Ms. Ref. No.: NIMA-D-07-00205
Title: G^0 Electronics and Data Acquisition
(Forward-Angle Measurements)
Nuclear Inst. and Methods in Physics Research, A

Dear Prof. Dr. Klanner,

Could you please address our answers and comments to the reviewer?
Thank you very much,
Best regards,
D. Marchand.

We thank the referee for a careful reading of our paper and for a number of useful suggestions and questions. We have identified several points which we had not adequately explained and we have made several corrections to the paper to make it clearer.

We apologize for the fact that the version sent to the referee didn't include the last figure (fig. 17) which shows the elastic asymmetry obtained for each detector using the North American and the French apparatus. Unfortunately we were not aware that this figure was missing until after we received your comments. This resulted from an error in the path of this figure within the Latex file which we submitted. Also several figures, which looked fine in our draft of the paper, suffered distortions in size or resolution due to the submission process. We apologize for the distraction caused by these problems and will try to ensure that all figures appear correctly in the resubmitted version of the paper.

Please find below our answers and comments corresponding to the different points you mentioned.

1) The paper describes a setup which due to the unfortunate choice of a DAQ system only supporting a limit number of crates and no pipe line electronics was much more involved than modern systems would have been.

The G0 experiment (E91-017) was first proposed at JLab in 1991. At that time the North American electronics scheme was proposed and financed. Also in this time frame, the choice of FastBus monitor electronics was made. More modern alternatives might have streamlined the monitor electronics. But that was only peripheral part of the experiment, used to monitor performance of the detectors and to determine small corrections to the deadtime of the main time-encoding electronics. For this purpose the FastBus electronics was adequate.

2) Using pipeline TDCs like the one developed for compass together with pipeline ADCs from Caen would have simplified things and would have allowed event by event information of all the relevant variables.
The instantaneous event rate is high, but the the event size of the experiment, due to the limited number of channels is very small. So also from this side no problem would have prevented an event by event readout.

Thanks to your comment we realize that we had not clearly explained the rates involved. We have tried to better specify the high rate we have to process. For example, in the abstract, we have replaced "to handle events a mean rate of 2 MHz per detector" by "to handle events from the 128 detector pairs at a mean rate of 2 MHz per detector pair …" . The introduction now explicitly states that the event rate would be over 250 MHz when explaining why we did not use event-by-event readout.

Also, when the monitoring electronics is introduced at the end of the Apparatus section we now explain that the data rate would be roughly 7 TeraBytes per hour from even a minimal event-by-event readout. (This corresponds to reading 2 Bytes of ADC and 2 Bytes of TDC information from just two of the four PMTs in a detector pair, giving 8 Bytes per hit or 16 MB/s from each of the 128 detector pairs.) We hope the reviewer will agree that this is prohibitive, especially since this is all valid data so no second-level trigger reduction would be possible. While it might, in principle, be possible to capture such a data rate (by using essentially 128 parallel DAQ systems writing to 128 parallel data stores, for example), it would be pointless since such information would not add appreciably to the usefulness of the data for the parity measurement.

The number of detector pairs is now also specified at the beginning of the "Electronics" section (section 4).

We also added this argument in the "Electronics" section : "At a mean rate of 2 MHz per detector pair (128 pairs of detectors), it was impossible to store full event-by-event information (timing and amplitude of the PMT signals) which would have led to a data rate in excess of 2 GB per second, even if data were recorded from only one detector of each pair."

It may be worth mentioning that this experiment is unusual among parity-violation experiments in that any digitization of the signal was done. Typically, because of the huge data rates involved, such experiments measure only an analog integral of the detector output with no attempt to identify individual hits, let alone recording a time spectrum.

3) As the choice of the readout of the experiment can not be reversed and the paper is written in relatively clear way I suggest to it for publication in NIM after the specific points listed below are done.

We hope our answers and the proposed modifications of the text will meet your expectation and that our explanations above have make it clear that the readout system was not so ill-conceived as it may have appeared.

4) The quality of some figures is not adequate for publication, fig 4, the last 2 and some don't fit on a page if they are printed or looked to in acrobat, like fig. 13 and 14.

Unfortunately the system at NIM which generated the pdf file from the Latex file somehow didn't take into account the magnification factors associated with the figures. For the revised version, we have resized the figures when exporting them in eps format. We hope that will fix the problem. We will also try to understand why figure 4 suffered a terrible loss of resolution when processed by the NIM system.

5) Please describe all symbols on the plots, for example fig 6 the pink symbols are not explained. hollow should be open symbols.

Concerning figure 6 (now 7), while we are not entirely sure of the source of confusion, we have modified the symbols and the caption and added a key to reduce the possibility of misreading the symbols. Also the choice of symbols has been changed to more closely match figures 15 and 16 (now 16 & 17).
We have also tried to improve the quality of fig. 15 and 16 (now 16 & 17) and modified slightly their captions.

6) *Please specify in more detail why two different schemes for reading 8 scintillator arrays has been chosen. If the different systematics of the 2 concepts is really so important support this by plots. Because currently a naive reader could think no agreement in the collaboration could be achieved for one of the 2 readouts.*
In general a flat concept for electronics and readout is advisable for a detector.

The history of the development of two readout systems is complex and involves financial, sociological, and political factors. Much of the discussion would be quite out of place in a scientific article. While we chose not to recount the reasons for the development of two parallel systems, it is quite true that we found it to be highly advantageous to have two systems to check against each other. This has been given in the paper as an advantage of having two such systems and we stand by that explanation.

We can recount some of the history for the referee's information, but the referee will probably agree that it would be inappropriate to include this in the paper:
The experiment was originally approved in 1991 with the North American scheme for high-rate time encoding. Subsequently, because of budget restrictions, the experiment was de-scoped. It was decided to instrument only four of the eight octants. Procurement of scintillator, PMTs, and electronics were all scaled down to match this new plan. Then the French groups (Orsay and Grenoble) expressed an interest in joining the collaboration. Since they had funding available, they proposed to instrument the remaining four octants. While the electronics had been designed by Carnegie Mellon University to be “fast and dumb” (too dumb to make mistakes or to introduce helicity-correlated biases) the Orsay group had tremendous expertise in development of sophisticated electronics. Rather than simply build more of this very simple design, they chose to try to make improvements. Most notably their design gave a smaller bin size, 250 ps instead of 1 ns. Also, while the CMU design encoded only the most interesting 23 ns of the 32 ns time spectrum, the new design would also encode the part of the spectrum in which no events were expected. The trade-off was that the system was far more complex, with multiple levels of data-handling going on in DSPs which required careful programming and analysis.

Once the design of the French electronics was well advance, the production of the CMU design could have been aborted and the more sophisticated electronics could have been used to instrument all channels. That was not the course of action chosen for a number of reasons:

- The CMU design had already been prototyped and thoroughly tested for the kind of tiny effects which could introduce false asymmetries at the part-per-million level at which the experiment was sensitive. (A helicity-correlated shift of the edge of a time bin by a few femto-seconds, for example, would have been unacceptable.) The French design was still being finalized.
- With beam time approaching, the CMU design was ready for production while the French electronics was still in prototyping. In fact, as it turned out, the French design was ready in time. But use of the CMU design ensured that at least half the electronics would be ready in time for beam.

- The CMU design was too dumb to fail. Some collaborators worried that the French design was sufficiently complex that it could fail. In fact the French design eventually was shown to work very well, but the fear was not groundless. Under unusual beam conditions, it was found that the instantaneous data rate could overwhelm the DSPs, causing an over-write type of error which corrupted further data until the module was reset. If the newly-developed 32 ns beam had been 'spiky' enough, the French electronics might have failed in this way at an unacceptable level.
- Ultimately, as explained in the paper, it was considered to be an advantage to have two systems to compare against each other. Time encoding at these rates was unusual and doing so under conditions which required part-per-million precision was unprecedented. There was no obvious advantage for one system over the other. (The greater resolution was expected to have little impact because of the intrinsic time width of the signal.)

So it was decided that the safest route was to use both systems. In worst case, if one system failed completely, it would be far better to have half the data (as anticipated after de-scoping of the project) than none. In best case, being able to compare results taken with both systems would increase our confidence that there were no unexpected systematic errors. Fortunately, this was the case and the final figure of the paper (which was accidentally omitted in the first version submitted) shows that the two systems give compatible results with very similar size of error bars.

In the end, the use of two electronics designs protected us from a completely unanticipated systematic error. There was an unexpectedly large asymmetry (several hundred ppm) in the very low-current leakage of beam intended to go to the 2 other experimental halls. Because the leakage beam didn't have the 32ns time structure, this caused a large false asymmetry across our entire time spectrum. Because it resulted from beam leakage, it changed rapidly and unpredictably. This couldn't have been adequately corrected if we had only used the NA electronics readout. The fact that the French electronics measured the spectrum in a region where no data was expected tuned out to be key to correcting for this false asymmetry. The study of this asymmetry is discussed in detail in ref. [2], [14] and in the PhD thesis of Jianglai Liu, University of Maryland, 2006.

Although it is unusual (and we wouldn't necessarily recommend it to other experiments) the use of two very different electronics systems cost us nothing in terms of data quality, increased the certainty (in advance of the experiment) that we would have a usable data set, gave us a way to cross-check for unexpected systematic errors and gave us complementary sets of features to use to our advantage. As mentioned above, one such feature, which wouldn't have been available if we simply duplicated our original design, saved us from a completely unanticipated problem.

7) till page 27 more than ones the timing resolution is mentioned but never a specific number is given.

This was an oversight. We thank the referee for drawing it to our attention. Furthermore, we had not explained that the elastic peak has a large intrinsic spread in ToF because of the length of the target. This is now explained in the apparatus section.

In the subsection “Time-Encoding Electronics”, we have added the following sentence :”The requirements on the TEE bin size were based on the width of the elastic peaks which is are intrinsically one to three nanoseconds wide due to the target length and the spectrometer optics. To select the elastic peak, the NA electronics has a 1ns time bin whereas the French electronics with 250 ps time bins allows more detailed studies of the distribution of asymmetry in background regions. As discussed below, both systems had adequate time resolution and provided a similar accuracy on the resulting elastic asymmetries.”

8) give the reason why not even the divider of the pmt and the splitter concept have been chosen to be the same.

The PMTs were chosen and purchased at different times in different countries. By the time the French purchased their tubes, it had become apparent that we would have to run all the tubes at very low gain for them to survive throughout the experiment with such high sustained rates. Given that, it made sense for them to switch to a tube with fewer stages which might be expected to run with greater stability at such low gain. The 8-stage tubes weren’t compatible with the same bases so bases were designed from scratch, incorporating features which the French groups had found useful in the past (and, again, taking advantage of the technical expertise which was available to them).

The splitters, which were located in the electronics room, interfaced to different front-ends for the two sets of electronics. While the NA splitter boards had been build to connect to commercial CFD modules, the French boards had to supply signals in a different format to the input of the DMCH. The NA splitters had been built by hand by students at CMU using coaxial cables because it wasn't worth the effort to design a carefully impedance matched board to do a relatively simple task. For the engineers and technicians at the French national labs, it was easier to design and build an active splitter board than to build similar passive splitters. Since the French cabling was not compatible with the NA, the new splitter board couldn't replace the existing NA splitters and there was no reason to design a comparable splitter board to replace the passive modules which already existed.

The differences in the splitters were accidentally over-emphasized in the paper because they were described one after the other in the combined discussion of monitoring systems. This made the differences sound more significant than they are. While we think it is appropriate to mention the splitters in a description of the electronics, we have now moved the description of each splitter to the discussion of the individual electronics systems so the differences between them are not given undue emphasis.

We hope the referee will agree that these details of how and why the systems evolved as they did are not worth including in the paper, but the actual systems used should be accurately described.

9) What is the fraction of pile-up events. please specify your statement on page 33 concerning pile up in more detail. How was this simulated.

The wording of this discussion inadvertently made the conclusion sound more profound than was intended (and failed to get the intended idea across). We did not intend to imply that the CFD was immune to all pile-up effects, such as deadtime. The discussion has now been reworded to

make it clearer that it refers just to one effect which had been observed (and corrected) in the NA CFD's and which was found not to be a problem for the French CFD's.

The effect referred to is one in which a CFD can fail to detect an above-threshold pulse because its shape is distorted by a preceding below-threshold pulse. This is not a deadtime in the conventional sense because the amount of signal-loss does not depend on the rate of firing of the CFD, rather it depends on the rate of sub-threshold pulses which are, individually, invisible to the CFD. The mechanism of the effect is that the combined pulses act as a distorted pulse with the unusual property that the constant fraction is reached before the arming threshold is reached.

While this caused baffling behaviour of the NA CFDs in an early test run, when the delay and fraction of the CFD had been set to poorly chosen values, it was shown by Cadence/Spice simulations to not be a problem for the French CFDs.

To answer the referee's question, at 2 MHz/detector there was a roughly 6% pileup of detected events in a 32 ns beam pulse. The actual rate of pileup was slightly higher because the CFD threshold was near the pion peak for most detectors, so the singles rate in each CFD was somewhat higher than the (4-fold coincidence) event rate. Furthermore, it was these pions which contributed the sub-threshold events which caused the effect described above.

10) please avoid internal experimental slang as buddy features

We have removed the experimental slang concerning the buddy feature and now refer to them as: “paired detectors” used to measure the correlation between hits in one detector and deadtime in a similar detector.

11) please show comparisons between the timing resolution achieved by the fastbus chain with attenuation corrections and so on, compared to the 2 non event type readout concepts.

Since the intrinsic ToF width of the elastic peaks is much greater than the bin width of the French electronics, it dominates over the effect of the bin size there as well as in the FastBus. As a result, the time resolutions obtained by FastBus and the French electronics are virtually identical. The resolution obtained by the NA electronics is only slightly wider. (As a diagnostic, the FastBus and NA timing spectra have been compared and were found to be consistent.) This has a minor effect on setting the ToF cut to optimize signal-to-noise contributions. Given the fact that the intrinsic width of the peak is the relevant factor, we hope the referee will agree that there is little point in showing the ToF spectrum as measured by FastBus.

The use of Constant Fraction Discriminators removed most of the 'walk' which might otherwise shift the time of an event depending on the pulse size. Studies done using the FastBus data showed no significant correlation of the ToF of the elastic peak with pulse height, indicating that the CFDs had successfully eliminated most walk. Again, the event-by-event FastBus data did not give an improvement in resolution over the time-encoding electronics. (And, of course, the time encoding electronics gave several orders of magnitude more statistics, which is the critical issue for measurement of a parity-violating asymmetry.)

The time resolution of the detectors themselves was monitored using the fastbus electronics. For that, the time difference between the meantime of the front scintillator and back scintillator in a

pair was measured. This gave a much narrower timing peak, as expected, but is of no significance in the selection of elastic protons from background.

12) please show a figure like fig 9 also for the US readout scheme

This figure has been added. We agree that it is useful for the reader to be able to compare spectra from the two systems. (We chose the same detector position, number 8, for both sample spectra for this reason.) It is figure 6 at the end of section 5.2.2.

13) what is the background of inelastic protons and pions in the elastic proton sample due to the limited timing resolution.

Energy-loss measurements from the fastbus data confirm that the background is almost entirely protons, with negligible contributions from pions and gammas, as expected due to the time separation.

The background under the elastic peak due to inelastic protons was a significant focus of the analysis of this data, especially because it had a large asymmetry. Protons from hyperon decay, in particular had a pronounced effect because of the huge parity-violating asymmetry of the decays. But it is important to note that this contamination was intrinsic to the design of the spectrometer and was not introduced by the electronics. The contamination was just as large in the fastbus data as in the TEE data!

As indicated in the PRL article published by the G0 collaboration (95, 092001 (2005), reference [2] of our paper), the fraction of the background (inelastic protons) ranges from 6% to 20% over the first 16 bins in Q^2. The final (highest) Q^2 bins are exceptional and have background fractions of 40% and 78%. The errors (and systematic errors) assigned to those bins reflect the large background subtraction.

14) was the achieved timing resolution adequate for the physics needs.

Yes, (See point 7) the intrinsic spread in proton arrival times dominated.

15) present tables summarizing the different contributions to the dead time and the timing resolution.

Probably the question of timing resolution is moot, given the previous answer.

Figure 6 (now renumbered as figure 7) gives the breakdown of the deadtime into the part due to the time-encoding electronics and the part due to the front-end electronics. The former is designed to dominate (but to be perfectly deterministic, so it is easily corrected) for events which fire all 4 PMTs on a pair of detectors. The front-end electronics, by design, don't contribute to the deadtime for such events.

The second correction, due to the front-end electronics, results entirely from 'singles' events in which 1, 2 or 3 of the PMTs fire. In the NA electronics, that is dominated by the deadtime of the CFD. In the French electronics, as explained in 6.2.1, the deatime depends on whether one or both inputs fire on a MT.

16) why are plots 15 and 16 are not also existing for the US electronics.

Figure 6 was the equivalent of figure 16 except that the result was not expressed in %.
The vertical axis Figure 6 (now figure 7) has been changed to % to more nearly match figure 16 (now figure 17). The former figure 15 (now figure 16) was included mainly for pedagogical reasons. It may be easier for the reader to understand how deadtime is measured by changing beam current. The ratio of asymmetry to charge asymmetry, however, is the more relevant quantity since it measures the slope at 40 microamp rather than averaging over many different currents.

17) please include plots showing the stability of the electronics vs rate.

The stability of the yield vs rate is given in the figures discussed in point 16.
The stability of time-encoding vs. rate is too small to measure and so does not make an informative plot. For the French electronics, no variation in the time was seen up to 2 MHz counting rate, with an upper limit of 50 ps. For the NA electronics, the positions of the edges of the time bins were mapped as a function of rate from a few kHz to more than 30 MHz. No detectable shift was seen. An upper limit on the rate-sensitivity was set at a few ps per MHz.
A plot would show only scatter, with no discernable slope.