DØ note 3629
June 13, 2000

$p(\not\!\!E_{\rm T})$
The missing transverse energy resolution of an event

Bruce Knuteson, Mark Strovink

University of California at Berkeley
Lawrence Berkeley National Laboratory

Bob Olivier, Ursula Bassler, Frederic Fleuret, Gregorio Bernardi

LPNHE
Universites Paris VI and VII, France

Abstract

The missing transverse energy resolution of the DØ detector is determined by the convolution of a number of effects: misvertexing, energy resolution, and hot cells are among the most prominent. Since the potential contribution of these effects to the missing transverse energy varies greatly from one event to the next, we consider the possibility of calculating for each event a probability distribution $p(\not\!\!E_{\rm T})$. This note describes how $p(\not\!\!E_{\rm T})$ can be calculated, and how it can be used once we have it.

Contents

1 Introduction

This is an idea that might be useful in understanding the dependence of the calculated missing transverse energy on z vertex position, energy resolution, and the existence of hot cells. The desired object $p(\not\!\!E_T)$, which we will calculate in section 3, is simply the probability density for the missing transverse energy in an event. That is, for a given event, the probability that the actual missing transverse energy in that event is between $\not\!\!E_T$ and $\not\!\!E_T+\delta\not\!\!E_T$ is given by $p(\not\!\!E_T)\delta\not\!\!E_T$.

2 Probability densities

We begin by calculating several probability densities. The first is $p(z)$, the probability density for the z position of the vertex. The remaining are $p(E_i)$, the probability densities for the energies of the physics objects in the event: jets, electrons, muons, and the unclustered energy.

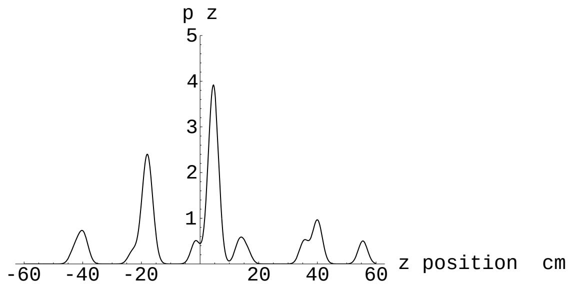

Figure 1: Plot of a probability density $p(z)$. This particular event has a primary vertex at $z \approx 5cm$ and a secondary vertex at $z \approx -18cm$. The resolution of the z-intercept for each track is assumed to be about $\sigma = 1.5cm$ for the purposes of this plot.

2.1 Probability density for the z-vertex: $p(z)$

The probability density for the z vertex position could be constructed in several ways, each of which require knowledge of the N_t tracks found in the event. For each track i one can calculate the position μ_i at which the track r-z projection crosses the z-axis. One can also calculate a resolution σ_i for each μ_i, which represents an uncertainty in the μ_i calculated. For example, one would expect the z-intercept of a forward-going track to be known less well than the z-intercept of a more central track. This resolution may be determined from errors on the fit of the track, using information contained in the ZTRK block in the μDSTs.[1]

Since the primary vertex is by definition the vertex with the most associated tracks, we might begin by taking

$$p(z) = \frac{1}{N_t} \sum_{i=1}^{N} K(\mu_i, \sigma_i, z), \tag{1}$$

where

$$K(\mu, \sigma, z) \equiv \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{(z-\mu)^2}{2\sigma^2}\right) \tag{2}$$

is a gaussian kernel centered at the position μ with width σ. We might expect $p(z)$ to look something like the graph shown in Figure 1.

A more sophisticated method of determining $p(z)$ might involve weighting each of the gaussian bumps in (1). We would like to weight a track associated with an electron or muon in an event more heavily than a track associated with a jet, which we would in turn weight more heavily than a track not associated with any physics object in the event. With this modification (1) becomes

$$p(z) = \frac{1}{N_t} \sum_{i=1}^{N_t} w(o_i) K(\mu_i, \sigma_i, z), \tag{3}$$

where o_i is the physics object associated with the track i, and $w(o_i)$ is the weight to be associated with that physics object. We might introduce four weights: w(electron), w(muon), w(jet), and w(none).

2.2 Probability densities for energies: $p(E_i)$

The remaining probability densities we will need are the densities $p(E_i)$ for the energy E_i of the i^{th} physics object in the event. Explicitly, we want to construct $p(E_i)$ such that the probability that the i^{th} physics object has an energy between E_i and $E_i + \delta E_i$ is $p(E_i)\delta E_i$.

2.2.1 Jets

We might handle the uncertainty in calculated energy μ_i for jets by smearing this value across a gaussian of width

$$\sigma_i = \frac{80\%}{\sqrt{\mu_i}} \mu_i, \tag{4}$$

so that

$$p(E_i) = K(\mu_i, \sigma_i, E_i). \tag{5}$$

The expression for σ_i should be chosen to roughly reflect the uncertainty in energy in the hadronic calorimeter at DØ.

Since our energy resolution also depends upon η_{det}, we can consider a function $\sigma_i(E_T, \eta)$ which depends upon both the transverse energy deposited and the location in the calorimeter. This study has been done for the measurement of the triple differential dijet cross section, and we should be able to use their results with little modification.[4]

2.2.2 Electrons

We should be able to apply the same algorithm to electrons in the event, though in this case the smearing should use a gaussian of a smaller width

$$\sigma_i = \frac{15\%}{\sqrt{\mu_i}} \mu_i. \tag{6}$$

2.2.3 Muons

We had rather poor muon momentum resolution in Run I, and we should take this into account as well. The muon resolution is typically described in terms of a gaussian in the reciprocal of the muon momentum, so we would have, for a muon with measured reciprocal momentum μ_i,

$$p(E_i) = K\left(\mu_i, \sigma_i, \frac{1}{\sqrt{E_i^2 - m_\mu^2}}\right), \tag{7}$$

where σ_i is the error on the reciprocal momentum.[1]

2.2.4 Hot cells

We might take hot cells into account in a similar fashion. A useful modification to AIDA would involve calculating a probability that a given cell is a hot cell. It could then be used as it is now, to simply flag cells which have a high probability of being a hot cell, or it could return in addition this calculated probability.

Assume that for the i^{th} physics object (call it a jet) we have calculated $p(E_i)$ as above without removing any hot cells. Now say that our modified AIDA determines that one of the cells within the jet is a hot cell with probability $P_{hotcell}$, and that the energy deposited within this cell is $E_{hotcell}$. Then we modify our probability distribution according to

$$p(E_i) \to p'(E_i) = (1 - P_{hotcell})p(E_i) + P_{hotcell}p(E_i + E_{hotcell}). \tag{8}$$

As a sanity check, we note that this procedure behaves nicely in the limits $P_{hotcell} \to 0$ and $E_{hotcell} \to 0$.

A sample graph of what $p(E_i)$ might look like, where a hot cell (with 30% probability) of 5 GeV/c^2 has been included, is shown in Figure 2.

It may be that there are very subtle correlations among hot cells in events which we should include in our calculation of the hot cell probability. At the very least, we expect some cells to be hot more frequently than others. We may determine a "prior probability" for a cell to be a hot cell by going back through the Run I data and counting the number of times each calorimeter cell was hit. A graph of the frequency of hits for each of the cells in the calorimeter might give us some insight into how to model this problem.

Another way to treat the hot cell problem is to see if we can directly improve upon the AIDA algorithm. Such an approach is described in [2]; the authors have graciously made the code and ntuples used in that analysis accessible for our studies.

[1] $\sigma = 0.18(p-2)/p^2 \oplus .008$, where the $\oplus$ means add in quadrature.[3]

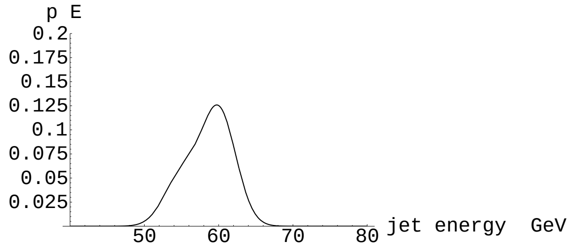

Figure 2: Example of a probability density $p(E)$ for a 60 GeV jet containing a hot cell (with 30% probability) of 5 GeV. The small hump to the left of the peak is the result of taking this hot cell into account.

2.2.5 Unclustered energy

Some of the energy in an event remains unassociated with physics objects, and is called *unclustered energy.* The distribution of unclustered energy in the calorimeter exists in STAs, but not in the μDSTs we plan to use in our analysis. We may obtain the transverse component $\vec{E}_T^x$ of the unclustered energy by simple subtraction:

$$\vec{E}_T^x = \vec{\not{E}}_{\rm T} - \sum_i \vec{E}_{T_i}, \tag{9}$$

where $\vec{E}_{T_i}$ is the transverse energy of the i^{th} physics object in the event. We can then treat this unclustered energy as an additional physics "object".

In particular, we can recalculate the transverse energy $\vec{E}_T^{x\prime}$ of the unclustered energy under a change in vertex $z \to z' = z + \delta z$ in the same way that we would recalculate the transverse energy of a standard physics object. Of course, $\vec{E}_T^x$ does not in itself contain enough information to compute the new $\vec{E}_T^{x\prime}$; this computation requires knowing the energy deposited in each unclustered calorimeter cell. But computing the new $\vec{E}_T'$ of any physics object requires knowing the distribution of energy deposited along the object's path, so some assumptions will need to made as to the form of all of these distributions.

We might apply the *maximum entropy* principle to estimate the distribution of unclustered energy. That is, we assume that the unclustered energy is distributed roughly uniformly about the center of mass of the event, subject to the constraint that the sum of unclustered transverse energy is equal to $\vec{E}_T^x$.

3 Probability density for $\not{E}_{\rm T}$: $p(\not{E}_{\rm T})$

Once we have determined $p(z)$ and the $p(E_i)$, we can use these to calculate the probability density $p(\not{E}_{\rm T})$ for the missing transverse energy. Explicity, we

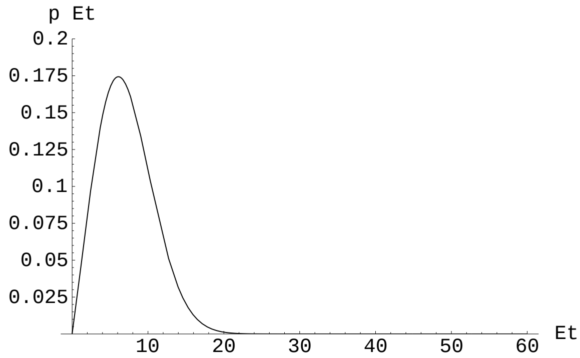

Figure 3: Graph of a sample probability density for the missing transverse energy for a dijet event. The corresponding two-dimensional probability density in the ${\not{E}_{\mathrm{T}}}_x$-${\not{E}_{\mathrm{T}}}_y$ plane is actually peaked at the origin; a Jacobian zero at the origin appears when one considers the one-dimensional scalar $\not{E}_{\mathrm{T}}$.

want this function to have the property that the probability that the missing transverse energy in the event is between $\not{E}_{\mathrm{T}}$ and $\not{E}_{\mathrm{T}} + \delta\not{E}_{\mathrm{T}}$ to be given by $p(\not{E}_{\mathrm{T}})\delta\not{E}_{\mathrm{T}}$.

This may be done by a Monte Carlo method as follows. Choose a value of z for the primary vertex randomly from the distribution $p(z)$. Choose values for the energies of the objects in the event randomly from the distributions $p(E_i)$. Calculate the missing energy in the event using these values of z and $\{E_i\}$. If we do this $N \approx 10^3$ times and stick the missing energies we calculate into a histogram, we will obtain an estimate of the probability density $p(\not{E}_{\mathrm{T}})$. Graphs of what this might look like for a dijet event and for a W event are shown in Figures 3 and 4. Just to emphasize, this is a probability density for the missing transverse energy for a single event.

4 Reducing $p(\not{E}_{\mathrm{T}})$ to variables of interest

Assume that with the procedure outlined above we have calculated N possible values of the missing transverse energy in some event; label these ${\not{E}_{\mathrm{T}}}_i$, with $i = 0, .., N$. If we fill a histogram with these values, that histogram may look something like the density shown in Figure 3 or 4.

4.1 Moments of the probability density

There are several ways we might then use these values ${\not{E}_{\mathrm{T}}}_i$. Two important features we can extract are the first two moments of the probability density we

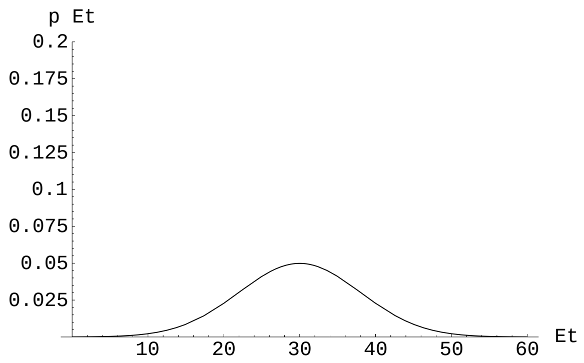

Figure 4: Graph of a sample probability density for the missing transverse energy for a W event.

have constructed. We could calculate the expected missing transverse energy in the event by

$$\langle \not{E}_{\mathrm{T}} \rangle = \frac{1}{N} \sum_{i=1}^{N} \not{E}_{\mathrm{T}i}, \tag{10}$$

or the uncertainty in the missing transverse energy by

$$\sigma^2_{\not{E}_{\mathrm{T}}} = \frac{1}{N-1} \sum_{i=1}^{N} |\not{E}_{\mathrm{T}} - \langle \not{E}_{\mathrm{T}} \rangle|^2. \tag{11}$$

This uncertainty, or spread, in the missing transverse energy could be a useful measure of how much we should trust the missing transverse energy calculated in the event.

4.2 $\not{E}_{\mathrm{T}}$ significance

We might also want some way of quantitatively determining the significance of the calculated $\not{E}_{\mathrm{T}}$: how likely is it that the missing transverse energy in an event is significantly different from zero? Mercedes-like events should have many of the $\not{E}_{\mathrm{T}i}$ close to zero; three-legged horse events will have very few of the $\not{E}_{\mathrm{T}i}$ close to zero. A quantity such as

$$p(\not{E}_{\mathrm{T}} > 10\,GeV) = \frac{1}{N} \sum_{i=1}^{N} \theta(\not{E}_{\mathrm{T}i} - 10\,GeV), \tag{12}$$

where $\theta(x)$ is the usual step function, would give the probability that the missing transverse energy in the event is greater than 10 GeV.

We could also use the likelihood

$$\mathcal{L} = \lim_{\not{E}_{T_1}\to 0} \frac{p(\not{E}_{T_2})/\not{E}_{T_2}}{p(\not{E}_{T_1})/\not{E}_{T_1}}, \tag{13}$$

where $\not{E}_{T_2}$ is the value of the missing transverse energy that maximizes $p(\not{E}_T)/\not{E}_T$, as a comparison of two hypotheses. The first hypothesis is that the $\not{E}_T$ is that given by the value at which $p(\not{E}_T)/\not{E}_T$ is maximum; the second hypothesis is that $\not{E}_T$ vanishes. The quantity $p(\not{E}_T)/\not{E}_T$ is used here in place of $p(\not{E}_T)$ to remove the Jacobian zero at the origin, noted previously in the caption of Figure 3. With this definition we have $1 \leq \mathcal{L} \leq \infty$.

The most convenient definition, however, uses the projection $\not{E}_{T_a}$ of the missing transverse energy along the axis a defined by the measured missing transverse energy, since this avoids the problems associated with the Jacobian zero at the origin. The definition of the likelihood used in our code and in the next section is

$$\mathcal{L} = \frac{p(\not{E}_{T_a})_{\max}}{p(\not{E}_{T_a} = 0)}. \tag{14}$$

5 Measures of performance

The question then arises how we know if what we have done is worthwhile.

5.1 Testing $p(\not{E}_T)$ for correctness

We would like to check that the probability density $p(\not{E}_T)$ that we construct really does have the desired property: that the probability that the missing transverse energy is between $\not{E}_T$ and $\not{E}_T + \delta\not{E}_T$ is $p(\not{E}_T)\delta\not{E}_T$. This may be checked using W+jets Monte Carlo, where the W decays to leptons. Each event i has some known parton-level $\not{E}_{T_i}$ corresponding to the neutrino in the event. For each of these events we calculate the quantity

$$\xi_i = \int_0^{\not{E}_{T_i}} p_i(\not{E}_T)\, d\not{E}_T, \tag{15}$$

where $p_i(\not{E}_T)$ is the $\not{E}_T$ probability density calculated for the i^{th} event. A histogram of the ξ_i then provides a very nice check, since $p(\not{E}_T)$ is an accurate estimate of the probability density for the missing transverse energy if and only if ξ has a uniform distribution between 0 and 1.

5.2 Testing $p(\not{E}_T)$ for usefulness

The ultimate goal of this effort is to use the function $p(\not{E}_T)$ to distinguish between signal events with large true missing E_T and background events with zero true missing E_T. To obtain a quantitative measure of "usefulness" we

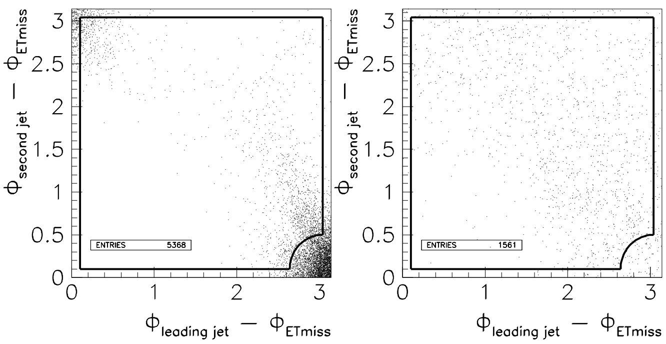

Figure 5: A standard topological cut to reject QCD background — events outside the solid lines are discarded. (Left) QCD events with ≥ 3 jets and $30 < \not\!\!E_{\mathrm{T}} < 40$ GeV, collected with the `jet_85` trigger. (Right) W+jets Monte Carlo events with ≥ 2 jets and $30 < \not\!\!E_{\mathrm{T}} < 40$ GeV. This topological cut clearly eliminates a greater fraction of QCD events than W+jets events, but can one do better?

look at two samples, one with true missing transverse energy, and one without. The first is a sample of $W+ \geq 2$ jets Monte Carlo events, and the second is a sample of QCD events containing ≥ 3 jets collected with the `jet_85` trigger, in which any missing transverse energy presumably arises from detector effects. We require the events in both samples to have $30 < \not\!\!E_{\mathrm{T}} < 40$.

Figure 5 shows the type of topological cut made by Adam Lyon in the search for squarks and gluinos in the $\not\!\!E_{\mathrm{T}}jjj$ channel [5] — all events outside the solid lines are removed by this cut. We find it convenient for purposes of comparison to introduce a variable s, defined such that an event lying on the solid line in Fig. 5 has $s = 1$ and an event lying on the boundary of the plot has $s = 0$; s is a measure of how far an event is from the boundary and the lower right-hand corner of the plot.

Figure 6 shows distributions of the topological variable s and the likelihood variable $\mathcal{L}$ for the W+jets and QCD samples. We find that the standard cut of $s > 1$ retains 83% of the W+jets Monte Carlo events, and we tune our cut on the likelihood in order to achieve the same efficiency on this sample. A cut of $\mathcal{L} > 3$ retains 82% of the W+jets events. Applying the same cuts to our QCD sample, we find that a cut of $s > 1$ keeps 29% of these unwanted events, while a cut of $\mathcal{L} > 3$ retains only 5.4% of these events. That is, for the same efficiency on events with true missing transverse energy, our likelihood cut appears to reject a factor of five more background events than does the topological cut that has been used in previous jets+$\not\!\!E_{\mathrm{T}}$ analyses.

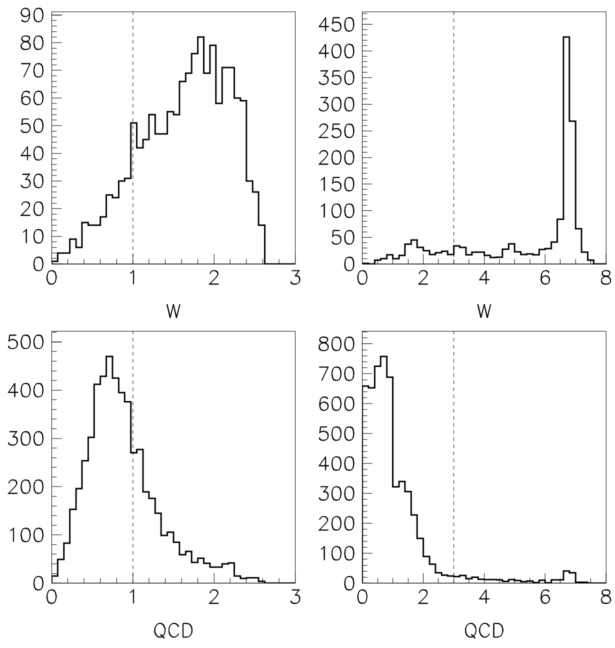

Figure 6: A comparison of the topological variable s and the likelihood $\mathcal{L}$ for W+jets Monte Carlo events and QCD data events. The W+jets Monte Carlo events are required to have ≥ 2 jets and $30 < \not{E}_T < 40$ GeV; the QCD events, collected from `jet_85`, are required to have ≥ 3 jets and $30 < \not{E}_T < 40$ GeV. (Top left) The topological variable s for the W+jets Monte Carlo events. A cut of $s > 1$ keeps 83% of this sample. (Top right) The likelihood variable $\mathcal{L}$ for these events. A cut of $\mathcal{L} > 3$ keeps 82% of this sample. (Bottom left) The topological variable s for the QCD data events. A cut of $s > 1$ keeps 29% of this sample. (Bottom right) The likelihood variable $\mathcal{L}$ for these events. A cut of $\mathcal{L} > 3$ retains only 5.4% of these unwanted events, less than one fifth of the number retained using the standard topological variable s for the same W+jets efficiency.

Figure 7 shows the distributions in s and $\mathcal{L}$ for a W+jets data sample, selected by requiring one good electron, ≥ 2 jets, $50 < M_T^{e\nu} < 90$ GeV, and $\not{E}_T > 30$ GeV. We see that the cuts $s > 1$ and $\mathcal{L} > 3$ perform comparably on real data, as well as on Monte Carlo.

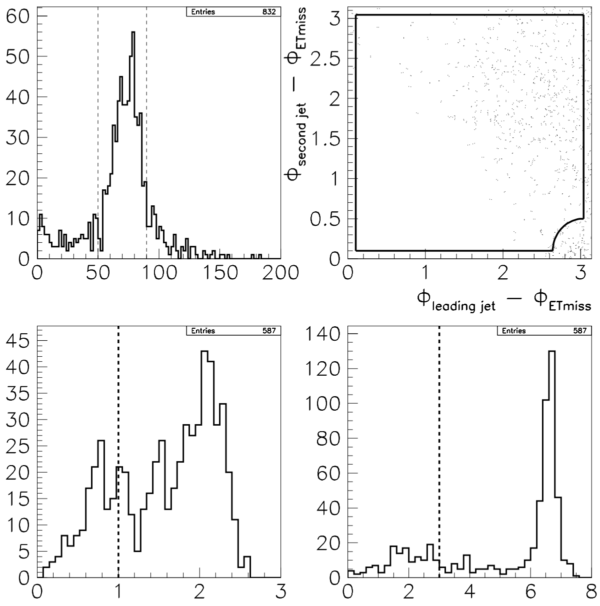

Figure 7: A comparison of the topological variable s and the likelihood $\mathcal{L}$ for W+jets data events. (Top left) The transverse mass distribution for the $W+\geq 2$ jets data sample; we keep events with $50 < M_T^{e\nu} < 90$ GeV. (Top right) Scatter plot of these events in the variable space considered by Adam Lyon and others. (Bottom left) The topological variable s for these events. A cut of $s > 1$ keeps 74% of the events. (Bottom right) The likelihood variable $\mathcal{L}$ for these events. A cut of $\mathcal{L} > 3$ keeps 76% of these events.

6 Application

Once we believe that we understand our method, which will be quite a lot of work in itself, we will be ready to apply it to previous New Phenomena analyses. I will soon begin to compile ntuples and code from recent DZero searches. We will need to ensure that we understand what was done originally, and will need to make sure that we are able to reproduce the previous results. Once this is done, we go through the analyses one at a time. We begin by loosening $\not\!\!E_T$-related cuts, and then use our improved knowledge of the resolution on missing transverse energy to make improved cuts.[2]

The final result of this effort will be better limits in a number of channels that have previously been searched. Since the tracking system in DZero will be very different in Run II, it is unlikely that our algorithm for determining $p(z)$ will be useful in Run II. On the other hand, our improved characterization of hot cells may be very useful in Run II. The general understanding of what is needed in order to make a good measurement of the missing transverse energy, which we will hopefully have gained from this effort, will certainly be of use in pulling physics out of Run II.

7 Conclusion

These are just some ideas that need to be made more concrete. We can of course apply this procedure by calculating $p(z)$ as above and assuming that we know the energies E_i precisely, and vice versa. That is, we can consider this proposed program as consisting of modular units, each of which accounts for a component of the final resolution of $\not\!\!E_{\mathrm{T}}$. Since there will be several people working on this analysis, it may be useful to treat these as components. They may also be discarded as components if a few of these ideas end up not being any good.

References

[1] I. Adam, V. Balamurali, S. Chopra, U. Heintz, and M. Narain. The Micro DST Format. *DZero note 2327*, 1995.

[2] L. Babukhadia et al. Hot Cells and Jets at DØ: AIDA Cell Restoration. *DZero note 3408*, 1998.

[3] DØ Collaboration. Studies of gauge boson pair production and trilinear couplings. *Phys. Rev. D*, 1997.

[4] G. DiLoreto and F. Nang. The Triple Differential Di-Jet Cross Section at $\sqrt{s}$=1.8 TeV. *http://msupc3.fnal.gov/gian/results/triple/triple_note_8-14-98.ps*, 1998.

[5] A.L. Lyon, *A Search for Squarks and Gluinos Using the Jets and Missing Energy Signature at DØ*, Ph.D. Thesis, University of Maryland, 1997.

[2] Improved in the sense of maximizing sensitivity to the signal. DZero note 3396 contains a useful definition of "sensitivity".