Polynomial systems for the determination of masses in events with missing energy

Bob McElrath
CERN

Les Houches 2009, June 22, 2009

What is the *BEST* thing to do?

Most studies approach the problem as: Given that I have these 4-vectors $\{p_i^\mu\}$ that came out of my Monte Carlo or Experiment, what function $f(\{p_i^\mu\}|\lambda)$ can I write down which will tell me some hypothesis (parameter) λ?

The answer is that *every* function $f(\{p_i^\mu\}|\lambda)$ depends on the parameters λ, and I'm left with the question: Which $f(\{p_i^\mu\}|\lambda)$ is "best" for my purpose?

I approach this from the other side: the most powerful statistic for differentiating two hypotheses λ and λ' is the ratio of two Likelihoods (Neyman-Pearson Lemma). Our Likelihood for $n = 1..N$ events is

$$L(\lambda|\{\{p_j^\mu\}_n\}) = \prod_{n=1}^{N} P_n(\{p_j^\mu\}_n|\lambda).$$

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

Now let me make systematic approximations to this ideal situation.

Polynomial Systems

In a hadron collider with missing energy, the PDF is defined as

$$P(p_i^\mu|\lambda) = \int \frac{(2\pi)^{4-3N}}{2^N F\sigma \prod_i E_i} \left|\mathcal{M}(p_0^\mu, p_i^\mu|\lambda)\right|^2 \delta^4\left(p_0^\mu - \textstyle\sum_i p_i^\mu\right) dx_1 dx_2 d^3p_1 d^3p_2$$

Now let us go into the narrow width approximation by replacing

$$\frac{1}{(q^2-M^2)^2 - M^2\Gamma^2/4} \to \frac{\pi}{M\Gamma}\delta(q^2-M^2)$$

in $\left|\mathcal{M}(p_0^\mu, p_i^\mu|\lambda)\right|^2$, for some hypothesis diagram (valid for $\Gamma \ll M$).

Alternatively, one can simply insert the appropriate delta functions corresponding to a diagram, and view this as a variable change.

Note that this integral is 4 dimensional at a hadron collider. Therefore, *by specifying 4 masses, the integral is reduced to a discrete set of solutions for the missing momenta*.

A pair of simultaneous quadratics is not guaranteed to have a solution!

The General Recipe for using Polynomial Systems

- Write down a hypothesis diagram describing the visible final state particles and missing energy you see.
- Combine resonances with entirely visible decay products and call it a single final state particle.
- Count the missing particles N and the intermediate, *on-shell* particles M with missing particles "down-stream".
- $M < 3N - 2$: ("underconstrained") use kinks or edges.
- $M = 3N - 2$: ("exactly constrained") one can change variables from the missing momenta into these masses. Each event defines a volume in mass space. See JHEP 0712:076,2007 and arXiv:0811.2138
- $M > 3N - 2$: ("overconstrained") it is possible to solve for discrete values of the masses, by constructing a larger polynomial system from multiple events, *under the assumption that they contain the same physics*. See: Phys.Rev.Lett.100:252001,2008.

The General Recipe for using Polynomial Systems

Once your polynomial system is constructed, one can ask the question if $M = 3N - 2$ (exactly constrained):

Is the Probability Density P zero or nonzero for a given set of hypothesis masses?

The nonzero answer defines a volume in mass space, which one must then devise an algorithm to extract the true mass from by combining events.

If $M > 3N - 2$ (overconstrained):

Is the n-event likelihood $L_n = \prod_{i=1}^{N} P_i$ zero or nonzero?

If these are nonzero, in the narrow width approximation, you have just solved for a set of 4-momenta consistent with the event (and therefore, all the intermediate masses too).

The General Recipe for using Polynomial Systems

These are systematic approximation to the "best" Likelihood method, accurate to $\mathcal{O}\left(\frac{\Gamma}{M}\right)$ and ignoring spin. The only thing better is to use a true Matrix Element Method, which also includes off-shell effects.

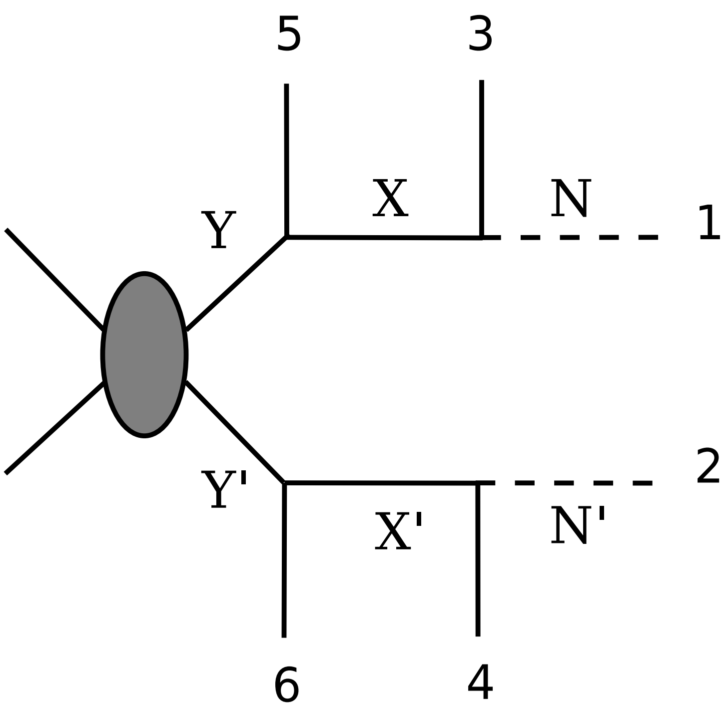
The overconstrained case is the *best* option for small data. In principle it works for as few as two events in SPS1a.

These methods need long chains: at least 5 on-shell intermediate particles is overconstrained, 4 is exactly constrained.

These methods are probably not useful with 3 or more missing particles: this needs 7 on-shell intermediate resonances.

The intermediate particles must be *on-shell*.

Exactly constrained example

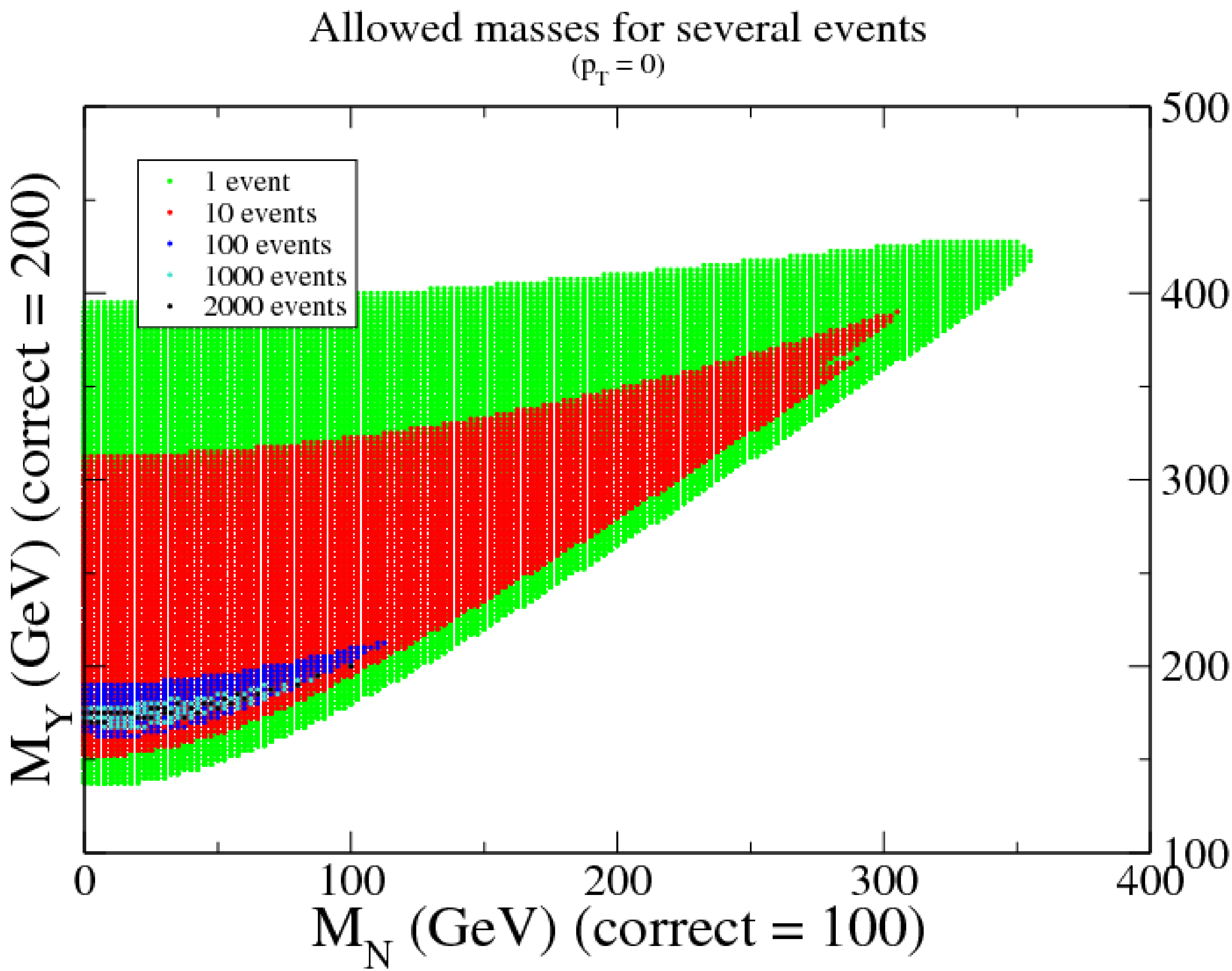


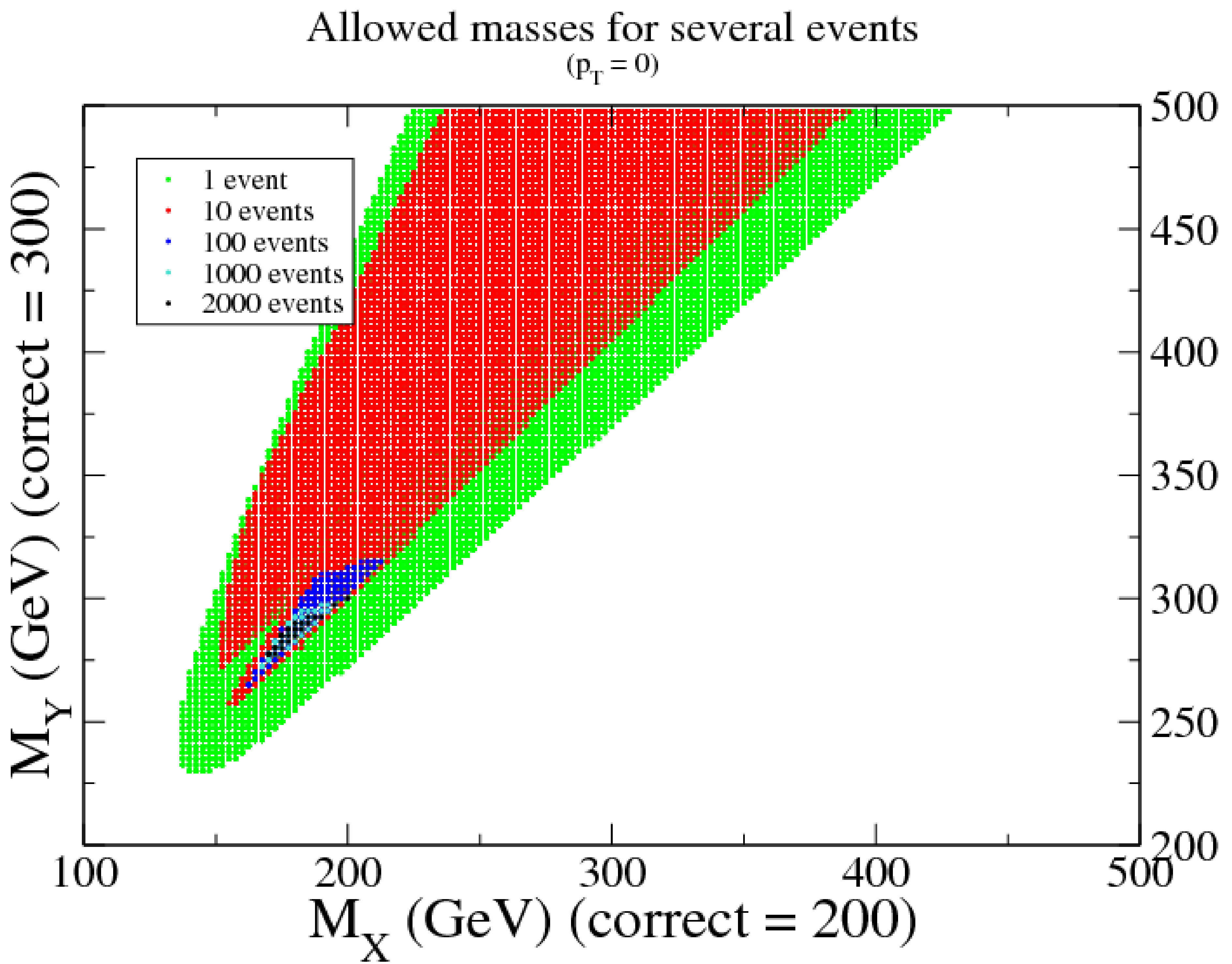
This topology can be applied to many processes with 4 visible and 2 invisible particles.

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:

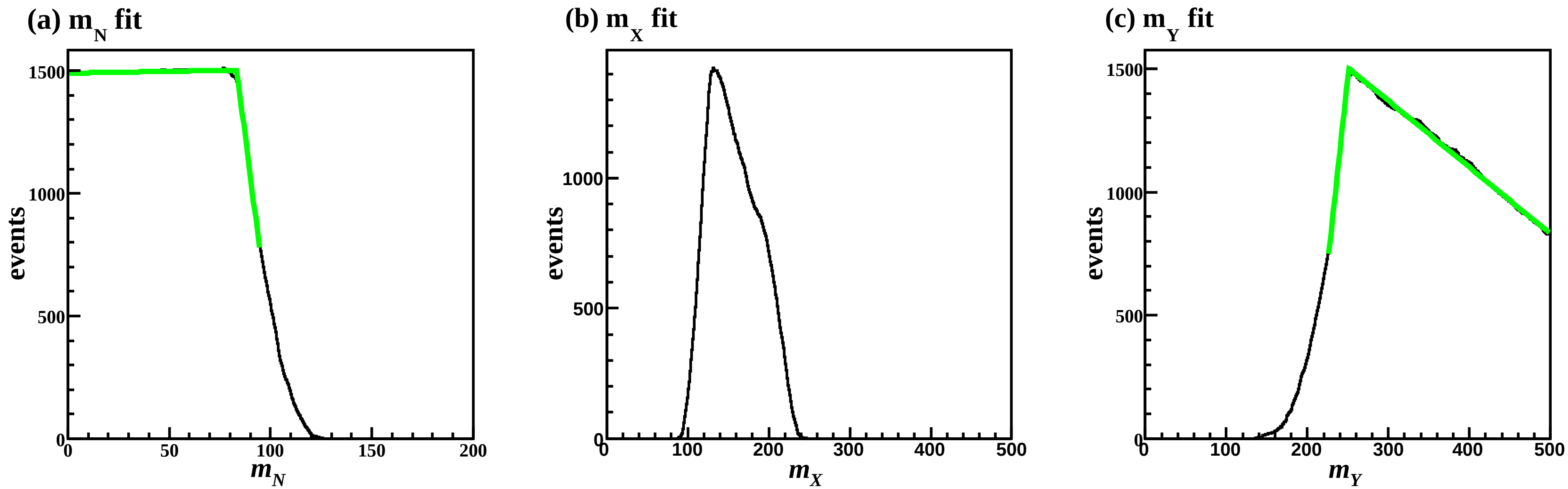
$$
\begin{aligned}
t\bar{t} &\rightarrow bW^+bW^- \rightarrow bl^+\nu bl^-\bar{\nu} \\
\tilde{\chi}_2^0\tilde{\chi}_2^0 &\rightarrow l\tilde{l}l\tilde{l} \rightarrow ll\tilde{\chi}_1^0 ll\tilde{\chi}_1^0 \\
\tilde{q}\tilde{q} &\rightarrow q\tilde{\chi}_2^0 q\tilde{\chi}_2^0 \rightarrow ql\tilde{l}ql\tilde{l}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


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

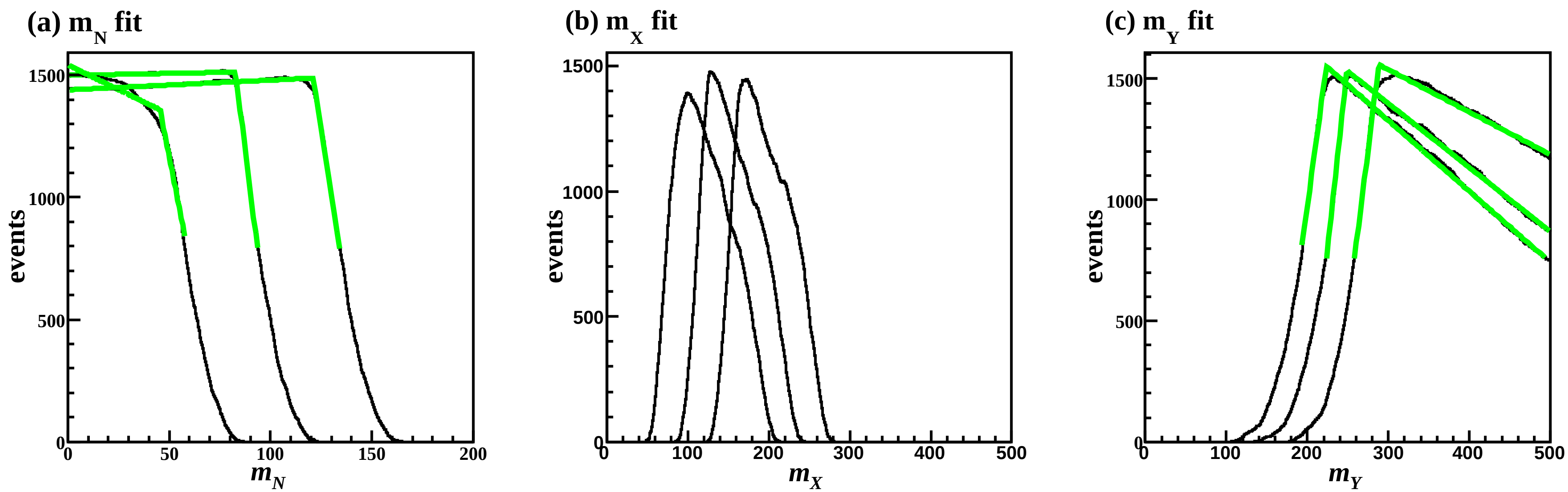


Fixing two of the masses, we scan in the third mass. Unfortunately an analytic expression for these curves is probably intractable to derive.

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

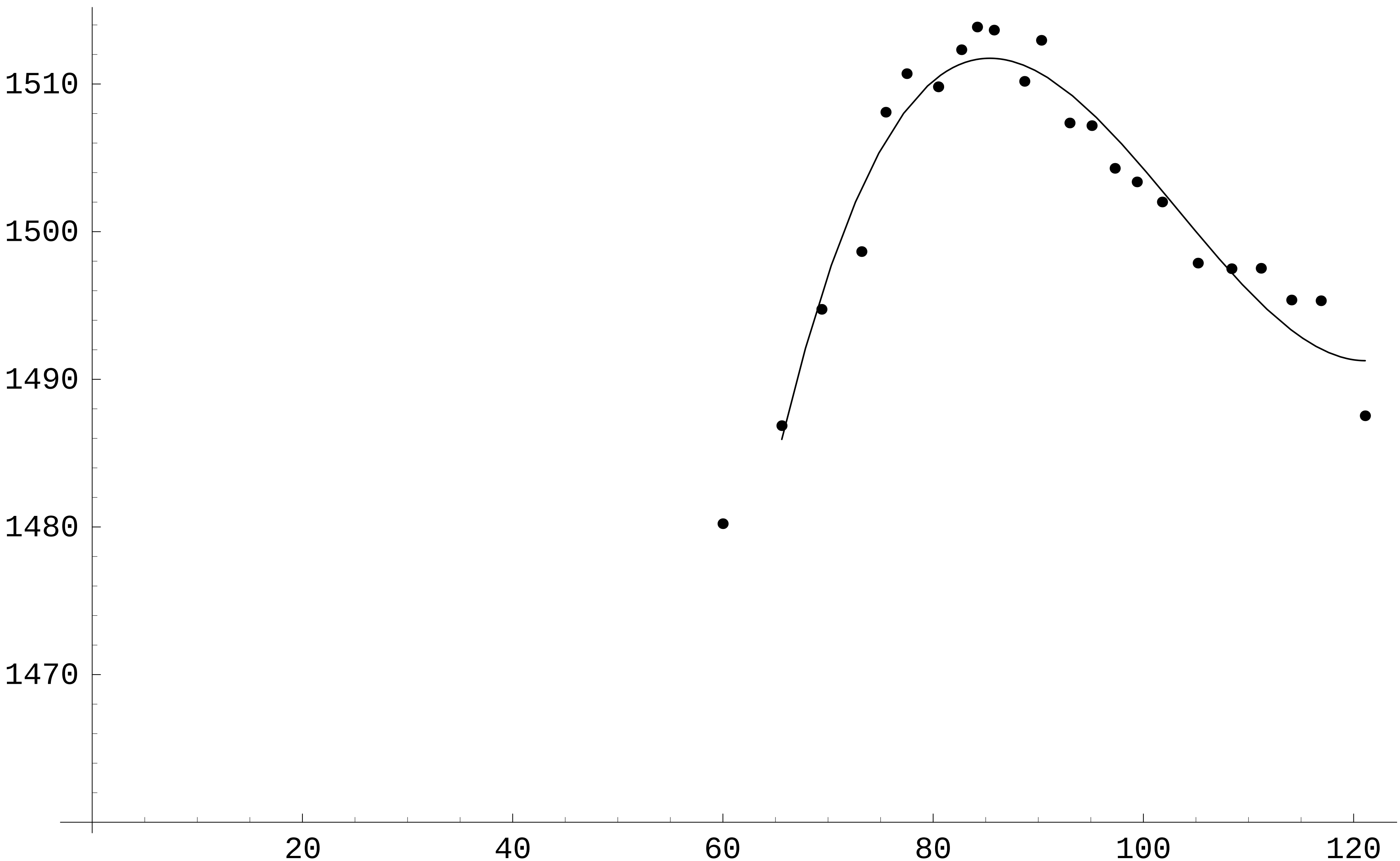


Iterate in each mass, fitting for each mass successively.

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

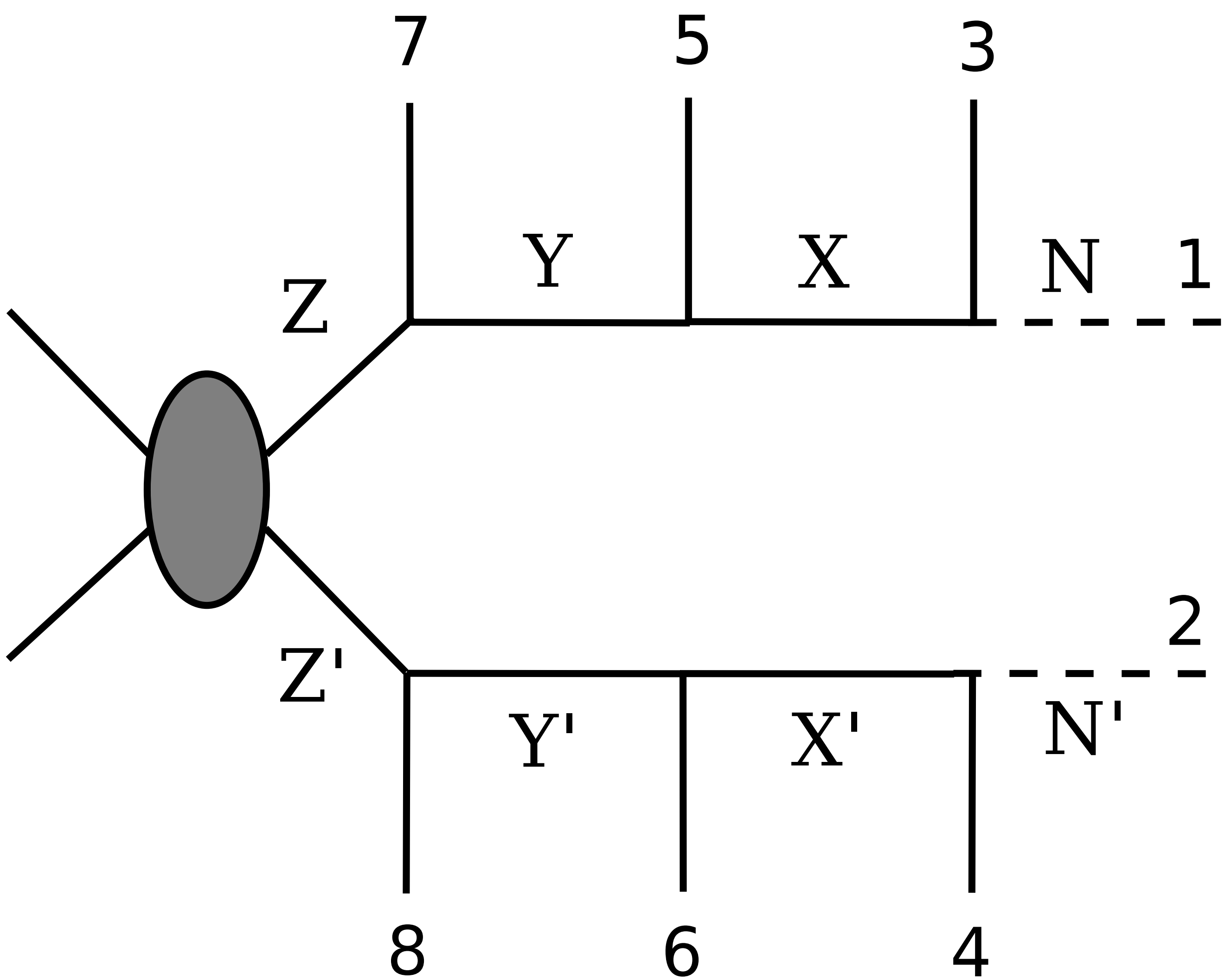


M_{3C}

A refinement was recently provided by Barr, Pinder and Serna, (arXiv:0811.2138) in which they use this topology, assume the mass differences are known, and perform a constrained min/max-imization over the missing energy.

This provides an upper/lower bound on the overall mass scale, and is a simple 1D variable.

Example Two: Overconstrained



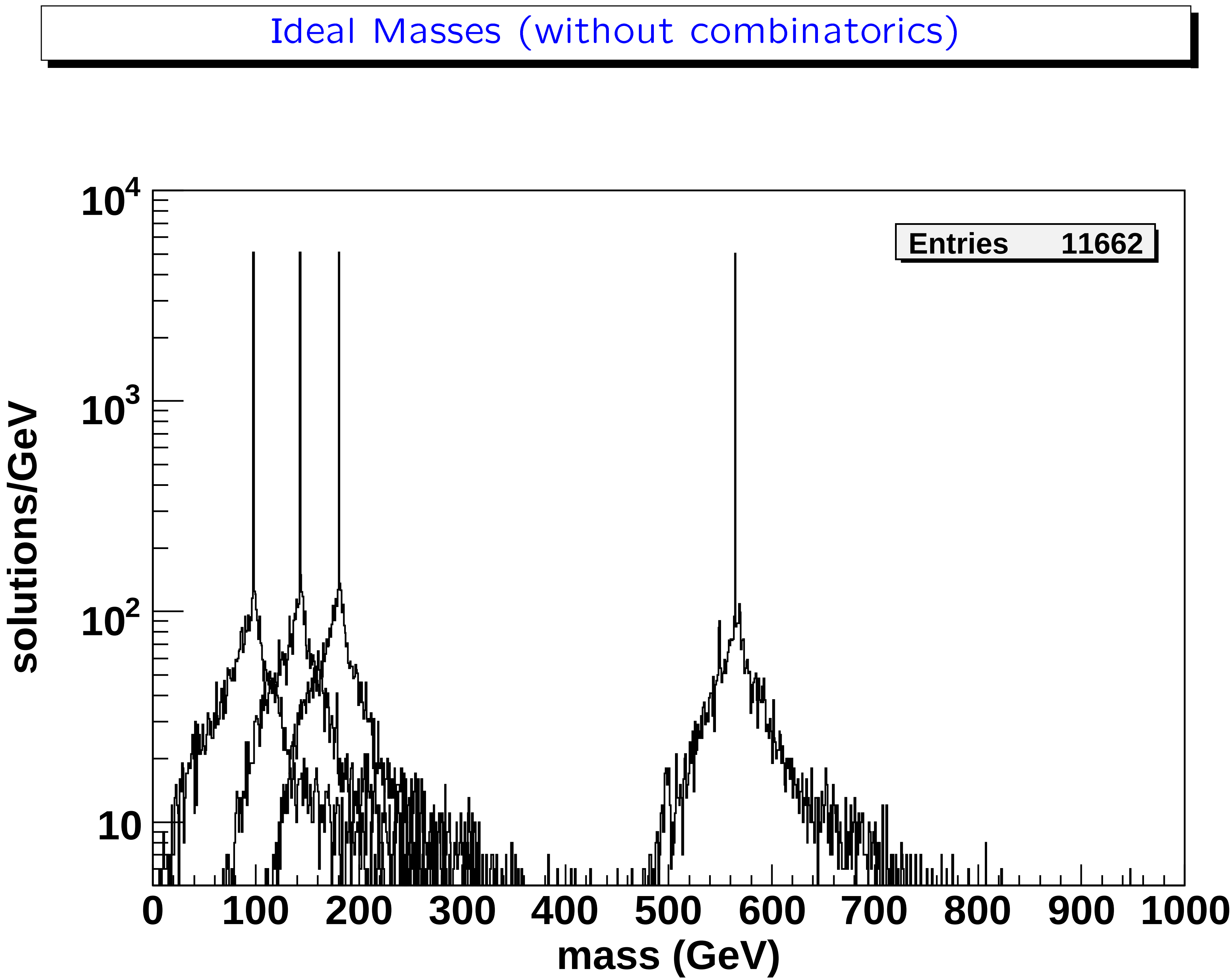
Constraint Equations

$$
\begin{array}{rrcl}
(M_Z^2 =) & (p_1+p_3+p_5+p_7)^2 & = & (p_2+p_4+p_6+p_8)^2, \\
(M_Y^2 =) & (p_1+p_3+p_5)^2 & = & (p_2+p_4+p_6)^2, \\
(M_X^2 =) & (p_1+p_3)^2 & = & (p_2+p_4)^2, \\
(M_N^2 =) & p_1^2 & = & p_2^2.
\end{array}
\tag{1}
$$

$$p_1^x + p_2^x = p_{miss}^x, \quad p_1^y + p_2^y = p_{miss}^y.$$

$$
\begin{array}{rcccl}
q_1^2 & = & q_2^2 & = & p_2^2, \\
(q_1+q_3)^2 & = & (q_2+q_4)^2 & = & (p_2+p_4)^2, \\
(q_1+q_3+q_5)^2 & = & (q_2+q_4+q_6)^2 & = & (p_2+p_4+p_6)^2, \\
(q_1+q_3+q_5+q_7)^2 & = & (q_2+q_4+q_6+q_8)^2 & = & (p_2+p_4+p_6+p_8)^2,
\end{array}
$$

$$q_1^x + q_2^x \;=\; q_{miss}^x, \qquad q_1^y + q_2^y \;=\; q_{miss}^y.$$



Ideal Masses (without combinatorics)
Entries 11662
solutions/GeV
mass (GeV)
10
10^2
10^3
10^4
0
100
200
300
400
500
600
700
800
900
1000

Application of Realism

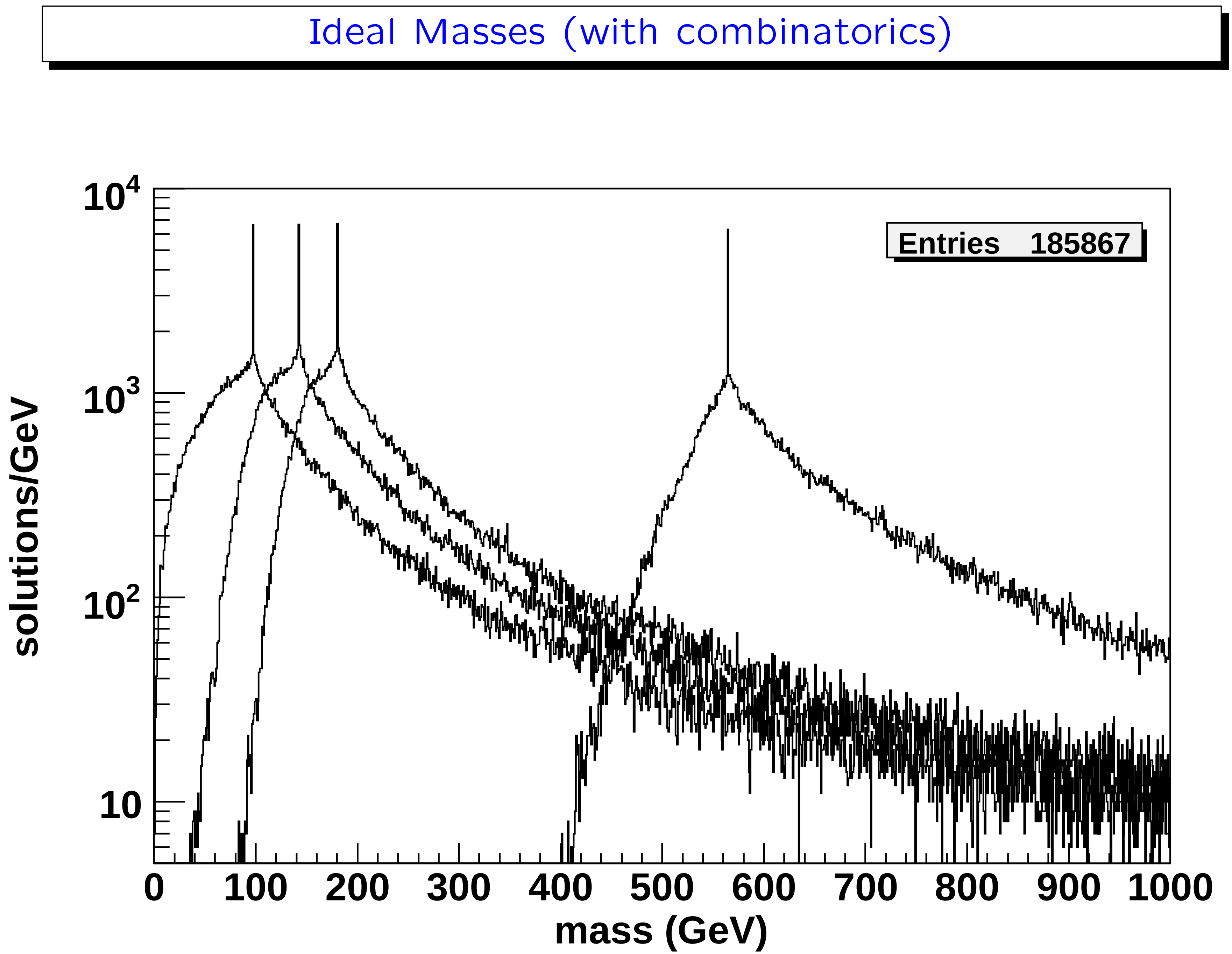
- Combinatorics: There are 16 choices of where to assign the leptons/jets per event for 4μ or $4e$, or 8 for $2\mu 2e$. Combinatorics are fundamental and *must* be taken into account. There is no magic cut which gets rid of them. Combinatorics also *carry information about mass*.

- Backgrounds: This signal has no real SM background. We include all SUSY backgrounds including $\widetilde{\tau}$ decays and $\widetilde{\chi}^0_2$ not from squark decay, and $\widetilde{g}$ events (which have extra hard jets).

- Finite widths: $\Gamma_{\widetilde{q}} = 5$ GeV, $\Gamma_{\widetilde{\chi}^0_2} = 20$ MeV, $\Gamma_{\widetilde{\ell}_R} = 200$ MeV.

- Mass splitting: Different flavor squarks have different masses by 6 GeV. Therefore, our squark mass result is an average of these signals.

Note that these techniques work with *very* few events (e.g. ten).



Ideal Masses (with combinatorics)
Entries 185867
solutions/GeV
10^4
10^3
10^2
10
0
100
200
300
400
500
600
700
800
900
1000
mass (GeV)

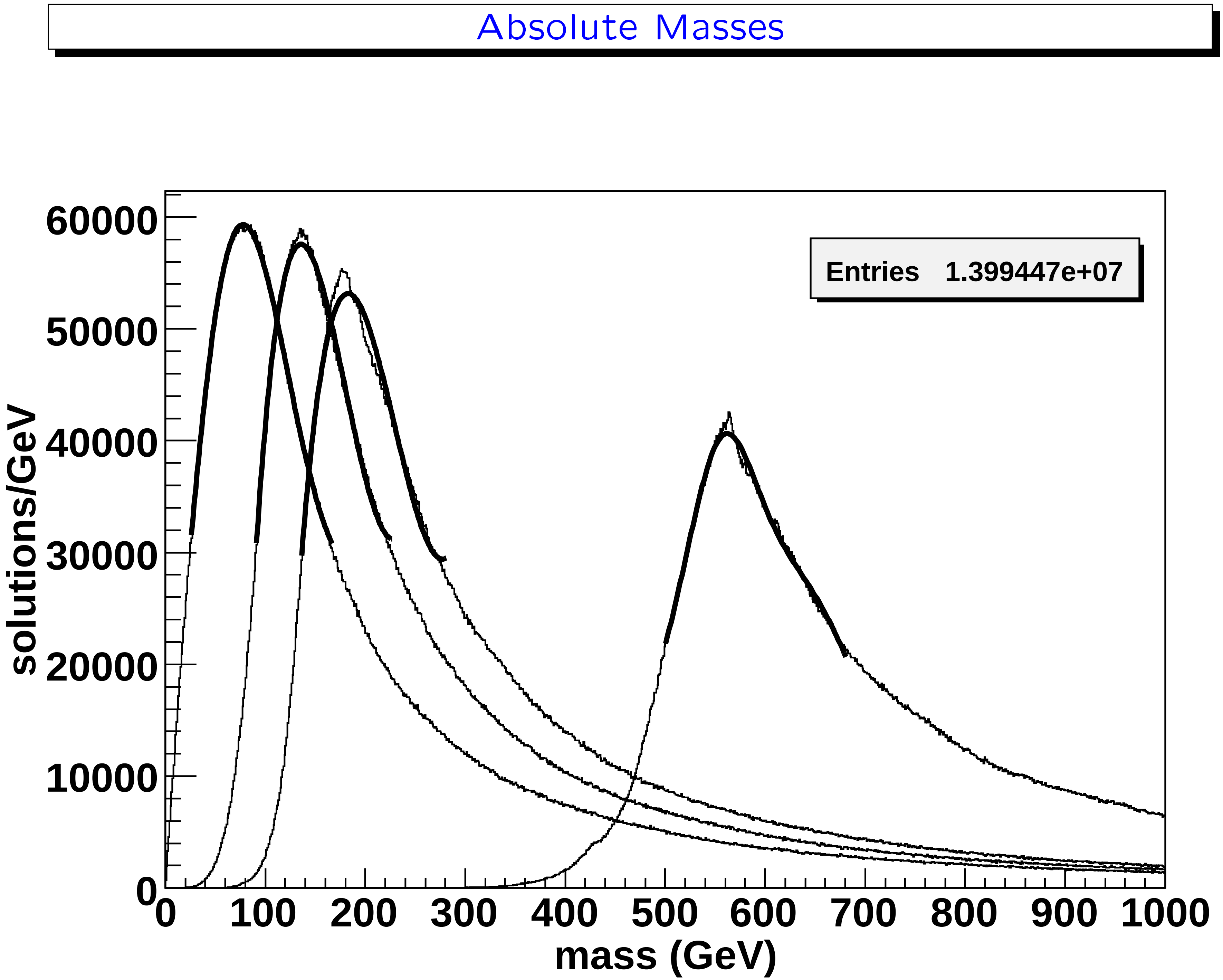
Application of Realism

We simulate all events with ATLFAST running in high-luminosity mode. We assume 300 fb^{-1} of luminosity. We require

- 4 isolated ($\Delta R < 0.4$) leptons with $p_T > 10$ GeV, $|\eta| < 2.5$. (flavors, charges chosen to match our $\tilde{\chi}_2^0 \to \tilde{\ell} \to \tilde{\chi}_1^0$ topology.

- no b-jets and ≥ 2 jets with $p_T > 100$ GeV, $|\eta| < 2.5$. The highest p_T jets are taken to be particles 7,8 (extra jets from parton shower/reconstruction are present).

- Missing $p_T > 50$ GeV.



Absolute Masses
Entries 1.399447e+07
solutions/GeV
60000
50000
40000
30000
20000
10000
0
0
100
200
300
400
500
600
700
800
900
1000
mass (GeV)

Extra Cuts

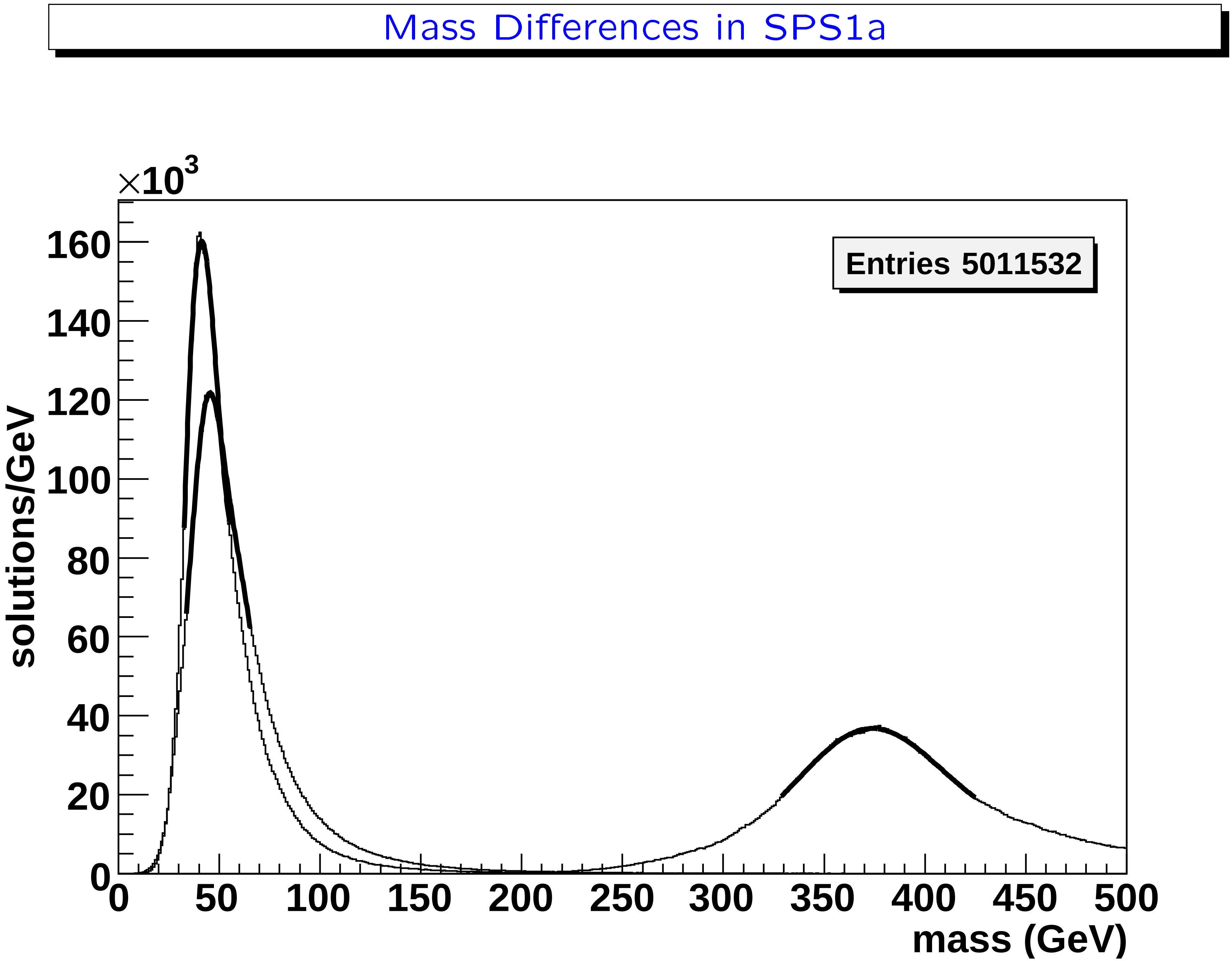
We add new cuts to improve S/B and decrease bias

- We require that each combination c in each event i have solutions with some combination in 75% of the other events. $N_{pair}(c,i) < 0.75 N_{events}$

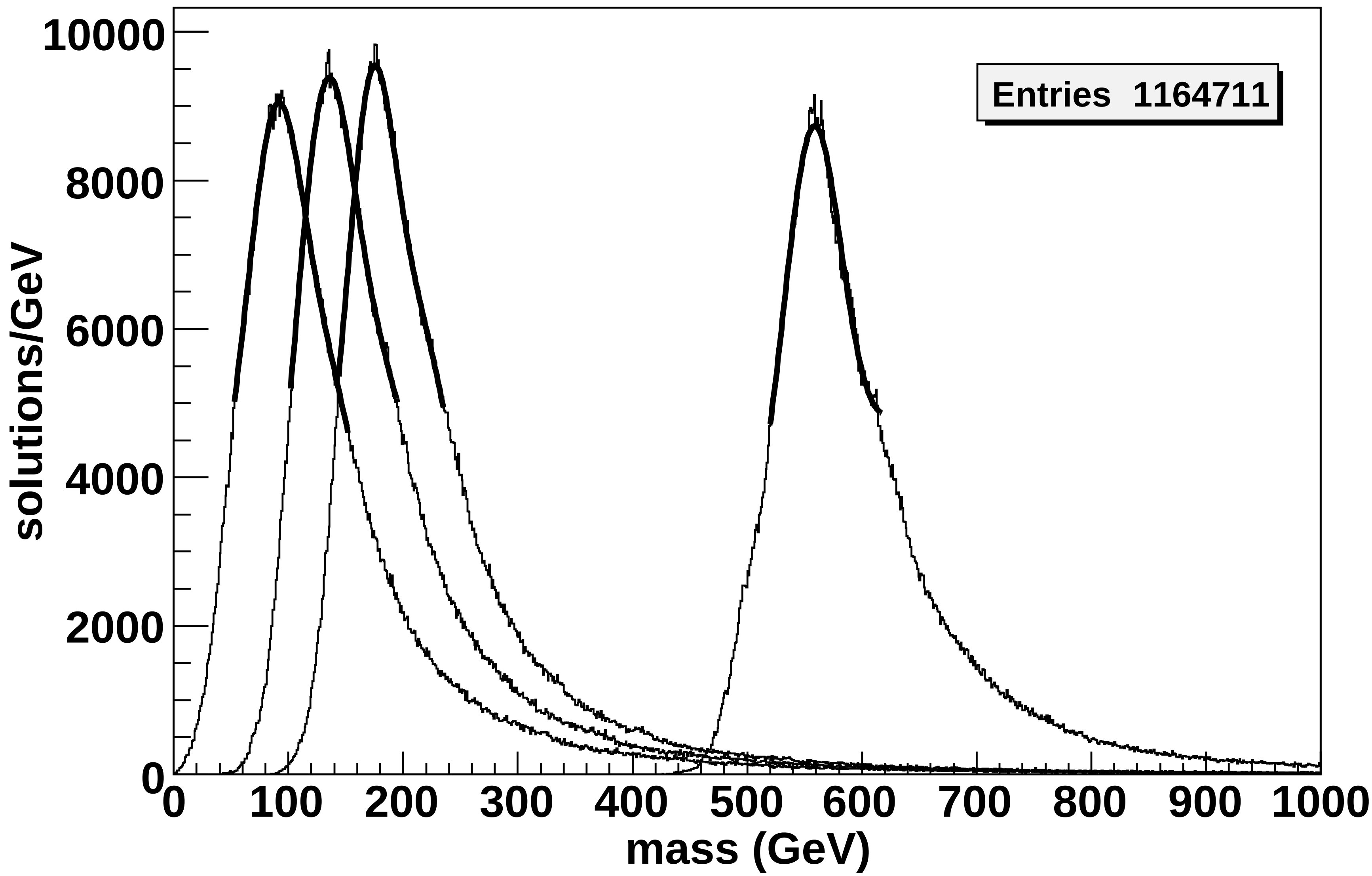
- We weight the final histogram by $1/N$ where N is the number of solutions in a given pair.

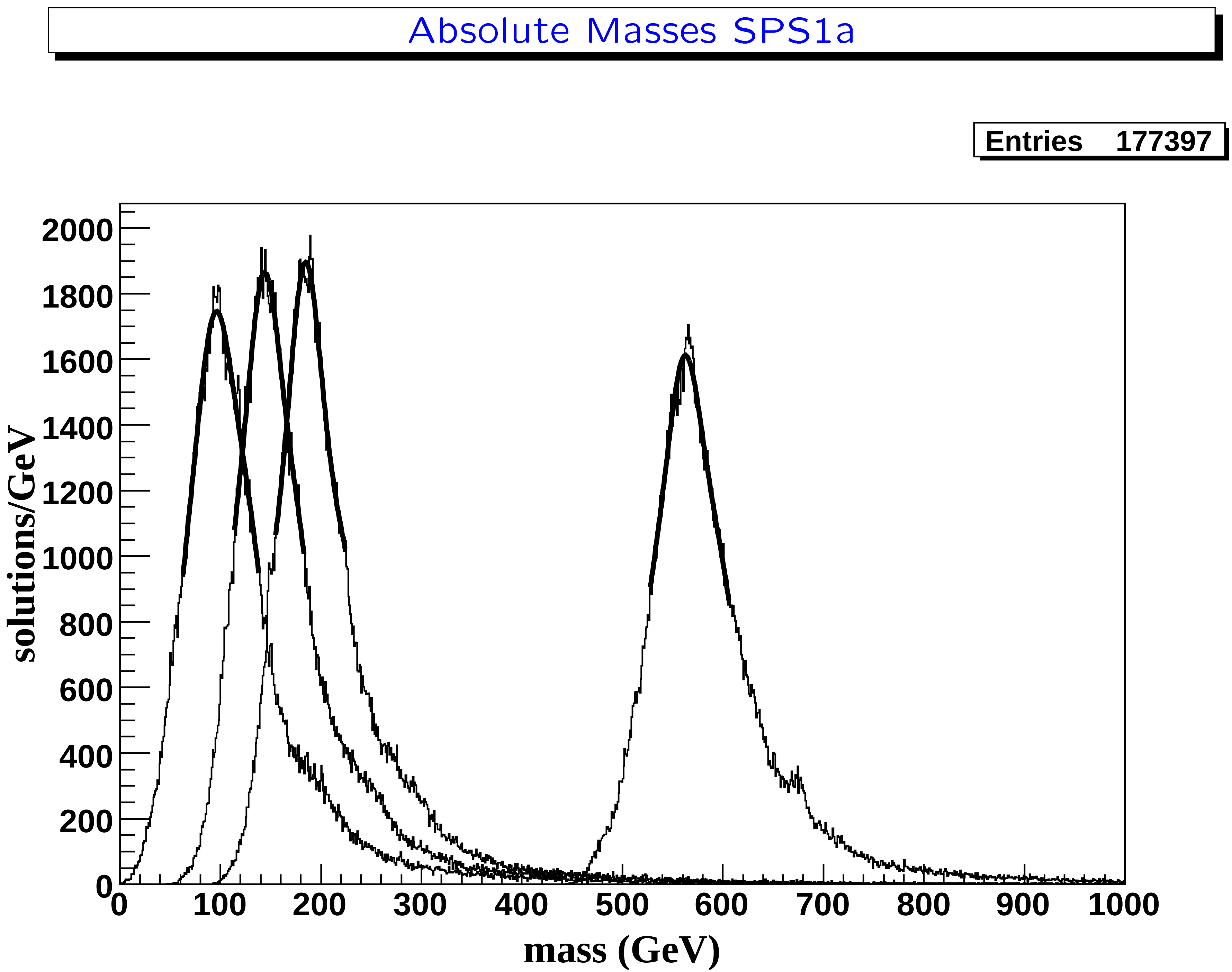
- We cut on the mass differences (window defined by 0.6 of peak height – e.g. Full Width at 0.6 Max)

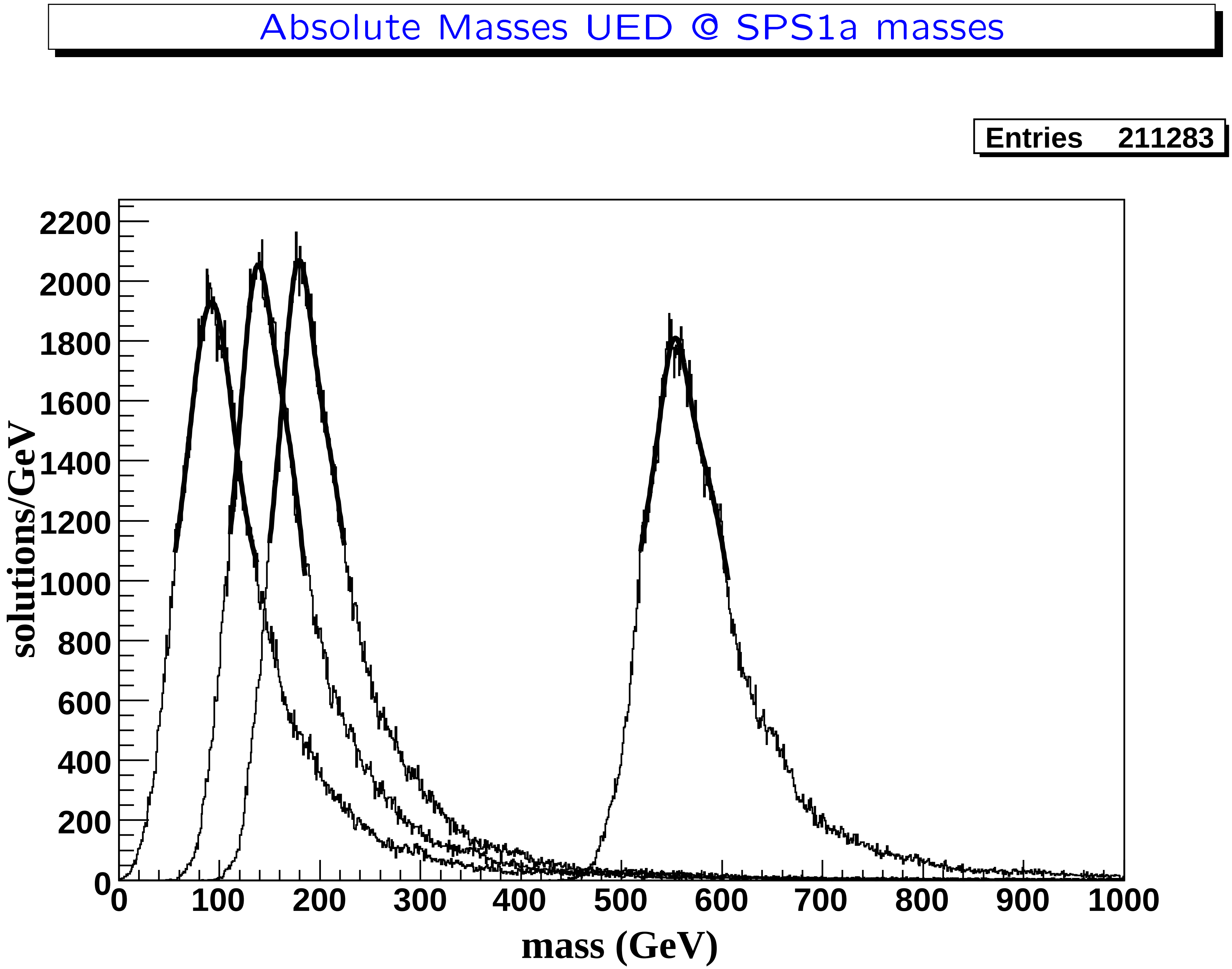
There are many other interesting manipulations one can do, that are quite different from cutting on physical observables.


Mass Differences in SPS1a
$\times 10^3$
Entries 5011532
solutions/GeV
0
20
40
60
80
100
120
140
160
0
50
100
150
200
250
300
350
400
450
500
mass (GeV)

Absolute Masses SPS1a (cuts on ΔM)




Absolute Masses SPS1a
Entries 177397
solutions/GeV
0
200
400
600
800
1000
1200
1400
1600
1800
2000
0
100
200
300
400
500
600
700
800
900
1000
mass (GeV)



Absolute Masses UED @ SPS1a masses
Entries 211283
solutions/GeV
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
0
100
200
300
400
500
600
700
800
900
1000
mass (GeV)

Results

We fit peaks using a gaussian+quadratic polynomial, and use the maximum as our mass estimator. This is a biased estimator, but can be used to estimate our statistical error by repeating the measurement. Using 10 independent sets of Monte Carlo, for the SPS1a point with masses {91.7,135.9, 175.7 558.0}

$$\begin{aligned} m_N &= 94.1 \pm 2.8\text{GeV}, \\ m_X &= 138.8 \pm 2.8\text{GeV}, \\ m_Y &= 179.0 \pm 3.0\text{GeV}, \\ m_Z &= 561.5 \pm 4.1\text{GeV}. \end{aligned} \tag{2}$$

There are 539 signal + 195 background events in this sample after all cuts.

Precision is degraded by our "bias reduction" procedure. This is great for getting the mass within 5% very quickly (without scanning in masses), but final errors using these techniques is about a factor 2 better.

Code Availability

For the construction of the polynomial system, the problem can be divided into two stages: a linear stage and a quadratic stage. (Don't spend a lot of time with equations in Mathematica/Maple, there's an easier way to do it, and it's just a matrix) Each missing particle mass-shell constraint provides one quadratic, and any resonance with two or more invisible particles downstream provides a quadratic.

Solving a system of 2 quadratics is straightforward (it can be reduced to a quartic, and solved analytically).

Solving systems of $n > 2$ quadratics is highly nontrivial.

We have packaged up our code to solve a 2-quadratic system and 3-quadratic system, and the construction of the quadratic systems described in our paper(s).

http://particle.physics.ucdavis.edu/hefti/projects/doku.php?id=wimpmass

I have some (unfinished) C++ classes which are very general and could be used for any process with any number of quadratics. (I need collaborators)

Summary

We *really can* make plots of mass!

Breit-Wigners appear in plots of mass, and the appearance of a Breit-Wigner is *real proof* of a new particle. Edges/slopes are far less convincing that one has discovered a new particle and not a detector effect (or a misinterpretation of a resonance as an edge!)

These techniques can be thought of as answering: Is the N-particle narrow-width likelihood L_N zero or non-zero?

These techniques require ≥ 4 resonances for 2 missing particles, or ≥ 7 resonances for 3 missing particles.

These techniques use *all* available data, (including missing p_T) and automatically take into account the fact that there are multiple solutions and combinatorics.

If the signal nature presents us is compatible with these requirements, this is really the the best, unambiguous variable to use.

What do I use? What needs to be done?

Variable	Ref	good for	fails for
M_{eff}			not specific
H_T			not specific
M_T	(ancient)	One missing	
Cascade Edges			calorimeter nonlinearity?
M_{T2}	hep-ph/9906349		
M_{T2} kink	0709.0288	large mass differences	
M_{2C}	0712.0943		
M_{3C}	0811.2138		
$\sqrt{s}_{min}$	0812.1042		
Overconstrained	0802.4290	$\geq$ 5 cascade resonances	
Exactly cons. poly	0707.0030	4 cascade resonances	
???		3 missing	