Analysis Note for 60H Dataset Relative Unblinding

Nick Kinnaird - Boston University

January 8, 2019

High Level Summary

- Lead Analysts: Nick Kinnaird, James Mott
- Positron Reconstruction Method: Recon West
- Software Release: V9_11_00
- Dataset: gm2pro_daq_full_run1_60h_5033A_withfullDQC
- Histogramming Method: Weighted Ratio
- Gain Correction Method: Default in reconstruction
- Pileup Correction Method: Asymmetic shadow window
- Lost Muon Spectrum Extraction: Triple coincidence, not included in ratio fit
- Models for CBO and VW: Exponential envelopes, frequency from tracking analysis
- $R = -19.something \pm 1.somethingppm$ (blinding with common string)
- $\chi^2/NDF = 4211/4200$something
- P value =

Final fit function:

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

$$f_\pm(t) = f(t \pm T_a/2)$$

$$f(t) = C(t)(1 + A\cos(\omega_a t + \phi))$$

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

Chapter 1

Analysis Procedures

1.1 Key parameters in reconstruction method

Find out procedures used in 60 hr production dataset

1.2 Analysis Data Preparation Procedure

- git branch: gm2analyses branch feature/KinnairdAnalyses
- Majority of code located under gm2analyses/macros/RatioMacro

1. Submit jobs to OSG to run the rootTreesAndLostMuons.fcl file which produces root trees of positron hits using the ClusterTree analyzer module and coincident MIP hits using the TestCoincidenceFinder analyzer module.
2. Submit jobs to Fermigrid to produce histograms from root trees using the ClusterTreeToHistsPileup.C macro in RatioMacro/HistMaking. Beyond standard threshold histograms this macro produces pileup and lost muon histograms all within the same root file.

1.3 Histogramming Procedure

Method: Weighted Ratio (threshold)

1. Loop through all clusters and apply an artificial deadtime (ADT) to combine hits within 6 ns into a single pulse using the same procedure and code that the pileup method uses (see below). Drop clusters with time $< 25\mu s$ or time $> 600\mu s$.
2. Histograms are constructed with ROOT's TH1F class with 149.15 ns bins from $0 - 699.96095\mu s$ corresponding to 4693 bins.
3. Randomize times by $\pm 149.15/2$ ns and fill histograms for energies > 1.7 GeV. Randomization uses ROOT's default TRandom3 class.
4. Fill one of the four histograms $\{u_+(t), u_-(t), v_1(t), v_2(t)\}$ as shown in Equation A.18 per cluster. The associated histogram is determined by generating a random number between 0 and 1, and comparing that number to the relative probabilities of the different weights.
5. Clusters filled into the $u_+(t)$ histogram have their times shifted by $t \rightarrow t - T/2$ and clusters filled into the $u_-(t)$ histogram have their times shifted by $t \rightarrow t + T/2$.

1.4 Gain Correction Procedure

Gain correction method: Default by the Italian Calibration Team

1. Long term gain is corrected using out-of-fill lasers included normalization from the Source Monitor.

2. In-fill gain is corrected using in-fill lasers including normalization from the Source Monitor.

3. Short-term double pulse (SDTP) effect is not included.

1.5 Pileup Correction Procedure

Pileup correction method: Asymmetric shadow window

1. Create a vector of clusters per calorimeter per fill. For each cluster look for a second cluster in a window from 12-18 ns after the time of the first cluster. This corresponds to a shadow dead time (SDT) of 6 ns and a shadow gap time (SGT) of 12 ns, equal to 1 and 2 times the applied ADT respectively.

2. Create shadow doublets with energies and times as:

$$E_{doublet} = E_1 + E_2$$
$$t_{doublet} = \frac{t_1 * E_1 + (t_2 - SGT) * E_2}{E_1 + E_2}$$

3. Randomize $t_{doublet}$ times by $\pm 149.15/2$ ns as in the histogramming procedure described above.

4. For each calorimeter construct a pileup spectrum P = doublets - singlets = D - S, where the singlets are subtracted at time $t_{doublet}$ as opposed to their individual times, and pulses are only added or subtracted if they are above 1.7 GeV. Subtract P off energy and threshold histograms.

5. For pileup subraction in the ratio method, randomly split associated doublets and singlets into 4 separate histograms as is done in the histogramming procedure described above, with times shifted accordingly. Subtract 4 pileup histograms off corresponding $\{u_+(t), u_-(t), v_1(t), v_2(t)\}$ histograms before forming the ratio.

6. The errors of the pileup corrected histogram were determined to be:

$$\sigma(N_{corrected}) = \sqrt{N_{corrected} + 2N_1 + 6N_4},$$

   where N_1 is the number of doublets where both singlets were below threshold, and N_4 is the number of doublets where both singlets were above threshold, and this is a quantity evaluated at each time bin. (Cite this? DocDB 14830. Derive this in the appendix?) A histogram of error multipliers was created by factoring out the $N_{corrected}$ term, which is then applied to the bin errors before fitting. This is true even for the ratio errors to good approximation. (Cite this? Derive it as JP did?) Note that I did not time randomize the N_1 and N_4 entries when constructing the correct errors, which should have a negligble effect.

7. The pileup correction at the triplet/contamination level is not included. The machinery exists to apply such a correction, but it requires more work to get it correct. It has been determined not be necessary for the 60H and Run 1 data.

1.6 Lost muon spectrum extraction procedure

Method: Triple coincidence of clusters

Note that the lost muons are not included in the ratio fit because the ratio method divides out such slow effects. This is reflected by the lack of a low frequency peak in the FFT of the fit residuals for the ratio fit, whereas such a peak exists for T method fits. I include here however my method for extracting the lost muon function for possible future systematic studies.

1. Triple coincidence of clusters in 3 consecutive calorimeters are made with an energy cut of 100 MeV < E < 250 MeV and 5 ns < dt < 8.5 ns.

2. A time histogram is made with the muon cluster in the first calorimeter.

3. The function that would be used in the final fit is:

$$\Lambda(t) = 1 - \kappa_{loss} \int_0^t L(t')e^{(-t'/\gamma\tau_\mu)}dt'$$

   where $L(t)$ is the triples histogram, and an arbitrary 10^{-6} factor has been absorbed into κ_{loss} in order to bring it to a more reasonable value (from $\mathcal{O}(10^{-10})$ to $\mathcal{O}(10^{-4})$).

1.7 Beam Dynamics: CBO Model

1. The CBO frequency as a function of time is taken from the tracking analysis, DocDB 14208. The form used is

$$\omega_{cbo}(t) = \omega_0(1 + \Delta\omega t + Ae^{-t/\tau_A} + Be^{-t/\tau_B})$$

   with parameters determined from station 12 in the 60H dataset and fixed in the fit as:

$$\begin{aligned} \Delta\omega &= 1.86 \times 10^{-8}\,\mathrm{ns}^{-1}, \\ A &= -0.0504, \\ \tau_A &= 73.3\,\mu\mathrm{s}, \\ B &= -0.131, \\ \tau_B &= 16.6\,\mu\mathrm{s}. \end{aligned}$$

   The parameter ω_0 is allowed to float in the fit and starts with a value of $2.3051\,\mathrm{rad}\,\mu\mathrm{s}^{-1}$. The ratio method has trouble with letting the other parameters float, and fixing them to various values does not change the fit results significantly.

2. Because the ratio method divides out the CBO partially (reduction by a factor of ~ 5 in the FFT cbo peak amplitude), the ratio fit has a hard time fitting the CBO lifetime. Therefore τ_{cbo} is fixed to $180\,\mu$s, determined from a T Method fit to the same data.

3. An exponential function is assumed for the CBO decoherence.

4. The N_{cbo} term is included in the fit, A_{cbo} and ϕ_{cbo} are excluded. The 2CBO term is excluded.

5. The final CBO function is:

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

1.8 Beam Dynamics: Vertical Waist Model

1. ω_a is senstive to the width of the beam, which is characterized by the frequency

$$\begin{aligned} f_{VW} &= f_{cyc} - 2f_y, \\ f_y &= f_{cbo}\sqrt{\frac{2f_{cyc}}{f_{cbo}} - 1}. \end{aligned}$$

In the 60H dataset $f_{VW} \approx 2.3\,\mathrm{MHz} \approx 10 \cdot \omega_a$, an even multiple of the g-2 frequency. Because of this, the vertical waist largely cancels in the numerator of the ratio. (show this explicitly?) This combined with the time randomization to remove the fast rotation (f_{cyc}) means the vertical waist does not need to be included in the ratio fit. This is justified by the lack of a vertical waist peak in the FFT of the residuals of the fit. (mention at all how there is a very small but observable peak at very early times or omit it?)

2. If included, an exponential function is assumed for the VW decoherence as in the CBO, and the final VW term is

$$V(t) = 1 + A_{VW} e^{-t/\tau_{VW}} \cos(\omega_{VW} t + \phi_{VW})$$

3. ω_{VW} is assumed to be a constant value even though the CBO frequency changes vs time.

1.9 Final Fit Function

The following function is used for the final fit for each calorimeter and for the calorimeter sum:

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

$$f_\pm(t) = f(t \pm T_a/2)$$

$$f(t) = C(t)(1 + A\cos(\omega_a t + \phi))$$

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

All parameters are floating except for terms in $\omega_{cbo}(t)$ and τ_{cbo} as described above.

Chapter 2

Analysis Results

2.1 Pre-corrected and corrected energy and time spectra

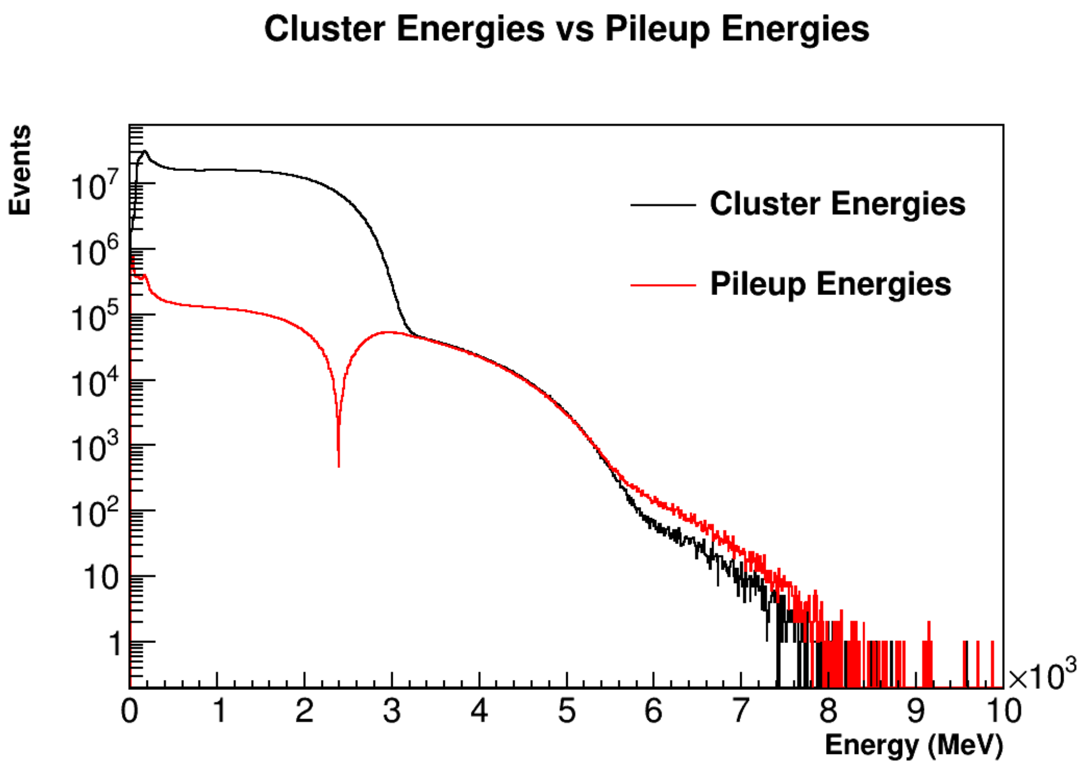
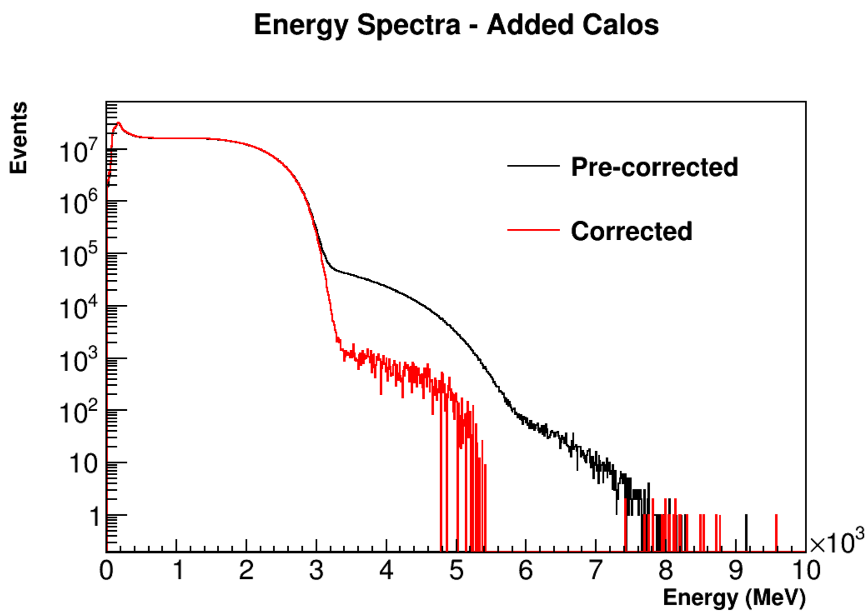
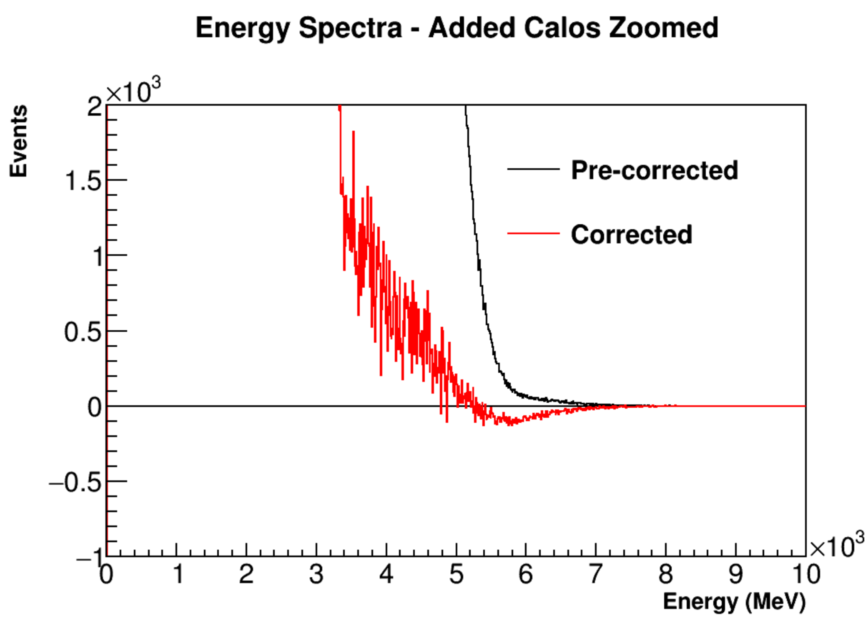


Figure 2·1: Cluster energies in black are plotted vs pileup energies in red, for all calorimeters added together. At energies below about 2.4 GeV the pileup spectrum goes negative. In this plot the absolute value of the pileup energies is plotted, and a spike at about 2.4 GeV can be seen as a consequence of this. Due to the triplets and contamination in the pileup spectrum, the red and black curves can be seen to diverge at high energies.



(a) Log scale - the corrected energy spectrum goes negative around 5 GeV.



(b) Linear scale - zoomed in to show the shape.

Figure 2·2: 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 exacltly along zero.

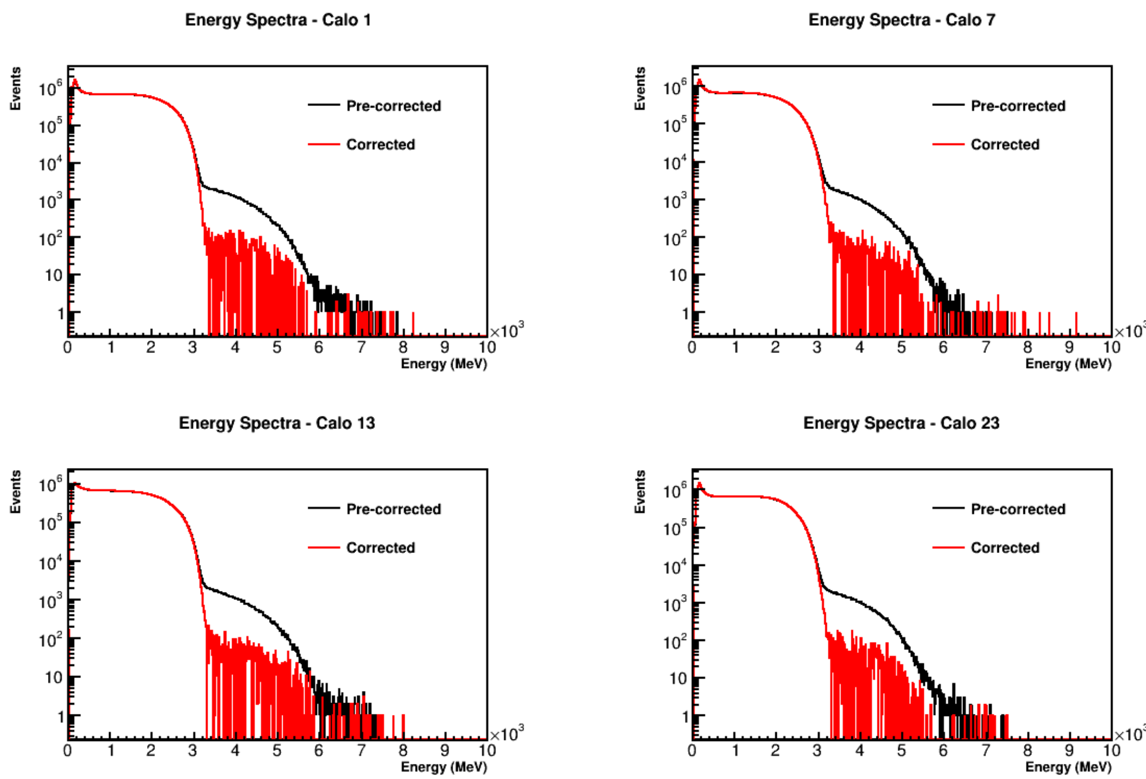


Figure 2·3: Pre-corrected and corrected energy spectra for calorimeters 1, 7, 13, and 23 plotted on a log scale.

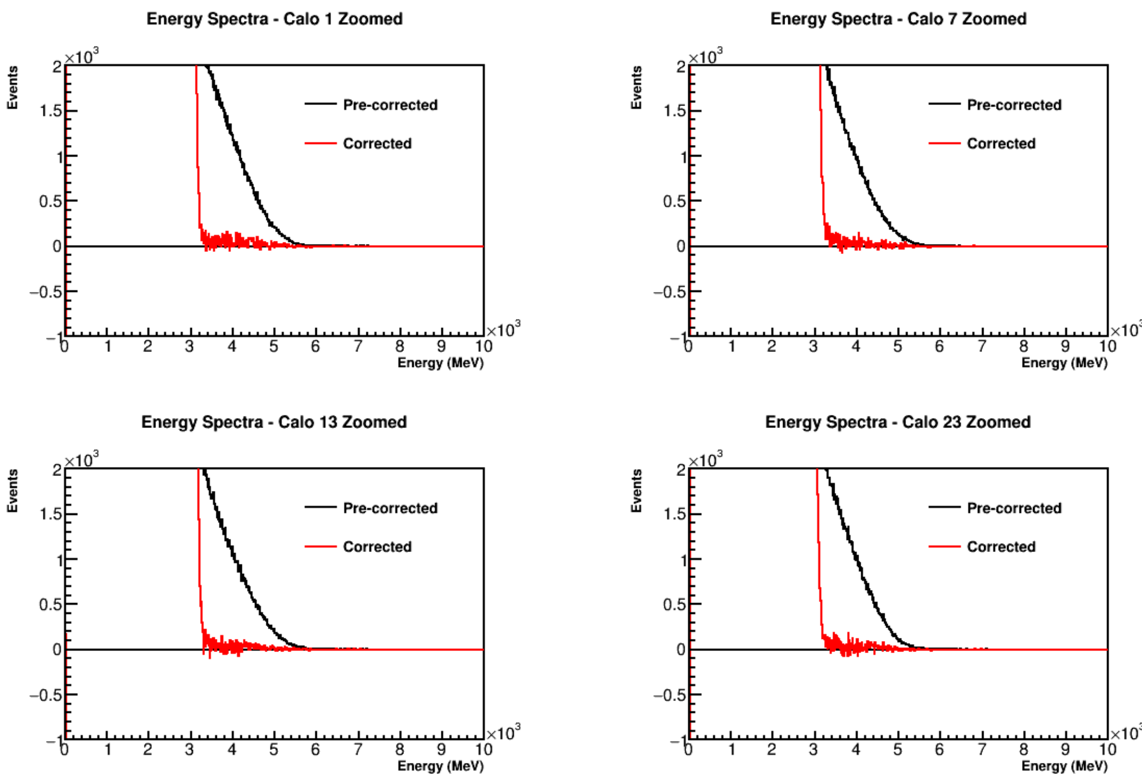


Figure 2·4: Pre-corrected and corrected energy spectra for calorimeters 1, 7, 13, and 23 plotted on a linear scale and zoomed in.

2.2 6 Parameter Ratio Fit

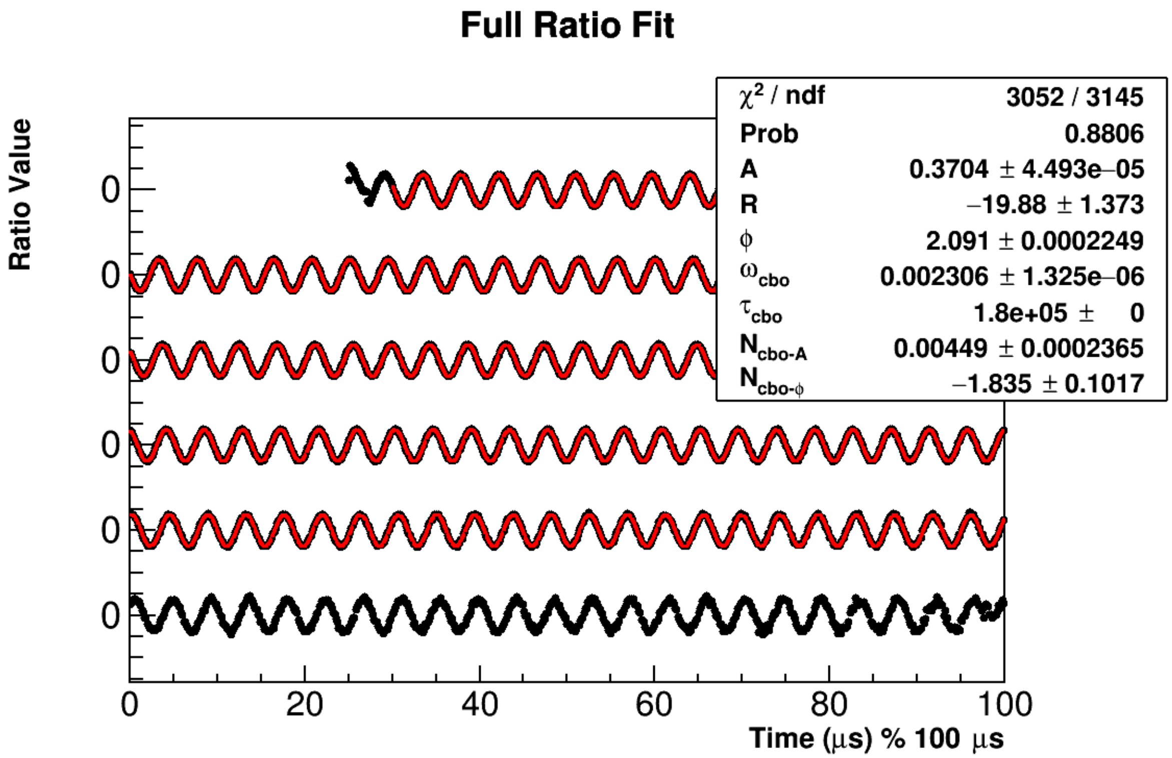
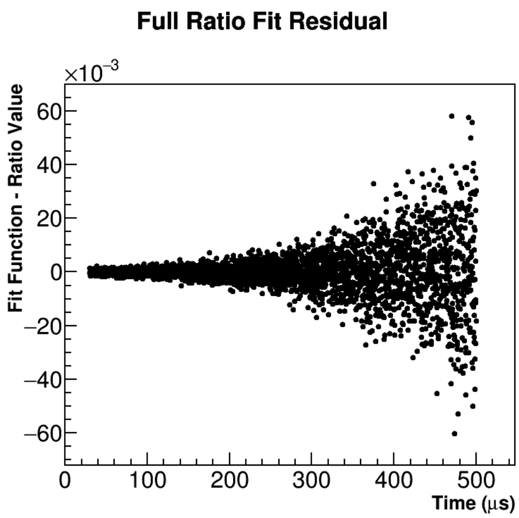
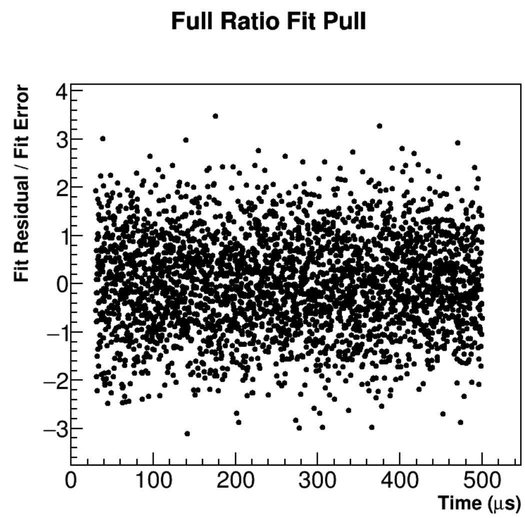


Figure 2·5: Final fit result for the 60 hour dataset. The fit includes 6 free parameters and one fixed. The x axis is in units of μs modulo 100 μs, with successive portions of the data points and fit shifted downwards on the plot. The parameter values in the stats box for the CBO frequency and lifetime are in units of ns. R is blinded locally. The fit ranges from 30 μs to 500 μs.

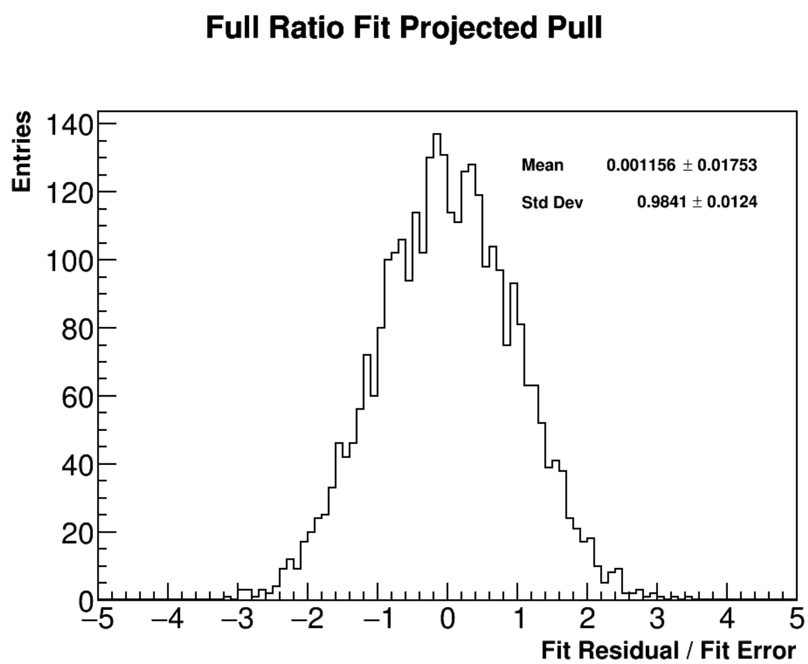
2.3 Residual and FFT



(a) Fit residuals.



(b) Fit pulls.



(c) Fit pulls projected onto the y axis. Note the Gaussian shape centered around 0 with unit width.

Figure 2·6: Residuals and pulls for the full ratio fit.

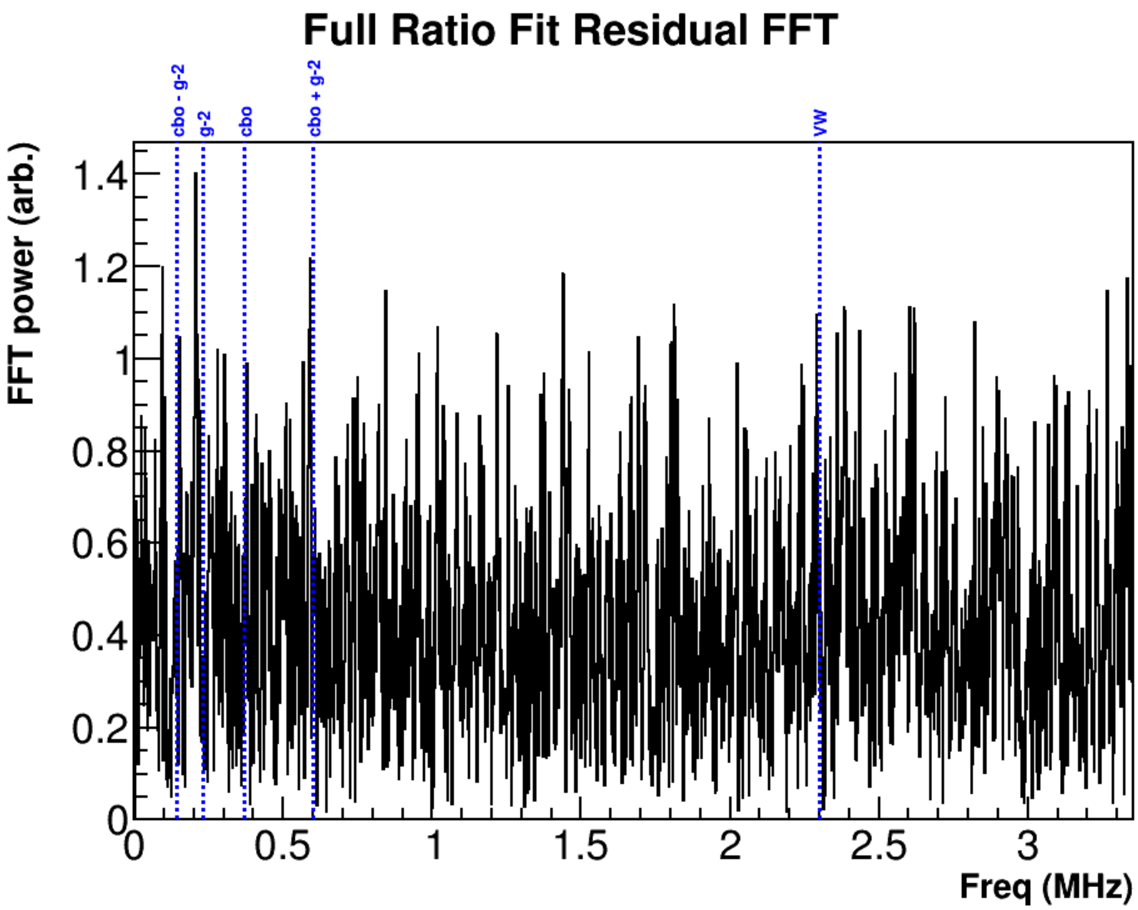


Figure 2·7: FFT of the residuals of the full ratio fit. No significant peaks remain in the ratio fit residuals after fitting with CBO terms. Overlayed are dotted lines for the $g-2$, CBO, and vertical waist frequencies. Peaks close to the lines are coincidental but don't line up when zoomed in.

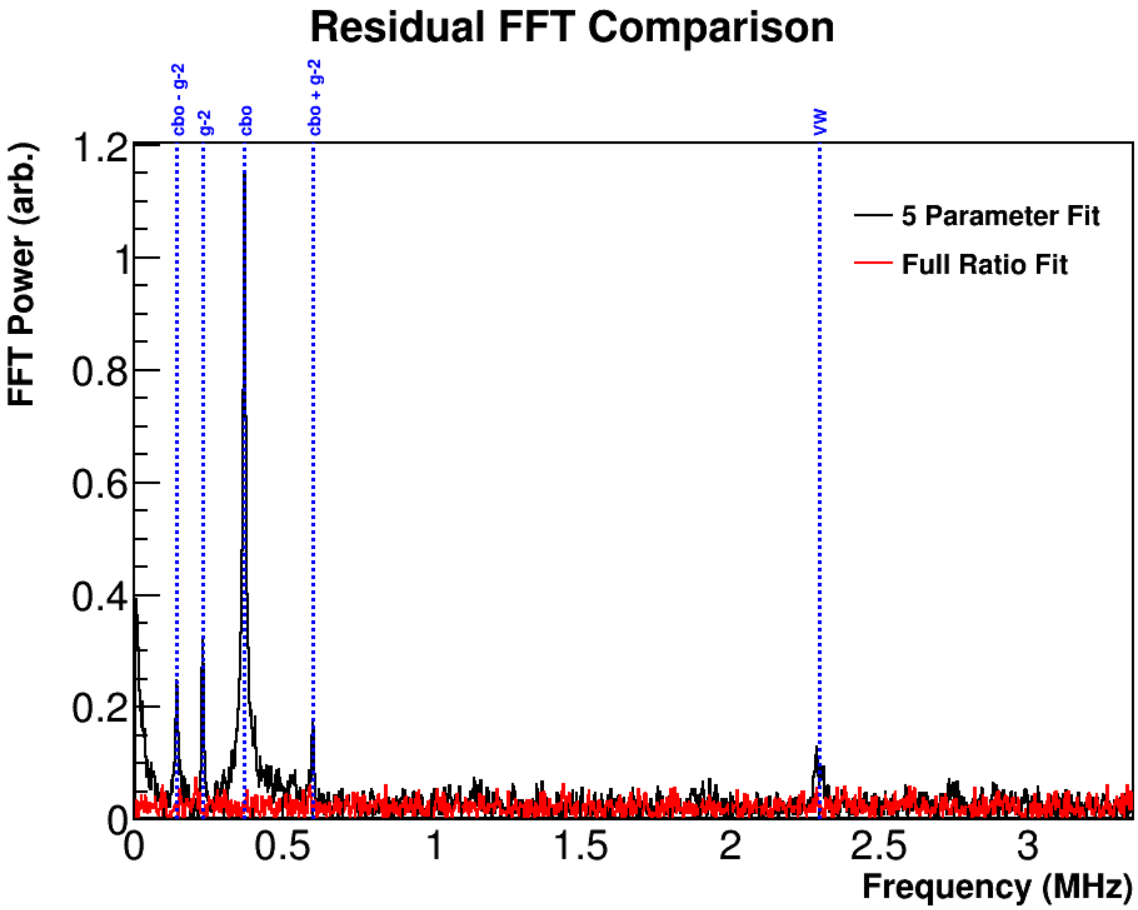
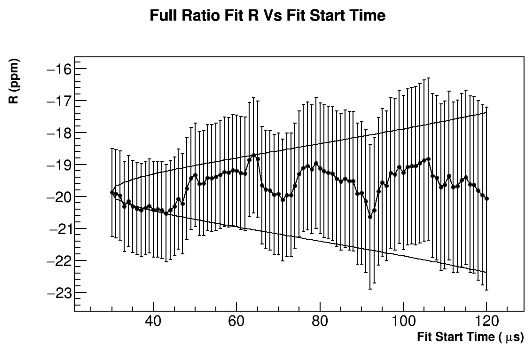
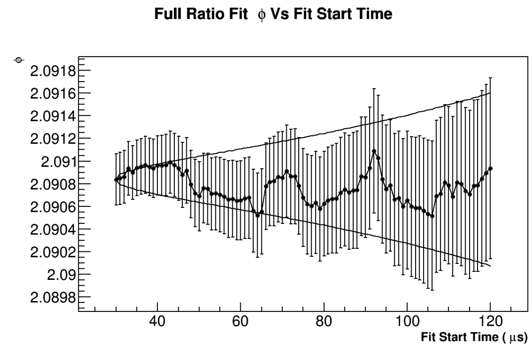


Figure 2·8: A plot of the FFT of the residuals of the fit for the five parameter fit compared to the ratio fit. In black is the FFT for a five parameter fit, where peaks for the CBO and vertical waist can be seen as well as the $g-2$ peak. In red is the FFT of the full ratio fit residuals, where it has been scaled up to be visible on this plot.

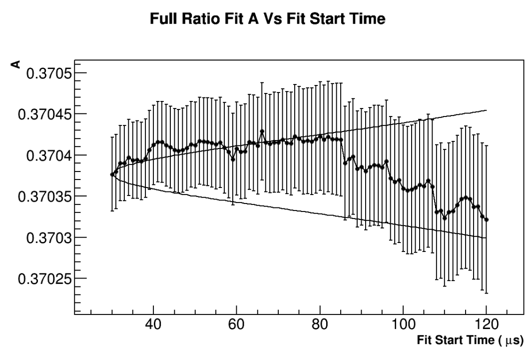
2.4 Start time scans



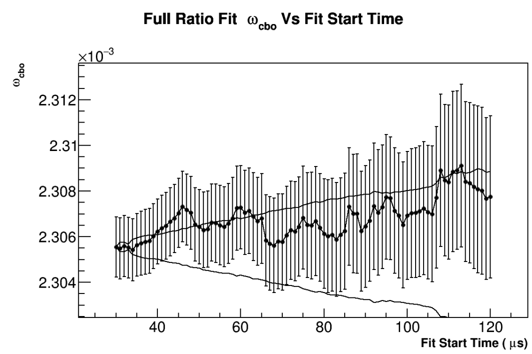
(a)



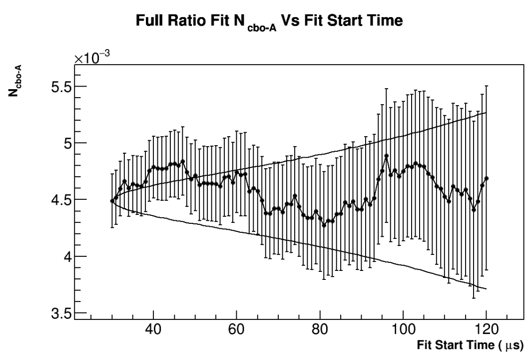
(b)



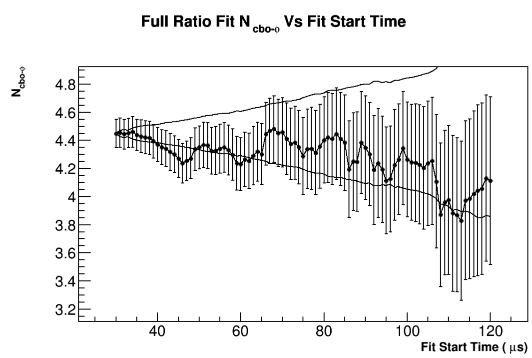
(c)



(d)



(e)



(f)

Figure 2·9

2.5 Results vs calorimeter

2.6 Correlation matrix for fit parameters

Chapter 3

Systematic Uncertainty Evaluations

3.1 Sensitivity of ω_a to gain corrections

3.2 Sensitivity of ω_a to pileup

3.3 Sensitivity of ω_a to lost muon function shape

3.4 Sensitivity of ω_a to CBO function

3.5 Sensitivity of ω_a to VW function

3.6 Sensitivity of ω_a to various effects

3.7 Final Systematic Uncertainty Table

3.8 Final Results

Should my appendix really be a section at the front of the report? Or should it be in the appendix as I've place it here?

Appendix A

Ratio Method Derivation and Fit Function

Consider the 5 parameter function:

$$N_5(t) = N_0 e^{-t/\tau}(1 + A\cos(\omega_a t + \phi)), \tag{A.1}$$

which describes some ideal dataset in histogram format. Here ϕ will be set to zero for simplicity. Now define the variables $u_+(t)$, $u_-(t)$, $v_1(t)$, and $v_2(t)$ as

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

where the 1/4 out front reflects randomly splitting the whole dataset into 4 equally weighted sub-datasets, and T is the g-2 period known to high precision, $\mathcal{O}(10^{-6})$. This corresponds to a weighting of 1:1:1:1 between the datasets. To be explicit here regarding the signs, the counts that are filled into the histogram described by u_+ have their times shifted as $t \to t - T/2$, which is what the function $N_5(t + T/2)$ describes, and vice versa for u_-. To form the ratio define the variables:

$$\begin{aligned} U(t) &= u_+(t) + u_-(t) \\ V(t) &= v_1(t) + v_2(t) \\ R(t) &= \frac{V(t) - U(t)}{V(t) + U(t)}. \end{aligned} \tag{A.3}$$

Plugging in and dividing the common terms $(N_0 e^{-t/\tau}/4)$,

$$R(t) = \frac{2(1 + A\cos(\omega_a t)) - e^{-T/2\tau}(1 + A\cos(\omega_a t + \omega_a T/2)) - e^{T/2\tau}(1 + A\cos(\omega_a t - \omega_a T/2))}{2(1 + A\cos(\omega_a t)) + e^{-T/2\tau}(1 + A\cos(\omega_a t + \omega_a T/2)) + e^{T/2\tau}(1 + A\cos(\omega_a t + \omega_a T/2))}. \tag{A.4}$$

Now set $\omega_a T/2 = \delta$, and note that T is really

$$\begin{aligned} T = T_{guess} &= \frac{2\pi}{\omega_a} + \Delta T, \\ \Delta T &= T_{guess} - T_{true}. \end{aligned} \tag{A.5}$$

Being explicit,

$$\delta = \frac{\omega_a}{2} T_{guess} = \frac{\omega_a}{2}(\frac{2\pi}{\omega_a} + \Delta T) = \pi + \pi \frac{\Delta T}{T_{true}} = \pi + \pi(\delta T), \tag{A.6}$$

and δ can be redefined as

$$\delta = \pi(\delta T), \tag{A.7}$$

by flipping the sign of any cosine terms that contain δ.

Then, using the trig identity

$$\cos(a \pm b) = \cos(a)\cos(b) \mp \sin(a)\sin(b) \tag{A.8}$$

so that

$$\begin{aligned} \cos(\omega_a t \pm \delta) &= \cos(\omega_a t)\cos\delta \mp \sin(\omega_a t)\sin\delta \\ &\approx \cos(\omega_a t)(1 - \delta^2) \mp \sin(\omega_a t)\delta \\ &\approx \cos(\omega_a t), \end{aligned} \tag{A.9}$$

since $\delta \sim O(10^{-5})$, the ratio becomes

$$R(t) \approx \frac{2(1 + A\cos(\omega_a t)) - (1 - A\cos(\omega_a t))(e^{-T/2\tau} + e^{T/2\tau})}{2(1 + A\cos(\omega_a t)) + (1 - A\cos(\omega_a t))(e^{-T/2\tau} + e^{T/2\tau})}. \tag{A.10}$$

Expanding

$$e^{\pm T/2\tau} = 1 \pm \frac{T}{2\tau} + \frac{1}{2}\left(\frac{T}{2\tau}\right)^2 \pm \ldots, \tag{A.11}$$

repacing and simplifying,

$$R(t) \approx \frac{A\cos(\omega_a t) - C(1 - A\cos(\omega_a t))}{1 + C(1 - A\cos(\omega_a t))}, \tag{A.12}$$

where

$$C = \frac{1}{16}\left(\frac{T}{\tau}\right)^2 \approx 2.87 * 10^{-4}. \tag{A.13}$$

Using the expansion

$$f(x) = \frac{1}{1+x} = 1 - x + x^2 - \ldots, \quad |x| < 1, \tag{A.14}$$

and since C is small, the denominator can be manipulated such that

$$\begin{aligned} R(t) &\approx (A\cos(\omega_a t)) - C(1 - A\cos(\omega_a t)))(1 - C(1 - A\cos(\omega_a t))) \\ &\approx A\cos(\omega_a t) - C + CA^2\cos^2(\omega_a t), \end{aligned} \tag{A.15}$$

after dropping terms of $\mathcal{O}(C^2)$ and higher. In practice the last term is ommitted since it has a minimal effect on the fitted value of ω_a [**cite**], and one arrives at

$$R(t) \approx A\cos(\omega_a t) - C, \tag{A.16}$$

the conventional 3 parameter ratio function.

In order to avoid approximations one can instead weight the counts in the histograms as

$$u_+(t) : u_-(t) : v_1(t) : v_2(t) = e^{T/2\tau} : e^{-T/2\tau} : 1 : 1, \tag{A.17}$$

so that

$$\begin{aligned} u_+(t) &= \frac{e^{T/2\tau}}{2 + e^{T/2\tau} + e^{-T/2\tau}} N_5(t + T/2) \\ u_-(t) &= \frac{e^{-T/2\tau}}{2 + e^{T/2\tau} + e^{-T/2\tau}} N_5(t - T/2) \\ v_1(t) &= \frac{1}{2 + e^{T/2\tau} + e^{-T/2\tau}} N_5(t) \\ v_2(t) &= \frac{1}{2 + e^{T/2\tau} + e^{-T/2\tau}} N_5(t). \end{aligned} \tag{A.18}$$

(These factors out front aren't so far off from 1/4 since $e^{\pm T/2\tau} \approx e^{\pm 4.35/2*64.4} \approx 1.034, .967$.) Then instead $R(t)$ becomes

$$R(t) = \frac{2(1 + A\cos(\omega_a t)) - (1 - A\cos(\omega_a t + \delta)) - (1 - A\cos(\omega_a t - \delta))}{2(1 + A\cos(\omega_a t)) + (1 - A\cos(\omega_a t + \delta)) + (1 - A\cos(\omega_a t + \delta))}, \tag{A.19}$$

where the $e^{\pm T/2\tau}$ terms out front now cancel. Using Equation A.9 again and this time avoiding approximations in δ,

$$R(t) = \frac{2A\cos(\omega_a t)(1 + \cos\delta)}{4 + 2A\cos(\omega_a t)(1 - \cos\delta)}, \tag{A.20}$$

after simplifying. In the limit that

$$\delta = \pi(\delta T) \to 0 \tag{A.21}$$

since δT is small,

$$R(t) \approx A\cos(\omega_a t), \tag{A.22}$$

with the only approximation being made at $\mathcal{O}(\delta^2) \sim \mathcal{O}(10^{-10})$.

Finally, while the 3 parameter ratio function suffices for fits to data containing slow modulations, it does not suffice for faster oscillation features. In that case it is more useful to fit with the non-approximated or simplified version of the ratio,

$$\begin{aligned} R(t) &= \frac{v_1(t) + v_2(t) - u_+(t) - u_-(t)}{v_1(t) + v_2(t) + u_+(t) + u_-(t)}, \\ &= \frac{2f(t) - f_+(t) - f_-(t)}{2f(t) + f_+(t) + f_-(t)}, \end{aligned} \tag{A.23}$$

where

$$\begin{aligned} f(t) &= C(t)(1 + A\cos(\omega_a t + \phi)) \\ f_\pm(t) &= f(t \pm T_a/2), \end{aligned} \tag{A.24}$$

and $C(t)$ can encode any other effects in the data that need to be fitted for, such as the CBO,

$$C(t) = 1 + A_{cbo} e^{-t/\tau_{cbo}} \cos(\omega_{cbo} t + \phi_{cbo}). \tag{A.25}$$

Additionally, any other fit parameters such as A or ϕ can be made a fuction of t. Using the non-approximated form for the final fit function gives greater confidence in the fit results for the high precision ω_a extraction necessary for the experimental measurement.

Appendix B

Ratio Method Errors