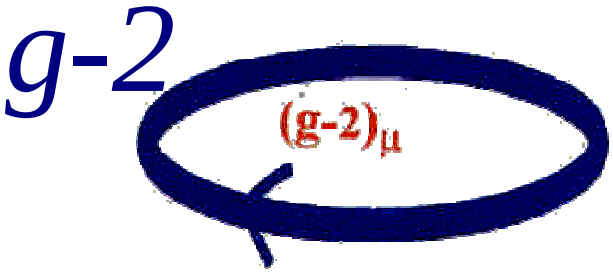


Brookhaven National Laboratory

Brookhaven Science Associates

Upton, New York 11973

Muon g-2 Note No. 453

Title: **Asymmetry-weighted analysis of the 2001, ω_a data**

Author: **Chris Polly**

Affiliation: **Indiana University**

Date: **Original written December 11, 2003**

Posted: **September 30, 2005**

Asymmetry-weighted analysis of the 2001 ω_a data

Chris Polly

11th December 2003

Abstract

An analysis of the g-2 frequency for the 2001 $\mu-$ dataset is presented. A multi-parameter fit to asymmetry-weighted data is the main focus. The evolution of a complete multi-parameter fit with an acceptable χ^2 at extremely early times (20 μs) is discussed. Results are compared to a traditional analysis and many consistency checks are applied. Topics also include run selection, gain correction, muon loss construction, pileup subtraction, and various simulation results.

Contents

1 Philosophy

The general approach of this analysis was to maximize the statistical power of the dataset while keeping systematics controlled at an acceptable level. As such, the data is fit with each event entering the histogram with a weight corresponding to its asymmetry. From a physical perspective, this type of weighting reflects the underlying information contained in the parity-violating decay of the muon. Specifically, higher-energy decay electrons carry more information about the underlying muon spin direction, so should be weighted preferentially in the fit. It can be shown that by choosing a weight proportional to the asymmetry, the statistical power of the dataset is maximized. A quick confirmation of this fact can be recognized by the equivalence of analyzing the data by energy band versus applying one fit to asymmetry-weighted data. An analysis by narrow energy bin will reproduce the results of performing a single asymmetry-weighted fit to the entire dataset. Both the g-2 frequency and its corresponding error are equivalent in the two approaches.

Traditional weighting, where all events enter the histogram with a weight of one, results in an easier statistical analysis at the expense of sacrificing some of the statistical power of the dataset. Uniform-weighting takes a low asymmetry signal sitting on a large background and treats it as though it were on equal footing with the high asymmetry data. This results in an increase in the error on the g-2 frequency and imposes an energy threshold below which adding more data only *decreases* the statistical power. In addition to this artificial threshold, the unnaturalness of uniform-weighting is further illustrated by the fact that analyzing the data by energy bin is no longer an equivalent treatment. In contrast, an asymmetry-weighted analysis can be performed with any choice of energy-binning and the results will be unaffected.

If the BNL g-2 experiment were a counting experiment (like earlier g-2 experiments) where only the times of events above a threshold could be recorded, then using w=1 would be the only way to analyze the data. Similarly, if it were an integrating experiment where only the integrated charge in each detector were recorded, then one would be forced to use the Q method which is equivalent to w=E weighting. Neither w=1 nor w=E produce an incorrect frequency, but they both reduce the statistical power. Because of the innovation of recording a pulseshape for each decay, E821 is able to treat the data in a manner consistent with the underlying parity violation. Any weighting scheme will work, but only the one that reflects the true physics, w=A, will minimize the error.

Given the equivalence of an analysis by energy bin, it is fair to ask why one would prefer to use the asymmetry-weighted fit. The answer stems from several technical points. As was shown in Fred Gray's analysis last year, subdividing the dataset makes it harder to obtain an acceptable χ^2 for each of the fits at early times. This is due to the fact that at earlier times there are many perturbations to the fitting function from gain effects, CBO, pileup, muon losses, etc. Most of these effects are corrected to some degree of precision, but there is always a residual effect that persists at some level. By fitting many detectors and energy bands simultaneously, these residual perturbations tend to be reduced or cancel. For instance when correcting for gain with the average energy data, one expects that it is as likely to over-correct as under-correct the data. If this is the case, then adding multiple detectors will reduce the net gain effect. Also, CBO effects on the acceptance are known to be largely out of phase between detectors on opposite sides of the ring. By adding these detectors the effect is reduced. Higher energy bands have fractionally larger pileup and much stronger CBO modulations, by adding in the low energy data the effects of both

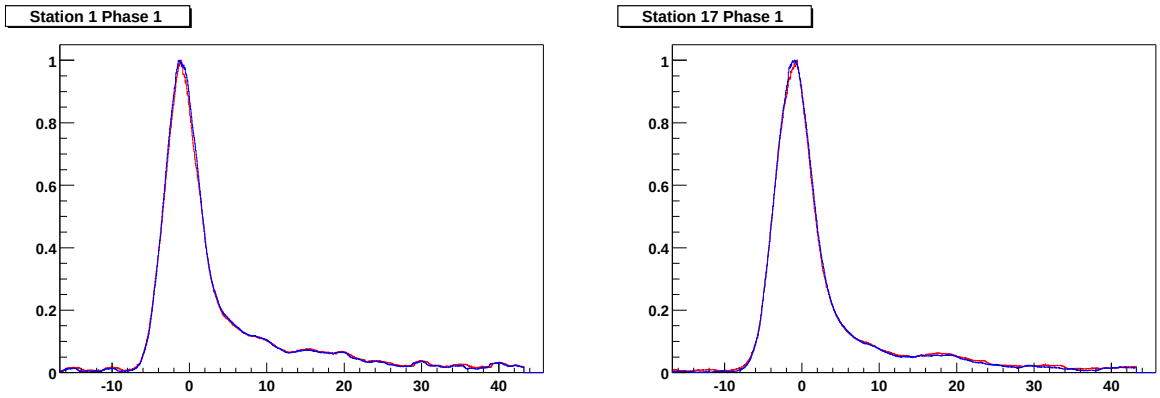
are reduced. At low energies, various experimental backgrounds (flashlets, protons, narrow pulses) start to influence the fit function, adding in the high energy data dilutes their significance.

As a final point with regard to fitting by energy band versus asymmetry-weighting, it should be recognized that combining the data allows the fit to extend to later times before the limit of Gaussian statistics per bin is reached. Although this contributes a fractionally insignificant amount of data, by fitting more g-2 periods, the correlation between the frequency and the phase is reduced, which results in yet another reduction in the statistical error on the frequency determination.

In the 2001 run, several major improvements in the way the experiment was run allowed the fit start times to be pushed much earlier. Most notably, the n values at which the storage ring was operated were chosen to avoid the overlap of primary CBO and g-2 harmonics. In addition, the scraping time was shortened, the pion flash was better regulated, detector gate-on times were earlier, and WFD thresholds were lowered. All of these changes in the *modus operandi* of the experiment result in an acceptable fit start times of 27 μs for a multi-parameter fit. By comparison, in the 2000 data the start times were extended to 50 μs in order to obtain an acceptable χ^2, and even then, most analyses required the inclusion of a 'mop' term to account for misconstructed slow effects. In the 2001 analysis, since some of the detectors still do not turn on until after 30 μs, it is necessary to analyze the data in two detector groups in order to take advantage of the early start times. For this reason, I deviate slightly from trying to fit all the detectors and energy bands simultaneously and fit two subsets of detectors in order to preserve the philosophy of minimizing the statistical error on the data.

2 Production

Other than a few minor redefinitions of fill quality status bits, the G2Too production of the 2001 data is identical to the 2000 production. In particular, the minimum area allowed for a T0 pulse was lowered in order to accommodate the smaller signal that arose from the degradation of the counter. Pulseshape libraries were constructed and compared for runs 9304, 9375, 9687, 10146, and 11229. The variation in the pulseshape was found to be small enough that the pulseshapes from run 11229 were used to fit the entire dataset. See Appendix A for a comparison of the pulseshapes from run 9304 and run 11229 for all detectors and WFD phases. Although these runs occur at the two extremes of the 2001 running period, for most detectors the pulseshapes are indistinguishable. Fig. 1 shows the two detectors with the largest deviations in the pulseshape. For a detailed discussion of the pulse fitting procedure, see Fred Gray's g-2 note 417 [1].



(a) Detector 1, phase 1 (b) Detector 17, phase 1

Figure 1: Comparison of the pulseshapes early (red) and late (blue) in the 2001 run.

3 Endpoint Calibration

3.1 Overview

Traditionally, the relative energy scales for the $g-2$ calorimeters have been determined by applying fits to the upper energy region of the relativistic Michel decay spectrum. While this procedure has proven to be robust enough to correct for changes in detector gain on a run by run basis, it is particularly sensitive to any effects that might distort the energy spectrum at high energies. In particular, pileup of decay electrons produces a greater distortion of the spectrum at high energies due to the dramatically decreasing singles rate as the energy approaches 3.1 GeV. One advantage of the Kolmogorov-Smirnov (KS) test is that it can be applied to a larger energy range than the standard linear extrapolation, thereby increasing the statistical power of the method. The test is also versatile in that it can be applied to smaller data sets from any available energy band without having to make assumptions about the shape of the underlying distribution. See Appendix B for a theoretical discussion of the KS test.

3.2 The Kolmogorov Shuffle

3.2.1 Notation and Overview

The basic problem of finding energy scale changes for the $g-2$ calorimeters can be treated as follows. For simplicity of notation, consider for now the case of a single calorimeter which measures a number of events, n_r, within a single run period. Let the observed energy distribution for that run period be given by,

$$N_r(E^0) = \{E_j^0 : j = 1,2,...,n_r\}, \tag{1}$$

where the superscript on E^0 indicates energy before any software calibrations have been applied.

The measured energy distribution can change from run to run due to instability in the high voltage control, degradation of the photocathode, changes in local magnetic fields, absorption of He in the photomultiplier tubes (PMTs), etc. All these effects result in a change in the PMT response or gain which affects the distributions in (1). In order to correct for these changes in scale, assume there exists a set of constants,

$$C = \{c_r : \forall\, r \in \text{runlist}\},$$

such that if the energies of the events in each run are scaled by these constants,

$$E_j^1 = c_r E_j^0, \tag{2}$$

then the resulting distributions,

$$N_r(E^1) = \{E_j^1 : j = 1,2,...,n_r\} = \{c_r E_j^0 : j = 1,2,...,n_r\},$$

will have the **same underlying parent distribution**.

Extending the notation to include all 24 detectors, the set of scaling constants can be written,

$$C = \{c_r^d : r = 9000, 9001, ...11384; d = 1,2,...24\},$$

where r ranges over run numbers from data acquired in 2001, and d ranges over all 24 detectors. Similarly, the number densities for each run and detector may be written,

$$N_r^d(E^1) = \{E_j^1 : j = 1,2,...,n_r^d\} = \{c_r^d E_j^0 : j = 1,2,...,n_r^d\}, \tag{3}$$

where n_r^d is the total number of events in the dth detector for a given run, r.

In the case of $g-2$, the parent distribution is the Lorentz-boosted Michel decay spectrum convolved with a detector response function and multiplied by an acceptance. The response function includes effects that persist in a stable manner from run to run, i.e. resolution of the detectors and shower leakage. Figure 2 shows a typical energy spectrum acquired by a single detector for a particular run. The electrical impulse for each event was recorded via 400 MHz waveform digitizers (WFDs) using an 8-bit analog-to-digital converter (ADC), and the energies were constructed offline by fitting each pulse to a known shape determined from the ensemble. The sharp edge at around 110 ADC counts is due to the hardware threshold of the WFDs. Events with energies smaller than the hardware threshold are only recorded if they happen to come within a few tens of ns of a pulse that has already triggered the WFD.

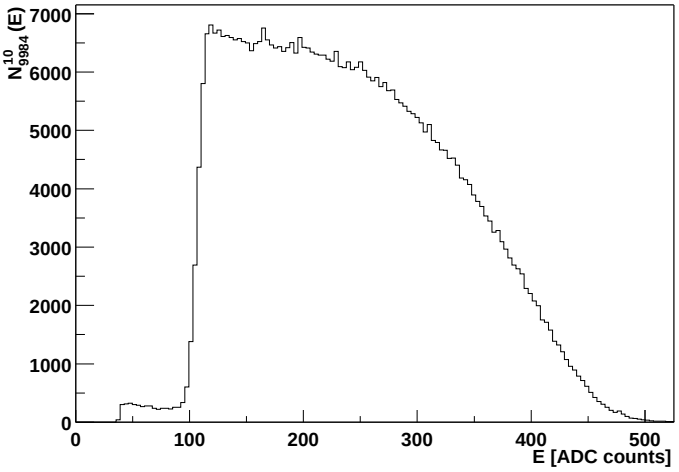


Figure 2: Histogram of measured energies for detector 10, run 9984.

The maximum theoretically allowed particle energy is 3.1 GeV which is why $N(E)$ drops rapidly for large E. The value in ADC counts that corresponds to a 3.1 GeV particle is defined to be the energy endpoint. An examination of Figure 2 reveals that for this particular run and detector the number of ADC counts corresponding to 3.1 GeV is around 460, however, the response function of the detector softens the cutoff such that an exact determination is not possible. Nevertheless, the goal of this study is to find the relative energy endpoint for each detector as a function of run number with the highest precision possible. Knowledge of the absolute conversion from ADC counts to particle energy is not as important as knowing the changes in scale from run-to-run and are not the emphasis of this study. For this reason, the scaling constants for each detector may yet need to be multiplied by an overall factor in order to set the absolute scale.

For a brief preview with specifics to follow, the general approach of the Kolmogorov shuffle is to generate a histogram from each $N_r^d(E^0)$ and compare it to a standard reference using the KS test. Once d_K has been calculated, the contents of the histogram are shifted by an integer number of bins and then d_K is recalculated for the same reference. This procedure is repeated for as many as 20 shifts to the left and 20 shifts to the right. The resulting d_K are plotted as a function of the shift and the minimum is used to calculate the optimal scaling constant, c_r^d. Since the act of shifting the data is a translation of the energies, while gain changes are a scaling effect, the resulting scale factors are only approximately correct. Therefore, the method must be applied iteratively until the optimal shift found from translating is sufficiently small.

3.2.2 First Iteration

The first step in applying the method, is to construct an initial set of energy densities, $N_r^d(E^0)$, from the processed data. Pileup-subtracted spectra were generated for each run and detector using symmetric time windows to construct the pileup events. To further reduce distortions to the ideal spectra from background and incomplete pileup subtraction, only events 50 μs after injection or later were included. Since it is more intuitive to work in units of energy rather than ADC counts, the initial energy endpoint was assumed to be 440 ADC counts, a reasonable starting guess based

on the calibration constants from prior years. Figure 3 shows the data for one detector summed over all runs in the runlist before and after pileup subtraction. Accurate pileup subtraction is important to ensure that trends with run number are not merely due to fluctuations in the AGS intensity throughout the year. For this study the runlist consists of approximately 1300 runs where short and special runs have been eliminated. Laser runs are present, but fills in which the laser fired have been removed. Any other fills marked as bad due to quad sparks, short quad on time, missing T0, etc. have also been removed using the G2Too status bits determined during data processing.

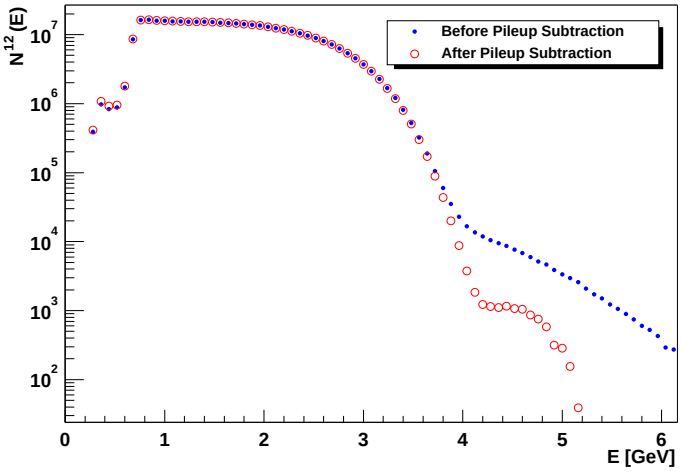


Figure 3: Energy spectrum for detector 12 summed over all runs in runlist before and after pileup subtraction.

A standard reference histogram was created by summing the individual $N_r^d(E^0)$ over all detectors and runs in the runlist, excluding detector 20 which has a dramatically different energy spectrum due to upstream material in its particle path. Let the sum of number densities over all runs for a particular detector be denoted by,

$$N^d(E^0) = \sum_{r \in \text{runlist}} N_r^d(E^0), \tag{4}$$

where the lack of a subscript on N indicates a summation over runs. The standard reference used for this iteration may then be written,

$$N_{\text{ref}}(E^0) = \sum_{d \neq 20} N^d(E^0).$$

Since no attempt has yet been made at an offline calibration, the sum of these energy densities represents the average of the ensemble. It should not be expected for the sum to be functionally equivalent to the individual histograms, since the shape of the energy distributions are known to vary by detector, but for this first rough iteration it is of little importance. In fact, it is beneficial for the first iteration to compare the data from all detectors to the same standard since this sets the detector to detector relative scales for all subsequent iterations. The standard reference used for this iteration is shown in Figure 4. When comparing each $N_r^d(E^0)$ to the standard, the whole range of E^0 cannot be used since there will be large deviations from the standard at energies below 1

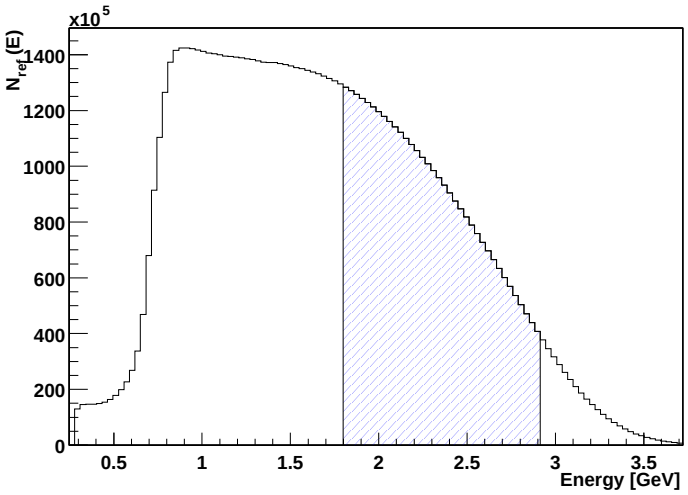


Figure 4: Standard reference, $N_{\rm ref}(E^0)$, used for first iteration. The shaded region shows the actual range of E^0 used in the calculations.

GeV where the hardware thresholds vary substantially from detector to detector. Energies above 3.1 GeV are also suspect since pileup events tend to have a greater effect in this region. Therefore, the comparison range is chosen rather conservatively to lie between 1.8 and 2.9 GeV, as indicated by the shading under the curve. The cdf used for the comparison is calculated only in this region.

The comparison is performed by taking an individual $N_r^d(E^0)$, truncating it in the same energy range as the standard reference, and finding its cdf. The sample cdf is compared to the cdf of $N_{\rm ref}(E^0)$ in order to find the Kolmogorov distance, d_K. The contents of the histogram of $N_r^d(E^0)$ are then shifted by an integer number of bins without changing the scale of the x-axis. This means that some events flow into the comparison region while others flow out, therefore a new truncation is performed from which a new d_K is derived. Figure 5 shows the d_K obtained as a function of the amount of shift applied as well as the correspond p-value calculated from equation (44). For this iteration, a 31 MeV bin width was used in the original construction of the energy distributions. The x-axis has been converted into energy units using the approximate conversion (440 ADC counts = 3.1 GeV) as before. From the plot it is apparent that a shift of about 270 MeV is required to make the shape of $N^3_{10147}(E^0)$ most closely approximate the shape of the standard reference. The actual minimum is determined by finding the center of the minimum bin and its amplitude along with the bin centers and amplitudes of the two neighboring bins. These three points define a parabola from which the minimum is analytically derived.

In principle, one could instead find the optimum shift by finding the maximum p-value returned by the KS test. However the difficulty is demonstrated in the second half of Figure 5. The p-value returned by ROOT is zero as soon as the probability falls to less than a part in 10^{15}, which is the case as soon as $N_r^d(E^0)$ is shifted by even one bin (31 MeV) from the optimum.

Figure 6 shows a histogram of the optimal shift required for all $N_r^d(E^0)$, detector 20 excluded. The spiky nature of the distribution is a testament to the overall stability of the calorimeters throughout the year. To make each detector closely approximate the ensemble a certain shift is required. That shift, although detector dependent, changes by very little as a function of run number. Since the standard reference was calculated from the sum over all $N_r^d(E^0)$, one would expect

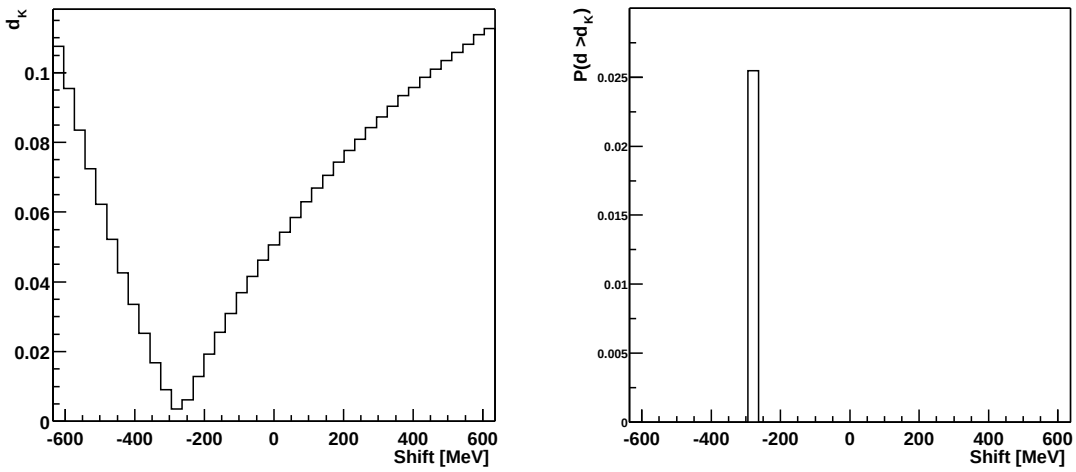


Figure 5: Kolmogorov distance and probability obtained when comparing $N^3_{10147}(E^0)$ to $N_{\text{ref}}(E^0)$ as a function of translation of $N^3_{10147}(E^0)$. The comparison was calculated in the 1.8 to 2.9 GeV range.

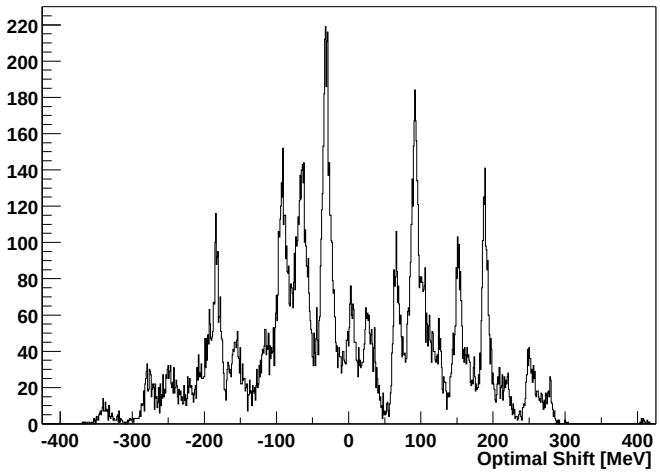


Figure 6: Histogram of energy shifts required to minimize d_K for all $N^d_r(E^0)$, detector 20 excluded.

the distribution of optimal shifts to be fairly symmetric about zero. In fact, the mean of the distribution in Figure 6 is -6.5 MeV despite shifts for some runs and detectors being as large as 400 MeV.

Once the optimal shifts have been obtained, the scaling constants must be calculated. Herein lies the ambiguity with trying to treat a translation as if it were a scaling phenomenon. Let the optimal shift for a particular $N^d_r(E^0)$ be denoted by $\Delta\varepsilon^d_r$. The energy transformation from translating $N^d_r(E^0)$ is,

$$E^1 = E^0 + \Delta\varepsilon^d_r. \tag{5}$$

By contrast, letting,

$$c^d_r = 1 + \varepsilon^d_r,$$

in equation (3), the energy transformation from scaling may be written,

$$\begin{aligned} E^1 &= c_r^d E^0 \\ &= (1+\varepsilon_r^d)E^0 \\ &= E^0 + \varepsilon_r^d E^0. \end{aligned} \tag{6}$$

Equating equation (5) with (6) and solving for c_r^d we find,

$$c_r^d = 1 + \frac{\Delta\varepsilon_r^d}{E^0}. \tag{7}$$

The shift is a fixed quantity and the scale factor must also be energy independent so the question becomes what to use for E^0. Since the comparison was performed in the energy range from 1.8 to 2.9 GeV,

$$E^0 \in [1.8, 2.9].$$

Which implies c_r^d must be chosen from the interval,

$$c_r^d \in \left[\frac{\Delta\varepsilon_r^d}{2.9}, \frac{\Delta\varepsilon_r^d}{1.8}\right]. \tag{8}$$

From Figure 6 we can see the largest shift is about 400 MeV. For this extreme example we get a continuum of c_r^d values ranging from 1.14 to 1.22, therefore the discrepancy is actually rather small. The value taken for the scaling constant is determined by letting $E^0 = 2.35$ GeV, the value at the middle of the comparison range, with the understanding that another iteration is required.

The proof that that the discrepancy must disappear with successive iterations is rather simple. Define the ambiguity on c_r^d to be the difference of the interval in 8,

$$\begin{aligned} \Delta c_r^d &= \left|\frac{\Delta\varepsilon_r^d}{E_{\min}} - \frac{\Delta\varepsilon_r^d}{E_{\max}}\right| \\ &= \left|\Delta\varepsilon_r^d\right| \times \text{constant} \end{aligned}$$

As long as the optimal shifts get smaller with each iteration, so too will the error on c_r^d.

Examples of the new endpoints obtained from the first iteration of c_r^d values are shown in Figure 7 for several detectors. The y-axes have been set to display a range +/- 4% of the median value. The first plot shows detector 5, where the effect of sitting near the injection region of the ring is clearly evident by the steady 4% decline in gain as the year progressed. This drop is clearly evident at various levels in all of the detectors that view the initial flash of particles at injection. The next plot of Detector 12 shows the last major iteration of fine gain adjustments at around run 9750. Detector 16 was the subject of a few repairs early on followed by a unexplained drop at around run 10500. The endpoint plot for detector 19 shows a typical well behaved detector where the resolution of the Kolmogorov shuffle is beginning to become apparent by the clear variations at later runs.

As a final means of quantifying the resolution of the 1st iteration, the point-to-point difference between successive endpoints divided by $\sqrt{2}$ was calculated for all detectors (excluding 20) and is plotted in Figure 8. If the intrinsic gain of the detectors was absolutely stable from run to run, then

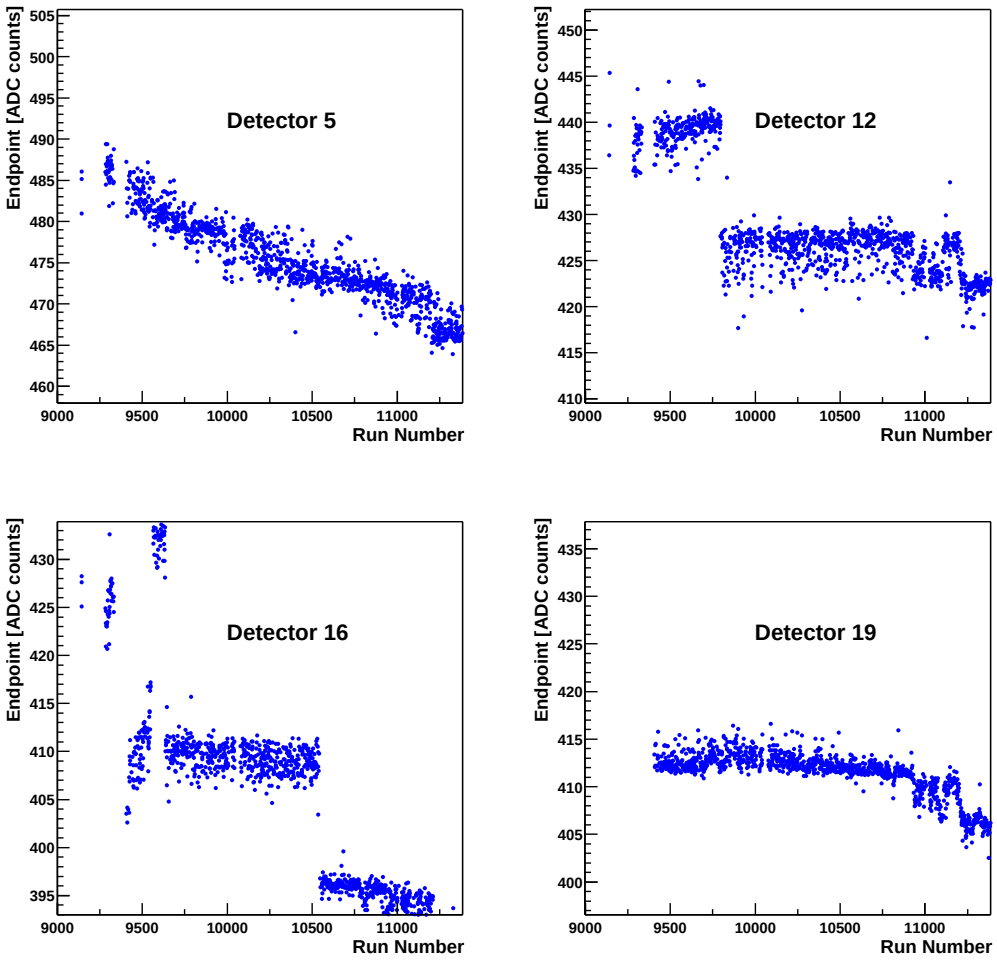


Figure 7: Endpoints obtained from the new values of c_r^d versus run number for select detectors.

the width of this difference plot would be a direct measurement of the resolution inherent to the procedure. Since the gains are not absolutely stable, there is also an irreducible term that adds to the width. Therefore, the width of this distribution provides an upper limit on the resolution from which we can compare results from subsequent iterations. Unfortunately, the distribution is not Gaussian so for comparison purposes the resolution will be defined by the RMS of the distribution. For the first iteration we get an upper limit on the resolution of the procedure of about 9 MeV or 0.29%. Note that the differences shown in the histogram were converted to energies using the new endpoint values.

3.2.3 Second Iteration

The first step in the second iteration is to go back to the raw data and rebuild the new energy densities for each run and detector with the scaling constants derived from the first iteration. This time the binning was chosen to be 10MeV, smaller by about a factor of three than the first iteration. For this iteration a standard reference was created for each of the 24 detectors as in equation (4).

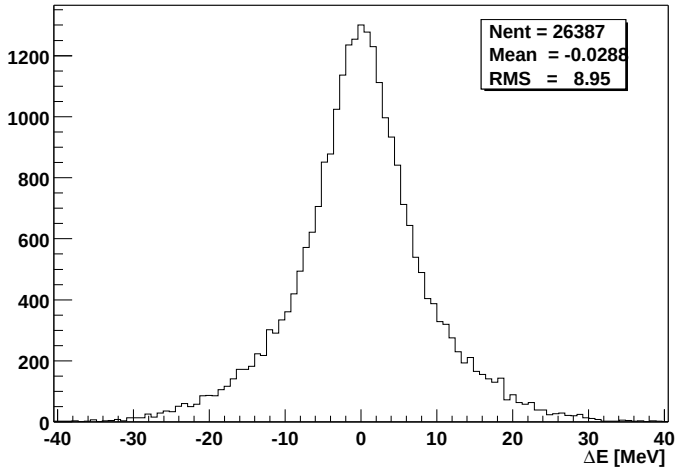


Figure 8: Endpoint resolution obtained after the 1st iteration.

Figure 9 shows the reference histograms normalized by the number of counts in the comparison region,

$$f^d(E^1) = \frac{N^d(E^1)}{\int_{1.8}^{2.9} N^d(E^1) dE^1}.$$

The need for separate references by detector is explicitly shown by changes in slope from detector to detector. The slope is a measure of how much mass is in the particle path as evidenced by the steadily decreasing slope of $f^5(E^1)$, $f^6(E^1)$, $f^7(E^1)$, $f^8(E^1)$ where the decay electrons for each detector see progressively more of the kicker plates as the detector number increases. The slope of the energy spectrum in detector 20 has been so affected by the mass of the traceback and vacuum chamber, that it cannot reasonably be calibrated with respect to the other calorimeters without a model. The fact that all of the other graphs intersect in the region between 2.3 and 2.6 GeV is a direct result of the first iteration attempting to make the curves look as close as possible to $N_{\text{ref}}(E^0)$.

In order to increase the statistical power, the comparison range was extended to include the 1.6 to 1.8 GeV data. It should also be noted when converting from ADC counts to energy, it is no longer appropriate to use the fixed conversion (440 ADC counts = 3.1 GeV). The new conversion is now determined by the endpoint data, as shown in Figure 7, for each run and detector.

To summarize the differences with this iteration:

A. New $N_r^d(E^1)$ were built with scaling constants, c_r^d, determined from the first iteration.

B. $N_r^d(E^1)$ were built with finer bin size of 10 MeV.

C. Standard references were constructed on a per detector basis according to eq. (4).

D. The comparison range was extended down by 200 MeV to $[1.6, 2.9]$.

The Kolmogorov distance and probability obtained when comparing $N^3_{10147}(E^1)$ to $N^3(E^1)$ are plotted as a function of applied shift in Figure 10. This is the same run and detector as shown in the first iteration where a 270 MeV shift was required to minimize d_K. Even though the reference has

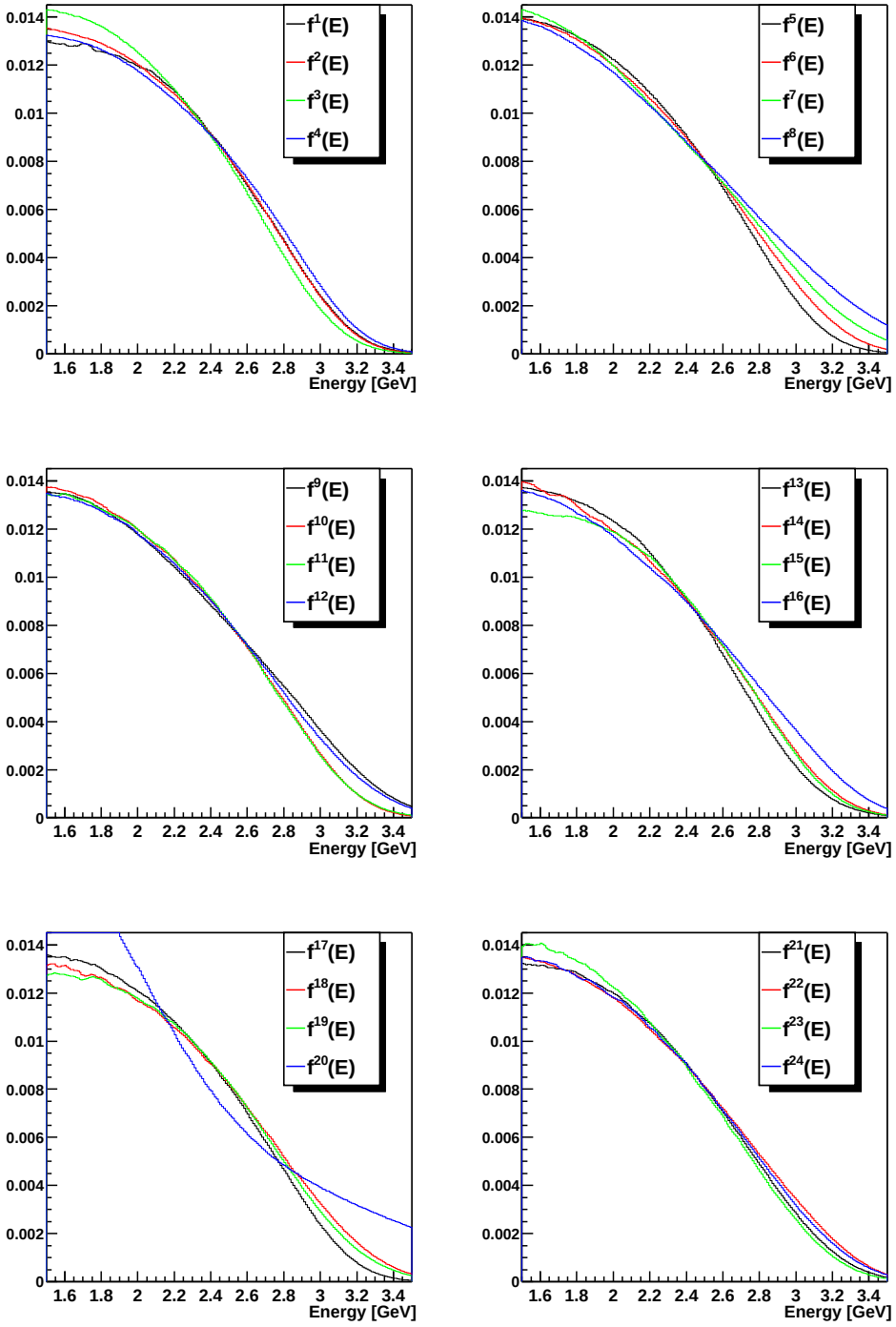


Figure 9: Reference distributions generated by detector and used for the 2nd iteration.

changed, the optimal shift for the second iteration is still quite small at around 11 MeV. Note that the probability is still not very useful with only two bins having a non-zero probability. Again a parabolic minimum was found using the three smallest d_K bins. The results for all runs and detectors are shown in Figure 11(a). Note the dramatic difference between this plot and the distribution

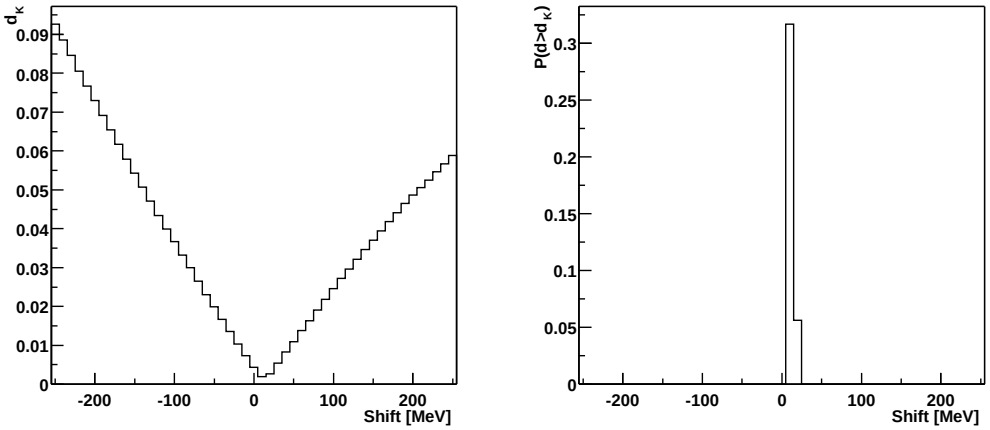
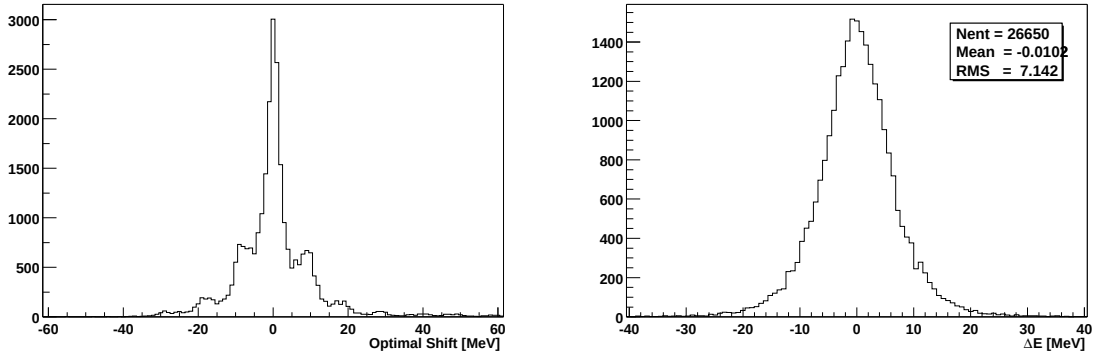


Figure 10: Second iteration Kolmogorov distance and probability obtained when comparing $N^3_{10147}(E^1)$ to $N^3_{\text{ref}}(E^1)$ as a function of translation of $N^3_{10147}(E^1)$. The comparison was calculated in the 1.6 to 2.9 GeV range.



(a) (b)

Figure 11: (a) Histogram of energy shifts required to minimize d_K for all $N^d_r(E^1)$, detector 20 excluded. (b) Endpoint resolution obtained after the 2nd iteration.

of shifts required in the first iteration shown in Figure 6. As in the 1st iteration, the point-to-point scatter in the endpoints is plotted in Figure 11(b) where it is apparent that the endpoint resolution has improved to 7.1 MeV.

3.2.4 Third Iteration

In the final iteration, new $N^d_r(E^2)$ were constructed using the product of scaling constants found in the last two iterations to determine the energies of each event. The bin width was further reduced to 1 MeV. Once again the standard references were built for each detector according to equation (4) and the optimal shift was found for each individual energy density. The comparison range was left unchanged at $[1.6, 2.9]$. An example of the Kolmogorov distance and probability versus applied shift is shown in Figure 12 for the same run and detector as shown in the last two iterations, $N^3_{10147}(E^2)$. The optimal shift is less than 1 MeV indicating that the best resolution has already been obtained in the last iteration.

The distribution of optimal shifts obtained is slightly better than in the last iteration, as shown in 13(a), where it is clear that the outliers in Figure 11(a) have been drawn into the central peak. However, the RMS of the point-to-point resolution is 7.4 MeV which is slightly worse than the 7.1 MeV attained in the second iteration. Part of the problem stems from the small 1 MeV binning that was used. This binning occasionally resulted in small local variations in the distributions of d_K values versus shift that would cause the minimum to be mislocated by a few MeV. Rebinning the shift data into 5 MeV bins actually improved the resolution from 7.4 to 7.2 MeV. Either way, the resolution did not improve in the third iteration indicating that the scaling constants have been determined as precisely as possible given the statistics of the data sets.

3.3 Further Considerations

3.3.1 Optimization

It is useful to explore the ultimate resolution of the procedure as a function of start time and the range of energies used to make the comparison. Although an energy endpoint resolution of 7 MeV out of 3.1 GeV is more than adequate for determining changes in gain on a run-to-run basis, the true benefit of optimizing the procedure lies in its subsequent application where shifts in the energy spectrum versus injection time are studied.

Rather than rebuild histograms and perform a 4th iteration, the 3rd iteration was repeated while changing the parameters of interest. In particular, the start time for histogram filling, the minimum energy used in calculating the cdf, and the maximum energy were varied. The figure of merit is again taken to be the RMS of the point-to-point variations in the endpoint. For the first step in the optimization, the start time for filling the $N_r^d(E^2)$ histograms was moved from 48 μs back to 30.6 μs in steps of 1 $g-2$ period. This is close to the earliest time after injection that all calorimeters are turned on. The energy range used for the calculation of the cdfs was 1.8 to 3.1 GeV. The results are shown in the first frame of Figure 14. Essentially, going to earlier start time results in a larger statistical ensemble and improves the resolution. The resolution gets slightly worse at the

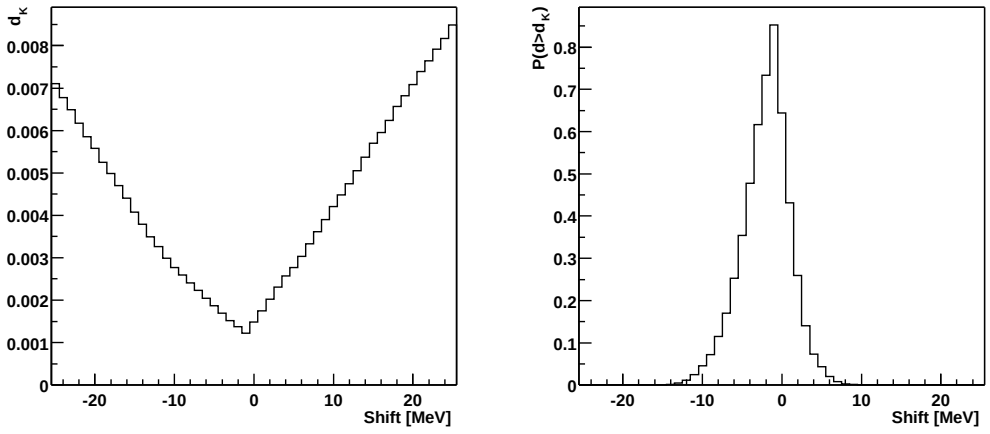
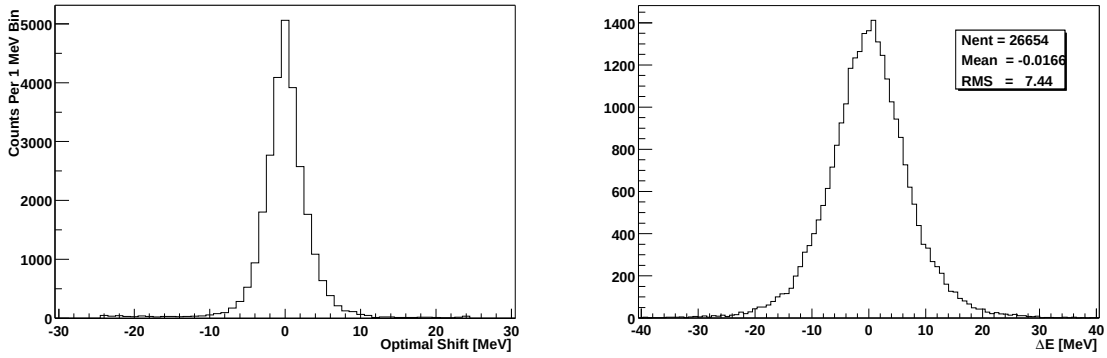


Figure 12: Third iteration Kolmogorov distance and probability obtained when comparing $N^3_{10147}(E^2)$ to $N^3_{\rm ref}(E^2)$ as a function of translation of $N^3_{10147}(E^2)$. The comparison was calculated in the 1.6 to 2.9 GeV range.



(a) (b)

Figure 13: (a) Histogram of all the run lengths in 2001, measured in AGS cycles. (b) Zoomed in region. The vertical, dashed line indicates where the cut on runlength was placed

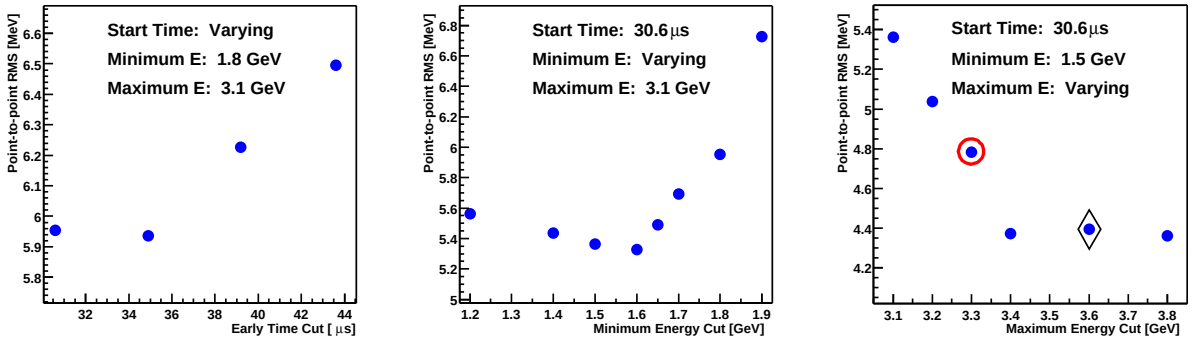


Figure 14: Plots of the resolution obtained via the KS shuffle versus start time and energy cut.

very earliest data point which indicates that perturbations to the energy spectrum at early times are canceling the benefit of higher statistics. With this optimization, the resolution drops from a little over 7 MeV to about 6 MeV. For subsequent parameter scans, the 30.6 μs start time was used.

For the next optimization, the minimum energy used in the comparison was varied. Lowering the minimum energy at first results in a better resolution, but at about 1.6 GeV the resolution starts slowly degrading. This can be attributed to a balance between higher statistics and the flat nature of the energy spectrum below 1.6 GeV as shown in Figure 4. Since the whole method is based on translating the energy spectrum, the KS shuffle is insensitive to regions where the distribution is flat. For instance, sliding a perfectly uniform distribution across the comparison region would result in the individual snapshots having identical distributions regardless of the shift applied. For the subsequent optimizations, the lower energy cut was set at 1.5 GeV which lowers the resolution from 6 to 5.4 MeV.

Next, the maximum energy was allowed to increased which at first results in a fairly dramatic improvement that flatlines at about 3.4 GeV. This is simply because there is hardly any data beyond 3.4 GeV to make the resolution either better or worse. One could argue that there should not be particles above 3.1 GeV other than pileup events, but this is not quite true since there has been no attempt at an absolute calibration. The assumption in the first iteration that 440 ADC counts is about 3.1 GeV is what sets the scale for all subsequent iterations and could easily be wrong by

10%. With this optimization, the resolution is reduced to 4.3 MeV out of 3.1 GeV.

As mentioned previously, the use of bins smaller than the resolution of the method results in sporadic errors in which the algorithm settles in a local minimum rather than finding the true shift which minimize d_K. To illustrate this, the two data points enclosed by the red circle and the black diamond in Figure 14 were repeated using a larger bin size. The results are shown in Figure 15 where as before, a 200 keV improvement is obtained. This brings the optimal resolution down to 4.2 MeV out of 3.1 GeV, or 0.135%, which is in line with what is required to study the gain stability versus start time.

3.3.2 Pileup Sensitivity

It is clear that including data from earlier times and extending the energy range results in higher precision in the endpoint determination, however the question of whether a systematic bias has been introduced remains. In particular, since pileup events occur in proportion to the square of the data rate and the data rate increased by a factor of two or so in the relevant run range, a systematic bias might have been introduced that varies from early in the year to later in the year. Each parameter extension can enhance a pileup effect in its own unique way. Moving to earlier times where the rate is higher creates relatively more overlapping pulses in the data set. The comparison range can be more sensitive to pileup at high energies where pileup events start to dominate the spectrum, or at energies lower than twice the hardware threshold (1.8 GeV) where the pileup algorithm used is known to be inefficient.

A priori, one would expect the bias to be small based on the fact that run to run variations in the rate would have caused the resolution to decrease rather than improve. Furthermore, by design the Kolmogorov shuffle is only sensitive to relative changes in the pileup fraction. The method does not care whether the underlying distribution has a pileup component, since gain affects pileup energies

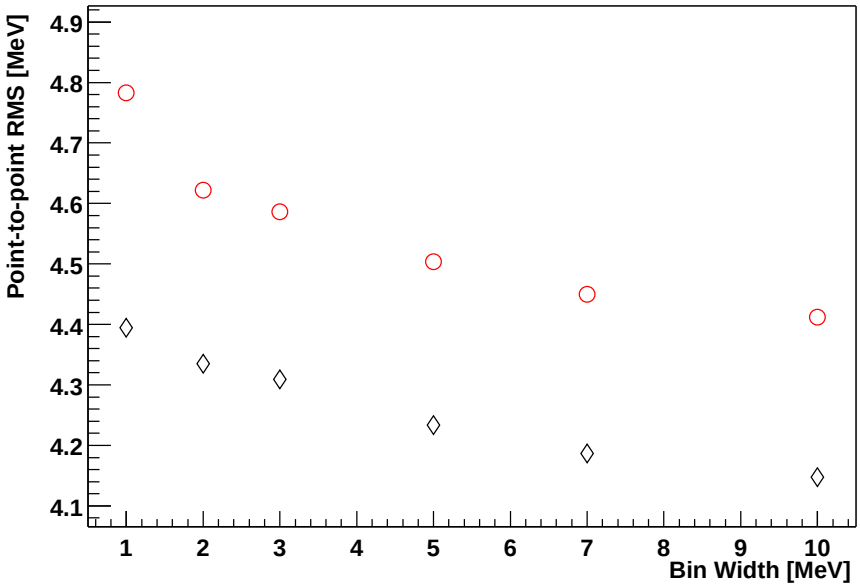


Figure 15: Improvement in resolution by implementing the last iteration of the KS shuffle with a larger bin width.

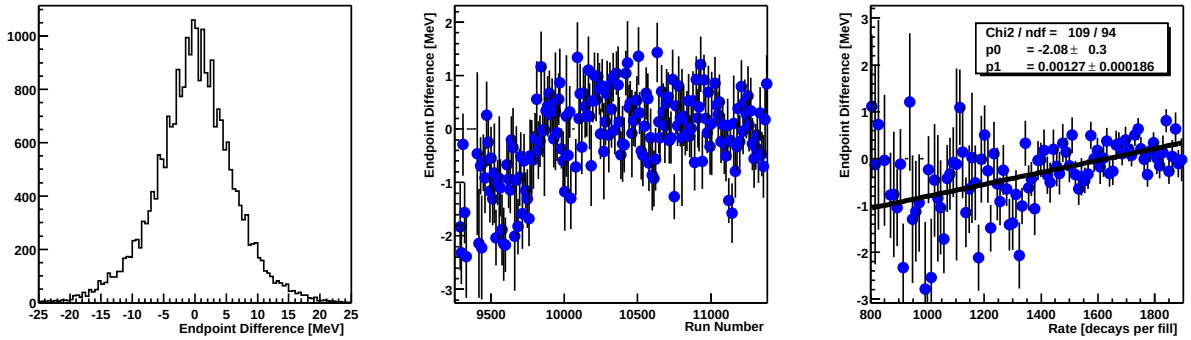


Figure 16: Plots of the difference in the endpoint values obtained via the KS shuffle when the energy range is extended and the start time moved forward.

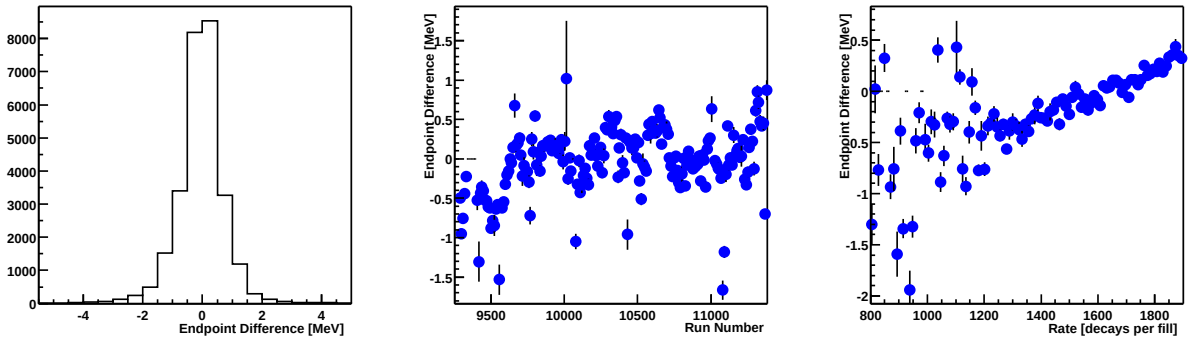


Figure 17: Plots of the difference in the endpoint values obtained via the KS shuffle with and without pileup subtraction.

with the same scaling as non-pileup events. It is only when the shape of the distribution changes as the relative amount of pileup increases or decreases that an effect could be present. To check the sensitivity to pileup, the endpoints obtained with conservative parameters (48.0 μs, 1.8 GeV, 2.9 GeV) were compared with those obtained with an earlier start time and broader comparison range (30.6 μs, 1.5 GeV, 3.6 GeV). The differences in the endpoints obtained are plotted in a histogram, and as a function of run number and rate in Figure 16. The rate was quantified by finding the average number of decay electrons per fill for each run. At lower run numbers there is a systematic shift of about 1 MeV that is confirmed by a fit to the difference versus rate. Therefore, one must conclude that the reduction in the spread of the endpoints from 7.4 MeV to 4.4 MeV comes at the cost of a 1 MeV bias. Although, it is apparent that after run 9800 the rate and consequently the relative endpoints are fairly stable.

Since the datasets compared above have a substantial amount of non-overlap it is perhaps more informative to compare the difference between endpoints using the same dataset with and without pileup subtraction. For this comparison a start time of 26 μs was used and an energy range from 1.5 to 3.2 GeV. The results are show in Figure 17. Since the overlap of the datasets only differs by the pileup events, the distribution of differences is much narrower. However, the systematic difference is much clearer both in the fluctuations with run number and in the overall trend in the endpoint difference versus rate. The conclusion is striking, if pileup subtraction is completely neglected,

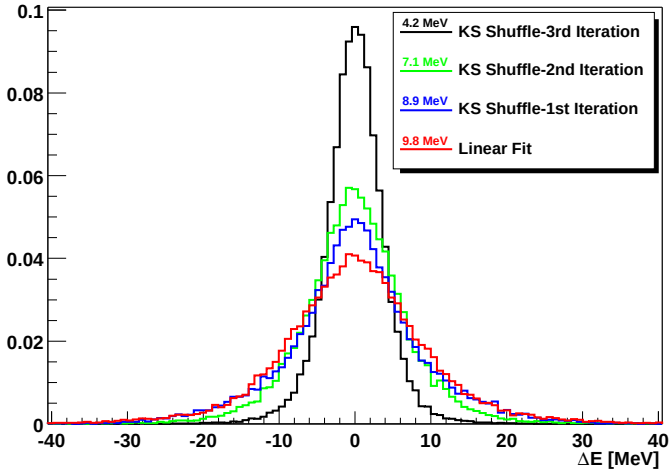


Figure 18: Comparison of energy resolution for the standard linear fit and the KS test.

then the resulting bias in the endpoint is still at the 1.5 MeV level. Since the vast majority of the data comes from fills with greater than 1200 events per fill, the effect is actually smaller.

3.3.3 Comparisons with G2OFF Endpoints

An independent effort to determine the endpoints for the 2001 run has been conducted by Xiaobo Huang using the standard method of fitting the linear region of the energy spectrum and extrapolating to the x-axis in order to find the energy endpoint. This data provides for a nice comparison of the resolutions of the two methods, as well as a general comparison of the G2Too and G2OFF productions. Using the same point-to-point resolution, the RMS of endpoints obtained by the linear fit method is about 9.8 MeV. Figure 18 shows the normalized histograms of the point-to-point differences for the linear fit method along with various iterations of the Kolmogorov shuffle.

Figure 19: Endpoints for detector 1 showing PMT sensitivity to field changes resulting from beamline instabilities. Neighboring detectors are also plotted for comparison.

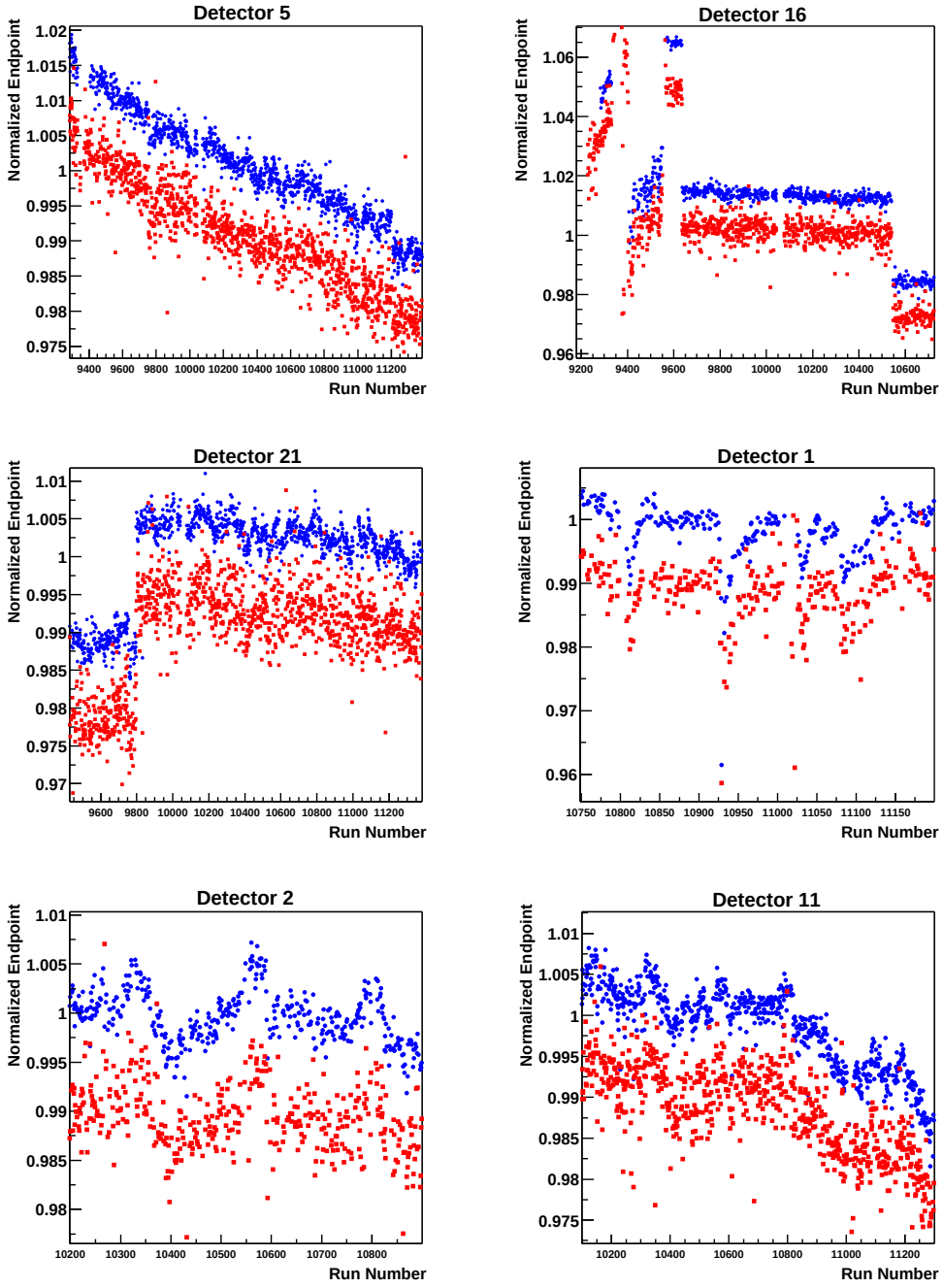


Figure 20: Comparison of endpoints obtained via the KS shuffle (blue circles) and the standard linear extrapolation (red squares) for select detectors.

Figure 20 shows snapshots of the endpoints for various detectors plotted versus run number. Blue circles show the results of the optimized Kolmogorov method, while red squares indicate the corresponding endpoints from the analysis of the G2OFF data. The data for each detector has been normalized by the endpoint averaged over all runs, and the G2OFF data has been reduced by 1% for demonstration purposes. As shown in the first 3 frames, the agreement in the productions is

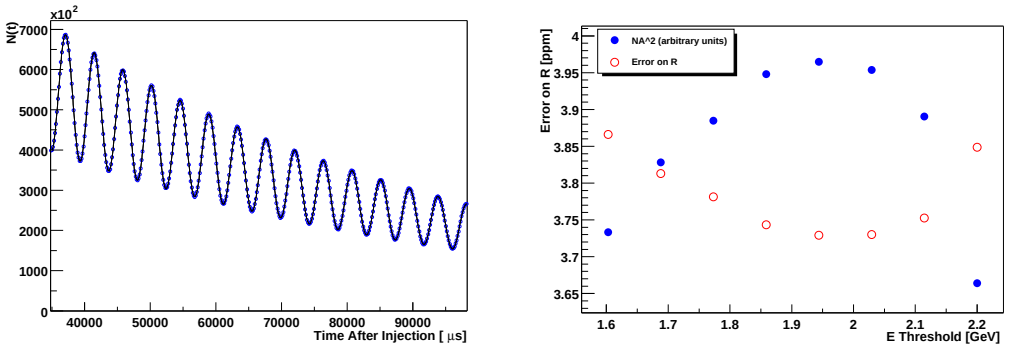
impressive. Both see the same degradation of detector 5, the intermittent jumps in detector 16, and the stability of detector 21. The last 3 frames show some of the detailed fluctuations made obvious by the enhanced resolution of the Kolmogorov method.

As an interesting aside, it was noticed that the endpoints plotted versus run for detector 1 look different than the other detectors in that there are sharp dips in the gain. It was hypothesized that such behavior could be due to changes in the local field surrounding the PMTs on that station which is the closest to the inflector and beamline. Tracking down the runs in the logbook revealed that quite a few of the sharp dips can in gain can be attributed to problems with the last set of quads as shown in Figure 19. The endpoints versus run for the two neighboring stations are also plotted where a slight perturbation can be seen in detector 24, but detector 2 appears to be completely insensitive.

3.3.4 Final Calibration

The Kolmogorov shuffle has provided a precise determination from run to run of the relative endpoints for each detector, but there still exists the issue of how to set the energy scales from one detector to the next, and then how to calibrate with respect to the true energy of the particles. Finding the absolute scale from the Michel spectrum would involve a rather complex model of the relative amounts of preshowering in mass upstream of each calorimeter, and in the end knowledge of the absolute scale does not turn out to be all that useful.

Instead of attempting an absolute calibration, it is more important to maximize the statistical power of each detector. When fitting for $g-2$ in a threshold based analysis, there are two competing factors that determine where to set the software threshold. Extending the range of energies included in the sample to lower energies increases the statistics of the sample, however low energy decay electrons have a greatly reduced asymmetry which dilutes the average asymmetry of the sample. Theoretical predictions [2, 3] show that the maximum statistical power is obtained by setting a threshold of about 1.8 GeV. Although the theoretical prediction does not include detector considerations that perturb the ideal energy distribution, it is still useful to find from the data the lower energy cut where NA^2 maximizes and define the energy at that point to be 1.8 GeV. With this definition, one can implement a 1.8 GeV threshold in subsequent analysis and be assured that the data is being used to its full extent. Alternatively, one could of course define a threshold for every



(a) (b)

Figure 21: (a) Time distribution for detector 12 with 5 parameter fit. (b) NA^2 and the error on R plotted versus lower energy cut.

Detector Number	NA^2 Maximum [Counts x 10^6]	NA^2 E Threshold [GeV]	Error on R Minimum [ppm]	Error on R E Threshold [GeV]	Error on R E Threshold/1.8
1	20.67	1.94	3.32	1.96	1.091
2	23.41	1.91	3.11	1.94	1.080
3	23.22	1.86	3.13	1.90	1.054
4	23.29	1.91	3.12	1.95	1.081
5	21.02	1.89	3.29	1.92	1.069
6	19.70	1.92	3.40	1.96	1.087
7	18.70	1.94	3.49	1.99	1.103
8	16.12	2.01	3.77	2.09	1.163
9	17.54	1.95	3.62	2.03	1.128
10	19.20	1.90	3.45	1.93	1.073
11	21.06	1.91	3.29	1.95	1.081
12	19.97	1.94	3.38	1.97	1.097
13	21.07	1.89	3.29	1.93	1.070
14	22.36	1.89	3.20	2.05	1.137
15	20.17	1.95	3.37	1.98	1.101
16	21.87	1.95	3.22	2.00	1.111
17	19.43	1.90	3.43	1.93	1.073
18	17.38	1.95	3.62	1.99	1.104
19	20.35	1.95	3.35	1.98	1.099
20	7.08	1.74	5.72	1.76	0.978
21	23.29	1.93	3.12	1.97	1.095
22	22.32	1.95	3.19	1.99	1.106
23	19.78	1.89	3.40	1.93	1.070
24	19.89	1.93	3.39	1.97	1.095

Table 1: Lower energy cuts required to maximize NA^2 and minimize the error on R for all 24 detectors. The last column shows the final calibration constant used to define the relative scale from detector to detector.

detector, but since the final calibration can not be performed absolutely it might as well be used to simplify the thresholds.

Using the endpoints obtained in the last iteration of the Kolmogorov shuffle, time distributions of the detected particles were created for various energy ranges. The upper limit was kept at 6.2 GeV while the lower threshold was stepped from 1.6 to 2.8 GeV in steps of 100 MeV. The time spectra for each detector were summed over all runs and fit in the range from 35 to 350 μs using a five parameter function of the form,

$$N(t) = Ne^{-t/\tau}\left[1 - A\cos\left(\omega t + \phi\right)\right]. \tag{9}$$

Figure 21(a) shows the time distribution constructed from the data and the fit for a typical detector (12 in this case). The reduced χ^2 of the fits were typically 1.2 with 2000 degrees of freedom. A rather poor fit, but good enough for the task at hand. From the fits to the data, NA^2 and the error on R were obtained as a function of the energy threshold. The results for detector 12 are

shown in Figure 21(b). As expected, the error on R is minimized when NA^2 is at a maximum.

The maximum NA^2 and minimum errors on R were found using a parabolic interpolation of the 3 central points for each detectors. The maximum NA^2 and minimum error on R, along with the corresponding energy thresholds are shown in Table 1. In general, the energy threshold that minimizes the error on R is systematically higher by about 40 MeV than the threshold that maximized NA^2. Since the goal is really to minimize the error on R in the analysis, the thresholds determined from the R data were used to define an overall scaling constant for each detector. The new endpoints were applied and the final results are summarized in Figures 22 and 23 where the error on R versus lower energy cut and the resulting energy spectra are shown. As an aside, the data from the 5 parameter fit was also used to find the time offsets required to align the $g-2$ phase for each detector. The resulting phase versus lower energy cut is summarized in Figure 24 for each detector.

3.4 Endpoint Summary

Combining the KS statistic with the sliding histogram method has resulted in a precise determination of the relative energy scales of the $g-2$ calorimeters to 0.14% for sample sizes typical of a single detector in a 2001 run. The resolution is considerably better than with traditional linear fitting and is sufficiently insensitive to effects from pileup. A final calibration was determined by finding the lower energy cut that yields the smallest error on R when fitting with a 5 parameter function, and then redefining the energy scale such that the cut is placed at 1.8 GeV. Future application to determining the gain versus time in spill looks promising, and extensions of the test could be developed to determine the efficiency for pileup construction.

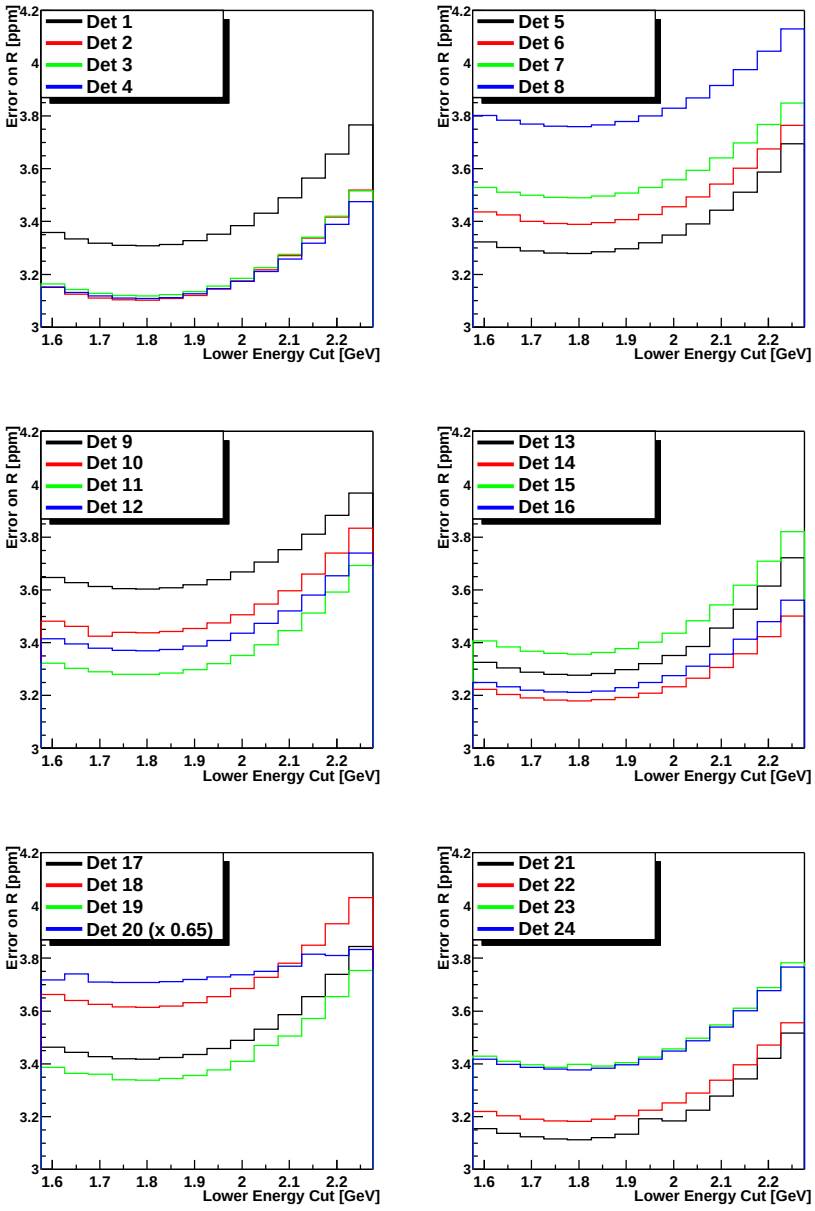


Figure 22: Error on R versus lower energy cut for each detector after applying scaling constants that define the minimum to be at 1.8 Gev.

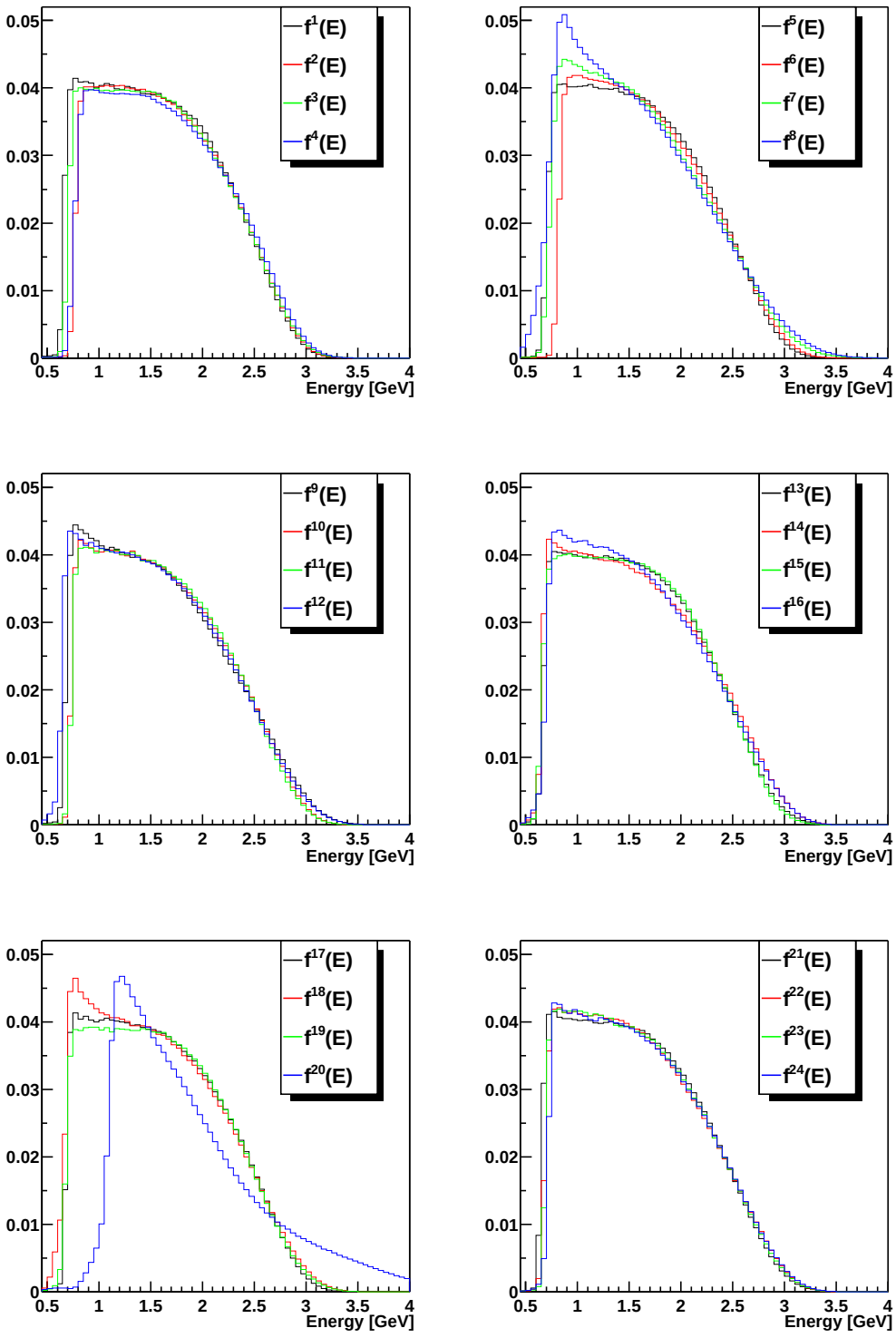


Figure 23: Energy spectrum for each detector with final scaling constants applied.

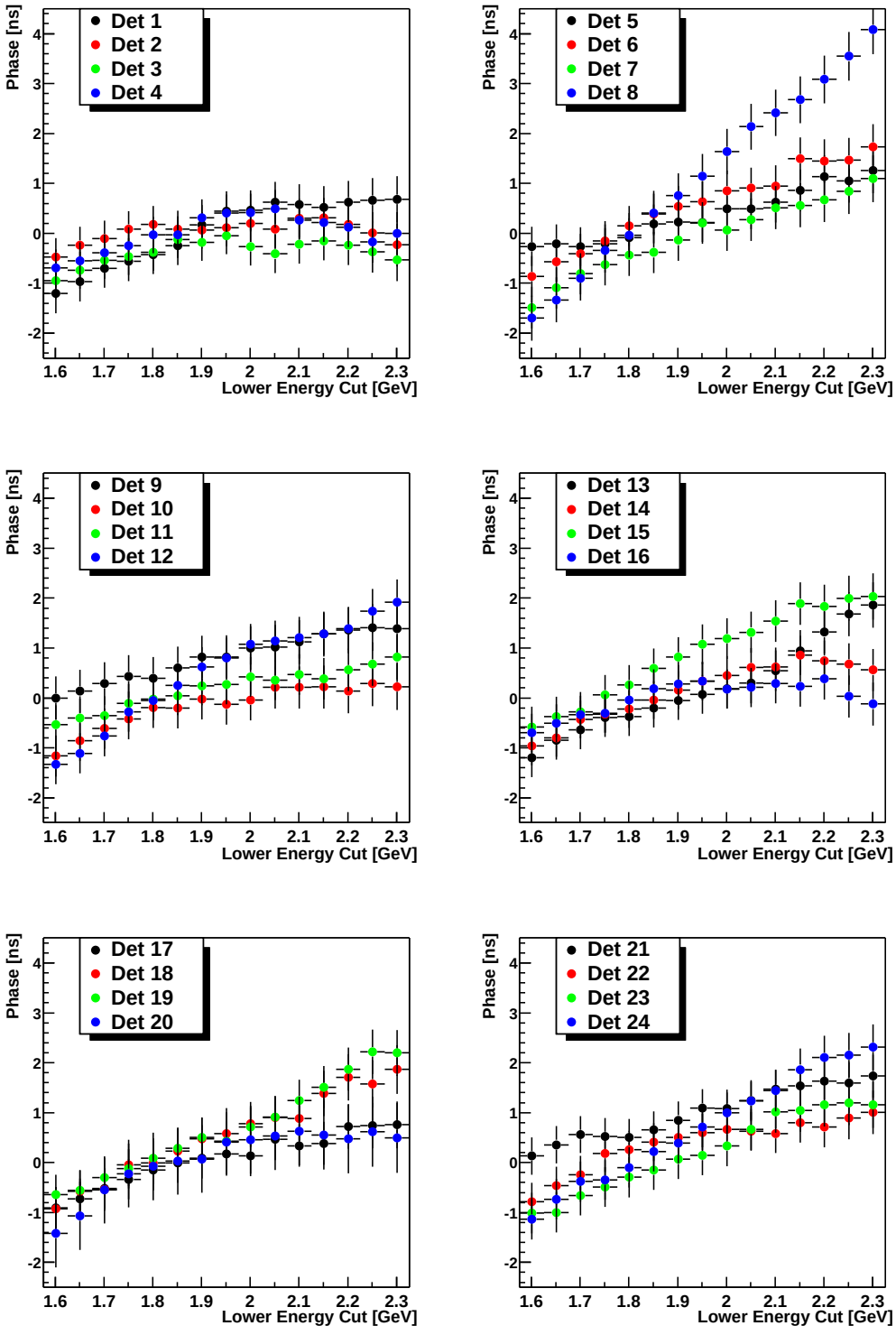


Figure 24: Phase from 5 parameter fit versus lower energy cut after time offsets have been applied.

4 Data Selection

4.1 Completeness

In 2001, the run numbers ranged from 9001 to 11384. The G2Too production was checked for the presence of all runs in this range and it was found that 96 run numbers were missing. The missing run numbers were cross-referenced with the online run log and the tape copying log. After some detective work, 41 were determined to be missing due to the run numbers never being assigned, 35 were too short to bother with tracking down and reprocessing, and 9 were not recorded to tape. The remaining 11 runs potentially useful for ω_a analysis were recovered from tape and processed. To facilitate comparison of the G2Off and G2Too datasets, the actual run numbers and more details on the detective work are given in **??**.

Prior to the production stage, G2Too breaks the raw data runs into chunks of 200 AGS cycles with the exception that the last chunk of each run contains however many cycles are left. Each chunk is treated as a separate entity during the pulse fitting, which opens the possibility for a particular chunk to become corrupted without corrupting the entire run from which it was derived. The integrity of the processed data was checked by opening each chunk and verifying that the data from the beginning, middle, and end could be accessed. Out of approximately 7700 chunks processed, 98 chunks failed this test. Most of these chunks were traced to a problem with the marker pulse that would result in a divide by zero error during the production. Fills with this problem were thrown out and 78 of the raw chunks were successfully reprocessed. The remaining 20 chunks seem to come from predominantly good runs but will not be pursued further at this point. This should not exclude their runs from the G2Too runlist since in most cases other chunks from the same run were processed without incident. Chunk numbers are given in **??**.

After checking that each run number exists and every chunk processed is accessible, there is still room for a small amount of data to be lost if individual chunks were misplaced during the production. A test was implemented to check that if chunk number i exists for a certian run, then is there also a chunk $(i-1),(i-2),\ldots,1$ from that run. Note that this does not catch chunks missing from the end of the run, but it does allow a measure of how often such an error might have occurred. A total of 6 chunks were found to be missing and are listed in **??** as 'magic', since they seem to have just disappeared.

4.2 The 'Golden' and 'Silver' Runlists

The following subsections discuss runs completely removed from the analysis due to instable running conditions, run length, special studies, and negative logbook entries.

4.2.1 Experimental Stability

The first cut applied to the 2001 data set was based on overall stability of the experiment. At the start of the experiment, calorimeter voltages and thresholds went through several iterations. The last major iteration took place at run 9385 where the HV was lowered considerably on det 4, and perhaps more importantly, all of the thresholds were set to values that were maintained throughout the run. After this run, a few fine tunings of the top/bottom response and energy endpoint were

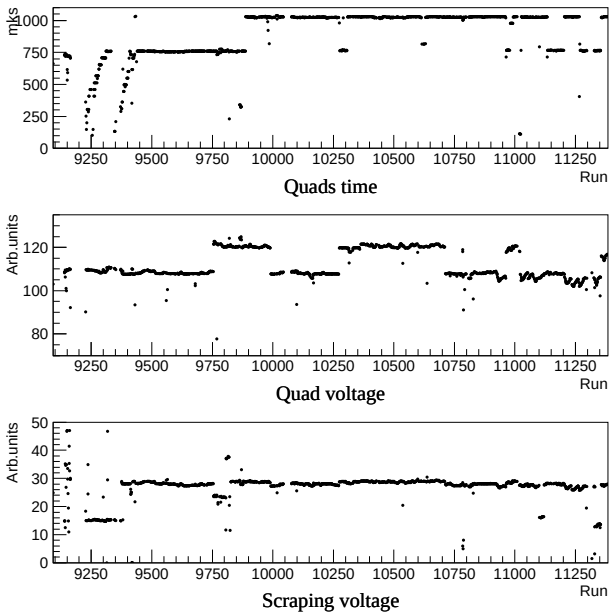


Figure 25: Vanya's web plot of quad parameters showing stability starts around run 9409.

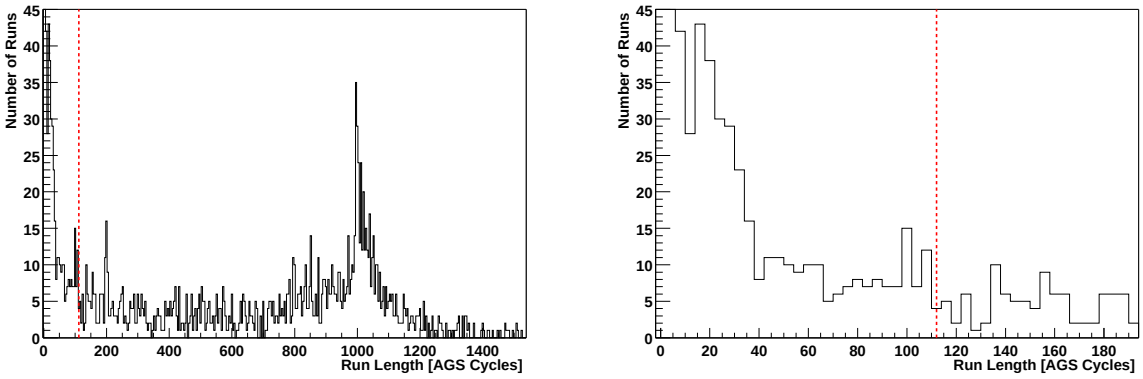
made. These changes were large enough that the energy endpoint in the $g-2$ analysis should be calculated on a per run basis, but small enough as to not require excluding the data.

As a result of the change to μ^- in 2001, the ability to maintain the quad voltage for the duration of the measurement period became another source of instability at the start of the run. It was not until run 9409 that the quads were able to run consistently out to 750 μs. About this same time, the scraping voltage was decided, see Figure 25. After run 9409, a series of scraping and radial field studies were performed, making run 9421 the first candidate for a starting run number.

4.2.2 Run Length

A second cut applied to the 2001 data set was based on run size. Short runs are often indicative of something gone awry in the experiment. If one allows the operator 5 minutes after the start of each run to look around the counting house and find a problem, then this implies a minimum of 100 AGS cycles for a run to be considered acceptable. Figure 26(a) shows the distribution of AGS cycles per run for all the runs in 2001. Given the slightly elevated number of runs ending around 100 AGS cycle, see Figure 26(b), the cut was imposed on runs smaller than 112 AGS cycles. A total of 842 runs were elimated, of which 513 were 0 length (DAQ misstarts) and 138 were already eliminated by the stability cut.

In addition to the run length in AGS cycles, the number of events per run is another indicator of run quality. Figure 27 shows a plot of the number of raw events per run summed over all detectors. The open red circles represent runs that have already been eliminated based on the 5 minute rule. A typical good run has about 2×10^7 raw events while the cut based on AGS cycles transitions at about 2×10^6 events. This in not surprising since 112 AGS cycles is about 10% of a normal run. The 2×10^6 cut was extended to include runs longer than 112 AGS cycles which eliminated an additional 158 runs from the runlist, 110 of which were already eliminated by the cut on stable run conditions. These runs are shown in 27 by the open black boxes.



(a) (b)

Figure 26: (a) Histogram of all the runlengths in 2001, measured in AGS cycles. (b) Zoomed in region. The vertical, dashed line indicates where the cut on runlength was placed

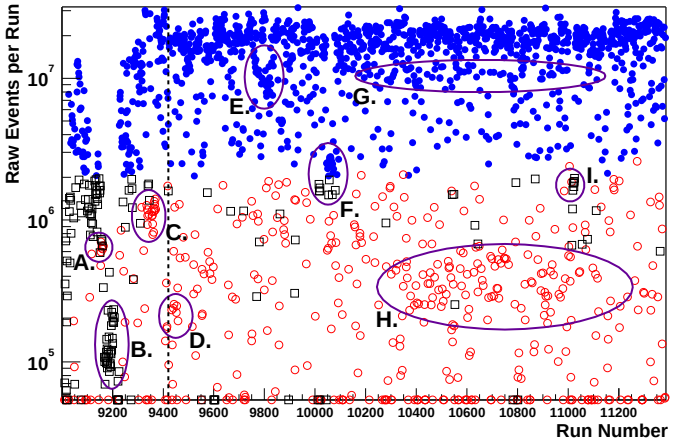


Figure 27: Number of raw events plotted versus run number. Closed blue circles represent runs that survived both the minimum 112 AGS cycles cut (open red circles), and the raw events $> 2 \times 10^6$ cut (open black boxes). The vertical dashed line is drawn at the point where run conditions became stable. The solid purple ellipses indicate clusters of runs taken under special conditions, see text.

A closer examination of Figure 27 reveals clusters of runs with an anomalously low number of counts. Cross checking with the online logbook showed the following special run conditions for a give cluster.

A. The runs in this region and immediately around it were special scraping studies.

B. The trolley got stuck so this cluster contains a series of laser runs (9167-9207) without beam.

C. Runs with the Cerenkov detector at various pressures (9335-9345) and a series of fiber beam monitor(FBM) runs (9354-9370) under different scraping conditions studies.

D. Period of time with erratic beam and multiple DAQ problems.

E. Runs in this ellipse were subject to a dip in the AGS intensity(9759-9772) followed by a series of scraping and radial field studies (9774-9790, 9805-9826) after the beam intensity

rose to normal.

F. FBM studies under various scraping conditions(10010-10019, 10045-10075).

G. The line of data points within this ellipse corresponds to laser runs which were typically half the length of a production run.

H. This cluster shows the effect of adding a warning system that alerted the operator when a run started without all of the WFDs read out.

I. More FBM studies(11014-11025).

Conditions	Runs Eliminated
4kV Scraping	11291-11354
Kicker $< 97\%$	9992-1002, 10111-10114, 10189-10198, 10770-10775, 10784-10797, 10962-10967, 10996-10998
Fiber Beam Monitor	9354-9370, 9600-9605, 9894-9895, 9965-9967, 10010-10019 10045-10075, 11014-11025
Cerenkov	9335-9345
Scraping and Radial Field	9102-9164, 9219-9239, 9312-9318, 9412-9420, 9774-9790, 9805-9824 10610, 10785-10787
Resonance	9820, 9863-9870
No Beam Laser	9168-9207
Sweeper Off	11002, 10811, 10831-10833
Scraping Off	11299-11324
Quad Problems	11330-11331

Table 2: Runs elimated due to special running conditions.

4.2.3 Online logbook entries

During the course of the run, an online logbook was kept where the run operator indicated whether or not a run was considered good. Owing to its subjective nature, this criterion was not used as a cut in selecting runs used in the data set. Many runs marked as bad were due to inappropriate quad parameters, lasers runs, DAQ crashes, short runs, etc. Nearly all of these runs are eliminated by the run length cuts discussed above or by the fill quality checks discussed in the next section. However, the online logbook was crucial in tracking down periods where systematic studies were performed or experimental running conditions changed significantly. See Table 2 for a list of runs eliminated due to special running conditions.

4.2.4 Other Runs Removed By Specific Tests

Many checks were performed on the integrity of each run by quantifying certain measurables and looking for outliers versus run number. I need to go back to the computer code and summarize the specifics of the checks, but for now I will outline the basic idea of each test.

Test	Runs Eliminated
High Losses	9549,9550,10935
Low KS Prob	9997,10004,10006,10366,10367,10369,10370,10372,10373,10402,, 10408,10410
Sneaky Laser	10534
Flashlets	10918,10928,11114
Reflections	9934,9991-10031,10360-10373,10381,10398-10412
Xiaobo Checks	9830,9833,9918,9968,10843,11267,11330
Field Concerns	9675,9677,9975,10463,10477,10708,10710,10780,10784,10839,10862, 10922,10925,10927,10928,10943,11082,11105,11110,11205,11208

Table 3: Runs elimated due to failed integrity test.

- Fractional muon losses were constructed for each run. By far most of the runs with excessive losses were traced to causes like FBM or no scraping runs. A few runs were removed on the basis of high losses with no apparent cause.

- After performing all of the endpoint calibrations using the KS compatibility test, the average probability that the energy spectrum of the detectors from a single run were consistent with the sum over all runs was plotted. Some runs had a low probability, but the power of the test was questionable. A lot of these same runs had moderately high losses. Runs that exhibited both a low KS probability and a moderate increase in losses were removed.

- Since laser runs are in the golden runlist, it is important to know if the photodiodes can be used reliably on a per fill basis. If for some reason the photodiodes do not fire or fills get misaligned between WFD crates, then laser event will sneak into the data. Laser runs were checked by building the number of n-fold coincidences between detectors in a particular crate for each run. Fills with high n-fold coincidences were cross-checked with the laser photodiode to find its efficiency. Only one run was found to be bad, and the problem was traced to the amplitude of the laser drifting down (nitrogen running out) to the point that the photodiode WFD was no longer above threshold.

- Quad off data was use to check for an excessive fraction of flashlets in each run. Some types of runs that had already been removed were able to mimic the signature of a high flashlet run, FBM runs for instance. The runs where the sweeper was turned off also showed up in this test. In addition to the high flashlet runs with a good explanation, there were still a few more runs that were removed.

- The time difference between the i and i+1th event in each detector was plotted. Some runs showed a spike in the dt spectra corresponding to a reflection. The odd thing about this reflection is that when it appeared, it would be in multiple detectors. The cause was assumed to be electronic and the runs were removed.

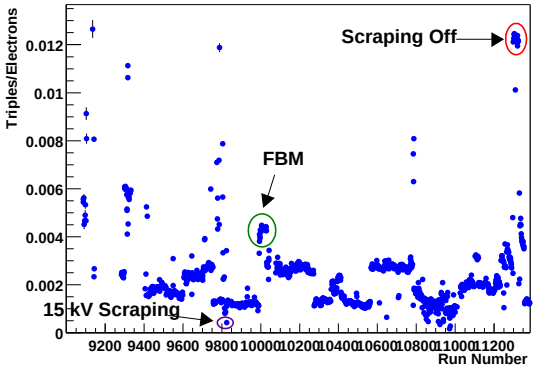
- The CBO frequency was found for each run, but no new run cuts were needed.

- In addition to the checks listed above, Xiaobo did some other types. In particular, he fit each run with a 5-parameter fit and looked for outliers. In addition to this, his cuts on minimum

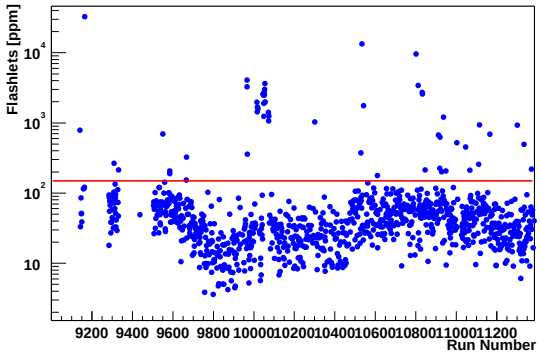
number of events were slightly different. A compromise was reached and the additional runs cut by Xiaobo's run selection were applied.

- The field analysis excluded several runs due to various concerns.

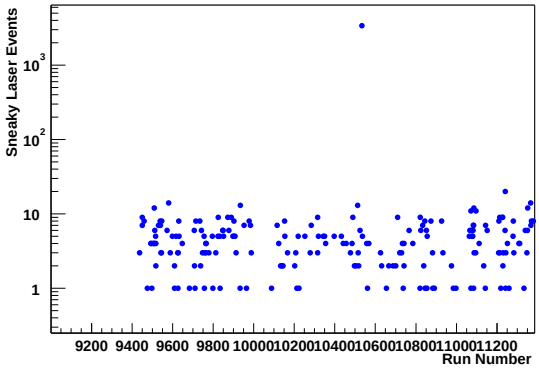
The runs removed by these tests are given in Table 3. Some of the measurables used are plotted in Figures 29, 28, and 30. The data in Fig. 30 was used to divide the dataset into low and high n running periods. The run period notation used in the rest of this document is shown in the figure.



(a) Normalized muon losses.



(b) Flashlet fraction



(c) Sneaky laser events

Figure 28: Measurables used for run cuts versus run number.

4.2.5 Run Cut Summary

Some of the systematic studies produce data in the calorimeters that is unusable under any circumstances for g-2 fitting, i.e. laser no beam, Cerenkov, resonance, and FBM studies. However, some of the more benign studies like varying scraping and radial field conditions, turning the sweeper off, and running without scraping or sweeping might be salvageable after checking relative muon loss levels, flashlet contamination, etc. In order to quantify the impact of these cuts on the data selection, Table 4 shows the fraction of raw events lost on each run cut. Raw events are fitted events in the WFD without any time or energy cut, therefore the total number can not be used to quantify the statistical error on $g-2$. However, the fraction of such events lost to each cut is still

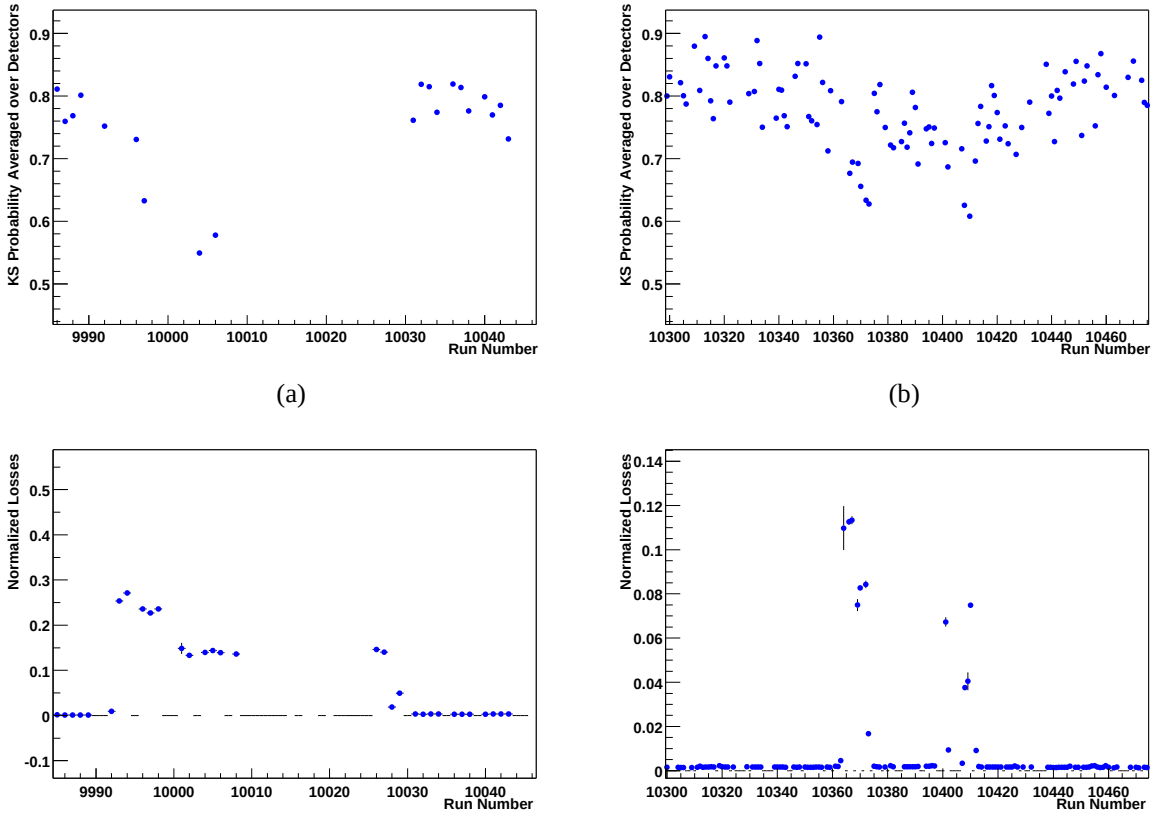


Figure 29: Regions with anomalously high losses and low KS probabilities.

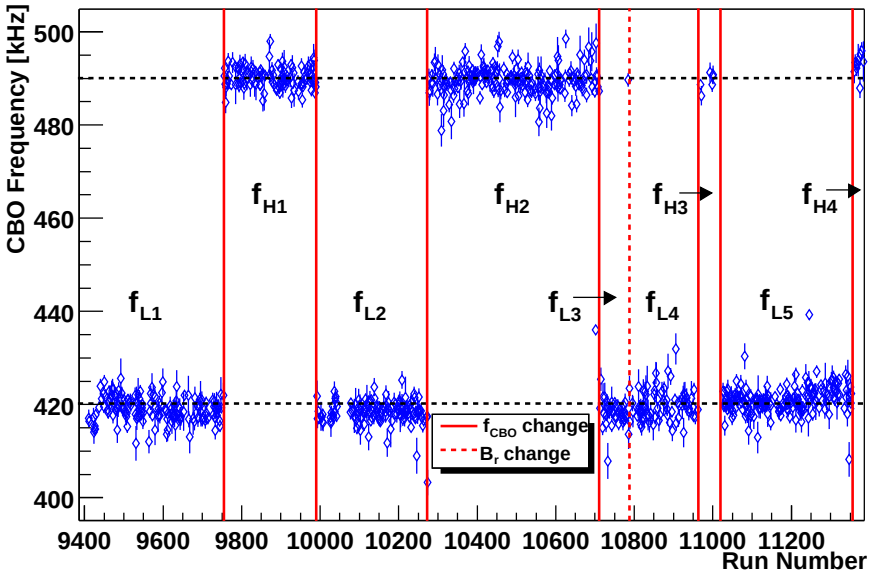


Figure 30: CBO frequency determined from FFT of 5 parameter fit residuals.

meaningful for monitoring the data lost on each cut. The amount of data lost from entries that do not have the 'Number of Events' quantified, was estimated by the number of runs removed.

Run Cut	Number of Events ($\times 10^9$)	Percent of Total
None	20.204	N/A
Unstable	1.515	7.5
Short/Low Counts	0.222	1.1
Scraping/B_r Studies	0.221	1.1
Scraping Off Study	0.464	2.3
Resonance Study	0.032	0.2
Sweeper Off	0.0095	0.05
Failed Loss Check		0.3
Poor KS Test and Loss		1.2
Failed Flashlet Check		0.2
Quad Problems		0.2
Failed Laser Check		0.1
Removed by Xiaobo		0.5
Removed by Field Analysis		2.2

Table 4: Fraction of data lost from run cuts.

The largest impact is due to the cut on experimental stability. It is estimated that about half of the 7.5% data loss could be salvaged if runs prior to 9421 were more carefully considered. However, runs in this region would have to be fit separately with a much earlier stop time due to the fluctuating quad on time. The endpoints in this region could be determined on a per run basis, but changes in the calorimeter top/bottom balance as well as the hardware thresholds could still affect the fit in more subtle manners. The scraping voltage was changed several times just prior to this run so the muon losses would also have to be constructed specifically for this run range. All in all the net benefit is small relative to the cost.

The above set of cuts constitutes the golden runlist. This runlist was intended to be aboslutely safe to debug and define the fitting function. A silver runlist was created by expanding the golden runlist to include all of the 4kV scraping runs and the kicker runs where the voltage was 95-95%. Comparisons between the golden and silver fit results are given in the section on fitting.

It is important to note that the runs in the golden runlist, which were used by all analyzers, were not checked for problems that should be caught by fill checks. For instance, runs where the quad shuts off at 500 μs are not eliminated. It is assumed that each analyzer will place their own cut on quad duration suitable for their fitting range. In the same vein, no run cuts are based on T0 or laser status.

4.3 Fill checking

In addition to run cuts, a quality check is implemented on a per fill basis at the time of production. In each fill the T0 counter, quad voltages, and the laser photodiodes are checked and a corresponding status bit is set. When building histograms, the status bit is checked and fills with an error status are automatically filtered out unless the G2Too user has specified otherwise.

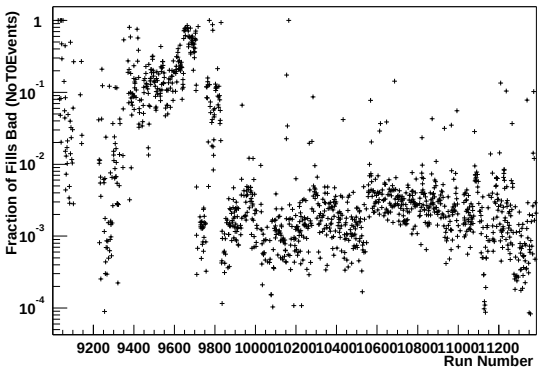
As particles enter the ring, a piece of acrylic placed at the entrance of the backleg of the inflector generates a Cerenkov signal that is readout via a PMT. This signal is analyzed and its time relative to the AGS prepulse is subtracted from the time of all detected decay electrons. This assures that the g-2 phase stays fixed from fill to fill even if the timing of the AGS prepulse is changes or becomes erratic. The NoT0Events status bit is set when the area of the To signal is found to be less than 60 counts. Fig. 31(a) shows the fractional number of times that fills were lost for each run due to the T0 status bit.

In addition to the T0 counter, the laser reference photodiodes are checked for events. If an event is found, then the corresponding fill is marked as a laser fill. Fig. 31(b) shows the fractional number of times that the laser fired. The runs with the fraction of fills at about 50% are the true laser runs. The runs with the much smaller fraction of laser events, are runs in which the laser would sporadically receive a trigger. The cause is not known, but the effect has been confirmed in non-laser runs by looking at the n-fold coincidence data.

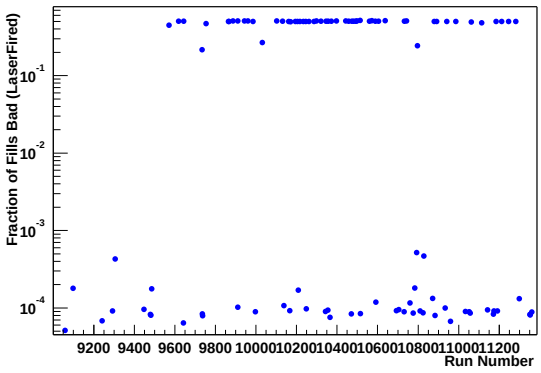
Finally, the quad traces are checked for a sharp decrease in voltage between 60 μs and 600 μs after injection which would correpsond to a quad spark or a short quad duration. From Fig. 31(c) it is apparent that quad sparks occur with a probablity of about 10^{-4}. The points where the fraction is much higher, unity in some cases, are from runs in which the quads were shut off before 600 μs. The quad traces are also checked to see if they were surpressed, as is the case if they are off for a fill in which flashlets are checked. Fig. 31(d) shows the fraction of times a fill was flagged as quad suppressed. The point in the run where the dial on Cenap's box was turned from supressing one in 25 fills to one in 50 is clearly evident. Runs in which the fraction is not 0.04 or 0.025 are probably a result of short runs affecting the fraction.

Fill Cut	Number of Events ($\times 10^8$)	Percent of Total
NoT0Events	6.75	3.6
LaserFired	3.18	1.7
NoQuadEvents	1.24	0.7
QuadSpark	1.06	0.6
QuadSuppressed	0.0084	0.005

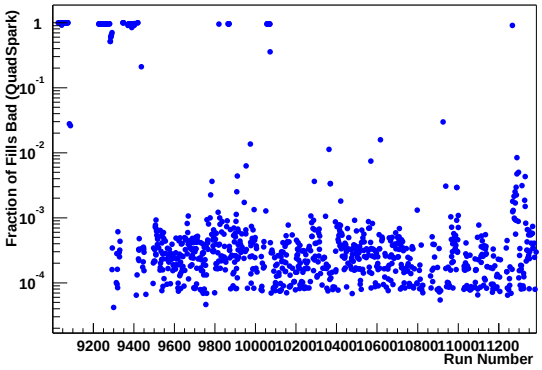
Table 5: Fraction of data lost from status bits.



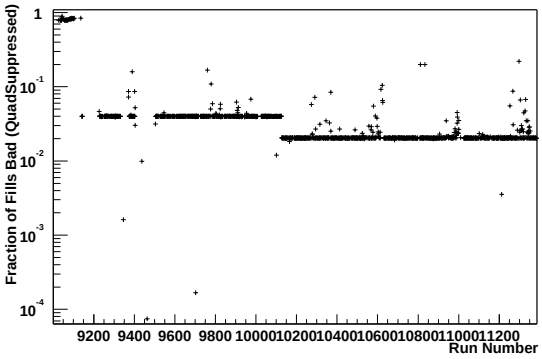
(a) T0 Found



(b) Laser Fired



(c) Quad Spark



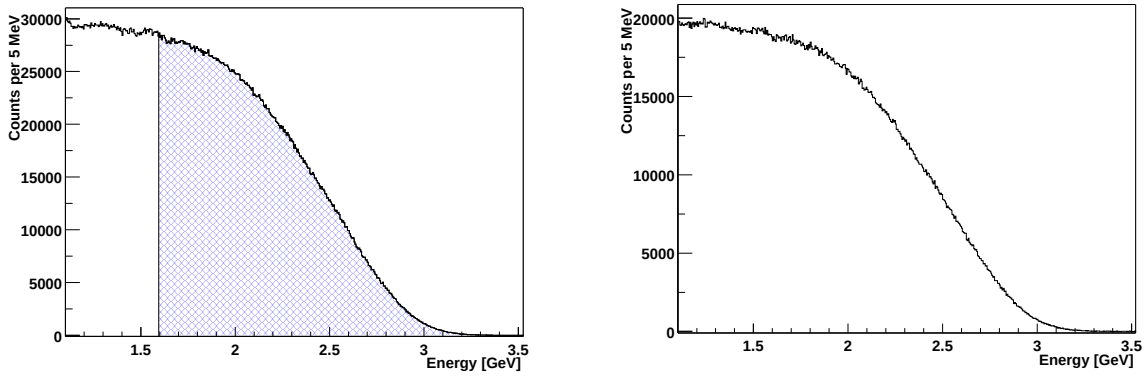
(d) Quad Suppressed

Figure 31: Fraction of fills in which various status bits were flagged.

5 Gain Correction

5.1 KS Gain Construction

The energy scale corrections versus time in spill were constructed using the same Kolmogorov-Smirnov (KS) method that was used to determine the endpoints, see section 3. A standard reference for the energy spectrum in each detector was constructed from pileup-subtracted data in the interval from 185 to 450 μs. Individual histograms were constructed for each g-2 period prior to 185 μs. An example of the reference for detector 1 and an individual histogram taken from the 7th g-2 cycle after injection are shown in Fig. 32.

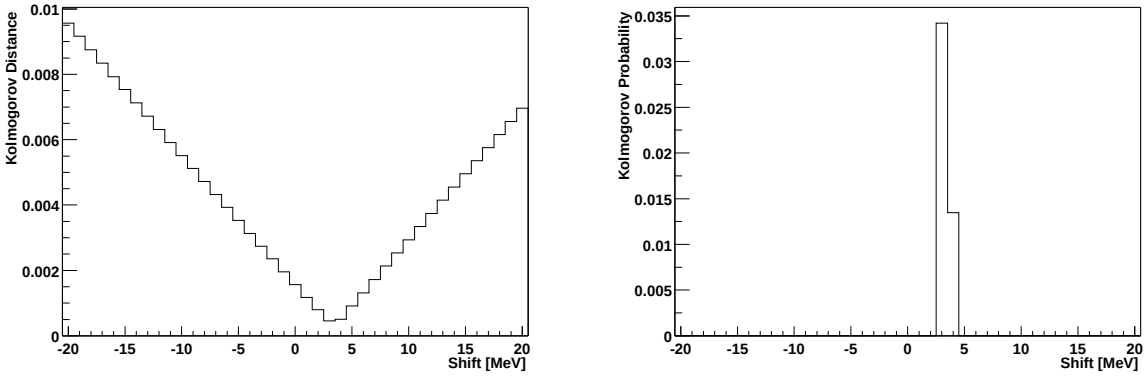


(a) (b)

Figure 32: Detector 1 example: (a) Standard reference constructed at late times. The hashed blue region indicates the energy range in which the KS test is applied. (b) Individual histogram taken from the 7th g-2 cycle after injection.

The individual histograms were compared to the standard reference in the energy range from 1.6 to 3.6 GeV and a KS distance and probability were calculated. The individual histograms were then translated +/- 50 MeV in 5 MeV steps, probabilities and distances recalculated, and the point of minimum KS distance was determined using an analytical parabolic minimization. Fig. 33 shows the KS distance and probability versus the amount of shift applied to the individual spectrum shown in Fig. 32(b). Note that as the Kolmogorov distance grows, the probability rapidly goes to 0, which is why the minimum distance was used in place of the maximum proability to find the optimal shift. Unlike the endpoints, it is assumed that one iteration of the KS method is sufficient for determining the gain correction.

Since the counts in the individual histograms fall exponentially with the muon lifetime, in order to get sufficient statistics at times later than 60 μs, multiples of g-2 cycles were used to construct the individual histograms. For instance, up to the 14th g-2 cycle only 1 period is used to build the individual histograms, but starting with period 15 two g-2 cycles are used, etc. The resulting gain versus time in spill for a few representative detectors is shown in Fig. 34. With a close look at the data points, one can see where they shift by one half of a g-2 cycle, corresponding to times when an extra g-2 cycle was added to the 'moving average'. The plots show the constructed gain before the correction is applied, the fit to that data that was used for the corrections, and the gain after the correction has been applied. In order to meet the condition of unity at late times and still have



(a) Minimum KS Distance (b) Maximum KS Probability

Figure 33: Plots of the Kolmogorov (a) distance and (b) probability as a function of the amount the individual histogram in Fig. 32(b) was shifted.

some degrees of freedom to describe the constructed gains, a fit of the form,

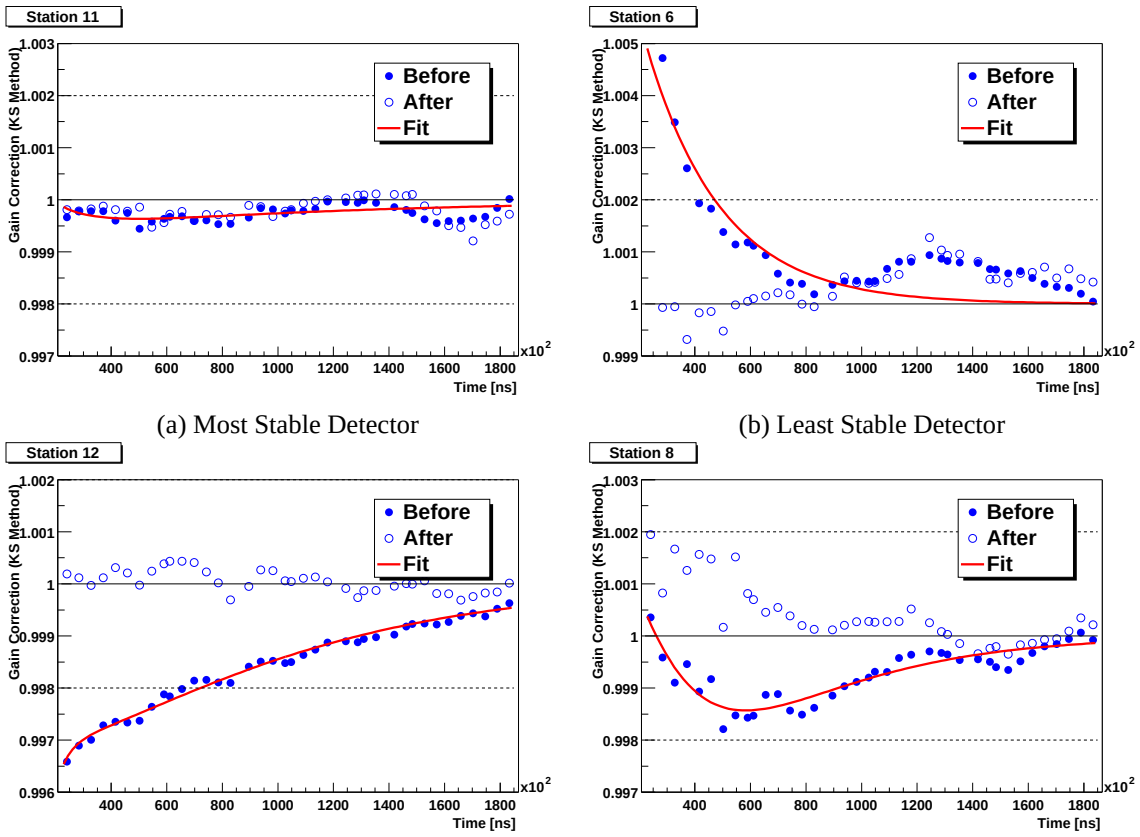
$$g(t) = 1 + N_1 e^{-t/\tau_1} + N_2 e^{-t/\tau_2} + N_3 e^{-t/\tau_3}, \tag{10}$$

was used. For most of the detectors, only two of the exponentials were necessary. The function was a little too constrained to completely describe the gain curve in a few detectors, see Fig. 34(b) for the worst example, but in general the fit and the data were in good agreement as shown in the other panels. The dashed lines in Fig. 34 are drawn at $\pm 0.2\%$ in order to guide the eye. The rule of thumb has always been that the detectors need to be stable at this level over a time period of 200 μs in order to prevent biasing R significantly. Most of the detectors already meet this requirement without applying a gain correction, but a few do exceed the 0.02% level. Almost all of the gain changes are large enough to affect the χ^2 significantly(NEED A PLOT). Plots similar to those in Fig. 34 can be found for all of the detectors in Appendix D.1. In general, a survey of all of the detectors shows that the gain correction using the KS method worked extraordinarily well. Indeed only one iteration was required.

5.2 Pileup Sensitivity

One of the primary motivations for developing an independent method of constructing the gain correction was to study the influence of pileup. In prior years, the average energy plots have trended upward at early times which seems counterintuitive from what one might expect from the photocathodes either recovering from the flash or gating off. This upward trend is also consistent with what one would see if pileup subtraction were incomplete. The KS method uses the entire shape of the distribution and should therefore be less sensitive then taking an average energy, where, by the nature of averaging, high energy pileup events are preferentially weighted.

There are a couple of ways to study the effect of pileup on the final result. One can leave out pileup subtraction or check the sensitivity to the maximum energy used in the comparison. The results of both studies are shown for a typical detector in Fig. 35. The KS test was iterated under several different conditions starting with no pileup subtraction and an energy range that extended up to 3.7 GeV. By changing the maximum energy to 3.1 GeV the upward trend in the gain curve



(a) Most Stable Detector

(b) Least Stable Detector

(c) Most Impressive Correction

(d) Least Impressive Correction

Figure 34: Plots of the constructed gain before correction, the fit from Eq. 10 used for the correction, and the data after correction.

is significantly lowered. When pileup subtraction is performed, after 20 μs the upper energy limit hardly has an effect on the constructed energy. Only at very early times is there a significant deviation. This means that in the energy range from 3.1-3.7 GeV, the pileup construction is quite accurate. The equivalence of the KS analysis in two different energy ranges also confirms that the method directy measures gain, as opposed to the average energy method which would require two different multipliers to extract the gain result.

For comparison, a completely independent analysis of the gain was performed using the more traditional average energy method. The conversion factors to go from average energy to gain for each detector were taken from Mario and can be found in his report. The same energy range was used, so the multipliers should be valid. In general, the results from the average energy and the KS method are in good agreement. Some examples are given in Fig. 36, while the plots for all of the detectors are in Appendix D.2. The biggest exception to the consistency is detector 8, shown in Fig. 36(d). For the average energy and the KS test to yield different results, it is likely that the shape of the energy distribution (not just the scale) is changing early-to-late.

It has been known since the 2000 analysis that detector 8 has a poor light yield. The effect was first discovered by observing the prepulses that come before an electron showers in the calorimeter. The working hypothesis is that the prepulse comes from the decay electron preshowering in upstream material, yielding a secondary particle, which then Cerenkov's in the light guides. The shorter path length to the PMT accounts for the apparent causality breaking. The logbooks from

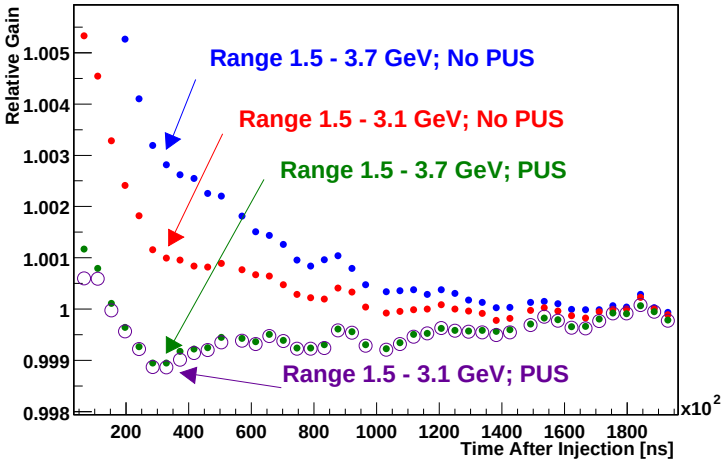
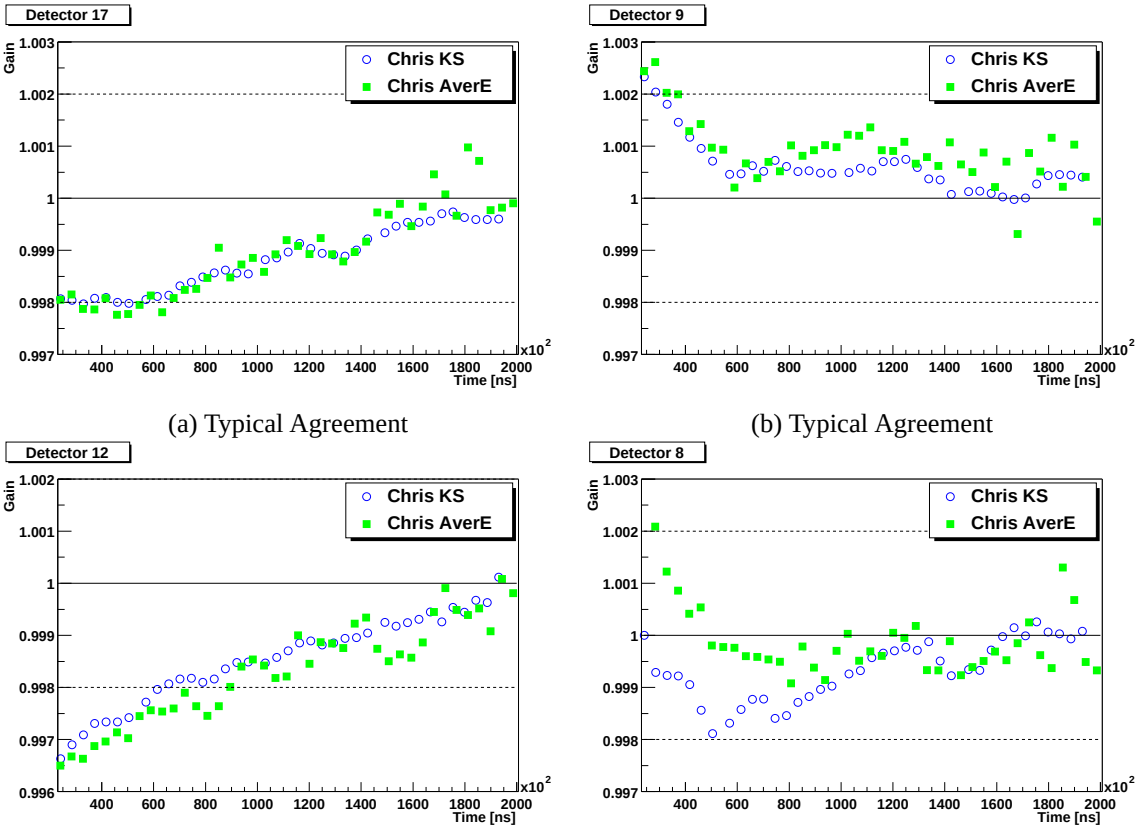


Figure 35: Sensitivitity of the KS metod to pileup.



(a) Typical Agreement

(b) Typical Agreement

(c) Typical Agreement

(d) Anomalously Poor Agreement

Figure 36: Examples of the consistency between the average energy and the KS method.

the test beam were checked and it was found that detector 8 did have a much lower photoelectron yield. This discrepancy in the average energy and KS results could be another symptom of the low light yield coupled to an early-to-late background. For instance, in the early days of the experiment (1997 maybe?) there was a huge bacground of 'narrow' pulses that were hypothesized to come

from neutrons converting in the concrete floor and emitting a gamma that interacted directly with the photomultiplier tubes. The problem was solved by increasing the light yield of the detectors via replacing some of the lightguides, adding reflectors to the ends of the calorimeter, and removing the yellow filters (originally intended to cut the Cerenkov signal in the light guides). Since detector 8 has the lowest light yield, it is possible that the narrow pulses are still getting into the data stream. The narrow pulses in detector 8 have been observed in an excellent study performed by Jimmy John at Boston, but I do not recall if his study proved the narrow pulses to all be of the prepulse variety (Cerenkov type) or if they were uncorrelated to the next pulse (neutron type). Here I am refering to the narrow pulses in his study that were not due to the phase alignment problem.

Back to the topic at hand, another way to check the sensitivity to pileup is to vary the applied deadtime. In G2Too, a deadtime is enforced on the dataset in order to remove the energy dependent deadtime that naturally arises in pulse fitting. Lengthening this deadtime increases the amount of pileup in the data, which can then be compensated by increasing the width of the windows used in the pileup construction. The resulting gain corrections for three different deadtimes are shown for detector 13 in Fig. 37. The gains were constructed with and without pileup subtraction to show the effect on average energy if pileup is neglected. The average energy method was used for this method since it is easier to apply and has been proven to be consistent with the KS results. Once again, the gain correction appears to be largely insensitive to pileup after the pileup subtraction has been performed. This is similar to the finding in Fig 35 with the exception that an increased deadtime includes any biases from all energies. Qualitatively one can conlude that the gain correction is not significantly influenced by residual pileup.

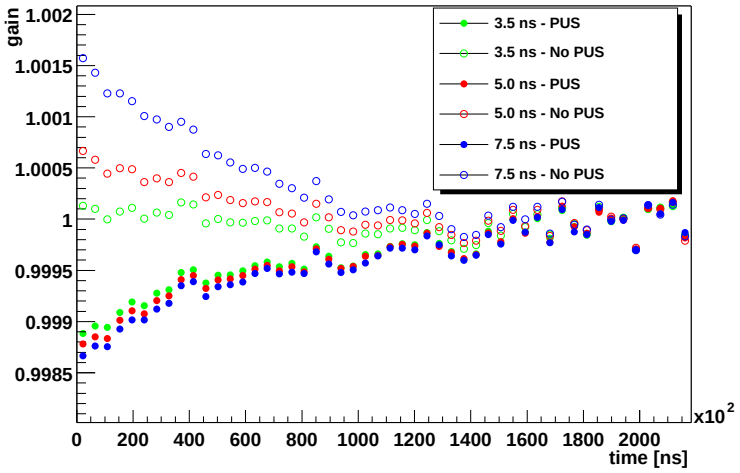


Figure 37: Gain construction with various deadtimes applied.

A final study of whether or not the average energy method is sensitive to pileup was performed by looking at the average energy versus time at the peaks and valleys of the $g-2$ cycle. The results for a particular detector are shown in Fig. 38, where there is no discernable difference in the average energies. This was verified for the other detectors as well.

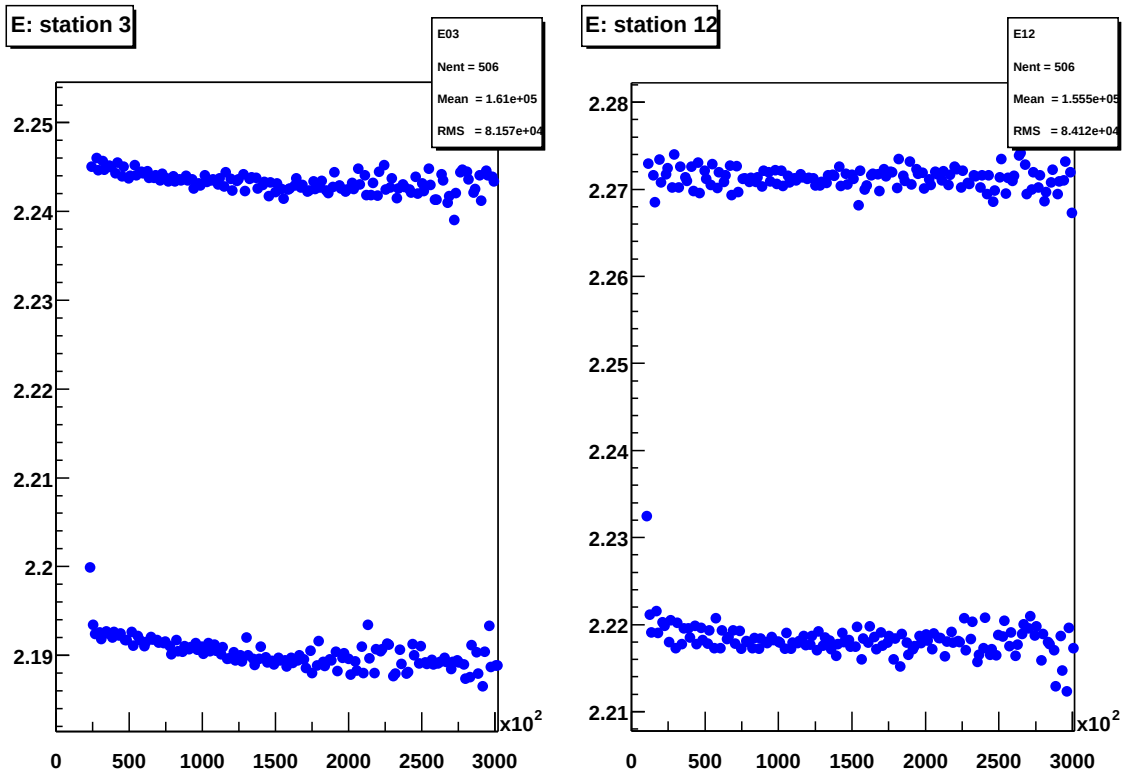


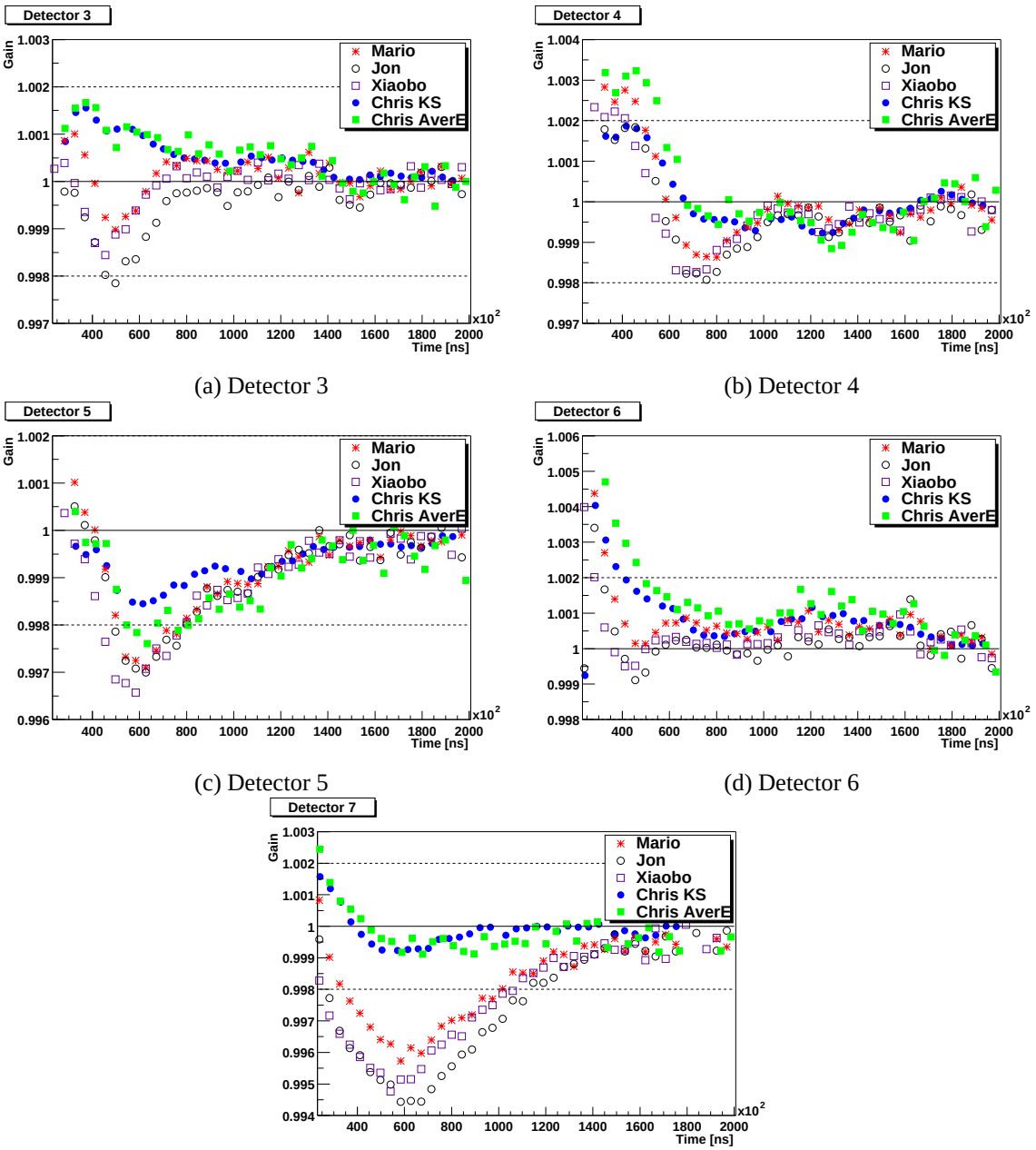
Figure 38: Average energy versus time computed at the peaks and valleys of each $g-2$ cycle.

5.3 Comparison With Other Analyzers and the Whammo

Gain effects that arise from hardware issues, like the photmultipliers sagging or a rate dependence in the WFD gain, should be the same for all analyzers and both productions. Differences are most likely to stem from problems with an analyzers construction of the gain or from a difference between the G2Off and G2Too productions. Several discrepancies exist, the worst of which has been termed the 'whammo' effect. For all analyzers using the G2Off data, there is a clear dip in the average energy spectrum for those detectors in the injection region, as shown in Fig. 39. It is clearly a difference in the production.

Many different theories were proposed as to the source of the difference between productions, including:

- An injection related background (like neutrons) that is somehow treated differently by the two productions.
- Something concerning small pulses and the different software thresholds used in the production.
- A pedestal related effect from the elevated pedestals in these detectors.
- Something related to the island breaking algorithm used in the two productions.



(a) Detector 3

(b) Detector 4

(c) Detector 5

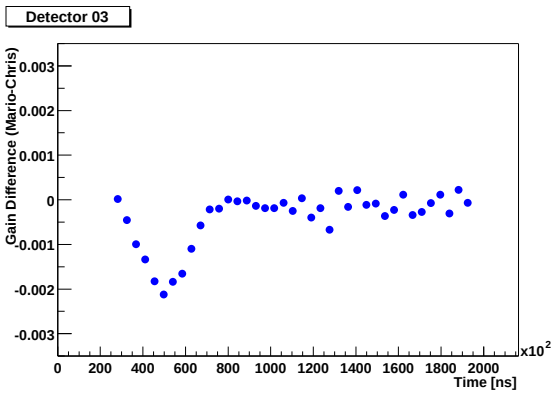
(d) Detector 6

(e) Detector 7

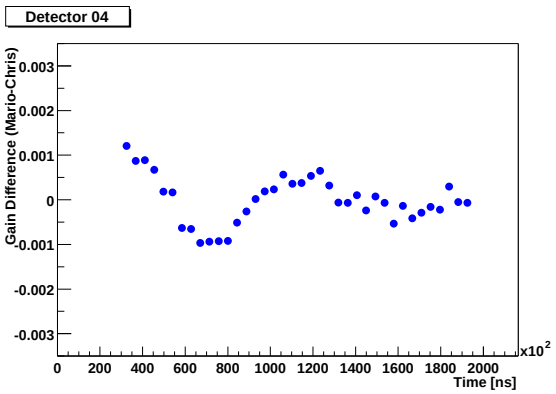
Figure 39: Gain comparison between analyzers for the whammo detectors.

- G2Off uses the entire island when performing the minimization, while G2Too only use ± 7 samples. Perhaps the difference is generated when the island length is long (like during continuous digitization).

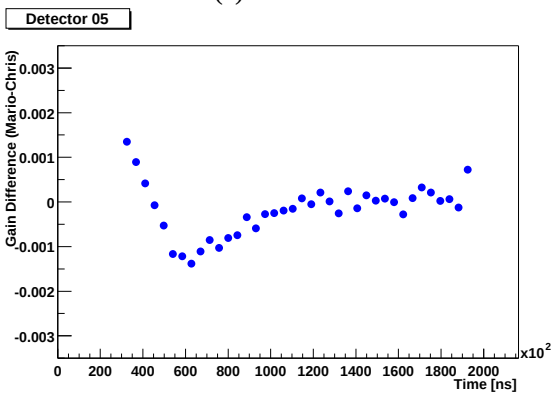
The source of the difference has not yet been determined, but recent work by Vanya has shown that the dips do seem to be in phase with the end of continuous digitization. Whatever the cause, the effect on the average energy is $> 0.2\%$, which makes it **larger** than the bias one would get if pileup were not subtracted, as is illustrated in Fig. 40 where the difference between Mario's gain curves and the KS results from above are plotted for the whammo detectors.



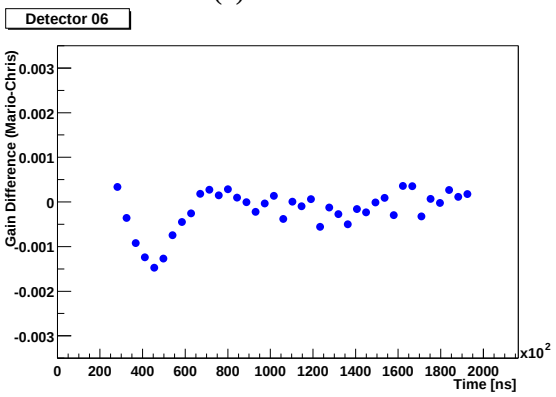
(a) Detector 3



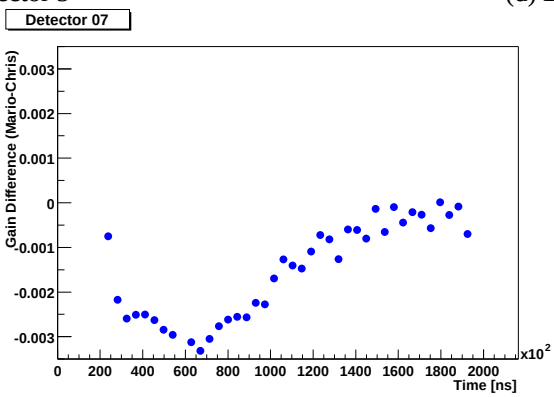
(b) Detector 4



(c) Detector 5



(d) Detector 6

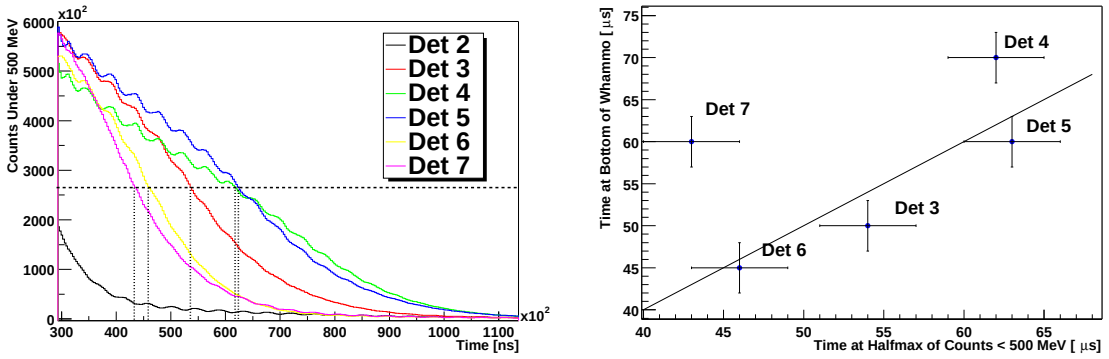


(e) Detector 7

Figure 40: Gain comparison between analyzers for the whammo detectors.

A quick study to confirm Vanya's observation of the phase alignment of the whammo and contiuous digitization was performed. The time when the whammo is maximally affecting the average energy was taken by finding the minima in the difference plots (Mario-Chris) shown in Fig. 40. The point where continuous digitization ends was determined by looking at the time distribution of pulses under 500 MeV. These pulses are only constructed if the WFD is continuously digitizing or if a neighboring pulse crosses threshold. The point where continuous digitization ends was taken to be the point where the flux of particles under 500 MeV crosses its half-maximum. The time distributions are shown in Fig. 41(a), along with lines illustrating where the times were

chosen. Plotting the time of the whammo minimum versus the time when continuous digitization is ending, Fig. 41(b), reveals a strong correlation in detectors 3-6. The exception to the rule is detector 7, which falls quite far from the $y = x$ line drawn in the figure. The whammo in detector 7 is has always been considered a little different due to the much longer timescale and the factor of two larger effect on the average energy. The time structure of the whammo indicates that the problem is not really the continuous digitization itself, but rather the end of the continous digitization. As an example of a type of model that would fit this description, consider the case where one phase of the WFD stops digitizing before the other. Perhaps events in this time range are reconstructed by G2Off (with missing energy due to the lacking phase), but are thrown out by G2Too. An update from Vanya, reveals that the problem actually has to do with a correlation between energy and island length in the G2Off fitter. It the WFD is not continuously digitizing, then the island length is determined by the WFD hardware and is relatively stable versus time in spill. During continuous digitization, the islands are broken up uniformly. The problem occurs at the end of continuous digitization where the island length was allowed to vary substantially.



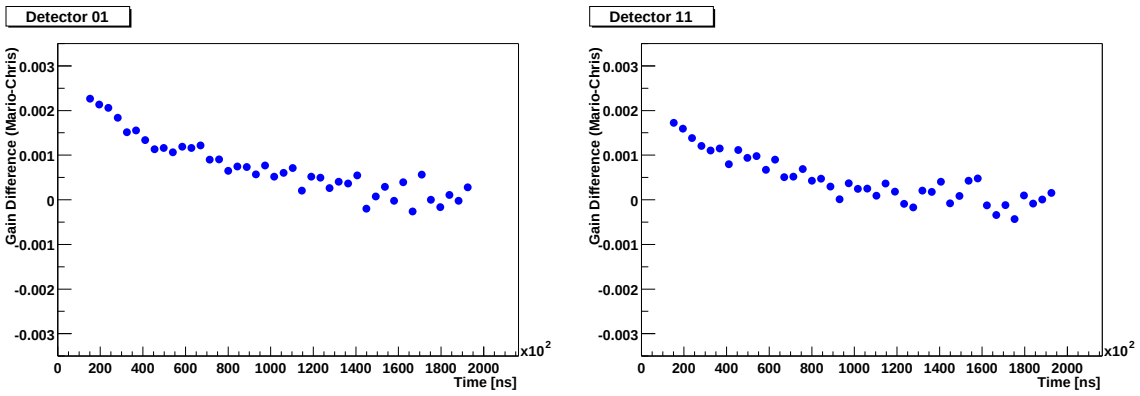
(a) (b)

Figure 41: (a) Plots of events under 500 MeV illustrating where continuous digitization ends. (b) Correlation between the time of the whammo and where continous digitization ends.

5.4 Comparison With Other Analyzers in Non-Whammo Detectors

The other difference in the gain corrections, although not as dramatic as the whammo, is a slow early-to-late difference. As an example, the plots in Fig. 42 show the difference in gain construction between Mario and I for a couple of the non-whammo detectors. I glance through Appendix D.4 will reveal that this 0.15% early-to-late discrepancy exists for nearly all detectors.

For a general comparison of the average analyzer result it does not make sense to average the gains and then compare the results since some of the upward trending gains will cancel with the downward trending. It is better to first compute the difference for a particular detector and then look at the average of those differences, as shown in Fig. 43. The 0.15% early-to-late difference between Mario and myself is evident (note the consistency between the average energy and KS results). There is also a 0.05% discrepancy between Mario and the other G2Off analyzers (Jon and Xiaobo). Note that these two differences imply that there is at least a 0.1% early-to-late difference between G2Off and G2Too, perhaps indicative of another production difference. As further evidence of this



(a) Detector 1 (b) Detector 11

Figure 42: Gain difference between Mario's average energy and my KS results.

problem, there was a third independent study of the gains in G2Too, performed by Tau, which was shown to be consistent with my KS results at a collaboration meeting. Unfortunately, I do not have access to his results to produce plots at the moment.

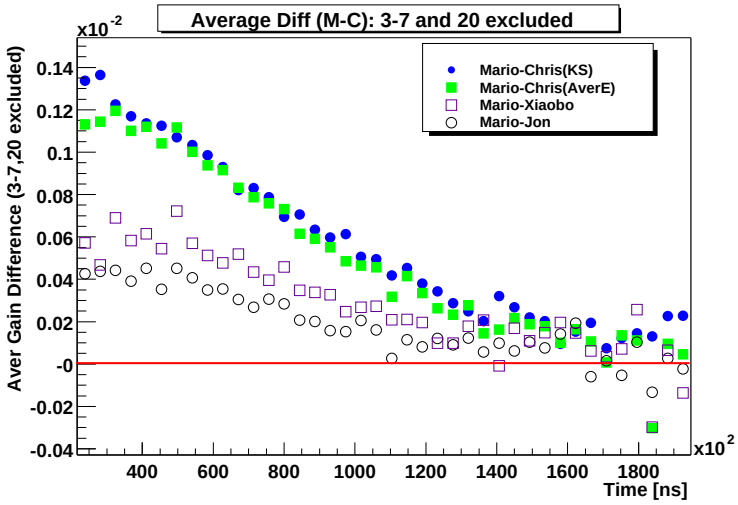
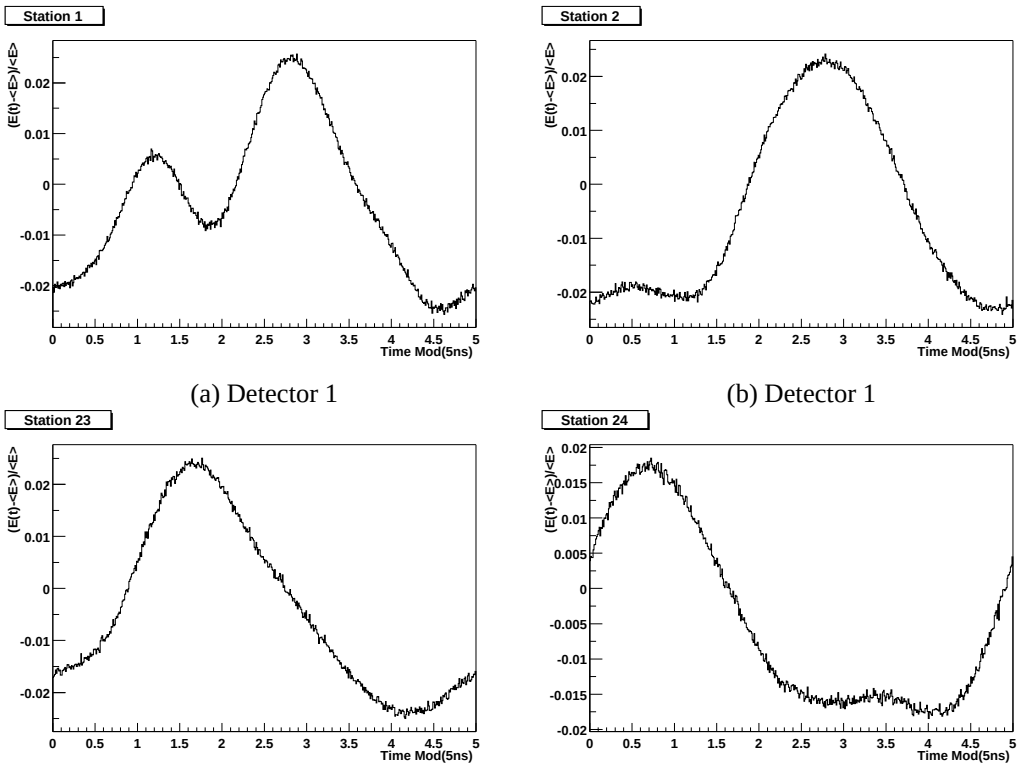


Figure 43: Average difference in constructed gains for non-whammo detectors.

5.5 Clock Correction

In addition ot the normal gain correction, it was found that there is an energy dependence with respect to the WFD clock phase. This dependence was constructed from the average energies, and the energy of each pulse was then corrected. Fig. 44 shows a few examples of the average energy versus clock phase. The rest of the plots fro each detector can be found in Appendix D.5.



(a) Detector 1 (b) Detector 1

(c) Detector 1 (d) Detector 1

Figure 44: Average energy with respect to WFD clock.

6 Muon Losses

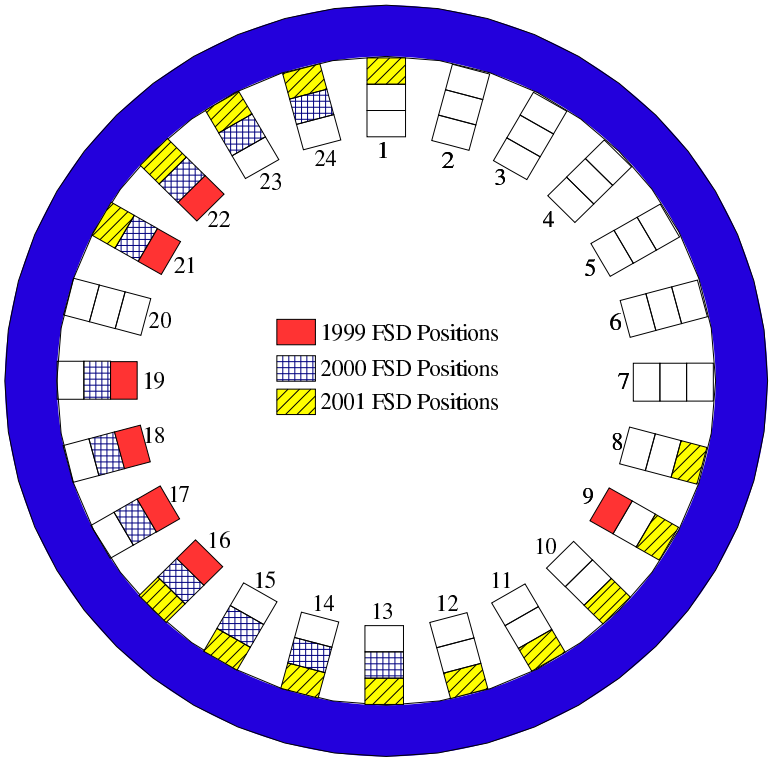


Figure 45: Distribution of FSDs around the ring in various years.

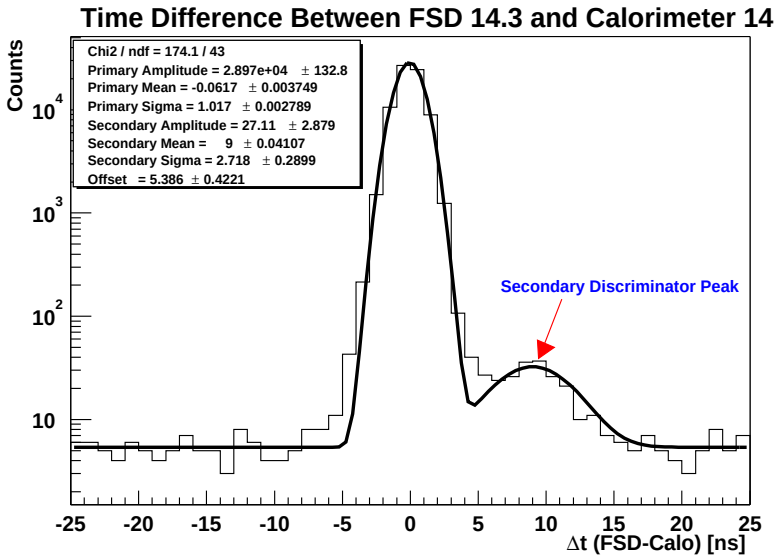


Figure 46: Example of discriminator refiring.

- Three-fold coincidences in the FSDs were used to determine the functional form of the lost

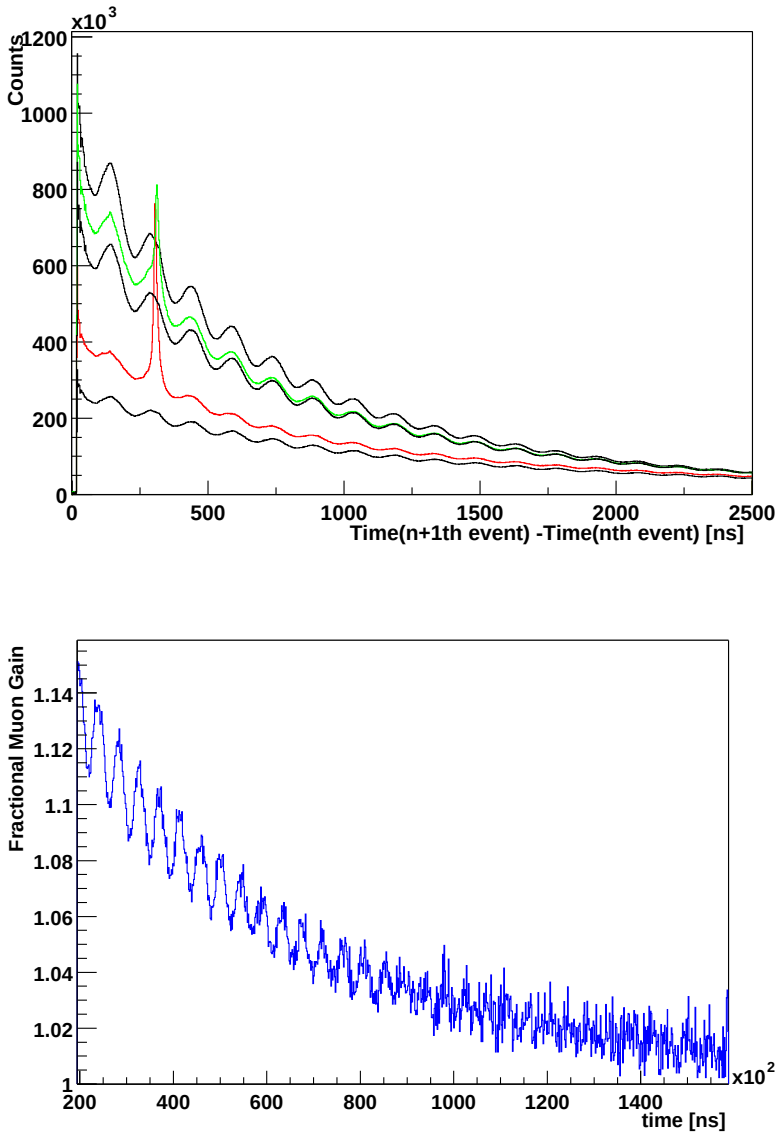


Figure 48: Total correction to muon losses in 2000.

muons.

- New corrections were built for the discriminator deadtimes and accidental calorimeter vetoes.
- A new implementation of the muon losses was derived and used by both Mario and myself.

7 Histogram Filling with w=A

The intention of this section is to document the procedure used when filling and fitting histograms with the asymmetry-weighted method. Hopefully, I will be able to publish a more theoretical version later as a g-2 note, but for now a few trivial examples are given to help intuit how the asymmetry-weighting works. More detailed discussions are given in g-2 notes 326 by J. Pretz and 327 by F. Farley.

7.1 Trivial Examples

The main advantage of the asymmetry-weighted method is that it takes full advantage of the statistical power of the dataset by not allowing the weak signal in the low energy data to contribute at the same level as the strong signal in the upper energy regime. One positive side effect relative to uniform weighting is that the requirement of a lower energy bound is removed. If it were not for hardware issues, even the data with negative asymmetry could be used. Pedagogically, it is helpful to think of the simple example of two sine waves 180 degrees out of phase.

Example 1

$$\begin{aligned} f_1(t) &= N_1 A_1 \sin(\omega t) \\ f_2(t) &= N_2 A_2 \sin(\omega t + \pi) \end{aligned} \tag{11}$$

Consider the function g(t) that results when these two functions are summed with arbitrary weights w_1 and w_2,

$$\begin{aligned} g(t) &= w_1 f_1(t) + w_2 f_2(t), \\ &= w_1 N_1 A_1 \sin(\omega t) + w_2 N_2 A_2 \sin(\omega t + \pi). \end{aligned} \tag{12}$$

Under a uniform-weighting scheme, all asymmetries, A_i, are treated equally,

$$w_i = 1 \Rightarrow$$

$$\begin{aligned} g_{w=1}(t) &= 1 \cdot N_1 A_1 \sin(\omega t) + 1 \cdot A_2 \sin(\omega t + \pi) \\ &= N_1 A_1 \sin(\omega t) + N_2 A_2 \sin(\omega t + \pi) \\ &= N_1 A_1 \sin(\omega t) - N_2 A_2 \sin(\omega t) \\ &= (N_1 A_1 - N_2 A_2) \sin(\omega t) \end{aligned} \tag{13}$$

In the case where $A_1 = A_2$ and $N_1 = N_2$, the signal is totally lost. However, rewriting the second sine wave as,

$$\begin{aligned} f_2(t) &= N_2 A_2 \sin(\omega t + \pi), \\ &= -N_2 A_2 \sin(\omega t), \\ &= N_2(-A_2) \sin(\omega t), \end{aligned} \tag{14}$$

we see that it can really be considered as a sine wave that is in phase with an amplitude $-A_2$. Adding the two functions weighted by their asymmetry yields,

$$\begin{aligned} g_{w=A}(t) &= A_1 \cdot N_1 A_1 \sin(\omega t) + (-A_2) \cdot N_2 (-A_2) \sin(\omega t) \\ &= N_1 A_1^2 \sin(\omega t) + N_2 A_2^2 \sin(\omega t + \pi) \\ &= (N_1 A_1^2 + N_2 A_2^2) \sin(\omega t). \end{aligned} \tag{15}$$

It is apparent that the sine waves no longer add destructively.

The last example was just a trivial example of the more general premise. Namely, when adding sinusoidal signals of varying amplitudes, the best frequency resolution is formed by using asymmetry-weighting. From the above example, it should be evident that any general weighting scheme must apply a negative weight to data with a negative asymmetry if it is to add coherently with positve asymmetry data. A more general proof that using $w = A$ minimizes the statistical error is given by J. Pretz in note 326.

The main drawback in using the asymmetry-weighting is that it complicates the error formalism. In a simple counting experiment, data is entered into a histogram with a weight of one. Technically, the probability to count ν events in a time period Δt is binomially distributed, but in the limit of a large number of particles with a small probability per Δt to decay the Poisson distribution is a good approximation,

$$P(\nu) = e^{-\mu} \frac{\mu^\nu}{\nu!}. \tag{16}$$

The parameter μ is the average number of events measured in the time Δt if the experiment were repeated many times. In the limit that μ is large (40 or so in most applications), the Poisson distribution can be approximated by a Gaussian with a mean of μ. The variance for the Poisson distribution in the Gaussian limit is then,

$$\sigma = \sqrt{\mu}. \tag{17}$$

Thus we have arrived at the familiar fact that in a counting experiment with sufficient statistics, if N events are measured in a given time interval, then the error is just $\sqrt{N}$.

The reasons for wanting to apply a different weighting scheme are widely varied, but perhaps the most common has to do with a known inefficiency in a system.

Example 2

Suppose that in a particular time interval 100 decays are counted, but the detector is known to be only 50% efficient during that interval. How many counts should be attributed to that time interval and with what error?

Since only 100 events are actually measured, the best estimate for the variance on those counts would be,

$$\sigma = \sqrt{100} = 10. \tag{18}$$

Taking the efficieny, $e = 0.5$, into account, the prediction for the actual number of decays is given by,

$$\begin{aligned} N_{decay} &= \frac{1}{e} N_{meas} \\ &= \frac{1}{e}(100 \pm 10) \\ &= 2(100 \pm 10) \\ &= 200 \pm 20 \end{aligned} \tag{19}$$

Note that the error is simply scaled along with the measured counts.

This simple example was really made to contrast with Example 3. Rather than modify the total number of measured counts at the end of the counting period, as was done in the last example, one can instead keep track of the number of real decays 'on the fly' by assigning a weight of $1/e$ to each event. The variance can also be continually maintained by keeping track of the sum of weights-squared,

$$\sigma^2 = \sum_{i=1}^{N} w_i^2, \tag{20}$$

as illustrated below.

Example 3

Show that filling each event with a weight of $1/e$ (where $e = 0.5$) and calculating the error as the sum of weights-squared reproduces the results of Example 2.

With 100 measured events the height of the histogram bin is,

$$\begin{aligned} h &= \sum_{i=1}^{100} w_i = \sum_{i=1}^{100} 1/e = 100(1/e) = 100(2) \\ &= 200. \end{aligned} \tag{21}$$

While the error is given by,

$$\begin{aligned} \delta^2 h &= \sum_{i=1}^{100} w_i^2 = \sum_{i=1}^{100} 1/e^2 = 100(2^2) \\ &= 400, \end{aligned} \tag{22}$$

which implies,

$$\begin{aligned} \delta h &= \sqrt{400} \\ &= 20. \end{aligned} \tag{23}$$

So we see that both the height, h, and the error on the height, δh, from Example 2 are reproduced.

The last couple of examples illustrated how one can keep track of the true number of counts and the error on that number in a system where the detector is known to have an inefficiency. The same formalism of keeping track of the sum of weights and the sum of weights-squared applies to the asymmetry-weighted analysis. Essentially, we purposefully make the counting of low energy events 'inefficient' relative to high energy decay electrons. Examples 2 and 3, treated the case where there was a constant efficieny, but it is not hard to imagine how to combine data from subsets of varying efficiency. The height of a histogram bin is just the sum of the individual heights and the error is just the individual errors added in quadrature. Now take the limit where the distribution of weights becomes continuous instead of discrete and the asymmetry-weighted method is obtained.

7.2 Procedural Details

Once the weighting procedure is understood, the implementation is trivial. Two histograms must be filled; one that contains the sum of all the weights in a particular bin and one that keeps track of the sum of weights-squared. In ROOT, there is no need to actually make a second histogram. By setting the histogram to include the Sumw2() option, the bookkeeping for the error is automatically performed. The only missing ingredient is the actual weighting function to be used. Since the asymmetry is a well defined function of the decay electron energy and the energy of each event is measured separately, we need only to find the function $A(E)$ in order to fill the histograms weighted on an event-by-event basis. In some sense, the fact that we can even use the asymmetry-weighted method can be heralded as yet another advance of the Brookhaven g-2 experiment. By placing WFDs on each detector, we have access to the continuum of asymmetry information. The end result is that the error on the g-2 frequency is reduced by about 10%. Our typically four month runs get the equivalent of an additional month for free.

The parameterization of the asymmetry could be done in many ways, but the procedure used for this analyis was to first fit the uniformly-weighted data with a five-parameter function in narrow energy bands for each detector. The data for this initial parameterization of the asymmetry was calibrated for the endpoints, corrected for gain, and pileup subtracted. It is not really important that these corrections be done, since all that is really needed is an approximation for $A(E)$. Any deviation from the true function will reduce the statistical power of the dataset slightly, but the effect is small once the shape has been roughly determined. The plots in Fig. 49 show the energy dependence of the asymmetry for all 23 detectors used in the analysis. The functional form of $A(E)$ is very similar for all detectors with the exception of 7, 8, and 9 where energy loss in the kicker plates smears the distribution.

The data in Fig. 49 were fit with a 5th-degree polynomial over the range from zero energy up to where the asymmetry maximizes. Beyond this point, there is not much data and what remains is some combination of the tail of the energy distribution and residual pileup. Events that do happen to occur at these high energies were assigned the maximal asymmetry obtained from the fit. Since energies greater than 3.3 GeV were not used when fitting for ω_a, this is not much of an issue. Fig. 50 shows the distribution of the energies at which the asymmetry was found to be maximal for all of the detectors. As can be seen from the red-dashed line, the upper threshold of the analysis resides close to this point. As expected from the plots in Fig. 49, the energy at which the asymmetry

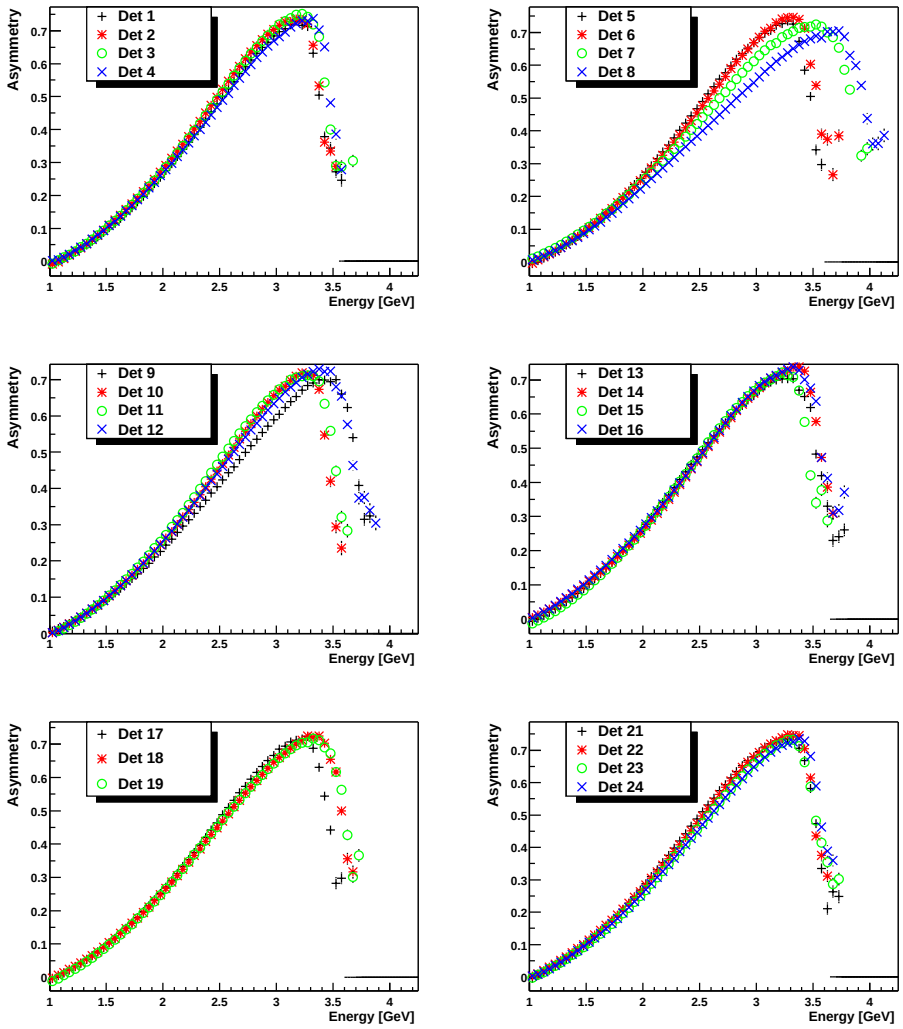


Figure 49: Asymmetry as a function of energy for all of the detectors.

maximizes is much higher for the detectors in the kicker region. This is probably just an artifact of how the energy scale was defined for each detector. Namely, the threshold at which the error on R is minimized in a uniformly-weighted analysis was defined to be 1.8 GeV. In reality, particles tend to lose energy in the kicker plates and the reconstructed energy is reduced. Therefore, the energy where R minimizes is also lowered. By defining that point to be 1.8 GeV, the energy distribution is shifted to the right relative to the other detectors.

Having obtained A(E) for each detector, the data was reanalyzed and the energy of each particle was used to determine the weight at which that event entered the histograms. Details of the modifications to the pileup algorithm are discussed in the next section.

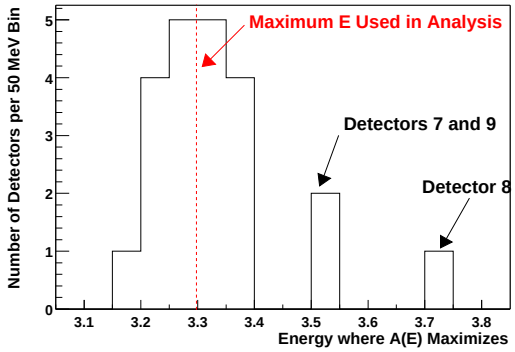


Figure 50: Distribution of energies at which the asymmetry is maximal for all detectors used in the analysis.

8 Pileup Construction

- Pileup was built using symmetric windows taken from -10 to -5 ns before the primary pulse and 5 to 10 ns after.
- For detectors with hardware thresholds greater than 750 MeV, a correction to the constructed pileup was applied to allow fitting down to a 1.5 GeV threshold.
- Pileup was subtracted from the data prior to fitting.

Pileup was constructed using symmetric pileup windows starting ± 5 ns to either side of a trigger pulse, as illustrated in Fig. 51. The 5 ns width of the pileup windows was determined by the applied deadtime of 5 ns used by all G2Too analyzers. Before histograms are filled, the data is looped through and any pulses that fall within 5 ns of each other are combined into one pulse using appropriate Logashenko coefficients, which are discussed more in a moment. This applied deadtime removes the energy dependence that naturally arises during the minimization procedure. This is in contrast to analyzers who use G2Off, where the width of the pileup windows is an energy-dependent function provided by Vanya.

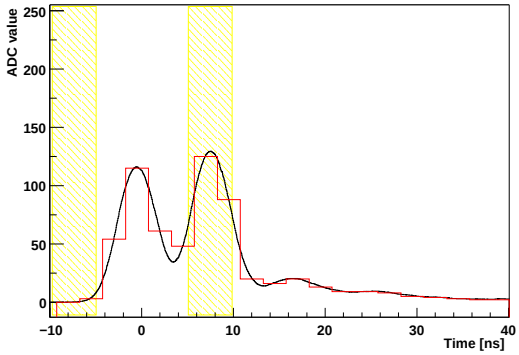
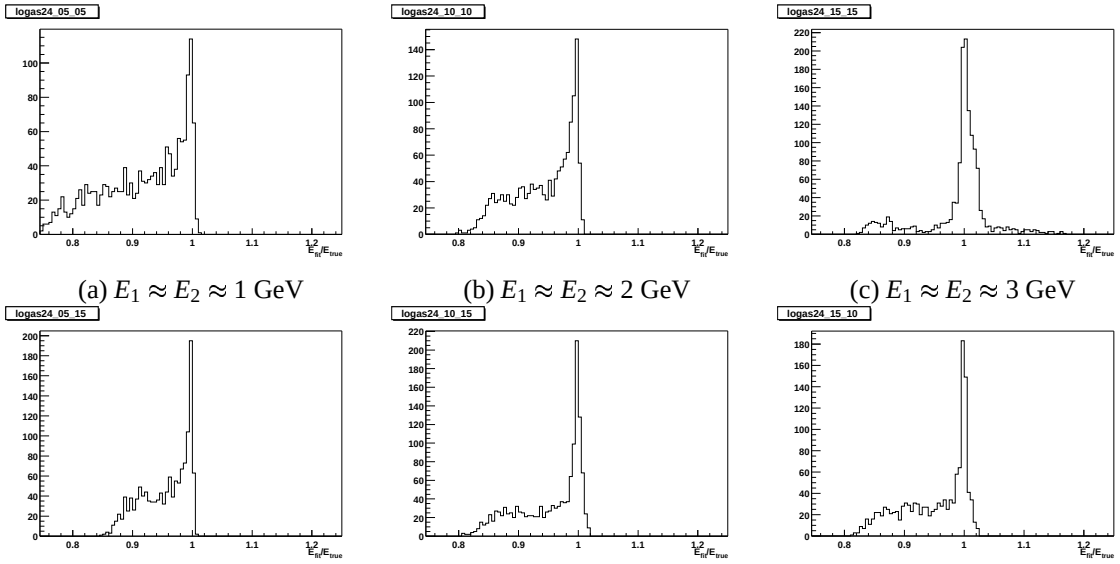


Figure 51: Example of two pulses with an 8 ns separation. The hashed areas indicate where pileup windows were placed.

It was found that the pulseshapes in 2001 hardly differed from the 2000 waveforms, so the same Logashenko coefficients were assumed. These coefficients were determined by simulating the pulsefitter's behavior when fitting two events that pileup in the detector. The reconstructed energy of the pileup event is generally smaller than the sum of the individual energies, with an average Logashenko coefficient of 0.94. The correction depends on the energy of the two pulses, so in practice, a two dimensional lookup table is used for determining the coefficient for each set of pulses. The correction is actually performed at two different stages of this analysis. The first is when the 5ns deadtime is being enforced, and the second occurs when constructing the energy spectrum of the doubles in the pileup construction. In simulation it was found that for a given set of energies of the primary and secondary pulse, there is actually a distribution of coefficients that arises. Most (all?) other analyzers use a fixed coefficient, wheras in this analysis a random point from a distribution of Logashenko coefficients was used. This was the same procedure used by Fred in the 2000 analysis, and a couple of examples of the distributions of Logashenko coefficients are given in Fig 52.



(a) $E_1 \approx E_2 \approx 1$ GeV (b) $E_1 \approx E_2 \approx 2$ GeV (c) $E_1 \approx E_2 \approx 3$ GeV

(d) $E_1 \approx 1$ GeV;$E_2 \approx 3$ GeV (e) $E_1 \approx 2$ GeV;$E_2 \approx 3$ GeV (f) $E_1 \approx 3$ GeV;$E_2 \approx 2$ GeV

Figure 52: Examples of Logashenko distributions.

One thing worthy of noting with regard to using symmetric windows is the possibility of double counting. Looking at Fig. 51, there is a trigger pulse along with an event in the pileup window to the right. If one imagines those pulses sliding to the left, then the pulse that was in the pileup window becomes the trigger and the original trigger pulse is now in the pileup window on the left. By construction, this happens everytime there are two pulses above the hardware threshold with a separation of 5-10 ns. Rather than double counting these pulses, they are only allowed to enter the S and D histograms once. This means that for the cases where both pileup pulses are above the hardware threshold, typically 800 MeV, the 'symmetric' pileup construction only use the 5 ns window to the right and is thus no different then the Mediterranean method. Considering that the applied deadtime is 5ns, it is good that only one of the two windows contributes, otherwise the total pileup window width would be 10 ns and we would be constructing twice as much pileup as

is really under the signal.

The symmetric pileup window really becomes useful when considering the cases where only one of the pileup pulses is above the threshold. The combinatorics of the pileup construction become quite interesting. Consider the time ordered pair (E_1, E_2). It is a familiar fact that in the Mediteranean method that if E_1 is less than the hardware threshold (with E_2 greater than the hardware threshold), then that type of pileup event is missed since E_1 does not trigger the WFD. It is true that E_2 fires the WFD, but since the pileup window is only drawn to the right, no pileup event is found. It is for this reason that analyzers using the Mediterranean method assume symmetry with respect to the times when E_2 is the member of the pair that is below the hardware threshold. Since E_1 triggers the WFD, there is no problem with finding these type of events, and double-counting them takes cares of the missing type of events. This requires a modification to the error in the S and D histograms due to the subset that is double counted. Note that this modification to the error bar can be done 'on the fly' by filling histograms for these type of events with a weight of 2 and keeping track of the sum of weights-squared.

In the case of using symmetric windows, it is apparent that the type of event missed in the Mediterranean method is no longer a problem. If E_2 is the pulse that is above the hardware threshold, then E_1 is found in the window to the left. Therefore, no modification to the error bar in the S and D histograms is required. Now one might ask, Ïf the deadtime is 5 ns and you are using 10 ns of pileup windows to construct the pileup, aren't you overconstructing pileup by a factor of two?Ïhe answer is no, because events where $E_1 > E_2$ are combinatorially unique from events where $E_2 < E_1$. Events consistent with the first inequality come only from a 5 ns window on the right and events consistent with the second inequality come only form the 5 ns window to the left, so 'my pileup porridge is just right'.

Of course, using the asymmetry-weighting complicates pileup construction a bit, but it is trivial to see how it must be done. If two pulses fall within the applied 5 ns deadtime, then they are recognized as one pulse of higher energy. From the $A(E)$ curves shown in the last section, these events are assigned an asymmetry and entered into the histogram with that weight. Of course, in reality the asymmetry is not correct since this event was really composed of two lower energy events. To reverse this mistake, it is necessary to subtract the high asymmetry event from the histogram and add in two lower energy(asymmetry) events. This means that when pileup is constructed, the events in D are assigned a weight of $A(L(E_1+E_2))$, in order to undo the first mistake. The individual events are then added to S with the correct weights $A(E_1)$ and $A(E_2)$ as they should have been if pileup were not an issue. Note that the L in the above equation is meant to indicate a Logashenko coefficient has been applied.

One of the main advantages to using the asymmetry-weighting is that it removes the need for a lower threshold. Fig. 53 shows N, A, and the statistical power NA^2 as a function of energy. The dashed lines indicate thresholds of 1.5 and 1.8 GeV. The 1.8 GeV threshold is the familiar threshold where the error on R is minimized in a uniformly-weighted analysis. Below this line there is still a fair amount of statistical power. By pushing the threshold down to 1.5 GeV, most of the statistical power is used. Fundamentally, there is no reason to not allow this threshold to be pushed even lower. However, as was explored carefully in the 2000 analysis, pileup subtraction becomes quite complicated when trying to use data from below twice the hardware threshold. Since quite a few of the hardware thresholds were set at 750 MeV, I chose 1.5 GeV as my minimum energy.

In constructing pileup, it is common to make the plot of the energy of the first pulse versus

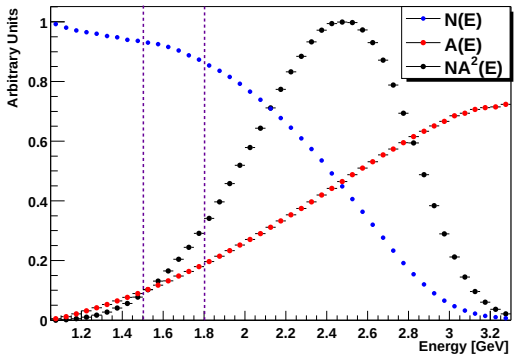


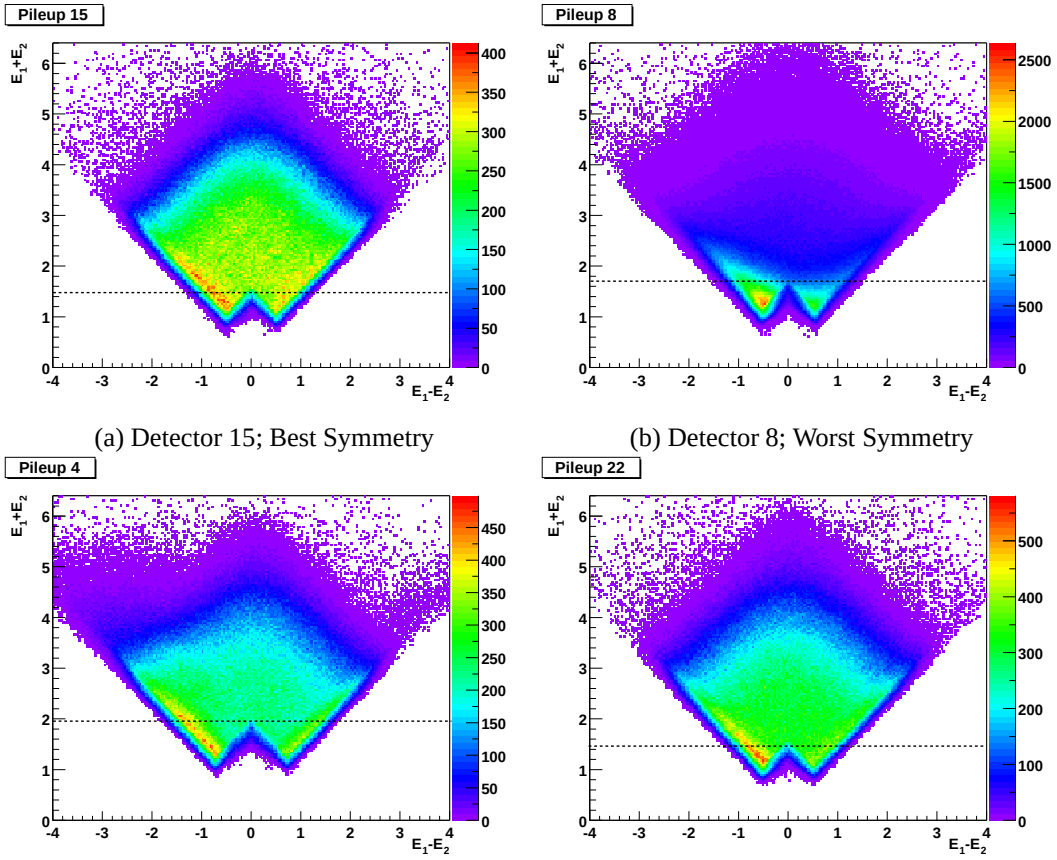
Figure 53: Energy dependence of N, A, and the statistical power NA^2. The dashed lines indicate the region of statistical power gained by moving the threshold down from 1.8 to 1.5 GeV.

the energy of the second pulse. From this plot, it is easy to visualize regions where the pileup construction is inadequate. For instance, in the Mediterranean method, the region where E_1 is less than the hardware threshold will show no counts The plots in Fig. 54 are very similar, with the exception that a coordinate transformation has been applied. The quantity $E_1 + E_2$ is plotted versus $E_1 - E_2$, which corresponds to a rotation of 45 degrees about the origin from the normal E_1 vs E_2 plots.

These plots are quite illustrative. Neglecting the changes in number of counts for the moment and just looking at the outline around the baseball diamond-like figures, we see that there is a fair degree of symmetry across the $E_1 = E_2$ axis. As discussed before, this is because the symmetric windows remove the threshold problem with E_1 discussed above. The eqivalent plot in a Mediterranean pileup subtraction would be missing a 750 MeV wide diagonal band running from third base to the home plate.

Now concentrating on distortions, the most dramatic difference is in detector 8, where there is five times as much pileup and the energy distribution of the pileup is very different. The difference in the shape is better visualized in Fig. 55. This difference is primarily due to the mass of the kicker plates. Apparently, quite a few particles shower in the kicker plates and send secondaries into detector 8. Furthermore, since there is a factor of five more pileup, there must be many more particles that shower than those that are detected by detector 8. Presumably these are electrons that would have been detected in detectors 9-12 if they had not preshowered in the kicker plates. Detector 8 is known to have the poorest light yield, so there is some contribution to these counts from Cerenkov prepulses in the lightguides. However, the elevation in pileup and distortion of the spectrum is also present in detectors 7 and 9, which indicates that the problem is not solely due to the prepulses in 8.

Another interesting feature is the wing like structures, observable in Fig. 55 that appear when one of the two pileup events has a low energy. These wings can also be seen as diagonal bands in the two dimensional plots of Fig. 54. These wings are simply a result of the underlying number density of particles exploding at low energies. The amount of pileup from these type of events increases exponentially around the software threshold, which was about 300 MeV in G2Too. Technically the software threshold was defined in terms of ADC counts. Perhaps, the fact that the wings are much smaller for detector 15 implies that the energy scale in that detector was such that 6 or 7 ADC counts corresponded to a higher energy decay electron than in other detectors where this



(a) Detector 15; Best Symmetry

(b) Detector 8; Worst Symmetry

(c) Detector 4; Most Missing Pileup

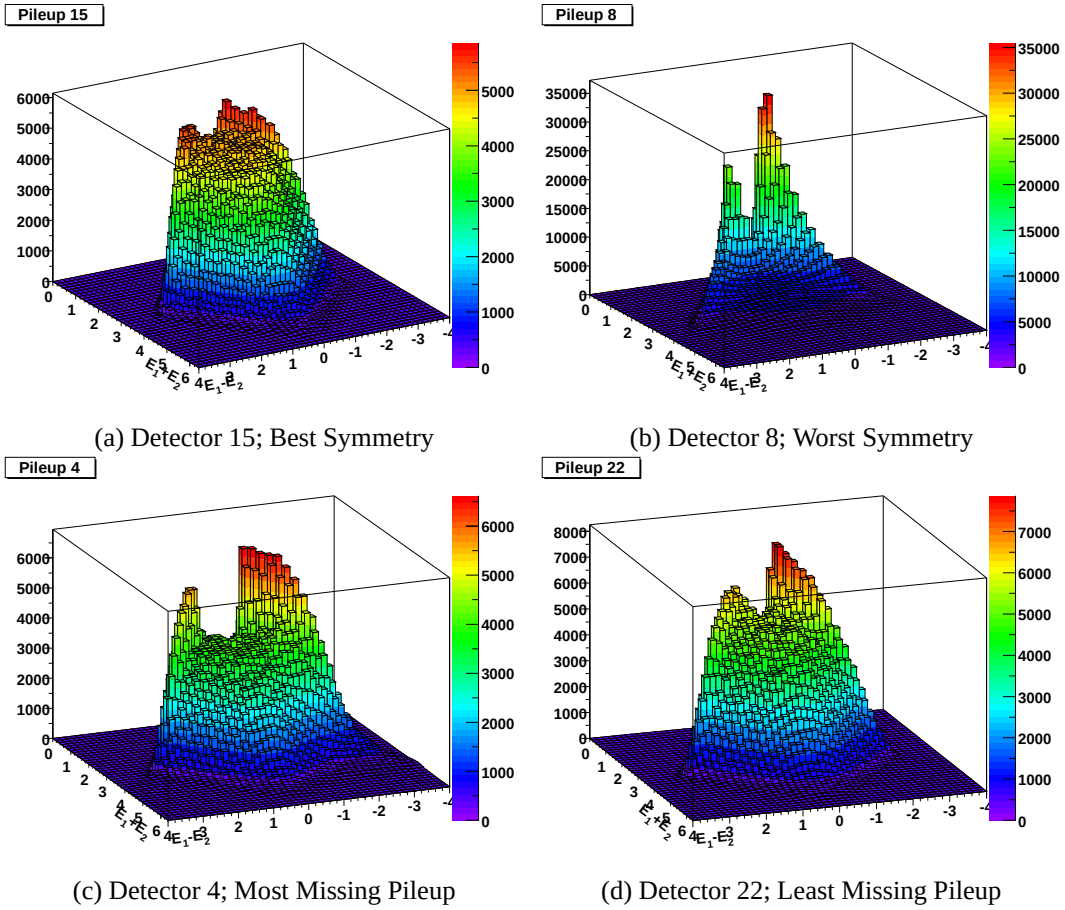
(d) Detector 22; Least Missing Pileup

Figure 54: Plot of $E_1 + E_2$ versus $E_1 - E_2$ for pairs of pulses used in constructing a pileup. Note that this is the same as the more standard E_1 versus E_2 plots; just rotated by 45 degrees.

explosion in counts is observed. However, this model is not very consistent with detector 4 whose gain was purposefully lowered, yet it still has large wings.

It is also apparent that there is an asymmetry in the wings. The left wing always has more counts than the right wing. This implies that we are much more efficient at finding pileup events containing a low energy decay electron, when that low energy electron is the first in the time-ordered pair. This simply means that the pileup window to the left finds more pulses around the software threshold than the window on the right. It seems likely that this is just an artifact of the pulse finder. Before the scintillator pulse starts, the island is very quiet and it is easier to construct a low energy pulse, while the pileup window on the right is buried under the tail of the scintillator pulse making it harder to construct these low energy decays. It is apparent that this is a substantial difference between the two pileup construction methods, since the Mediterranean construction assumes symmetry about the $E_1 = E_2$ axis and double counts the wing on the right. Although this difference looks large in the plots, one should keep in mind that only events where $E_1 + E_2 > E_{\text{threshold}}$ affect the data that is actually fit. Most of the asymmetry is below this threshold, although by looking at Fig. 54, one can see that there is still some asymmetry above the threshold.

As was mentioned before, the main problem with pushing the energy threshold lower is the inefficiency of the pileup construction below twice the hardware threshold. This inefficiency is what cause the notch at the bottom of the diamonds in Fig. 54. To further illustrate this point,



(a) Detector 15; Best Symmetry

(b) Detector 8; Worst Symmetry

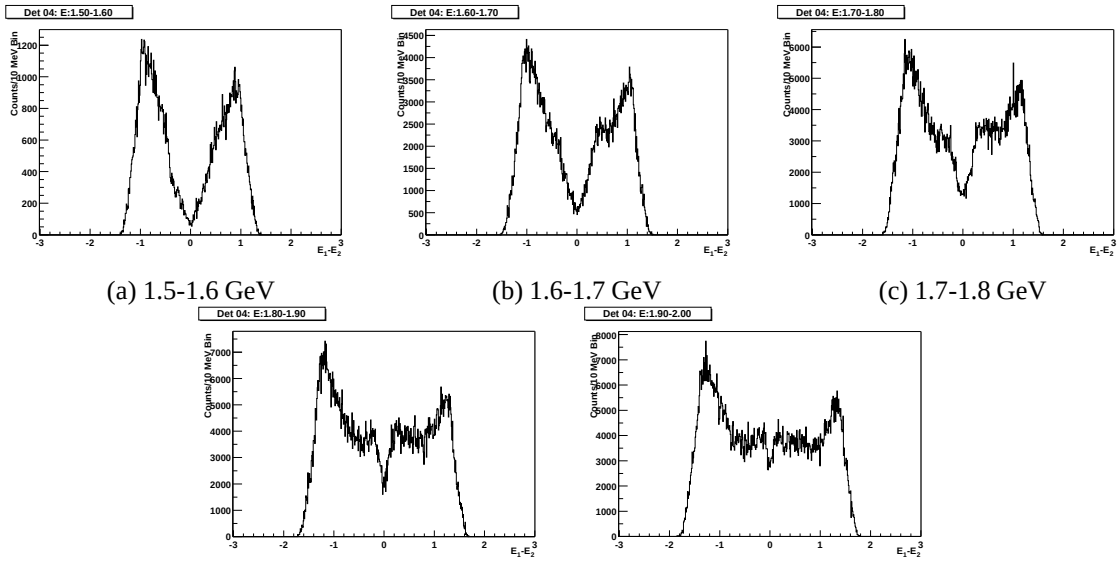
(c) Detector 4; Most Missing Pileup

(d) Detector 22; Least Missing Pileup

Figure 55: Three dimensional views of the same data shown in Fig. 54

horizontal dashed lines were drawn at the point where finding doubles events of energy $E_1 + E_2$ starts to lose efficiency. For detector four, the inefficiency start at about 2 GeV, while in detector 22 pileup construction is robust down to 1.5 GeV. In order to correct for the missing pileup horizontal slices of the plots in Fig. 54 were projected onto the x-axis and the fraction of missing pileup was interpolated. In fact, obtaining these projection was the whole motivation for the coordinate transformation. For detector 4, which has the highest hardware threshold by far, a series of these projections is shown for five 100 MeV wide energy bands covering the range from 1.5-2.0 GeV in Fig. 56.

The missing V-shaped notch at the centers of the distributions in Fig. 56 is the result of the hardware threshold inefficiency. In order to interpolate the fraction of missing pileup, reference spectra were built for detectors with thresholds set at 750 MeV. These reference histograms were scaled to match the spectra with missing pileup by matching the right side of the $E_1 - E_2$ distributions, i.e. using the region of $E_1 - E_2$ space where both are efficient. The fraction of missing pileup as a function of the energy of the double was extracted for each detector. With the corrections known, as pileup was being constructed, weights were assigned to each double event and its corresponding singles to account for the missing pileup. The plots in Fig. 57 show the weights for each detector as a function of the energy of the double.



(a) 1.5-1.6 GeV (b) 1.6-1.7 GeV (c) 1.7-1.8 GeV

(d) 1.8-1.9 GeV (e) 1.9-2.0 GeV

Figure 56: Missing pileup regions in detector 4 for various bands in the doubles energy.

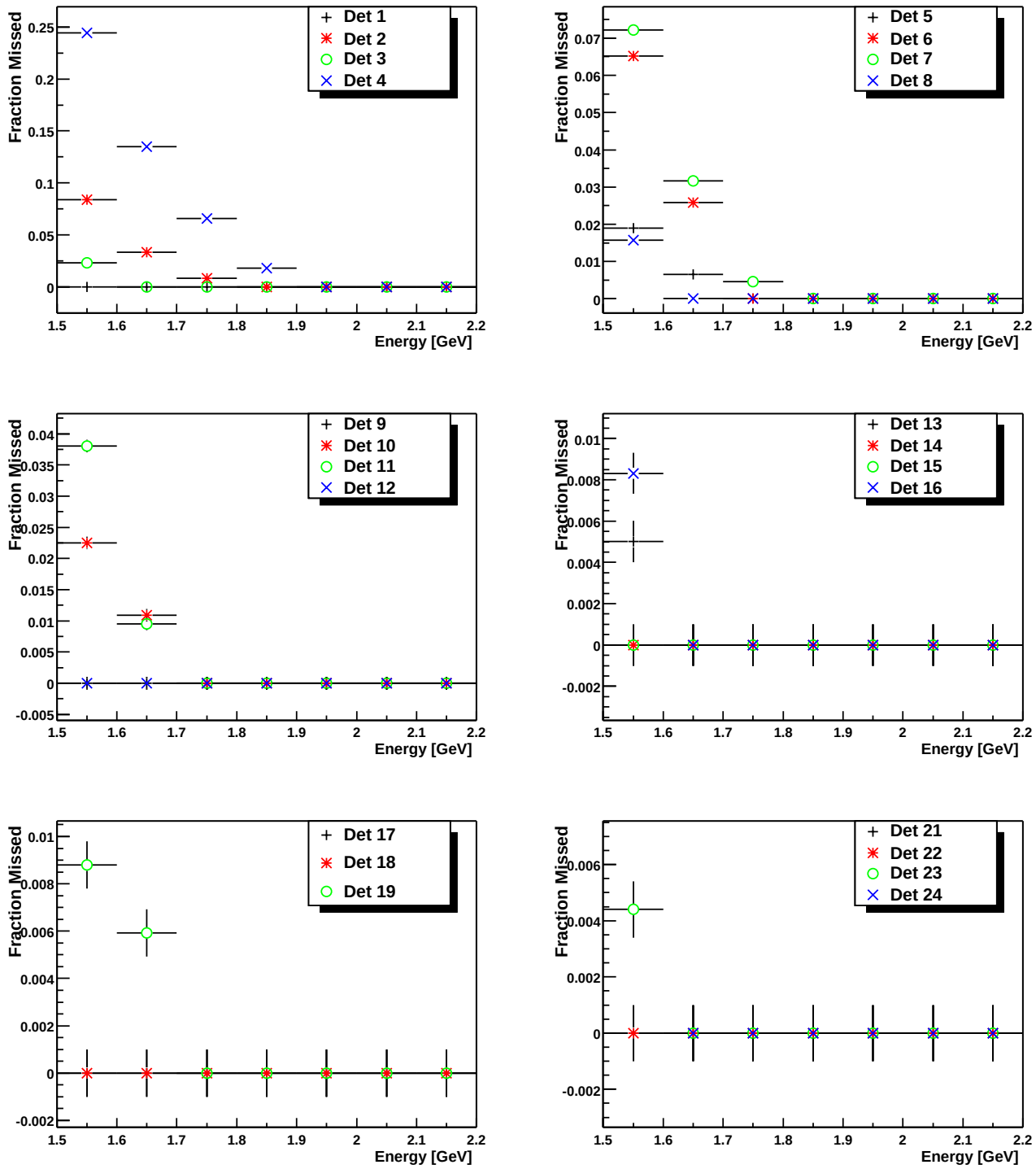


Figure 57: Fraction of pileup missing at low energies due to hardware threshold.

9 Evolution of the Fitting Function

In a multi-parameter fit, looking for improvements in the reduced χ^2 and examining the residuals are the primary methods for determining what terms are required in the fitting function. A basic 5-parameter fit (24),

$$N(t) = \frac{N_0}{\tau} e^{-t/\tau}[1 - A\cos(\omega_a t + \phi)], \tag{24}$$

characterizing the exponential decay of the muon and the $g-2$ oscillation, yields a reduced χ^2 that is clearly unacceptable. Looking at the Fourier transform of the residuals from such a fit to data that has been pileup subtracted, Fig 58, one can see a significant amount of very low frequencies corresponding to slow terms along with a dominant peak in the 400 kHz range that arises from the beating between the cyclotron frequency and the horizontal betatron oscillations. We will refer to this peak as the primary CBO peak since it has by far the largest amplitude. Satellite peaks appear to either side of the primary CBO peak due to the beating with the g-2 frequency. The only other peak present occurs due to the vertical cyclotron motion which we will return to later. The blue lines in the insets of Fig. 58 indicate where the g-2 harmonics reside. The plot from the 2000 data shows the primary CBO peak nearly centered on the second g-2 harmonic, while with the indices of refraction chosen in 2001 it is apparent that the peak has moved off of the harmonic. This greatly reduces the correlation between CBO terms and the g-2 frequency, thus making it possible to safely fit the data much earlier. It is interesting to note that the satellites and vertical terms are greatly reduced in the low n 2001 data relative to both the high n data and the data from 2000.

To account for the primary CBO peak, an overall multiplicative term is included in the fit,

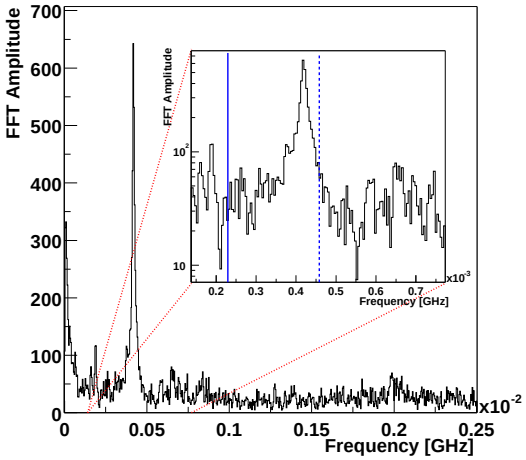
$$N(t) = \frac{N_0}{\tau} C(t)[1 - A\cos(\omega_a t + \phi)], \text{ where} \tag{25}$$

$$C(t) = 1 - e^{-t/\tau_{cbo}}[1 - A_1 cos(\omega_{cbo} t + \phi_1)]. \tag{26}$$

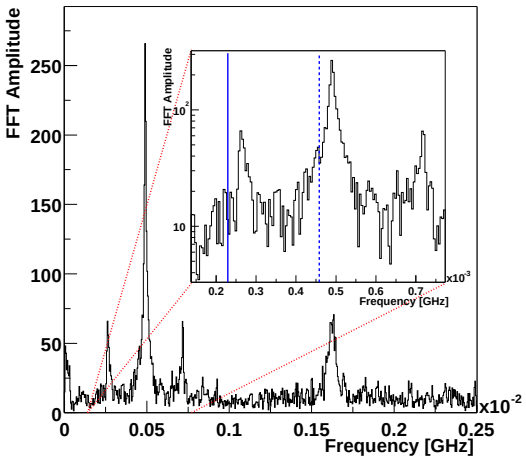
This terms accounts for the modulation in counts as the beam oscillates horizonatally off of the end of the calorimeter. It can be shown mathematically that this term removes both the primary peak and the satellites. In the 2000 data analysis it was necessary to build the CBO envelope, but it has been shown in Mario's report that the constructed envelope is closely approximated by an exponential in 2001. Fig. 59 shows the χ^2 versus fit start time after a series of modifications to the data and the fitting function. Including the CBO term has a much larger effect on χ^2 in the high data. In order to extend the fit earlier than 30 μs, detectors which are not gated on (3-6) are not included in the fits. It should be noted that the data used for generating Figures 58 and 59, as well as the rest of the analysis presented in this section, has already been corrected for gain and uses uniform weighting in the 1.8 to 3.3 GeV range. The plots are for one random seed, so the asymptotic value of the reduced χ^2 at late times need not be one.

After correcting the data for gain and performing pileup subtraction there is still a significant deviation from the expected decay lifetime of 64 μs due to muons being lost from the storage ring. Despite intentionally losing particles from the periphery of phase space during scraping, there is still a significant early-to-late change in the flux of lost muons. Details of the lost muon construction and the mathematical derivation of the implementation in the fitting function are given in the section on muon losses. With the loss term in, the fitting function can be described as,

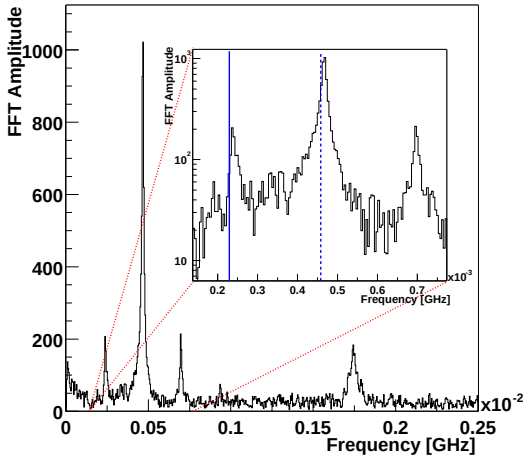
$$N(t) = \frac{N_0}{\tau} \Lambda(t) C(t)[1 - A\cos(\omega_a t + \phi)], \text{ where} \tag{27}$$



(a) 2001 Low n Value (n=0.121)



(b) 2001 High n Value (n=0.141)



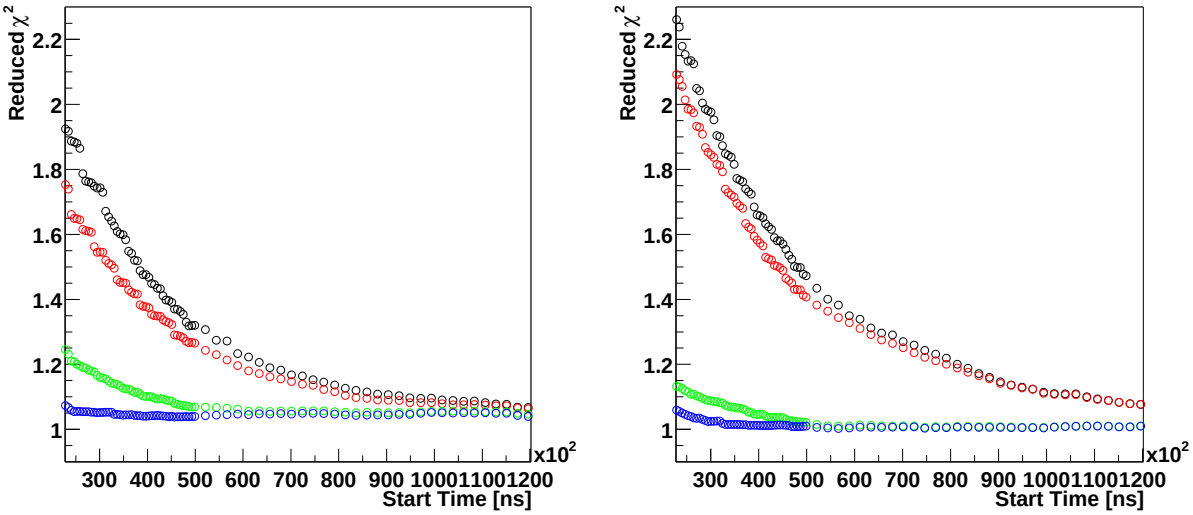
(c) 2000 Period 4 (n=0.131)

Figure 58: Fourier transform of the residuals after a 5-parameter fit starting at 25 μs after injection with data that has been pile-up subtracted and corrected for gain. In the enlarged regions, a blue solid line indicates the g-2 frequency, while the dashed blue line is drawn at twice the g-2 frequency.

$$\Lambda(t) \quad = \quad 1 - A_{loss} \exp \frac{-t_0}{\tau} \int_{t_0}^{t} L(t') \exp \frac{t'}{\tau} dt'. \tag{28}$$

From Fig. 59 it is evident that the muon loss term significantly improves the quality of the fit at early times. With this correction to the fitting function, the stability of the χ^2 is close to being acceptable.

In analyzing the 2000 data, it was recognized both in the data analysis and simulation work that CBO affects more than just the number of counts in the histogram. Because of the radial dependence of the decay electron energy and the correlation of that energy to the asymmetry of the muon decay, the average g-2 asymmetry is also modulated at the CBO frequency. The average time of flight also changes due to the stored beam oscillating between then inner and outer radii of the ring, and also due to the direction of the muon spin tipping radially in and out. Both result in



(a) Low n Value (n=0.121)

(b) High n Value (n=0.141)

Figure 59: Improvement in the χ^2 versus start time. Initial 5-parameter fit to gain corrected data (black), followed by inclusion of pileup subtraction (red), incorporation of a CBO term (green), and correcting for muon losses (blue).

a modulation in the g-2 phase at the primary CBO frequency. The fitting function must therefore include these effects by introducing a time dependence to A and ϕ,

$$N(t) = \frac{N_0}{\tau}\Lambda(t)C(t)[1-A'(t)\cos(\omega_a t+\phi'(t))], \text{ where} \quad (29)$$

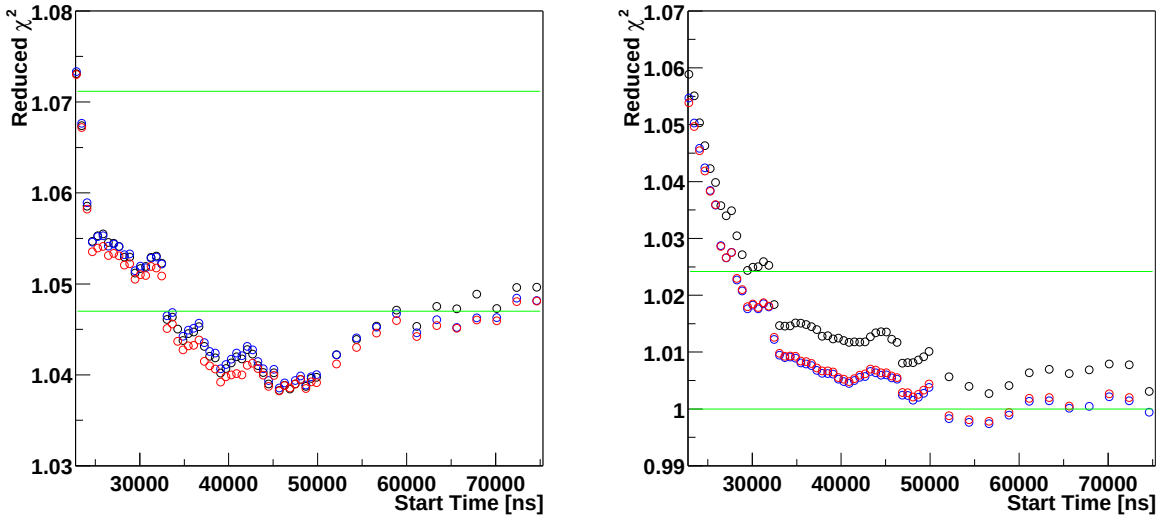
$$A'(t) = A(1-e^{-t/\tau_{cbo}}[1-A_2\cos(\omega_{cbo}t+\phi_2)]), \text{ and} \quad (30)$$

$$\phi'(t) = \phi(1-e^{-t/\tau_{cbo}}[1-A_3\cos(\omega_{cbo}t+\phi_3)]). \quad (31)$$

As can be see from Fig. 60, only the amplitude modulation has a significant effect on the χ^2, and then only in the high n data. Although the other terms do not have much of an impact on the χ^2 when fitting a large number of detectors simultaneously, for completeness they are included in the fit. Their effect becomes more pronounced when trying to fit individual detectors due to the lack of cancellation. The green lines in Fig. 60 indicate the statistically allowed deviation in the reduced χ^2. When fitting from 25 to 550 μs after injection with 150 ns bins, there are approximately 3500 degrees of freedom in the fit. The allowed deviation in the reduced χ^2 is,

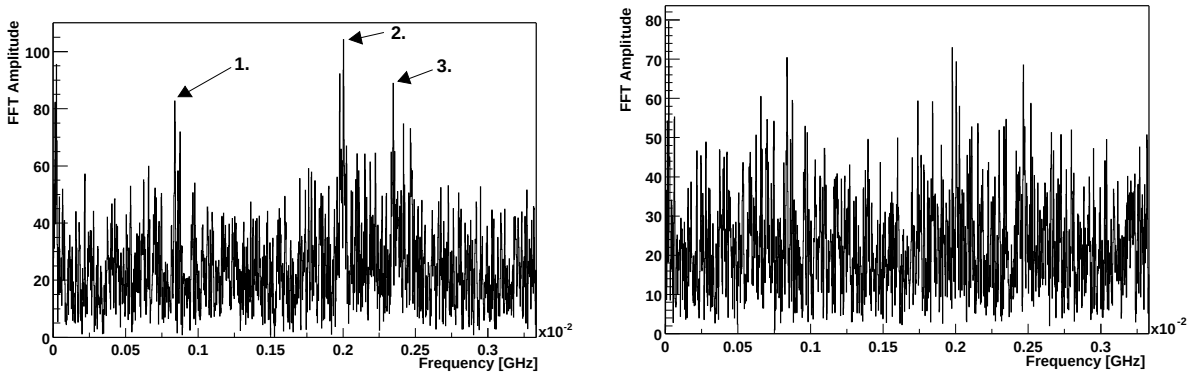
$$\begin{aligned}\sigma_\chi &= \sqrt{2/d}, \\ &= \sqrt{2/3500}, \\ &\approx 0.024.\end{aligned} \quad (32)$$

At this point, the fit function is what one would consider to be the traditional 'full physics' fit from 2000. The early-to-late χ^2 stability is already quite good at 30 μs, but at earlier times there is a clear upward trend. To determine the source of the problem we can once again turn to the residuals after fitting with the full physics function, as shown in Figures 61 and 62 where



(a) Low n Value (n=0.121) (b) High n Value (n=0.141)

Figure 60: Improvement in the χ^2 versus start time. Initial fit without $g-2$ amplitude and phase modulations, followed by inclusion of the asymmetry modulation (blue), and with both asymmetry and phase modulations (red).

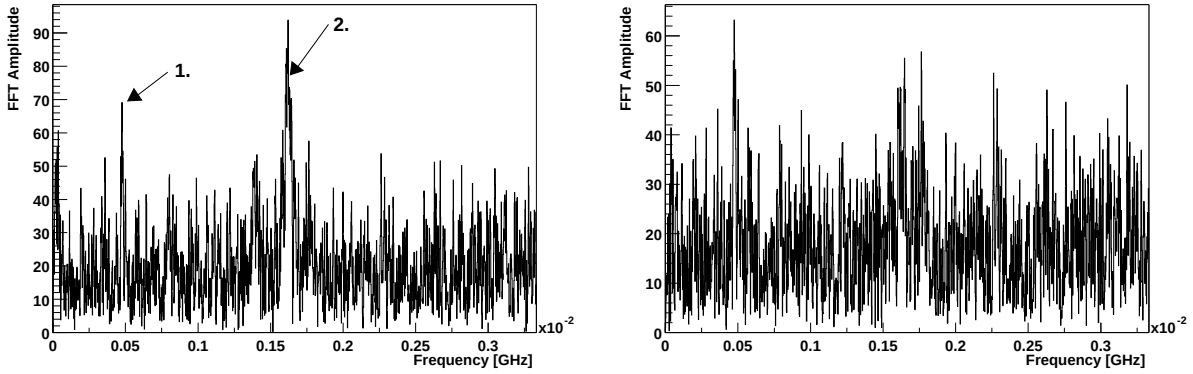


(a) Low n, Start time = 23 μs (b) Low n, Start time = 30 μs

Figure 61: Fourier transform of the residuals left after fitting the low n data with the full physics function. The remaining peaks are double CBO (1.), vertical waist (2.), and vertical betatron (3.).

the dominant peaks have been labeled. When the residuals are calculated starting at 23 μ, it is apparent the peak labeled 2 in both n values has grown relative the the FFT taken at 30 μs. This peak corresponds to the difference between the cyclotron frequency and twice the vertical betatron frequency. Commonly referred to as the vertical waist peak, it has long been thought to arise from a vertical breathing mode of the beam. Another explanation is that it arises from the vertical motion itself and is manifested at twice the vertical betatron frequency as each vertical oscillation drops the decay electrons off the top and then the bottom of the calorimeter. Regardless of its source, the effect of including a term in the fit of the form,

$$N(t) = \frac{N_0}{\tau}\Lambda(t)V(t)C(t)[1-A'(t)\cos(\omega_a t+\phi'(t))], \text{ where} \quad (33)$$

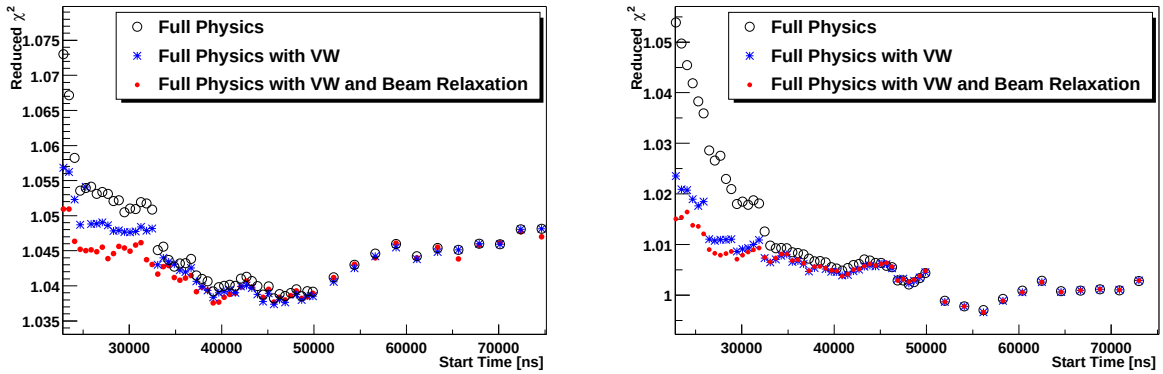


(a) High n, Start time = 23 μs (b) High n, Start time = 30 μs

Figure 62: Fourier transform of the residuals left after fitting the high n data with the full physics function. The remaining peaks are primary CBO (1.) and vertical waist (2.)

$$V(t) \quad = \quad (1 - e^{-t/\tau_{vw}}[1 - A_{vw}\cos(\omega_{vw}t + \phi_{vw})]), \tag{34}$$

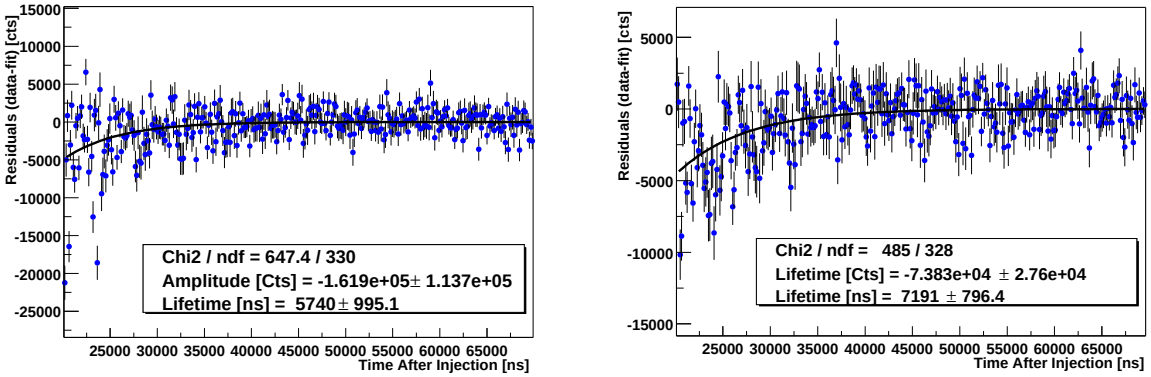
reduces the χ^2 at early times, as shown in Fig. 63.



(a) Low n Value (n=0.121) (b) High n Value (n=0.141)

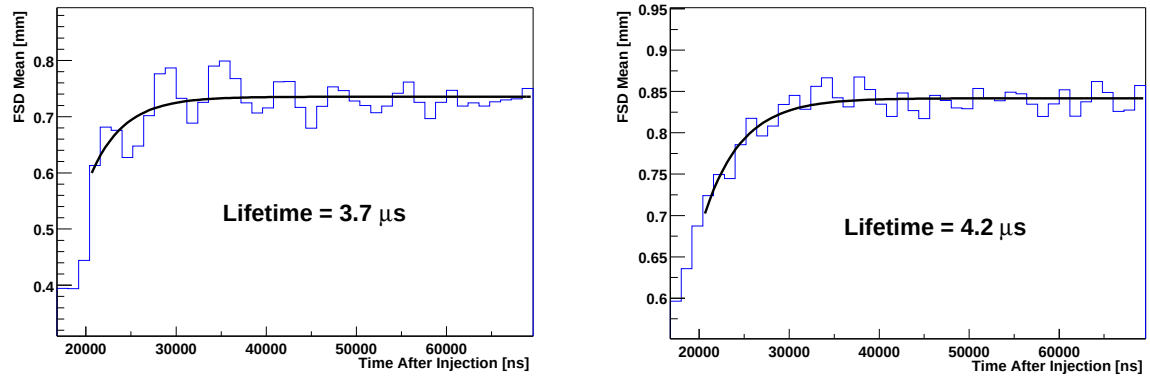
Figure 63: Improvement in the χ^2 versus start time. Initial fit with the traditional full physics function (black), followed by including the vertical waist term (blue), and finally with a term accounting for the relaxation of the beam after scraping (red).

Also shown in Fig. 63, is the improvement in the χ^2 from including a term in the fitting function that accounts for the beam relaxation after scraping. After scraping is turned off, the beam will relax vertically back to the center of the storage ring with an expected lifetime of about 5 μs. This can be illustrated by looking directly at the residuals or by looking at the movement of the beam on the FSDs. Fig. 64, shows the residuals after fitting a traditional full physics function with vertical terms to the data. Applying an exponential fit to the residuals yields a lifetime of 6 and 7 μs for the low and high n values, respectively, with an error of about 1 μs. The drop at early times implies an absence of expected events, which is consistent with the decay electrons missing the calorimeter before the beam has settled. A similar behavior is observed by looking at the change in the mean vertical position of the decay electrons on the FSDs, shown in Fig. 65, where it is clear that up



(a) Low n Value (n=0.121) (b) High n Value (n=0.141)

Figure 64: Residuals after fitting a traditional full physics function with vertical terms. An exponential fit to the residuals shows a 6 μs time constant for the beam relaxation after scraping.



(a) Low n Value (n=0.121) (b) High n Value (n=0.141)

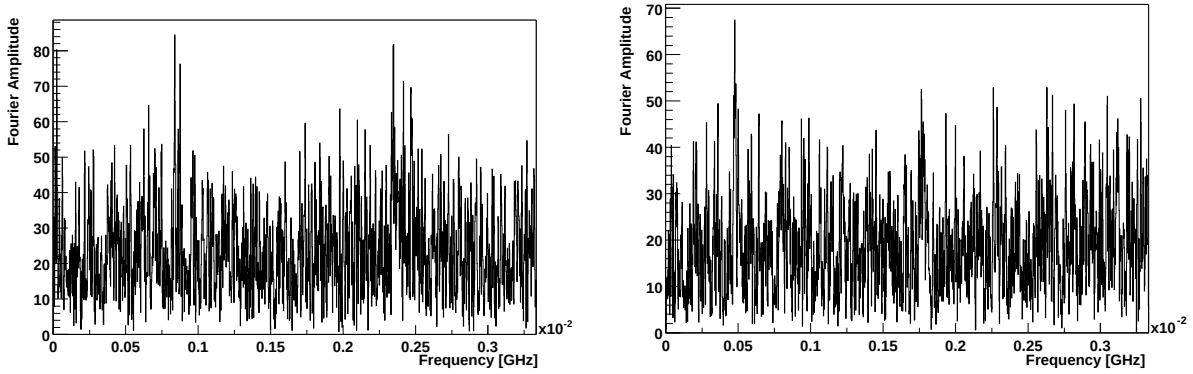
Figure 65: Mean position of the decay electrons striking the FSDs. An exponential fit yields a time constant of about 4 μs for the beam relaxation after scraping.

until 30 μs after injection the beam is still moving. The fitted lifetime comes out at about 4 μs for both n values. To account for the beam relaxation, a term with a fixed 5 μs lifetime is incorporated into the fit,

$$N(t) = \frac{N_0}{\tau}\Lambda(t)V(t)B(t)C(t)[1-A'(t)\cos(\omega_a t+\phi'(t))], \text{ where} \quad (35)$$

$$B(t) = 1-A_{br}e^{-t/\tau_{br}} \text{ with } \tau_{\rm br} = 5\mu{\rm s}. \quad (36)$$

Fitting with equation (35), the early-to-late stability of the χ^2 is now acceptable, see Fig. 63. Furthermore, a Fourier transform or the residuals taken at a very early time of 23 μs after injection shows only a few small peaks remain. Looking at Fig.66, the only peaks discernible in the low n data correspond to double CBO (twice the primary CBO frequency) and the direct vertical betatron oscillation at 2.3 kHz. The high n data still has some structure around the primary CBO and vertical waist peaks. This could be an artifact of not having quite the right CBO envelope. All of these remaining peaks are small enough to be neglected in the fit and assigned as a systematic error



(a) Low n, Start time = 23 μs (b) High n, Start time = 23 μs

Figure 66: Fourier transform of the residuals after fitting the data with the early physics function.

later. To avoid confusion, equation (29) will be referred to as the full physics function, as was the terminology in the 2000 analysis. The expanded function in equation (35) will be referred to as the early physics function.

10 Fitting Results and Consistency Checks

10.1 Start Time Optimization

Even though the function in equation (35) is capable of providing an acceptably stable reduced χ^2 at early times, there are still many other consistency checks that need to be satisfied. Start time scans on other parameters should be examined, along with fits by detector and energy band which might elucidate other problems. Data can also be broken up by run number under many different schemes, i.e. n value, run period, pre vs post radial field change, etc. Since equation (35) only contains one CBO frequency, it is necessary to split the data by n value.

Furthermore, if one wants to take full advantage of the statistical power of the dataset, the different gate-on times of the detectors imposes a second division of the data. Table 6 shows the gate-on times for the detectors in 2000 and 2001. In both years there were 6 different gating groups, but in 2001 the latest two gating groups were turned on much earlier. This enables a fit to all of the detectors to start as early as 31.6 μs after injection, but with the early fitting function there exists the possibility to fit much earlier. One could argue that each gating group or detector could just be fit separately. In addition to the overhead of working with so many different fits, one loses the natural reduction or cancellation of systematic deviations from the ideal fitting function that comes with summing detectors. Furthermore, less g-2 cyclse can be fit which increases the correlation between R and ϕ. With that philosophy, my personal decision was to split the dataset into at most two detector groups.

It should be pointed out that the times in Table 6 are actually the time of the first g-2 zero crossing after the detectors are turned on. As has been shown in every past g-2 analysis, slow terms result in phase pulling of the g-2 frequency. Due to imperfections in the muon loss construction, gain correction, and pileup subtraction, residual slow terms are unavoidable, but the frequency is least biased if results are only quoted at g-2 crossings. In fact, it is this phase pulling of R that

2000			2001	
Gate-On Time [μs]	Detectors	Group	Gate-On Time [μs]	Detectors
46.9	4,5	a	31.6	4,5
33.8	3,6	b	25.1	3,6
22.9	7,8	c	22.9	2,7,8
14.2	1,9-12	d	12.0	1,9-12
9.8	13-16	e	9.8	13-16
5.4	17-24	f	7.6	17-24

Table 6: Earliest times at which each detector can be used when fit is constrained to start at g-2 zero crossings.

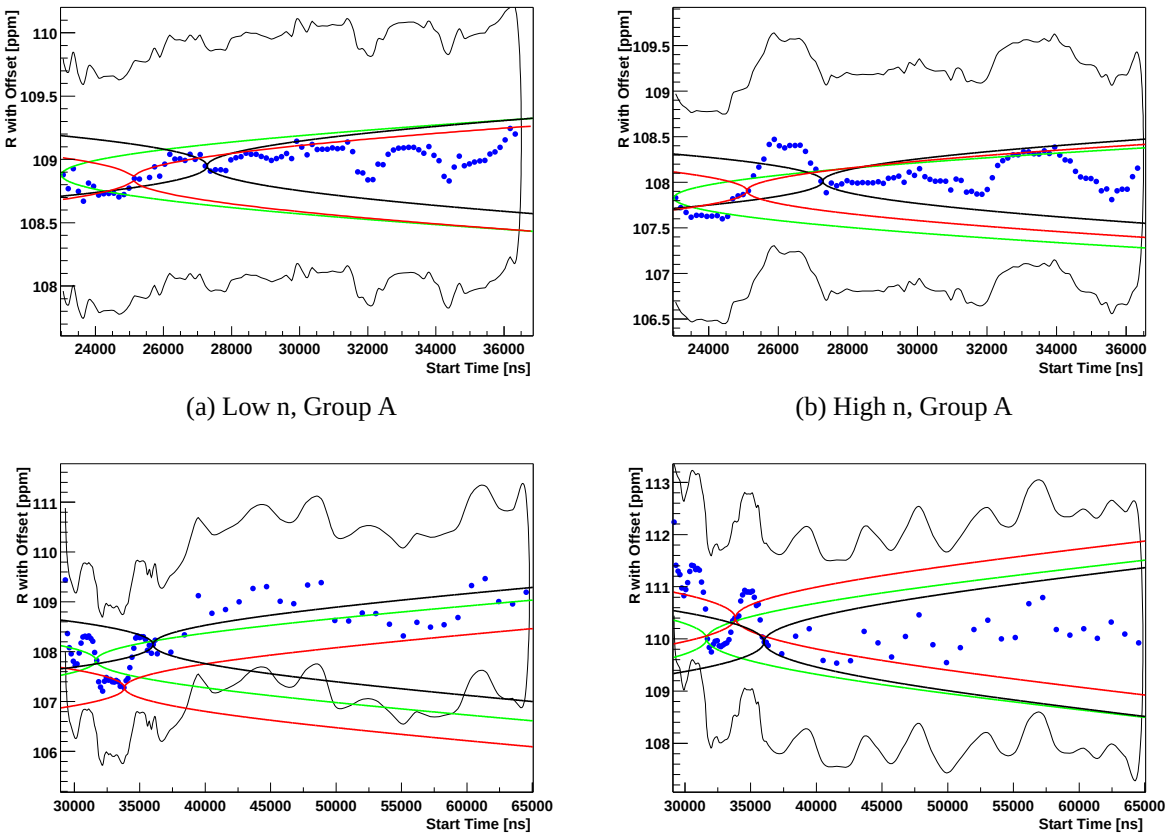
in the end determined the choice of detector subgroups. At first, the goal was to fit ab separately from cdef, see Table 6 for a definition of which detectors are included. However, the severe phase pulling precluded this division. Instead, the detectors were divided into the groups abc and def. The optimized gating groups abc and def, will be referred to respectively as gating groups B and A. Even with these subgroups, it is still the phase pulling that limits the start time in a multi-parameter fit. Although the χ^2 is stable enough to quote A as early as the zero crossing at 22.9 μs, it can be seen from the start time scan in Fig. 67 that R is erratic. The Kawall bands are drawn at the g-2 zero crossings, and R is not really stable until the black Kawall band drawn at 27.3 μs. As can be seen in Fig. 67, the stability of R for detector group B is much worse. This is probably caused by a combination of the late gates, the flash, and the enhanced CBO amplitude in detector 8 which is included in the group. Due to the phase pulling, the start time for group B was actually moved out by one g-2 cycle from the start time of 31.6 μs achievable if all of the detectors are fit simultaneously. One way to look at this optimization procedure is to say one g-2 cycle of data was sacrificed from seven detectors in order to gain one g-2 cycle in the remaining 16 detectors. It is assumed that any residual phase pulling will be taken care of in the systematic errors assigned to gain, pileup subtraction, and muon losses.

10.2 Fitting Procedure and Results

With the fitting function, detector subgroups, energy range, and start times well defined, we can now start looking at the numerical results from the fits. The R start time plots in Fig. 67 were generated from data that has been asymmetry-weighted, but in general a traditional fit will be maintained for comparison purposes. The particular details of the asymmetry-weighted and traditional fits are given in Table 7. Both fits use the same runlist, energy endpoints, gain corrections, and muon loss construction.

The procedure used in fitting the data is very similar for both fits. The only difference is in the definition of the weight with which each event enters the histogram, and the corresponding error. In general, the procedure used was:

1. Runs from the 'silver' runlist are analyzed.
2. A 5 ns deadtime was enforced on each detector.
3. Logashenko coefficients were used when combining events for either the enforced deadtime or constructed pileup doubles.
4. Times are randomized using fill randomization with a 149.197 ns period.



(a) Low n, Group A

(b) High n, Group A

(c) Low n, Group B

(d) High n, Group B

Figure 67: R start time scans with Kawall bands plotted at various g-2 zero crossings.

	A-Weighted	Traditional
Weighting	A	1
Number of Detector Groups	2	1
Detectors In A	2-8	N/A
Detectors In B	9-19,21-24,1	N/A
Start Times	27.3(A) and 36.0(B) μs	31.6 μs
Stop Time	650 μs	650 μs
Energy Range	1.5-3.3 GeV	1.8-3.3 GeV
Vertical Terms	Yes(A) No(B)	No
Beam Relax Terms	Yes(A) No(B)	No
Pileup Subtracted	Yes	Yes
Gain Corrected	Yes	Yes

Table 7: Details of the two fits performed on the data.

5. The energy of each event is modified by a gain factor determined for each detector versus time in spill.
6. The gain factor consists of the usual factor times a factor that corrects for the gain versus WFD clock tick dependence.

7. Histograms for N, S, and D are filled with a 149.2 ns bin width.
8. Events are entered into the N histogram with a weight of one (traditional) or a weight of A (asymmetry-weighted).
9. A correction factor f(E) for pileup events with energies less than twice the hardware threshold was constructed, and appropriate event in S and D were filled with a weight f(E)*A. Only applicable to asymmetry-weighted data where energies down to 1.5 GeV are used.
10. The error for each bin is kept track of via recording the sum of weights-squared.
11. Pileup is constructed from symmetric pileup windows in the range [-10 ns,-5 ns] and [5 ns,10 ns] around the trigger pulse.
12. The error on the pileup-subtracted histogram is taken by adding the errors from N, S, and D in quadrature. This is not quite correct since it does not take correlations into account.
13. The χ^2 is calculated using the error on each histogram bin as the variance.
14. Minuit is used with the following sequence of commands:

```
gMinuit->Command("SET STRATEGY 1");
gMinuit->Command("MIGRAD 20000 0.1");
gMinuit->Command("SET STRATEGY 2");
gMinuit->Command("HESSE");
gMinuit->Command("MIGRAD 20000 0.0001");
gMinuit->Command("HESSE");
gMinuit->Command("IMPROVE 20000");
gMinuit->Command("HESSE");
gMinuit->Command("MINOS 20000 5");
```

15. Equation (35) is used on all fits with start times prior to 33 μs. For later times, the vertical waist and beam relaxation terms are neglected.
16. Only the lifetime of the beam relaxation (5μs) is fixed.

The R values obtained by applying the above procedure are summarized in Table 8. The results have been averaged over 5 random seeds. The agreement between the regular and the asymmetry-weighted analysis methods is quite good. In the simulation results section, **??**, it is shown that the allowed difference is about one half of the error. This allowed deviation is due to asymmetry-weighting and lowering the threshold to 1.5 GeV, but it does not included the allowed deviation from the different start times. Moving the start time by 1 g-2 cycle, as was the case here, has an allowed deviation of one fourth of the error. Adding the two allowed deviations in quadrature, and assuming an error of 0.63 ppm, implies a total allowed deviation of 0.54 ppm. Therefore, the results are consistant at the 0.7σ level.

	Asymmetry Weighted						Traditional	
	Group A		Group B		Combined Groups			
n Value	R	δR	R	δR	R	δR	R	δR
Low n	109.11	0.95	107.93	1.53	108.78	0.81	108.44	0.90
High n	107.91	1.17	109.72	1.90	108.41	1.00	108.03	1.12
Combined n					108.63	0.63	108.28	0.70

Table 8: Fit results for R and δR, all units are in ppm with the common offset applied.

	Asymmetry Weighted					
	Group A		Group B		Traditional	
n Value	χ^2	$\Delta\chi^2$	χ^2	$\Delta\chi^2$	χ^2	$\Delta\chi^2$
Low n	1.040	0.0047	1.012	0.012	1.023	-0.0015
High n	1.020	0.012	1.046	0.013	1.029	0.0116

Table 9: Reduce χ^2 and the change in the χ^2 when the start time is extended by 64 μs.

See Tables 23 and 24 in Appendix C for the correlation coefficients from the fits.

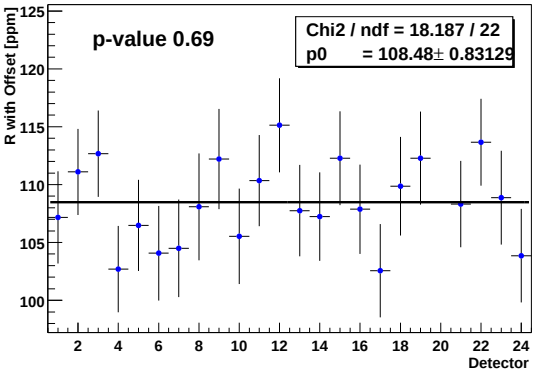
10.3 Results By Detector

The consistency of fits by detector was studied for both weighting schemes at a start time of 31.7 μs. The asymmetry-weighted plots include the extra energy range from 1.5-1.8 GeV. One important point to remember about these type of plots is that the random seed can cause fluctuations of 10% of the error bar. Therfore, a 10% error should be added in quadrature to each statistical error before checking the consistency of a fit to a constant. The other approach, which is used here, is to average the results from multiple random seeds. Here I use 5 random seeds and plot only the statistical error. There is still a possiblity that the errrors could be slightly inflated by including systematic effects, i.e. CBO is much higher in detectors 7, 8, and 9.

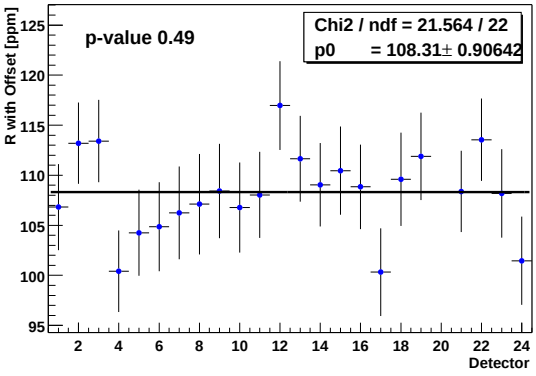
Fig. 68 shows the R plots versus detector number for the n values separately and then combined. The p-value shown on each plot is the probability that one would have found a worse χ^2 when testing the hypothesis that the data by detector can be described by a fit to a constant. In the limit of a large number of degrees of freedom, this can be consider the integral of the one-sided tail of a Gaussian distribution. The combined plots are not the results from a fit to the combined data, rather they are the statistical union of the plots for the high and low n data.

The consistency looks pretty good for all of the fits to separate n values, with the exception of the w=A fit to the high n data. Here we would expect to have had a worse χ^2 a mere 4% of the time. Excluding the possiblity of a statistical fluke for the moment, there are two possiblities. The w=A method might be systematically biased or there is an underlying bias and the lower error in the asymmetry-weighted analysis is more sensitive. Of course the other possiblity is that it is just a statistical fluke. The first question, one might ask is how well the w=A data agrees with the w=1 data in the high n run period. Fig. 69 shows the overlay of the two high n plots and their difference. The errors on the difference are taken to be half of the asymmetry-weighted errors, which was confirmed by simulation. The χ^2 from a fit to a constant is nearly one, which indicates that the w=1 and w=A analyses are statistically compatible. Just to be fair, the overlay for the low n data is also shown in Fig. 70.

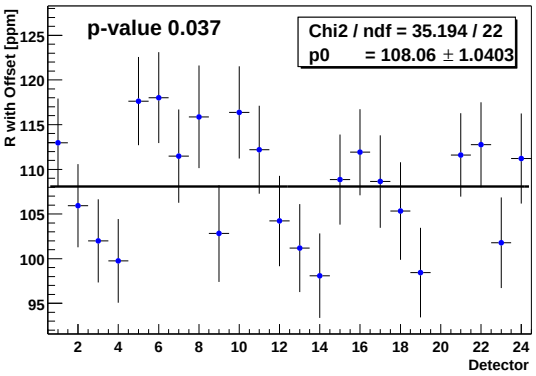
It is interesting to note that when the n values are combined, the consistency for both n values increases substantially. It could just be statistics or it could indicate an anti-correlation between the n values. To explore this hypothesis, the difference between the high and low n R valuse is plotted versus detector for both weighting schemes in Fig. 71. In the w=A analysis, the probability of consistency between the n values is only 3%. Interestingly, in the w=1 analysis, the probability is also fairly low at 4.6%.



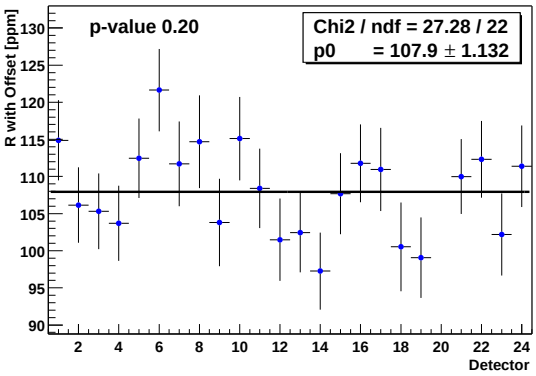
(a) w=A, Low n



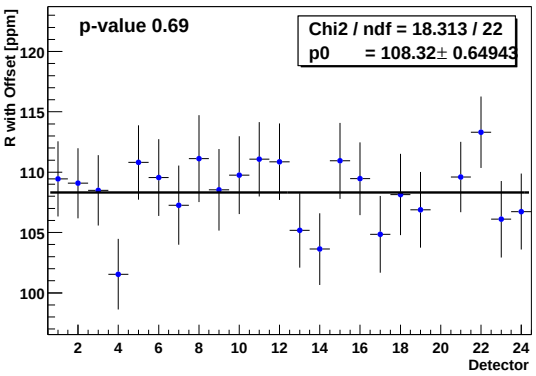
(b) w=1, Low n



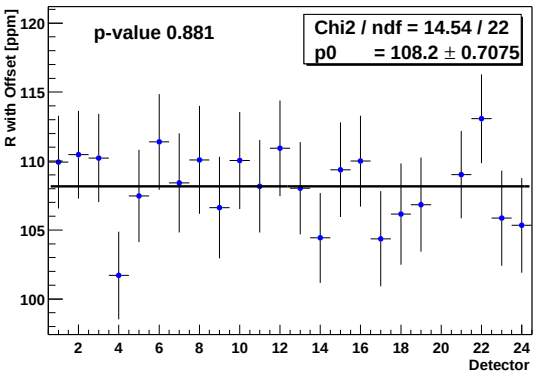
(c) w=A, High n



(d) w=1, High n



(e) w=A, Combined n

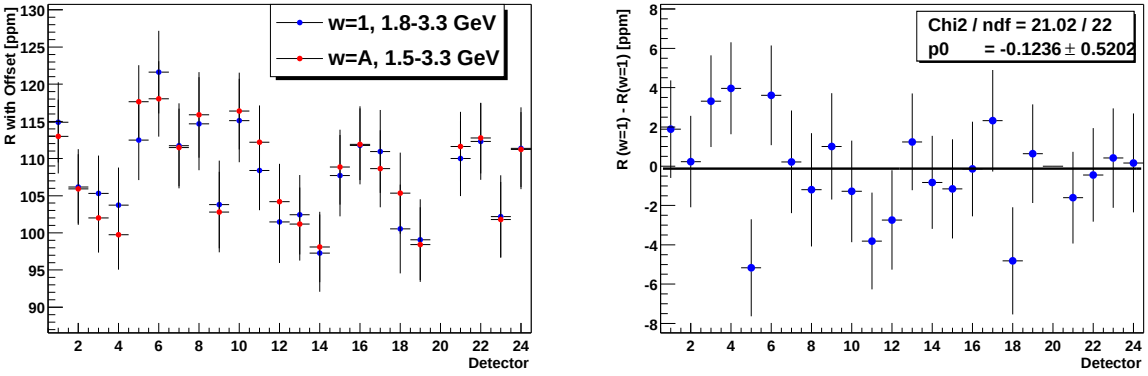


(f) w=1, Combined n

Figure 68: R versus detector for asymmetry and uniform-weighted data. Plots are shown for each n value separately and combined.

10.4 Results By Energy

The consistency by energy band is shown in Fig. 72 for both weighting schemes. As in the plots by detector, an average of 5 random seed is used. Due to the widely fluctuating parameter versus energy, it is hard to ensure that every fit converges. Therefore, some points in the plots are actually averaged over fewer random seeds. As has been shown, in the past the consistency in the high n



(a) w=A, w=1 Overlay (b) w=A, w=1 Difference

Figure 69: Overlay of w=A and w=1 R versus detector plots and their difference for the high n data.

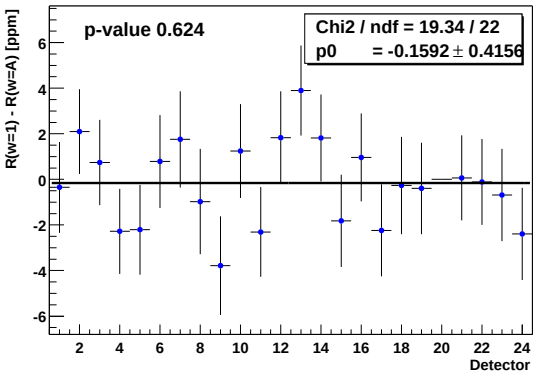
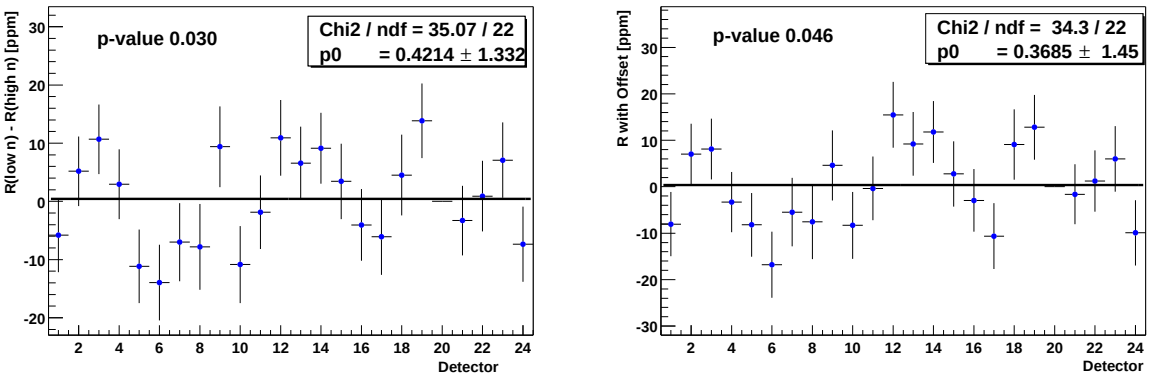


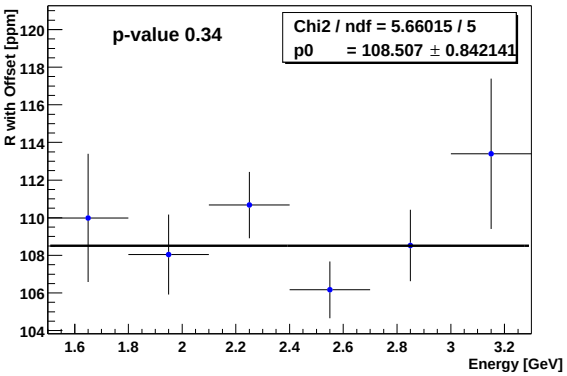
Figure 70: Overlay of w=A and w=1 results for the low n run period.

data is not as good as in the low n data. Based on previous results, the inconsistency does not look as dramatic as it has in the past. This occurs as a result of having a smaller number of degrees of freedom in the histogram, using the p-value instead of assuming a Gaussian, and averaging over more random seeds.

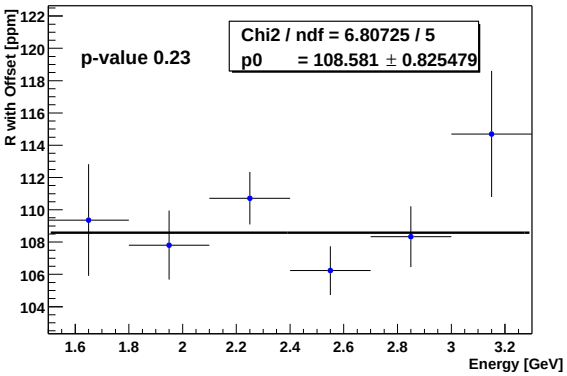


(a) w=A (b) w=1

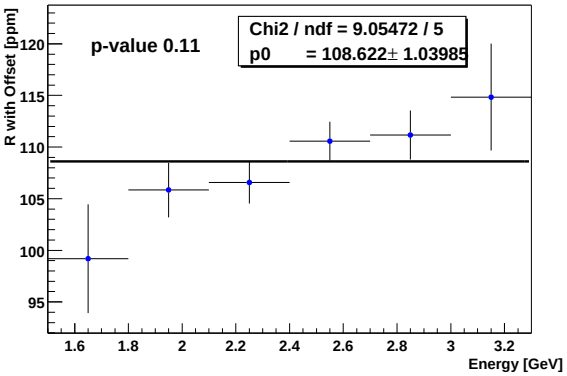
Figure 71: Difference between high and low n R values by detector.



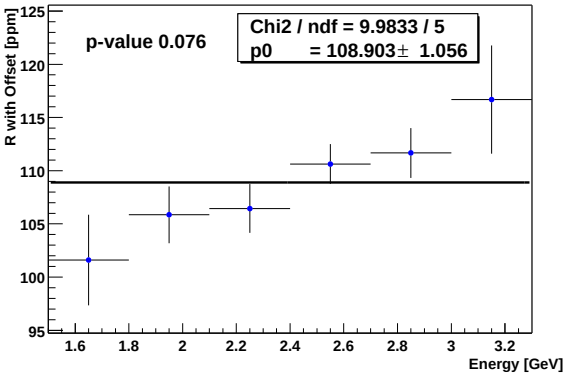
(a) w=A, Low n



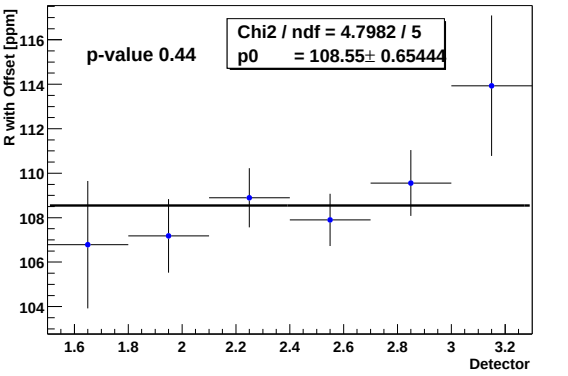
(b) w=1, Low n



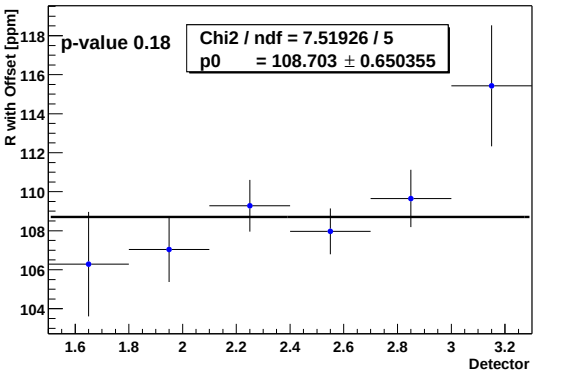
(c) w=A, High n



(d) w=1, High n



(e) w=A, Combined n



(f) w=1, Combined n

Figure 72: R versus energy for asymmetry and uniform-weighted data. Plots are shown for each n value separately and combined.

11 Systematic Studies

Systematic studies were performed for an asymmetry-weighted and a traditional analysis. The details of the two methods are given in Table 7, but the main differences are summmarized below.

- Histograms are filled with w=1 or w=A.
- The asymmetry-weighted analysis use the 1.5-1.8 GeV range.
- The traditional analysis fits all detector simultaneously at 31.7 μs, while the asymmetry-weighted analysis pushes the start times later by one $g-2$ cycle for detectors 2-8 and earlier by one $g-2$ for the remaining detecotors.

	Low n	High n	Combined	Assumptions
Pileup	0.048	0.069	0.055	+/-15%, +/-2 ns
Gain	0.072	0.033	0.056	+/-40%
CBO	0.06	0.15	0.09	gaus vs expo
Random Seed	0.03	0.03	0.03	5 seeds
Muon Losses	0.08	0.08	0.08	loss shape, Jon sim
Vertical Waist	0.02	0.04	0.03	+/-15μs,+/-8 kHz
Beam Relaxation	0.05	0.05	0.05	on/off
Fitting Procedure	0.03	0.03	0.03	Simulation
Flashlets	0.02	0.02	0.02	from Xiaobo
Bin Width	0.06	0.06	0.06	from Fred
Other CBO	0.03	0.03	0.03	from Jon
Bin Width	0.1	0.1	0.1	Vanya/Bill
Total			0.20	quadrature

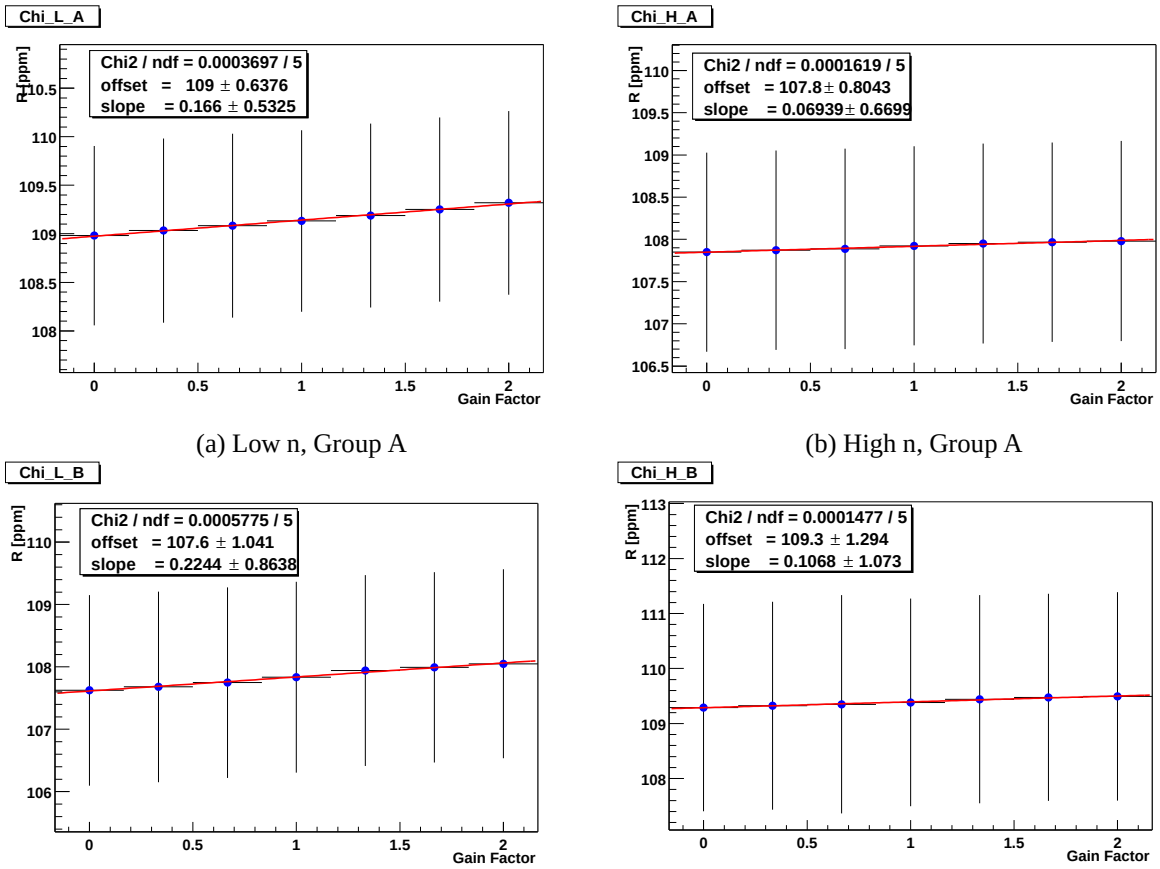
Table 10: Summary of systematic errors for the asymmetry-weighted analysis.

11.1 Gain Systematic

The equation for the gain corrections applied to the data may be written as,

$$g(t) = 1 + f \cdot \delta(t), \tag{37}$$

where f is a factor that modifies the magnitude of the gain correction. If $f = 1$, then the applied gains are just those constructed from the data, while $f = 0$ would imply that no gain correction has been applied. The gain factor was incremented in steps of one third from 0 to 2 and the impact on R and the reduced χ^2 from the fits to each detector subgroup (A and B) are shown in Figs. 73 and 74 for both n values. To get the overall sensitivities, the R plots in Fig. 73 were combined



(a) Low n, Group A

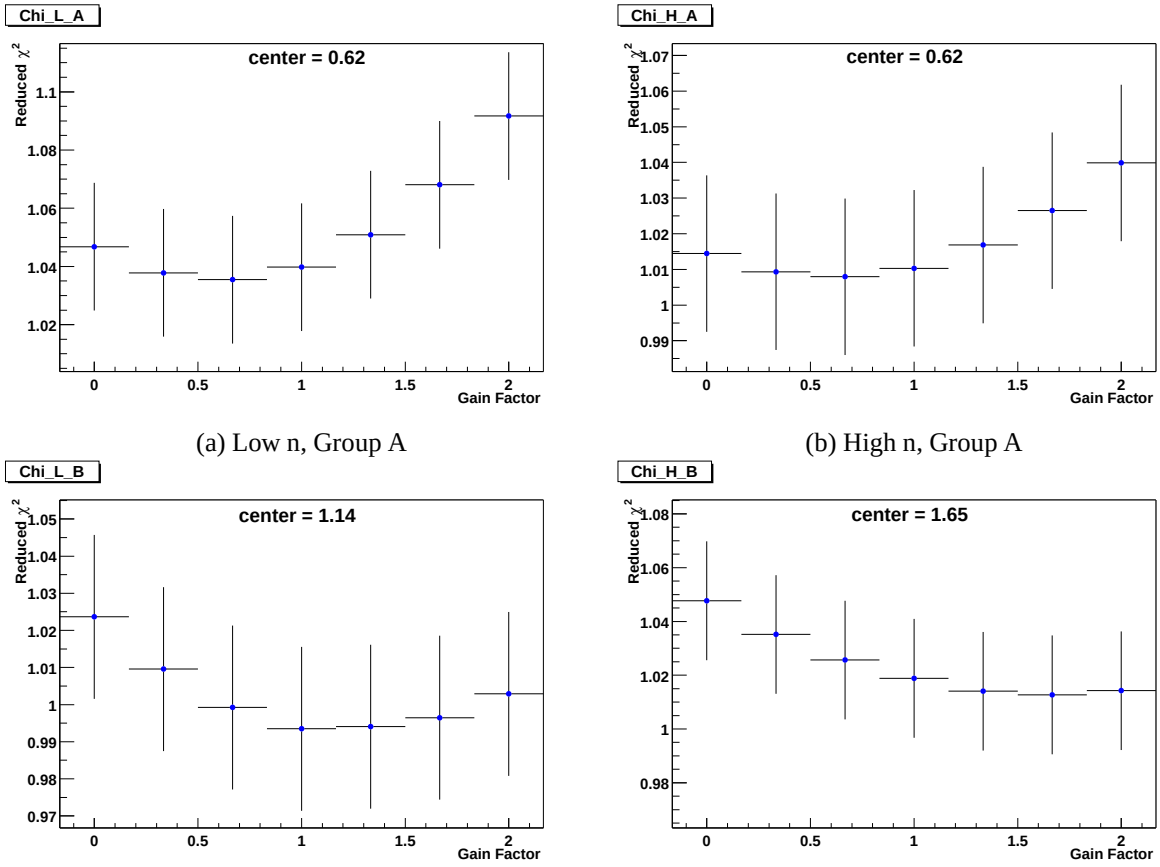
(b) High n, Group A

(c) Low n, Group B

(d) High n, Group B

Figure 73: Impact of the gain multiplier on R for both detector groups and n values.

using standard error propagation and then fit for a slope. The slopes were determined separately for each n value, and then the two R versus gain multiplier plots were combined and fit for the overall sensitivity. In addition to the asymmetry-weighted analysis, the systematics were also explored for a traditional fit, see Table 7 for details regarding energy ranges, start times, etc. It should be noted that the histograms for analyzing the gain systematic for the traditional fit were accidentally constructed using an energy range of 1.5-3.3 GeV, correcting the error will require another pass through the data. The sensitivities are given in Table 11.



(a) Low n, Group A

(b) High n, Group A

(c) Low n, Group B

(d) High n, Group B

Figure 74: Impact of the gain multiplier on χ^2 for both detector groups and n values.

	Slope (ppm/unit gain)		
Method	Low n	High n	Combined
w=A	0.178	0.083	0.139
w=1	0.055	-0.012	0.027

Table 11: Sensitivity of R to gain.

	Systematic Error (ppm)		
Method	Low n	High n	Combined
w=A	0.072	0.033	0.056
w=1	0.022	-0.005	0.011

Table 12: Systematic error due to the gain factor assuming a $\pm$40% error in the construction.

The only remaining question is how to estimate how well the gains are constructed. The agreement between various analyzers is at about the 30% level, however the reduced χ^2 plots versus gain multiplier for the quite detectors (group A) are minimized at a gain multiplier of 0.6 which would suggest the gains are overconstructed, as shown in Fig. 74. Therefore, a conservative $\pm$40% is used to calculate the error due to the gain.

11.2 CBO Systematic

Typically, in order to get all of the CBO parameters to behave, analyzers have fixed the CBO frequency and lifetime and then manually scanned them in order to quote a systematic. At the start times used in this analysis, the fit does not have any problem latching on to these parameters so the CBO frequency and lifetime were allowed to float. Unlike manually scanning them, this procedure automatically accounts for all of the correlations in the fit. Floating the parameters reduces the steepness of the χ^2 which naturally increases the error on R. Fitting with the CBO lifetime and frequency floating increases the error on R as though a 0.03 ppm term were added in quadrature.

One remaining source of error is the shape of the CBO envelope. It was shown by Mario that the empirically constructed envelope closely resembles an exponential, so that is what is used in this analysis. In order to get a feel for how sensitive the R value is to the shape of the CBO envelope, a Gaussian envelope centered at $t = 0$ was also tried.

	Asymmetry Weighted						Traditional	
	Group A		Group B		Combined Groups			
n Value	R	δR	R	δR	R	δR	R	δR
Low n	109.182	0.948	107.782	1.529	108.793	0.806	108.402	0.901
High n	107.927	1.179	109.539	1.892	108.378	1.001	108.151	1.119
Combined n					108.629	0.628	108.303	0.702

Table 13: Fit results for R and δR with an exponential envelope.

	Asymmetry Weighted						Traditional	
	Group A		Group B		Combined Groups			
n Value	R	δR	R	δR	R	δR	R	δR
Low n	109.128	0.947	107.701	1.528	108.732	0.805	108.378	0.901
High n	107.740	1.177	109.501	1.890	108.232	0.999	108.121	1.118
Combined n					108.535	0.627	108.277	0.702

Table 14: Fit results for R and δR with a Gaussian envelope.

	Asymmetry Weighted				Traditional	
	Group A		Group B			
n Value	χ^2 expo	χ^2 gaus	χ^2 expo	χ^2 gaus	χ^2 expo	χ^2 gaus
Low n	1.059	1.061	1.006	1.005	1.010	1.010
High n	1.015	1.025	1.035	1.038	1.007	1.008

Table 15: Reduced χ^2 from fits with an exponential or Gaussian envelope.

The difference between using an exponential or a Gaussian envelope is 0.094 ppm for the asymmetry-weighted data and 0.026 ppm for the traditional analysis. The systematic error due to the envelope is taken to be this difference. It is interesting to note that in asymmetry-weighted data the impact on R is more significant on the high n than the low n data. There is a 0.06 ppm diffference in the low n, while in the high n data the difference is 0.15 ppm. Perhaps this can be

considered as more evidence that the envelope in the high n data is not quite as well understood. This has already been suspected due to the remaining peaks around 490 kHz in the Fourier transform of the high n residuals after a full physics function fit. Mario also shows that the χ^2 from an exponential fit to his empirically constructed envelope is not as good for the high n as for the low n data. In general, the χ^2 from the fits indicates that the exponential from is the same or better than the fits to a Gaussian envelope, see Table 15.

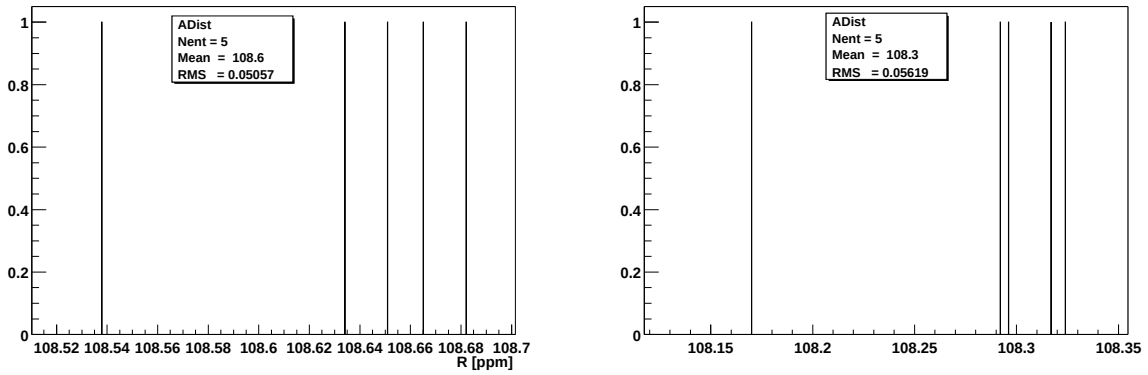
In addition to changing the envelope from exponential to Gaussian, the Rob and Jim terms were allowed to take on dramatically different lifetimes. The Rob and Jim terms were constrained to be one half, the same, and a factor of two larger than $/tau_{CBO}$, where $/tau_{CBO}$ is the fitted value for the primary CBO envelope. The lifetime was allowed to float in each fit. The results for an asymmetry-weighted analysis, fitting the sum of all detectors at 31.7 μs in the 1.5-3.3 GeV energy range, are

11.3 Random Seed

The R values for various random seeds are given in Table 16, and the distributions of those values are in Fig. 75.

	Asymmetry-weighted		traditional	
Seed	R	δR	R	δR
1	108.682	0.628	108.296	0.702
2	108.651	0.627	108.317	0.703
3	108.665	0.628	108.292	0.702
4	108.634	0.623	108.324	0.702
5	108.538	0.627	108.170	0.705

Table 16: Random seed results for R and δR.



(a) Asymmetry-Weighted (b) Traditiona

Figure 75: Distribution of R values with 5 different seeds.

The analysis of the effect of fill randomization has been studied by many analyzers in the past where it was conluded that the error from using a single random seed is about 10% of the error

bar. The errors in Table 16 and the RMS of the distributions in Fig. 75 are consistent with this estimation for both the asymmetry-weighted and the traditional analyses. Therefore, since five random seeds were used, the estimated error is,

$$\frac{0.10 \cdot \delta R}{\sqrt{5}}, \tag{38}$$

which works out to about 0.028 ppm for the asymmetry-weighted analysis and 0.032 ppm for a traditional fit. Not that it quantitatively matters, but pedagogically the question of rather to use a square root of 4 or 5 in the denominator came up. I use 5 because I am assuming that the width of the distribution is already known to be 10% of the error on R based on analyses in the past that have looked at more random seeds.

11.4 Pileup Systematic

The systematic due to pileup is determined for both weighting schemes by looking at sensitivities to the pileup phase and the amplitude.

11.4.1 Pileup Amplitude

For this study the pileup in all energy and time bins was scaled by a pileup multiplier that ranged from zero to two and then subtracted from the data. The error on the each energy and time bin was calculated as,

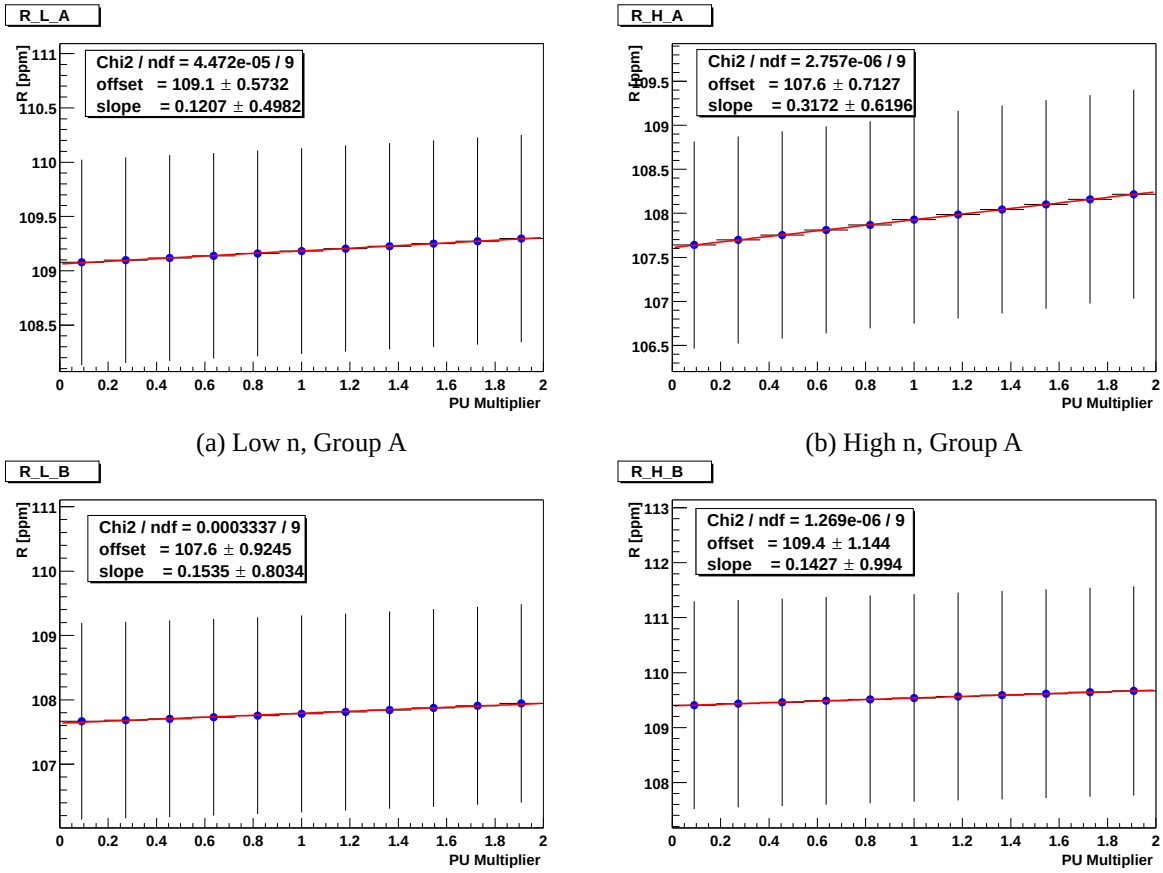
$$\sigma_i = \sqrt{N_i + m^2 S_i + m^2 D_i}, \tag{39}$$

where m is the pileup multiplier. To be clear, a pileup mutliplier of zero indicates no pileup subtraction, while a multiplier of two would be subtracting twice the constructed pileup. The plots in Fig. 76 show the sensitivity of the R value to the pileup multiplier in the asymmetry-weighted analysis. As in the gain multiplier study, these plots were combined for each n-value using standard error propagation and then the slope was extracted. For the overall sensitivity, all 4 plots were combined and then fit for the slope. The summary of the sensitivities is given in Table 19. Results are also shown for the traditional method. Surprisingly, the asymmetry-weighted analysis is less sensitive to the gain multiplier even thought it weights the higher energies preferentially with most detectors starting earlier. Since pileup has a non-wiggling component, one possible explanation could be the precision at which the starting bin actually resides at a zero crossing.

	Slope (ppm/unit pileup)		
Method	Low n	High n	Combined
w=A	0.129	0.269	0.184
w=1	0.199	0.261	0.223

Table 17: Sensitivity of R to the pileup multiplier.

To confirm this result, the uniformly weighted fit was applied to the same detector subgroups with the same start times. The comparison is shown in Fig 77, where several interesting conclusions can be drawn. First of all, the reduction in the error bar with the asymmetry-weighting is clearly evident. Secondly, the slopes are practically identical. Finally, the difference in R between



(a) Low n, Group A

(b) High n, Group A

(c) Low n, Group B

(d) High n, Group B

Figure 76: Impact of the pileup multiplier on *R* for both detector groups and n values.

	Systematic Error (ppm)		
Method	Low n	High n	Combined
w=A	0.023	0.048	0.033
w=1	0.036	0.047	0.040

Table 18: Systematic error due to the pileup multiplier assuming a ±18% error in the construction.

the two analysis methods is 0.18 ppm. Simulations show that the allowed difference is about half of the error on the asymmetry-weighted data, which comes to an allowed difference of 0.31 ppm. Therefore, the two analyses agree at the 0.6σ level. This also implies that half of the 0.35 ppm difference shown in Table 8 comes from the earlier start times. Note that the other analyzer to use an earlier start time, Jon Paley, also gets a R value that is slightly higher than Xaiobo. The comparison with Mario is a little more difficult since he uses an extended runlist.

Now that the sensitivites are known, as in the case of gain, an estimation of the pileup efficiency is required. The plots in Fig. 78 show the reduced χ^2 for the 4 asymmetry-weighted fits. Taking a simple average of the 4 parabolic centers yields an optimal pileup multiplier of 1.18, therefore it is assumed that pileup subtraction is performed with an efficiency of ±18%. The systematic errors given in Table 18 are taken as a product of the sensitivy and the efficiency.

One possible flaw with the way this systematic error is determined can arise from assuming

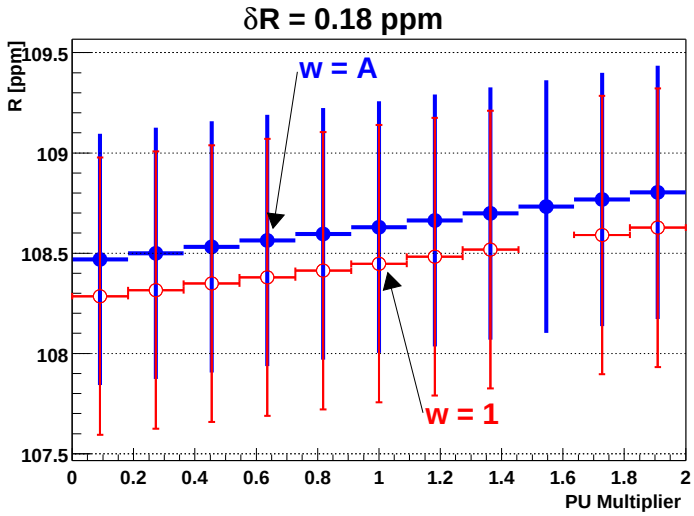
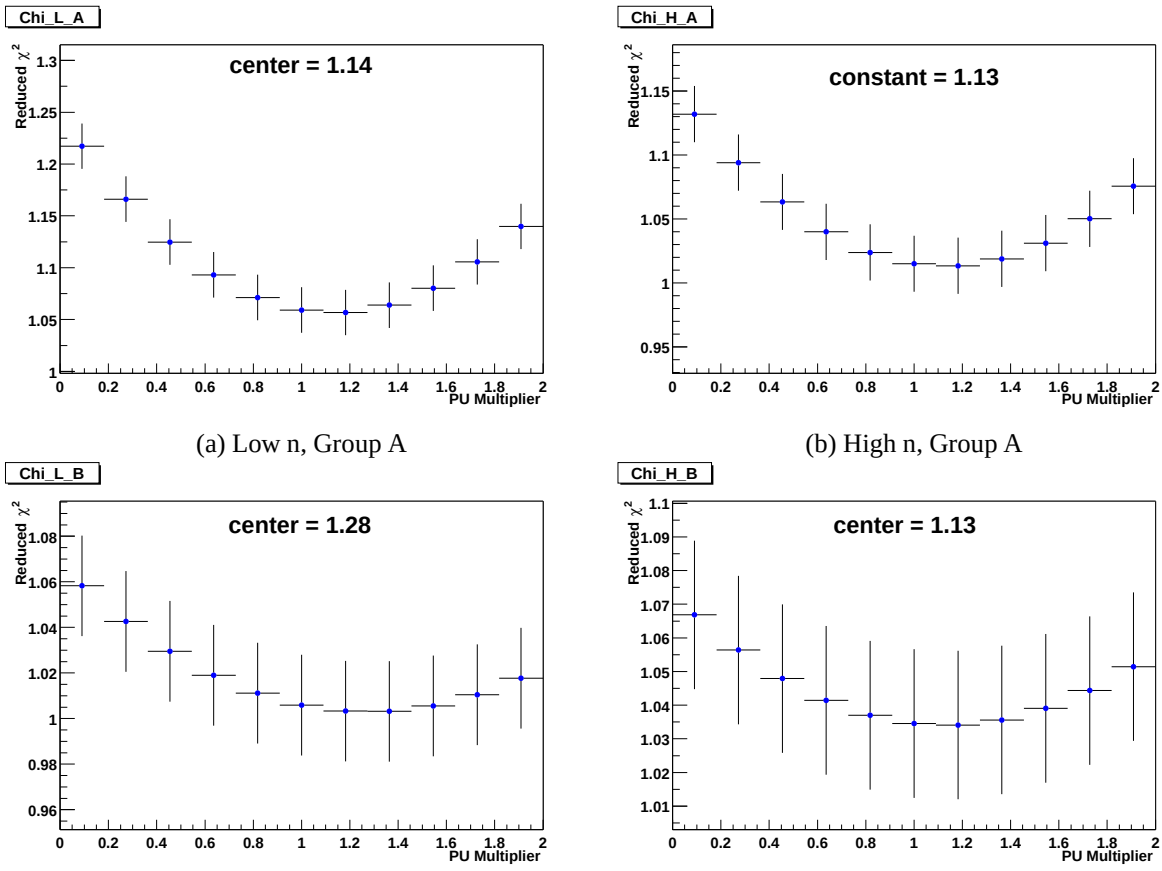


Figure 77: Comparison of the sensitivity to the pileup multiplier with w=A and w=1 at the same start times.



(a) Low n, Group A

(b) High n, Group A

(c) Low n, Group B

(d) High n, Group B

Figure 78: Impact of the pileup multiplier on χ^2 for both detector groups and n values.

that the overall sensitivity can just be multiplied by the efficiency. The inherent assumption is that the efficiency is the same in all energy bands. If the efficiency for pileup construction is worse

at the low energies, then the traditional method underestimates the systematic error. Conversely, if the efficiency is problematic in the high energies, then the asymmetry-weighted error is underestimated. Fig. 79 shows the optimal pileup multiplier, determined from where the reduced χ^2 minimizes, as a function of energy. There is no conclusive evidence that the 18% estimation is particularly bad at either end of the energy range.

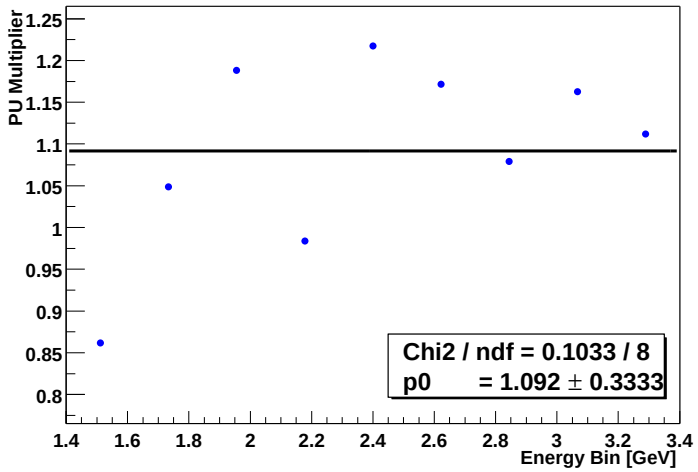


Figure 79: Optimum pileup multiplier (determined from where the χ^2 minimizes), analyzed in separate energy bands.

11.4.2 Pileup Phase

As has been pointed out before, pileup can really be considered as a vector with an amplitude and a phase. The pileup multiplier study determines the error due to the amplitude, but a separate study is required for the phase.

	Slope (ppm/ns)		
Method	Low n	High n	Combined
w=A	-0.021	-0.025	-0.022
w=1	-0.027	-0.029	-0.028

Table 19: Sensitivity of R to the pileup phase.

	Systematic Error (ppm)		
Method	Low n	High n	Combined
w=A	-0.042	-0.050	-0.044
w=1	-0.054	-0.058	-0.056

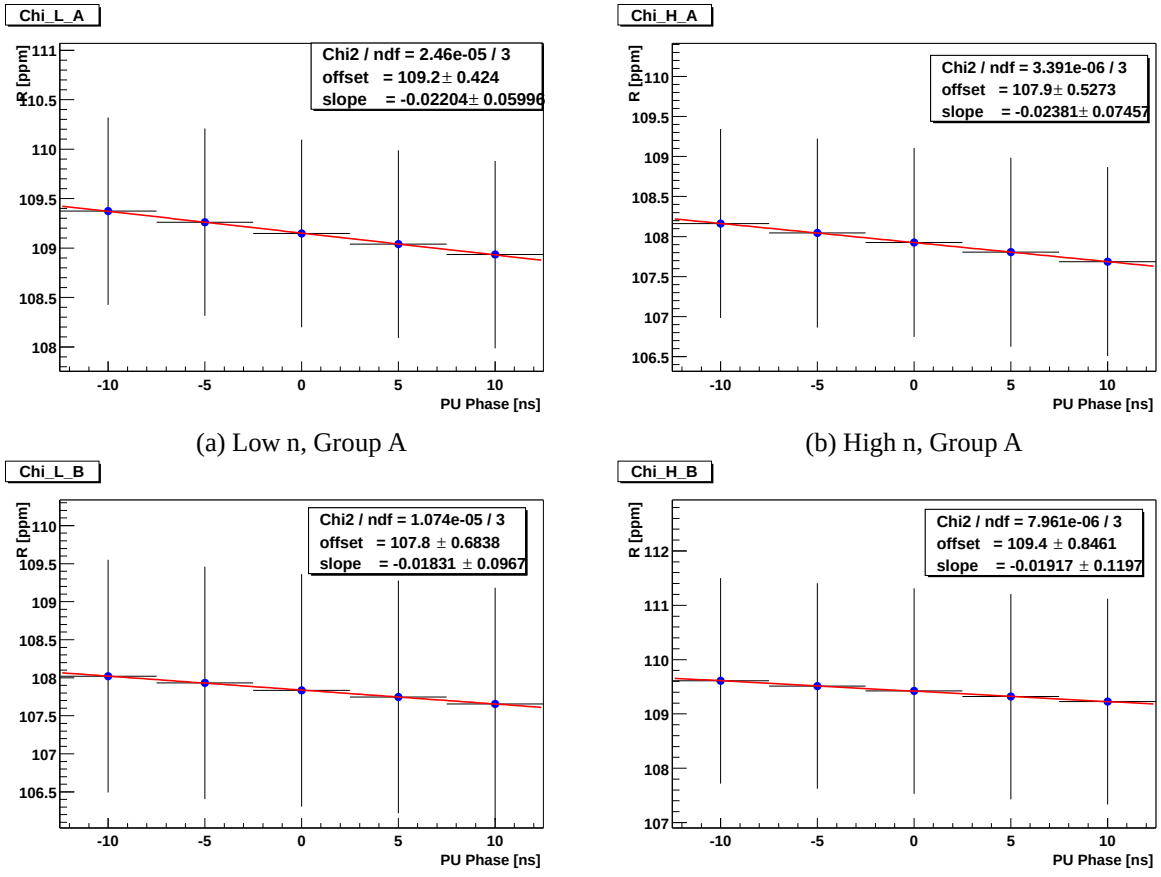
Table 20: Systematic error due to the pileup phase assuming a 2 ns error in the construction.

Normally when pileup is constructed using symmetric pileup window the times of the S events are just the times determined from the pulse fitter, while the time of the D events is taken to be

	Systematic Error (ppm)		
Method	Low n	High n	Combined
w=A	0.048	0.069	0.055
w=1	0.064	0.075	0.069

Table 21: Systematic errors from pileup phase and amplitude combined in quadrature.

the energy-weighted time. In order to explore the sensitivity of the fit to the pileup phase, the times of S and D were shift in increments of 5 ns ranging from -10 to +10 ns. For those who think in angles, this corresponds to shifting the pileup phase by ± 14.4 mrad. The impact on R for the 4 asymmetry-weighted fits is shown in Fig. 80. The same procedure as used for the gain and pileup multipliers was used to combine the results first by n value and then in total. The slopes are summarized in Table 19. The asymmetry-weighted analysis is slightly less sensitive to pileup phase than the traditional analysis.



(a) Low n, Group A

(b) High n, Group A

(c) Low n, Group B

(d) High n, Group B

Figure 80: Impact of the pileup multiplier on *R* for both detector groups and n values.

Once again the sensitivities are now known, but an estimation of how poorly the phase is constructed is still needed. Unfortunately, the χ^2 is fairly insensitive to the phase, as shown in Fig. 81, so it can not be used as an estimate of the phase error. Jon Paley had an interesting way of estimating the efficiency where he fit the phase of the pileup after dropping the highest or

lowest 200 MeV energy bin. He found the phase was most affected by dropping the lowest energy bin, which resulted in a 2.8 mrad (1.9 ns) phase shift. Until a better method of quantifying the efficiency is developed, I will use Jon's estimate of about 2 ns. The resulting systematic errors are summarized in Table 20, and the quadratically combined errors due to amplitude and phase are in Table 21.

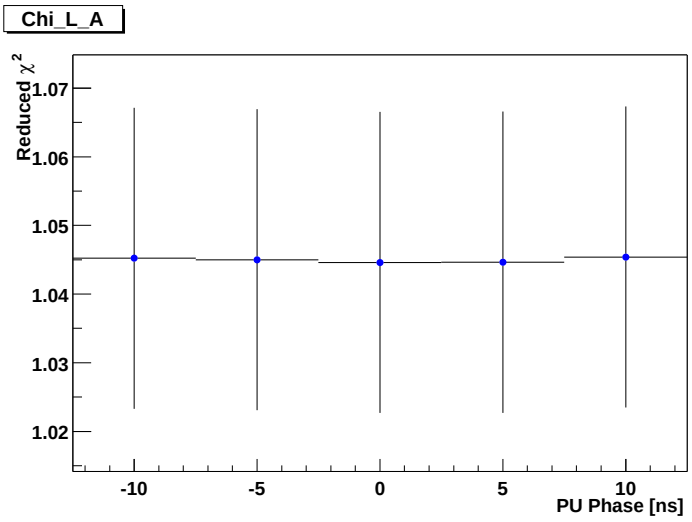


Figure 81: Reduced χ^2 as a function of phase.

11.4.3 Threshold Correction

In the asymmetry-weighted analysis, the threshold is extended down to 1.5 GeV. Most of the detector have a hardware threshold around 750 MeV and do not really require a correction in order to construct the pileup down to 1.5 GeV. However, a few do require the correction. As is evident in Table 22, the difference between analyzing with and without the correction is at the 0.01 ppm level.

	Applied		Not Applied	
n Value	R	δR	R	δR
Low n	108.452	0.819	108.472	0.819
High n	108.232	1.017	108.232	1.017

Table 22: Change in R based when the correction for pileup below twice the hardware threshold is not applied.

11.5 Fit Procedure Systematics

To check for obvious bug in the fitter, several thousand data sets were randomly generated and fit with a 5 parameter function. It was found that the output and input R values were consistent at a 0.03 ppm level. In addition to this, the effect of fitting with $\sqrt{N}$ or $\sqrt{f}$ as the error was found to be < 0.01 ppm. Fitting the data out to the point where there were 80 counts in a bin or out to the

point where there was only one count left did not produce a measurable difference in R. The error due to the fit procedure is taken to be 0.03 ppm. For details of the simulation see the section on simulation results.

11.6 Muon Loss Systematic

The systematic error due to muon losses can stem from two known sources. The first has to do with the effect on the fitting function. Like gain, muon losses are a slow term that can result in phase pulling. By staying near $g-2$ zero crossings the pulling of R is minimized. However, unlike gain, the muon loss normalization floats in the fit and is therefore already included in the statistical error. Therefore, like the CBO envelope, the systematic error arises only from a mistake in the shape of the constructed losses. Shape-space is infinite and hard to quantify, but a couple of tests were performed. The simplest is to fit with and without the muon losses. The difference in R is 0.18 ppm, but this is a fairly extreme change in shape. Mario performs a more subtle test where he fits separately with Chris and Jon's constructed losses. He finds the difference to be < 0.01 ppm. This is probably conservative since Jon and I are using the same data to construct losses and the largest source of error stems from not having full coverage of the FSDs around the ring. Looking at the shape of the losses in each set of triples, it is not unreasonable to find a 10% early-to-late difference, as shown in Fig. 82, where the ratio of the triples nearest the inflector $(8\cdot 9\cdot 10)$ to the triples 'furthest' from the inflector $(23\cdot 24\cdot 1)$ is plotted versus the time in spill. Note that there is also a factor of one third in the ratio to set the scale close to one, i.e. the detectors near the inflector see three times the losses. I take the error due to the fitting function to be 20% of the 0.18 ppm difference found when leaving the muon losses out altogether. This 0.036 ppm difference is small relative to estimates for the second source of error.

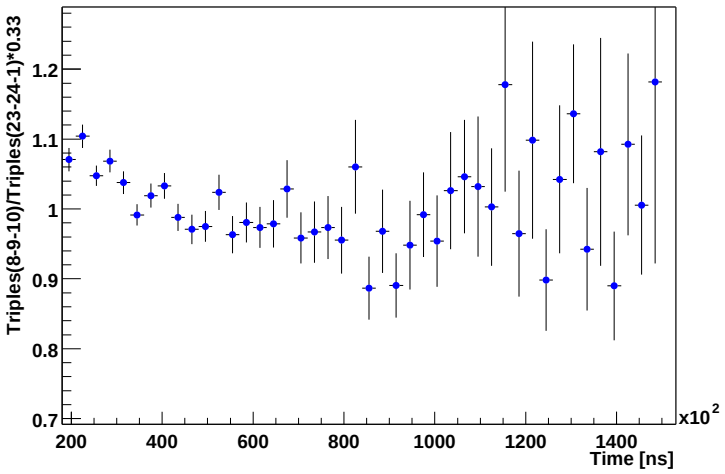


Figure 82: Ratio of triples constructed from $(8\cdot 9\cdot 10)$ to $(23\cdot 24\cdot 1)$, normalized to one.

The second source of systematic error arises if the lost muons carry a phase that is different from the stored muons. Simulation results by Jon have set a limit on this error of 0.08 ppm. Combining this error in quadrature with the 0.036 above yields a total systematic error of 0.09

ppm.

11.7 Vertical Waist Systematic

The vertical waist frequency and lifetime were manually varied over a range that is twice the size of the actual fitted error. It was found that R is stable at the 0.03 ppm level.

11.8 Beam Relaxation Systematic

The term accounting for the beam relaxation after scraping was turned on and off in the fit. The difference in R was found to be 0.05 ppm.

11.9 Flashlet Systematic

Flashlets occur predominantly at low energies in the area around the injection point. Relative to a traditional analysis, my asymmetry-weighted analysis weights in the low energy data less preferentially and the detectors in that region are not fit until later times. On the other hand, I use the 1.5-1.8 GeV range which does have more flashlets. All in all, I assume the effects will cancel and adopt Xiaobo's estimation of 0.02 ppm.

11.10 Bin Width Systematic

The effect of the bin width on the fit has been explored by many analyzers in the past. It should not be different in 2001, so I adopt Fred's rather conservative 0.06 ppm.

11.11 Other CBO Systematic

The effect of higher order CBO terms like double CBO and VHF have been explored by many analyzers with an error < 0.03 ppm.

11.12 Unseen Pileup Systematic

This is still a subject of debate. Without a real good method of getting an experimental handle on it, the current consensus is to use 0.1 ppm which is in between Vanya and Bill's estimate.

12 Simulation Results

The plots below are the start of my summary of simulation results. The first plot, Fig. 83, shows the allowed difference in R when moving from a 1.8-3.3 GeV traditional analysis to a 1.5-3.3 GeV asymmetry-weighted analysis. The allowed difference is 0.33 ppm. For normalization, the errors on R for the simulated data sets are shown in Fig. 84. The error on the asymmetry-weighted data is 0.72 ppm, which implies an allowed difference of about half of the errror bar.

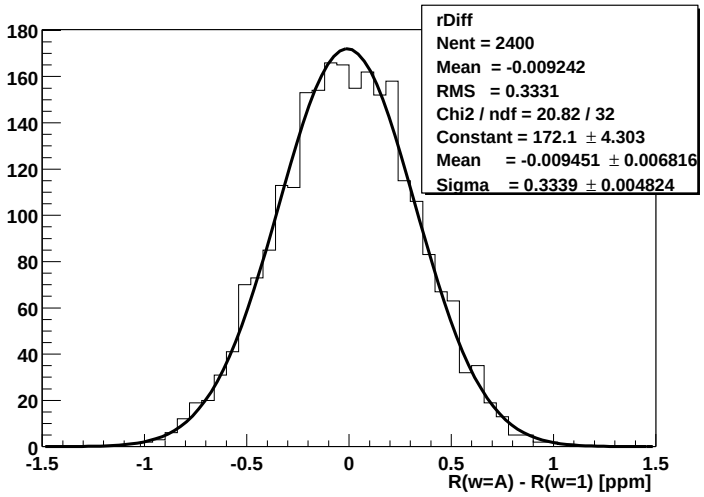


Figure 83: Difference in R when changing from w=1 to w=A and adding the 1.5-2.8 GeV range.

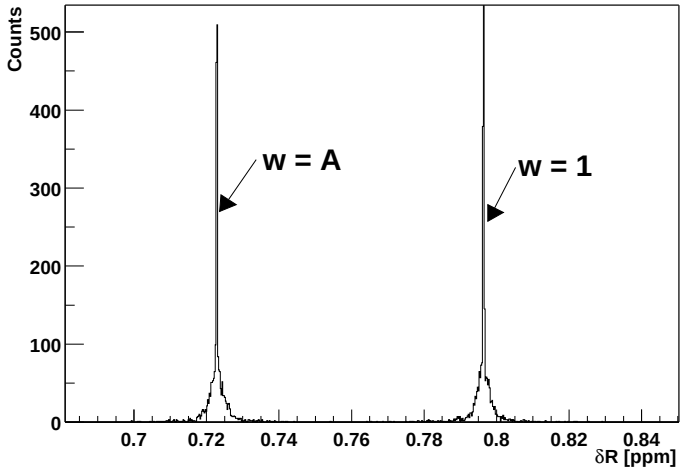


Figure 84: Error on R for the simulated data sets.

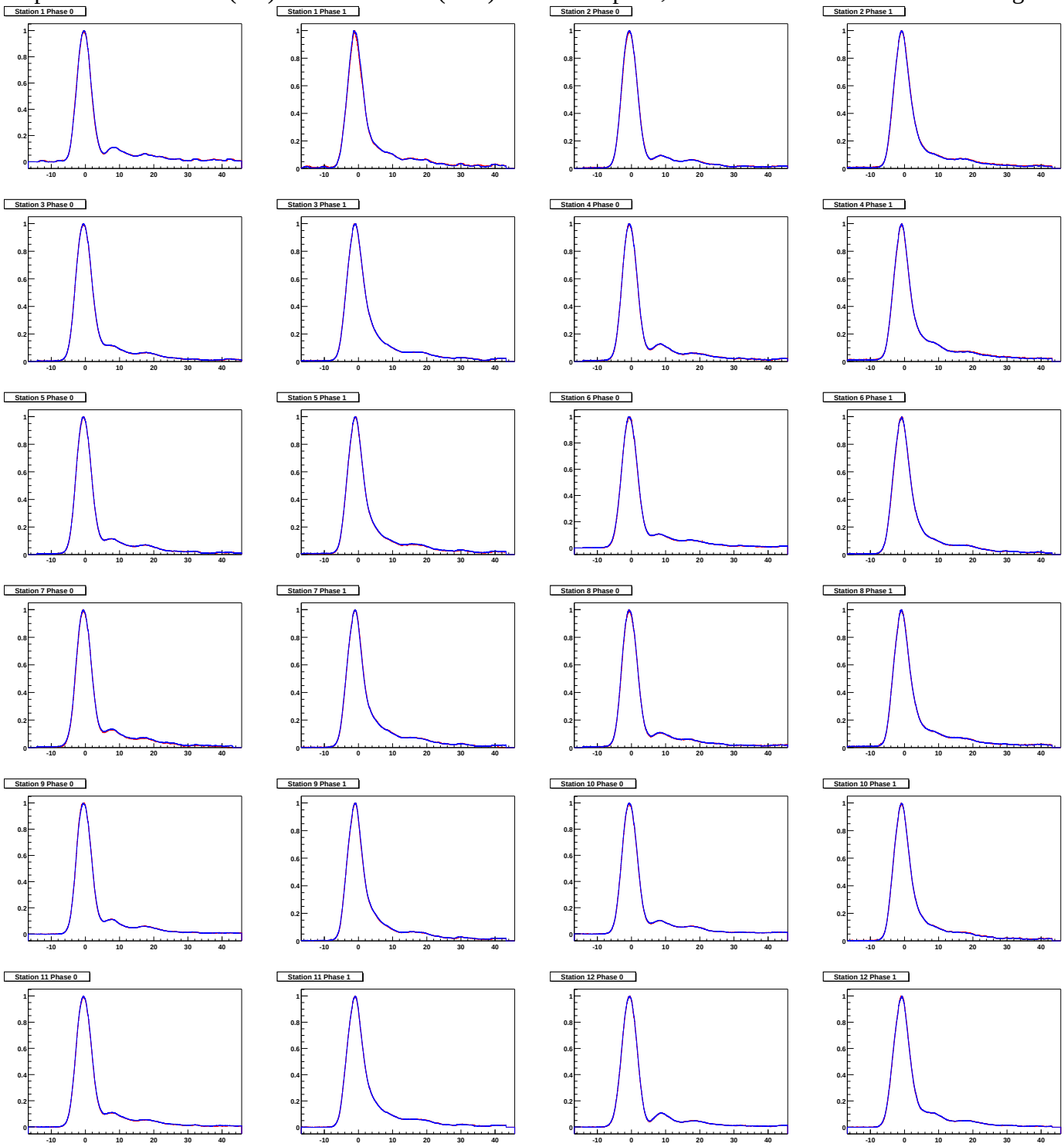
13 To-Do List

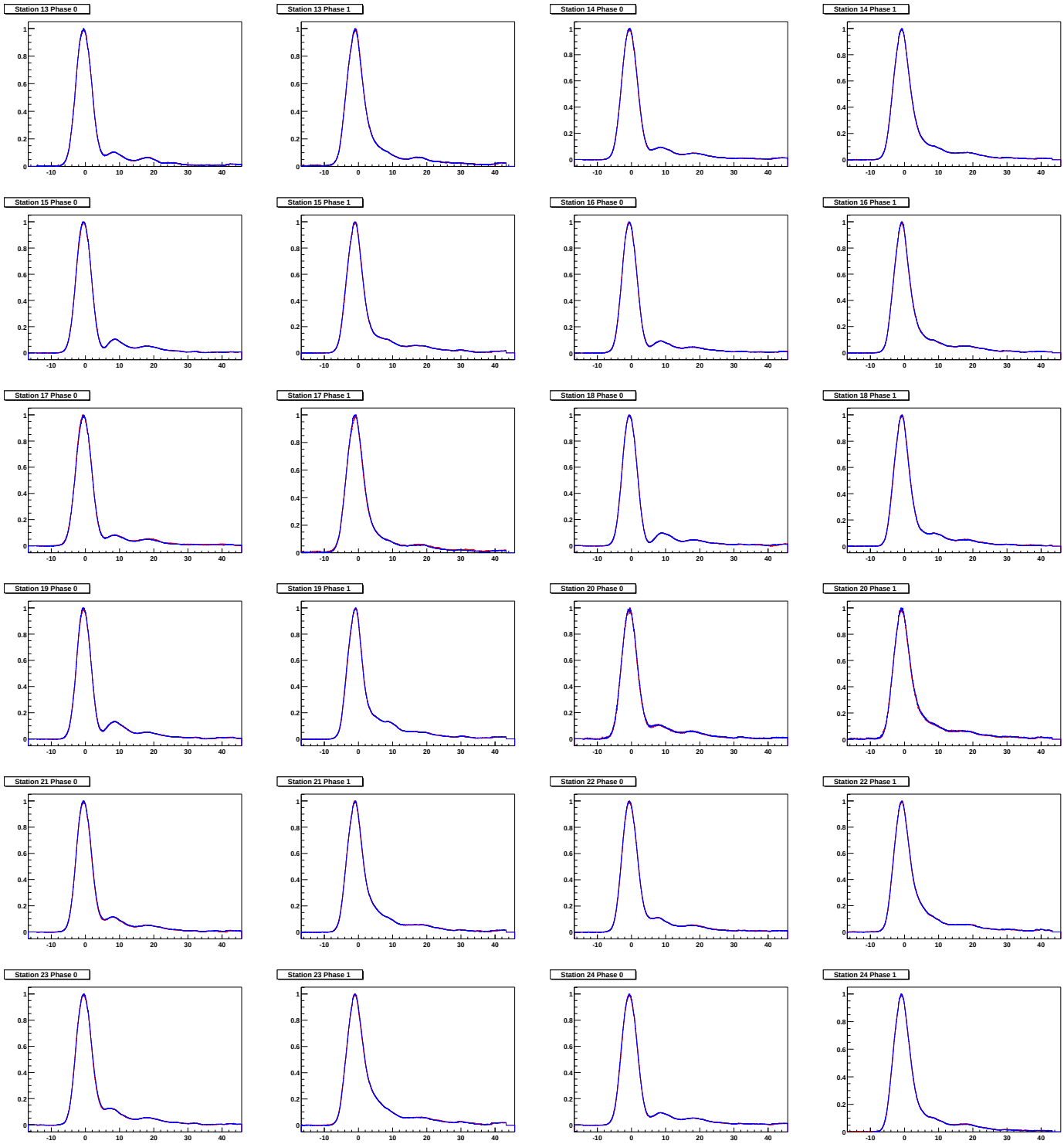
If I were to spend more time on the analysis, this is what I would do...

- Complete and independent flashlet study.
- Perform a few more consistency checks on the muon loss construction.
- Have a closer look at the correlation between the Rob and Jim terms and pileup subtraction.
- Try to come up with some new studies with regard to the consistency in the high n results versus detector and energy.

A Pulseshapes

The pulseshapes for each detector and phase are plotted below. Each plot is actually an overlay of the pulseshapes from run 9304 (red) and run 11229 (blue). For most plots, the red and blue lines are indistinguishable.





B The Kolmogorov-Smirnov Histogram Compatibility Test

Loosely speaking, the Kolmogorov-Smirnov compatibility test is a statistical procedure used to determine whether or not two samples have the same underlying parent distribution. A standard

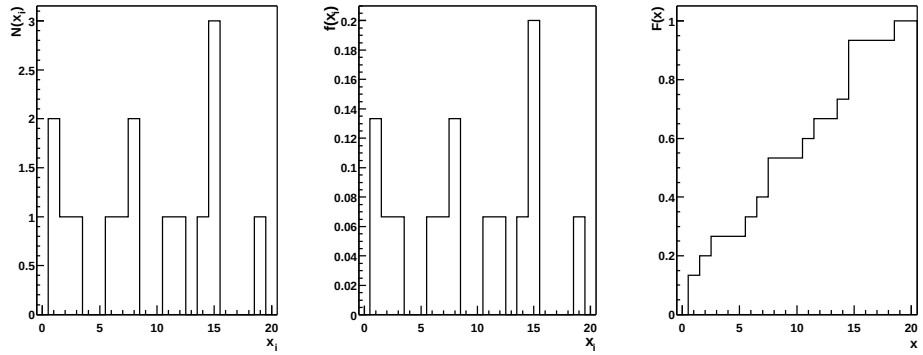


Figure 85: Plot of the number density, probability distribution function, and cumulative distribution function for the sample data set X.

implementation is available in PAW where the subroutine is called HDIFF, and in ROOT under the method TH1::KolmogorovTest.

To understand the basis of the KS test, consider two ordered sets of data, $X = x_1, x_2, \ldots x_n$ and $Y = y_1, y_2, \ldots y_m$, where n and m can take on any integer value. The KS test is a test of the null hypothesis,

$$H_o : P(x \leq c) = P(y \leq c), \; \forall \, c, \tag{40}$$

where P is a probability function. This is just a compact way of indicating that the populations of the X and Y values have the same cumulative distribution which then implies that they have the same parent distribution. The first step in applying the test is to calculate a cumulative distribution function (cdf) for each sample. In any experiment the number of measurements is finite and each measurement is discrete, i.e. no measurement device has infinite resolution. Therefore, the number density (frequency distribution), $N(x_i)$, is also discrete-valued since it simply counts the frequency with which x_i occurs in the data set. The cdf is calculated by first finding the probability distribution function (pdf) for the sample data,

$$f(x_i) = N(x_i)/n, \tag{41}$$

where n is the number of data points in the sample. This is just a scaled version of the number density such that the sum of $f(x_i)$ over all values of x_i is normalized to 1. The cdf is found by summing the pdf up to a particular value of the range,

$$F(z) = \sum_{x_i \leq z} f(x_i). \tag{42}$$

Since $f(x_i)$ is discrete-valued and normalized to unity, $F(z)$ is a step function whose height starts at 0 and ends at 1.

As a trivial example consider the data set,

$$X = \{1, 1, 2, 3, 6, 7, 8, 8, 11, 12, 14, 15, 15, 15, 19\},$$

composed of 15 events. The number density, pdf, and cdf for the sample data are shown in Figure 85. Note that the pdf and cdf discussed here are calculated directly from the data and should not be confused with the distribution functions associated with the underlying population.

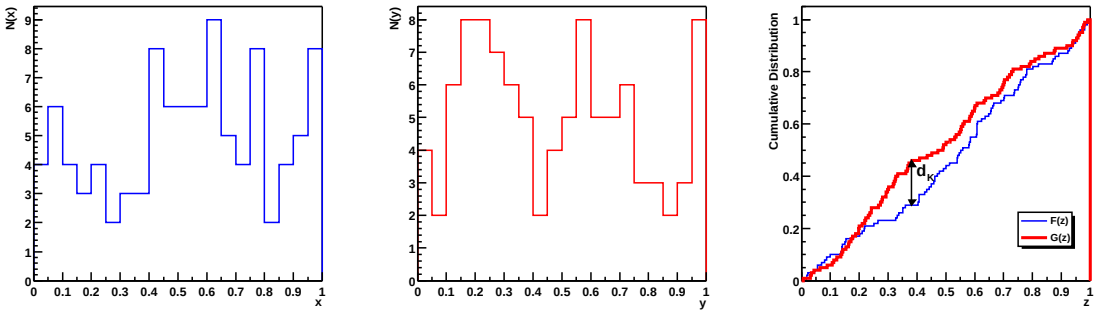


Figure 86: Plots showing the number density functions for the X and Y samples followed by a plot of their cumulative distribution functions. The Kolmogorov distance indicated occurs where the vertical difference in the cdfs is maximal.

Once the sample cumulative distribution functions for X and Y have been obtained, the KS two-sided test takes as its statistic the maximum vertical distance between the two graphs. The Kolmogorov distance, d_K, is

$$d_K = \max |F(z) - G(z)|, \tag{43}$$

where F(z) is the cdf for X and G(z) is the cdf for Y. It was shown by Kolmogorov [4] that when comparing two samples drawn from the same parent distribution, d_K follows a distribution that is independent of the parent distribution. Furthermore, in the limit of infinite data sets, the distribution of d_K values can be expressed in closed form as,

$$P(d > d_K) = 2\sum_{k=1}^{\infty}(-1)^{k-1}\exp\left[-2k^2 d_K^2\left(\frac{mn}{m+n}\right)\right]. \tag{44}$$

This is referred to as the p-value of the test and it is simply the probability to obtain a Kolmogorov distance greater than the one derived from the two samples assuming H_o is true. Once $\min(m,n)$ exceeds 80, the probability returned by Equation 44 is sufficiently accurate for most applications. Tables are available for tabulating $P(d > d_K)$ for smaller data sets. As an aside, it is common practice in the literature to see a change of variable similar to the following. Let,

$$\lambda^2 = d_k^2\left(\frac{mn}{m+n}\right), \tag{45}$$

so that Equation 44 may be written,

$$P\left(d > \lambda\sqrt{\frac{m+n}{mn}}\right) = 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2 d_K^2\lambda^2}. \tag{46}$$

For a more detailed derivation or further properties of the KS test, the reader is referred to [5, 6, 7, 8].

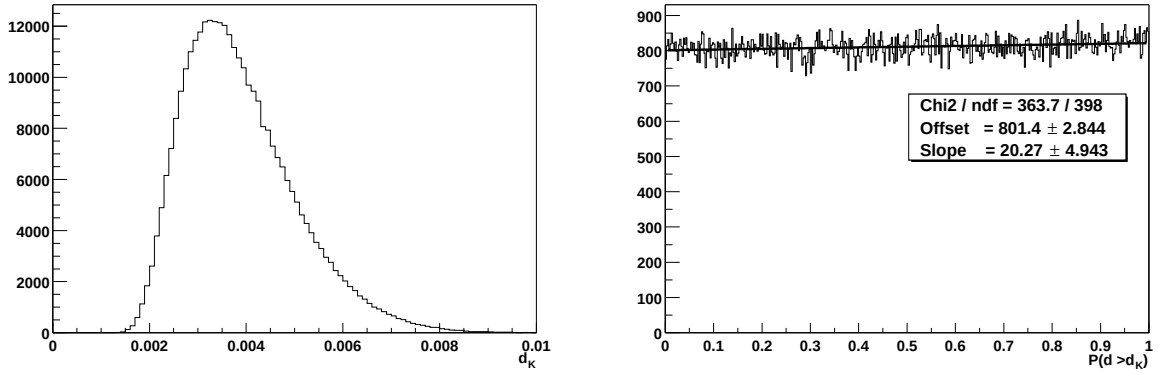
As an example, consider the following samples X and Y which are both composed of 100 data points drawn from a uniform distribution on the interval (0,1). Figure 86 shows the distributions for the two samples, $N(x)$ and $N(y)$, with an increased bin width for demonstration purposes. The third plot shows the overlaid cdfs constructed from the finely binned number densities and the Kolmogorov distance associated with this particular comparison. In this case, $d_K = 0.170$ and the associated p-value calculated from equation (44) is 0.11. The interpretation is stated below.

If these two samples have the same underlying parent distribution, then given the statistics of the samples that we obtained, one would expect to get a d_K greater than 0.17 approximately 11% of the time.

The decision of when to reject H_o depends on the chosen value for the significance level of the test. In general, the significance level of the test, α, is defined to be the p-value for which the null hypothesis is rejected. This implies the existence of a critical value, d_α, given by,

$$P(d > d_\alpha) = 2\sum_{k=1}^{\infty}(-1)^{k-1}\exp\left[-2k^2 d_K^2\left(\frac{mn}{m+n}\right)\right] = \alpha, \tag{47}$$

for which the null hypothesis is rejected if $d_K > d_\alpha$. The choice of α is unfortunately subjective in nature and depends on the relative risk versus benefit of making a Type I or a Type II error. A Type I error occurs when you reject a H_o that was in fact true, while a Type II error arises if a false hypothesis is inadvertently accepted. It should be evident that the significance level in equation (47) is the probability of making a Type 1 error. The author is not aware of any method other than simulation for obtaining, β, the probability of making a Type II error with the KS test.



(a) Distribution of d_K values.

(b) Probability distribution, $P(d > d_K)$.

Figure 87: Plot of the distribution of the KS test statistic, d_K, obtained from comparing samples drawn from a uniform distribution; and the corresponding distribution of p-values, $P(d > d_K)$.

If the null hypothesis was rejected in the above example, then there is an 11% chance that a mistake was made and the samples really did come from the same distribution. To illustrate this concept, the KS test statistic, d_K, was computed for 300,000 independent pairs of samples drawn from the same uniform parent distribution. The samples were each composed of approximately 100,000 events and the resulting Kolmogorov distances are shown in Figure 87(a). The probability of obtaining a larger Kolmogorov distance was calculated using equation (44) and is shown in Figure 87(b). As expected, the distribution of probabilities is nearly uniform. This implies that if the KS test is used as a means of quality control in selecting runs for a dataset, then one should expect to lose, in addition to the bad runs, a fraction of good runs equivalent to the chosen value of α.

A fit to the distribution in Figure 87(b) shows that $P(d > d_K)$ has a slightly positive slope. This can be partially attributed to the fact that the data was binned. The KS test was designed to be performed on ungrouped data. Applying the test to binned data means that the cumulative distribution

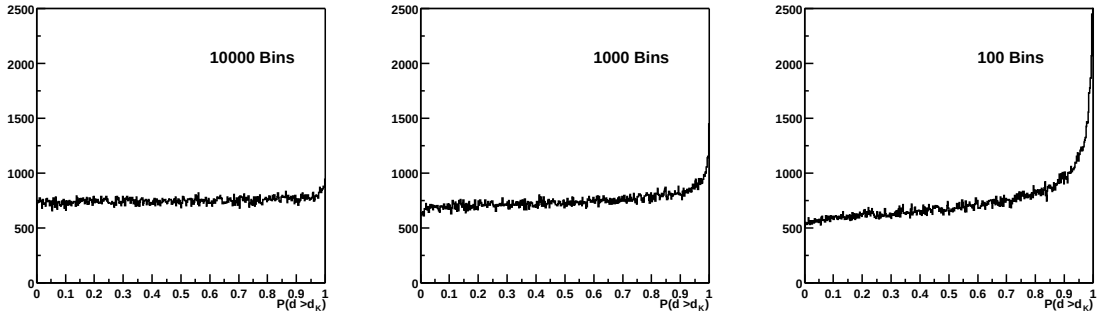


Figure 88: Probability distributions, $P(d > d_K)$, showing the effect of applying the KS test to binned data. All samples contained 100,000 events drawn from a uniform distribution.

function, equation (42), can only be calculated at bin boundaries. Therefore, the maximum difference obtained between the cdfs is potentially underestimated and the corresponding probability is consequently overestimated. Figure 88 shows results for the same data as in Figure 87 where only the bin size of the number density, $N(x_i)$, was changed. It is apparent that coarser binning results in the probability distribution being skewed towards higher values.

One final word of caution about applying the KS test concerns the range of the random variable under study. The KS test is not equally sensitive for all values of the random variable. To study this effect, datasets composed of 100,000 events were again drawn from a uniform distribution, only this time a spike of 1000 events was added at a particular value of x_i. These datasets were compared to datasets without the perturbation.

The spike was systematically moved from 10% up to 90% of the maximum value of x in steps of 10%. Figure 89 shows the average d_K value obtained and the corresponding p-value as a function of where the spike was placed. It is clear that in this particular case,the KS test is less sensitive to an anomaly placed near the middle of the x range. It should be noted that if α were chosen to be as small as 0.03, then the null hypothesis would have been rejected in nearly all instances, i.e. the test would be 100% efficient in rejecting a perturbation of this magnitude.

Multi-dimensional extensions of the test exist [9, 10, 11] and have been used in the 2000 data analysis to aid in run selection.

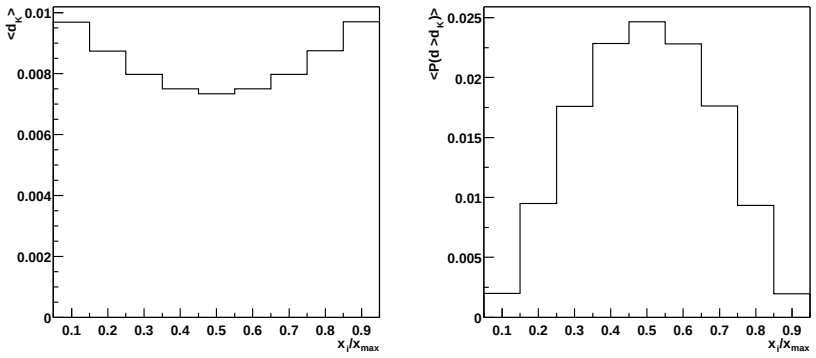


Figure 89: KS test results from comparing 100,000 event samples drawn from a uniform distribution with an additional spike of 1000 events added at 0.10 fractional increments of the range.

C Correlation Coefficients

The correlation coefficients from the traditional and asymmetry-weighted fits for both n values are given in Tables 23 and 24. The fit details can be found in section 10.2, Table 7. For the asymmetry-weighted fits, the correlation coefficients are only given for fits to the detector group with an early enough start time (group A) for vertical and beam relaxation terms to be included.

Low n Value														
Par	τ	N_0	A	ϕ	ω_a	ω_{cbo}	τ_{cbo}	a_1	ϕ_1	A_2	ϕ_2	A_3	ϕ_3	A_{loss}
No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.000	0.877	-0.026	-0.008	-0.006	-0.022	-0.025	0.030	-0.026	0.013	-0.022	-0.036	-0.040	0.892
2	0.877	1.000	-0.038	-0.011	-0.008	-0.032	-0.037	0.045	-0.038	0.020	-0.031	-0.052	-0.058	0.995
3	-0.026	-0.038	1.000	-0.005	-0.004	-0.003	0.002	-0.006	-0.006	-0.074	-0.031	-0.054	0.043	-0.037
4	-0.008	-0.011	-0.005	1.000	0.833	-0.026	-0.042	0.050	-0.030	-0.045	0.092	0.022	0.062	-0.011
5	-0.006	-0.008	-0.004	0.833	1.000	-0.019	-0.031	0.037	-0.022	-0.034	0.066	0.014	0.045	-0.008
6	-0.022	-0.032	-0.003	-0.026	-0.019	1.000	0.305	-0.284	0.917	-0.076	-0.045	0.054	0.097	-0.031
7	-0.025	-0.037	0.002	-0.042	-0.031	0.305	1.000	-0.920	0.283	-0.031	-0.079	0.202	0.024	-0.036
8	0.030	0.045	-0.006	0.050	0.037	-0.284	-0.920	1.000	-0.264	0.044	0.073	-0.196	-0.030	0.043
9	-0.026	-0.038	-0.006	-0.030	-0.022	0.917	0.283	-0.264	1.000	-0.072	-0.056	0.045	0.108	-0.037
10	0.013	0.020	-0.074	-0.045	-0.034	-0.076	-0.031	0.044	-0.072	1.000	-0.003	-0.019	-0.028	0.019
11	-0.022	-0.031	-0.031	0.092	0.066	-0.045	-0.079	0.073	-0.056	-0.003	1.000	-0.009	-0.005	-0.030
12	-0.036	-0.052	-0.054	0.022	0.014	0.054	0.202	-0.196	0.045	-0.019	-0.009	1.000	0.007	-0.050
13	-0.040	-0.058	0.043	0.062	0.045	0.097	0.024	-0.030	0.108	-0.028	-0.005	0.007	1.000	-0.056
14	0.892	0.995	-0.037	-0.011	-0.008	-0.031	-0.036	0.043	-0.037	0.019	-0.030	-0.050	-0.056	1.000

High n Value														
Par	τ	N_0	A	ϕ	ω_a	ω_{cbo}	τ_{cbo}	a_1	ϕ_1	A_2	ϕ_2	A_3	ϕ_3	A_{loss}
No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1.000	0.871	-0.023	0.009	0.010	0.021	0.010	0.010	-0.026	-0.007	-0.014	-0.029	-0.006	0.887
2	0.871	1.000	-0.035	0.010	0.010	0.029	0.014	0.014	-0.037	-0.010	-0.022	-0.044	-0.009	0.995
3	-0.023	-0.035	1.000	-0.008	-0.006	0.007	-0.011	-0.014	-0.009	0.068	-0.025	0.018	-0.072	-0.034
4	0.009	0.010	-0.008	1.000	0.835	0.018	0.022	0.024	-0.028	-0.046	-0.094	-0.119	-0.026	0.010
5	0.010	0.010	-0.006	0.835	1.000	0.006	0.014	0.017	-0.014	-0.035	-0.065	-0.088	-0.019	0.010
6	0.021	0.029	0.007	0.018	0.006	1.000	0.246	0.220	-0.899	-0.005	-0.205	0.016	-0.073	0.028
7	0.010	0.014	-0.011	0.022	0.014	0.246	1.000	0.893	-0.220	-0.135	-0.078	-0.069	0.016	0.013
8	0.010	0.014	-0.014	0.024	0.017	0.220	0.893	1.000	-0.195	-0.122	-0.082	-0.068	0.030	0.014
9	-0.026	-0.037	-0.009	-0.028	-0.014	-0.899	-0.220	-0.195	1.000	-0.009	0.188	0.003	0.079	-0.036
10	-0.007	-0.010	0.068	-0.046	-0.035	-0.005	-0.135	-0.122	-0.009	1.000	0.008	0.007	-0.013	-0.010
11	-0.014	-0.022	-0.025	-0.094	-0.065	-0.205	-0.078	-0.082	0.188	0.008	1.000	0.017	0.012	-0.021
12	-0.029	-0.044	0.018	-0.119	-0.088	0.016	-0.069	-0.068	0.003	0.007	0.017	1.000	0.008	-0.042
13	-0.006	-0.009	-0.072	-0.026	-0.019	-0.073	0.016	0.030	0.079	-0.013	0.012	0.008	1.000	-0.008
14	0.887	0.995	-0.034	0.010	0.010	0.028	0.013	0.014	-0.036	-0.010	-0.021	-0.042	-0.008	1.000

Table 23: Correlations coefficients from traditional fit to the sum of all detectors.

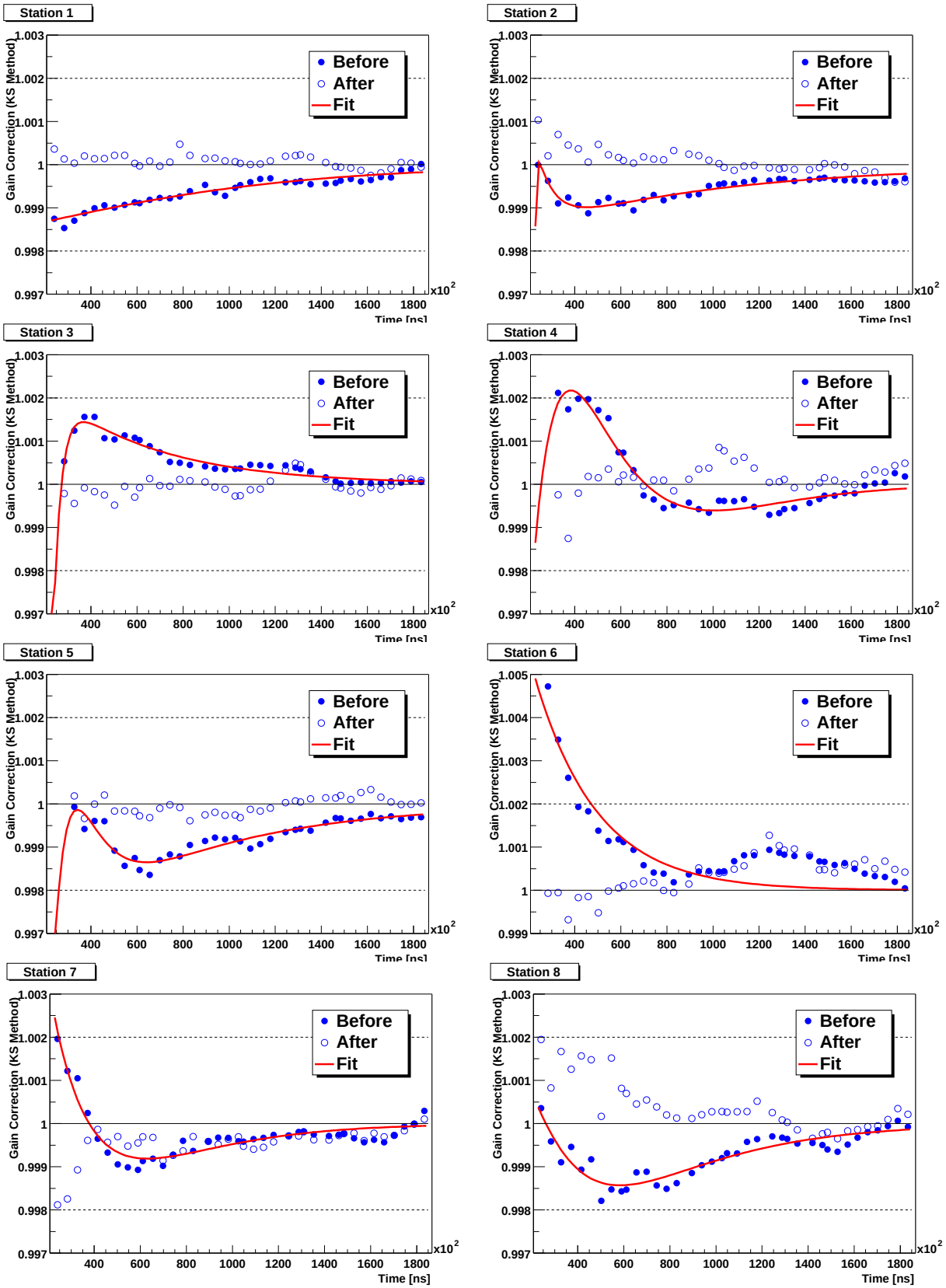
Low n Value																			
Par Name	τ	N_0	A	ϕ	ω_a	ω_{cbo}	τ_{cbo}	a_1	ϕ_1	A_2	ϕ_2	A_3	ϕ_3	A_{loss}	ω_{vw}	τ_{vw}	a_{vw}	$\phi_1 vw$	A_{br}
Par No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1.000	0.897	0.019	0.024	0.019	-0.017	-0.028	-0.037	0.023	0.016	0.024	-0.017	0.050	0.910	-0.005	0.003	-0.001	0.007	-0.610
2	0.897	1.000	0.029	0.033	0.025	-0.027	-0.038	-0.050	0.034	0.021	0.033	-0.022	0.069	0.997	-0.007	0.002	0.001	0.009	-0.791
3	0.019	0.029	1.000	-0.001	-0.000	0.002	-0.008	-0.014	-0.001	-0.080	-0.006	0.064	-0.035	0.025	-0.003	-0.003	0.004	0.003	-0.062
4	0.024	0.033	-0.001	1.000	0.821	-0.022	0.016	0.023	0.011	-0.024	0.103	0.053	0.053	0.032	0.001	0.008	-0.010	-0.002	-0.056
5	0.019	0.025	-0.000	0.821	1.000	-0.036	0.008	0.013	0.026	-0.019	0.077	0.039	0.041	0.024	-0.001	0.007	-0.008	0.000	-0.027
6	-0.017	-0.027	0.002	-0.022	-0.036	1.000	0.186	0.171	-0.909	0.000	-0.147	0.014	-0.180	-0.026	0.026	-0.063	0.059	-0.027	0.069
7	-0.028	-0.038	-0.008	0.016	0.008	0.186	1.000	0.904	-0.169	-0.117	-0.066	0.162	-0.052	-0.037	-0.001	-0.026	0.026	0.001	0.070
8	-0.037	-0.050	-0.014	0.023	0.013	0.171	0.904	1.000	-0.156	-0.107	-0.056	0.152	-0.060	-0.048	-0.002	-0.028	0.028	0.003	0.088
9	0.023	0.034	-0.001	0.011	0.026	-0.909	-0.169	-0.156	1.000	0.002	0.138	0.001	0.174	0.033	-0.025	0.053	-0.049	0.027	-0.085
10	0.016	0.021	-0.080	-0.024	-0.019	0.000	-0.117	-0.107	0.002	1.000	0.001	-0.031	-0.014	0.021	0.001	0.001	-0.002	-0.001	-0.017
11	0.024	0.033	-0.006	0.103	0.077	-0.147	-0.066	-0.056	0.138	0.001	1.000	-0.007	0.045	0.032	-0.002	0.011	-0.011	0.001	-0.076
12	-0.017	-0.022	0.064	0.053	0.039	0.014	0.162	0.152	0.001	-0.031	-0.007	1.000	-0.078	-0.022	-0.009	-0.019	0.023	0.011	0.017
13	0.050	0.069	-0.035	0.053	0.041	-0.180	-0.052	-0.060	0.174	-0.014	0.045	-0.078	1.000	0.067	-0.005	0.014	-0.014	0.005	-0.139
14	0.910	0.997	0.025	0.032	0.024	-0.026	-0.037	-0.048	0.033	0.021	0.032	-0.022	0.067	1.000	-0.007	0.002	0.001	0.008	-0.772
15	-0.005	-0.007	-0.003	0.001	-0.001	0.026	-0.001	-0.002	-0.025	0.001	-0.002	-0.009	-0.005	-0.007	1.000	-0.043	0.042	-0.976	0.014
16	0.003	0.002	-0.003	0.008	0.007	-0.063	-0.026	-0.028	0.053	0.001	0.011	-0.019	0.014	0.002	-0.043	1.000	-0.972	0.042	0.009
17	-0.001	0.001	0.004	-0.010	-0.008	0.059	0.026	0.028	-0.049	-0.002	-0.011	0.023	-0.014	0.001	0.042	-0.972	1.000	-0.042	-0.015
18	0.007	0.009	0.003	-0.002	0.000	-0.027	0.001	0.003	0.027	-0.001	0.001	0.011	0.005	0.008	-0.976	0.042	-0.042	1.000	-0.016
19	-0.610	-0.791	-0.062	-0.056	-0.042	0.069	0.070	0.088	-0.085	-0.017	-0.076	0.017	-0.139	-0.772	0.014	0.009	-0.015	-0.016	1.000

High n Value																			
Par Name	τ	N_0	A	ϕ	ω_a	ω_{cbo}	τ_{cbo}	a_1	ϕ_1	A_2	ϕ_2	A_3	ϕ_3	A_{loss}	ω_{vw}	τ_{vw}	a_{vw}	$\phi_1 vw$	A_{br}
Par No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	1.000	0.889	0.021	0.013	0.010	0.005	0.020	0.027	-0.007	-0.012	0.000	-0.022	0.014	0.903	-0.014	0.001			
2	0.889	1.000	0.033	0.018	0.013	0.007	0.028	0.037	-0.011	-0.018	-0.000	-0.031	0.020	0.996	-0.019	-0.001			
3	0.021	0.033	1.000	-0.007	-0.006	0.022	0.003	0.003	-0.025	-0.049	-0.065	-0.076	-0.023	0.028	-0.004	0.000			
4	0.013	0.018	-0.007	1.000	0.822	0.000	-0.035	-0.043	0.006	0.097	-0.055	0.070	-0.105	0.017	0.013	0.004			
5	0.010	0.013	-0.006	0.822	1.000	-0.000	-0.025	-0.031	0.004	0.071	-0.038	0.053	-0.075	0.013	0.010	0.003			
6	0.005	0.007	0.022	0.000	-0.000	1.000	-0.132	-0.123	-0.880	-0.015	-0.306	0.022	-0.061	0.007	0.015	0.003			
7	0.020	0.028	0.003	-0.035	-0.025	-0.132	1.000	0.899	0.117	-0.233	0.046	-0.086	0.014	0.027	-0.015	0.000			
8	0.027	0.037	0.003	-0.043	-0.031	-0.123	0.899	1.000	0.109	-0.215	0.039	-0.084	0.034	0.036	-0.020	0.000			
9	-0.007	-0.011	-0.025	0.006	0.004	-0.880	0.117	0.109	1.000	0.008	0.274	-0.006	0.058	-0.010	-0.019	-0.006			
10	-0.012	-0.018	-0.049	0.097	0.071	-0.015	-0.233	-0.215	0.008	1.000	0.000	0.036	-0.004	-0.017	0.001	0.004			
11	0.000	-0.000	-0.065	-0.055	-0.038	-0.306	0.046	0.039	0.274	0.000	1.000	-0.017	0.035	0.000	-0.003	0.000			
12	-0.022	-0.031	-0.076	0.070	0.053	0.022	-0.086	-0.084	-0.006	0.036	-0.017	1.000	-0.041	-0.030	-0.009	-0.010			
13	0.014	0.020	-0.023	-0.105	-0.075	-0.061	0.014	0.034	0.058	-0.004	0.035	-0.041	1.000	0.019	-0.021	-0.002			
14	0.903	0.996	0.028	0.017	0.013	0.007	0.027	0.036	-0.010	-0.017	0.000	-0.030	0.019	1.000	-0.018	-0.001			
15	-0.014	-0.019	-0.004	0.013	0.010	0.015	-0.015	-0.020	-0.019	0.001	-0.003	-0.009	-0.021	-0.018	1.000	-0.125			
16	0.001	-0.001	0.000	0.004	0.003	0.003	0.000	0.000	-0.006	0.004	0.000	-0.010	-0.002	-0.001	-0.125	1.000			
17	-0.000	0.002	-0.000	-0.005	-0.004	-0.004	0.000	0.001	0.008	-0.004	0.000	0.011	0.003	0.002	0.123	-0.990	1.000	-0.120	-0.014
18	0.016	0.021	0.004	-0.015	-0.011	-0.016	0.017	0.022	0.021	-0.001	0.003	0.010	0.024	0.020	-0.991	0.123	-0.120	1.000	-0.043
18	-0.556	-0.751	-0.074	-0.025	-0.019	-0.010	-0.060	-0.079	0.018	0.047	-0.001	0.078	-0.054	-0.729	0.038	0.011	-0.014	-0.043	1.000

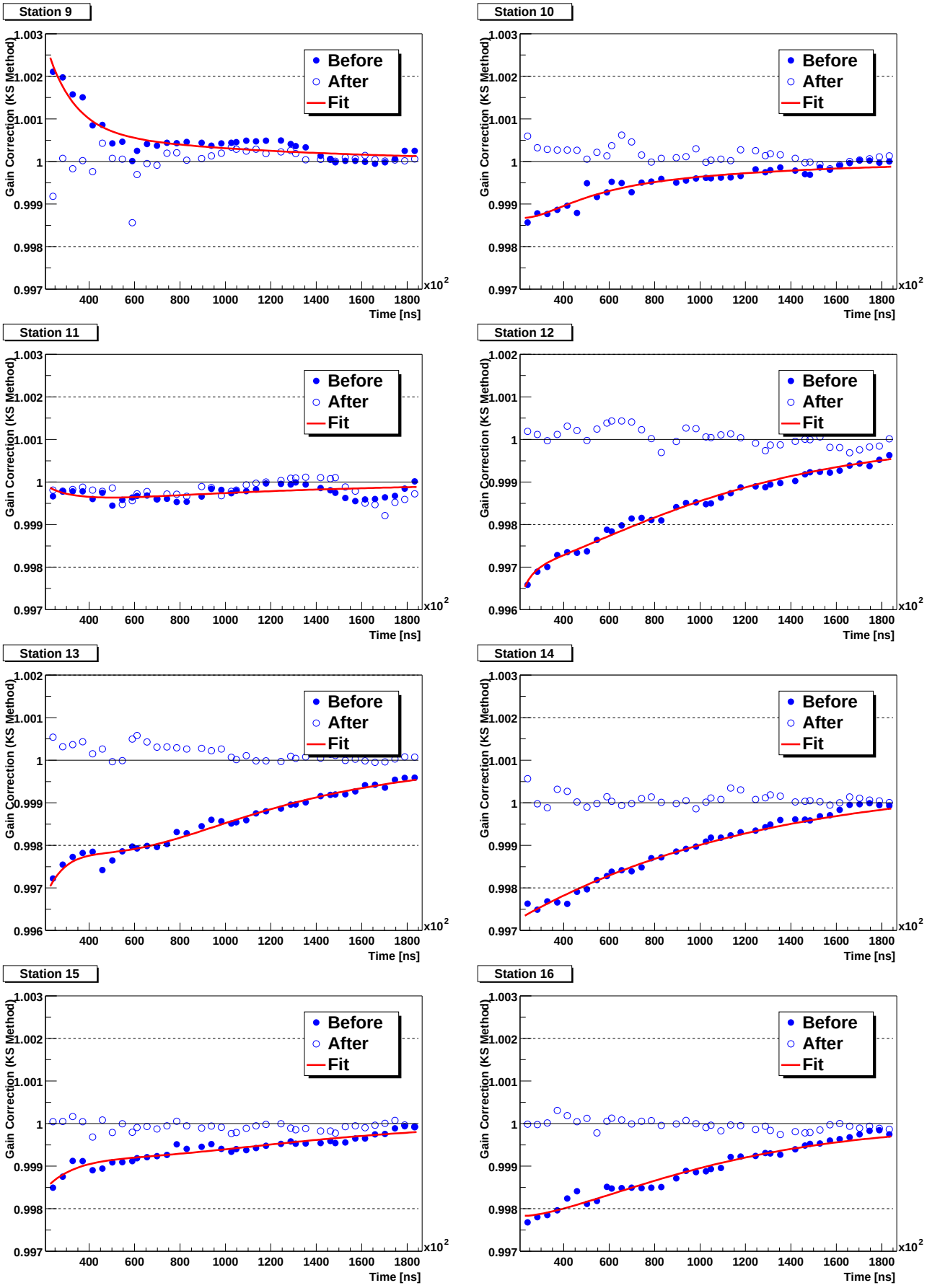
Table 24: Correlations coefficients from the asymmetry-weighted fit to the early detector group (A).

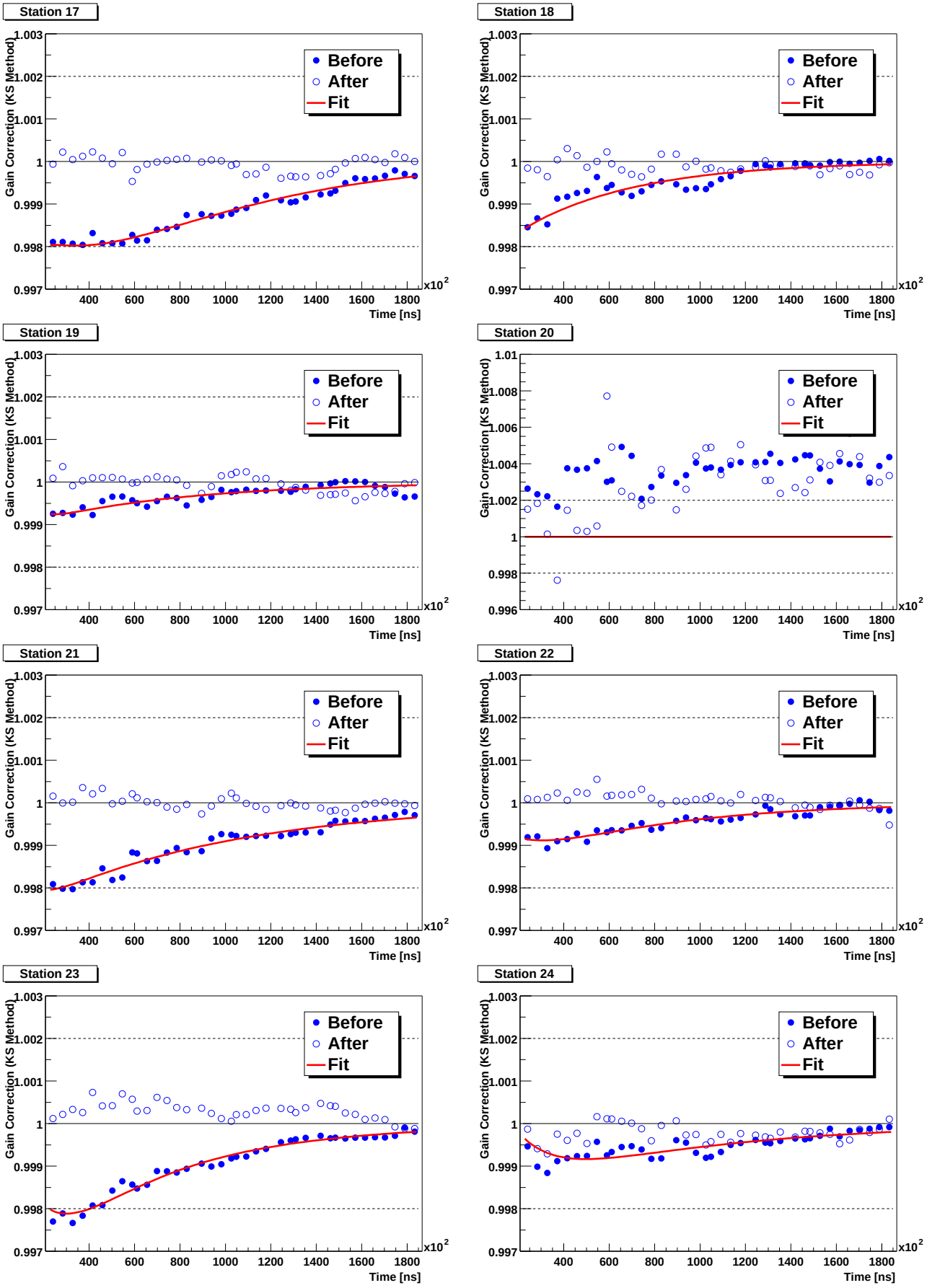
D Gain Related Plots

D.1 Gain Curves Constructed from KS Method

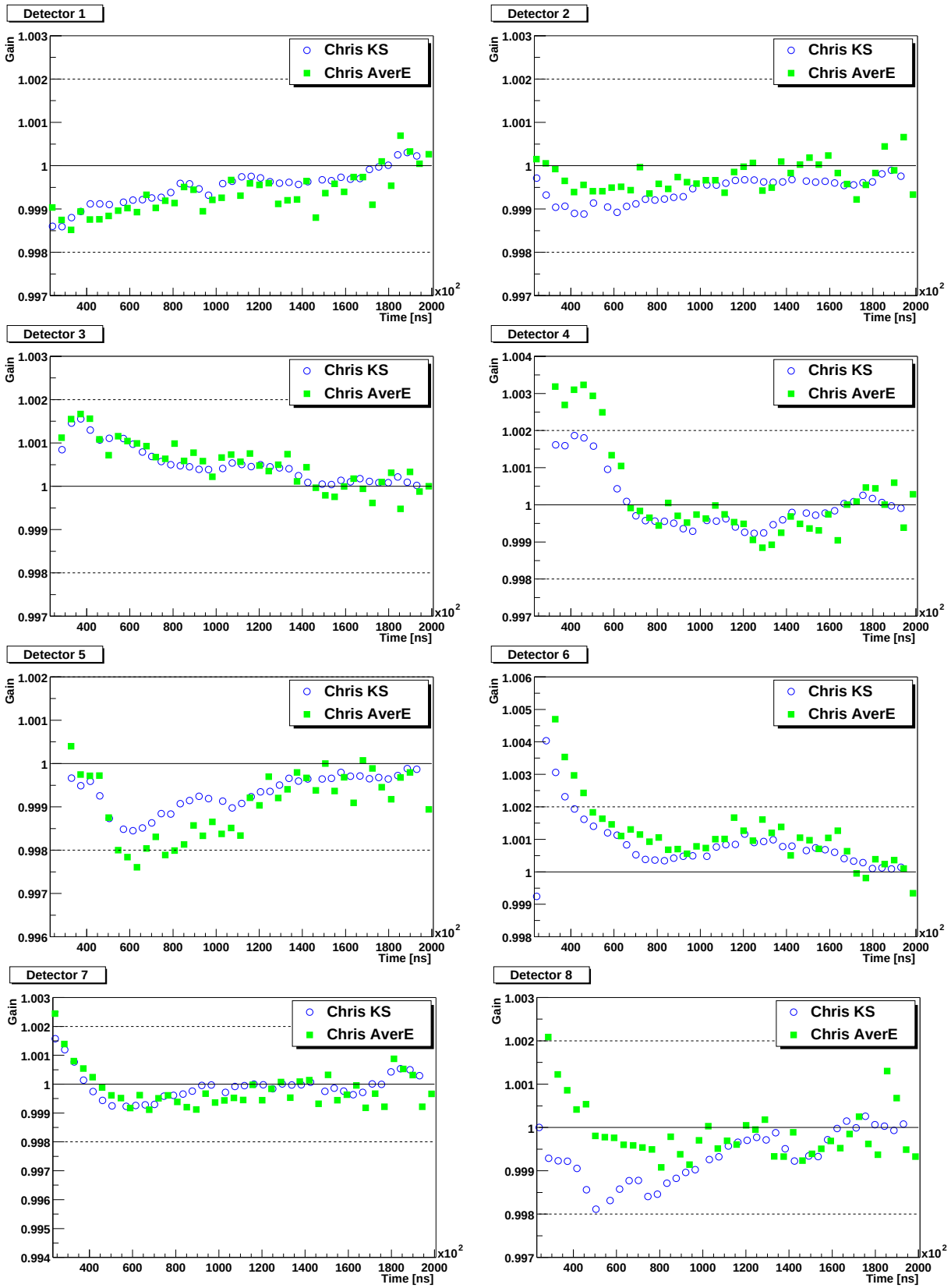


The plots in this section show the gain curve constructed using the KS method (blue closed circles), the fit to the gain curve (red line), and the gain curve constructed after the correction has been applied (blue open circles).

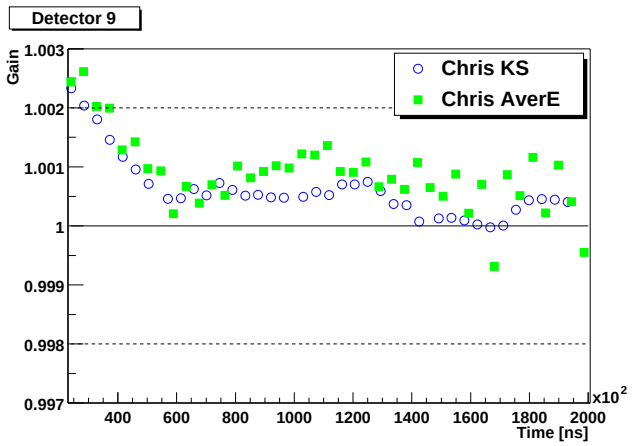

Station 9
Station 10
Station 11
Station 12
Station 13
Station 14
Station 15
Station 16
Before
After
Fit
Gain Correction (KS Method)
Time [ns]
x10²

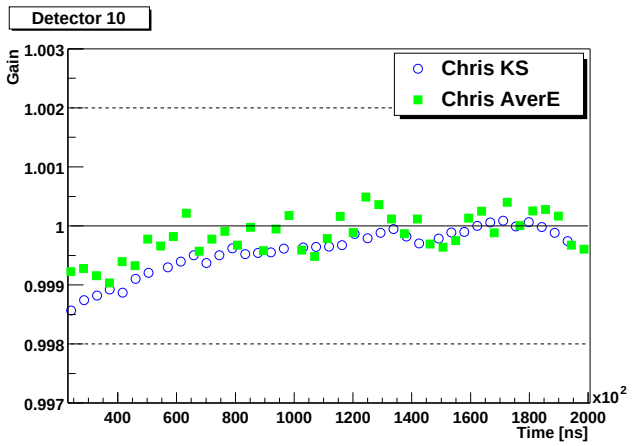

Station 17
Station 18
Station 19
Station 20
Station 21
Station 22
Station 23
Station 24
Before
After
Fit
Gain Correction (KS Method)
Time [ns]
x10²

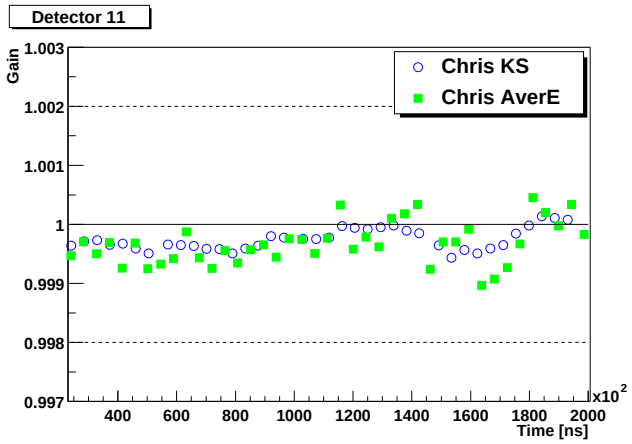
D.2 Comparison of KS and Average Energy Method

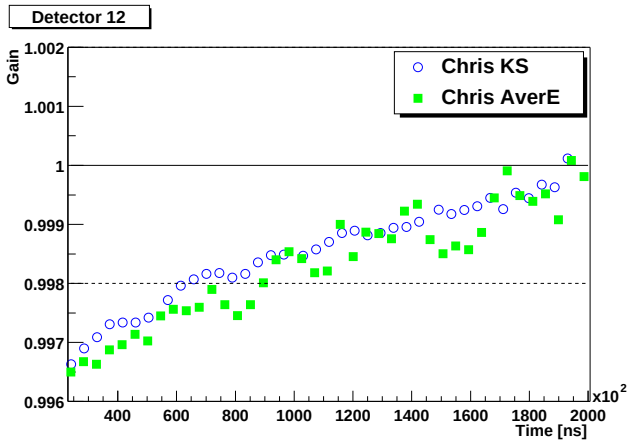


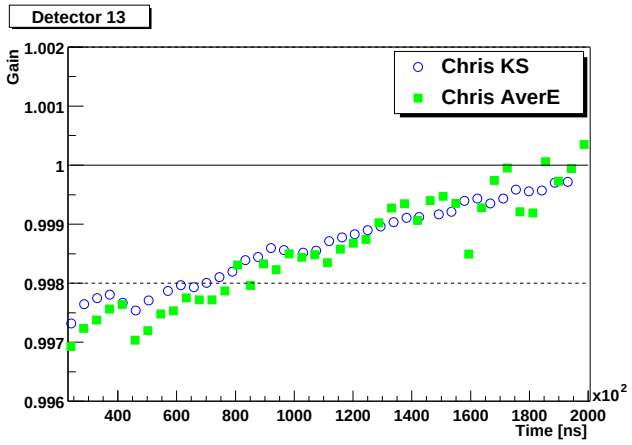
These plots show the gain correction constructed using the KS method (blue open circles) and average energy (green closed squares).

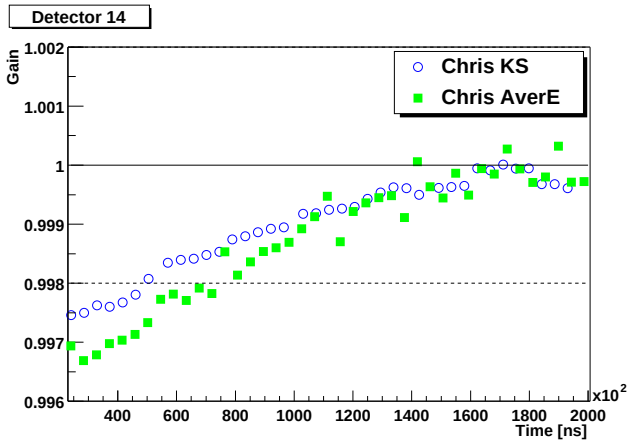

Detector 9
Gain
Chris KS
Chris AverE
Time [ns]
x10²

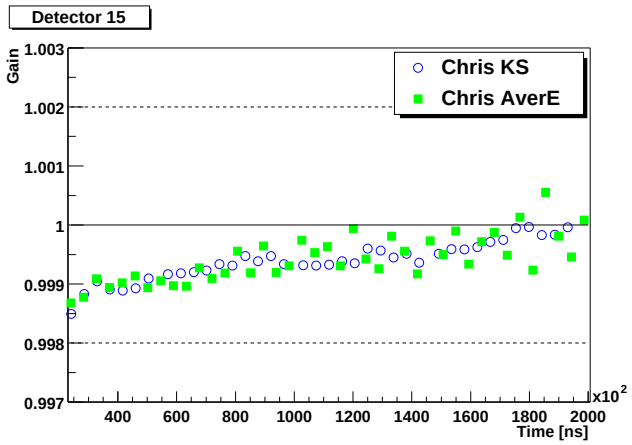

Detector 10
Gain
Chris KS
Chris AverE
Time [ns]
x10²

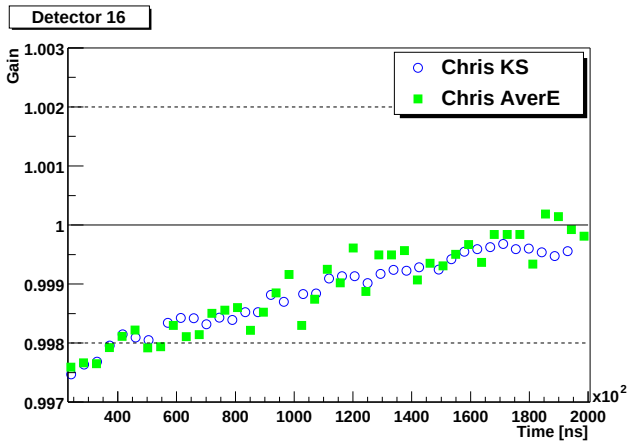

Detector 11
Gain
Chris KS
Chris AverE
Time [ns]
x10²

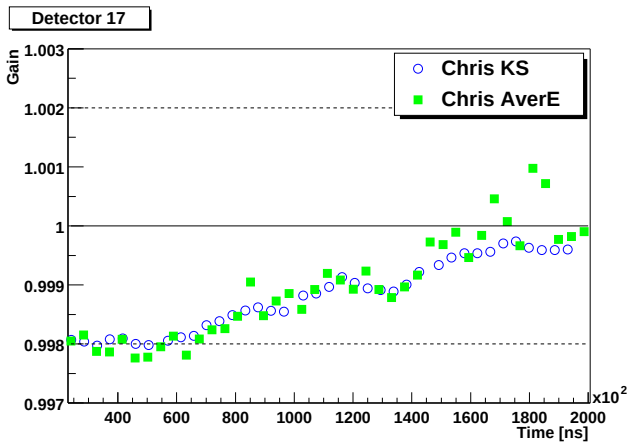

Detector 12
Gain
Chris KS
Chris AverE
Time [ns]
x10²

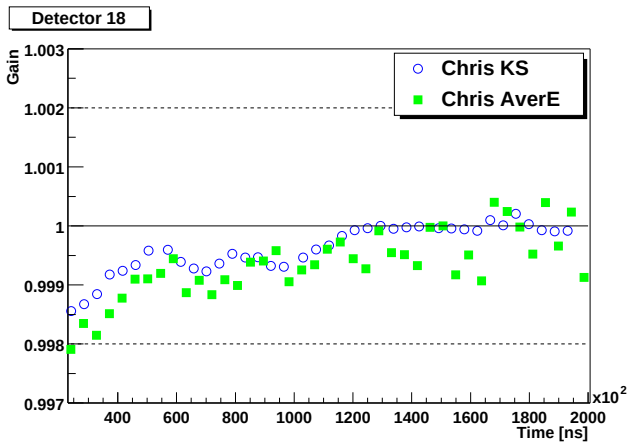

Detector 13
Gain
Chris KS
Chris AverE
Time [ns]
x10²

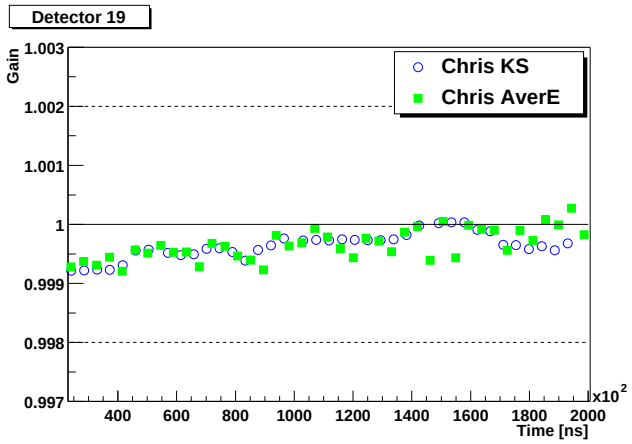

Detector 14
Gain
Chris KS
Chris AverE
Time [ns]
x10²

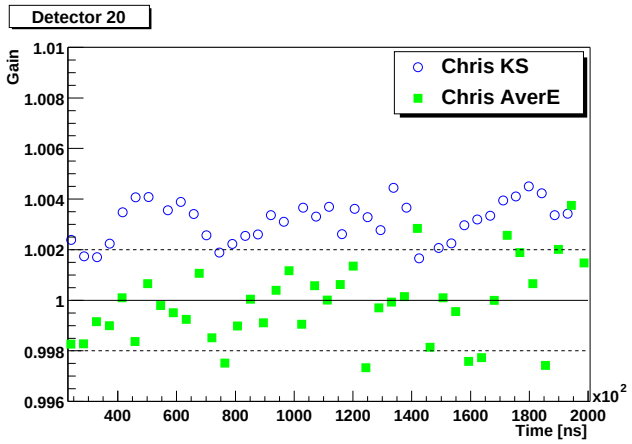

Detector 15
Gain
Chris KS
Chris AverE
Time [ns]
x10²

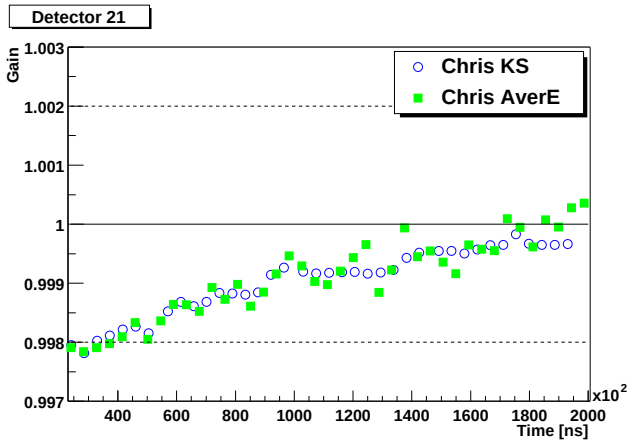

Detector 16
Gain
Chris KS
Chris AverE
Time [ns]
x10²

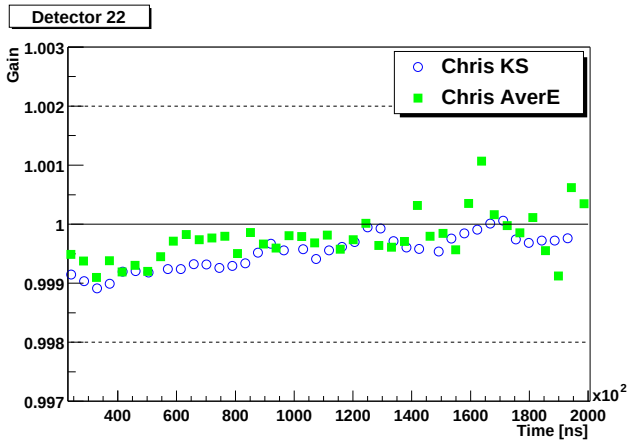

Detector 17
Chris KS
Chris AverE
Gain
Time [ns]
x10²

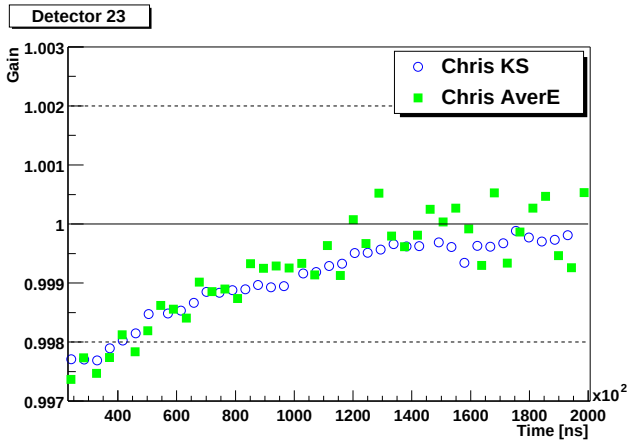

Detector 18
Chris KS
Chris AverE
Gain
Time [ns]
x10²

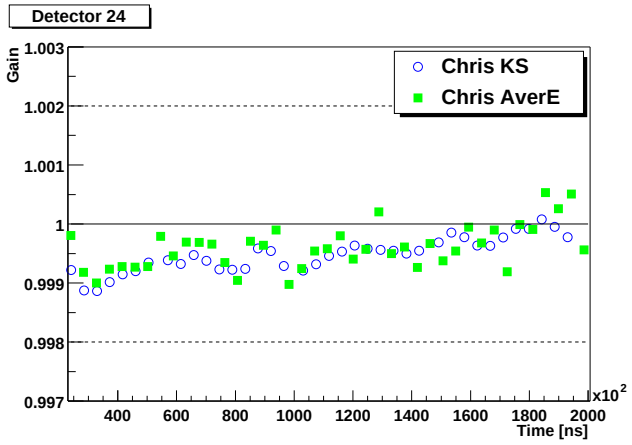

Detector 19
Chris KS
Chris AverE
Gain
Time [ns]
x10²


Detector 20
Chris KS
Chris AverE
Gain
Time [ns]
x10²

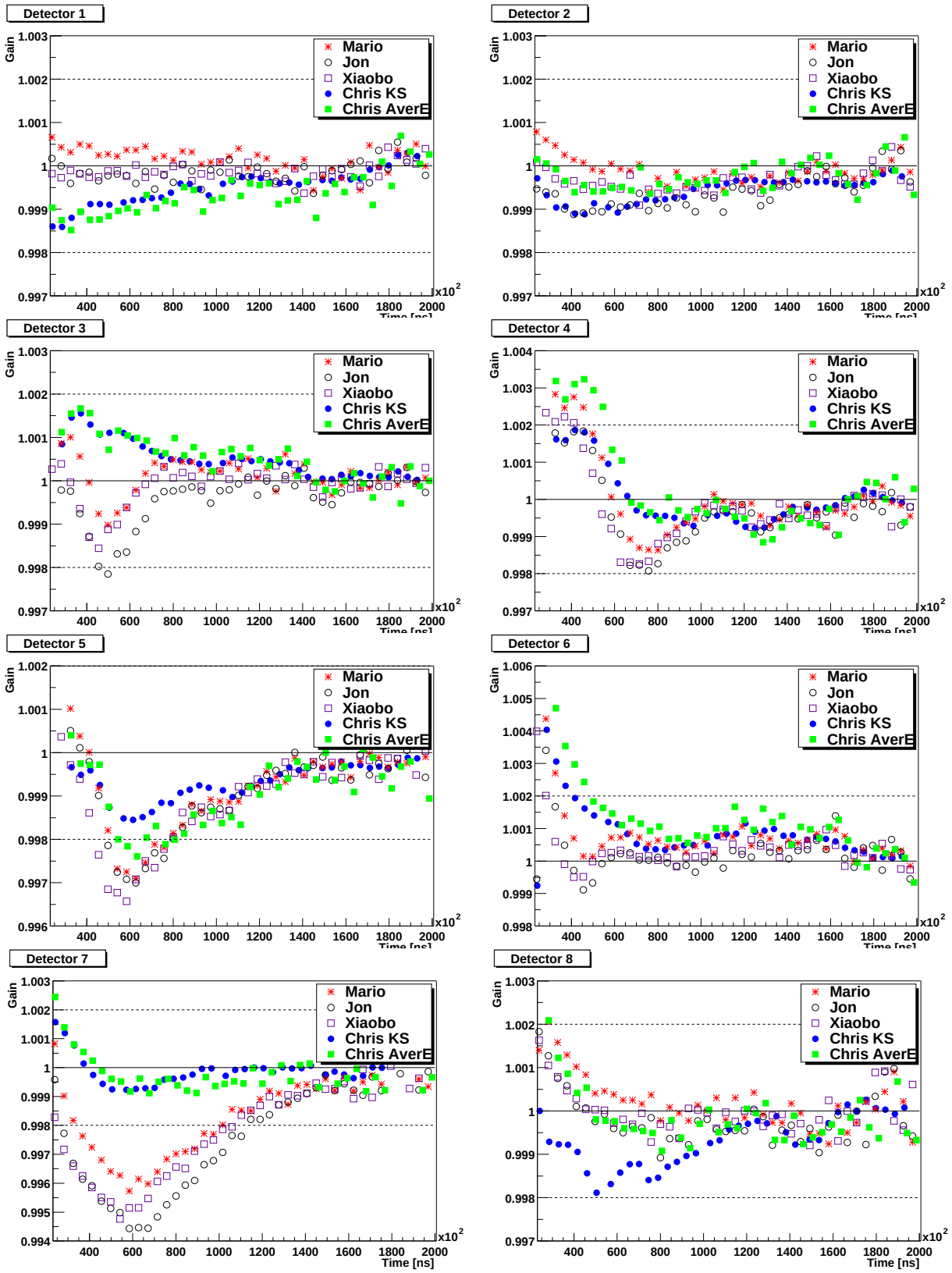

Detector 21
Chris KS
Chris AverE
Gain
Time [ns]
x10²


Detector 22
Chris KS
Chris AverE
Gain
Time [ns]
x10²

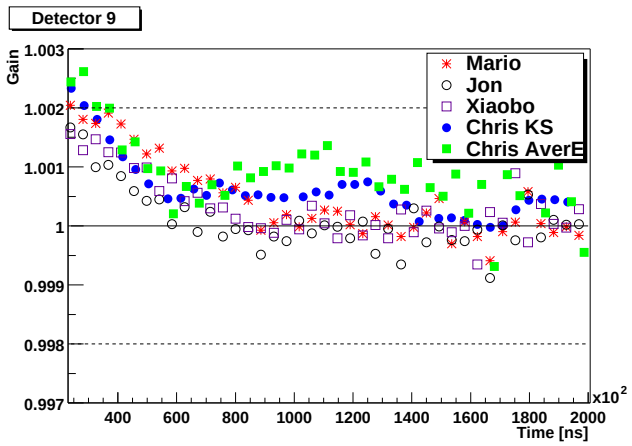

Detector 23
Chris KS
Chris AverE
Gain
Time [ns]
x10²

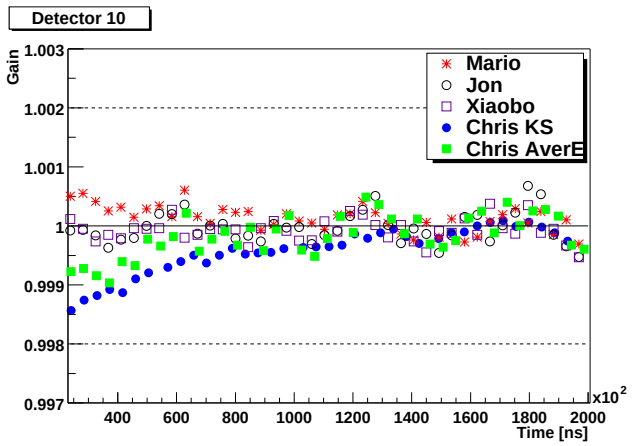

Detector 24
Chris KS
Chris AverE
Gain
Time [ns]
x10²

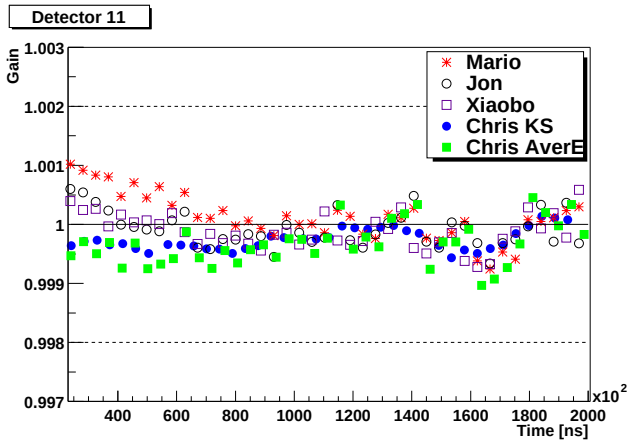
D.3 Gain Comparison Between Analyzers

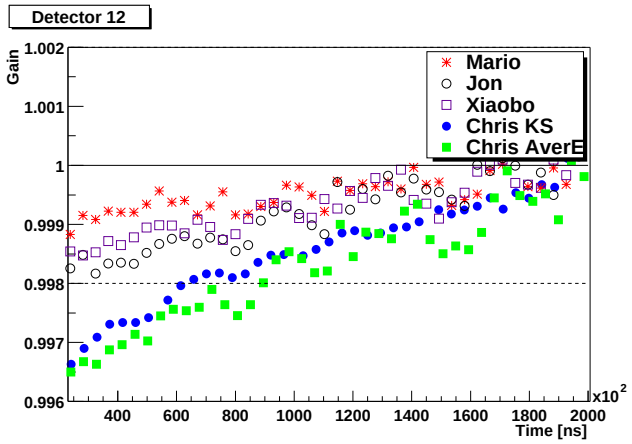


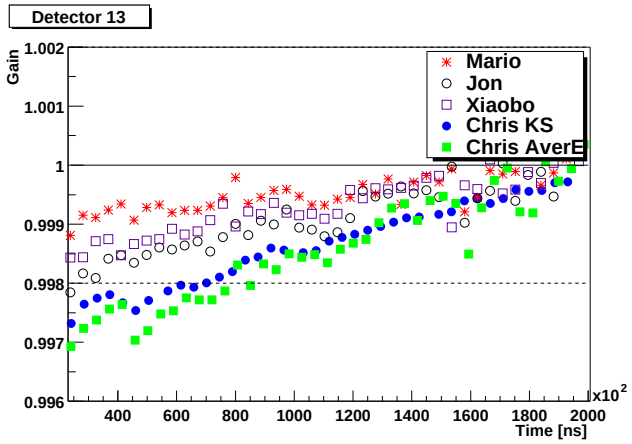
These plots show the gain versus time in spill constructed by all of the analyzers.)

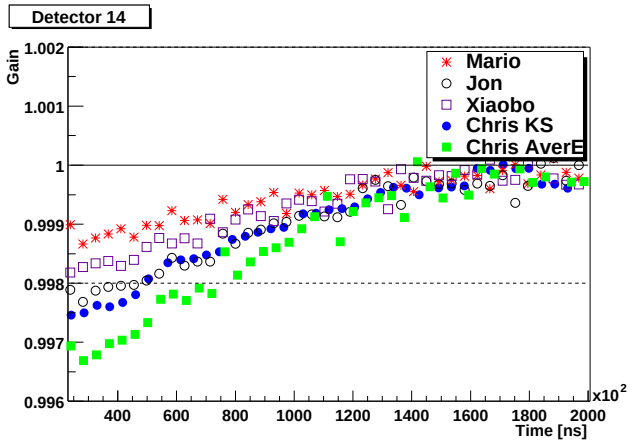

Detector 9
Gain
Time [ns]
x10²
Mario
Jon
Xiaobo
Chris KS
Chris AverE

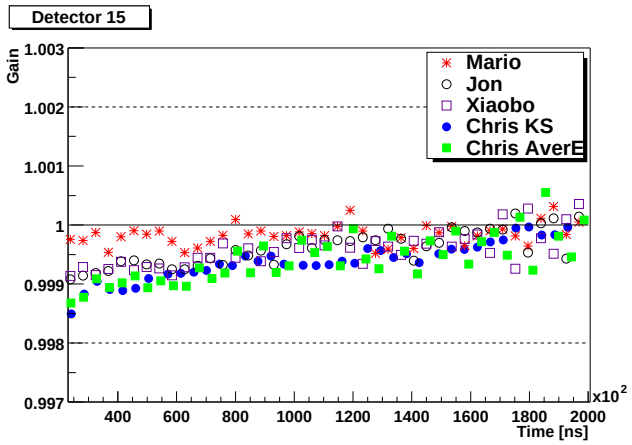

Detector 10
Gain
Time [ns]
x10²
Mario
Jon
Xiaobo
Chris KS
Chris AverE

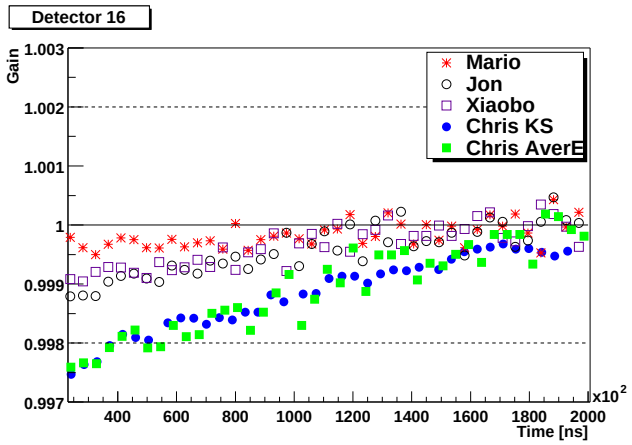

Detector 11
Gain
Time [ns]
x10²
Mario
Jon
Xiaobo
Chris KS
Chris AverE

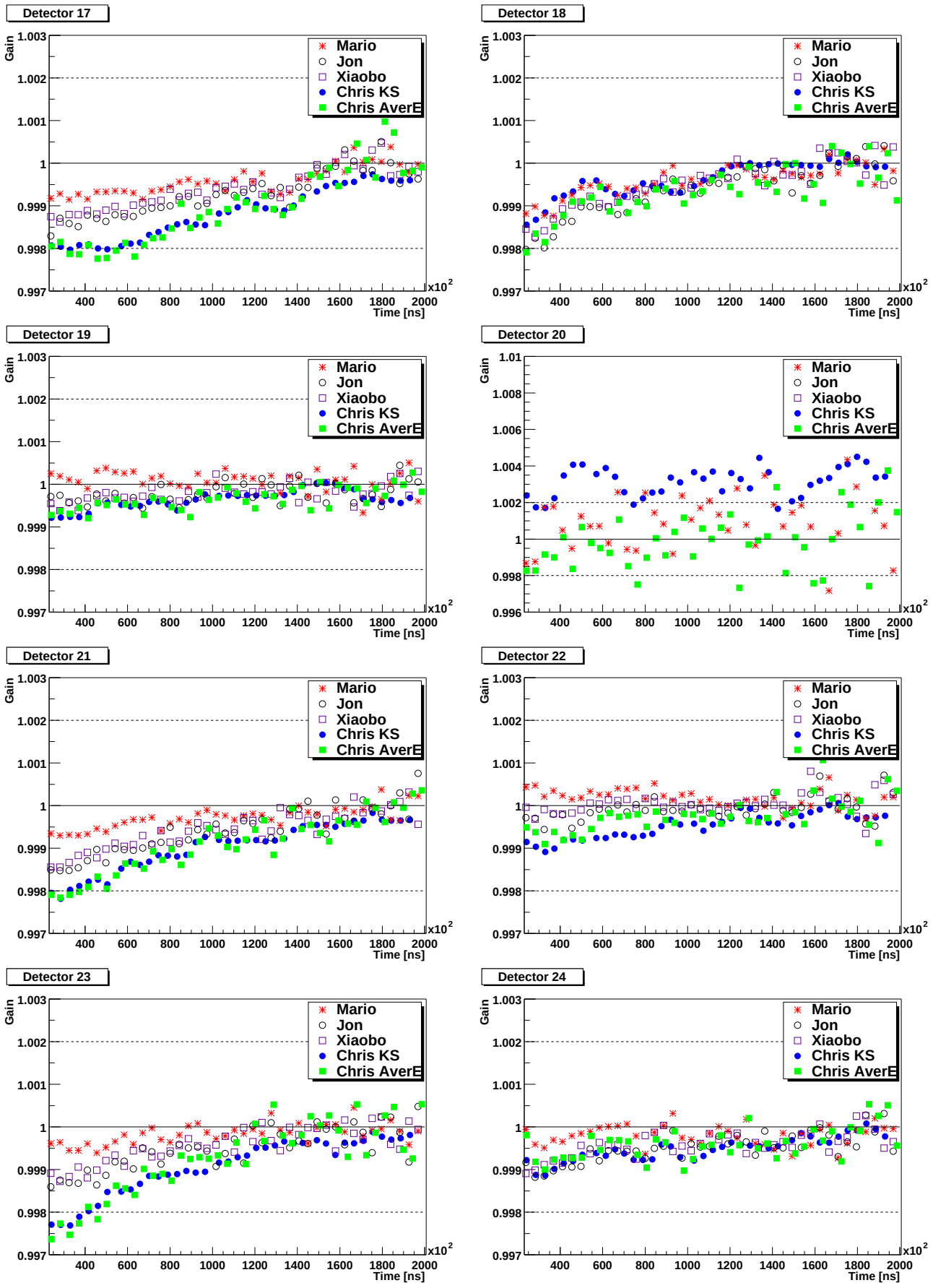

Detector 12
Gain
Time [ns]
x10²
Mario
Jon
Xiaobo
Chris KS
Chris AverE


Detector 13
Gain
Time [ns]
x10²
Mario
Jon
Xiaobo
Chris KS
Chris AverE

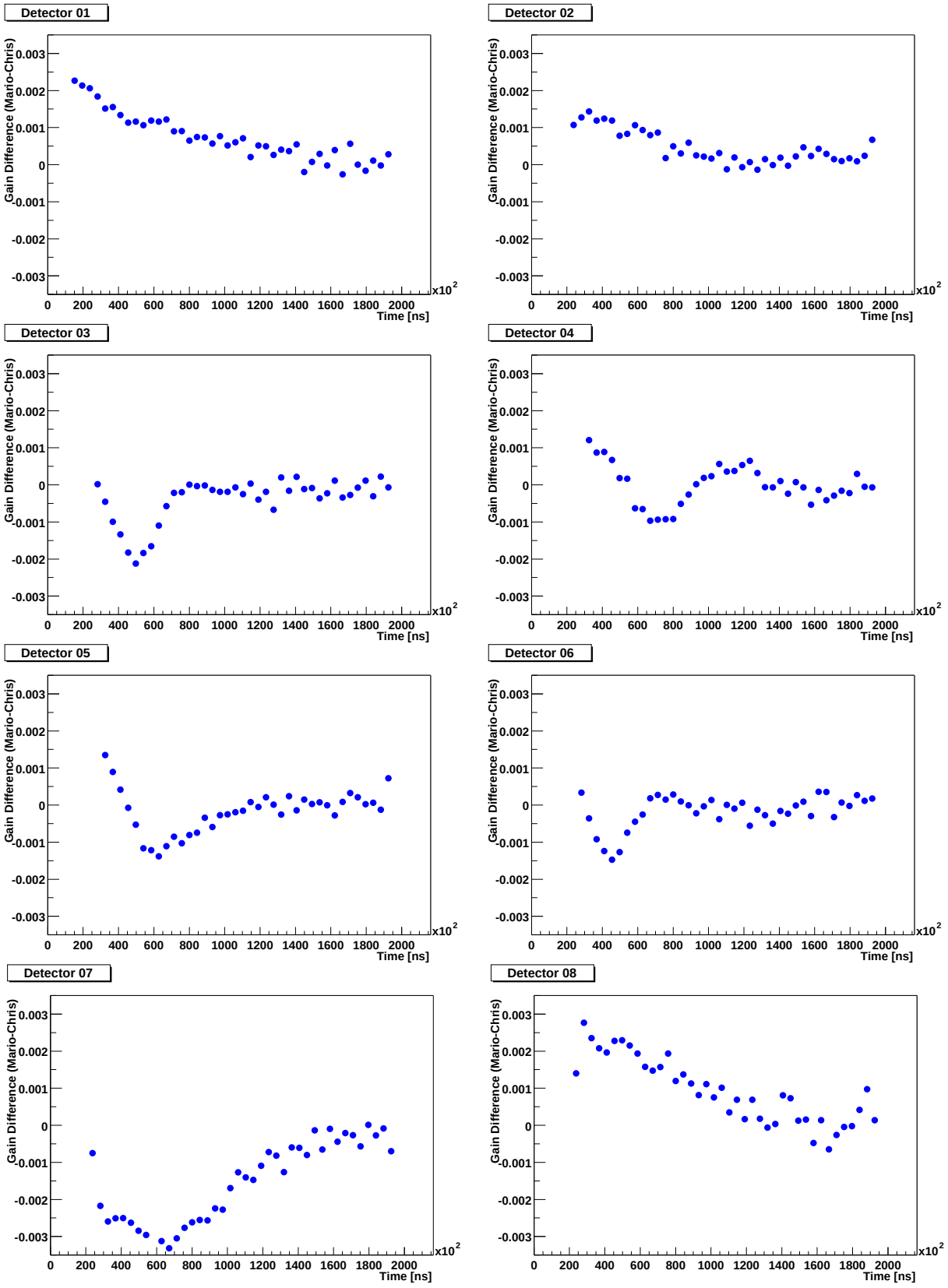

Detector 14
Gain
Time [ns]
x10²
Mario
Jon
Xiaobo
Chris KS
Chris AverE


Detector 15
Gain
Time [ns]
x10²
Mario
Jon
Xiaobo
Chris KS
Chris AverE

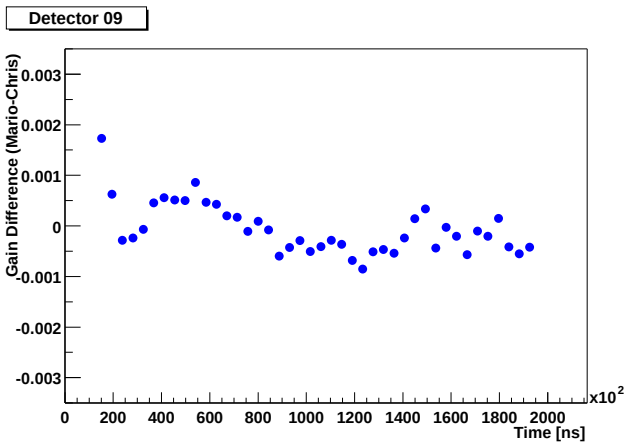

Detector 16
Gain
Time [ns]
x10²
Mario
Jon
Xiaobo
Chris KS
Chris AverE

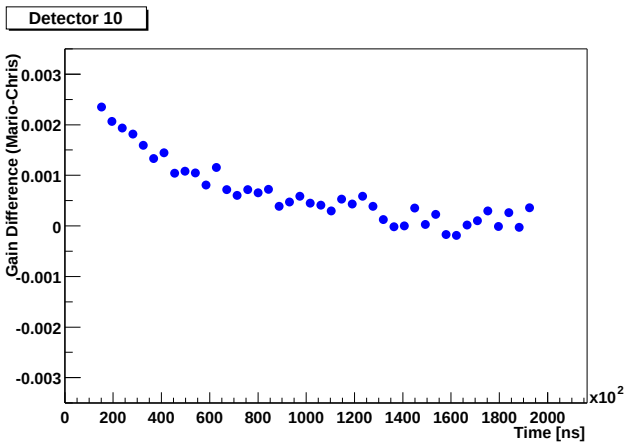

Detector 17
Detector 18
Detector 19
Detector 20
Detector 21
Detector 22
Detector 23
Detector 24
Gain
Time [ns]
x10^2
Mario
Jon
Xiaobo
Chris KS
Chris AverE

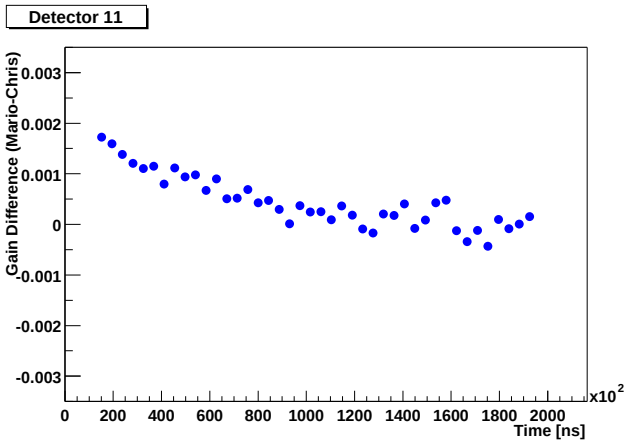
D.4 Difference in G2Too/G2Off Gains

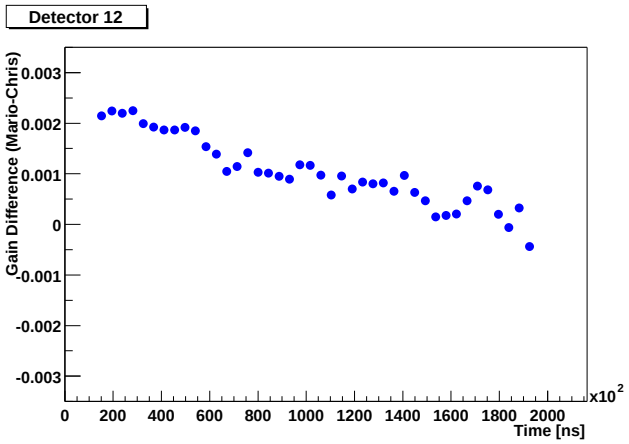


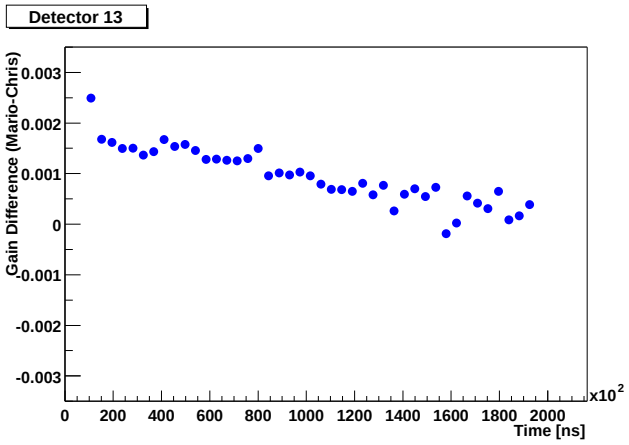
These plots illustrate the difference betweem Marios's gain construction for G2Off and my KS results for G2Too.

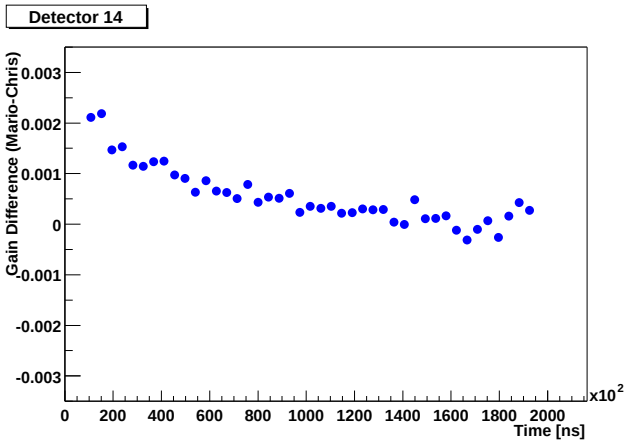

Detector 09
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

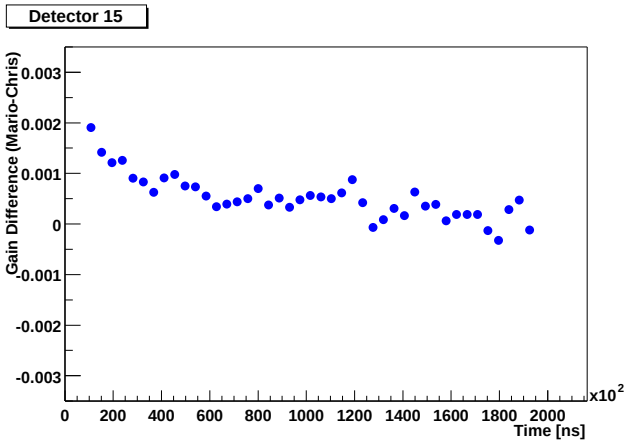

Detector 10
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

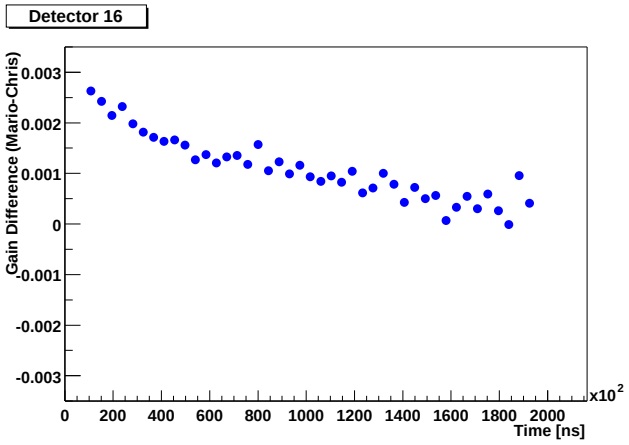

Detector 11
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

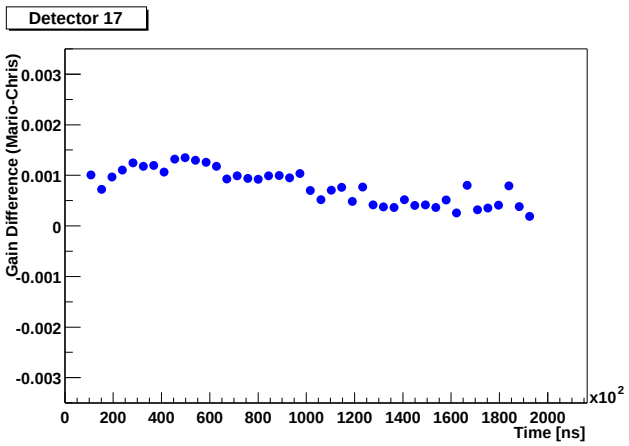

Detector 12
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

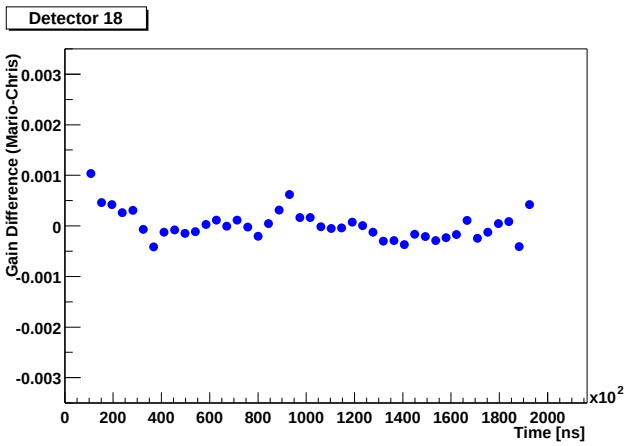

Detector 13
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

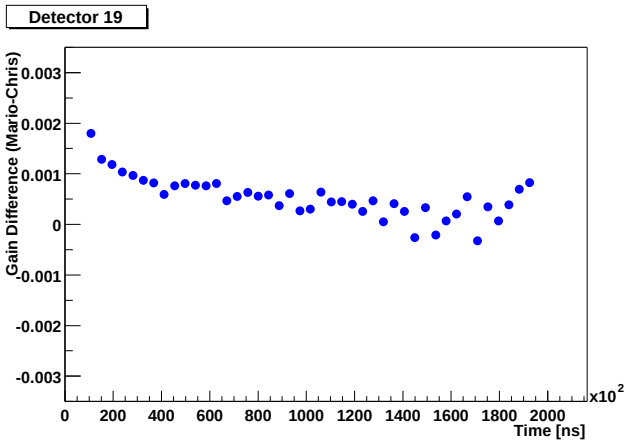

Detector 14
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

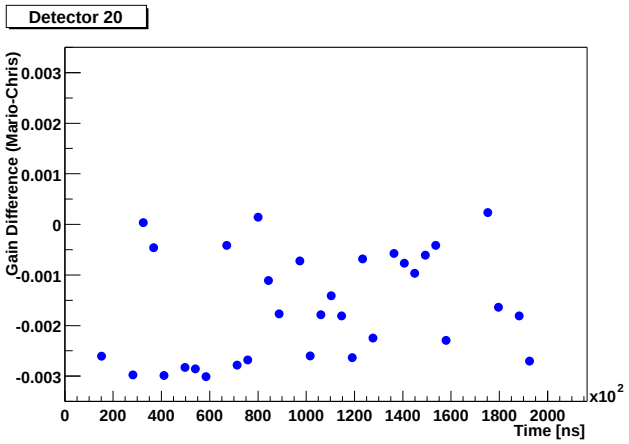

Detector 15
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

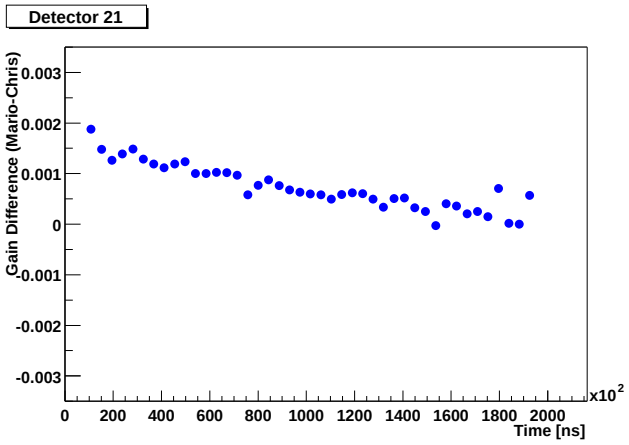

Detector 16
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

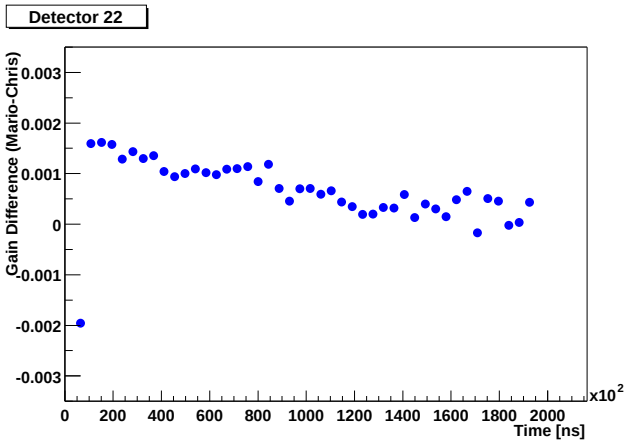

Detector 17
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

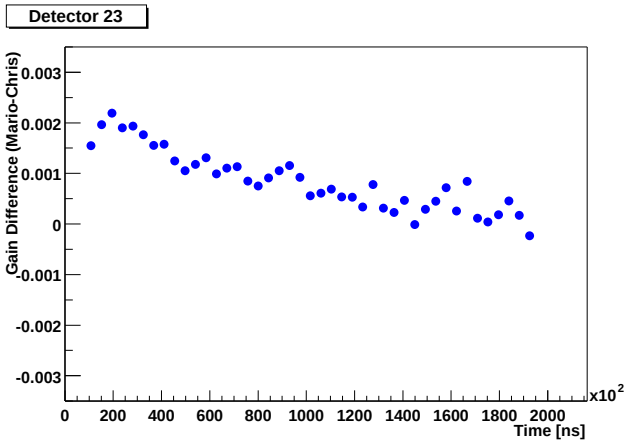

Detector 18
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

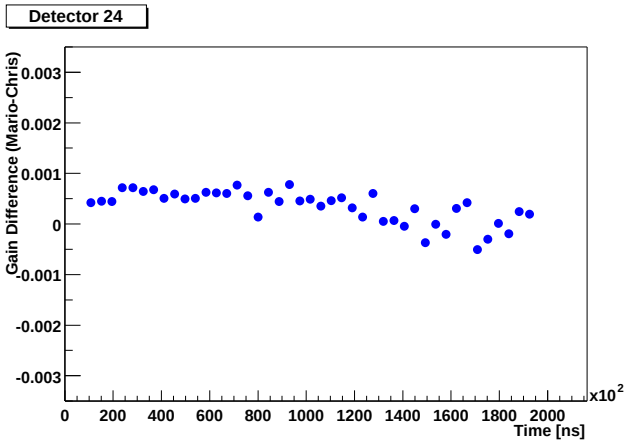

Detector 19
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]


Detector 20
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

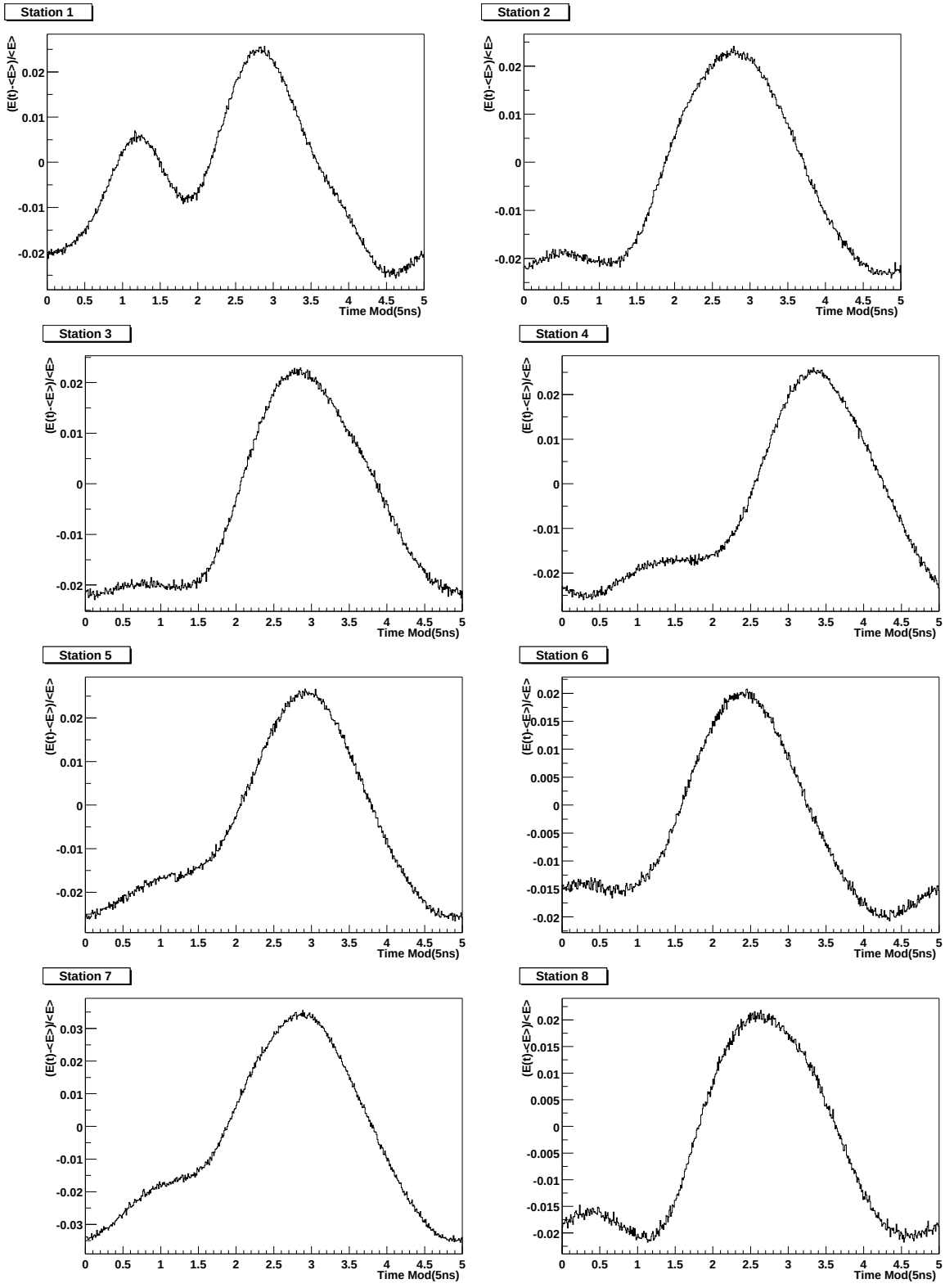

Detector 21
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

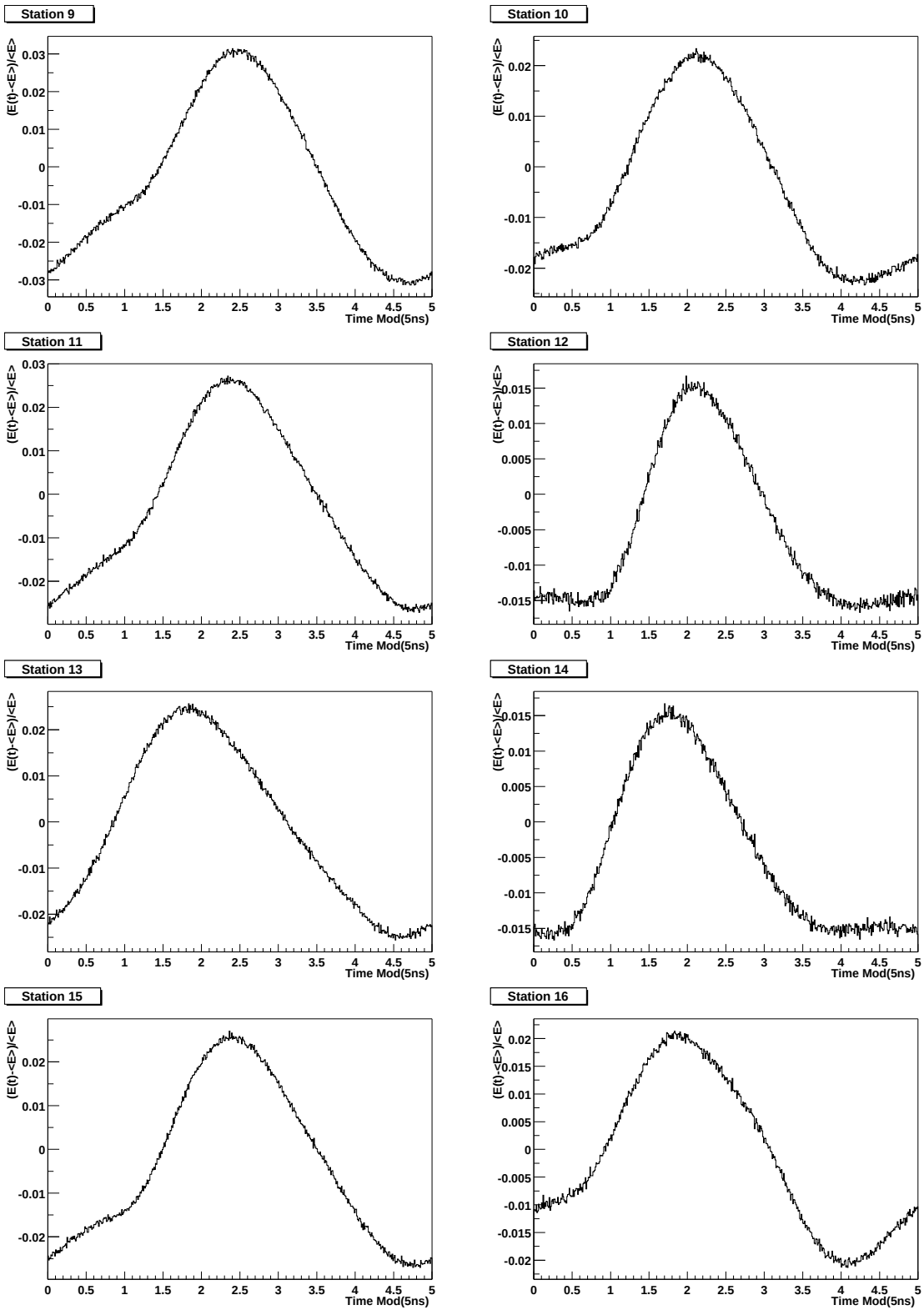

Detector 22
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

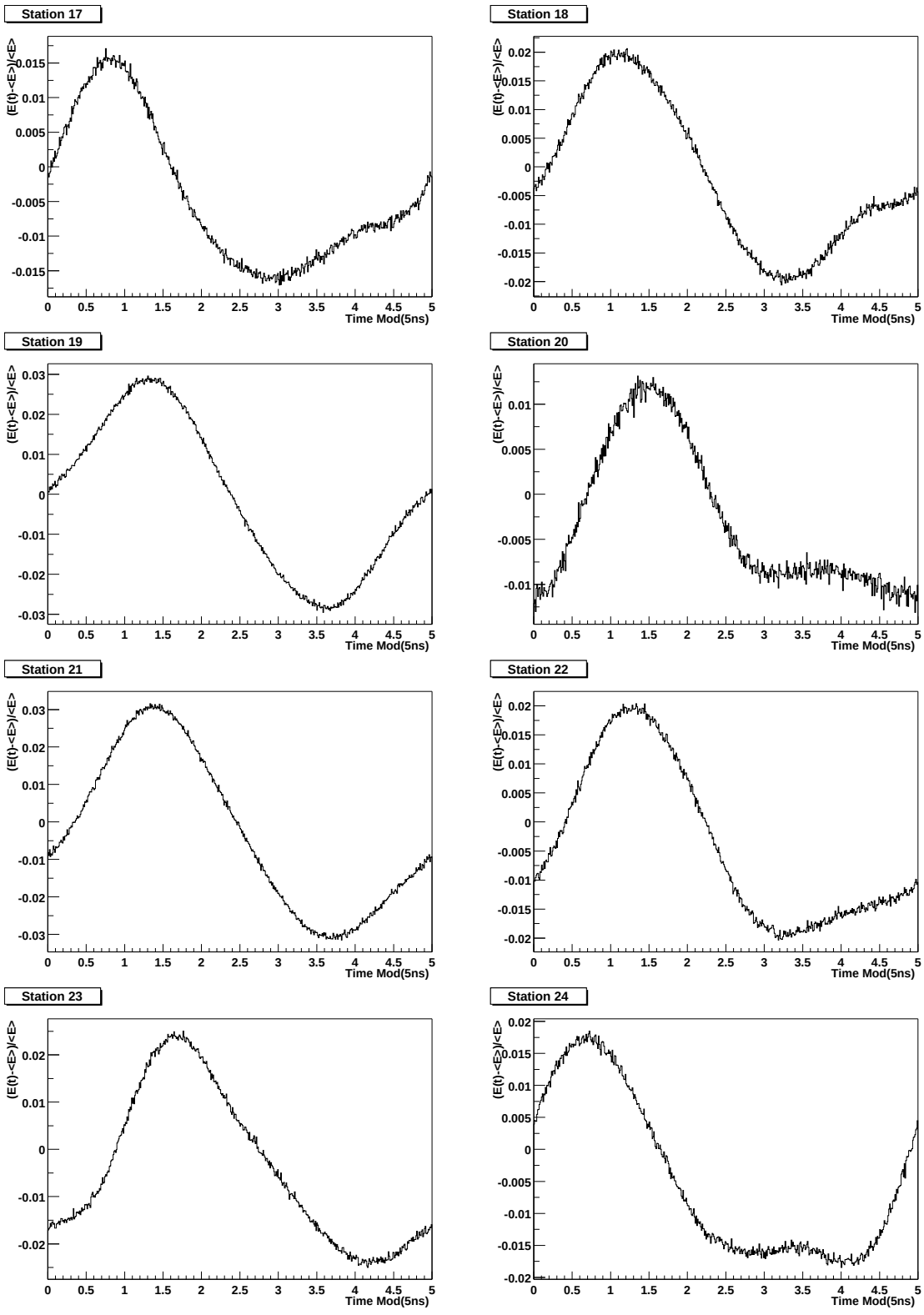

Detector 23
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]


Detector 24
Gain Difference (Mario-Chris)
0.003
0.002
0.001
0
-0.001
-0.002
-0.003
0 200 400 600 800 1000 1200 1400 1600 1800 2000
x10^2
Time [ns]

D.5 Energy Dependence on Clock Boundary

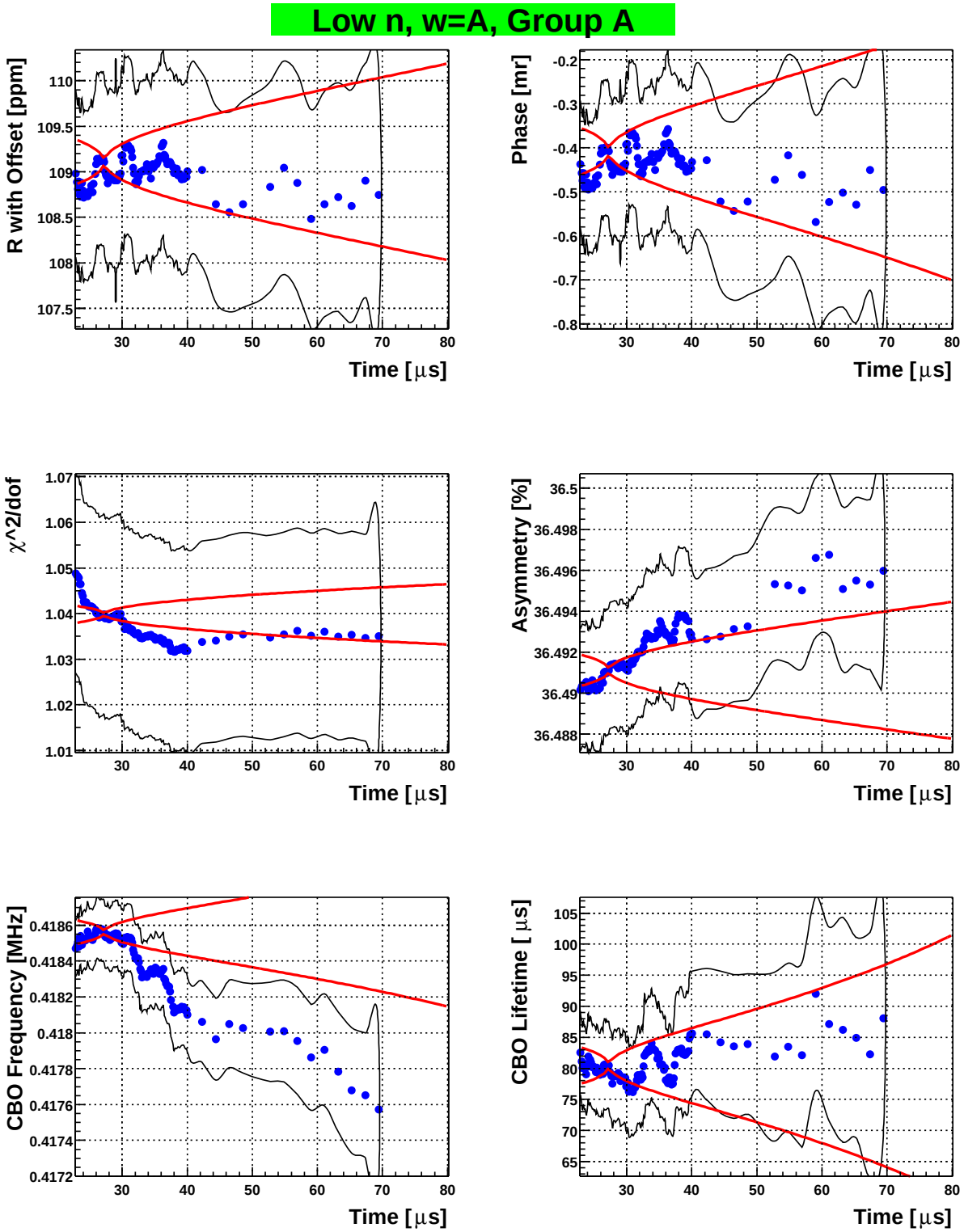


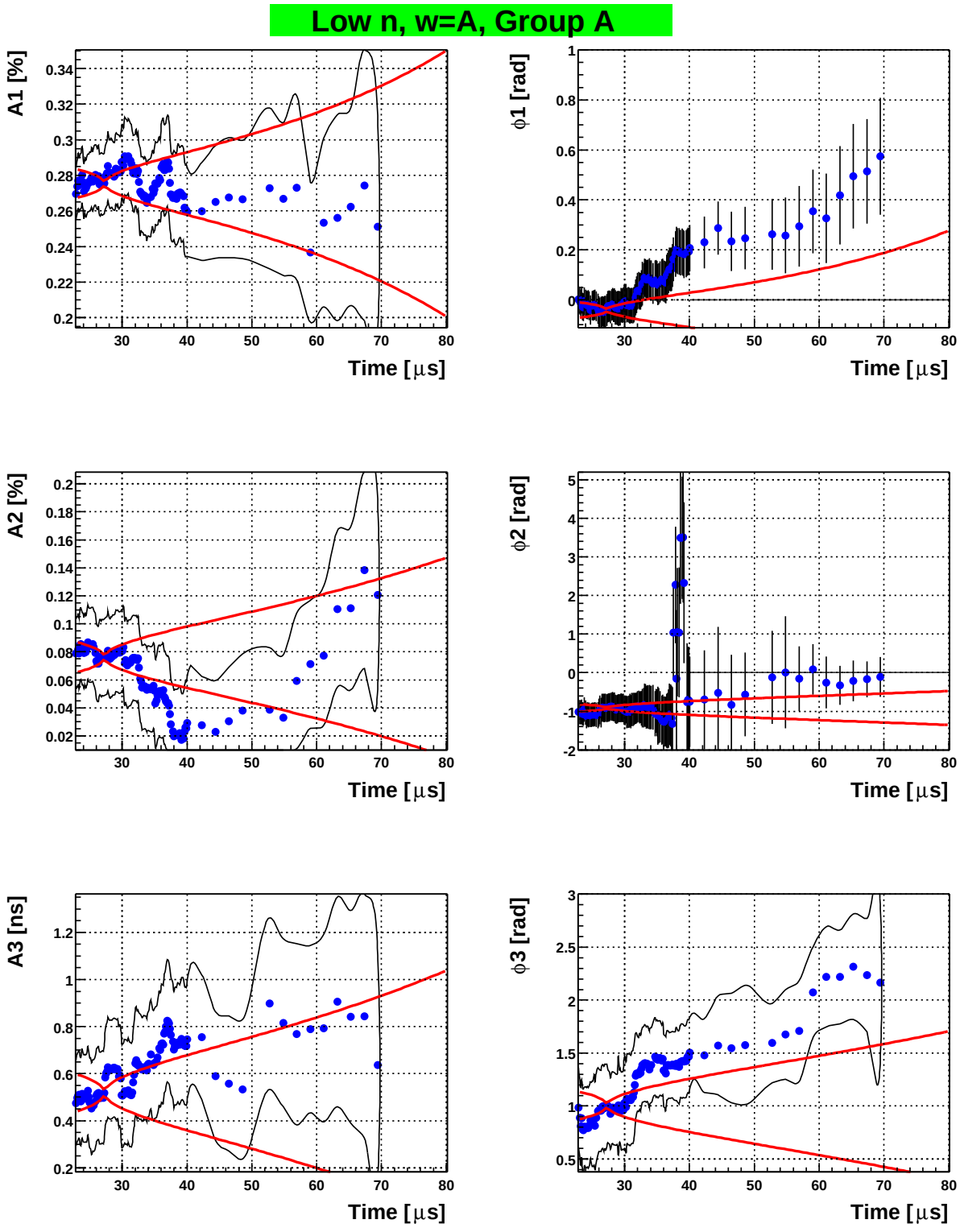

Station 9
Station 10
Station 11
Station 12
Station 13
Station 14
Station 15
Station 16
(E(t)-<E>)/<E>
Time Mod(5ns)


Station 17
Station 18
Station 19
Station 20
Station 21
Station 22
Station 23
Station 24
(E(t)-<E>)/<E>
Time Mod(5ns)

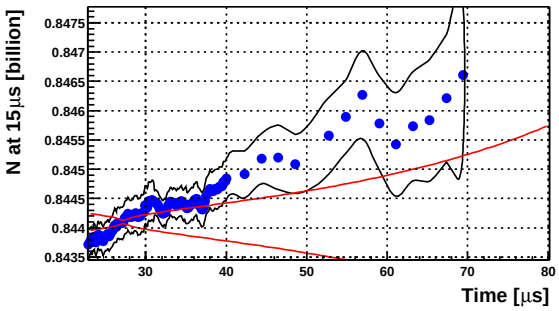
E Start Time Scans

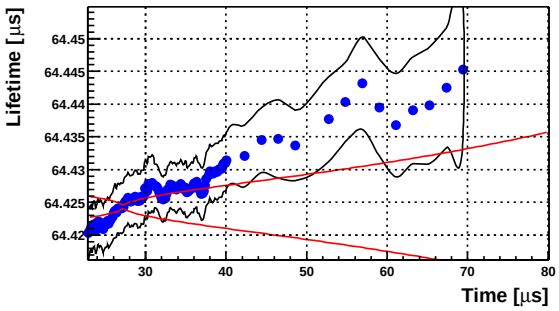
E.1 Asymmetry Weighting; Two Detector Groups

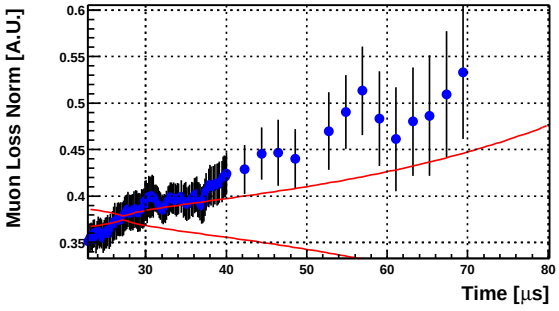


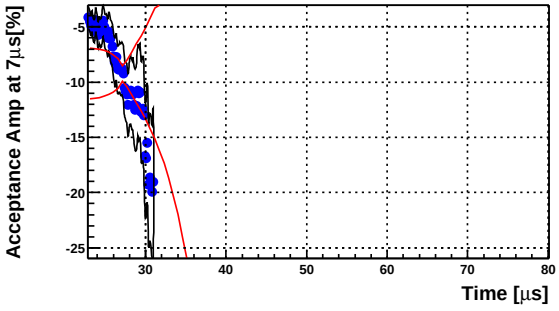

Low n, w=A, Group A
A1 [%]
Time [μs]
φ1 [rad]
Time [μs]
A2 [%]
Time [μs]
φ2 [rad]
Time [μs]
A3 [ns]
Time [μs]
φ3 [rad]
Time [μs]

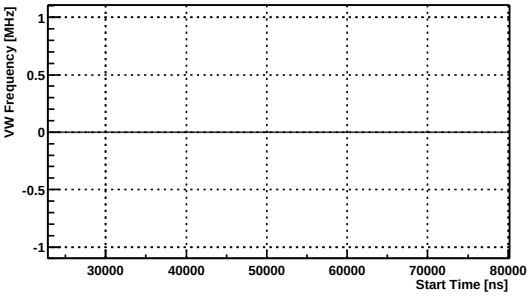
Low n, w=A, Group A

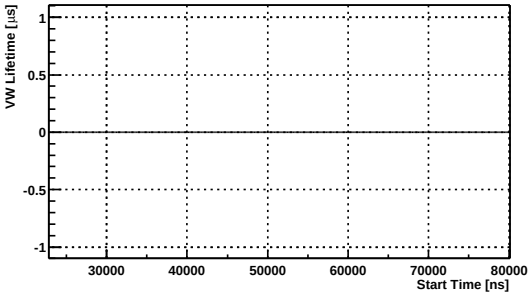


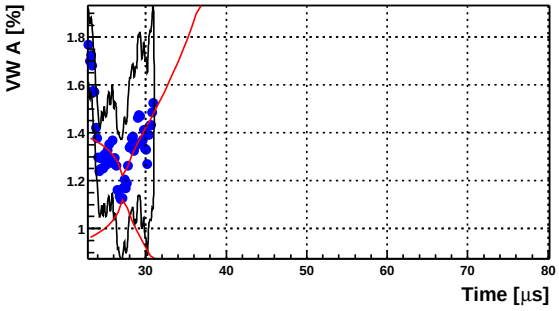


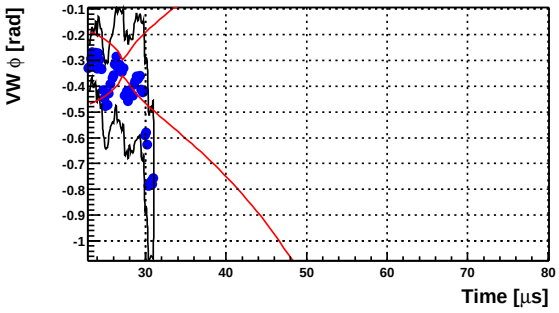


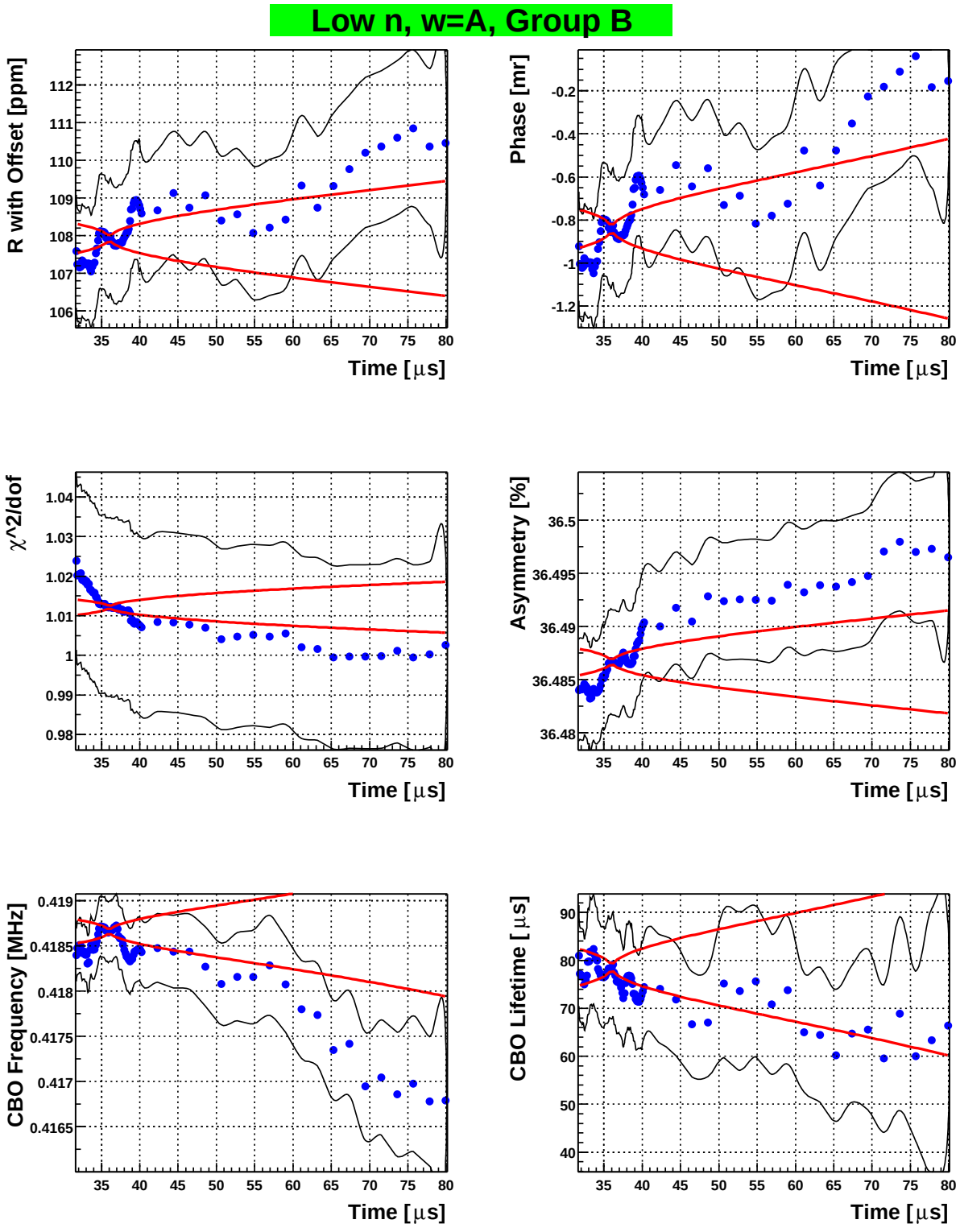


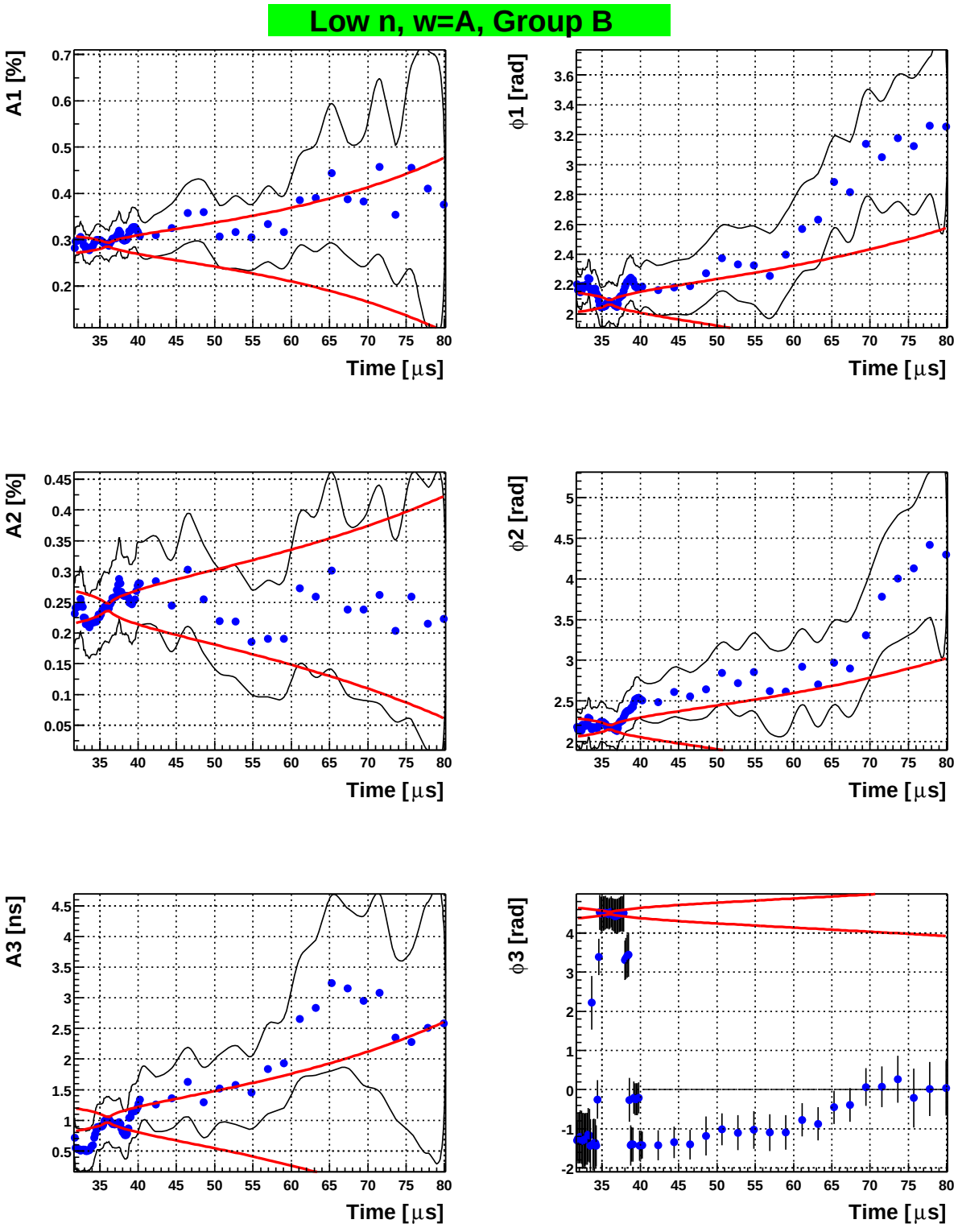




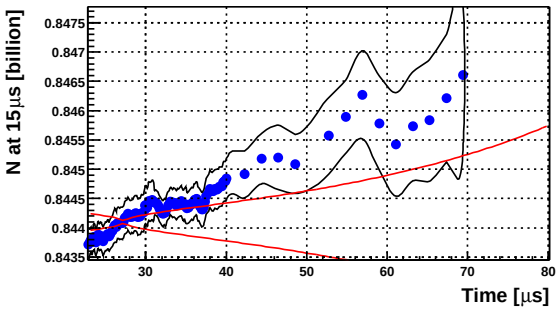


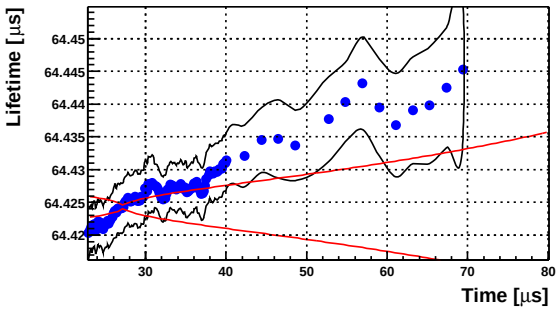


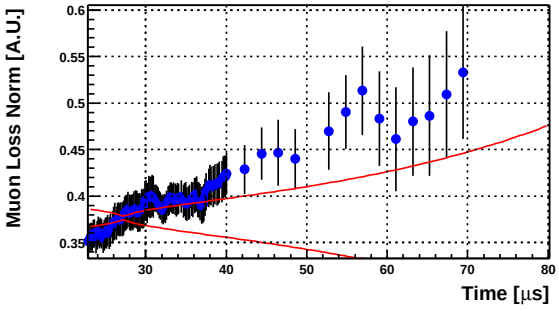

Low n, w=A, Group B
R with Offset [ppm]
Phase [mr]
χ^2/dof
Asymmetry [%]
CBO Frequency [MHz]
CBO Lifetime [μs]
Time [μs]

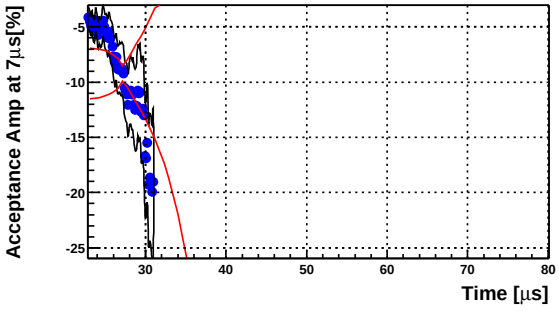

Low n, w=A, Group B
A1 [%]
ϕ1 [rad]
A2 [%]
ϕ2 [rad]
A3 [ns]
ϕ3 [rad]
Time [μs]

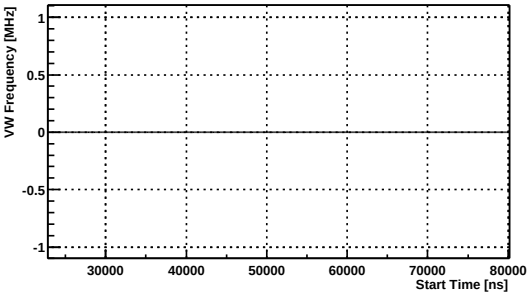
Low n, w=A, Group A

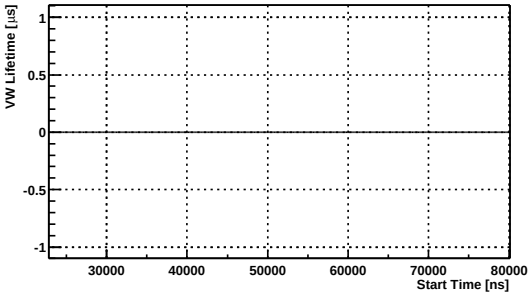


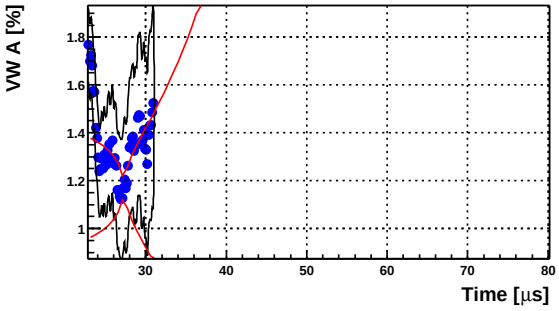


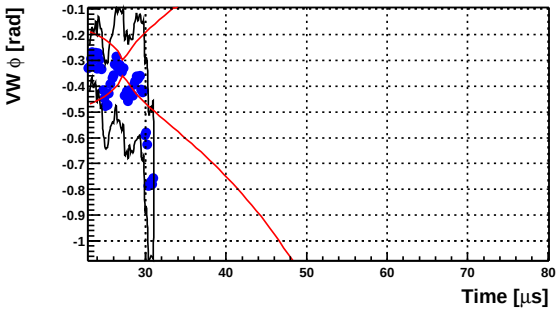


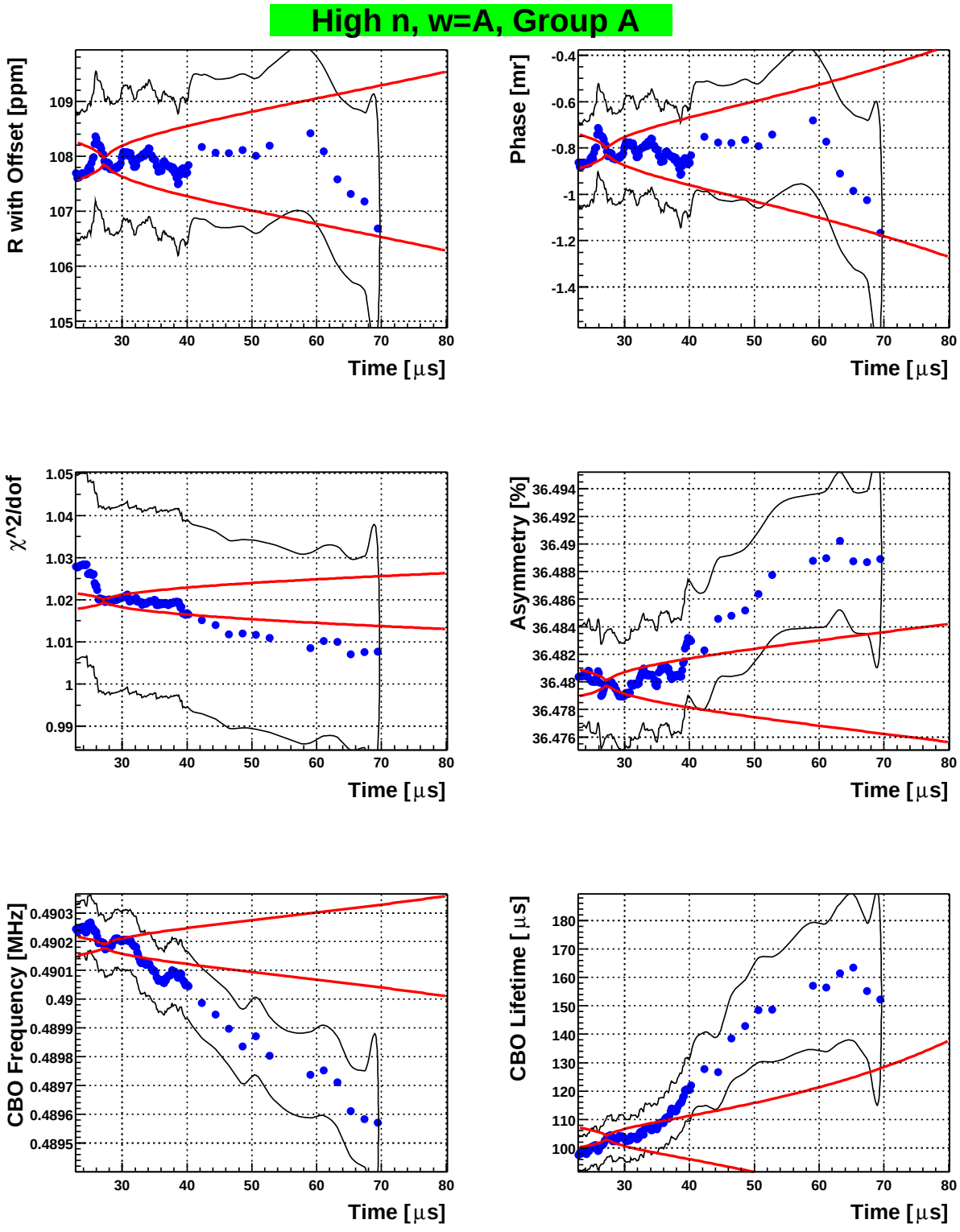


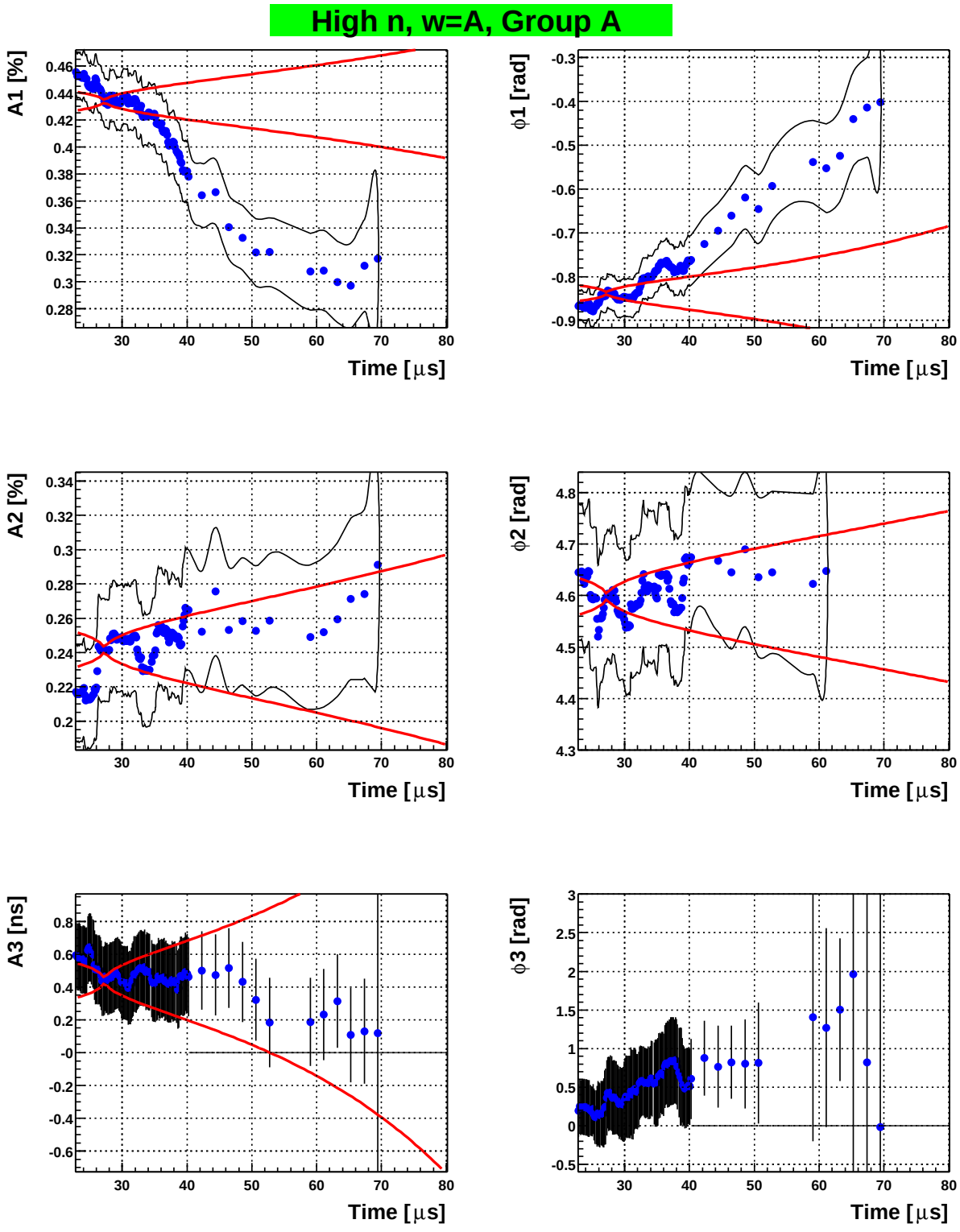




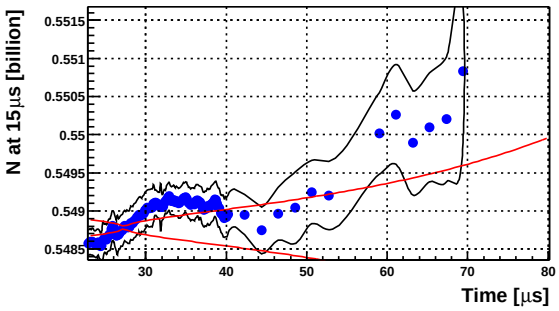


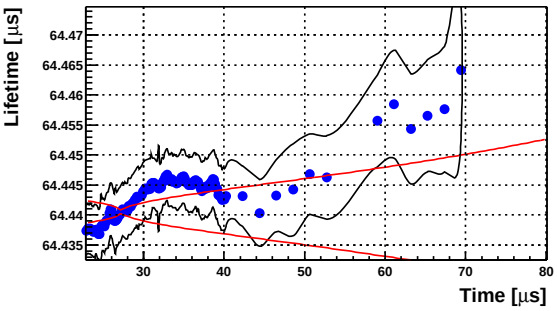


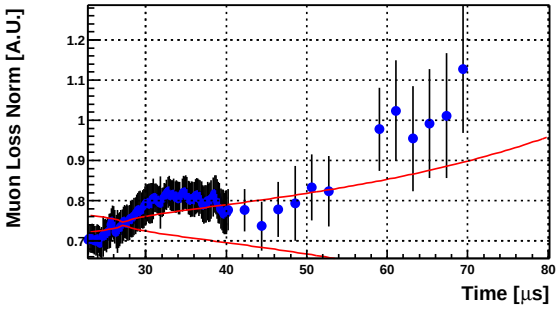

High n, w=A, Group A
R with Offset [ppm]
Phase [mr]
χ^2/dof
Asymmetry [%]
CBO Frequency [MHz]
CBO Lifetime [μs]
Time [μs]

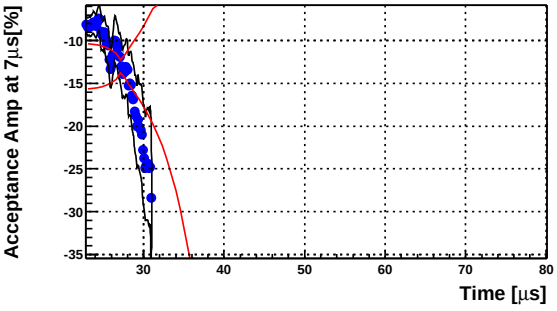

High n, w=A, Group A
A1 [%]
ϕ1 [rad]
A2 [%]
ϕ2 [rad]
A3 [ns]
ϕ3 [rad]
Time [μs]

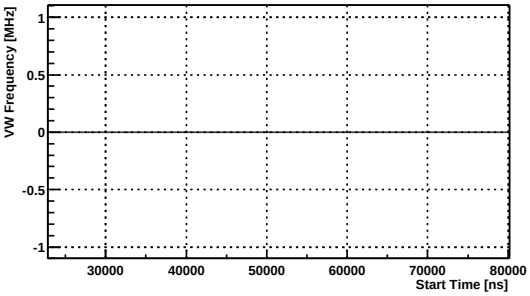
High n, w=A, Group A

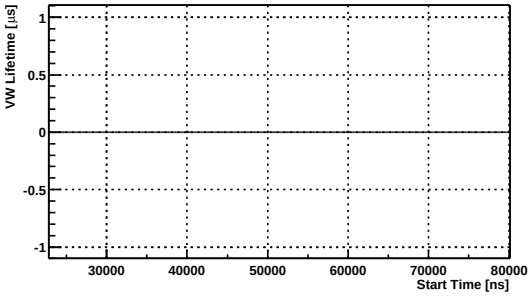


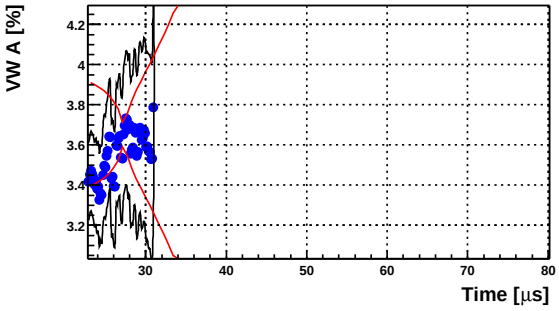


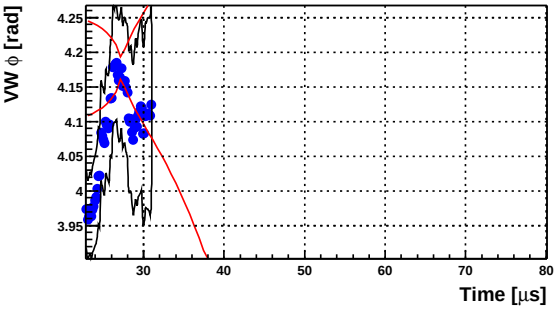


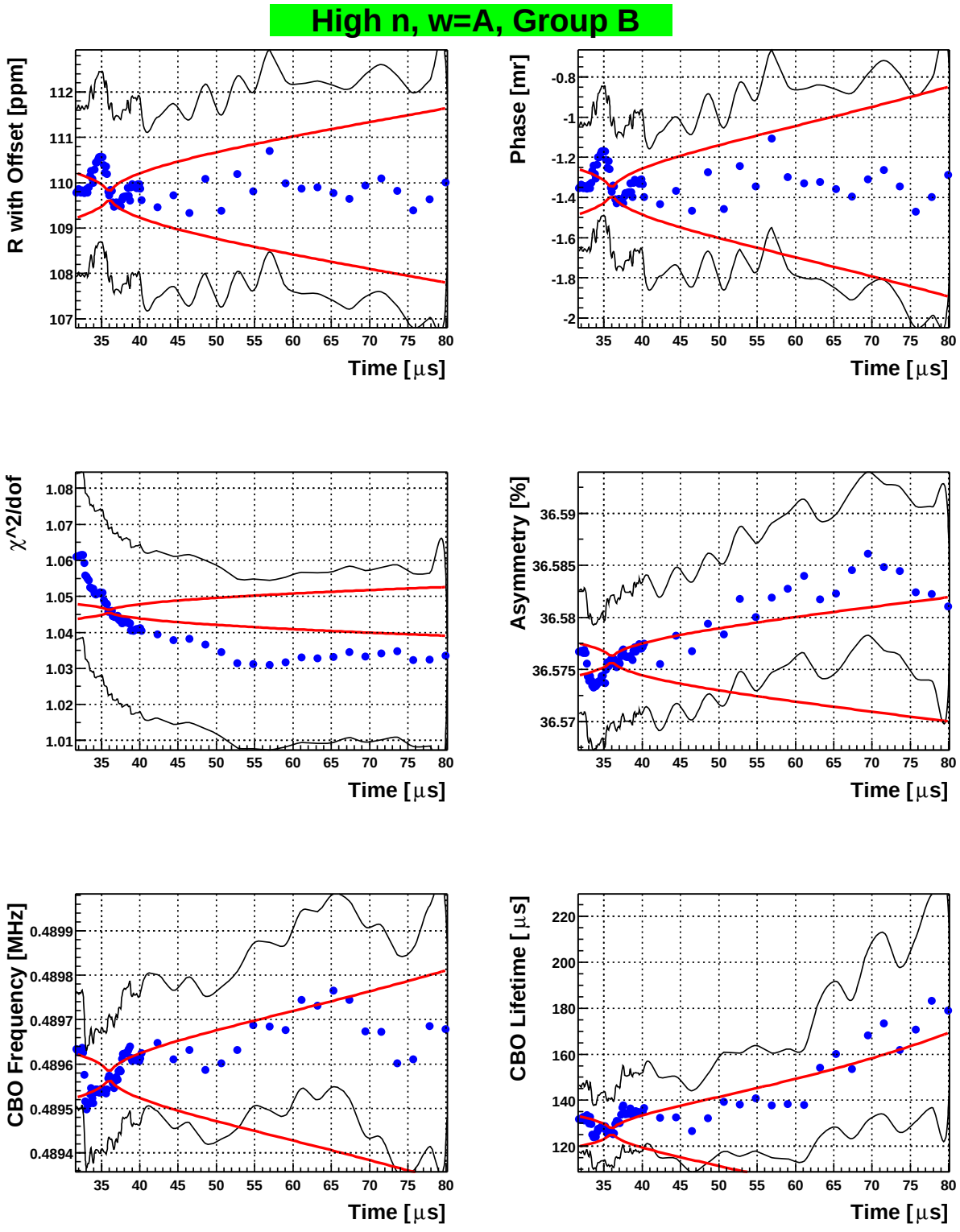


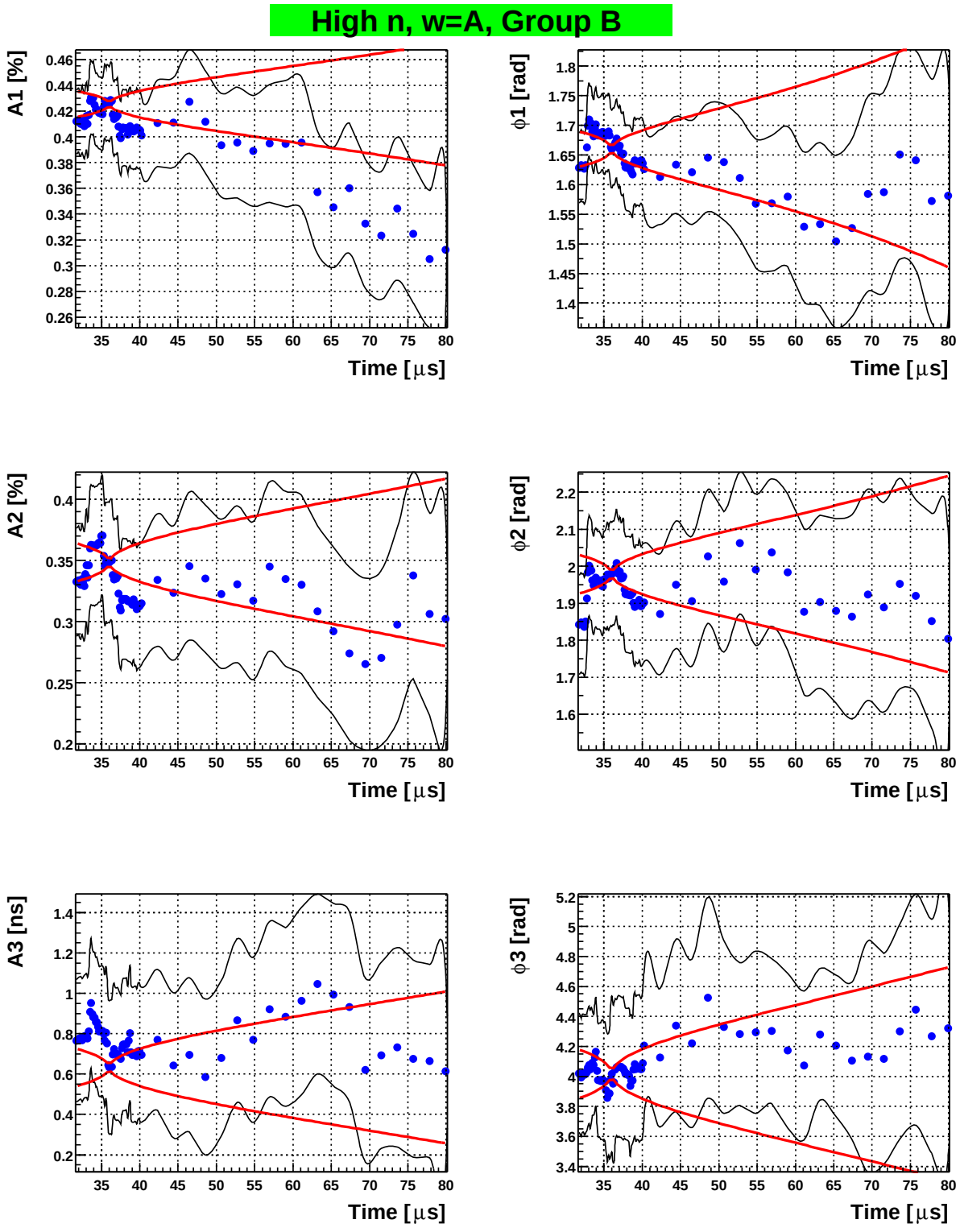




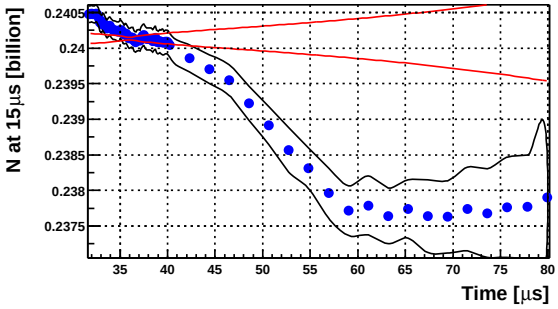


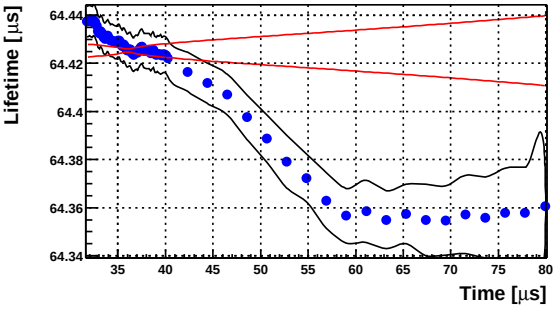


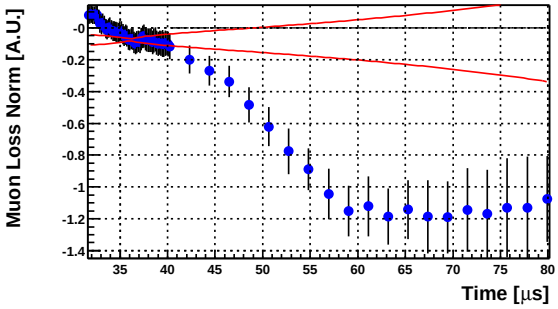

High n, w=A, Group B
R with Offset [ppm]
Phase [mr]
χ^2/dof
Asymmetry [%]
CBO Frequency [MHz]
CBO Lifetime [μs]
Time [μs]

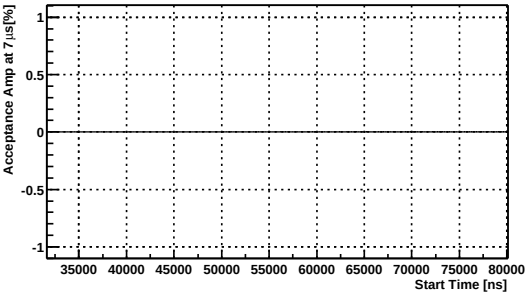

High n, w=A, Group B
A1 [%]
ϕ1 [rad]
A2 [%]
ϕ2 [rad]
A3 [ns]
ϕ3 [rad]
Time [μs]

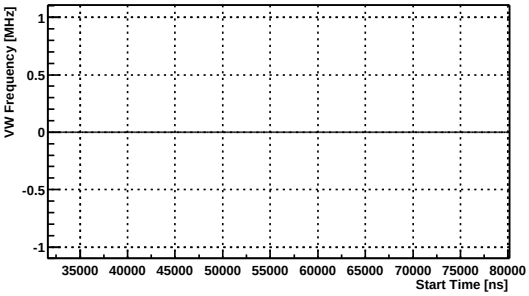
High n, w=A, Group B

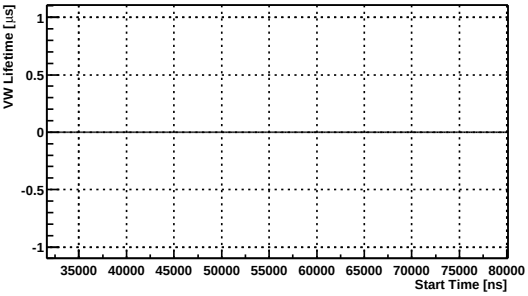


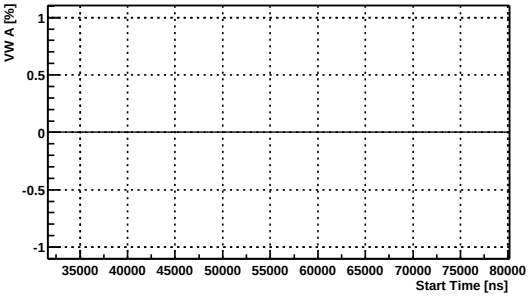


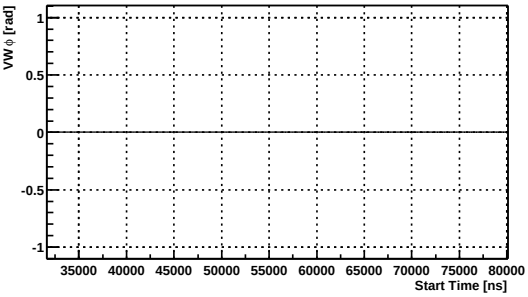




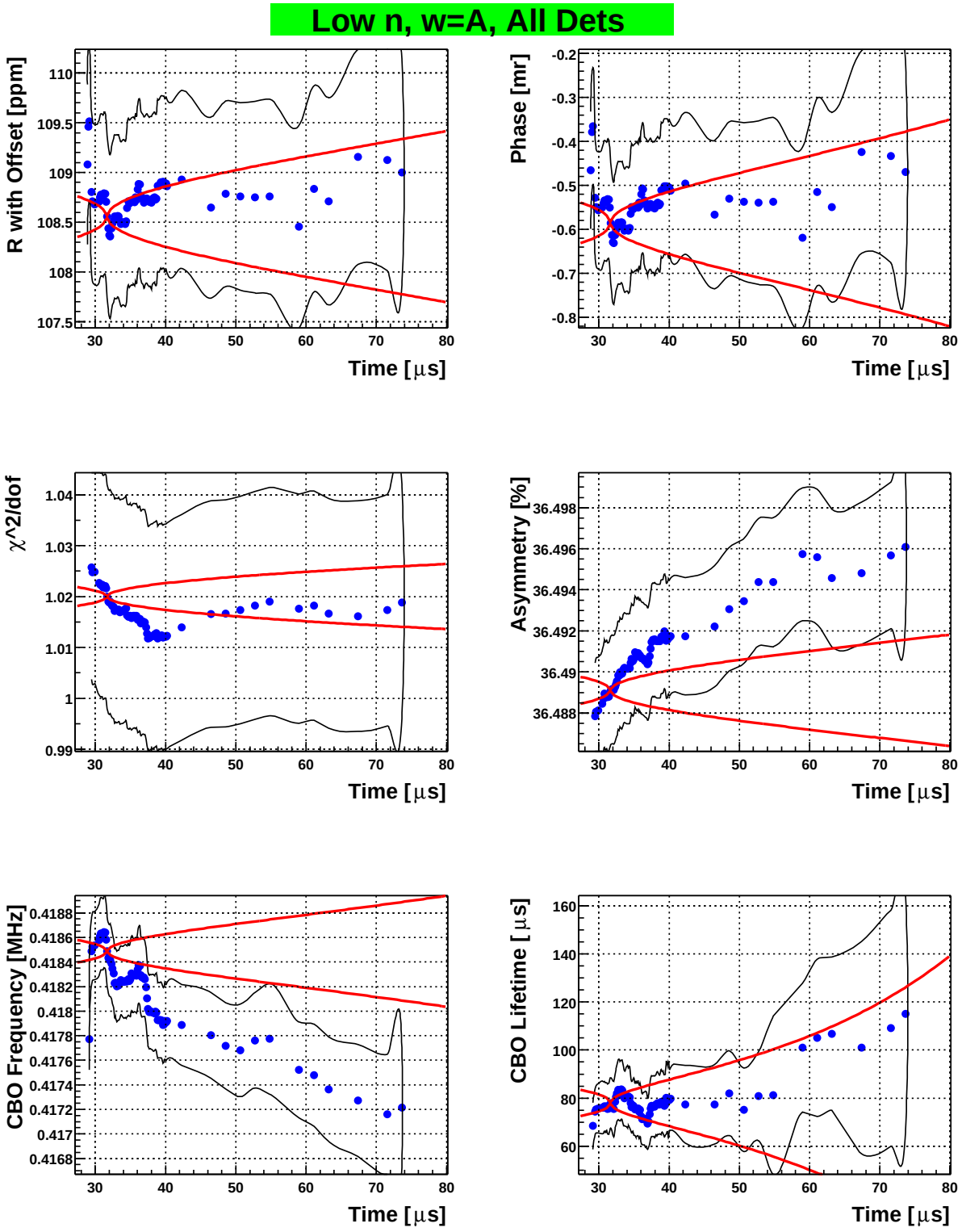


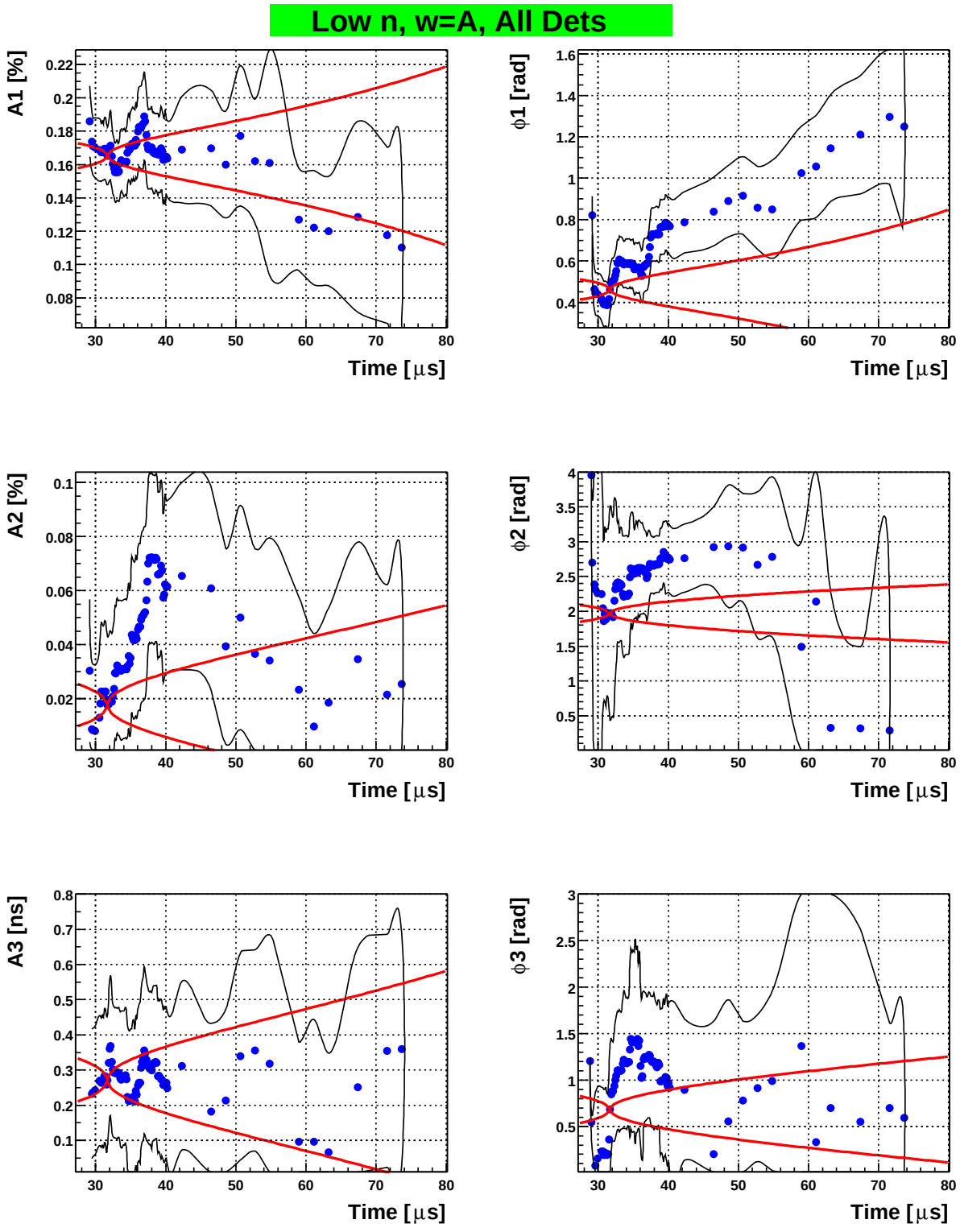


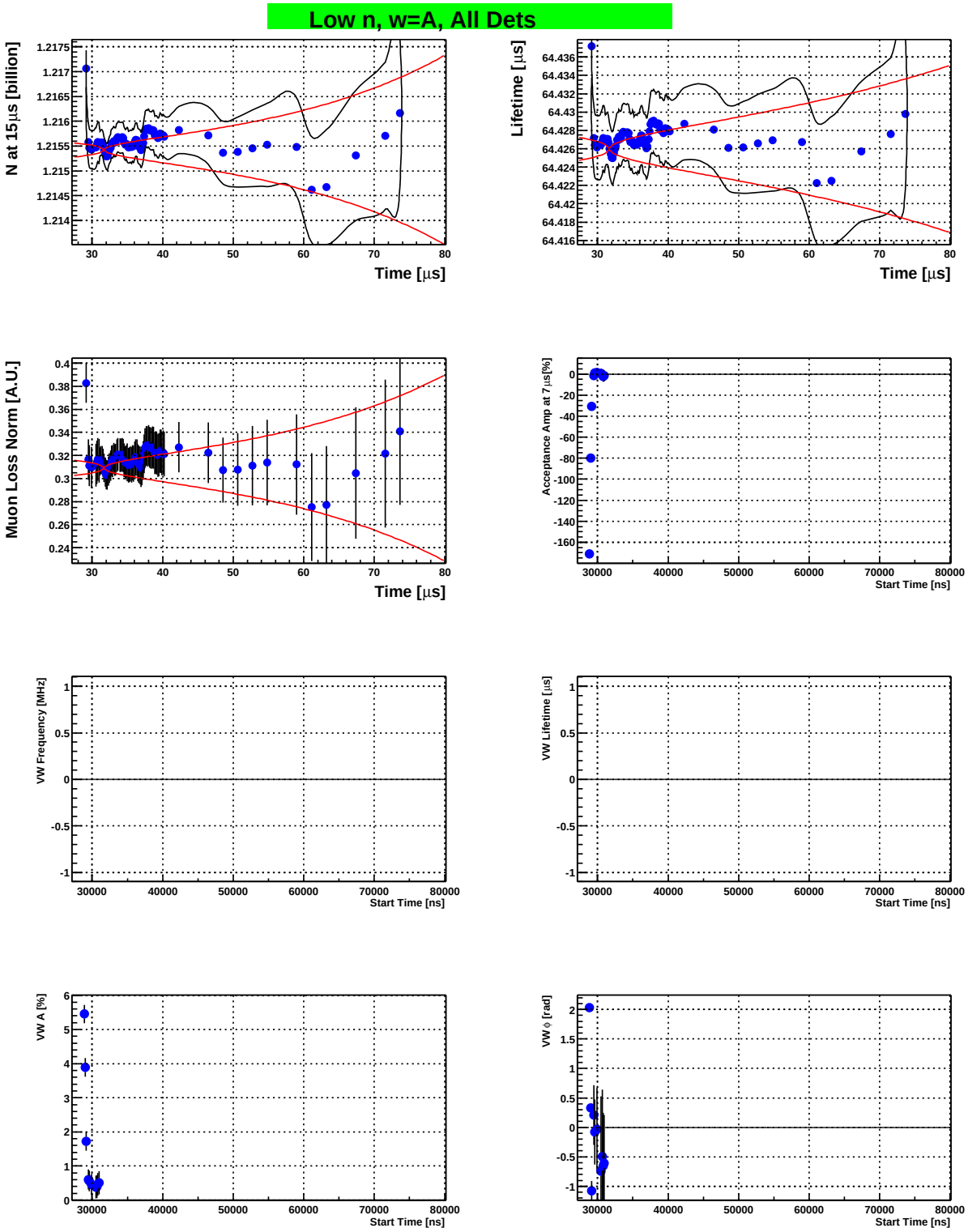


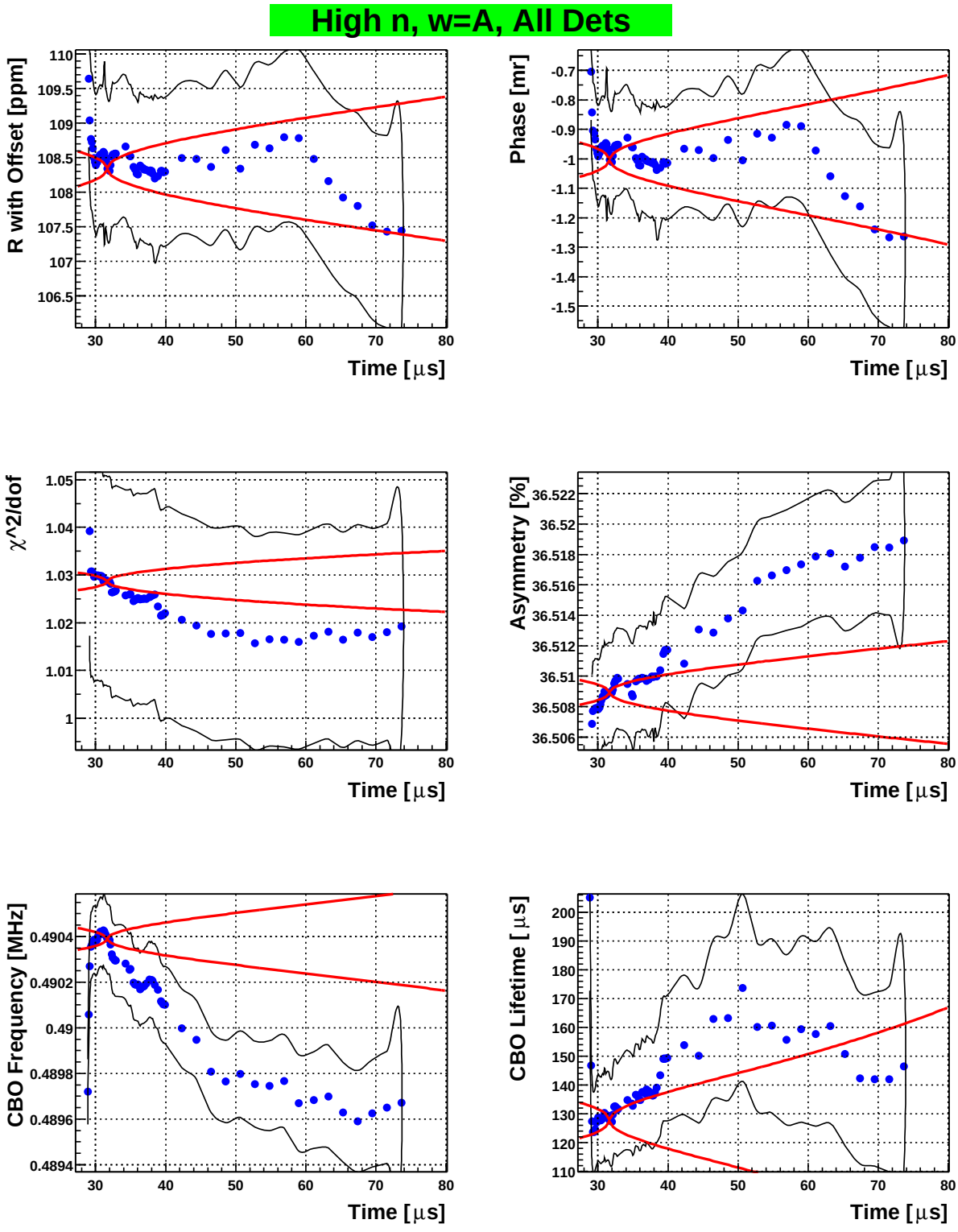


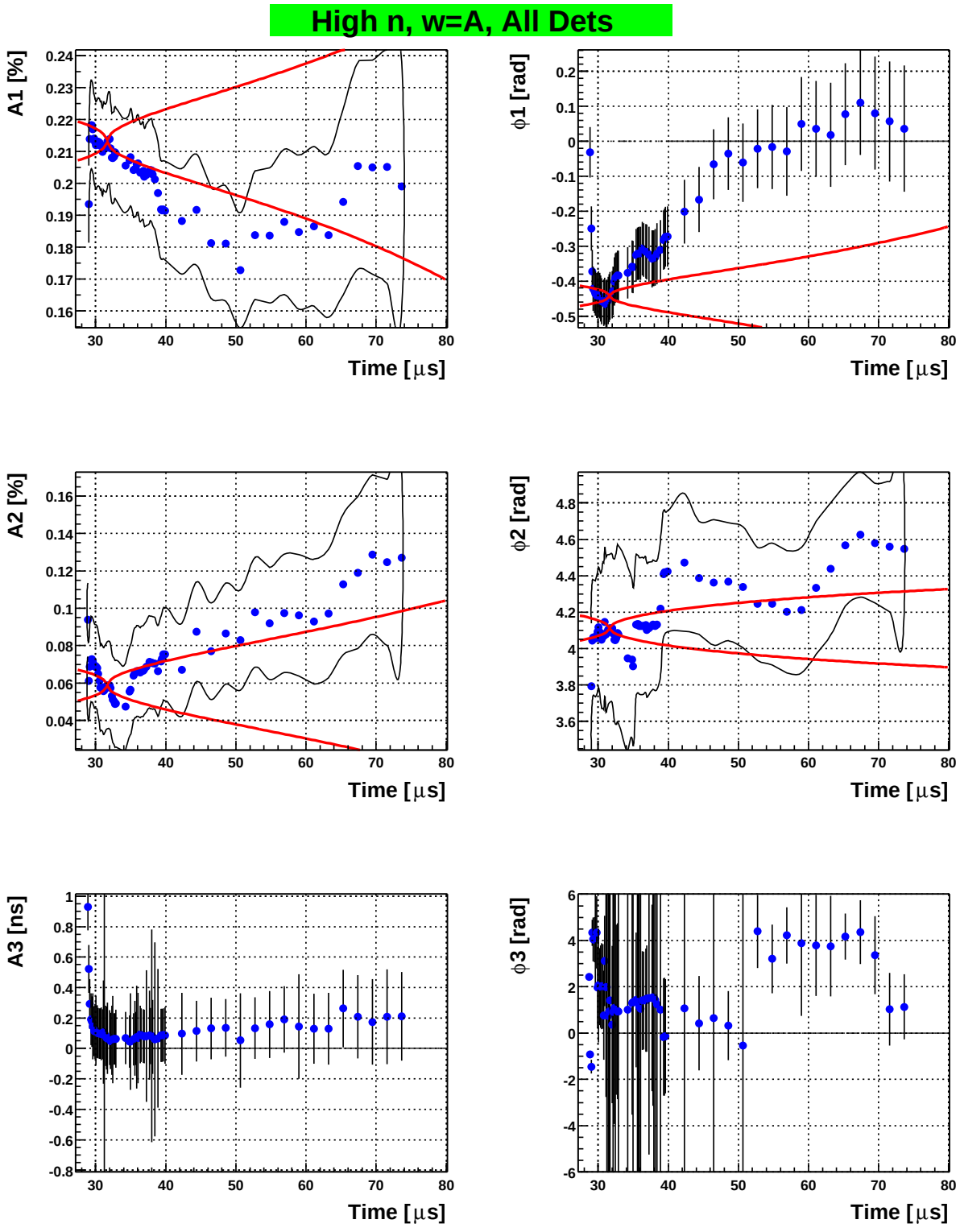
E.2 Asymmetry Weighting; All Detectors



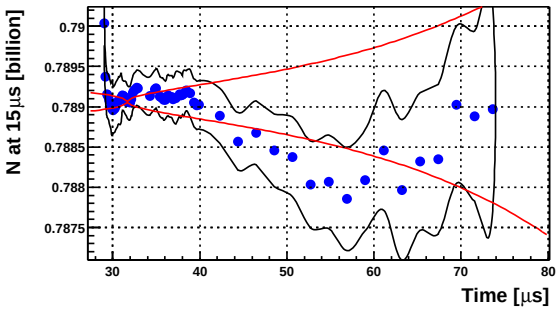

Low n, w=A, All Dets
A1 [%]
Time [μs]
ϕ1 [rad]
Time [μs]
A2 [%]
Time [μs]
ϕ2 [rad]
Time [μs]
A3 [ns]
Time [μs]
ϕ3 [rad]
Time [μs]

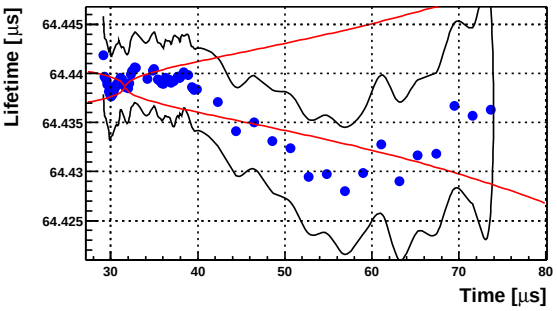

Low n, w=A, All Dets
N at 15μs [billion]
Time [μs]
Lifetime [μs]
Time [μs]
Muon Loss Norm [A.U.]
Time [μs]
Acceptance Amp at 7 μs[%]
Start Time [ns]
VW Frequency [MHz]
Start Time [ns]
VW Lifetime [μs]
Start Time [ns]
VW A [%]
Start Time [ns]
VW ϕ [rad]
Start Time [ns]

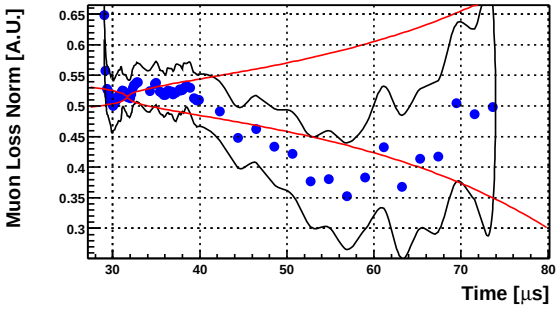

High n, w=A, All Dets
R with Offset [ppm]
Phase [mr]
χ^2/dof
Asymmetry [%]
CBO Frequency [MHz]
CBO Lifetime [μs]
Time [μs]

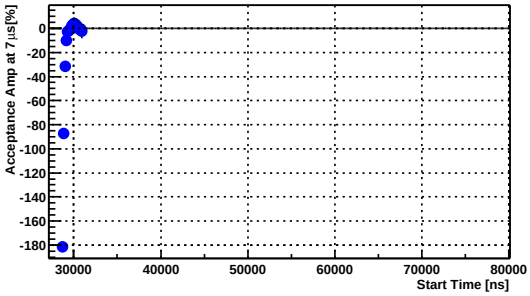

High n, w=A, All Dets
A1 [%]
Time [μs]
ϕ1 [rad]
Time [μs]
A2 [%]
Time [μs]
ϕ2 [rad]
Time [μs]
A3 [ns]
Time [μs]
ϕ3 [rad]
Time [μs]

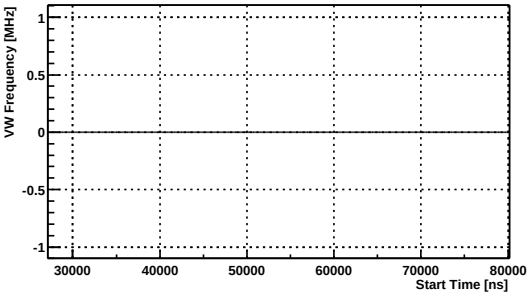
High n, w=A, All Dets

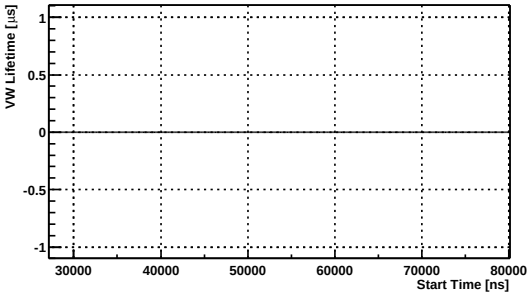


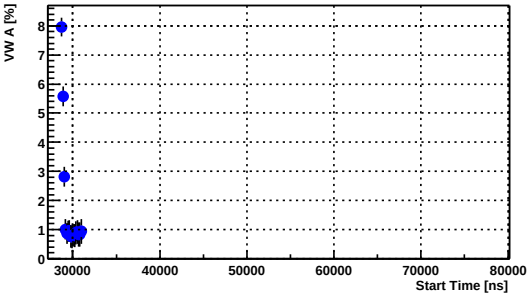


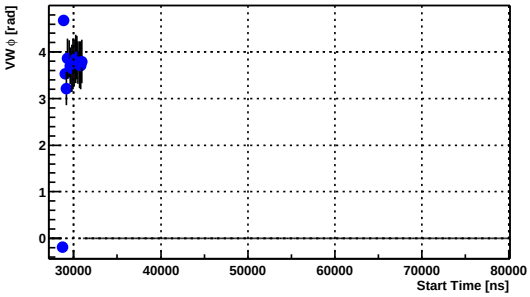




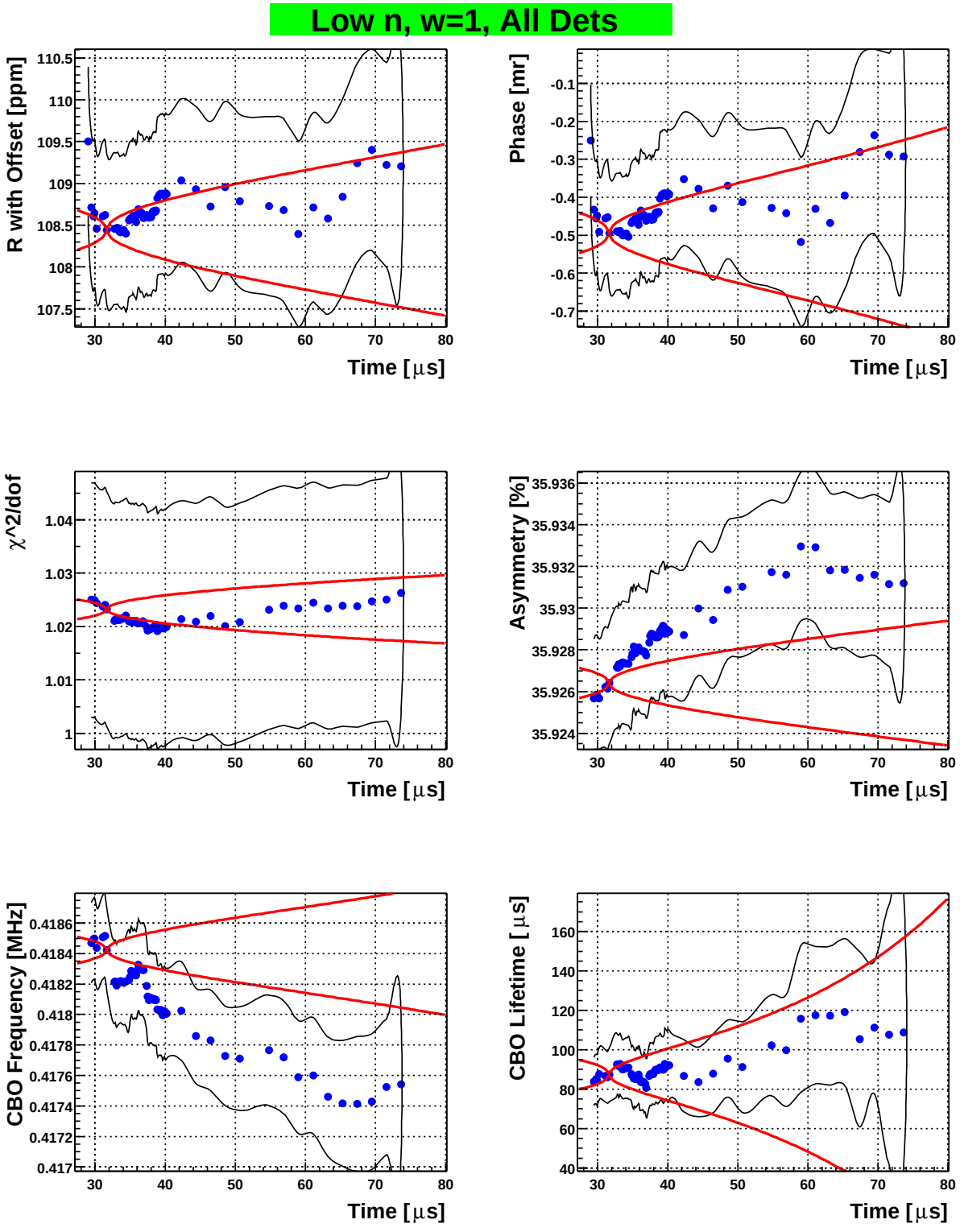


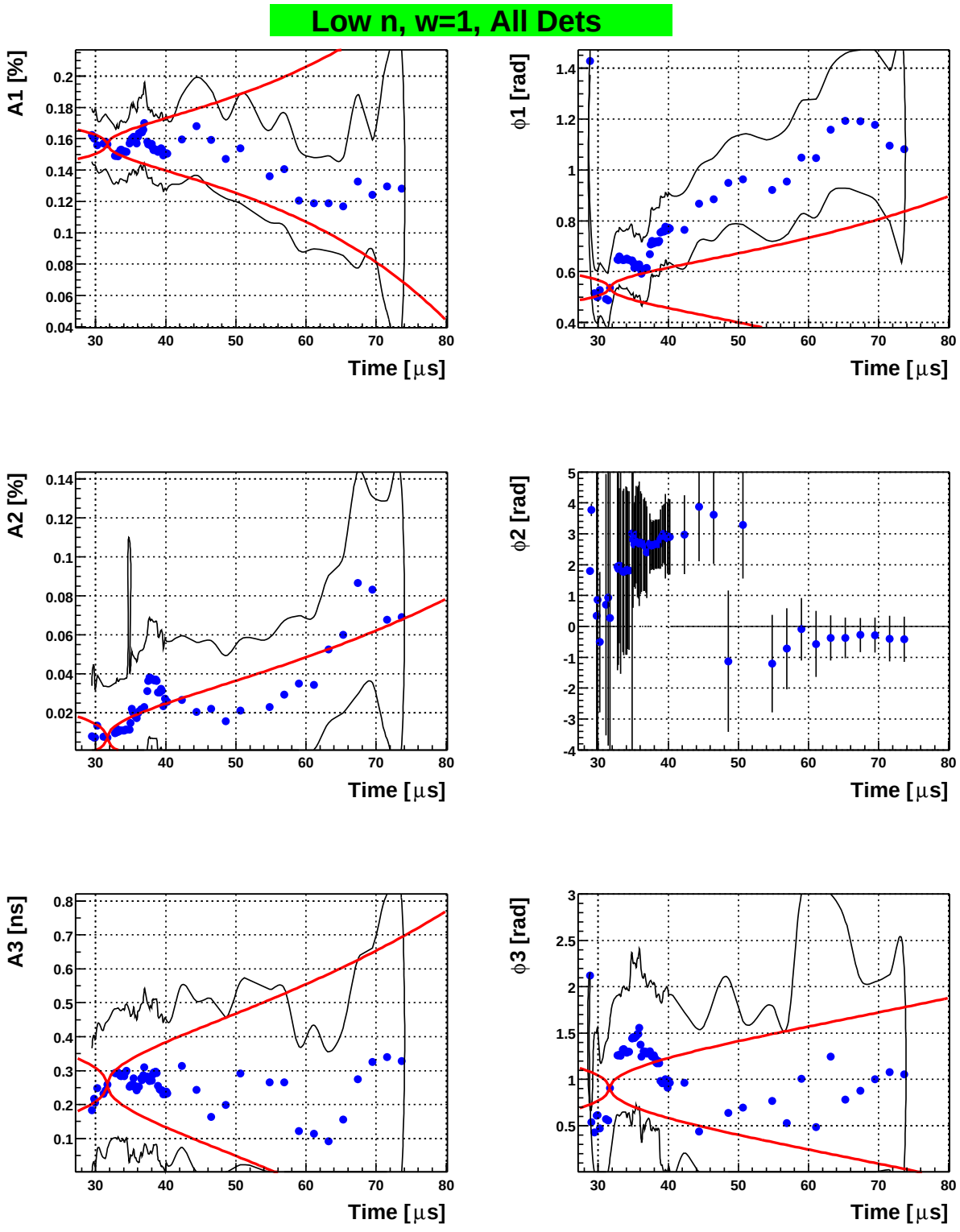


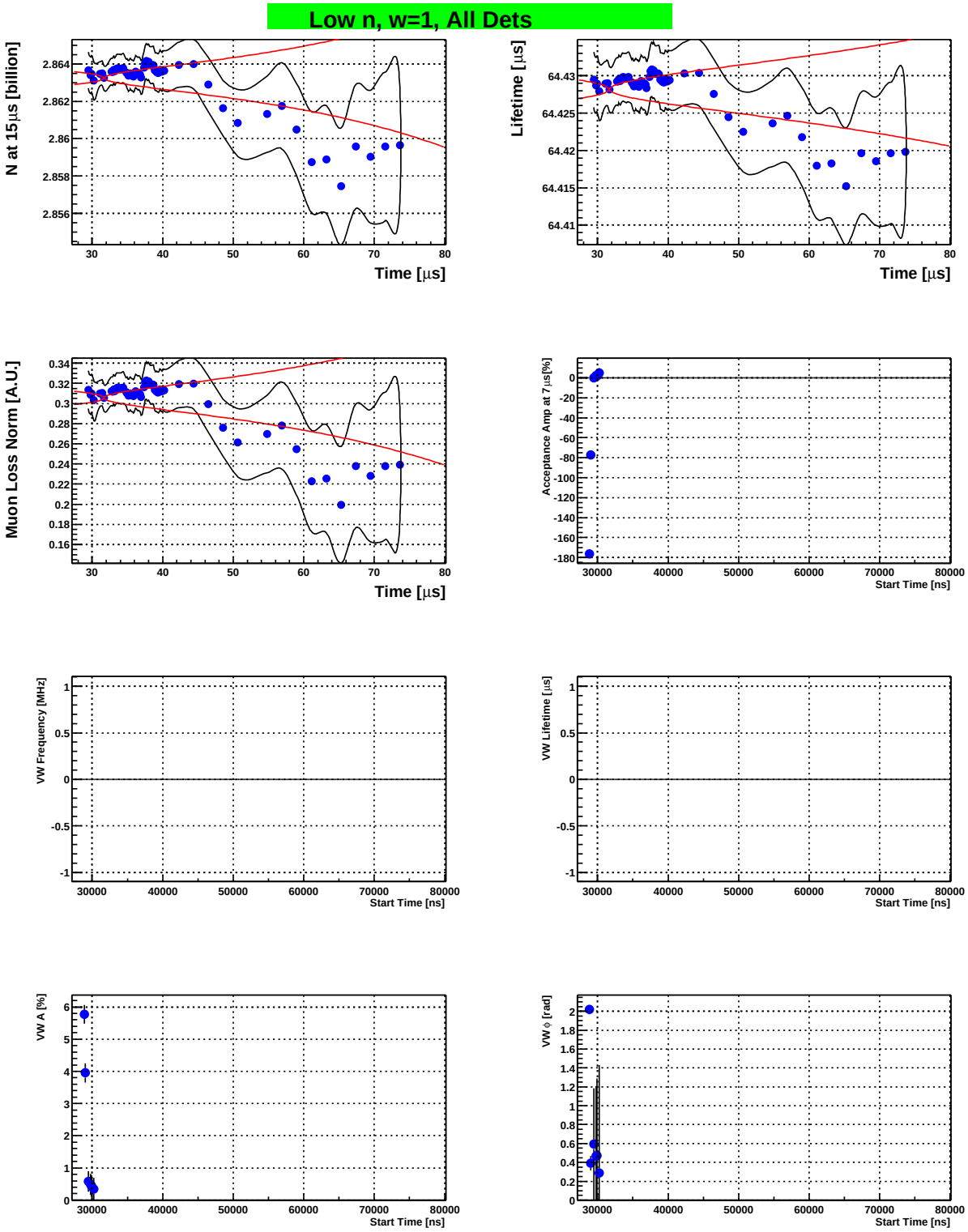


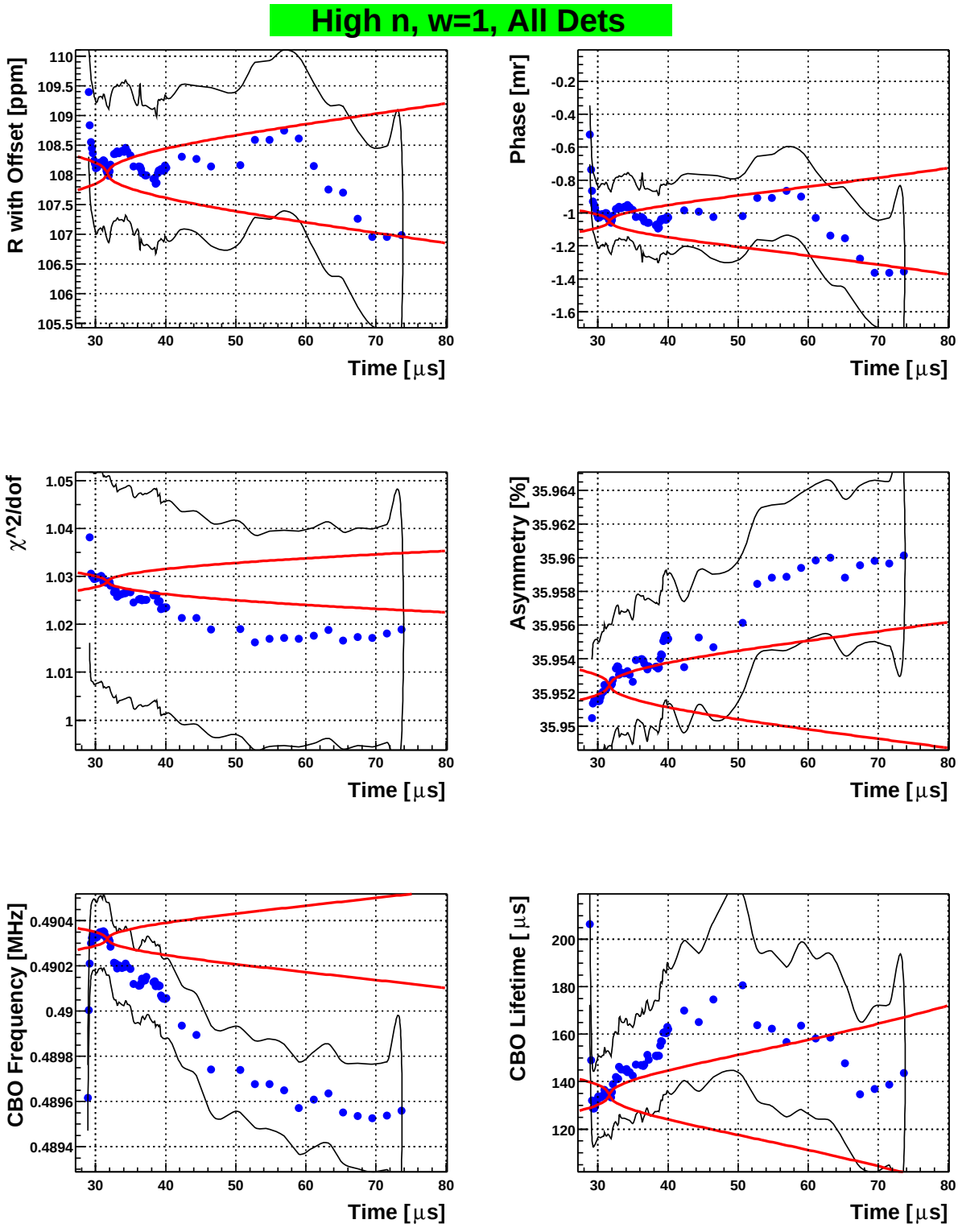


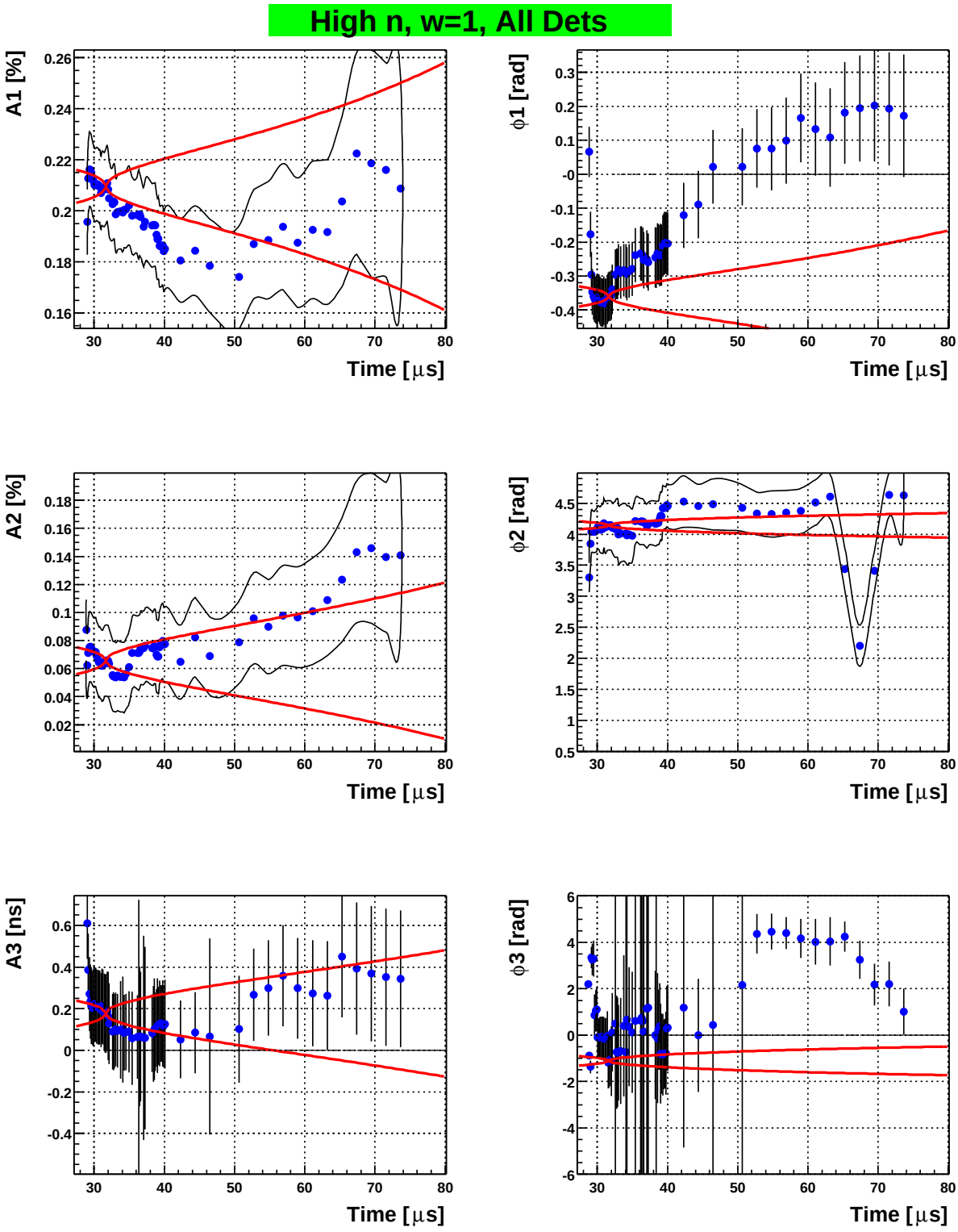
E.3 Uniform Weighting; All Detectors

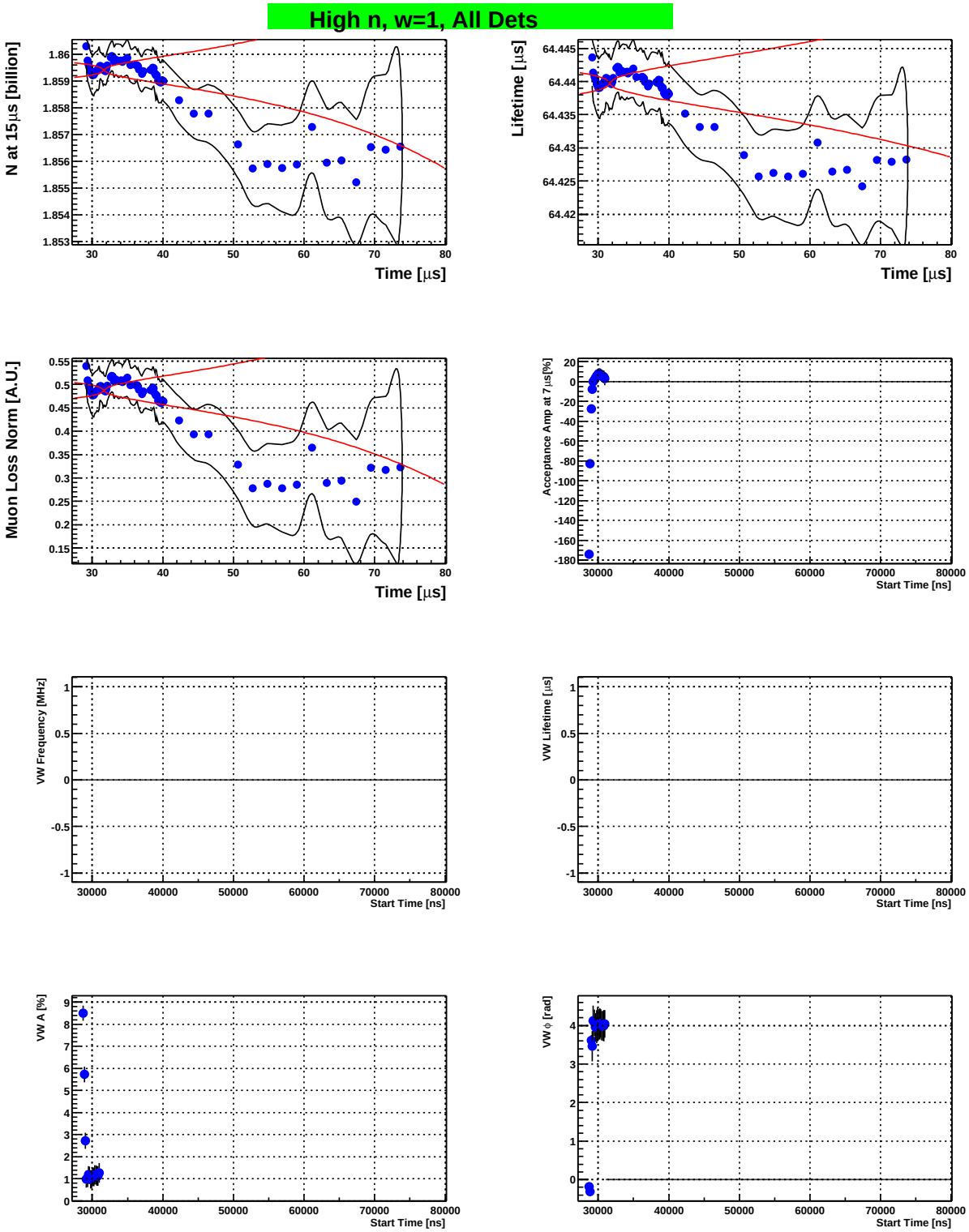



Low n, w=1, All Dets
A1 [%]
Time [μs]
ϕ1 [rad]
Time [μs]
A2 [%]
Time [μs]
ϕ2 [rad]
Time [μs]
A3 [ns]
Time [μs]
ϕ3 [rad]
Time [μs]


Low n, w=1, All Dets
N at 15μs [billion]
Time [μs]
Lifetime [μs]
Time [μs]
Muon Loss Norm [A.U.]
Time [μs]
Acceptance Amp at 7 μs[%]
Start Time [ns]
VW Frequency [MHz]
Start Time [ns]
VW Lifetime [μs]
Start Time [ns]
VW A [%]
Start Time [ns]
VW ϕ [rad]
Start Time [ns]


High n, w=1, All Dets
R with Offset [ppm]
Phase [mr]
χ^2/dof
Asymmetry [%]
CBO Frequency [MHz]
CBO Lifetime [μs]
Time [μs]


High n, w=1, All Dets
A1 [%]
Time [μs]
ϕ1 [rad]
Time [μs]
A2 [%]
Time [μs]
ϕ2 [rad]
Time [μs]
A3 [ns]
Time [μs]
ϕ3 [rad]
Time [μs]

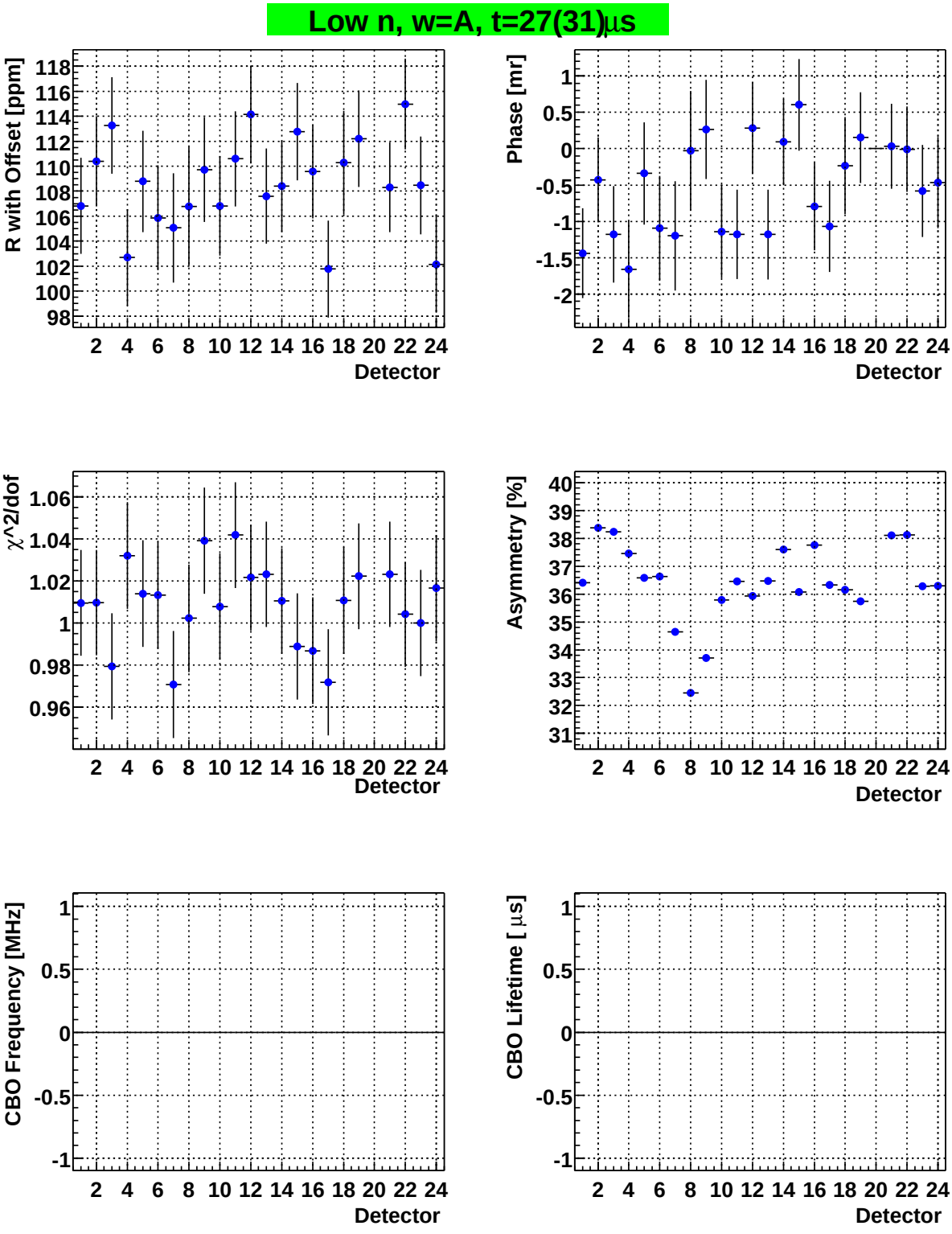

High n, w=1, All Dets
N at 15μs [billion]
Time [μs]
Lifetime [μs]
Time [μs]
Muon Loss Norm [A.U.]
Time [μs]
Acceptance Amp at 7μs[%]
Start Time [ns]
VW Frequency [MHz]
Start Time [ns]
VW Lifetime [μs]
Start Time [ns]
VW A [%]
Start Time [ns]
VW ϕ [rad]
Start Time [ns]

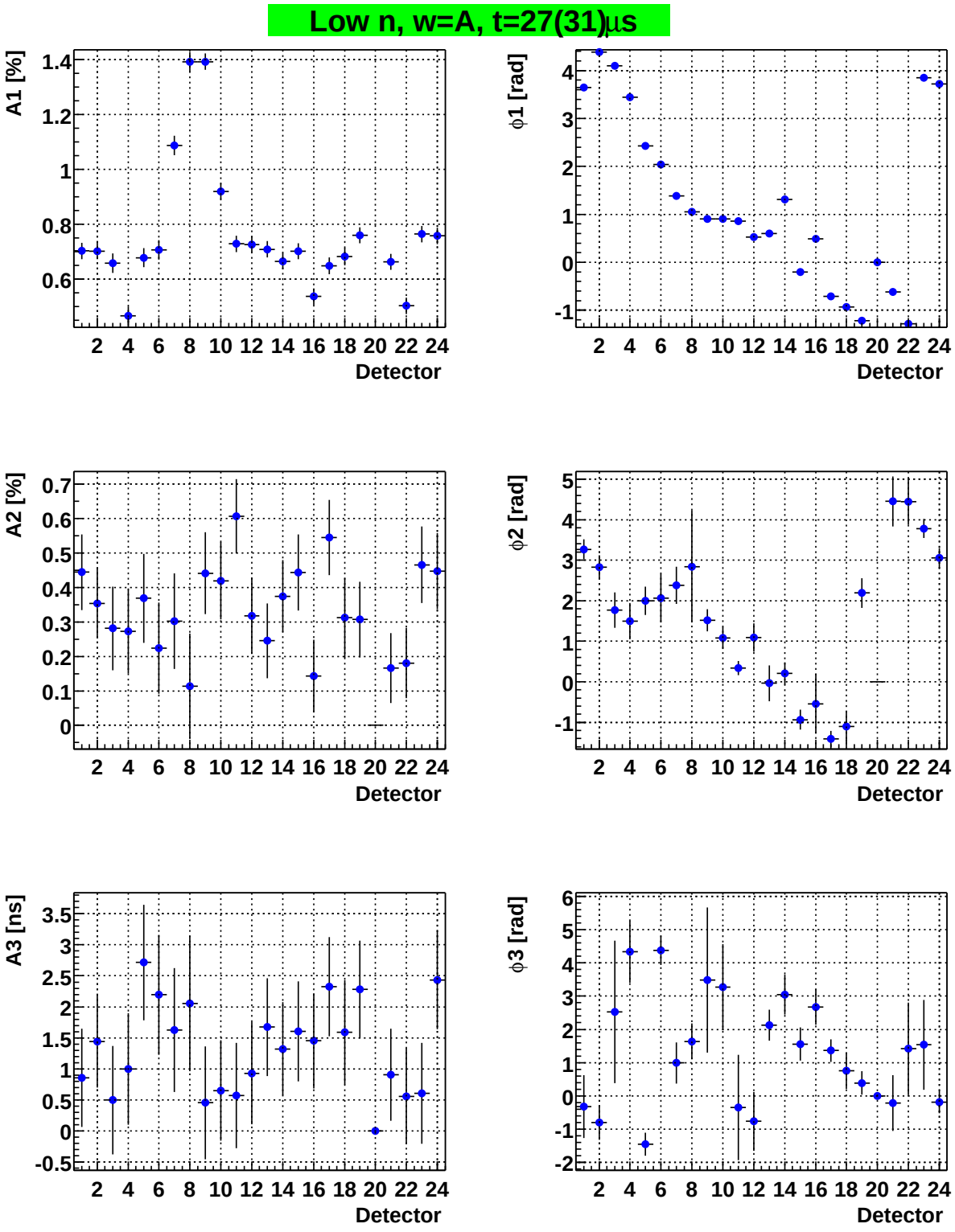
F Plots By Detector

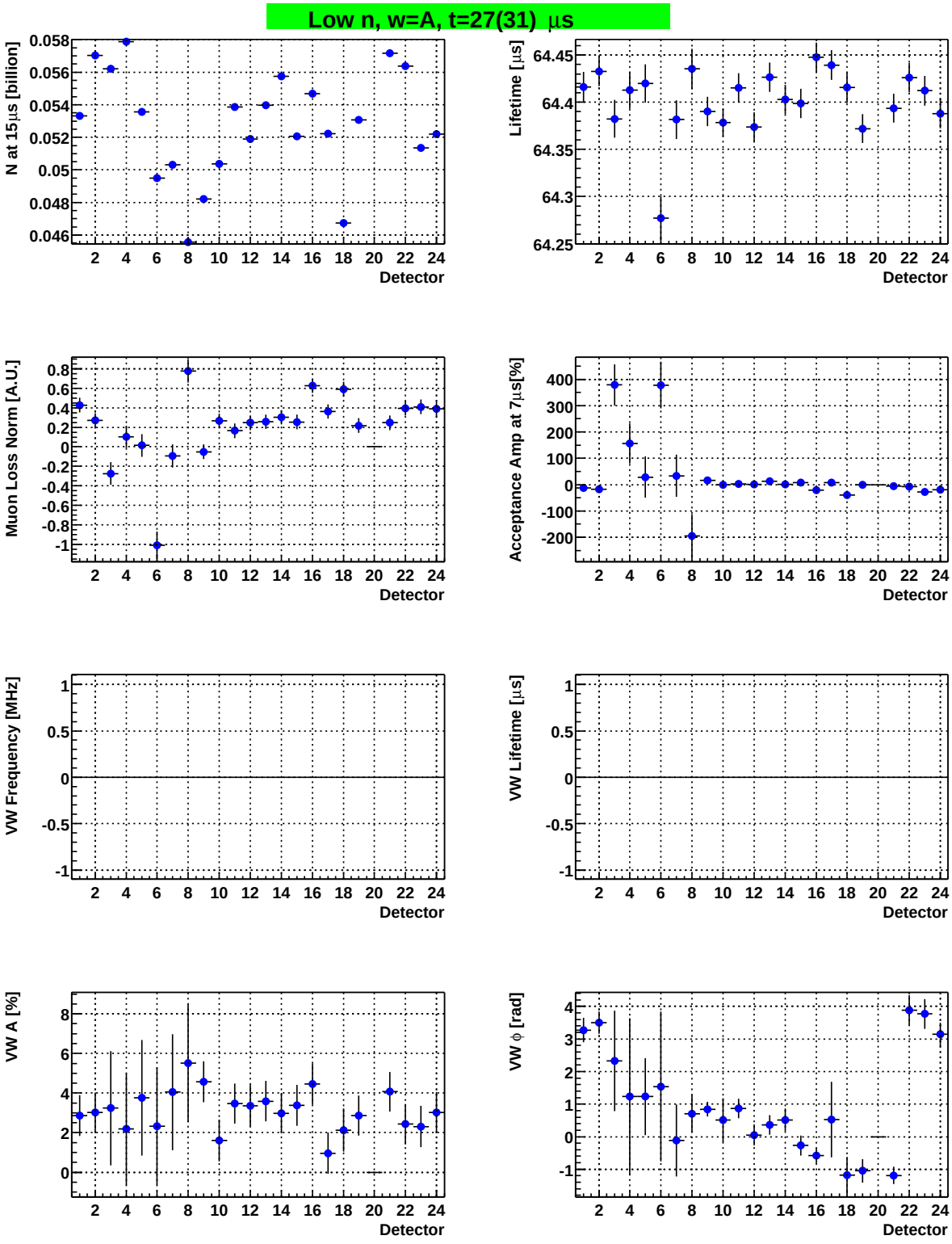
All fit parameters are plotted from separate fits to each detector.

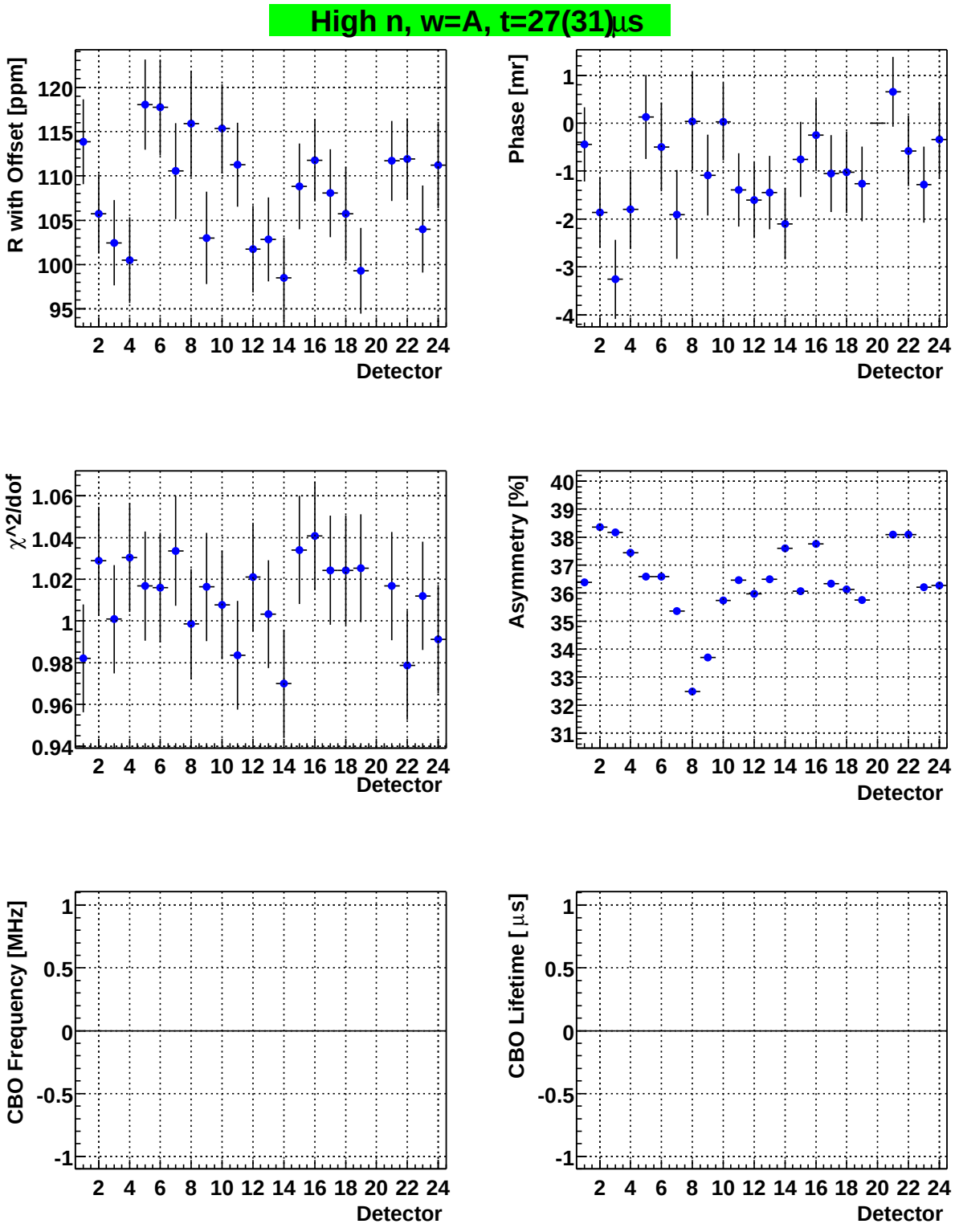
F.1 Asymmetry Weighting; Start Times at 27 and 31 μs

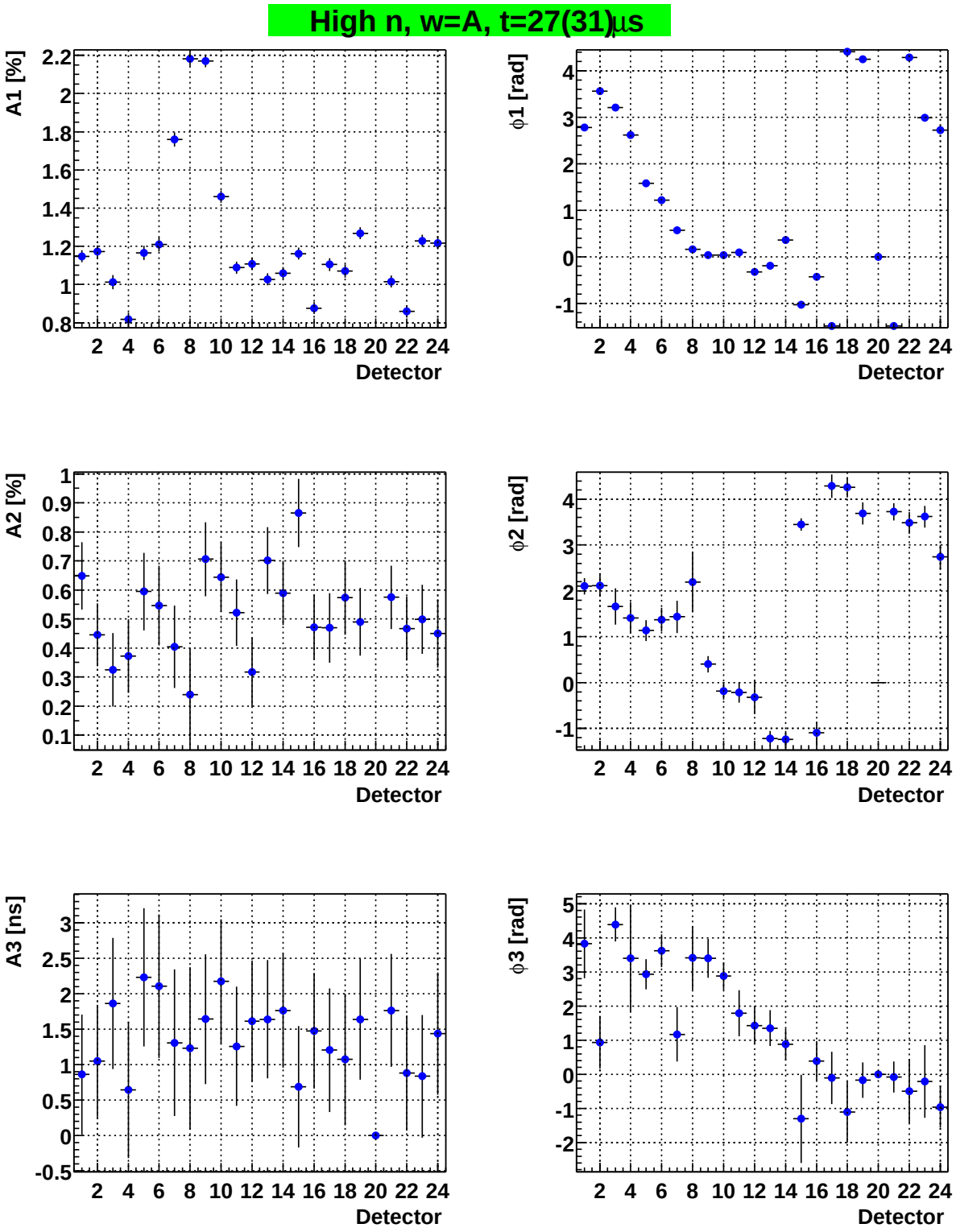
Detectors 3-8 start at 31 μs all others at 27 μs.

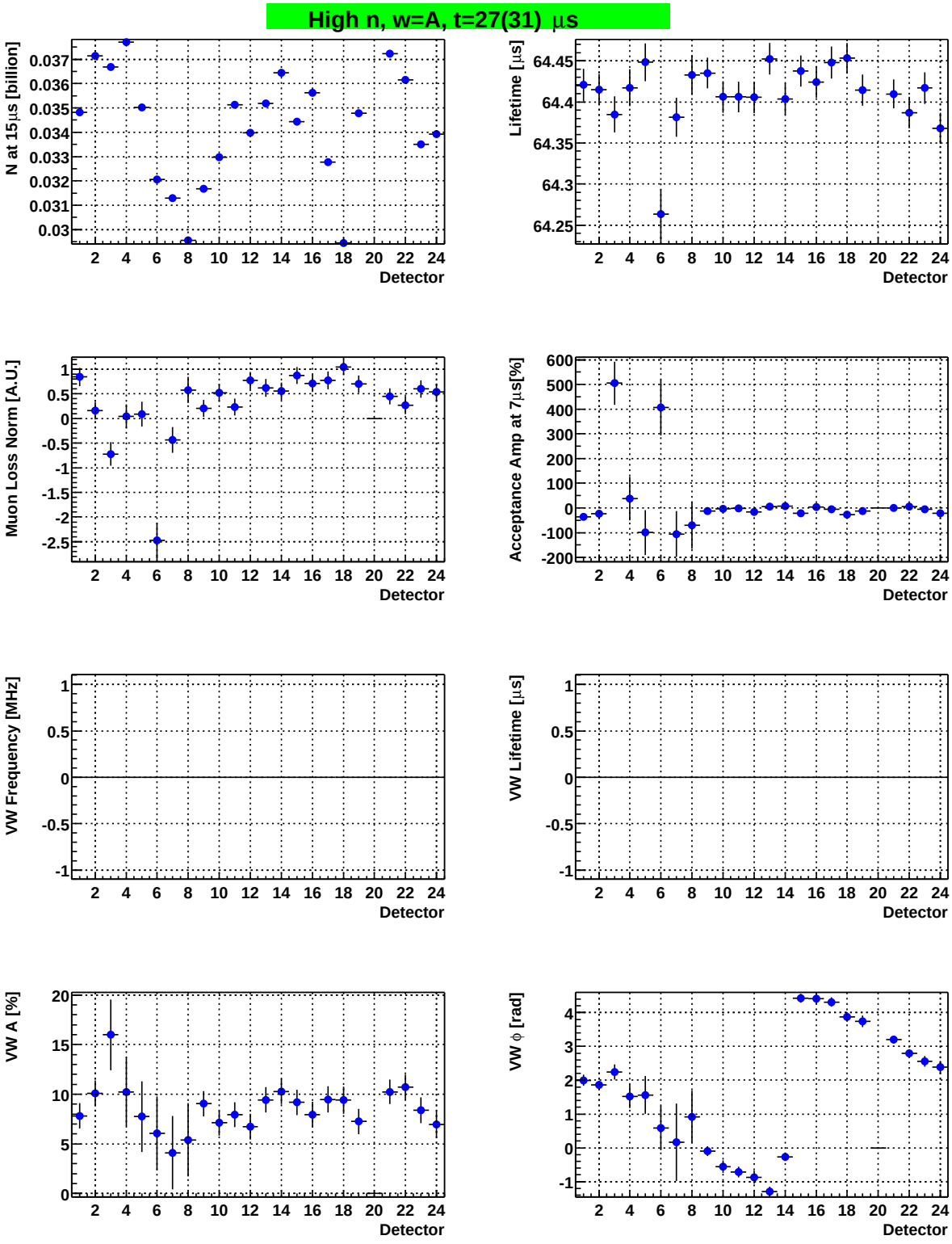



Low n, w=A, t=27(31)μs
A1 [%]
φ1 [rad]
A2 [%]
φ2 [rad]
A3 [ns]
φ3 [rad]
Detector

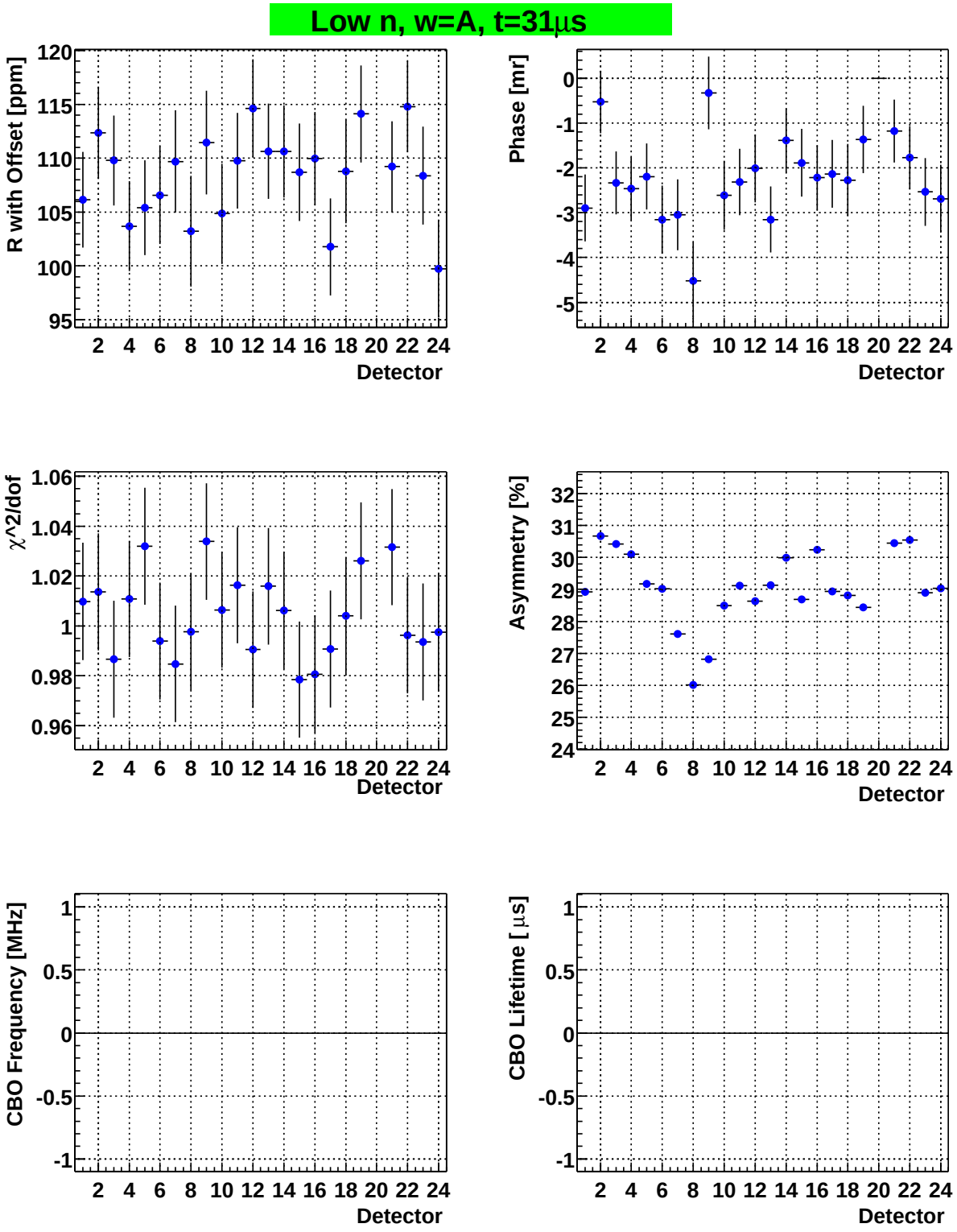

Low n, w=A, t=27(31) μs
N at 15μs [billion]
Lifetime [μs]
Muon Loss Norm [A.U.]
Acceptance Amp at 7μs[%]
VW Frequency [MHz]
VW Lifetime [μs]
VW A [%]
VW ϕ [rad]
Detector

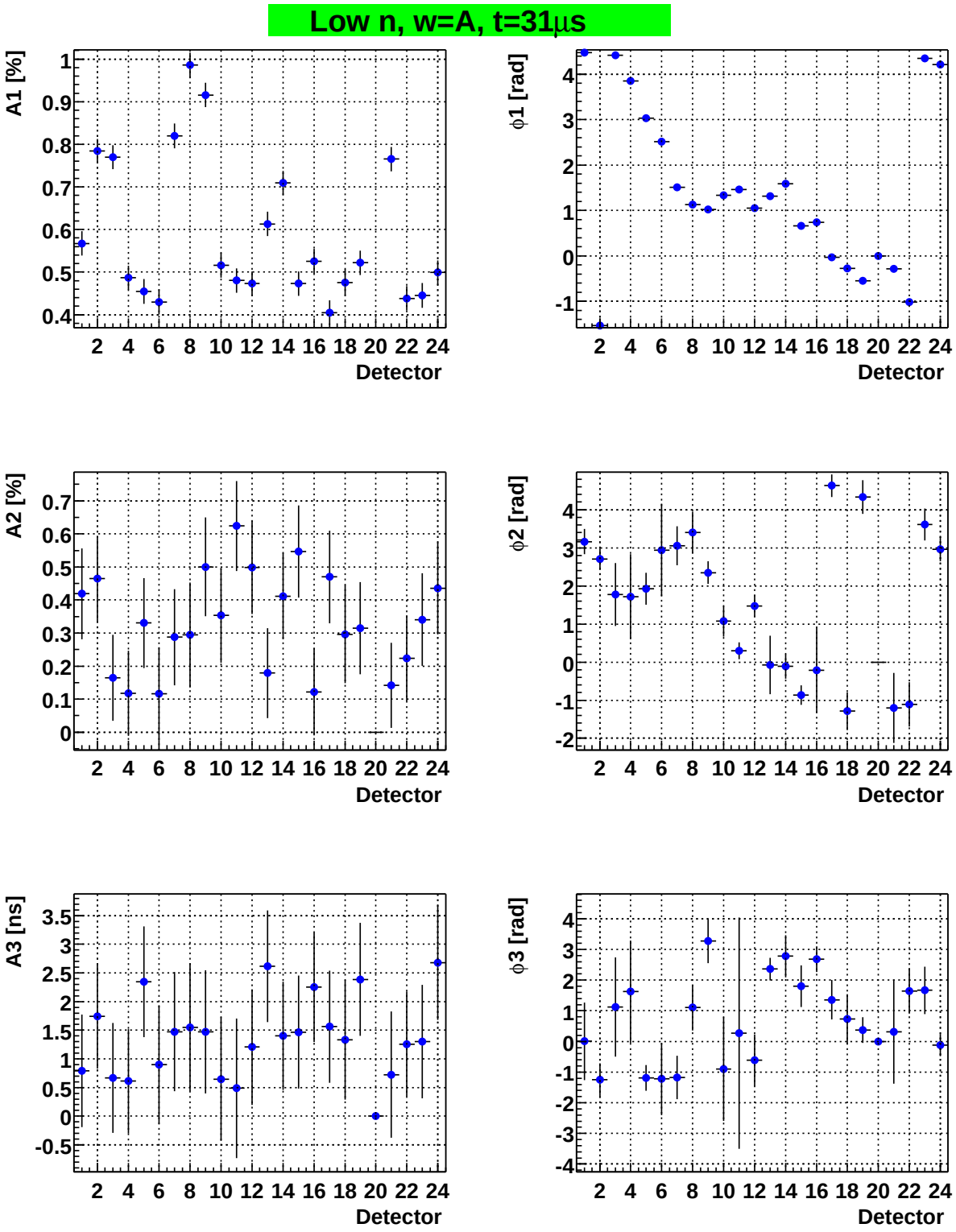

High n, w=A, t=27(31)μs
R with Offset [ppm]
Phase [mr]
χ^2/dof
Asymmetry [%]
CBO Frequency [MHz]
CBO Lifetime [μs]
Detector

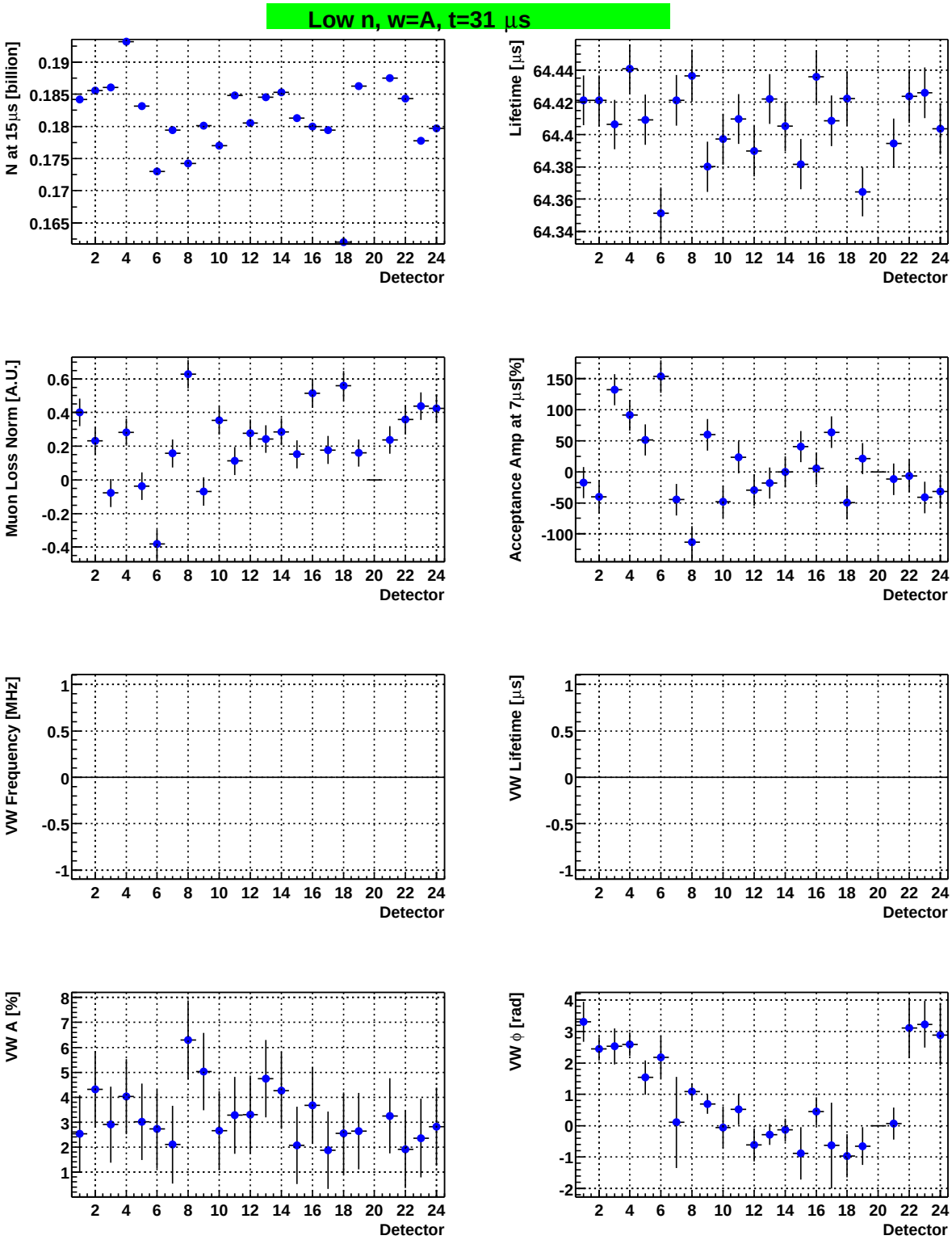

High n, w=A, t=27(31)μs
A1 [%]
Detector
ϕ1 [rad]
Detector
A2 [%]
Detector
ϕ2 [rad]
Detector
A3 [ns]
Detector
ϕ3 [rad]
Detector

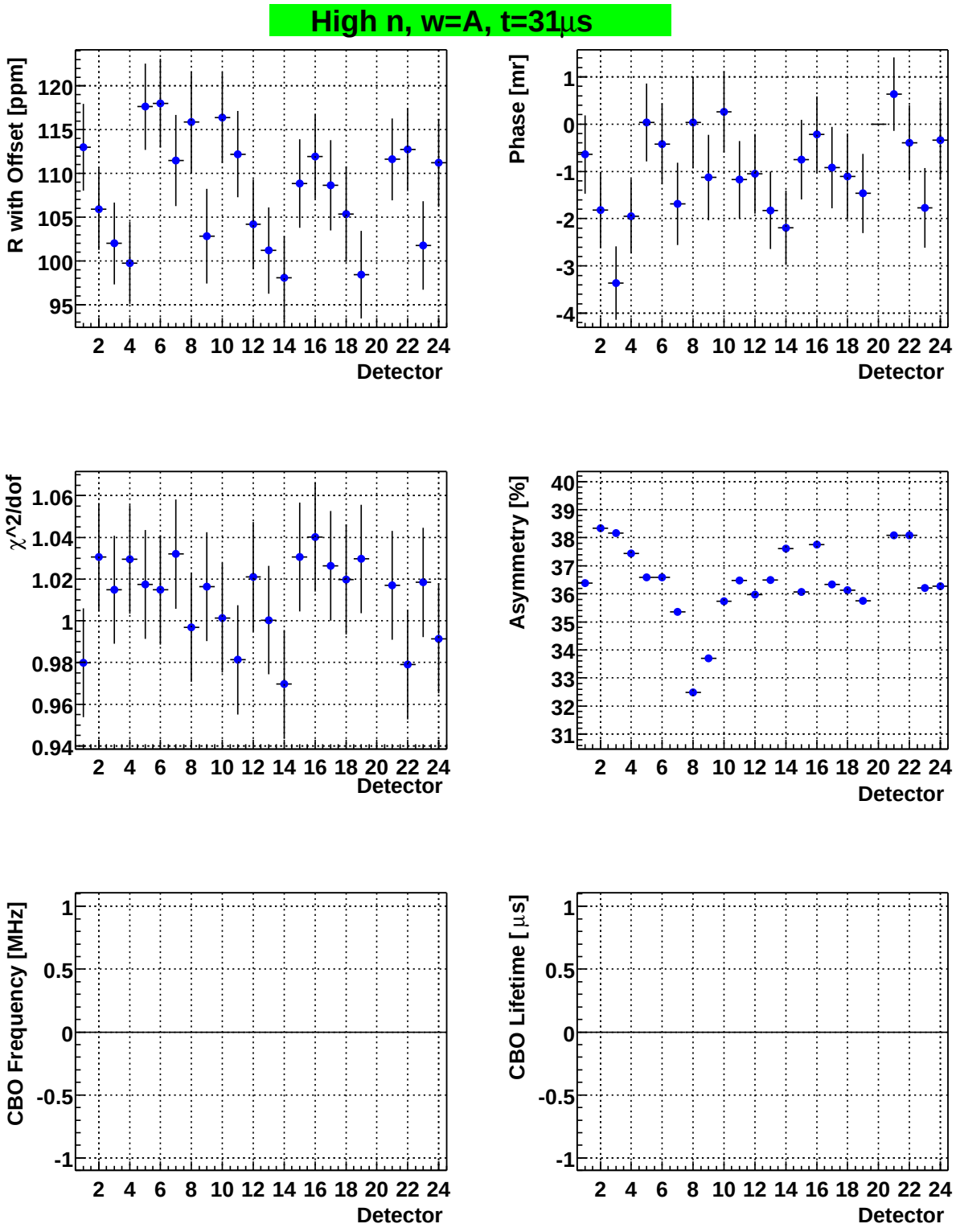

High n, w=A, t=27(31) μs
N at 15μs [billion]
Lifetime [μs]
Muon Loss Norm [A.U.]
Acceptance Amp at 7μs[%]
VW Frequency [MHz]
VW Lifetime [μs]
VW A [%]
VW ϕ [rad]
Detector

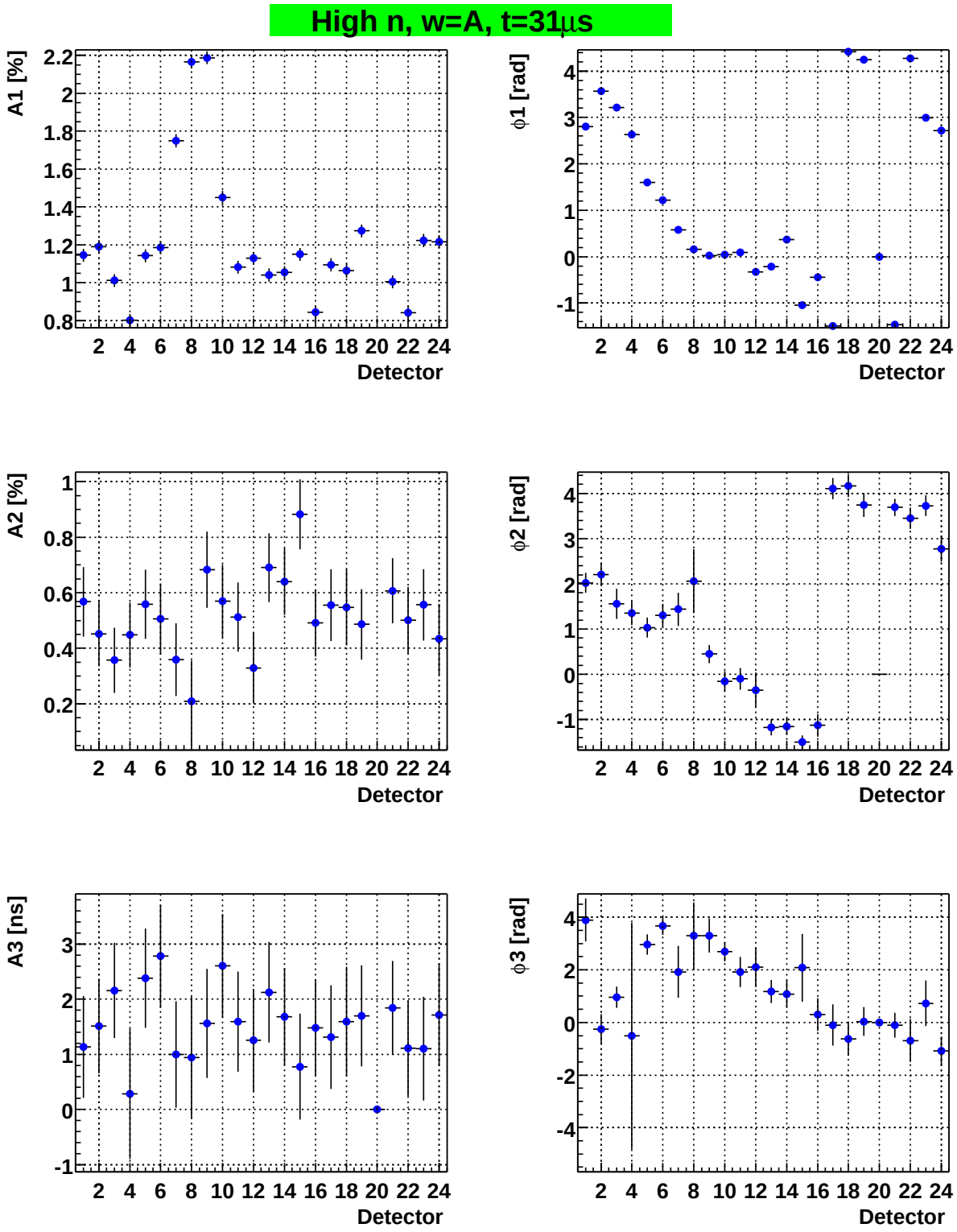
F.2 Asymmetry Weighting; Start Times at 31 μs

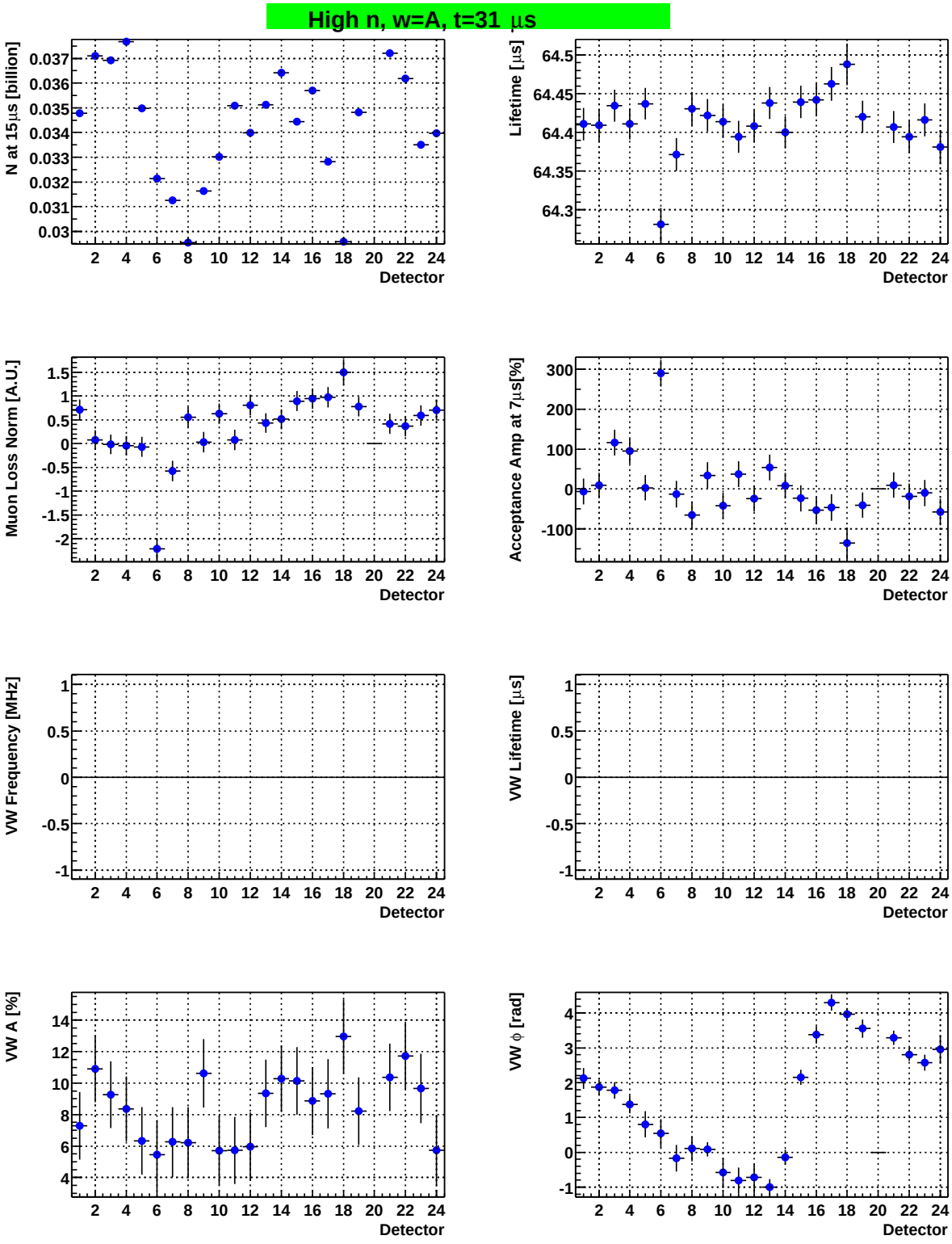



Low n, w=A, t=31μs
A1 [%]
φ1 [rad]
A2 [%]
φ2 [rad]
A3 [ns]
φ3 [rad]
Detector

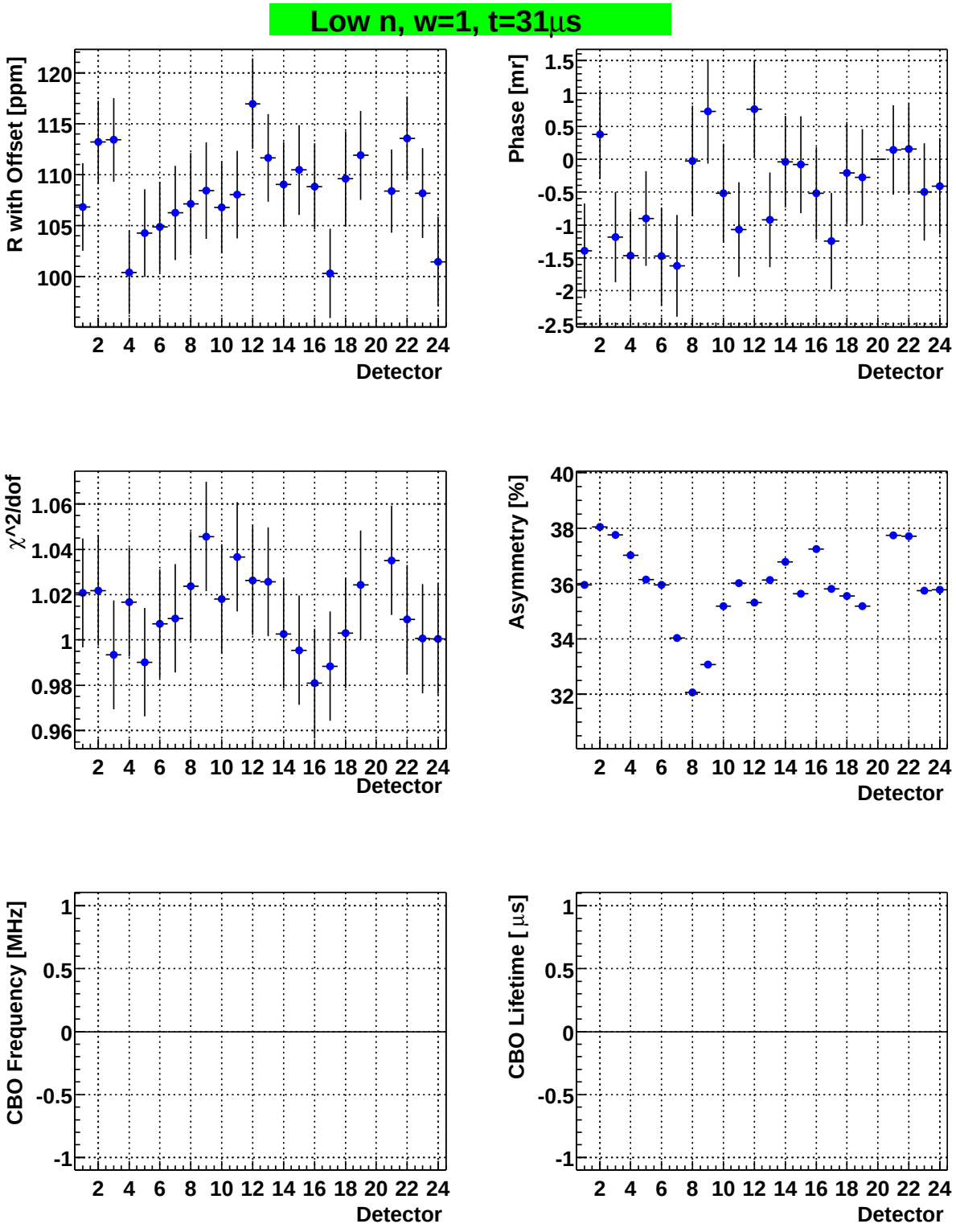

Low n, w=A, t=31 μs
N at 15μs [billion]
Lifetime [μs]
Muon Loss Norm [A.U.]
Acceptance Amp at 7μs[%]
VW Frequency [MHz]
VW Lifetime [μs]
VW A [%]
VW ϕ [rad]
Detector

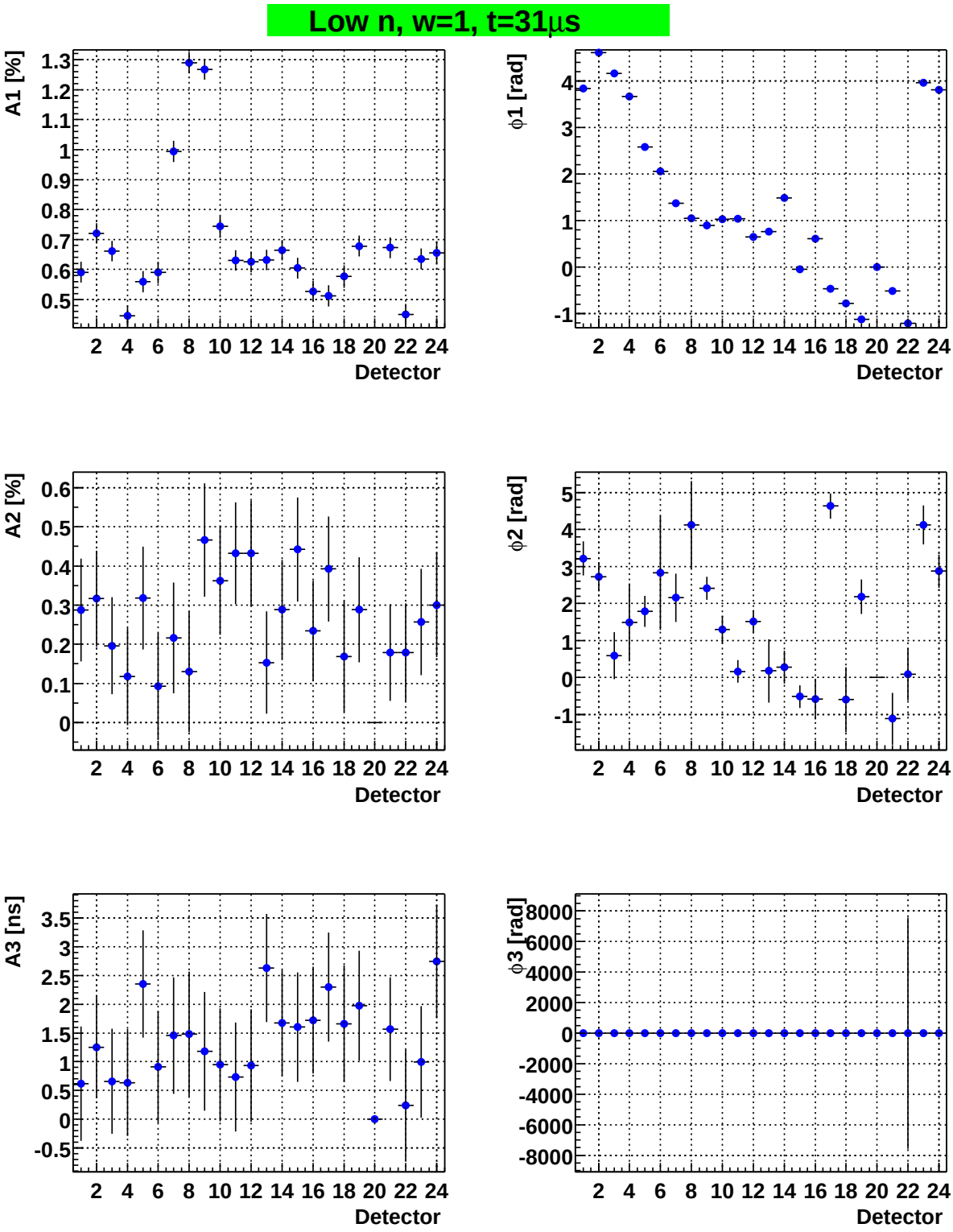

High n, w=A, t=31μs
R with Offset [ppm]
Phase [mr]
χ^2/dof
Asymmetry [%]
CBO Frequency [MHz]
CBO Lifetime [μs]
Detector

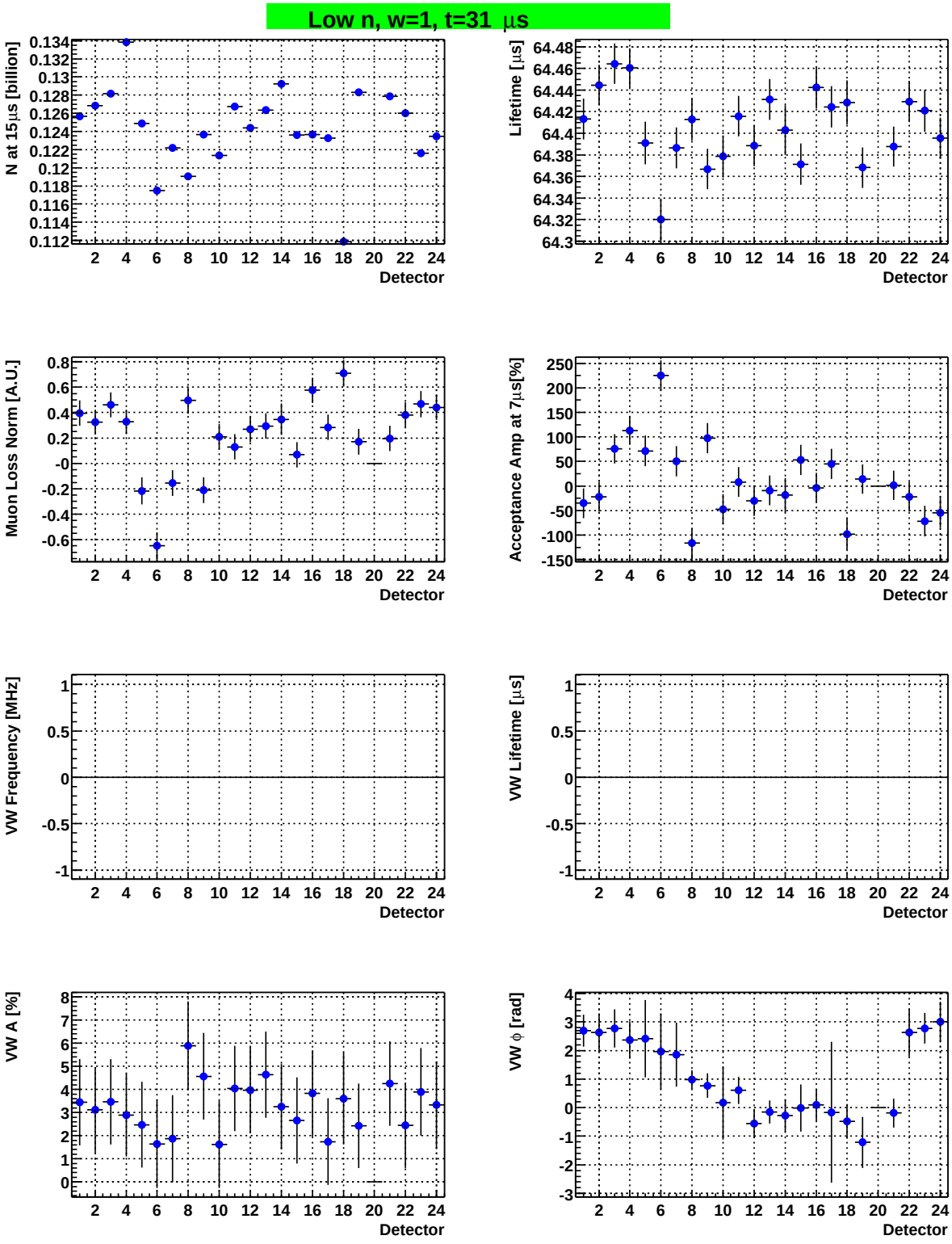

High n, w=A, t=31μs
A1 [%]
ϕ1 [rad]
A2 [%]
ϕ2 [rad]
A3 [ns]
ϕ3 [rad]
Detector

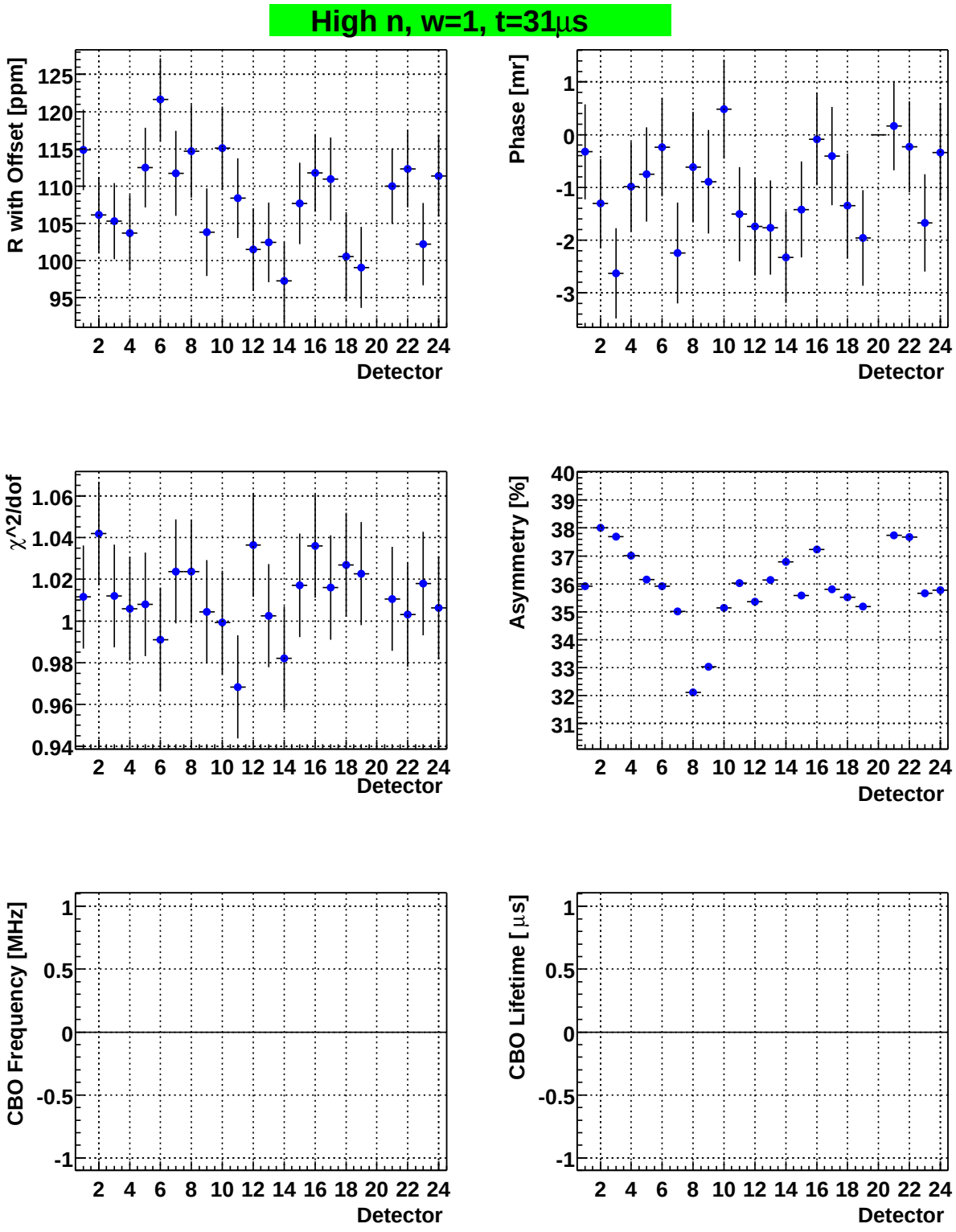

High n, w=A, t=31 μs
N at 15μs [billion]
Lifetime [μs]
Muon Loss Norm [A.U.]
Acceptance Amp at 7μs[%]
VW Frequency [MHz]
VW Lifetime [μs]
VW A [%]
VW ϕ [rad]
Detector

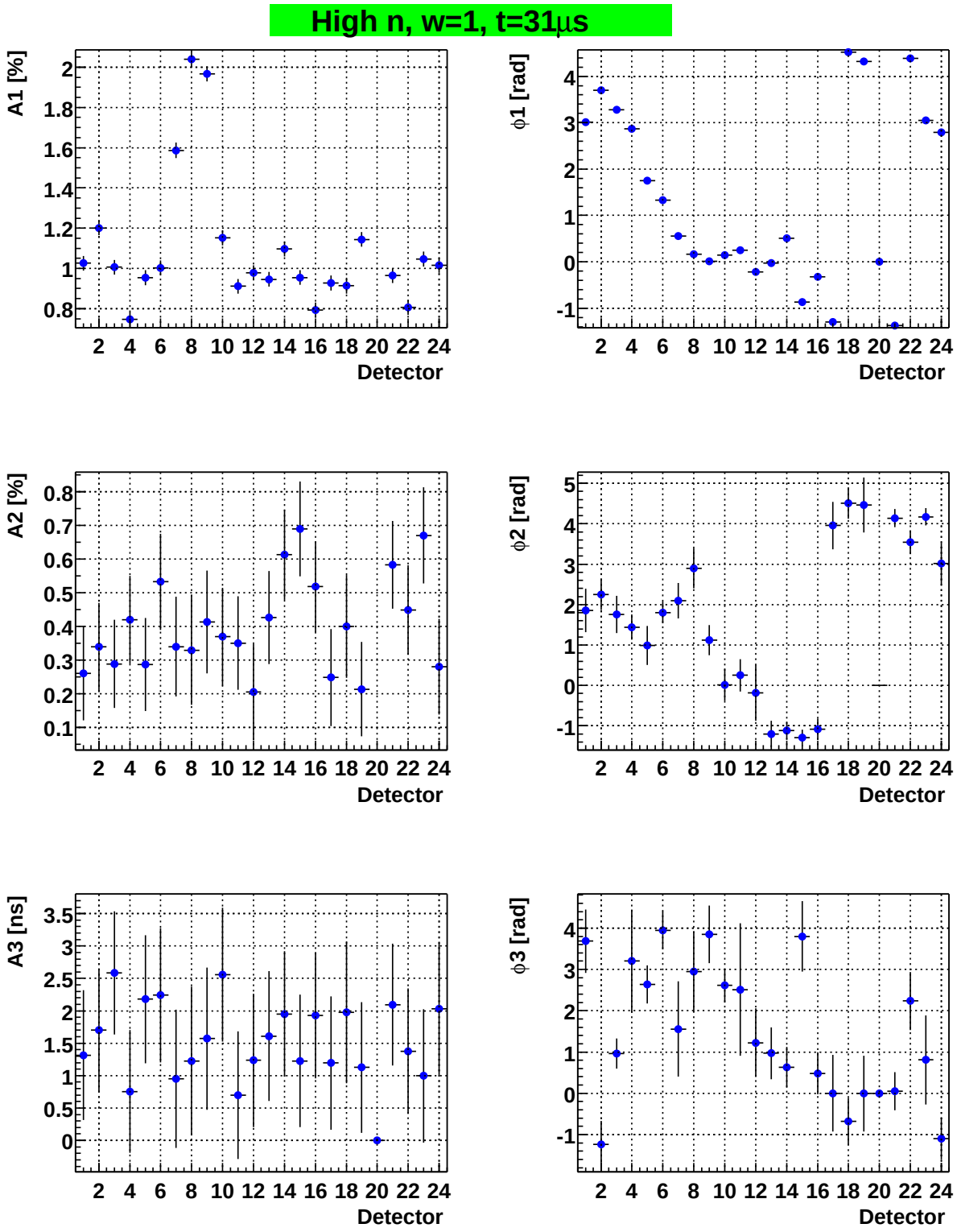
F.3 Uniform Weighting; Start Times at 31 μs



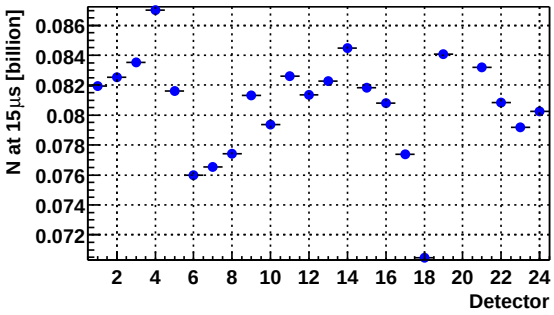

Low n, w=1, t=31μs
A1 [%]
Detector
ϕ1 [rad]
Detector
A2 [%]
Detector
ϕ2 [rad]
Detector
A3 [ns]
Detector
ϕ3 [rad]
Detector

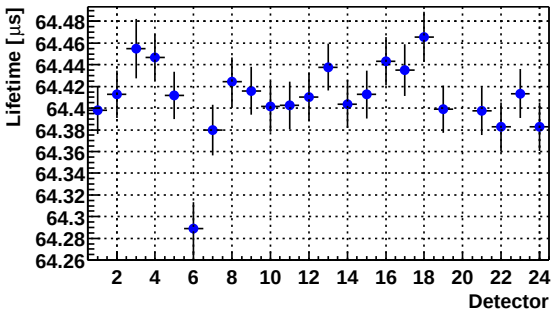

Low n, w=1, t=31 μs
N at 15μs [billion]
Lifetime [μs]
Muon Loss Norm [A.U.]
Acceptance Amp at 7μs[%]
VW Frequency [MHz]
VW Lifetime [μs]
VW A [%]
VW φ [rad]
Detector

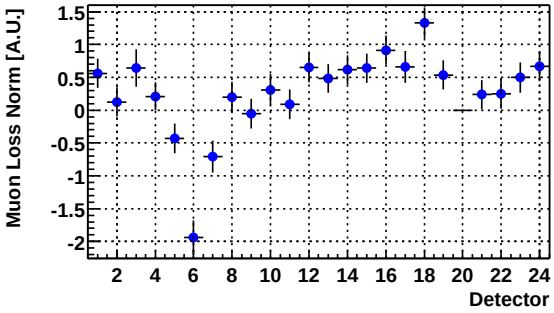

High n, w=1, t=31μs
R with Offset [ppm]
Phase [mr]
χ^2/dof
Asymmetry [%]
CBO Frequency [MHz]
CBO Lifetime [μs]
Detector

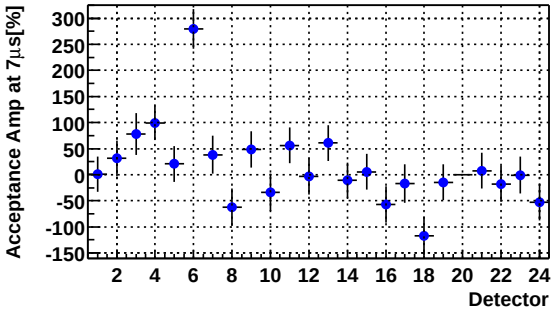

High n, w=1, t=31μs
A1 [%]
Detector
ϕ1 [rad]
Detector
A2 [%]
Detector
ϕ2 [rad]
Detector
A3 [ns]
Detector
ϕ3 [rad]
Detector

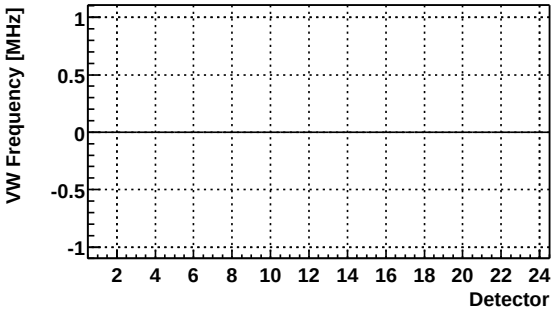
High n, w=1, t=31 μs

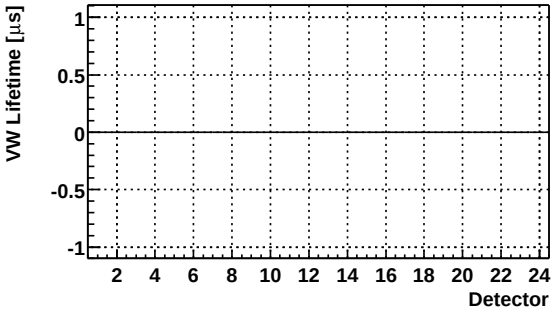


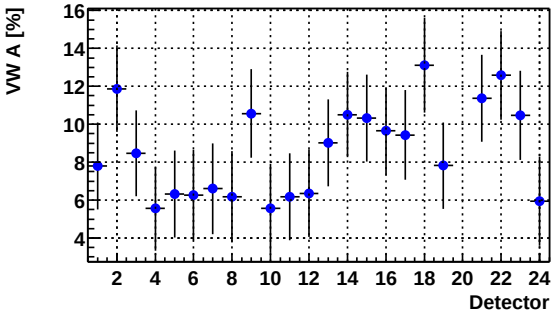


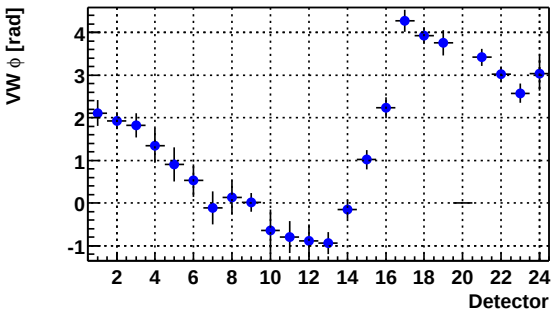










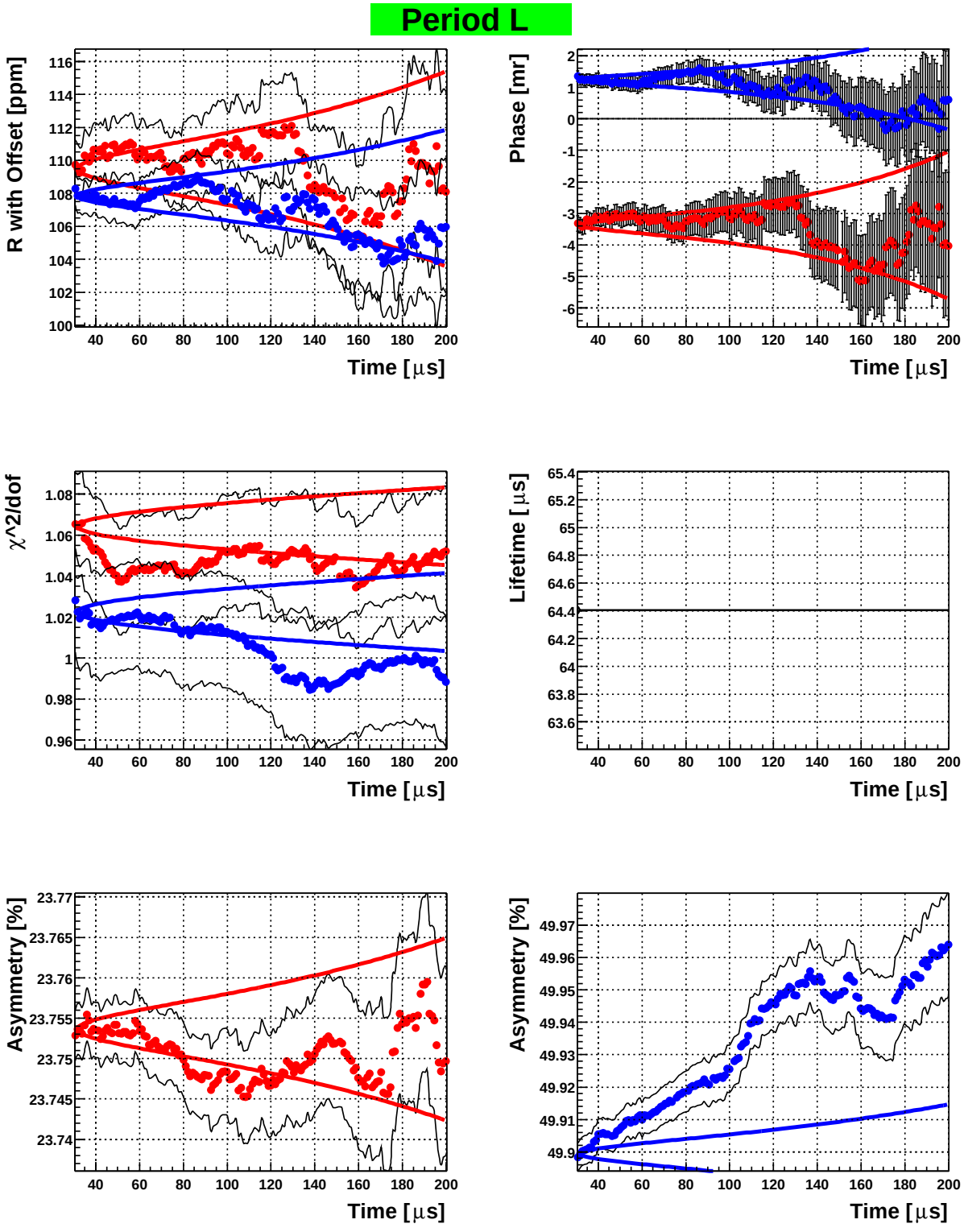


G Plots By Energy Band

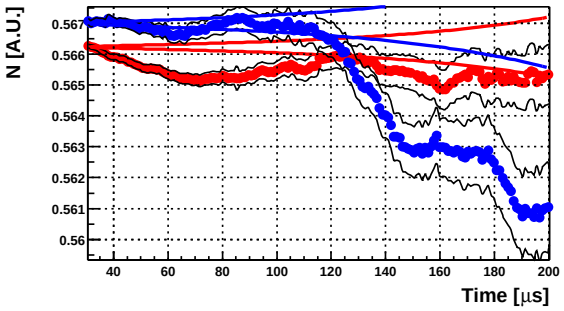
The plots in this section show start time scans for two different energy bands, 1.5-2.25(red) and 2.25-3.3(blue) GeV, for all of the run periods.

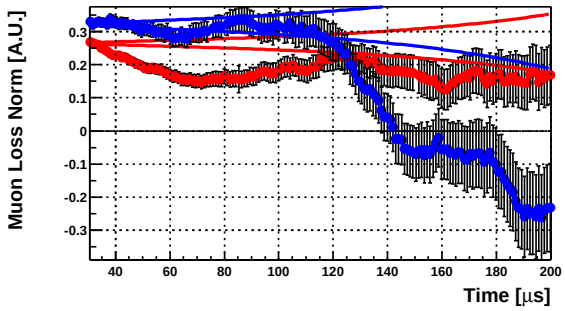
H Plots By Energy Band and Run Period

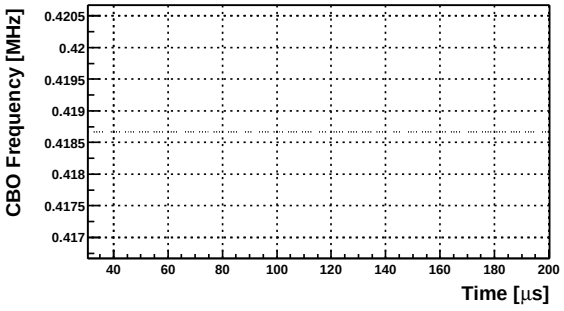
H.1 Period L

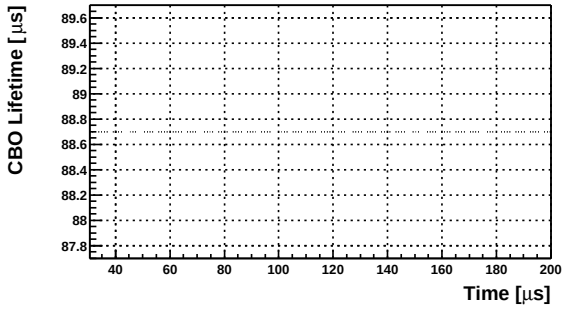


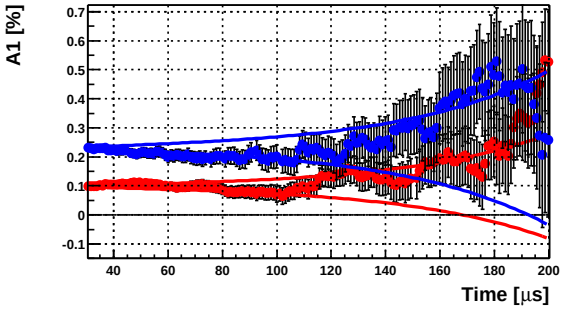
Period L

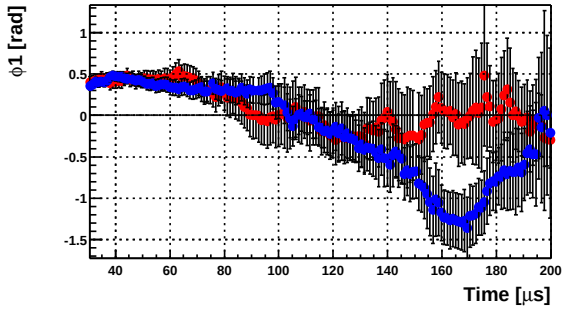

N [A.U.]
Time [μs]

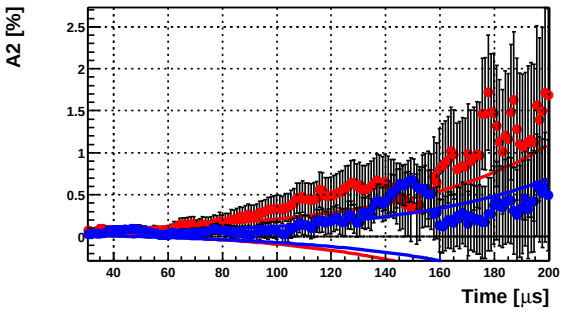

Muon Loss Norm [A.U.]
Time [μs]

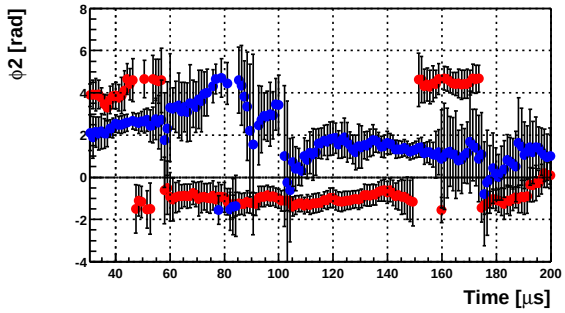

CBO Frequency [MHz]
Time [μs]


CBO Lifetime [μs]
Time [μs]

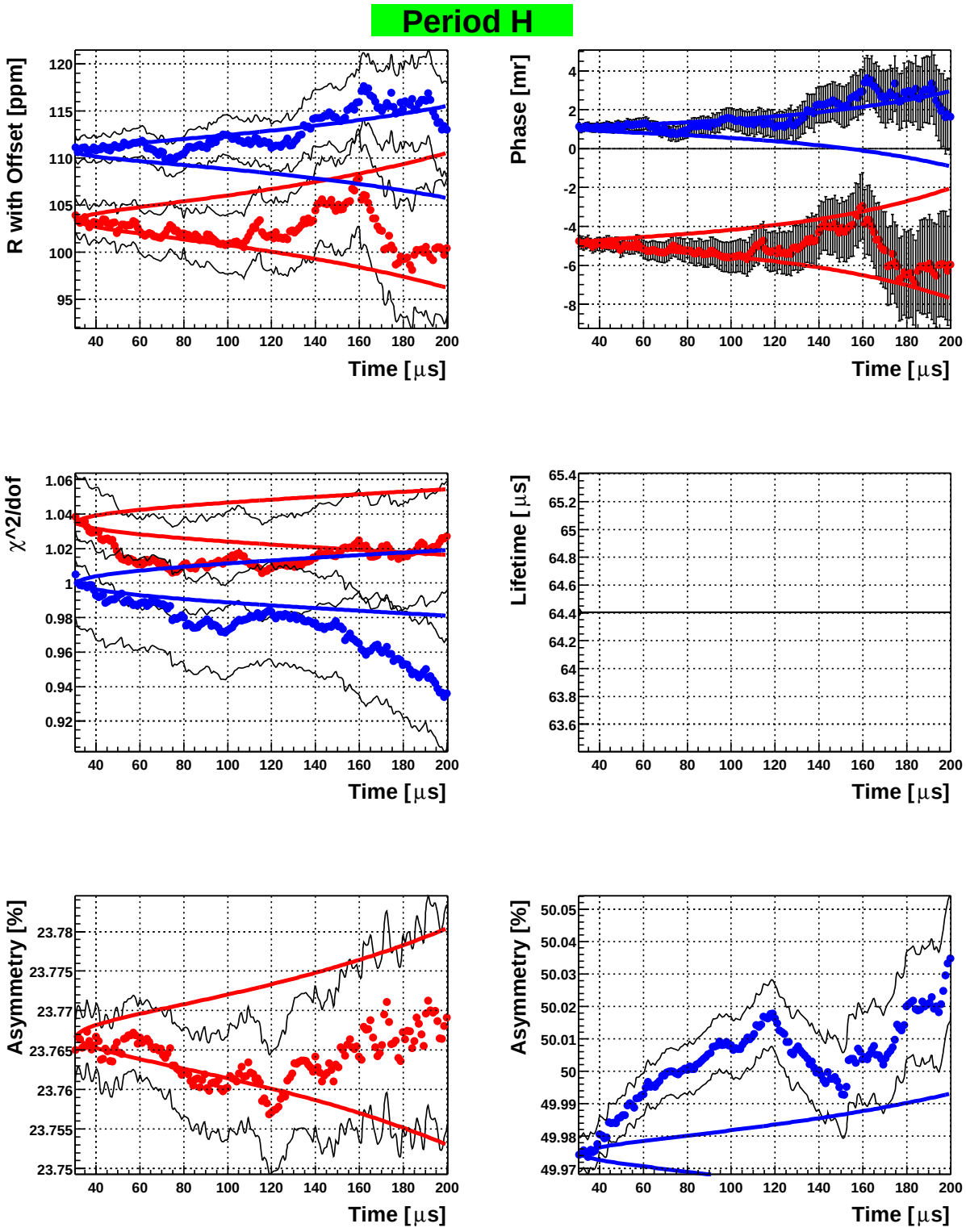

A1 [%]
Time [μs]

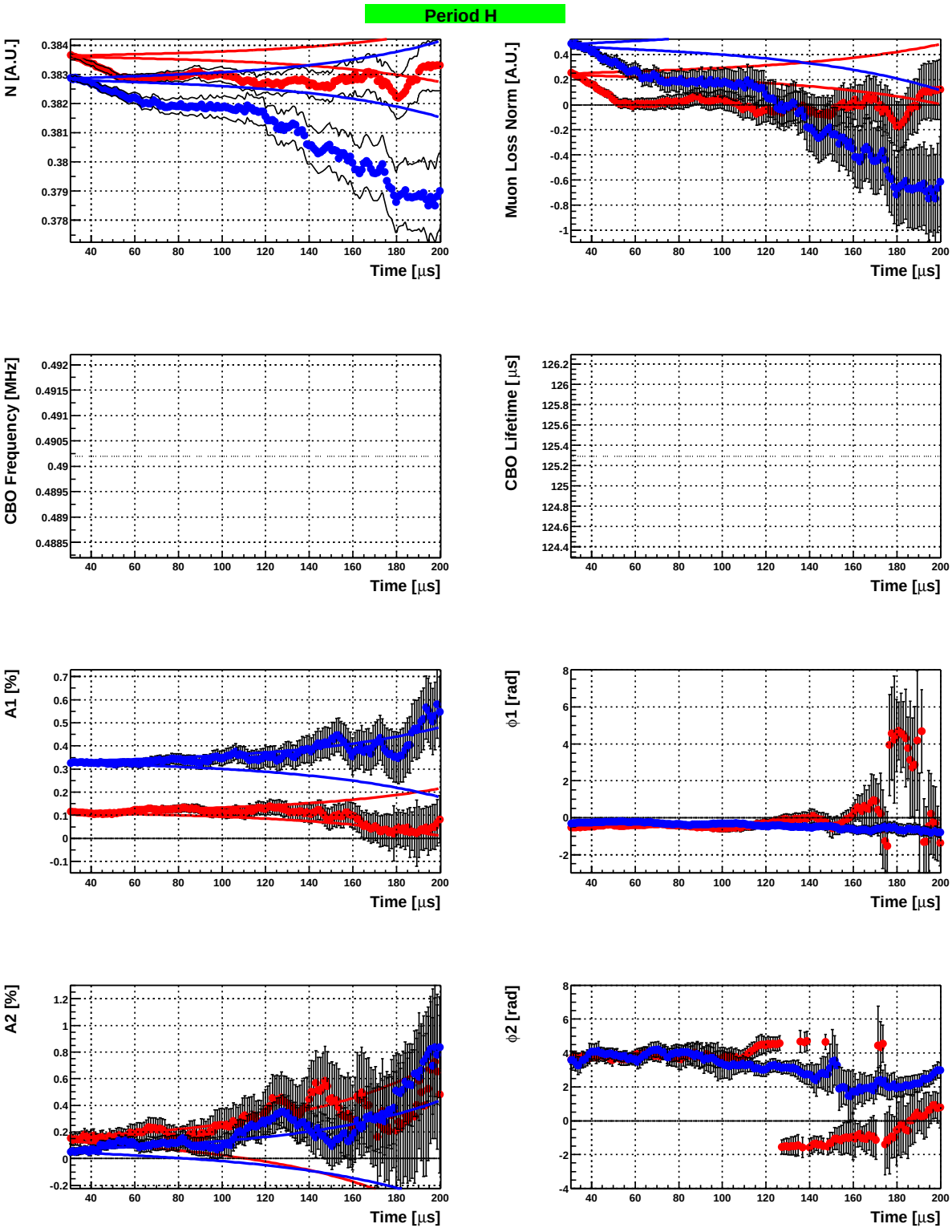

ϕ1 [rad]
Time [μs]


A2 [%]
Time [μs]

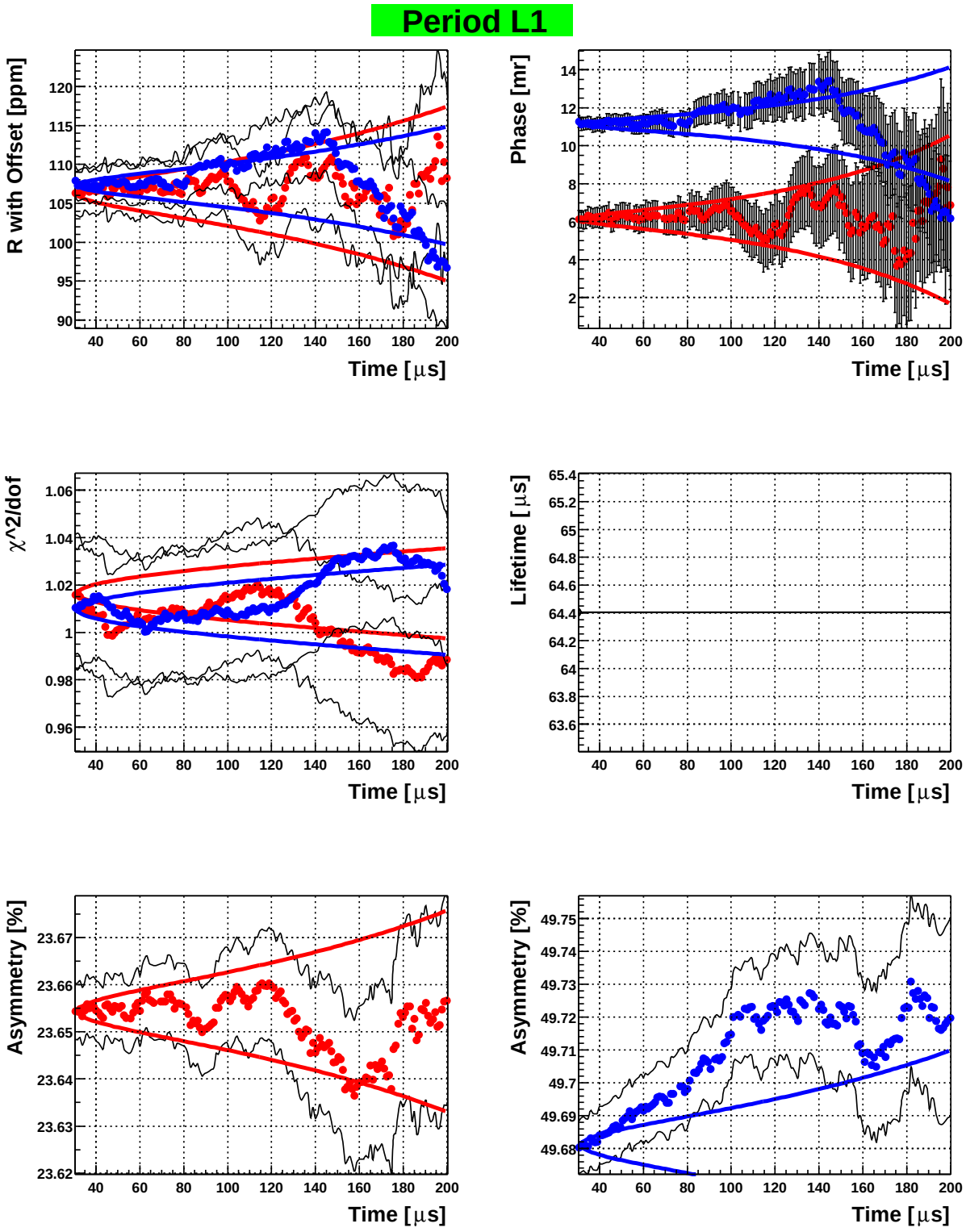

ϕ2 [rad]
Time [μs]

H.2 Period H

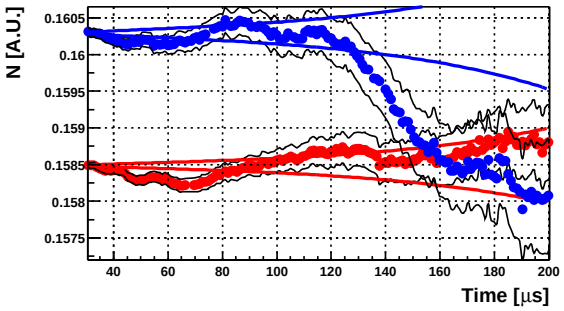

Period H
R with Offset [ppm]
Time [μs]
Phase [mr]
Time [μs]
χ^2/dof
Time [μs]
Lifetime [μs]
Time [μs]
Asymmetry [%]
Time [μs]
Asymmetry [%]
Time [μs]

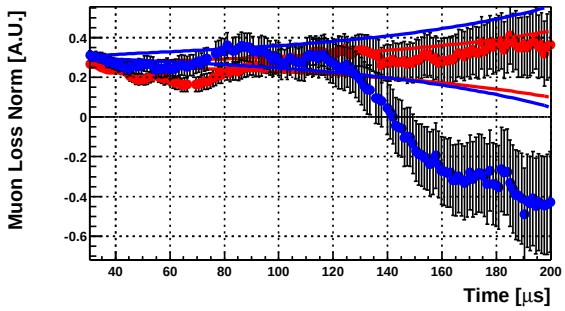

Period H
N [A.U.]
Muon Loss Norm [A.U.]
CBO Frequency [MHz]
CBO Lifetime [μs]
A1 [%]
ϕ1 [rad]
A2 [%]
ϕ2 [rad]
Time [μs]

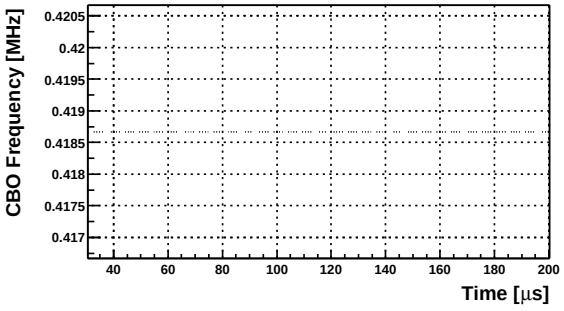
H.3 Period L1

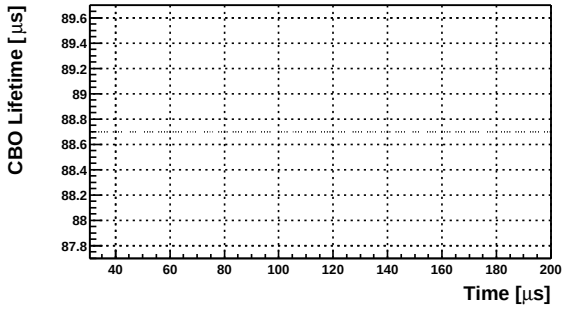

Period L1
R with Offset [ppm]
Time [μs]
Phase [mr]
Time [μs]
χ^2/dof
Time [μs]
Lifetime [μs]
Time [μs]
Asymmetry [%]
Time [μs]
Asymmetry [%]
Time [μs]

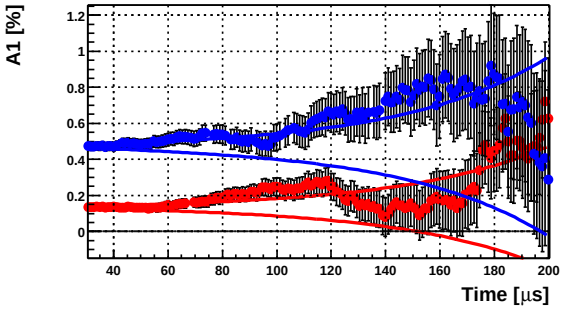
Period L1

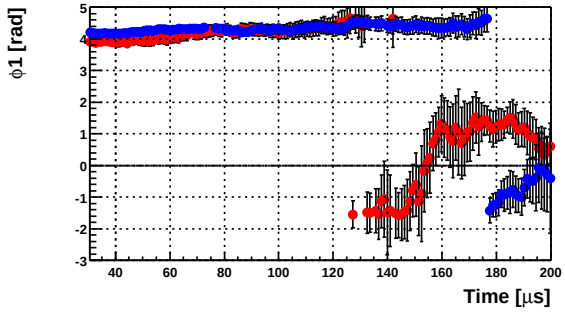

N [A.U.]
Time [μs]

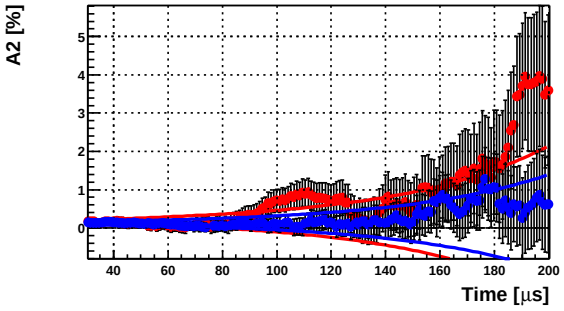

Muon Loss Norm [A.U.]
Time [μs]

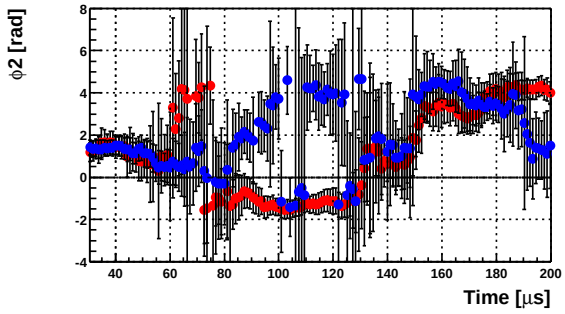

CBO Frequency [MHz]
Time [μs]


CBO Lifetime [μs]
Time [μs]

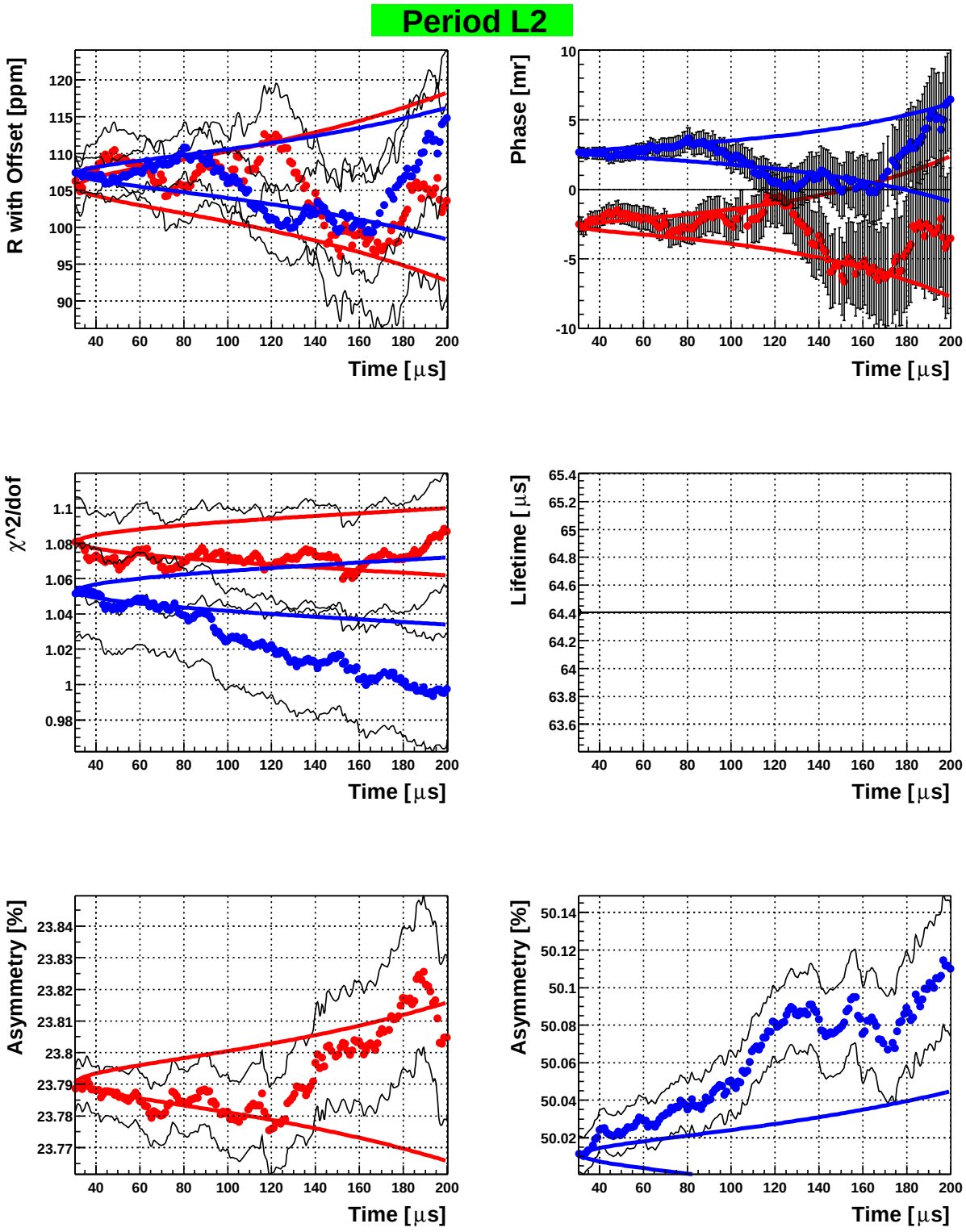

A1 [%]
Time [μs]


ϕ1 [rad]
Time [μs]

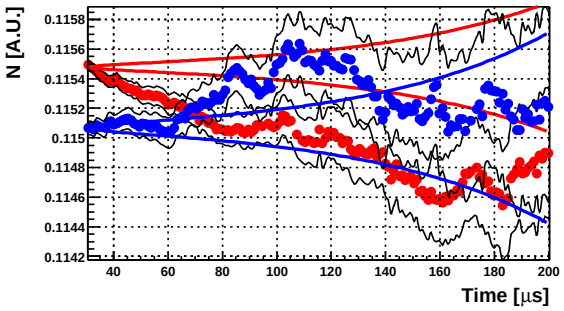

A2 [%]
Time [μs]

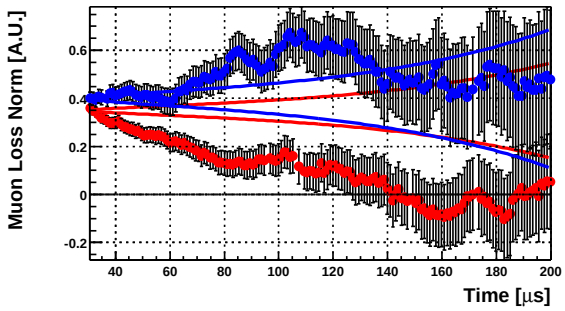

ϕ2 [rad]
Time [μs]

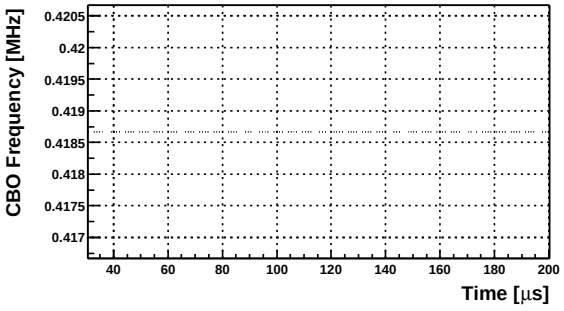
H.4 Period L2

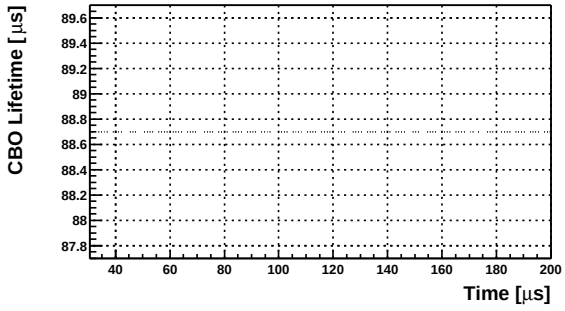

Period L2
R with Offset [ppm]
Time [μs]
Phase [mr]
Time [μs]
χ^2/dof
Time [μs]
Lifetime [μs]
Time [μs]
Asymmetry [%]
Time [μs]
Asymmetry [%]
Time [μs]

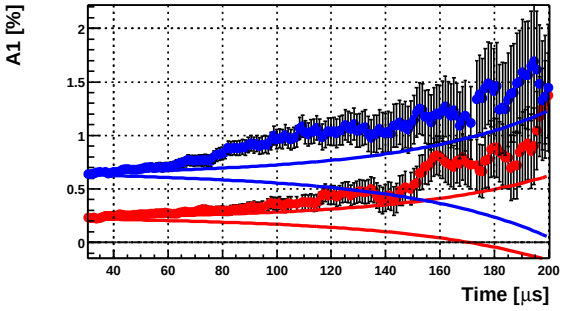
Period L2

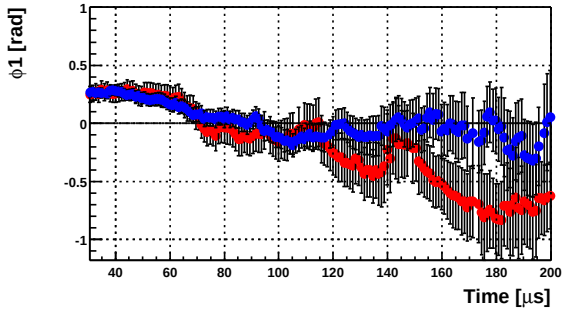

N [A.U.]
Time [μs]

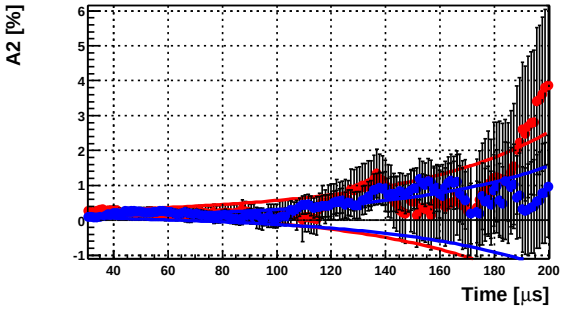

Muon Loss Norm [A.U.]
Time [μs]

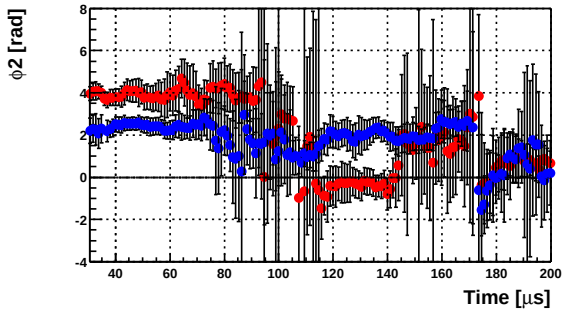

CBO Frequency [MHz]
Time [μs]


CBO Lifetime [μs]
Time [μs]

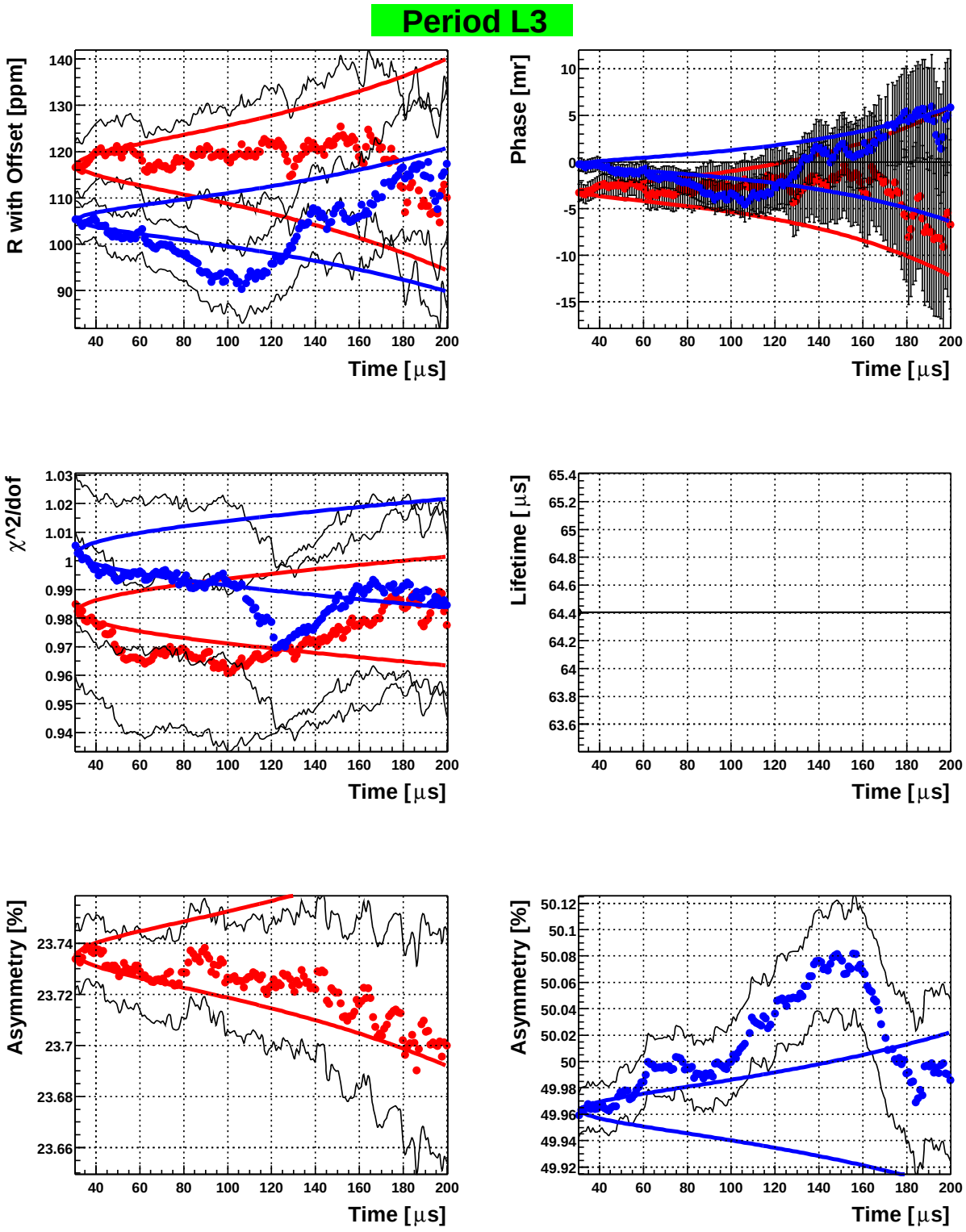

A1 [%]
Time [μs]


ϕ1 [rad]
Time [μs]

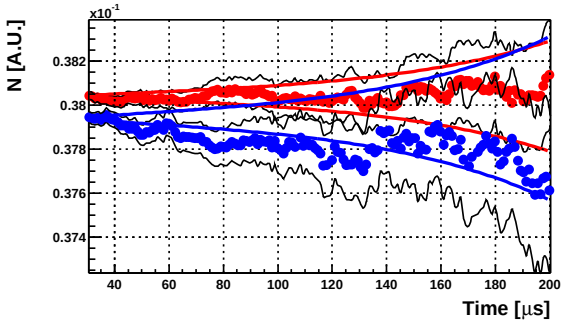

A2 [%]
Time [μs]

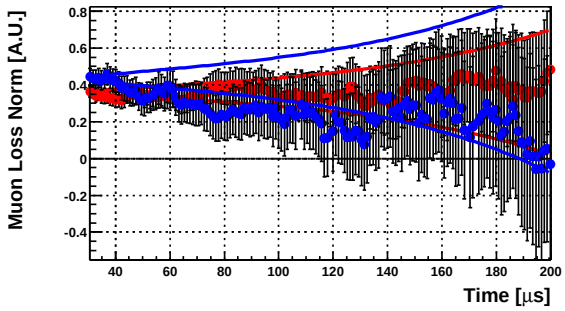

ϕ2 [rad]
Time [μs]

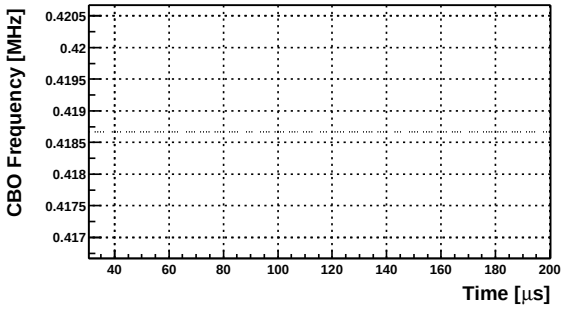
H.5 Period L3

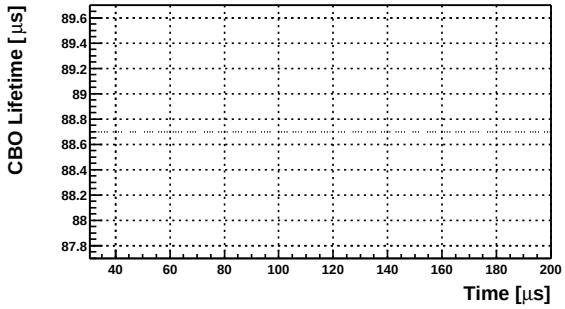

Period L3
R with Offset [ppm]
Time [μs]
Phase [mr]
Time [μs]
χ^2/dof
Time [μs]
Lifetime [μs]
Time [μs]
Asymmetry [%]
Time [μs]
Asymmetry [%]
Time [μs]

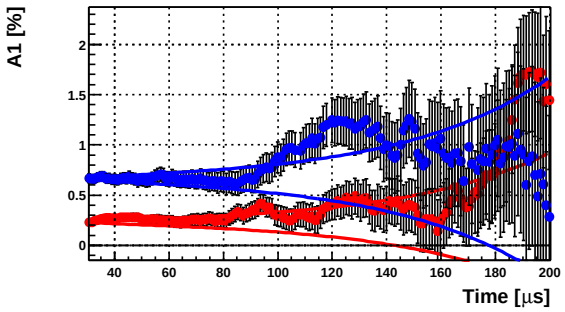
Period L3

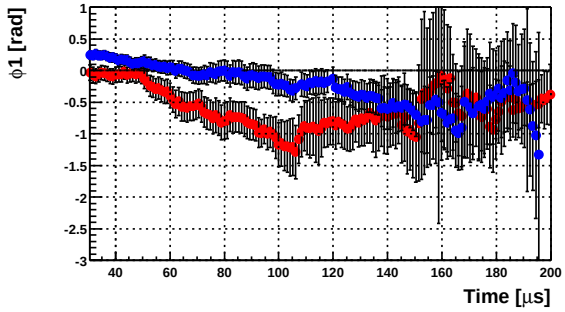

x10^-1
N [A.U.]
0.382
0.38
0.378
0.376
0.374
40 60 80 100 120 140 160 180 200
Time [μs]

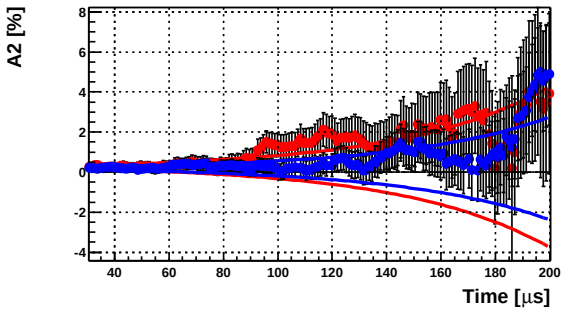

Muon Loss Norm [A.U.]
0.8
0.6
0.4
0.2
0
-0.2
-0.4
40 60 80 100 120 140 160 180 200
Time [μs]

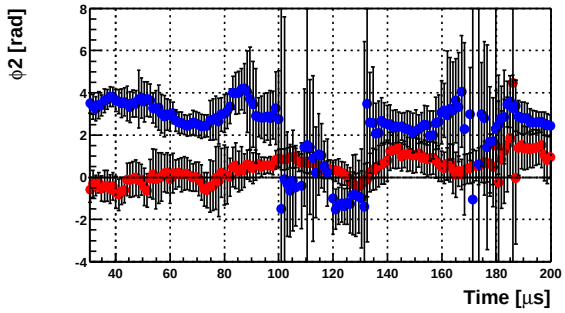

CBO Frequency [MHz]
0.4205
0.42
0.4195
0.419
0.4185
0.418
0.4175
0.417
40 60 80 100 120 140 160 180 200
Time [μs]


CBO Lifetime [μs]
89.6
89.4
89.2
89
88.8
88.6
88.4
88.2
88
87.8
40 60 80 100 120 140 160 180 200
Time [μs]

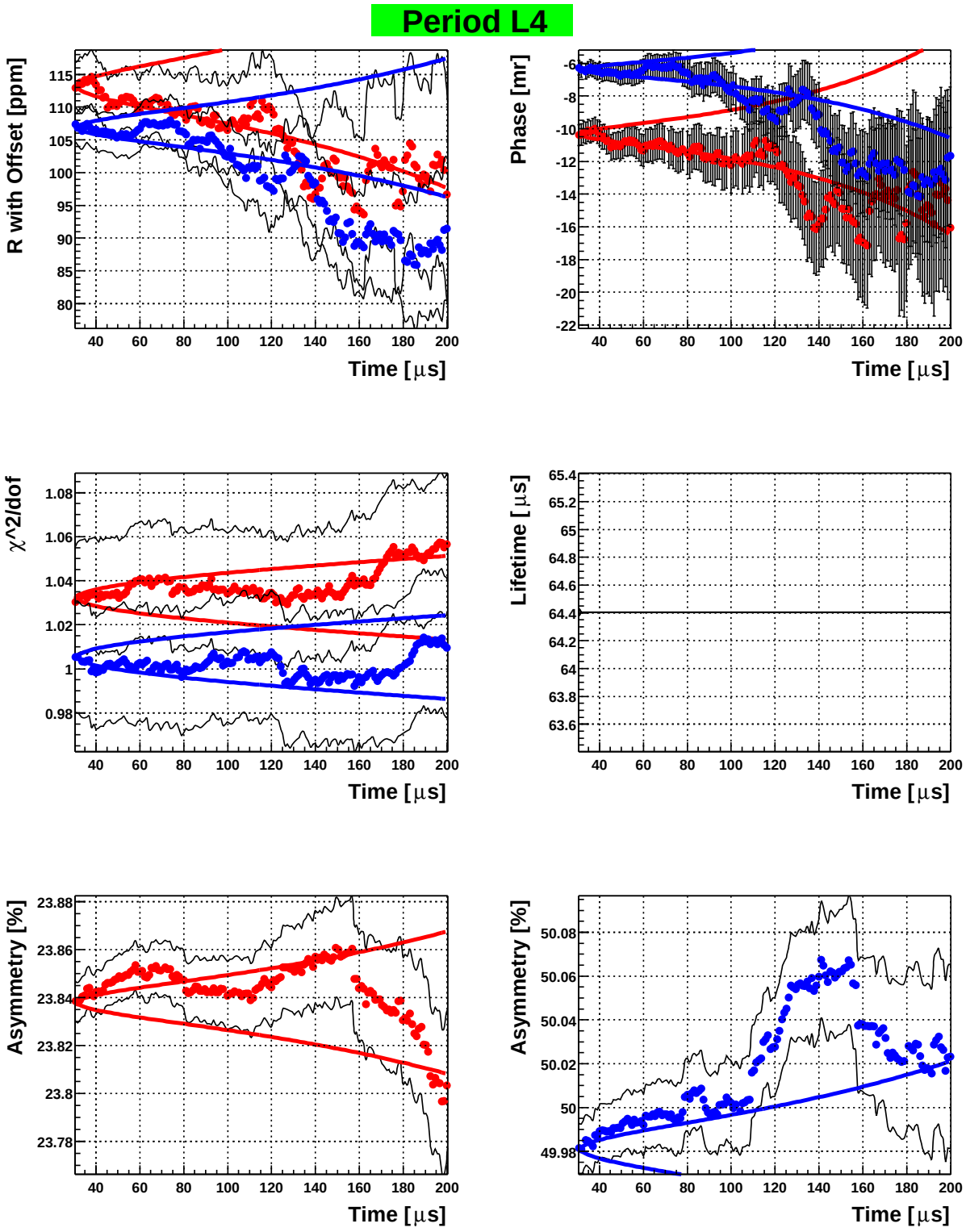

A1 [%]
2
1.5
1
0.5
0
40 60 80 100 120 140 160 180 200
Time [μs]


ϕ1 [rad]
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
40 60 80 100 120 140 160 180 200
Time [μs]

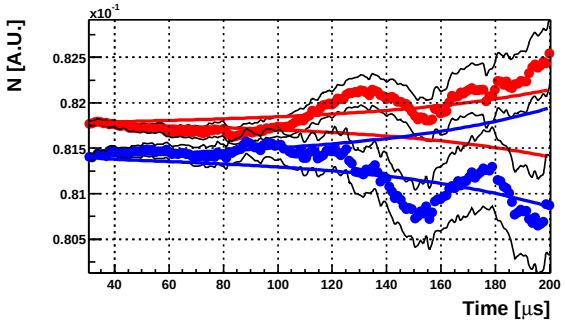

A2 [%]
8
6
4
2
0
-2
-4
40 60 80 100 120 140 160 180 200
Time [μs]

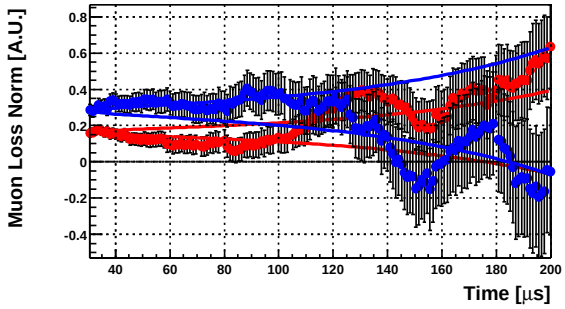

ϕ2 [rad]
8
6
4
2
0
-2
-4
40 60 80 100 120 140 160 180 200
Time [μs]

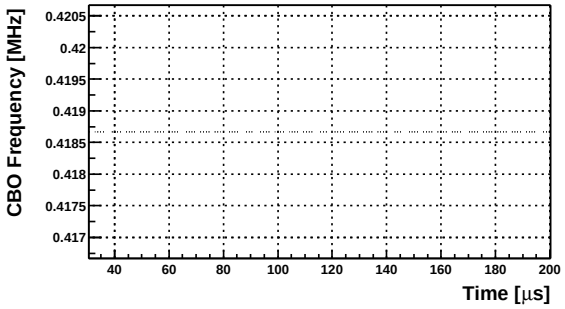
H.6 Period L4

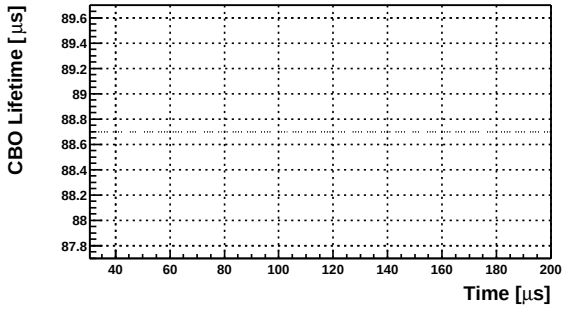

Period L4
R with Offset [ppm]
Time [μs]
Phase [mr]
Time [μs]
χ^2/dof
Time [μs]
Lifetime [μs]
Time [μs]
Asymmetry [%]
Time [μs]
Asymmetry [%]
Time [μs]

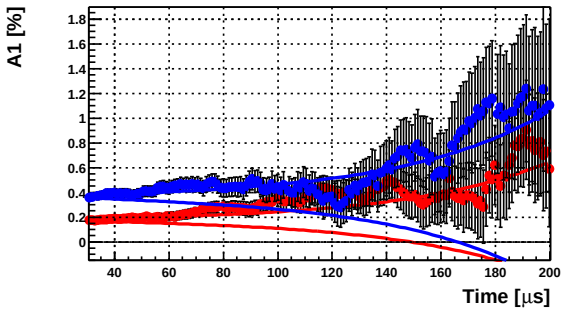
Period L4

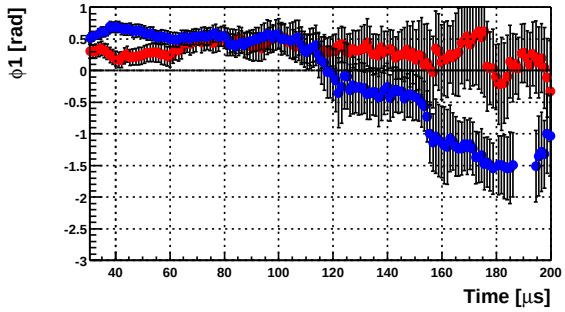

x10^-1
N [A.U.]
0.825
0.82
0.815
0.81
0.805
40 60 80 100 120 140 160 180 200
Time [μs]

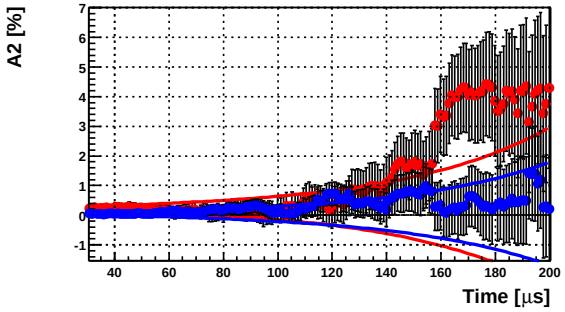

Muon Loss Norm [A.U.]
0.8
0.6
0.4
0.2
0
-0.2
-0.4
40 60 80 100 120 140 160 180 200
Time [μs]

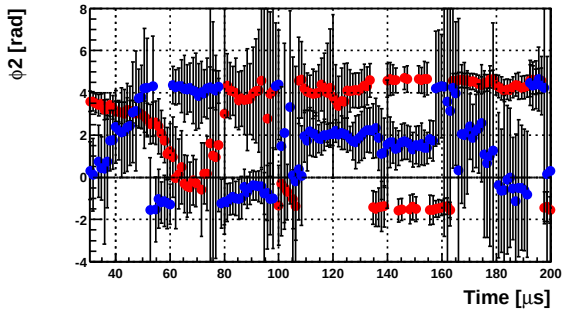

CBO Frequency [MHz]
0.4205
0.42
0.4195
0.419
0.4185
0.418
0.4175
0.417
40 60 80 100 120 140 160 180 200
Time [μs]


CBO Lifetime [μs]
89.6
89.4
89.2
89
88.8
88.6
88.4
88.2
88
87.8
40 60 80 100 120 140 160 180 200
Time [μs]

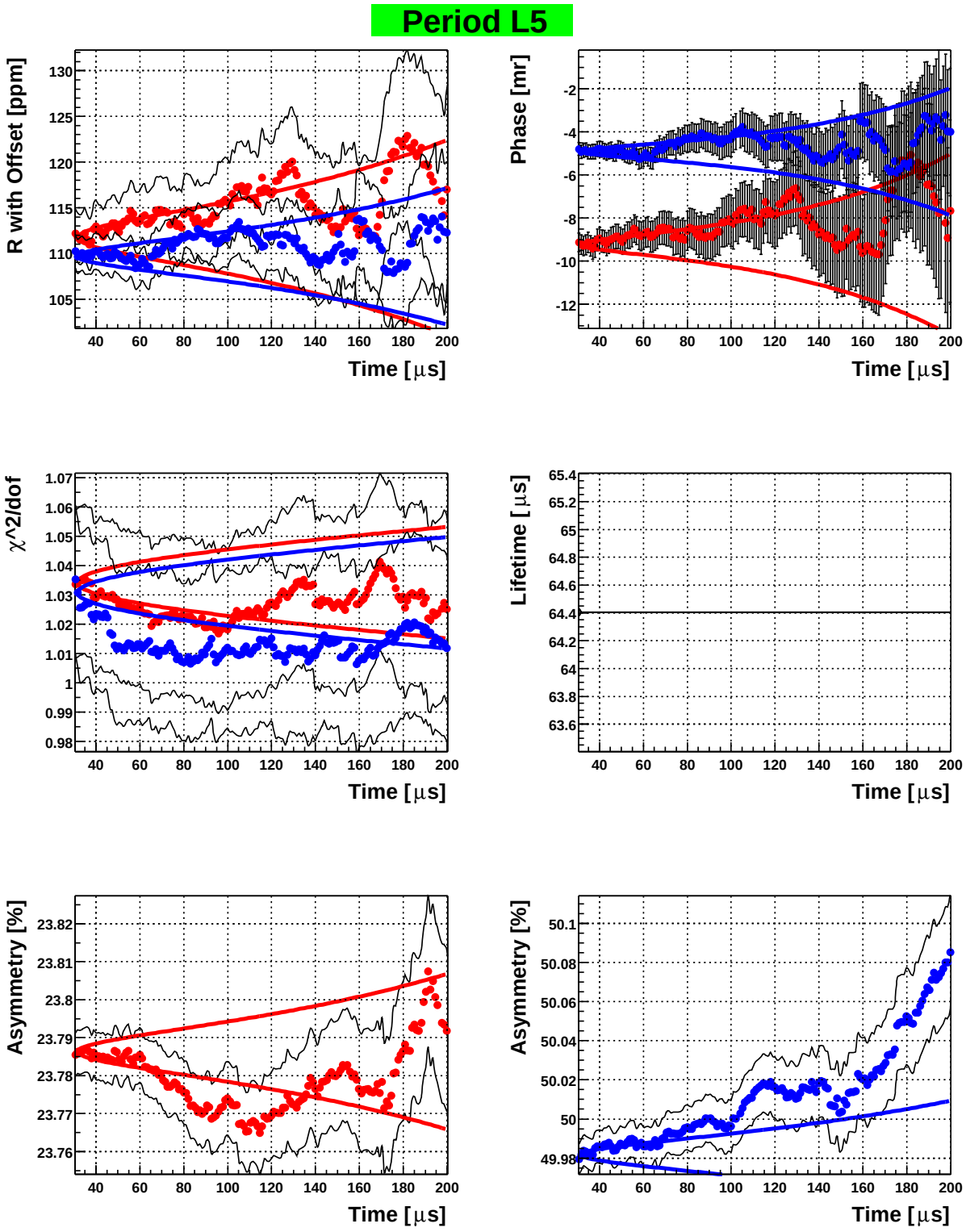

A1 [%]
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
40 60 80 100 120 140 160 180 200
Time [μs]


ϕ1 [rad]
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
-3
40 60 80 100 120 140 160 180 200
Time [μs]

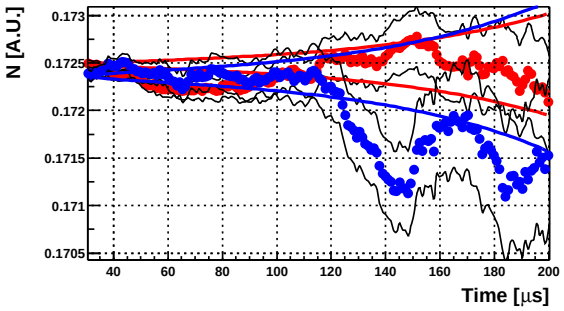

A2 [%]
7
6
5
4
3
2
1
0
-1
40 60 80 100 120 140 160 180 200
Time [μs]

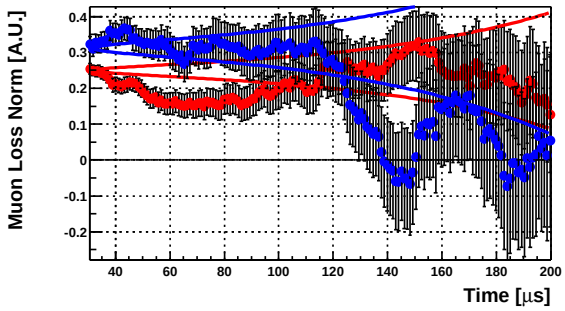

ϕ2 [rad]
8
6
4
2
0
-2
-4
40 60 80 100 120 140 160 180 200
Time [μs]

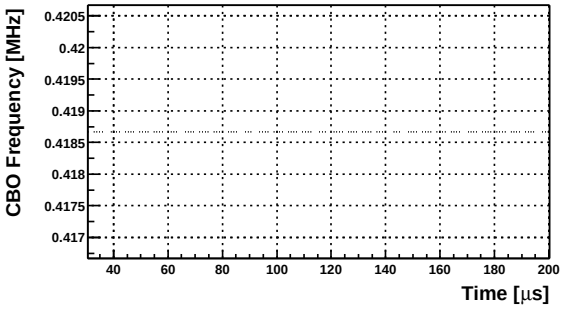
H.7 Period L5

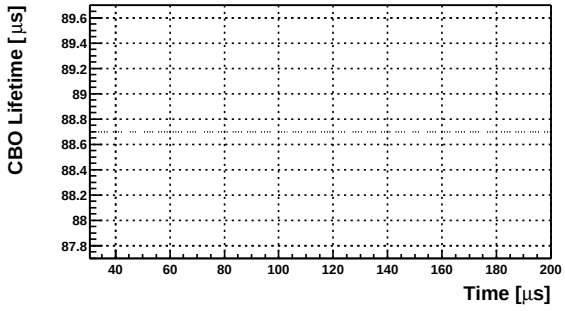

Period L5
R with Offset [ppm]
Phase [mr]
χ^2/dof
Lifetime [μs]
Asymmetry [%]
Asymmetry [%]
Time [μs]

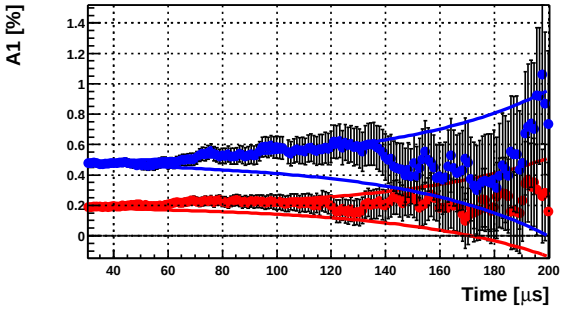
Period L5

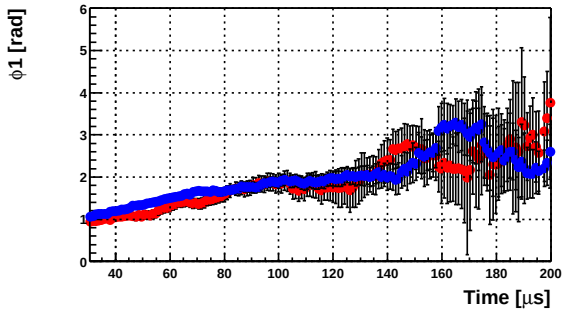

N [A.U.]
Time [μs]

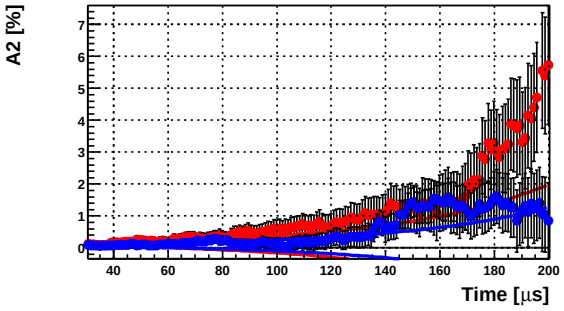

Muon Loss Norm [A.U.]
Time [μs]

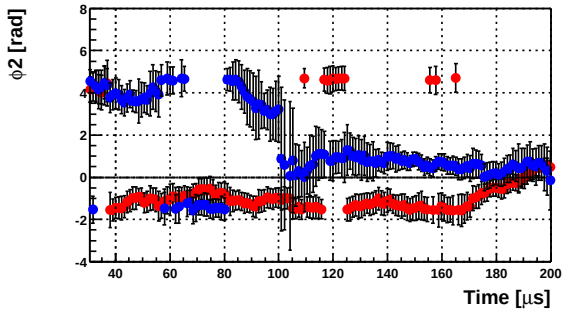

CBO Frequency [MHz]
Time [μs]


CBO Lifetime [μs]
Time [μs]

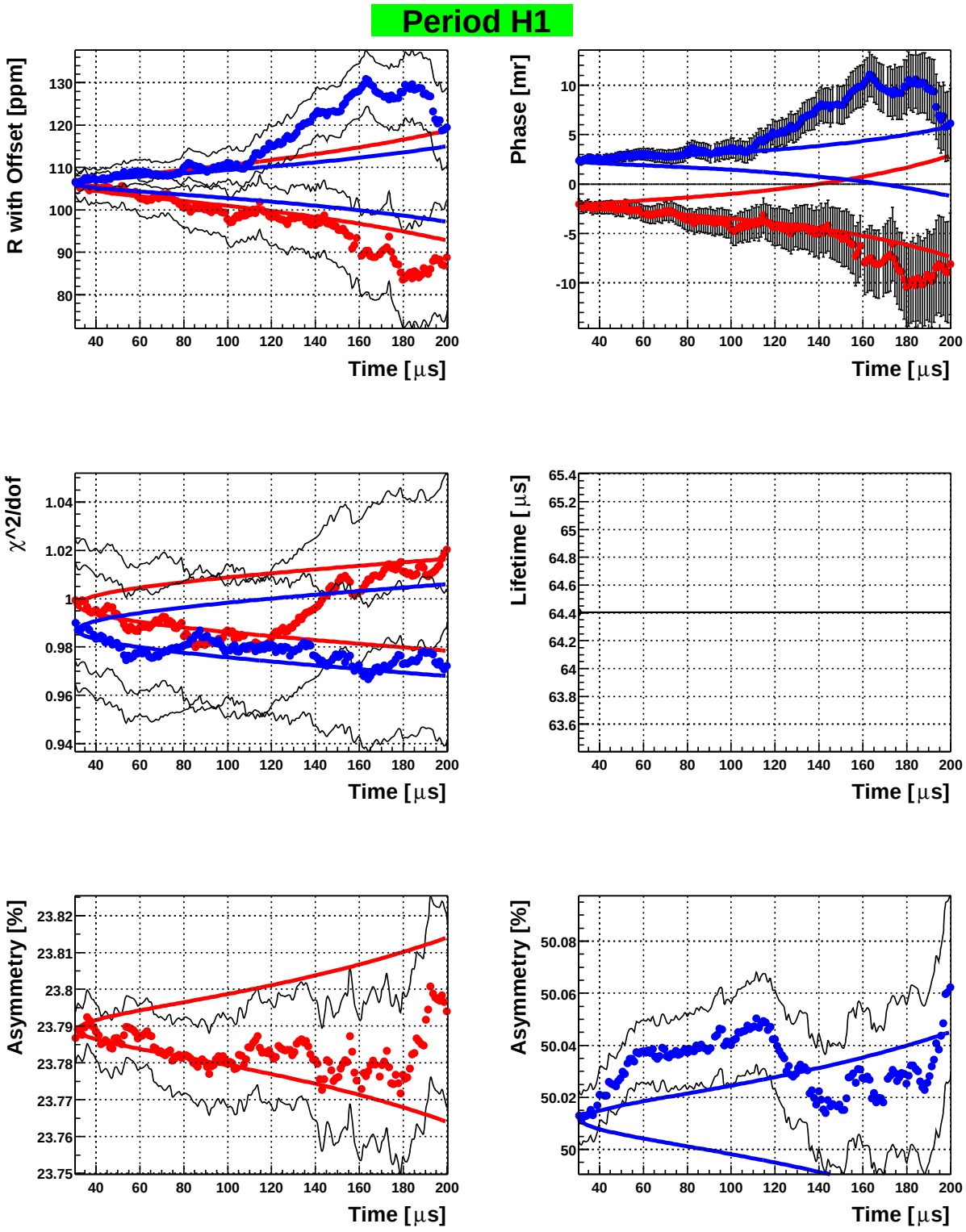

A1 [%]
Time [μs]


ϕ1 [rad]
Time [μs]

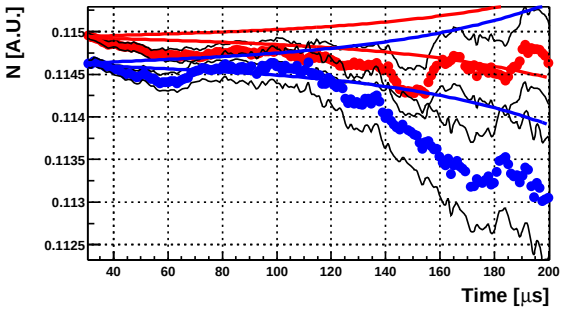

A2 [%]
Time [μs]

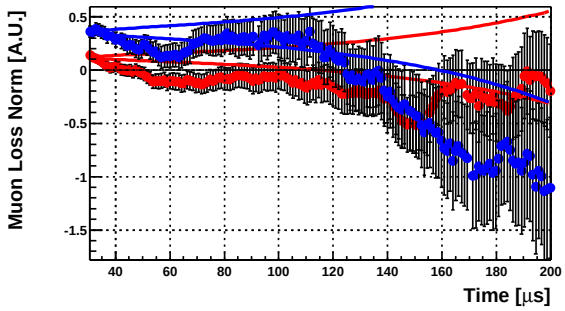

ϕ2 [rad]
Time [μs]

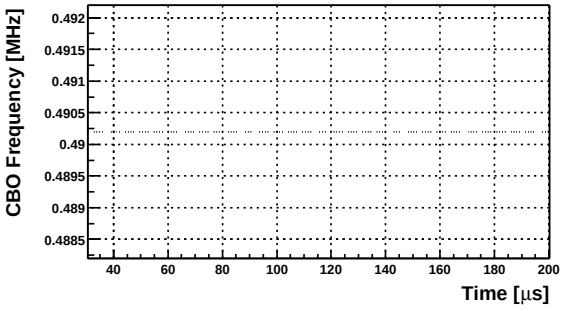
H.8 Period H1

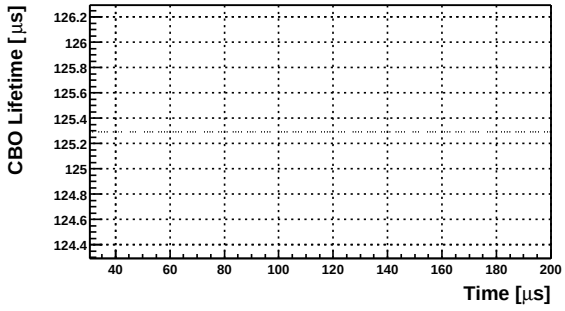

Period H1
R with Offset [ppm]
Time [μs]
Phase [mr]
Time [μs]
χ^2/dof
Time [μs]
Lifetime [μs]
Time [μs]
Asymmetry [%]
Time [μs]
Asymmetry [%]
Time [μs]

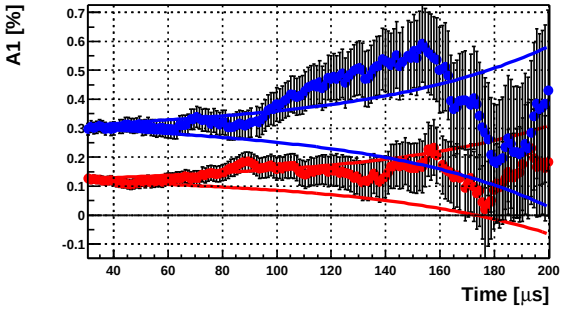
Period H1

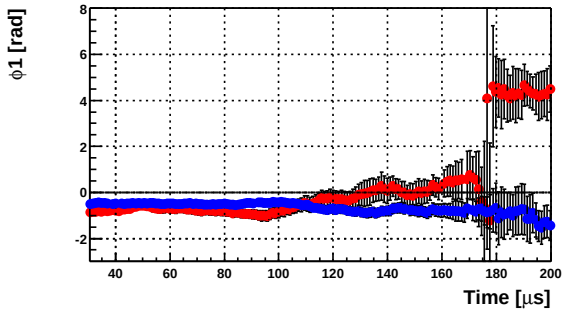

N [A.U.]
Time [μs]

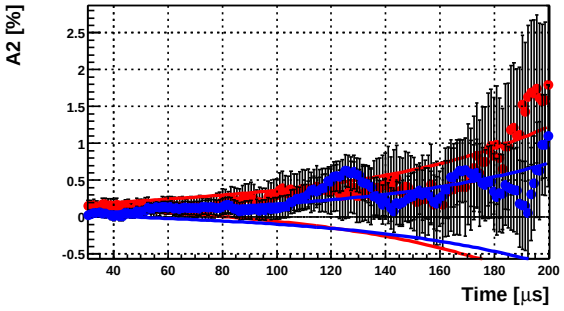

Muon Loss Norm [A.U.]
Time [μs]

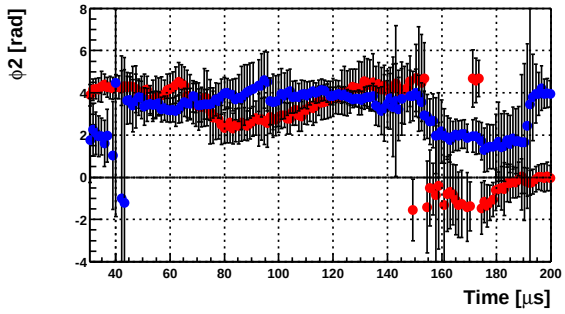

CBO Frequency [MHz]
Time [μs]


CBO Lifetime [μs]
Time [μs]

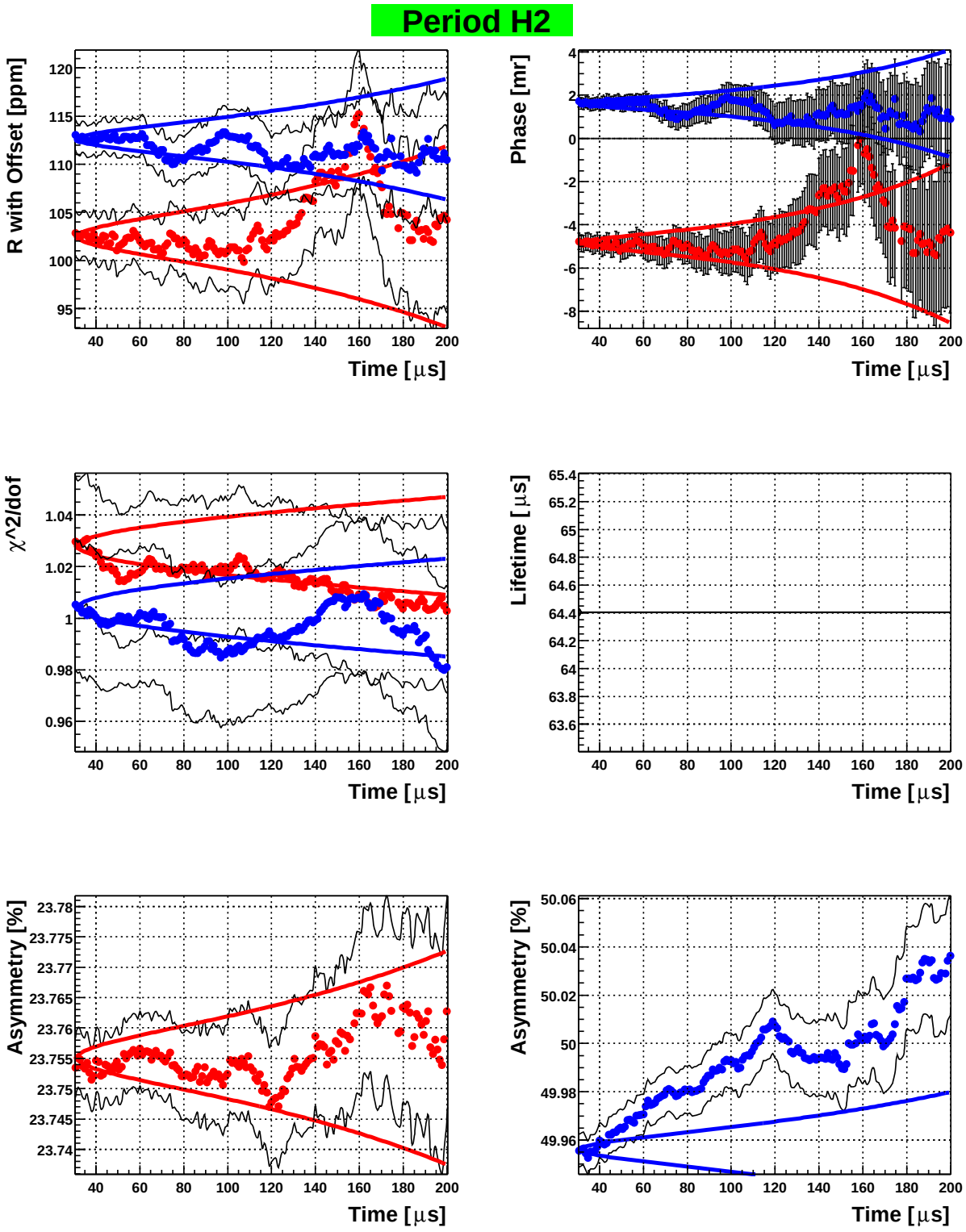

A1 [%]
Time [μs]


ϕ1 [rad]
Time [μs]

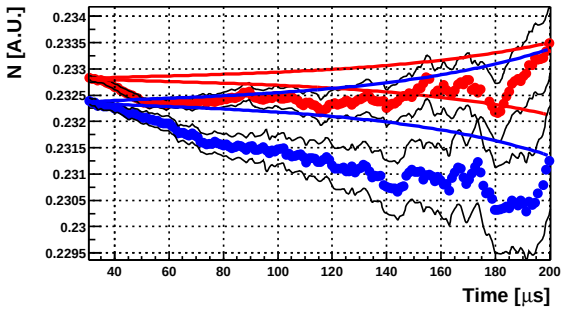

A2 [%]
Time [μs]

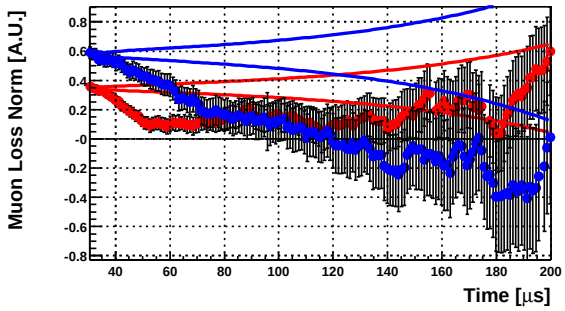

ϕ2 [rad]
Time [μs]

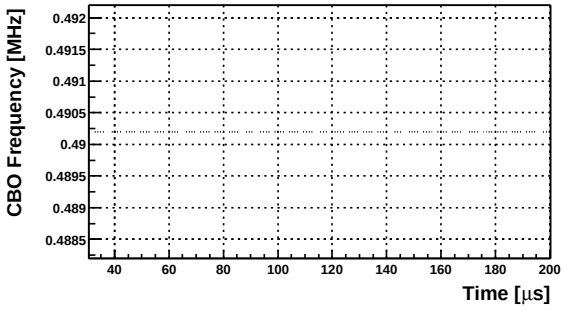
H.9 Period H2

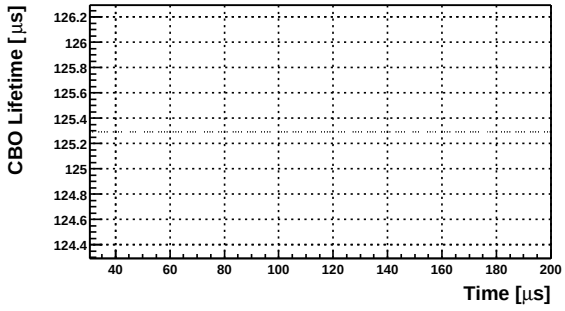

Period H2
R with Offset [ppm]
Time [μs]
Phase [mr]
Time [μs]
χ^2/dof
Time [μs]
Lifetime [μs]
Time [μs]
Asymmetry [%]
Time [μs]
Asymmetry [%]
Time [μs]

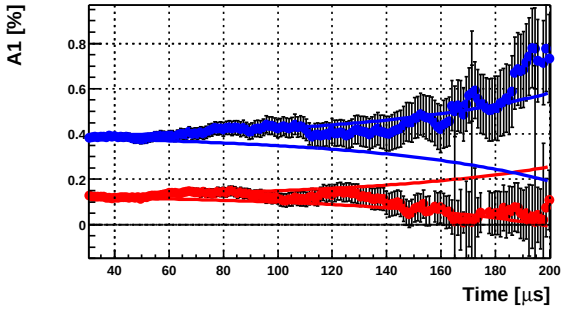
Period H2

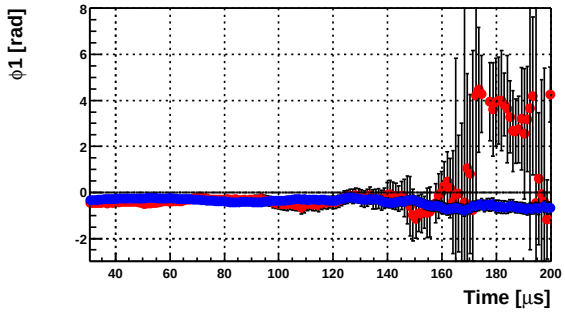

N [A.U.]
Time [μs]

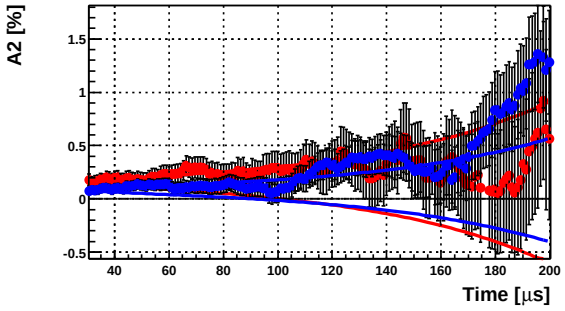

Muon Loss Norm [A.U.]
Time [μs]

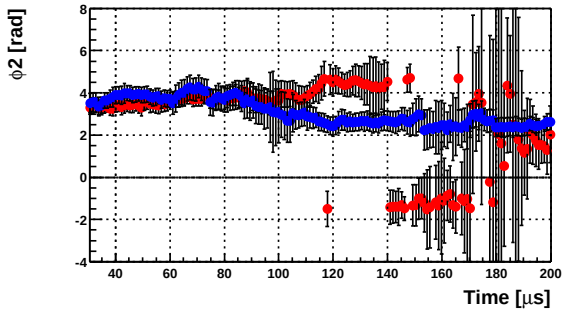

CBO Frequency [MHz]
Time [μs]


CBO Lifetime [μs]
Time [μs]

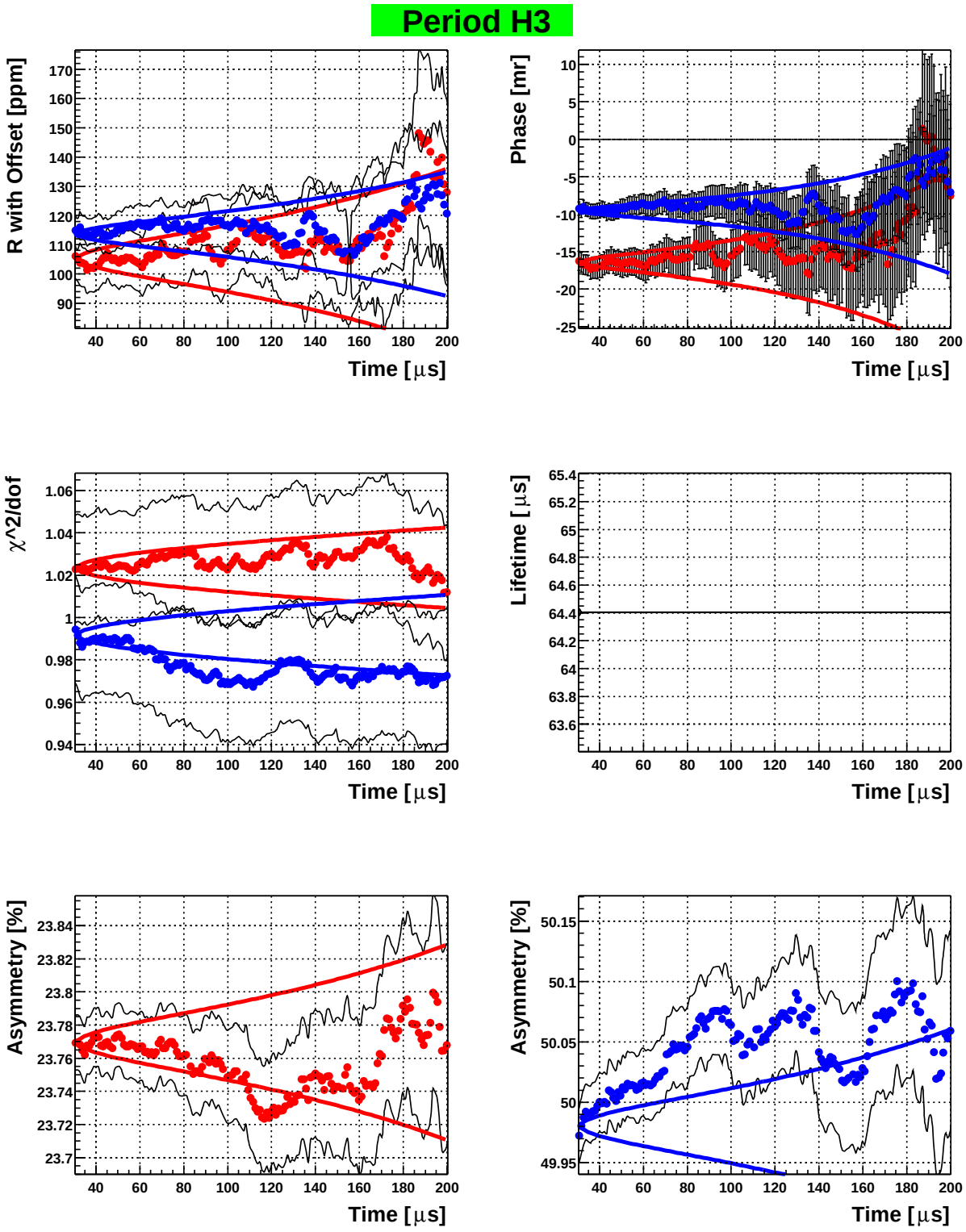

A1 [%]
Time [μs]


ϕ1 [rad]
Time [μs]

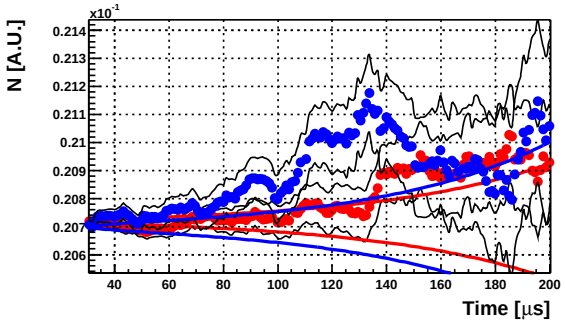

A2 [%]
Time [μs]

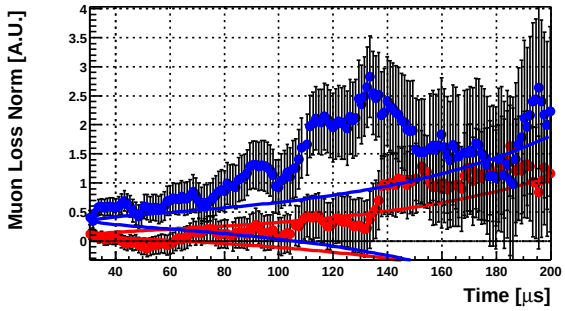

ϕ2 [rad]
Time [μs]

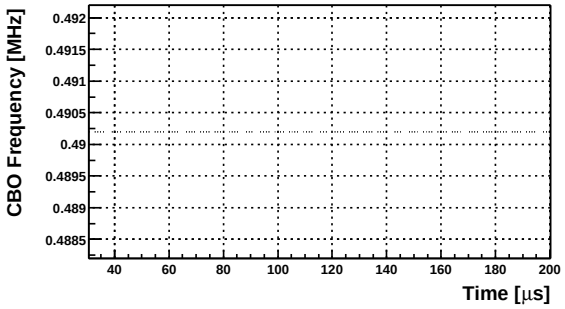
H.10 Period H3

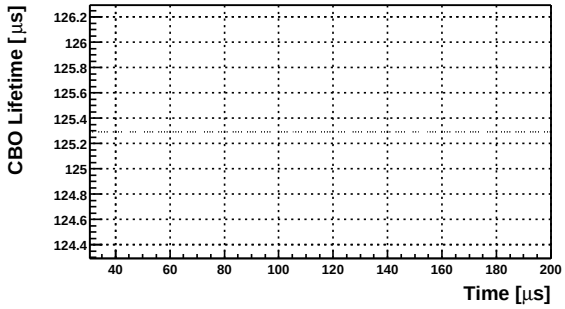


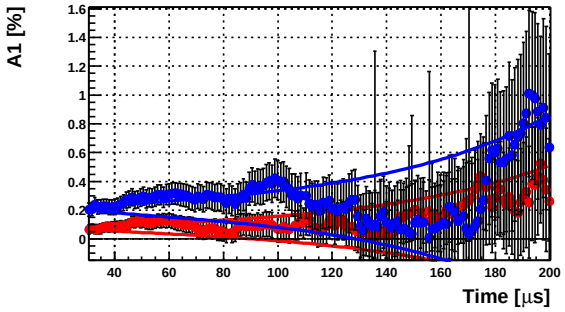
Period H3

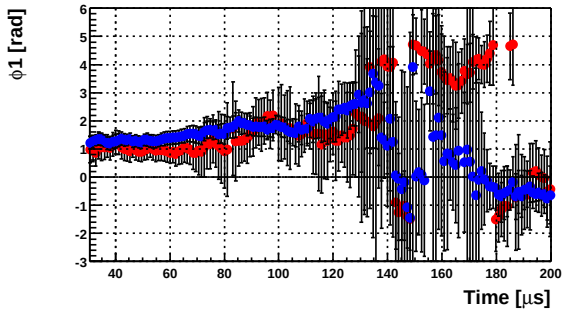

N [A.U.]
x10^-1
Time [μs]

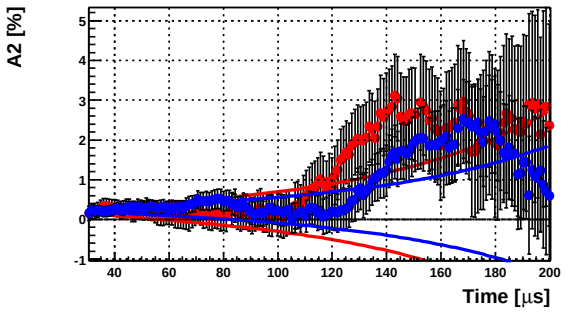

Muon Loss Norm [A.U.]
Time [μs]

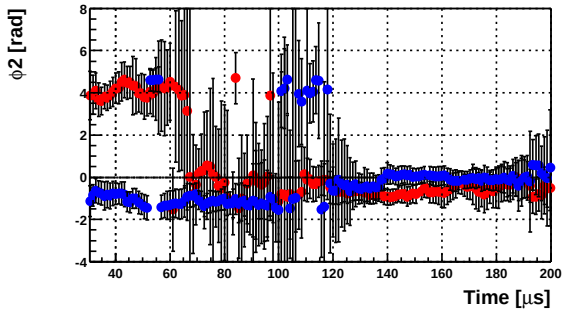

CBO Frequency [MHz]
Time [μs]


CBO Lifetime [μs]
Time [μs]

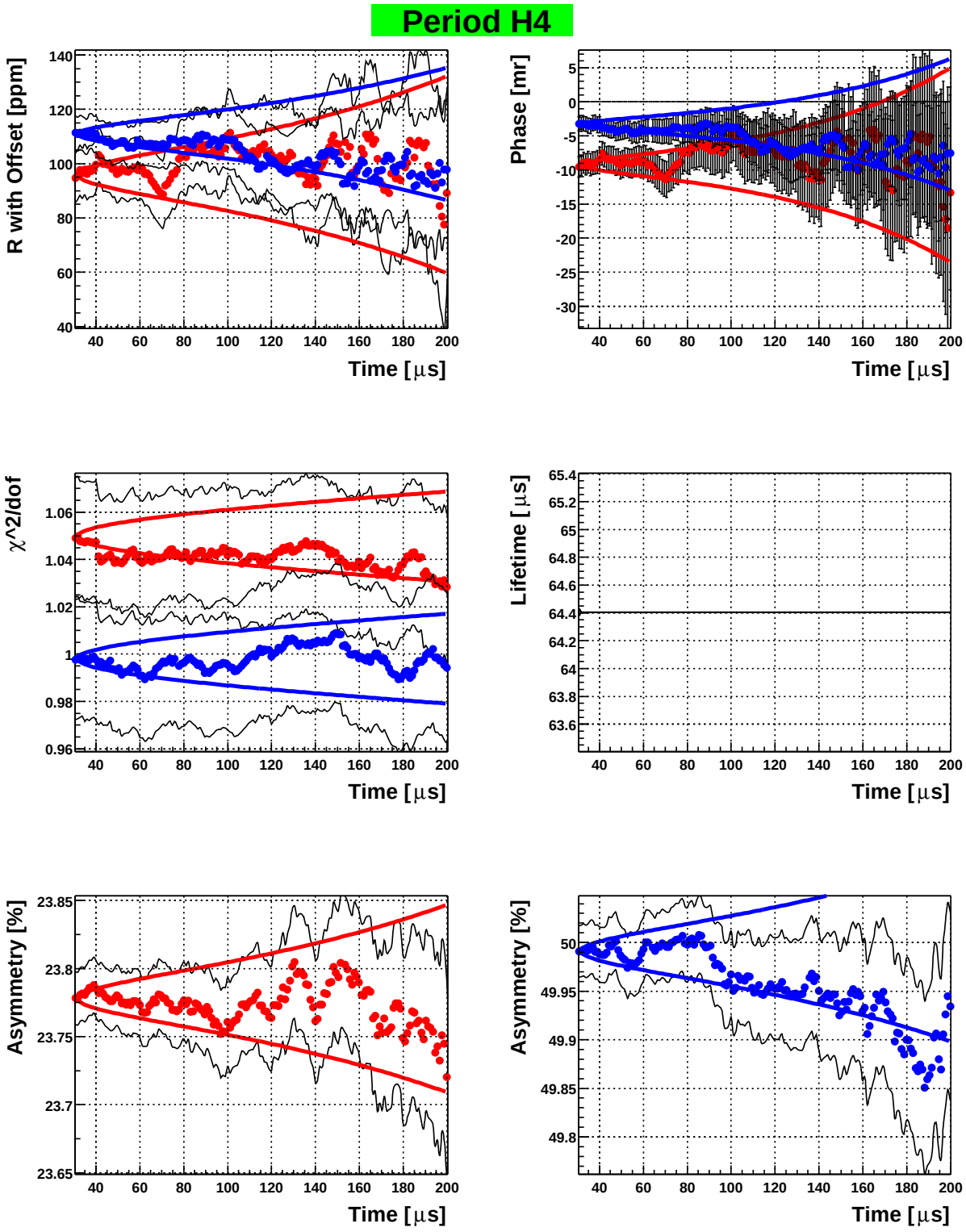

A1 [%]
Time [μs]


ϕ1 [rad]
Time [μs]

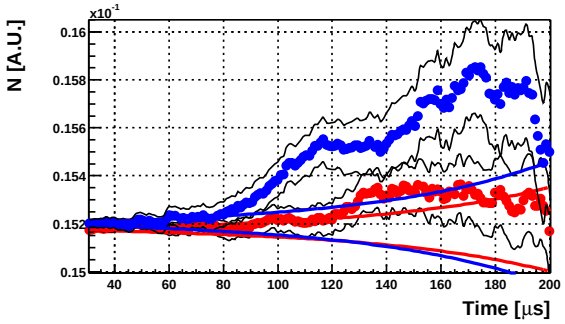

A2 [%]
Time [μs]

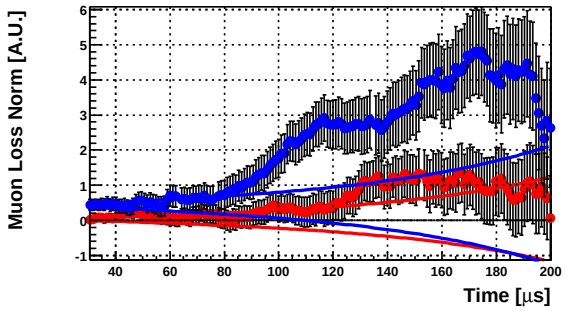

ϕ2 [rad]
Time [μs]

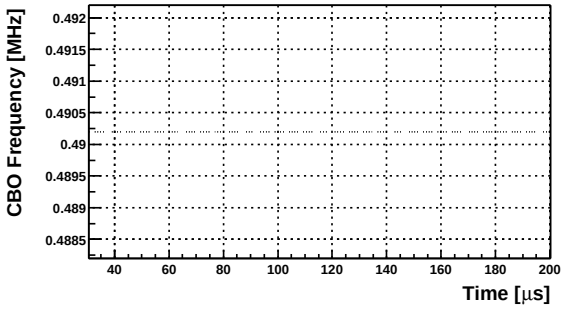
H.11 Period H4

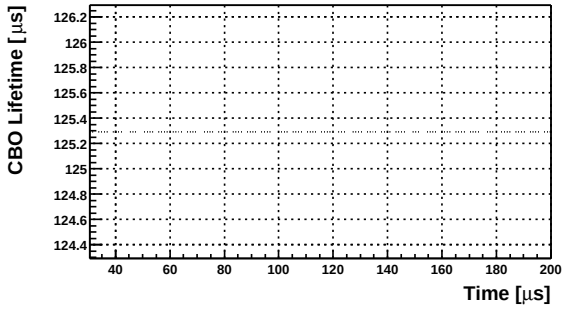


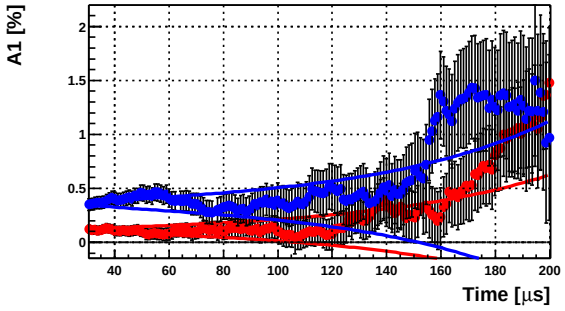
Period H4

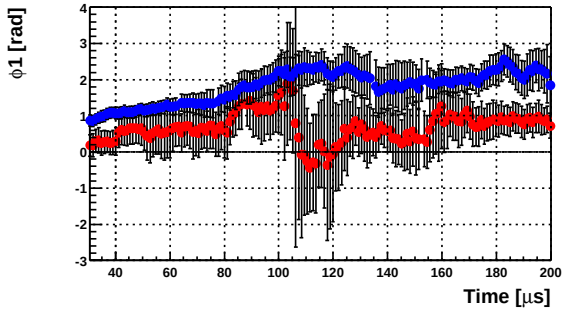

N [A.U.]
x10^-1
Time [μs]

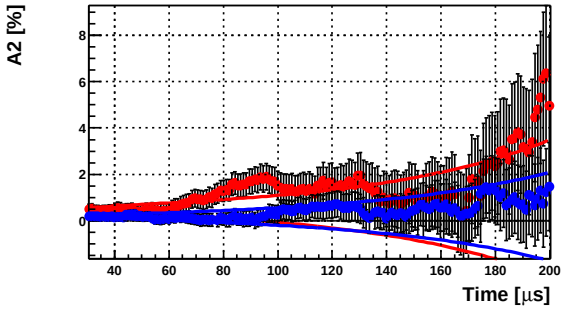

Muon Loss Norm [A.U.]
Time [μs]

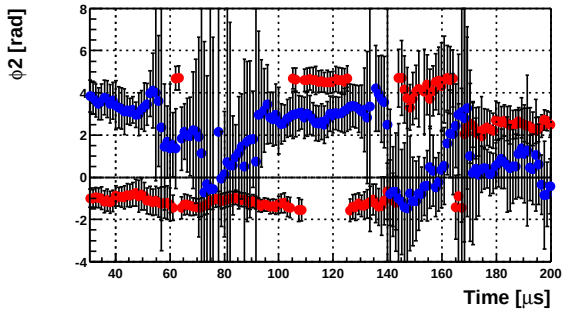

CBO Frequency [MHz]
Time [μs]


CBO Lifetime [μs]
Time [μs]


A1 [%]
Time [μs]


ϕ1 [rad]
Time [μs]


A2 [%]
Time [μs]


ϕ2 [rad]
Time [μs]

References

[1] F. Gray and G. Onderwater, Technical Report No. 417 (unpublished).

[2] J. Miller, g-2 internal note 68, E821 Collaboration (unpublished).

[3] J. Miller, g-2 internal note 69, E821 Collaboration (unpublished).

[4] A. Kolmogorov, *Foundations of the Theory of Probability* (Chelsea Publishing Company, New York, 1950).

[5] R. von Mises, in *Mathematical Theory of Probability and Statistics* (Academic Press, New York, 1964), Chap. IX, pp. 490–492.

[6] M. G. Kendall and A. Stuart, in *The Advanced Theory of Statistics* (Griffin, London, 1979), Chap. 30, pp. 468–477.

[7] A. Frodesen, O. Skjeggestad, and H. Tofte, in *Probability and Statistics in Particle Physics* (Oxford University Press, Oxford, 1979), Chap. 14, pp. 617–622.

[8] M. Hollander and D. Wolfe, in *Nonparametric Statistical References* (Jon Wiley and Sons, New York, 1973), Chap. 10, pp. 219–228.

[9] G. Fasano and A. Franceschini, Mon. Not. R. astr. Soc. **225**, 155 (1987).

[10] E. Gosset, Astron. Astrophys. **188**, 258 (1987).

[11] R. Smallwood, Phys. Med. Biol. **41**, 125 (1996).