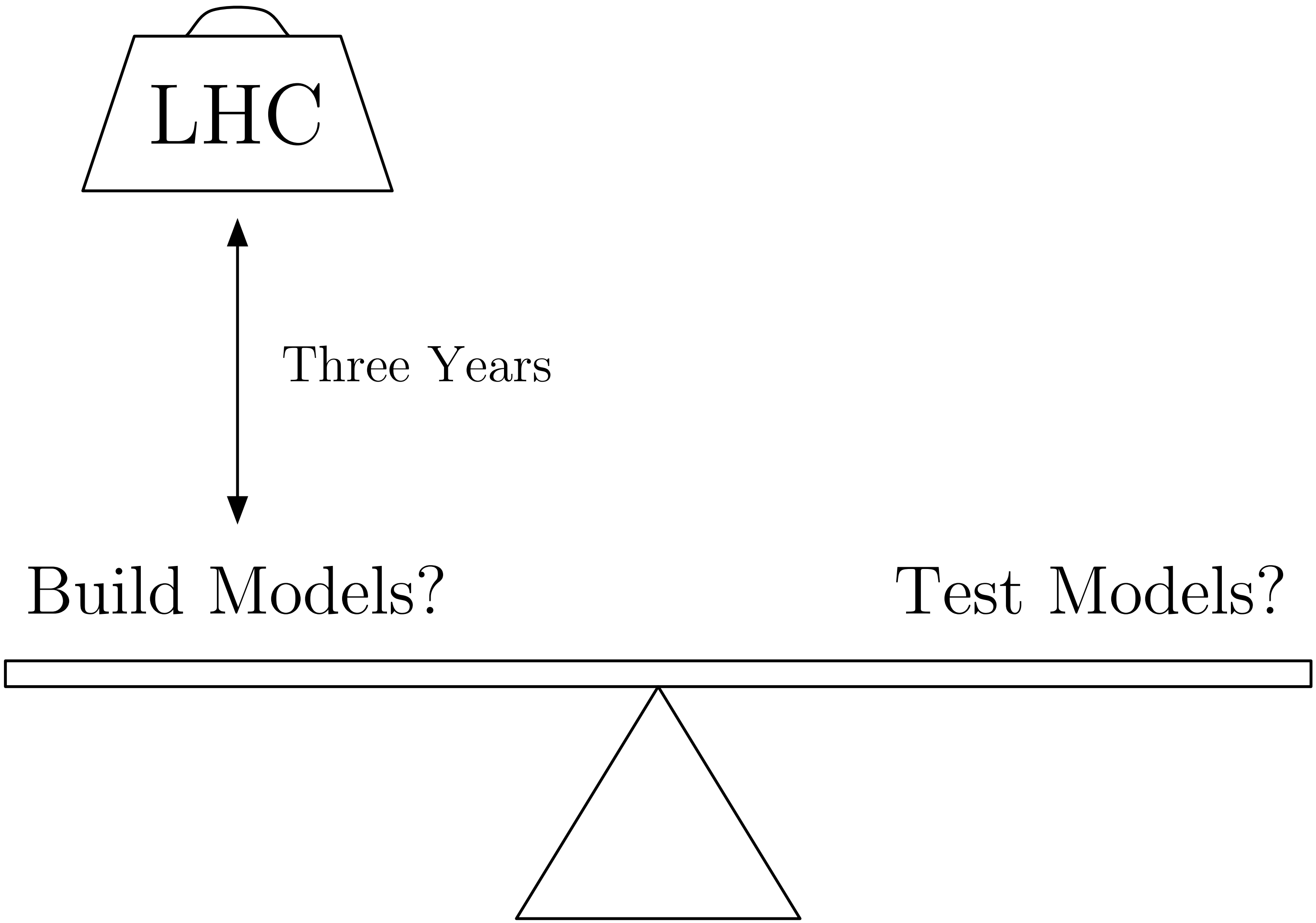
Supersymmetry and the LHC Inverse Problem

Jesse Thaler

with N. Arkani-Hamed, G. Kane, and L.-T. Wang

Particle Theory Circa 2005



LHC

Three Years

Build Models?

Test Models?

Models Still Worth Building

1. “Best SUSY Ever” (or Best Composite Higgs Ever?)

   e.g. 6 × Nomura, *et al.*

2. Proof of Concept

   e.g. Twin Higgs (Chacko, Goh, Harnik)

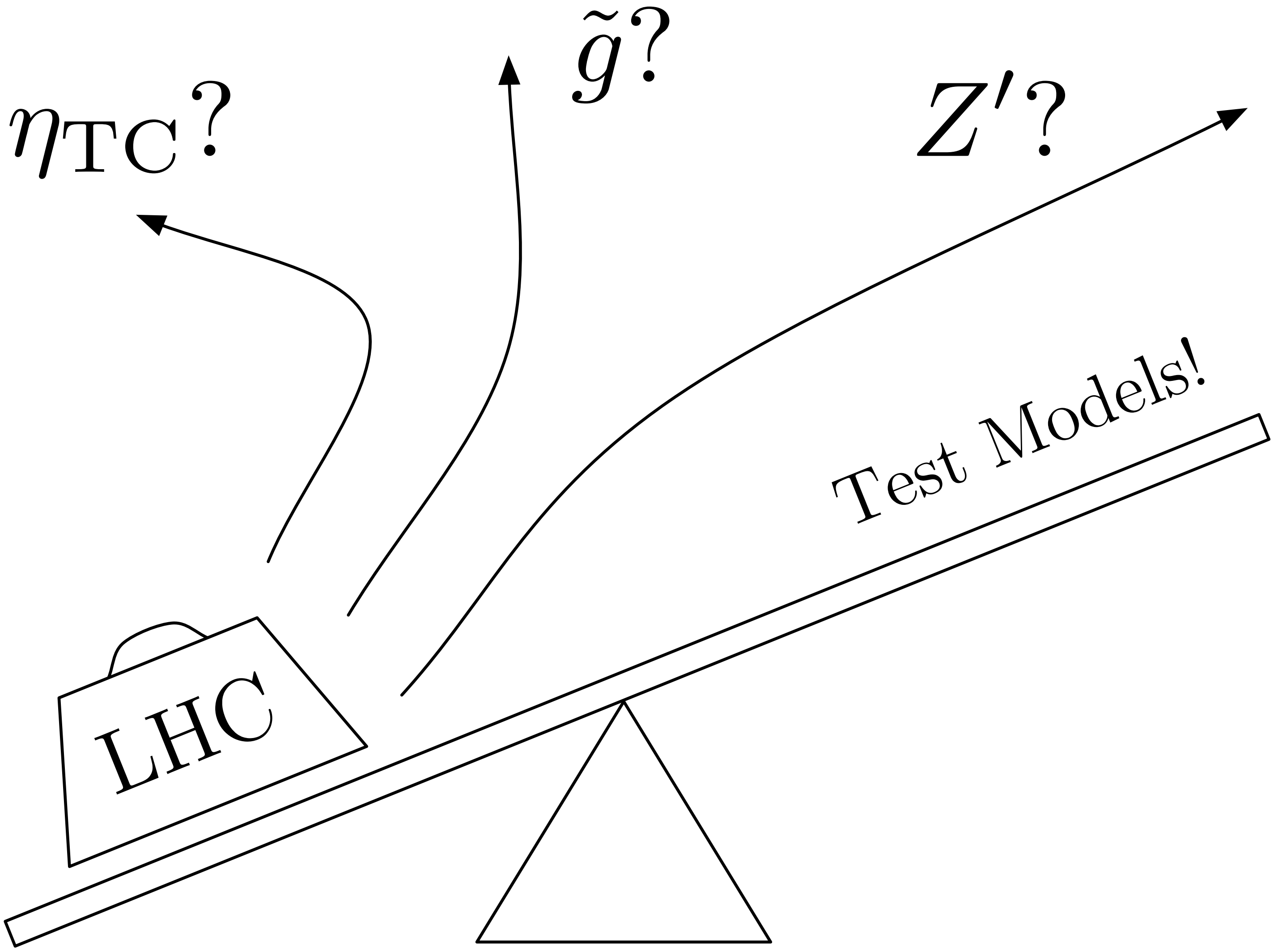
3. Unnatural Theories

   e.g. Minimal Dark Matter (Mahbubani, Senatore)

4. Canonical Models

   e.g. Little M-theory (Cheng, JKT, Wang)

Particle Theory Circa 2008



One Year at the LHC

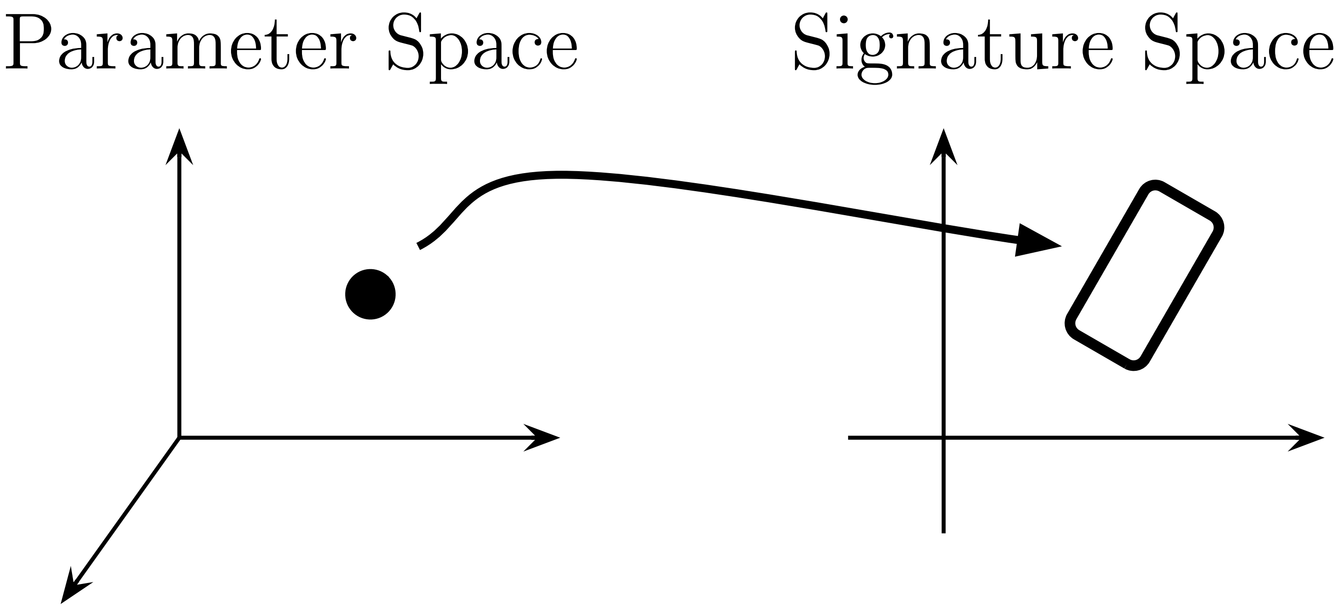
$$10 \text{ fb}^{-1}$$

Experimental Data $\longrightarrow$ Theoretical Models?

Assume it is the MSSM. Without forcing constraints on soft parameters, can we "rule in":

Gaugino Unification?
Weak Scale Dark Matter?
$\tan\beta$?

Standard Method

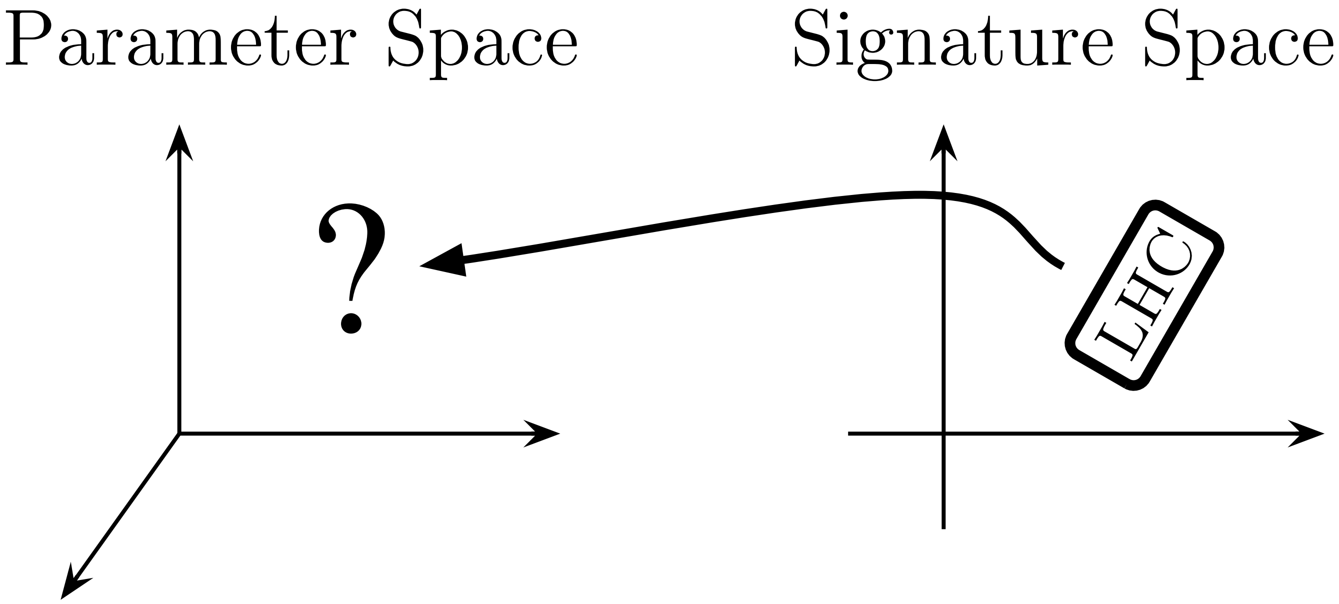


TDR, Benchmark studies, LEP/Tevatron bounds, etc.

Reduced parameter set (mSUGRA, AMSB, GMSB):

Experimental Data ⟶ Precision Measurements

The Inverse Problem

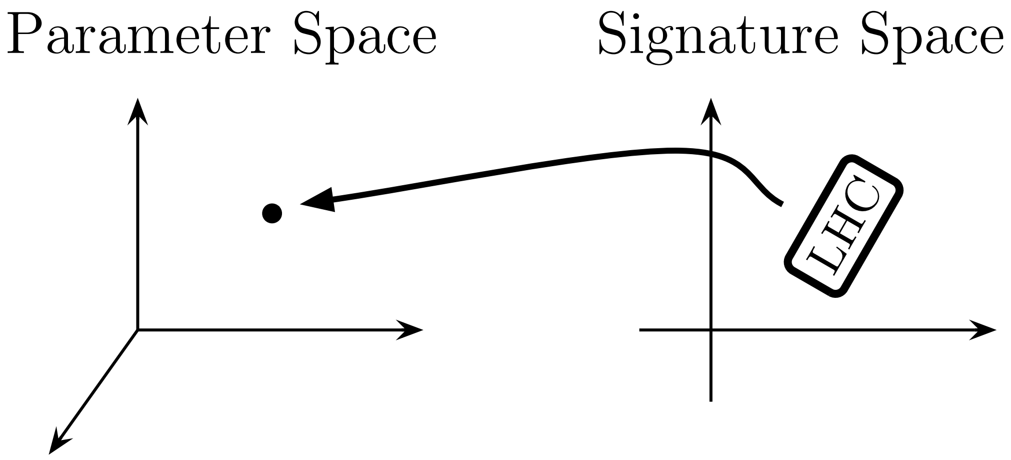


Much more interesting!

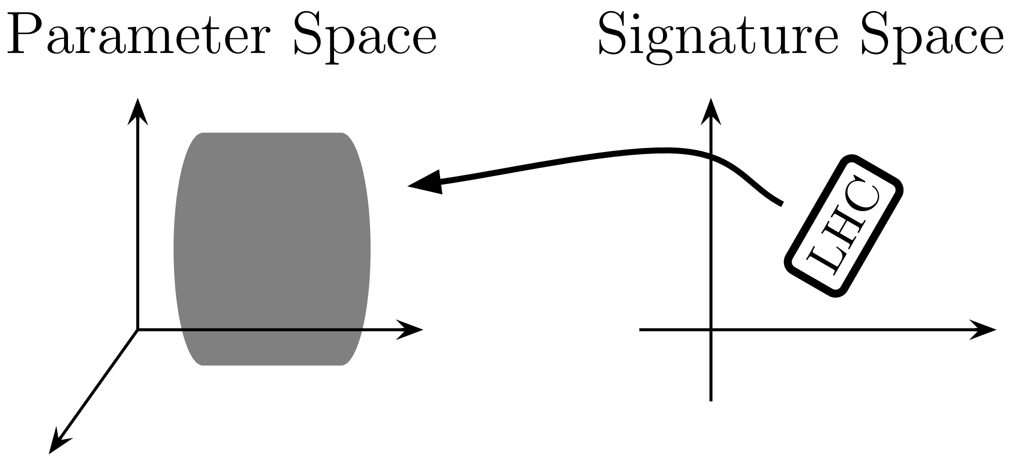
$$\text{Data} \xrightarrow{?} \text{gaugino unification, dark matter, ...}$$

Much more important! (500 GeV ILC in 10 years?)

Best of all Possible Worlds

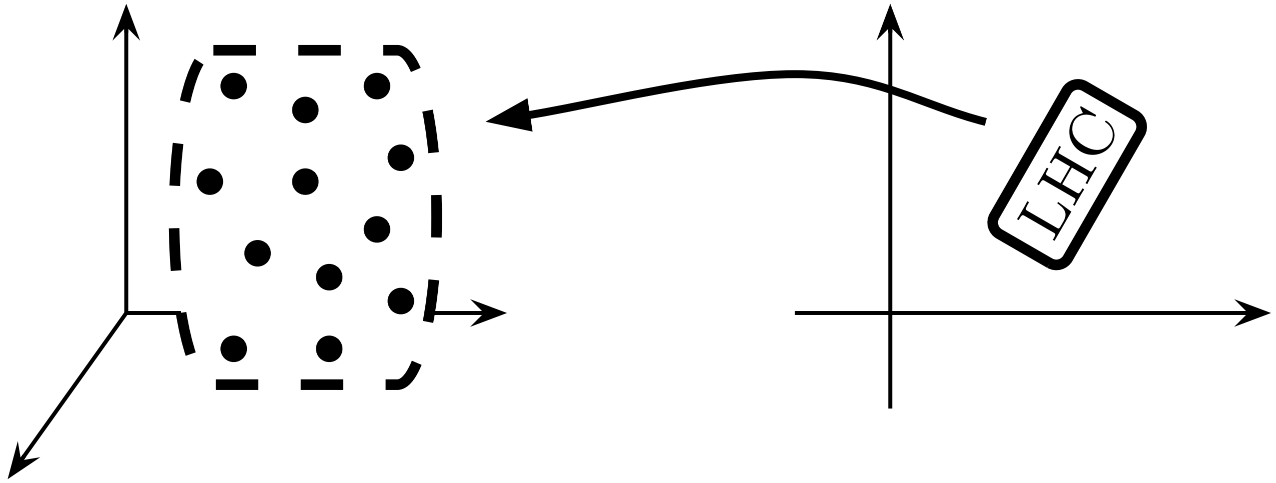


Worst of all Possible Worlds



Our Picture of the Inverse Map

Parameter Space Signature Space



Degeneracies!

Many small footprints in a large overall region.

Advertisements

Do LHC Experimentalists Need Our Help?

No $\left\{ \begin{array}{l} \text{A long-lived gluino at 300 GeV...} \\ \text{A 2 TeV } Z'\text{...} \\ \text{A Higgs and nothing else...} \end{array} \right.$

Yes $\left\{ \begin{array}{c} \text{Almost any model of TeV scale} \\ \text{physics with a } \mathbf{Z}_2 \text{ symmetry} \\ \text{and heavy-ish colored particles.} \end{array} \right.$

The L(ittle) H(ierarchy) C Problem

Precision EW: $\Lambda_{\text{higher dim}} \sim 5 - 10$ TeV

Direct Bounds (if not hidden):

$$\tilde{\chi}_0 \gtrsim 100 \text{ GeV}, \quad h \gtrsim 115 \text{ GeV}, \quad \tilde{g} \gtrsim 300 \text{ GeV}$$

Flavor-independent mediation scheme SUSY:

Colored Particles > Electroweak Particles

LHC is a *pp*-collider: Smallish σ_{SUSY} if heavy gluino, huge SM background to “jets plus missing energy.”

So What?

LHC is supposed to be a “discovery” machine...

Eventually, a next generation collider (ILC?)...

But model builders are impatient!

“What will we know
after one year?”

Concerns About Uniqueness

Sfitter, Fittino: Observables → SUSY parameters

Without ILC observables, algorithms don't coverge.

Miller, Osland, Gjelsten: Detailed study of SPS 1a

Edges/Endpoints → Various Mass Scenerios

Lester, Parker, White: Markov Chain Technique

Some discussion of decay chain ambiguity.

Back to the Inverse Problem...

The Inverse Problem: A Statistical Approach

LHC Data → Theoretical Models

Too many models, impossible to analyze them all!

But, if mapping is not one-to-one, all we need is one degeneracy to prove so.

In fact, “birthday problem” gives us a simple statistical technique to measure the degree of one-to-one-ness!

Outline: The LHC Inverse Problem

1. What is our statistical technique?

   Generalized Pigeon Hole Principle

2. What kind of models are we simulating?

   15 Parameter MSSM Scan

3. What kind of data are we using?

   1153 LHC Observables

4. What kind of degeneracies are there?

   Flippers, Sliders, and Squeezers

Statistics of the MSSM

Simple Statistical Question

$$N_{\text{models}} \overset{?}{\gtreqless} N_{\text{sig bins}}$$

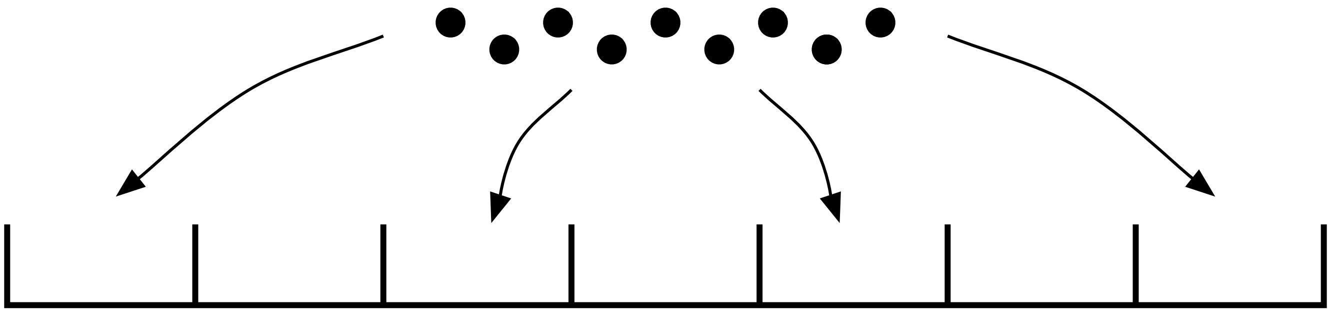
N_{models}
↑
different MSSMs
measured in desired
parameter accuracy

$N_{\text{sig bins}}$
↑
distinguishable
experimental
outcomes at the LHC

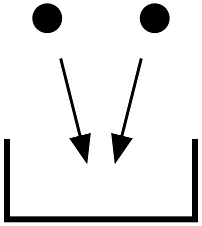
If $N_{\text{models}} > N_{\text{sig bins}} \Rightarrow$ Pigeon hole principle!
Guaranteed degeneracies at the LHC.

How to Count Degeneracies

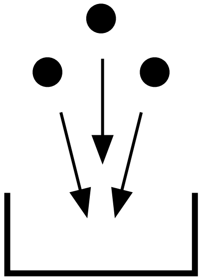
m balls



N bins



$$= \frac{1}{2!}\,\frac{m^2}{N}$$



$$= \frac{1}{3!}\,\frac{m^3}{N^2}$$

As of Last Night...

We simulated $m = 39137$ MSSMs (+ 3918 waiting...)

Number of pairs with matching LHC signatures = 2120

$$N_{\text{sig bins}} \sim m^2/d \sim 10^6$$

$$N_{\text{models}} \sim 10^8 \text{ (We'll return to this later.)}$$

Probability of choosing correct model at LHC = 1%

Equivalently, for every model, there are 99 other (well motivated?) models with the same LHC signatures.

More Realistically

Demonic Bins:

$$x \text{ of models} \mapsto \lambda \text{ of signature bins}$$

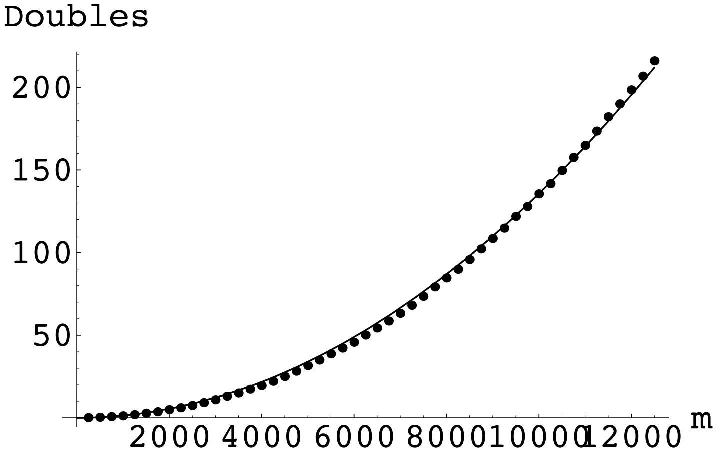
Normal Bins:

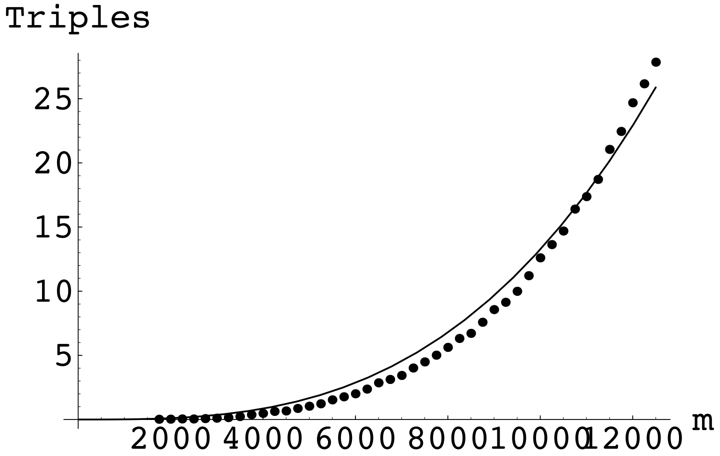
$$1 - x \text{ of models} \mapsto 1 - \lambda \text{ of signature bins}$$

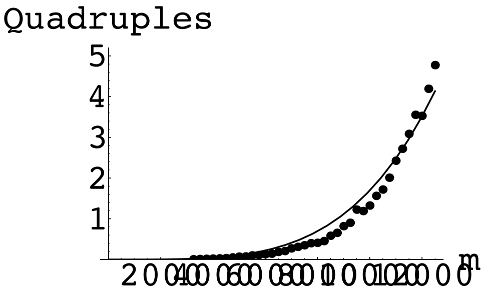
$$k\text{−tuples} = \frac{m^k}{k!N^{k-1}} \left(\frac{x^k}{\lambda^{k-1}} + \frac{(1-x)^k}{(1-\lambda)^{k-1}} \right)$$

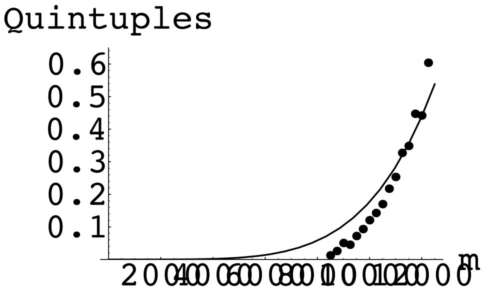
Fitting k−tuples will tell us if degeneracies are generic.

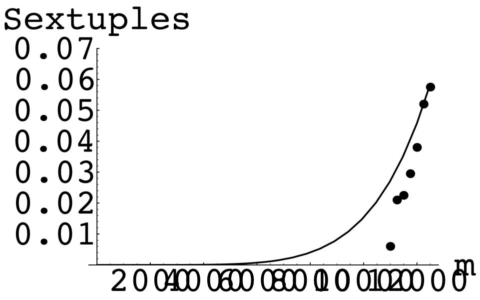
$N_{\text{models}} = 775000$







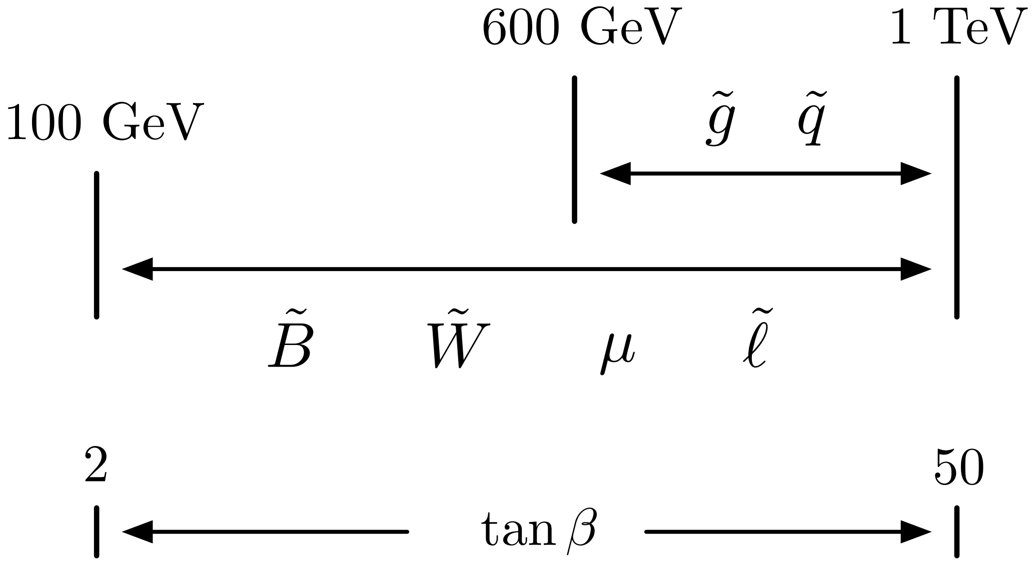




2.8% of models map to 0.07% of signature bins

What Kind of Models?

Choosing 39,137 MSSMs



$$\tilde{q}\left\{\begin{array}{ccc} \tilde{q}_L & \tilde{q}_{Ru} & \tilde{q}_{Rd} \\ \tilde{q}^3_L & \tilde{q}^3_{Ru} & \tilde{q}^3_{Rd} \end{array}\right.$$

$$\tilde{\ell}\left\{\begin{array}{cc} \tilde{\ell}_L & \tilde{\ell}_R \\ \tilde{\ell}^3_L & \tilde{\ell}^3_R \end{array}\right.$$

Constraints: Nothing more than 50 GeV decoupled

Max Colored > Max Electroweakino > Max Slepton

No slepton LSP.

Generating 39,137 MSSMs

Thanks to “Opus” at UMich, “Sauron” at Harvard:

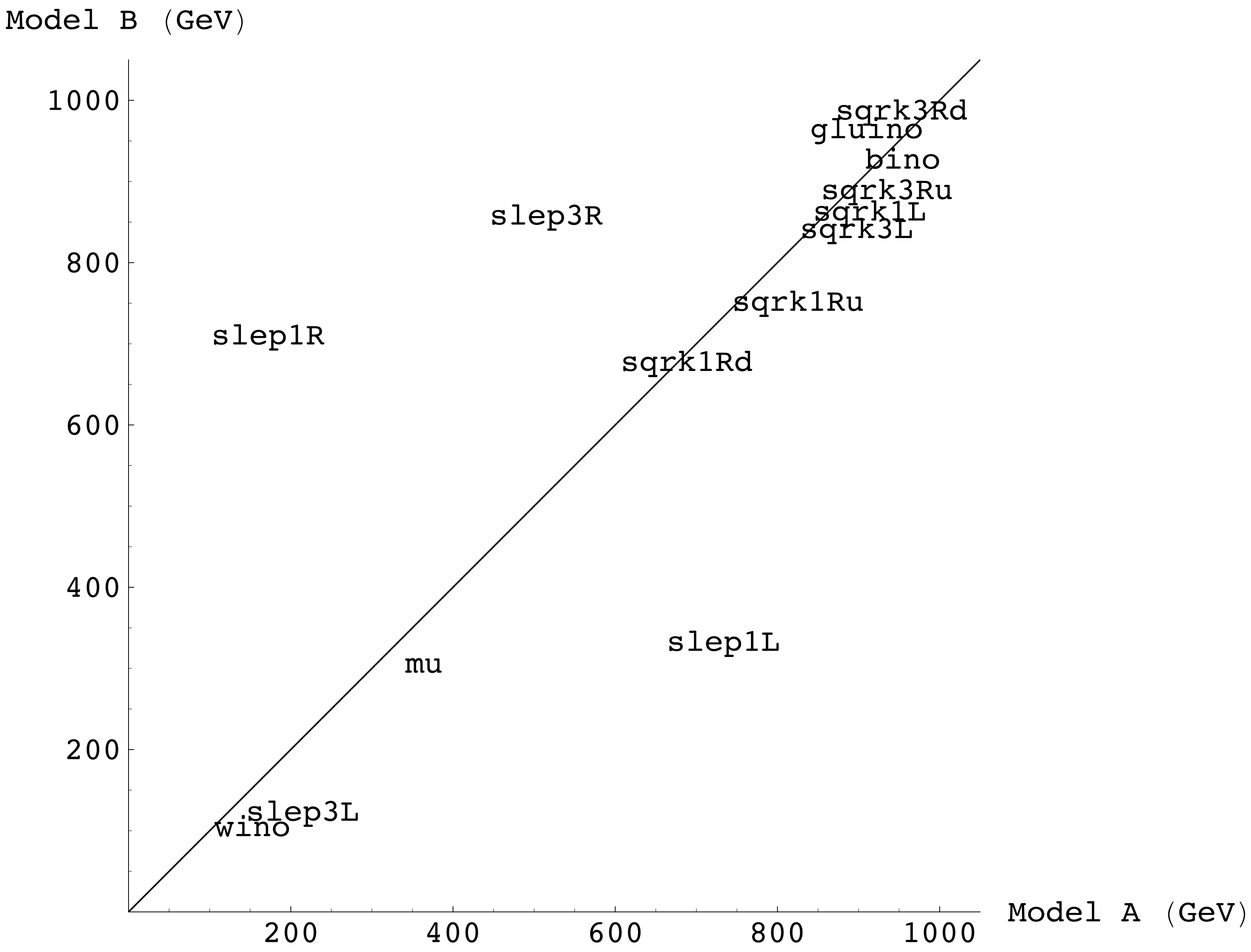
1 Model / CPU / hr

PYTHIA → PGS (Pretty Good Simulation)

Two terabytes of storage for raw data like:

```
#typ      eta    phi       pt     jmas ntrack   btag
1  3    0.816  0.492    36.20     1.00    3.0    0.0
2  4    2.674  5.539    69.72     7.54    0.0    0.0
3  4    1.794  5.499    45.57     5.65    9.0    0.0
4  4   -0.932  1.546    34.33     6.02    9.0    0.0
5  4    1.622  4.151    70.02    23.54   10.0    0.0
6  4    2.636  3.328    22.96     5.84    0.0    0.0
7  4    0.340  2.154    65.68     2.60    9.0    0.0
8  6    0.000  5.441    23.70     0.00    0.0    0.0
```

A Typical Pair: 2005 vs. 32444



Some Peculiarities

Non-decoupling constraint tends to decouple particles!

Very rarely three light inos.
Very rarely cascade decays ino-slepton-ino.
Most leptons in final states from W and Z.

Future run with guaranteed cascade decays.

(Is the degeneracy story different when the spectra looks more like mSUGRA, AMSB, GMSB?)

What Kind of Data?

Blackboard Interlude

LHC Observables

Triggers: $\not{E}_T > 125$ GeV

$$\sum_{3j\; w/\; 0\ell} P_T > 300 \text{ GeV}, \qquad \sum_{3j\; w/\; 1^+\ell} P_T > 225 \text{ GeV}$$

Categorize all events by the number of:

j, bj, e, μ, hadronic τ, γ

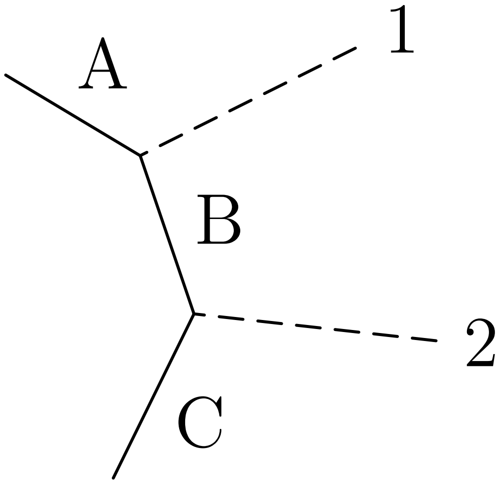
Count NEV, everything else ratios.

Errors: $\sqrt{N}$ error (times scale factor) plus percent error (1%, except 15% for NEV).

Kinematic Histograms

$$m_{\text{eff}} = \sum |p_T|, \qquad m_{\text{inv}}^2 = \left(\sum p_\mu\right)^2$$

Edges/endpoints of distributions.



Offshell B (Endpoint):

$$m_{\text{inv}}^{12} = m_A - m_C$$

Onshell B (Edge):

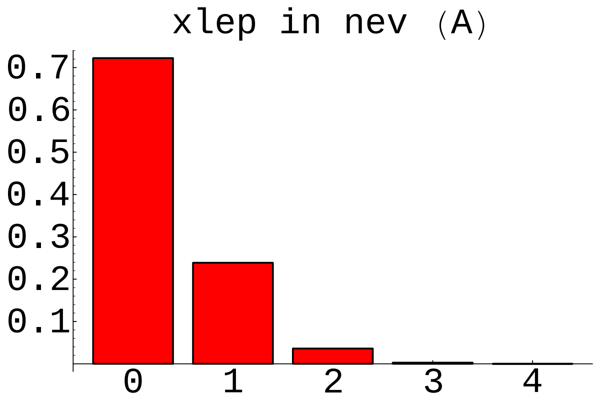
$$m_{\text{inv}}^{12} = \sqrt{(m_A^2 - m_B^2)(m_B^2 - m_C^2)/m_B^2}$$

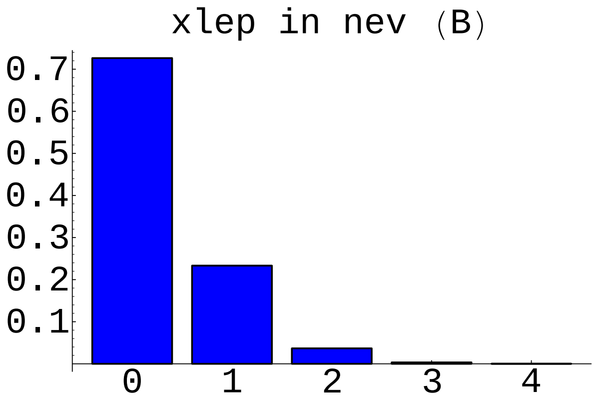
Distributions in quantiles. (Deciles, m_{eff}; 20-tiles, m_{inv})

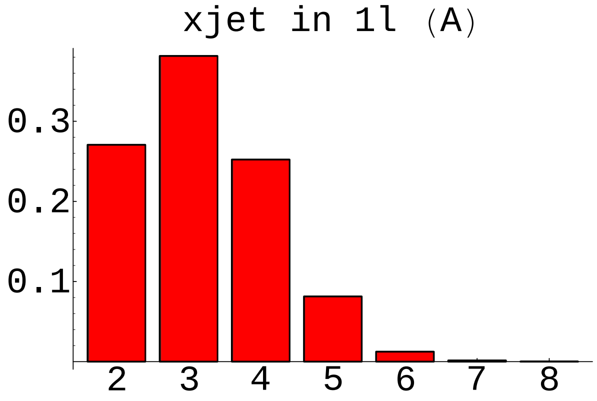
Match distribution, not just endpoint/edge.

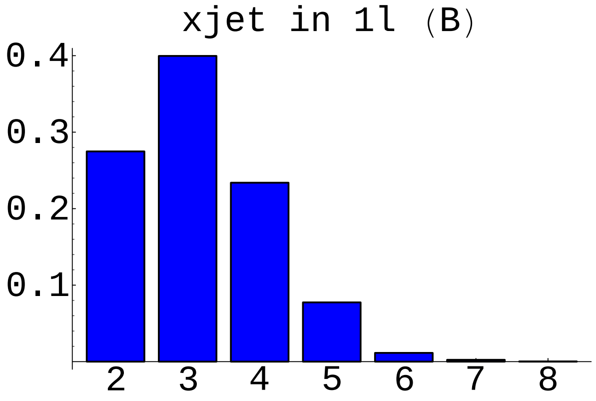
Typical Data: 2005 vs. 32444

$$\mathrm{NEV}_A = 30253, \quad \mathrm{NEV}_B = 32251$$

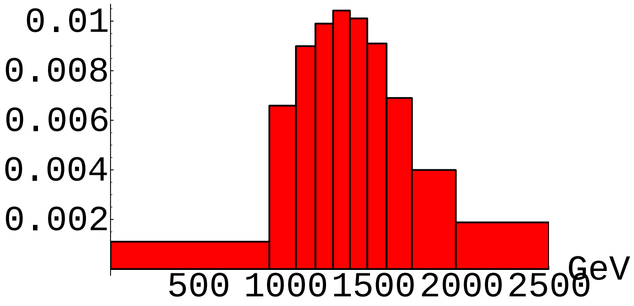




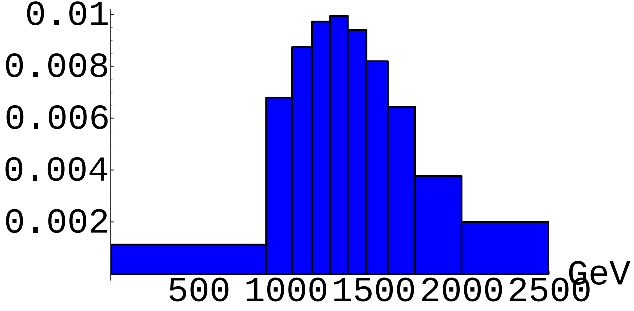




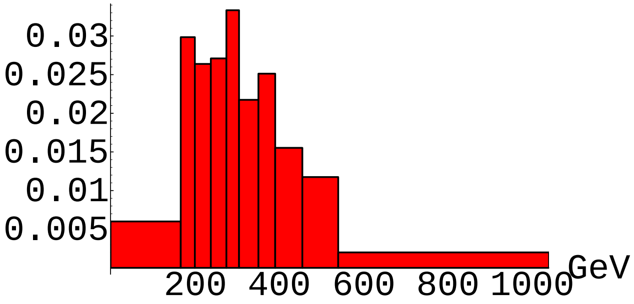
Meff 0l (A)



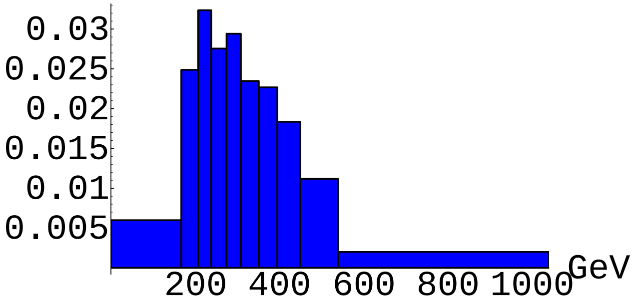
Meff 0l (B)



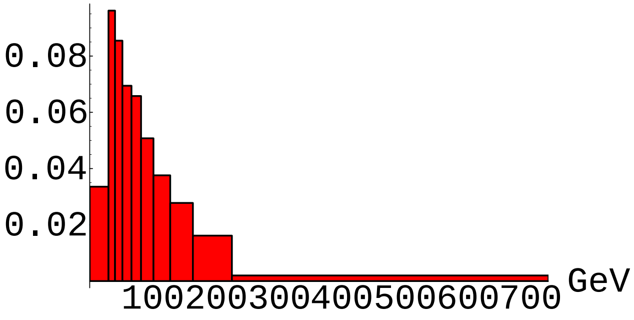
MissingPT 2l (A)



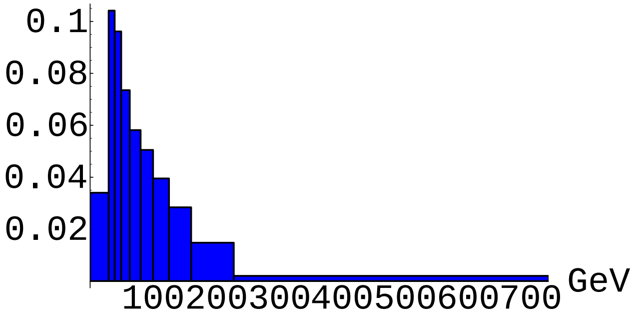
MissingPT 2l (B)

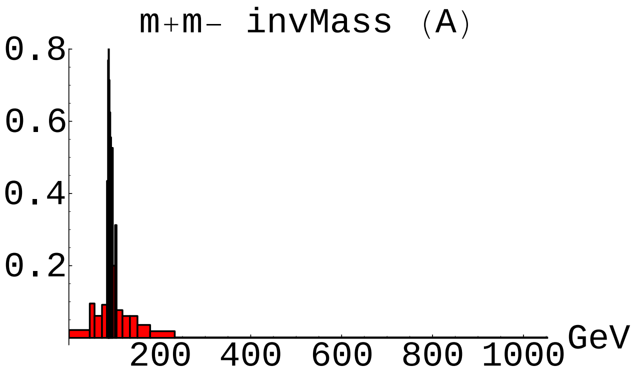


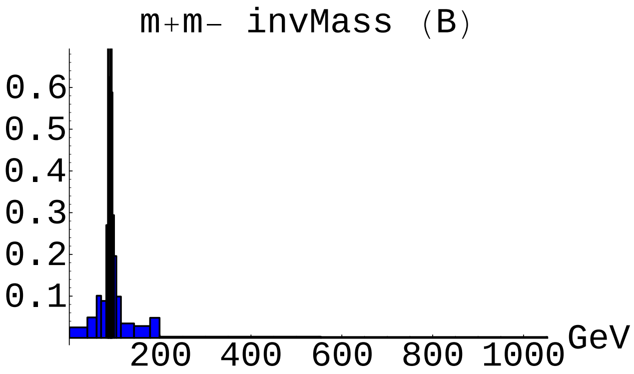
elec PT (A)

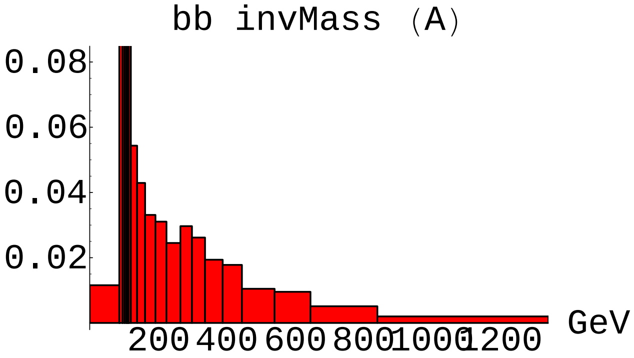


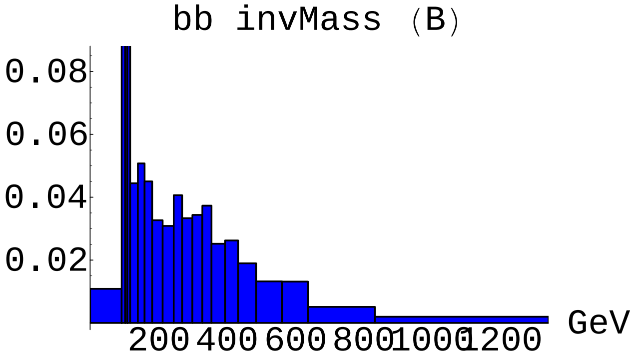
elec PT (B)

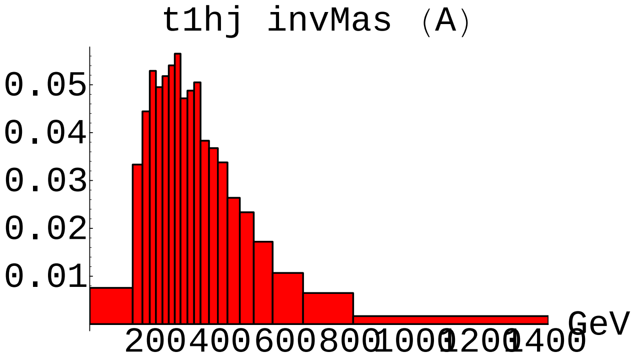


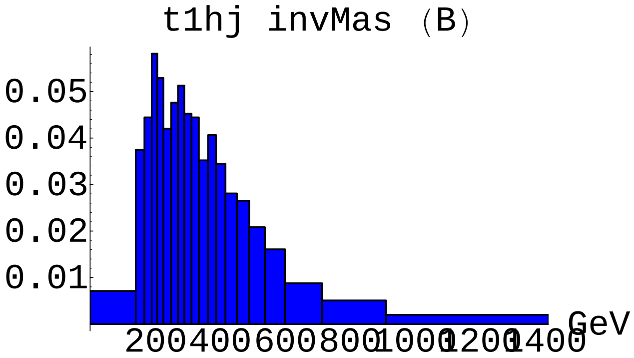

m_+m_- invMass (A)


m_+m_- invMass (B)


bb invMass (A)


bb invMass (B)


t1hj invMas (A)


t1hj invMas (B)

Understanding Degeneracies

Parameter vs. Signature Distance

Pairwise comparison.

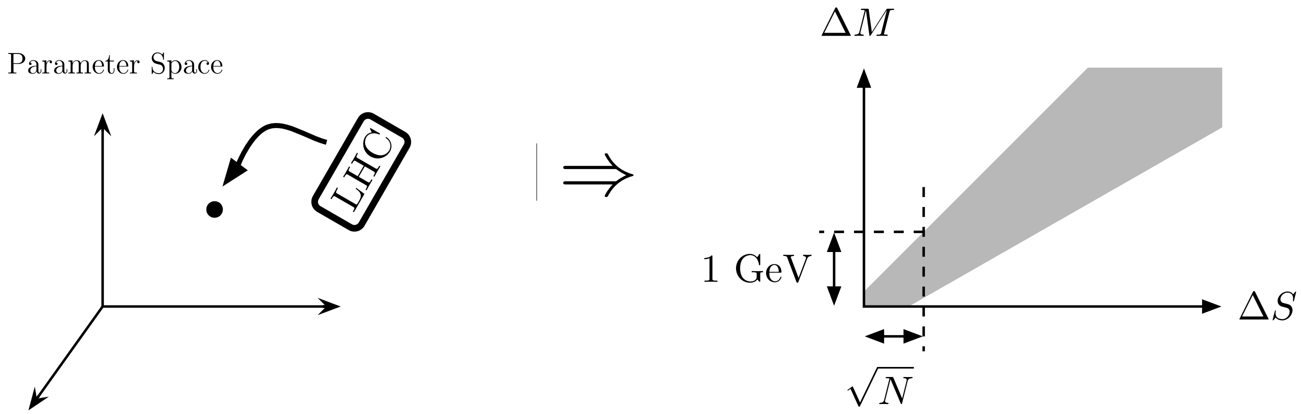
$$\Delta S = \tilde{\chi}^2_{\text{sig}} = \frac{1}{n_{\text{sig}}} \sum_i \left(\frac{\Delta s_i}{\sigma_i} \right)^2$$

$$\Delta M = \text{“}\chi^2_{\text{para}}\text{”} = \sum_j \Delta m_j$$

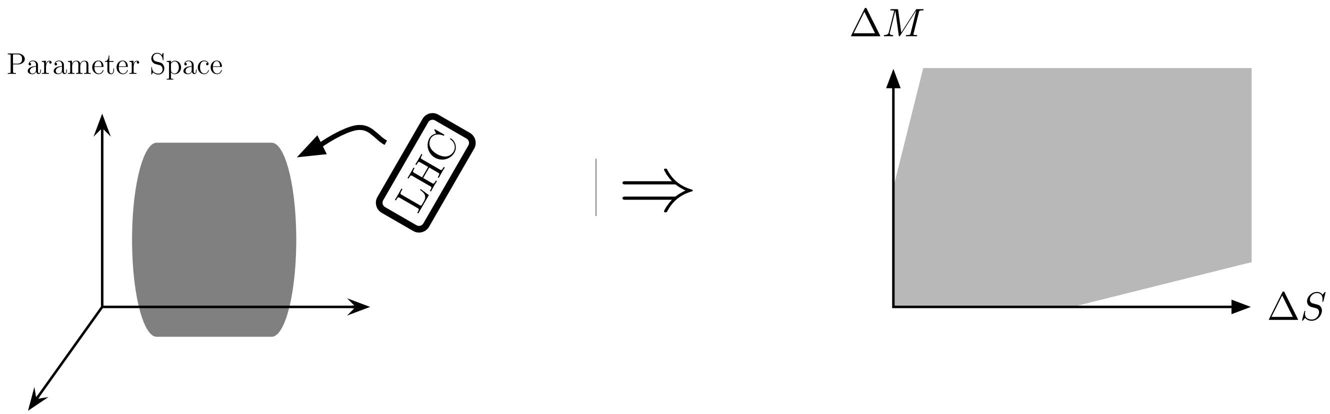
Ignore a signature if for either model, signature is consistent with zero.

Same model over and over again: $\tilde{\chi}^2_{\text{sig}} < 0.3$.

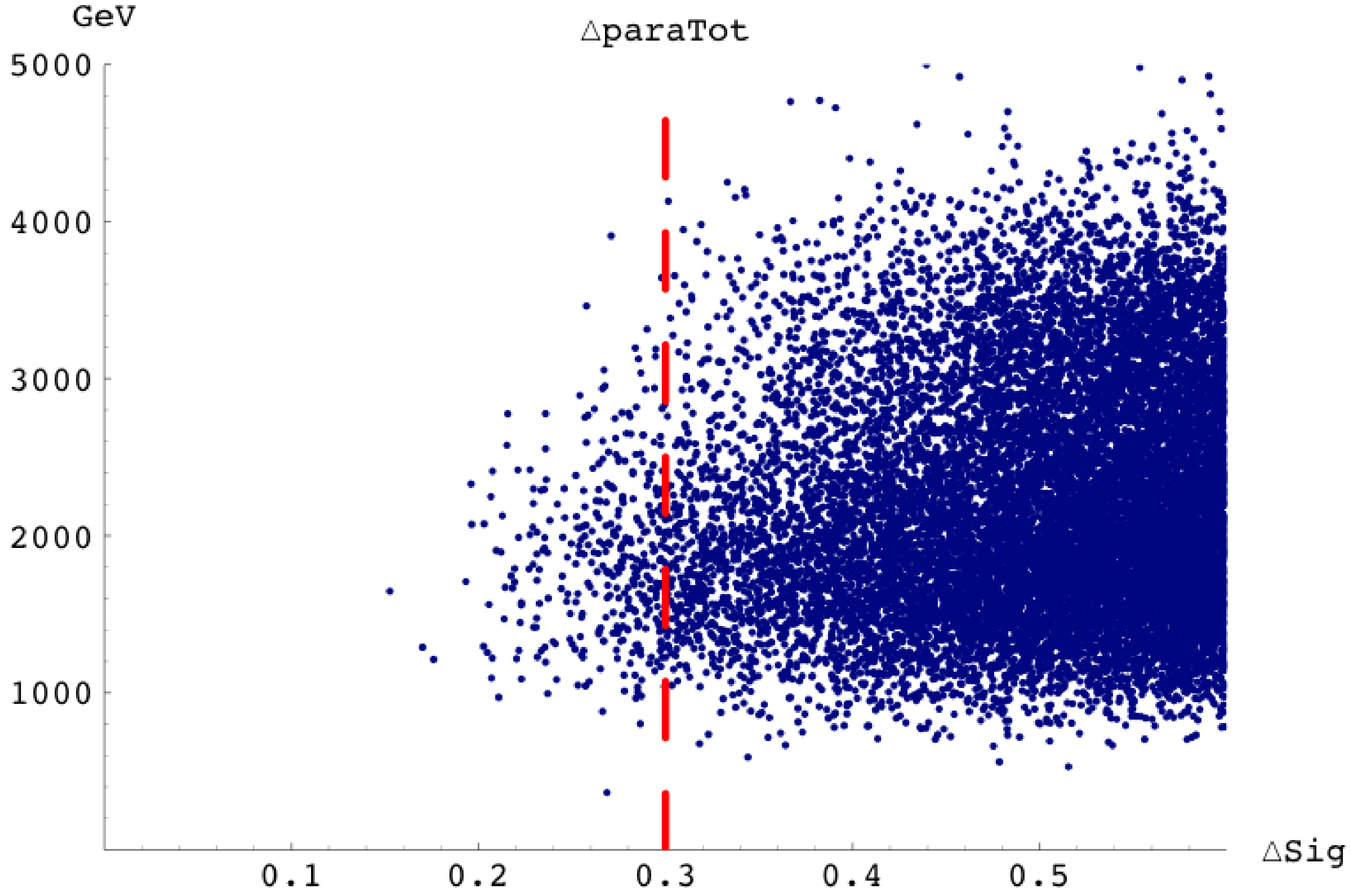
Best of all Possible Worlds



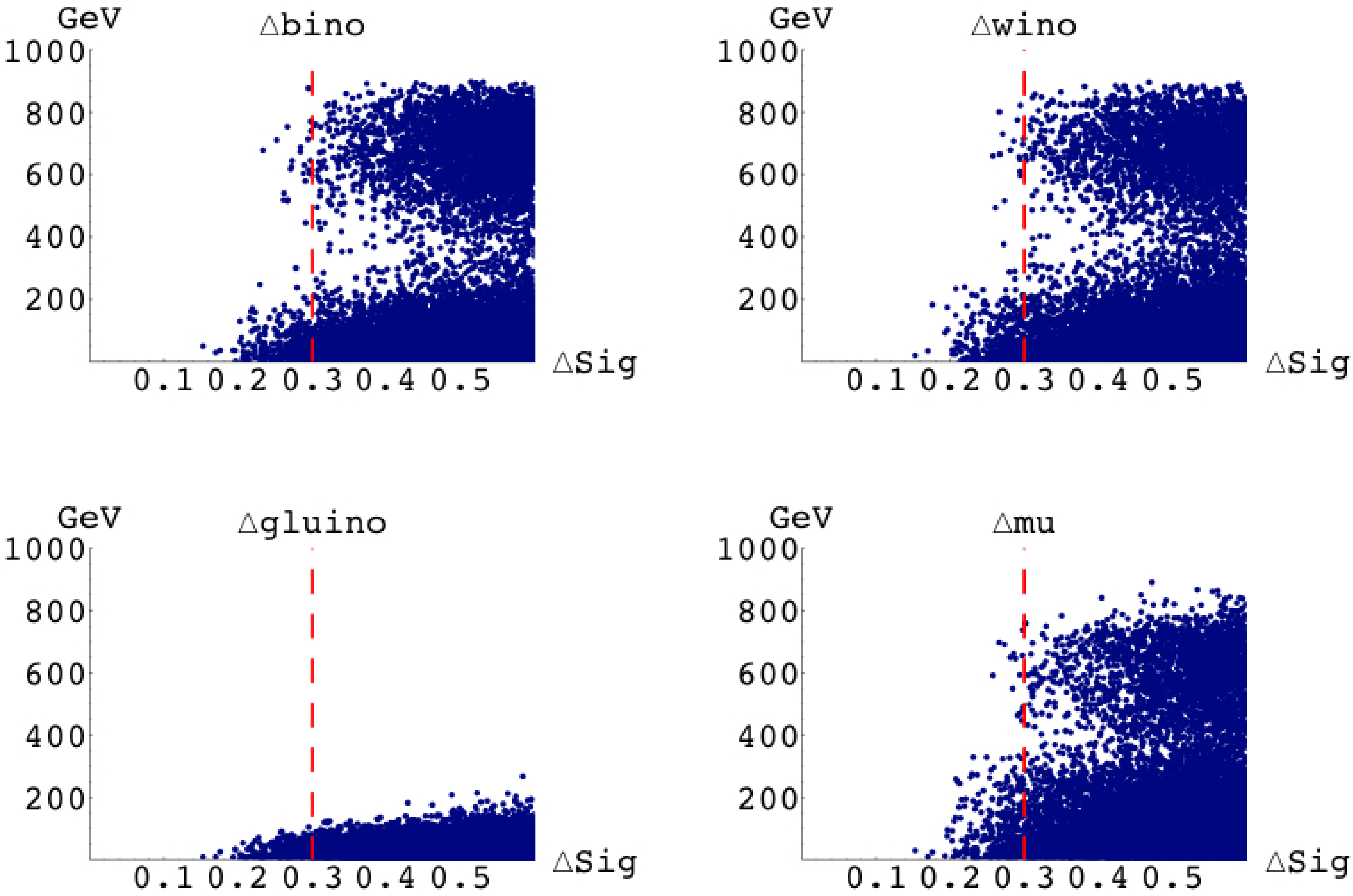
Worst of all Possible Worlds



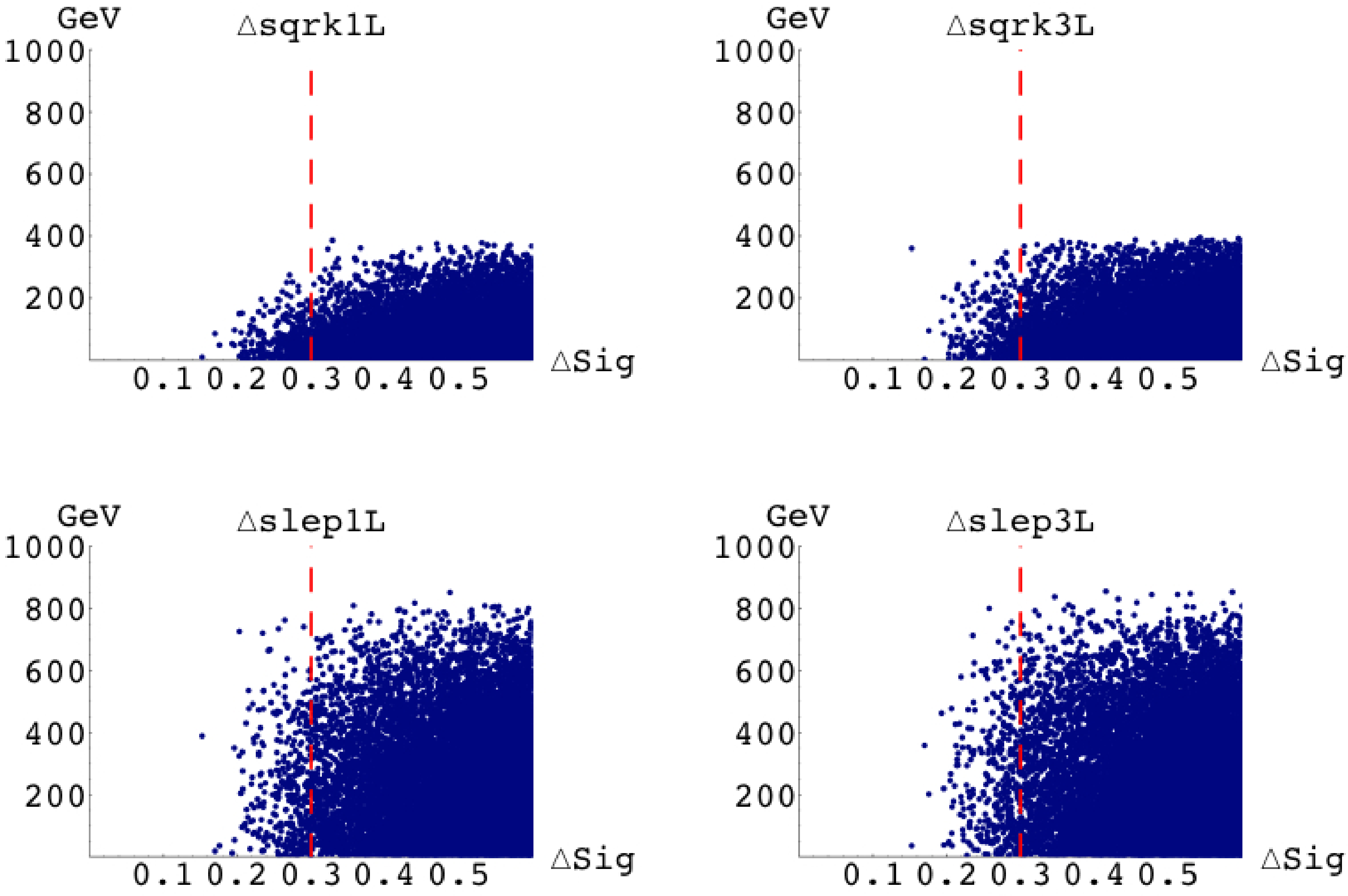
Total Parameter Distance



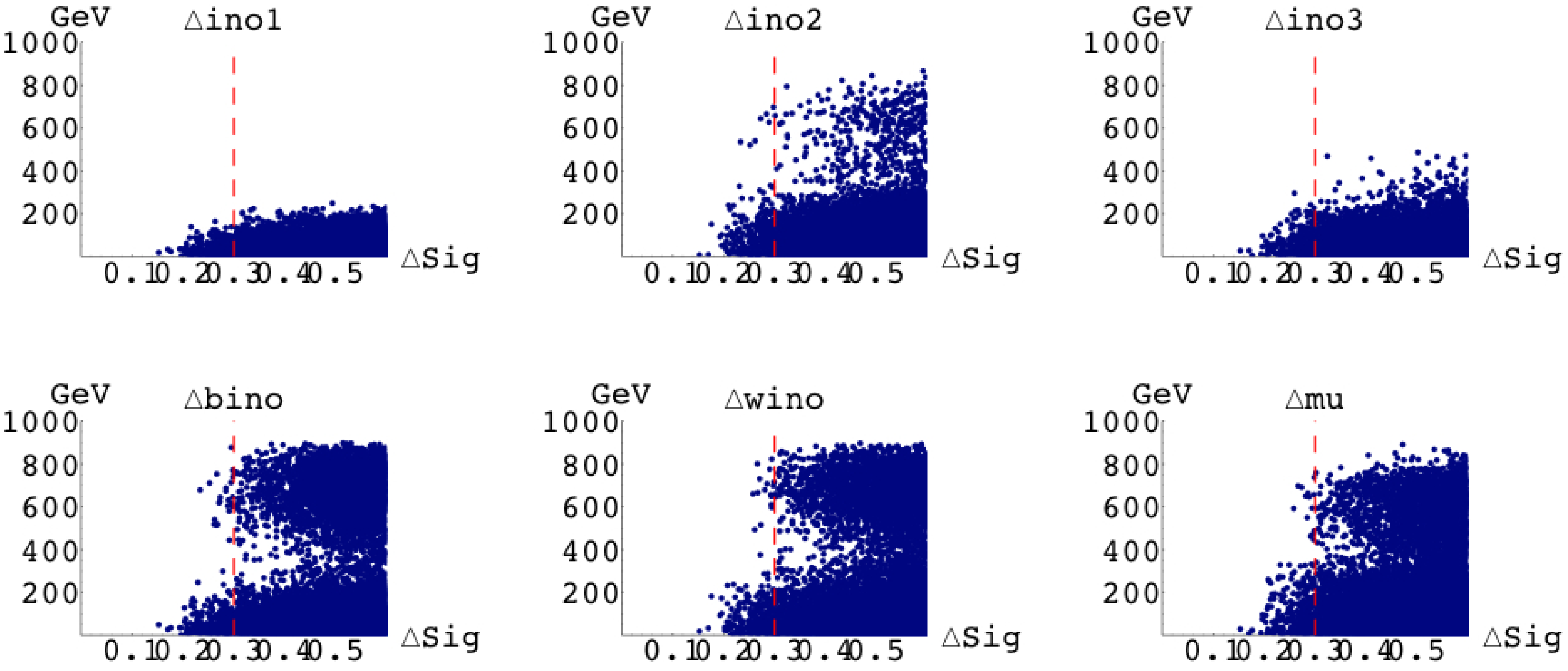
Inos


GeV
Δbino
1000
800
600
400
200
0.1 0.2 0.3 0.4 0.5
ΔSig
GeV
Δwino
1000
800
600
400
200
0.1 0.2 0.3 0.4 0.5
ΔSig
GeV
Δgluino
1000
800
600
400
200
0.1 0.2 0.3 0.4 0.5
ΔSig
GeV
Δmu
1000
800
600
400
200
0.1 0.2 0.3 0.4 0.5
ΔSig

Sfermions


GeV
Δsqrk1L
1000
800
600
400
200
0.1 0.2 0.3 0.4 0.5
ΔSig
GeV
Δsqrk3L
1000
800
600
400
200
0.1 0.2 0.3 0.4 0.5
ΔSig
GeV
Δslep1L
1000
800
600
400
200
0.1 0.2 0.3 0.4 0.5
ΔSig
GeV
Δslep3L
1000
800
600
400
200
0.1 0.2 0.3 0.4 0.5
ΔSig

Eigenvalues Not Eigenvectors



Evidence for discrete choices for ino degeneracies, not large flat directions. (Similar story for squarks, but not sleptons.)

Expected Number of Degeneracies

We can calculate $\langle d \rangle$ in a simple way.

$$\langle d \rangle = \frac{\text{number of matching pairs}}{\text{number of ``good'' pairs}}$$

The way we choose “good” effectively sets N_{Models}.

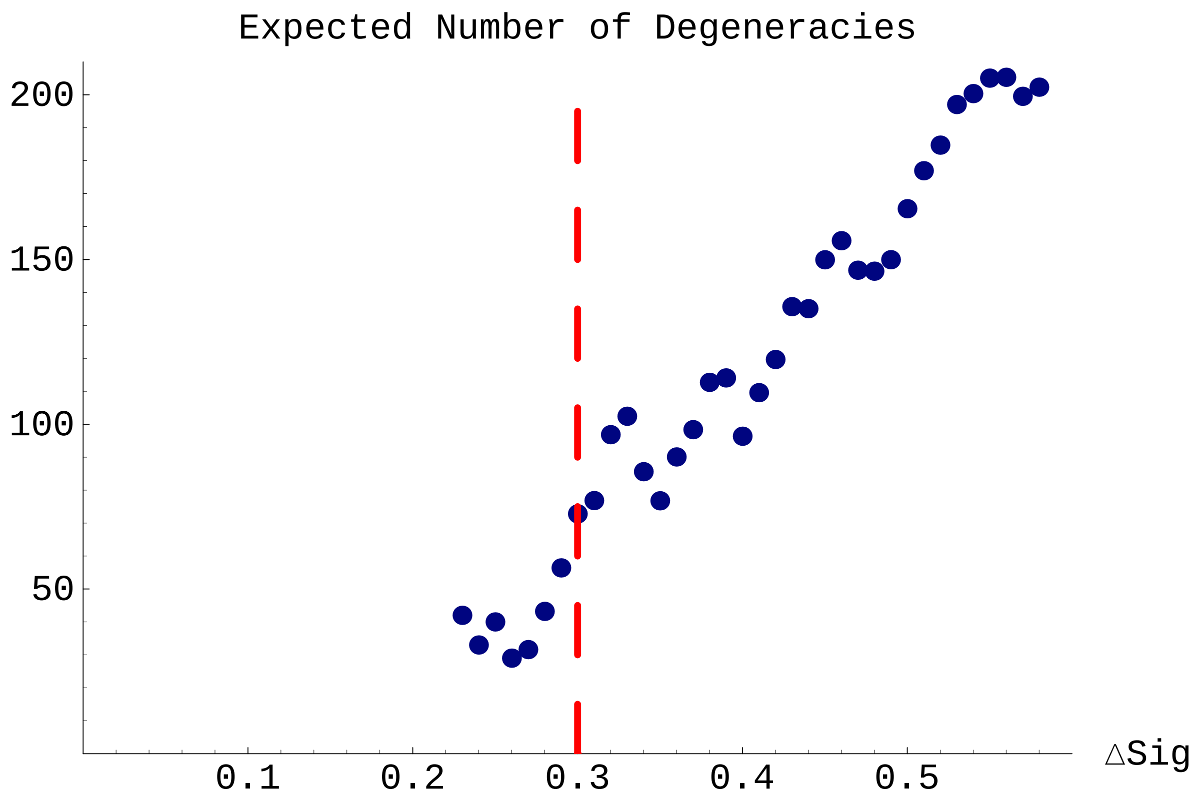
I want to know the $\tilde{B}$, $\tilde{W}$, μ and $\tilde{g}$ masses to 10%.

I want to know the lightest $\tilde{\ell}$ and $\tilde{q}$ masses to 10%.

I want to know the mean $\tilde{\ell}$ and $\tilde{q}$ masses to 10%.

I can live without knowing $\tan\beta$.

Around 100 Degeneracies



This justifies our ball/bin counting from before.

Confronting Degeneracies

Types of Degeneracies

We have 39137 models, many degenerate pairs.

Open Question:

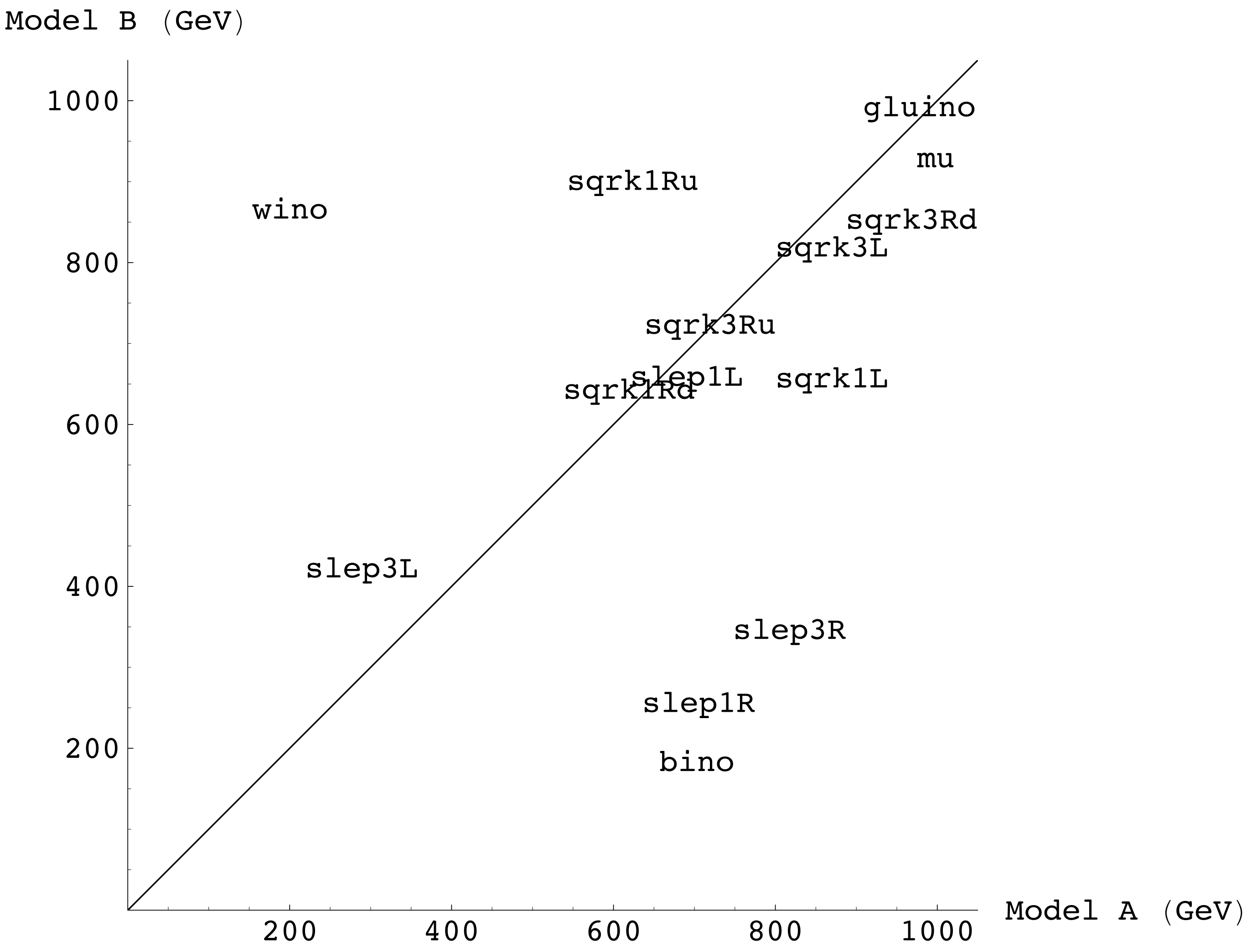
Fundamentally Indistinguishable @ LHC?

New Clever Observables?

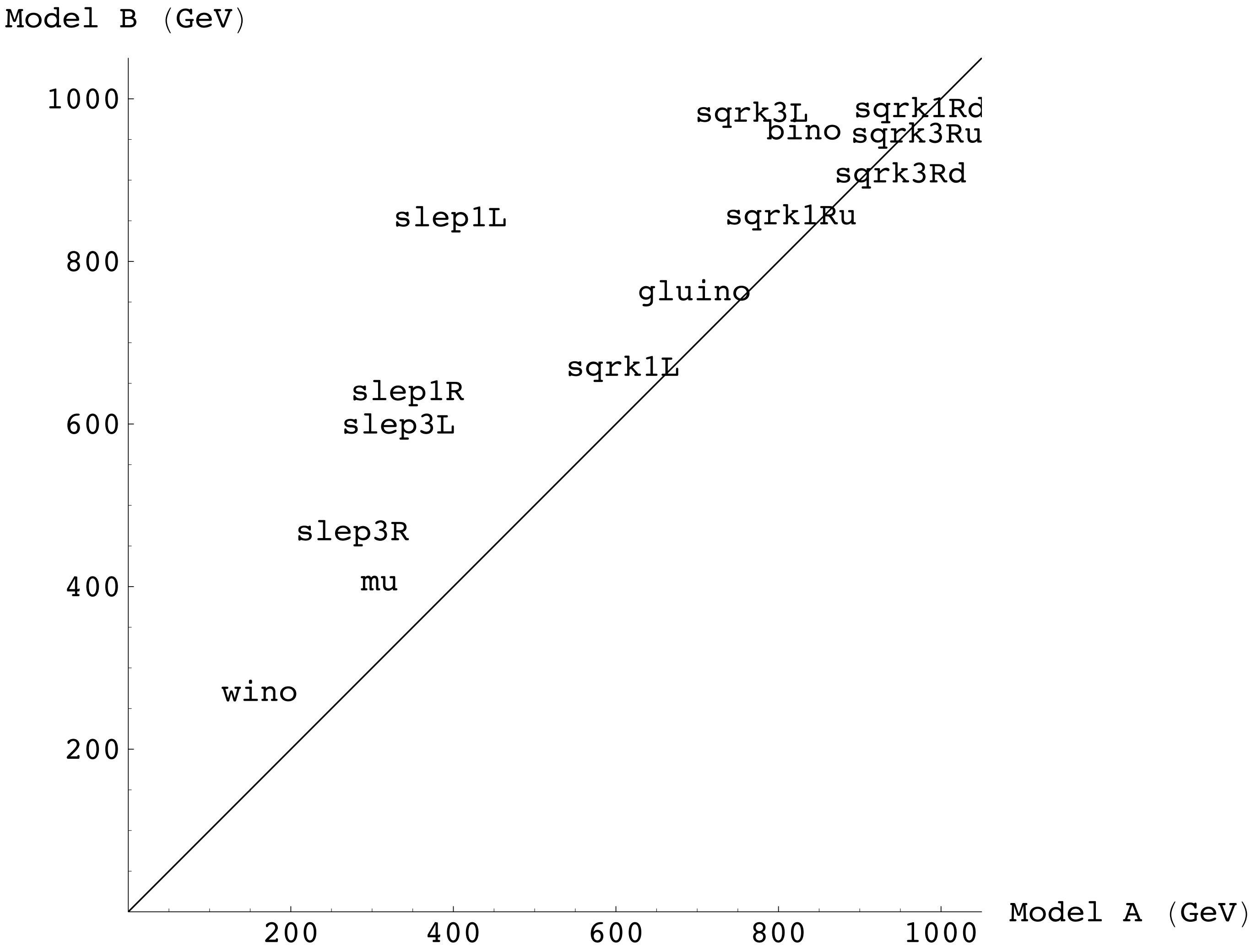
First Step: Catalog...

Electroweakino Sector: Flippers, Sliders, and Squeezers

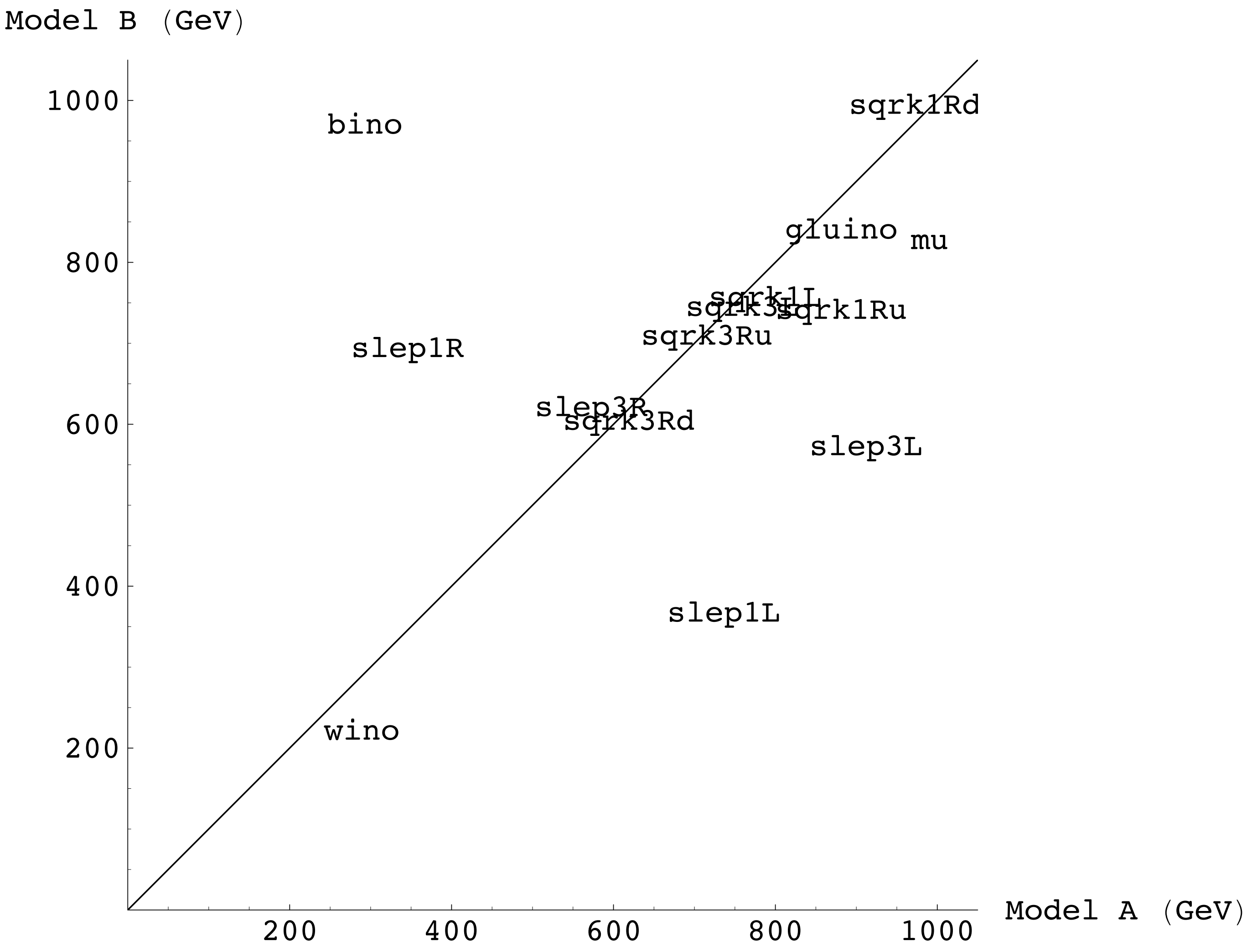
Flippers: 17653 vs. 20026



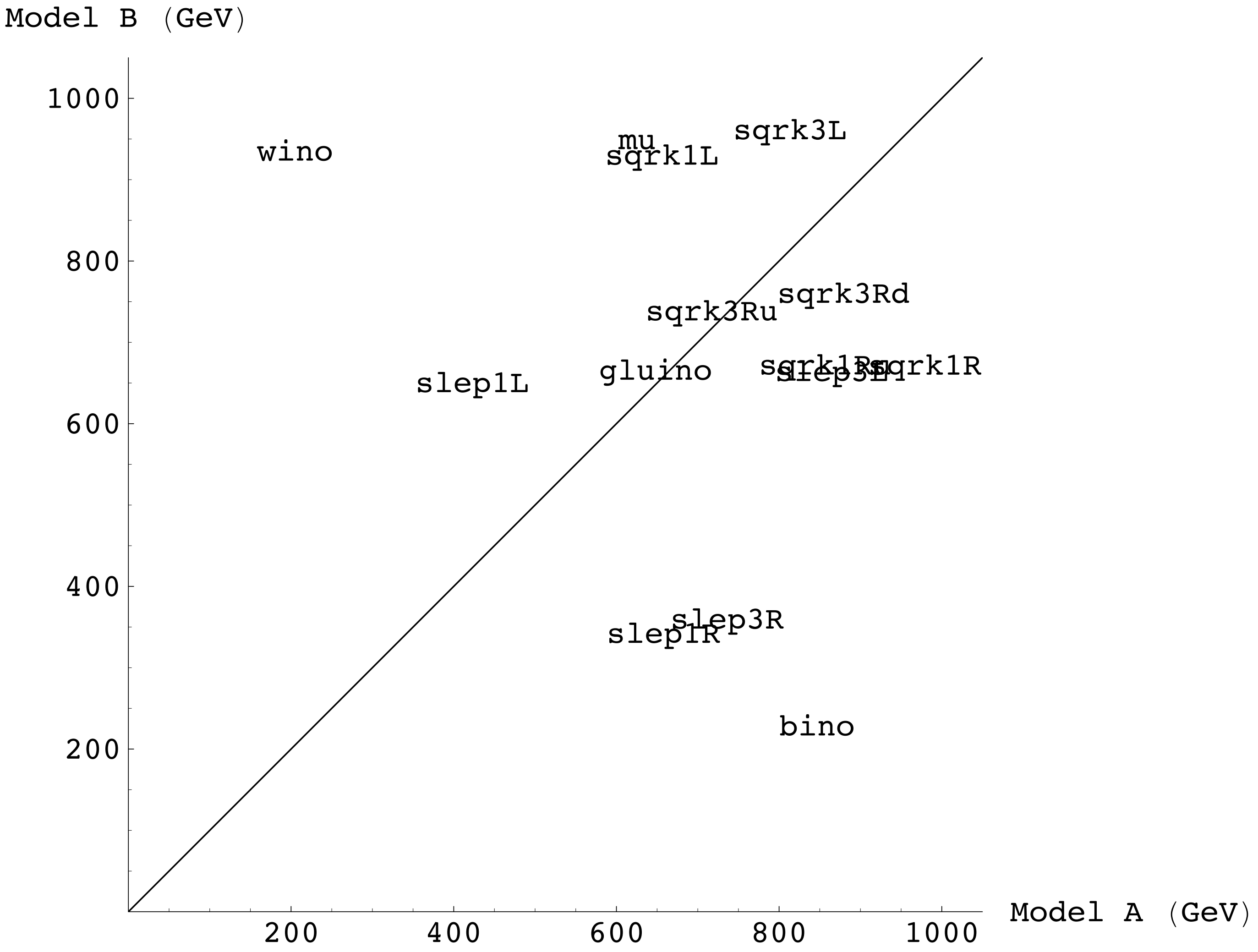
Sliders: 24230 vs. 25628



Squeezers: 24119 vs. 35520

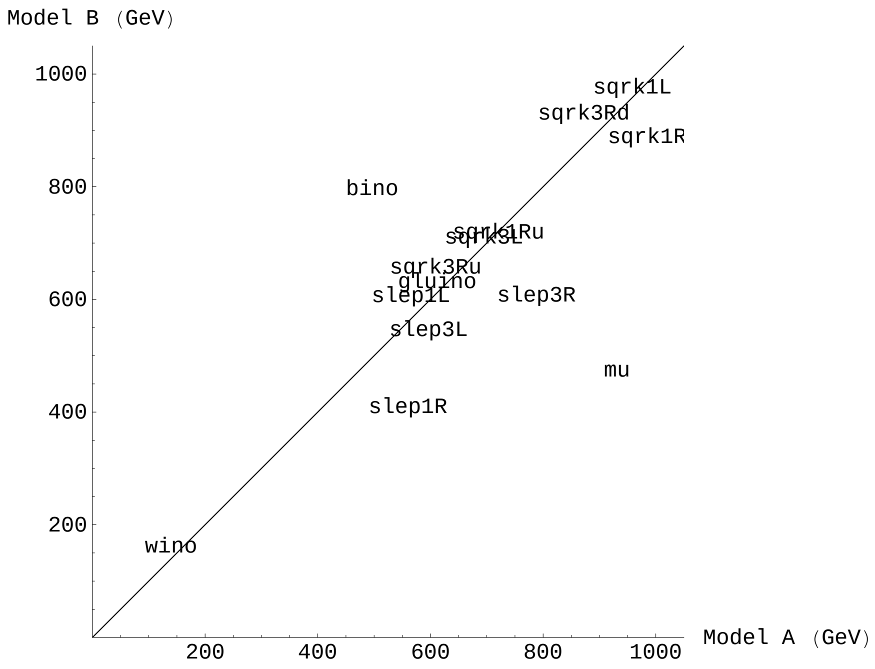


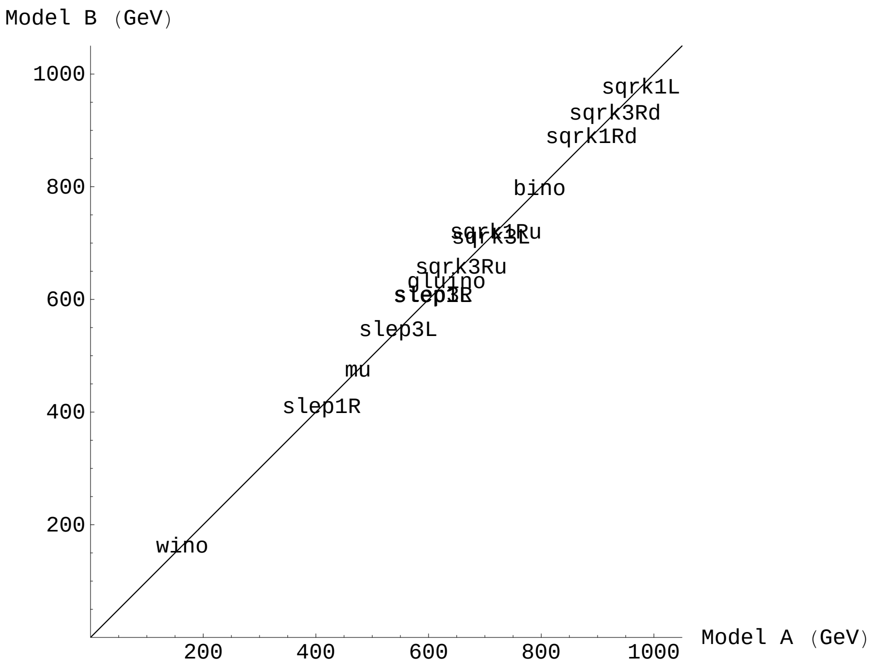
And So On...: 17544 vs. 27775



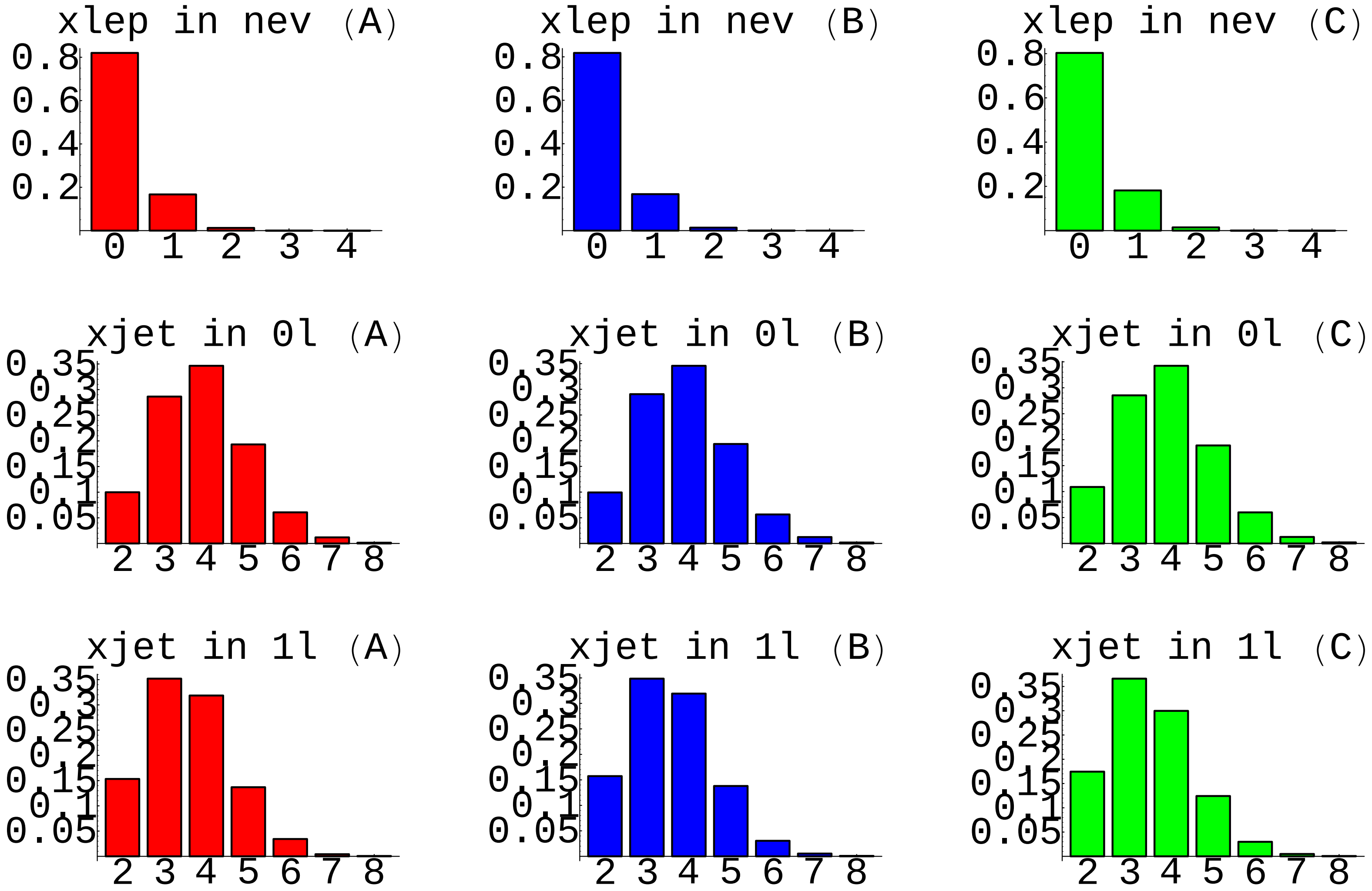
“Are There Really Degeneracies?”

Case Study: MSSM 26 vs. 6226 vs. 26

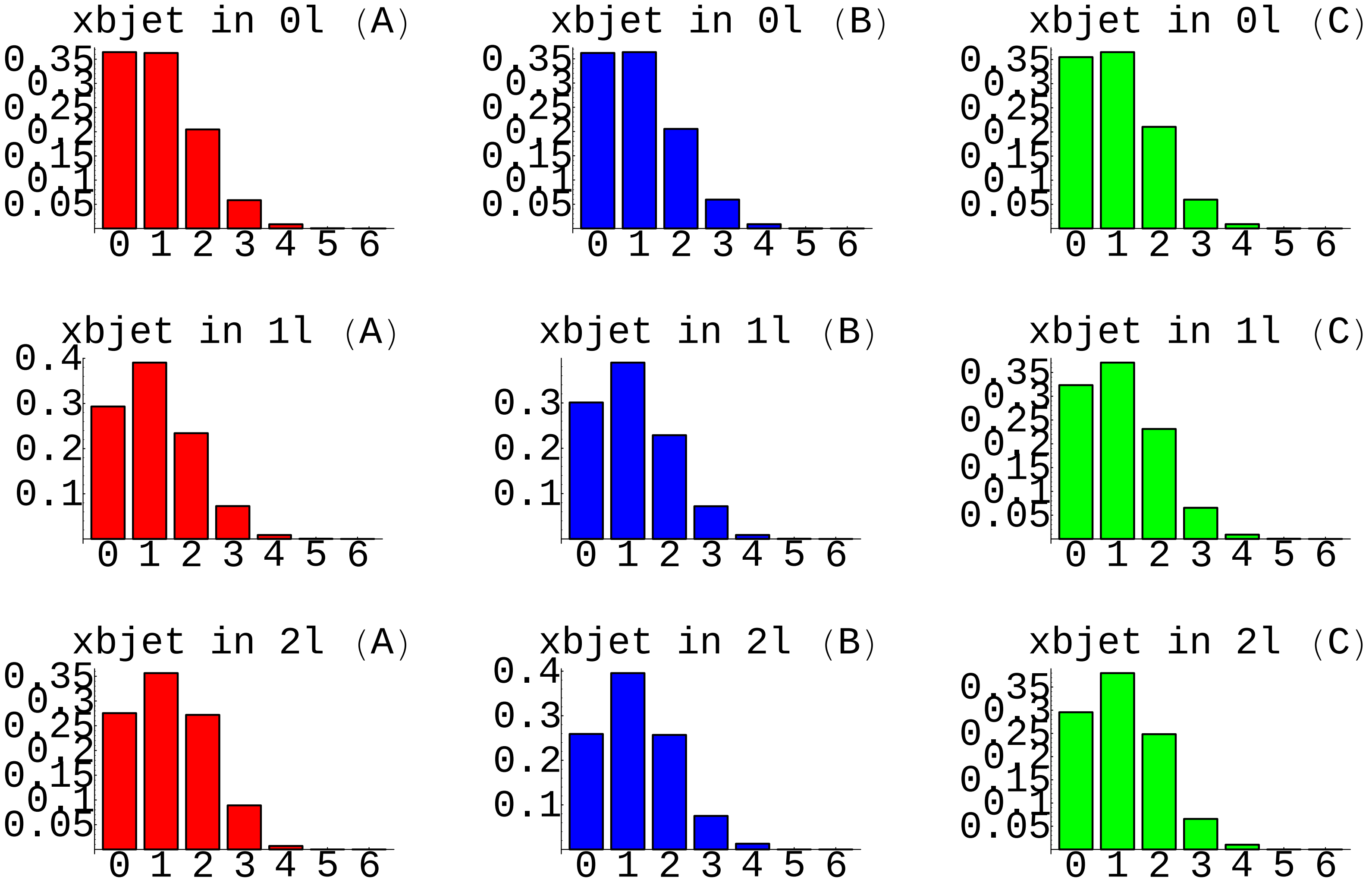




$\mathbf{A = B}$ or $\mathbf{B = C}$?

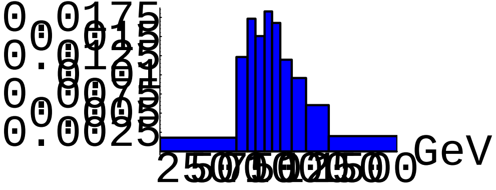


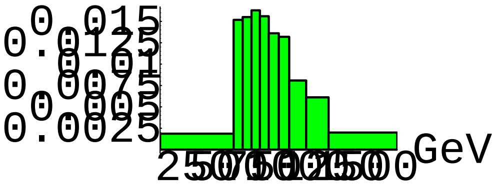
$\tilde{W} < \tilde{B} < \mu$ **or** $\tilde{W} < \mu < \tilde{B}$ **?**

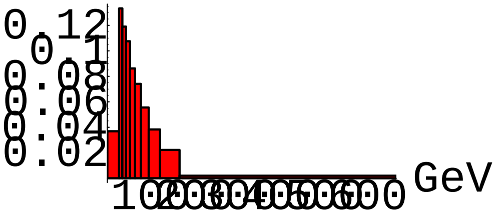


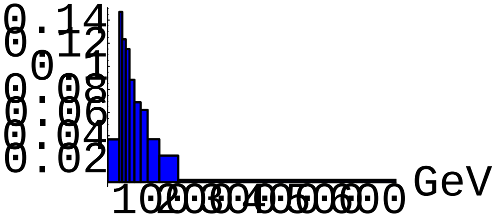
Wait Until ILC?

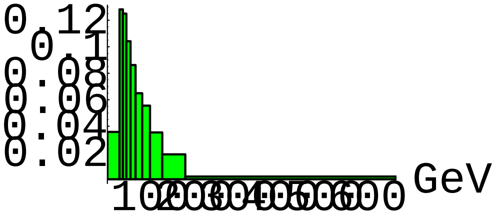

Meff 2l (A)

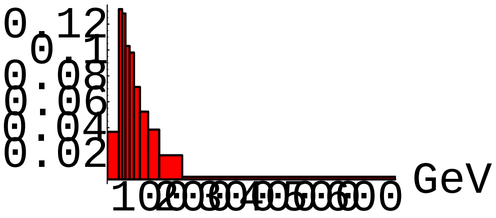

Meff 2l (B)

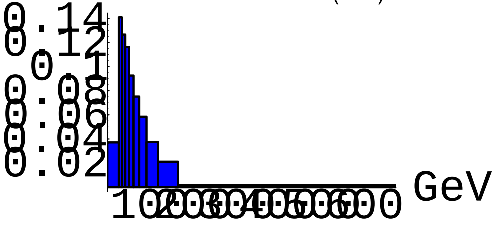

Meff 2l (C)

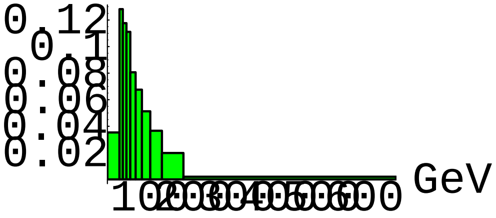

elec PT (A)


elec PT (B)

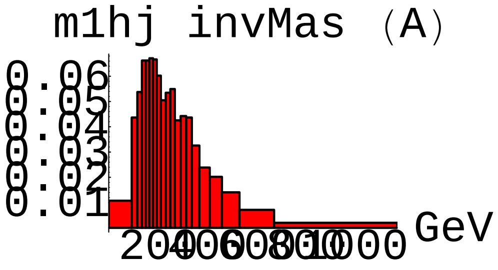

elec PT (C)

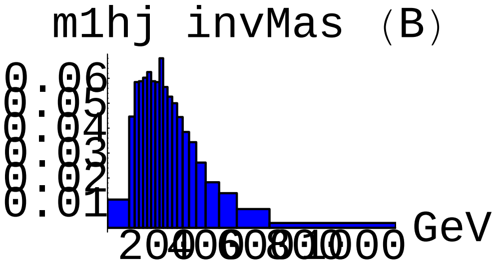

muon PT (A)

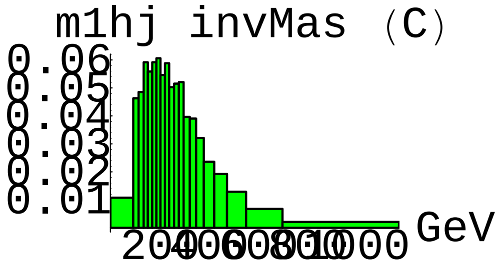

muon PT (B)

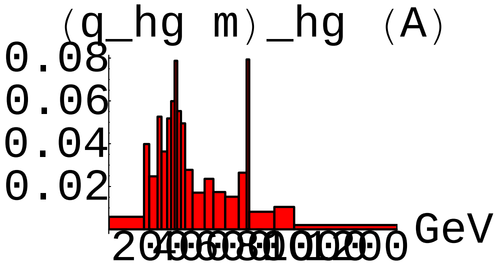

muon PT (C)

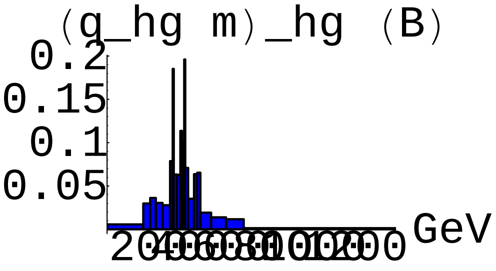
A = B or B = C (or A = C) ?

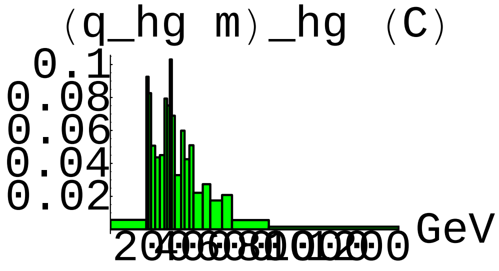

m1hj invMas (A)

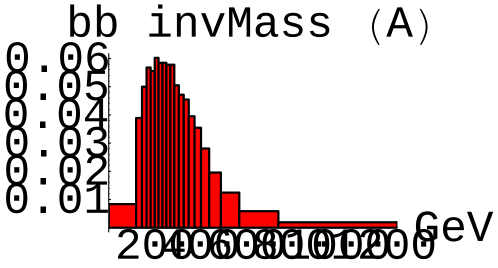

m1hj invMas (B)

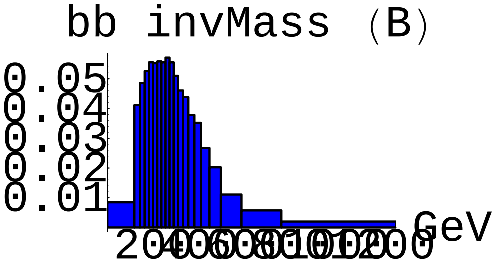

m1hj invMas (C)

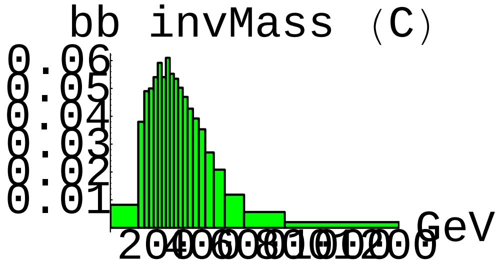

(q_hg m)_hg (A)


(q_hg m)_hg (B)

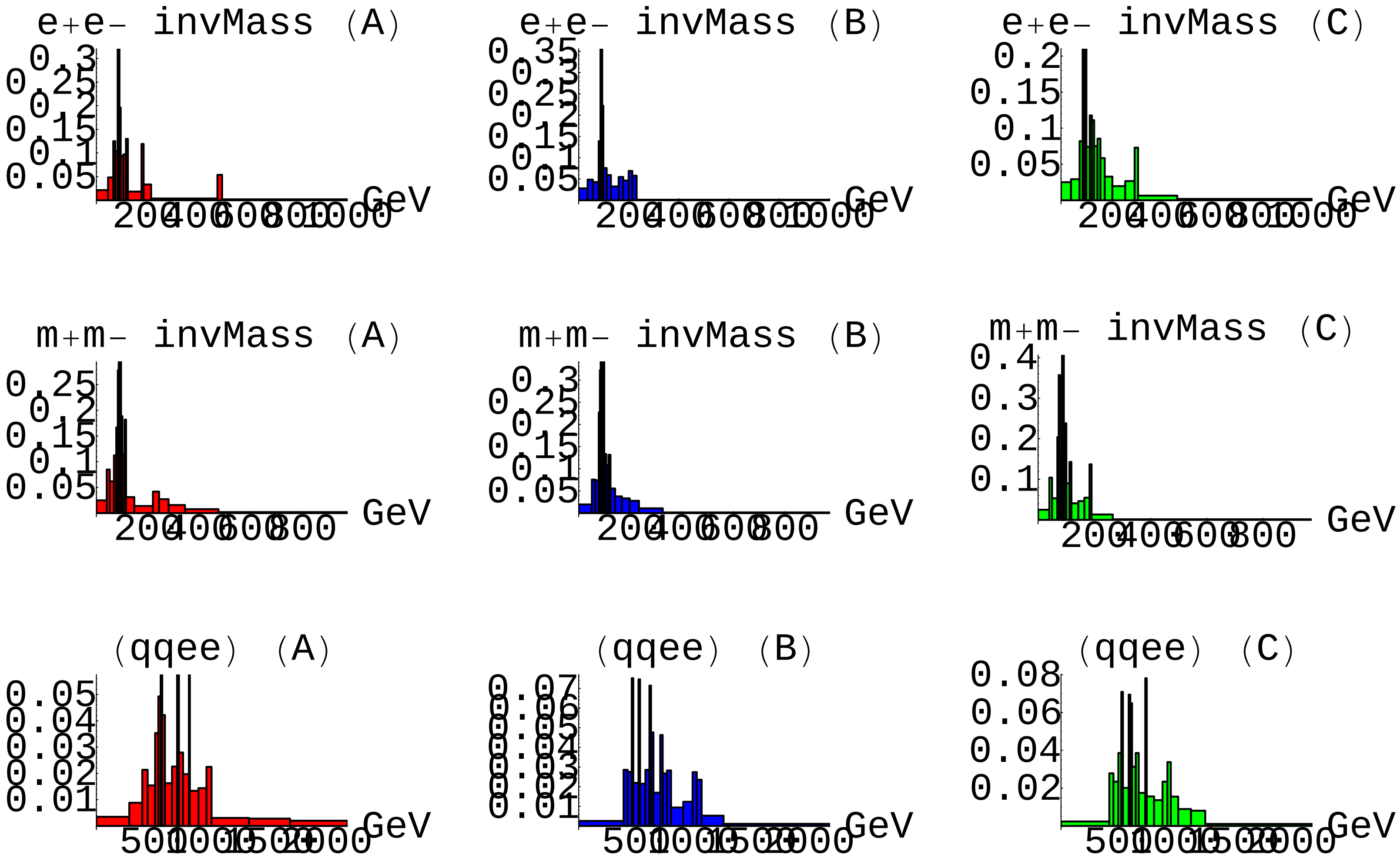

(q_hg m)_hg (C)


bb invMass (A)


bb invMass (B)

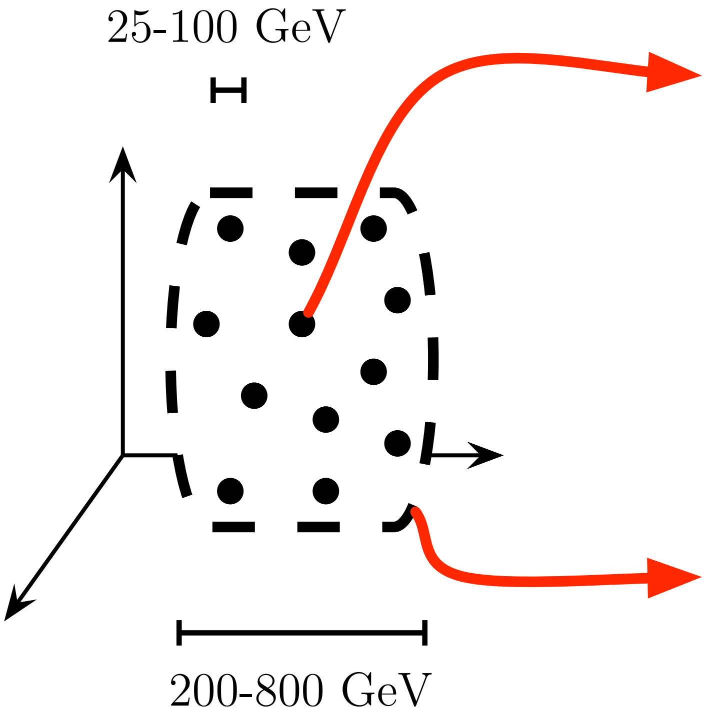

bb invMass (C)

The Answer is A = B (Computer: B = C)



The Outlook

If You Have a Favorite Theory...



If LHC data is consistent with mSUGRA, you will measure mSUGRA parameters very accurately...

...but there at least $\mathcal{O}(100)$ other (well-motivated?) MSSMs consistent with the same experimental data.

200 GeV can change gaugino unification, LSP identity, etc.

Reconciling Two Different Attitudes

SUSY at LHC is easy!

If a constrained model is consistent with LHC data, then you can make reasonably accurate measurements at 10 fb^{-1} (assuming no SM background).

SUSY at LHC is hard!

There are $\mathcal{O}(100)$ different models that also match the data, and there is no systematic way to find them. Do they come from nice UV theories?

Lessons from the Degenerate MSSM

To nail the MSSM at the LHC, we must drastically increase the number of independent LHC signatures.

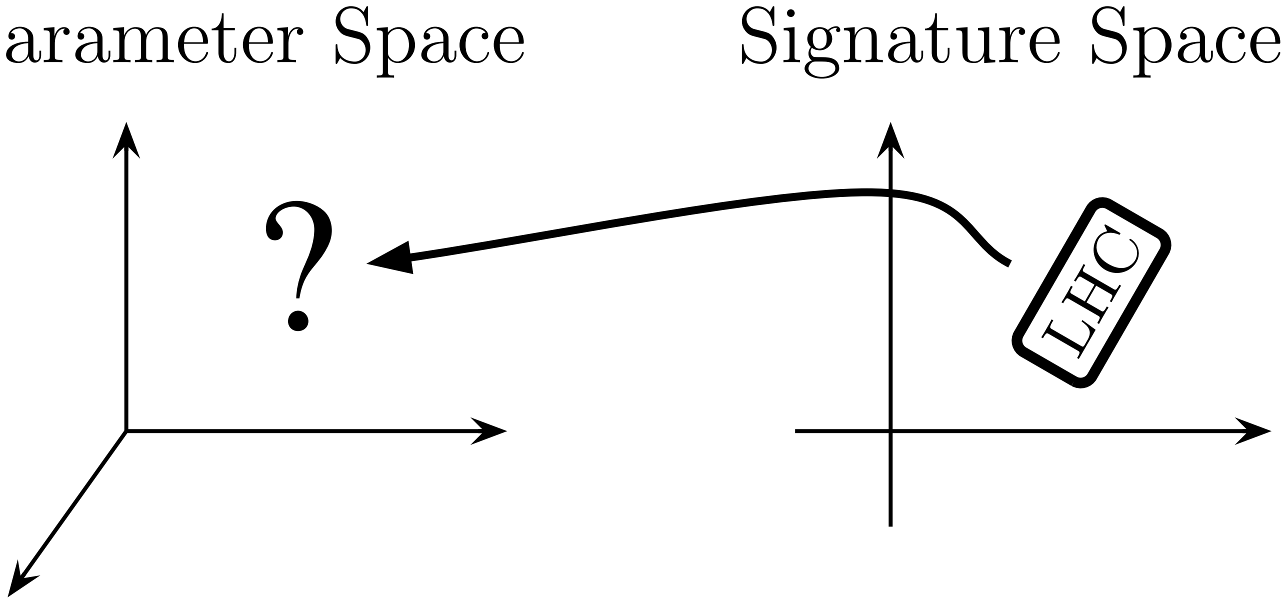
(Is this still true for leptophilic models? At 300 fb^{-1}? With SM background?)

Easy to check if a new signature works: calculate $\langle d \rangle$.

Statistical techniques useful for diagnosis, but once we have data, how will we actually "rule in" models?

The LHC Inverse Problem

Very little work in this direction. This question is perfect for particle physicists trained as model builders.



Parameter Space

Signature Space

MSSM vs. NMSSM vs. UED vs. T-parity LH ?

An Invitation to the LHC Olympics

Competition as motivation for interpreting data.

The First LHC Olympics: CERN, July 2005

University of Michigan Blackbox
(Partially Analyzed by University of Washington)

The Second LHC Olympics: CERN, February 9-10, 2006

Three New Blackboxes (Harvard, UMich, UWash)

Our tools: experience, intuition, consensus, and luck.