CDFNote 6928
March 15, 2004
Version 0.1

A Method of Using the HADTDC System to Reject Large Met Events with Out-of-Time Energy

David Toback
Texas A&M University

Abstract

Until the EMTiming system is fully instrumented the HADTDC system is one of the most powerful tools to reject events where “°ou-of-time”± ener gy fr o mbea -gas interactions, cosmics, beam-halo etc. contributes fake Met. In this note we describe a methodology for using the HADTDC system which was designed for use in the GMSB-SUSY $\gamma\gamma$ + Met search, but is directly applicable to any analysis which suffers from any of these of backgrounds. In particular we need to be both sensitive to events with large amounts of hadronic energy deposited out-of-time, as well as be sensitive to small hints in the HAD that there may be large fake energy in the EM. We begin with a discussion of the limitations of the HADTDC system and of using a Run I style, fixed timing window cut analysis. We then discuss a new Sleuth-like methodology for identifying out-of-time events, and quantify the efficiency of our cuts. Finally we show how the method works on a photon+Met sample which is dominated by fake Met and conclude with the impact of our cut on the final diphoton+Met GMSB-SUSY analysis.

1. Introduction and Overview

The measured time-of-arrival of energy in the calorimeter is one of the most powerful tools to identify events where “°ou-of-time”±sources such as beam-gas interactions, cosmics, beam-halo etc. have an impact on the missing energy. Ensuring that all the energy in the event is from the primary collision is of the utmost importance in any search for new physics with Met in the final state. Unfortunately, timing for the entire calorimeter is not yet instrumented. Until the EMTiming system [1] comes online the HADTDC is often our best defense. In this note we describe a methodology for using the HADTDC system which was designed for use in the GMSB-SUSY $\gamma\gamma$ + Met search [2], but is directly applicable to any analysis which suffers from any of these of backgrounds. In particular we need to be both sensitive to events with large amounts of hadronic energy deposited out-of-time (for example in SUSY=> Jets + Met searches [3]), as well as be sensitive to small hints in the HAD that there may be fake energy in the EM (any search with final state photons and Met).

The HADTDC system has been improved for Run II [4], and now is active for the entire CHA, WHA and PHA. For every tower the time of arrival is measured with a TDC and this raw time is corrected for discriminator energy slewing as well as a t_0 of the crossing relative to the TDC t_0. For highly relativistic particles which promptly come from the collision (“ï -time”) the final corrected time has a mean of zero, but to first order is smeared by the (energy and detector dependent) system resolution[1]. It has long been known that this methodology can be very helpful for rejecting cosmics, beam halo, beam-gas interactions etc. which can deposit lots of energy “°ou-of-time”± and cause large MET [5]. In the following sections we discuss the limitations of the HADTDC system as well as the problems with some of the Run I style, fixed timing window cut analyses. In particular, searches for events with photons and Met require that we look at low energy towers, and doing so has implications both in terms of the resolution capabilities of the detector as well as dealing with the statistics of identifying problems when large statistics can cause large fluctuations which mimic the out-of-time signature. We then discuss a new Sleuth-like methodology for identifying out-of-time events, and quantify the efficiency of our cuts. Finally we show how the method works on a photon+Met sample which is dominated by fake Met and conclude with the impact of our cut on the final diphoton+Met GMSB-SUSY analysis. For more details on how to get the corrected HADTDC time for each tower see Appendix A.

[1] As discussed in more detail later, the timing resolution becomes much worse as the energy goes down. In addition, at low energies the timing distribution becomes more and more asymmetric.

2. Why not use a Run I Style fixed-window HADTDC Analysis?

In Run I timing was implemented only on the CHA and WHA, and the mean of the corrected timing distribution was centered at 10nsec (the typical time of flight from the collision point to the face of the CHA) with a resolution of about 5 nsec. While there were a number of methods of using the HADTDC system to reduce fake backgrounds, a typical analysis [5] picks a fixed timing window around the mean of the distribution (-25 nsec < time < 35 nsec) and every tower is classified as being either in-time (in the timing window) or out-of-time. The E_T of all the out-of-time towers is then summed and if this is a significant amount of energy the event is thrown away. This method works especially well if one is able to only consider towers with large amounts of energy such as in the SUSY Jets+Met analysis [2]. However, these conditions are not always met, in particular in searches with final state photons, and in this case there are three problems with this simple analysis style for Run II.

2.1 Some backgrounds can produce large amounts of fake energy in the EM, but only small amounts in the HAD

In searches for new physics with Met as well as photons in the final state one can classify out-of-time energy into two categories: Type 1) Cosmic/Beam-halo-like in which large amounts of energy are deposited in the EM, sometimes as fake photons, but only small amounts are deposited in the HAD; and Type 2) Beam-gas-like interactions in which large amounts of fake energy can be deposited in both the EM and HAD. In type 1, to first order the photons are back-to-back with the missing energy and can often (assuming the event is clean enough such that the Met-phi isn’¯t smeared, or you can make such a cut) be removed using back-to-back cuts between the photon and Met [2]. Events with two photons from beam-halo can safely be rejected if both photons are in the same wedge, and have a vector E_T which is equal and opposite the Met. HADTDC cuts provide additional rejection for these types of events. Type 2, and type 1 with two photons from the cosmics are more insidious. For type 2 the energy out-of-time is uncorrelated with the photons (the photon candidates are from the primary collision) and can make searches such as the SUSY-GMSB $\gamma\gamma$+Met search very vulnerable until the EM towers are directly timed. In particular, in order to safely remove both types of events we need to search for the presence of “°hit”¡± th even a small amount of energy was deposited in a single tower. Figure 1 shows a spectacular example of an event from the $\gamma\gamma$ +Met analysis which is likely to be from a cosmic ray in which there is large EM energy and only a single, low energy HAD tower which is out-of-time with the collision[3].

[3] In Section 4 (see

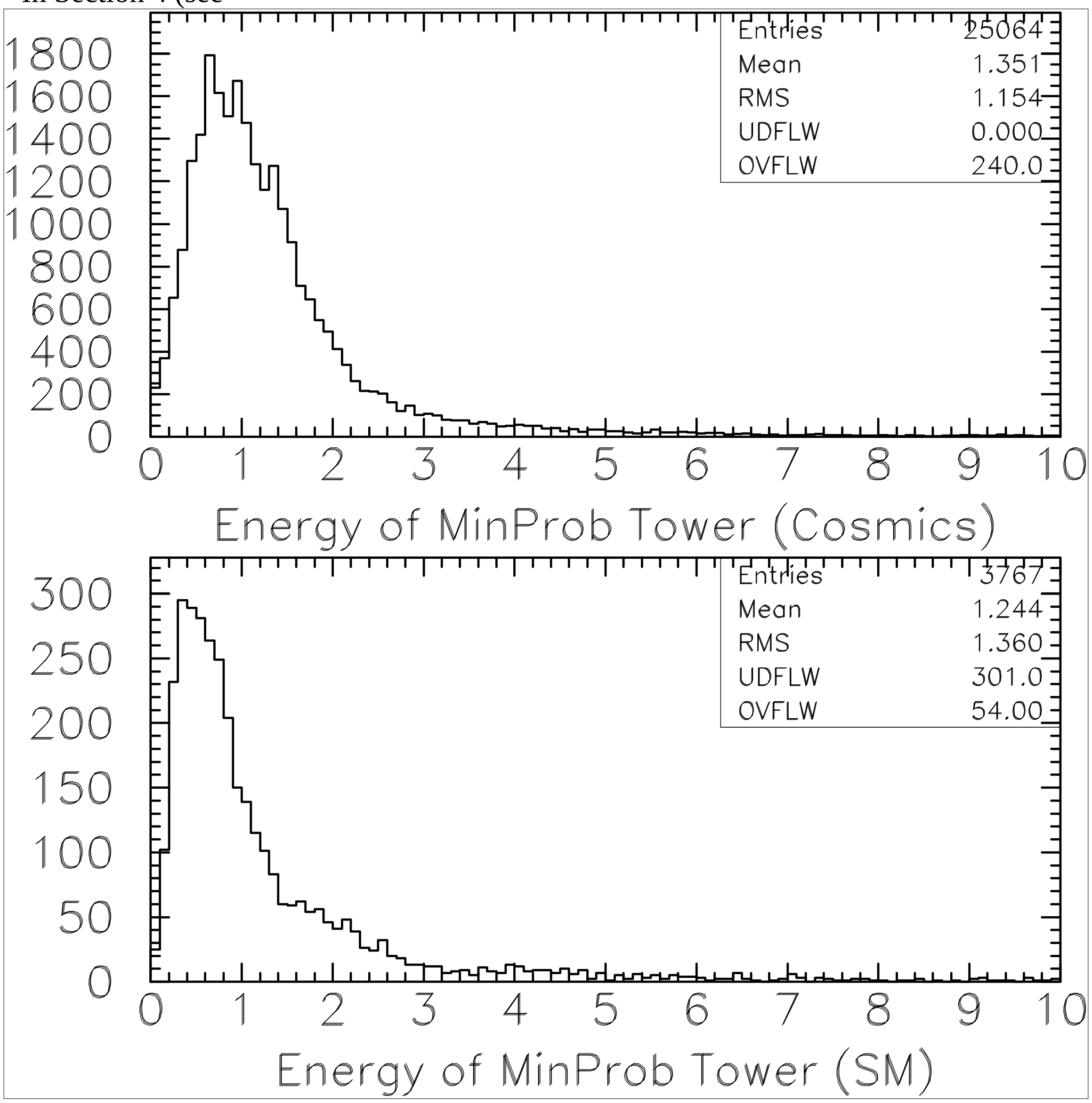


Figure 9) we confirm that the majority of out-of-time towers in cosmics events are low energy.

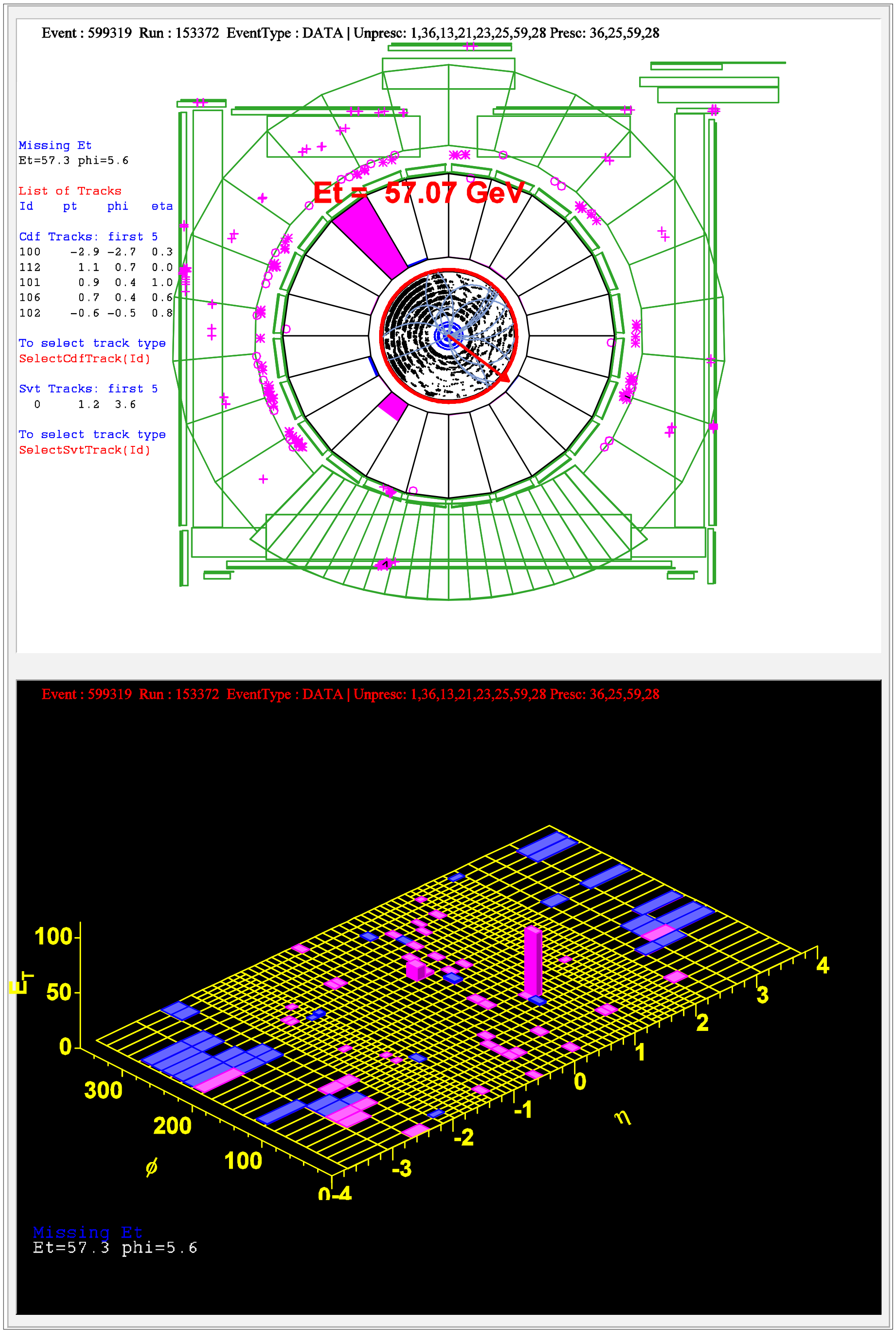


Figure 1: This figure shows an event display of R153372/E599319 from the γγ+Met sample before the HADTDC cuts. This event contains two high E_T photons (75 GeV and 15 GeV) and 59 GeV of Met. Next to the high energy photon is a single hadronic tower with 1.5 GeV of energy and 85 nsec out-of-time. It is directly in line with both photons (not behind either) and is a possible indication that a cosmic ray passed trough the detector and deposited all three objects. There is no indication of any other towers out-of-time in the event, and there is no timing information for either photon.

2.2 Dealing with HADTDC Response to Low Energy Deposits

Looking for low energy “hits” in the HAD calorimeter isn’t in itself a problem, the calorimeter can easily measure the energy down there, unfortunately the HADTDC system response worsens as the energy goes down. Specifically, there are two problems: 1) The HAD timing resolution gets worse, and 2) the distribution develops long tails and becomes asymmetric.

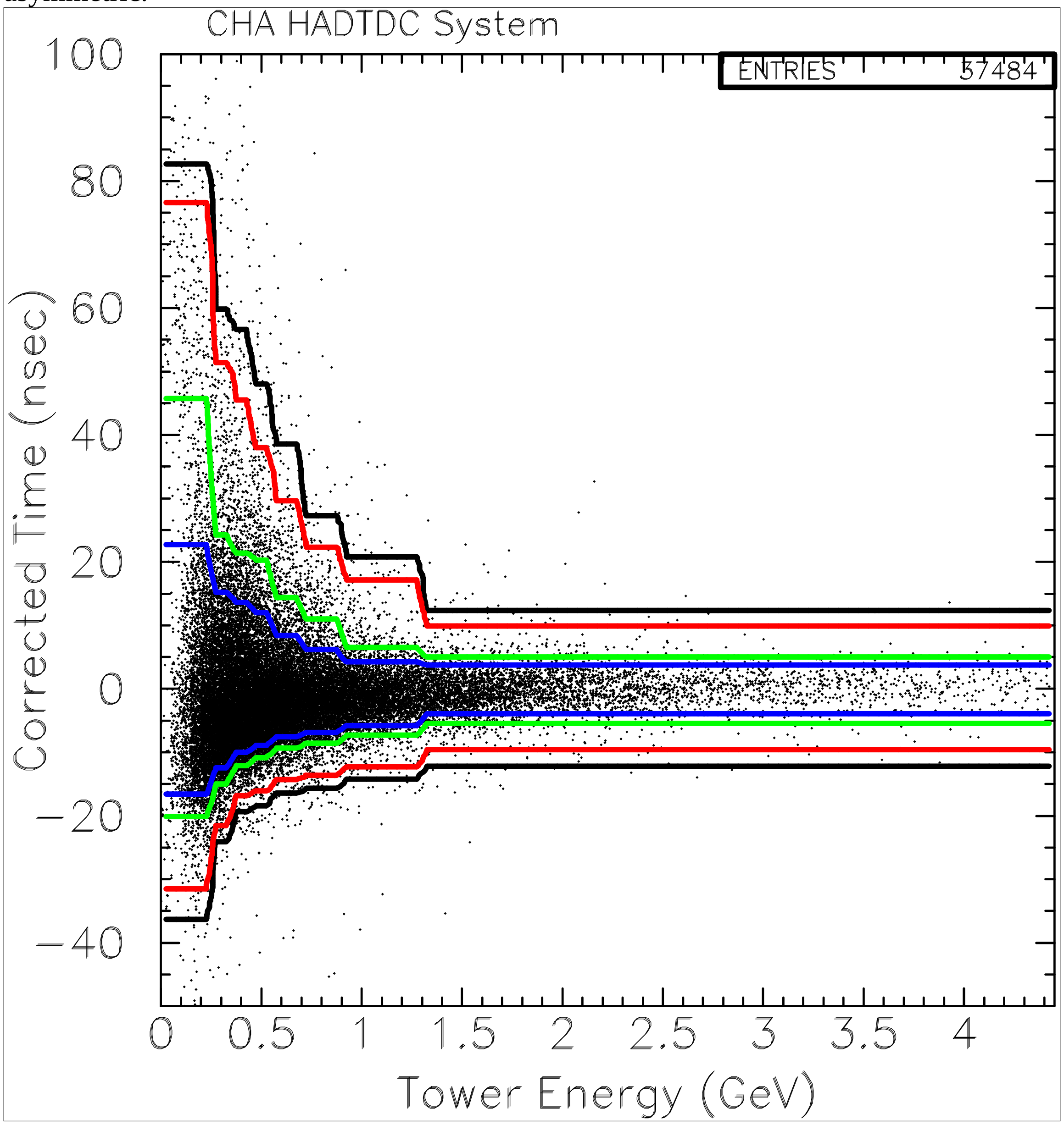
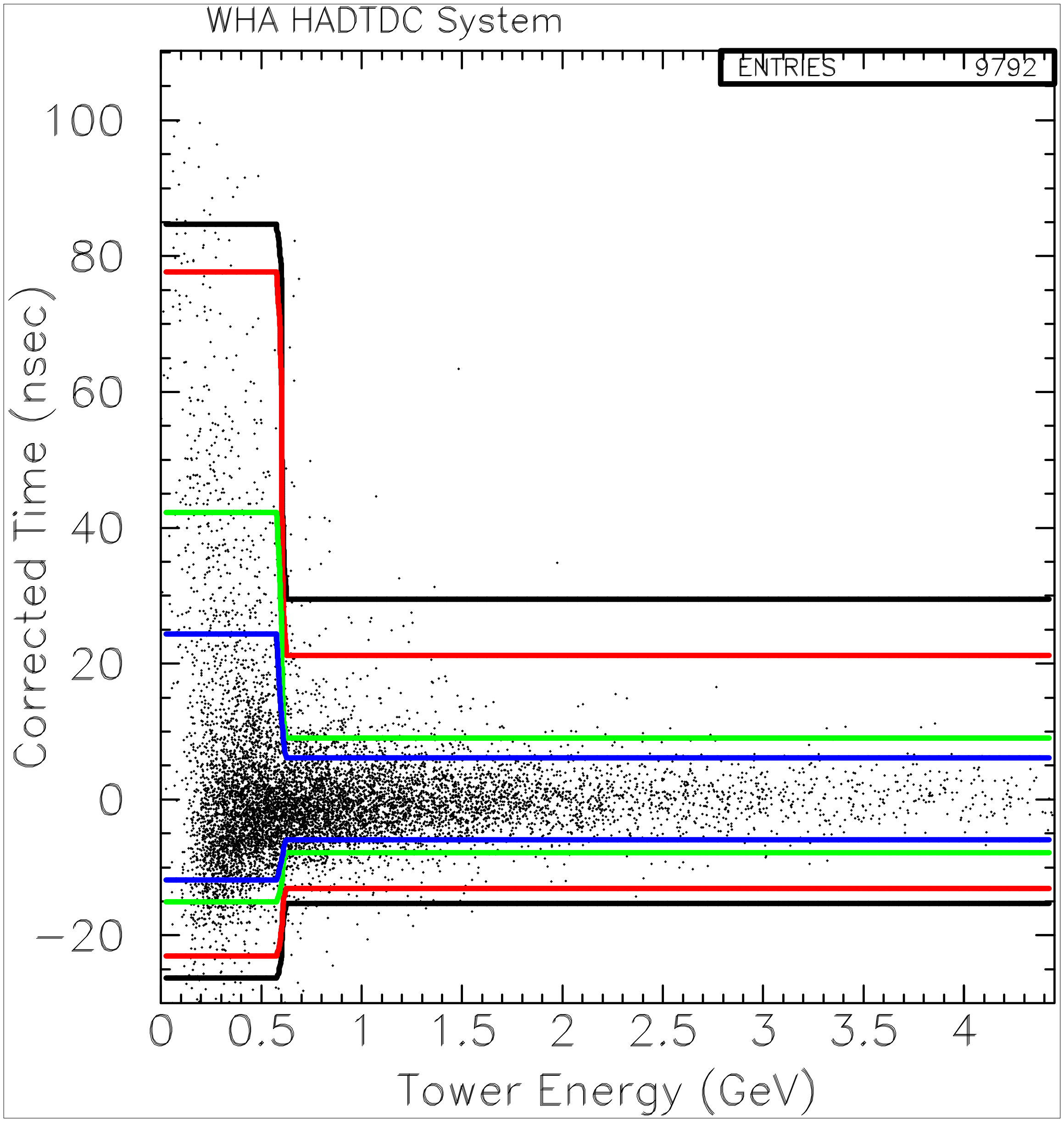


Figure 2,



WHA HADTDC System
ENTRIES 9792
Corrected Time (nsec)
100
80
60
40
20
0
−20
0
0.5
1
1.5
2
2.5
3
3.5
4
Tower Energy (GeV)

Figure 3 and

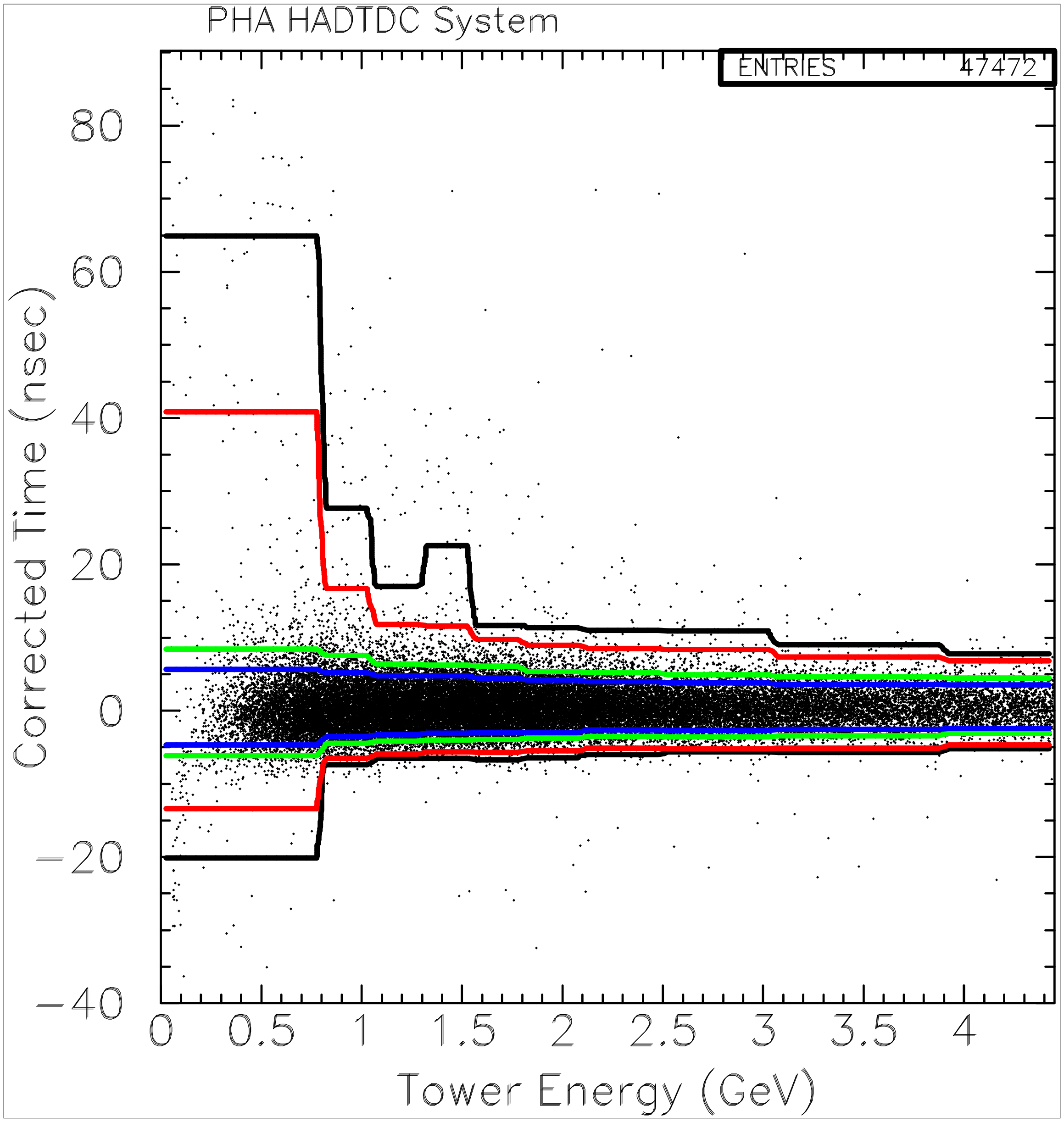


Figure 4 show the corrected time distributions as a function of energy for the CHA, WHA and PHA respectively.

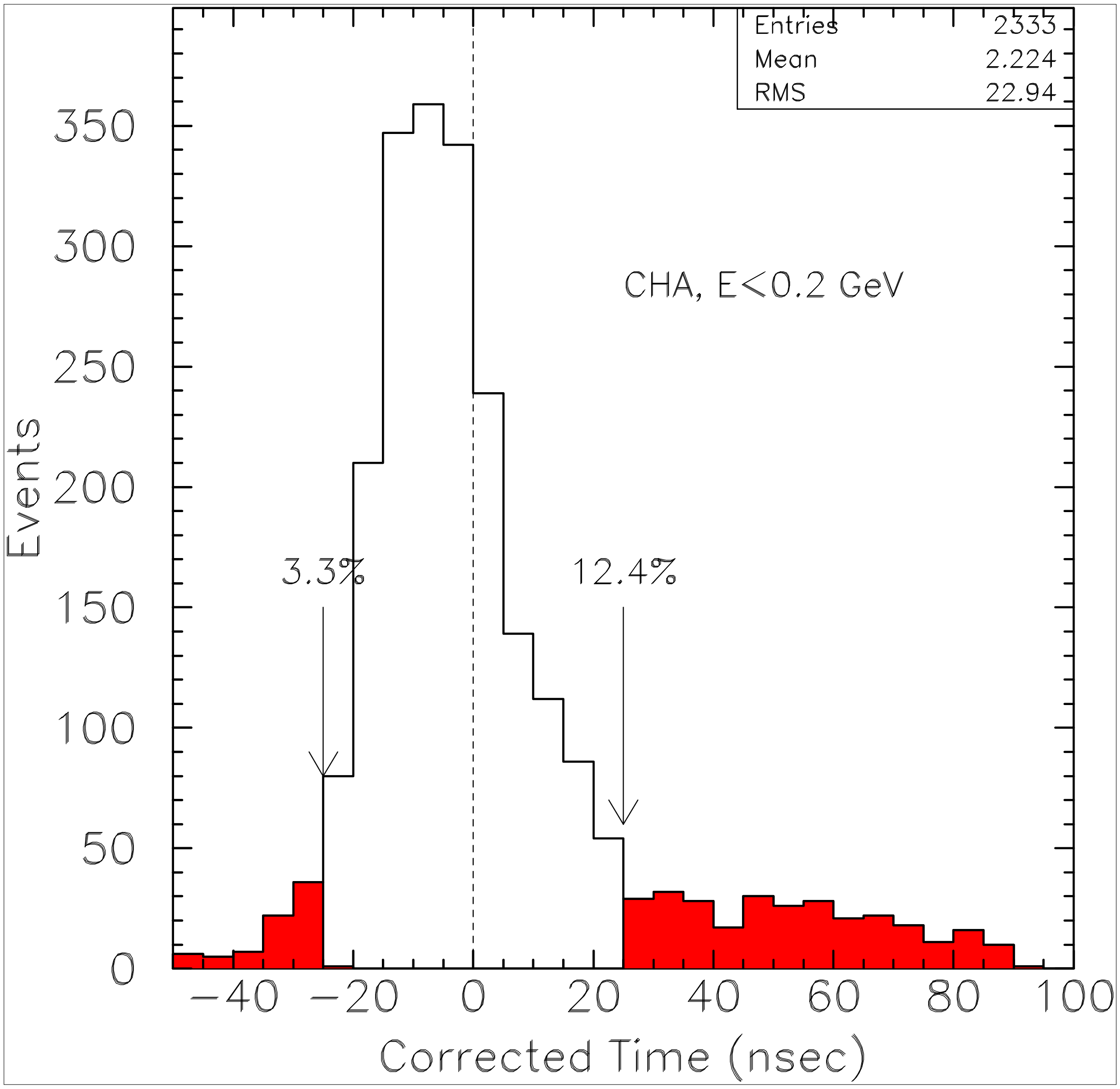


Figure 5 shows an example of the time distribution for very low energies in the CHA. Thus, if one wants to look at low energy towers, E<2 GeV, one needs to take these detector and energy dependent distributions into account. It is possible to create fixed-efficiency cuts, but they would be both energy and detector dependent, as shown in the Figures.

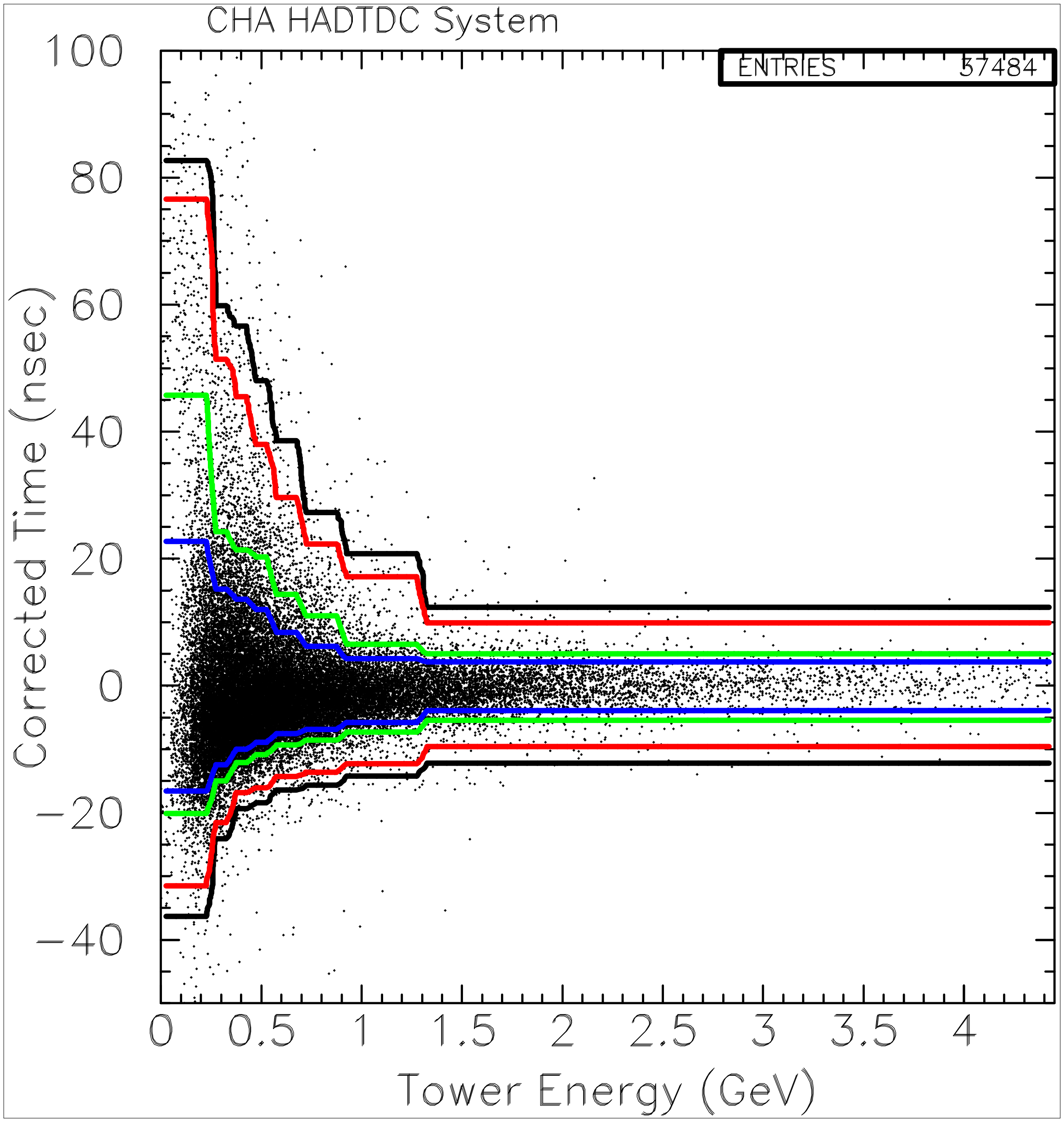


Figure 2: This figure shows the corrected time vs. energy for CHA hits from a Z=>ee control sample. Note that for low energies the width of the distribution increases, and become asymmetric. The solid lines show the contours of constant probability for events to be within a timing window. From innermost to outermost they are 20%, 5%, 1% and 0.5% respectively.

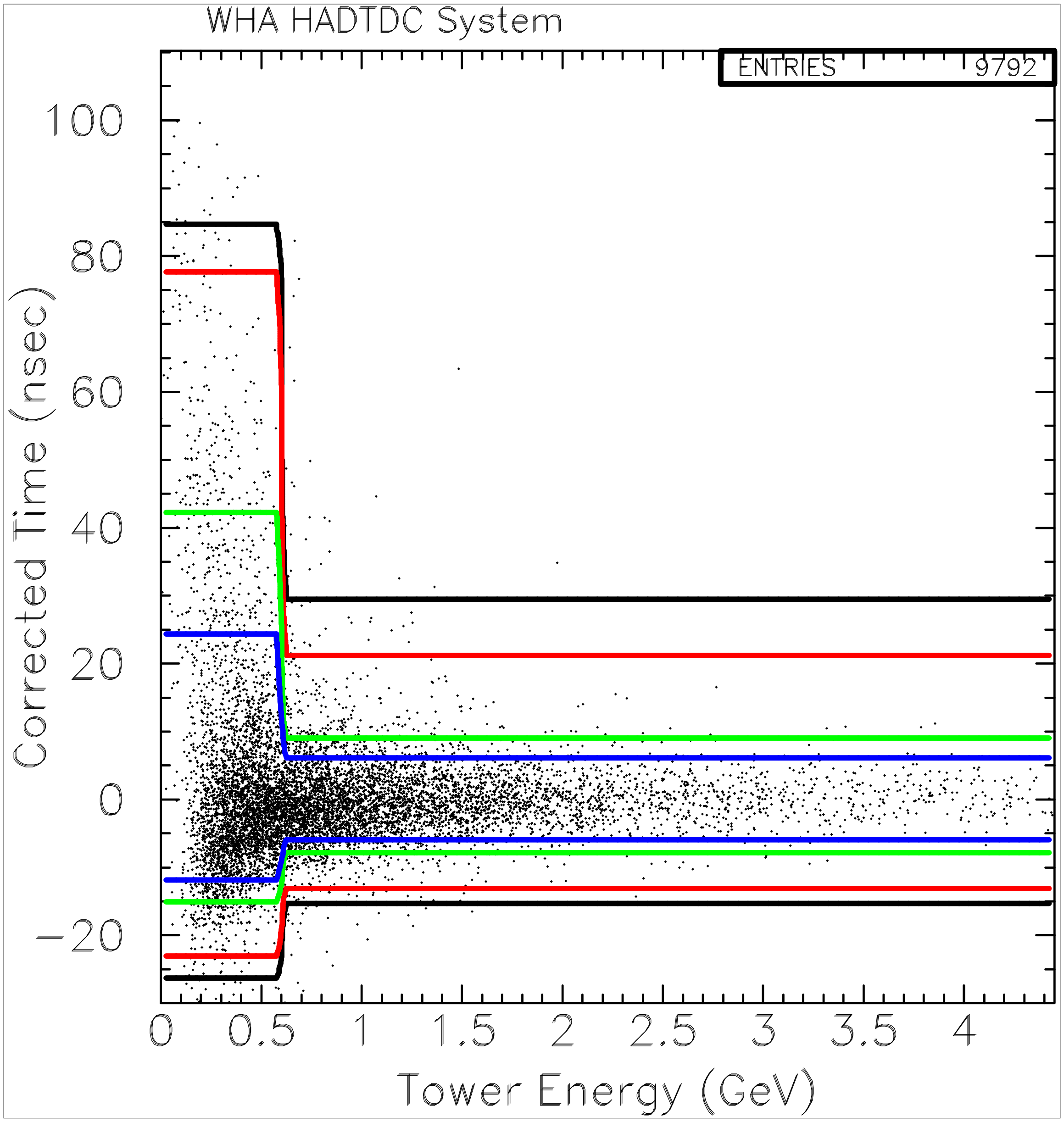


Figure 3: This figure shows the corrected time vs. energy for WHA hits from a Z=>ee control sample. Note that for low energies the width of the distribution increases, and become asymmetric. The solid lines show the contours of constant probability for events to be within a timing window. From innermost to outermost they are 20%, 5%, 1% and 0.5% respectively.

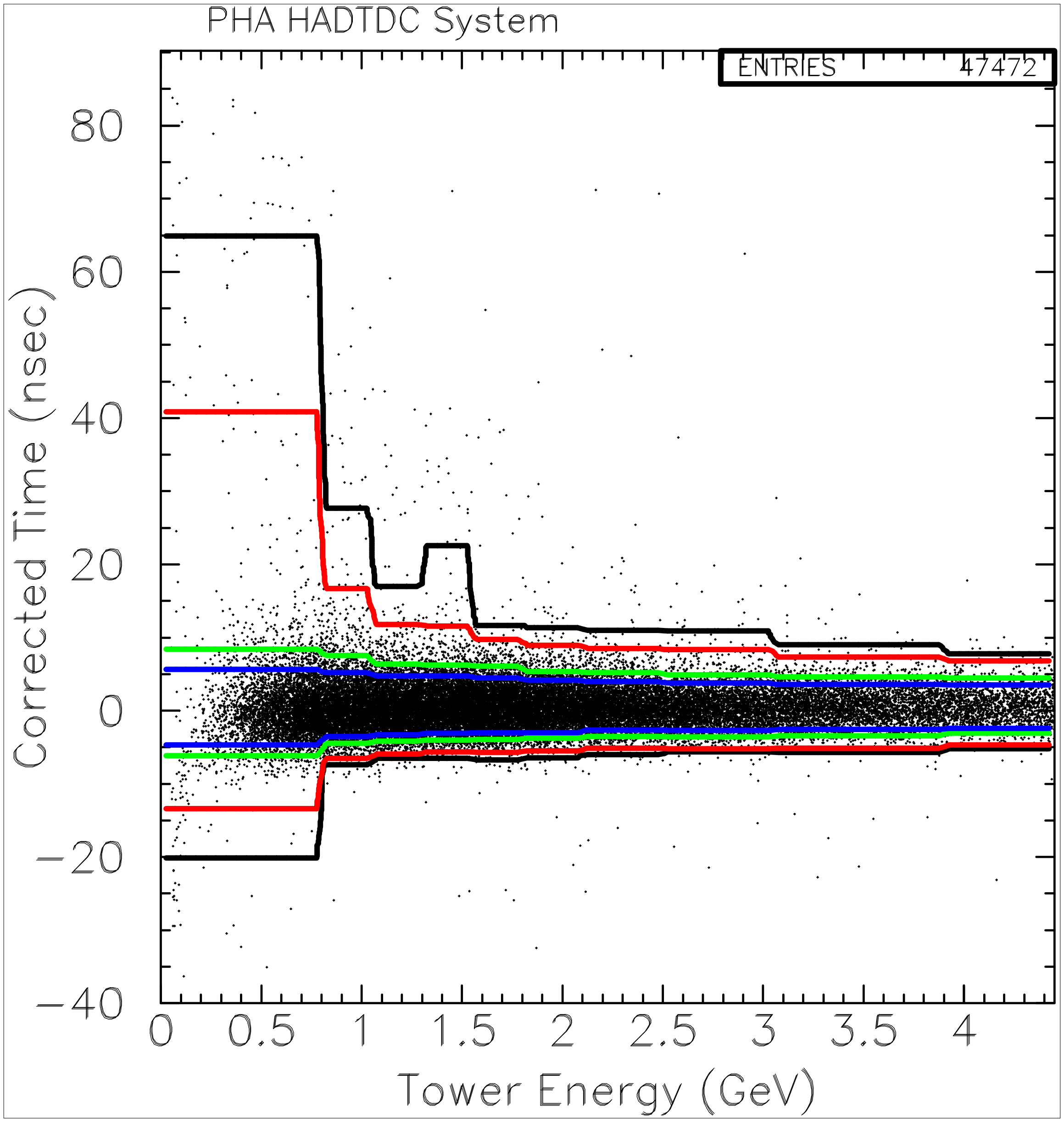


Figure 4: This figure shows the corrected time vs. energy for PHA hits from a Z=>ee control sample. Note that for while at low energies the width of the distribution increases, but the distribution is not nearly as asymmetric as for the CHA and WHA. The solid lines show the contours of constant probability for events to be within a timing window. From innermost to outermost they are 20%, 5%, 1% and 0.5% respectively.

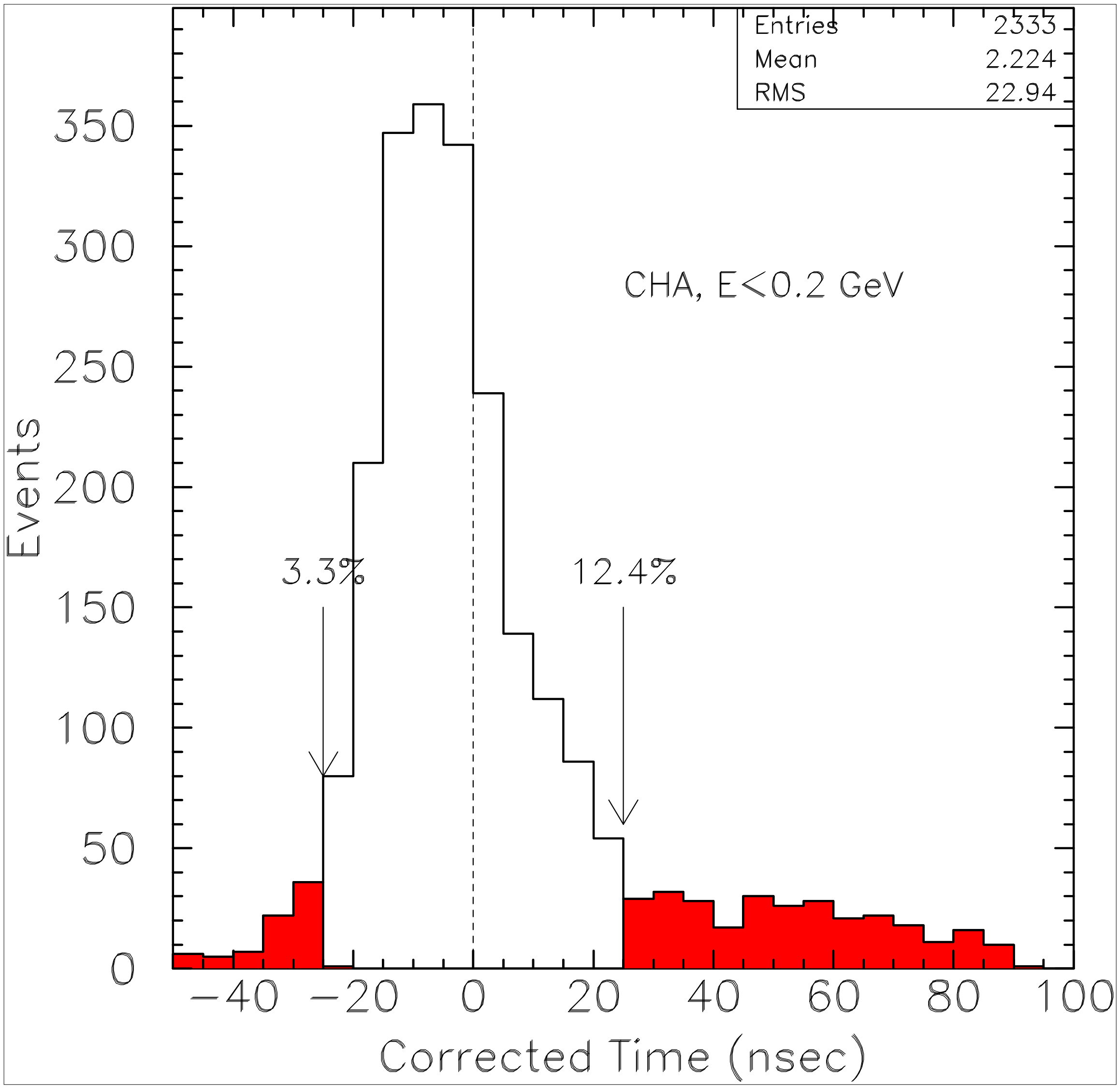


Figure 5: This figure shows the timing distribution for CHA towers with very low energy (E<0.2 GeV). In this example, the mean and the median are not at the same place. The corrections have largely centered the mean, which is at about 2.2 nsec, but the median is at about -5 nsec. The tails on both sides extend to large absolute times.

2.3 Dealing with Events with Large Numbers of Towers with Timing Hits

A third problem with the canonical high energy fixed efficiency/fixed timing window cut method is a bias against events with lots of towers (as might come from large q^2 interactions from Squarks and Gluinos) since there is a larger probability for a tower to fluctuate out of the timing window and fail the cut. The number of towers in an event can vary greatly.

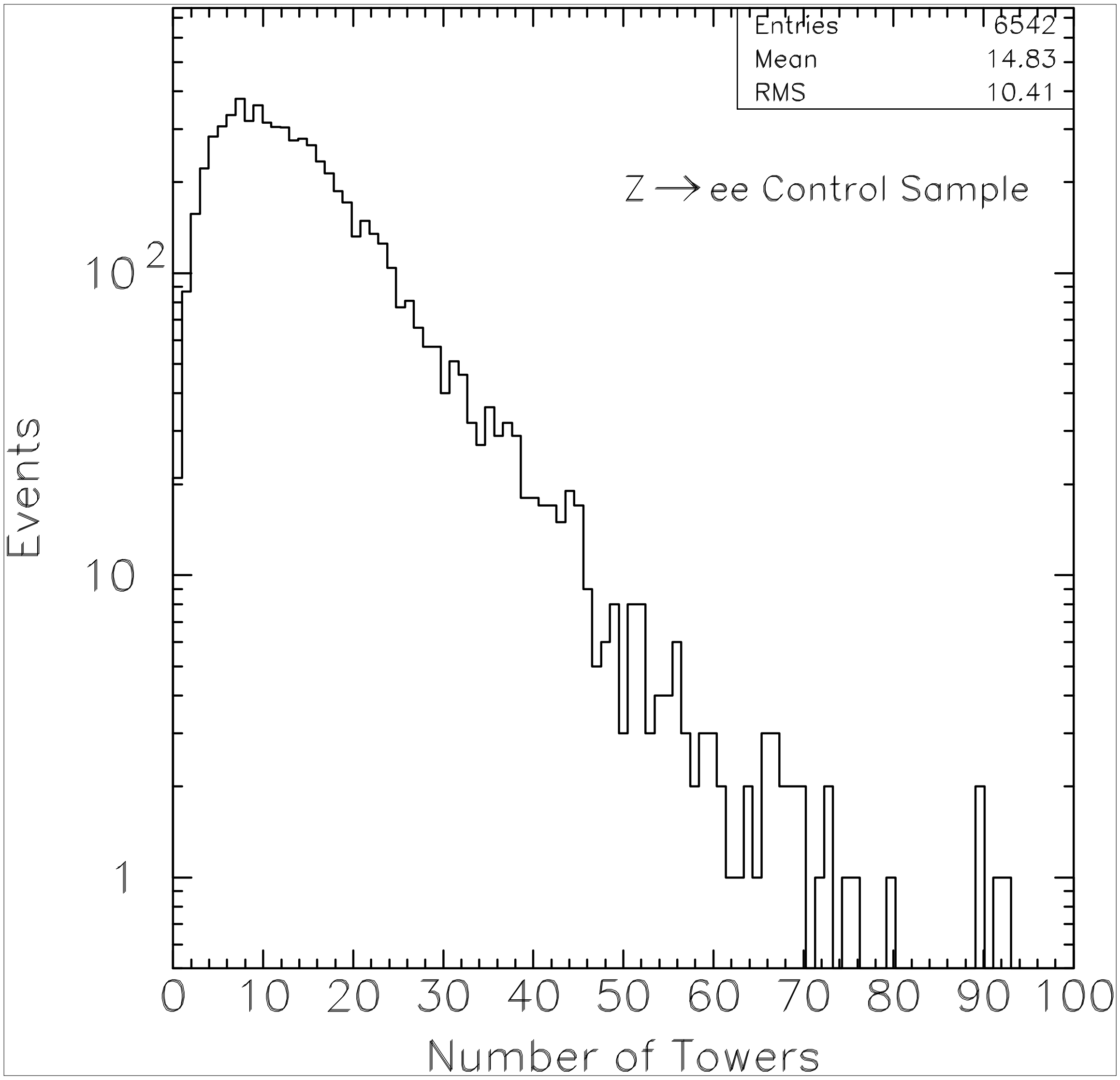


Figure 6 shows the variation for a Z=>ee sample, which has a mean of 10 as well as an event with 100 towers. For example, a 99% efficient timing window cut with 10 towers has a ~10% chance that one of them will fluctuate to be outside the timing window. Similarly 100 towers has a >60% inefficiency. The bottom line is that while we can create fixed efficiency cuts, doing so gives an event-by-event or model dependent efficiency, and it is less efficient as the SUSY mass goes up (i.e., there are more towers). The easiest ways to get around this are to raise the energy threshold to reduce the number of towers per event (which is unacceptable for the reasons listed above), or to make the timing window so wide that the difference in efficiency between 10 towers and 100 towers is negligible (also unacceptable as it make the timing window useless in separating out backgrounds). We elect to simply take into account the statistics of the number of towers hit when determining whether or not a potential hint of out-of-time energy is statistically significant.

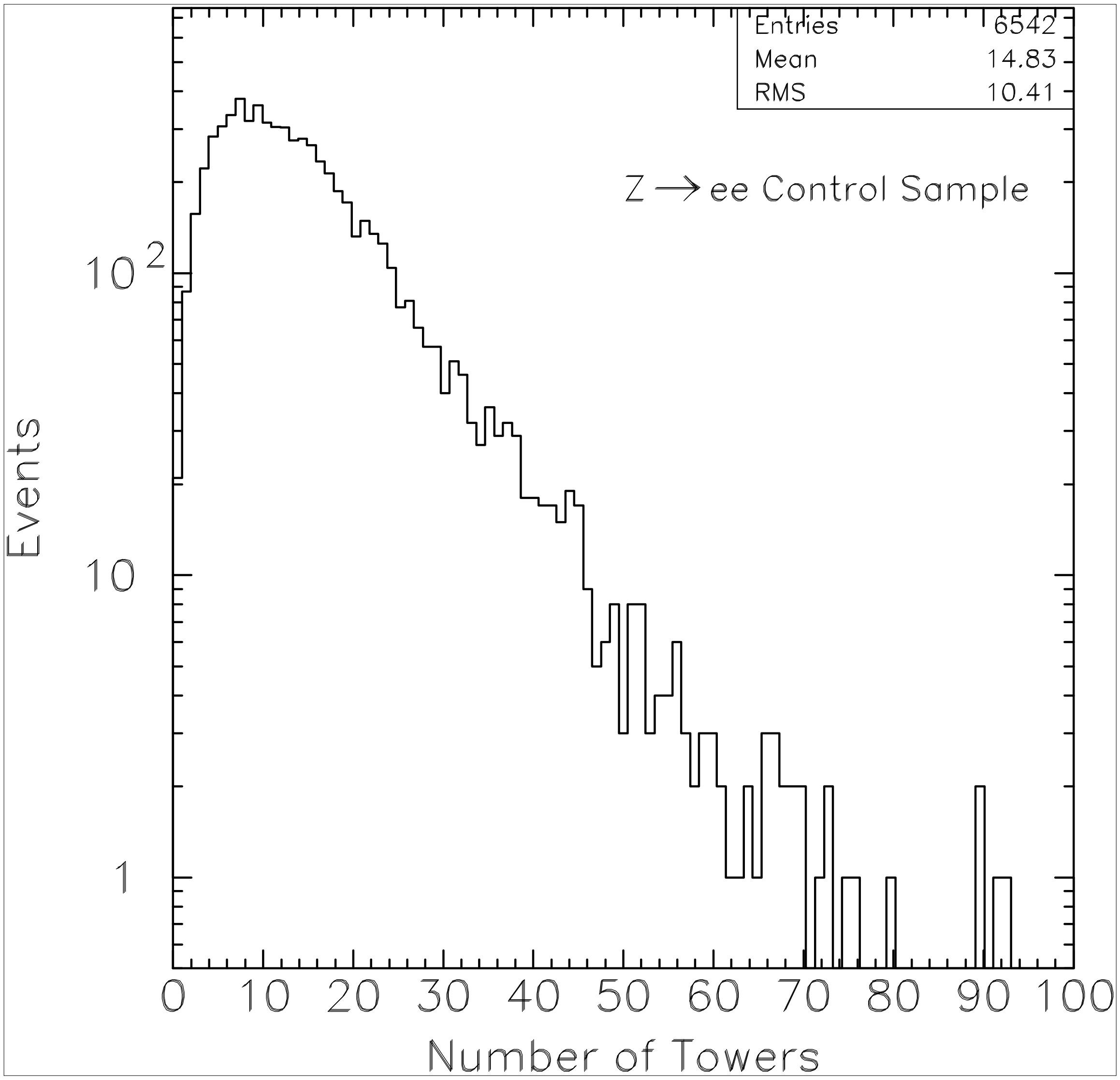


Figure 6: This figure shows the number of towers with a timing hit in each event in a sample of Z=>ee events. Note that while the mean is low, ~10 towers/event, the variation and tails are large and there are events with almost 100 towers.

3. A New Method of Rejecting Out-of-Time events

In this section we propose a method of using the HADTDC to search for hints that there is out-of-time energy which allows us to deal with the particulars of the low-energy timing distributions as well as take into account the number of towers hit in the event. To do this we use a Sleuth-like [6] methodology and algorithm which has been shown to be successful at identifying events which are unlike the expected backgrounds (which in this case means "unlike events from the primary collision"). For each event we find the tower which is the most unlikely (lowest probability) to be from a SM collision (estimated from Z=>ee events), and then find the fraction of Hypothetical Similar Experiments (HSE) which would produce this low a probability hit or lower. In other words, we are looking for the single tower which is most likely to indicate that this event is a cosmic (pick the

“° wors”¡± t ove, and then ask how likely is this tower to be a fluctuation. To do this we take into account the energy dependent timing distributions as well as the number of towers hit on an event by event basis. The advantage of this methodology is that it allows us to have a fixed efficiency for each event which eliminates bias against models of new physics such as SUSY.

1) The first step in the algorithm is to take each hit in the event and using its energy, the corrected time and the detector it is in find the fraction of SM hits which would arrive at this time or later. In other words, convert the time of arrival into a probability. As an example, consider a tower with E=0.2 GeV in the CHA with corrected time = 25 nsec. This situation is shown in

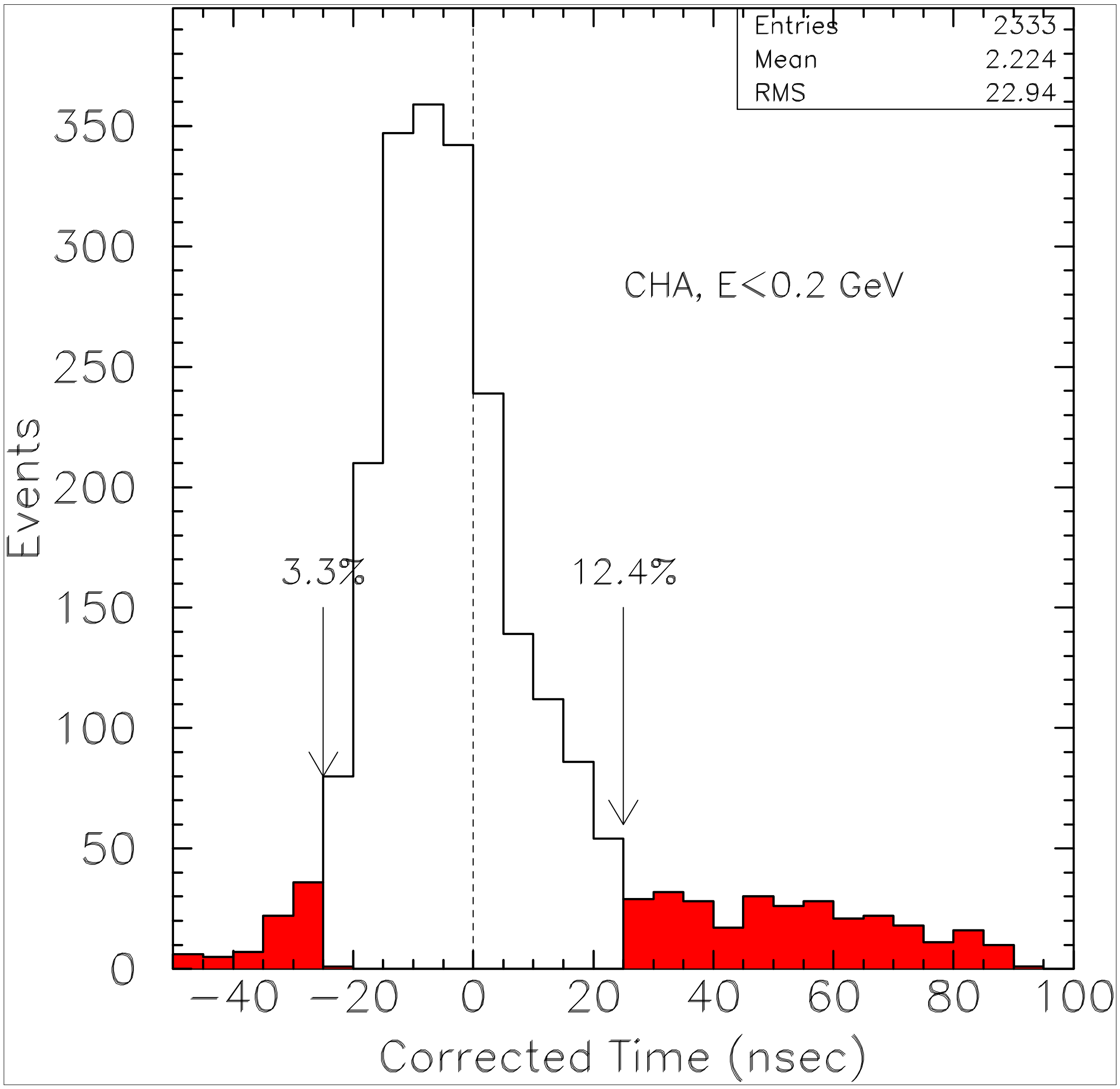


Figure 5. As shown in the figure 12.4% of hits from SM sources are at this time or later. Similarly, if the corrected time were -25 nsec, only 3% of hits have this time or earlier. Thus, any tower from a SM source, by construction, has a flat probability between 0 and 50%.

2) Loop over all towers to find the one which is most unlikely to be from the SM i.e., find the minimum probability. It is important to again note that this probability is very dependent on the number of towers in the event with timing hits. The top part of

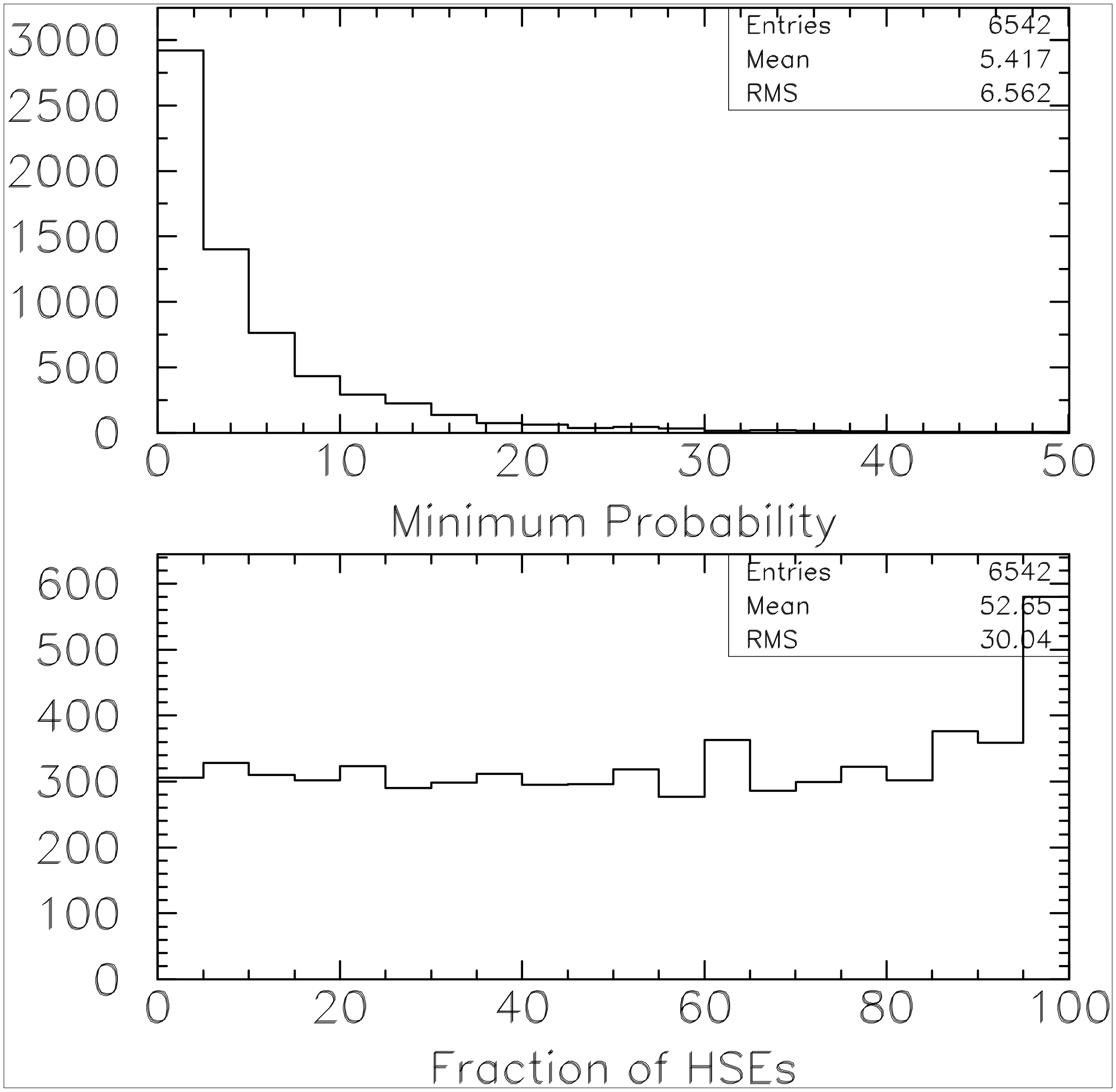


Figure 7 shows the distribution of the minimum probability tower hit in the event from a sample of Z=>ee events which is dominated by SM collisions. This distribution is very dependent on the number of towers in the event (see

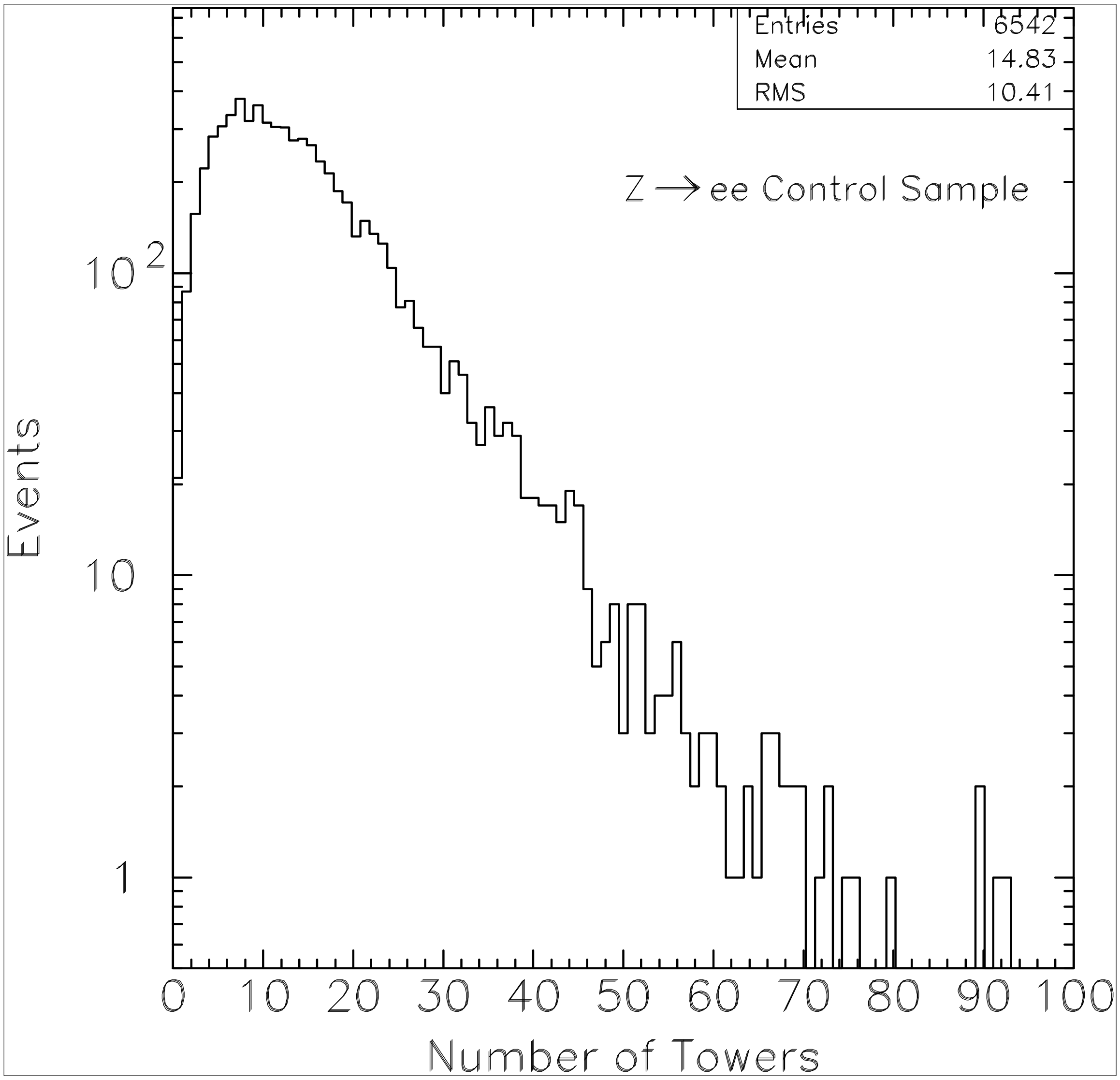


Figure 6) and shows this bias by peaking at low probability.

3) Next we remove this bias on an event-by-event basis by asking the question "° what fraction of Hypothetical Similar Experiments (HSE'¯s, with the number of observed towers in the event, would produce the observed minimum probability or lower." Since the expected probability distribution is flat between 0% and 50% we can answer the question by running a toy MC to simulate the probability and the number of towers. After this "°un wei ghti n"¡± o"°t aki ng i nt o account of t he tri als factor"± he result for the Z=>ee sample is shown in

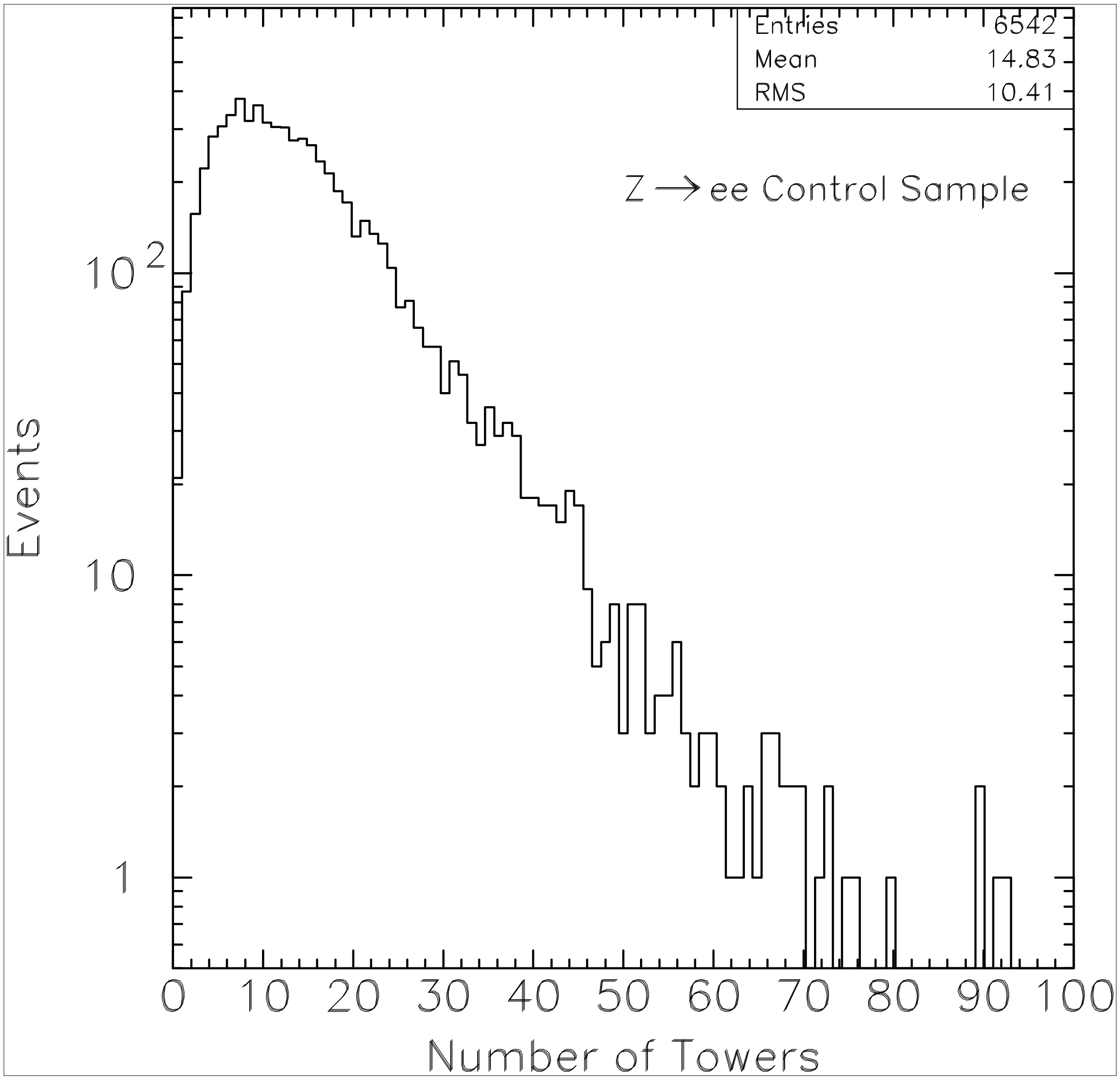


Figure 6. As expected it effectively gets rid of NTower dependence which is shown as the distribution is essentially flat for SM between 0%-100%.

This construction has the advantage that a cut on the fraction of HSE’¯s gives the efficiency of the cut. Specifically a cut at 3% is ~97% efficient and is true on an event-by-event basis, and is not model dependent. It is true that our sensitivity to out-of-time energy goes down as a function of the number of towers (in other words, as the number of towers goes up, a tower has to be more out and more out-of-time for it to be identified as anomalous) however, this has to be true; we cannot keep the single tower efficiency and the event efficiency fixed. We have chosen to make the later true for simplicity in estimating limits, as well as not being biased against high q^2 production.

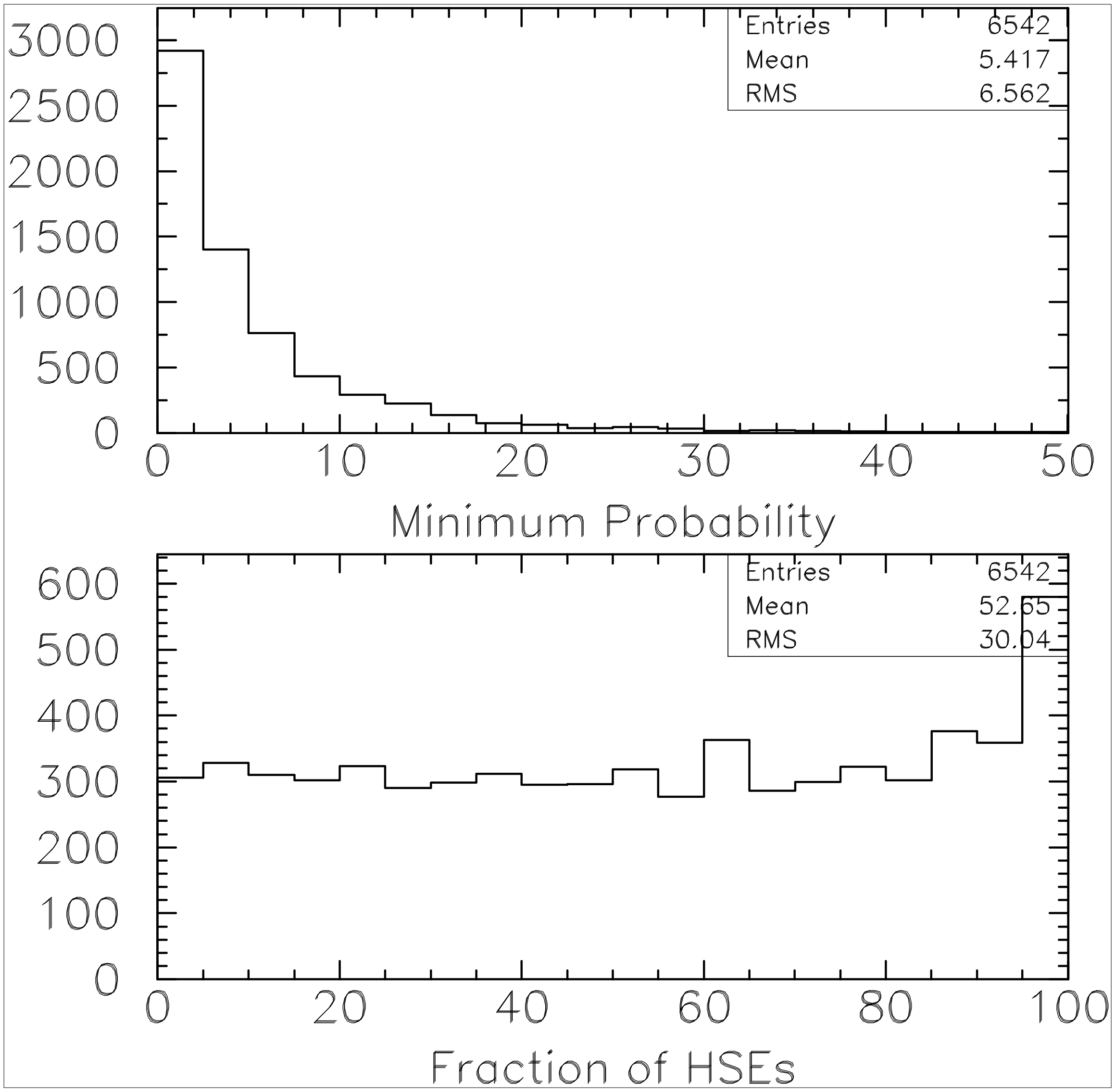


Figure 7: The top figure shows the distribution of the minimum probability tower hit in the event from a sample of Z=>ee events. This distribution is very dependent on the number of towers in the event (see

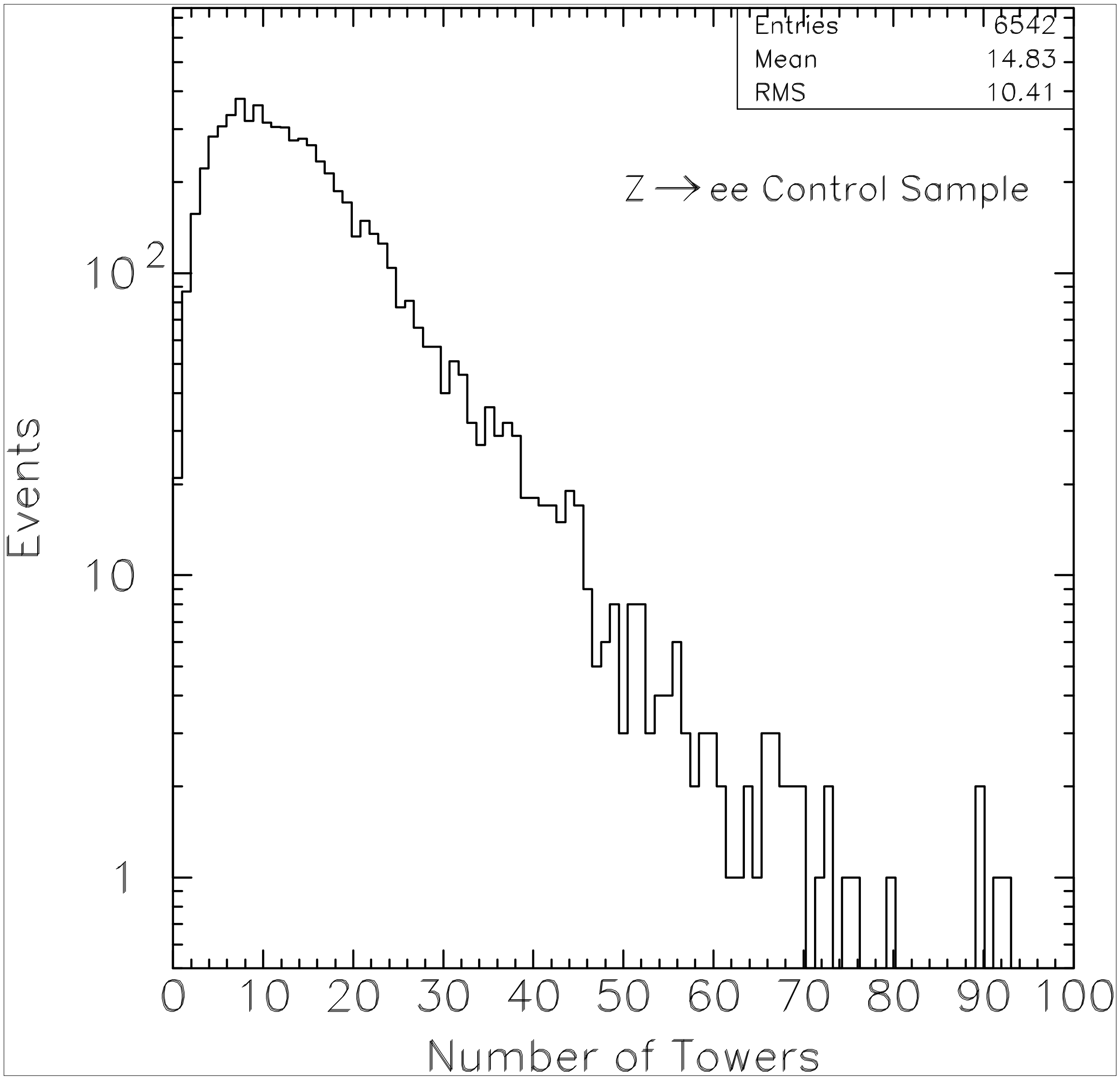


Figure 6) and thus peaks at low probability. The bottom figure shows the fraction of Hypothetical Similar Experiments (HSE’¯s) which, for the same number of towers, has the observed probability in the event or lower. This effectively takes into account the “°trials facto” of the event, and makes the probability roughly flat in fraction of HSE’¯s.

4. Results

While the algorithm described in Section 3 appears to work as constructed for SM-type events it remains to shown how it works on samples which supposedly contain the backgrounds we are trying to reject. To do this we use a sample of isolated photon+Met events with photon E_T>25 GeV and photon E_T/Met>0.9 which is dominated by out-of-time backgrounds.

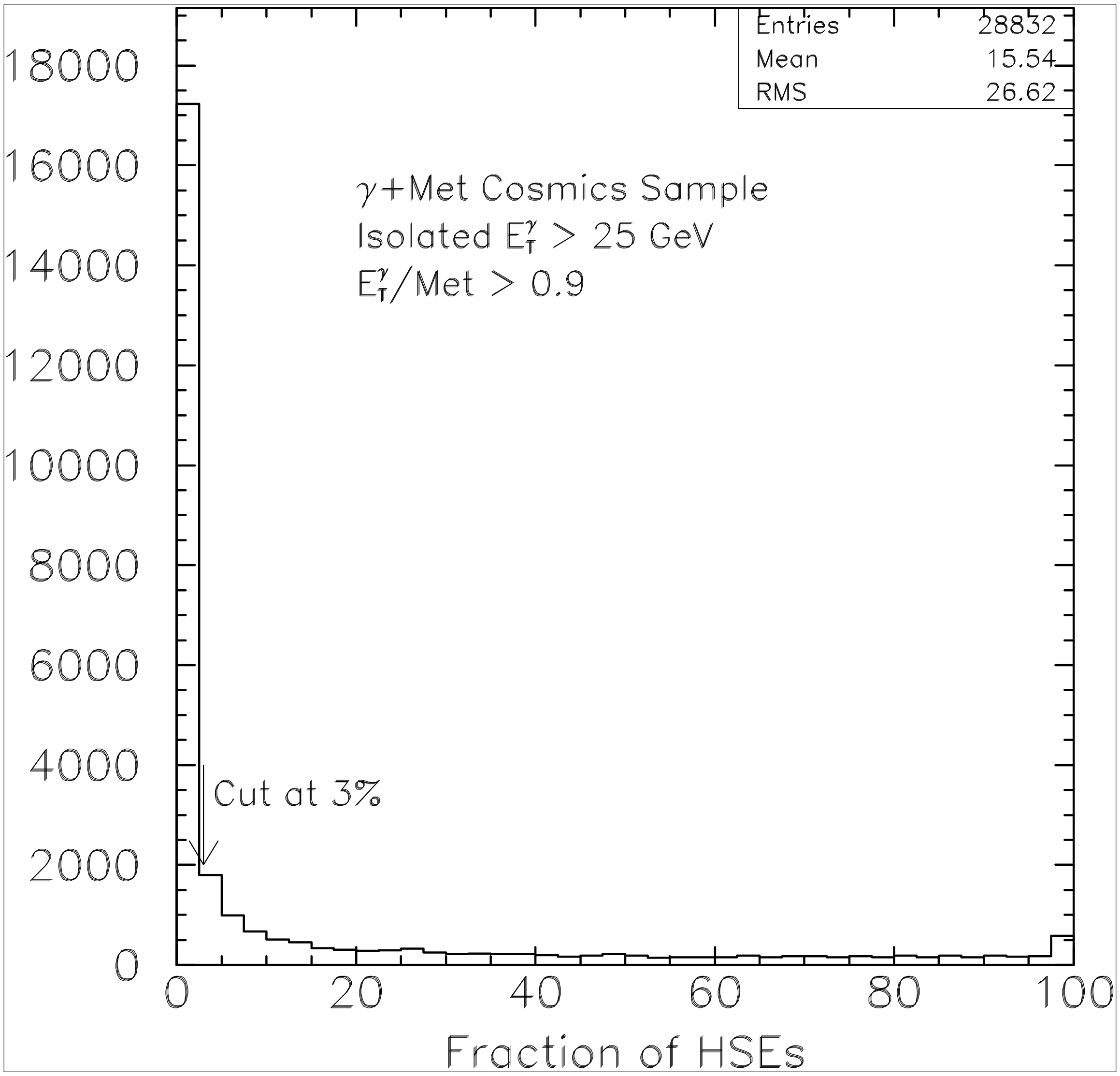


Figure 8 shows the fraction of HSE'¯ for the sample. The distribution is dominated by events with low probability as expected for events with real out-of-time energy; 80% of the events have a value less than 3%. While there is a tail to high probabilities (consistent with SM contamination in the sample), the algorithm and a 3% cut does an excellent job of efficiently separating events with real out-of-time energy and those events from prompt production. We also use this sample to test our hypothesis that we need to consider low energy towers in the sample.

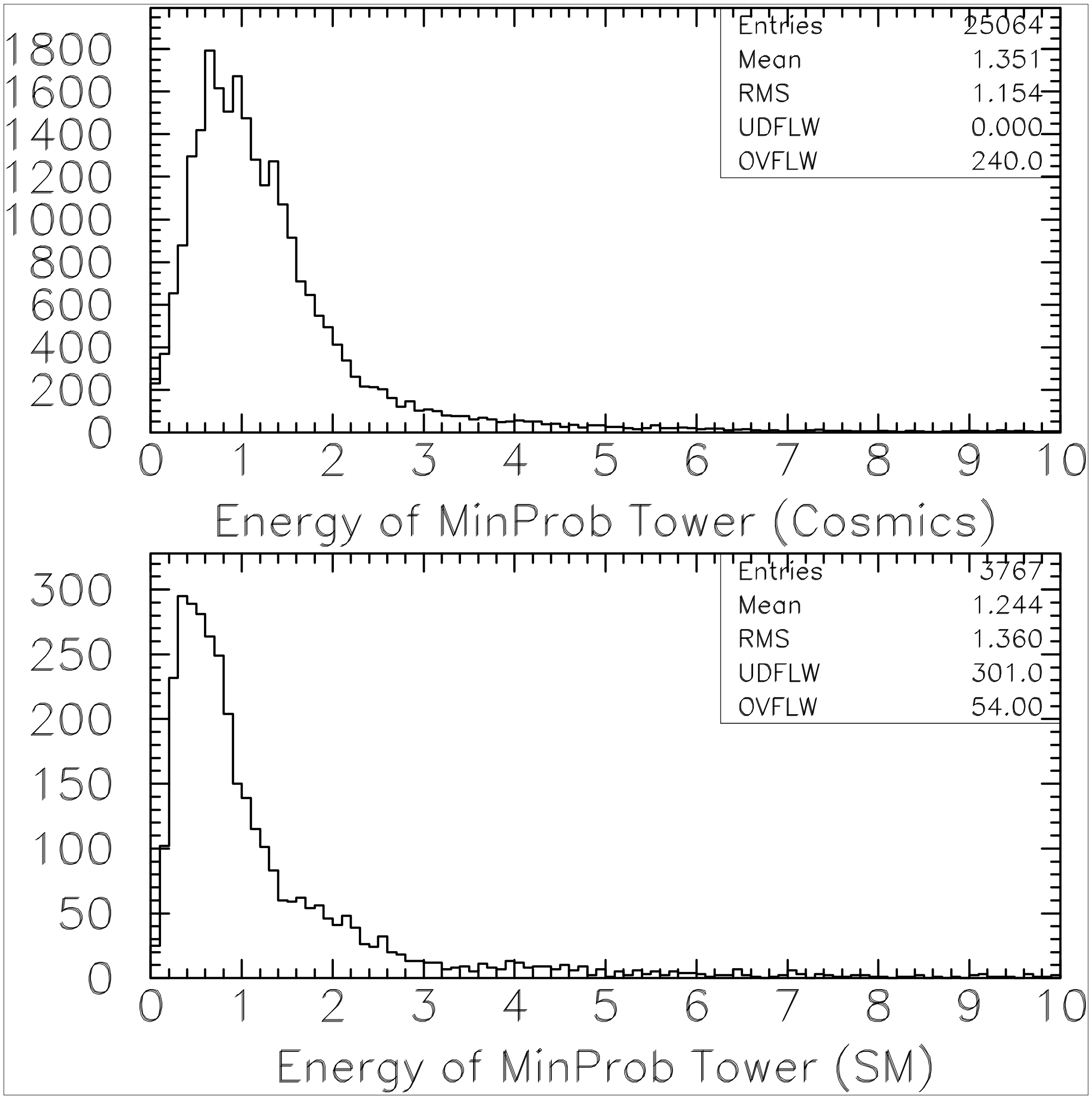


Figure 9 shows the energy of the tower selected by the algorithm as being the most statistically likely to be out-of-time. The top plot, from the events with less than 3%, are dominated by towers with E<2 GeV which justifies our efforts to consider low energy towers.

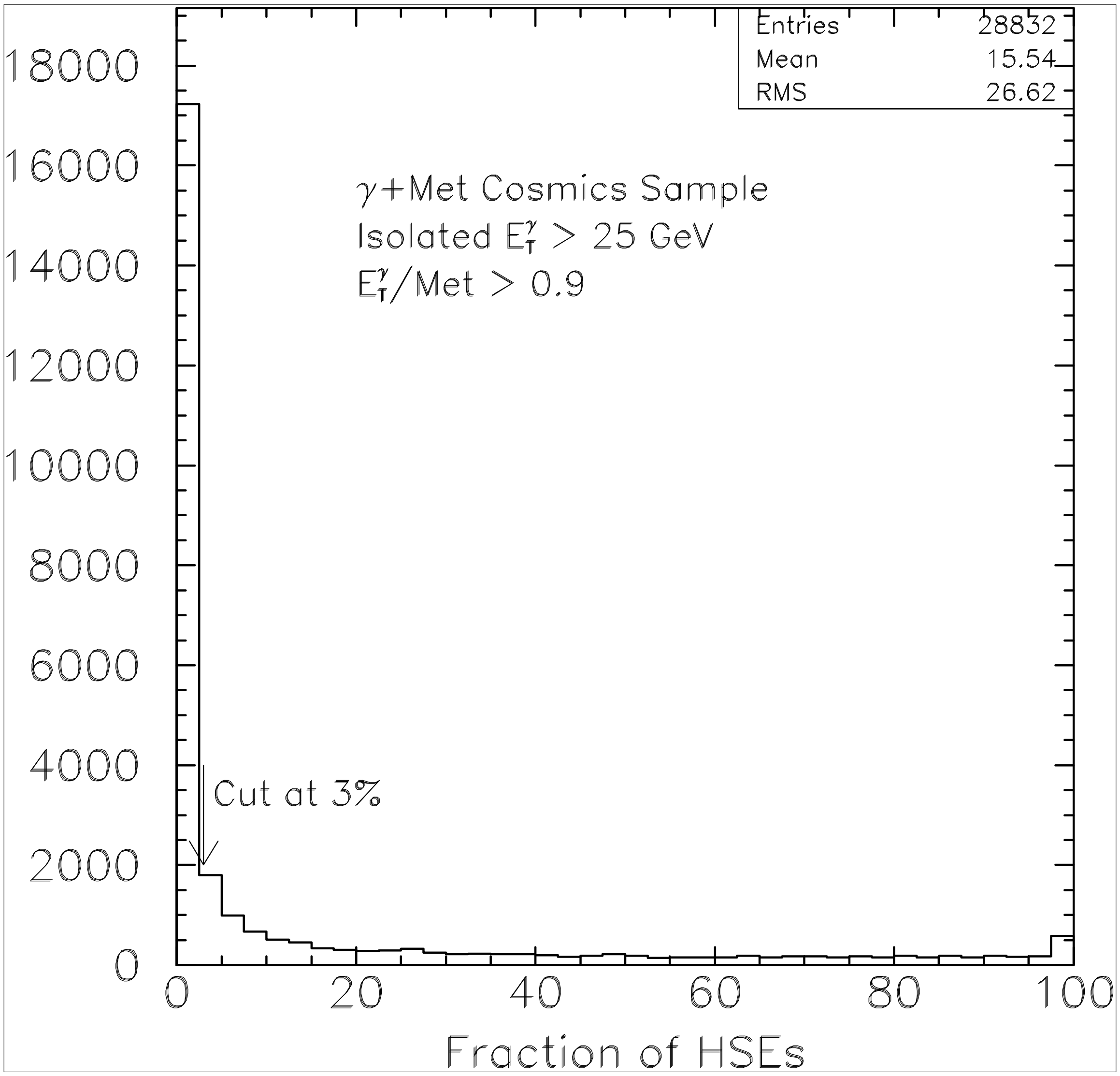


Figure 8: This figure shows the fraction of HSE’s, as estimated by the algorithm for our photon+ Met sample which is dominated by out-of-time events. The distribution is dominated by events with low probability as expected for events with real out-of-time energy; 80% of the events have a value less than 3%. We also note that long tail to high probabilities is consistent with SM contamination in the sample.

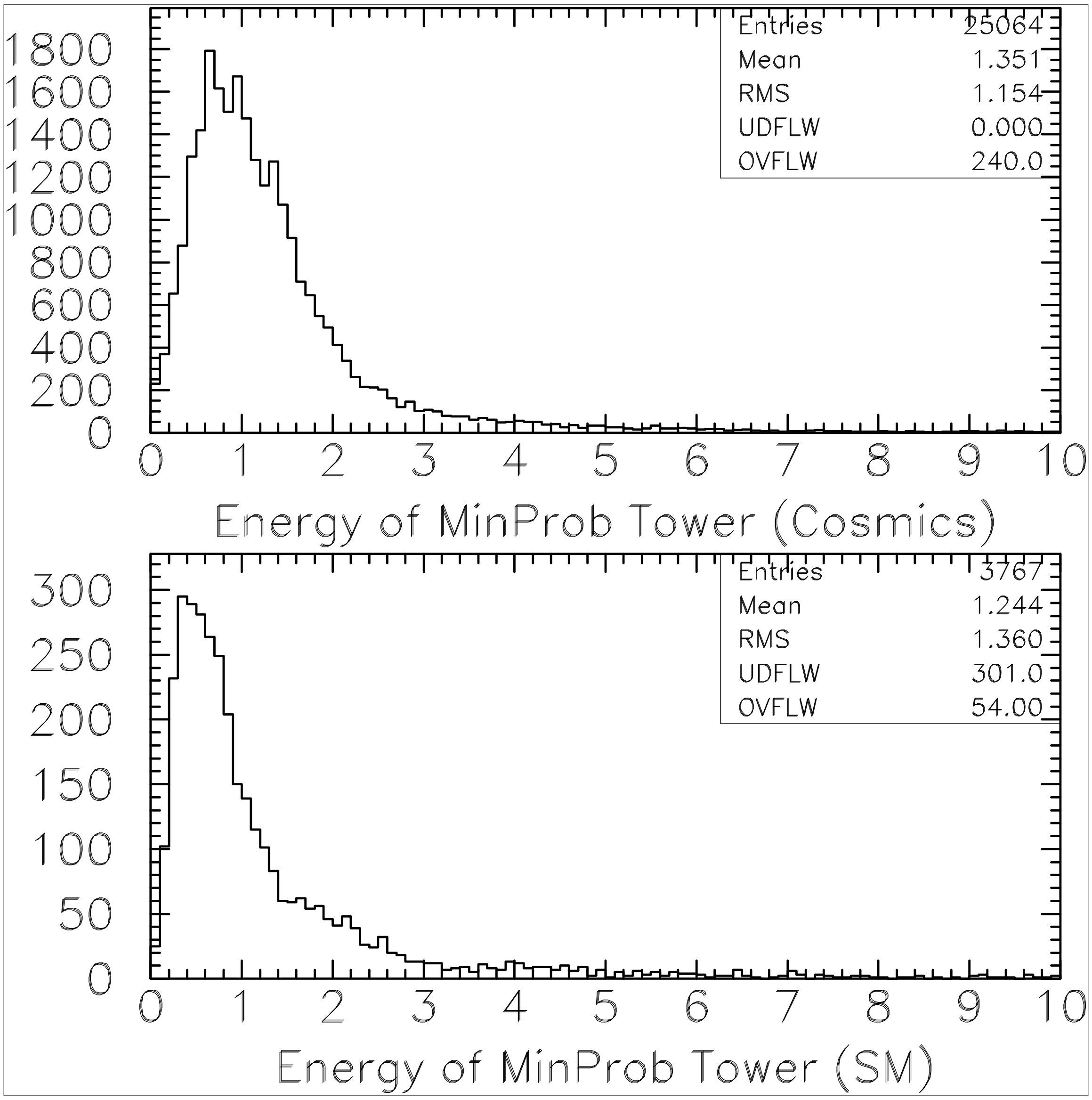


Figure 9: This figure shows the energy of the minimum probability tower identified by the algorithm in the photon+Met sample. The upper histogram is for events which are identified as cosmics (<3%) and the lower is for the rest of the events which are likely to be SM-like. The distribution clearly justifies including low energy (E<2 GeV) towers in our search for indications that events with photos+Met are contaminated by cosmics.

4.1 Using the algorithm for the GMSB-SUSY γγ +Met search

Having shown that the algorithm appears to work as expected we apply it to the “°si debnd”± control sample used to estimate the SM backgrounds in the GMSB-SUSY search in the γγ +Met channel [2].

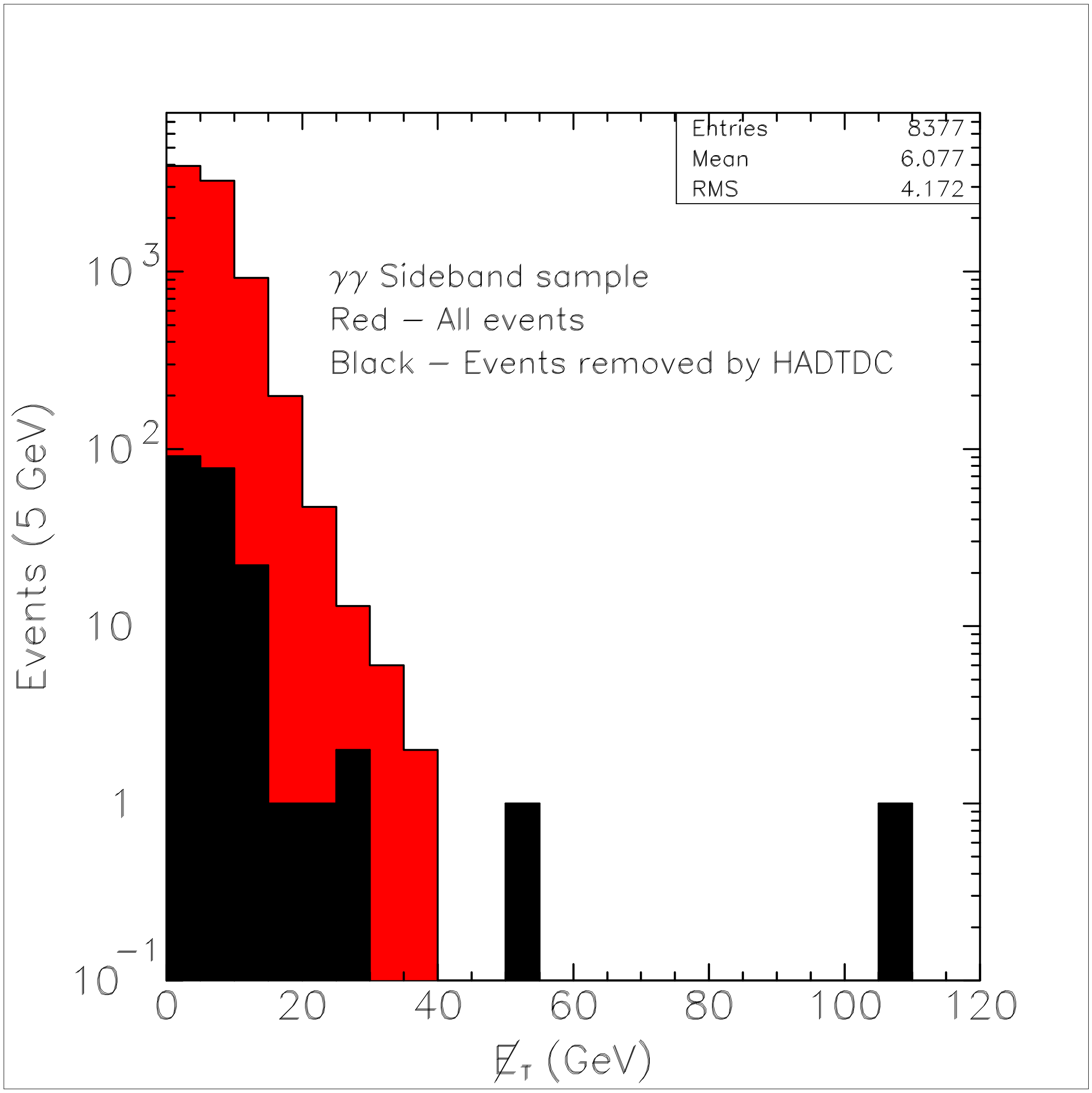


Figure 10: The Met distribution for the “°si deban”ûcontrol sample used to estimate the background shape for in the GMSB-SUSY γγ +Met search [2]. This sample is composed of events which just barely fail the photon ID cuts and is thus dominated by jet-jet and γ+jet events, and as such is expected to have the same Met resolution as the signal region events. The black distribution shows the events which fail the 3% cut. We note that the two largest Met events, which appear to be anomalous compared to the smooth resolution curve, are removed as having out-of-time energy.shows the Met distribution for a sample composed of events which just barely fail the photon ID cuts and is thus dominated by jet-jet and γ+jet events which are expected to have the same Met resolution as the signal region events. The two largest Met events, which appear anomalous, are removed by the 3% cut and the remaining distribution appears well modeled by a smooth Met resolution in the detector. This gives us reason to believe that the sideband sample, after the out-of-time cuts, is largely contamination free and is useful as a background estimate for the search. Since the search results themselves are outside of the scope of this note, we do not show results but

note that there are no events in the final $\gamma\gamma$+Met sample which appear anomalous and/or are from cosmics/beam halo/beam-gas etc.

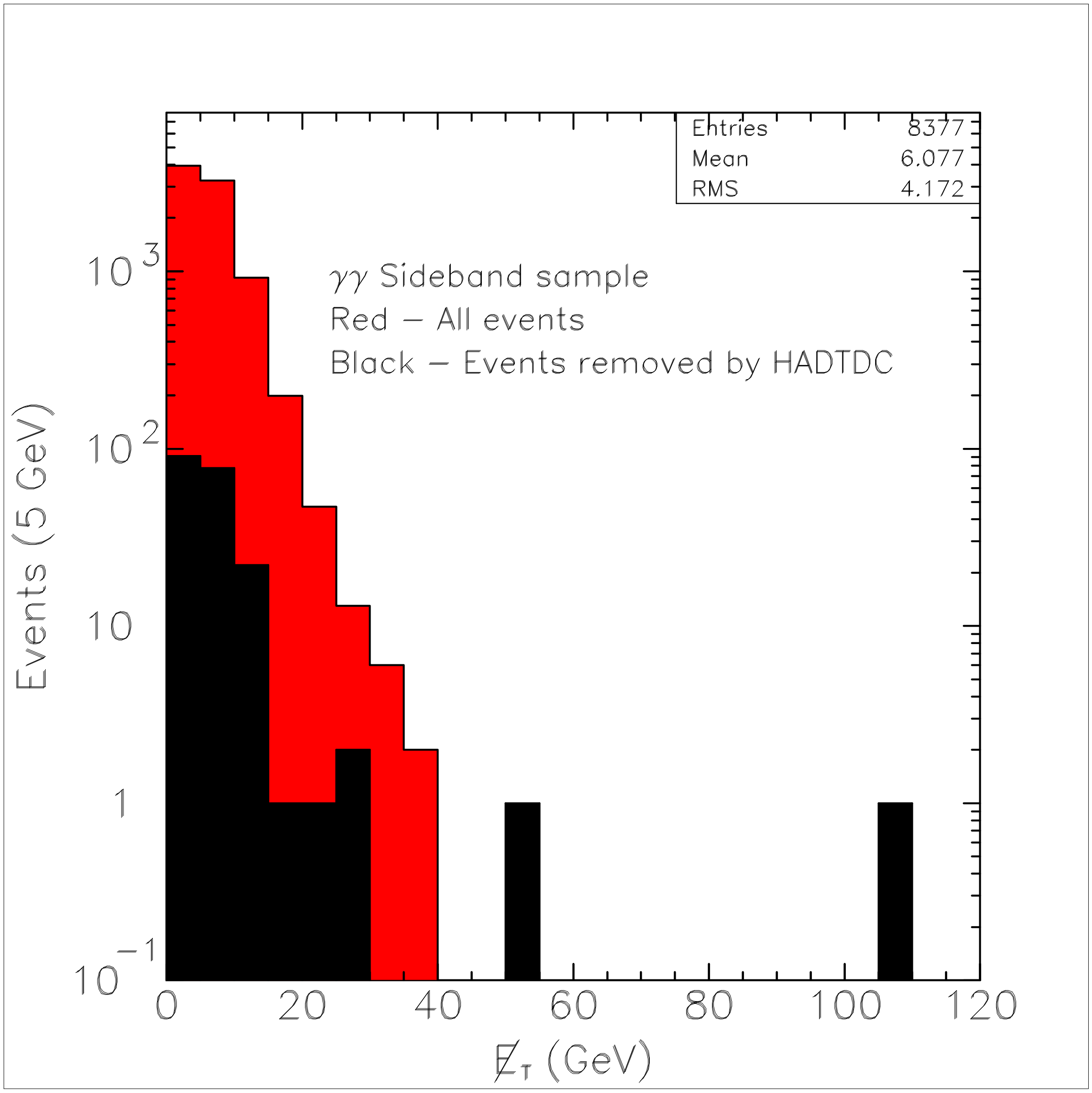


Figure 10: The Met distribution for the “°si deban”ïcontrol sample used to estimate the background shape for in the GMSB-SUSY γγ +Met search [2]. This sample is composed of events which just barely fail the photon ID cuts and is thus dominated by jet-jet and γ+jet events, and as such is expected to have the same Met resolution as the signal region events. The black distribution shows the events which fail the 3% cut. We note that the two largest Met events, which appear to be anomalous compared to the smooth resolution curve, are removed as having out-of-time energy.

5. Conclusions

The HADTDC system can be a very powerful tool in rejecting out-of-time backgrounds to many kinds of large Met searches. Unfortunately, until the EMTiming system is installed any attempt to use this tool for searches with final state photons must look for hints rather than looking directly at the large energy in the event because it is possible that large out-of-time is deposited in the EM. Doing so requires an analysis which studies lots of low energy hits in the detector where the response has poorer resolution as well as asymmetric tails. We have presented a methodology which treats all hits at all energies, in every detector, on an equal probability footing and uses a Sleuth-like methodology to systematically search for out-of-time hits and, by taking into account the number of hits in the event, quantifies the probability that the event is likely to be from a SM event. We have shown that the algorithm behaves as expected on SM samples, giving a 97% efficiency, and that it does an excellent job of rejecting events with known out-of-time contamination; rejecting 80% of photon+Met events. We note that when EMTiming is fully instrumented, expected Fall 2004, we can stop looking for hints as any large energy in the calorimeter will be timed. Thus, future versions of this analysis will only need consider towers with high energy which have excellent timing resolution, and by reducing the number of towers we same algorithm effectively has a more restrictive timing window (better rejection) for the same efficiency.

Acknowledgements

The author would like to thank Ray Culbertson for providing the datasets and run/event lists for samples used in this analysis, Slava Krutelyov for stripping off data to make small and easily useable subsets, and Vadim Khotilovich for running the analysis jobs which unpacked the energy and timing information for each event. The author would like to thank Fotis Ptohos and Toni Munar for their help understanding the data, as well as Teruki Kamon and Max Goncharov for helping interpret it and providing comments.

Appendix A: Technical details on getting to the HADTDC data as well as applying the proper corrections

In this Appendix we outline some of the technical details for getting the corrected time for each HADTDC tower as well as how we use it in the analysis. Specifically we:

1) Use Pass 13 of the calorimeter calibration constants
2) Remove towers with the BADTDC bit not set to zero. This bit is now in the HADTDC database and reflects whether there is a good χ^2 in the efficiency calculation as a function of energy which is an excellent measure of if the tower is functioning properly.
3) Ignore towers with zero energy (no energy correction possible)
4) Use version 5.3.1pre2 or later of CalData to get proper correction of towers with only one PMT firing (e.g. Chimney in the CHA and dead PMTs in the WHA)
5) We only consider times which have been shown to be consistent with being during the energy integration window [7], as energy deposited outside the integration window can't affect the Met or any of the objects. The times ar
    a) -50 nsec <CHA< 100 nsec
    b) -30 nsec < WHA< 110 nsec
    c) -40 nsec < PHA < 90 nsec
6) We always use the corrected time which is closest to 0.0 and ignore any other hits in the integration window. It is true that we would like to reject events where there is an indication of getting energy deposited twice in the same integration window. Unfortunately, there are reflections on the transition boards which cause fake double-firing (centered at ~70 nsec) for large, but not too large energies. This is measured to be 1% for CHA and 6% for WHA which is too large an inefficiency to take for our analyses. This has been fixed for EMTiming.

Appendix B: Questions which we are still investigating, or thinking about investigating

Question: What is the peak at large fraction of HSE’¯s n

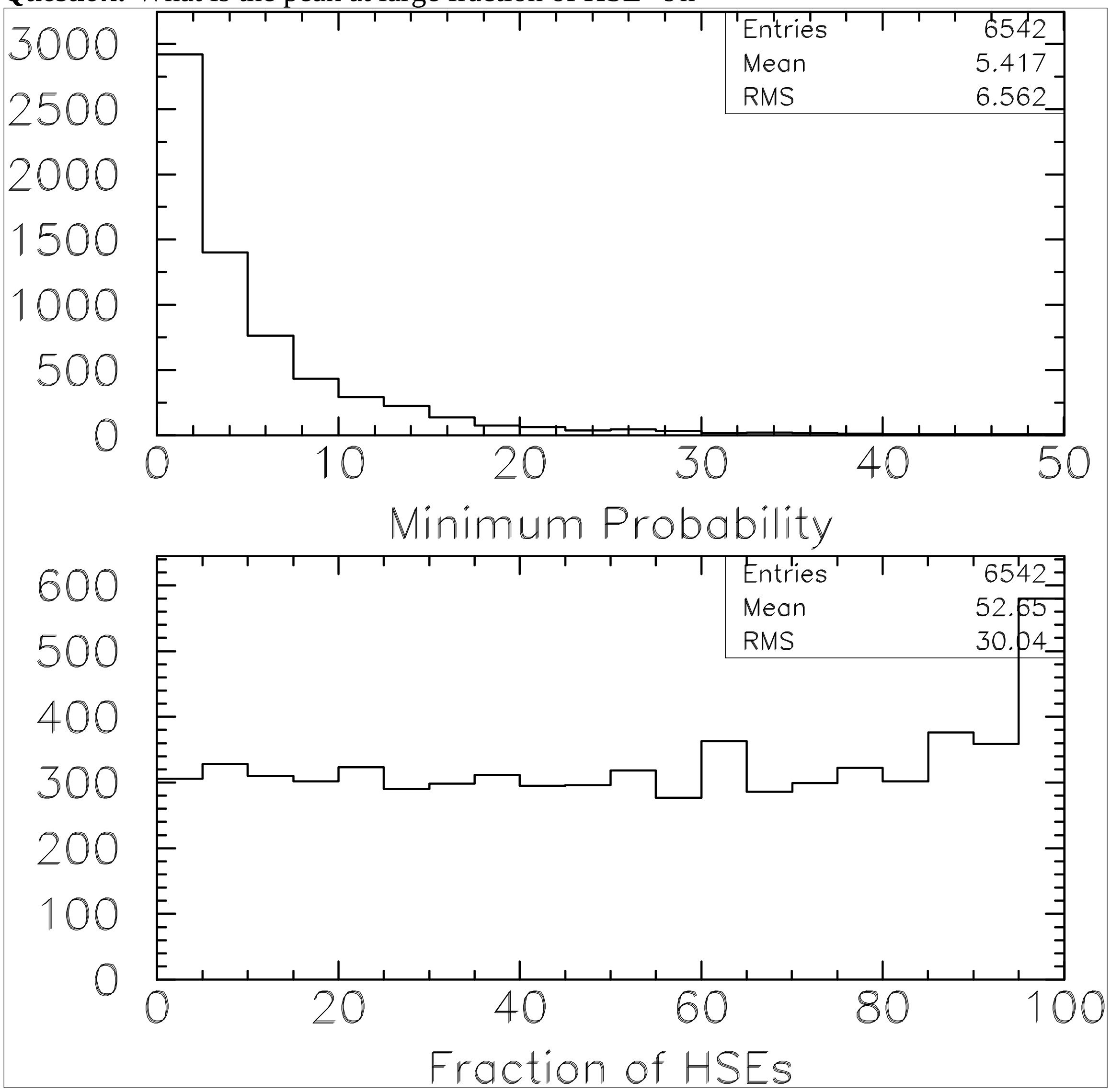


Figure 7?

Answer: We’¯re not completely sure. Just to be clear, this is an excess of events with a set of towers which are so far from the edges of the effective timing window that greater than 99% of HSEs give a tower with a lower probability than the one that is observed. We have noticed that there is some variation in the shape of the HSE as a function of run number which indicates that that the HADTDC system is not uniform as a function of time. For example, if there were a number of noisy towers at the beginning which were fixed, that would mean that the times (probabilities) from those towers would be systematically more probable in later runs, thus making them less likely to give a tower which has a small probability and gives a bias for large fraction of HSEs.

Question: Why don’t you do a likelihood, Kolmogorov or Chi2 test on the entire event rather than just use the tail? I could imagine that there are events with lots of towers which are barely out-of-time, but none which is significantly out-of-time. Don’t you want to reject those as well?
Answer: This is a good question and it is under study. We believe there is nothing wrong with what we’ve done, but the proposed idea could certainly work more effectively in certain cases and might provide additional rejection power, albeit with currently unknown efficiency costs.

References

[1] D. Toback et al., CDF Note 5518. For more up-to-date details see http://hepr8.physics.tamu.edu/hep/emtiming
[2] M.S. Kim et al., CDF Note 6389.
[3] M. Spiropulu, CDF Note 5448.
[4] M. Cordelli et al., CDF Note 5856.
[5] While there are lots of analyses which use this type of analysis. See, for example, D. Toback, CDF Note 4582, or S. Kopp CDF Note 2541 for two examples with large EM fraction.
[6] B. Knuteson and D. Toback DZero Note 3686. See B. Abbot et al., (DZero collaboration) Phys.Rev.D 62, 092004 (2000) hep-ex/0006011.
[7] Taken from Tony Munar. Details at http://www.hep.upenn.edu/~munar/cdf/out_time/out.html