

Chapter 5

vs. time

ω_a Measurement

The measurement of ω_a is determined by counting the number of detected positrons in the calorimeters above some energy threshold [as described in Section 2.1]. Doing so results in a histogram of counts ~~that~~ which is modulated by ω_a , Figure 2.4. Fitting for the frequency allows ω_a to be extracted. The ω_a measurement therefore consists of the steps needed to construct the histogram of counts, the fitting of that histogram, and any systematic studies ~~done in the analysis~~.
the necessary

5.1 Reconstruction of decay positron hits

incident ?

The calorimeters measure hit times and energies of impacting particles, where these hit times and energies are determined from the raw SiPM signals and a reconstruction procedure. In E989 there are two overall separate reconstruction algorithms, ReconWest and ReconEast, both written in the *art* framework similar to the tracking reconstruction. Each of these reconstruction algorithms is modularized, and the steps of the reconstruction process can be switched in and out at will. Using separate reconstruction methods gives confidence in any final results by removing single points of failure. The reconstruction method used in this analysis is ReconWest. A summary of its details will be presented here. A more thorough description is detailed by A. Fienberg [87].

The raw data are digitized waveforms, which are voltage versus time traces output from each SiPM for each calorimeter crystal hit. Due to the incredible amount of

data coming in with the high muon fill rate, only those pulses which exceed some threshold are saved to disk. An online processing system checks the traces against this pre-configured threshold by passing all of the data through a GPU farm [88]. If any trace is found above threshold, then the data ~~is~~ ^{are} saved from every SiPM in every calorimeter, for a time range around the over-threshold trace. This time range is called a time island, similar to that in the tracking reconstruction, and typically has a width of 40 ns [87].

The traces are then fit with templates in order to extract the area and peak times of any present pulses. Each SiPM has its own templates, one for positrons and one for laser pulses. These templates are extracted from data, where each template is determined by collecting many single pulse traces from a SiPM, normalizing by pulse area, aligning in time, and averaging them. These templates were checked against many systematic effects in order to make sure that the constructed templates did not bias the energy or time measurements, such as hit angle, energy (pulse size), position, and rate, as well as aging effects [60, 87]. Each trace is fit using a χ^2 minimization algorithm with the corresponding SiPM templates in order to determine the time and energy of the hit. In order to fit for multiple pulses in a single time island, the fitting procedure first fits with a single template, and then checks the residuals for any remaining peaks. If peaks exist above some threshold, then the fitting is repeated until all pulses have been fit. The time measurement performance in the pulse finding was found to be unaffected by the number of pulses in a time island, and there is 100% pileup separation at 5 ns [87]. See Figure 5.1 for a typical single template fit to a SiPM trace.

Once a pulse has been fit with a template, the pulse area needs to be converted to real energy units using an energy calibration procedure. *Several* A couple of different techniques exist that can be used, including a method that counts photo-statistics seen in

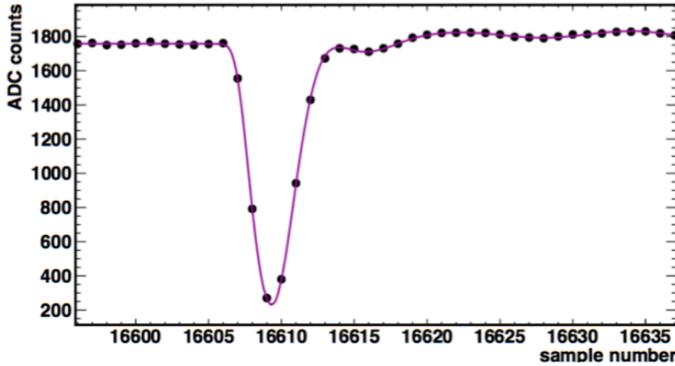


Figure 5.1: A template fit in purple to a SiPM trace delineated by the black points which is in units of ADC counts. Plot courtesy of Aaron Fienberg.

the SiPMs [87]. The default method used is a comparison of lost muon energy signatures in the calorimeters. As described in Reference [89], muons lost from the storage ring can spiral inward and hit consecutive calorimeters with a specific time separation between calorimeter hits. These lost muons are minimum-ionizing particles, and thus leave a very distinct energy signature in the crystals, see Section [5.3.3]. Selecting on the time signature allows hits corresponding to lost muons to be isolated, and the energy signature can be used to determine the appropriate conversions from area to energy¹.

The energy calibration for positron hits as compared to lost muon hits then needs to be determined. Again there are several couple of different techniques, including a comparison of endpoint energies for high energy positrons which tail off at the magic momentum of 3.094 GeV, and comparison with simulation. The default technique is to calibrate the energies such that the optimal energy threshold for the ω_a analysis is near 1.7 GeV [87]. Ultimately the energy calibration is not very important because it is not the energy units that really matter. Most important is the number of positrons above some energy threshold, where that threshold can be optimized em-

¹Different channels can also be equalized based on the energy signatures.

What does this statement mean?

pirically. In fact, the entire ω_a analysis could be done without even considering the energy of the incident positrons, and only considering the area of the SiPM pulses².

Each pulse fit now has an associated energy and time. Because the measurement of ω_a depends heavily on the time reconstruction since the analysis is a frequency extraction, pulse times must be corrected for various effects in order to reach the precision goal. The fitted times for each pulse need to be aligned on a fill-by-fill basis relative to the injection time of the beam, corrected for any channel differences due to differing pulse shapes or fiber lengths, and corrected for any calorimeter time misalignments due to the use of different laser system components. The fill-by-fill alignment is corrected for using the T0 detector as described in Section 3.1.1. The channel differences are corrected by aligning calorimeter channels in time using signals from islands with large simultaneous pulses in neighboring crystals. Calorimeters are time aligned using lost muon coincident events as described before. Once the times of the pulse fits or crystal hits have been determined, the energies can be corrected appropriately for gain effects measured by the laser system. As described in Section 3.2.3, the laser calibration system corrects for out-of fill, STDP, and in-fill gain effects, in that order. Figure 5-2 shows an in-fill gain function fit to data for a single calorimeter. Systematic effects for corrected gain effects are studied in Section 5.5.2.

The last part of the calorimeter reconstruction is the clustering. Clustering is the stage which takes the individual template fit results from separate crystals, and turns them into the times, energies, and positions of decay positron impacts. For a time island with a single positron impact, the procedure is straightforward. The energy for the positron hit cluster is the sum of the individual hit crystal energies. The time for the cluster is taken as the time of the maximum energy hit in the island. This works

²This statement ignores the effects of pileup which must be accounted for, and applies for a threshold style analysis, and not for other analysis methods which depend on the energy of the pulses.

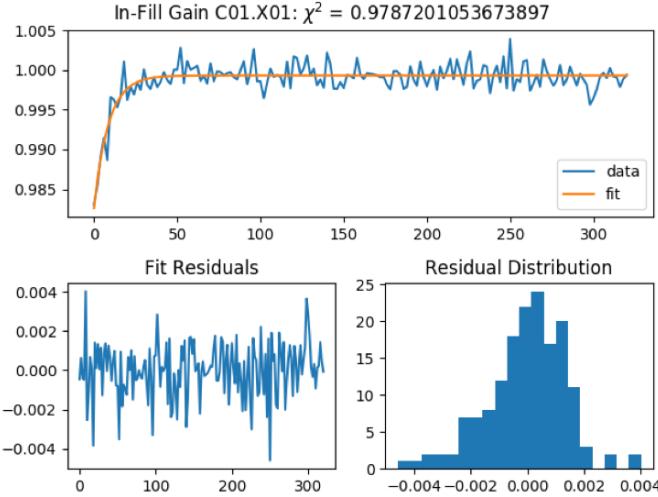


Figure 5.2: In-fill gain function fit for a single calorimeter crystal (top) and fit residuals (bottom). Each crystal has its own in-fill and STDP gain function parameters. Plot courtesy of Matthias Smith.

because most of the deposited energy from a hit is localized to a single crystal. The position of the cluster is determined with a logarithmic weighting function between crystal hits, which for a 2 GeV positron in the E989 calorimeters results in a resolution of 2 mm [87]. See Figure 5.3 for a single calorimeter cluster from a positron hit in the calorimeter. For a time island with multiple positron impacts, the individual crystal hits are separated in time, where the time partitioning separates hits that are 2.5 ns apart, and the clustering proceeds as before. For hits which are within this time window, a pileup event has occurred. If the pileup event happens within the same crystal, then the multiple hits are measured as a single hit, and this needs to be corrected for using a pileup subtraction technique, as described in Section 5.2.1. For hits that occur in separate crystals, the pileup can be resolved using the spatial separation of the calorimeters. This is an ongoing area of work, and one technique is described in Reference [87]. For this analysis the spatial separation was turned off, which simplifies the analysis somewhat. This increases the amount of pileup seen in the data, which then needs to be handled by the pileup subtraction technique. For

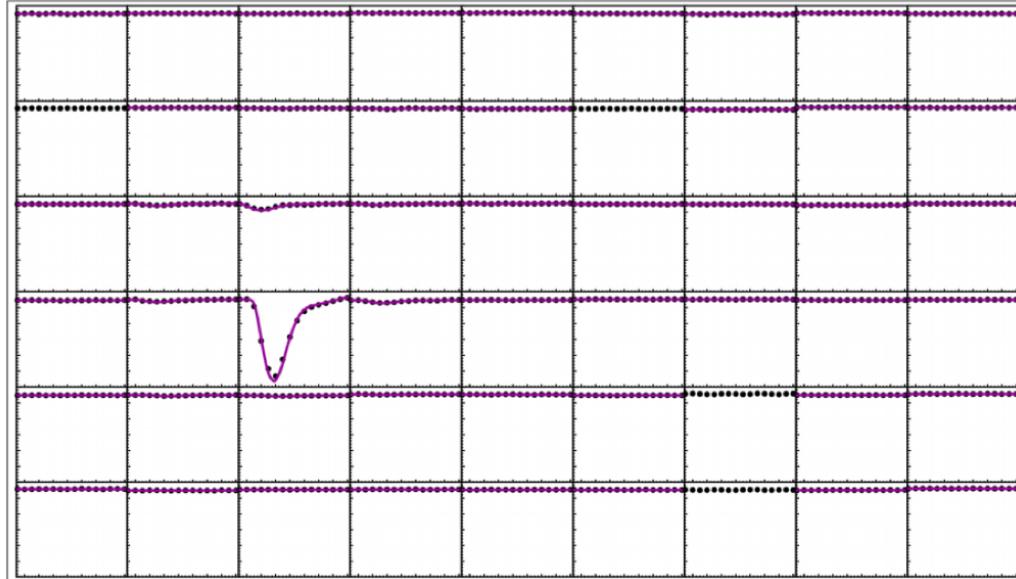


Figure 5.3: A single positron hit in the calorimeter, which resulted in a reconstructed calorimeter cluster. Each box is a crystal in the calorimeter, where the contained trace is the SiPM output fit with a template. The positron hit the crystal three from the left and three from the bottom, where it deposited most of its energy. Some of the energy was deposited in the neighboring crystals. Plot courtesy of Aaron Fienberg.

the precision of the Run 1 analysis result, this was found to be acceptable.

5.2 Construction of positron hit energy and time spectra

Once the reconstruction has processed all calorimeter hits into clusters, the energy and time spectra histograms are made. At the very last stage of the reconstruction procedure, an *art* module takes the produced clusters and puts them into **ROOT TTree** formats, where individual data members include the energies, times, calorimeter numbers, etc. of the individual clusters. There is of order 20,000–140,000 cluster data files per dataset, which are combined down to order 200–1,400 **ROOT TTree** files. These **ROOT TTrees** are then passed through a **ROOT** macro to produce **ROOT** files with the histograms defined by the **TH1F** class, one **ROOT** histogram file per tree file.

Time Spectra Parameters	
Parameter	Value
Energy threshold (E_{th})	1700 MeV
Bin width (T_c)	149.2 ns
Artifical dead time (ADT)	5 ns
Shadow dead time (SDT)	5 ns
Shadow gap time (SGT)	10 ns
Pileup energy scaling (C)	1
$g - 2$ period (T_a) in Ratio Method	4.365 411 μ s
Muon lifetime (τ_μ) in Ratio Method	64.44 μ s

Table 5.1: Parameters used in the construction of ω_a time spectra. **fill this table out more once I've gone through the various parts**

It should be noted that some of the parameter choices for the constructed histograms were informed by analysis results. All analysis parameters were chosen to be identical between the distinct analyzed datasets, in order to simplify both the comparison and combination of different dataset results. This section describes the justification for the different histogram parameters chosen. A table of the histogram parameters is shown in Table 5.1.

Energy and time histograms are made for each individual calorimeter. These are summed together to form histograms of all hit times and energies. Figure 5.4 shows a sample energy spectrum for the Endgame dataset. An energy threshold is applied to the clusters before filling the time histograms. As described at the end of Section 2.1, the optimal energy threshold is where the quantity NA^2 reaches the maximum, at least in the case of a five parameter fit³. By scanning over the choice of energy threshold and fitting the resulting time spectra with Equation 2.13, the optimal energy threshold can be determined as seen in Figure 5.5. The optimal choice of energy threshold was determined to be 1700 MeV, in accordance with the

³Using the final fit function and looking at the error directly on the fitted ω_a frequency, a slightly better estimate can be found.

In E821 our simulations showed
that most of the losses
were in the first 20 to
30 turns.

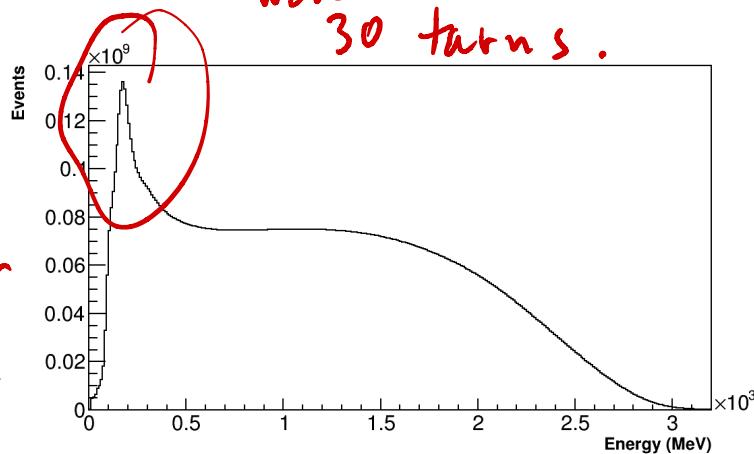


Figure 5.4: Energy spectrum for hits in all calorimeters for the Endgame dataset on a linear scale. The peak at about 170 MeV corresponds to lost muons.

cluster reconstruction energy calibration.

The optimal bin width for the time histograms was determined to be 149.2 ns, the peak of the cyclotron period distribution determined from fast rotation analyses to the data⁴. As described in Section 2.6.2, this bin width combined with a time randomization on each cluster over a range of $\pm T_c/2 = 149.2 \text{ ns}/2$ serves to eliminate the fast rotation signal in the data⁵. This randomization is done using ROOT's TRandom3 class. As will be described in Section 5.3.2, the cluster times are also randomized by the half the vertical waist period, $\pm T_{VW}/2$, where T_{VW} is determined from the inverse of the VW frequency, Equation 2.26. Putting the frequency just in terms of the n value and the cyclotron frequency f_c the VW period is given by

$$T_{VW} = \frac{1}{f_{VW}} = \frac{1}{(1 - 2\sqrt{n}) \cdot f_c}. \quad (5.1)$$

⁴A change in the bin width of order 0.1 ns has very little effect on the extracted ω_a frequency as seen in Section 5.5.7

⁵Some analyzers randomize all times in a single fill by half the cyclotron period as opposed to each individual pulse.

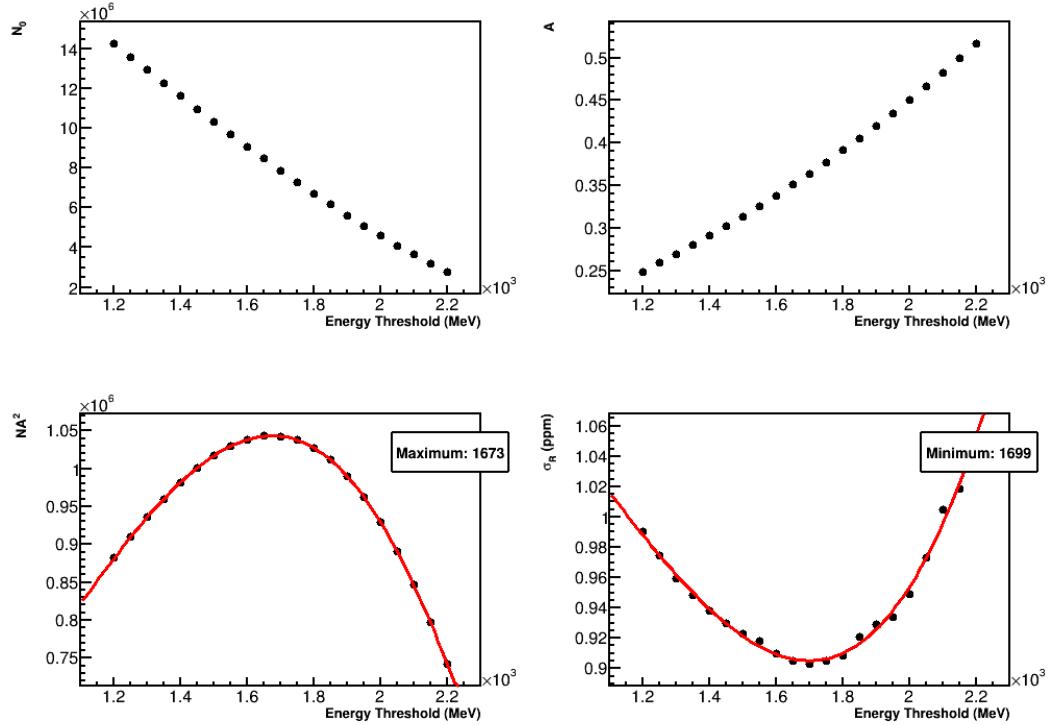


Figure 5.5: The optimal energy threshold can be determined from the NA^2 quantity as described in Section 2.1 from five parameter fits to the data (bottom-left) or from the calculated error with the final fit function (bottom-right). Fitted N and A parameters are also shown (top) which satisfy Equations 2.14 and 2.15. The maximum and minimum are determined using seven parameter polynomial fits to the relevant quantities. The optimal threshold varies slightly per calorimeter and per dataset. Since the region of the minimum is relatively flat, a single energy threshold of 1700 MeV was chosen. Data from the 9d dataset.

This randomization is done in order to remove the effects of the VW in the data⁶. The default random seed for each histogram ROOT file is the hash of the input file name using C++'s standard hash class. Histograms are defined with a time range of 0–699.8972 μs (the closest integer multiple of the bin width to 700 μs), corresponding to 4691 bins. Clusters with times less than 25 μs or greater than 660 μs are dropped, corresponding to 4256 bins containing data.

5.2.1 Pileup subtraction

As described in Section 3.2 there will be a certain amount of pileup in the detectors. Pileup again is the term for when multiple particles hit a calorimeter within the dead time of the detector such that they are registered as a single hit or cluster. The measured energy and time spectra for all observed clusters will include this pileup background. For the energy threshold time histogram, the number of counts will be wrong for cases where two below-threshold particles are registered as a single cluster above threshold, and where two above-energy threshold particles are registered as a single cluster. In the former, an extra count is added into the histogram, and in the latter a count is missed. The case where two lower energy positrons are registered as a single higher energy cluster will have a different $g - 2$ phase than an actual single cluster at the same energy. This is because the lower energy positrons on average decay from muons which have travelled further around the ring, due to acceptance effects. These muons which have travelled further around the ring have spent more time in the magnetic field, and thus their spins have precessed more. See Figure 5.6. Clusters which originate from pileup events therefore have a different $g - 2$ phase than non-pileup events.

don't you mean 'constant in time'

If pileup was a constant effect, then the phase of the time histogram would be

⁶Even though the VW frequency was found to be changing over the course of the fill, this constant time randomization was found to remove all residual traces of the VW in the data.

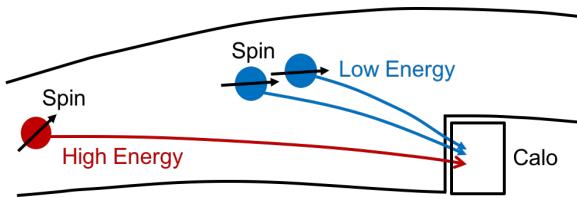


Figure 5.6: Pileup example, where two low energy positrons are registered as a single high energy positron. The black arrows indicate the (exaggerated) direction of the muon spins at the time of decay. Because of acceptance effects the lower energy decay positrons typically come from muons which have traveled further around the ring, and thus the muon spins have precessed more in the magnetic field, leading to a different measured $g - 2$ phase for pileup events.

shifted by some constant amount, and the extracted ω_a frequency would be unaffected. However, the rate of pileup in the detectors changes over the time of a fill, as muons decay away. The rate of double pileup events in the detectors, where the word double indicates cases where two hits are registered as a single cluster, will go approximately as half the rate of single hit events⁷, and similarly for triple and higher orders of the pileup effect. Because the rate of hits in the detectors oscillates at the $g - 2$ frequency, pileup will increase and decrease accordingly leading to oscillations in the pileup time spectra at ω_a and $2\omega_a$. The lifetime of the overall pileup effect is approximately half the lifetime at which clusters are registered in the detectors, at τ_μ , since double pileup is the dominant contribution. In order to extract the correct ω_a frequency, the pileup effect thus needs to be included in the fit function or subtracted out of the data. The former is challenging due to the non-linear nature of the dead time of the detectors, and would in the end include another phase in the argument of the cosine term in the fit function, thus worsening the statistical precision of the extracted ω_a frequency. All analyzers thus construct an approximation of the pileup effect and subtract it from the data before fitting.

There are various methods to construct pileup spectra which are then subtracted

⁷It is not exactly half when including the non-linear dead time of the detectors

don't you mean 'Construct a pileup spectrum and then subtract it from the data before fitting.'?

From

off the main time and energy spectra. The method used in this analysis is called the ‘asymmetric shadow method’, originally developed in E821 [90]. This method statistically constructs an approximation for the pileup from the data by assuming that the probability of observing a pileup pulse is the same as the probability that two pulses will be offset by some small amount of time, such as 10 ns. The method works by looking in time windows after trigger pulses to see if a ‘shadow’ pulse exists. If such a pulse exists, then a shadow doublet is created, see Figure 5.7. The width of the time window, and the time offset from the trigger pulse to the window, are called the shadow dead time (SDT) and shadow gap time (SGT) respectively. The times and energies of the constructed pileup doublets are taken as

$$E_{\text{doublet}} = C \cdot (E_1 + E_2), \quad (5.2)$$

$$t_{\text{doublet}} = \frac{t_1 \cdot E_1 + (t_2 - SGT) \cdot E_2}{E_1 + E_2}, \quad (5.3)$$

where the energy of the doublet is the sum of the two singlet pulses $E_{1,2}$ times some calibration constant C , with a default value of 1, and the time of the doublet is the energy-weighted time of the two singlets $t_{1,2}$. The procedure for constructing the pileup spectra is as follows:

- Put each hit into a vector corresponding to a specific fill and a specific calorimeter
- Time order the hits
- Loop through the hits; for each hit look within a window of width SDT a time SGT later to see if a shadow pulse exists
- If a shadow pulse exists, construct a shadow doublet with energies and times as defined in Equations 5.2 and 5.3

- Randomize t_{doublet} over the range $\pm T_c/2$ (to remove fast rotation as before, Section 2.6.2)
- Per calorimeter, construct pileup energy and time spectra as $P = D - S$, where D is the sum of doublets and S is the sum of singlets used in the construction of the doublets, with the times of the singlets set as t_{doublet} ; when constructing the pileup time spectra, only include those doublets and singlets above the energy threshold

*"do you mean
"before subtraction"
then from
calo sum
histograms"*

Thus pileup energy and time spectra are constructed for each calorimeter, which can then be subtracted off the calorimeter cluster energy and time histograms. When combining the data, the individual pileup histograms are simply added together before subtraction off the calorimeter sum histograms. ?

In order to produce an estimate of the pileup spectra which best matches the data, an artificial dead time (ADT) is applied to the data before time randomization. This is done because the true dead time of the detectors depends on the energies and spatial separation of the incoming hits. While this is a small effect, by applying an artificial deadtime and matching the shadow window time, the pileup estimation is improved slightly. The construction of the artificial pileup is handled in the same way as the construction of the shadow pileup, with SGT set to 0 ns. The constructed artificial doublets replace the singlets in the data. The value for the ADT and SDT is set at 5 ns, the time threshold at which pileup is 100% resolved.

The value of the SGT is simply set to twice the SDT, in order to push the shadow window out to times well beyond the dead time of any pileup events, but not so far that an appreciable fraction of muons have decayed. The value of the doublet energy scaling factor C is set to 1, which is a fine approximation as the spatial separation in the reconstruction is turned off⁸. The values for each pileup parameter is shown in

⁸With the spatial separation turned off, ‘pileup’ events can occur in crystals that are easily

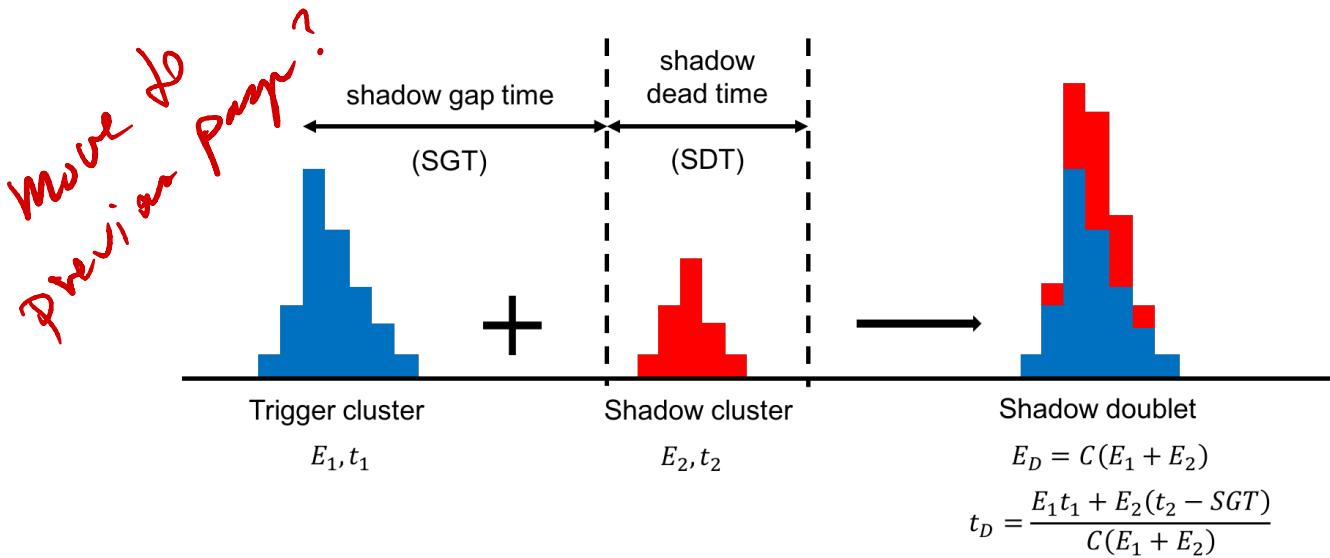


Figure 5·7: The shadow pileup method looks for shadow clusters within a time window SDT, a small time away (SGT) from trigger clusters. If a shadow cluster is found, an artificial doublet is formed and included in the pileup spectra if it exceeds the chosen energy threshold.

Table 5.1. See Section 5.5.1 for systematic studies on the effect on ω_a due to these chosen parameters.

The pileup energy spectra ~~as compared to~~ the cluster energy ~~spectra~~ ^{Spectrum} and ~~are~~ ^{um are} reasonable in Figure 5·8. In general, the two lobes starting at approximately 3 GeV and 6 GeV consist of double and triple pileup events respectively⁹. It can be seen that the shadow method of pileup construction produces a pileup energy spectra which is a decent approximation of the cluster energies above the maximum energy that a single decay positron would have at (3.094 GeV) + detector resolution, for cases of double and even triple pileup. The shape difference arises from two factors: First, the shadow method is only written to construct doublets, and does not consider cases of triple or higher orders of pileup. Second, the real pileup in the data contaminates the construction of the shadow pileup spectra, such that a shadow doublet can be constructed from separated by eye. While this increases the level of pileup seen in the data, the pileup approximation method also does not consider the spatial separation, and thus handles the level of pileup accordingly.

⁹All orders of pileup fill out the whole energy range, but certain areas consist of mostly one or the other.

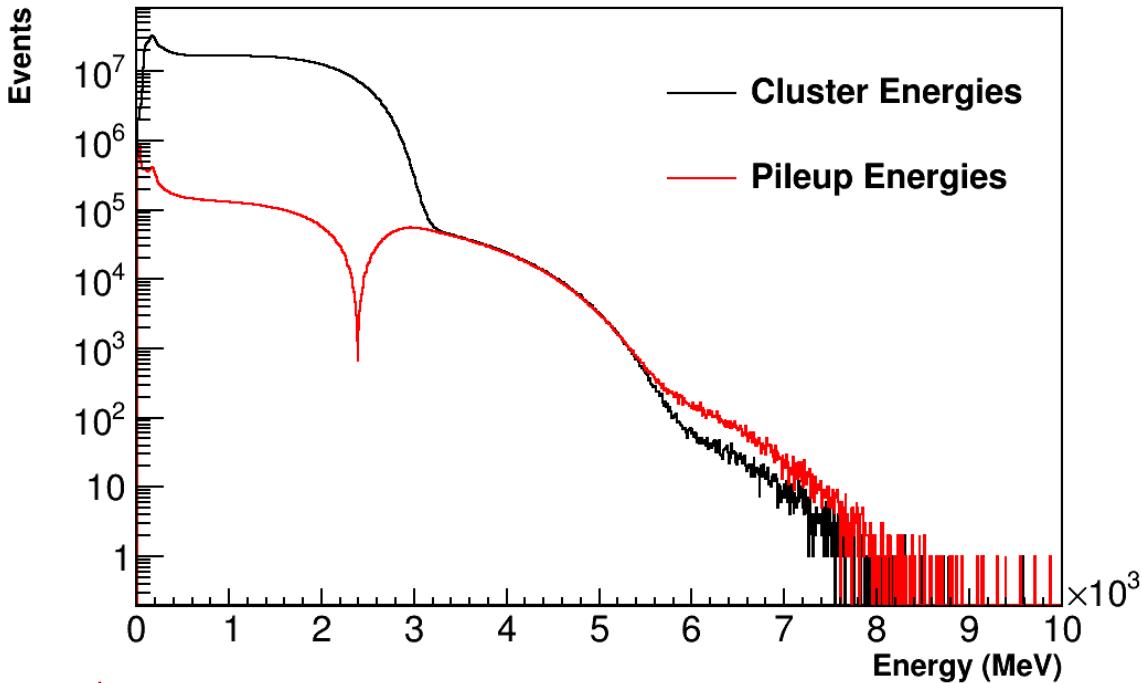


Figure 5.8: Cluster energies in black are plotted vs pileup energies in red, for all calorimeters added together, plotted on a log scale. At energies below about 2.4 GeV the pileup energy spectrum goes negative. In this plot the absolute value of the pileup energies is plotted, and a spike at about 2.4 GeV can be seen as a consequence of this. The shapes do not match perfectly for the constructed pileup spectra, which can be seen at high energies. It should be noted that for energies above 3.094 GeV there is still a shape mismatch even though the red and black curves overlap due to the plotting scale. Data from the 60h dataset.

This statement is confusing

difference

real pileup pulses. While this alleviates the triplet problem slightly, it means that the doublet pileup spectrum is slightly wrong. The corrected energy spectra (cluster energies minus pileup energies), can be seen in Figure 5.9. The shape mismatch is even more apparent as the corrected energy spectrum is high for energies above the expected tail of the true energy distribution, and then goes negative before tailing off to zero.

In order to produce a slightly better estimate of the pileup, a multiplier can be applied to the pileup energy and time spectra. By taking the ratio of cluster energies

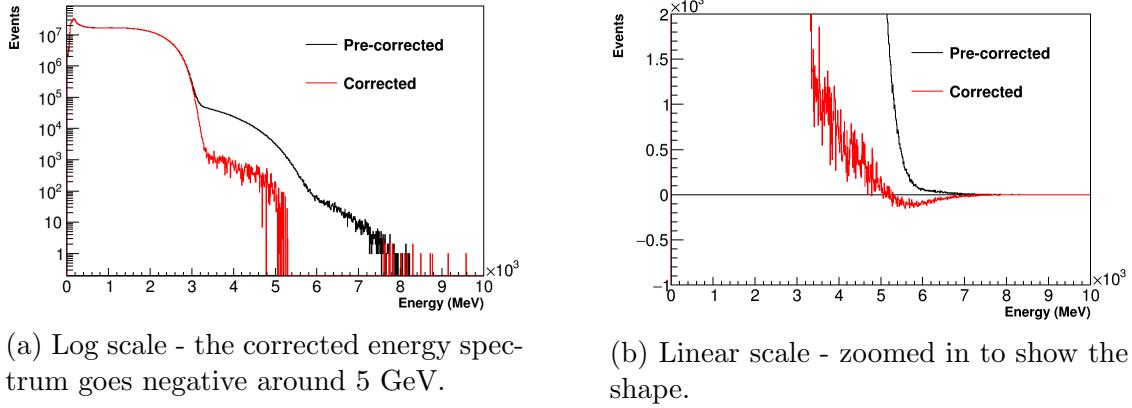


Figure 5.9: Plots for the pre-corrected and corrected energy spectra are shown, all calorimeters added together. Because the triplets and contamination are not accounted for, the corrected energy spectrum does not lie exactly along zero above the energy response of the detectors. Data from the 60h dataset.

over pileup energies and fitting the region where the energies are dominated by real pileup doublets, a correction factor of approximately 3-4% is found, as shown in Figure 5.10. The default multipliers for the Run 1 datasets are 1.03210, 1.03413, 1.03387, and 1.03819 for the 60h, HighKick, 9d, and Endgame datasets respectively. Similarly, the cluster times can be examined for cluster energies over 3500 MeV, where the clusters consist purely of pileup pulses, Figure 5.11. By taking the ratio of the pileup corrected times over all times, the level of residual pileup can be determined. Just as in the ratio of the energies, an approximately ~~3-4%~~ 3-4% factor is found. When applying this multiplier, the cluster times above 3.5 GeV can be seen to be eliminated as in Figure 5.11. As will be shown in Section 5.5.1, the scale of this multiplier is well within 1σ of the pileup multiplier error. The final pileup time spectrum for those pileup pulses above energy threshold is shown in Figure 5.12.

It has been determined that regardless of any residual shape mismatch in the cluster times below 3.5 GeV, the systematic error on the extracted ω_a frequency due to the pileup is within the target uncertainty for the level of statistics in the Run 1

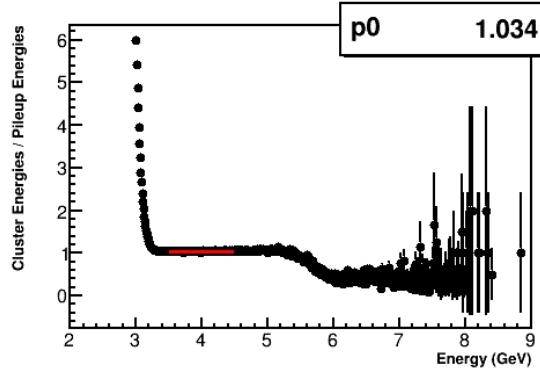


Figure 5.10: Cluster energies over pileup energies. A region from 3500–4500 MeV is fit to a straight line, where the doublets dominate the energy distribution. Data from the 9d dataset.

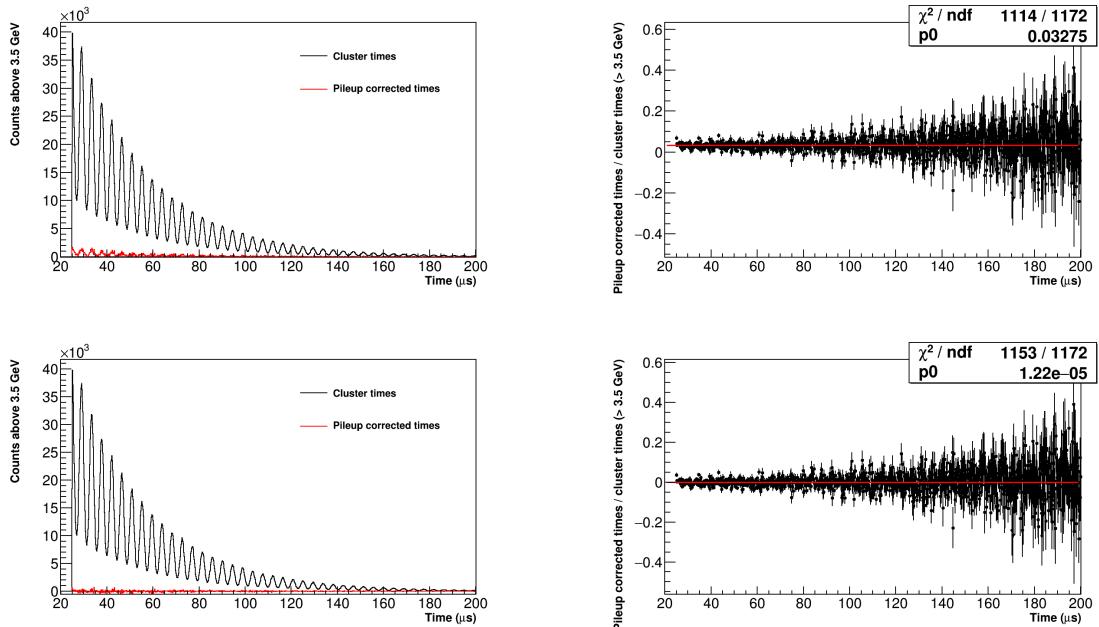


Figure 5.11: Cluster times and pileup corrected times for counts above 3.5 GeV (left) and their ratio (right). The top two plots are used to determine the approximate level of residual pileup left in the data, coming out to about 3%. The bottom two plots show the application of that factor and the resulting removal of the remaining pileup. Data from the 9d dataset.

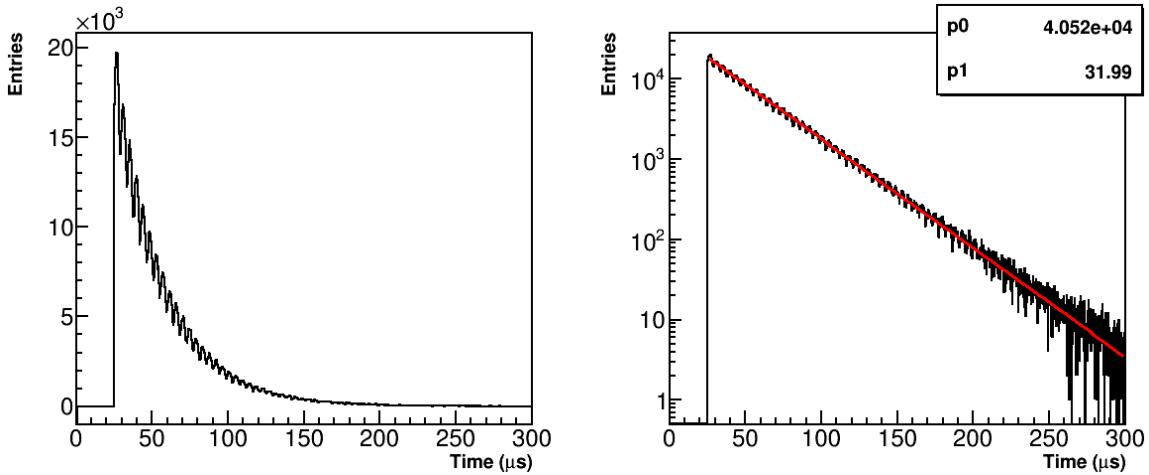


Figure 5.12: Plotted is constructed pileup time spectrum on a linear (left) and log (right) scale. The histogram on the right is fit to a simple two parameter exponential to get an idea of the lifetime of the pileup, calculated here as $31.99 \mu\text{s}$, which is close to half of the muon lifetime at about $64.44 \mu\text{s}$. Data from the 60h dataset.

dataset, Section 5.5.1] For analyses past Run 1 where the error budget is reduced, it may be necessary to improve the shadow method to account for triplets and the contamination. Finally, since the pileup is statistically constructed and then subtracted from the data, the errors on the final time histogram are no longer Gaussian. The proper calculation of the errors is detailed in Appendix C

5.2.2 Ratio Method

The method used in this analysis to extract ω_a is called the “Ratio Method,” or sometimes “R-Method.” It is a technique that modifies the data in such a way that the exponential decay in the time histogram is removed, and slow effects are reduced. It was used successfully in the E821 experiment [91, 92, 93]. A full derivation of the equations in the method is given in Appendix B; here is given a short summary.

Figure 5.13 provides a pictorial representation of how the method works.

The method works by dividing the data into four separate datasets, one with the

randomly

any slowly varying

times of all clusters shifted up by half a $g - 2$ period, $+T_a/2$, one with cluster times shifted down by half a $g - 2$ period, $-T_a/2$, and two unchanged. Assuming the data

~~are~~ is described by the five parameter function described in Section [2.1] and shown in Figure [5.13a].

$$N_d(t, E_{th}) = N_0(E_{th}) \cdot e^{-t/\gamma\tau_\mu} \cdot [1 + A(E_{th}) \cos(\omega_a t + \phi(E_{th}))], \quad (5.4)$$

and that the data ~~is~~ ^{are} equally split into four subsets, then the new four datasets are given as¹⁰:

$$\begin{aligned} u_+(t) &= \frac{1}{4}N_5(t + T/2) \\ u_-(t) &= \frac{1}{4}N_5(t - T/2) \\ v_1(t) &= \frac{1}{4}N_5(t) \\ v_2(t) &= \frac{1}{4}N_5(t) \end{aligned} \quad (5.5)$$

In order to time shift the data as such, T_a needs to be known a priori to high precision. The value used is taken from the E821 result, and its value is taken as $1/f_a$, where f_a is 0.229 073 5 MHz:

$$T_a \approx 4.365\,411\,\mu\text{s} \quad (5.6)$$

This value for f_a was determined by averaging column 2 of Table XV of the E821 Final Report [29], which consists of the f_a results for the different run periods in that experiment. A systematic error on the choice of this parameter is calculated in Section [5.5.6]

¹⁰When handling the pileup in the ratio method, the pileup time spectra are split into four datasets and time-shifted in the same way as the cluster hit times. Associated doublets and singlets are kept together in the same individual dataset, and the four pileup datasets are subtracted off their respective ratio datasets before forming the ratio.

The datasets are then combined as

$$\begin{aligned} U(t) &= u_+(t) + u_-(t), \\ V(t) &= v_1(t) + v_2(t), \end{aligned} \quad (5.7)$$

both of which are shown in Figure 5.13b. It is immediately apparent that the $U(t)$ data are shifted 180° out of phase from the $V(t)$ data. The ratio is then defined as¹¹

$$R(t) = \frac{V(t) - U(t)}{V(t) + U(t)} \quad (5.8)$$

where the numerator and denominator are plotted in Figures 5.13c and 5.13d respectively. The numerator is an exponentially decaying cosine, while the denominator is a simple exponential, both of which can be seen as originating from the difference and sum of the $U(t)$ and $V(t)$ data respectively. The resulting ratio spectrum can be seen in Figure 5.13e, where the exponential has been eliminated. The fit function is then reduced from five parameters down to three:

$$R(t) \approx A \cos(\omega_a t) - C, \quad (5.9)$$

where

$$C = \frac{1}{16} \left(\frac{T}{\tau} \right)^2 \approx 2.87 * 10^{-4}, \quad (5.10)$$

and these functions have been determined from the time-shifted five parameter function plugged into the $U(t)$ and $V(t)$ variables. In addition to the exponential being eliminated, any slow terms in the data get time-shifted and divided as well, such that the amplitude of said slow effects are reduced. For faster effects, the degree of cancellation of the effect is dependent on the frequency. Effects at frequencies which

¹¹The ratio can also be defined with $U(t) - V(t)$ in the numerator, however then the phase of the ratio spectrum is shifted 180° from the original $N_5(t)$ spectrum.

are an odd multiple of ω_a are preserved while effects at an even multiple of ω_a are completely cancelled out. An example is shown in Figure 5.14. While this makes fitting the data easier in some cases, in others it is a downside that effects which still need to be included in the fit function now have their amplitudes reduced, making them harder to fit.

In order to eliminate the constant C at the end of Equation 5.9, a different weighting scheme can be used as described in Reference [37]:

$$\begin{aligned} u_+(t) &= \frac{e^{T/2\tau}}{2 + e^{T/2\tau} + e^{-T/2\tau}} N_5(t + T/2) \\ u_-(t) &= \frac{e^{-T/2\tau}}{2 + e^{T/2\tau} + e^{-T/2\tau}} N_5(t - T/2) \\ v_1(t) &= \frac{1}{2 + e^{T/2\tau} + e^{-T/2\tau}} N_5(t) \\ v_2(t) &= \frac{1}{2 + e^{T/2\tau} + e^{-T/2\tau}} N_5(t) \end{aligned} \quad (5.11)$$

Here $\tau = \gamma\tau_\mu$, and the factors out front are each close to 1/4 and account for the degree of muon decay over a time period of $T_a/2$. Similar to T_a , the muon lifetime must be known a priori. Its value is taken as 64.44 μs , determined from fits to the data. A systematic study regarding this parameter is described in Section 5.5.6. The ratio spectrum is then almost exactly described by just the cosine term,

$$R(t) \approx A \cos(\omega_a t), \quad (5.12)$$

in the absence of other effects in the data.

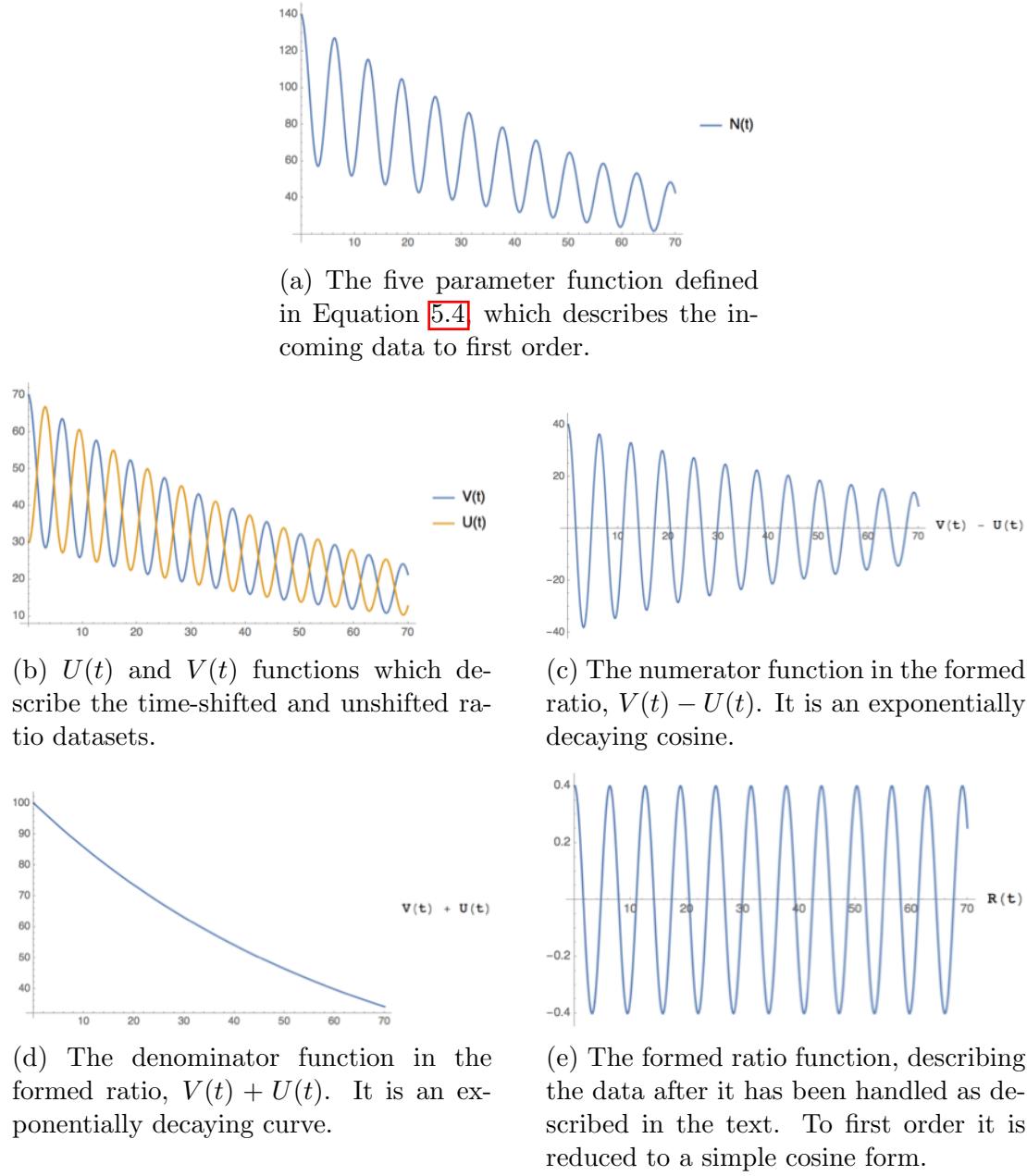


Figure 5.13: Functions describing the formation of the ratio in the data.

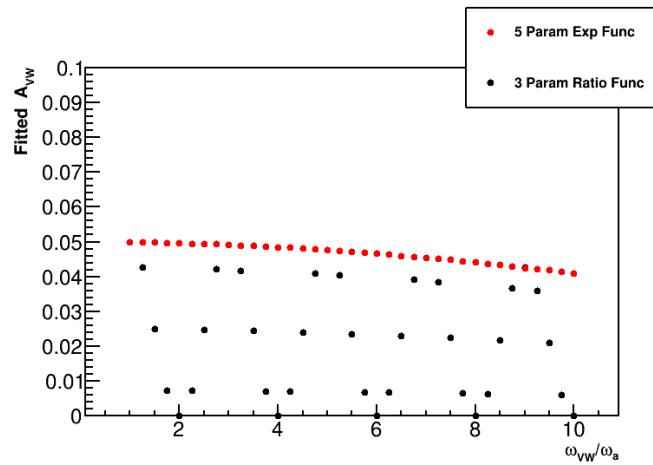


Figure 5.14: Fitted amplitude for a VW effect as a function of frequency in units of ω_a in a Toy MC simulation, with a five parameter function in red and a three parameter ratio function in black. The input amplitude was 0.05. The fall off of red points is due to the high frequencies relative to the bin widths leading to an underestimate of the amplitude; performing an integral fit removes this trend. As shown, the amplitude of the effect goes to zero for frequencies which are an even multiple of ω_a .

5.3 Fitting the data

The basic five parameter function used to fit the data as described before is given as¹²

$$f(t) = \mathbf{N}_0 \cdot e^{-t/\tau} \cdot (1 + \mathbf{A} \cdot \cos(\omega_a t + \phi)), \quad (5.13)$$

where the fit parameter for ω_a is recast in terms of a ppm level shift \mathbf{R} on a reference frequency,

$$\omega_a = 2\pi \cdot 0.2291 \text{ MHz} \cdot (1 + \mathbf{R} \times 10^{-6}). \quad (5.14)$$

This reference frequency of 0.2291 MHz was the same reference frequency used in E821, and \mathbf{R} is blinded at the hardware and software levels [94, 95]. Fitting the data with Equation 5.13 however is insufficient to properly describe the data. Figure 5.15 shows there are peaks in the FFT due to beam dynamics frequencies corresponding to the CBO, VW, and some beat frequencies with ω_a . In order to properly account for these effects, additional terms need to be added to the fit function.

Equation 5.13 can be expanded to

$$f(t) = \Lambda(t) \cdot V(t) \cdot N_{cbo}(t) \cdot \mathbf{N}_0 \cdot e^{-t/\tau} \cdot (1 + A_{cbo}(t) \cdot \cos(\omega_a t + \phi_{cbo}(t))), \quad (5.15)$$

where many additional terms have been added in order to account for effects in the data. The various additional terms $\{N_{cbo}(t), A_{cbo}(t), \phi_{cbo}(t), V(t), \Lambda(t)\}$ are described in the following sections. Fitting the data with this function, referred to as the “Threshold Method” or just “T-Method,” while not the subject of this dissertation, was done in this analysis as a diagnostic and informative tool for the Ratio Method analysis.

¹²In all cases here and onwards the actual fit parameters are in bold.

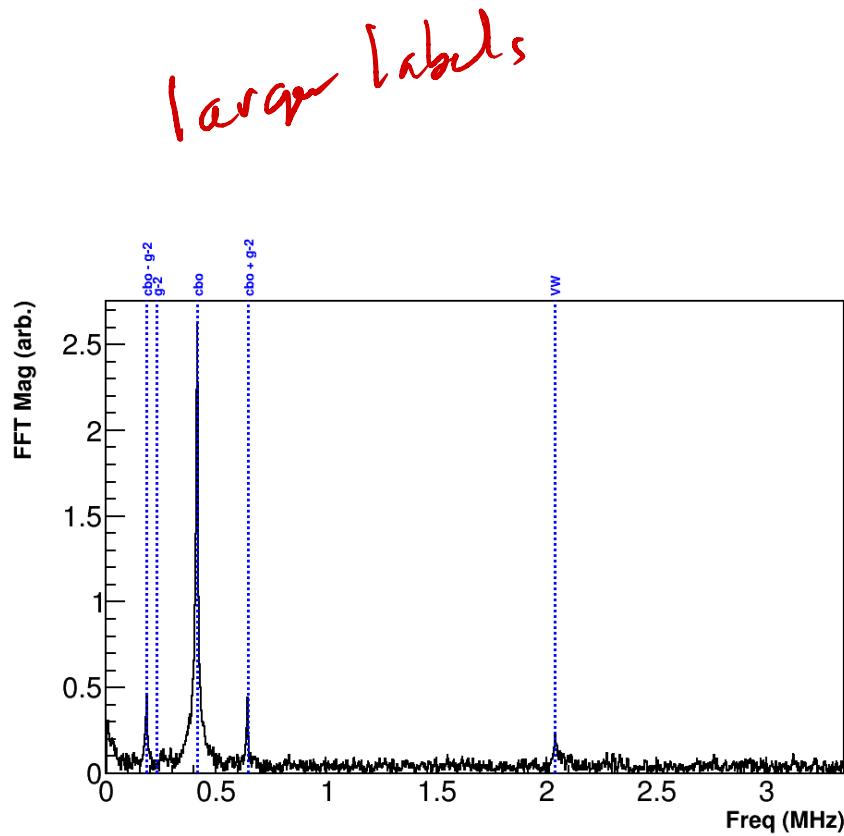


Figure 5.15: FFT of five parameter fit residuals. Peaks corresponding to beam dynamics frequencies of the CBO and VW and some beat frequencies with ω_a are readily apparent. A rise at low frequencies corresponds to the effects of the lost muons in the data. The VW peak disappears with the additional level of time randomization at that frequency. From the 60h dataset.

This needs to be explained more clearly

would be

In order to fit the ratio time spectra as constructed in Section 5.2.2, a different function is used. While the immediate inclination is to use an expansion of Equation 5.12 with included additional effects similar to the T-Method fit function, instead the fit function used is a return to the explicit definition of the construction of the ratio time spectra in Equations 5.7 and 5.8. Including the additional effects previously mentioned, the fit function goes as *becomes*

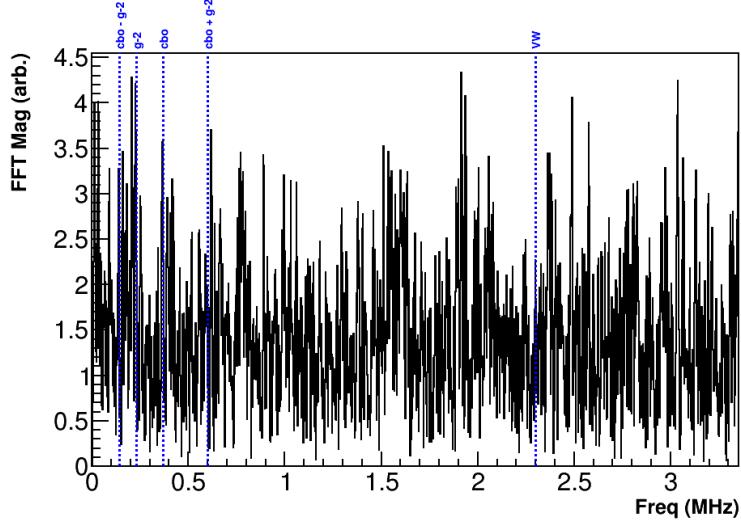
$$R(t) = \frac{2f(t) - f_+(t) - f_-(t)}{2f(t) + f_+(t) + f_-(t)}, \quad (5.16)$$

$$f_{\pm}(t) = f(t \pm T_a/2), \quad (5.17)$$

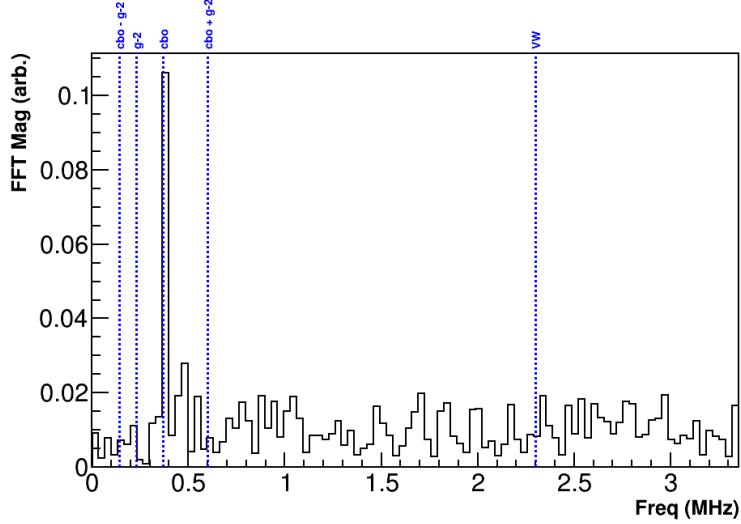
$$f(t) = \Lambda(t) \cdot V(t) \cdot N_{cbo}(t) \cdot (1 + A_{cbo}(t) \cdot \cos(\omega_a t + \phi_{cbo}(t))). \quad (5.18)$$

The $f(t)$ given here differs from that in Equation 5.15 in that the $N_0 \cdot e^{-t/\tau}$ terms have divided out, thus reducing the number of fit parameters necessary to model the data. Using this function as opposed to an expansion of the three parameter ratio function eliminates any approximations made in that three parameter function derivation, and any fit parameters should be consistent in value between the T-Method and Ratio Method results, barring adjustments due to the application of the Ratio Method.

Because the Ratio Method reduces the sensitivity of the ω_a ~~extraction~~ *determination* to various effects in the data, peaks that appear in the FFT of the five parameter fit residuals ~~do not~~ *don't* appear in the FFT of the three parameter ratio fit residuals, Figure 5.16. Indeed except for bad fits, unless one looks at the FFT over the early part of the fit (first 30 μs) or at the shape of the Ratio Method denominator, one might not know the *Never the less,* effects even exist in the data at all. ~~However~~ those effects typically still need to be included in the fit function for a proper estimation of ω_a , ~~and~~ *B* because of the reduction in sensitivity, there are some parameters which the ratio has trouble fitting by itself. Using the T-Method fit function is then a useful tool for constraining these specific parameters.



(a) FFT of fit residuals over all times within the fit range. There is no immediately apparent structure for residual effects left out of the fit function.



(b) FFT of fit residuals over the first 30 μ s. The CBO peak can be seen above the noise, though the corresponding beat frequencies do not appear.

Figure 5.16: FFT of three parameter ratio fit residuals. Dashed blue lines indicate various beam dynamics frequencies and their beat frequencies with ω_a . Data from the 60h dataset.

5.3.1 CBO terms

amplitude

As described in Section 2.6.1, the CBO modulates the ω_a oscillation. This shows up as a modification on the five parameter function parameters $\{N_0, A, \phi\} \rightarrow \{N_0 \cdot N_{cbo}(t), A_{cbo}(t), \phi_{cbo}(t)\}$ where these terms are given to first order as

$$N_{cbo}(t) = (1 + A_{cbo-N} \cdot e^{-t/\tau_{cbo}} \cdot \cos(\omega_{cbo}(t) \cdot t + \phi_{cbo-N})) \quad (5.19)$$

$$A_{cbo}(t) = A \cdot (1 + A_{cbo-A} \cdot e^{-t/\tau_{cbo}} \cdot \cos(\omega_{cbo}(t) \cdot t + \phi_{cbo-A})) \quad (5.20)$$

$$\phi_{cbo}(t) = \phi_0 + A_{cbo-\phi} \cdot e^{-t/\tau_{cbo}} \cdot \cos(\omega_{cbo}(t) \cdot t + \phi_{cbo-\phi}) \quad (5.21)$$

In Equation 5.19 the N_0 term is left out since the ratio fit includes the $N_{cbo}(t)$ term but not N_0 , and in Equation 5.21, $\phi_{cbo}(t)$ has an additive phase instead of a multiplicative one since ϕ_0 is not an amplitude and can be equal to zero. Each of the terms then includes additional fit parameters in an extra amplitude and phase, as well as one shared CBO lifetime and frequency. As described in Section 4.4, the default model for the CBO modulation is assumed as an exponentially decaying envelope. The CBO frequency, $\omega_{cbo}(t)$, was time-dependent for Run 1 as found in Section 4.4. The function for the CBO frequency shown in Figure ?? is given in the fit function as

$$\omega_{cbo}(t) = \omega_{cbo} \cdot \left(1 + \frac{Ae^{(-t/\tau_A)}}{\omega_0 t} + \frac{Be^{(-t/\tau_B)}}{\omega_0 t} \right), \quad (5.22)$$

where ω_{cbo} is the free fit parameter, and the model parameters $\{\omega_0, A, \tau_A, B, \tau_B\}$ are fixed from the tracking analysis. These parameters for the various datasets and two tracker stations are given in Table 5.2.

It should be noted that Equations 5.20 and 5.21 are not necessarily needed in order to get good fits to the data (whereas Equation 5.19 always is). This is typically dataset or random seed dependent. While some datasets had certain parameters with large errors relative to their amplitudes, for this analysis all terms were successfully

CBO Frequency Model Parameters						
Dataset	Tracker Station	ω_0 (rad/μs)	A (rad)	τ_A (μs)	B (rad)	τ_B (μs)
60h	12	2.3389	2.9	81.8	5.12	7.7
	18	2.3387	2.82	81.1	5.08	8.2
HighKick	12	2.6145	3.27	52.8	6.96	6.6
	18	2.6137	3.23	46.2	6.61	6.8
9d	12	2.6106	2.86	72.8	5.50	8.5
	18	2.6110	2.89	79.2	5.44	9.2
Endgame	12	2.3377	7.43	95.1	4.71	9.0
	18	2.3379	7.44	95.2	4.90	9.2

Table 5.2: Fixed parameters in the CBO frequency model [96, 97].

Here I source personal communication with James, as I don't think there's a source for the HighKick numbers.

more clearly

included in all dataset fits with appropriate tuning of the starting parameters with well converging fits. *This sentence needs to be stated*

When considering higher order CBO modifications to the fit function, the only term that was found to be *fittable* was the second order CBO modulation on the N_0 term,

is there a better word?

$$N_{2cbo}(t) = (1 + A_{2cbo-N} \cdot e^{-2t/\tau_{cbo}} \cdot \cos(\omega_{cbo}(t) \cdot t + \phi_{2cbo-N})). \quad (5.23)$$

This stands to reason as the $N_{cbo}(t)$ is the largest CBO effect. The form is assumed to be the same as the first order CBO terms, except the lifetime of the effect is half the CBO lifetime, $\tau_{cbo}/2$. This is due to the fact that the $N_{2cbo}(t)$ is reasoned to come from the width of the oscillating beam, as opposed to the oscillating mean. Indeed as will be shown in Section 5.4.1, the inclusion of this term is necessary to get good fits to the per calorimeter data, where the CBO effect is stronger compared to in the sum of the calorimeter data. For this reason, and again because the $N_{2cbo}(t)$ term is *fittable* in the calorimeter sum data, this term is included in fits to each of the datasets.

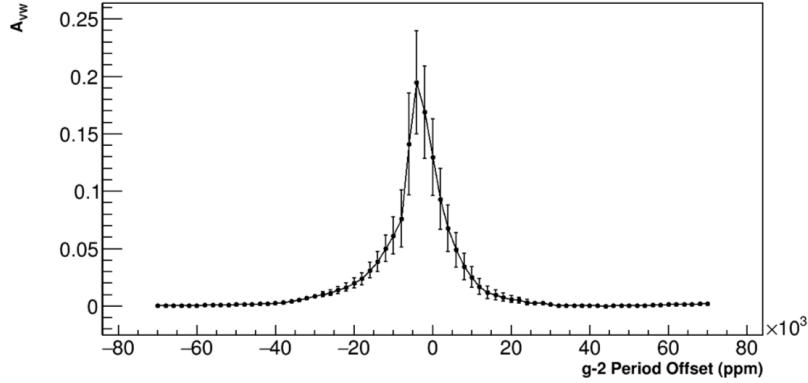


Figure 5.17: Fitted VW amplitude as a function of the choice of offset from T_a in units of thousands of ppm for the 60h dataset. The default time-shift lies at 0 on this plot, right on the resonance where the VW amplitude blows up. Only by time-shifting by a drastically different amount (which negatively affects R), can the resonance be avoided.

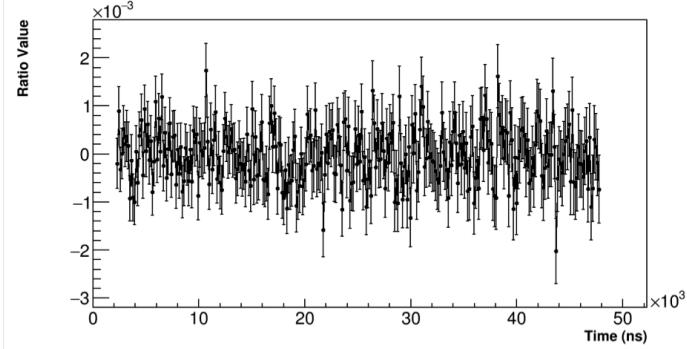
Systematic studies relating to the choice of envelope and choice of fixed parameters in the frequency model are explored in Section 5.5.3. For future runs beyond Run 1, it may be necessary to include the higher order modifications to the \mathbf{A} and ϕ terms.

5.3.2 VW term

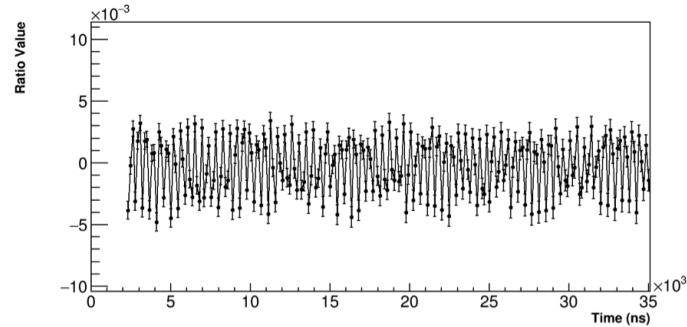
As mentioned briefly at the end of Section 5.2, the VW effect is time-randomized out of the data, such that $V(t) = 1$. This is done due to complications with the Ratio Method. In the 60h and Endgame datasets, the VW frequency was found to be nearly $10 \cdot \omega_a$, on a potential resonance. While to first order this even multiple frequency implies the VW effect should completely cancel out in the Ratio Method, Figure 5.14, the FR effect in combination with the VW leads to a modified envelope for the VW effect in the Ratio Method fits and inflated VW amplitudes [98]. See Figures 5.17 and 5.18.

do you need this chart?

For the 9d dataset which has a VW frequency that's nearly $9\omega_a$ and avoids said resonance (and by extension the HighKick), it was found that the Ratio Method flattened out the VW amplitudes as a function of calorimeter number leading to a



(a) Without the FR effect included.



(b) With the FR effect included.

Figure 5.18: Ratio data with and without the FR effect from a Toy MC simulation, with a VW effect with a frequency $\omega_{VW} = 10 \cdot \omega_a$. The ω_a wiggle itself has been removed, and the lifetime of the VW was set to a large number. The top plot shows ratio data which is consistent with 0 after all effects have been removed and the VW has divided out. The bottom plot shows ratio data inconsistent with 0, with oscillations at the VW frequency, and an interesting beating structure. Note the different scales.

is this the change in R or the
change in the error on R?

145

Change in Asymmetry due to VW Randomization				
Dataset	A no randomization	A with randomization	ΔA	$\Delta\sigma_R$ (ppb)
60h	0.3697	0.3637	-0.0060	22.7
9d	0.3714	0.3639	-0.0075	18.1
Endgame	0.3747	0.3686	-0.0061	10.7

Table 5.3: Asymmetry values in three of the Run 1 datasets with and without the VW randomization, and the corresponding change in the statistical error on R. An energy cut of 1700 MeV was applied to the data. The HighKick dataset has the lowest asymmetry and the least statistics, so the increase in the statistical error would be negligible as it is for the other datasets.

systematically smaller VW amplitude in the calorimeter sum fit. The simplest solution to remove both of these problems was to randomize out the VW effect entirely. Table 5.3 gives the change in asymmetry and corresponding change in the statistical error on R due to this added time randomization for three of the Run 1 datasets. It was found that the added randomization increased the statistical error on R by a negligible amount for the Run 1 analysis. It was also found that the added level of time-randomization changed the mean value of R for many random seed fits to the data by a small amount, statistically consistent without the extra randomization [98]. Going forward for the future runs, it may be necessary to include the proper VW envelope in the fits using a functional form of the FR instead of randomizing out the effect.

Barring the envelope changes in the Ratio Method, here is still given the form for the VW as it is used in the T-Method fits. The form for the VW term is taken identically to the CBO terms,

$$V(t) = 1 + \mathbf{A}_{VW} \cdot e^{-t/\tau_{VW}} \cos(\omega_{VW}(t) \cdot t + \phi_{VW}), \quad (5.24)$$

with an exponentially decaying envelope, and an additional amplitude and phase

parameter. The VW frequency ω_{VW} is given in Equation 2.26 where it is seen to be dependent on the cyclotron frequency and vertical betatron frequency. Using Equations 2.20 and 2.25, the dependence on the CBO frequency is seen as

$$\begin{aligned}\omega_{VW}(t) &= 2\pi(f_c - 2f_{y_{BO}}), \\ &= 2\pi\left(f_c - 2f_{cbo}(t)\sqrt{2f_c/f_{cbo}(t) - 1}\right),\end{aligned}\quad (5.25)$$

where $f_{cbo}(t) = \omega_{cbo}(t)/2\pi$ is determined in the tracking analysis as described in Section 5.3.1 and given by Equation 5.22. While Equation 5.25 is the theoretical frequency for the VW effect, it was found in the tracking analysis that including an adjustment factor on the CBO frequency $f_{cbo} \rightarrow \kappa f_{cbo}$ on the order of about a percent resulted in better agreement with the directly measured VW frequency [86, 99]. In the fitting function itself, the VW frequency is then taken as

$$\omega_{VW}(t) = 2\pi\left(f_c - 2 \cdot \kappa_{VW} \cdot f_{cbo}(t)\sqrt{2f_c/(\kappa_{VW} \cdot f_{cbo}(t)) - 1}\right), \quad (5.26)$$

where now the VW frequency fit parameter is κ_{VW} . The origin of this extra factor is unclear, whether it is a tracker measurement issue, something to do with the frequency function approximation, or if it has something to do with the electrostatic quadrupoles. The inclusion of the extra factor however provides better fits and so it is kept in.

left in the fit function.

5.3.3 Lost muons

Muons lost from the storage ring during the frequency analysis portion of each fill will distort the observed decay positron spectrum. These hits show up as a rise at low frequencies in the FFT of the fit residuals due to the slow nature of the effect, Figure 5.15. These muon losses typically originate from those muons with large betatron amplitudes which hit material near the edge of the storage ring, or those

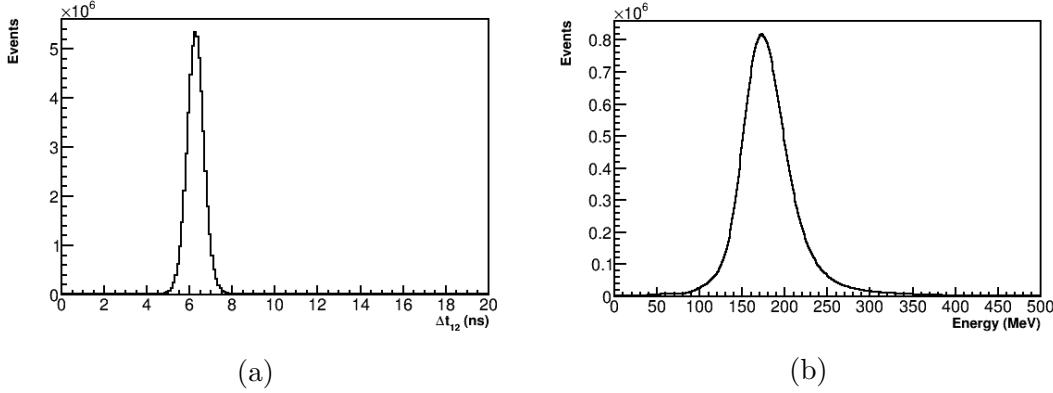


Figure 5.19: Δt and energy deposition distributions for lost muons passing through adjacent calorimeters. A typical flight time is 6.5 ns and energy deposition is 170 MeV. Data from the Endgame dataset.

muons which experience local field perturbations one too many times. In both cases the muons will lose energy and spiral inward out of the ring, a fraction of which will then pass through multiple calorimeters. Because lost muons are minimum-ionizing particles (MIPs), they are relatively easily identified by their small energy deposition signature in hit calorimeters, around 170 MeV as shown in the left peak in Figure 5.4. These lost muons typically have a flight time between adjacent calorimeters of $\Delta t_{12} = 6.5$ ns [89, 100]. The Δt and energy deposition distributions are shown in Figure 5.19. By looking for coincidences between three adjacent calorimeters, or triples, and then applying cuts and subtracting backgrounds, a pure sample of lost muons can be constructed. This lost muon spectrum $L(t)$ can then be implemented into the fit function in order to account for the positrons that would have been observed in the absence of losses.

The cuts used for the lost muon selection are given in Table 5.4. Triple coincidences are only included where every cluster consists of three or less crystals hit, with 80% of the energy deposited in one crystal. Δt and energy deposition ranges are taken as $5 \text{ ns} \leq \Delta t_{12,23} \leq 7.5 \text{ ns}$ and $100 \text{ MeV} \leq E_{1,2,3} \leq 250 \text{ MeV}$, where these ranges come from inspection of Figure 5.19. The Δt_{12} distribution as a function of time in-fill is

Lost Muon Cuts	
Parameter	Value or Range
Cluster size	≤ 3 crystals
Cluster energy fraction	≥ 0.8 in main crystal
Time of flight between adjacent calorimeters	$5 \text{ ns} \leq \Delta t_{12,23} \leq 7.5 \text{ ns}$
Energy deposition	$100 \text{ MeV} \leq E_{1,2,3} \leq 250 \text{ MeV}$
Time of flight between separated calorimeters	$\Delta t_{13} \leq 14.4 \text{ ns}$

Table 5.4: Lost muon selection cuts. In a triple coincidence the subscripts of 1, 2, and 3 correspond to the three calorimeters hit clockwise around the ring.

shown in Figure 5.20. By examining this distribution in the range $2 \text{ ns} \leq \Delta t_{12} \leq 4 \text{ ns}$, and averaging the contained counts, an approximation for the accidental background can be determined and subtracted off the triples spectrum. The accidental background typically comes from either double coincidences and a real positron hit, or a particle shower induced by an incident positron which hits an adjacent calorimeter.

Also shown in Figure 5.20 are two bands of stable beam contaminants corresponding to stored deuterons and protons. These particles have different times of flights between calorimeters due to their larger masses. By looking at the Δt_{13} distribution for times greater than $300 \mu\text{s}$, the deuteron population is easily isolated, Figure 5.21. While the deuteron population is mostly removed by the Δt_{12} cut, an additional cut of $\Delta t_{13} \leq 14.4 \text{ ns}$ helps remove any remaining deuteron contamination. The proton population, due to its nearness to the real lost muon population, is harder to remove. The simplest solution is to simply cut on the negative side of the Δt_{12} or Δt_{13} distributions. See Section 5.5.4 for the results using this additional cut. It was found that the proton contamination makes almost no difference to the fitted value of R. The default choice then is to use the previously specified cut ranges in order to increase the amount of statistics in the lost muons distribution with which to fit. A study into the exact rate of these beam contaminants is included in Reference [101].

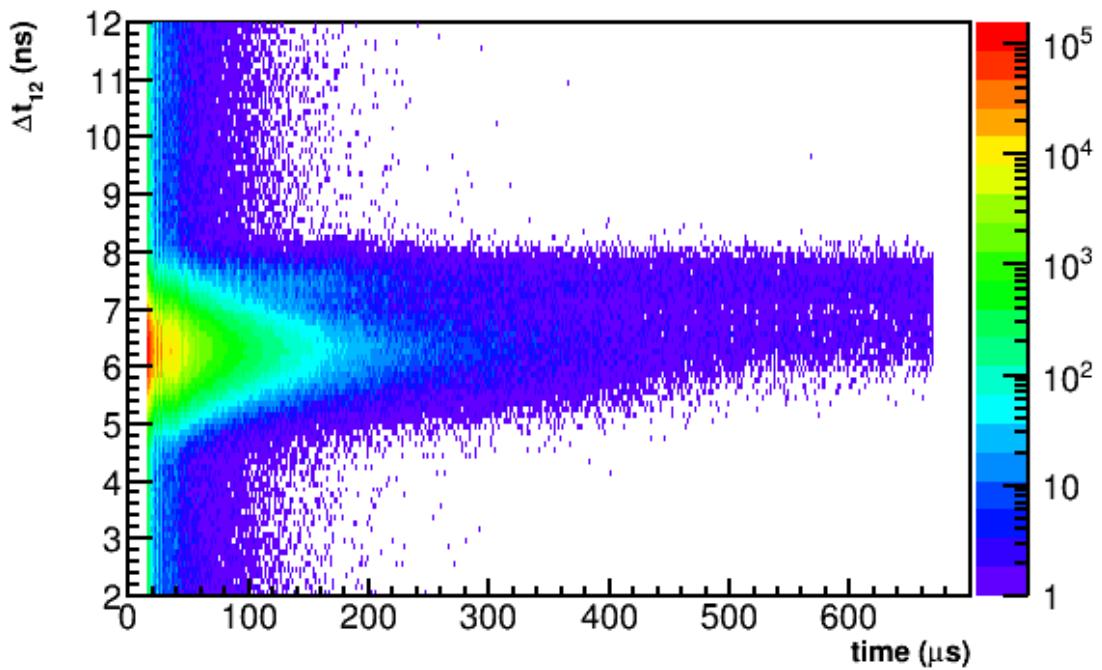


Figure 5.20: Δt_{12} distribution as a function of time in-fill before any cuts. Note the log scale. Lost muons have a Δt_{12} distribution centered at 6.5 ns. The accidental background can be seen as counts out at Δt 's far from the center of the distribution. Color striations in the core of the distribution correspond to CBO periods. There are two bands of hits that do not fall off as severely with time as the lost muons do. The band contained mostly between 7 and 8 ns corresponds to deuterons, while the band contained mostly between 6 and 7 ns corresponds to protons. Data from the Endgame dataset.

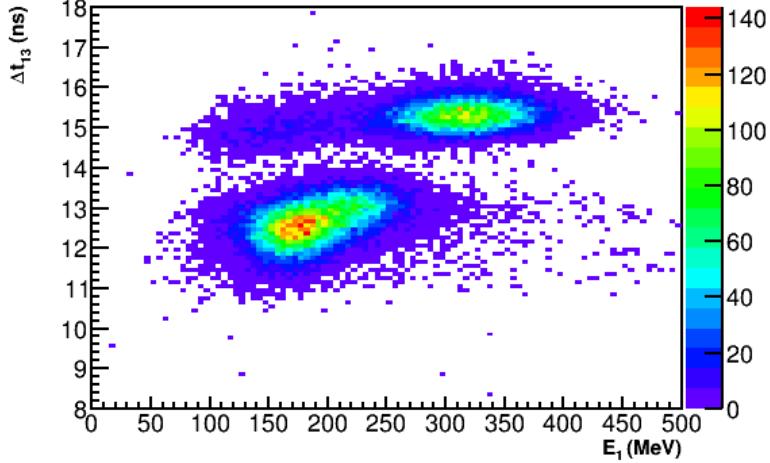


Figure 5.21: Δt_{13} distribution as a function of energy for times greater than 300 μ s. The lost muons can be seen as a blob centered at 170 MeV and $\Delta t_{13} \approx 12.5$ ns, while the deuterons can be seen as the oblong blob at $\Delta t_{13} \geq 14.4$ ns. While the deuterons have a preferentially larger energy deposition, they can be seen to extend to low energies, making cutting on energy unrealistic. Though not easily separated by eye, the stored protons are contained within the upper right portion of the lost muons blob. Data from the Endgame dataset.

The last background is the quadruples spectrum. Due to how the triple coincidences are constructed, real quadruples will be counted as two separate triples. While the quadruples spectrum can be used instead of the triples spectrum for a purer sample of lost muons, the amount of statistics is much reduced, and similarly for higher order coincidences. The quadruples spectrum is constructed in the same was as the triples with the same cut ranges. The quadruple background is removed by subtracting off those triples which originated from quadruple coincidences.

Figure 5.22 shows the final Δt_{12} distribution with cuts as a function of time infill for selected lost muons. The final 1D triple losses spectrum $L(t)$ is shown in Figure 5.23. Once we have the lost muon distribution we need to include it in the fit function in order to account for the changing number of hits over the course of the fill. The true lost muon rate will be given by $L(t)/\epsilon$, where ϵ is the loss detection

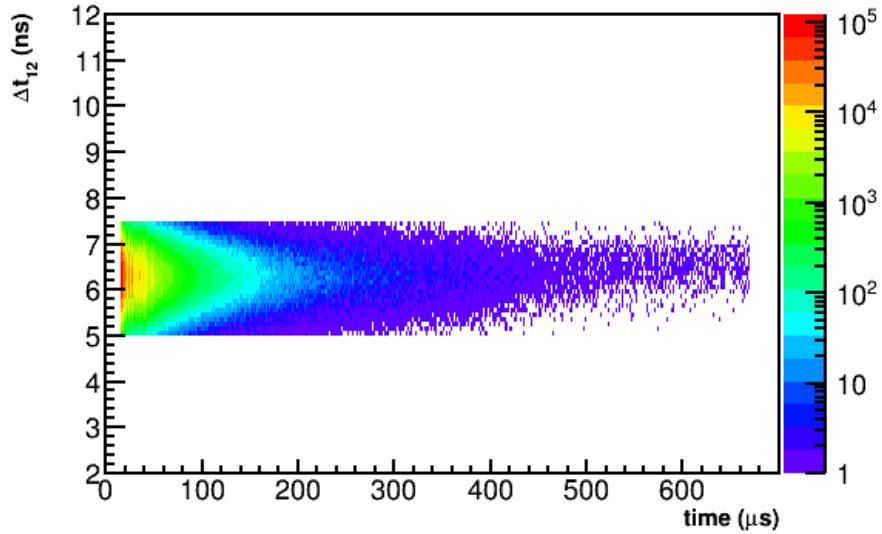


Figure 5.22: Final Δt_{12} distribution for selected lost muons as a function of time in-fill. Data from the Endgame dataset.

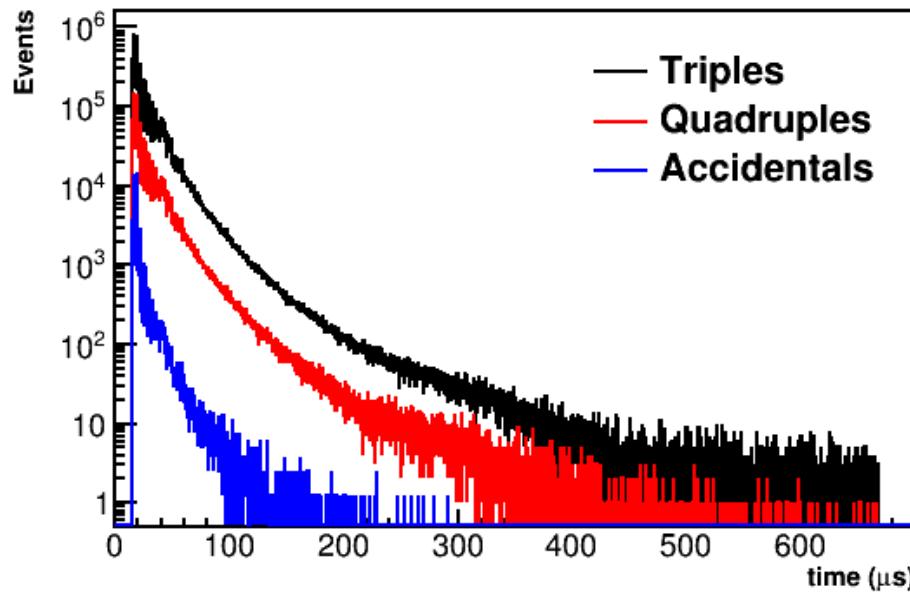


Figure 5.23: Lost muon triples spectrum $L(t)$ as a function of time in-fill, after all cuts and background subtractions. Also shown are the quadruples and accidental background distributions, which can be seen to be much smaller than the triples. The shape of the triple and quadruple spectra can be seen to be nearly the same. The shape of the triples spectrum comes from beam dynamics (CBO) and lost muon acceptance effects. Data from the Endgame dataset.

efficiency [87]. The change in the number of muons within the storage ring can be written as

$$dN = -\frac{N}{\tau}dt - \frac{L(t)}{\epsilon}dt, \quad (5.27)$$

where the solution can be determined by inspection as

$$N(t) = N_0 \cdot e^{-t/\tau} \cdot \left(1 - \frac{1}{\epsilon N_0} \int_{t_0}^t L(t')e^{(t'/\tau)}dt'\right). \quad (5.28)$$

The parameter τ is the same muon lifetime as in the T-Method fit function, except here it is set as the default value of 64.44 μs ¹³. The value for t_0 can be taken as any time at or before the start of the fitting range, as it simply changes the scale of N_0 which is largely unimportant. The modification $\Lambda(t)$ as listed in Equation [5.15] is then taken as

$$\Lambda(t) = 1 - \kappa_{loss} \int_{t_0}^t L(t')e^{(t'/\tau)}dt' \quad (5.29)$$

where $\kappa_{loss} = 1/\epsilon N_0$ is taken as the fit parameter¹⁴.

It should be noted that the Ratio Method is largely insensitive to the slow lost muons effect which divides out. No rise appears in the FFT of the ratio fit residuals, Figure [5.16] and letting κ_{loss} float in the fit does not produce fit nice convergence. It was found however that R changes on the order of tens of ppb when the lost muon term is included with κ_{loss} fixed to the value determined from a T-Method fit, so by default it is included. See Section [5.5.4] for more discussion on this.

¹³The fitted ω_a frequency is largely insensitive to this parameter, and especially so in the Ratio Method fits, so this is acceptable.

¹⁴By construction κ_{loss} is a very small parameter, $\mathcal{O}(10^{-10})$. This factor is absorbed into the parameter in such that the fit parameter is $\mathcal{O}(1)$.

5.3.4 Fit procedures and parameters

χ^2 fits are done to the pileup corrected positron time spectra and ratio data in `ROOT` using the standard `TH1F` and `TGraph` fit methods with a strategy level of two. Fits are performed in stages where groups of associated parameters are freed, fit, and then fixed before fitting the next set of parameters. In the final fit all parameters are freed and fit at once. Each of the four Run 1 precession frequency analysis datasets are fit with 14 parameters. 13 parameters are free in fits to the 60h, 9d, and Endgame datasets, with the κ_{loss} parameter determined and fixed from a T-Method fit. In the HighKick dataset τ_{cbo} is also fixed from a T-Method fit as the lifetime is relatively small and the Ratio Method can't successfully fit it, resulting in 12 free fit parameters. In the T-Method fits there are two extra free fit parameters, those being the N_0 and τ terms which are left out of the Ratio Method fits.

Fits to calorimeter sum spectra are done from 30.2–650 μs , corresponding to 4155 bins. The choice of fit start time was made to lie directly on a $g - 2$ zero crossing, which from E821 was shown to reduce some systematic errors [cite this?](#). It was also made at a late enough time to allow for storage of the muon beam after the scraping procedure. The choice of fit end time was made to be just before the quad storage field turned off, in order for the fill time to fit within the timing structure of the experimental beam. Individual calorimeter fits were performed out to 400 μs , in order to stay within the region of Poisson statistics such that χ^2 fits are satisfied.

Table 5.5 gives the various fit procedure parameters.

As a reminder R is blinded at the hardware and software levels [94, 95]. The software blinding string used for the 60h dataset was different than that used when fitting the HighKick, 9d, and Endgame datasets, each of which used the same blinding string. This was done in order to perform a software-level relative unblinding exercise between different analyzers with the 60h dataset in order to determine if there were any

Fit Procedure Parameters	
Parameter	Value
Fit strategy level	2
Fit start time	30.2 μ s
Fit end time (calorimeter sum)	650 μ s
Fit end time (calorimeters)	400 μ s
Bins in fit (calorimeter sum)	4155
Number of fit parameters	14
Number of free fit parameters (60h, 9d, Endgame)	13
Number of free fit parameters (HighKick)	12

Table 5.5: Various parameters used in the fit procedure.

obvious problems in the analyses [102]. **Cite a summary talk somewhere? Can't find any.** Therefore in the following results, R values between different datasets are comparable between the datasets except for the 60h, barring any differences due to field conditions which are not discussed in this dissertation.

5.4 Fit results

Figure 5.24 shows fits to the four Run 1 precession frequency analysis datasets. Table 5.6 gives all fit parameters and their errors. In each dataset case the χ^2/NDF is acceptable as evidenced by the p value included in the table results. Fit pulls and the FFT of the fit residuals for the 60h dataset are provided in Figure 5.25. As shown all structure has been eliminated within the fit residuals implying that all effects in the data have properly been accounted for in the fit function. The same checks were made for the HighKick, 9d, and Endgame datasets and in each case no residual structure remained. Figure 5.26 shows the correlation matrix for the fit to the 60h dataset. The only fit parameter that is significantly correlated with R is the $g - 2$ phase. This increases the confidence in the ω_a extraction, as effects in the data which might potentially be mis-modelled will only weakly correlate with the final fitted R .

value. The various different CBO parameters are self-correlated to different degrees depending on the parameter and the dataset that is being fit. Typically either the phases and frequencies are correlated, or the lifetimes and amplitudes. Appendix D provides the correlation matrices for the HighKick, 9d, and Endgame datasets.

The $g - 2$ phases for the different datasets showed small differences, due primarily to upstream beam adjustments before injection into the storage ring. Similarly, slightly different asymmetries can be attributed to very small acceptance differences in the stored beams. As described in Section 5.3.3, the value for κ_{loss} was determined and fixed from a T-Method fit to the data. The values themselves don't directly correspond to the level of losses, as each dataset has it's own loss function $L(t)$ which is left out of the picture here.

The CBO frequencies for the 60h and Endgame datasets with n values of 0.108 were found to be 2.338 and 2.339 rad/ μ s respectively, corresponding to approximately 0.37 MHz. For the HighKick and 9d datasets with n values of 0.120, the CBO frequencies were found to be 2.559 and 2.615 rad/ μ s respectively, corresponding to approximately 0.415 MHz. These frequencies correspond to the expected frequencies as described in Section 2.6, with some slight deviations due to statistics and the reduced sensitivity in the Ratio Method¹⁵. The CBO lifetimes between the different datasets are relatively consistent, barring the HighKick dataset for which a smaller CBO lifetime was measured. Ratio Method fits typically converge with lifetimes with large errors compared to T-Method fits, due to the reduction in sensitivity in the Ratio Method. In the HighKick dataset, the CBO lifetime did not like to converge nicely in the ratio fits, and was therefore fixed to that from a T-Method fit. The main CBO amplitudes A_{cbo-N} for the different datasets were on the order of 0.3–0.4%, while the higher order CBO amplitudes were in general an order of magnitude less. The

¹⁵The VW frequencies, though time-randomized out in the analysis presented here, were found to be approximately 2.30 and 2.04 MHz for the datasets with $n = 0.108$ and $n = 0.120$ respectively.

strength of the various higher order CBO amplitudes fluctuated for different datasets, with one parameter being large compared to another in one dataset and vice versa in a different dataset. In some cases, the errors on the higher order CBO term amplitudes were of the order the amplitude itself. While this implies these terms can be dropped from the fit function, all terms were included for analysis uniformity among the different datasets. These relatively large errors, while making some of the fits slightly more challenging to achieve fit convergence for, were nonetheless handled appropriately.

The final statistical errors on R for the 60h, HighKick, 9d, and Endgame datasets are 1.358, 1.411, 0.903, and 0.639 ppm respectively. The single seed R results for the HighKick, 9d, and Endgame datasets, all of which used the same blinding string, are all well within 1σ of each other. The average R value for fits to 50 different random seeds are provided in Section 5.4.5.

Beyond looking at single fit residuals to evaluate the integrity of the fits, other checks were made to verify consistency. In general this consisted of slicing up the data in different ways and fitting the subsets. These tests and scans included fitting individual calorimeters, modifying the fit start and end times, adjusting the applied energy thresholds, and fitting individual beam bunches.

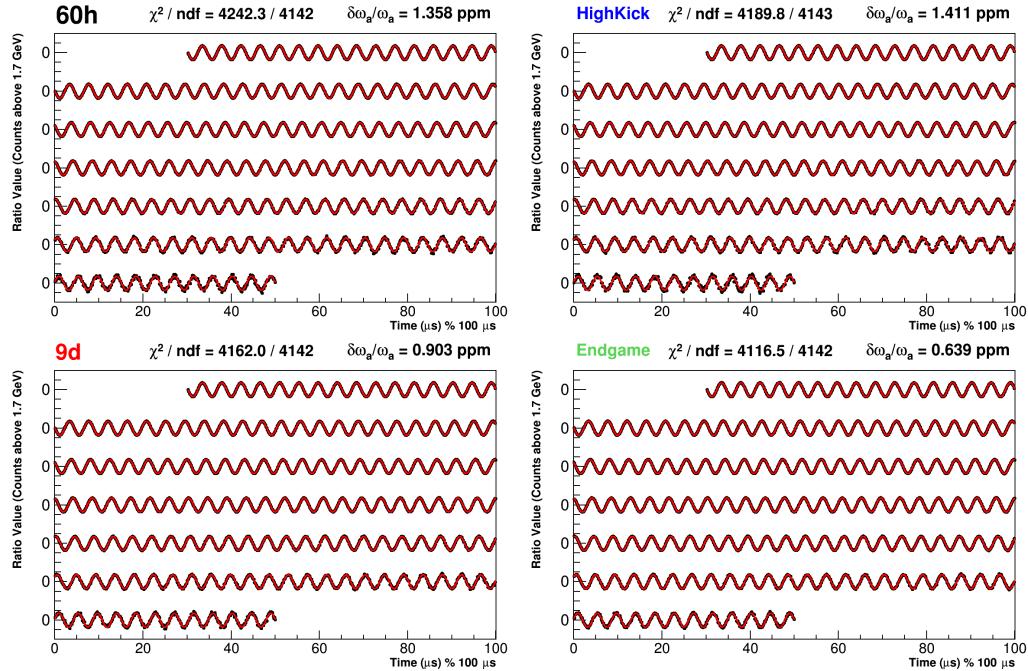


Figure 5.24: Single random seed fits to calorimeter summed data of Run 1 precession frequency analysis datasets. Data is in black and the fits are in red. The x axis is in units of μs modulo $100 \mu\text{s}$, with successive portions of the data points and fit shifted downwards on the plot. The fit ranges from $30.2\text{--}650 \mu\text{s}$. Dataset names are given in the upper left corners of the figures, alongside the χ^2 per degree of freedom and relative error on ω_a or R .

Ratio Method Fit Results								
	60h		HighKick		9d		Endgame	
χ^2/NDF	4242/4142		4190/4143		4162/4142		4116/4142	
p value	0.1356		0.3018		0.4104		0.6079	
Parameter	Value	Error	Value	Error	Value	Error	Value	Error
A	0.3637	4.4×10^{-5}	0.3632	4.6×10^{-5}	0.3639	2.9×10^{-5}	0.3686	2.1×10^{-5}
R (ppm, blinded)	-20.8479	1.3581	-17.5433	1.4112	-17.8214	0.9033	-17.5674	0.6393
ϕ	2.091	2.2×10^{-4}	2.081	2.3×10^{-4}	2.080	1.5×10^{-4}	2.076	1.1×10^{-4}
ω_{cbo} (rad/ μ s)	2.338	1.4×10^{-3}	2.599	6.6×10^{-3}	2.615	5.6×10^{-3}	2.339	0.8×10^{-3}
τ_{cbo} (μ s)	175.2	46.8	99.4	0	137.4	62.0	200.3	33.5
A_{cbo-N} ($\times 10^{-4}$)	43.1	5.0	42.8	9.9	39.3	9.7	32.3	2.0
ϕ_{cbo-N}	-2.343	0.107	3.817	0.446	3.302	0.374	-0.710	0.062
A_{2cbo-N} ($\times 10^{-4}$)	1.9	1.3	4.9	4.5	2.2	2.7	1.2	0.5
ϕ_{2cbo-N}	3.331	0.638	5.665	1.274	-4.936	1.127	0.322	0.448
A_{cbo-A} ($\times 10^{-4}$)	5.5	3.9	9.5	4.1	6.4	2.5	2.7	1.9
ϕ_{cbo-A}	-0.271	0.737	-2.073	0.600	1.750	0.561	-2.825	0.686
$A_{cbo-\phi}$ ($\times 10^{-4}$)	8.0	4.2	5.7	4.4	8.8	3.1	1.9	1.9
$\phi_{cbo-\phi}$	-1.183	0.533	1.227	0.920	4.313	0.415	-1.576	0.995
κ_{loss}	8.974	0	5.651	0	2.510	0	2.345	0

Table 5.6: Fit parameters for the four Run 1 precession frequency analysis datasets. The bold row highlights the final fitted R values and their respective errors. As a reminder the 60h dataset has a different blinding to the rest. The κ_{loss} parameter is fixed in each dataset fit corresponding to the 0 value in the error column, and similarly for τ_{cbo} in the fit to the HighKick dataset.

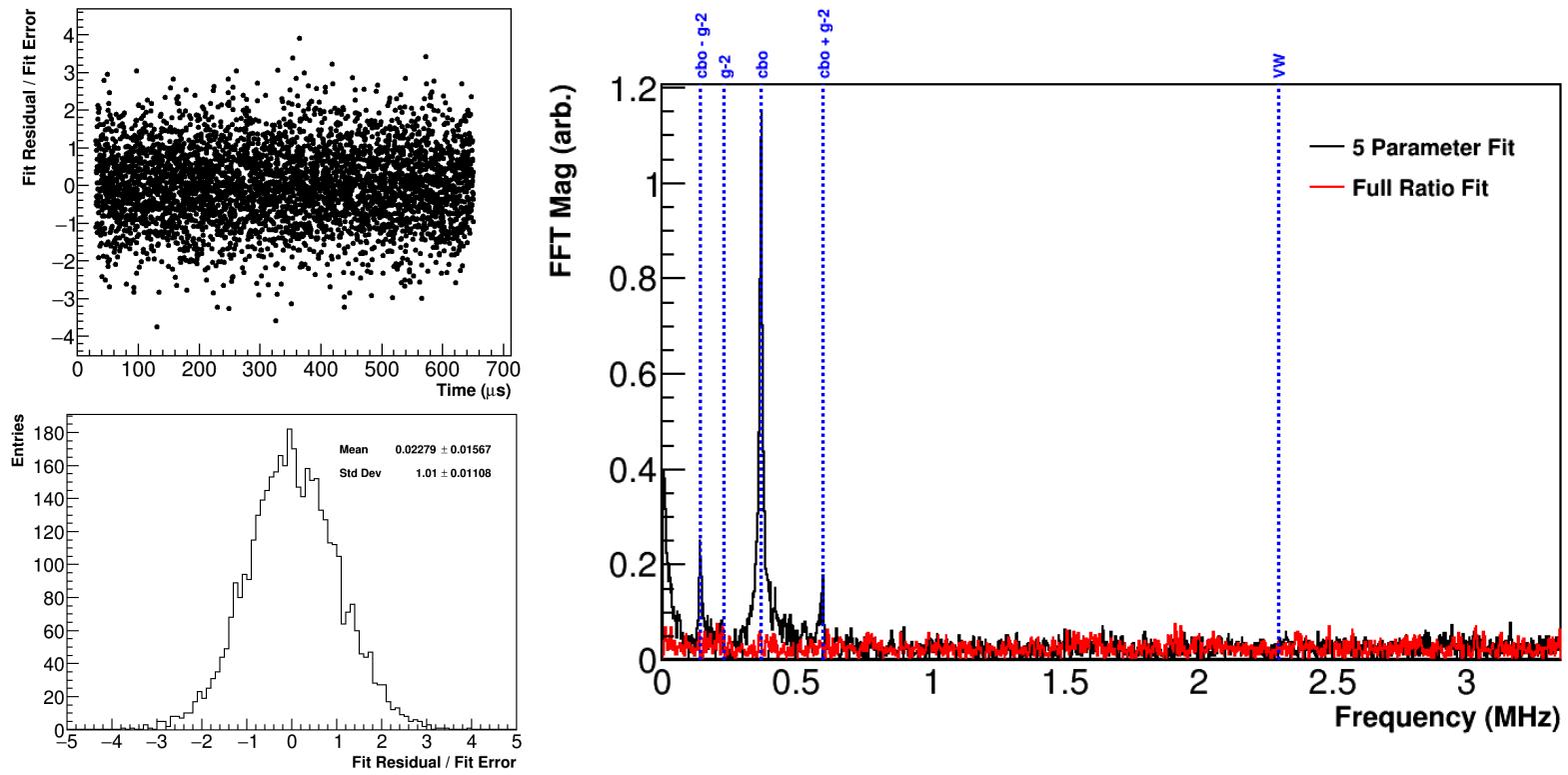


Figure 5.25: Fit pulls (top-left), their projection on to the y axis (bottom-left), and the FFT of the fit residuals (right) for the ratio fit to the 60h dataset. Note the pull projection has a Gaussian shape centered around zero with unit width. In the FFT the results from a five parameter fit to the data are overlayed, along with blue dashed lines for the main beam dynamics peaks which appear in the data. There is no obvious structure in the pulls and no remaining peaks above the noise in the ratio fit FFT.

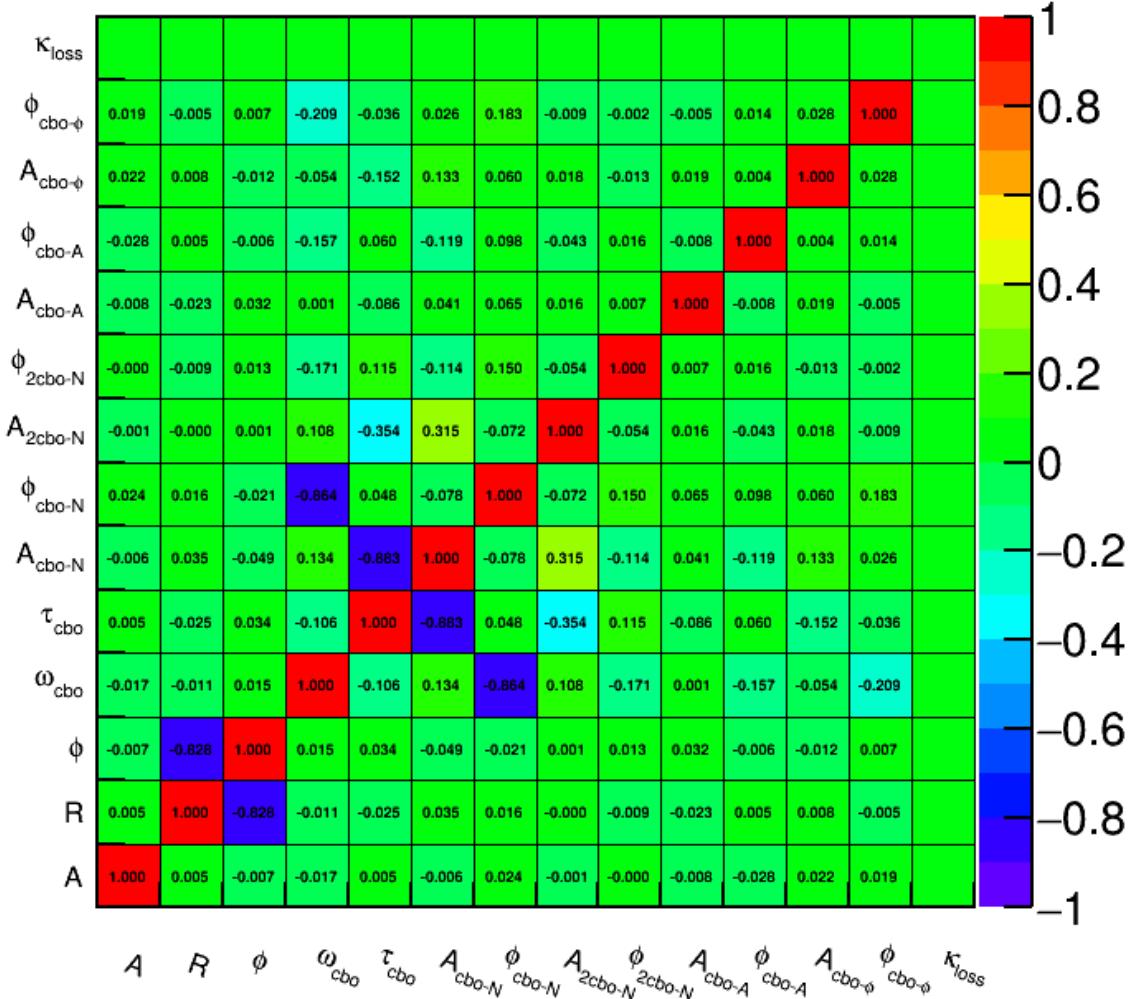


Figure 5.26: Correlation matrix for the single seed ratio fit to the 60h dataset. The only significant correlation with R is the $g-2$ phase. κ_{loss} is fixed, hence the corresponding empty row and column.

5.4.1 Individual calorimeter fits

Fits to all 24 individual calorimeters for each of the datasets were performed with the same number of free fit parameters as used in the calorimeter summed fits. Figure 5.27 shows the χ^2/NDF 's for the calorimeter fits which are nicely spread around 1. Figure 5.28 shows the fitted R values as a function of calorimeter number. Straight line fits were performed to the R values, and the fitted constant returned a value in each case that was consistent with the calorimeter sum fit R values. Examining the R values as a function of calorimeter between datasets, particular calorimeter numbers do not tend to lie above or below the fitted line uniformly. The spread in R values for each calorimeter then can be said to be driven statistically, though it should be noted that with the larger error bars on the individual calorimeter fits it's hard to tell if there are any preferences one way or another.

Figures 5.29 and 5.30 show calorimeter fit results for the other free parameters in the fit for the Endgame dataset. The $g - 2$ phases are relatively consistent among the different calorimeters, barring calorimeters 13 and 19 which lie lower on the plot. These two calorimeters sit behind the tracker stations, implying a different acceptance and therefore a slightly different $g - 2$ phase. Any correlated effects on R are not immediately observed, and might potentially be hidden behind the large errors of the fit. Similarly, the different calorimeters have different fit asymmetries, once again due to their different acceptances. The CBO parameters are in general consistent with some spread due to acceptance, with the phases running from $0-2\pi$ around the ring as expected. As one might notice, the amplitudes of the CBO parameters are an order of magnitude higher than in the calorimeter summed fits. Because the phases vary around the ring, when adding up all the calorimeters the CBO effect becomes reduced. In fact, while it is not always necessary to include the higher order CBO terms for good fits to the calorimeter summed data, there are many calorimeters

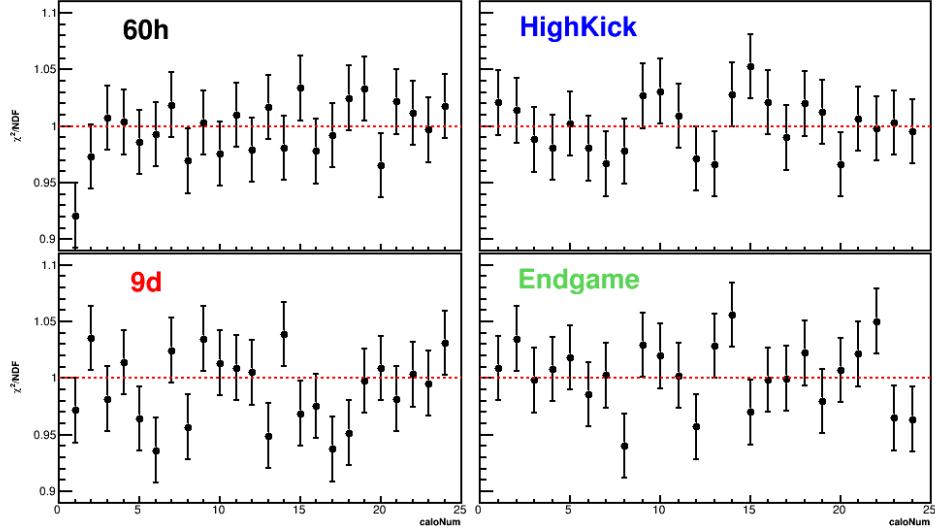


Figure 5.27: χ^2/NDF versus calorimeter for the Run 1 precession frequency analysis datasets. Red dashed lines are placed at $\chi^2/\text{NDF} = 1$ to aid the eye. No individual calorimeter fits are preferentially low or high when comparing across datasets.

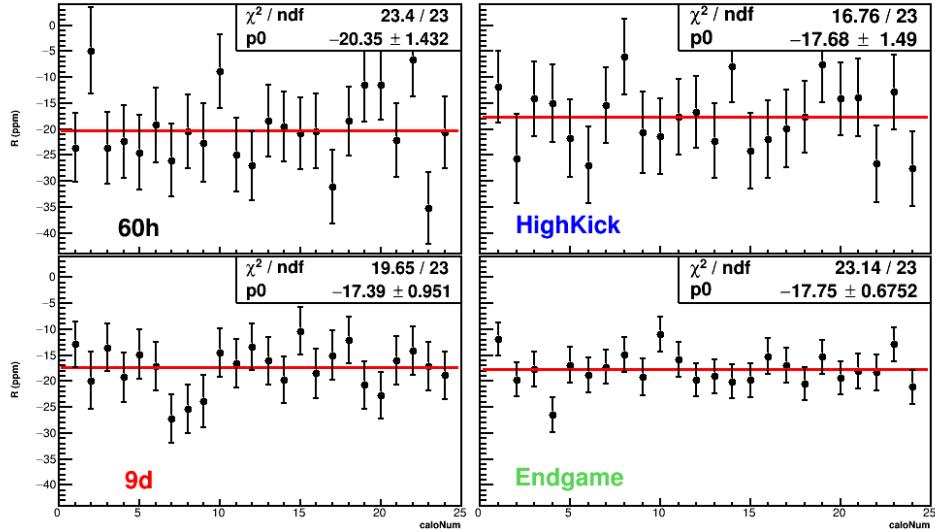


Figure 5.28: R versus calorimeter for the Run 1 precession frequency analysis datasets. The scale is the same on each of the plots. A straight line fit was performed on the fitted values, with the fit result shown in the upper right box as parameter p_1 in units of ppm. The different blinding in the 60h dataset is readily observed, along with the higher precision fits in the 9d and Endgame datasets with their correspondingly smaller error bars.

which need the higher order terms for good fits.

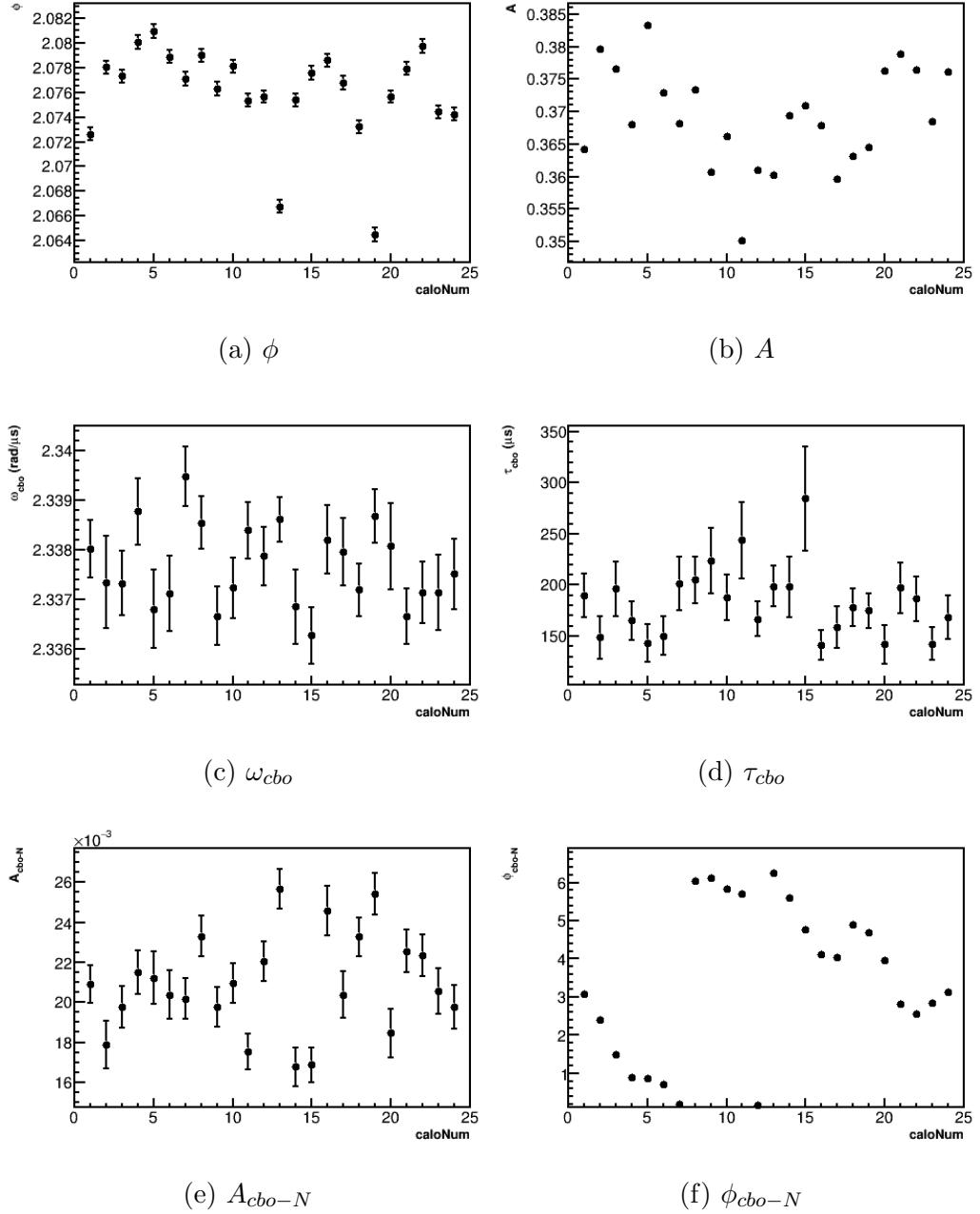


Figure 5.29: Endgame fit parameters versus calorimeter number. In the $g - 2$ phase ϕ (top-left) two points lie below the others, corresponding to those calorimeters which sit behind tracker stations. The CBO phase ϕ_{cbo-N} (bottom-right) runs from $0 - 2\pi$ around the ring. These plots are typical of all datasets, with small variations in the final fitted parameters.

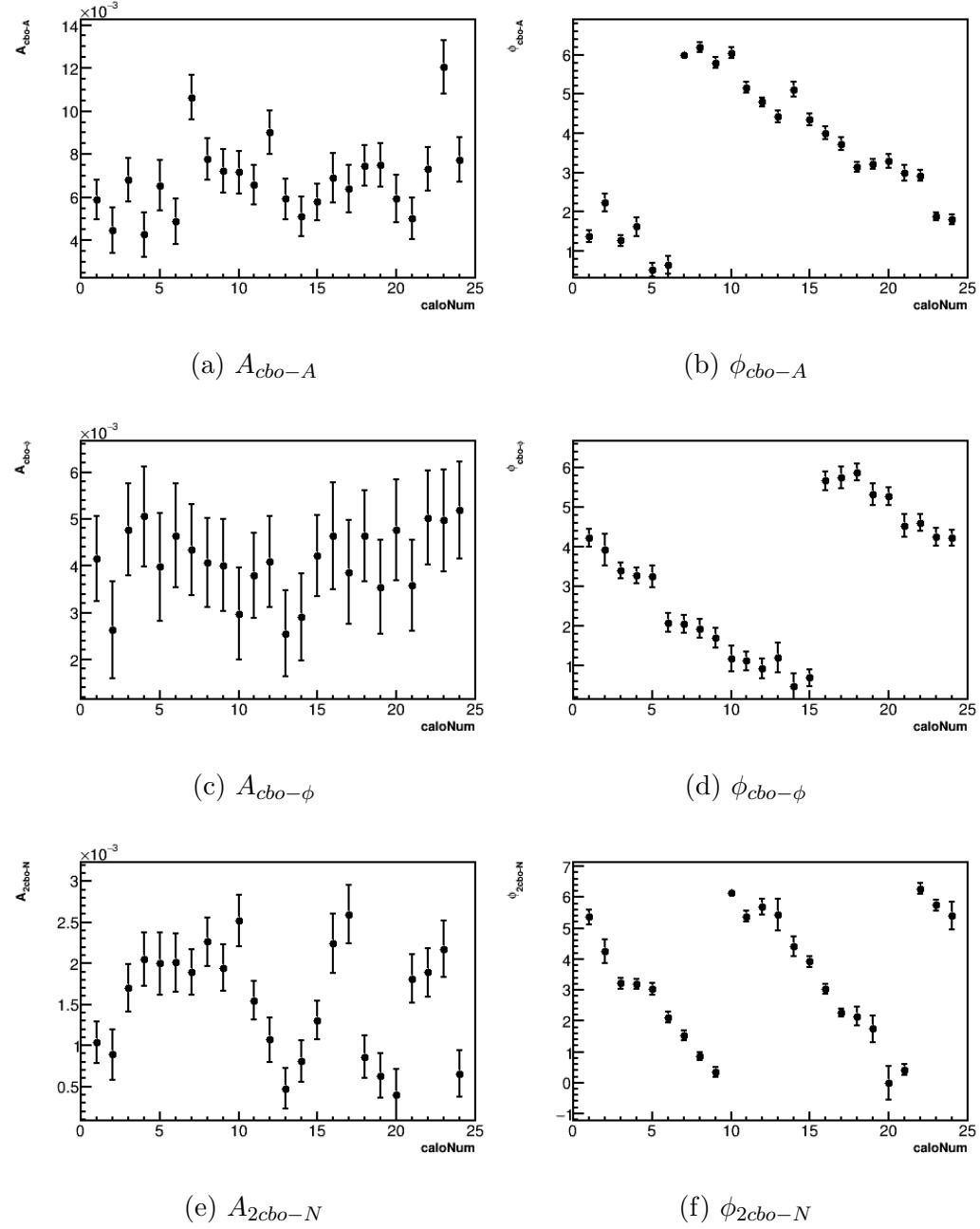
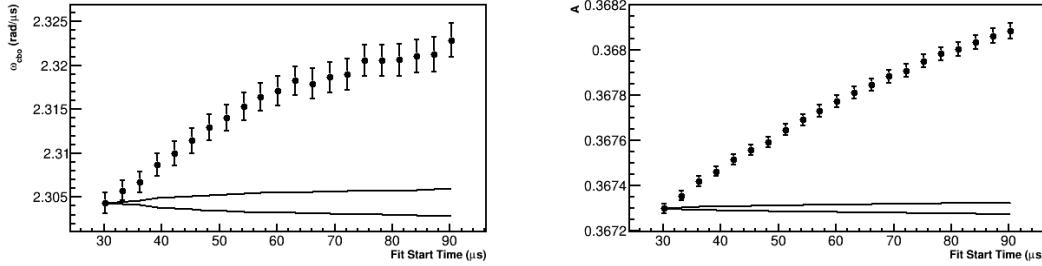


Figure 5.30: Endgame fit parameters versus calorimeter number. The CBO phases ϕ_{cbo-A} and $\phi_{cbo-\phi}$ run from $0-2\pi$ around the ring, while ϕ_{2cbo-N} runs around twice. These plots are typical of all datasets, with small variations in the final fitted parameters.



(a) The fit start scan for the fitted CBO frequency when a constant frequency model is included. The p value of the first ratio fit in this case is 0.191, indicating a good fit. Only by performing a fit start scan or examining the fit residuals closely is the deficiency in the fit observed.

(b) The fit start scan for the fitted asymmetry when the pileup is not subtracted out. If the pileup is not properly accounted for, then decay positrons with lower asymmetries contaminate the observed decay time spectrum leading to a lower fitted asymmetry value. As the pileup diminishes the fitted asymmetry tends to its true value.

Figure 5.31: Examples of fit start scans when effects are improperly accounted for in fitting. Data from the Endgame dataset.

5.4.2 Fit start and end scans

In order to determine if there are any deficiencies as a function of time, for instance if the CBO was modelled incorrectly, fits are performed with varying fit start and end times. If a parameter is incorrectly modelled, it will wander away from the statistically allowed deviation as the mismodelled effect grows stronger or weaker. Two examples of this are given in Figure 5.31. In general, doing a fit start scan is a very useful tool in the precession frequency analysis beyond just verifying consistency of a fit parameter as a function of time, as it also provides hints as to what might be wrong with how the data is being handled.

The statistically allowed deviation between two sets of data, where one is a subset of the other, is given by [29]

$$\sigma_{diff} = \sqrt{\sigma_2 - \sigma_1(2\frac{A_1}{A_2} \cos(\phi_1 - \phi_2) - 1)}, \quad (5.30)$$

where the subscript 2 stands for the larger dataset while the subscript 1 stands for the smaller sub-dataset. This statistically allowed deviation depends both on the size of the datasets as well as their “analyzing powers,” which come from the asymmetries and phases of the datasets. For fit start scans the analyzing powers are in general the same, such that the approximation

$$\sigma_{diff} \approx \sqrt{\sigma_2 - \sigma_1} \quad (5.31)$$

can be made.

It should be noted that for fit start scans where much of the data is dropped at late fit start times, certain CBO parameters start to become unstable as the CBO effect becomes diminished in the data¹⁶. Fits with start times at 100 μ s are half a lifetime or more along the CBO effect, meaning that amplitudes either start to go to 0 or become unfittable. In particular, the CBO lifetime itself become very hard to fit and tends to converge to whatever the upper bound of the fit limit is while sending the amplitude to zero. For those parameters which were found to be unreliable, they were fixed to their starting fit values. These typically included most of the higher order CBO amplitudes and phases, though in some cases they can still be fit out to late times depending on the dataset.

Fit start time scans were performed from the default value of 30.2 μ s up to 100.2 μ s in steps of 1 μ s corresponding to 71 separate fits. In order to assist fit convergence, the final fit parameters from one fit were passed on as the starting parameters of the next. The χ^2/NDF 's and fitted R values as a function of fit start time for the four Run 1 precession frequency analysis datasets are shown in Figure 5.32. The error on

¹⁶The same applies to the VW effect when not time-randomized out.

the individual points in the χ^2 plots are given as

$$\sigma_{\chi^2} = \sqrt{2/NDF}, \quad (5.32)$$

where NDF changes as the fit start time is pushed later in time and bins are left out of the fit. The black parabolic bands indicate the 1σ statistically allowed deviation as given by Equation 5.31^{[17](#)}. As shown the goodness of fit for the four datasets are all consistent with fit start time, only wandering in and near the bands without diverging.

R is similarly consistent as a function of fit start time. The only dataset for which R goes noticeably outside the bands more than the others is that for the HighKick. Since the deviation is less than 2σ and it appears to have a trend back towards the bands however, it does not appear particularly indicative of unaccounted effects in the data. In general R is very insensitive to changing fit start time, most likely simply due to the small correlations with other fit parameters.

Figure 5.33 shows the fit start scan results for the other free parameters for the Endgame dataset. In all cases the fit parameters only wander in and near the bands, showing that all effects in the ratio data are properly accounted for.

¹⁷These statistical deviation bands are sometimes referred to as Kawall bands.

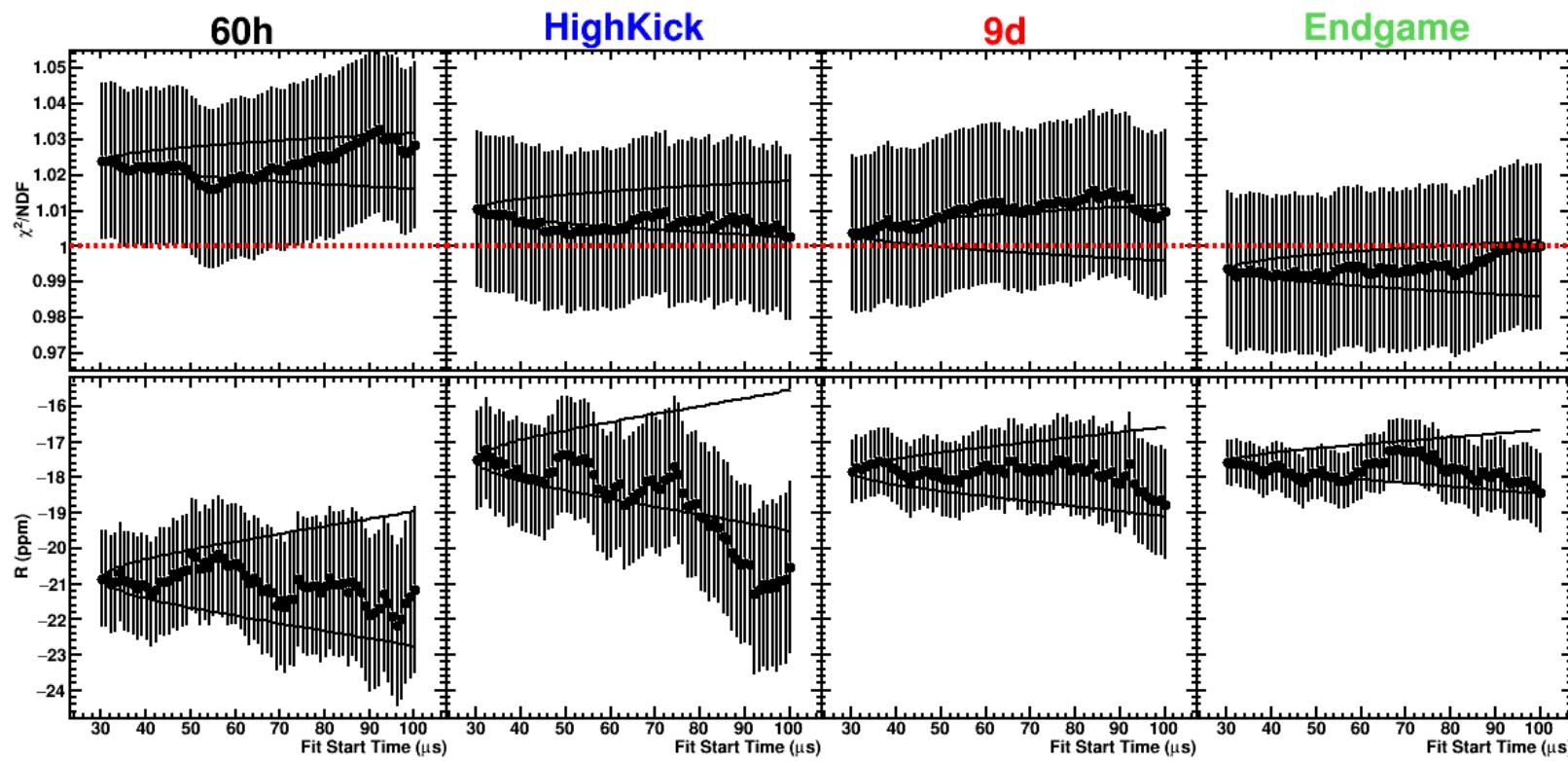


Figure 5.32: χ^2/NDF (top) and R (bottom) versus fit start time for the Run 1 precession frequency analysis datasets. Plots in the top and bottom rows are on the same scale respectively. The black parabolic bands represent the 1σ statistically allowed deviation. The χ^2 's are all consistent with 1, for which a dashed red line has been overlayed. In all cases the fit points lie in and around the 1σ statistical bands.

For fit end scans, all of the same methods and conclusions apply. In general fit end scans are both less dangerous and more stable than fit start scans, as the amount of data being removed from the fit is relatively small. While these fit end scans in a T-Method fit might be able to be completely ignored in favor of the fit start scans, it's nice to check that they satisfy the statistical deviations in the Ratio Method as the ratio data errors grow larger with less data [102]. Fit end time scans were performed from 650 μs to 400 μs in steps of 10 μs , corresponding to 26 separate fits. As in the fit start time scan, fit results from the end of one fit were passed on as the starting parameters to the next. R values for fit end scans for the Run 1 precession frequency analysis datasets are shown in Figure 5.34. As shown the R values are comfortably within and near the bands. The only dataset where R wanders a little more than the others is in the 60h dataset, however it veers back towards the bands at the end of the scan and is thus deemed acceptable.

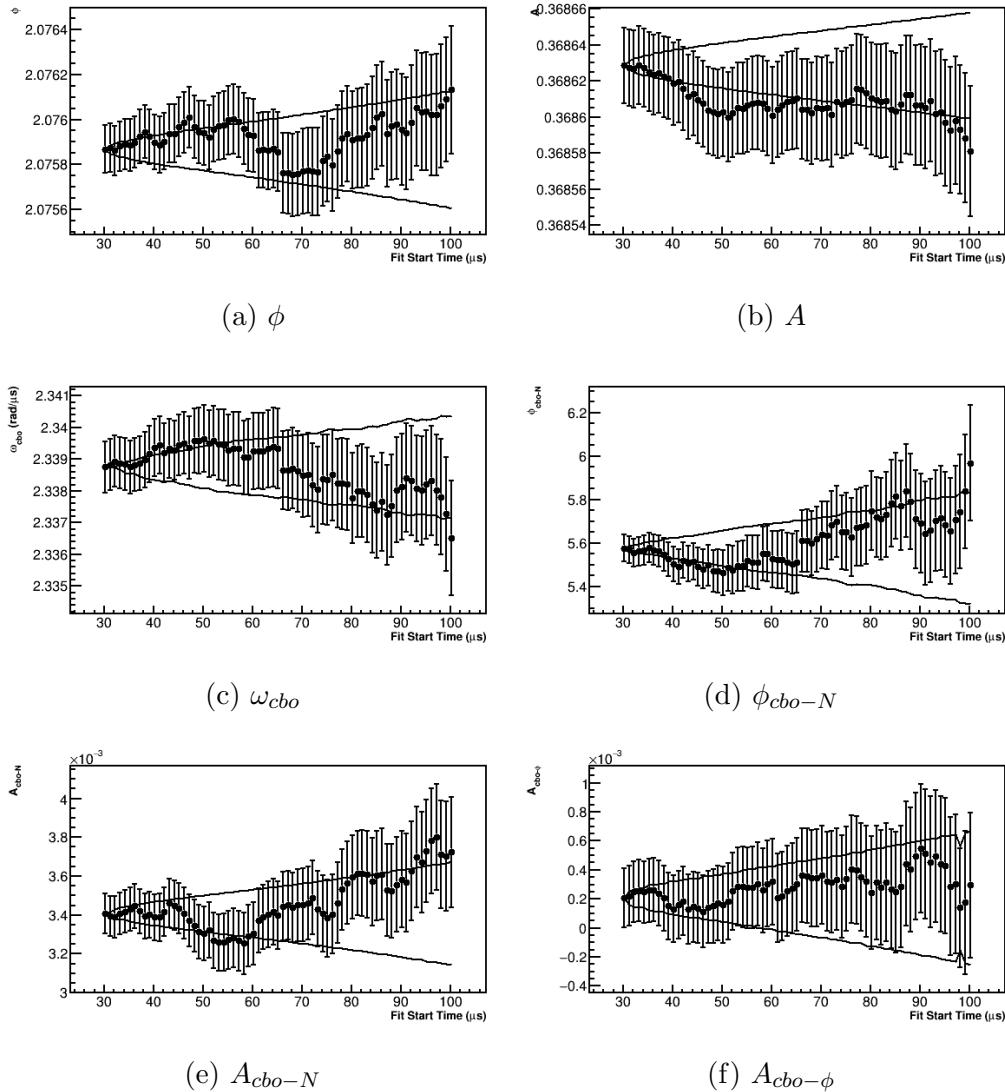


Figure 5.33: Fit start scans for free parameters in the Endgame dataset. Those parameters not shown here are fixed to their starting values over the course of the scan, as at late times they can be unstable as the CBO effects die away.

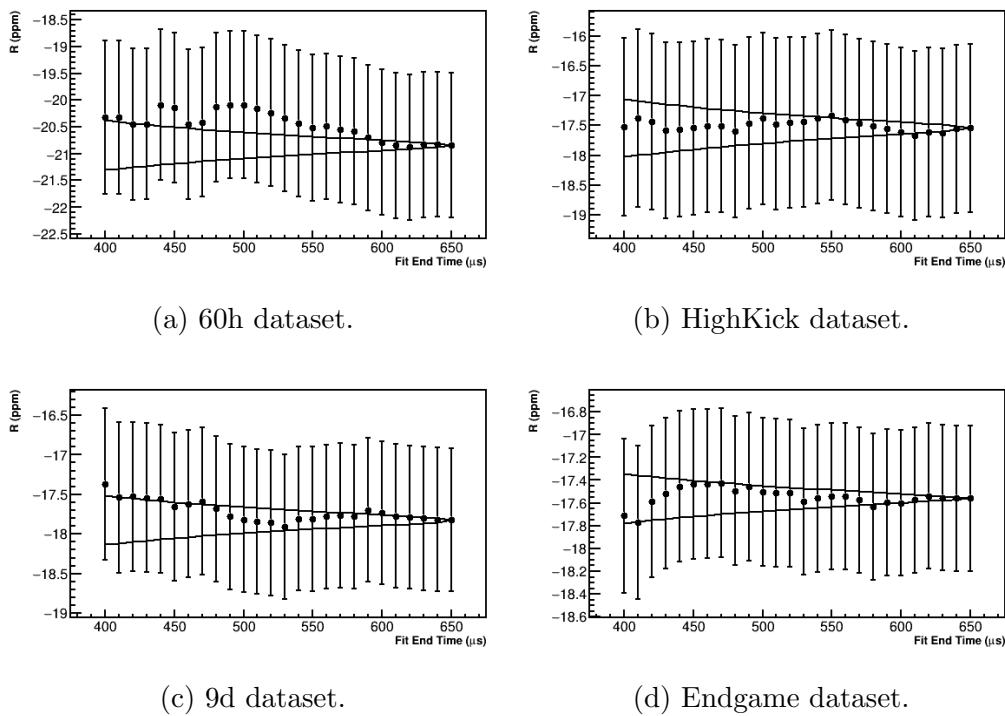


Figure 5.34: R versus fit end time for the Run 1 precession frequency analysis datasets. The fit points lie in and around the 1σ statistical bands.

5.4.3 Energy threshold scans

Similarly to fit start and end time scans, it is worthwhile to verify that R is consistent regardless of the energy threshold applied to the decay positron time spectrum. The energy threshold was varied from 1.2 GeV to 2.2 GeV in steps of 50 MeV corresponding to 21 separate fits. The fitted R values for the four Run 1 precession frequency analysis datasets are shown in Figure 5.35. The statistical bands are the same as those defined in Equation 5.30, where now the analyzing power part of the equation plays a larger role as the asymmetries and phases of the different fit points are significantly different. As shown there are no major deviations in the fitted R values.

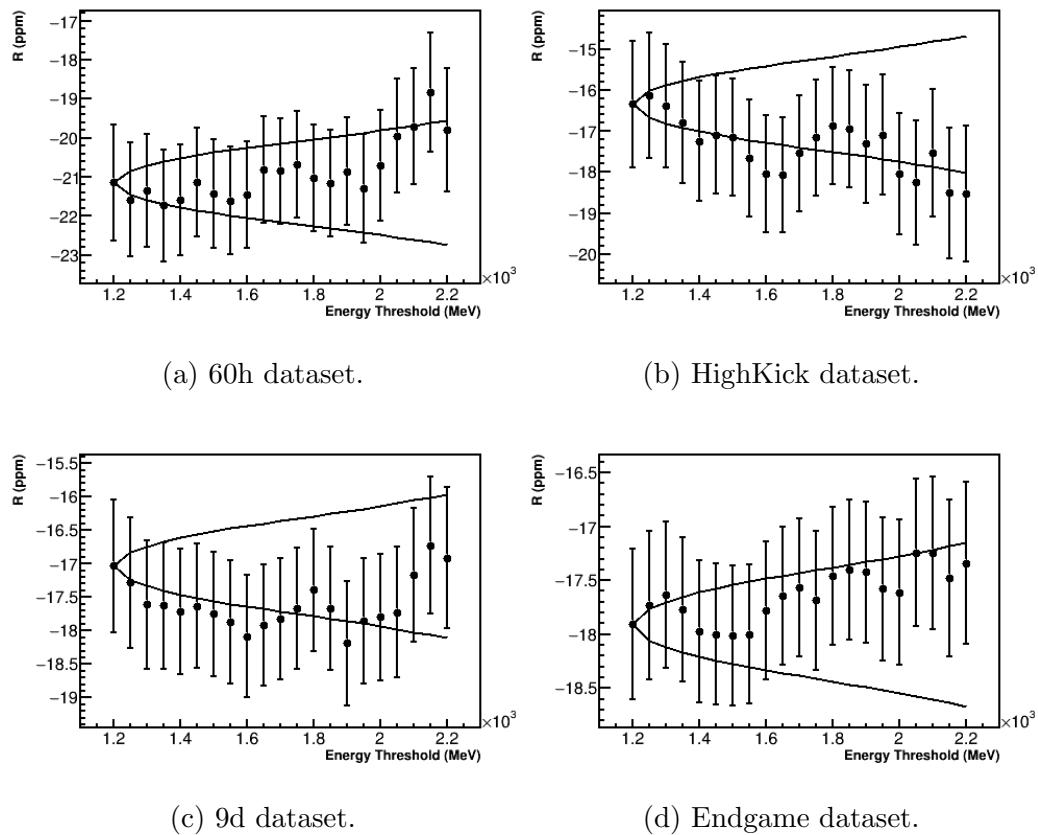


Figure 5.35: R versus energy threshold for the Run 1 precession frequency analysis datasets. The fit points lie in and around the 1σ statistical bands.

5.4.4 Fits to bunch number

As described in Section 2.3, eight distinct and separate bunches of muons are sent to the E989 experiment within the accelerator timing structure. If different bunches have different momentum spreads due to how the particles are transported down the various beamlines and injected into the storage ring, then muons might live for different times and therefore their spins might precess more or less within the magnetic field. This difference in $g - 2$ phase between the different bunches would then affect the final fitted R value, and might be a systematic error in the precession frequency measurement if certain bunches have significantly more statistics than the others. In order to check if there were any differences in the final fitted R values per bunch, the bunches were fit individually. Figure 5.36 shows the fitted R values for the eight individual bunches alongside the bunch-sum result, for the Run 1 precession frequency analysis datasets. In all cases there appear no systematically different R values per bunch. When the eight individual bunches were fit to a straight line, they were found to be very consistent with the bunch-sum result.

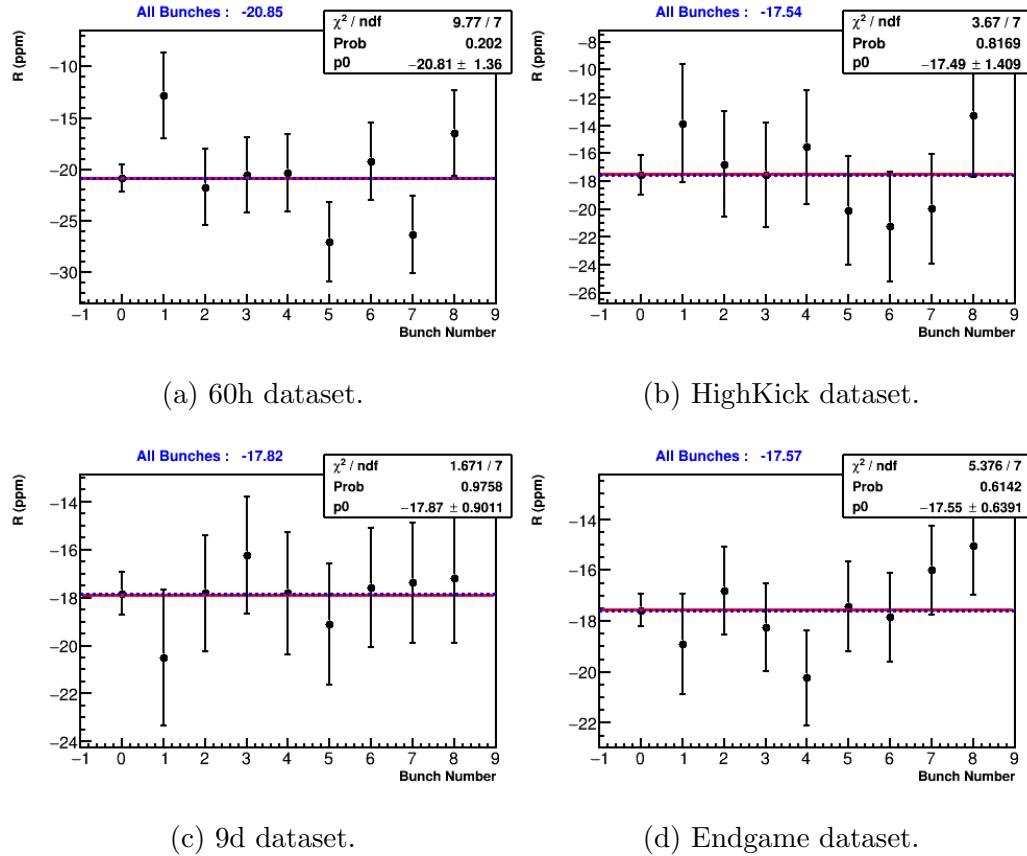


Figure 5.36: R versus bunch number for the Run 1 precession frequency analysis datasets. Bunch number 0 corresponds to the data from all bunches added together. The blue dashed line intersects the bunch number 0 point, and its value is displayed in the top left of each plot. The red line corresponds to a fit to bunches 1–8, with p_0 being the fit parameter. In all cases the fitted R value to bunches 1–8 is well within 1σ of the all-bunches result.

Random Seed Fit Results				
Dataset	χ^2 Mean	R Mean	R RMS	R Error on Mean
60h	0.999	-20.5562	0.3443	0.0487
HighKick	1.001	-17.4755	0.4226	0.0598
9d	0.999	-17.7182	0.2118	0.0300
Endgame	1.002	-17.3406	0.1249	0.0177

Table 5.7: Random seed fit results to the four Run 1 precession frequency analysis datasets. The χ^2 means are consistent with 1. As a reminder the 60h dataset used a different blinding than the other three, hence the significantly different R mean. Units are in ppm.

5.4.5 Fits to many random seeds

While the single seed fit results presented earlier indicate good fits and well understood parameters, it is always a good idea to fit other random seeds in case the single seed results ended up on an outlier. Doing so not only improves the confidence of the result, but also gives a more central R value to quote as being closer to the ‘true’ R of the dataset. Figures 5.37 and 5.38 give the χ^2 and R distributions for fits to 50 different random seeds for the four datasets. As shown the χ^2 distributions are nicely centered around 1 as they should be for fits to the properly distributed data. Table 5.7 compares the random seed fit results between the different datasets. The means for the R distributions of the datasets which shared the same blinding string {HighKick, 9d, Endgame} are statistically consistent with one another, with differences on the order of several hundreds of ppb, well within the statistical errors of 600 ppb or greater¹⁸.

¹⁸The errors on the mean are calculated as $\sigma_\mu = \text{RMS}/\sqrt{N}$, where N is the number of random seeds. These errors come from the randomization of cluster times before fitting, and have no bearing on statistical consistency when comparing different datasets.

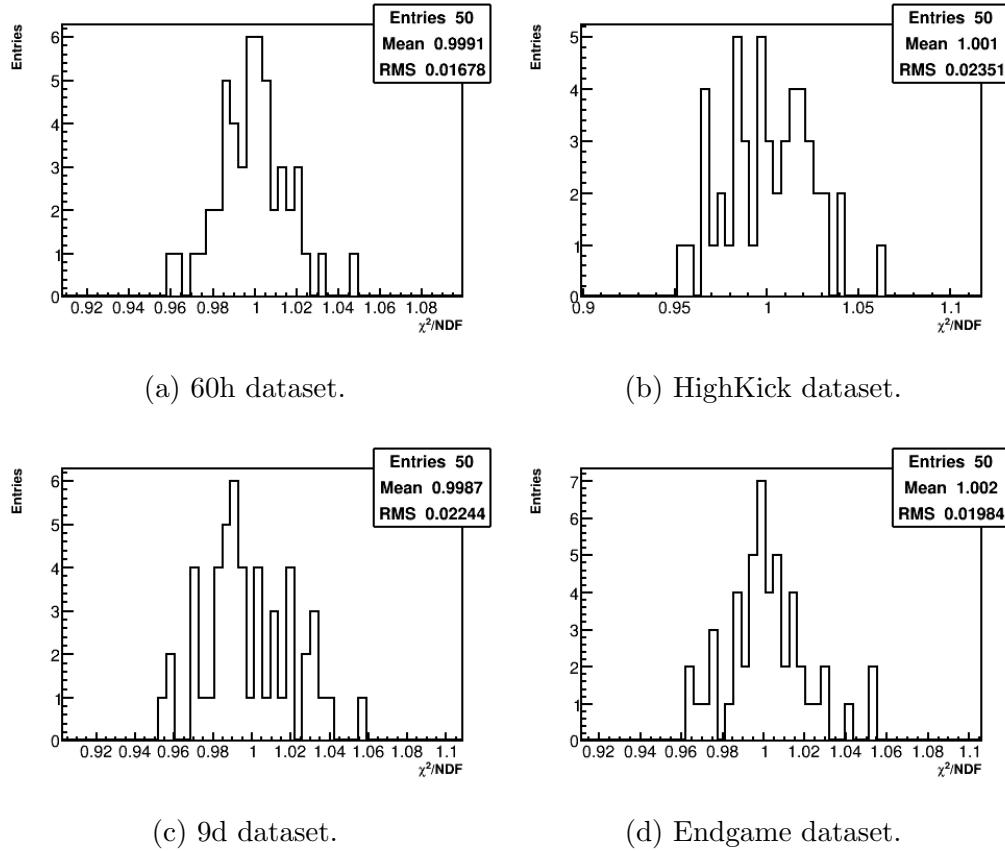


Figure 5.37: χ^2 's for fits to 50 different random seeds for the Run 1 precession frequency analysis datasets. The distributions are nicely centered around 1 which is to be expected if the randomized data is properly distributed and fit correctly.

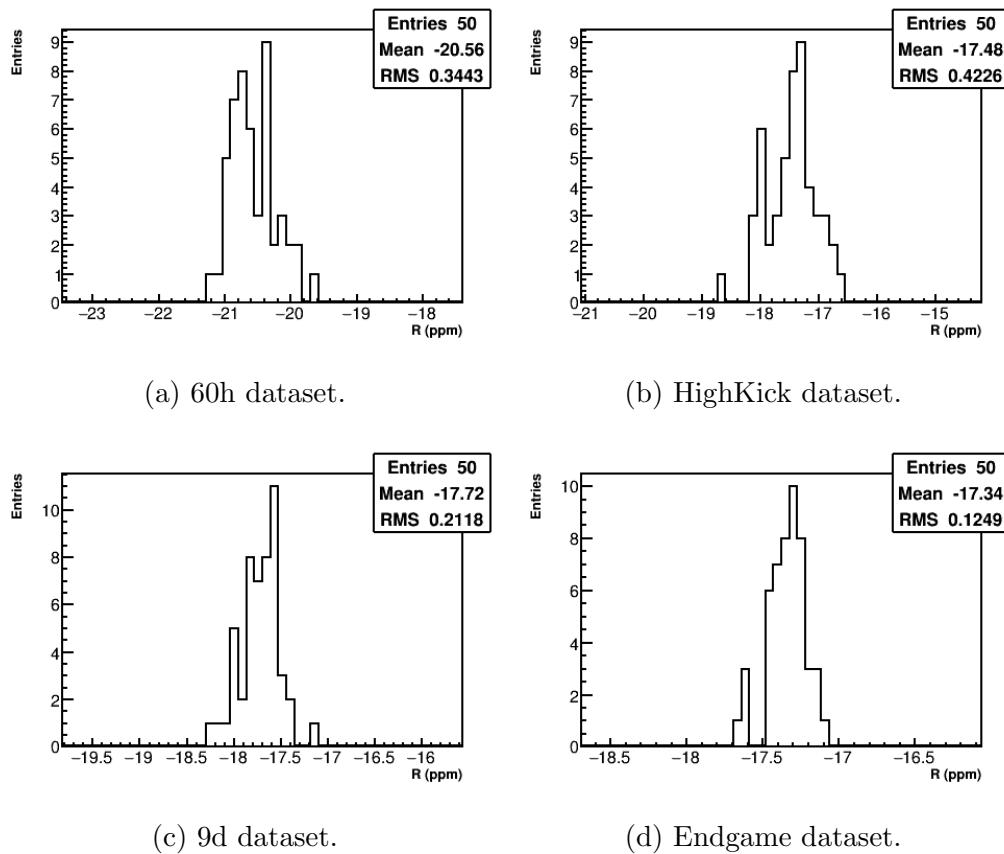


Figure 5.38: R values for fits to 50 different random seeds for the Run 1 precession frequency analysis datasets. The statistics box has units of ppm for the mean and RMS of the distributions.

5.5 Systematic errors

Evaluating the systematic errors which bias the ω_a extraction is a major part of the precession frequency analysis. Systematic errors typically stem from sources which introduce time-dependent phases over the course of a fill,

$$\cos(\omega_a t + \phi(t)) \rightarrow \cos((\omega_a + \phi_1)t + \phi_0 + \phi_2 t^2 + \dots), \quad (5.33)$$

where $\phi(t)$ can in general be expanded to higher orders. Because the phase is highly correlated to the extracted frequency, and because more decay positrons are detected at early times within a fill, a time-dependent phase pulls the extracted frequency away from its real value. Sources of systematic uncertainty can be separated into several general categories. These include systematic errors in the pileup subtraction, gain corrections, CBO fit model, muon losses, and E-field and pitch corrections, among others. In each case early-to-late shifts are a danger in the analysis and must be evaluated appropriately. Table 5.8 gives the target systematic uncertainties for the E989 experiment, which has the goal of a 70 ppb systematic error on the precession frequency measurement. This represents a three to four-fold improvement over the E821 experiment systematic errors which were of order 210 – 310 ppb depending on running period [29]. talk about reason for improvement at all here? seems kind of late to be doing so

Properly evaluated systematic errors should in general be static with respect to the amount of statistics. Run 1 of E989 consisted of more than four separate datasets with distinct running conditions, necessitating the need for independent systematic uncertainty evaluations. Such uncertainty evaluations for four of the Run 1 datasets are presented here. Systematic errors for the lost muon phase error and the the E-field and pitch corrections are dependent on analyses from other working groups. The numbers given here for them are preliminary estimations.

Target ω_a Measurement Uncertainties	
Sources of uncertainty	E989 Goal (ppb)
Pileup	40
Gain changes	20
CBO	30
Lost muons	20
E-field and pitch corrections	30
Other	30
Quadrature sum	~ 70

Table 5.8: Target systematic errors in the precession frequency measurement for the E989 experiment.

5.5.1 Pileup systematic errors

As described in Section 5.2.1, the pileup background oscillates at ω_a which by extension means a strong effect on the final fitted R value. If the subtracted pileup spectrum is mis-constructed in any way, there will be a systematic error on R . In general the pileup systematic error can be separated into two parts, the error on the amplitude and the error on the phase. In order to estimate the two parts, the uncertainties on the pileup amplitude and phase need to be estimated along with the sensitivities of R to them. Table 5.1 gives the default values used for the pileup construction parameters {ADT, SDT, SGT, C}. How these parameters feed into the amplitude and phase systematic errors will be discussed in turn, and the overall errors calculated for the different Run 1 datasets.

As a reminder the default values used for the ADT and SDT were 5 ns each, and a default automatic pileup amplitude multiplier of ~ 1.03 was applied to the pileup spectra. In order to calculate the systematic dependence on the choice of ADT or SDT, the SDT parameter was scanned over from 5 ns to 10 ns in steps of 1 ns. This was done with and without the same automatic pileup amplitude scaling procedure as described in Section 5.2.1. The results of the study for the 9d dataset are shown

in Figures 5·39 and 5·40. In the case where there was no automatic scaling applied there is a clear minimum in the χ^2 results and a steep slope in R corresponding to a large sensitivity of R to the choice of SDT. In the case where the automatic scaling was applied however, the minimum in the χ^2 results has disappeared, while the sensitivity of R has become much reduced to the point of no longer being a clear trend¹⁹. The fact that applying the automatic pileup amplitude scaling produces nearly identical pileup spectra with no clear trend in R regardless of the choice of SDT (and by extension ADT), any systematic error due to the choice of these two parameters can be subsumed into the direct pileup amplitude error itself, discussed down below. It should be noted in fact that the choices of ADT and SDT are largely irrelevant barring statistics, as the automatic amplitude scaling procedure can always account for any differences between the two.

In order to calculate the systematic dependence on the choice of SGT (default value of 10 ns), the SGT parameter was scanned over from 10 ns to 20 ns in steps of 1 ns. The results of the study with the automatic pileup amplitude scaling applied is shown in Figure 5·41. Just as in the SDT scan with the automatic pileup amplitude scaling, there is no minimum in the χ^2 results, the sensitivity of R to the value of SGT not so clear, and the pileup spectra for the various choices of SGT are nearly identical. Therefore again any systematic error due to the choice of SGT is subsumed into the pileup amplitude error.

The pileup amplitude systematic error is the error on R assuming the scale of the pileup was incorrectly constructed. In order to evaluate this error, multipliers were applied to the pileup spectra from 0.9 to 1.1 in steps of 0.01 (dropping the default automatic pileup scaling of ~ 1.03 mentioned before). The data was then re-fit to find the change in R . The results of the study for the 9d dataset are shown in Figure 5·42.

¹⁹This slope in R varies between positive and negative values based on dataset, so there is no real clear trend in R .

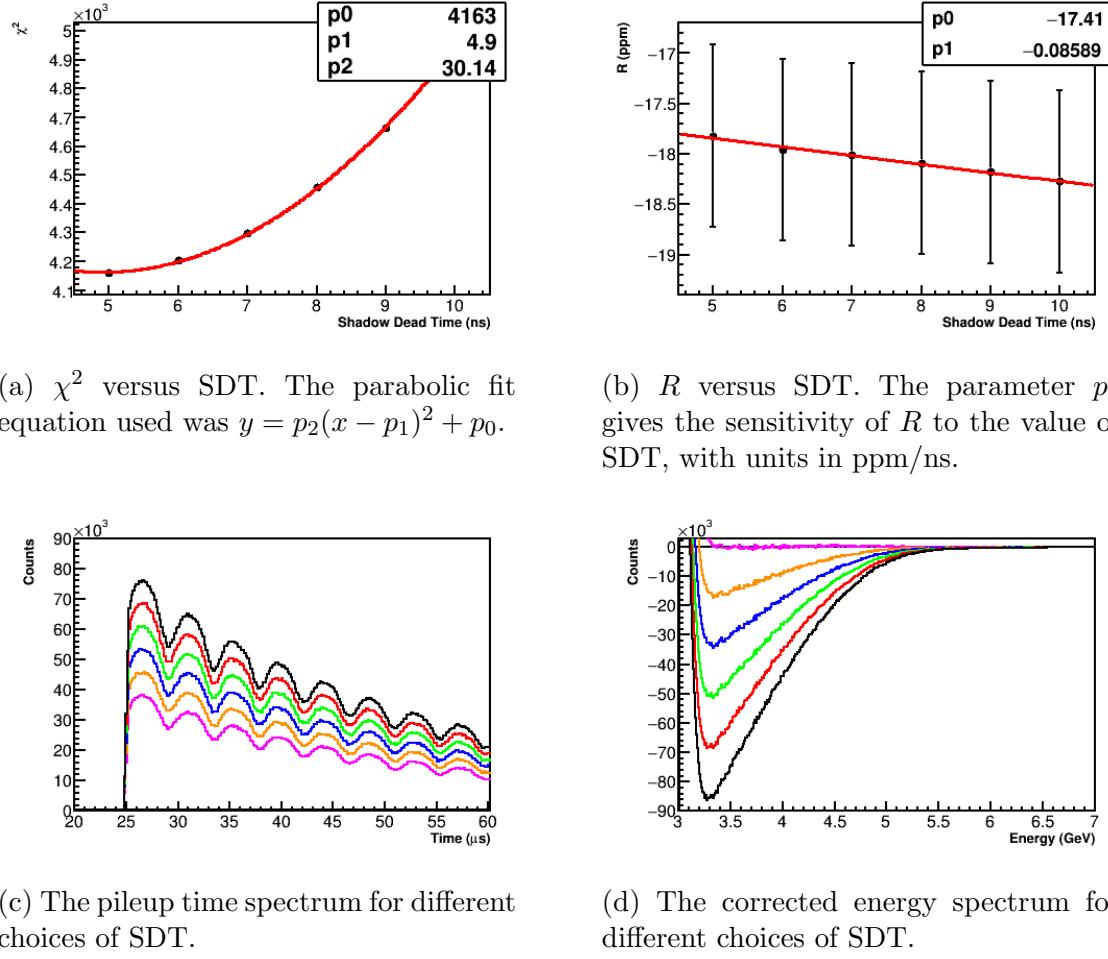


Figure 5.39: Shadow dead time scan results without automatic pileup amplitude scaling. A clear minimum in the χ^2 plot is seen near 5 ns corresponding to the choice of ADT, and a large sensitivity for R is observed. In the bottom two spectra plots the magenta curve corresponds to a choice $SDT = 5$ ns while the black curve corresponds to $SDT = 10$ ns. The larger choice of SDT leads to a greater estimation of the pileup, which as shown in the energy spectra plot leads to a corresponding over-subtraction at energies where hits consist mostly or purely of pileup pulses. Data from 9d dataset.

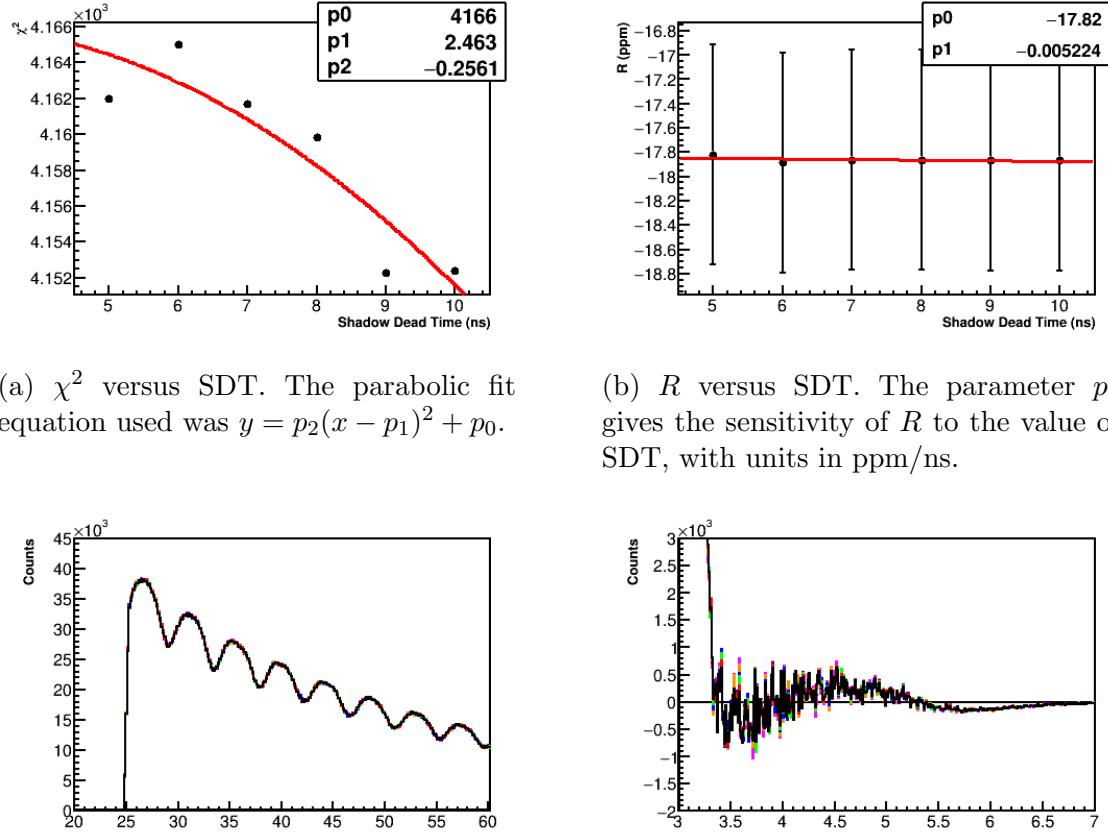


Figure 5.40: Shadow dead time scan results with automatic pileup amplitude scaling. No clear minimum is observed in the χ^2 plot, and the sensitivity for R is small. In the bottom two spectra plots the magenta curve corresponds to a choice $SDT = 5\text{ ns}$ while the black curve corresponds to $SDT = 10\text{ ns}$. With the automatic amplitude scaling applied, the time and energy spectra are nearly identical and lie on top of each other. That combined with the lack of clear minimum in the χ^2 plot and no clear sensitivity in R indicate that there is no real systematic error due to the choice of SDT. Data from 9d dataset.

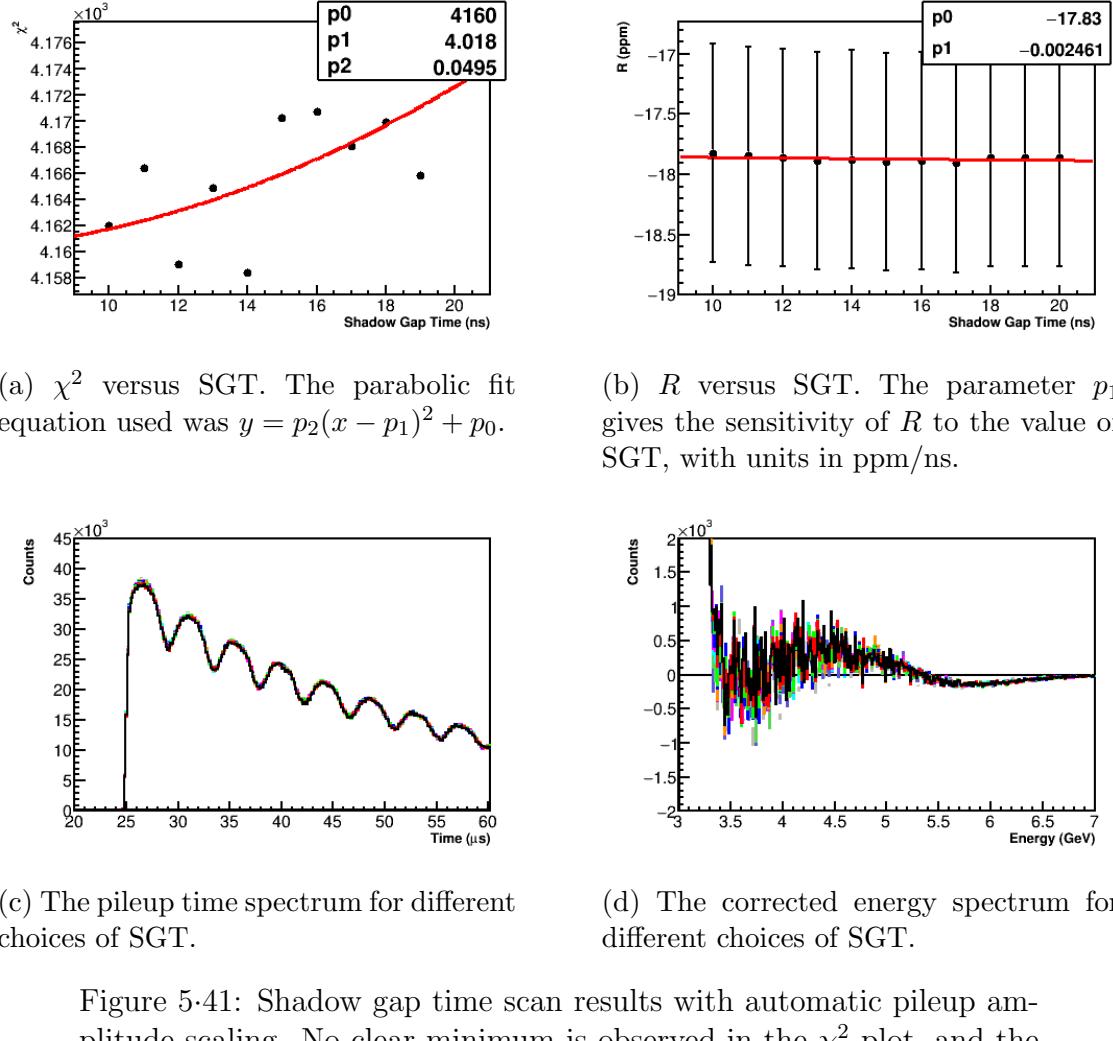


Figure 5.41: Shadow gap time scan results with automatic pileup amplitude scaling. No clear minimum is observed in the χ^2 plot, and the trend for R isn't clear, with points fluctuating above and below the fit curve. In the bottom two spectra plots one of the grey curves (hidden) corresponds to a choice $\text{SGT} = 10 \text{ ns}$ while the black curve corresponds to $\text{SGT} = 20 \text{ ns}$. With the automatic amplitude scaling applied, the time and energy spectra lie on top of each other. That combined with the lack of clear minimum in the χ^2 plot and small sensitivity in R indicate that there is no real systematic error due to the choice of SGT. Data from 9d dataset.

As shown there is a clear minimum near 1 in the χ^2 results and a large sensitivity of R to the multiplier. The systematic error on R is calculated as

$$\delta R = \sigma_{P_m} \times \frac{dR}{dP_m}, \quad (5.34)$$

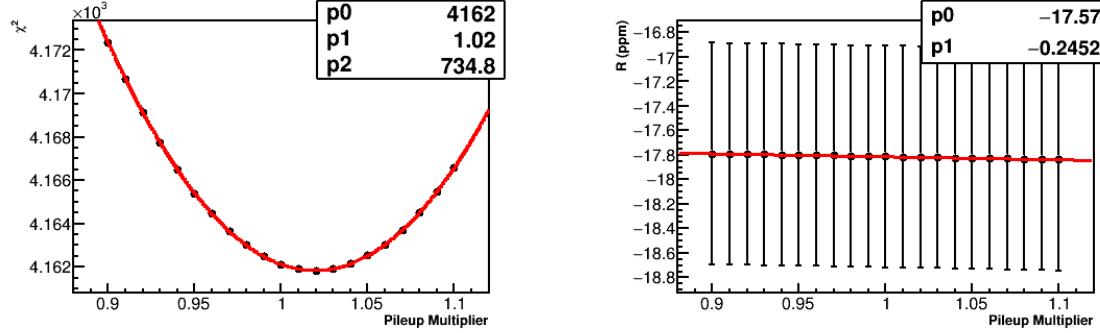
where P_m is the value of the pileup multiplier. The error σ_{P_m} is calculated as the width of the fitted parabola in the χ^2 plot, defined as the change in P_m from the minimum for the χ^2 to increase by 1. This is calculated as

$$\sigma_{P_m} = \sqrt{\frac{2}{f''(\chi^2)}} = \frac{1}{\sqrt{p_2}}, \quad (5.35)$$

where p_2 is the fit parameter as given in the top right of the χ^2 plot. The sensitivities of R to the pileup multiplier, uncertainties in the pileup amplitude, and final corresponding systematic errors for the Run 1 precession frequency analysis datasets are given in Table 5.9. As shown in the table, the uncertainties on the pileup amplitude are of order 2 to 5%, while the systematic errors on R are on the order of 10 to 20 ppb depending on dataset. It should be noted that the default automatic pileup multiplier of ~ 1.03 does not necessarily correspond to the minimum in the χ^2 plot, but is within 1σ of 1 or the minimum (except the Endgame which is closer to 2σ)²⁰

The pileup phase error is the error on R assuming the phase of the pileup was incorrectly constructed. This is separated into two parts. The first part is calculated by applying time-shifts to t_{doublet} as given in Equation 5.3. Doing this artificially applies a phase shift to the pileup time spectrum. The data is then re-fit with the different pileup spectra and the change in R is calculated. Figure 5.43 shows the study results for the 9d dataset with time-shifts applied between -10 ns and 10 ns in steps of 1 ns. A clear sensitivity of R to the value of time-shift is observed, however

²⁰Monte-Carlo tests with various random seeds showed this minimum fluctuating above and below 1. The distance from 1 therefore is not a good measure for the uncertainty in the pileup amplitude compared to the width of the χ^2 parabola fit.



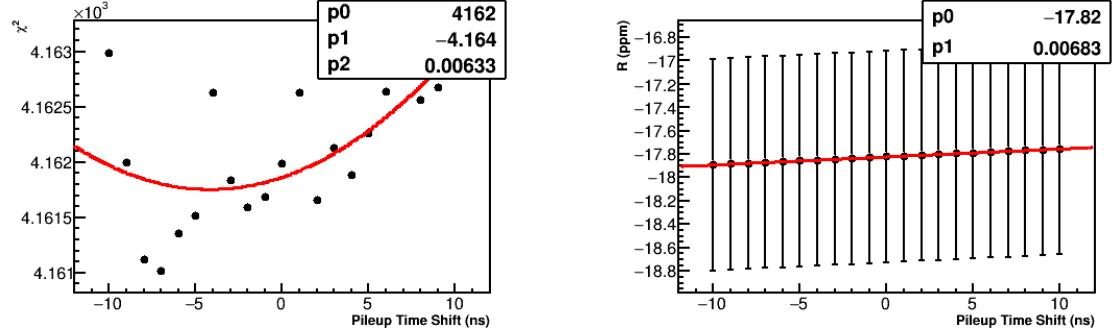
(a) χ^2 versus pileup multiplier. The parabolic fit equation used was $y = p_2(x - p_1)^2 + p_0$.

(b) R versus pileup multiplier. The parameter p_1 gives the sensitivity of R to the value of the pileup multiplier, with units in ppm.

Figure 5.42: Pileup multiplier scan. Data from 9d dataset.

Systematic Error due to Pileup Amplitude				
Dataset	dR/dP_m	σ_{P_m}	$P_{m\min}$	δR
60h	-419.3	0.053	0.993	22.2
HighKick	-372.8	0.051	0.997	19.0
9d	-245.2	0.037	1.020	9.0
Endgame	-335.3	0.028	0.985	9.4

Table 5.9: Systematic error due to the pileup amplitude in the Ratio Method fits for the Run 1 precession frequency analysis datasets. The bold column gives the systematic error on R . Units for dR/dP_m and δR are in ppb.



(a) χ^2 versus pileup time-shift. There is no clear minimum in the plot.

(b) R versus pileup time-shift. The parameter p_1 gives the sensitivity of R to the value of the pile time-shift, with units in ppm/ns.

Figure 5.43: Scan over pileup time-shift. Data from 9d dataset.

there is no clear minimum in the χ^2 results. Because the width of the χ^2 parabolic fit cannot be taken as the uncertainty in the pileup time-shift parameter , instead the uncertainty is taken conservatively at half the ADT at 2.5 ns. The systematic error is calculated in the same way as for the pileup amplitude uncertainty,

$$\delta R = \sigma_{P_t} \times \frac{dR}{dP_t}, \quad (5.36)$$

where P_t is the value of the pileup time-shift. The sensitivities of R to the pileup time-shift and corresponding systematic errors for the Run 1 precession frequency analysis datasets are given in Table 5.10.

The second part of the pileup phase error comes from the choice of constant C in the calculation of E_{doublet} as given in Equation 5.2. If the energy of the pileup pulses are systematically mis-constructed, then pileup shadow doublets will be added or lost near the applied energy threshold when constructing the pileup spectrum. This leads to an error on the pileup phase since it is energy-dependent. In order to calculate the systematic error from the energy construction, the parameter C was scanned over

Systematic Error due to Pileup Time Shift		
Dataset	dR/dP_t	δR
60h	7.0	17.6
HighKick	7.6	19.0
9d	6.8	17.1
Endgame	5.7	14.3

Table 5.10: Systematic error due to the pileup time-shift parameter P_t in the Ratio Method fits for the Run 1 precession frequency analysis datasets. The bold column gives the systematic error on R . Units for dR/dP_t and δR are in ppb/ns and ppb respectively. The error on the P_t is by default taken to be 2.5 ns as described in the text. **fix spacing of table**

from 0.9 to 1.1, in steps of 0.01. The results of the study for the 9d dataset are shown in Figure 5.44. The systematic error on R is calculated as

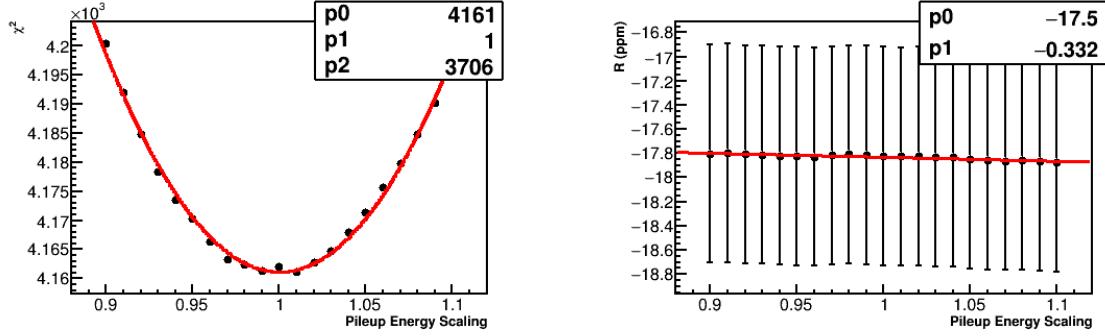
$$\delta R = \sigma_C \times \frac{dR}{dC}. \quad (5.37)$$

Similarly to the pileup amplitude error, there is a clear minimum in the χ^2 results which can be used to estimate the uncertainty in the pileup energy scale²¹. Table 5.11 gives the sensitivities of R to the pileup energy scale, uncertainties in the pileup energy scale, and the corresponding final systematic errors for the Run 1 precession frequency analysis datasets. As shown the uncertainties on the pileup energy scale are of order 1 to 2%, and the value for C which produces the minimum in the χ^2 results is consistent with 1²². Interestingly, the sensitivity of R to C in the 60h dataset is noticeably larger than in the rest of the datasets, though the origin of this is currently unknown. **what to do about this?**

Table 5.12 gives the quadrature sum for the total pileup systematic errors for the Run 1 precession frequency analysis datasets. As shown for each dataset the total

²¹The trend isn't as clean as in the scan over the pileup amplitude multiplier, but that is acceptable.

²²Because the spatial separation is turned off in the clustering portion of the reconstruction, a value of $C = 1$ is to be expected.



(a) χ^2 versus pileup energy scale. The parabolic fit equation used was $y = p_2(x - p_1)^2 + p_0$.

(b) R versus C . The parameter p_1 gives the sensitivity of R to the value of C , with units in ppm.

Figure 5.44: Scan over pileup energy scale. Data from 9d dataset.

Systematic Error due to Pileup Energy Scale				
Dataset	dR/dC	σ_C	C_{\min}	δR
60h	-835.1	0.023	0.997	19.4
HighKick	-167.7	0.022	0.995	3.7
9d	-332.0	0.016	1.000	5.5
Endgame	-431.4	0.012	0.982	5.3

Table 5.11: Systematic error due to the fixed pileup energy scale parameter C in the Ratio Method fits for the Run 1 precession frequency analysis. The bold column gives the systematic error on R . Units for dR/dC and δR are in ppb.

Total Pileup Systematic Errors					
Type of Error	Parameter	60h	HighKick	9d	Endgame
Amplitude	P_m	22.2	19.0	9.0	9.4
Phase	P_t	17.6	19.0	17.1	14.3
Phase	C	19.4	3.7	5.5	5.3
Quadrature sum		34.3	27.1	20.1	17.9

Table 5.12: Total pileup systematic errors for the Run 1 precession frequency analysis datasets.

error is below the target final error of 40 ppb in spite of the contamination in the pileup shadow method. For future runs of the experiment with increased rate and therefore increased pileup, these errors may grow. In that case either the pileup shadow method might need to be improved to account for the contamination and pileup triplets, or discarded in favor of a different method.

5.5.2 Gain systematic errors

As described in Section 5.1, the energies of the positron hits are gain-corrected for in-fill, short-term double pulse, and out-of-fill effects. The latter occurs over time scales much longer than a fill. This does not bias the precession frequency measurement as the phase is not time-dependent over the course of a fill. For the cases of in-fill or STDP gain variations, any uncorrected fluctuations in the gain causes acceptance changes over the course of a fill which then modifies the average measured phase of the detected positrons above threshold, and thus causes a systematic shift in the extracted R value.

The IFG function as measured by the laser calibration system, Section 3.2.3, describes the measured energies of the individual crystal hits as a function of time in-fill. It is given by [58]

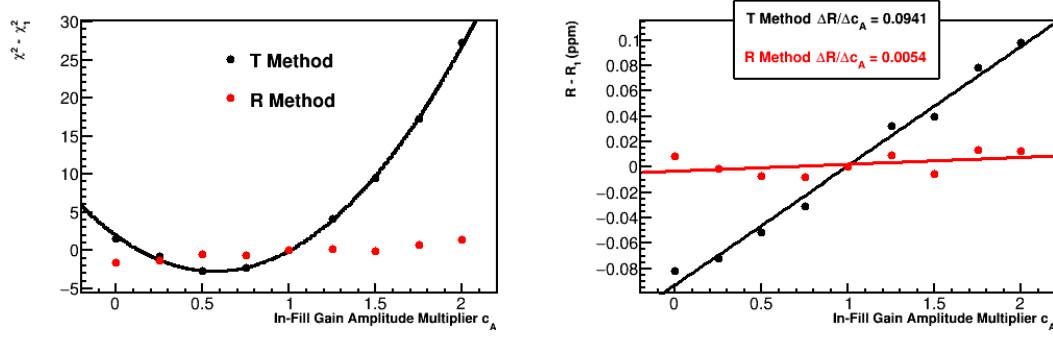
$$E = E_0(1 - C_A e^{-t/\tau_g}), \quad (5.38)$$

where E_0 is the ‘true’ energy of the detected positron and E is the as-measured energy. Constants C_A and τ_g are measured with the laser calibration system, and then the function given in Equation 5.38 is used to convert the measured energies back to their real energies. In order to determine a systematic error from the applied IFG function, the IFG function was un-applied and then re-applied to the crystal hits with modified parameters. Specifically, the amplitude of the IFG function C_A was scanned over in multiplicative steps, such that all crystal energies included in a cluster hit were adjusted by a multiplicative factor before re-summing²³. The multipliers applied to the IFG amplitude were from 0 to 2 in steps of 0.25. The results of the scan for the 60h dataset is shown in Figure 5.45, with a comparison to the results from fits with the T-Method. It is immediately apparent that the sensitivity of the T-Method to the IFG multiplier is significantly greater than that of the Ratio Method. The Ratio Method’s insensitivity to slowly varying effects which it divide out is one of its primary strengths. In fact, while there is an observable minimum in the T-Method χ^2 results, no such minimum exists for the Ratio Method results.

Table 5.13 gives the results for the IFG amplitude scans for the different datasets. Sensitivities for both the T-Method and Ratio Method fits along with the T-Method minimum for C_A are included. The uncertainties on the IFG amplitudes from fits to the laser data were in general of order 25% [103]. This uncertainty multiplied against the as-measured Ratio Method sensitivities results in systematic errors of <6 ppb for the various datasets. The individual numbers are given in column 5 of Table 5.13.

Examining the table in more detail however, some interesting features can be seen. The T-Method results, while not the primary subject of this analysis, nevertheless vary between the datasets, in that the scan minima are not consistent with one another. For the Ratio Method results, the HighKick and Endgame datasets exhibit

²³It was decided that scanning over the parameter τ_g was unnecessary as it in general shifts the energies of the positrons in the same way as C_A .



(a) Normalized χ^2 vs IFG multiplier. The T-Method points fit nicely to a parabolic curve while the Ratio Method points do not.

(b) R values vs IFG multiplier. Both points are fit to straight lines, and the slopes are included in the text box in units of ppm.

Figure 5.45: Fitted χ^2 's and R values for the 60h dataset as a function of IFG amplitude multiplier. Results with ratio fits (red) are compared to T-Method fits (black). Values are normalized to their $C_A = 1$ results in order to put the curves on the same scale. As shown the sensitivity in the T-Method to the IFG amplitude is significantly larger than that of the Ratio Method. This is one of the primary strengths of the Ratio Method.

greater sensitivities to the IFG amplitude than the 60h and 9d datasets. The HighKick dataset can be seen to have a negative sensitivity compared to the positive sensitivities of the other datasets, and the T-Method sensitivities for all datasets. The Endgame dataset has a greater sensitivity stemming from the fact that fits to the multiplicative factors of $C_A = 1.75$ and $C_A = 2$ produced noticeably better χ^2 's and higher R values, $\mathcal{O}(40 \text{ ppb})$. These fits pulled the slope up from what is otherwise a very flat fit to the rest of the points, and imply that the default IFG multiplier should be greater than 1.

These various peculiarities imply there are imperfections in the gain corrections which have not been fully accounted for. Indeed recent investigations into the applied gain corrections show that for the HighKick and Endgame datasets the out-of-fill corrections were neglected [104]. While this does not introduce a systematic uncertainty as a function of time in-fill as described before, it does shift the scale of the correction and by extension the extracted R value. In order to determine more conservative bounds on the systematic errors in order to account for these deficiencies, the absolute changes in R for the 0x and 2x multiplier fits were examined. These values are included in Table 5.13. The maximum of the two values was then taken as the systematic error on R due to the IFG amplitude. In general the systematic errors as calculated from the maximum of the 0x or 2x fits are an order of magnitude larger than those errors calculated directly from the sensitivities. Because the scale of these numbers is still small compared to the Run 1 statistical errors however, they have been deemed acceptable for this analysis. The confidence in the scale of the Ratio Method results is ultimately preserved from the fact that the Ratio Method fits are not as sensitive to the gain variations as the T-Method fits, as the gain effects divide out. Going forward towards the end of the Run 1 analysis, once the outstanding gain issues have been resolved the systematic errors will be calculated directly from the

sensitivities.

Systematic Error due to IFG Amplitude							
Dataset	T-Method dR/dC_A	T-Method $C_{A_{\min}}$	R-Method dR/dC_A	R-Method $\delta R_{\sigma=0.25}$	R-Method ΔR_{0x}	R-Method ΔR_{2x}	R-Method $\delta R_{\max - 0x/2x}$
60h	94.1	0.57	5.4	1.4	8.2	12.4	—
HighKick	69.5	0.70	-23.6	5.9	28.6	-16.0	28.6
9d	64.9	0.27	1.4	0.4	-2.5	-3.9	—
Endgame	96.2	0.19	22.7	5.7	-6.9	43.6	43.6

Table 5.13: Sensitivities and systematic errors for the IFG amplitude. T-Method sensitivities are included for comparison, along with the χ^2 minima. Also included are changes in R for fits with IFG amplitude multipliers of 0x and 2x. Systematic error columns are in bold, where the one on the left corresponds to the Ratio Method sensitivity multiplied by a 25% uncertainty in the amplitude, and the one on the right corresponds to the absolute value of the maximum change in R with the 0x and 2x multipliers applied. Only the HighKick and Endgame values are used from the column on the far right. Units for errors and sensitivities are in ppb.

The STDP correction is in many ways the same as the IFG correction [57]. This correction is applied to pulses close in time, $\mathcal{O}(\text{ns})$, and is applied to the energies of the pulses before the in-fill gain correction. For the systematic error on R due to the application of the STDP gain correction, a new production of the 60h dataset was processed without the STDP correction applied. Because of how the time-randomization is applied to the clusters in the author's analysis, a new version of the same dataset by default uses a different default randomization to the cluster times and the ratio histogram filling. As shown in Section 5.4.5, the width in the fitted R values for many random seeds is of order $\mathcal{O}(100 \text{ ppb})$ making direct ratio fit comparisons between the two datasets un-informative. In order to avoid this difficulty, fits were done with the T-Method, and cluster times were randomized per-fill rather than per-cluster. Since the per-fill randomization uses fill ID's as the seeds for the randomization, and because the T-Method does not split the data into sub-datasets as the Ratio Method does, the randomization between the 60h dataset with and without the STDP correction is identical. The change in R for T-Method fits with and without this correction is then taken as the upper bound on the systematic error for the inclusion of the STDP in the Ratio Method fits. This is reasonable as the previous IFG systematic studies showed a reduction in sensitivity to gain effects which is reasonably extended to the STDP correction error. Table 5.14 gives the T-Method fit results with and without the STDP, along with the change in R . This difference was found to be 11.0 ppb, which is then taken as the systematic error for all datasets.

Table 5.15 gives the quadrature sum for the total gain systematic errors for the Run 1 precession frequency analysis datasets. As shown for the 60h and 9d datasets the total error is below the target final error of 20 ppb. For the HighKick and Endgame datasets, the errors are larger than 20 ppb but are nevertheless at the same scale, and should reduce once their known issues have been resolved. It should be noted that

Systematic Error due to STDP			
Fit Type	R with STDP (ppm)	R without STDP (ppm)	δR (ppb)
T-Method	-20.1619	-20.1729	11.0

Table 5.14: T-Method fit results with and without the STDP gain correction on the 60h dataset. T-Method fits were done instead of Ratio Method fits in order to force the cluster-time randomization to be consistent between the two dataset productions. The change in R in the bold column is taken as the upper bound on the systematic error in the Ratio Method due to the STDP gain correction.

Total Gain Systematic Errors				
Type of Error	60h	HighKick	9d	Endgame
IFG	1.4	28.6	0.4	43.6
STDP	11.0	11.0	11.0	11.0
Quadrature sum	11.1	30.6	11.0	45.0

Table 5.15: Total gain-related systematic errors for the Run 1 precession frequency analysis datasets.

there is further evidence of the need for a correction for some type of residual gain effect which is somehow unmeasurable by the laser system [87, 105]²⁴. This combined with the previously mentioned issues in the Endgame and HighKick datasets warrant re-evaluations of the gain systematic errors once the final datasets are available. It is expected that the errors given here are nevertheless reasonable approximations of what the final errors will end up, especially due to the Ratio Method's insensitivity to the gain relative to the T-Method. As a reminder, the errors given here even with the crude approximations are small compared to the statistical errors of the various datasets.

²⁴This implies the effect is very small.

Change in R with station 18 CBO parameters	
Dataset	δR
60h	7.5
HighKick	0.4
9d	-2.0
Endgame	8.0

Table 5.16: Changes in the fitted R values with tracker station 18 CBO frequency model parameters instead of tracker station 12. The systematic errors are conservatively taken as the absolute value in the changes in R . Units are in ppb.

5.5.3 CBO systematic errors

If the CBO is mis-modeled then there will be a systematic error on R since there is a an early-to-late change in both the frequency and the scale of the CBO. The CBO model is largely constrained by tracker measurements, but systematic errors can be evaluated by modifying the fixed frequency function given in Equation 5.22, and the decoherence envelope of the CBO.

Table 5.2 gives the CBO frequency model parameters for both tracker stations. The station 12 values are by default used in all fits to the data. Fits were performed with the station 18 values, and the changes in R are given in Table 5.16. The absolute values of the changes for the different datasets are conservatively taken as the systematic error on R due to the choice of fixed CBO frequency model parameters. While the CBO parameters in the tracking analysis fits do have errors on the parameters, they are tiny compared to the systematic errors between the two tracker stations [96]. Some few fits were made by varying the fixed frequency parameters by 1σ in their individual respective errors, and the changes in R were found to be negligible. For this reason the systematic errors from using station 18 values are taken as the errors.

The shape of the CBO, or the decoherence envelope, is also similarly constrained by the tracking analysis. The envelope by default is taken to be an exponential

as given in Equation 5.19 and shown in Figure ???. The other envelope which could reasonably exist in the data upon inspection in the tracking analysis is an exponential plus a constant

$$e^{-t/\tau_{cbo}} \rightarrow e^{-t/\tau_{cbo}} + C, \quad (5.39)$$

where C is some constant CBO amplitude which persists over the course of the fill. In order to assess this systematic error, this new envelope was introduced into the $N_{cbo}(t)$ fit term. Fits were done with C floating, where the starting values for C in the Ratio Method fits were taken from T-Method fits to the data²⁵. While in T-Method fits the C parameter converged to values with errors about half the value, in general in the Ratio Method fits the C parameters had relatively large errors. In spite of the large errors however, the fits converged properly with the floating C parameter. In general the final fit parameters are largely the same, with the exception being the fitted CBO lifetime which about halves. This is unsurprising as the lifetime is highly correlated to the amplitude. Only in the 9d dataset did some small complications arise, where the CBO lifetime had to be fixed in order to get the fit to converge properly. In spite of this constraint it was found that the C parameter converged to a negative value in the 9d dataset with a larger change in R . A reasonable hypothesis for this is that the correlations in the fitted parameters allows for good fits with varying fit values, and this fit just so happened to end up on a negative value. Regardless, though peculiar, the fit preference was clear and converged properly so result was taken to be acceptable.

The fitted values for the C constants and the changes in the final fitted R values are given in Table 5.17. As shown the changes in R are of order 10s of ppb for some of the datasets, with R varying both positively and negatively. These changes in R

²⁵T-Method starting values for C were taken to be 0 which resulted in well-converging fits.

Systematic Error due to CBO Envelope			
Dataset	$C \times 10^{-4}$	$\sigma_C \times 10^{-4}$	δR
60h	10.7	8.3	17.6
HighKick	11.6	10.4	-18.0
9d	-13.5	12.2	28.7
Endgame	10.7	4.0	-4.3

Table 5.17: Systematic error on R due to the choice of CBO envelope. The fitted floating parameter C and it's error are given along with the change in R compared to the standard exponential envelope in units of ppb.

Total CBO Systematic Errors				
Type of Error	60h	HighKick	9d	Endgame
CBO Frequency Model	7.5	0.4	2.0	8.0
CBO Decoherence Envelope	17.6	18.0	28.7	4.3
Quadrature sum	19.1	18.0	28.8	9.1

Table 5.18: Total CBO-related systematic errors for the Run 1 precession frequency analysis datasets.

are conservatively taken as the systematic errors on R for the different datasets.

Table 5.18 gives the quadrature sum for the total CBO systematic errors for the Run 1 precession frequency analysis datasets. As shown for each dataset the total error is below the target final error of 30 ppb. For future runs of E989 with more statistics it will be a question of whether the scale of these errors remains consistent or not. The changing frequency of the CBO in Run 1 was eliminated in Run 2, implying a reduction in any associated systematic errors. The increased statistics combined with the tracking analysis should constrain measurements of the CBO decoherence envelope even further, similarly reducing any associated systematic errors. On the other side of the coin, increased statistics might bring out higher order CBO effects which require improved modeling.

5.5.4 Lost muon systematic errors

Systematic errors from lost muons arise due to the fact that muons are preferentially lost at early times. Uncertainties in the lost muon spectrum itself, the fixed κ_{loss} parameter in the Ratio Method fits, and any error due to phase differences between lost and stored muons are evaluated here.

As mentioned in Section 5.3.3, the triples spectrum is made with cuts as defined in Table 5.4. Various backgrounds are subtracted off the triples spectrum in order to generate a clean sample of lost muons. Table 5.19 gives the change in R for various sets of cuts and background subtractions for the 9d and Endgame datasets. As shown the various different backgrounds and cuts ultimately make very little difference in the final fitted R value. Similarly, stable beam contaminants in the form of deuterons and protons contaminate the lost muon spectrum. The former are largely removed by straightforward Δt_{12} and Δt_{13} cuts. The latter can be mostly removed by cutting on the negative side of the Δt_{13} distribution with $\Delta t_{13} \leq 12.5$ ns which separates the populations more readily, Figure 5.46. While this does largely remove the protons at the cost of statistics, the fitted κ_{loss} parameter simply grows larger to compensate. Ultimately the effect on R is negligible, with $\Delta R = -0.3$ ppb. The sum of these separate types of errors is conservatively taken at 0.5 ppb for all datasets²⁶.

Because κ_{loss} is fixed in the Ratio Method fits from corresponding T-Method fits, the statistical correlation to R is neglected and a systematic error must be estimated. The systematic error can be determined by scanning over the value of κ_{loss} as in Figure 5.47. The error on κ_{loss} is taken as the statistical error determined in the T-Method fit to the same data. Table 5.20 gives the systematic errors on R for the Run 1 precession frequency analysis datasets. As shown the systematic errors as calculated from the sensitivities are all small, $\delta R < 5$ ppb. However, as mentioned

²⁶For the T-Method fits, though the changes in R are noticeably larger with the various cuts, they are still the same order of magnitude and the error is conservatively below 1 ppb.

ΔR with Various Lost Muon Cuts				
Type of fit or cut	9d Dataset		Endgame Dataset	
	ΔR (ppb)	κ_{loss}	ΔR (ppb)	κ_{loss}
Default cuts	0	2.510	0	2.345
No quadruple subtraction	0.2	1.811	-0.1	1.717
No accidental subtraction	0.1	2.503	< 0.1	2.339
$\Delta t_{13} \leq 12.5$ ns	0.1	4.469	-0.3	4.248
5 ns $\leq \Delta t_{12,23} \leq 8.5$ ns	0.1	2.507	< 0.1	1.709
100 MeV $\leq E_{1,2,3} \leq 500$ MeV	< 0.1	2.253	< 0.1	2.090

Table 5.19: Effect on the fitted R value in the 9d and Endgame datasets with various cuts used or backgrounds subtracted. Also included are the various corresponding κ_{loss} values which compensate for the level of statistics contained within $L(t)$ due to the differing cuts. Ultimately how the muon loss spectrum $L(t)$ is created has little bearing on the final fitted R value.

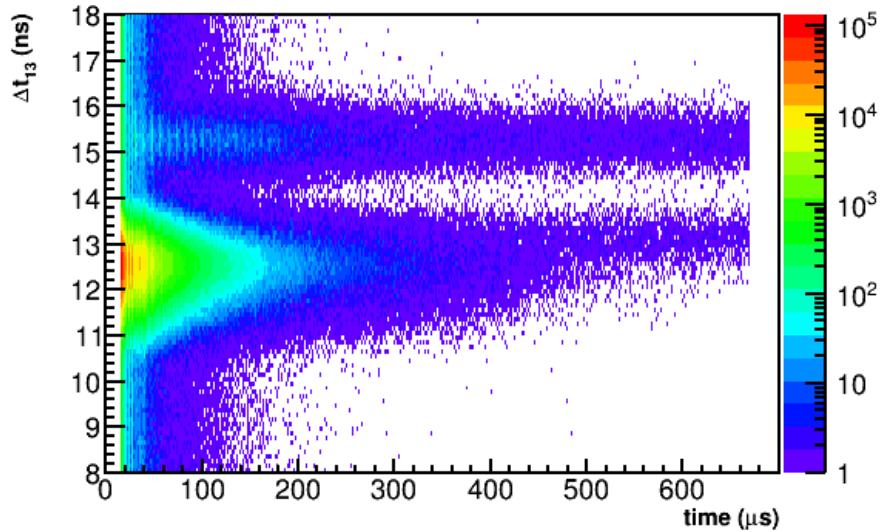


Figure 5.46: Lost muon Δt_{13} distribution as a function of time in-fill. The band of hits centered above 15 ns corresponds to the deuteron contamination, while the band of hits centered above 13 ns corresponds to the proton contamination.

Systematic Error due to Fixed κ_{loss}					
Dataset	$dR/d\kappa_{loss}$	$\sigma_{\kappa_{loss}}$	δR	$\Delta R_{(w/ - w/o)}$	δR_{2x-IFG}
60h	-3.5	0.338	1.2	-31.4	-
HighKick	-7.1	0.697	4.9	-40.1	39.1
9d	-18.3	0.170	3.1	-45.7	-
Endgame	-2.6	0.038	0.1	-6.1	28.9

Table 5.20: Systematic error due to the fixed κ_{loss} parameter in the Ratio Method fits for the Run 1 precession frequency analysis datasets. All units are in ppb except for the $\sigma_{\kappa_{loss}}$ parameter which is unit-less. $\sigma_{\kappa_{loss}}$ comes from the T-Method fit results and scales with the number of statistics. The bold columns gives the systematic errors on R , where the one on the left corresponds to the systematic error as calculated from the sensitivity, and the one on the right corresponds to the systematic error for the HighKick and Endgame datasets using the κ_{loss} value determined from fits with 2x the IFG multiplier.

in Section 5.5.2, there were recently discovered issues in the applied gain corrections for the HighKick and Endgame datasets. Because the κ_{loss} parameter is correlated to the gain, this implies the fitted κ_{loss} values are wrong by a some amount for those datasets. In order to determine conservative errors due to the fixed κ_{loss} parameter in those two cases, the κ_{loss} values for the fits with 2x the IFG multiplier were used and compared to the default fits with 1x the IFG multiplier. The κ_{loss} values for the HighKick and Endgame datasets with the 2x multiplier were 9.510 and 2.766 respectively. Compared to the default fit results of 5.651 and 2.345, these κ_{loss} values are 11.1σ and 5.5σ away in terms of the T-Method statistical error. Multiplying the Ratio Method sensitivities by these larger errors gives the systematic errors in the far right column of Table 5.20. These alternative errors are an order of magnitude or so greater than the errors determined from the sensitivities, but are nevertheless still small compared to the Run 1 statistical errors. Once the final datasets have been created with the fixed gain parameters, the systematic errors will be re-evaluated.

The most dangerous potential systematic error has to do with the phases of the

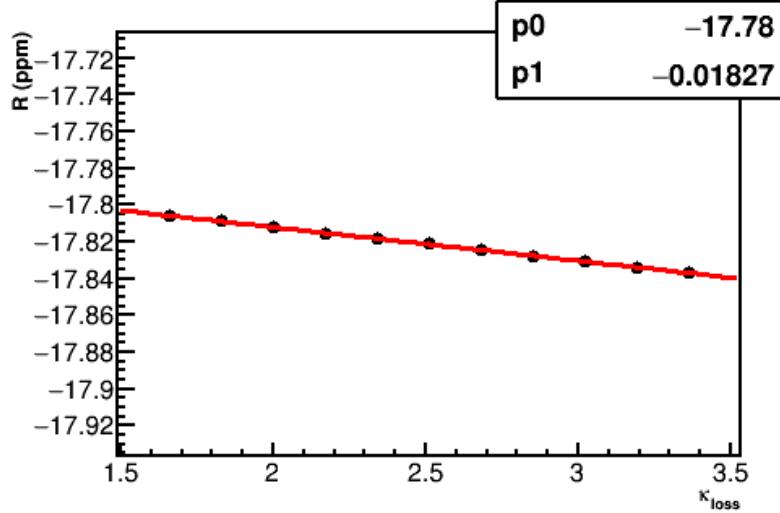


Figure 5.47: The sensitivity of R to the fixed κ_{loss} parameter. Error bars have been removed from the plot. Units are in ppm, data from the 9d dataset.

lost muons. If muon losses come from a population of muons with a different average phase than the stored muons, then there will be a phase shift over the course of a fill. This combined with the fact that muons are preferentially lost at early times implies the extracted ω_a frequency will be systematically pulled. Simulation and data have shown a correlation between phase and momentum at injection [106, 107]. If the losses are momentum dependent and the phase-momentum correlation is preserved after injection, then there will be a systematic error. From simulation losses are momentum dependent [108], the general idea being that the low kick applied in Run 1 resulting in an off-center stored muon beam results in more muons at greater radii with larger momentum near the edge of the storage aperture, compared to those muons on the inside of the storage region. The phase-momentum correlation originates from the fact that muons are born at different points within the accelerator beam-line before injection. Muons born at earlier locations in the beam-line will live longer within the

magnetic field, during which their spins will precess and their $g - 2$ phases change. Such muons could reasonably exist within a different injection parameter phase space, such that they exhibit larger betatron amplitudes once stored in the ring. Because losses most likely come from muons which oscillate to high betatron amplitudes, and for those muons at greater radii with larger momenta, such a systematic effect certainly exists, and thus an error needs to be estimated. Further simulation efforts are underway, along with data-driven studies, to understand and quantify the phase-momentum correlation and momentum-dependent loss probabilities [108, 109]. Here is given an estimation of the muon loss phase bias systematic error using the author's own calculations for the losses for the different Run 1 datasets.

The systematic shift in ω_a can be written as

$$\frac{\Delta\omega_a}{\omega_a} = \frac{1}{\omega_a} \frac{d\langle\phi\rangle}{dt}, \quad (5.40)$$

where $d\langle\phi\rangle/dt$ is the change in average phase over the course of a fill. With some assumptions made with regards to stored and lost muon populations this can be written as [108]

$$\frac{d\langle\phi\rangle}{dt} = f_s \cdot f_l \cdot \Delta\phi_{s-l}, \quad (5.41)$$

where f_s is the fraction of stored muons to all muons, f_l is the fractional loss rate, and $\Delta\phi_{s-l}$ is the phase difference between the stored and lost muon populations. The simulation studies in the afore-referenced document gives estimates of $f_s \approx 0.9$ and $\Delta\phi_{s-l} \approx 1$ mrad. The fractional losses (the integral in Equation 5.29 times the final fitted κ_{loss} parameter, assuming no losses are included in $L(t)$ before the fit start time) for the Run 1 precession frequency analysis datasets as determined in this analysis are shown in Figure 5.48. As shown the losses rise sharply at early times before leveling off to a constant rate around 100 μ s into the fill. Taking the

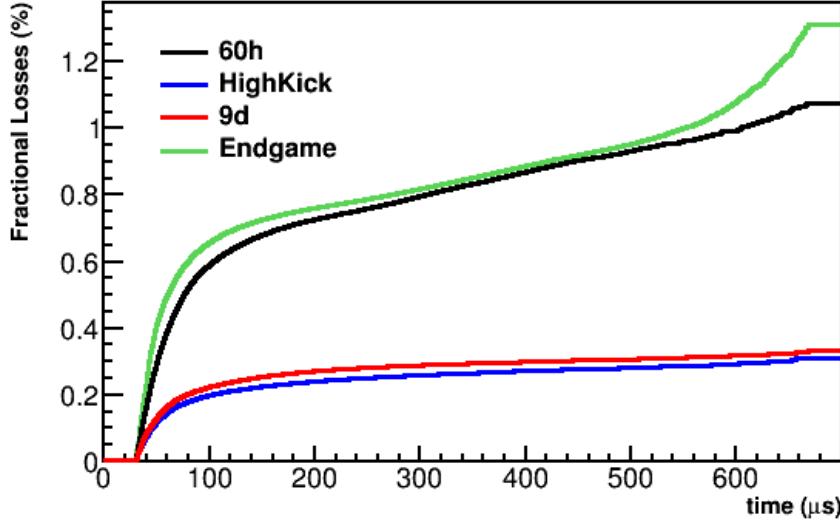


Figure 5.48: Fractional losses for the Run 1 precession frequency analysis datasets. The curves begin at $30.2\text{ }\mu\text{s}$ which is where the fit begins. A value of 1% at a specific time t indicates that there are 1% fewer stored muons at that time than there would be if there were no losses at all. The Endgame and 60h datasets can be seen to have the most losses, while the 9d and HighKick have less. This is due to the higher kicks in the latter datasets which put the stored muon beam on a more central orbit. The upward tail at the end of the Endgame dataset corresponds to the remnant proton contamination.

first $70\text{ }\mu\text{s}$ of the fit and assuming the loss rate is constant over the course of the fill, the approximate loss rates for the different datasets can be determined²⁷. These are given in Table 5.21. Applying the approximate loss probabilities into Equation 5.41, along with $\omega_a \approx 1.44\text{ rad}/1.44\text{ }\mu\text{s}$, the systematic shifts in the precession frequency are found. They are included in Table 5.21

Taking the shifts in ω_a as the systematic errors, the approximations here are on the order of 50 – 60 ppb for the 60h and Endgame datasets, and on the order of 20 ppb for the HighKick and 9d datasets. These estimates are the same order of magnitude

²⁷While the loss rate levels off after $100\text{ }\mu\text{s}$, losses occur mostly at early times, so this is a fine approximation.

Systematic Error due to Lost Muon Phase Bias		
Dataset	f_l (%/70 μ s)	$\Delta\omega_a/\omega_a$ (ppb)
60h	0.60	-53.6
HighKick	0.20	-17.9
9d	0.25	-22.3
Endgame	0.65	-58.0

Table 5.21: Loss rates and associated shift in ω_a for the Run 1 precession frequency analysis datasets. Loss rates are determined by inspection of the respective curves for the different datasets, and are approximate. The systematic shifts here are negative, due to the fact that the average phase of the stored muons is greater than that of lost muons, as determined in simulation.

as determined in other evaluations with slightly different parameters [107]. Estimates on the absolute upper bound of the systematic error is of order 125 ppb [87, 108]. As simulations improve and more studies are undertaken, the uncertainties here may be replaced in favor of actual corrections, in which case the systematic errors would then be the uncertainties in the evaluation of those corrections. While not below the target goal of 20 ppb, the systematic errors listed here are nevertheless small compared to the Run 1 statistical errors, expected to decrease going forward in future runs.

5.5.5 Systematic error from differential decay

Muons are injected into the storage ring with a range of momenta. Muons with larger momenta will in general live longer than muons with smaller momenta due to their time-dilated lifetimes. Muons at smaller radii therefore decay more often, leading to an increasing beam radial position and average momentum over the course of a fill. Decay positrons at greater radii with larger momenta in general take longer to reach calorimeters than those at small radii with smaller momenta. In the precession frequency measurement, the hit times of the detected positrons naturally correspond to offset times from when the muons decayed equal to the drift times. Because the

average drift time changes over the course of a fill due to the average momentum distribution increase, there is thus a changing phase over the course of a fill. To add to that, as discussed in Section 5.5.4, there is a phase-momentum correlation for muons at injection. This only compounds upon the changing phase even further. The systematic error due to this effect is called “differential decay.”

Underway beam-line simulations are needed to fully measure the phase-momentum correlation and fully estimate this systematic error. A calculation for the previous E821 experiment resulted in a shift on ω_a of $\langle \omega_a \rangle / \omega_a \approx -36.8$ ppb [110]. An estimation of the error neglecting the phase momentum correlation yielded a shift of -3.8 ppb [111]. Finally, an estimation by a first principles approach gave a value of -12 ppb [87]. The level of the fully evaluated systematic shift may not be small compared to the final target systematic error goal, but for Run 1 the scale of these estimates is deemed to be acceptable. For the purposes of this document, an upper bound of 40 ppb is taken as the systematic error. Because the differential decay effect only increases the average stored momentum of the beam and can’t decrease it, a correction can be applied if any final uncertainty is deemed to be too large.

5.5.6 Ratio construction systematic errors

In the construction of the ratio data, when filling the four sub-datasets as in Equation 5.11, the parameters T_a and τ_μ for the $g - 2$ period and muon lifetime need to be known a priori. If these parameters are incorrectly chosen, then it is possible there will be a systematic shift on R . This is especially important when considering T_a , because the quantity which the E989 experiment is measuring must be used in the analysis, creating a sort of self-dependence. The question then naturally arises as to how well the T_a parameter needs to be known. As described in Section 5.2.2, the input value for T_a is nominally taken as the result from the E821 experiment.

In order to determine systematic errors from these two fixed quantities, they were

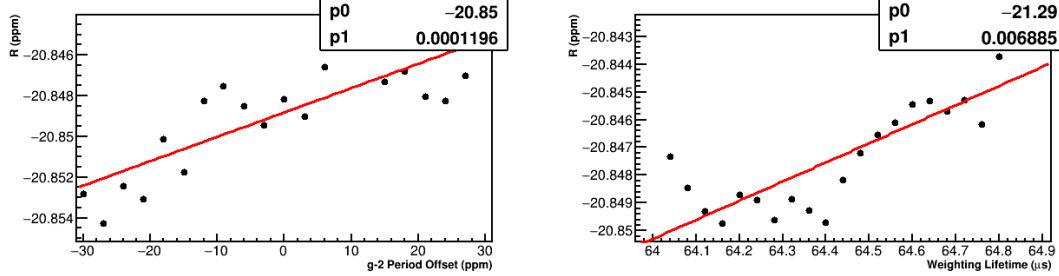
scanned over when forming the ratio data before fitting. The input value for T_a was varied from -30 ppm to 30 ppm around the default value in steps of 3 ppm. The input value for τ_μ was varied from $64.04\ \mu\text{s}$ to $64.84\ \mu\text{s}$ in steps of $0.04\ \mu\text{s}$. See Figure 5.49 for scan results for the 60h dataset. See Table 5.22 for the sensitivities determined from the scans for all datasets. As shown the sensitivities vary both positively and negatively for the different datasets, and are extremely small. The positive and negative variations imply there is no real systematic effect at play, and that as long as a reasonable choice for these two parameters is made, then the ratio data is very insensitive to the exact values chosen.

In order to be conservative however, a scale for the changes in R is given. Since the measured $g - 2$ period in the data is already modified by the hardware blinding, it is technically the hardware shifted $g - 2$ period that we want to use. The calorimeter digitizers use a “40” MHz clock which has been blinded to a value in the range of 39.997 to 39.999 MHz [94]. This corresponds to a 75 ppm range in the frequency, in a uniform distribution. Calculating the uncertainty from the uniform distribution and adding it in quadrature with a conservative 10 ppm uncertainty in the guess on the true $g - 2$ period from the E821 result,

$$\delta T_a = \sqrt{(75)^2/12 + 10^2} = 23.8 \text{ ppm.} \quad (5.42)$$

This results in a change in R on the order of 2.4 ppb for the 60h and HighKick datasets, and less for the 9d and Endgame datasets.

Similarly, the sensitivities of R to the chosen muon lifetime are very small, order ppb/ μs . Since the uncertainties in the muon lifetime are of order ns from fits to the data, any systematic errors from this parameter would be completely negligible even if the sensitivities all had the same sign.



(a) R versus input value for T_a , where the x axis is given in units of a ppm level shift of the default choice for T_a .

(b) R versus the input value for τ_μ . Parameter p_1 gives the sensitivity in ppm/ μ s.

Figure 5.49: Scans over ratio construction parameters for the 60h dataset. Error bars have been removed from the plots. In general the points are randomly spread around, and sensitivities are very small.

Sensitivity to Ratio Construction Parameters		
Dataset	dR/dT_a	$dR/d\tau_\mu$
60h	0.1	6.9
HighKick	-0.1	-4.1
9d	< 0.1	-1.1
Endgame	< 0.1	0.6

Table 5.22: Sensitivities of R to ratio construction parameters. dR/dT_a is in units of ppb/ppm, while $dR/d\tau_\mu$ is in units of ppb/ μ s. In both cases the sensitivities are both extremely small, and vary negatively and positively for the different datasets. **fix spacing of table**

5.5.7 Binning systematic errors

When constructing the time spectra to be fit, bin widths and the starting edge of the bin are by default chosen to be 149.2 ns and 0 ns respectively. In order to verify that no systematics arise from the choice of these parameters, the values were scanned over. The bin width was scanned from 148.7 ns to 149.7 ns in steps of 0.1 ns, while the bin edge was scanned from 0 ns to 149.2 ns in steps of 14.92 ns. Figure 5.50 shows scan results for the 9d dataset. Table 5.23 gives the sensitivities of R to both parameters.

Upon general inspection of the fit points themselves, it was found that the trends weren't so convincing as the points varied relatively widely. Still a line was fit to the points to assess the scale of the changes. For the bin edge scan, it was verified that a shift of one bin width returned the same fit results as the default shift of 0 ns. This combined with the negligible sensitivities and varying points implies no systematic effects on R from the choice of bin edge. For the sensitivities to the choice of bin width, not only did the points vary, but the trends were both positive and negative depending on the dataset. In general the choice of bin width should be optimized to be equal to the peak of the cyclotron period distribution of the stored muons, which from the fast rotation analysis informed the choice of 149.2 ns. Therefore it is reasonable to quote no systematic error for the choice of bin width. If one wanted to be conservative and quote a systematic error however, the sensitivities could be multiplied against the uncertainty in the optimal bin width. This uncertainty from the fast rotation analysis would certainly be less than 0.1 ns, which would correspond to uncertainties of 2.5, 0.6, 2.3, and 4.2 ppb for the 60h, HighKick, 9d, and Endgame datasets respectively, all of which are practically negligible.

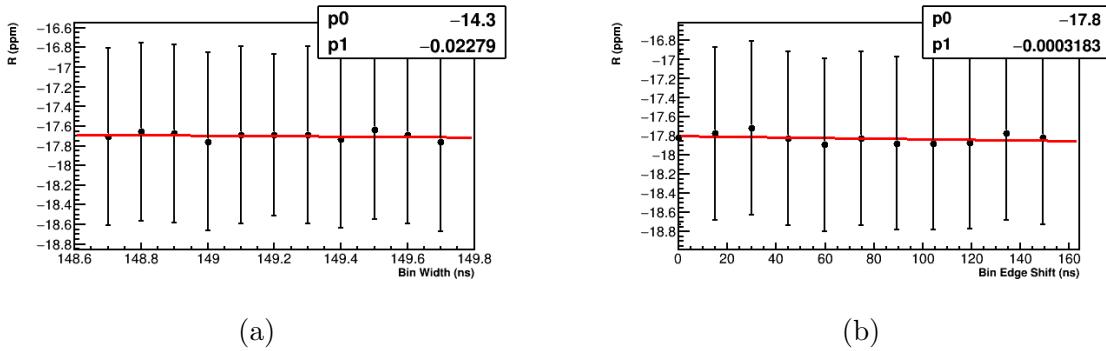


Figure 5-50: Scans over binning parameters for the 9d dataset. In general the points are randomly spread around, indicating no real systematic effect.

Sensitivity to Binning Parameters		
Dataset	$dR/d_{\text{bin width}}$	$dR/d_{\text{bin edge}}$
60h	24.5	-0.1
HighKick	6.0	-0.7
9d	-22.8	-0.3
Endgame	-41.7	-0.6

Table 5.23: Sensitivities of R to binning parameters. Units are in ppb/ns. While some of these values may appear significant, inspection of the actual plots reveals that the actual trends are not quite so convincing. **fix spacing of table**

5.5.8 Systematic errors in the E-field and pitch corrections

The electric field and pitch corrections modify the final value extracted for ω_a as described in Section 2.7. Any errors in the estimation of these corrections is by extension an error in the precession frequency measurement. The evaluation of these errors is performed by separate working groups. Given here are preliminary estimates of the errors.

The pitch correction is evaluated by the tracking analysis and is dependent on the vertical width of the muon beam. A preliminary analysis of the 60h dataset calculated an estimated correction and systematic error of [112, 113]

$$C_P \sim -160 \pm 15 \text{ ppb.} \quad (5.43)$$

This error is split into track reconstruction errors of which there are many ($\mathcal{O}(10 \text{ ppb})$), a tracker-calorimeter acceptance correction with a small expected error, and uncertainty in simulation models used in various parts of the tracking. While the actual correction itself may differ between datasets, the error in the correction is expected to be relatively consistent. The pitch correction error given in Equation 5.43 is taken as the systematic error for all datasets.

The E-field correction is evaluated with the fast rotation analysis and is dependent on the equilibrium radius of the muon beam. Analyses to the 60h, 9d, and Endgame datasets by the Cornell fast rotation analysis group yielded the results shown in Table 5.24. While not the final numbers for Run 1, the estimates and errors given will change only slightly from further DQC cuts. For the HighKick dataset which has not been evaluated at the time of writing, a systematic error on the E-field correction is taken as the largest error from the other datasets, that being 36 ppb for the 9d dataset.

The above corrections make certain assumptions with regards to the conditions

Electric field corrections			
Dataset	Correction	Statistical error	Systematic error
60h	-463	1.0	27.0
HighKick	N/A	N/A	N/A
9d	-463	1.0	36.0
Endgame	-467	1.0	20.0

Table 5.24: Electric field corrections estimated by the Cornell fast rotation analysis group [114, 115, 116]. Units are in ppb.

which influence the stored beam dynamics. Further errors can be attributable to non-linearities in the electrostatic quadrupole field, misalignment of the quadrupole plates, the changing voltage on the quadrupole plates due to the bad resistors, and any residual radial magnetic field. Preliminary analyses using simulation have estimated additional systematic errors on the pitch and E-field corrections of ~ 5 ppb and ~ 20 ppb respectively [117, 118]. These errors are conservatively added in quadrature with their respective errors evaluated above. Simulation efforts are continuing in order to identify and nail down these errors even further. While the errors listed here are greater than the target quadrature sum of 30 ppb, they are nonetheless small compared to the statistical error on R and are deemed acceptable for Run 1.

5.5.9 Systematic error due to beam motion

As described in Section 4.4, the vertical position of the stored muon beam changes over the course of a fill. Because the $g - 2$ phase is position dependent, this leads to a changing average phase and thus a systematic error. Preliminary studies fitting the average $g - 2$ phase with the tracker data have calculated systematic shifts in ω_a for the 60h and Endgame datasets of 115 and 180 ppb respectively [119]. Since these are one-directional shifts in ω_a , similar to the differential decay effect, they can be applied as corrections to the final extracted precession frequency. As systematic

studies mature and the shifts are better quantified, this is the intended action. For now these numbers are taken as limits on the systematic errors for their respective datasets, with the larger number from the Endgame dataset being applied to the HighKick and 9d datasets as well.

5.5.10 Systematic error summary

Table 5.25 gives all evaluated systematic errors in this analysis as described in the preceding sections. The final total quadrature sums of the systematic errors including the various preliminary and conservative estimates for some errors lie within the range 140 – 200 ppb. These are in comparison to the final target goal of 70 ppb for the precession frequency measurement. Once the gain issues are resolved and the other working groups improve their systematic uncertainty evaluations, the systematic errors for Run 1 and future runs of E989 will almost certainly reach the target goal. Indeed just removing the stored muon beam systematic error results in systematic errors in the range of 70 – 100 ppb, much closer to the target of 70 ppb. As the systematic errors stand currently in this analysis, they are all small compared to the Run 1 statistical errors of each respective dataset, making the presented analysis statistics-limited.

Run 1 Systematic Errors				
Error	60h	HighKick	9d	Endgame
Pileup amplitude	22.2	19.0	9.0	9.4
Pileup phase - time-shift	17.6	19.0	17.1	14.3
Pileup phase - energy-scale	19.4	3.7	5.5	5.3
In-fill gain amplitude	1.4	28.6	0.4	43.6
STDP On/Off	~ 11.0	~ 11.0	~ 11.0	~ 11.0
CBO frequency model	7.5	0.4	2.0	8.0
CBO decoherence envelope	17.6	18.0	28.7	4.3
Lost muon cuts	< 0.5	< 0.5	< 0.5	< 0.5
Fixed κ_{loss}	1.2	39.1	3.1	28.9
Ratio construction T_a	2.4	2.4	< 2.4	< 2.4
Ratio construction T_μ	< 0.1	< 0.1	< 0.1	< 0.1
Bin width	2.5	0.6	2.3	4.2
Quadrature Sum	41.0	59.4	37.1	57.3
Lost muon phase bias	53.6	17.9	22.3	58.0
Differential decay	< 40	< 40	< 40	< 40
Pitch correction	15.8	~ 15.8	~ 15.8	~ 15.8
E-field correction	33.6	~ 41.2	41.2	28.3
Stored beam motion	~ 115	~ 180	~ 180	~ 180
Quadrature Sum	138.1	190.4	190.9	196.0
Total Quadrature Sum	144.1	199.5	194.5	204.2

Table 5.25: Final systematic errors evaluated in the Run 1 precession frequency analysis to the 60h, HighKick, 9d, and Endgame datasets. All units are in ppb. The table is split into two sections. The upper section consists of systematic errors directly evaluated by the author while the lower section consists of preliminary systematic estimates by other working groups. The E-field and pitch correction errors have been added in quadrature with the quadrupole errors. The publication errors for the Run 1 datasets will change from these as the final DQC cuts are made and analyses improved, however the scale of these errors will remain consistent.

Chapter 6

Conclusion

Four datasets from Run 1 of E989 have been analyzed for this dissertation, those being the 60h, HighKick, 9d, and Endgame respectively. In each case the datasets are the not-quite-final datasets for Run 1. Precession frequency analysis was done using the Ratio Method, an analysis technique for fitting the decay positron time spectra which divides out the exponential decay and slow terms in the data. The final results for the blinded frequency R values for the different datasets, along with their total statistical and systematic errors, are given in Table 6.1. The sum total error for the four datasets analyzed in this report is 475.7 ppb assuming completely uncorrelated systematic errors, a reasonable approximation considering the different run conditions between datasets. The analysis is statistics limited, even with the conservative preliminary estimates for certain systematic errors as given in Table 5.25. The R values given here can be converted back into the precession frequency ω_a using Equation 5.14, once the datasets have been unblinded at the hardware and software levels. Lastly the E-field and pitch corrections can be applied, where preliminary estimates have been given in Section 5.5.8 for some of the datasets.

The expected error in the magnetic field measurement for Run 1 is $\mathcal{O}(140 \text{ ppb})$. That combined with the total error determined in this analysis would result in an error on a_μ of $\mathcal{O}(500 \text{ ppb})$, improving upon the uncertainty in the E821 measurement of 540 ppb. For the final Run 1 production datasets with the last DQC cuts and gain improvements, the final errors are expected to be improved upon slightly. An

Run 1 Final Results				
Dataset	<i>R</i>	$\sigma_{\text{stat.}}$	$\sigma_{\text{sys.}}$	$\sigma_{\text{tot.}}$
60h	-20.5562	1.3581	0.1441	1.3657
HighKick	-17.4755	1.4112	0.1995	1.4252
9d	-17.7182	0.9033	0.1945	0.9240
Endgame	-17.3406	0.6393	0.2042	0.6711
Total		0.4605		0.4757

Table 6.1: Run 1 final results for the precession frequency analysis datasets. The *R* values given here are mean values of fits to 50 different random seeds. The 60h dataset has a different software blinding than the rest, shown by the different mean *R* value. Statistical and systematic errors are included alongside the total error for each dataset. In each dataset case the error is statistics dominated. In the bottom row the total combined errors for the different datasets are shown, where the systematic errors are assumed to be completely uncorrelated. Units are in ppm.

independent measurement of a_μ that is statistically consistent with the E821 result would go a long way towards increasing the confidence in the discrepancy between theory and experiment of nearly four standard deviations. The Run 1 publication is expected to be complete in the spring of 2020. Data ~~has~~^{have} already been gathered for Run 2 of E989 and Run 3 will begin in the late fall of 2019, with Run 4 planned for 2021. With the rate improvements seen in Run 2 and the expected increases for Run 3 and Run 4, the target uncertainty goal of 140 ppb is a likely reality.

Appendix A

Tracking Derivations

A.1 Straw measurement angular correction

The tracker straws don't measure U and V coordinates directly, but instead measure the DCA radii deriving from measured hit times. In order to utilize the minimization procedure described in Section 4.2.2 on measured track parameters these radii must first be converted to U and V parameters, and similarly for the U and V errors. To first order the measured DCAs can be used identically as the U and V positions, but it was found that there were slight biases in the truth pulls due to this.

In order to improve the results, angular corrections were made to the DCAs to give more accurate estimates of the “measured” positions. It was found that for the error correction, assuming a straight particle path was sufficient for ideal results. For the position correction, it was found that assuming a circular particle path (constant field) correction for the curved tracks was sufficient. These corrections are dependent on the angle of the track, so it's important to note that during each successive iteration of the track fitting, the “measured” parameters are adjusted by the latest “predicted” momenta which change the angle of the track. The correction depends on whether the track went to the left or right side of the wire. Note that the momentum perpendicular to the straw measurement axis can be ignored since it doesn't affect the U or V value. A summary of the calculation of the right side correction follows, with the left side correction being calculated in a similar manner. See Figure A.1.

To solve for the measured U (or V) value, first use the following trigonometric