CLAS-NOTE 2002-016

Reliablility and Limitations of GSIM.

Matt Bellis
Rensselaer Polytechnic Institute

Abstract

One must have a detailed understanding of the angular dependance of the acceptance of CLAS for most measurements. The main tool used for acceptance corrections of CLAS data is GSIM. I have used an empirical method to determine the reliablility and limitations of GSIM for measuring the acceptance. I used data from the reaction $\gamma p \rightarrow p\pi^+\pi^-$. The data set comes from the g1c run period with a toroid current at half the nominal maximum value. The photon beam energy ranged from 0.5 - 2.4 GeV/c^2.

1 Empirical Acceptance Method

The analysis involves three final state particles. If I detect any two, the third can always be reconstructed from missing momentum, whether or not it has been detected in CLAS. This principle can be used to map out an acceptance for CLAS using either real world or simulated data. We can then compare these two acceptance functions.

For example, suppose I detect the π^+ and π^- in CLAS. I calculate the missing mass off the two pions and see if this corresponds to the mass of a proton. I then loop over the particles in the PART bank to see if CLAS was able to detect this proton.

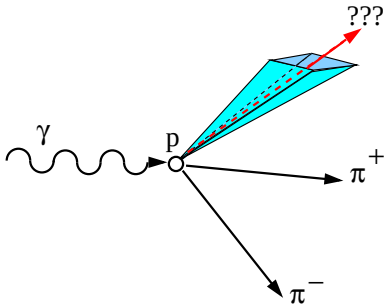

Figure 1: Representation of looking for a missing particle where I "expect" it to be based on my momentum reconstruction.

Using this information, I can now calculate some efficiency for detecting a particle based on it's momentum and angle.

$$\epsilon(\rho,\theta,\phi) = \frac{\#\text{ of times CLAS ``found'' the particle}}{\#\text{ of identified ``missing'' particles}}$$

where ρ, θ and ϕ are the spherical components of the momentum. Note that this "efficiency" combines both acceptance and detector response and *can be determined separately for either real or simulated data.*.

The next step is to determine the size of the mass and momentum "window" (the shaded volume in Fig.1) I want to look in to see if I see the particle. If the simulation and real data exhibit resolutions that differ significantly, it is important to remove the effect of this on our comparison of values of ϵ. This procedure and the results are detailed in Appendix A.

Now I can determine the efficiency as a function of some kinematic variable. For example, I look at the π^+. I choose some narrow momentum bin, and I integrate over all ϕ. I can then plot my "efficiency" over $\cos(\theta)$. (See Fig.2.) We found that our comparisons are sensitive to the differences in the real world momentum distributions, and the phase space that we run through the simulation. For this reason it is important that we bin finely in both momentum and $\cos(\theta)$.

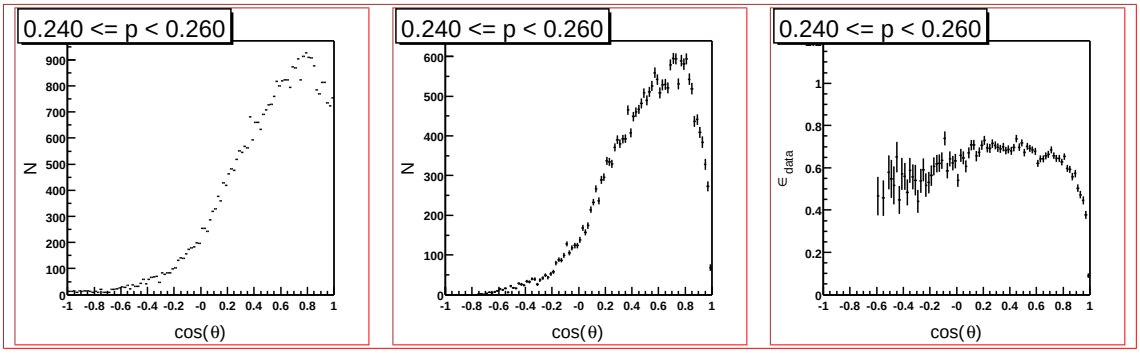

Figure 2: The first plot shows the distribution of π^+ lab angle $\cos(\theta)$ as reconstructed from missing mass for the momentum bin $0.24 < p < 0.26$ GeV. The fine binning is important for obtaining a valid result. The second plot shows the distribtion of π^+ where they are detected in CLAS. The third plot is the second divided by the first plot. This gives us our efficiency. This particular example uses real world data.

2 Characteristics of the Data and Tuning GSIM.

The Monte Carlo data was processed as follows. Three-body phase space events were generated using the TGenPhaseSpace class found in the ROOT package. They were generated roughly uniformly over the beam energy range 0.5-2.5 GeV/c^2. Following that, they were run through `gsim`, `gpp` and `a1c` with the `ffread` file and flags found in Appendix B.

After this there were still some obvious discrepancies between the data and the Monte Carlo.

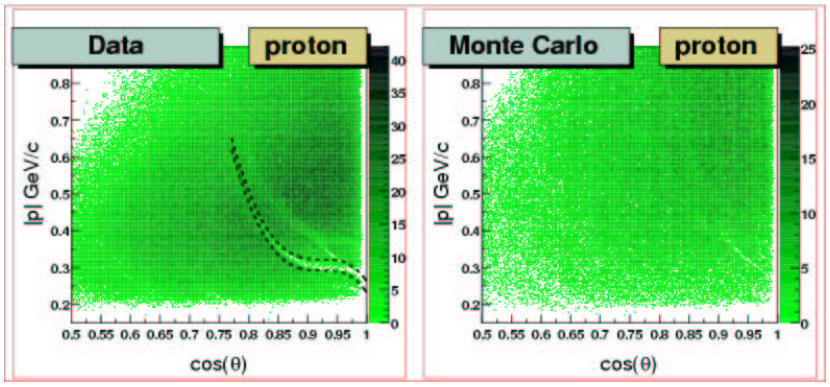

Figure 3: Plots showing the effect of the support bar in the Cerenkov detector on the distribution of protons. The dashed line in the data plot outlines the region that will be cut out of both real world and simulated data.

When the Cerenkov detectors were installed, there was a support bar used that was not put into GSIM at the time of this document. Protons that strike this bar are sometimes scattered or absorbed and are not

detected. If I plot momentum vs. $\cos(\theta)$ for the protons I can see a depletion region which is not visible in the Monte Carlo. (See Fig.3.)

At some point in the analysis, I apply energy loss corrections to account for the particles passing through the target walls. If I plot the momentum after this correction is applied, I will not see these depletion regions very clearly. These are a function of the momentum of the particle after it has passed out of the target and entered CLAS.

The other depletion structures are due to the carbon rods which support the main torus, or junctures between two time-of-flight walls and seem to be modeled properly in GSIM. I impose a momentum cut to remove the Cerenkov support bar depletion region in both my data and Monte Carlo. I determined the cut by taking some points in this plot and fitting them to a curve. Note that this cut is dependent on target position and magnetic field strength, so a different function may have to be used for different runs.

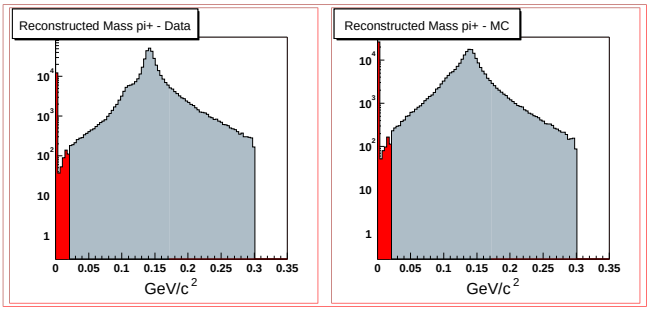

Figure 4: These plots show the mass of π^+'s as calculated from time-of-flight measurements. The first plot is real data and the second plot is Monte Carlo. Both are plotted on a log scale to better see "electrons" at the low end of the plot. The sharp cutoff at 0.3 GeV is imposed by CLAS software's particle identification routine.

Now if I look at the time-of-flight mass of the pions, I see what looks to be misidentified electrons. (See Fig.4.)I want to cut them out of both MC and data.

I also want to make sure that I have knocked out dead paddles in the Monte Carlo. After I have run it through GPP I can compare the distribution of hits for a particle in a particular sector over all the TOF paddles. If I look at Fig.5 I can see that paddle #44 is dead in the data, but has not been properly accounted for in the Monte Carlo.

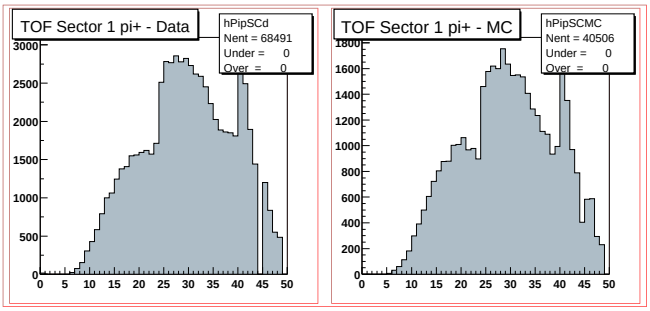

Figure 5: Hits in the time-of-flight paddles for π^+'s in sector 1. The first plot is real data and the second plot is Monte Carlo.

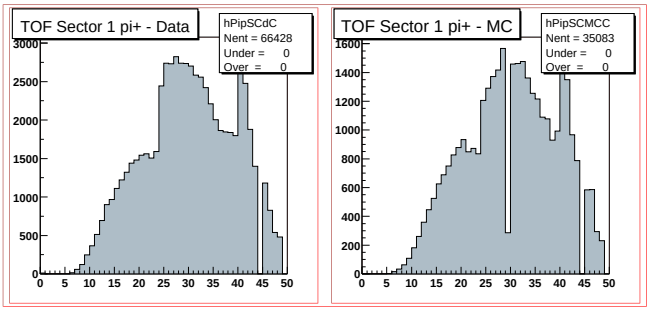

Figure 6: Hits in the time-of-flight paddles for π^+'s in sector 1 after I have cut out misidentified electrons. The first plot is real data and the second plot is Monte Carlo.

Fig.6 shows this distribution after I have cut out what I called "misidentified" electrons. Paddle #44 shows no reconstructed pions now in the Monte Carlo, but paddle #29 has been depleted. The Monte Carlo data has been filled with a status for each TOF paddle. But it does not always act the same as data when it comes to reconstruction the time and so I get more "electrons" in the Monte Carlo distribution. My solution was to impose a strict requirement on the health of the TOF paddles. First I only take hits that have a completely healty status. I also go through and identify what I will call "bad" paddles in the data and knock them out in my analysis code in both data and Monte Carlo. The procedure that I use is discussed in Appendix C.

There may also be slight discrepancies in the drift chamber positions between the real CLAS and in GSIM. To deal with this I cut in very slightly in the forward region and in phi. I chose three planes from the TBLA bank which correspond to each of the Regions. Then I look at where the hit position is in the sector coordinate system. These cuts and their results are detailed in Appendix C.

3 Comparison

The data was binned in 20 MeV/c momentum bins and 0.02 bins in $\cos(\theta)$. I integrated over all ϕ. When I'm able to reconstruct a particle from missing mass, I'll refer to it as a "reconstructed" particle. When CLAS detects the particle, I'll refer to it as a "found" particle.

To calculate the error on the efficiency, I assumed a binomial distribution for the "found" particles and used that to give me the error for them.

$$\begin{aligned}
\sigma^2_{found} &= N_{\text{found}} * p_{\text{found}} * p_{\text{missed}} && (1)\\
&= N_{\text{reconstructed}} * \frac{N_{\text{found}}}{N_{\text{reconstructed}}} * (1 - \frac{N_{\text{missed}}}{N_{\text{reconstructed}}}) && (2)\\
&= N_{\text{found}} * (1 - \frac{N_{\text{missed}}}{N_{\text{reconstructed}}}) && (3)
\end{aligned}$$

I also assume zero error on the reconstructed particles. The error on the efficiency then just comes out to be

$$\sigma_{\text{efficiency}} = \frac{\sigma_{\text{missed}}}{N_{\text{reconstructed}}}$$

I require that there be at least 30 reconstructed particles in a bin to calcuate an efficiency. I can now look at the data (black) and Monte Carlo (red) overlaid on the same plot, or as a ratio. (See Fig.7.)

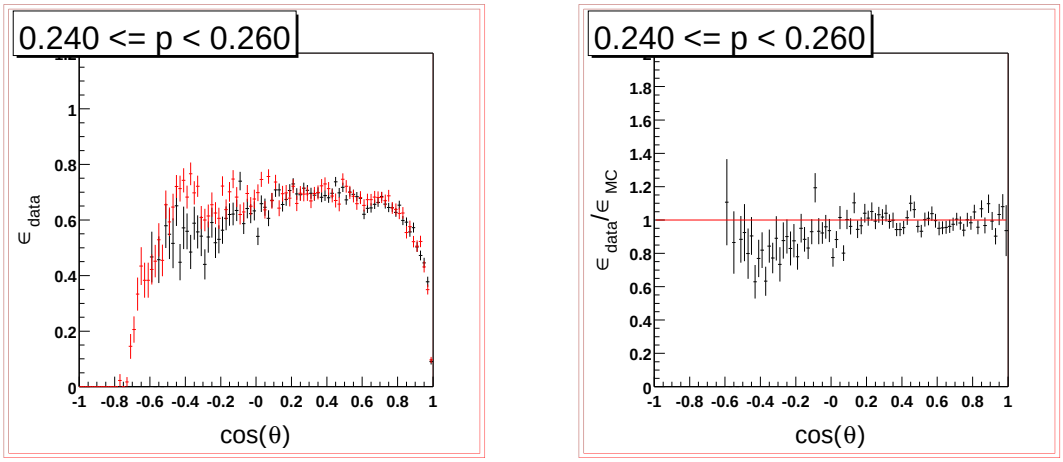

Figure 7: The first plot shows the comparison in efficiency for real-world data (black) and Monte Carlo data (red). The second plot shows the ratio of the two data sets.

There is a huge amount of data, so I needed some way to determine the validity of my cuts. I used a χ^2 per degree of freedom to see how much my ratio deviates from 1. I calculate this for each momentum bin where the degrees of freedom are the number of $\cos(\theta)$ bins that have values for both data and Monte Carlo.

$$\chi^2 = \frac{(1-R)^2}{\sigma_R^2}$$

$$R = \frac{\epsilon_{\text{data}}}{\epsilon_{\text{MonteCarlo}}}$$

I can look at this χ^2 for each momentum bin and see how the cuts improve the agreement. (See Fig.8-10.)

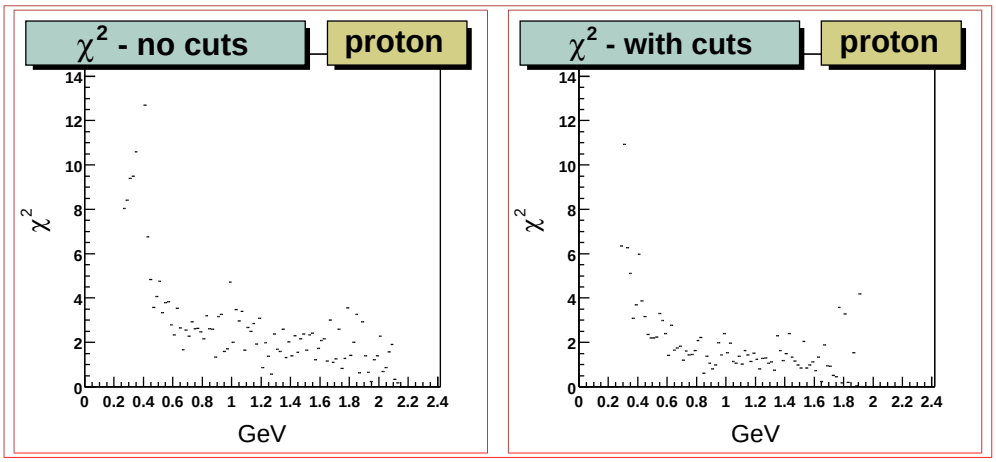

Figure 8: χ^2 per degree of freedom for the proton and pions before and after the cuts.

There is a separate document which contains plots of all the momentum bins generated for comparison in this analysis.

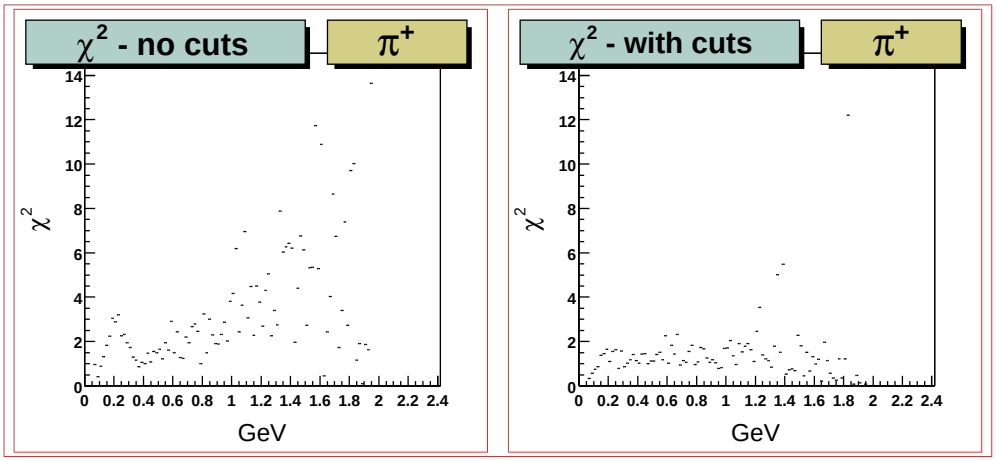

Figure 9: χ^2 per degree of freedom for the proton and pions before and after the cuts.

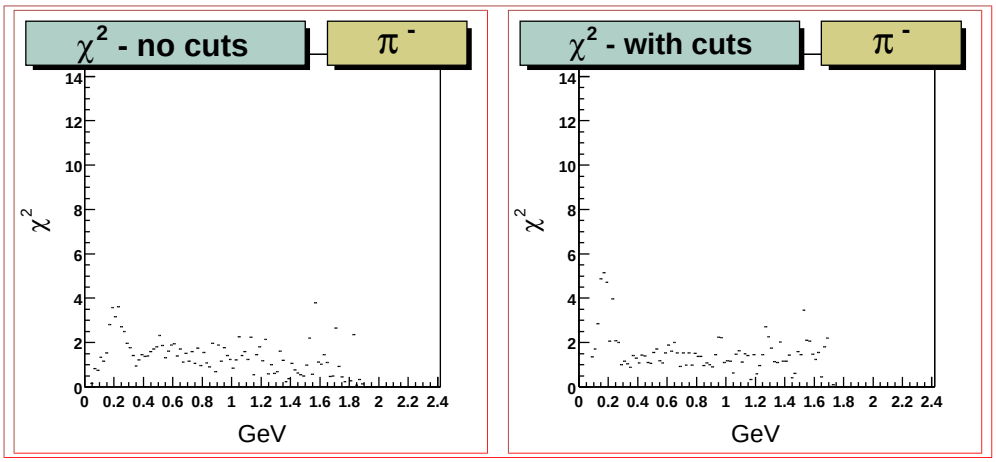

Figure 10: χ^2 per degree of freedom for the proton and pions before and after the cuts.

4 Conclusions

To summarize, here is the procedure we used to process and compare the data.

- Monte Carlo
  - `gsim`
  - `gpp`
  - `a1c`
- CLAS data and Monte Carlo
  - Energy loss and momentum corrections.
  - Vertex cuts.
  - Misidentified electron cuts.
  - Bad TOF paddles knocked out.
  - Region in proton momentum distribution affected by the Cerenkov support bar is knocked out.

– Fiducial cuts applied to track position in the drift chambers.

The combination of all the cuts seem to improve the agreement in the empirical efficiency calculation. In the course of the study it was noted that the bulk of the improvement in the middle $\cos(\theta)$ regions comes from the TOF paddle knock outs. The fiducial cuts seem to primarily improve the agreement for forward going particles.

Using the χ^2 plots as a guide, we decided to perform only low momentum cuts. We are somewhat motivated by the fact that the majority of the data lies in the momentum ranges where there is the best agreement between the data and Monte Carlo. We wind up cutting out protons below 400 MeV/c and π^+ and π^- below 300 MeV/c.

A Momentum window determination

To determinte the momentum window where I look to see if a particle has been identified, I look at events where all three particles have been detected. First for each particle A, calculate the missing mass off of particles B and C. Using this I set the following limits on MM^2. (See Fig.11.)

	low (GeV^2/c^4)	high (GeV^2/c^4)
proton	0.65	1.10
$\pi^{\pm}$	-0.04	0.08

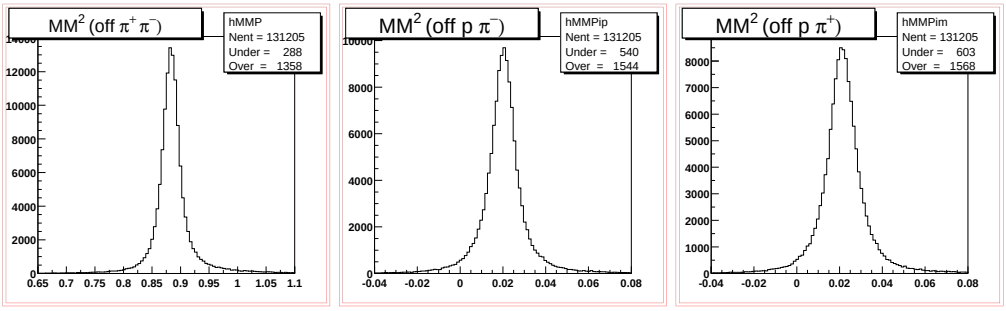

Figure 11: The first plot is the missing mass off the π^+ and π^- when all three particles have been identified exclusively. The second is the missing mass off the proton and π^-, and the third is the missing mass off the proton and π^+

I want to do the same for the spherical components of the momentum as well. I take the difference between the missing component off of A and B and the same component of C. I get the following limits for the particles. (See Fig.12-14.)

proton Δp_r	$\pm$ 0.1 (GeV/c)
proton Δp_θ	$\pm$ 0.2 rad
proton Δp_ϕ	$\pm$ 0.4 rad
$\pi^{\pm}$ Δp_r	$\pm$ 0.1 (GeV/c)
$\pi^{\pm}$ Δp_θ	$\pm$ 0.2 rad
$\pi^{\pm}$ Δp_ϕ	$\pm$ 0.4 rad

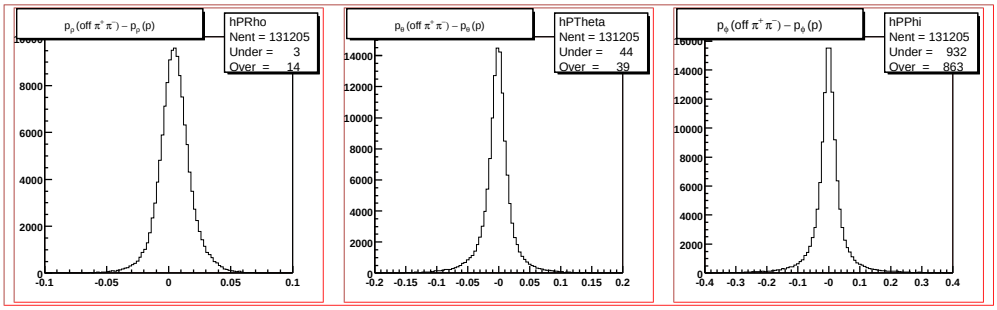

Figure 12: Difference between missing momentum off the π^+ and π^- and the measured momentum of the proton. Units are GeV and radians.

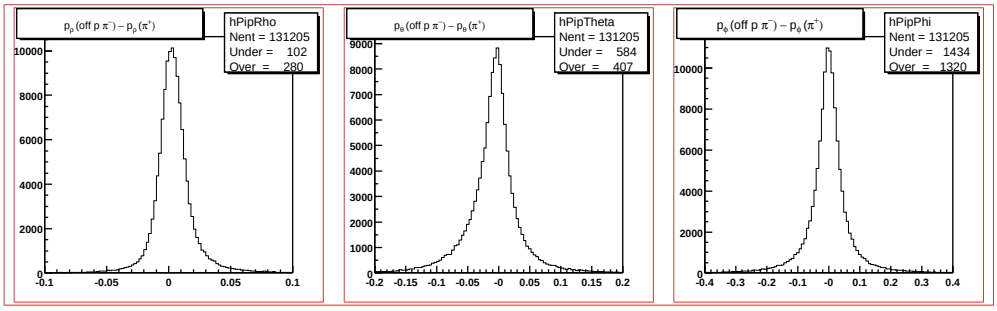

Figure 13: Difference between missing momentum off the proton and π^- and the measured momentum of the π^+. Units are GeV and radians.

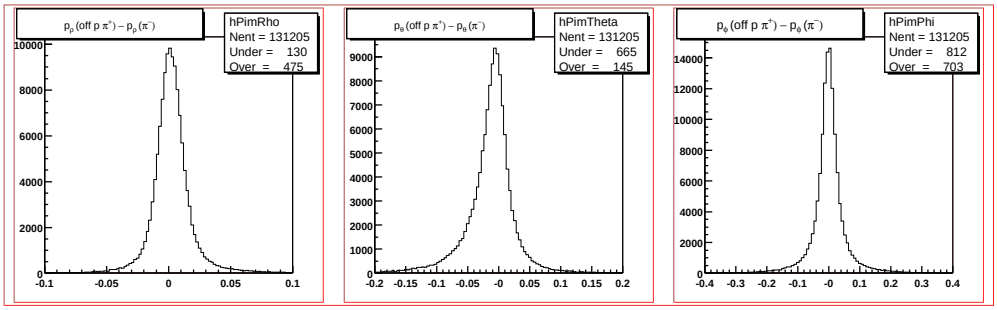

Figure 14: Difference between missing momentum off the proton and π^+ and the measured momentum of the π^-. Units are GeV and radians.

B Process flags

B.1 GSIM flags

My `ffread.in` file is shown in Table 1

AUTO	1				
CUTELE	1.e-4	1.e-4	1.e-4	1.e-4	1.e-4
CUTGAM	1.e-4	1.e-4	1.e-4	1.e-4	1.e-4
CUTHAD	1.e-4	1.e-4	1.e-4	1.e-4	1.e-4
CUTMUON	1.e-4	1.e-4	1.e-4	1.e-4	1.e-4
CUTNEU	1.e-4	1.e-4	1.e-4	1.e-4	1.e-4
MAGTYPE	2				
MAGSCALE	0.4977	0.0			
SIGBEAM	-1				
BEAM	2.445				
TARGET	'g1c'				
NOGEOM	'MINI'	'PTG'			
NOSEC	'OTHE'	'FOIL'			
RUNG	21785				
GEOM	'ST'				
STOP					

Table 1: The ffread file I use to process the Monte Carlo data.

I run with this on the command line
`-KINE 1`

B.2 GPP flags

`-R21785 -f0.75 -a1.8 -b1.8 -c1.8`

B.3 A1C flags

Note that the tagged version of A1C was used for cooking the Monte Carlo data.

```
-v -F -T1 -E2445 -ct1921.37 -cm0 -X0 -R21785 -D0x102d -P0x1fff
```

C Cuts

C.1 Vertex cuts

I made the vertex cuts on the z-component returned by the `mvrt` routine. The cuts are consistent with the target used in the g1c experiment. I include data consistent with

$$-9 < z < 9$$

C.2 Proton Cerenkov detector support bar cuts

To account for the Cerenkov support bar which is not properly modeled in GSIM I impose a momentum and angle cut on the protons. I exclude protons consistent with

$$p_{proton} < -25.76*\cos(\theta)^6+24.1*\cos(\theta)^5+22.03*\cos(\theta)^4-.8796*\cos(\theta)^3)-19.07*\cos(\theta)^2-15.77*\cos(\theta)+15.62$$

and

$$p_{proton} > -25.76*\cos(\theta)^6+24.1*\cos(\theta)^5+22.03*\cos(\theta)^4-.8796*\cos(\theta)^3)-19.07*\cos(\theta)^2-15.77*\cos(\theta)+15.57$$

and

$$p_{proton} < 0.60$$

C.3 Pion cuts

To cut out the misidentified "electrons" in the pion data I looked at the `qpid` entry in the PART bank. This is a mass reconstructed from time-of-flight measurements and is used in the `pid` routine for particle identification. I include data consistent with

$$\texttt{qpid} > .02$$

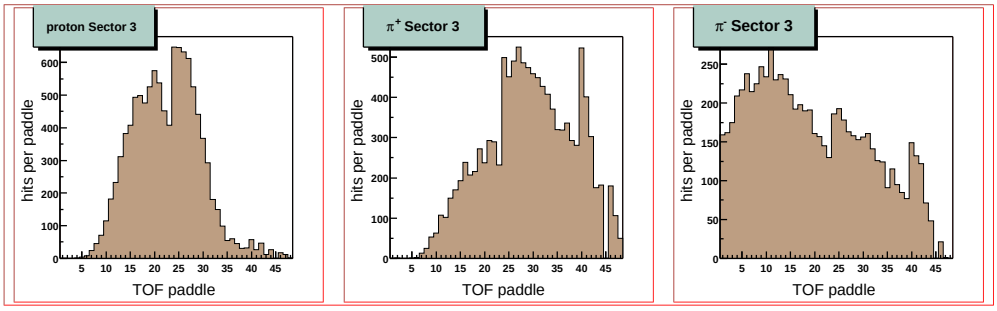

Figure 15: The hits in the time-of-flight paddle in Sector 3 for the proton, π^+ and π^-. In these plots there is no requirement on the status of the paddle. Real data is shown.

C.4 Bad time-of-flight paddle cuts

I want to determine a criteria for cutting out bad TOF paddles. I do this by looking at the percentage of times I get a non-status 15 for the paddles. Then I require this percentage to be higher or lower than some number for the proton, π^+ and π^-. Having both positive and negative particles in the test assures me that the depletion is not so dependant on the path the particles traverse.

I notice that for the first few paddles, the percentage of status 15 hits tends to be lower for the π^-. This may be a path dependance but I get very few hits in these paddles from the positively charged particles so it is difficult to say. For paddles 1-5, if the percentage of status 15 is below 80% for all three particles, I cut it out. For paddles 6-48 I require the percentage to be below 90% for it to be a bad paddle. (See Fig. 16.)

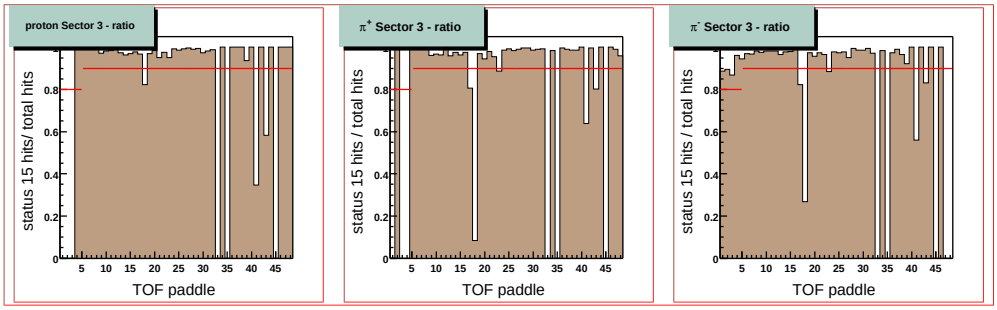

Figure 16: This is the ratio of hits when the counter shows a completely healthy status 15 to hits without any status requirement. The red lines are the levels that I consider to be a potentially bad TOF paddle. The ratio must be below this line for all three particles for me to label it as unreliable.

An explaination of what the TOF status indicates is shown in Table 2.

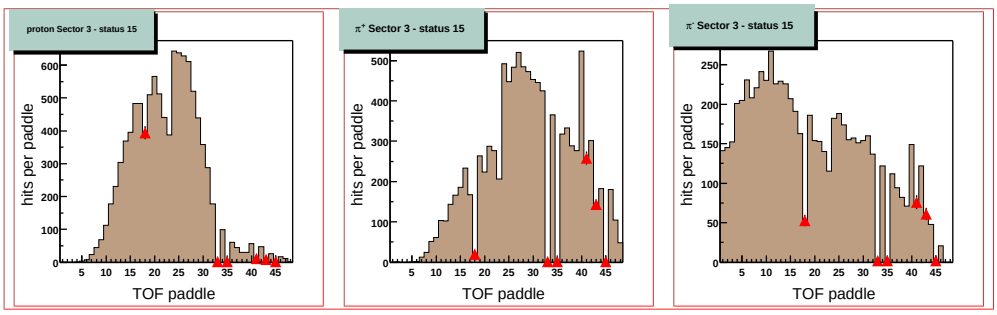

Figure 17: These are the hits when I require a status 15 for the paddle. The red markers indicate what I have determined to be bad paddles and will be cut out of subsequent analysis in both the data and Monte Carlo.

Value:	Mneumonic:	Meaning:
1	TDCL_GOOD	Only tdcl
2	ADCL_GOOD	Only adcl. Discarded.
3	LEFT_OK	tdcl,adcl
4	TDCR_GOOD	Only tdcr
5	TDCS_ONLY	tdcl,tdcr
6	ADCL_TDCR	Self-evident!
7	LEFT_OK_NO_ADCR	tdcl,adcl,tdcr
8	ADCR_GOOD	Only adcr. Discarded.
9	TDCL_ADCR	Self-evident!
10	ADCS_ONLY	adcl,adcr. Discarded.
11	LEFT_OK_NO_TDCR	tdcl,adcl,adcr
12	RIGHT_OK	tdcr,adcr
13	RIGHT_OK_NO_ADCL	tdcl,tdcr,adcr
14	RIGHT_OK_NO_TDCL	adcl,tdcr,adcr
15	BOTH_OK	tdcl,adcl,tdcr,adcr.

Table 2: Explanation of the status found in the SCRC bank.

C.5 Drift chamber volume cuts

I make cuts on where the particle hits in the drift chambers on the assumption that near the edges of the torus and in the forward region the magnetic field may be known with less accuracy. These cuts are given in either the Cartesian (x, y, z) or spherical (ρ, θ, ϕ) coordinates of the locations of the hits in the drift chambers as returned by the TBLA bank. I chose an arbitrary plane for each region in which to make my cuts.

To make my forward cut, I imagine some plane displaced from the beam position in the x-direction. To make the edge cuts I move in from $\phi = 60^o$ along $\cos(\theta)$ some distance. These cuts are listed below.

Region 1 - Plane 6

$$\rho * \sin(\theta) * (.5236 - |\phi|) > 5cm$$

$$x > 14$$

Region 2 - Plane 18

$$\rho * \sin(\theta) * (.5236 - |\phi|) > 20cm$$

$$x > 44$$

Region 3 - Plane 30

$$\rho * \sin(\theta) * (.5236 - |\phi|) > 22cm$$

$$x > 71$$

Figures 18-26 show the effect of the cuts. The red shows the hits that fall outside the fiducial cuts. Note that the cuts seem to fall along different location on the cuts when I look at data or Monte Carlo.

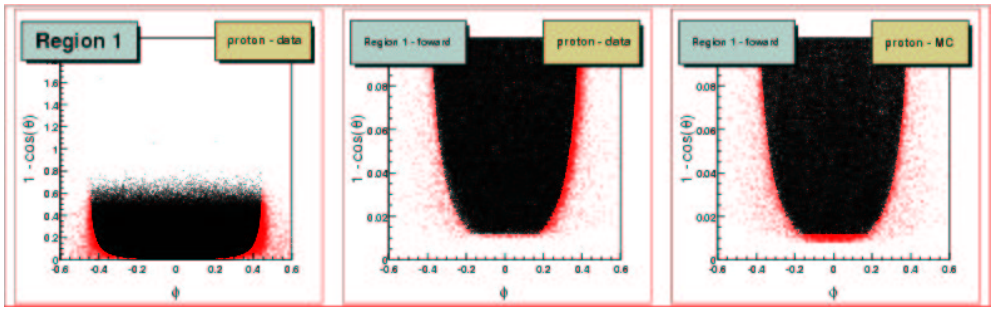

Figure 18: The hits for the protons in the Region 1 drift chamber.

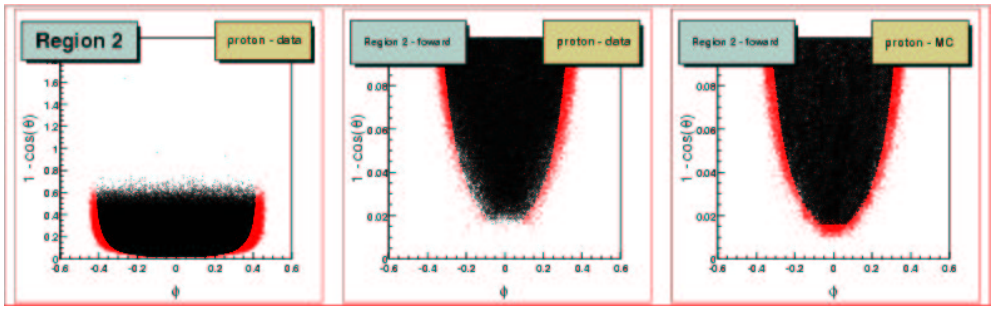

Figure 19: The hits for the protons in the Region 2 drift chamber.

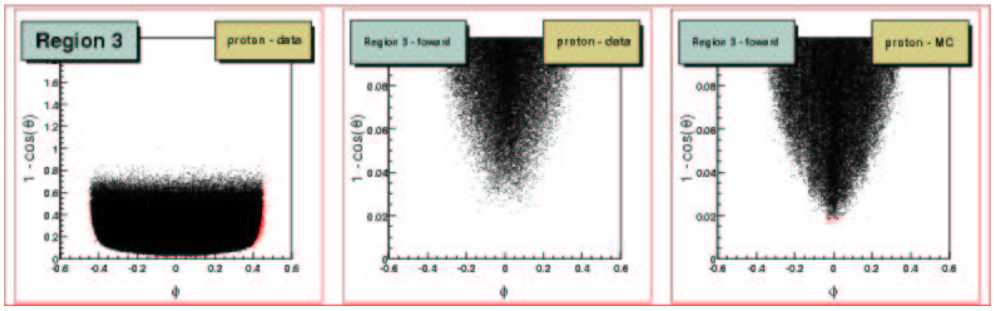

Figure 20: The hits for the protons in the Region 3 drift chamber.

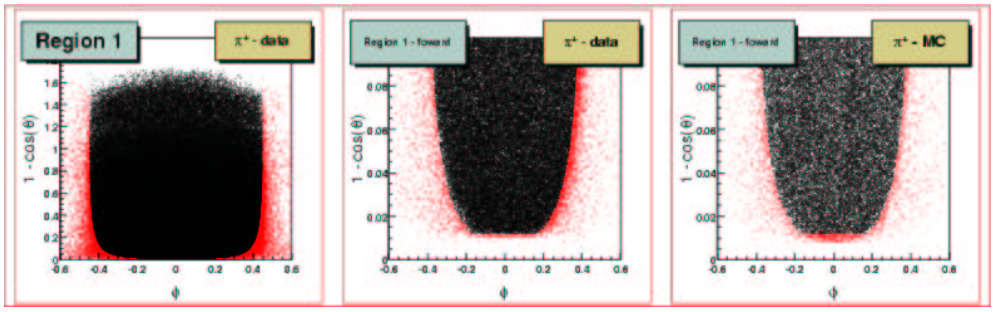

Figure 21: The hits for the π^+'s in the Region 1 drift chamber.

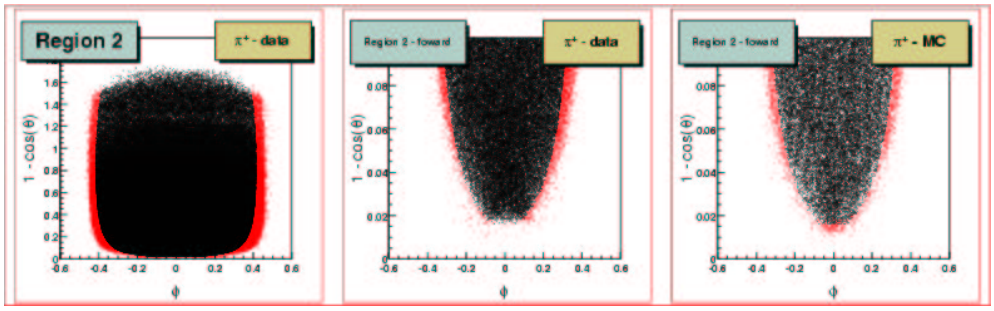

Figure 22: The hits for the π^+'s in the Region 2 drift chamber.

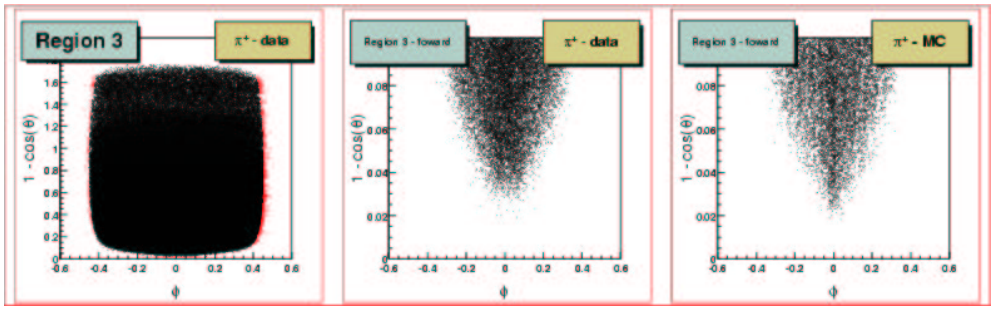

Figure 23: The hits for the π^+'s in the Region 3 drift chamber.

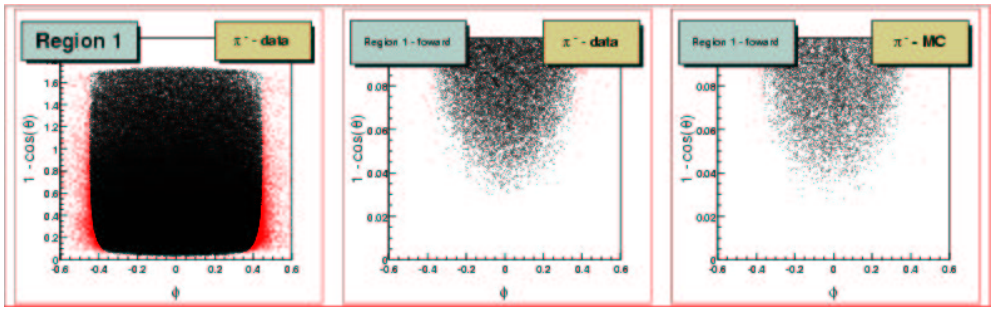

Figure 24: The hits for the π^-'s in the Region 1 drift chamber.

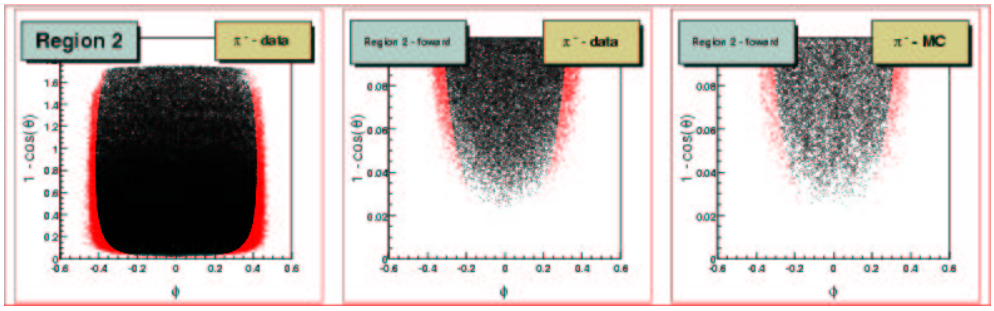

Figure 25: The hits for the π^-'s in the Region 2 drift chamber.

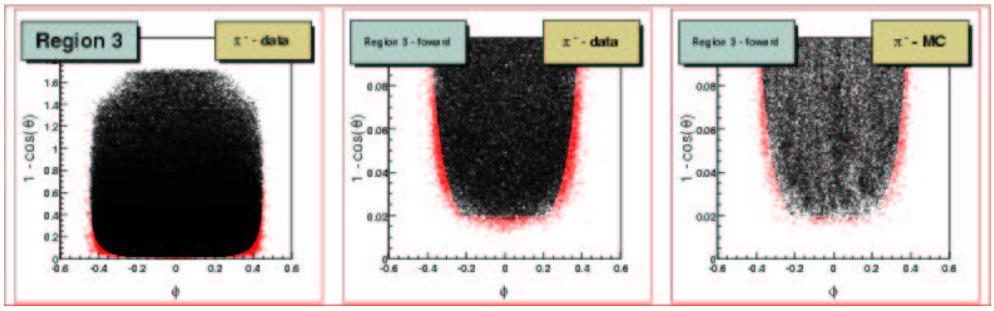

Figure 26: The hits for the π^-'s in the Region 3 drift chamber.