Mass Determination in Events with Missing Energy

Bob McElrath

University of California, Davis

Work(s) in progress with Zhen-Yu Han, Hsin-Chia Cheng, Jack Gunion, and Guido Marandella

Fermilab Dark Matter Workshop, May 10, 2007

Motivation

If our current particle picture of Dark Matter is correct, the LHC is likely to be a Dark Matter factory. Realistic models containing a Dark Matter particle tend to be very similar.

- A symmetry is added to keep Dark Matter stable $\rightarrow$ Dark Matter is produced in pairs.

- Symmetries which keep Dark Matter stable are often taken from other sources (because we prefer as simple a model as possible), such as:

  – Proton Stability (R-Parity in SUSY)

  – Custodial Symmetry (solving Little Hierarchy Problem)

  – 5D momentum conservation (KK number conservation in UED)

  "Other Sources" for the symmetry generically means "Other Particles".

Other Particles

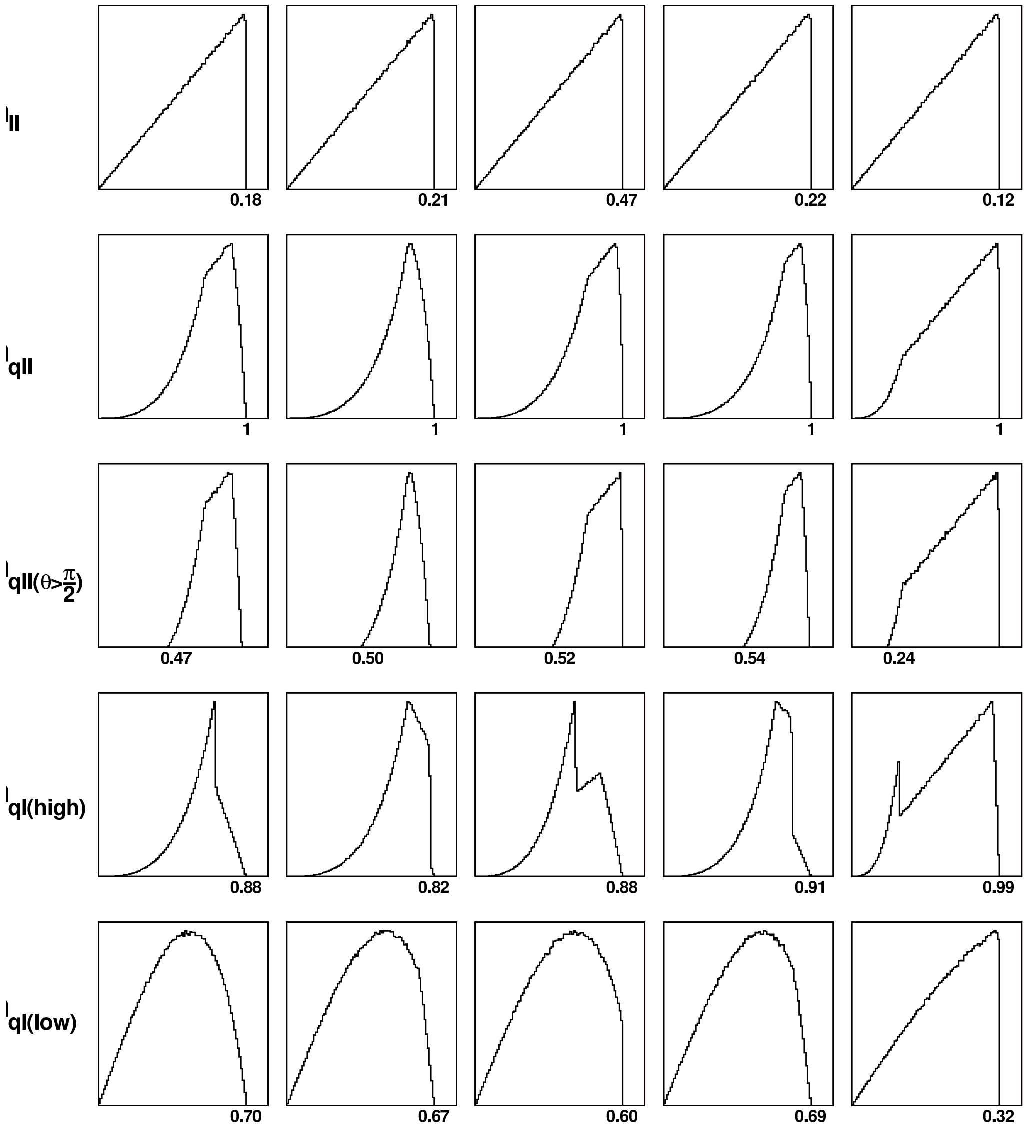
Other Particles means that the Dark Matter is generically produced from the decay of a heavier particle which has SM quantum numbers. (such as a squark, slepton, T-Parity-odd fermion, etc)

All masses require one to add the missing particle to a visible particle. Therefore expected signatures of new physics contain *no way to directly obtain the mass of a particle, from the 4-vectors in the event*.

e.g. In a visible decay $Z \to l^+l^-$, it is trivial to get a consistent estimator for the Z mass: $m^2 = (p_{l^+} + p_{l^-})^2$. This estimator m^2 has the property that its mean, $\langle m^2 \rangle$ converges to the Lagrangian parameter M_Z^2. This provides the strongest available indicator (to cut on) indicating that this is a Z.

For events with missing particle, no such thing is possible. (yet!)

Existing Studies (Barr Topology)



Existing studies all on fitting distribution in a variable correlated to masses. If we must rely on such things, this is troublesome

- Mass determinations are very sensitive to a small number of events (those occuring at inflection points and endpoints).
- Detector resolution makes all (Barr-type) distributions look similar.

[Gjelsten, Miller, Osland `hep-ph/0410303`]

Existing Studies: Cross Sections as Probability Densities

What are these studies doing, from a theoretical/statistics perspective? First let us define a probability distribution for an event. A cross section generally is given by

$$\sigma = \frac{1}{F} \int \left| \mathcal{M}(p_0^\mu, p_i^\mu, \lambda) \right|^2 \left(\prod_i \frac{d^3 \vec{p}_i}{(2\pi)^3 2E_i} \right) (2\pi)^4 \delta^4(p_0^\mu - \sum_i p_i^\mu)$$

for some initial state momenta p_0^μ and final state momenta p_i^μ. This is a zero-dimensional projection of a high-dimensional phase space, and contains very little information! Buried in here somewhere is all the information that is to be had. Let us do a little rearrangement to retain all information in the high-dimensional space.

$$P(\vec{p}_1, \ldots, \vec{p}_N) = \frac{1}{\sigma} \frac{d\sigma}{\prod_i d^3 \vec{p}} = \frac{(2\pi)^{4-3N}}{2^N F \sigma \prod_i E_i} \left| \mathcal{M}(p_0^\mu, p_i^\mu, \lambda) \right|^2 \delta^4 \left(p_0^\mu - \sum_i p_i^\mu \right).$$

this is a *probability density* expressing the probability of a particular configuration of momenta. For N external particles, it is a $3N - 4$ dimensional space.

Cross Sections as Probability Densities II

$$P(p_i^\mu|\lambda) = \frac{1}{\sigma}\frac{d\sigma}{\prod_i d^3\vec{p}} = \frac{(2\pi)^{4-3N}}{2^N F\sigma\prod_i E_i}\left|\mathcal{M}(p_0^\mu, p_i^\mu, \lambda)\right|^2 \delta^4\left(p_0^\mu - \textstyle\sum_i p_i^\mu\right). \quad (1)$$

In principle, one could directly compare this PDF (*Probability* Density Function) between simulated events and data. But, high-dimensional spaces require a lot of data to map out.

- Project onto lower dimensional space (e.g. Breit-Wigner, end-point/edge techniques)

- Use a Likelihood or "Matrix Element" method

The Neyman-Pearson lemma tells us that the most powerful statistic for differentiating two hypotheses λ and λ' is the ratio of two Likelihoods. Our Likelihood is

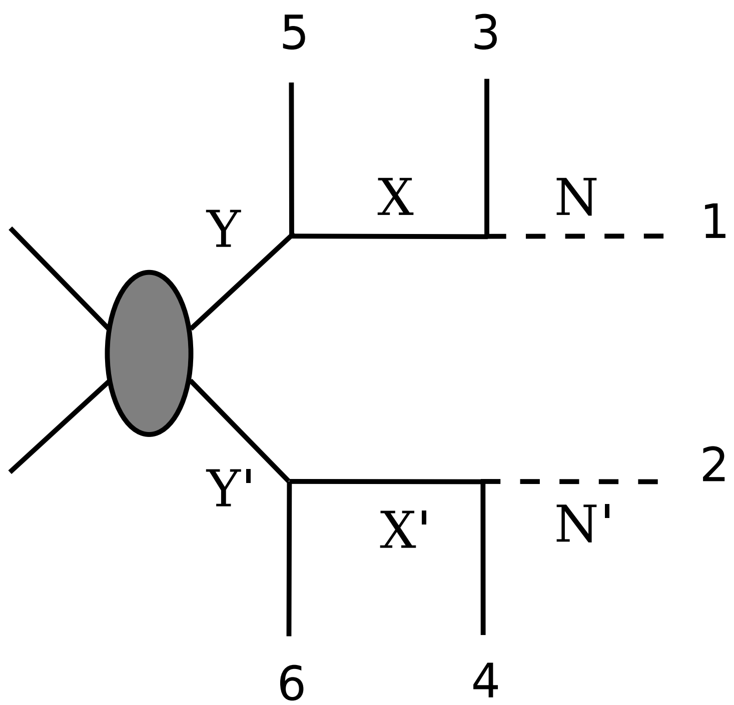
$$L(\lambda|\{p_i^\mu\}) = \prod_{i=1}^{N} P(p_i^\mu|\lambda).$$

Topology

Matrix Element Methods are powerful, but also complicated. And, we don't know the Lagrangian!

Let's try something simpler. What can we do with just phase space? (assuming the Matrix Element is a constant)

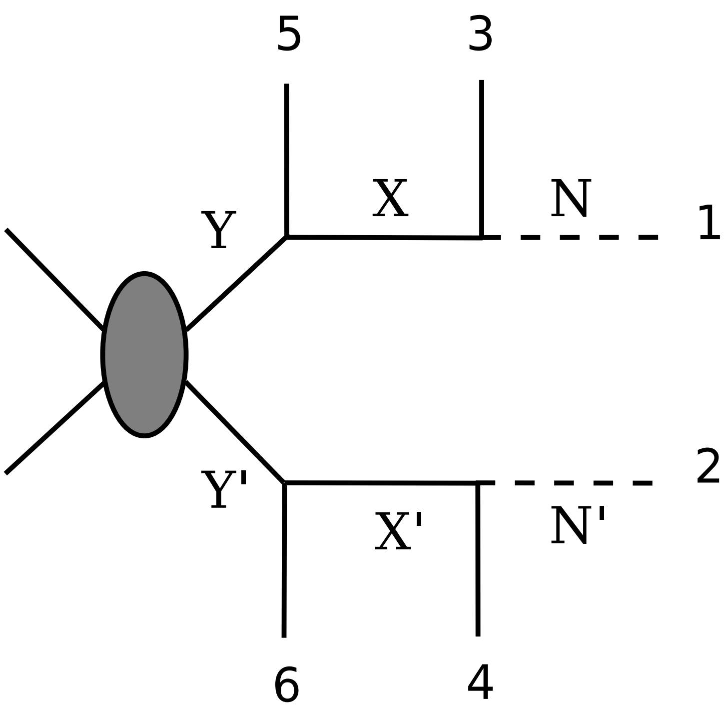
To examine this, let us choose the $t\bar{t}$ di-lepton topology, which is identical to many interesting SUSY decay topologies.



We have generalized the methods I will describe to any process with exactly 2 missing particles, and 2 or more visible particles.

The vector p_0 is the initial state. This diagram is kinematic. e.g. this also works for t-channel production.

Topology ctd. . .



This topology can be applied to many processes with 4 visible and 2 invisible particles.

For simplicity in analysis we will further assume $M_Y = M_{Y'}$, $M_X = M'_X$, and $M_N = M'_N$.

Examples that fit this:

$$\begin{aligned}
t\bar{t} &\rightarrow bW^+bW^- \rightarrow bl^+\nu bl^-\bar{\nu} \\
\tilde{\chi}_2^0\tilde{\chi}_2^0 &\rightarrow l\tilde{l}l\tilde{l} \rightarrow ll\tilde{\chi}_1^0 ll\tilde{\chi}_1^0 \\
\tilde{q}\tilde{q} &\rightarrow q\tilde{\chi}_2^0 q\tilde{\chi}_2^0 \rightarrow ql\tilde{l}ql\tilde{l} \rightarrow qll\tilde{\chi}_1^0 qll\tilde{\chi}_1^0 \\
\tilde{t}\bar{\tilde{t}} &\rightarrow b\tilde{\chi}^+\bar{b}\tilde{\chi}^- \rightarrow bW^+\tilde{\chi}_1^0\bar{b}W^-\tilde{\chi}_1^0
\end{aligned}$$

Changing Variables

If we want to talk about masses, the first thing we had better do is change variables.

The $t\bar{t}$ di-lepton topology at the LHC contains 4 kinematic unknowns, which is nice because it also has 4 unknown masses.

$$\begin{aligned}
a &= (p_2 + p_4 + p_6)^2 \\
b &= (p_2 + p_4)^2 \\
c &= (p_1 + p_3 + p_5)^2 \\
d &= (p_1 + p_3)^2 \\
0 &= \not{p}_x - p_{1x} - p_{2x} \\
0 &= \not{p}_y - p_{1y} - p_{2y} \\
0 &= \sqrt{s}\sigma - p_{vz} - p_{1z} - p_{2z} \\
0 &= \sqrt{s}\tau - E_v - E_1 - E_2 \\
M_1^2 &= E_1^2 - \vec{p}_1^{\,2} \\
M_2^2 &= E_2^2 - \vec{p}_2^{\,2}
\end{aligned}$$

This variable change is non-linear, and incurs a Jacobian J (important if you want to integrate your Probability Density in the mass basis!)

Changing Variables, ctd...

Writing the same thing in integral form, first write the PDF with the Dark Matter's mass constraint explicitly.

$$\begin{aligned} P(\{p_i^\mu\}|\lambda) &= f(\{p_i^\mu\},\lambda) \int |\mathcal{M}(\lambda, p_0^\mu, \ldots, p_N^\mu)|^2 \\ &\quad \times \delta^4(p_0^\mu - \textstyle\sum_i p_i^\mu) \\ &\quad \times 2E_1\, \delta(E_1^2 - m_1^2 - |\vec{p}_1|^2)\, 2E_2\, \delta(E_2^2 - m_2^2 - |\vec{p}_2|^2) \\ &\quad \times d\tau\, d\sigma\, d^3\vec{p}_1\, d^3\vec{p}_2\, dE_1\, dE_2 \end{aligned}$$

Next expand the dimensionality by 4 and add 4 delta functions, corresponding to the 4 propegators.

$$\begin{aligned} P(\{p_i^\mu\}|\lambda) &= f(\{p_i^\mu\},\lambda) \int |\mathcal{M}(\lambda, p_0^\mu, \ldots, p_N^\mu)|^2 \\ &\quad \times \delta^4(p_0^\mu - \textstyle\sum_i p_i^\mu) \\ &\quad \times 2E_1\, \delta(E_1^2 - m_1^2 - |\vec{p}_1|^2)\, 2E_2\, \delta(E_2^2 - m_2^2 - |\vec{p}_2|^2) \\ &\quad \times \delta((p_2 + p_4 + p_6)^2 - a)\, \delta((p_2 + p_4)^2 - b) \\ &\quad \times \delta((p_1 + p_3 + p_5)^2 - c)\, \delta((p_1 + p_3)^2 - d) \\ &\quad \times d\tau\, d\sigma\, d^3\vec{p}_1\, d^3\vec{p}_2\, dE_1\, dE_2\, da\, db\, dc\, dd \end{aligned}$$

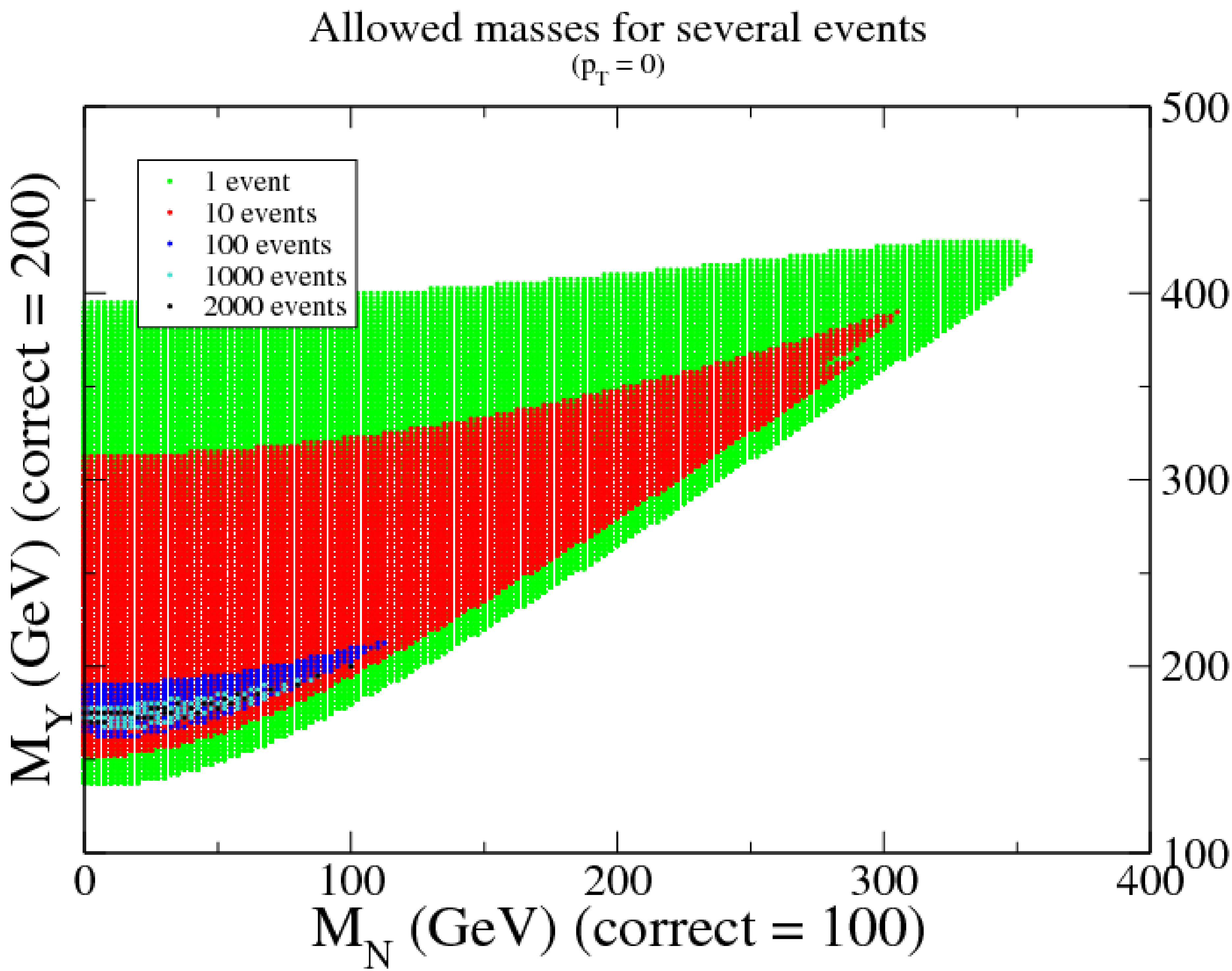
Simplified Likelihood

Let us try to characterize what's going on in mass space, by constructing the likelihood $L(\lambda|\{p_i^\mu\}) = \prod_i P_i(\{p_i^\mu\}|\lambda)$ in the narrow width approximation. Note that our $\delta(a-(p_1+p_3)^2)$ (etc) is exactly what would arise from a Matrix Element containing narrow widths.

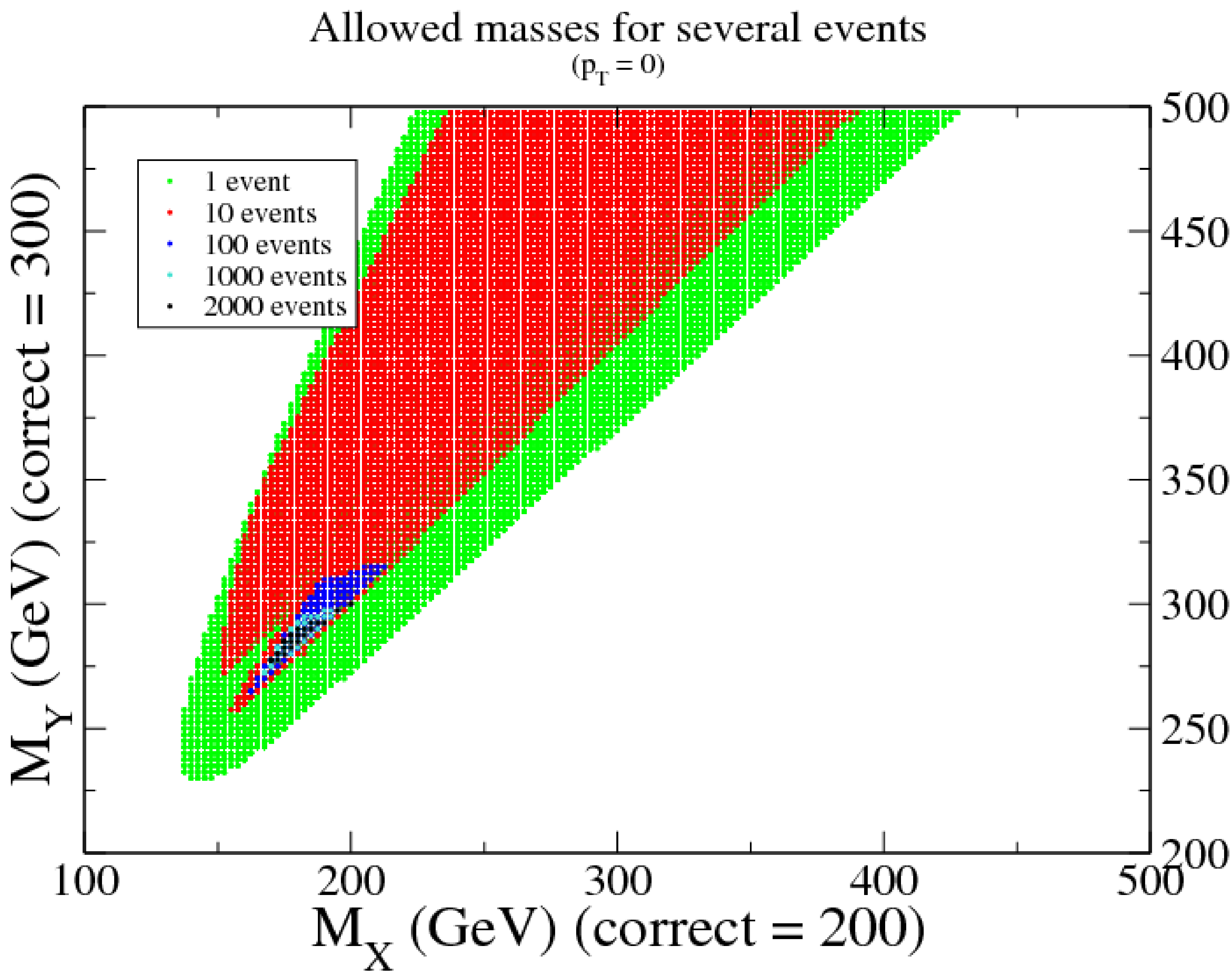
This is identical to taking the integrand to be 1 after our variable change. The $P(\{p_i^\mu\}|\lambda)$ is zero in regions where the variable change cannot be performed (would result in complex E, p).

It's also equivalent to answering the question: Given two events, *what is the region in mass space that is compatible with both events?*

We extend to N events and ask what region is consistent with *all* events. (no resolution, detector simulation, or combinatorics here)

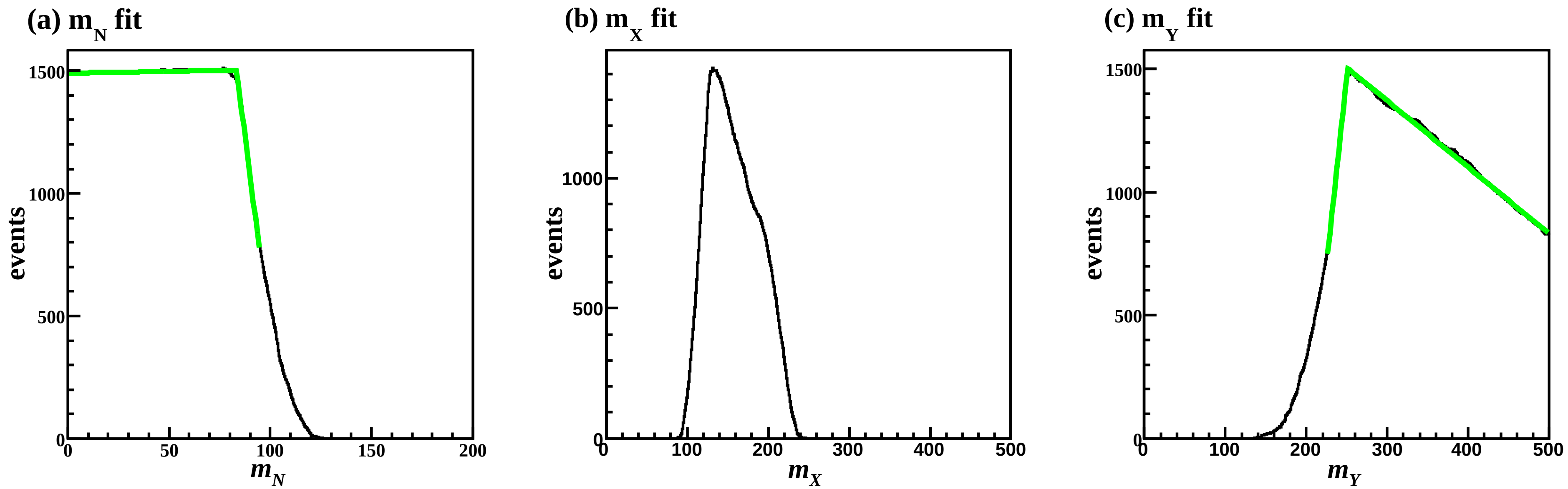


Allowed masses for several events
(p_T = 0)
1 event
10 events
100 events
1000 events
2000 events
M_Y (GeV) (correct = 200)
M_N (GeV) (correct = 100)
0
100
200
300
400
100
200
300
400
500



Allowed masses for several events
(p_T = 0)
1 event
10 events
100 events
1000 events
2000 events
500
450
400
350
300
250
200
100
200
300
400
500
M_Y (GeV) (correct = 300)
M_X (GeV) (correct = 200)

Graphical Algorithm

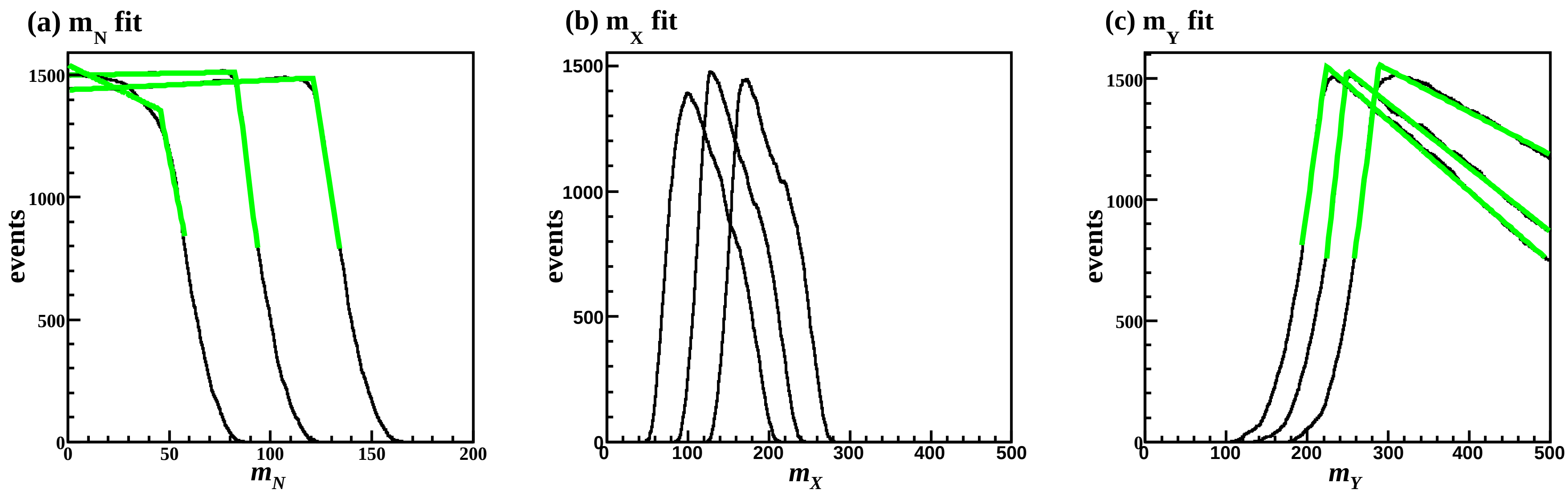


Fixing two of the masses, we scan in the third mass. Unfortunately an analytic expression for these curves is probably intractable to derive.

For a large number of events, we want the *largest* M_N compatible with the event. Large p_T cuts off the zero mass solution, but the high mass solution converges to the correct value faster, and our understanding of p_T in hadron colliders is poor. (e.g. M_T used to measure M_W is designed to be p_T insensitive)

But! Features are simple. We fit a line to the "corner" to determine its location.

Iterate

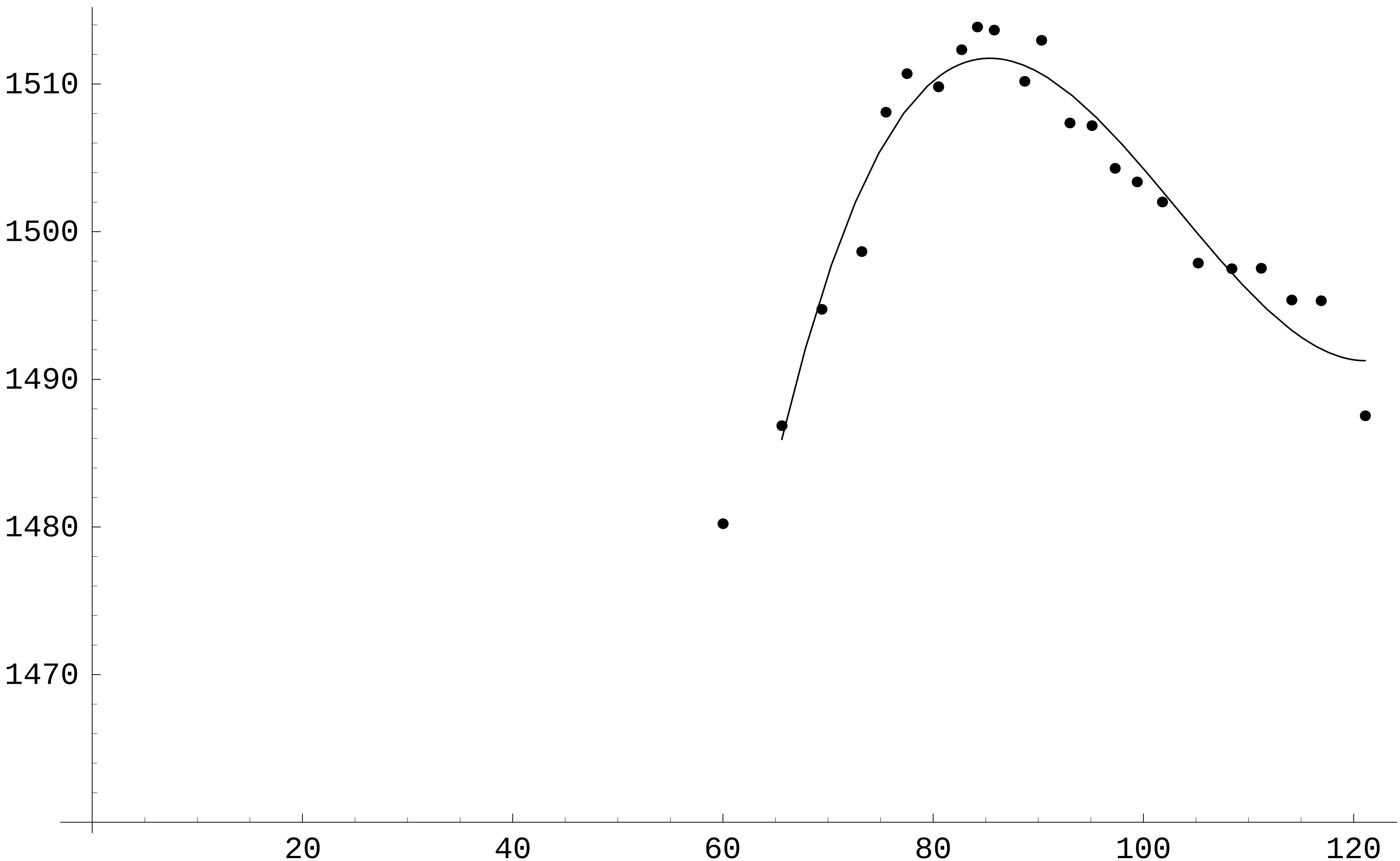


Iterate in each mass, fitting for each mass successively.

This procedure "walks up" the mass space, increasing the over mass scale, and is not convergent. (e.g. there still exists a solution at $M_N = \infty$ for most events)

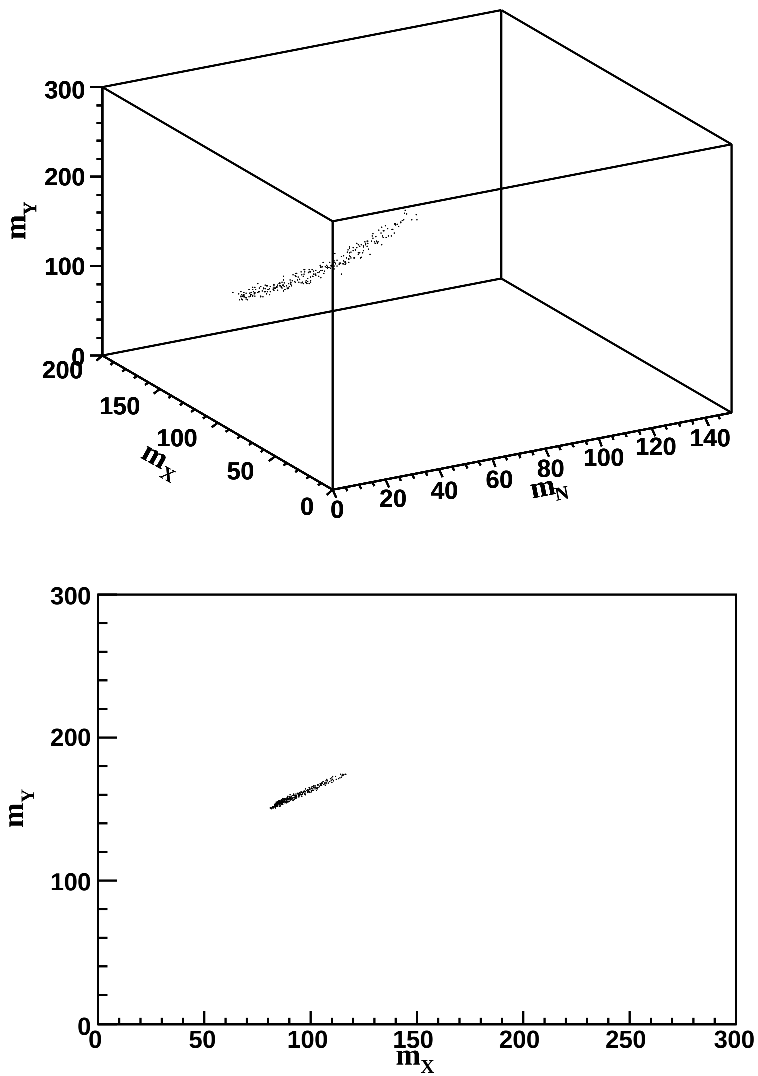
But! We have not yet used the total number of events fit.

M_N vs. Number of Events



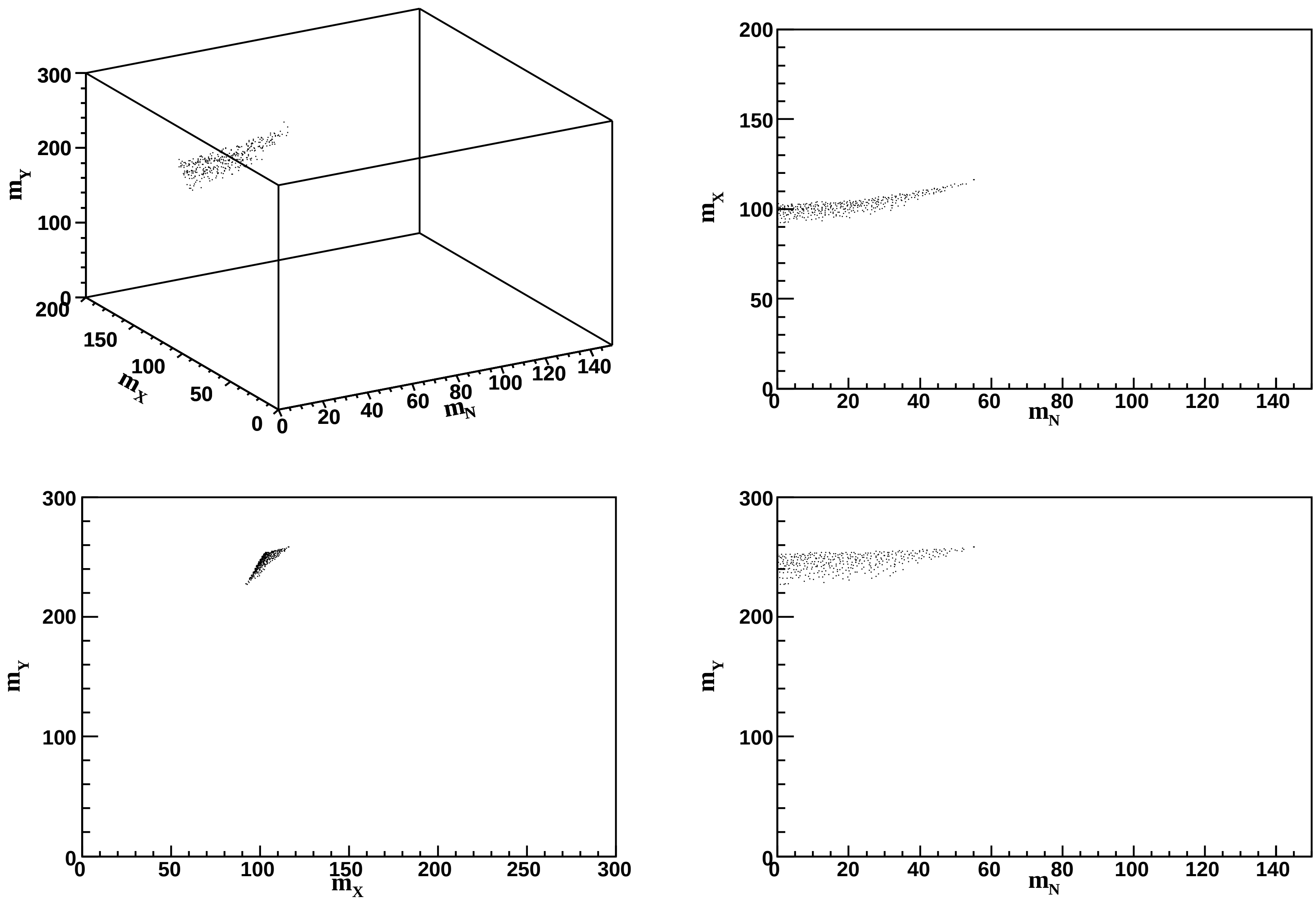
3D projections with no smearing

$$(m_Y, m_X, m_N) = (180.8, 147.1, 85.2)$$



3D projections with smearing

$$(m_Y, m_X, m_N) = (246.6, 128.4, 85.3)$$



Fit Results

In the ideal case (no resolution or combinatorics), given inputs:

$$(m_Y, m_X, m_N) = (246.6, 128.4, 85.3)$$

we reconstruct

$$(m_Y, m_X, m_N) = (251.5, 130.7, 85.4).$$

For different sets of 1000 events, the fitted masses vary by 2-3 GeV.

We smear momenta using ATLFAST's muon resolution, and a missing momentum resolution given by a gaussian with width 18 GeV.

When resolution is included, the final fit point is taken as where the *slope* in M_N vs. # events changes. (not peak)

Some Comments

Anything you do generally results in a *biased*, *inconsistent* estimator. But, due to detector resolution, all or estimators are generally biased and inconsistent anyway, and we know how to deal with this. (e.g. the "template" methods used in M_W extraction at the Tevatron) We have determined the mass within 1 GeV, but this must be corrected by the systematic bias of the method itself.

There are *many* possible estimators. I have worked on at least 4. We have shown that M_N *can* be determined. The full probability density (e.g. Matrix Element Method) *does have sensitivity to the overall mass scale*.

This should be used as a crude tool to determine the mass scale and kinematic structure, with relatively low statistics.

A secondary analysis should enumerate different spin hypothesis for the various particles, and use a full Matrix Element (rather than narrow width).

Why does this work?

What is cutting off the high-mass solutions and what is cutting off the zero mass solutions?

High Mass: Some events have "holes". e.g. a region in the middle of the mass space that is *disallowed*. When intersected with a non-hole event, a very small volume remains. The allowed volume is given by an 8th order polynomial in mass. An 8th order polynomial can be genus-1. (e.g. torus-like) Two events can also go to infinite mass in different directions.

Low Mass: Events with sizeable p_T can cut off the 0 mass solution. But, reasonable p_T spectra still have the high-M_N solution closer to the true value than the low M_N cutoff. If we maximize $\frac{dP}{dM_N}$ in a *one missing particle* process, this gives an estimator for M_N that is simply $|p_T|$.

Solving the Equations: High Mass Cutoff

These ellipses can also be written using the vector $x_i = (E_1, E_2, 1)$ as

$$A = x_i f^{ij} x_j = 0, \qquad B = x_i g^{ij} x_j$$

using the tensors

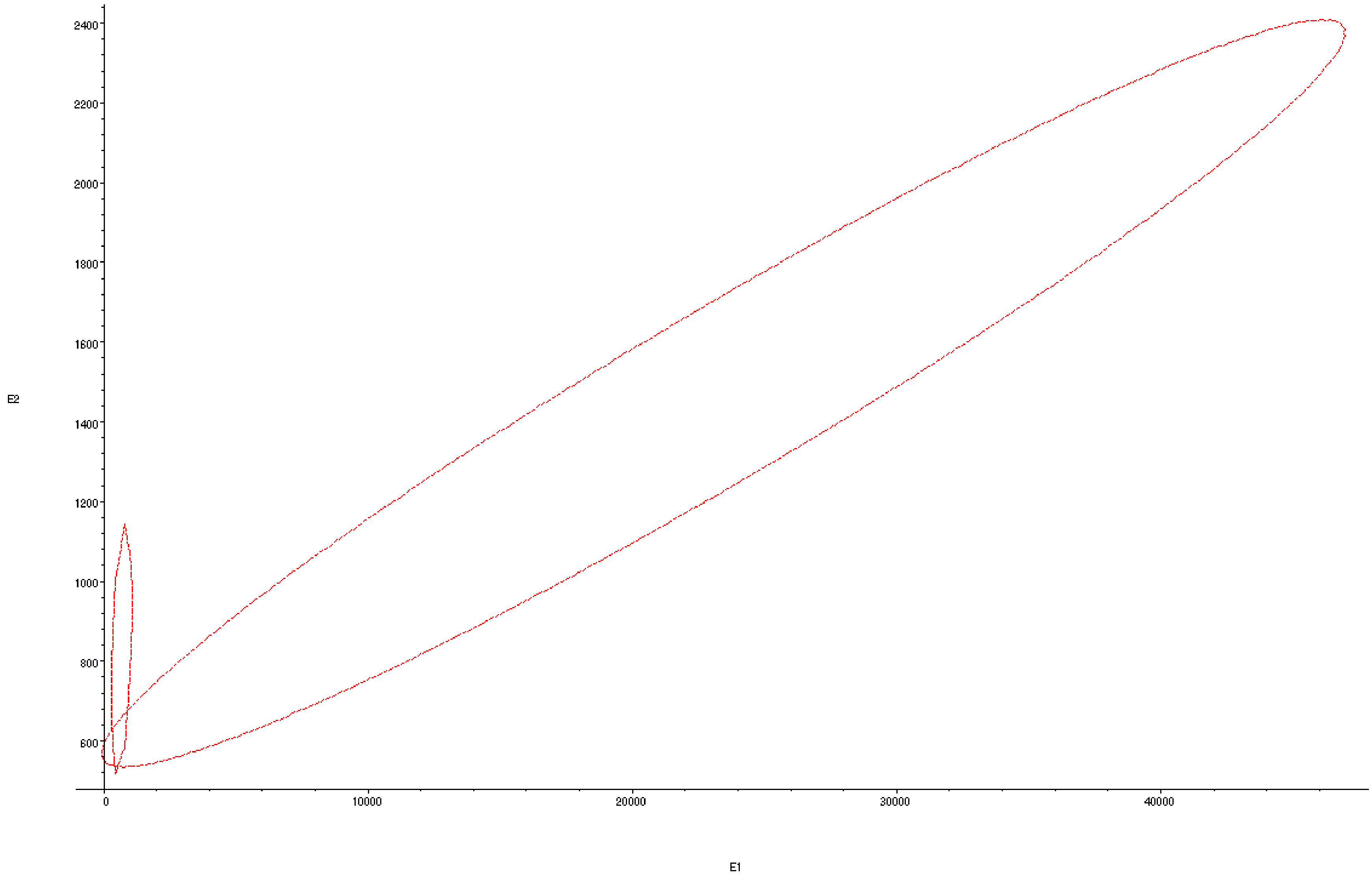
$$f^{ij} = \begin{bmatrix} a_{11} & a_{12} & a_1(m) \\ a_{12} & a_{22} & a_2(m) \\ a_1(m) & a_2(m) & a_0(m, m^2) \end{bmatrix} \qquad g^{ij} = \begin{bmatrix} b_{11} & b_{12} & b_1(m) \\ b_{12} & b_{22} & b_2(m) \\ b_1(m) & b_2(m) & b_0(m, m^2) \end{bmatrix}$$

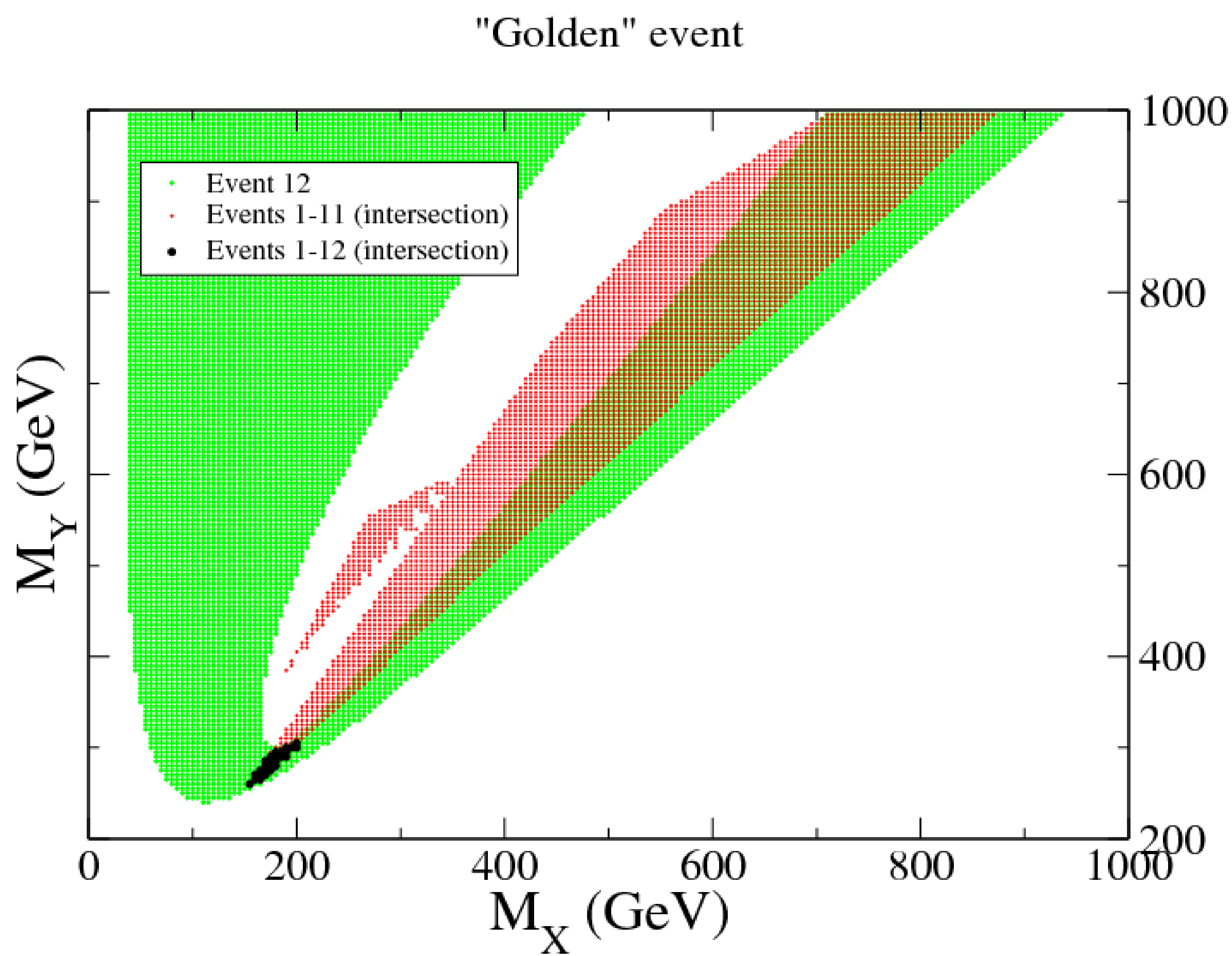
For determining masses, we only care about *whether* a solution exists, not the actual values of E_1, E_2. Therefore, without loss of generality, we are free to rotate, translate, and scale the vector x_i.

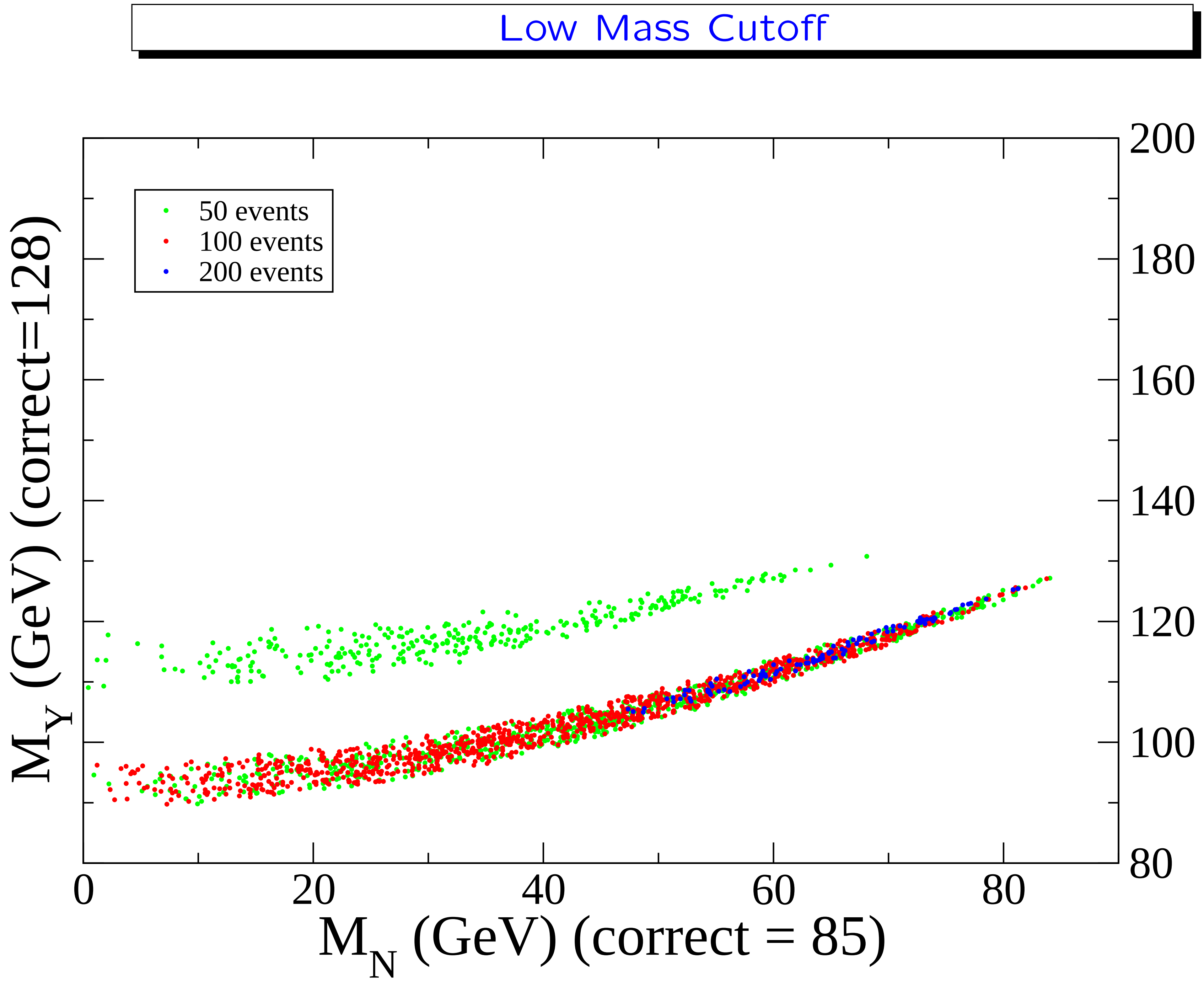
$$f^{ij} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -r^2(m, m^2) \end{bmatrix}, \quad \begin{array}{c} x^2 + y^2 - r^2(m, m^2) = 0 \\ \frac{(x - x_0(m))^2}{a^2} + \frac{(y - y_0(m))^2}{b^2} - R^2(m, m^2) = 0 \end{array}$$

$$g^{ij} = \begin{bmatrix} 1/a^2 & 0 & -x_0(m)/a^2 \\ 0 & 1/b^2 & -y_0(m)/b^2 \\ -x_0(m)/a^2 & -y_0(m)/b^2 & -R^2(m, m^2) + x_0(m)^2/a^2 + y_0(m)^2/b^2 \end{bmatrix}$$

J. Random Event




"Golden" event
Event 12
Events 1-11 (intersection)
Events 1-12 (intersection)
M_Y (GeV)
M_X (GeV)
0
200
400
600
800
1000


Low Mass Cutoff
50 events
100 events
200 events
M_Y (GeV) (correct=128)
M_N (GeV) (correct = 85)
0
20
40
60
80
100
120
140
160
180
200

Conclusions

We have shown that the overall mass scale (and therefore, the magnitude of all intermediate masses) can be determined at a hadron collider, for a $t\bar{t}$-dilepton-like topology at the $\mathcal{O}(1\text{GeV})$ level with $\mathcal{O}(1000)$ events.

This can be extended to any process with 2 or more visible particles and 2 missing particles.

There are *many* possible techniques that follow from intersections of subsets, maximizing the probability density for a single event, the "centroid" of the allowed region, or using different fit functions in the technique just presented. Which is optimal?

A matrix element method *does* have sensitivity to the overall mass scale.