $k_{LM}$. For this reason, they are kept fixed to the values found at the nominal start time. The start time scans have been produced for all the relevant parameters for all datasets. Figure 7.2 shows the fitted main parameters of interest for the start time scans of the Run-1C dataset.

All the parameters, and in particular the $\omega_a$ value, are found to be stable within the $1\sigma$ allowed statistical drift for the datasets Run-1A, Run-1B, and Run-1C. However, the start time scans of parameters $N$ and $\tau_\mu$ of dataset Run-1D show a significant deviation
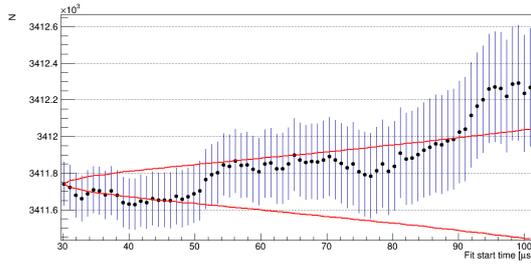
(a) Number of positrons $N$.
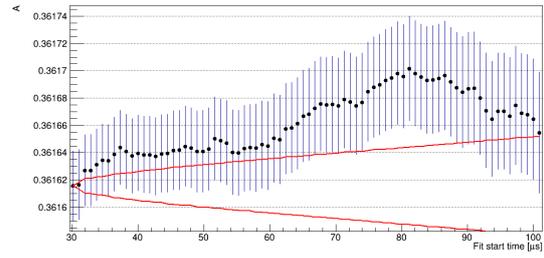
(b) Asymmetry $A$.
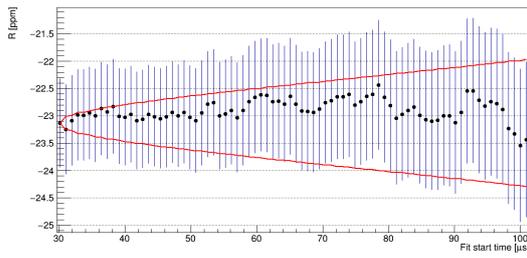
(c) Blinded $\omega_a$ value $R$.
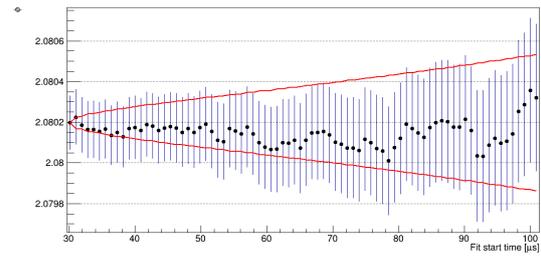
(d) Phase $\phi$.
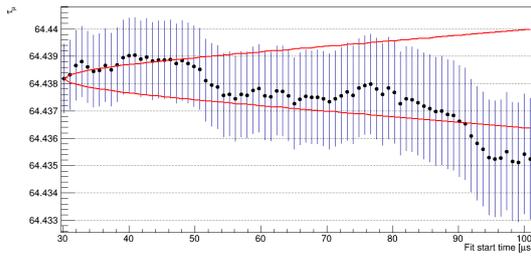
(e) Boosted muon lifetime $\tau_\mu$.
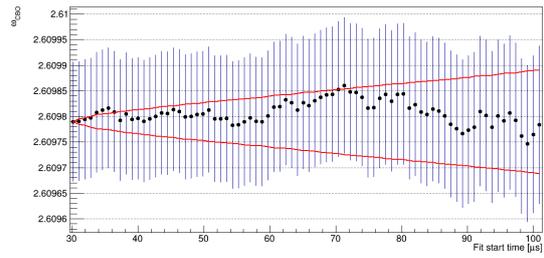
(f) CBO oscillation frequency $\omega_{CBO}$.

Figure 7.2: Fit start time scan for the Run-1C dataset. The fit parameters of the *f5par* and the CBO oscillation frequency $\omega_{CBO}$ are shown. The red lines represent the $1\sigma$ allowed statistical deviation from the nominal starting point of $t = 30$ $\mu$s.
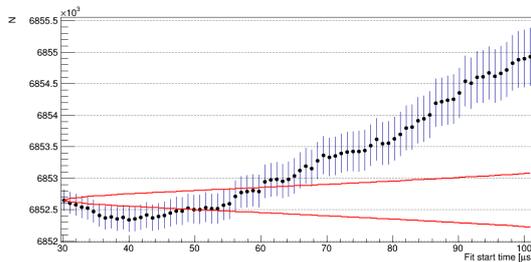


(a) Number of positrons $N$.
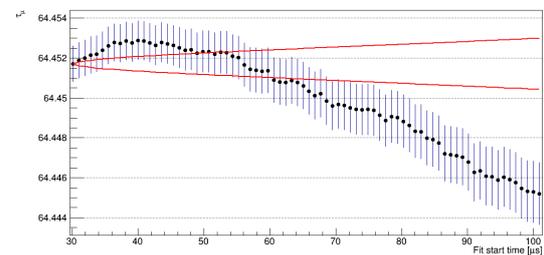
(b) Boosted muon lifetime $\tau_\mu$.

Figure 7.3: Fit start time scan for the Run-1D dataset. Only the $N$ and $\tau_\mu$ parameters are shown, which are the ones that significantly deviate away from the allowed $1\sigma$ bands (red lines).

from the $1\sigma$ bands (Figure 7.3). This deviation is caused by a residual early-to-late effect that will be discussed in the next section.

## 7.2.2   Energy binned scans

Another consistency check is the fit scan over different positron energies. Systematic effects on the detectors can cause energy and time dependent phase shifts that can bias the measured $\omega_a$ value at different energies, as anticipated in Section 5.4. Uncorrected detector-based systematic effects are expected to create biases on the order of 100 ppb [49], which are higher than the precision goal of the experiment. Beam dynamics effects can also cause energy dependent biases via the energy dependence of the calorimeter acceptance. Therefore, it is necessary to verify the consistency of the measured $\omega_a$ and the other parameters versus positron energy. In particular, energies toward the endpoint of the positron spectrum (3.1 GeV) are more sensitive to pileup or energy-related effects. The reason is that the relative impact of such effects is enhanced by the small number of positrons in high energy bins.

The energy scans are performed by slicing the pileup-corrected time-energy histogram (Figure 6.1a) into energy bins 30 MeV wide. Each slice is then individually fitted using the same function as the complete dataset. The same procedure is also used to extract the AMethod weighting function described in Section 6.4. Due to the low statistics of the highest energy bins, the fit end time is adjusted to $\approx 400\ \mu$s to avoid very low-populated bins. In addition, the fit routine is performed with the Maximum Likelihood method included in the `TMinuit` package [62]. Figure 7.4 shows the fitted main parameters of interest for the energy scan of the Run-1A dataset. Similar results are found for all Run-1 datasets. The fitted $\omega_a$ value is sufficiently stable at high energies where the shorter fit is performed.

The $g-2$ phase, $\phi$, shows a variable trend which can be explained by the fact that high energy positrons travel further in space before hitting a calorimeter than low energy positrons, because of the larger radius of curvature (Figure 7.5). The longer time of flight means that the high energy positrons observed by a calorimeter are generated further away than low energy positrons, thus from muons of a different phase. Complementary, the calorimeter acceptance depends on the positron energy [53].

The muon loss scaling parameter, $k_{LM}$, shows an unexpected deviation at high ener-

(a) Number of positrons $N$.



(b) Asymmetry $A$.



(c) Blinded $\omega_a$ value $R$.



(d) Phase $\phi$.



(e) Boosted muon lifetime $\tau_\mu$.



(f) Muon loss normalization $k_{LM}$.
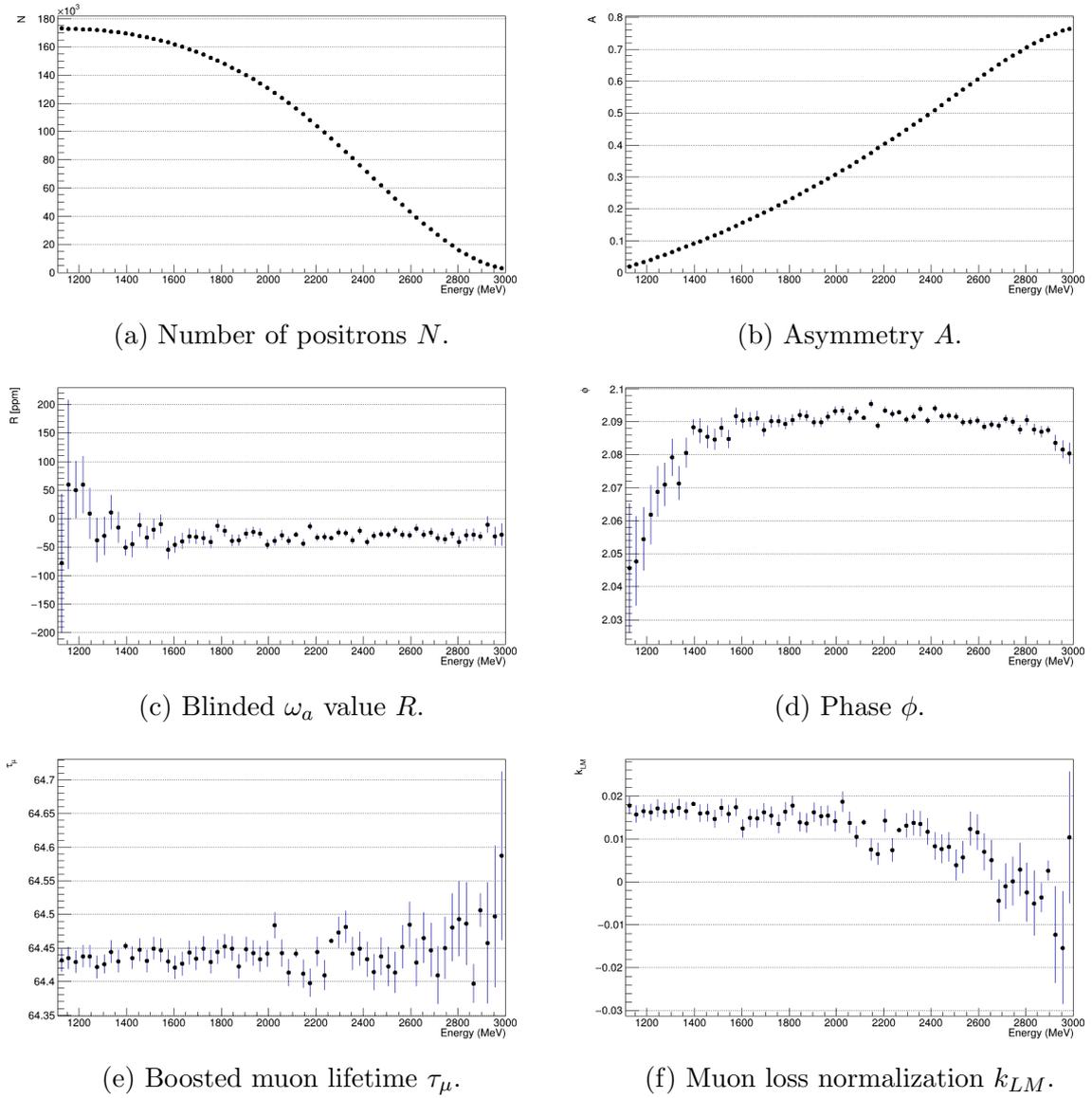
Figure 7.4: Energy binned scan for the Run-1A dataset. The fit parameters of the *f5par* function and the muon loss normalization $k_{LM}$ are shown.

gies. In fact, the parameter is introduced to scale the lost muons function $L(t)$ extracted empirically to the actual number of muons that are lost before decaying. By construction, the parameter should not depend on the positron energies and should always be positive; a negative value would mean that muons are *added* to the storage ring during the fill, which is not physical. Figure 7.4f shows that $k_{LM}$ is not stable across the energy spectrum and deviates to negative values for high energy positrons. This is a symptom that could indicate an imperfect gain correction, a time dependence of the lost muons acceptance for the empirical extraction of $L(t)$, a rate dependent difference in the
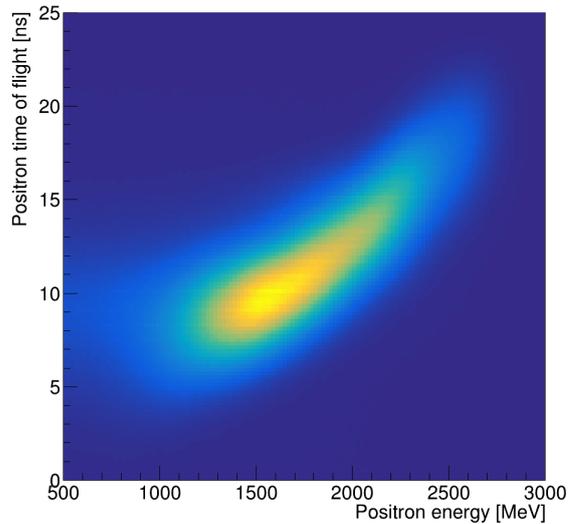
Figure 7.5: Time of flight between the decay vertex and the calorimeter hit versus positron energy. High energy positrons travel longer than lower energy positrons because of the larger radius of curvature.

way positrons are reconstructed, or a combination of these and other unknown effects. Moreover, the same residual systematic effect can account also for the trends in the start time scans shown in Figure 7.2. The nature of this residual effect is not yet completely understood, although the current Run-2 and Run-3 analyses point toward the positron reconstruction as the main contribution. For the Run-1 analysis, the effect is treated as an additional systematic uncertainty for the final value of $\omega_a$, and will be discussed in Section 7.3.4.

### 7.2.3   Calorimeter fits

The analysis presented in this Chapter was conducted on the wiggle plot containing the contribution from all the 24 calorimeters. It is possible to fit each calorimeter individually for additional consistency checks. Individual calorimeters are more sensitive to the motion of the beam with respect to the combined data, as close-orbit beam dynamics effects average out when integrating over the full azimuth angle. Some parameters, such as $\tau$ and $k_{LM}$ are expected to be the same across all calorimeters (Figure 7.6), while the beam oscillation parameters and $A$, $\phi$, and $N$ may differ depending on the different materials present in front of the calorimeters. Imperfect per-calorimeter energy calibrations could also have an effect, and, as mentioned in Section 6.2, the bunched structure of the

beam produces an effect on the fitted $\omega_a$ measured by single calorimeters. The latter is mitigated by the hit randomization technique.

Figure 7.7 shows the fitted $\omega_a$ value for each calorimeter compared to the fit obtained with all calorimeter combined. The individual fits, as well as their average, are in excellent agreement with the combined result. The parameter consistency between calorimeters validates the treatment of beam oscillations, particularly the CBO and the VW terms.



(a) Boosted muon lifetime $\tau_\mu$.     (b) Muon loss normalization $k_{LM}$.
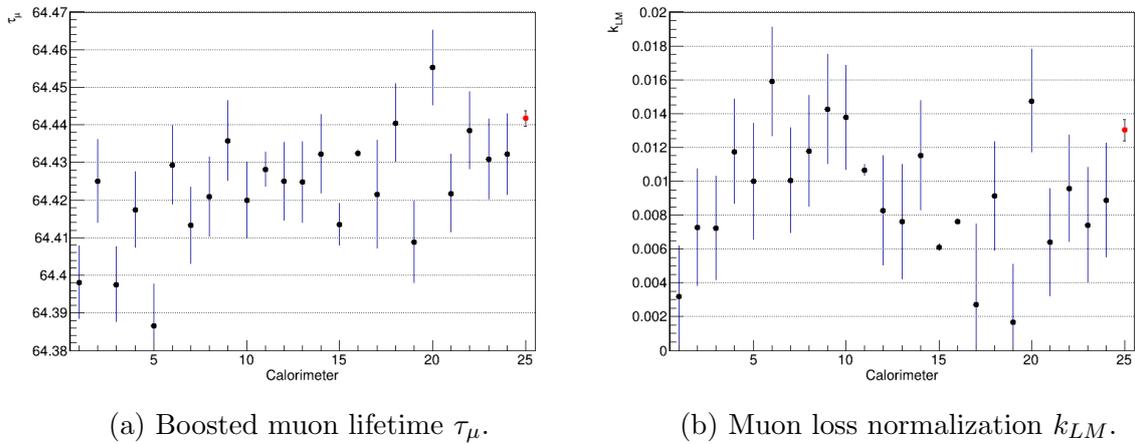
Figure 7.6: Calorimeter scan of the Run-1D dataset. The two parameters $\tau_\mu$ and $k_{LM}$ show sufficient agreement between the independent calorimeters. The nominal value (red point on the right) is shown for comparison.

## 7.2.4 Long term checks

Finally, an additional consistency check is performed to search eventual long-term fluctuations due to environmental factors. Temperature fluctuation of the hall, as well as long term instabilities of the magnet and beam conditions, could have a detectable effect on the $\omega_a$ result. The Run-1D dataset, which is Run-1 biggest one, was sliced into multiple independent subsets according to different criteria:

- Long-term fluctuations: the dataset is sliced on a per-day basis, where each subset starts at 7 a.m. Fermilab time and ends at 7 a.m. of the next day;

- Daily fluctuations: the same subsets of the previous item are further divided into day (7 a.m. - 7 p.m.) and night (7 p.m. - 7 a.m.) slices;

(a) Run-1A

(b) Run-1B

(c) Run-1C
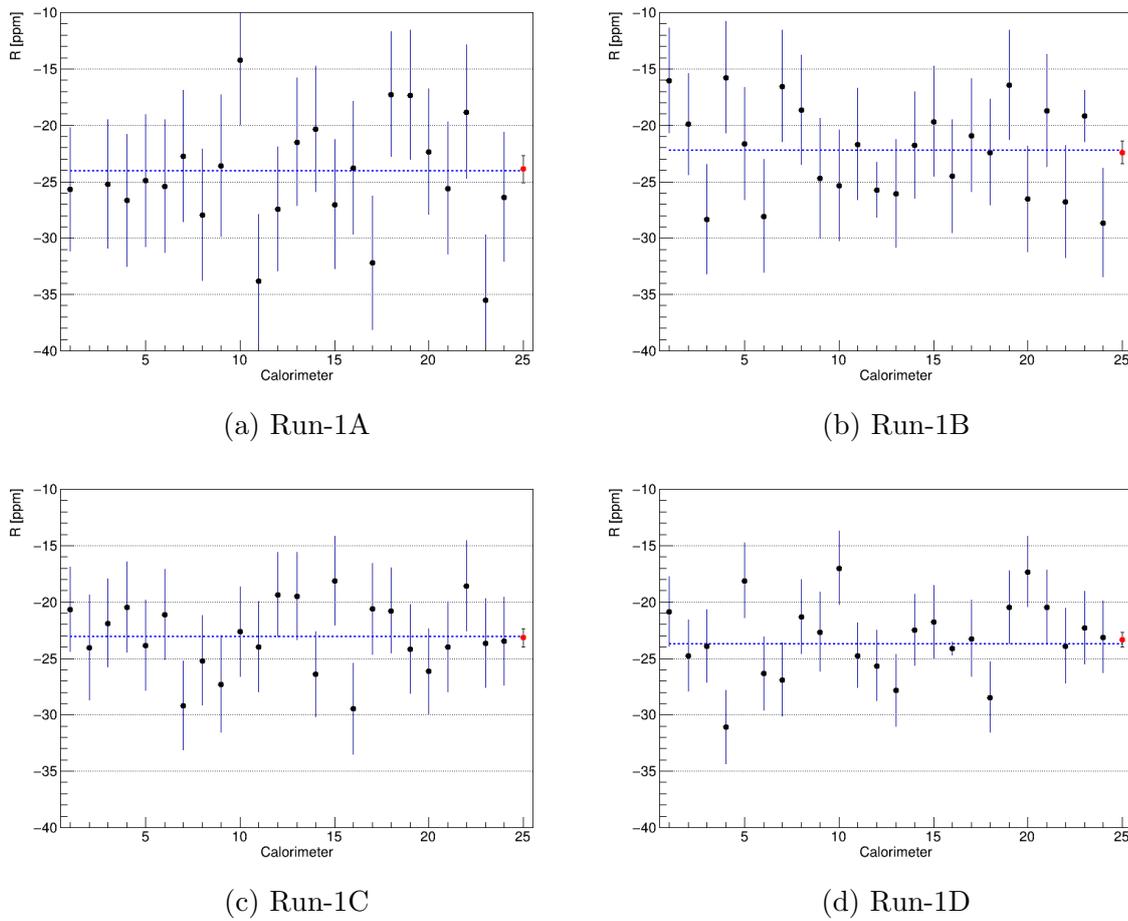
(d) Run-1D

Figure 7.7: Blinded $\omega_a$ values for the calorimeter scans of the four Run-1 datasets. The points are obtained from AMethod fits for each calorimeter separately. The average of the 24 calorimeters (blue dashed line) is compared to the nominal result (red point on the right).

- Magnet measurements: the dataset is sliced in sections bounded by the *trolley* measurements of the magnetic field described in Section 2.4;

Figure 7.8 shows the fitted $\omega_a$ for each subset compared to the nominal fits. The individual fits, as well as their average, are in excellent agreement with the combined result, without showing any trend as a function of the acquisition time. This adds further confidence on the long-term gain corrections applied to the reconstructed data described in Section 5.4.3.

(a) Full day fits



(b) Half day fits
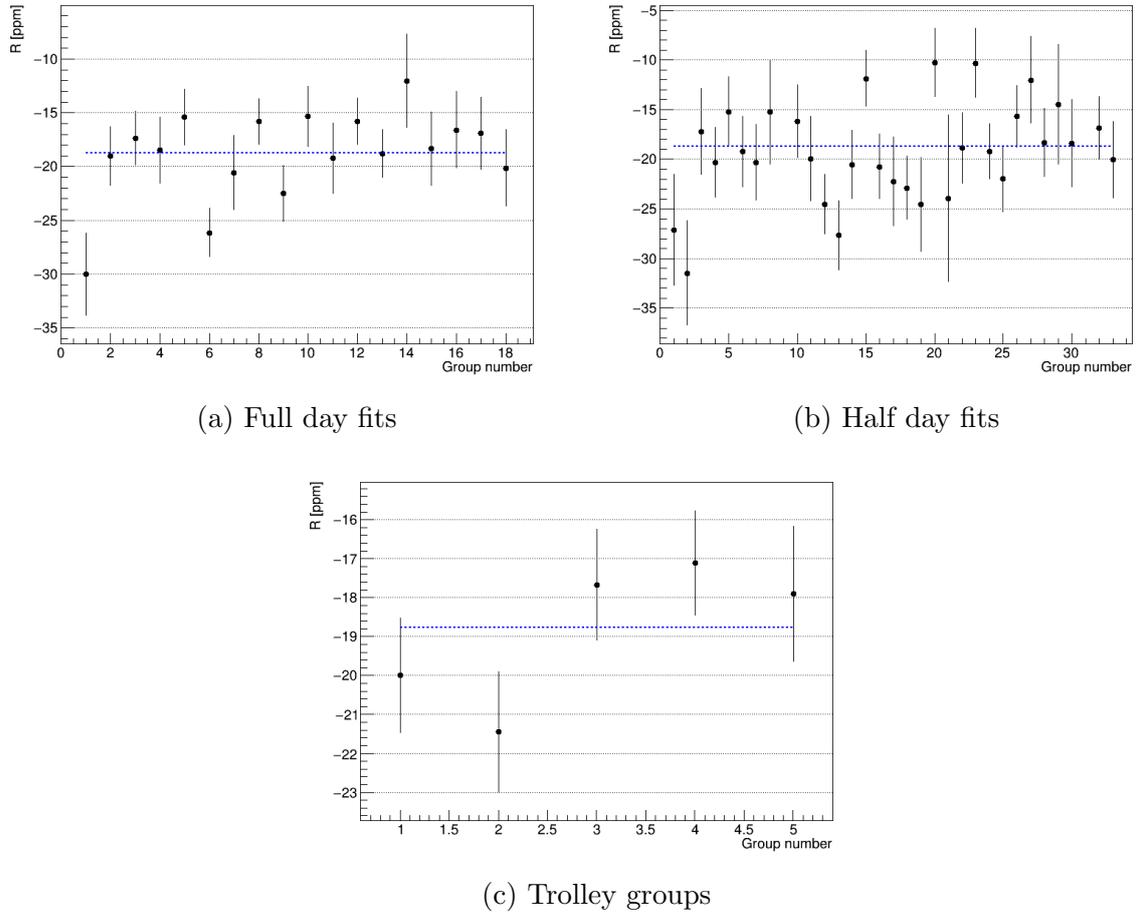


(c) Trolley groups

Figure 7.8: Blinded $\omega_a$ values for the long-term scans of the Run-1D dataset. The data is separated into groups depending on the day (a), the time of the day (b) and the trolley measurements (c). The average is shown as a blue dashed line.

## 7.3    Evaluation of systematic uncertainties

The previous sections presented the central values for the Run-1 anomalous precession frequency analysis, together with the statistical uncertainties. This section will give an evaluation of the systematic uncertainties related to the measurement of $\omega_a$. While, as we will see, the Run-1 measurement uncertainty is completely dominated by the statistical error, it is of crucial importance to determine and evaluate all the possible biases that can affect the measured value of $\omega_a$. The final goal of the Muon $g - 2$ Experiment is to contain the systematic uncertainties of $\omega_a$ under 70 parts per billion.

Sources of systematic errors are the gain corrections, the pileup subtraction, the phase acceptance of the detectors, the beam dynamics model, and the electric field. Finally,

any uncorrected effect that arises from the internal consistency checks must be evaluated.

In this section I will mainly focus on the systematic uncertainties related to the reconstruction and the gain corrections of the positron data. For each systematic study, in order to compare the results from different histograms, the hit time randomization described in Section 6.2 has not been applied unless explicitly specified.

## 7.3.1 Gain corrections

As anticipated in Section 5.4, a high gain stability of the detectors, especially in the timescale within a single fill, is required. Any time-dependent change in the energy scale of the calorimeter will introduce a time dependency of the energy acceptance. Therefore, any uncorrected fluctuation could result in a direct bias in the measurement of $\omega_a$.

The long-term gain correction described in Section 5.4.3 does not have a direct effect on $\omega_a$ as it is applied on much longer timescales than the 700 $\mu$s of the measurement period. However, the consistency checks shown in Section 7.2.4 demonstrate the reliability of such correction.

The technique employed to evaluate the gain systematics is to re-apply the gain corrections scaled by a certain multiplying factor $\alpha$. When the factor is $\alpha = 0$ it is equivalent as not applying the gain corrections at all, while for $\alpha = 1$ the nominal corrections are applied. The difference between these two scenarios gives the overall $\omega_a$ sensitivity to the gain corrections. A scan over values close to $\alpha = 1$ will give the final estimated systematic uncertainty according to the estimated precision of the gain correction parameters.

**In-Fill gain correction**

The In-Fill gain correction described in Section 5.4.2 applies a correction to the SiPM energies of the form:

$$g_{IFG}(t) = \frac{1}{1 - Ae^{-t/\tau}} \, , \tag{7.5}$$

in order to correct for the gain sag caused by the splash of particles after beam injection.

In order to evaluate the systematic effects of this correction a scan of the amplitude

parameter $A$ is performed. For each crystal, the actual gain correction is removed by applying the inverse function to the crystal energy, and then a new gain correction, with an amplitude scaled by an $\alpha$ multiplier, is applied instead:

$$E' = E\frac{g'_{IFG}(t)}{g_{IFG}(t)} = E\frac{1 - Ae^{-t/\tau}}{1 - \alpha Ae^{-t/\tau}} \tag{7.6}$$

The new positron energy is computed by summing all the modified crystal energies of the original cluster. The clustering algorithm is not re-applied at this stage. After recalculating all the positron energies, the pileup correction routine is run again and the modified wiggle plot is built. The same procedure is replicated over a range of scaling factors from $\alpha = 0$ to $\alpha = 2$ in steps of 0.25.

Each wiggle plot is then fitted with the standard routine described in Section 6.5.
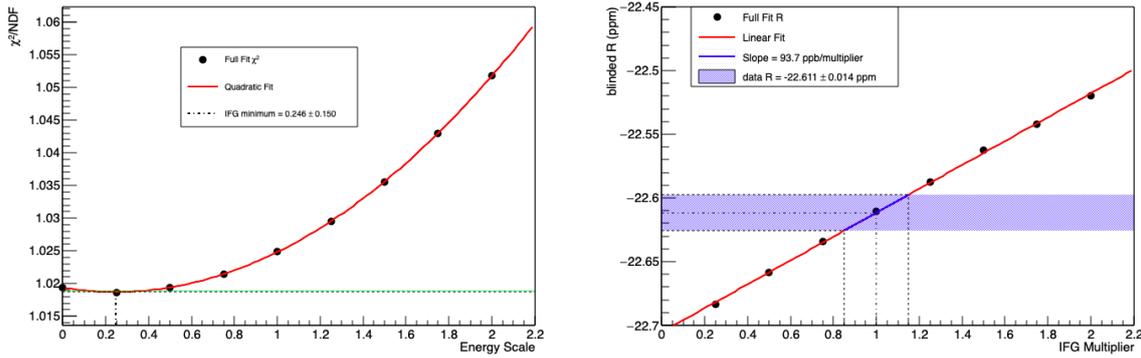


Figure 7.9: In Fill Gain amplitude scan: on the left the $\chi^2/NDF$ as a function of the $\alpha$ multiplier, where the green band represents a $\chi^2_{min} + \chi^2_{Unit}$ variation. On the right the fitted $R$ value. The blue band represents a change in $R$ corresponding to one $\chi^2$ unit variation around the expected minimum.

Figure 7.9 shows the fit reduced chi-square $\chi^2_{red}$ and the blinded $\omega_a$ value as a function of the multiplier factor $\alpha$ for the Run-1D dataset. The multiplier giving the minimum $\chi^2$ value is expected to be $\sim 1$, since $\alpha = 1$ is the nominal correction as measured by the laser system. A value different than 1 may suggest that other effects are involved in this gain terms. Similar results are found for all the datasets in Run-1. The $\chi^2$ value, however, might be affected by some of the 22 parameters of the fitting function (Table 7.1) that could change in order to balance out the effect of a wrong In-Fill gain correction.

For this reason, the systematic uncertainty on the In-Fill gain amplitude is extracted

by multiplying $\omega_a$ sensitivity to the correction with the average uncertainty of the amplitude $A$ parameter as measured by the laser system:

$$\sigma_{IFG} = \frac{dR}{d\alpha}\langle \sigma_A^{IFG}\rangle. \tag{7.7}$$

As described in section 5.4.2, the In-Fill gain lifetime, $\tau$, is determined from the long-term double pulse (LTDP) studies. This allows to measure the lifetimes for each crystal with a precision of $\sim 2\%$. Since the time constant variation can be linearly related to the amplitude variation when the former is small, all the IFG systematic contribution is assumed to be inside the amplitude parameter variation and the IFG time constant uncertainty is set to be zero. Table 7.4 summarizes the systematic uncertainties extracted for Run-1.

|  | $dR/d\alpha$ [ppb] | | $\sigma_{IFG}$ [ppb] | |
|---|---|---|---|---|
| Dataset | TMethod | AMethod | TMethod | AMethod |
| Run-1A | 110.4 | 100.1 | 23.6 | 21.4 |
| Run-1B | 102.5 | 99.6 | 6.7 | 6.5 |
| Run-1C | 65.0 | 59.4 | 9.9 | 9.1 |
| Run-1D | 4.2 | 104.1 | 0.6 | 3.9 |

Table 7.4: Summary of the Run-1 sensitivities and systematic uncertainties related to the In-Fill gain correction for the TMethod and AMethod.

**Short-Term gain correction**

The Short-Term gain correction described in section 5.4.1 is of the following form:

$$g_{STDP}(t) = \frac{1}{1 - E_1 A e^{-\Delta t/\tau}}, \tag{7.8}$$

where $E_1$ is the energy of the hit preceding the one for which the STDP correction is applied, and $\Delta t$ depends on both hits. For this reason, it is not trivial to re-scale the correction on the reconstructed data.

In order to evaluate the systematic effect of the correction, each dataset has been produced again without applying the STDP correction. This is equivalent to having the $\alpha$ scaling parameter set to $\alpha = 0$. The difference $\Delta\omega_a = \omega_a - \omega_a^{No-STDP}$ is taken as the $\omega_a$ sensitivity to the STDP correction. In analogy to the In-Fill gain, the systematic uncertainty is extracted by multiplying $\Delta\omega_a$ with the average uncertainty $\langle \sigma_A^{STDP}\rangle$ of

the amplitude parameter as measured by the laser system. The results for the Run-1 datasets are reported in Table 7.5. The systematic is very small and mostly < 1 ppb.

Similarly to the IFG systematics, since a small variation in the time constant can be related to a variation in the amplitude parameter, it has been assumed that all the STDP systematic contribution lies in the amplitude parameter.

|  | $\sigma_{STDP}$ [ppb] | |
| --- | --- | --- |
| Dataset | TMethod | AMethod |
| Run-1A | 0.1 | 0.1 |
| Run-1B | 0.1 | 0.1 |
| Run-1C | 0.3 | 0.2 |
| Run-1D | 1.2 | 0.1 |

Table 7.5: Summary of the Run-1 systematic uncertainties related to the Short-Term gain correction for the TMethod and AMethod.

## 7.3.2   Pileup correction

To determine the systematic uncertainty related to the pileup correction, a series of measurements were performed varying the pileup parameters discussed in section 6.3.

**Pileup amplitude**

In order to study the systematics related to the overall magnitude of the pileup correction, a scan similar to the gain correction studies described in Section 7.3.1 was performed. The pileup correction for Run-1 is of the form (Equation 6.6):

$$N(t, E) = N'(t, E) + \alpha(S_D(t, E) - D(t, E)),$$ (7.9)

where the optimal value for the scaling factor $\alpha$ is determined from a scan of the parameter in a small range around 1, typically $0.9 \leq \alpha \leq 1.1$. For each scaling parameter value, a wiggle plot is built and then fitted with the full fitting model. The fit reduced chi-square, $\chi^2/NDF$, is plotted as a function of the scaling value $\alpha$ and the data points are fitted with a quadratic function in order to find the minimum. The $\alpha$ value used for the actual pileup correction is the one that gives the minimum $\chi^2/NDF$ value.

The same routine is used to extract the related systematic uncertainty, shown in Figure 7.10 for the Run-1D dataset. By calling $\delta_\alpha$ the shift in $\alpha$ which increases the $\chi^2$ value by one unit, the systematic uncertainty related to this parameter is defined to be the sensitivity of $\omega_a$ to the pileup scale multiplied by $\delta_\alpha$. The $\omega_a$ sensitivity $dR/d\alpha$ is extracted with a linear fit as shown in Figure 7.10. Table 7.6 summarizes the uncertainties for the four Run-1 datasets.
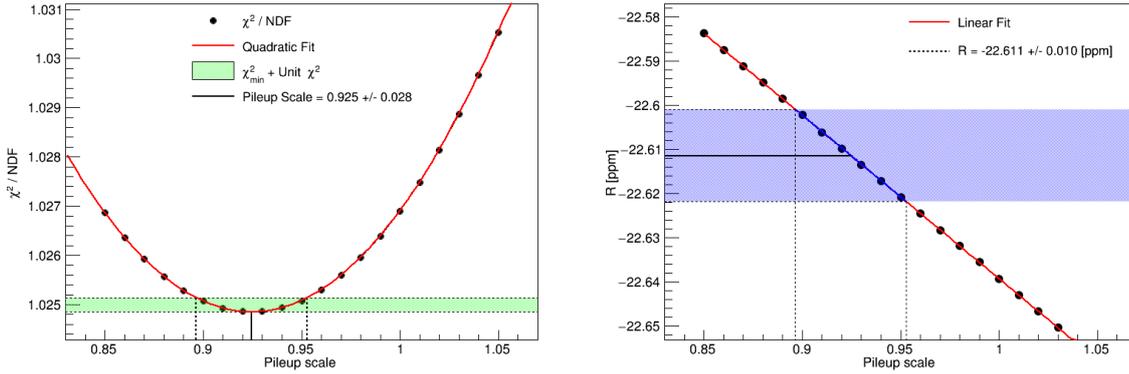


Figure 7.10: $\chi^2/NDF$ (left) and R value (right) for each pileup scale factor for dataset Run-1D. The green band represents a unit of $\chi^2$ variation from the minimum value and the blue band the corresponding $\omega_a$ variation.

| Dataset | $dR/d\alpha$ [ppb] | | $\delta_\alpha$ | | $\sigma_{PU}^\alpha$ [ppb] | |
|---------|---------|---------|---------|---------|---------|---------|
|         | TMethod | AMethod | TMethod | AMethod | TMethod | AMethod |
| Run-1A  | 334.2   | 461.9   | 0.0516  | 0.0351  | 17.2    | 16.2    |
| Run-1B  | 147.6   | 495.9   | 0.0450  | 0.0283  | 6.6     | 14.0    |
| Run-1C  | 136.0   | 391.6   | 0.0387  | 0.0249  | 5.3     | 9.8     |
| Run-1D  | 233.2   | 463.7   | 0.0445  | 0.0190  | 10.4    | 8.8     |

Table 7.6: Summary of the Run-1 sensitivities and systematic uncertainties related to the pileup correction $\alpha$ parameter for the TMethod and AMethod.

## Triple pileup contamination

The $\omega_a$-*Europe* pileup correction of Run-1 does not include the effect of triple pileup events. As the probability of a double coincidence is $\approx 0.01$, the expected triple pileup contamination in the data is a factor $\approx 100$ smaller than the one from the double pileup. The systematic uncertainty for not considering the triple pileup contamination is estimated by scaling the $\omega_a$ sensitivity to the double pileup correction, i.e. the $dR/d\alpha$

measured by the scans, by a factor 0.01. The bias on $\omega_a$ is small and summarized in Table 7.7 for each of the four Run-1 datasets.

| Dataset | $\sigma_{PU}^{Triples}$ [ppb] | |
|---|---|---|
| | TMethod | AMethod |
| Run-1A | 3.3 | 4.6 |
| Run-1B | 1.5 | 5.0 |
| Run-1C | 1.4 | 3.9 |
| Run-1D | 2.3 | 4.6 |

Table 7.7: Run-1 systematic uncertainties due to the triple pileup contamination in the data.

**Pileup cluster energy**

A source of systematic uncertainty related to the pileup correction is the energy assignment to the doublet event constructed with the shadow method described in Section 6.3.1. When a double coincidence is found, the associated pileup doublet is constructed by summing the energies of the two single positrons (Equation 6.3b). However, when the reconstruction operates on a true pileup event it could result in a different outcome depending on the energy and time separation of the two underlying positrons.

To assign an uncertainty to this effect, an artificial multiplier is applied to the doublet's energy:

$$E_{doublet} = k \cdot (E_1 + E_2). \tag{7.10}$$

A scan over the $k$ multiplier is performed in an interval of $\pm 0.1$ around the nominal value $k = 1$. Similarly to the pileup amplitude scaling, the systematic uncertainty for this parameter is defined to be the variation of $\omega_a$ in the interval identified by a $\chi^2$ unit variation around the minimum.

Figure 7.11 shows the scan performed for the Run-1D dataset. The minimum of the $\chi^2/NDF$ scan is compatible with $k = 1$, indicating that the pileup correction is sufficiently accurate. The sensitivity of the fitted $\omega_a$ to the energy scale factor, the slope $dR/dk$, is found by fitting the right plot of Figure 7.11 with a linear function. Table 7.8 summarizes the systematic uncertainty extracted for each of the four Run-1 datasets.
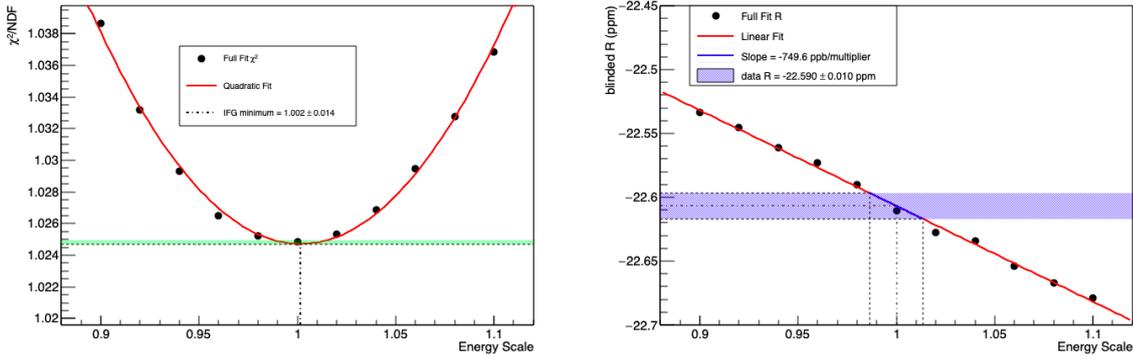
Figure 7.11: Scan of the energy scale parameter in the Run-1D dataset. The green band represents one $\chi^2$ unit variation above the minimum and the blue band corresponds to the $\pm\sigma$ uncertainty on the $R$ value. Similar uncertainties are found for all the datasets in Run-1.

| Dataset | $dR/dk$ [ppb] | | $\delta_k$ | | $\sigma_{PU}^E$ [ppb] | |
|---|---|---|---|---|---|---|
| | TMethod | AMethod | TMethod | AMethod | TMethod | AMethod |
| Run-1A | 485.8 | 389.2 | 0.024 | 0.018 | 11.7 | 7.0 |
| Run-1B | 621.6 | 618.5 | 0.018 | 0.013 | 11.2 | 8.0 |
| Run-1C | 100.9 | 544.6 | 0.017 | 0.012 | 1.7 | 6.5 |
| Run-1D | 383.6 | 660.8 | 0.022 | 0.001 | 8.4 | 7.0 |

Table 7.8: Summary of the Run-1 sensitivities and systematic uncertainties related to the pileup energy scale for the TMethod and AMethod.

**Pileup cluster time**

When a pileup doublet is constructed from a shadow window coincidence, the resulting doublet assigned time is (Equation 6.3a):

$$t_{doublet} = \frac{t_1 E_1 + (t_2 - t_{gap})E_2}{E_1 + E_2} + \frac{t_{gap}}{2} \,, \tag{7.11}$$

where, for Run-1, $t_{gap} = T_c$. The last term, $+\frac{t_{gap}}{2}$, is introduced to account for the different pileup rate between the trigger and the shadow window.

This time is assigned both to the doublet and to the singlets when building the pileup histograms. The $\omega_a$ sensitivity related to the pileup assigned time is evaluated with a scan performed by shifting $t_{doublet}$ by an offset $t_S$ in the range $\left[-\frac{T_c}{2}, \frac{T_c}{2}\right]$.

The systematic uncertainty for the pileup cluster assigned time is set to be the change

in the measured $\omega_a$ corresponding to a variation:

$$t_S = \pm \frac{t_W}{\sqrt{12}}, \tag{7.12}$$

where $t_W = 5$ ct is the shadow window width, and where $1/\sqrt{12}$ is the variance of an uniform distribution of width 1.

The result for the Run-1D dataset is shown in figure 7.12. Table 7.9 summarizes the systematic uncertainties extracted for each of the four Run-1 datasets.
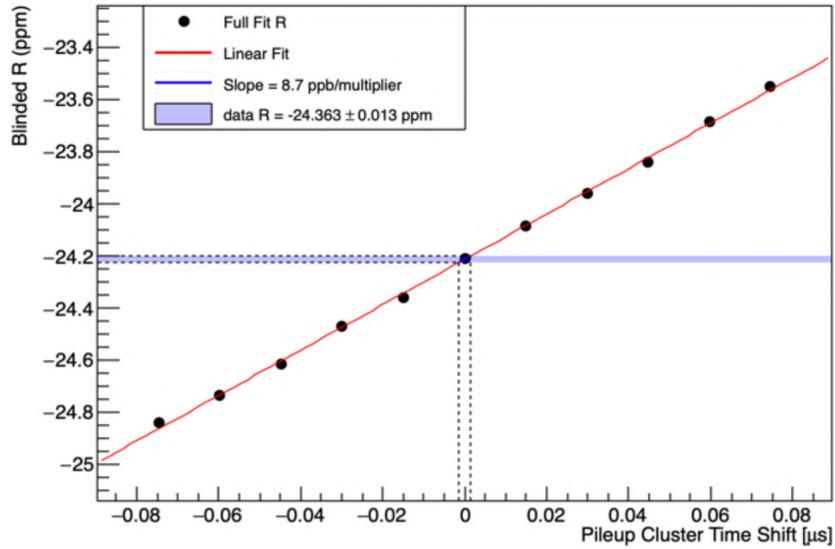


Figure 7.12: The fitted R values for a range of different time shift values. The dotted vertical lines correspond to the $\pm \frac{5}{\sqrt{12}}$ ct interval, and the horizontal dotted lines show consequent effect on R.

| Dataset | $dR/dt_S$ [ppb/ct] | | $\sigma_{PU}^t$ [ppb] | |
| --- | --- | --- | --- | --- |
| | TMethod | AMethod | TMethod | AMethod |
| Run-1A | 8.66 | 40.82 | 12.5 | 58.9 |
| Run-1B | 9.85 | 44.45 | 14.2 | 64.2 |
| Run-1C | 8.54 | 39.62 | 12.3 | 57.2 |
| Run-1D | 5.21 | 32.91 | 7.5 | 47.5 |

Table 7.9: Summary of the Run-1 sensitivities and systematic uncertainties related to the pileup cluster time model for the TMethod and AMethod.

**Unseen pileup events**

An additional systematic associated to the pileup correction is related to very low energy positrons. The reconstruction of the calorimeter events applies a low energy threshold of $\sim 50$ MeV. Lowering this threshold further would result in a high number of fake pulses extracted from the noise of the SiPM traces. When a low energy ($< 50$ MeV) positron hits a calorimeter at the same time of a higher energy positron in a pileup event, the resulting pileup event will be registered nonetheless. The shadow method is not capable of correcting for these events, as both the trigger and shadow positrons must be $> 50$ MeV in order to be detected. The pileup events that cannot be corrected with the shadow method technique, for this reason, are named *unseen* pileup events.

In order to study this effect and to assign a systematic uncertainty, an artificial threshold $E_{th} = 100$ MeV has been imposed to the trigger and shadow window clusters, so that clusters with $E < E_{th}$ are not used to build the pileup correction.
As the artificial threshold is twice the one of the calorimeter reconstruction, the systematic uncertainty is extracted as the difference between the nominal $\omega_a$ value and the one measured with the threshold applied. It is assumed that the difference is the same one between the nominal $\omega_a$ value and the one that would be measured if the detector energy acceptance could reach 0 MeV. The measured systematic effect is small, and the results for all the Run-1 datasets are summarized in Table 7.10.

| Dataset | $\sigma_{PU}^{Unseen}$ [ppb] | |
|---------|---------|---------|
|         | TMethod | AMethod |
| Run-1A  | 5.3     | 1.1     |
| Run-1B  | 0.5     | 2.9     |
| Run-1C  | 0.6     | 1.1     |
| Run-1D  | 0.7     | 0.8     |

Table 7.10: Run-1 systematic uncertainties due to the unseen pileup contamination in the data.

**Artificial Dead Time**

The Artificial Dead Time (ADT) parameter described in Section 6.3.2 was introduced to equalize the clustering performance on pileup separation to a box function of size $t_{ADT}$ and height 1. This way, the coincidences found in the window of the shadow method

match the pileup events more accurately.

However, there are two considerations to be made when applying such parameter. First, as the clustering performance separates fewer pileup events, the magnitude of the pileup correction increases and so does the related systematic errors. Then, the clustering algorithm might not be well represented with a box function. The assumption that the pileup separation efficiency is exactly 1 for $\Delta t > t_{ADT}$ might not be true.

In order to estimate the sensitivity of $\omega_a$ to the choice of the $t_{ADT}$ parameter, several wiggle plots are produced with different $t_{ADT}$ values. In order for the pileup correction to work properly, the shadow window width $t_W = t_{ADT}$ is scaled accordingly. Each histogram is then fitted and the $\omega_a$ sensitivity extracted with a linear fit.

The result of this scan is shown in Figure 7.13. The histograms built with different $t_{ADT}$ values are partially correlated and an estimate of the uncorrelated statistical uncertainty of the fitted $\omega_a$ values is given by:

$$\sigma_{uncorr}^i \approx \sqrt{\sigma_i^2 - \sigma_{ref}^2}\,, \tag{7.13}$$

for the $i$-th data point, where $\sigma_{ref}$ is the first point $t_{ADT} = 3.75$ ns.
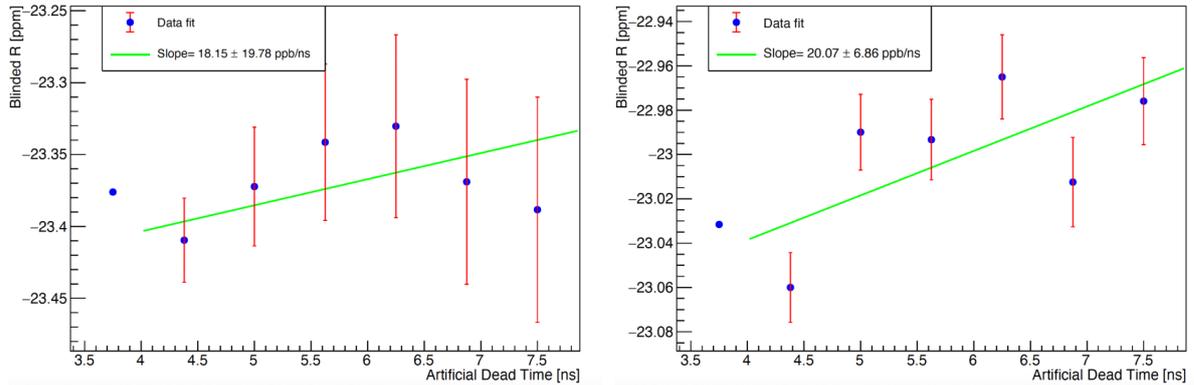


Figure 7.13: Scan of the Artificial Dead Time parameter for the Run-1C dataset for the TMethod (left) and AMethod (right). The uncertainties are the uncorrelated errors calculated according to Equation 7.13. The green line is a linear fit, and the first point is not included since it has zero error. Courtesy of M. Sorbara.

The systematic uncertainty is evaluated as the variation in the measured $\omega_a$ for an ADT shift of $\pm 1.25$ ns ($= 1$ ct) around the nominal value of $t_{ADT} = 5$ ct. This shift

of ADT values corresponds to the reasonable range of choices based on the clustering algorithm time partitioning parameters described in Section 5.5.

The sensitivities found in this way appear to vary for different datasets but, since the uncorrelated statistical uncertainties on the $\omega_a$ values are roughly estimated, the significance of the fitted slopes is uncertain. The systematic uncertainty related to the ADT is therefore evaluated by averaging $\sigma_{ADT}$ over the four datasets, separately for the TMethod and AMethod techniques. The estimated systematic contributions are $\sigma_{ADT} = 14$ ppb for the TMethod and $\sigma_{ADT} = 19$ ppb for the AMethod.

### 7.3.3   Muon Losses

The effect of the lost muons in the functional form of the $\omega_a$ fitting function was described in Section 6.5.2. An imperfect empirical distribution $L(t)$ can be a source of systematic uncertainty. The scaling parameter $k_{LM}$ is highly correlated to the fitted number of positrons $N$, so any distortion in the lost muons function can be absorbed in the fit by these two parameters. The muon loss systematics are evaluated by measuring how much the fitted $\omega_a$ value changes when varying some of the lost muons function parameters.

First of all, the $L(t)$ function is based on a set of selection cuts applied to extract a pure sample of muons hitting multiple calorimeters. In order to estimate the systematic uncertainty related to these cuts, the energy and time selection criteria have been changed slightly by relaxing or tightening the cuts. For each set of cuts, the $L(t)$ function was extracted and applied in the $\omega_a$ fitting function. The effect on the fitted $\omega_a$ is always smaller than 0.1 ppb and therefore considered completely negligible.

Then, a source of systematic uncertainty could be the number of coincidences required in the selection of the muon sample. As described in Section 6.5.2, the function $L(t)$ is built with the muons that hit three, four, and five calorimeters in a row. The systematic bias given by excluding the double coincidences is evaluated by building the $L^D(t)$ function using the double coincidences only. Figure 7.14 shows this integrated $L^D(t)$ function compared to the nominal $L^{TQQ}(t)$ one. The measured systematic bias on $\omega_a$ is $\mathcal{O}(1)$ ppb for all the Run-1 datasets. Table 7.11 summarizes the uncertainties for all the four Run-1 datasets.
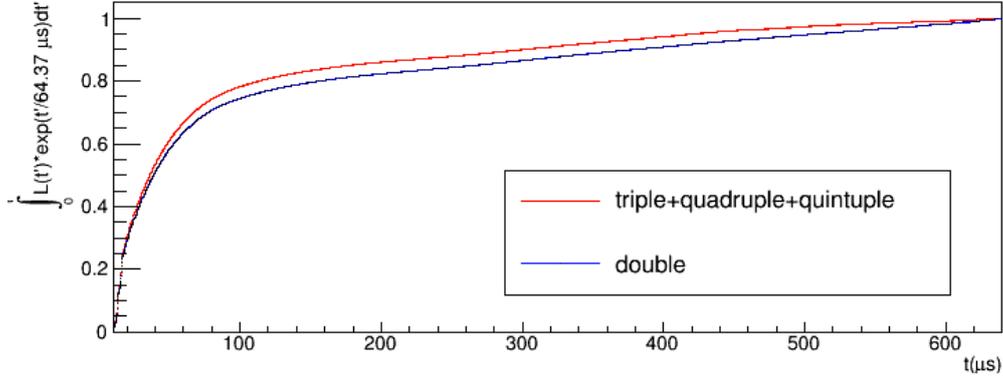
Figure 7.14: Comparison between two different lost muons selections based on the number of calorimeters in a coincidence.

|          | $\sigma_{LM}^{Cuts}$ [ppb] | | $\sigma_{LM}^{Doubles}$ [ppb] | |
| Dataset | TMethod | AMethod | TMethod | AMethod |
|---|---|---|---|---|
| Run-1A | 0 | 0 | 0.7 | 0.7 |
| Run-1B | 0 | 0 | 1.1 | 0.6 |
| Run-1C | 0 | 0 | 0.8 | 0.7 |
| Run-1D | 0 | 0 | 1.2 | 0.4 |

Table 7.11: Run-1 systematic uncertainties due to the empirical $L(t)$ function used for modeling the lost muons distortion of the wiggle plot.

## 7.3.4    Residual slow-term

The fit start-time scans of Section 7.2.1 suggest the presence of a residual time-varying effect which is unaccounted for in the data. The energy-binned scans showed in Section 7.2.2 also reveal an energy dependence of the lost muon correction scaling parameter $k_{LM}$ which should not depend on energy. Both of these effects appear to go away when applying a time-dependent scale correction to the positron energies of the following form:

$$E' = \frac{E}{1 - \delta_g \cdot e^{-t/\tau_\mu} \cdot \left[1 + A_g \cdot \cos(\omega_a t + \phi)\right]} \, , \tag{7.14}$$

where $\tau_\mu$, $\phi$, and $\omega_a$ parameters are taken from the $\omega_a$ full fit, while $\delta_g \approx 5 \cdot 10^{-4}$ and $A_g \approx 0.2$ are specific to this correction. Figure 7.15 show the energy binned scan with and without the correction applied to the positron energies for the dataset Run-1A.

The functional form of Equation 7.14 takes inspiration from a study which evaluated
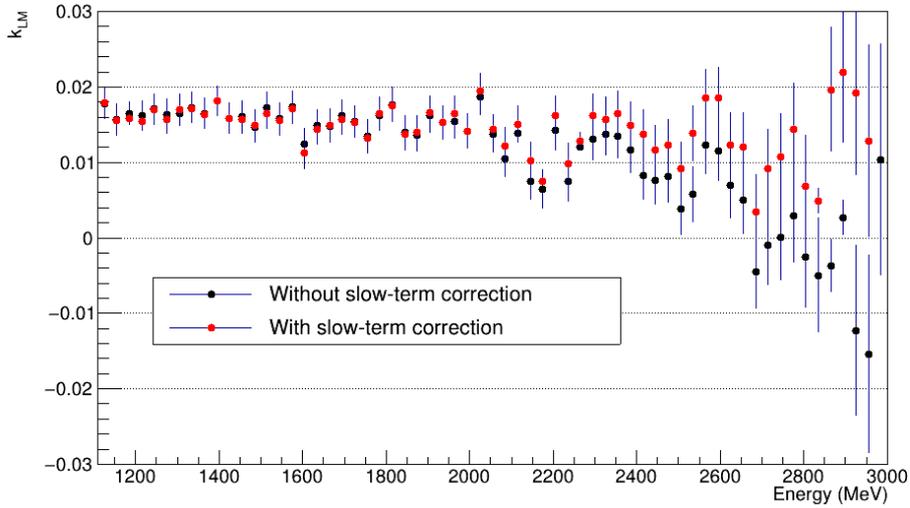
Figure 7.15: Muon loss normalization parameter $k_{LM}$ as a function of the positrons energy with (red) and without (blue) the slow-term correction of Equation 7.14.

the possibility of an uncorrected rate-dependent gain fluctuation of the calorimeters, described in [50]. The gain stability of the calorimeters in the measurement window is monitored by the In-Fill gain correction described in Section 5.4.2. The systematic study performed to evaluate the In-Fill gain related uncertainty in Section 7.3.1 also suggests that the slow-effect might be related to the calorimeter gain corrections.

The possibility of an inaccurate modeling of the In-Fill gain variation is tested by accumulating the fit residual of the In-Fill gain functions to the laser pulses on the SiPMs. While the In-Fill gain sag amplitude depends on the calorimeter location with respect to the beam injection point, any rate-dependent effect should be equal in all calorimeters. However, as the positron hit distribution on the calorimeter surface is skewed toward the beam storage region, such an effect would be enhanced in the outer-most crystals. Figure 7.16 shows the accumulated fit residual obtained by summing over the 1296 SiPMs for all the 21 datasets of Run-4. The contribution from the highest-rate crystals is also shown.

Run-4 is the biggest acquisition period currently reconstructed, with more than twice of the statistics of the Run-3 datasets, thus providing the most precise evaluation for these residuals. However, the plot shows that the residuals are well contained within $\pm 10^{-4}$ bands, thus rejecting the hypothesis of an uncorrected rate-dependent gain fluctuation.
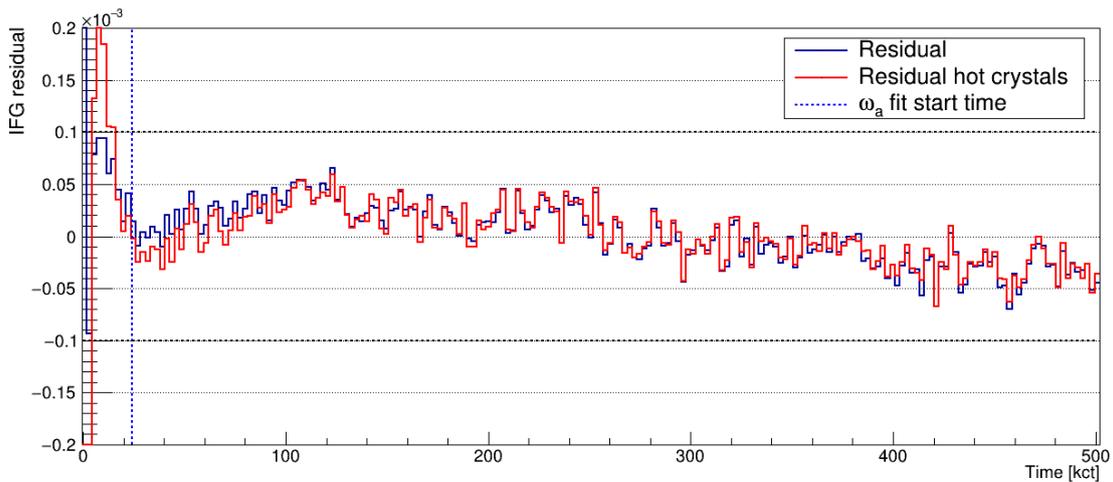
Figure 7.16: In-Fill fit residuals accumulated over the entire Run-4 acquisition period. The red histogram shows the contribution from the grid of $4 \times 4$ crystals closest to the beam storage region, nicknamed as *hot* crystals due to the higher positron rate. The vertical dashed line represents the $\omega_a$ fit start time of $t \approx 30~\mu$s.

As the nature of this slow-term effect is, at this point, not completely understood, the correction term of Equation 7.14 is used to evaluate the relative systematic uncertainty.

To do so, the $\delta_g$ parameter is varied from $1 \times 10^{-4}$ to $1 \times 10^{-3}$, and the asymmetry $A_g$ is varied from 0 to 0.4. The wiggle plot is built using the modified positron energies and the usual pileup and fit routine is performed.
Figure 7.17 shows the $\chi^2/NDF$ and $\omega_a$ scans for the Run-1D dataset. The values $A_g \approx 0.2$ and $\delta_g \approx 9.5 \cdot 10^{-4}$ minimize the fit $\chi^2/NDF$ value.
The four Run-1 data sets prefer different values of $\delta_g$, but all in the order of $0 \div 1 \times 10^{-3}$.

The systematic uncertainty is evaluated as the $\omega_a$ variation given by the one $\chi^2$ unit around the best $\delta_g$ value. A more conservative estimate would be the difference in the measured $\omega_a$ between $\delta_g = 0$ and $\delta_g^{best}$. The results are listed in Table 7.12.

## 7.3.5    Beam oscillations

The complete $\omega_a$ fitting function described in Section 6.5.5 contains 27 free parameters, 21 of which are related to the dynamics of the beam in the storage region. The slightly reduced functions used for Run-1 analysis, listed in Table 7.1, still contain more than 14 beam-related parameters. These parameters do not strongly correlate with the muon pre-
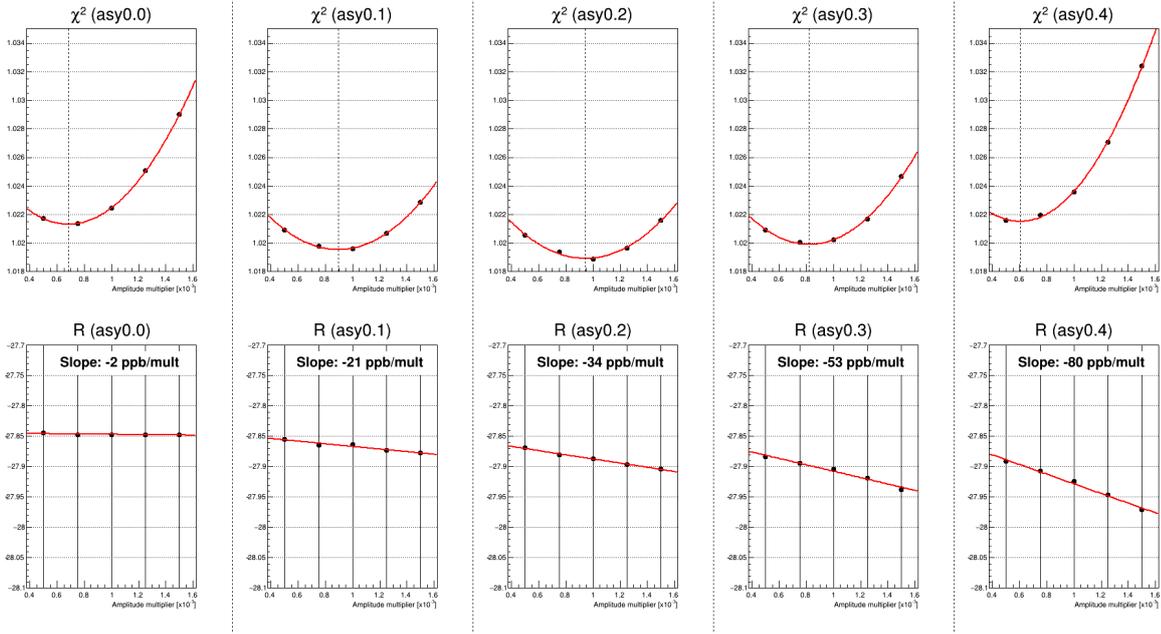
Figure 7.17: Fit $\chi^2/NDF$ (top) and $\omega_a$ values (bottom) as a function of $\delta_g$ for the Run-1D data set. The five vertical sections of the figure correspond to different values of $A_g$ from 0 (left) to 0.4 (right).

| | $\delta_g$ [$10^{-4}$] | | $dR/d\delta_g$ [ppb / $10^{-4}$] | | $\sigma_{slow-term}$ [ppb] | |
|---|---|---|---|---|---|---|
| Dataset | TMethod | AMethod | TMethod | AMethod | TMethod | AMethod |
| Run-1A | 4.9 | 8.2 | 14 | -11 | 12 | 5 |
| Run-1B | 5.6 | 6.0 | -35 | -57 | 33 | 17 |
| Run-1C | -0.7 | 5.0 | 6 | -38 | 18 | 17 |
| Run-1D | 9.7 | 8.7 | -52 | -43 | 13 | 6 |

Table 7.12: $\omega_a$ sensitivities and systematic uncertainties due to the residual slow-term effect for the Run-1 datasets.

cession parameter $\omega_a$ (Appendix A), but any inaccuracy of the model can systematically affect its measurement.

**CBO Frequency change**

As discussed in Section 6.5.1, the damaged quadrupole resistors introduce a time-dependent variation of the field index, and, as a consequence, variable CBO-related frequencies. This frequency change is modeled using data from the tracker stations. The uncertainties on these parameters are obtained directly from the fits to the tracker data and are propagated into an error on $\omega_a$ by randomly sampling the $A$, $B$, $\tau_A$ and $\tau_B$ parameters

in equation 6.44 according to a Gaussian distribution [53]. The widths of the distributions are given by the uncertainties of the parameters as extracted from the tracker data. The extraction is repeated for 100 times, and for each set of values the wiggle plot is fitted with the modified function. The variance of the distribution of the $\omega_a$ values obtained with this method is taken as the systematic uncertainty related to the variable CBO frequency model. Table 7.13 summarizes the uncertainties for the four datasets of Run-1.

| Dataset | $\sigma_{CBO}^{Var}$ [ppb] | |
| --- | --- | --- |
|  | TMethod | AMethod |
| Run-1A | 9.8 | 11.6 |
| Run-1B | 11.2 | 12.2 |
| Run-1C | 15.0 | 17.3 |
| Run-1D | 1.4 | 25.0 |

Table 7.13: Run-1 systematic uncertainties due to the variable CBO frequency parameters obtained from the tracker data.

**CBO lifetime**

The CBO oscillations enter the $\omega_a$ fit via three multiplicative terms:

$$N_{CBO}(t) = 1 + A_{CBO}\cos(\omega_{CBO}(t) + \phi_{CBO})e^{-\frac{t}{\tau_{CBO}}} ,$$

$$A_{BO}(t) = 1 + A_A\cos(\omega_{CBO}(t) + \phi_A)e^{-\frac{t}{\tau_{CBO}}} ,$$

$$\phi_{BO}(t) = 1 + A_\phi\cos(\omega_{CBO}(t) + \phi_\phi)e^{-\frac{t}{\tau_{CBO}}} .$$

The decay lifetime $\tau_{CBO}$ is assumed to be the same for all three contributions. In principle, however, the CBO model can have three different decay time constants: $\tau_{CBO}^N$, $\tau_{CBO}^A$, and $\tau_{CBO}^\phi$, for each of the equations above respectively. The systematic uncertainty associated to the assumption that the three time constants are equal is evaluated by letting them float in the $\omega_a$ fit. The difference in the fitted $\omega_a$ between floating and fixed parameters is taken as the uncertainty.

The fit with floating parameters fails to converge for the Run-1A and Run-1B datasets because of the lower statistical power. For these two datasets the $\tau_{CBO}^\phi$ has been set equal to $\tau_{CBO}^A$, and then the fit managed to converge. This choice is motivated by the best fit parameters for Run-1C and Run-1D datasets, which show $\tau_{CBO}^\phi \approx \tau_{CBO}^A \approx \tau_{CBO}^N/2$.

Table 7.14 summarizes the lifetime-related uncertainties for the four datasets of Run-1. The lower sensitivities found for dataset Run-1D can be explained by the later fit start time at $t \approx 50 \ \mu$s.

|          | $\sigma_{CBO}^{\tau}$ [ppb] | |
| -------- | ------- | ------- |
| Dataset  | TMethod | AMethod |
| Run-1A   | 2.7     | 14.2    |
| Run-1B   | 3.5     | 2.6     |
| Run-1C   | 35.0    | 15.9    |
| Run-1D   | 0.8     | 0.6     |

Table 7.14: Run-1 systematic uncertainties due to the fixed CBO decoherence lifetime parameters.

**CBO decoherence model**

The beam oscillation amplitudes decrease through the beam storage time because of the decoherence effect mentioned in Section 6.5.1. Analysis of the tracker data shows that this decoherence can be modeled with a multiplicative factor of exponential form:

$$D(t) = e^{-\frac{t}{\tau_i}} , \tag{7.15}$$

where $i$ represents any of the beam related oscillations. Similar terms are applied to the fitted asymmetry and the phase (Equations 6.37a-6.37b). The systematic uncertainty related to the CBO lifetime parameter $\tau_{CBO}$ has been summarized in Table 7.14.

However, the model itself can a source of systematic uncertainty, as the exponential term might be an approximation to the real phenomenon. Simulations of the $g - 2$ storage ring show that the betatron oscillations can recohere [53]. A wrong decoherence model could directly affect the fitted beam frequencies, and therefore indirectly bias the measured value of $\omega_a$.

To test for this, an alternative decoherence model function has been used:

$$D(t) = e^{-\frac{t}{\tau_{CBO}}} \left[ 1 + C \cdot \cos \left( \frac{2\pi t}{T} - \varphi \right) \right] , \tag{7.16}$$

with $C \sim 0.135$ and $T \sim 700 \ \mu$s and free to float in the fit. The systematic uncertainty is assessed as the difference between the $\omega_a$ values fitted with the two alternative decoher-

ence models. Table 7.15 summarizes the uncertainties related to the CBO decoherence model for the four datasets of Run-1.

| Dataset | $\sigma_{CBO}^{D}$ [ppb] | |
|---------|---------|---------|
|         | TMethod | AMethod |
| Run-1A  | 22.0    | 39.7    |
| Run-1B  | 13.7    | 7.1     |
| Run-1C  | 9.5     | 15.7    |
| Run-1D  | 9.6     | 5.0     |

Table 7.15: Run-1 systematic uncertainties due to the CBO decoherence model.

### 7.3.6   Hit randomization

The effect of the bunched structure of the beam can be efficiently suppressed with the hit randomization technique described in Section 6.2. For each seed provided to the random number generator, a value of $\omega_a$ is extracted. The final values for $\omega_a$ (Table 7.3) have been obtained by averaging the results of several random seeds. The systematic uncertainty associated to the application of this procedure is evaluated as the standard error of the mean of the distribution of the $\omega_a$ values. The number of random seeds over which the routine is executed has been chosen so that the final uncertainty is compatible with the other systematic sources at the $\mathcal{O}(2)$ ppb level. The number of seeds was limited by the computational resources available. Table 7.16 summarizes these uncertainties for the four datasets of Run-1.

| Dataset | $\sigma_{rand}$ [ppb] | |
|---------|---------|---------|
|         | TMethod | AMethod |
| Run-1A  | 31.4    | 26.0    |
| Run-1B  | 26.3    | 23.6    |
| Run-1C  | 18.9    | 18.3    |
| Run-1D  | 10.9    | 12.6    |

Table 7.16: Run-1 systematic uncertainties due to the hit randomization procedure.

### 7.3.7   Complete systematics table

All the systematic uncertainties described in the previous sections are summarized in Table 7.17 for the $\omega_a$-*Europe* analysis of Run-1.

| Systematic source | Run-1A | | Run-1B | | Run-1C | | Run-1D | |
|---|---|---|---|---|---|---|---|---|
| | T | A | T | A | T | A | T | A |
| Gain corrections | 23.7 | 21.5 | 6.8 | 6.6 | 10.2 | 9.3 | 1.8 | 4.0 |
|    In-fill gain amplitude | 23.6 | 21.4 | 6.7 | 6.5 | 9.9 | 9.1 | 0.6 | 3.9 |
|    In-fill gain time constant | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
|    STDP gain amplitude | 0.1 | 0.1 | 0.1 | 0.1 | 0.3 | 0.2 | 1.2 | 0.1 |
|    STDP gain time constant | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Pileup | 50.0 | 87.8 | 34.0 | 94.1 | 21.3 | 78.5 | 29.3 | 68.7 |
|    Pileup amplitude | 17.2 | 16.2 | 6.6 | 14.0 | 5.3 | 9.8 | 10.4 | 8.8 |
|    Pileup cluster time model | 12.5 | 58.9 | 14.2 | 64.2 | 12.3 | 57.2 | 7.5 | 47.5 |
|    Pileup cluster energy model | 11.7 | 7.0 | 11.2 | 8.0 | 1.7 | 6.5 | 8.4 | 7.0 |
|    Unseen pileup | 5.3 | 1.1 | 0.5 | 2.9 | 0.6 | 1.1 | 0.7 | 0.8 |
|    Triple pileup correction | 3.3 | 4.6 | 1.5 | 5.0 | 1.4 | 3.9 | 2.3 | 4.6 |
| Artificial dead time | 14.0 | 19.0 | 14.0 | 19.0 | 14.0 | 19.0 | 14.0 | 19.0 |
| Lost muons | 0.7 | 0.7 | 1.1 | 0.6 | 0.8 | 0.7 | 1.2 | 0.4 |
|    Muon loss time cuts | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
|    Higher-order coincidences | 0.7 | 0.7 | 1.1 | 0.6 | 0.8 | 0.7 | 1.2 | 0.4 |
| Residual slow-term | 12.0 | 5.0 | 33.0 | 17.0 | 18.0 | 17.0 | 13.2 | 6.0 |
| Beam dynamics | 34.5 | 65.5 | 28.4 | 21.9 | 60 | 48.9 | 11.8 | 30.6 |
|    CBO frequency change | 9.8 | 11.6 | 11.2 | 12.2 | 15.0 | 17.3 | 1.4 | 25.0 |
|    CBO time constants | 2.7 | 14.2 | 3.5 | 2.6 | 35.5 | 15.9 | 0.8 | 0.6 |
|    CBO decoherence model | 22.0 | 39.7 | 13.7 | 7.1 | 9.5 | 15.7 | 9.6 | 5.0 |
| Time randomization seed | 31.4 | 26.0 | 26.3 | 23.6 | 18.9 | 18.3 | 10.9 | 12.6 |
| Total | 74.7 | 116.3 | 63.1 | 102.9 | 71.0 | 98.1 | 38.6 | 78.9 |

Table 7.17: Full systematic table of the $\omega_a$-*Europe* Run-1 analysis. For each category, the first line represents the total uncertainty and the various contributions are shown as indented lines. The systematic uncertainties belonging to the same category are considered 100% correlated and summed linearly. The different categories are considered uncorrelated and summed in quadrature in the final *Total* line. All values are in ppb.

## 7.4    Current and future analyses

After the Run-1 results were published in April 2021, the collaboration started analyzing the next datasets. In particular, Run-2 and Run-3 are now being analyzed together, and a new publication with the results is expected in 2023. As of December 2022, Run-4 production is being finalized and Run-5 production is in progress, while Run-6 production is happening in parallel with the data taking.

After Run-1 analysis, many groups started working to improve the reconstruction and analyses techniques in order to attempt to reduce the largest systematic uncertainties affecting the Run-1 results. One of the highest contributions to the $\omega_a$ systematic uncertainty is the presence of pileup, as visible in Table 7.17. To tackle this, and to try to improve the positron reconstruction in general, I lead the development of the *ReconITA* reconstruction described in detail in Chapter 5. I am also taking part of the Run-2 and Run-3 analyses together with the $\omega_a$-*Europe* group. This section will describe the analysis improvements for Run-2 and Run-3 as well as the impact of the new *ReconITA* reconstruction.

### 7.4.1    ReconITA

The *ReconITA* reconstruction has been thoroughly described in Chapter 5. The main differences between the *ReconITA* approach and the *ReconWest* reconstruction of Run-1 are the calorimeter hit pulse fitting procedure and the hit clustering algorithm. However, the new *ReconITA* pulse fitting software development was completed after both Run-2 and Run-3 were already produced, and entered the official production workflow starting from Run-4.

For Run-1 analysis, only two positron reconstructions procedures were available: *ReconWest* (RW) and *ReconEast* (RE), both described in Section 5.1. Each reconstruction was coupled with its own clustering algorithm, *Time Partitioning* (TP) and *Energy Partitioning* (EP) respectively. The reconstruction-clustering used for the $\omega_a$ analysis were RW-TP and RE-EP. The Run-1 analysis presented in this Chapter was based on the RW-TP combination.

The TP clustering algorithm is characterized by two parameters: $t_{low}$ and $t_{high}$, as described in Section 5.5. For Run-1, the parameters were set to 3 and 5 clock ticks respectively. The TP clustering was modified starting from Run-2, by lowering those

parameters to 2 and 3 clock ticks respectively. For convention, the two different configurations will be named TP1 and TP2:

$$\text{TP1} \rightarrow (t_{low}, t_{high}) = (3, 5) \text{ ct}, \tag{7.17a}$$

$$\text{TP2} \rightarrow (t_{low}, t_{high}) = (2, 3) \text{ ct}. \tag{7.17b}$$

Starting from Run-2, the *ReconITA* (RI) clustering is now part of the analysis path. The $\omega_a$-*Europe* analysis of Run-2 and Run-3 is conducted on the RW-RI reconstruction-clustering combination. This will be referred to as RITA1 in the following sections. In addition, another clustering algorithm was also developed by the University of Washington group referred to as UW [71].

Starting from Run-4, the complete *ReconITA* (RI) reconstruction was inserted in the production and analysis paths, and the $\omega_a$-*Europe* group will conduct the analysis on the RI-RI combination, which will be referred to as RITA2. Table 7.18 provides a summary of the techniques and the combinations for the various run groups.

| Dataset | Reconstruction | Clustering | $\omega_a$-*Europe* analysis |
|---------|----------------|------------|------------------------------|
| Run-1 | RW, RE, Q | TP1, EP | RW-TP1 |
| Run-2/3 | RW, RE, Q | TP2, UW, RI, EP | RW-RI (RITA1) |
| Run-4/5 | RW, RE, RI, Q | TP2, UW, RI, EP | RI-RI (RITA2) |

Table 7.18: List of reconstructions and clustering algorithms deployed for the analysis of the various run groups.

## 7.5    Run-2/3

The analysis of Run-2 and Run-3 is currently in progress and is being finalized by all the participating groups with the goal of publishing the new $a_\mu$ value in the first half of 2023. The software blinding has not been removed yet, and the central values of the different analysis groups cannot be compared.

This section will discuss the improvements of the new ReconITA clustering with respect to the standard TP2 algorithm. The same pileup correction algorithm, described in Section 6.3, has been applied in both cases.

### 7.5.1 Pileup

The main goal of the *ReconITA* clustering is to improve the separation of the pileup events in order to reduce the systematic uncertainty associated to the pileup correction.

Figure 7.18 shows the positron energy spectrum before and after applying the pileup correction. The ReconITA clustering manages to reduce the pileup contamination by a factor of 2 with respect to TP2. This value can be estimated by looking at the number of events with energy greater than 3500 MeV. This region is populated by pileup events only, as the kinematic limit for positrons is about 3100 MeV. The TP2 algorithm itself is an improvement over the Run-1 TP1, with a reduction of ∼1.5 thanks to the new parameter values. Thus, with respect to the analysis procedure used in Run-1, the RITA1 algorithm reduces pileup events by a factor $\sim 3$. As a consequence, it is expected that the pileup related systematics listed in Table 7.17 for Run-1 will be greatly reduced in the Run-2/3 results. As seen in Section 6.3, the pileup correction for Run-2 and Run-3 can be written as (Equation 6.20):

$$N = N' + f_{scale} \left[ f_D \left( S_D - D \right) + f_T \left( S_T - T \right) \right] . \tag{7.18}$$

By varying the scaling factor $f_{scale}$ it is possible to test the sensitivity of the fitted $\omega_a$ value to the pileup correction. Figure 7.19 shows a scan in which $f_{scale}$ has been varied from 0.7 to 1.4. Each modified wiggle plot is fitted with the full function; the difference in $\omega_a$ is extracted with respect to the nominal value at $f_{scale} = 1$. The sensitivity in the RITA1 analysis (245 ppb/mult) is reduced with respect to RW-TP2 (141 ppb/mult) by a factor $\sim 1.7$.

### 7.5.2 Gain systematics

The gain corrections applied in Run-1, and described in Section 5.4, have been improved in the Run-2/3 analysis both in terms of software analysis and in terms of the hardware. In particular:

- The STDP correction for Run-2 includes the sensitivity of the STDP gain sag to the temperature of the SiPMs, as discussed in Section 5.4.1.

- In the Summer between Run-2 and Run-3, a new air conditioning system has been installed in the experimental hall which has reduced the temperature fluctuations around the average value from ±2.5 °C in Run-2 to ±0.3 °C in Run-3 (Figure 5.22).
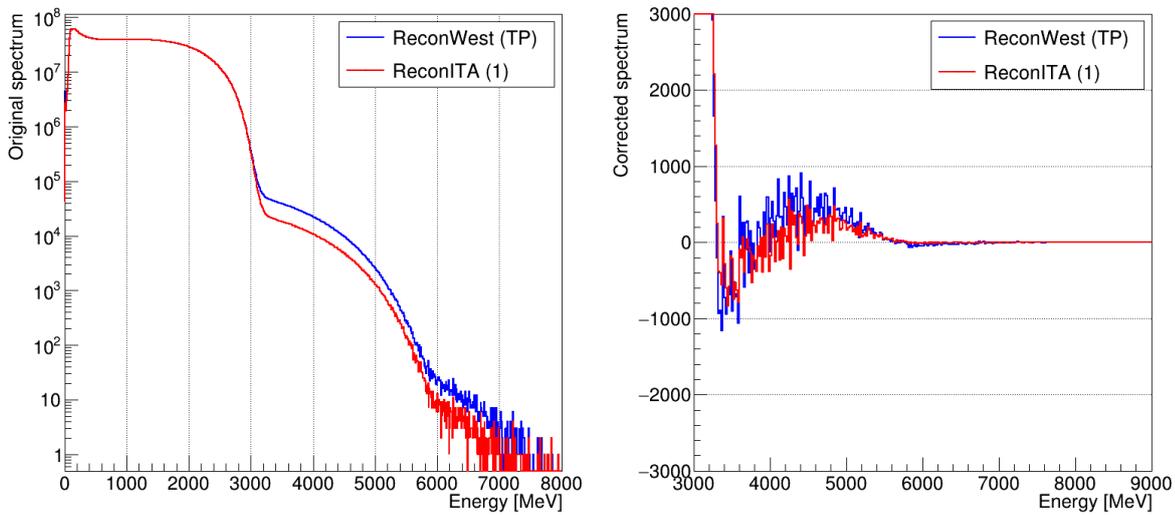
Figure 7.18: (left) Positron energy distribution for dataset Run-2C before any pileup correction is applied. (right) The same energy distribution after the pileup correction has been applied, zoomed in the $E > 3000$ MeV region. The blue curves correspond to the TP2 clustering, while the red curves correspond to the RITA1 clustering.
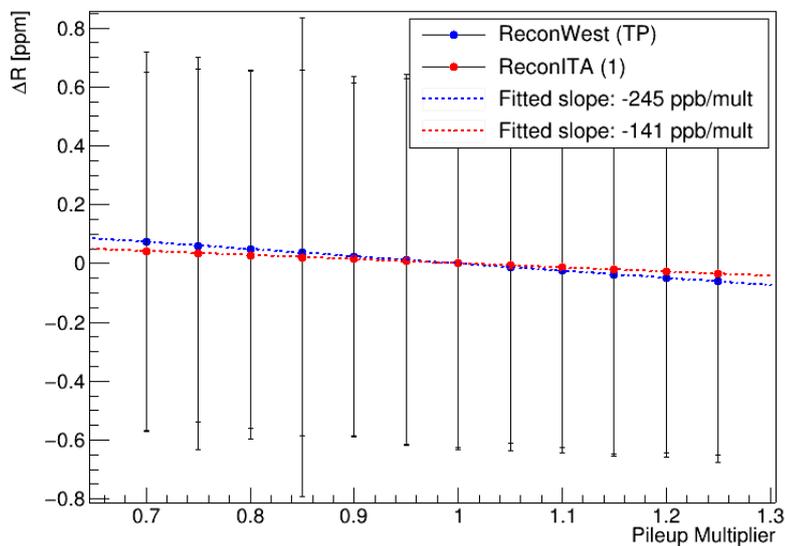


Figure 7.19: Pileup amplitude scan for the Run-2C dataset, with both TP2 (blue) and RITA1 (red) clustering compared. Each point is obtained with a full fit on the AMethod wiggle plot. The dashed lines are linear fits to the data points. Note that the error bars are strongly correlated across the different points.

This has removed the need for any STDP temperature correction and has greatly reduced the effect of the OOF correction, which depends on temperature too.

- The In-Fill gain correction was extracted with higher precision thanks to the LTDP dedicated studies described in Section 5.4.2.

The determination of the systematic uncertainty related to the IFG and STDP corrections is evaluated with a technique which is different from Run-1 as described below. Both the amplitude and lifetime parameters of the two corrections are scaled independently and the $\omega_a$ sensitivity to the scaling factor is extracted. The functional forms of the two corrections are:

$$g_{IFG}(t) = N(1 - \boldsymbol{a}e^{-t/\boldsymbol{\tau}}),  \tag{7.19}$$

$$g_{STDP}(\Delta t, E_1, T) = 1 - E_1 \cdot \boldsymbol{P_1} \cdot (1 + \alpha\,(T - T_{ref})) \cdot e^{-\Delta t/\boldsymbol{\tau}},  \tag{7.20}$$

where the parameters to be artificially scaled are highlighted in bold.

In Run-1, the parameters $a$ and $\tau$ are scaled by a factor $\alpha$ from 0 to 2, with $\alpha = 1$ being the nominal correction.

$$X' = \alpha \cdot X,  \tag{7.21}$$

The systematic uncertainty is then obtained by multiplying the $\omega_a$ sensitivity $\frac{dR}{d\alpha}$ by the average relative error of the $a$ and $\tau$ parameters as obtained from the gain correction extraction. In Run-2 and Run-3, instead, each parameter is shifted by a multiple of the measured error:

$$X' = X + \alpha \cdot \sigma_X,  \tag{7.22}$$

where $\sigma_X$ is the error which is derived from the laser studies. By applying this formula to all individual SiPMs, a more realistic systematic uncertainty is extracted. Figures 7.20 and 7.21 show the IFG and STDP scans respectively for Run-2. The scaling parameter $\alpha$ was varied from -2 to +2 except for the IFG lifetime parameter $\tau$ (Figure 7.20b). This parameter is at the exponent of the correction function, thus the correction is highly sensitive to this parameter and its variation has been limited to $\pm 1$. By applying this procedure, the systematic uncertainty to be assigned to the correction corresponds directly to the value of the slope $\frac{dR}{d\alpha}$. Table 7.19 summarizes the preliminary gain-related uncertainties for Run-2.

The amplitude and lifetime parameters are highly correlated. The correlation has been evaluated by extracting the IFG correction without constraining the lifetime pa-

| Method | IFG | | STDP | |
| --- | --- | --- | --- | --- |
| | Amplitude | Lifetime | Amplitude | Lifetime |
| TMethod | 2.6 | 2.7 | 0.1 | 0.2 |
| AMethod | 2.8 | 2.8 | 0.1 | 0.3 |

Table 7.19: Preliminary systematic uncertainties related to the gain corrections in Run-2 for the $\omega_a$-*Europe* analysis. All values are in ppb.

rameter to the LTDP values. The correlation $\rho$ is calculated as:

$$\rho = \frac{cov(a, \tau)}{\sigma_a \sigma_\tau} \, , \tag{7.23}$$

from the covariance matrix provided by the `ROOT` fit routine. The correlation between the IFG parameters is -92% on average. The total IFG systematics uncertainty can then be written as:

$$\sigma_{IFG}^{tot} = \sqrt{(\sigma_{IFG}^A)^2 + (\sigma_{IFG}^\tau)^2 + 2\rho\sigma_{IFG}^A\sigma_{IFG}^\tau} \, , \tag{7.24}$$

and equal to $\approx 1$ ppb. The correlation between the STDP parameters has not been evaluated as their effect on $\omega_a$ is already negligible.



(a)                                                                  (b)
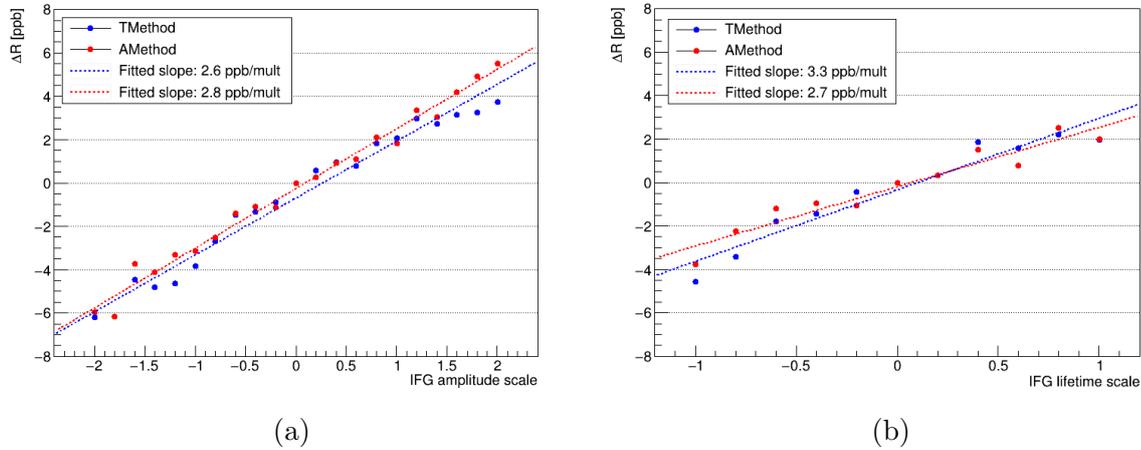
Figure 7.20: In-Fill Gain amplitude scan for the combined Run-2 dataset. (a) Scan of the amplitude multiplier, and (b) scan of the lifetime multiplier. Both the TMethod and AMethod results are fitted with a linear function to extract the sensitivity.

## 7.5.3   Outlook

The Run-2 and Run-3 $\omega_a$ analysis is now at the final stages and the various groups are finalizing the estimation of the systematic uncertainties. The statistical uncertainty
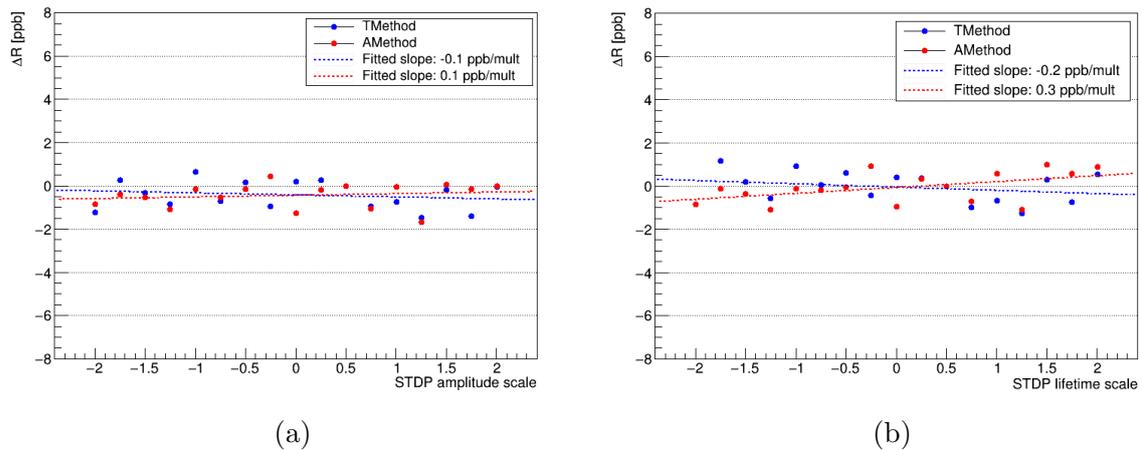
Figure 7.21: STDP Gain amplitude scan for the combined Run-2 dataset. (a) Scan of the amplitude multiplier, and (b) scan of the lifetime multiplier. Both the TMethod and AMethod results are fitted with a linear function to extract the sensitivity.

is expected to be $\sim$200 ppb for the AMethod, as the number of observed positrons is more than twice the Run-1 ones (Table 6.1). All the analysis teams are focusing on the reduction of the systematic uncertainties. The improvements from the $\omega_a$-*Europe* group with respect to Run-1 are:

- Improved gain uncertainty estimation, as discussed in this section;

- Reduced pileup correction and related uncertainties thanks to the new ReconITA clustering described in Chapter 5;

- Removal of the Artificial-Dead-Time and the Triple pileup uncertainties thanks to the new pileup correction technique described in Section 6.3;

- The residual slow-term is artificially corrected with a gain-like correction term as in Equation 7.14;

- More accurate modeling of the CBO terms in the $\omega_a$ fit, including a new CBO-VW beating term (Section 6.5.4) measured thanks to the higher statistics.

Overall, the systematic uncertainties associated to $\omega_a$ are expected to decrease by almost a factor of two, from the Run-1 value of 56 ppb (presented in Section 8.6) to a projected preliminary value of $\mathcal{O}(30)$ ppb.

# 7.6   Run-4/5

The analysis of the Run-4 and Run-5 has not started yet. However, the Run-2C dataset has been reconstructed also by applying the RITA2 procedure in order to evaluate the additional improvements of RITA2 with respect to RITA1 in view of the final publication with the full statistics. This section discusses some preliminary results.

## 7.6.1   Slow-term correction

The main improvement of RITA2 with respect to RITA1 is the new pulse fitting technique. As described in Chapter 5, the algorithm tries to further separate pileup clusters that overlap spatially and generate multiple hits on the same SiPM. In addition, even if two positrons hitting the same calorimeter were already clearly separated, the algorithm manages to recover some previously missed hits belonging to the second positron. These missed hits are proportional to the rate of pileup, and are partially responsible of the residual slow-term effect observed in Run-1 and discussed in Section 7.3.4. This section will compare the RITA2 reconstruction with the RW one used for Run-2 and Run-3. It has to be noted that work is in progress also on the RW reconstruction by its developers in view of Run-4 and Run-5 in order to overcome some of the limits observed in the current version that RITA2 tries, and manages, to solve.

To investigate the impact of the new ReconITA reconstruction on the residual slow-term effect, an energy binned analysis has been performed. The RITA2 analysis shows an improvement in the fitted value of $k_{LM}$ at high energies (Figure 7.22, keeping it positive to almost 2500 MeV. A reduction of $\sim 66\%$ on the negative deviation is measured up to 2900 MeV. Consequently, the peak at low frequencies in the Fast Fourier Transform of the wiggle fit residuals is removed as well (Figure 7.23). While the slow-term effect is not totally resolved, it is limited to the last few energy points, in which the statistics decreases down to zero at the kinematic limit of 3.1 GeV, and the related systematic uncertainty is expected to be reduced.

## 7.6.2   Start-time scans

One of the standard cross checks performed to identify effects that have not been accounted for is the *start time scan*, first introduced in Section 7.2.1. Figure 7.24 shows the comparison between the RW-TP2 and RITA2 reconstruction for the TMethod analysis. Any significant deviation from the allowed $1\sigma$ bands would suggest an unaccounted-for
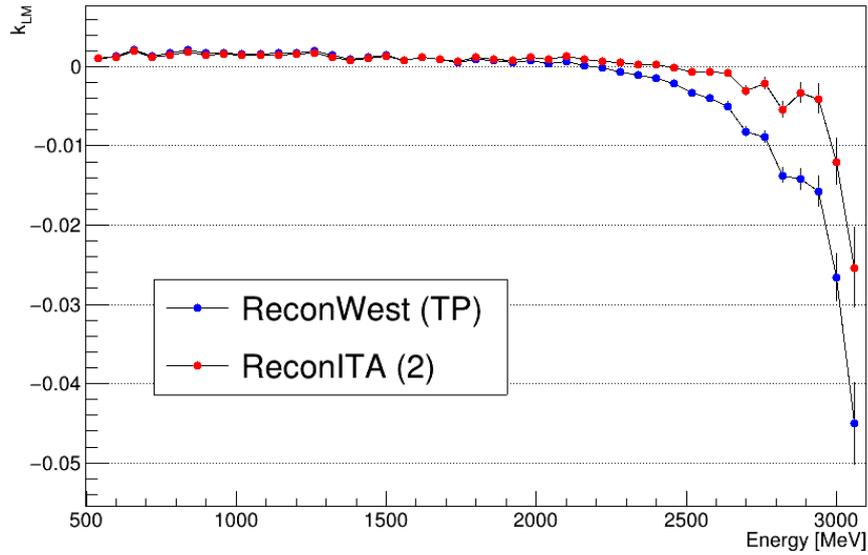
Figure 7.22: Energy binned analysis of the Run-2C dataset with the RW-TP2 (blue) and RITA2 (red) reconstructions. The negative deviation of the Muon loss normalization parameter $k_{LM}$ at high energies is reduced in RITA2 by $\sim 66\%$.
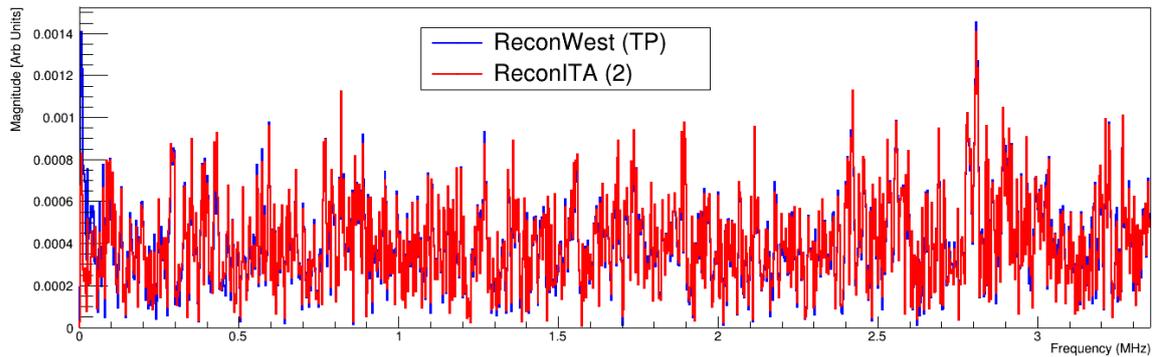


Figure 7.23: Fast-Fourier-Transform of the residuals of the wiggle fit for the TMethod analysis. The peak at very low frequencies visible in the RW-TP2 analysis (blue curve) is removed by the RITA2 reconstruction (red curve).

effect present in the data. The $N$, $\tau$, and $k_{LM}$ parameters show deviations in RW-TP2 that are greatly reduced by the RITA2 reconstruction. We have already shown that the $k_{LM}$ parameter behavior improves with the new reconstruction (Figure 7.22) and its fitted value goes from a nonphysical negative value to a positive one. The fitted $\omega_a$ value is stable for both reconstructions, indicating that these slow-term effects have a limited impact, which however will have to be quantified, on the measured muon precession frequency.

(a) Blinded $\omega_a$ value $R$.



(b) Number of positrons $N$.



(c) Boosted muon lifetime $\tau$.



(d) Muon loss normalization $k_{LM}$.
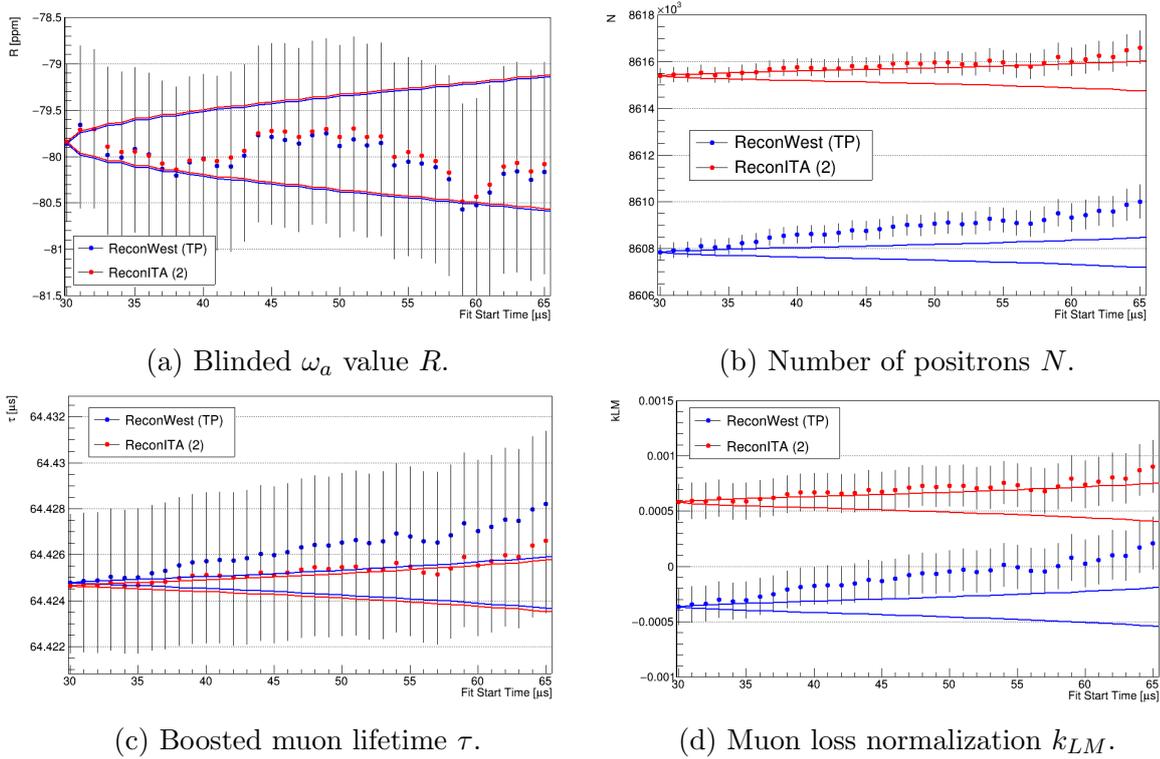
Figure 7.24: Fit start time scan for the Run-2C dataset reconstructed with RW-TP2 (blue) and RITA2 (red). The four fit parameters $\omega_a$, $N$, $\tau$, and $k_{LM}$ are shown. The solid lines represent the $1\sigma$ allowed statistical deviation from the nominal starting point of $t = 30$ $\mu$s.

### 7.6.3   Outlook

The Run-4 and Run-5 $\omega_a$ analysis has not started yet. Run-6 will be possibly analyzed together as well, as soon as data collection is finished. The statistical uncertainty is expected to be another factor of two smaller than the combined Run-2 + Run-3, with a final value of $\mathcal{O}(100)$ ppb for the AMethod. Production of Run-4 is now finishing and the full ReconITA reconstruction has been incorporated. The expected improvements for the $\omega_a$-Europe analysis with respect to Run-2 and Run-3 are:

- Greatly reduced residual slow-term and related uncertainties thanks to the new complete ReconITA reconstruction described in Chapter 5;

- Possible inclusion of the fills containing In-Fill laser pulses thanks to an improved rejection algorithm.

This work is currently in progress and it is expected to be concluded at the end of 2024, with the final publication at the beginning of 2025. The total uncertainty with

the full statistics is expected to reach the Technical Design Report [49] goal of 140 ppb with the concrete possibility of performing even better. This will be an important and awaited result in view of the current puzzle between different theoretical predictions.

# Chapter 8

# Run-1 $a_\mu$ measurement

In the previous chapter I have mostly discussed the determination of the muon precession frequency $\omega_a$. This chapter presents the complete $a_\mu$ determination from the first year of data taking, Run-1, which was published in April 2021. Additional details can be found in the four published papers [11, 68, 53, 51]. A similar procedure is currently going on for the analysis of Run-2 and Run-3, for which the unblinding and the publication are expected in the first half of 2023. The preliminary results from these two runs are currently under embargo and cannot be described at the time of publication of this thesis.

## 8.1 Complete $a_\mu$ formula

The muon anomaly is calculated with the equation described in Section 2.1:

$$a_\mu = \frac{\omega_a}{\tilde{\omega}'_p(T_r)} \frac{\mu'_p(T_r)}{\mu_e} \frac{m_\mu}{m_e} \frac{g_e}{2} \,, \tag{8.1}$$

where the first fraction is the quantity measured by the Muon $g-2$ Experiment. The other terms are known with high precision from other experiments. The proton-to-electron magnetic moment ratio $\mu'_p(T_r)/\mu_e$ is taken from [73] at the reference temperature $T_r = 34.7\ °\text{C}$. The electron $g$-factor is taken from [16], while the CODATA-2018 result is used for the $m_\mu/m_e$ mass ratio [14]. These factors are known with a total uncertainty of 25 ppb.

The $\tilde{\omega}'_p$ term can be written in the form:

$$\tilde{\omega}'_p = f_{calib} \langle \omega_p(x,y,\phi) \times M(x,y,\phi) \rangle \,, \tag{8.2}$$

representing the convolution between the magnetic field strength, expressed in terms of the Larmor precession of the proton $\omega_p$, and the muon distribution $M$ inside the storage region, all multiplied by the absolute calibration factor $f_{calib}$.

In addition to the muon anomalous precession frequency, $\omega_a$, and the muon-weighted proton Larmor precession frequency, $\tilde{\omega}'_p(T_r)$, additional correction factors have to be included to take into account beam effects, transient magnetic and electric fields, and the residual effect to the quadrupole field. The first fraction of Equation 8.1 can be expressed as:

$$\mathcal{R}'_\mu \equiv \frac{\omega_a}{\tilde{\omega}'_p(T_r)} \approx \frac{f_{clock}\omega_a^m(1 + C_e + C_p + C_{ml} + C_{pa})}{f_{calib}\langle\omega_p(x,y,\phi) \times M(x,y,\phi)\rangle(1 + B_k + B_q)} \, . \tag{8.3}$$

The next sections will describe each of the terms of Equation 8.3 and their measured values.

## 8.2  $\omega_a$ combination

The results presented in Chapter 7 refer to the measurement of the muon precession frequency performed by the $\omega_a$-$Europe$ group of which I am part of. However, five other analysis groups participated in the $\omega_a$ measurement with different reconstructions, corrections, and analysis techniques, for a total of 11 determinations of $\omega_a$ for each Run-1 dataset (Table 8.1).

| Analysis group | Reconstruction | Weighting method |
|---|---|---|
| $\omega_a$-$Europe$ | ReconWest | T, A |
| Boston University | ReconWest | T, R |
| Cornell University | ReconEast | T, A |
| University of Kentucky | QMethod | Q |
| Shanghai Jiao Tong University | ReconWest | T, A |
| Washington University | ReconWest | T, A |

Table 8.1: List of the Run-1 $\omega_a$ analysis groups and the relative reconstructions and methods used.

The combination of the different analyses into a single $\omega_a$ value is not trivial, as the various reconstructions and methods are highly correlated as the majority of the data is

| Analysis | RE-T | RE-A | RW-T | RW-A | RW-R | Q |
|----------|------|------|------|------|------|------|
| RE-T | 1.00 | 0.91 | 0.95 | 0.91 | 0.95 | 0.51 |
| RE-A |      | 1.00 | 0.90 | 0.99 | 0.90 | 0.58 |
| RW-T |      |      | 1.00 | 0.91 | 1.00 | 0.51 |
| RW-A |      |      |      | 1.00 | 0.90 | 0.57 |
| RW-R |      |      |      |      | 1.00 | 0.50 |
| Q    |      |      |      |      |      | 1.00 |

Table 8.2: Statistical correlations calculated with Monte Carlo simulations for the different types of reconstructions and analysis technique for the determination of $\omega_a$ [68].

shared among them. These correlations have been determined from Monte Carlo simulations that incorporate the major differences between the various reconstruction and analysis techniques. The correlation coefficients are listed in Table 8.2.

Given these correlation coefficients, the allowed statistical deviation $\Delta\sigma_{12}$ between the fitted $\omega_a$ values for two different analyses can be calculated as:

$$\Delta\sigma_{12} = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\,. \tag{8.4}$$

The combination of two positively correlated results has a maximum variance for [74]:

$$\rho_{crit} = \frac{min(\sigma_1/\sigma_2)}{max(\sigma_1/\sigma_2)}\,, \tag{8.5}$$

and drops to zero when the correlation is $\rho \to 1$. For very high correlations, the weight of the less precise measurement becomes negative, affecting the stability of the BLUE (best linear unbiased estimator) procedure used to calculate the $\omega_a$ average [68].
Because of this, the following decisions have been made for the final combination of $\omega_a$:

- For each analysis, the TMethod result is discarded in favor of the more precise AMethod value. This has been demonstrated in Section 6.4.

- The QMethod results are not included in the final average because they have a much higher statistical uncertainty and show a moderate systematic shift with respect to the RW/RE analyses. The impact of the damaged quadrupole resistors and the different fit stop times can explain differences up to $\mathcal{O}(200)$ ppb.

- The Ratio Method result is not included since, for Run-1, it has the same statistical precision of the TMethod. The new analyses of Run-2 and Run-3 will apply the

asymmetry weighting to the RMethod too.

- The three analyses using the same ReconWest reconstruction are averaged together with equal weights.

- The ReconWest combined and ReconEast results are averaged with equal weights.

Overall, four analyses have been combined, three of which are based on the ReconWest reconstruction, to provide the final value of $\omega_a$. The analysis presented in Chapters 6 and 7 conducted by the author as part of the $\omega_a$-*Europe* group is one of the combined ones.

Finally, the systematic uncertainties from the various analyses are combined assuming 100% correlation within the same systematic category. The blinded combination average for the four datasets in Run-1 and the relative systematic uncertainties are summarized in Table 8.3.

| Quantity | Run-1A | Run-1B | Run-1C | Run-1D |
|---|---|---|---|---|
| $\omega_a$ (blinded) | -28371 | -26986 | -27596 | -27644 |
| Statistical uncertainty | 1207 | 1022 | 823 | 675 |
| Systematic uncertainty | 64 | 70 | 54 | 49 |
|    Time synchronization | 4 | 1 | 1 | 1 |
|    Gain corrections | 12 | 9 | 9 | 5 |
|    Pileup | 39 | 42 | 35 | 31 |
|    Beam dynamics | 42 | 49 | 32 | 35 |
|    Residual slow-term | 21 | 21 | 22 | 10 |
|    Hit randomization | 15 | 12 | 9 | 7 |
| Total uncertainty | 1209 | 1025 | 825 | 676 |

Table 8.3: Combined $\omega_a$ values and uncertainties for the four Run-1 datasets. The blinded $\omega_a$ values of different datasets cannot be directly compared because they are related to different measured magnetic field strengths. All values are in ppb.

## 8.2.1   Clock blinding factor

As discussed in Section 7.1, the Muon $g - 2$ Experiment employs a hardware blinding factor to the clock frequencies that govern the calorimeter digitizers. The first term of the numerator of Equation 8.3, $f_{clock}$, represents the factor needed to convert the measured

$\omega_a^m$ to the actual unshifted $\omega_a$ value. The stability of the detuned clock frequency is continuously monitored and the uncertainty on the $f_{clock}$ factor is negligible.

## 8.3 Corrections to $\omega_a$

There are four corrections to be applied to the measured value of $\omega_a$. These terms cannot be directly extracted from the wiggle plot fit and dedicated studies have been performed. If $\omega_a^m$ is the measured precession frequency, then:

$$\omega_a \approx \omega_a^m \left(1 + C_e + C_p + C_{lm} + C_{pa}\right). \tag{8.6}$$

This section will briefly describe the four corrections applied to the Run-1 data. Table 8.4 lists the corrections applied to the four Run-1 datasets.

| Correction term | Symbol | Run-1A | Run-1B | Run-1C | Run-1D |
|---|---|---|---|---|---|
| Electric field | $C_e$ | 471 ± 53 | 464 ± 54 | 534 ± 54 | 475 ± 53 |
| Pitch | $C_p$ | 176 ± 12 | 199 ± 14 | 191 ± 14 | 166 ± 12 |
| Muon loss | $C_{ml}$ | -14 ± 6 | -3 ± 2 | -7 ± 4 | -17 ± 6 |
| Phase acceptance | $C_{pa}$ | -184 ± 96 | -165 ± 74 | -117 ± 60 | -164 ± 80 |

Table 8.4: List of corrections to the measured $\omega_a$ value for the four Run-1 datasets. All values are in ppb.

### 8.3.1 Electric field correction $C_e$

The momentum of the muons has been specifically chosen to minimize the effect of the focusing electrostatic quadrupole on the precession frequency of the muons. Equation 2.4 for the anomalous precession frequency reads:

$$\vec{\omega}_a = -\frac{e}{mc}\left[a_\mu\vec{B} - \left(a_\mu - \frac{1}{\gamma^2 - 1}\right)\vec{\beta}\times\vec{E} - a_\mu\frac{\gamma}{\gamma+1}\left(\vec{\beta}\cdot\vec{B}\right)\vec{\beta}\right], \tag{8.7}$$

the term due to the electric field $\vec{E}$ can be minimized by choosing a relativistic gamma such that:

$$a_\mu - \frac{1}{\gamma^2 - 1} = 0 \rightarrow \gamma = \sqrt{\frac{1}{a_\mu} + 1} \approx 29.3\,, \tag{8.8}$$

which corresponds to a *magic* momentum of 3.094 GeV/c. The distribution of momenta of the muon beam, however, has a certain spread around the central value of 3.094

GeV/c. The muons with higher and lower momenta are then affected by the electric fields and so does the measured $\omega_a$. The electric field term vanishes at the *magic* orbital radius $R_0$, as a result of the design of the quadrupoles. However, the muon beam is not perfectly centered, and the momentum spread implies a spread in the distribution of radial positions.

The mean radial electric field experienced by a muon oscillating around an equilibrium radius $x_e$ in an ideal electric quadrupole is

$$\langle E_r \rangle = k x_e = \frac{n \beta c B_y}{R_0} x_e \,, \tag{8.9}$$

where $k$ is the electric field gradient, $n$ is the electric field index, and $B_y$ is the vertical component of the magnetic field. The electric field correction can be expressed in terms of the beam radial distribution:

$$C_e \approx 2n(1-n)\beta_0 \frac{\langle x_e^2 \rangle}{R_0^2} \,, \tag{8.10}$$

where $\beta_0 \approx 0.9994$ is the *magic* speed of the muons.

The beam distribution $\langle x_e^2 \rangle$ is measured by analyzing the cyclotron motion of the muons with a Fourier transform analysis of the calorimeter data [53]. Figure 8.1 shows the measured distribution for the four different datasets.
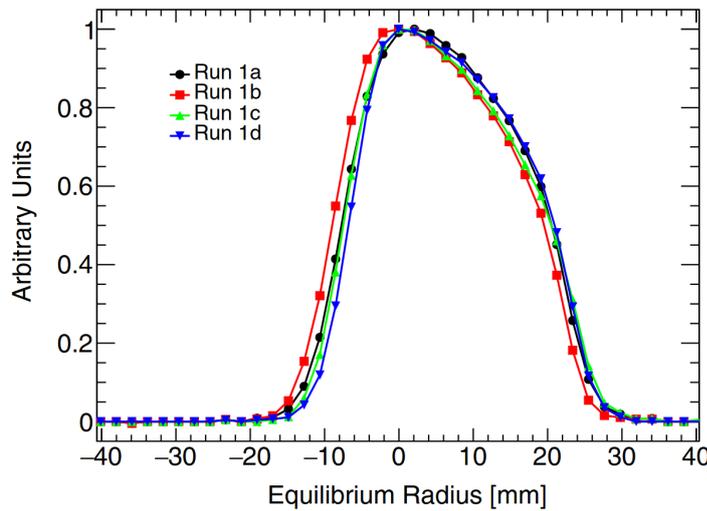


Figure 8.1: Radial distribution of the muon beam as measured from the cyclotron frequencies determined with a Fourier transform analysis of the calorimeter data. The equilibrium radius is defined to be 0 mm at the *magic* momentum of 3.094 GeV/c. Figure from [53].

## 8.3.2 Pitch correction $C_p$

The last term of Equation 8.7 represents the contribution to the anomalous precession frequency given by the component of the muon momentum parallel to the magnetic field. The quadrupole system used to provide the vertical focusing to the muon beam introduces vertical betatron oscillations, as described in Section 6.5.1. As the muons oscillate, their vertical component of the momentum will introduce a shift in the precession frequency. The correction can be expressed as:

$$C_p \approx \frac{n}{4} \frac{\langle A_y^2 \rangle}{R_0^2} \tag{8.11}$$

where $A_y$ is the amplitude of the vertical oscillations. The correction is extracted from the tracker measurements of the beam motion. Figure 8.2 shows the measured distribution of oscillation amplitudes.
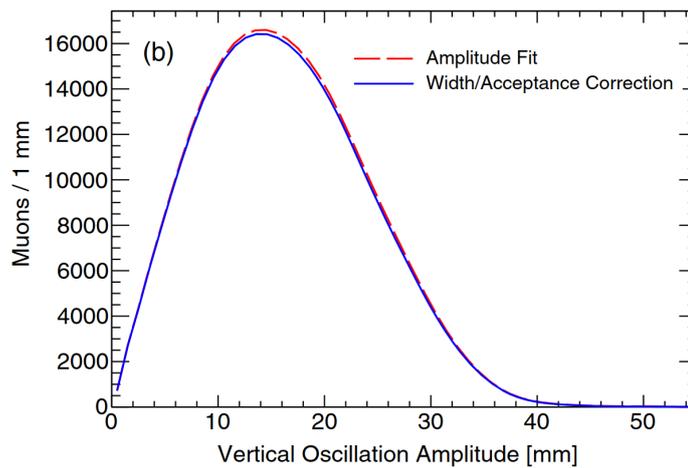


Figure 8.2: The distribution of beam vertical oscillation amplitudes, before (red) and after (blue) averaging over the azimuthal angle. Figure from [53].

## 8.3.3 Muon loss correction $C_{ml}$

Some of the muons that circulate in the storage ring are lost before decaying into positrons. In general, a muon will be scattered out of the storage region after it strikes one the collimators that limit the transverse region and the momentum acceptance. These collimators have an aperture of radius $r_0 = 45$ mm and are centered on the *magic* orbit. A muon hitting a collimator loses some energy, and starts curling inward until it exits the storage region.

The time distribution of lost muons induces a distortion in the exponential decay shape as observed by the calorimeters. This has been discussed already in Section 6.5.2 and is part of the $\omega_a$ analysis procedure.

Monte Carlo beamline simulations and simple analytical calculations both predict that a correlation exists between the injected muon average spin phase and the particle momentum [53]. As the probability of hitting a collimator depends on the muon momentum, the population of muons that are lost is correlated to the $\omega_a$ phase observed by the calorimeters. Since the muons are not lost uniformly across the storage time, a slow drift in the $\omega_a$ phase is induced. A time-dependent phase can be written as a Taylor expansion:

$$\phi(t) = \phi_0 + \frac{d\phi}{dt}t + \mathcal{O}(t^2)\,, \tag{8.12}$$

where $\frac{d\phi}{dt}$ is a direct bias on the measured $\omega_a^m$. The magnitude of the effect is directly proportional to the number of lost muons, which was especially high during Run-1 because of the damaged quadrupole resistors.

The correction for this effect has been evaluated by taking three special acquisition runs where the magnetic field was changed by $\pm 0.67\%$ from its nominal value. Different magnetic field strengths result in different momentum acceptances of the storage ring, allowing for the selection of a different portion of the original beam momentum distribution. For each of these three acquisitions, the $\omega_a$ analysis is performed to extract the average phase of the muon beam. The relation between the phase variation and the mean muon momentum is then extracted. Figure 8.3b shows the result for Run-1, with a measured dependence of:

$$\frac{d\phi}{dp} = (-10.0 \pm 1.6)\,\text{mrad}/\% \tag{8.13}$$

The time dependent phase can be expressed as:

$$\frac{d\phi}{dt} = \frac{d\phi}{dp} \cdot \frac{dp}{dt}\,. \tag{8.14}$$

The first term is the result of Equation 8.13, while the second term was determined by measuring the muon loss rate as a function of time in special systematic runs where muon beam momentum was selected using collimators in the delivery ring, before being

injected into the E989 ring.

The phase angle $\phi(t)$ is parameterized using a polynomial function, which is then used to generate a wiggle plot according to the five parameter formula

$$N(t) = Ne^{-t/\tau_\mu}(1 + A\cos(\omega_a t - \phi(t))) . \tag{8.15}$$

A second wiggle plot was generated by using a fixed $\phi(t) = \phi(0)$. The muon loss correction $C_{ml}$ is finally calculated as the difference in the fitted $\omega_a$ between the two artificial wiggle plots.



(a)                                                    (b)
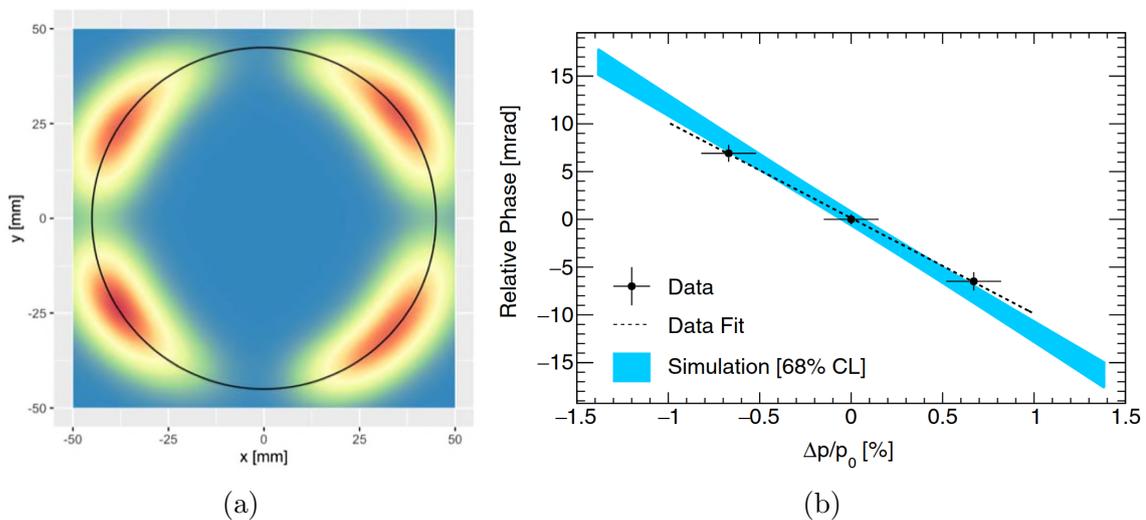
Figure 8.3: (a) Distribution of muon hits on the collimators. (b) Phase-momentum correlation from simulation (blue band) and from a data-driven approach (black). The three data points are obtained by fits to muon precession frequency data at nominal, reduced, and increased central magnetic field values. Figures from [53].

### 8.3.4   Phase acceptance correction $C_{pa}$

The $g - 2$ phase $\phi$ in the five-parameter expression of Equation 6.29 represents the average phase of the muons at injection. The measured phase, instead, depends both on the energy of the decay positron and the position of the decay vertex.

The energy dependence is due to the fact that higher energy positrons have a larger radius of curvature and therefore a longer trajectory path toward the calorimeters and a higher time-of-flight. Considering two positrons of different energy hitting a calorimeter at the same time, the higher energy one was produced further back in the storage ring from a muon with a different spin angle than the one that produced the lower energy

positron. A typical time-of-flight for a positron is ~12 ns, which corresponds to a phase change of ~17 mrad.

The position dependence of the phase has two explanations. The first one is related again to the time-of-flight of the decay positron: if a muon decays closer to the calorimeters, the resulting positrons will travel a shorter path on average with respect to the ones that decay more radially outward. The second reason has to do with the direction of the muon's spin that maximizes the calorimeter acceptance. As the calorimeters are placed to the inside of the muon trajectory, the spin orientation that maximizes acceptance into the calorimeters is not parallel to its momentum but rotated slightly radially inward.

These effects are a function of the transverse coordinates $x$ and $y$, and the positron energy $E$, and generate an effective phase shift $\phi_{pa}(x, y, E)$. The energy dependence was already observed in the energy-binned $\omega_a$ fit scans discussed in Section 7.2.2. Figure 8.4 shows the transverse "phase map" averaged over the azimuth angle as measured by the asymmetry weighted (AMethod) $\omega_a$ analysis. The map is generated from a Geant4 simulation of the storage ring [53].
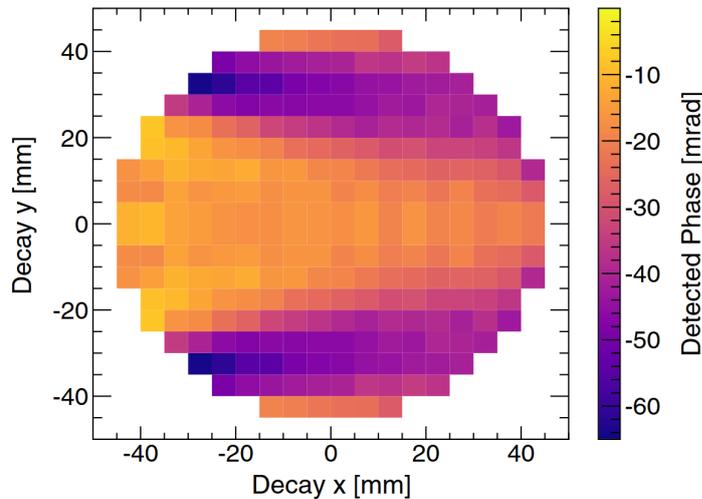


Figure 8.4: The map of the shift in the measured $\phi$ depending on the decay vertex position $(x, y)$, averaged over the azimuth angle. Figure from [53].

If the muon beam did not have any motion inside the storage ring, these phase-acceptance effects would not bias the measured value of $\omega_a$. However, the betatron oscillations and the beam movements due to the quadrupoles generate a time-varying phase that cannot be accounted for during the $\omega_a$ analysis.

The phase-acceptance correction is extracted by measuring the beam distribution with the two tracker stations and extrapolating the phase maps to all the azimuth angles corresponding to the calorimeter locations with a Geant4 simulation [53]. For each calorimeter, the phase variation $\phi_{pa}^C(t)$ is extracted as a function of time. Similar to the $C_{ml}$ technique, a wiggle plot is generated with the full $\omega_a$ function of Equation 6.45, including the phase variation $\phi_{pa}^C(t)$. The correction $C_{pa}$ is extracted as the difference between the fitted $\omega_a$ and the value used for generating the wiggle plot.

This correction is particularly relevant for Run-1 due to the additional vertical beam movements caused by the damaged quadrupole resistors.

## 8.4 $\tilde{\omega}_p'$ measurement

The experimental steps for measuring the magnetic field are outlined in Section 2.4. The field strength is measured by a suite of Nuclear Magnetic Resonance probes, which measure the Larmor precession of the proton [51] with a sample of petroleum jelly. The probes themselves cannot be positioned in the storage region while the muon beam is present, as they would disrupt it otherwise. So, every ∼3 days, a cylindrical "trolley", approximately 30 cm long and with a diameter of 9 cm, corresponding to the storage region cross section, containing 17 NMR probes is inserted in the storage region to provide a set of 2D field maps at about 9000 azimuthal locations (one every ∼0.5 cm).

To measure the field variations during the three days separating two trolley measurements, a set of 378 NMR probes are located 7.7 cm above and below the storage volume, continuously monitoring the field at 72 azimuthal locations. A subset of these probes is also used to provide a feedback to the magnet supply to actively stabilize the field strength across time.

The analysis of the probes data provides a 3D map of the average magnetic field inside the storage region, $\omega_p(x, y, \phi)$, over the measurement time of a $\omega_a$ dataset. Figure 8.5a shows the map averaged over the azimuth angle.

### 8.4.1 Absolute calibration

Both the trolley and fixed probes are calibrated with a water-sample probe that is installed on a translation stage in the ring vacuum. The probe can be positioned in the

same locations as the trolley probes, determining the calibration offsets between the trolley measurements and the actual $\omega_p$ values. The offsets are caused by the diamagnetic shielding effects of the petroleum jelly samples, and the trolley body and shape. The trolley probe calibration offsets are determined with an average uncertainty of 30 ppb [51].

In addition, dedicated studies in a Magnetic Resonance Imaging (MRI) solenoid provide an absolute calibration by comparing the water calibration probe response to a $^3$He magnetometer [75]. Finally, corrections are required to relate the calibration probe measurements to the ones expected from a proton in water at the reference temperature of $T_r = 34.7$ °C in order to be able to use the external measurement of the proton-to-electron magnetic moment ratio $\mu_p'(T_r)/\mu_e$ in the final $a_\mu$ computation (Equation 8.1).

## 8.4.2   Muon weighting

The calibrated field map, $\omega_p'(x, y, \phi)$, represents the average absolute magnetic strength inside the storage region. However, the circulating muons have a non-uniform distribution in the transverse plane of the orbit, as shown in Figure 8.5b. For this reason, the field maps have to be weighted by the distribution of muons $M(x, y, \phi)$ that are detected by the calorimeters in order to obtain the effective final field measurement $\tilde{\omega}_p'$.

The muon distribution $M$ is calculated by multiplying the spatial distribution as measured by the trackers with the positron rate as measured by the calorimeters. As the trackers are present in two azimuthal positions only, the complete beam distribution along the entire storage region is extrapolated using a beam dynamics Monte Carlo simulation.
The interpolated field maps are averaged over periods of roughly 10 s and weighted by the number of detected positrons during the same interval.

The combined total uncertainty of $\tilde{\omega}_p'$ from probe calibrations, field maps, tracker alignment and acceptance, calorimeter acceptance, and beam dynamics modeling is 56 ppb for the Run-1 measurement [51].

(a)                  (b)
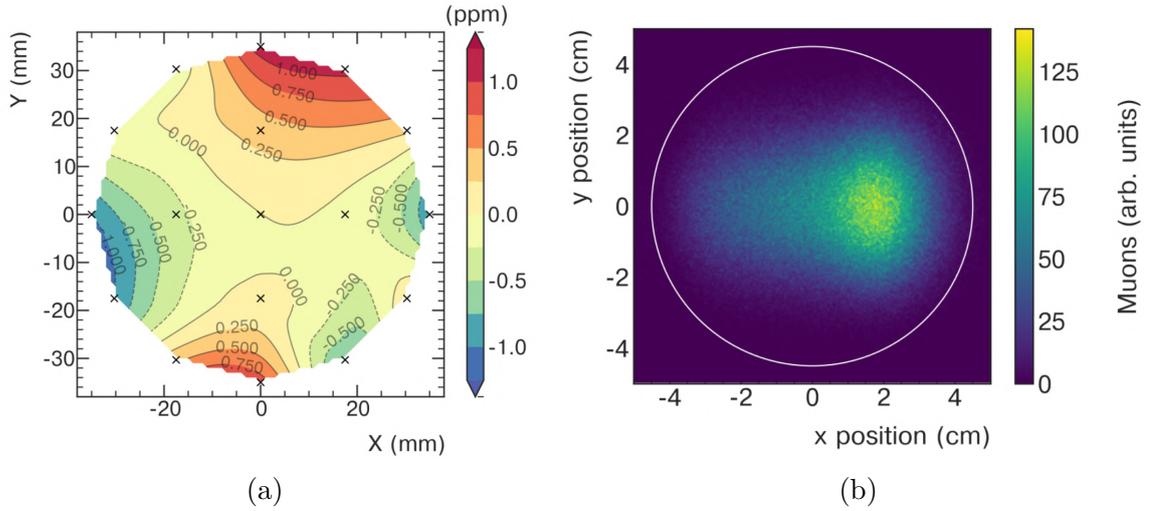
Figure 8.5: (a) Variations in the azimuthally averaged proton Larmor precession frequency $\omega_p(x, y)$ relative to the central probe. The locations of the 17 trolley probes are indicated by ($\times$). (b) Transversal muon distribution $M(x, y)$ as measured by the trackers. Figures from [51].

## 8.5    Corrections to $\tilde{\omega}'_p$

The magnetic field produced by the storage ring is very uniform and stable across time. However, any fast variation of the field strength within the storage time would influence the muon precession frequency and cannot be detected by the NMR probes. Two fast field transients have been observed, one introduced by the electrostatic quadrupoles charging process, which induces mechanical vibrations to the quadrupole plates, and one by the transient currents induced by the kicker magnets. The two correction terms are determined by dedicated measurements and are introduced into the final $a_\mu$ formula:

$$\tilde{\omega}_p \approx \omega_p^m (1 + B_q + B_k). \tag{8.16}$$

The values are listed in Table 8.5.

| Correction term | Symbol | Value | Uncertainty |
|---|---|---|---|
| Quadrupole vibration | $B_q$ | -17 | 92 |
| Kicker transient | $B_k$ | -27 | 37 |

Table 8.5: List of corrections to the measured $\omega_p$ value for Run-1. All values are in ppb.

### 8.5.1 Quadrupole correction $B_q$

The first correction, $B_q$, accounts for the mechanical vibrations of the charged quadrupole plates induced by pulsing them. The NMR probes cannot directly measure this effect to the required precision, primarily due to the skin depth effect of the ~1 cm Aluminum vacuum chamber width that separates the probes from the storage region. The perturbation was measured by a customized set of NMR probes inserted in the quadrupoles region while sealed inside polyether ether ketone (PEEK) plastic tubes for vacuum compatibility. The quadrupole system is composed of four short and four long sections, covering roughly ~43% of the storage ring in total. The NMR measurement time takes 1.2 s, which is almost as long as the full sequence of 16 muon bunches of a booster supercycle. The perturbation time dependence was measured by varying the delay time between the trigger to the quadrupoles and the NMR measurement. Figure 8.6 shows the measured variation of the magnetic field as a function of time.

The measurement has been performed at several positions inside one of the quadrupoles and at the center of the others, in order to determine an average correction. The correction $B_q$ is extracted by weighting the perturbation by the muon decay rate and averaging over the whole ring [51].



|         (a)         |         (b)         |
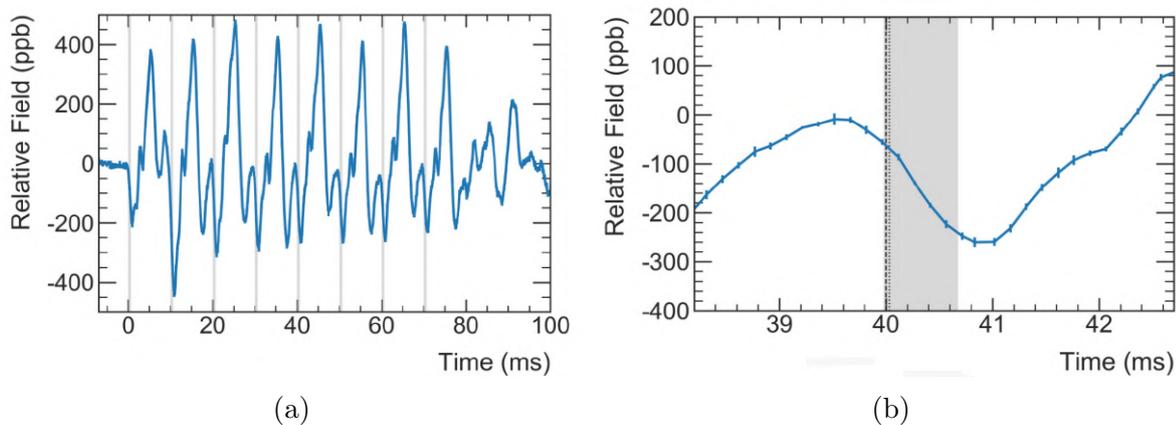
Figure 8.6: (a) Field perturbation induced by the quadrupole vibrations over the course of eight subsequent muon bunches. (b) The same time structure zoomed in to a single beam bunch. Figures from [51].

### 8.5.2 Kicker correction $B_k$

A set of three fast pulsed kicker magnets are placed in the storage ring, roughly 90° after the beam injection point, with the role of shifting the beam center of curvature so

that the correct orbit is achieved after the first turn. The kicker system is composed of three pairs of thin curved aluminum plates, each 1.27 m long, covering 8.5% of the storage ring. The kick itself lasts ∼150 ns and reduces the 1.45 T magnetic field locally by roughly 22 mT. The pulse, however, induces eddy currents in the surrounding metal that last for several milliseconds. These currents generate field perturbations during the beam storage time, affecting the muon precession frequency.

This effect has been measured by two independent Faraday magnetometers, which exploit the rotation of the polarization angle of light when passing through a dielectric material submerged in a magnetic field. If the field is parallel to the direction of propagation of the light, the polarization rotates by:

$$\Delta\theta(t) = V B(t) L \,, \tag{8.17}$$

where $V$ is the Verdet constant of the dielectric material, $B(t)$ is the (varying) magnetic field, and $L$ is the length of the material. The magnetometers consisted of a terbium gallium garnet (TGG) crystal placed vertically between the kicker plates and supported by a structure built without any metal parts. The polarized light was brought to the crystals with optical fibers for one magnetometer and via open space and mirrors for the other magnetometer.

The measurement result is shown in Figure 8.7. Field perturbation is obtained by fitting the data with an exponential function between 30 and 700 $\mu$s:

$$\Delta B(t) = \Delta B(t_0) e^{-(t-t_0)/\tau_B} \,. \tag{8.18}$$

The fractional effect on the muon anomalous precession frequency measurement for a fit starting at $t = t_0 = 30$ $\mu$s is

$$B_k = \frac{\Delta\omega_a}{\omega_a} \approx \frac{\Delta B(t_0) \cdot k}{B(t_0)} \left( \frac{\tau_B}{\tau_B + \tau_\mu} \right)^2 \,, \tag{8.19}$$

where $k$ is a scaling factor equal to 0.085 (corresponding to the kicker coverage) multiplied by 0.94 (corresponding to the estimated spatial dependence of the transient field).

<div align="center">(a)                                                        (b)</div>
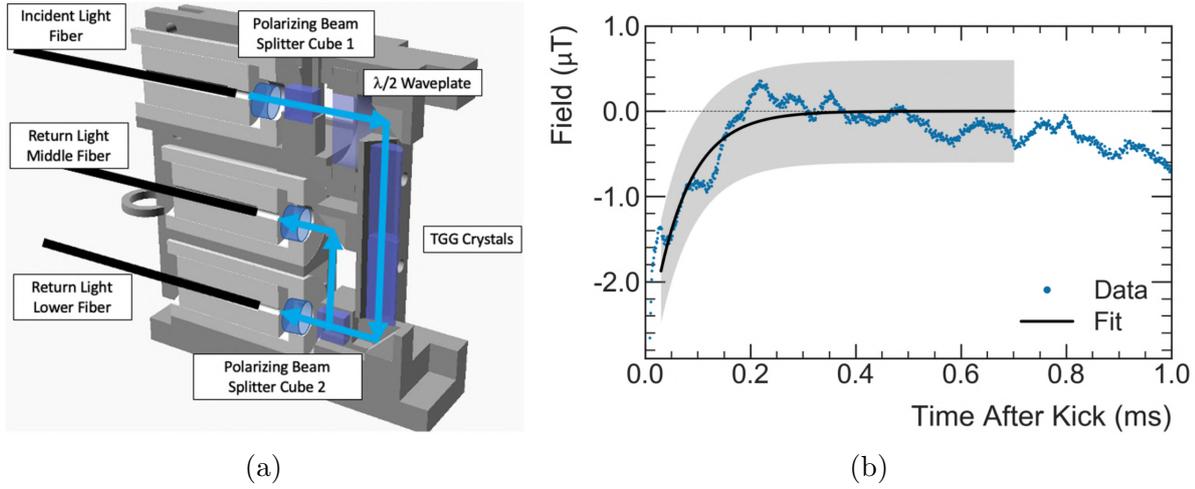
Figure 8.7: (a) Schematic of the fiber magnetometer inserted between the kicker plates. (b) The field perturbation caused by the eddy currents as measured by the fiber magnetometer. The gray shaded band represents the associated uncertainty of $\pm 0.6$ $\mu$T. Figures from [51].

## 8.6   Final Run-1 $a_\mu$ result

The fully corrected and unblinded Run-1 values for $\omega_a$, $\tilde{\omega}'_p$, and their ratio $\mathcal{R}'_\mu$ are summarized in Table 8.6. The weighted average value for $\mathcal{R}'_\mu$ is also computed. The quoted errors are the combination of statistical and systematic uncertainties. A constant fit of the four different $\mathcal{R}'_\mu$ shows a $\chi^2/NDF = 6.8/3$ that corresponds to a $P(\chi^2) = 7.8\%$, indicating a small tension among the datasets. However, this probability is still considered to be a plausible statistical outcome and not indicative of incorrectly estimated uncertainties [11]. A comprehensive list of corrections and uncertainties of all the terms of Equation 8.3 is presented in Table 8.7.

| Dataset  | $\omega_a/2\pi$ [Hz] | $\tilde{\omega}'_p/2\pi$ [Hz] | $\mathcal{R}'_\mu$ |
|----------|----------------------|-------------------------------|--------------------|
| Run-1A   | 229 081.06(28)       | 61 791 871.2(7.1)             | 0.003 707 300 9(45) |
| Run-1B   | 229 081.40(24)       | 61 791 937.8(7.9)             | 0.003 707 302 4(38) |
| Run-1C   | 229 081.26(19)       | 61 791 845.4(7.7)             | 0.003 707 305 7(31) |
| Run-1D   | 229 081.23(16)       | 61 792 003.4(6.6)             | 0.003 707 295 7(26) |
| Combined |                      |                               | 0.003 707 300 3(17) |

Table 8.6: Run-1 final results for $\omega_a$, $\omega_p$, and their ratio $\mathcal{R}'_\mu$

Combining the final value of $\mathcal{R}'_\mu$ with the various corrections and external factors described in this chapter we obtain the Run-1 measurement of the anomalous magnetic

| Quantity | Correction [ppb] | Uncertainty [ppb] |
|---|---|---|
| $\omega_a$ (statistical) | - | 434 |
| $\omega_a$ (systematic) | - | 56 |
| $C_e$ | 489 | 53 |
| $C_p$ | 180 | 13 |
| $C_{ml}$ | -11 | 5 |
| $C_{pa}$ | -158 | 75 |
| $f_{calib}\langle \omega_p'(x,y,\phi) \cdot M(x,y,\phi) \rangle$ | - | 56 |
| $B_q$ | -17 | 92 |
| $B_k$ | -27 | 37 |
| $\mu_p'/\mu_e$ | - | 10 |
| $m_\mu/m_e$ | - | 22 |
| $g_e$ | - | 0 |
| Total systematic | - | 157 |
| Total external factors | - | 25 |
| Total | 544 | 462 |

Table 8.7: Values and uncertainties for the correction terms and uncertainties of the other quantities of Equation 8.3 [11].

moment of the muon:

$$a_\mu(\text{FNAL, Run-1}) = 0.001\,165\,920\,40(54)\,, \tag{8.20}$$

with a final uncertainty of 460 ppb dominated by the statistical error. The result is compatible with the previous BNL measurement[1] [41, 10]:

$$a_\mu(\text{BNL}) = 0.001\,165\,920\,89(63)\,. \tag{8.21}$$

The E989 collaboration has carefully assessed any and all possible correlations between the Fermilab and Brookhaven experiments and has concluded that there are no important correlations that would impact a weighted average to obtain a correct combined result [11]. The combination of the two values gives the new experimental world average

$$a_\mu(\text{Exp, 2021}) = 0.001\,165\,920\,61(41)\,, \tag{8.22}$$

---

[1]The result reported by the BNL collaboration in 2006 has been updated in Equation 8.21 to the newest value for the absolute muon-to-proton magnetic ratio $\lambda = 3.183\,345\,142(71)$ [14]. The change induced in $a_\mu$ with respect to the old BNL value amounts to $+10 \times 10^{-11}$.

with a precision of 350 ppb.

This value increases the discrepancy with the data-driven prediction of the Standard Model quoted by the Theory Initiative [19] to $4.2\sigma$. The tension, however, is significantly smaller between the experimental value and the lattice-QCD prediction provided by the BMW collaboration [27], at the level of $1.5\sigma$. As discussed in Section 1.3.3, independent lattice-QCD groups seem to validate the BMW result, contributing to the new puzzle in the theoretical prediction. These values and discrepancies are shown in Figure 8.8.
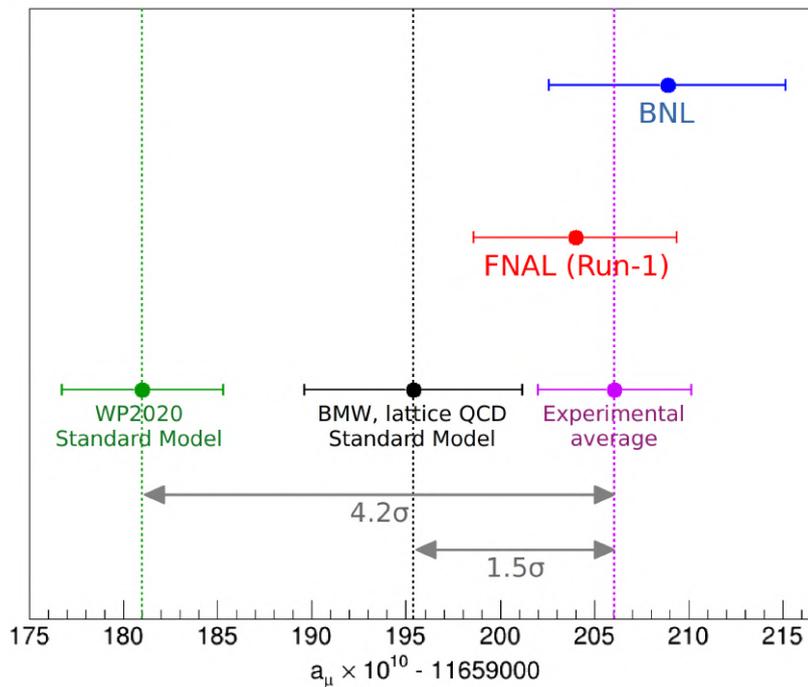


Figure 8.8: The Run-1 measurement [11] of $a_\mu$ compared to the BNL E821 experiment [41] and the Standard Model predictions provided by the Theory Initiative Group [19] and the BMW collaboration [27].

# Conclusions

The anomalous magnetic moment of a particle, $a = \frac{g-2}{2}$, is a fundamental property that encodes all the possible virtual interactions allowed by nature. Its precise measurement, together with the continuously evolving theoretical prediction, has represented an important test of the Standard Model for the last 75 years. The anomalous magnetic moment of the muon, $a_\mu$, is particularly interesting due to its enhanced sensitivity to massive particles participating in the virtual quantum loops. Interactions with not-yet-discovered massive particles could manifest as a discrepancy between a high-precision measurement and the Standard Model prediction. The Muon $g - 2$ experiment at Brookhaven National Laboratory (BNL), concluded in 2001, observed a discrepancy with the theoretical prediction which later grew, thanks to further improvements on the theory side, to $3.7\sigma$. The new E989 experiment at Fermilab was built to improve the experimental measurement by a factor of four and to confirm or disprove the long-standing discrepancy.

The E989 collaboration, of which I am part, published the first $a_\mu$ measurement on April 7th, 2021, relative to the first year of data taking, Run-1, that took place in the months between March and July 2018. The measured value of $a_\mu = 0.001\,165\,920\,40(54)$, with a precision of 460 parts per billion, is in excellent agreement with the BNL measurement, and when combining the two experimental values, the discrepancy with the data-driven theoretical prediction increases to $4.2\sigma$. Recent improvements on the theoretical side, however, are starting to question this long-standing discrepancy. Many independent groups are developing lattice-QCD calculations that provide ab-initio, i.e. without any external input from experimental data, predictions of the hadronic contributions to the muon anomaly. While the hadronic light-by-light $a_\mu^{HLbL}$ term is in agreement with the data-driven estimation, the lowest-order hadronic vacuum polarization $a_\mu^{HVP,LO}$ term is creating a tension between the two methods and is moving the overall value of $a_\mu$ closer to the experimental one. The Fermilab E989 experiment is collecting its last run of data taking which will allow for a reduction of the overall uncertainty to 140 ppb. On the other hand, further improvements on the theoretical front are expected to shine light on the current discrepancies.

The experimental technique of the E989 experiment consists of injecting a polarized beam of 3.1 GeV positive muons in a magnetic storage ring and measuring the muon

spin precession frequency $\omega_a$. The field is measured with a set of Nuclear Magnetic Resonance probes and the muon precession is measured by observing the rate of decay positrons with a set of electromagnetic calorimeters. In addition, two tracker stations are capable of measuring the beam dynamics of the muons as they rotate in the storage ring. The E989 experiment is currently collecting muons for the sixth consecutive year of data taking, Run-6, and has accumulated ∼20 times the statistics of the BNL experiment. My involvement in the key role of *production manager* for the production of these petabyte-sized datasets contributed to speeding up the process by a factor of ∼4 for Run-4 with respect to Run-2, as described in Chapter 4.

As the statistical error continues to improve, increasing attention is being dedicated to reducing systematic uncertainties as much as possible. Many of the highest uncertainties of the Run-1 $\omega_a$ measurement are related to the reconstruction of the positrons hitting the calorimeters. Chapter 5 of this dissertation described the new ReconITA reconstruction, mainly developed by me to provide an independent reconstruction technique and to improve the related systematics. The new clustering algorithm improves the separation of overlapping positrons by a factor of ∼3 with respect to Run-1, greatly reducing the dominant systematic uncertainty of $\omega_a$. This algorithm is used by the $\omega_a$-*Europe* group for the analysis of the upcoming measurement of Run-2 and Run-3. The new pulse fitting technique improves the separation of two positron signals hitting the same calorimeter crystal at the nanosecond timescale, while also recovering missed low-energy pulses in the long tails of the SiPMs waveforms. This is now part of the official production of Run-4, Run-5, and Run-6 and will reduce the slow-term effect, first observed in Run-1, by a factor ∼3. The calorimeter gain corrections extracted with the precise INFN laser calibration system, that I have developed, have been presented in Chapter 5. New refinements including the temperature dependency of the short-term correction and the determination of the in-fill gain recovery time constant with dedicated studies improve the final gain stability.

Finally, the complete $\omega_a$ determination from the analysis of the reconstructed positrons and the estimation of all the related systematic uncertainties has been presented in Chapters 6 and 7, and it relies on the reconstruction algorithm I have developed. A first version of the ReconITA program has been used for Run-2 and Run-3 reconstructions whose result is expected to be published in Spring 2023. The final publication, with the complete statistics, is foreseen for early 2025 with the realistic goal of performing even better than the design expected uncertainty of 140 ppb.

# Acknowledgments

This thesis marks the end of a very exciting and formative journey that has characterized my academic and personal growth over the last three years. It was not always a smooth ride, as distance and isolation were the recurring themes over many of these months. However, everything paid off on the day that we released the first measurement: a small piece was added to the puzzle of human knowledge and I can say I was one of the many hands that helped put it into place. Many people accompanied me during this period: some made this work possible at all, and others always helped me push through the entirety of it. First of all, I would like to thank Marco Incagli, my supervisor and mentor, for his support, motivation, and faith in my abilities. He encouraged me throughout my academic career in g-2 and he always managed to get me back on track when the work does not give the expected results. This work would not have been possible without his support. I want to thank the entire *g-2it* Italian group and in particular Graziano for his endless enthusiasm and precious advice. I want to thank Anna, Elia, Lorenzo, Matteo, and Riccardo, for being much more than colleagues, for all the great experiences we had together in Italy and at Fermilab, and for being always available when I need some help, both at work and in life. Elia, Riccardo: sembrava impossibile, ma ce l'abbiamo fatta! I thank Adam, Liang, Tammy, Yuri, and the rest of the offline team for the very formative and fun year I had with you producing the g-2 data. I would like to thank the entire g-2 collaboration for this fantastic experience. The next part, dedicated to my family, will be in Italian. Vorrei ringraziare di cuore la mia famiglia, che non mi ha mai fatto mancare nulla e che mi ha sempre supportato durante la mia vita, dandomi sempre l'opportunità di esplorare le mie passioni. Un pensiero speciale va alla nonna Elde, che purtroppo oggi non c'è più, la cui bavarese al cioccolato rimarrà sempre nel mio cuore. Infine, ma non per importanza, il ringraziamento più grande va a Martina, la mia compagna di vita e il mio grande amore, che rende tutto questo meritevole di essere vissuto. Mi hai sempre supportato (e sopportato) nei momenti di difficoltà e di stress, soprattutto quando grandi distanze ci hanno separato. Mi ritengo veramente fortunato ad averti al mio fianco e mi auguro di riuscire ad azzerare presto queste distanze, permettendoci di vivere ogni giorno come piace a noi: mangiando, esplorando il mondo e, come quando ci siamo conosciuti, guardando le stelle. Grazie.

# Appendix A

# Run-1 results

This appendix reports the fit values produced by the $\omega_a$-*Europe* analysis on the four Run-1 datasets. The fitting procedure is described in Chapter 6. Only the results from the AMethod weighting technique are reported here, as it is the technique used for the final Run-1 $\omega_a$ measurement combination. However, these results correspond to a single randomization seed and the $\omega_a$ values are different from the ones reported as the final results. This is because the final result is the average of several fits made using different seeds in the hit time randomization algorithm.

Together with the central values and statistical uncertainties, correlation matrices produced with the `ROOT` method `TFitResult::GetCorrelationMatrix` are reported as well.

## A.1  Dataset Run-1A

| Parameter | Unit | Result | Uncertainty |
|:---------:|:----:|:------:|:-----------:|
| $N$ | $-$ | 1617963. | 369. |
| $\tau$ | $\mu$s | 64.4388 | 0.0033 |
| $A$ | $-$ | 0.360934 | 0.000038 |
| $R$ | ppm | $-23.73$ | 1.19 |
| $\phi$ | rad | 2.09052 | 0.00020 |
| $A_{CBO}$ | $1 \times 10^{-3}$ | 4.65 | 0.13 |
| $\tau_{CBO}$ | $\mu$s | 186. | 13. |
| $\phi_{CBO}$ | rad | 4.239 | 0.027 |
| $\omega_{CBO}$ | rad/$\mu$s | 2.33816 | 0.00035 |
| $A_{VW}$ | $-$ | 0.0110 | 0.0044 |
| $\omega_{VW}$ | rad/$\mu$s | 14.392 | 0.011 |
| $\phi_{VW}$ | rad | 5.61 | 0.43 |
| $k_{LM}$ | $-$ | 0.01255 | 0.00040 |
| $A_A$ | $1 \times 10^{-3}$ | 0.63 | 0.23 |
| $\phi_A$ | rad | 4.91 | 0.36 |
| $A_\phi$ | $1 \times 10^{-3}$ | 0.39 | 0.24 |
| $\phi_\phi$ | rad | 5.23 | 0.61 |
| $A_{2CBO}$ | $1 \times 10^{-3}$ | 0.185 | 0.088 |
| $\phi_{2CBO}$ | rad | 4.66 | 0.47 |
| $\tau_{VW}$ | $\mu$s | 17.2 | 3.1 |

Table A.1: Fitted AMethod parameters from the full 20 parameters fit of the Run-1A dataset.
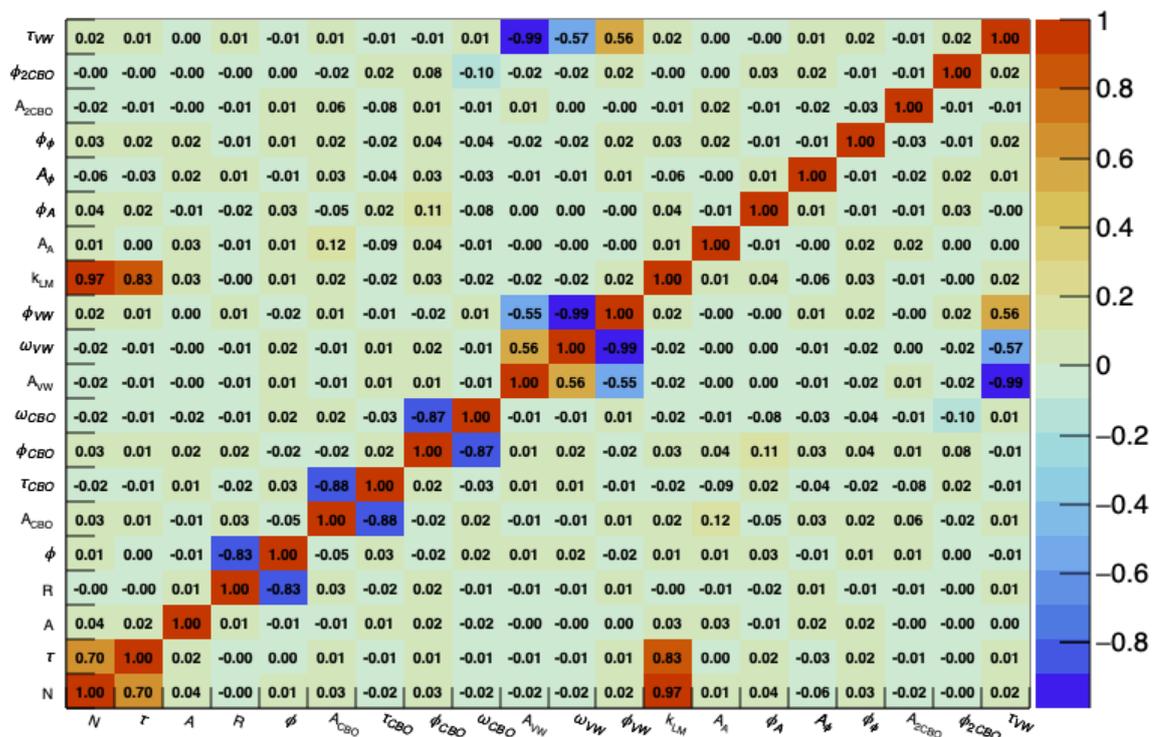
Figure A.1: Correlation matrix from the Run-1A dataset AMethod fit.

## A.2   Dataset Run-1B

| Parameter | Unit | Result | Uncertainty |
|-----------|------|--------|-------------|
| $N$ | – | 2251357. | 1038. |
| $\tau$ | $\mu$s | 64.4275 | 0.0027 |
| $A$ | – | 0.360798 | 0.000032 |
| $R$ | ppm | $-22.39$ | 1.01 |
| $\phi$ | rad | 2.08272 | 0.00017 |
| $A_{CBO}$ | $1 \times 10^{-3}$ | 4.09 | 0.15 |
| $\tau_{CBO}$ | $\mu$s | 114.9 | 7.3 |
| $\phi_{CBO}$ | rad | 3.650 | 0.036 |
| $\omega_{CBO}$ | rad/$\mu$s | 2.61497 | 0.00052 |
| $A_{VW}$ | – | 0.0110 | 0.0034 |
| $\omega_{VW}$ | rad/$\mu$s | 12.8019 | 0.0071 |
| $\phi_{VW}$ | rad | 1.99 | 0.28 |
| $k_{LM}$ | $1 \times 10^{-3}$ | 7.36 | 0.59 |
| $A_A$ | $1 \times 10^{-3}$ | 0.66 | 0.24 |
| $\phi_A$ | rad | 4.40 | 0.36 |
| $A_\phi$ | $1 \times 10^{-3}$ | 0.17 | 0.25 |
| $\phi_\phi$ | rad | 0.94 | 1.51 |
| $A_{2CBO}$ | $1 \times 10^{-3}$ | 0.14 | 0.11 |
| $\phi_{2CBO}$ | rad | 4.64 | 0.77 |
| $\tau_{VW}$ | $\mu$s | 18.0 | 2.6 |

Table A.2: Fitted AMethod parameters from the full 20 parameters fit of the Run-1B dataset.

Figure A.2: Correlation matrix from the Run-1B dataset AMethod fit.

## A.3   Dataset Run-1C

| Parameter | Unit | Result | Uncertainty |
|:---------:|:----:|:------:|:-----------:|
| $N$ | — | 3433721. | 1009. |
| $\tau$ | $\mu$s | 64.4378 | 0.0021 |
| $A$ | — | 0.361882 | 0.000026 |
| $R$ | ppm | $-22.99$ | 0.81 |
| $\phi$ | rad | 2.08020 | 0.00013 |
| $A_{CBO}$ | $1 \times 10^{-3}$ | 4.253 | 0.083 |
| $\tau_{CBO}$ | $\mu$s | 217. | 11. |
| $\phi_{CBO}$ | rad | 3.933 | 0.019 |
| $\omega_{CBO}$ | rad/$\mu$s | 2.60999 | 0.00023 |
| $A_{VW}$ | $1 \times 10^{-3}$ | 4.40 | 0.69 |
| $F_y$ | — | 1.01563 | 0.00025 |
| $\phi_{VW}$ | rad | 1.34 | 0.14 |
| $k_{LM}$ | $1 \times 10^{-3}$ | 6.59 | 0.38 |
| $A_A$ | $1 \times 10^{-3}$ | 0.36 | 0.15 |
| $\phi_A$ | rad | 2.92 | 0.41 |
| $A_\phi$ | $1 \times 10^{-3}$ | 0.64 | 0.15 |
| $\phi_\phi$ | rad | 4.36 | 0.24 |
| $A_{2CBO}$ | $1 \times 10^{-3}$ | 0.126 | 0.055 |
| $\phi_{2CBO}$ | rad | 4.54 | 0.44 |
| $\tau_{VW}$ | $\mu$s | 29.6 | 3.1 |
| $A_y$ | $1 \times 10^{-3}$ | 0.39 | 0.09 |
| $\phi_y$ | rad | 4.83 | 0.22 |

Table A.3: Fitted AMethod parameters from the full 22 parameters fit of the Run-1C dataset.
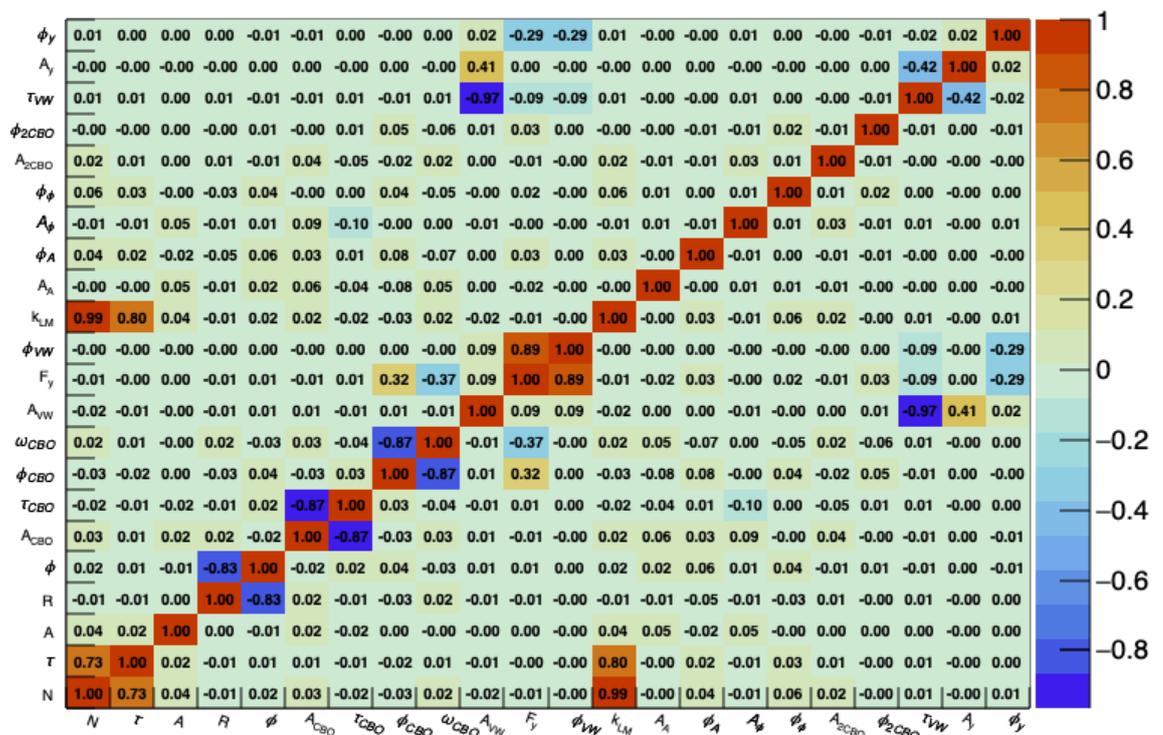
Figure A.3: Correlation matrix from the Run-1C dataset AMethod fit.

## A.4   Dataset Run-1D

| Parameter | Unit | Result | Uncertainty |
|:---:|:---:|:---:|:---:|
| $N$ | — | 6878558. | 3322. |
| $\tau$ | $\mu$s | 64.4415 | 0.0021 |
| $A$ | — | 0.364085 | 0.000021 |
| $R$ | ppm | $-23.08$ | 0.67 |
| $\phi$ | rad | 2.07557 | 0.00013 |
| $A_{CBO}$ | $1 \times 10^{-3}$ | 3.83 | 0.10 |
| $\tau_{CBO}$ | $\mu$s | 198. | 10. |
| $\phi_{CBO}$ | rad | 0.219 | 0.025 |
| $\omega_{CBO}$ | rad/$\mu$s | 2.33591 | 0.00025 |
| $A_{VW}$ | $1 \times 10^{-3}$ | 1.61 | 0.44 |
| $F_y$ | — | 1.01545 | 0.00034 |
| $\phi_{VW}$ | rad | 2.46 | 0.30 |
| $k_{LM}$ | — | 0.01314 | 0.00063 |
| $A_A$ | $1 \times 10^{-3}$ | 0.48 | 0.14 |
| $\phi_A$ | rad | 5.40 | 0.28 |
| $A_\phi$ | $1 \times 10^{-3}$ | 0.25 | 0.14 |
| $\phi_\phi$ | rad | 2.75 | 0.57 |
| $A_{2CBO}$ | $1 \times 10^{-3}$ | 0.197 | 0.058 |
| $\phi_{2CBO}$ | rad | 2.68 | 0.29 |
| $\tau_{VW}$ | $\mu$s | 49.9 | 9.7 |
| $A_y$ | $1 \times 10^{-3}$ | 0.103 | 0.060 |
| $\phi_y$ | rad | 3.35 | 0.57 |

Table A.4: Fitted AMethod parameters from the full 22 parameters fit of the Run-1D dataset.
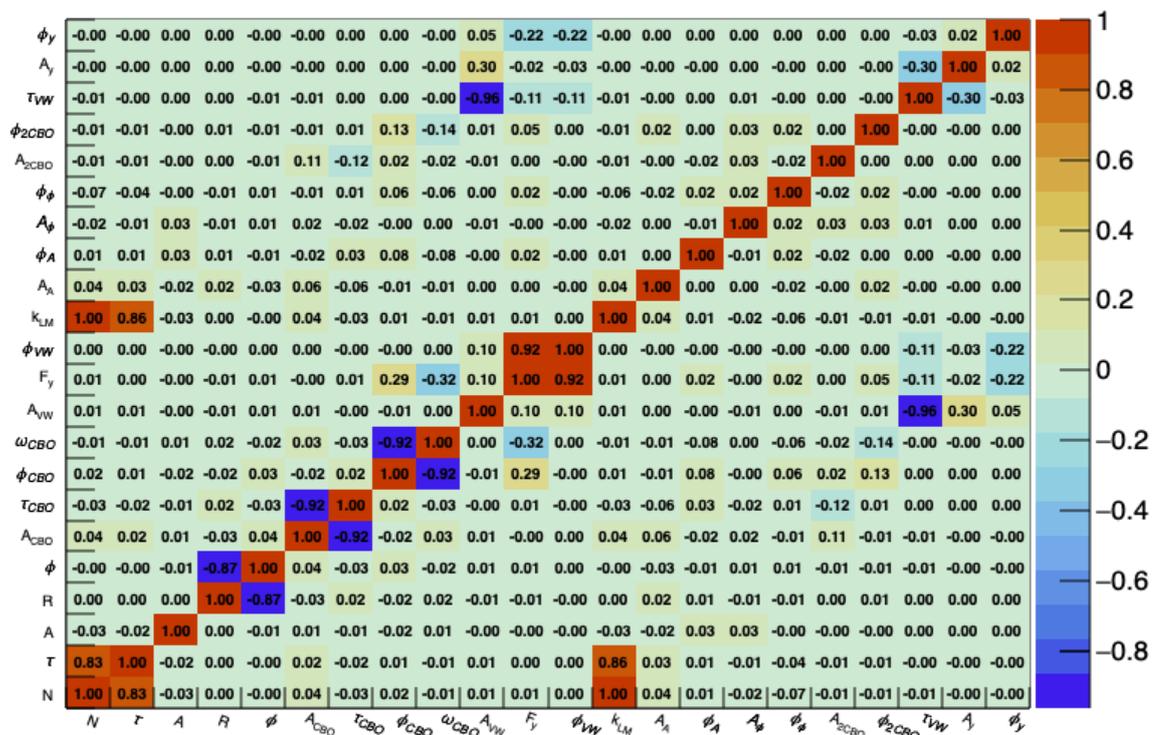
Figure A.4: Correlation matrix from the Run-1D dataset AMethod fit.

# Bibliography

[1] Paul Adrien Maurice Dirac and Ralph Howard Fowler. **The quantum theory of the electron**. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character, **117**(778):610–624, 1928.

[2] Matthew D. Schwartz. Quantum Field Theory and the Standard Model. Cambridge University Press, 3 2014.

[3] P. Kusch and H. M. Foley. **The Magnetic Moment of the Electron**. Phys. Rev., **74**:250–263, Aug 1948.

[4] Richard P. Feynman. The character of physical law. Cambridge, MA: The MIT press, 1967.

[5] Tatsumi Aoyama, Masashi Hayakawa, Toichiro Kinoshita, and Makiko Nio. **Tenth-order electron anomalous magnetic moment: Contribution of diagrams without closed lepton loops**. Phys. Rev. D, **91**:033006, Feb 2015.

[6] Tatsumi Aoyama, Masashi Hayakawa, Toichiro Kinoshita, and Makiko Nio. **Erratum: Tenth-order electron anomalous magnetic moment: Contribution of diagrams without closed lepton loops [Phys. Rev. D 91, 033006 (2015)]**. Phys. Rev. D, **96**:019901, Jul 2017.

[7] Richard H. Parker, Chenghui Yu, Weicheng Zhong, et al. **Measurement of the fine-structure constant as a test of the Standard Model**. Science, **360**(6385):191–195, 2018.

[8] Léo Morel, Zhibin Yao, Pierre Cladé, and Saïda Guellati-Khélifa. **Determination of the fine-structure constant with an accuracy of 81 parts per trillion**. Nature, **588**(7836):61–65, 2020.

[9] X. Fan, T. G. Myers, B. A. D. Sukra, and G. Gabrielse. **Measurement of the Electron Magnetic Moment**. Phys. Rev. Lett., **130**:071801, Feb 2023.

[10] Particle Data Group, R L Workman, V D Burkert, et al. **Review of Particle Physics**. Progress of Theoretical and Experimental Physics, **2022**(8), 08 2022. 083C01.

[11] B. Abi, T. Albahri, S. Al-Kilani, et al. **Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm**. Phys. Rev. Lett., **126**:141801, Apr 2021.

[12] M. Davier, S. Eidelman, A. Höcker, and Z. Zhang. **Confronting spectral functions from ee annihilation and $\tau$ decays: consequences for the muon magnetic moment**. The European Physical Journal C, **27**(4):497–521, apr 2003.

[13] S. Eidelman and M. Passera. **THEORY OF THE $\tau$ LEPTON ANOMALOUS MAGNETIC MOMENT**. Modern Physics Letters A, **22**(03):159–179, 2007.

[14] Eite Tiesinga, Peter J. Mohr, David B. Newell, and Barry N. Taylor. **CODATA recommended values of the fundamental physical constants: 2018**. Rev. Mod. Phys., **93**:025010, Jun 2021.

[15] Stephen Gasiorowicz and Jonathan L. Rosner. **Hadron spectra and quarks**. American Journal of Physics, **49**(10):954–984, 1981.

[16] D. Hanneke, S. Fogwell Hoogerheide, and G. Gabrielse. **Cavity control of a single-electron quantum cyclotron: Measuring the electron magnetic moment**. Phys. Rev. A, **83**:052122, May 2011.

[17] Fred Jegerlehner and Andreas Nyffeler. **The muon g-2**. Physics Reports, **477**(1):1–110, 2009.

[18] Tatsumi Aoyama, Masashi Hayakawa, Toichiro Kinoshita, and Makiko Nio. **Complete Tenth-Order QED Contribution to the Muon $g-2$**. Phys. Rev. Lett., **109**:111808, Sep 2012.

[19] T. Aoyama, N. Asmussen, M. Benayoun, et al. **The anomalous magnetic moment of the muon in the Standard Model**. Physics Reports, **887**:1–166, 2020. The anomalous magnetic moment of the muon in the Standard Model.

[20] Richard H. Parker, Chenghui Yu, Weicheng Zhong, et al. **Measurement of the fine-structure constant as a test of the Standard Model**. Science, **360**(6385):191–195, 2018.

[21] Roman Jackiw and Steven Weinberg. **Weak-Interaction Corrections to the Muon Magnetic Moment and to Muonic-Atom Energy Levels**. Phys. Rev. D, **5**:2396–2398, May 1972.

[22] C. Gnendiger, D. Stöckinger, and H. Stöckinger-Kim. **The electroweak contributions to $(g-2)_\mu$ after the Higgs-boson mass measurement**. Phys. Rev. D, **88**:053005, Sep 2013.

[23] Bouchiat, Claude and Michel, Louis. **La résonance dans la diffusion méson méson le moment magnétique anormal du méson**. J. Phys. Radium, **22**(2):121, 1961.

[24] Loyal Durand. **Pionic Contributions to the Magnetic Moment of the Muon**. Phys. Rev., **128**:441–448, Oct 1962.

[25] Stanley J. Brodsky and Eduardo de Rafael. **Suggested Boson-Lepton Pair Couplings and the Anomalous Magnetic Moment of the Muon**. Phys. Rev., **168**:1620–1622, Apr 1968.

[26] M. Gourdin and E. De Rafael. **Hadronic contributions to the muon g-factor**. Nuclear Physics B, **10**(4):667–674, 1969.

[27] Sz. Borsanyi, Z. Fodor, J. N. Guenther, et al. **Leading hadronic contribution to the muon magnetic moment from lattice QCD**. Nature, **593**(7857):51–55, apr 2021.

[28] G. Colangelo, A.X. El-Khadra, M. Hoferichter, et al. **Data-driven evaluations of Euclidean windows to scrutinize hadronic vacuum polarization**. Physics Letters B, **833**:137313, 2022.

[29] Marco Cè, Antoine Gérardin, Georg von Hippel, et al. **The hadronic running of the electroweak couplings from lattice QCD**. In Proceedings of 41st International Conference on High Energy physics — PoS(ICHEP2022), **414**, page 823, 2022.

[30] Marco Cè, Antoine Gérardin, Georg von Hippel, et al. **Window observable for the hadronic vacuum polarization contribution to the muon $g - 2$ from lattice QCD**, 2022.

[31] C. Alexandrou, S. Bacchio, P. Dimopoulos, et al. **Lattice calculation of the short and intermediate time-distance hadronic vacuum polarization contributions to the muon magnetic moment using twisted-mass fermions**, 2022.

[32] Richard L. Garwin, Leon M. Lederman, and Marcel Weinrich. **Observations of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: the Magnetic Moment of the Free Muon**. Phys. Rev., **105**:1415–1417, Feb 1957.

[33] J M Cassels, T W O'Keeffe, M Rigby, et al. **Experiments with a Polarized Muon Beam**. Proceedings of the Physical Society. Section A, **70**(7):543, jul 1957.

[34] D. P. Hutchinson, J. Menes, G. Shapiro, and A. M. Patlach. **Magnetic Moment of the Positive Muon**. Phys. Rev., **131**:1351–1362, 1963.

[35] G. Charpak, F. J. M. Farley, R. L. Garwin, et al. **Measurement of the Anomalous Magnetic Moment of the Muon**. Phys. Rev. Lett., **6**:128–132, Feb 1961.

[36] G. Charpak, F.J.M. Farley, R.L. Garwin, et al. **A new measurement of the anomalous magnetic moment of the muon**. Physics Letters, **1**(1):16–20, 1962.

[37] Georges Charpak, Francis J M Farley, Richard L Garwin, et al. **The anomalous magnetic moment of the muon**. Nuovo Cimento, **37**:1241–1363, 1965.

[38] J. Bailey, W. Bartl, G. Von Bochmann, et al. **Precision measurement of the anomalous magnetic moment of the muon**. Physics Letters B, **28**(4):287–290, 1968.

[39] J. Bailey, K. Borer, F. Combley, et al. **New measurement of (g-2) of the Muon**. Physics Letters B, **55**(4):420–424, 1975.

[40] J. BAILEY, K. BORER, F. COMBLEY, ET AL. **The anomalous magnetic moment of positive and negative muons**. Physics Letters B, **67**(2):225–230, 1977.

[41] G. W. BENNETT, B. BOUSQUET, H. N. BROWN, ET AL. **Final report of the E821 muon anomalous magnetic moment measurement at BNL**. Phys. Rev. D, **73**:072003, Apr 2006.

[42] K HAGIWARA, A.D MARTIN, DAISUKE NOMURA, AND T TEUBNER. **The SM prediction of g-2 of the muon**. Physics Letters B, **557**(1):69–75, 2003.

[43] CHISLETT, REBECCA. **The muon EDM in the g-2 experiment at Fermilab**. EPJ Web of Conferences, **118**:01005, 2016.

[44] M ABE, S BAE, G BEER, ET AL. **A new approach for measuring the muon anomalous magnetic moment and electric dipole moment**. Progress of Theoretical and Experimental Physics, **2019**(5), 05 2019. 053C02.

[45] G ABBIENDI. **Letter of Intent: the MUonE project**. Technical report, CERN, Geneva, 2019. The collaboration has not yet a structure, therefore the names above are for the moment an indication of contacts.

[46] RICCARDO NUNZIO PILATO. **Status of the MUonE experiment**. In Proceedings of 41st International Conference on High Energy physics — PoS(ICHEP2022), **414**, page 788, 2022.

[47] C.M. CARLONI CALAME, M. PASSERA, L. TRENTADUE, AND G. VENANZONI. **A new approach to evaluate the leading hadronic corrections to the muon g-2**. Physics Letters B, **746**:325–329, 2015.

[48] L.H. THOMAS B.A. **I. The kinematics of an electron with an axis**. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, **3**(13):1–22, 1927.

[49] J. GRANGE, V. GUARINO, P. WINTER, ET AL. **Muon (g-2) Technical Design Report**, 2015.

[50] AARON T. FIENBERG. Measuring the Precession Frequency in the E989 Muon $g-2$ Experiment. PhD thesis, 1 2019.

[51] T. ALBAHRI, A. ANASTASI, K. BADGLEY, ET AL. **Magnetic-field measurement and analysis for the Muon $g-2$ Experiment at Fermilab**. Phys. Rev. A, **103**:042208, Apr 2021.

[52] A. Anastasi, A. Basti, F. Bedeschi, et al. **The laser-based gain monitoring system of the calorimeters in the Muon g.2 experiment at Fermilab**. Journal of Instrumentation, **14**(11):P11025, nov 2019.

[53] T. Albahri, A. Anastasi, K. Badgley, et al. **Beam dynamics corrections to the Run-1 measurement of the muon anomalous magnetic moment at Fermilab**. Phys. Rev. Accel. Beams, **24**:044002, Apr 2021.

[54] A.P. Schreckenberger, D. Allspach, D. Barak, et al. **The fast non-ferric kicker system for the Muon g-2 Experiment at Fermilab**. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **1011**:165597, 2021.

[55] J. Kaspar, A.T. Fienberg, D.W. Hertzog, et al. **Design and performance of SiPM-based readout of PbF2 crystals for high-rate, precision timing applications**. Journal of Instrumentation, **12**(01):P01009, jan 2017.

[56] David Allen Sweigart. A Measurement of the Anomalous Precession Frequency of the Positive Muon. PhD thesis, Cornell University, 2020.

[57] S. Agostinelli, J. Allison, K. Amako, et al. **Geant4—a simulation toolkit**. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **506**(3):250–303, 2003.

[58] Nathan Seth Froemming. Optimization of Muon Injection and Storage in the Fermilab g-2 Experiment: From Simulation to Reality. PhD thesis, University of Washington, 2019.

[59] Kyoko Makino and Martin Berz. **COSY INFINITY Version 9**. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **558**(1):346–350, 2006. Proceedings of the 8th International Computational Accelerator Physics Conference.

[60] D. Sagan. **Bmad: A relativistic charged particle simulation library**. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **558**(1):356–359, 2006. Proceedings of the 8th International Computational Accelerator Physics Conference.

[61] Bokeh Development Team. Bokeh: Python library for interactive visualization, 2018.

[62] FRED JAMES AND MATTHIAS WINKLER. **MINUIT User's Guide**. 6 2004.

[63] MENGEL, MARC, WHITE, STEPHEN, PODSTAVKOV, VLADIMIR, ET AL. **Production Operations Management System (POMS) for Fermilab Experiments**. EPJ Web Conf., **245**:03024, 2020.

[64] C GREEN, J KOWALKOWSKI, M PATERNO, ET AL. **The art framework**. Journal of Physics: Conference Series, **396**(2):022020, dec 2012.

[65] KENNETH RICHARD HERNER, DENNIS BOX, JOE BOYD, ET AL. **The FIFE Project at Fermilab: Computing for Experiments**. In Proceedings of 38th International Conference on High Energy Physics — PoS(ICHEP2016), **282**, page 176, 2017.

[66] A MOIBENKO, J BAKKEN, E BERMAN, ET AL. **The status of Fermilab Enstore Data Storage System**. 2005.

[67] LAURA KELTON. AN ENERGY-INTEGRATED ANALYSIS FOR MEASURING THE ANOMALOUS PRECESSION FREQUENCY FOR THE MUON g - 2 EXPERIMENT AT FERMILAB. PhD thesis, University of Kentucky, 2022.

[68] T. ALBAHRI, A. ANASTASI, A. ANISENKOV, ET AL. **Measurement of the anomalous precession frequency of the muon in the Fermilab Muon $g-2$ Experiment**. Phys. Rev. D, **103**:072002, Apr 2021.

[69] JASON HEMPSTEAD. Measuring the anomalous precession frequency $\omega_a$ for the Muon $g-2$ experiment. PhD thesis, U. Washington, Seattle (main), 2021.

[70] PAOLO GIROTTI. Study of the systematics on the precession frequency measurement at the Muon g-2 experiment at Fermilab. Master's thesis, University of Bologna, 2019.

[71] HANNAH BINNEY. Precession frequency analysis for Run-2 and Run-3 of the Muon $g-2$ experiment. PhD thesis, University of Washington, 2022.

[72] G.W. BENNETT, B. BOUSQUET, H.N. BROWN, ET AL. **Statistical equations and methods applied to the precision muon (g-2) experiment at BNL**. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **579**(3):1096–1116, 2007.

[73] WILLIAM D PHILLIPS, WILLIAM E COOKE, AND DANIEL KLEPPNER. **Magnetic Moment of the Proton in H2O in Bohr Magnetons**. Metrologia, **13**(4):179, oct 1977.

[74] VALASSI, ANDREA AND CHIERICI, ROBERTO. **Information and treatment of unknown correlations in the combination of measurements using the BLUE method**. Eur. Phys. J. C, **74**(3):2717, 2014.

[75] MIDHAT FAROOQ, TIMOTHY CHUPP, JOE GRANGE, ET AL. **Absolute Magnetometry with $^3$He**. Phys. Rev. Lett., **124**:223001, Jun 2020.