Assessment of Hall A Vertical Drift Chamber Analysis Software Performance Through Monte Carlo Simulation

Amy Orsborn

Office of Science, SULI Program

Case Western Reserve University

Thomas Jefferson National Accelerator Facility

Newport News, Virginia

August 5, 2005

Prepared in partial fulfillment of the requirements of the Office of Science, U.S. Department of Energy Science Undergraduate Laboratory Internship (SULI) Program under the direction of Jens-Ole Hansen in the Hall A division of the Nuclear Physics Program at the Thomas Jefferson National Accelerator Facility (TJNAF).

Participant: _____
Signature

Research Advisor: _____
Signature

Table of Contents

1 ABSTRACT

Assessment of Hall A VDC Analysis Software Performance through Monte Carlo Simulation. AMY ORSBORN (Case Western Reserve University, Cleveland, OH 44106) JENS-OLE HANSEN (Thomas Jefferson National Accelerator Facility, Newport News, VA, 23606).

The High Resolution Spectrometers (HRS) employed by Hall A at the Thomas Jefferson National Accelerator Facility (JLab) rely on vertical drift chambers (VDCs) for particle tracking. In order to reconstruct particle paths, data from these VDCs are analyzed using custom analysis software (‘the analyzer’) that builds upon the ROOT data analysis framework [1]. To test the software's abilities and find methods for improvement, it is useful to examine simulated data. In this paper, a Monte Carlo simulation of the VDC wire chambers was carried out. The simulation was based on an existing framework [2], which was expanded and improved as part of this project. Realistic effects were incorporated into the simulation to test the analyzer's resolution and tracking limits. Data from this simulation were analyzed, and the results were compared with the simulated data to assess analyzer performance. The effects of angle estimation in the track reconstruction algorithm, VDC drift time resolution, random wire firing, and coincident tracks on analyzer performance were examined. It was found that track angle estimation reduces error in track position and slope linear fits by approximately 10% with a drift time resolution of 4.5 ns. In simulations with a drift time resolution of 4.5 ns, a 0.5% probability of random wire firing, 99.999% wire efficiency, and angle estimation incorporated, the overall tracking efficiency remained above 90% for all triggering rates ranging from 2 kHz to 700 kHz. Tracking efficiency in reconstructed multi-track events was shown to be below 90% for all rates, falling below 80% for trigger rates above 500 kHz. The simulation may still be improved via the addition of more

natural effects such as delta and cosmic rays. Preliminary simulation results show a high level of confidence in trigger track reconstruction but significant failures in the analyzer's processing and identification of multiple tracks. Further analysis of analyzer performance with reconstructed and generated multi-track events is necessary to identify problems in the algorithm's treatment of multiple track events.

2 INTRODUCTION

Drift Chambers are devices commonly used for the tracking of particles. Jefferson Lab's (JLab) Hall A uses vertical drift chambers (VDCs) to determine the direction of the particles passing through their High Resolution Spectrometers (HRS). One VDC has two chambers with two orthogonal planes (u and v) of 368 sense wires in each. The chambers are filled with easily ionized gas. As charged particles travel through the chambers, gas molecules are ionized, and freed electrons drift through the chamber along the electric field lines formed between the electrically grounded sense wires and two cathode planes, which are held at –4kV and are located above and below the wires. As a freed electron moves towards the sense wire, it ionizes other gas molecules, causing an avalanche effect. These incoming electrons generate electric signals in the wires that are amplified and used as a start signal for time-to-digital Converters (TDCs). The TDCs measure the time interval between the first received wire signal and a common stop signal, which is the delayed trigger signal of the spectrometer generated from a coincidence between two scintillator planes above the VDCs. An event is defined as the occurrence of a trigger. The TDCs are capable of recording multiple start signals (multiple hits) before the common stop. More detailed descriptions of the Hall A VDCs and their components can be found in [3].

Hall A's custom analysis software uses information from the VDCs to reconstruct a particle's path. TDC times of activated wires are first organized into clusters within each plane. Cluster TDC times are converted into a drift-time—how long the particle took to drift from its originating location to the sense wire. These drift times are converted into a drift distance using a nonlinear, track angle dependent, conversion method described in detail in [4]. A linear fit of these drift distances is performed within each cluster, giving the approximate track angle and cross-over point (position at which the particle path intersects the wire plane). See figure 1 for a diagram of this process. The analyzer calculations are based in the detector coordinate system, which is centered in the middle of the VDC plane, with the x-and y-axes along the length and width of the chambers (respectively) and the z-axis along the vertical. The u and v wire coordinates are transformed into this detector coordinate system, as described in detail in [5].

Because the time-to-distance conversion depends upon the track angle, which is unknown, a track angle is first guessed. Thus, the process described above is iterated until the reconstructed angle converges.

The analyzer matches the u and v clusters of each plane to reconstruct tracks. If more than one cluster is found in each plane, the analyzer may reconstruct multiple tracks. Clusters are combined to achieve best agreement between the track angles fitted within each cluster and the overall track angle calculated from the cross-over positions in the two planes. The track reconstruction algorithm has been extensively documented; see [3], [4], and [5] for further detail.

VDCs are subject to many phenomena that affect the accuracy of track reconstruction. The TDCs are LeCroy model 1877 Fastbus modules with a time resolution of 0.5 ns/channel [6]. The limited TDC resolution limits the accuracy of the measured drift times. Similarly, the track

reconstruction algorithm requires an initial estimation of the track angle, described above, which introduces possible error.

Natural phenomena also influence the accuracy of track reconstruction. Drift times are subject to random fluctuations due to the random walk-like nature of ions traveling within the gas, which limit the time resolution of the VDCs [2]. Random wire firing unprompted by a particle track is also possible, creating noise within the chambers. Wires might also fail to fire when hit by a track. Furthermore, multiple tracks may enter the chambers within an event, potentially interfering with the tracking of the primary (trigger) track. The analyzer's performance during multiple track events is of particular interest, because it will highlight possible flaws in the tracking algorithm, such as improper determination of the trigger track and mismatching of clusters.

Hall A's VDCs have been thoroughly tested using experimental data [2], [4]. However, the effects of the phenomena described above on tracking accuracy cannot easily be assessed using experimental input. Using a Monte Carlo simulation to create input data that incorporates the natural effects described above will thus facilitate a thorough analysis of the track reconstruction algorithm.

3 MATERIALS AND METHODS

3.1 Algorithm of Track Data Generation

A basic Monte Carlo simulation of randomized particle tracks was created by K. Rossato in August of 2004 [2]. This simulation generates a random origin and momentum (direction) of a particle. The generated distributions are chosen to approximate experimental data that cover the area of the VDCs uniformly. The geometry of the chambers is then used to calculate the index numbers of wires hit within each plane and the particle drift-distance for each hit wire. Using an

inverse of the time-to-distance conversion found in the analyzer, this distance is converted into a drift-time. These drift-times are converted into TDC times for each wire. Hit information—wire number and TDC time—are read into the analyzer using a specially developed decoder program. Details regarding this simulation may be found in [2].

The basic simulation [2], however, does not include the realistic effects described above. This existing framework was modified and expanded in this inquiry.

3.2 Simulation of Realistic Effects

The major realistic effects incorporated into the simulation include random wire failure (imperfect wire efficiency), random wire noise within the chamber, VDC resolution limitations, and coincident tracks (multiple particle tracks occurring within one event). Events were simulated via random numbers. Generation of evenly distributed and Gaussian random numbers were done using ROOT's random number generators gRandom::Rndm and gRandom::Gaus, respectively [7].

Chamber noise was simulated with a flat distribution of both wire numbers and TDC times. Chamber resolution limitations were simulated by adding Gaussian random numbers to all hit TDC values. A mean of 0 was always used to generate noise agreeing with experimental results found in [3]. The sigma (i.e. chamber resolution) was an adjustable parameter. Studies of the actual VDC performance yielded σ = 4.5 ns [3]. These studies also found a wire efficiency above 99% [3].

Coincident track occurrence was determined assuming that target particle emission was a random process following Poisson statistics. Coincident tracks were required to occur after their predecessors, which was accomplished by the addition of a random time offset to all TDC values

of hits in the track. Time offsets were increased for each additional track. Accordingly, the probability of an additional track was recalculated after each track using the formula:

$$1 - e^{-\lambda(T - t_{off})} \quad (1)$$

Where T is the TDC time window and t_{off} is the time offset of the most recently occurring track. Trigger tracks were given a time offset of 0.0 ns.

In order to assess the effects of the analyzer's track angle estimation, simulation files were generated fixing the track angle parameter used in the time-to-distance conversion to 45°. This data was then analyzed, fixing the same parameter within the analyzer, thus eliminating the angle estimation normally required. Fixing this parameter causes inaccurate calculations of drift-times or drift-distances in the simulator and analyzer (respectively), but the inaccuracies cancel out exactly, and angle estimation is eliminated.

A flow chart outlining data simulation and analyzer processing of information can be found in figure 2.

3.3 Analysis Methods

All Monte Carlo simulations generated 100,000 events. File series A, B, and C were generated and analyzed to assess the effects of chamber time resolution, chamber noise, and coincident tracks, respectively. In series A, chamber resolution was varied, while chamber noise and coincident tracks were turned off. Series B and C both included a 4.5 ns chamber resolution but turned off coincident tracks and chamber noise, respectively. These file series were analyzed using angle estimation. File series D was generated and analyzed using the method described in section 3.2 to assess the impact of angle estimation. Simulation files in the series had varying chamber resolutions, with all other effects turned off. All file series had wire efficiencies of 100%, thus eliminating any wire efficiency effects.

File series F was generated to assess the analyzer's performance with all natural effects present. Simulation files were generated with VDC resolutions of 4.5 ns, wire efficiencies of 99.999%, 0.5% probabilities of chamber noise, and varying emission rates. Analysis of these files was performed using track angle estimation.

The analyzer results were evaluated using ROOT. The errors in track position and slope linear fits were examined, thus assessing the analyzer's ability to properly locate cluster cross-over points. The analyzer's track reconstruction accuracy was examined by looking at the error distribution for the position coordinates *x* and θ (in the detector coordinate system). Gaussian fits of these distributions were performed to quantify coordinate resolutions (distribution σ values). The *y* resolution was assumed identical to that of *x*, and the φ resolution was assumed identical to that of θ.

The analyzer's ability to find the trigger track amongst chamber noise and/or coincident tracks was also assessed. A track error parameter was defined as

$$\frac{1}{4}\left(\frac{(x_a - x_t)^2}{(x_{res})^2} + \frac{(y_a - y_t)^2}{(y_{res})^2} + \frac{(\theta_a - \theta_t)^2}{(\theta_{res})^2} + \frac{(\varphi_a - \varphi_t)^2}{(\varphi_{res})^2}\right) = \frac{\chi^2}{df} \qquad (2)$$

Where x_a is the analyzer track *x*-coordinate, x_t is the simulated trigger track *x*-coordinate, x_{res} is the VDC *x*-coordinate resolution, and so on.

The analyzer track with the smallest error was said to be the best fit to the trigger track. Fits with an error of 5.0 or less were considered acceptable. This cutoff value was determined empirically by inspection of error parameter distributions. In events where the analyzer was unable to locate a track, an arbitrary large error (10^6) was assigned. Similarly, 10^6 was assumed as an upper limit of relative errors. The index numbers of best fit analyzer tracks were recorded along with their relative errors.

Proper location of the trigger track requires an acceptable fit, as well as storage of the best fit track in the first analyzer track position (i.e. an index number of 0). The overall tracking efficiency was defined as the ratio of the number of events with proper trigger track location and the total number of events. The "regenerated multi-track efficiency" (RMTE) was defined in the same fashion but only considered events where the analyzer produced two or more tracks; the "generated multi-track efficiency" (GMTE) only considered events where the simulator generated two or more tracks. Statistical errors in these ratios were calculated assuming proportionality to the inverse square root of the number of trials observed.

4 RESULTS

3.1 Track Angle Estimation and Chamber Resolution Effects

The errors of position and slopes found in local cluster fits are plotted as functions of chamber resolution (in figures 3 and 4 respectively) for both file series A and D. Angle estimation was found to decrease both position and slope fitting errors. The benefit of angle estimation was more significant for larger chamber timing uncertainties. For the realistic 4.5 ns resolution, angle estimation decreased position and slope fit errors by $11.2 \pm 0.5\%$ and $15.1 \pm 0.5\%$, respectively.

Histograms of track x and θ coordinate error distributions with Gaussian fits for a series A file with a realistic 4.5 ns chamber resolution can be seen in figures 5 and 6, respectively. The x error distribution was found to have a full width half maximum (FWHM) of 204 μm; the θ error distribution was found to have a FWHM of 832 μrads.

3.2 Chamber Noise Analysis

A plot of the overall tracking efficiency and RMTE as functions of the probability of random wire firing can be seen in figure 7. Chamber noise was found to have a significant

impact on track reconstruction accuracy. The probability of acceptable trigger track fits to be stored in the correct slot (i.e. have an index of 0) as a function of chamber noise can be found in figure 8. It was found that as noise levels increase, the analyzer was less likely to properly store trigger track fits. Figure 9 shows the probability of regenerated multi-track events as a function of chamber noise. The likelihood of regenerated multi-track events was shown to increase dramatically with chamber noise level.

Sample analysis of analyzer performance for a series B file with a realistic 0.5% probability of random wire firing can be seen in figure 10. This includes 3 plots: a comparison of the number of tracks found and the number of tracks generated (a), a histogram of track errors (b), and a histogram of track indices of acceptable track fits (c). The analyzer was shown to frequently find one or two more tracks than are generated. Though the majority of errors were small, indicating proper trigger tracking, they were very widely distributed. Good trigger fits were almost always properly stored at this noise level.

3.3 Coincident Track Analysis

A plot of the overall tracking efficiency, RMTE, and GMTE as functions of triggering rate can be seen in figure 11. The overall tracking efficiency decreased with emission rate, though it remained above 90% for all rates below 1MHz. RMTE and GMTE were both significantly lower and showed little dependence upon emission rate. Tracking efficiency as a function of the number of generated tracks in an event can be seen in figure 12. The analyzer showed significant failures in analyzing multi-track events, falling below 90% efficiency for events with 3 or more tracks.

Sample analysis of analyzer performance for a series C file with an emission rate of 100 kHz is included in figure 13. The analyzer never found more tracks than were generated, and

only infrequently regenerated all generated tracks. The majority of track errors were small, and their distribution was narrow. Acceptable fits were stored properly for the majority of events.

3.4 Combined Effects Analysis

A plot of the overall tracking efficiency and RMTE as functions of emission rate can be seen in figure 14. Overall tracking efficiencies remained above 90% for all rates below 700 kHz. Multi-track efficiencies, however, were always below 90% and fell below 80% at rates above 500 kHz. Figure 15 shows the probability of regenerated multi-track events as a function of emission rate. As expected, the analyzer was more likely to find multiple tracks at higher emission rates.

Figure 16 contains a sample analysis of analyzer performance for a series F file with an emission rate of 100 kHz. The analyzer was shown to find more tracks than generated in a significant number of events. Track errors were shown to have a wide distribution. Acceptable trigger fits were stored properly the majority of the time, though less frequently than in files with only chamber noise or coincident tracks.

5 DISCUSSION AND CONCLUSION

5.1 Track Angle Estimation and Chamber Resolution Effects

The decrease in linear fit errors seen with the introduction of angle estimation into analysis methods indicates that the algorithm's iterative estimation method functions properly.

The 204 μm FWHM obtained in the *x*-coordinate error distribution agrees with experimental data reported in [3]. This signifies proper simulation of chamber resolution.

5.2 Chamber Noise Effects

The influence of chamber noise on tracking efficiency has several implications. As shown in figure 10, the analyzer frequently found multiple tracks despite being given only one

track and small amounts of noise (0.5% probability of random wire firing, which equates to approximately 2 misfiring wires per plane per event). Because noise was simulated with evenly distributed random wire numbers and TDC values, formation of a legitimate cluster at low noise levels was highly unlikely. Thus, the analyzer's fictional track location suggests flaws in its cluster formation criteria.

The trend of increasing improper storage of acceptable trigger track fits with more noise seen in figure 8 also suggests that the analyzer's method of trigger track determination needs improvement. This increase is attributable to the more frequent occurrence of multiple regenerated tracks. However, at these relatively low noise levels, all tracks created by random hits would have very poor linear fits (both in locating the cross-over point and the overall track). Yet, the analyzer sometimes deemed these fictional tracks as "better" than the actual track.

The tracking efficiency decreases seen in figure 7 may also be caused by incorrect matching of cluster cross-over points during track formation. The significant difference between overall tracking efficiency and RMTE might be a result of such mismatching. The analyzer was less likely to correctly identify the trigger track if it found multiple tracks. Thus, it is possible that the presence of more than one cluster per plane (which would be necessary for multiple track regeneration), caused significant difficulties in finding the correct combination of points. However, the decreased RMTE may also be caused by improper storage of the trigger track. These multi-track events require more study to determine exact causes of tracking failure.

Additional evidence of such mismatching can be seen in figure 10b. The tracking error parameter had a very wide distribution, with a significant portion of track errors between 5 and 100. These intermediate errors may be caused by improperly mixing nearby clusters, thus giving a fit that is close, but not completely accurate.

5.3 Coincident Track Effects

As seen in figure 11, coincident tracks did not have a dramatic impact on overall tracking efficiency. However, there were significantly reduced efficiencies when the analyzer reconstructed and was given multiple tracks. The decrease in overall efficiency at higher rates can thus be attributed to the increased probability of regenerated and generated multi-track events. Figure 12 shows the impact of multiple track events on tracking efficiency. When given 3 or more tracks, the analyzer had significant difficulties properly locating the trigger, with efficiencies falling well below 90%.

Figure 13b suggests that identifying a coincident track rather than the trigger may be a cause of coincident track related failures. The track error parameter distribution had a very narrow peak with a long tail, indicating that the trigger was either identified properly or wasn't found at all. This may be due to recreating only one track corresponding to a coincident track, rather than the trigger.

The decreased efficiency may also relate to improper storage of trigger track fits. As seen in figure 13c, a significant number of trigger fits were improperly stored. This is further evidence of flaws in the analyzer's method for determining the trigger among multiple tracks.

5.3 Combined Effects

As seen in figure 14, with all realistic effects included, tracking efficiency remained above 90% for emission rates below approximately 700 kHz. Higher rates, however, were less accurate, which can be attributed to the analyzer's difficulty in processing multi-track events. The RMTE was significantly lower and showed a dependency upon emission rate. This dependency (not seen in files with coincident tracks alone) may be due to mismatching of noise clusters and track clusters, which would be more likely with more coincident tracks. The wide

track error distribution seen in figure 16b, much like the distribution seen in figure 10b, indicates that such mismatching may occur. Further investigation is necessary to determine why RMTE shows a dependence upon emission rate with the presence of coincident tracks and chamber noise.

As this analysis shows, the analyzer has a high efficiency in track reconstruction at lower triggering rates. However, there are significant failures in processing multiple track events and chamber noise. Further work is necessary to determine what algorithm flaws contribute to these failures. Areas of focus for such analysis should include the analyzer's method of cluster formation, criteria used to locate the trigger track amongst multiple regenerated tracks, and methods used to match cluster groups in track formation. Such an investigation may lead to improvements in the tracking algorithm and overall tracking efficiency.

6 ACKNOWLEDGEMENTS

I would like to recognize and extend thanks to my mentor, Jens-Ole Hansen. Without his guidance and help, this project would not have been possible. I would also like to thank the Department of Energy and the Office of Science for providing me with the opportunity to conduct research at JLab. Finally, thank you to JLab's Science Education Department for their support throughout my internship.

7 REFERENCES

[1] R. Brun and F. Rademakers, "The Root System Home Page," October 7, 2004, http://root.cern.ch/

[2] K. Rossato, "Analysis of Wire Tracking Algorithms Through Use of Simulated Data," Summer Undergraduate Laboratory Internship Project, Thomas Jefferson National Accelerator Facility, August 2004 (unpublished).

[3] K. G. Fissum et al., "Vertical Drift Chambers for the Hall A High Resolution Spectrometers at Jefferson Lab," *Nuclear Instruments and Methods in Physics Research, Section A*, vol. 474(2), pp. 108-131, December 2001.

[4] R. Wechsler, “Drift-Time Properties of CEBAF Hall A Vertical Drift Chambers,” B.S. thesis, Massachusetts Institute of Technology, Cambridge, MA, U.S., 1996.

[5] J. Alcorn et al., “Basic Instrumentation for Hall A at Jefferson Lab,” *Nuclear Instruments and Methods in Physics, Section A*, vol. 522(3), pp. 171-628, April 2004.

[6] LeCroy Website, “1877 Mutlihit Time-to-Digital Converter Data Sheet”, April 1996, http://www.lecroy.com/lrs/dsheets/1877.htm.

[7] R. Brun and F. Rademakers, “Root Reference Guide, ROOT Version 4.04/02”, July 19 2004, http://root.cern.ch/root/Reference.html.

8 Figures

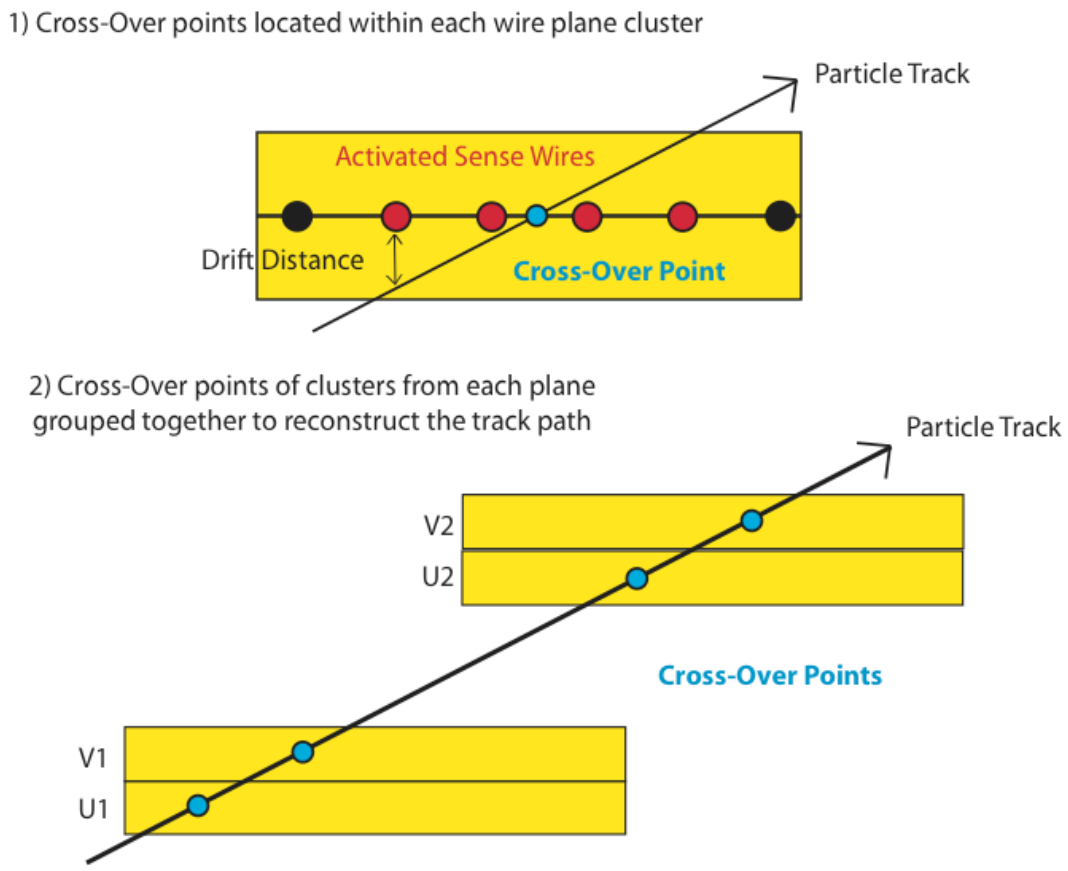


Figure 1. Diagram of analyzer track reconstruction

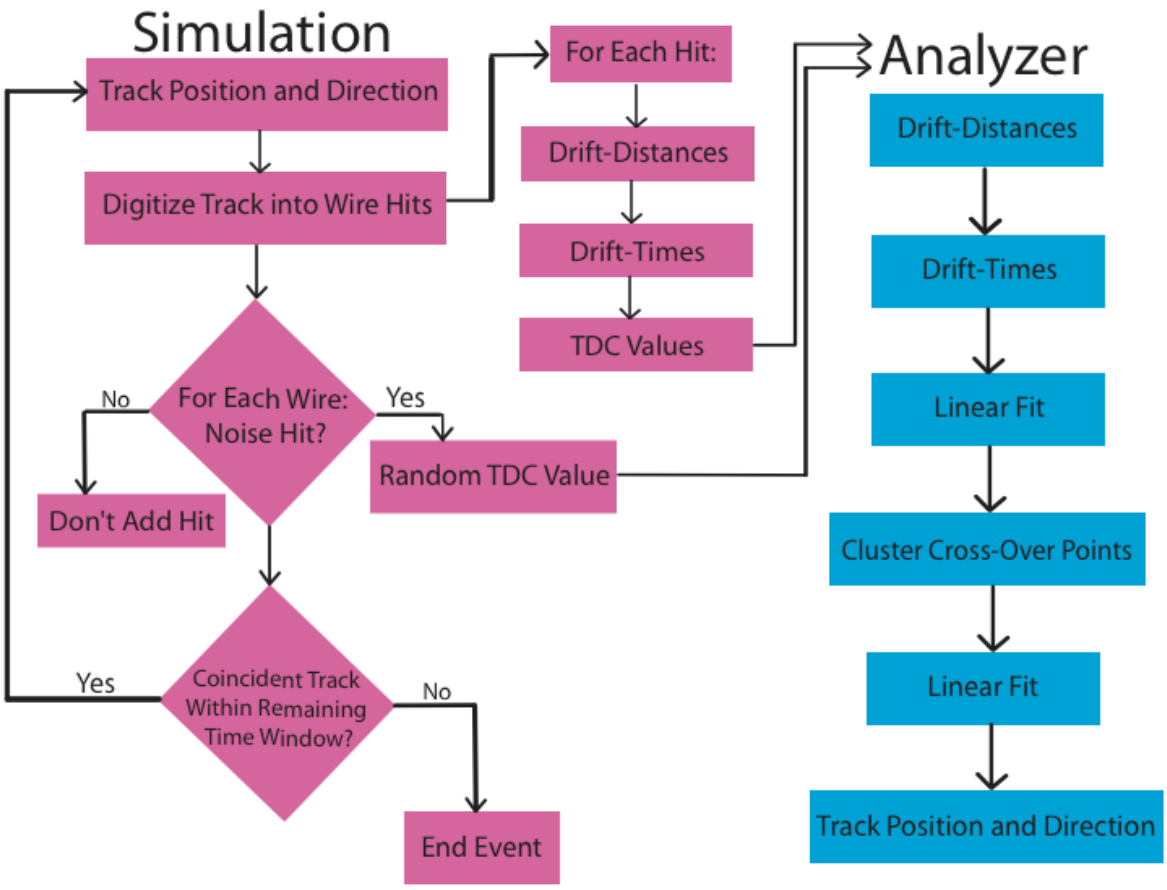


Figure 2. Flow chart outlining information generation in the simulation and information processing in the analyzer

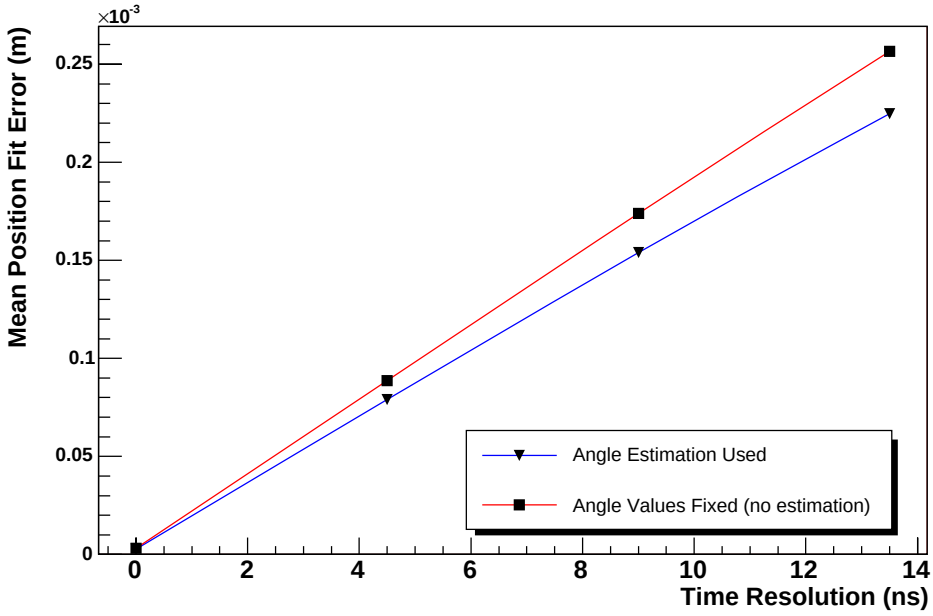


Figure 3. Average position fit error (among the four planes) vs. chamber resolution

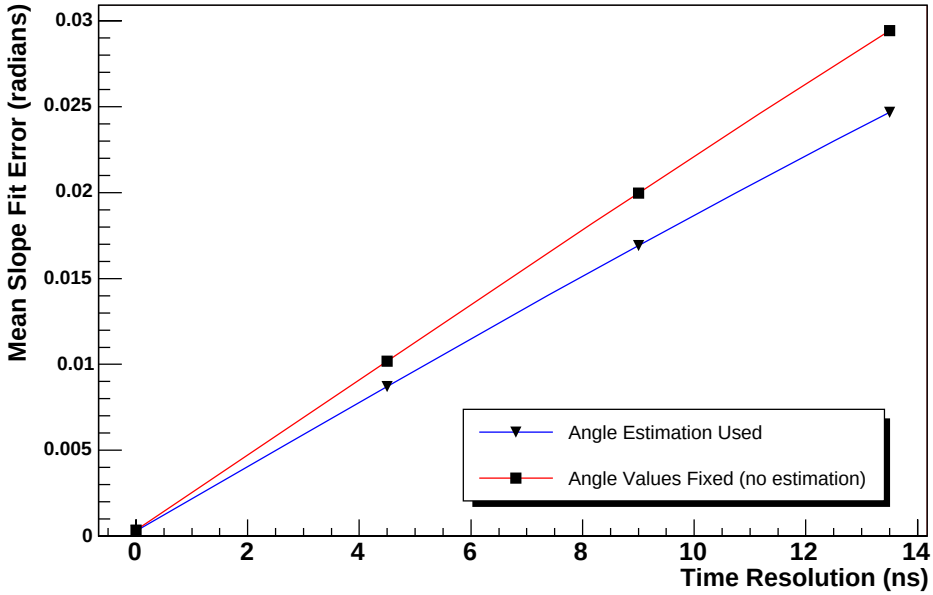


Figure 4. Average slope fit error (among the four planes) vs. chamber resolution

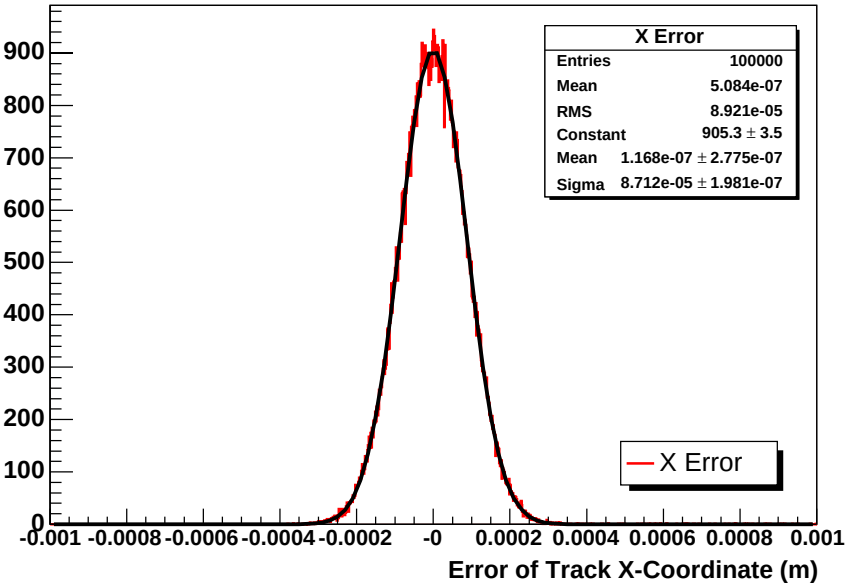


Figure 5. Error distribution of track *x*-coordinate for a file with a simulated 4.5 ns chamber resolution. Gaussian fit yields an x resolution of 87 μm (204 μm FWHM)

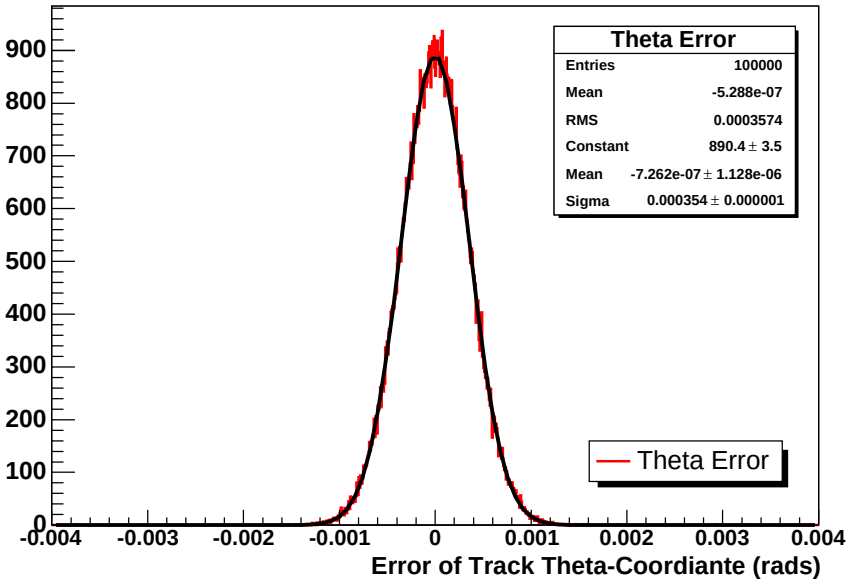


Figure 6. Error distribution of track θ-coordinate for file with a simulated 4.5 ns chamber resolution. Gaussian fit yields a θ resolution of 350 μrad (830 μrad FWHM)

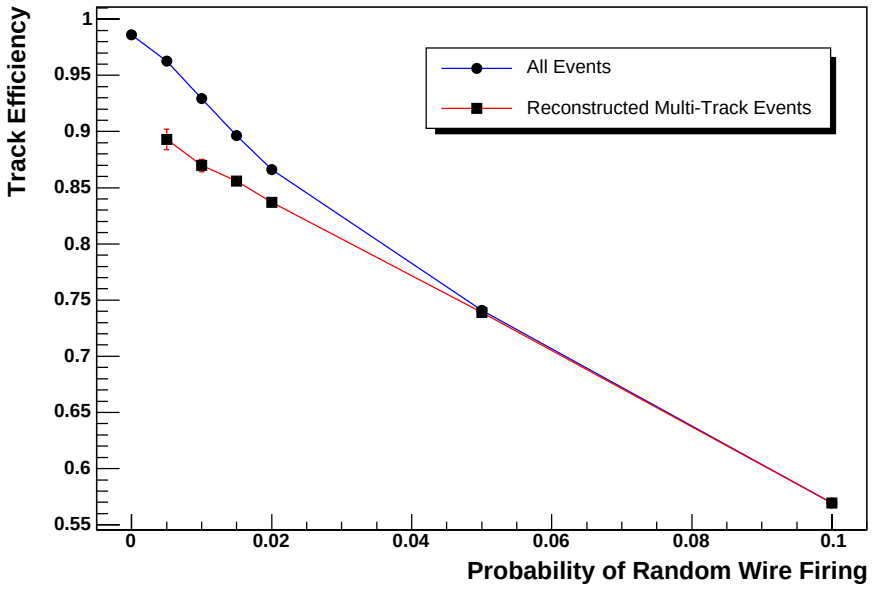


Figure 7. Overall and reconstructed multi-track efficiencies as a function of chamber noise level

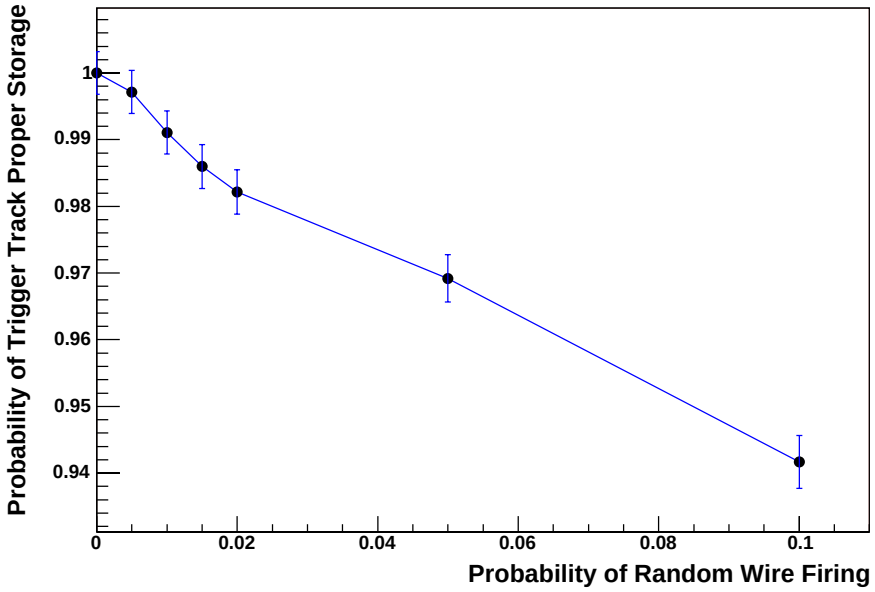


Figure 8. The probability of acceptable fits being stored with an index of 0 vs. level of chamber noise

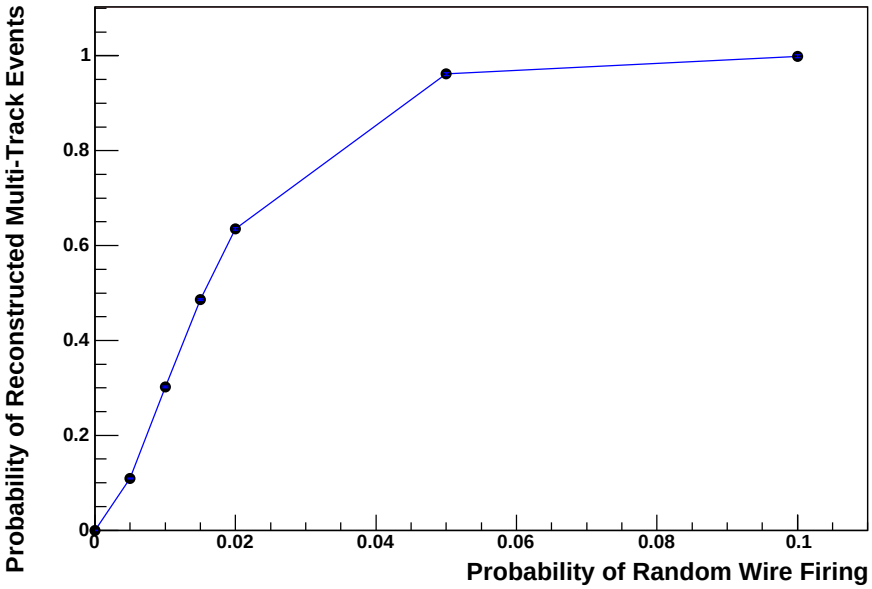
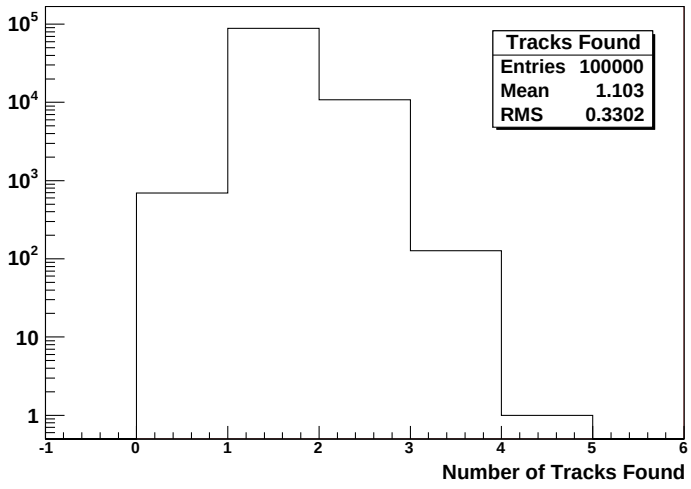
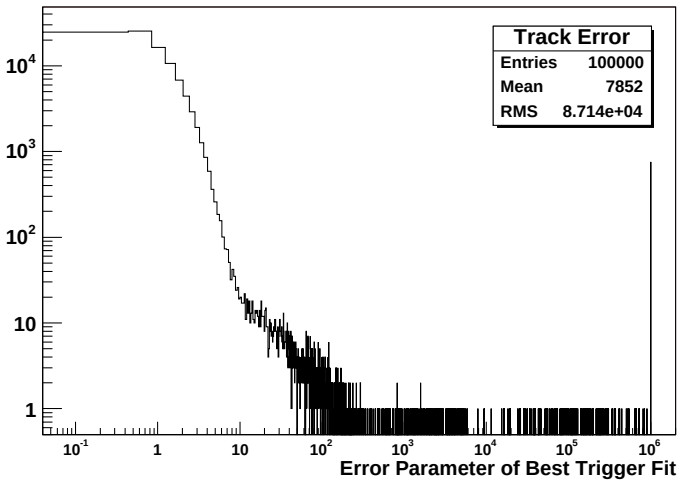


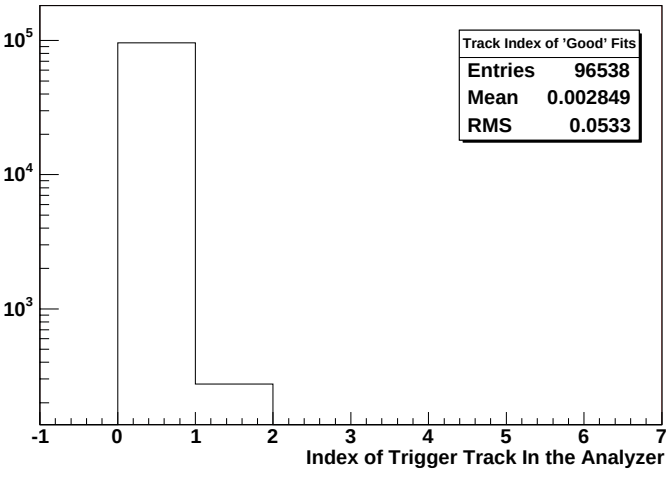
Figure 9. Probability of multiple track regeneration as a function of chamber noise level



a.



b.



c.

Figure 10. Sample analysis of a file containing 0.5% probability of random wire firing and a 4.5 ns chamber resolution. In figure a, note that all events contained only one generated track. Error parameter (figure b) defined in equation 2.

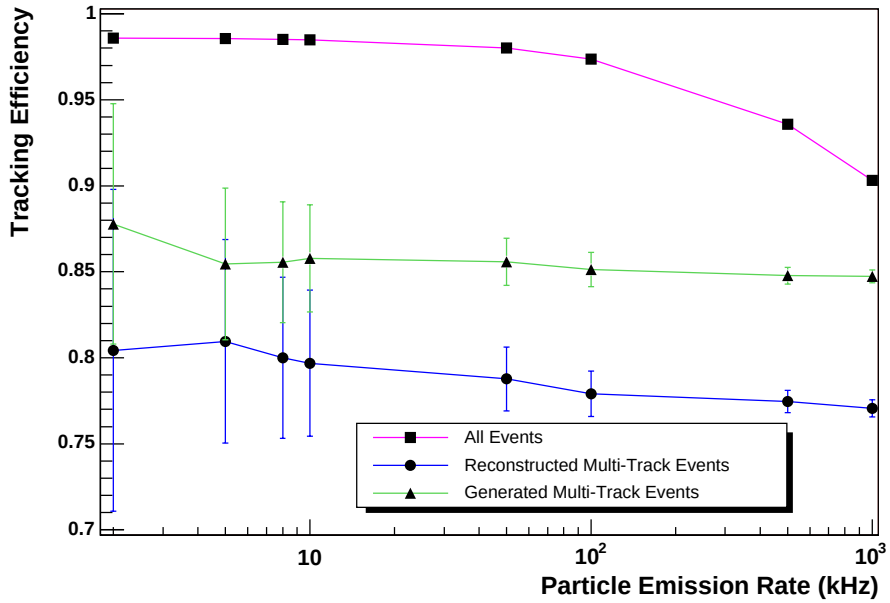


Figure 11. Overall tracking efficiency, RMTE, and GMTE as a function of triggering rate

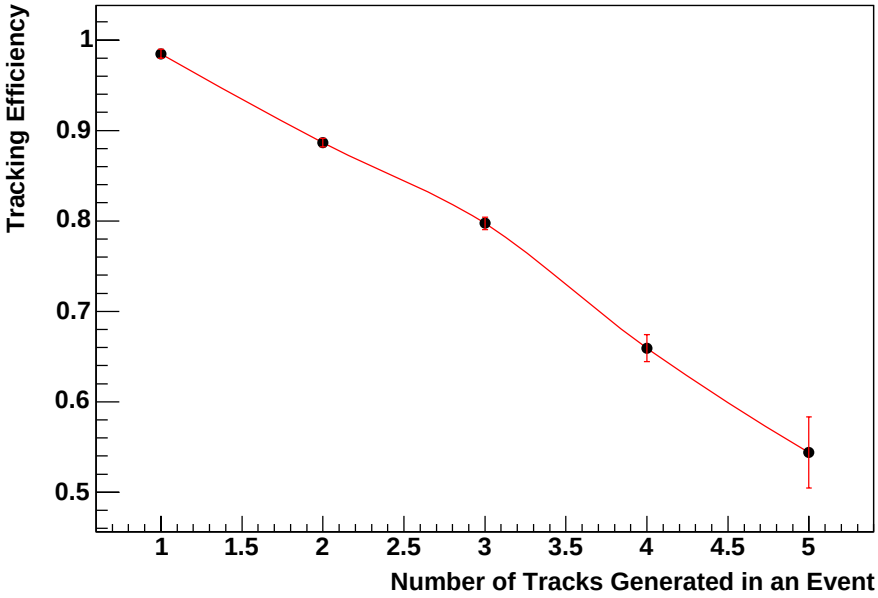
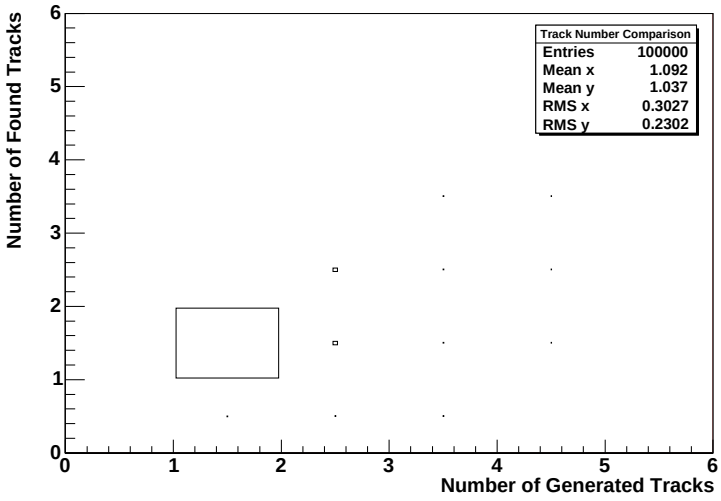
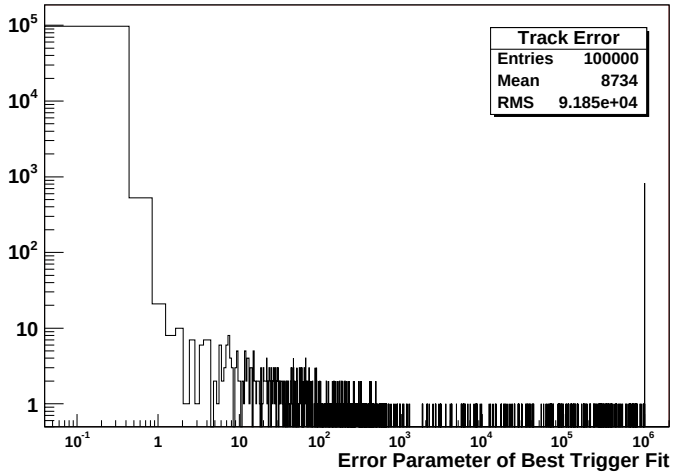


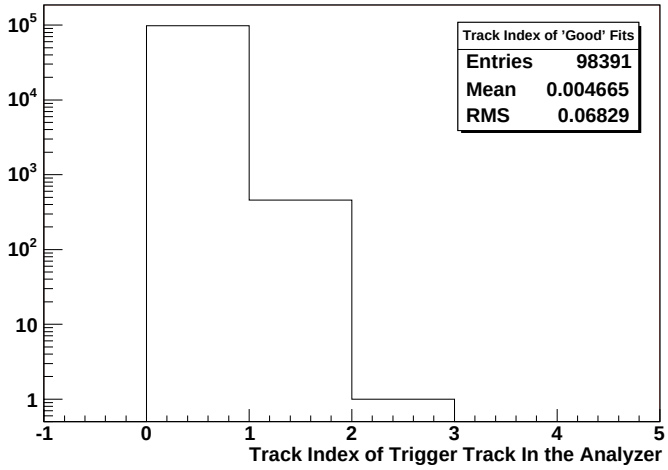
Figure 12. Overall tracking efficiency as a function of the number of tracks generated



a.



b.



c.

Figure 13. Sample analysis of a file with a 100 kHz triggering rate and 4.5 ns chamber resolution. The error parameter (figure b) is defined in equation (2)

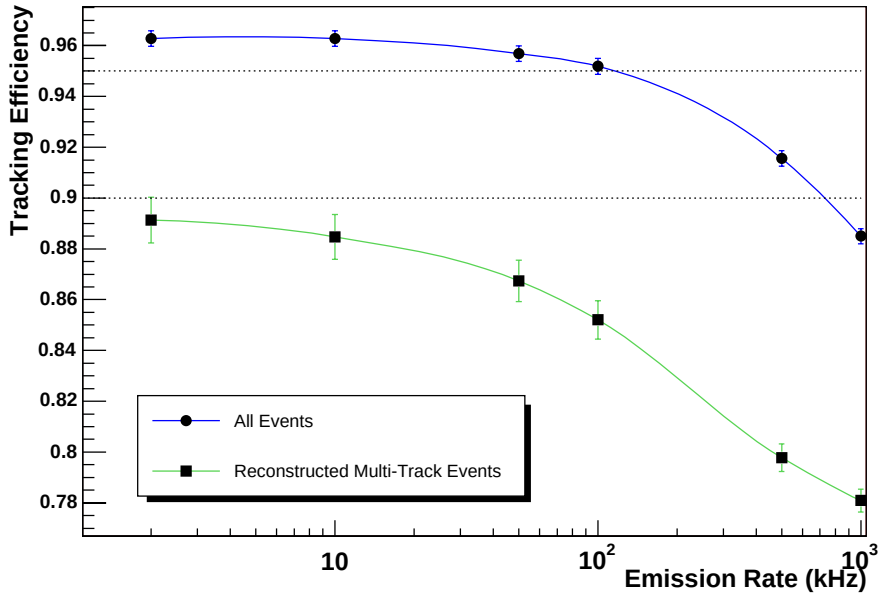


Figure 14. The overall and reconstructed multi-track efficiencies as a function of emission rate. All realistic effects (4.5 ns chamber resolution, 99.999% wire efficiency, 0.5% probability of random wire firing) were included.

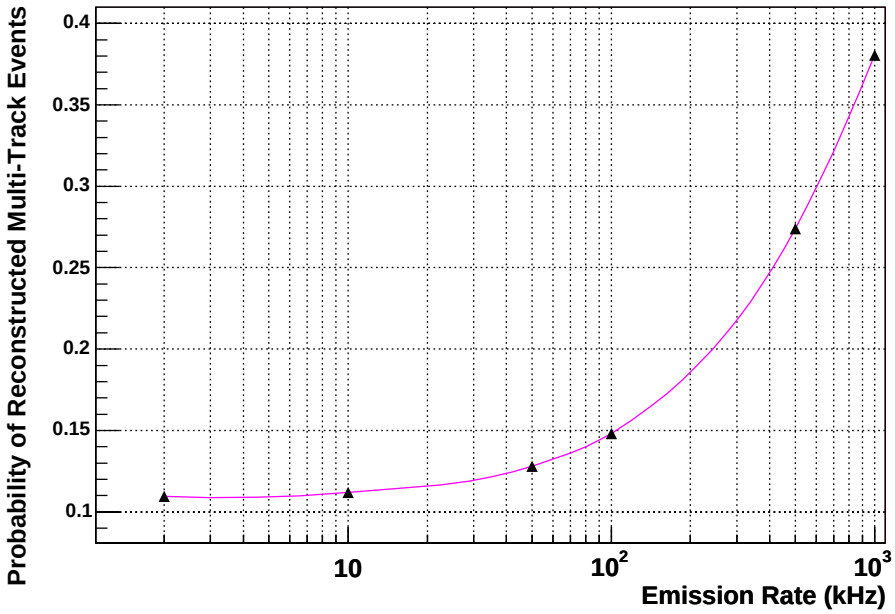
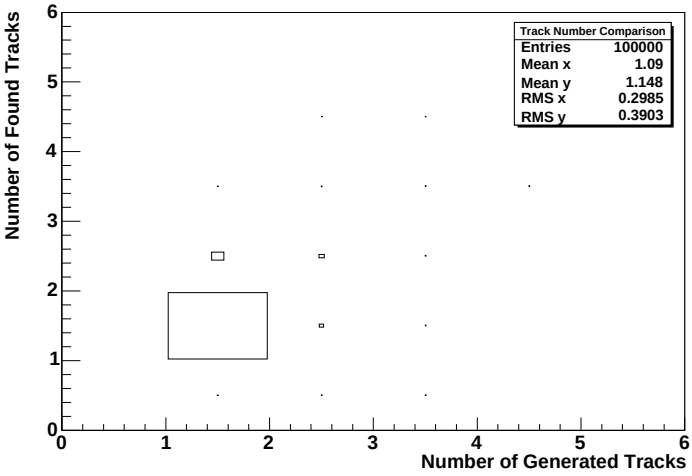
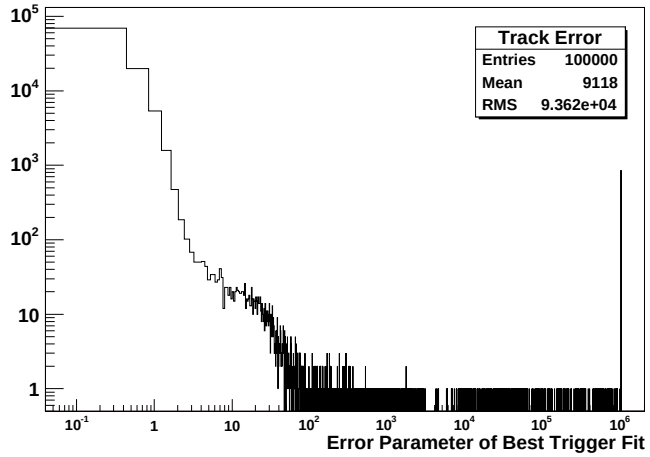


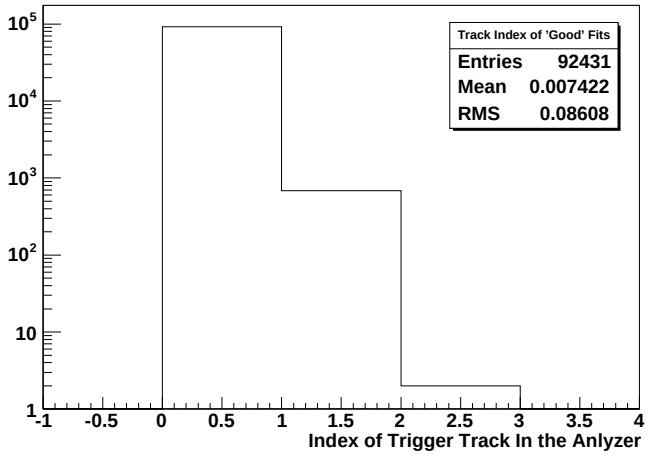
Figure 15. Probability of reconstructing multiple tracks as a function of emission rate



a.



b.



c.

Figure 16. Sample analysis of analyzer results of a simulation file with 4.5 ns chamber resolution, 99.999% wire efficiency, 0.5% probability of random wire firing, and 100 kHz triggering rate. The error parameter (figure b) is defined in equation (2).