MARMOSET:

Signal-Based Monte Carlo for the LHC

Jesse Thaler (Berkeley)

with Philip Schuster, Natalia Toro, Lian-Tao Wang,
Johan Alwall, Matthew Baumgart, Liam Fitzpatrick, Tom Hartman, Jared Kaplan,
Nima Arkani-Hamed, Bruce Knuteson, Steve Mrenna.

MARMOSET:

Mass And Rate Modeling for On-Shell Effective Theories

en.wikipedia.com

www.themanwhofellasleep.com

MARMOSET:

Mass And Rate Modeling for On-Shell Effective Theories

A Monte Carlo Tool

Approximate MC generation
for (almost) any model. (OSET)

An Analysis Strategy

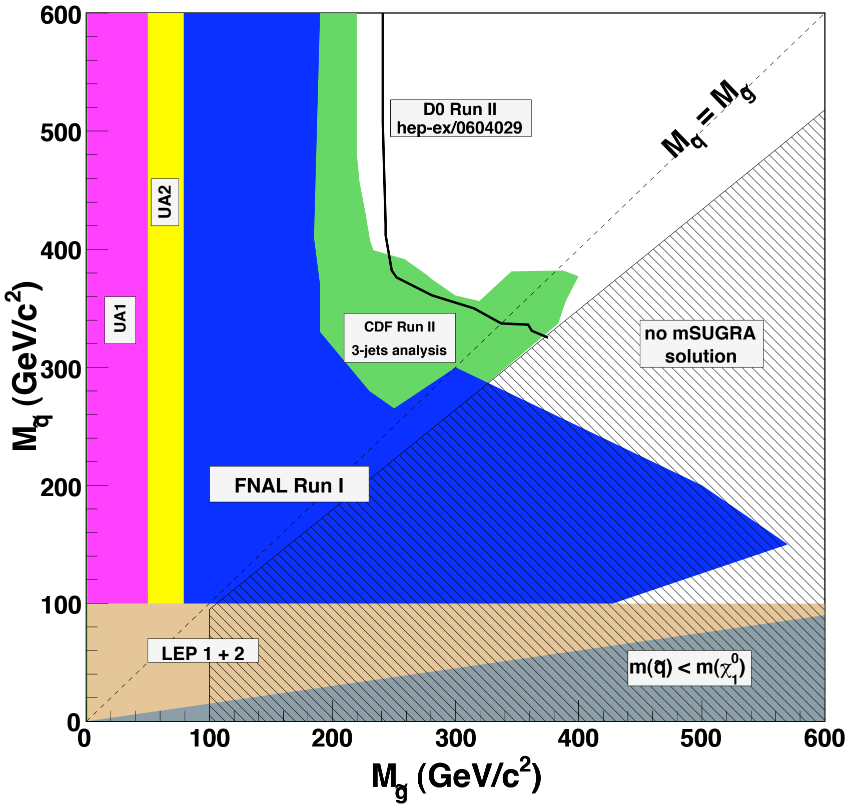
Inclusive observables for
mass/rate matching. (MARM)

www.themanwhofellasleep.com

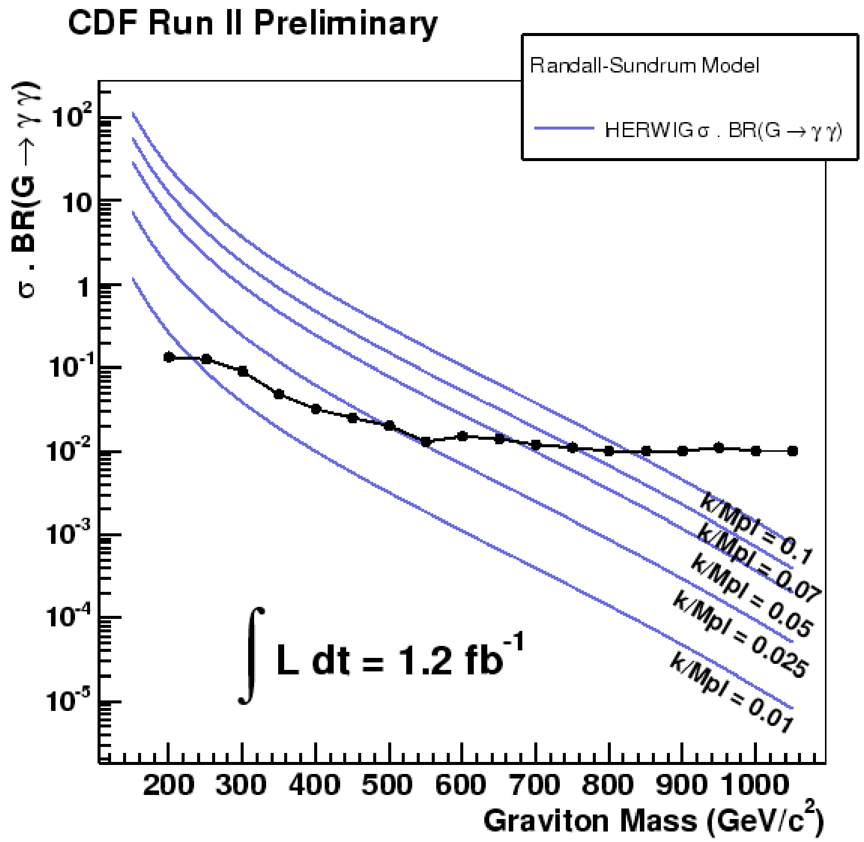
("Effective" in the "it works!" sense, not always in the Wilsonian sense.)

A Monte Carlo Tool:

Can you generate MC for an unknown model?



vs.



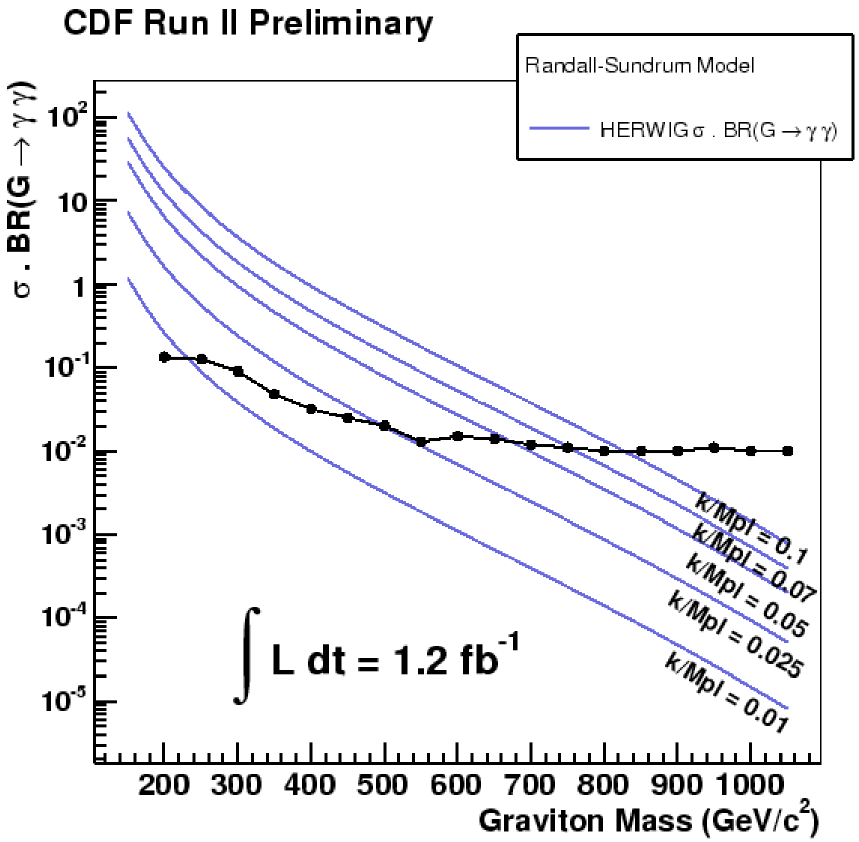
A Monte Carlo Tool:

Can you generate MC for an unknown model?

MARMOSET:

Meaningful exclusion plots for non-resonant production and complicated (e.g. SUSY-like) decay topologies.

Model-agnostic language for characterizing new physics.



An Analysis Strategy:

How should we characterize LHC excesses?

Excess of

$4b\ 4\ell\ \not{E}_T$

SUSY with $\tilde{g}\tilde{g} \longrightarrow t\bar{t}t\bar{t}\tilde{N}\tilde{N}$

(Through off-shell stop.)

OSET with $C_8 C_8 \longrightarrow t\bar{t}t\bar{t}C_0 C_0$

(Through three-body decay.)

Easier (necessary?) to ascertain Topology and then address Spin (especially with BTSM sources of missing energy).

Do we need to assume a stop to make a discovery?

An Analysis Strategy:

How should we characterize LHC excesses?

Excess of $4b\ 4\ell\ \not{E}_T$ $\longrightarrow$ OSET with $C_8 C_8 \to t\bar{t}t\bar{t}C_0 C_0$

(Through three-body decay.)

Wilson!

MARMOSET:

Reports results in terms of

$$\underbrace{\sigma \qquad \mathrm{Br}}_{\text{“Cheap”}} \qquad \underbrace{m}_{\text{“Expensive”}}$$

Strongly suggests global (inclusive) approach to signal analysis.

Outline

- The Physics Behind MARMOSET
  Approximate Monte Carlo Using (Only)
  Narrow Width / Phase Space Matrix Elements
- MARMOSET as a Monte Carlo Tool
  Trilepton Possibilities at the TeVatron
- MARMOSET as an Analysis Strategy
  Example Use of MARMOSET in LHC Olympics

MC: σ Br m $\cancel{|\mathcal{M}|^2}$

The Physics Behind MARMOSET

Approximate Monte Carlo Using (Only)
Narrow Width / Phase Space Matrix Elements

What Do Models Actually Look Like?

New Particles		In ATLAS or CMS
(Meta-)Stable (Neutral)	→	Missing Energy
(Meta-)Stable (Charged/Colored)	→	Cool Tracks/Out of Time Signals
Unstable	→	SM Particles + (Meta-)Stables

Assuming Dedicated Searches for (Meta-)Stable Charged/Colored Particles (and Black Holes)...

(and assuming the new physics has a description in term of relatively narrow resonances)

What Do Models Actually Look Like?

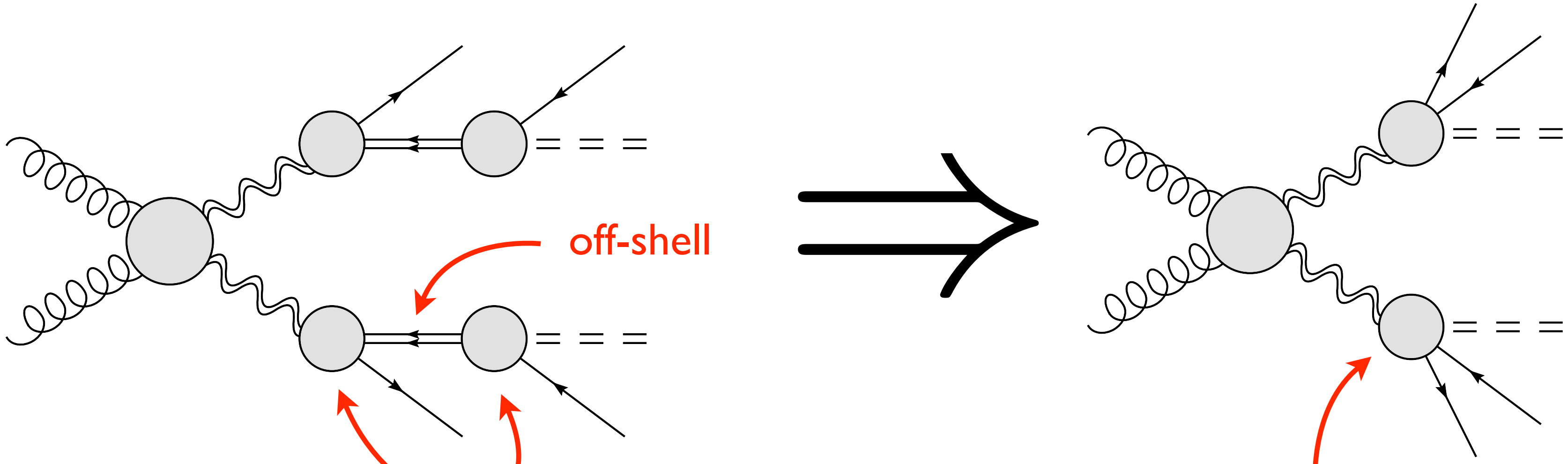
New Particles		In ATLAS or CMS
(Meta-)Stable (Neutral)	→	Missing Energy
Unstable	→	SM Particles + (Meta-)Stables

$$pp \to n\ \text{SM particles} + m\ \text{neutral stables}$$
$$\text{with some Matrix Element}$$

The Wilsonian Approach

$$pp \to n \text{ SM particles} + m \text{ neutral stables}$$

$$\text{with some Matrix Element}$$



off-shell

marginal interactions

irrelevant interaction

Use narrow width approximation.

Integrate out off-shell particles at each decay stage.

The Effective* Approach

$$pp \to n \text{ SM particles} + m \text{ neutral stables}$$
$$\text{with some Matrix Element}$$

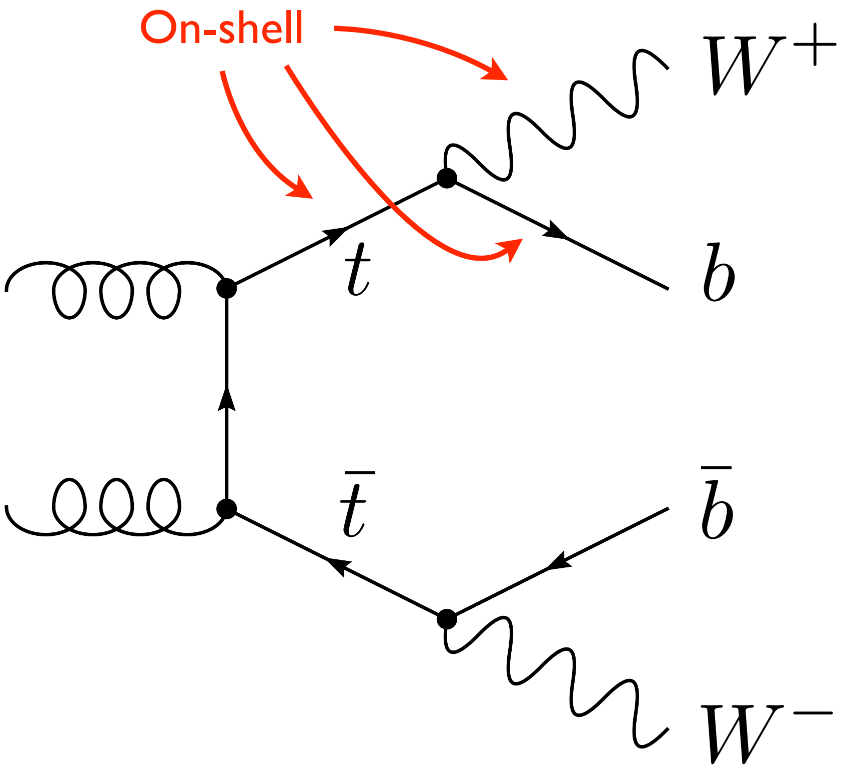
Key Point: For almost all models, Matrix Elements well-approximated by only considering Phase Space and Narrow Widths.

Dominant kinematic structures independent of Quantum Amplitudes.

Not only can we integrate out off-shell particles à la Wilson,
but we can often ignore detailed vertex structure.
Reinsert vertex structure as series expansion later...

E.g.: Top Quark

Masses, Rates, and Topology vs. Amplitudes



Dominant Top Properties:

$\sigma(gg \to t\bar{t})$

$\mathrm{Br}(t \to bW)$

m_t, m_W, m_b

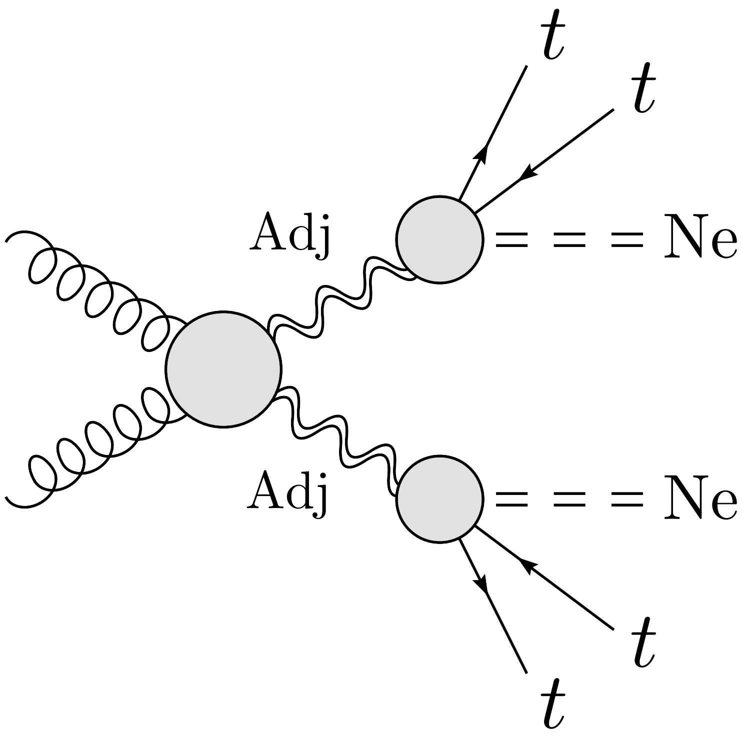
Detailed Top Properties:

$d\sigma/d\hat{t}$

W helicity

t charge

On-Shell Effective Theories



New Physics Properties:

$$m_{\mathrm{Adj}}, m_{\mathrm{Ne}}$$

$$\sigma(gg \to \mathrm{Adj\ Adj})$$

$$\mathrm{Br}(\mathrm{Adj} \to t\ t\ \mathrm{Ne})$$

Characterize New Physics In Term of
Production/Decay Topologies, Rates, and Masses

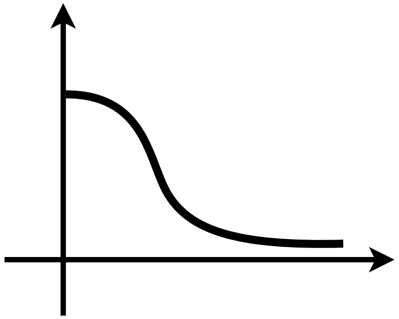
MC: σ Br m $\cancel{|\mathcal{M}|^2}$

Differential Cross Sections?

$$|\mathcal{M}|^2 = f_0(s) + f_1(s)z + f_2(s)z^2 + \ldots \qquad z = \cos\theta$$

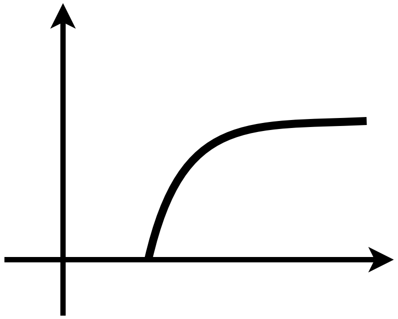
$$\frac{d\sigma}{d\hat{t}} = \int$$

Parton Luminosity



$$\times$$

Phase Space (Threshold)

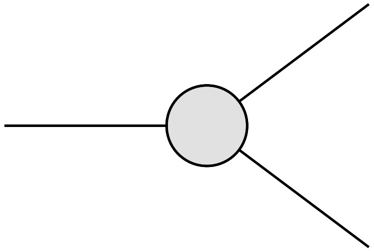


$$\times\, |\mathcal{M}|^2$$

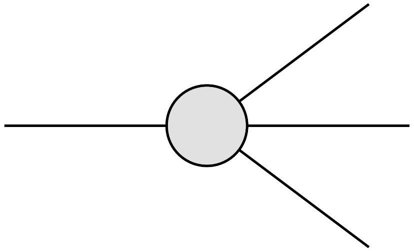
Cross Sections Dominated by Thresholds!

(Amplitude can be treated as systematic error or “measured” in Laurent expansion.)

Decay Kinematics?



Two-Body Decays:
At most, lose angular correlations with other parts of the topology. (Kinematics correct.)

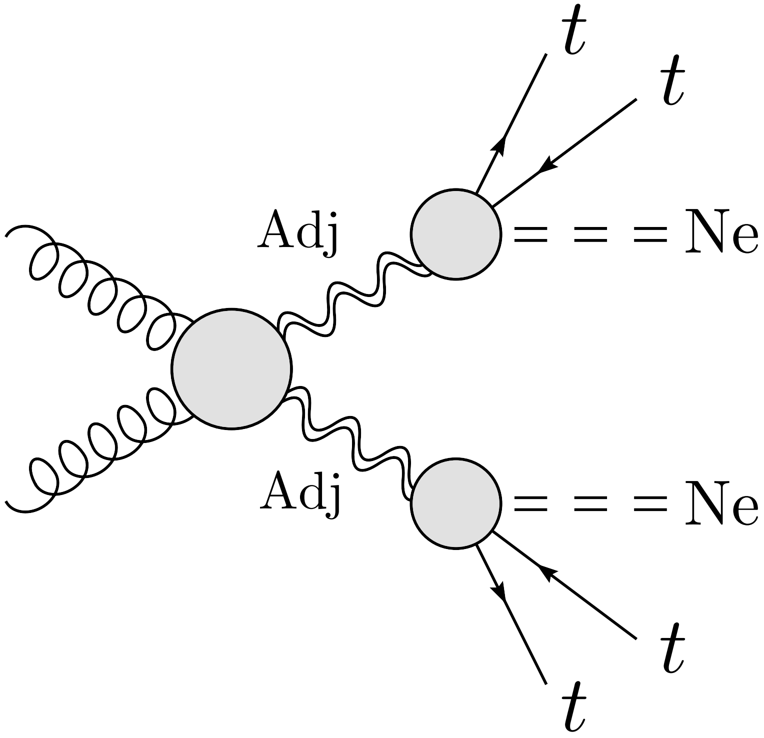


Multi-Body Decays:
Lose kinematic correlations among decay products. (Energy/momentum conserved.)

Pair-wise invariant masses have correct thresholds (i.e. edge/endpoint locations) but incorrect shapes.

(Use observable less sensitive to correlations, like single particle p_T.)

MARMOSET Input



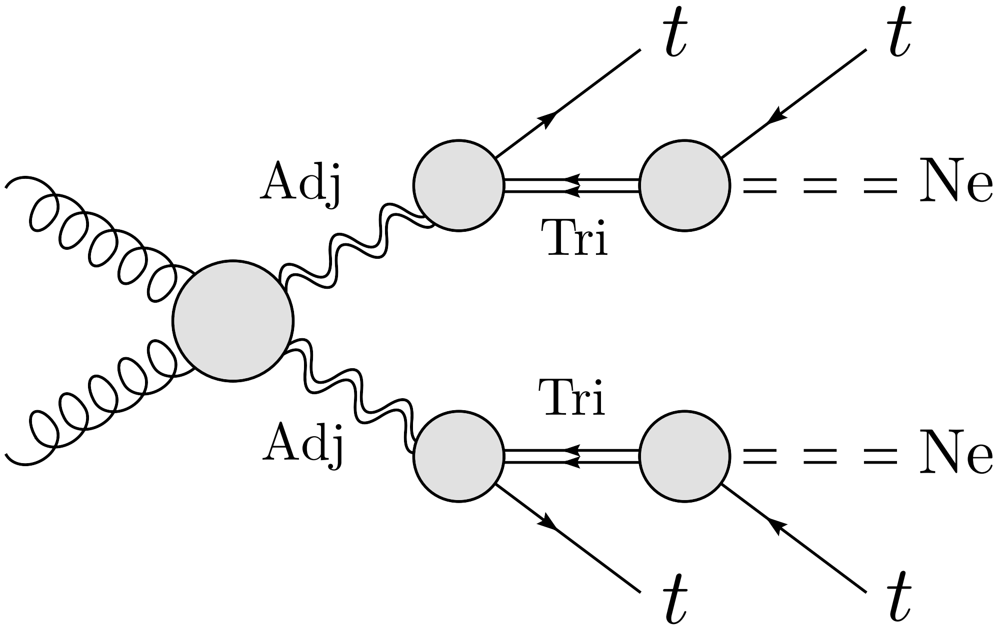
No Amplitudes Means Vast Simplification of MC Input!

```
Adj         : m=700 EM=0 SU3=8
Ne          : m=200 EM=0 SU3=0


Adj > t tbar Ne   : matrix=const


g g > Adj Adj     : matrix=const


g g > ( Adj > t tbar Ne ) ( Adj > t tbar Ne )
```

(Cross Sections / Branching Ratios stored for later reweighting.)

MARMOSET Input

Easy to Extend/Modify Models. Reusable MC.



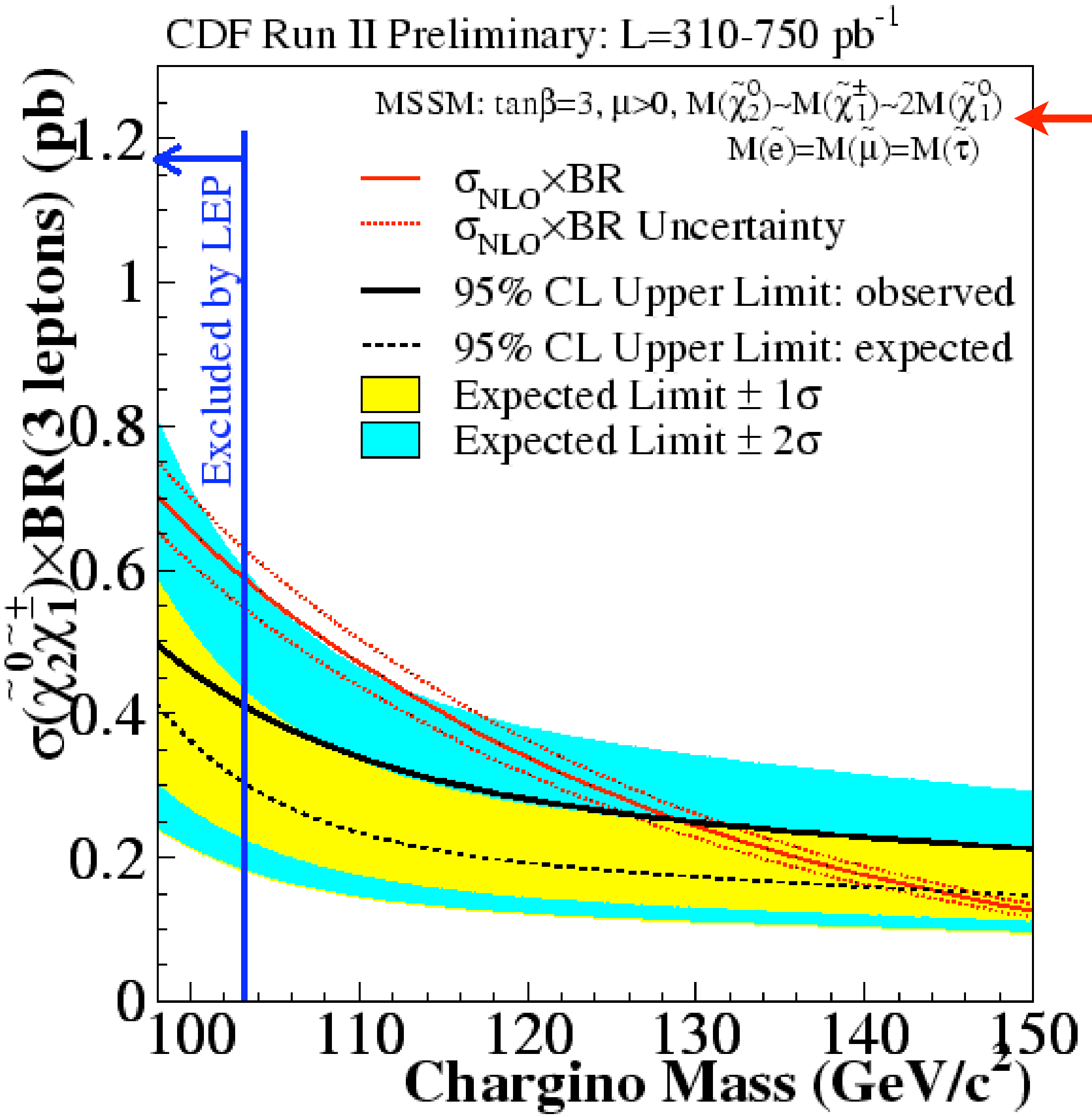
```
Adj       : m=700 EM=0 SU3=8
Ne        : m=200 EM=0 SU3=0
Tri Tri~  : m=500 EM=2 SU3=3

Adj > Tri tbar    : matrix=const
Tri > Ne  t       : matrix=const

g g > Adj Adj     : matrix=const
g g > Tri Tri~    : matrix=const

g g > ( Adj > ( Tri > Ne t ) tbar ) ( Adj > ( Tri~ > Ne tbar ) t )
```

(Monte Carlo generation with Pythia, output in StdHEP XDR format.)

MARMOSET as a Monte Carlo Tool

Using MARMOSET to Study Trileptons at the TeVatron

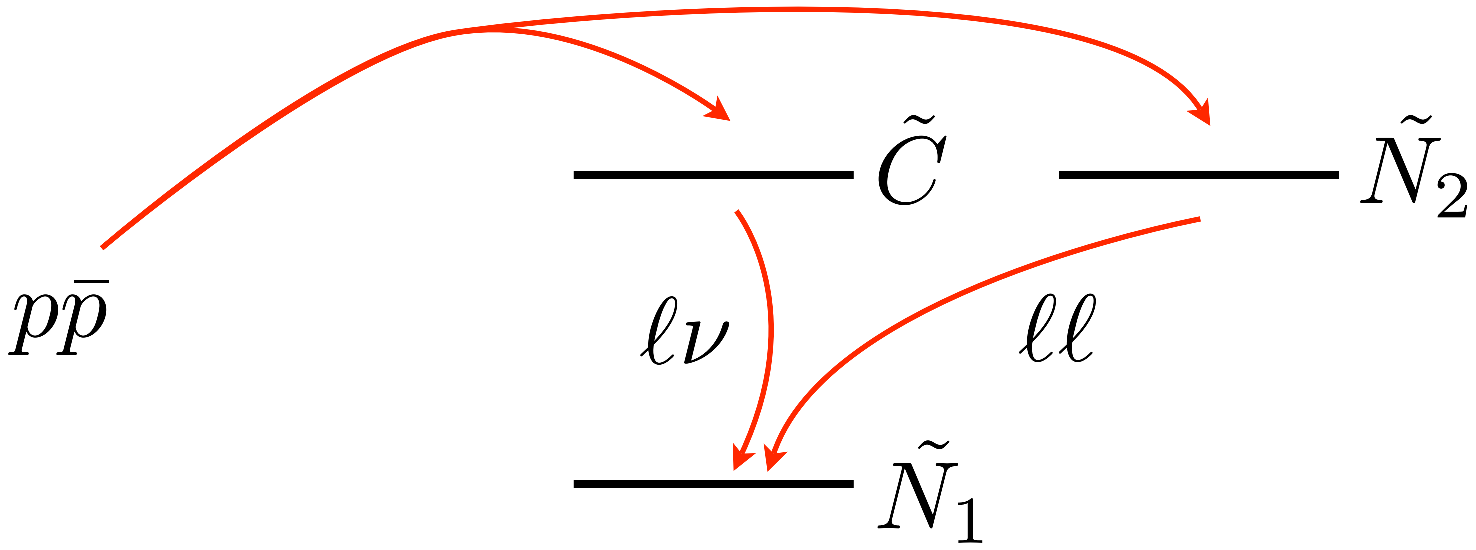
Trileptons at the TeVatron



Why?

This is fundamentally a counting experiment, so detailed kinematics are not very important.

Trileptons at the TeVatron



$p\bar{p}$

$\tilde{C}$

$\tilde{N}_2$

$\ell\nu$

$\ell\ell$

$\tilde{N}_1$

mSUGRA (4.1 parameters)

$m_0,\ m_{1/2},\ A_0,$

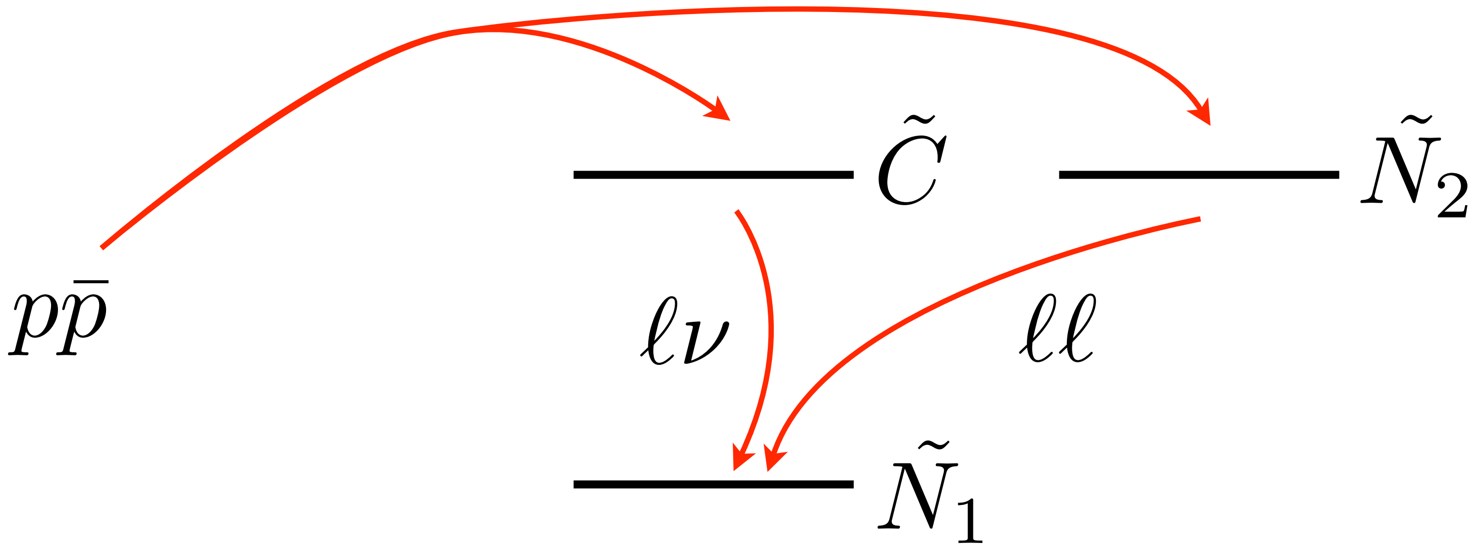
$\mathrm{sign}\ \mu,\ \tan\beta$

Small number of parameters at the expense of complicated correlations among rates, cross sections, and masses.

$m_0 \rightarrow m_{\tilde{\tau}} \rightarrow \mathrm{Br}(\tilde{C} \rightarrow \tilde{N}_1 \ell\nu)$

$m_0 \rightarrow m_{H_u} \rightarrow \mu \rightarrow \tilde{C}, \tilde{N}\ \text{mixing}$

Trileptons at the TeVatron



OSET (8 parameters)

$$\sigma(q\bar{q} \to \tilde{C}\tilde{N}_2)$$

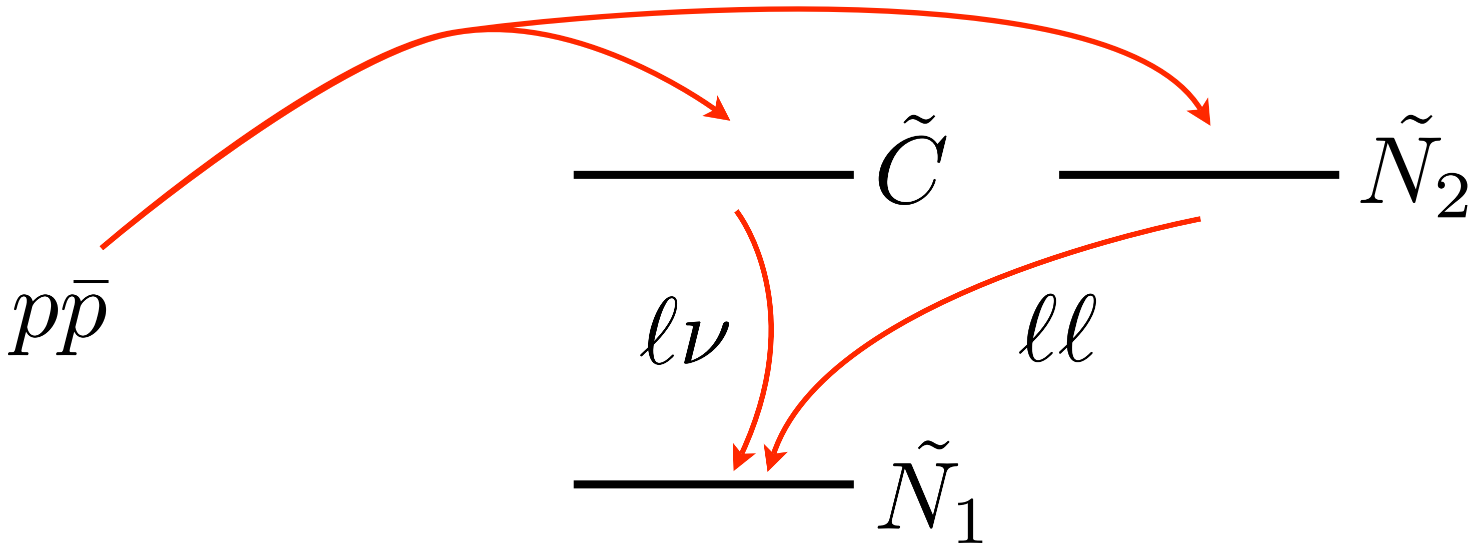
$$\left.\begin{array}{l}\mathrm{Br}(\tilde{C} \to \tilde{N}_1 \ell\nu) \\ \mathrm{Br}(\tilde{N}_2 \to \tilde{N}_1 \ell\ell)\end{array}\right\} \quad \ell = e, \mu, \tau$$

$$m_{\tilde{C}}, m_{\tilde{N}_2}, m_{\tilde{N}_1}$$

More information from same data!

E.g. : How does exclusion depend on heavy-light splitting?

Trileptons at the TeVatron



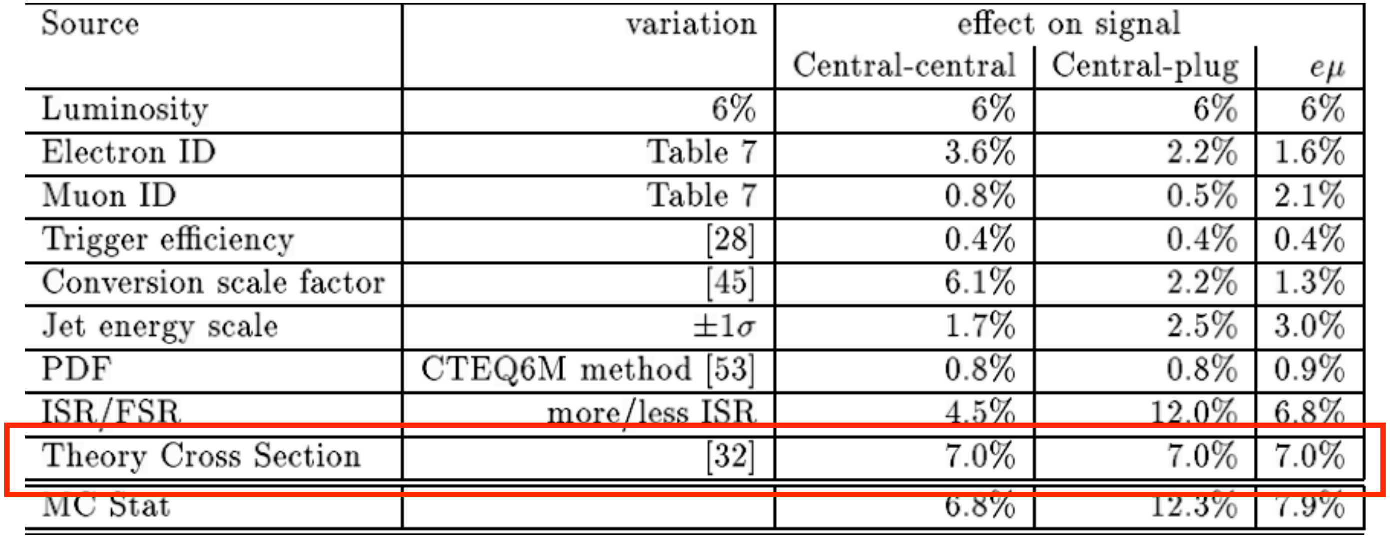
Search Optimized OSET (3 parameters)

$$\sigma(q\bar{q} \to \tilde{C}\tilde{N}_2) \times \mathrm{Br}(\tilde{C} \to \tilde{N}_1 \ell\nu) \times \mathrm{Br}(\tilde{N}_2 \to \tilde{N}_1 \ell\ell)$$

$$\ell = e/\mu \text{ universal, ignore } \tau$$

$$m_{\tilde{C}} = m_{\tilde{N}_2}, m_{\tilde{N}_1}$$

Trileptons at the TeVatron



Source	variation	effect on signal		
		Central-central	Central-plug	$e\mu$
Luminosity	6%	6%	6%	6%
Electron ID	Table 7	3.6%	2.2%	1.6%
Muon ID	Table 7	0.8%	0.5%	2.1%
Trigger efficiency	[28]	0.4%	0.4%	0.4%
Conversion scale factor	[45]	6.1%	2.2%	1.3%
Jet energy scale	$\pm 1\sigma$	1.7%	2.5%	3.0%
PDF	CTEQ6M method [53]	0.8%	0.8%	0.9%
ISR/FSR	more/less ISR	4.5%	12.0%	6.8%
Theory Cross Section	[32]	7.0%	7.0%	7.0%
MC Stat		6.8%	12.3%	7.9%

In mSUGRA, 7% systematic uncertainty on theoretical cross section.

In OSET, total cross section is output of analysis, but systematic uncertainty in differential cross section (e.g. error in distribution of events in central-central vs. central-plug regions).

Differential cross section systematic can be modeled by trying different hard scattering matrix elements. Are they ~7%?

OSETs vs. MSSM?

“I don’t believe in mSUGRA anyway. Why not use the full MSSM instead of mSUGRA?”

- MSSM still has a parameter correlation problem, though less severe. E.g. squark masses affect production cross sections, even though squarks aren’t produced directly.

“Can’t you use SUSY amplitudes but use an OSET bookkeeping scheme?”

- Yes! With reasonable assumptions about the SUSY spectrum (i.e. decoupled squarks for trilepton searches), you can use the SUSY vertex structure. Trade-off between model-independence and model realism.

MARMOSET as an Analysis Strategy

Using MARMOSET to Solve an LHC Olympics Black Box

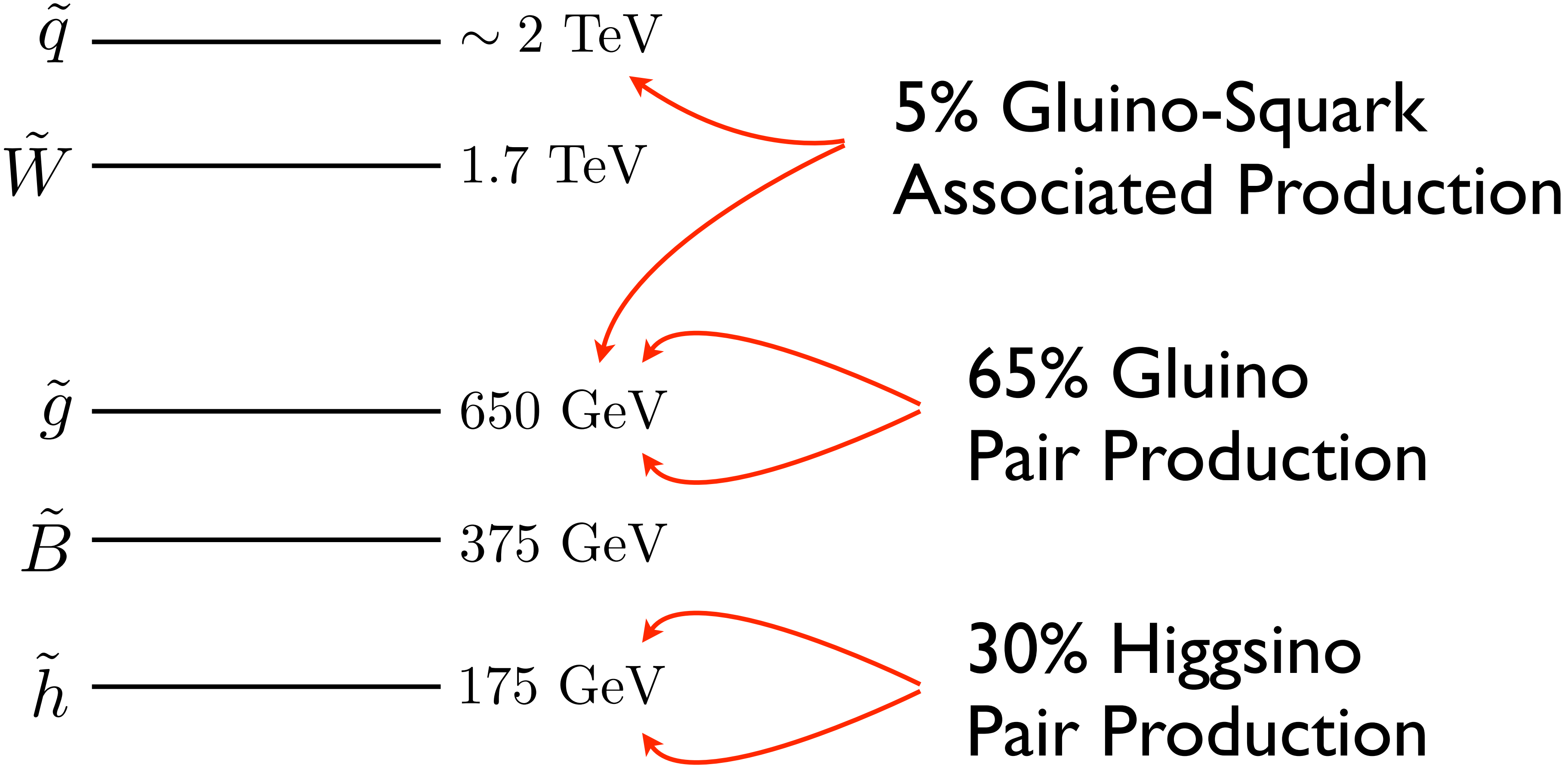
The Michigan Black Box

1st LHC Olympics (Geneva, July 2005)

- Gordy Kane’s string-inspired model that yields the MSSM at low energies.
- Lesson from the LHC Olympics: Easy to get a sense for what is going on (with no SM background). UWash group identified dominant mass scales, decay modes.
- Really hard to make statements about particular models without explicitly simulating them.
- At the 2nd LHC Olympics, Harvard used 3000 CPU/hours to “scan” SUSY models. Lesson: Correlations among SUSY parameters make this very hard. Where’s the physics?

The Michigan Black Box

1st LHC Olympics (Geneva, July 2005)



$\tilde{q}$ ~ 2 TeV

$\tilde{W}$ 1.7 TeV

$\tilde{g}$ 650 GeV

$\tilde{B}$ 375 GeV

$\tilde{h}$ 175 GeV

5% Gluino-Squark Associated Production

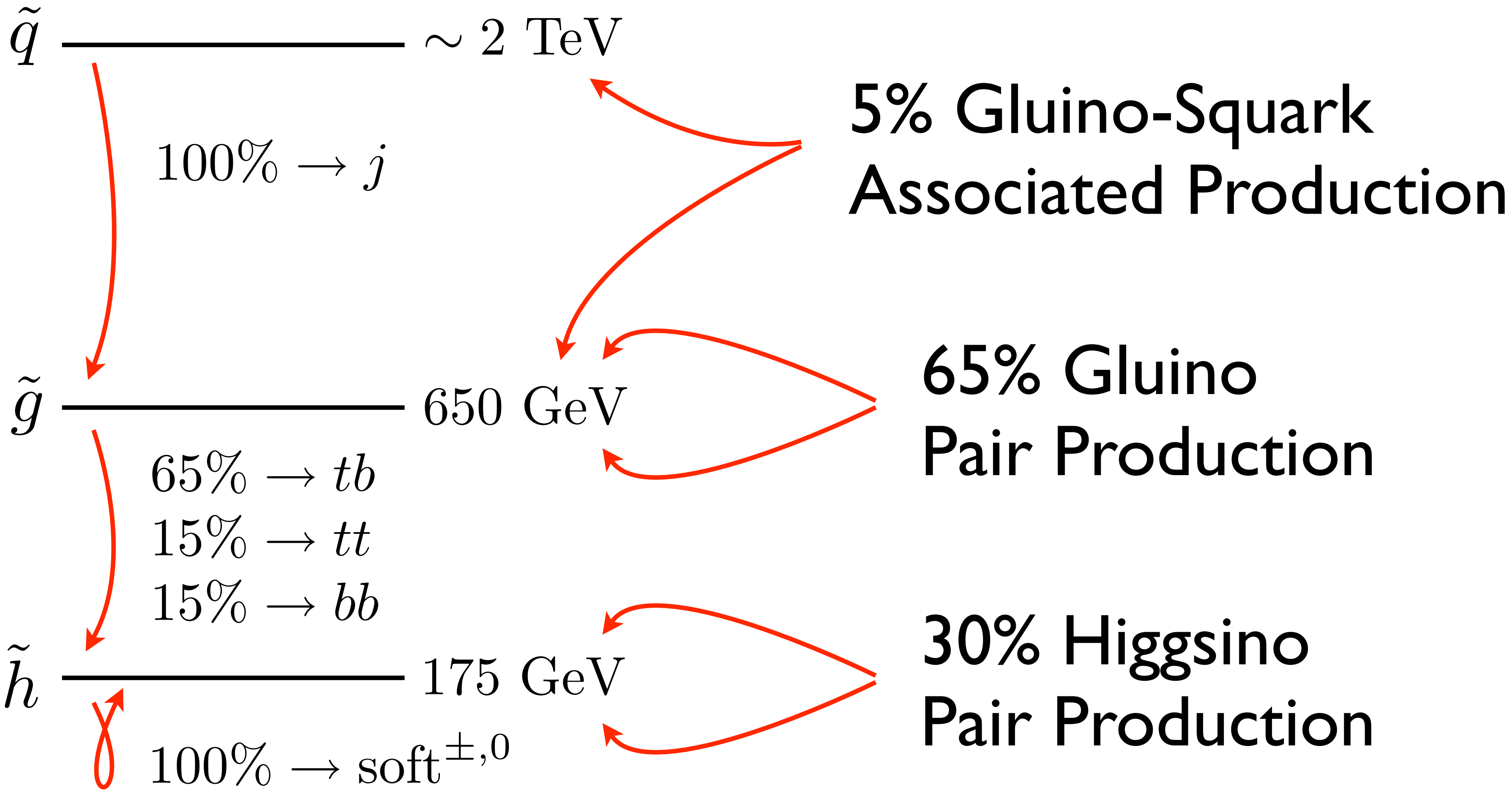
65% Gluino Pair Production

30% Higgsino Pair Production

(This is not the original Michigan Black Box; it is a "v2". My apologies...)

The Michigan Black Box

1st LHC Olympics (Geneva, July 2005)



$\tilde{q}$ ~ 2 TeV

$100\% \rightarrow j$

$\tilde{g}$ 650 GeV

$65\% \rightarrow tb$
$15\% \rightarrow tt$
$15\% \rightarrow bb$

$\tilde{h}$ 175 GeV

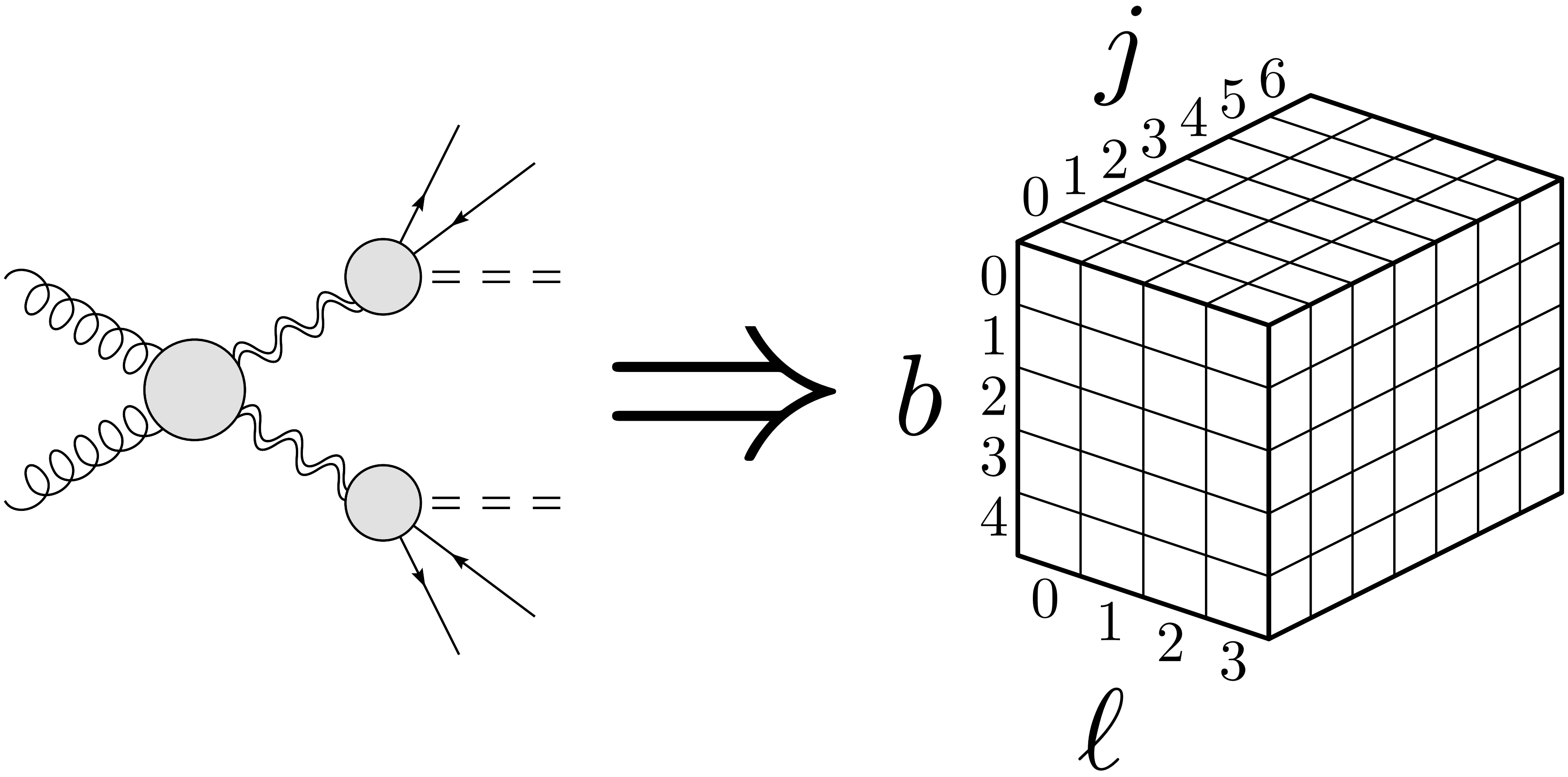
$100\% \rightarrow \text{soft}^{\pm,0}$

5% Gluino-Squark Associated Production

65% Gluino Pair Production

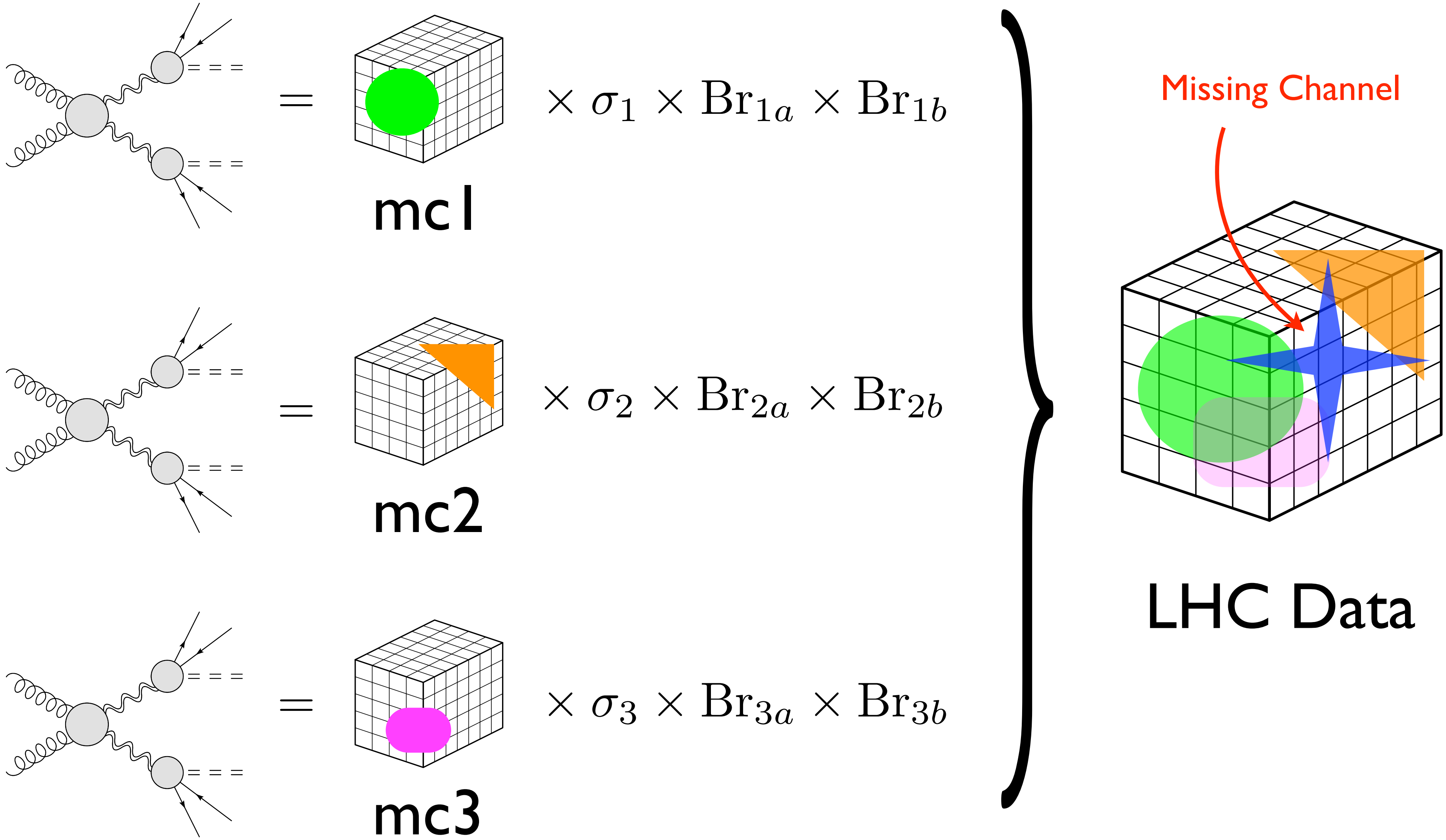
30% Higgsino Pair Production

Simplistic Inclusive Data



Assign every topology to a set of signatures.

Matching Rates to Data



$= \ \times \sigma_1 \times \mathrm{Br}_{1a} \times \mathrm{Br}_{1b}$

mc1

$= \ \times \sigma_2 \times \mathrm{Br}_{2a} \times \mathrm{Br}_{2b}$

mc2

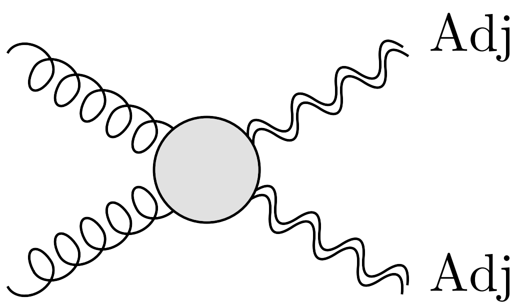
$= \ \times \sigma_3 \times \mathrm{Br}_{3a} \times \mathrm{Br}_{3b}$

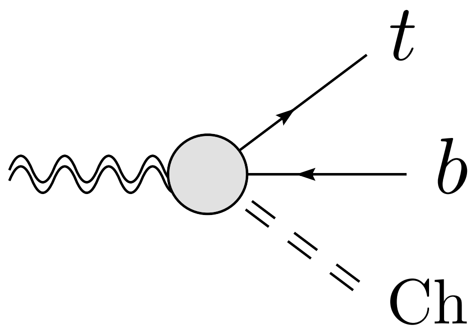
mc3

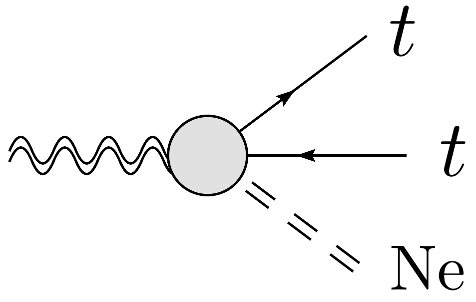
Missing Channel

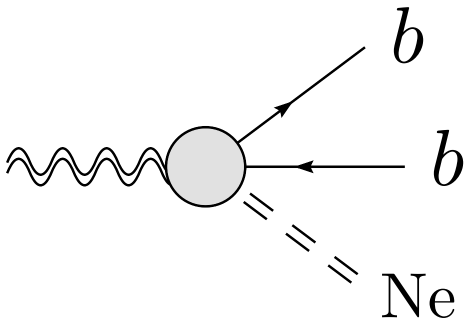
LHC Data

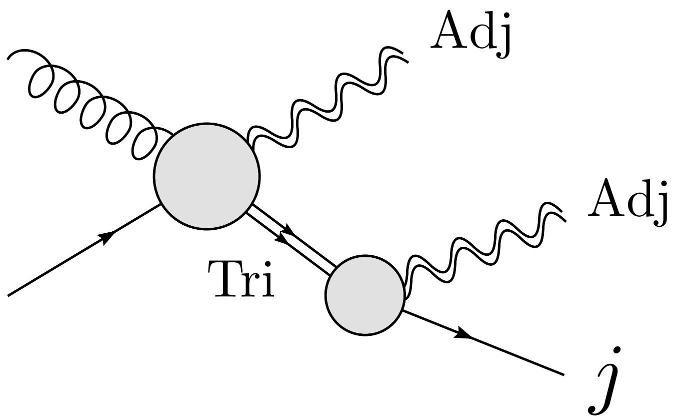
The Michigan OSET

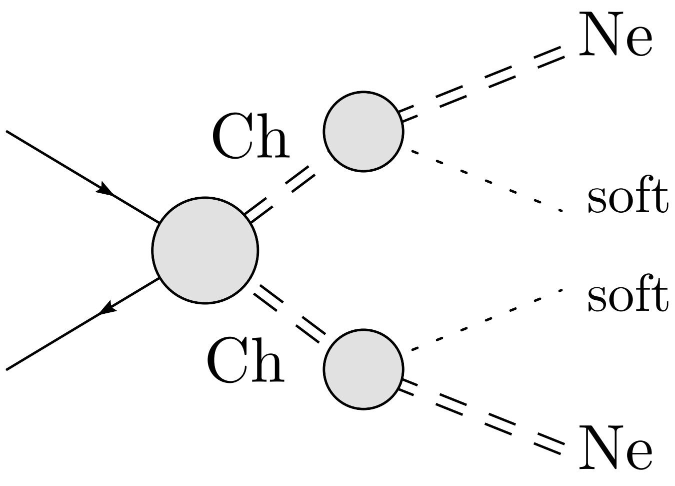






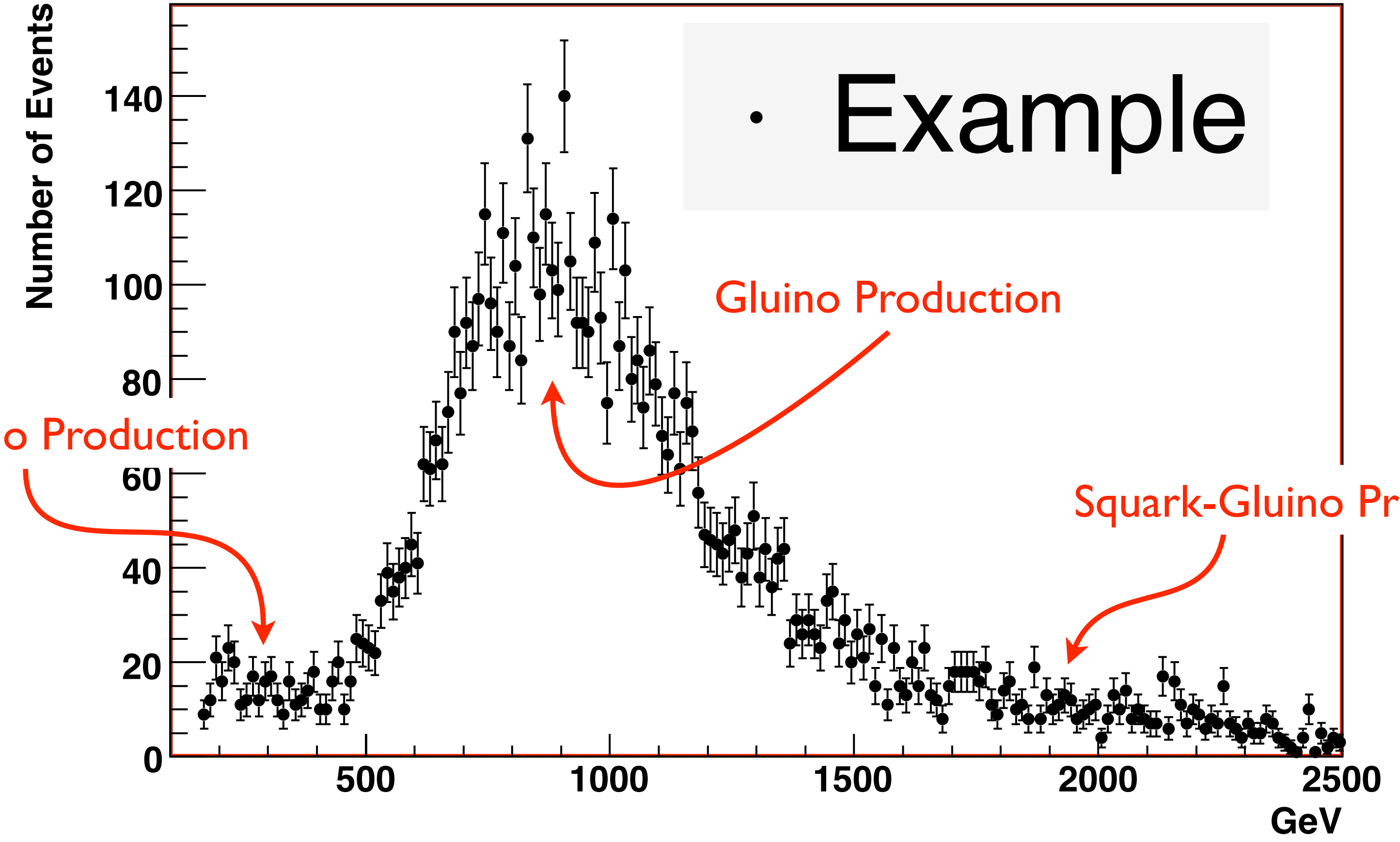






Example Distribution

$$m_{\text{eff}} = \sum_i p_T^i$$



Higgsino Production

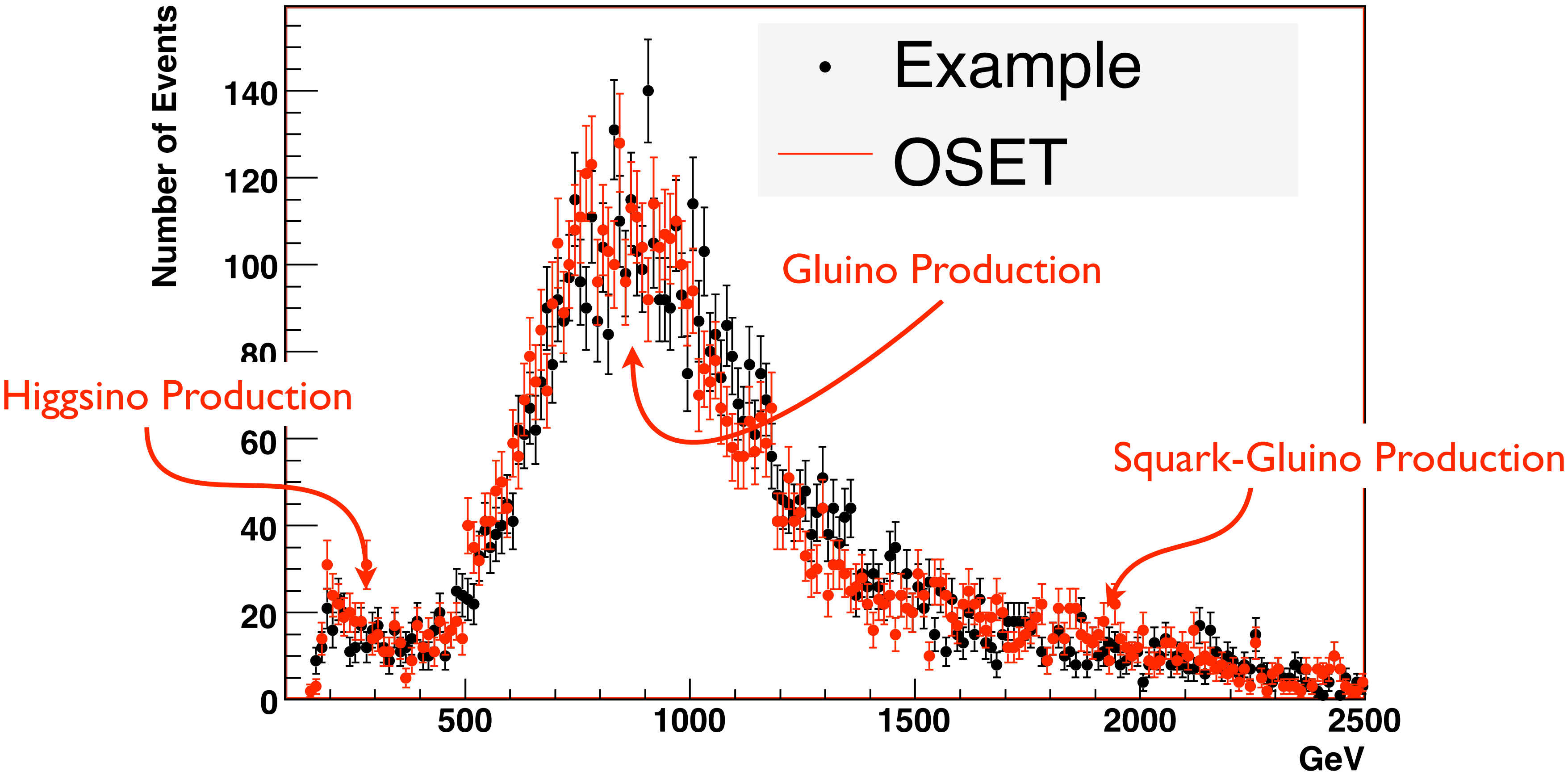
Gluino Production

Squark-Gluino Production

Results of a Global Fit

An OSET with All Three Production Modes

Masses are Fixed at Correct Values for Simplicity



Results of a Global Fit

An OSET with All Three Production Modes

```
                                      Target   Best   Error   +******|******+
l=1  b=2  j=4   (  500<pT< 1300)     59.0   66.5    10.0          |*
l=1  b=2  j=6   (  500<pT< 1300)     76.0   92.2    11.5          |*
l=1  b=2  j=6   ( 1300<pT<14000)     20.0   17.0     5.1         *|
l=1  b=2  j=10+ (  500<pT< 1300)      5.0    7.2     3.6          |*
l=1  b=2  j=10+ ( 1300<pT<14000)      6.0    2.1     3.1         *|
```

```
  Param      Low       Best       High       Name
  total    1.3134    1.3278    1.3422      Sum Sigma
     s0    0.0661    0.0692    0.0723      Sigma( g u > Tr Ad )
     s1    0.4692    0.4757    0.4822      Sigma( g g > Ad Ad )
     s2    0.4489    0.4551    0.4613      Sigma( u ubar > Ch~ Ch  )
   b0_0    0.0356    0.0780    0.1204      Br( Ad > Ne tbar t )
   b0_1    0.0962    0.1237    0.1512      Br( Ad > Ne bbar b )
   b0_2    0.0000    0.0005    0.0765      Br( Ad > Ne ubar u )
   b0_3    0.7240    0.7926    0.8611      Br( Ad > Ch~ t bbar )
   b0_4    0.0000    0.0052    0.0862      Br( Ad > Ch~ u dbar )
   b1_0    0.0000    0.0000    0.0089      Br( Ch > nu_e e+ Ne )
   b1_1    0.9911    1.0000    1.0000      Br( Ch > Ne u dbar )
   b2_0    1.0000    1.0000    1.0000      Br( Tr > u Ad )
```

Could this be done blind?

- At the 3rd LHC Olympics, Harvard made progress on the Rutgers Blackbox using similar techniques. (With MARMOSET, you find a basin of attraction in days, not months.)
- Tools like Sleuth provide a way to make automated cuts to increase signal/background purity, so SM background is probably just a nuisance, not a show-stopper.
- (Other Experimental Caveats)
- Some Harvard/SLAC/Berkeley folks are trying to solve an internal blackbox devised by Nima and Natalia.
- We have an OSET that fits the data reasonably well. But we can't find a theoretical model that would yield that OSET. Are we in a local minimum? Or is Nima just clever?

MARMOSET

MC: σ Br m $\cancel{|\mathcal{M}|^2}$

- As a Monte Carlo Tool, MARMOSET could be used right now at the TeVatron.
  Experimentalist can make their own TeV-athropic models!
- As an Analysis Strategy, MARMOSET requires many correlated excesses.
  Is this experimentally feasible? Trigger stream normalizations? Background estimation in every channel? Global view of the data? Sensitivity? Bias? Systematics?
- (Merging with MadGraph!)

MARMOSET

MC: σ Br m $\cancel{|\mathcal{M}|^2}$

- Factorizes Interpretation Problem
  $\mathcal{L} \longleftrightarrow \mathrm{OSET} \longleftrightarrow \mathrm{LHC}$
- Invariant Characterization of LHC Data with Real Physics Meaning
  OSET language is accessible to theorists outside of the experimental collaborations.
- Evolving OSETs Facilitate Model Building
  Model-independent results suggest new model-dependent searches.

MARMOSET

MC: σ Br m $\cancel{|\mathcal{M}|^2}$

- Is this an “after the champagne” or “before the champagne” tool?

  MARMOSET motivates model-independent discoveries, not just model-independent interpretation.

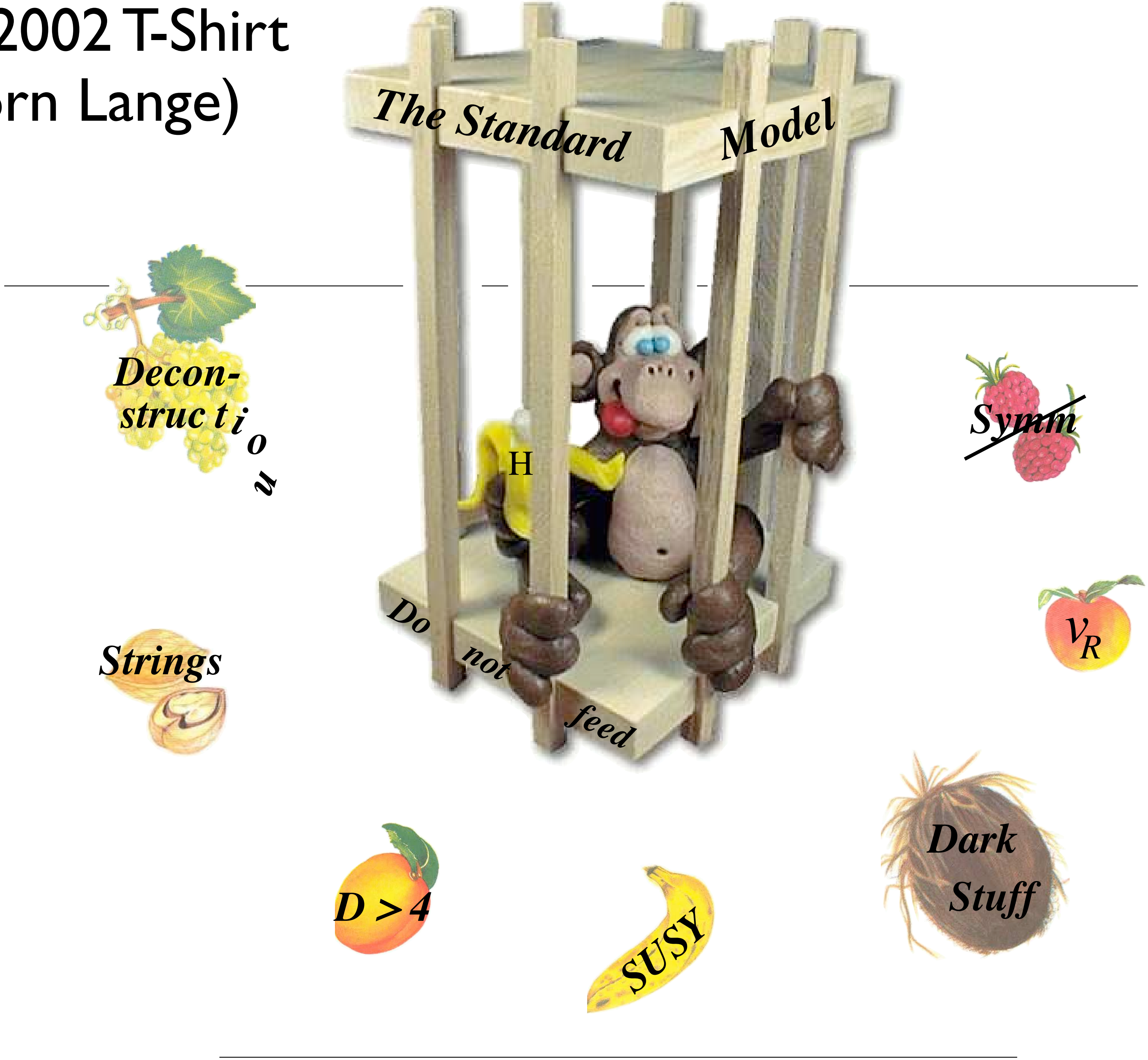
- MARMOSET Needs a Human Operator

  Who will use it? Theorists? Experimentalists? Theorists Looking over Experimentalists Shoulders? Vice Versa?

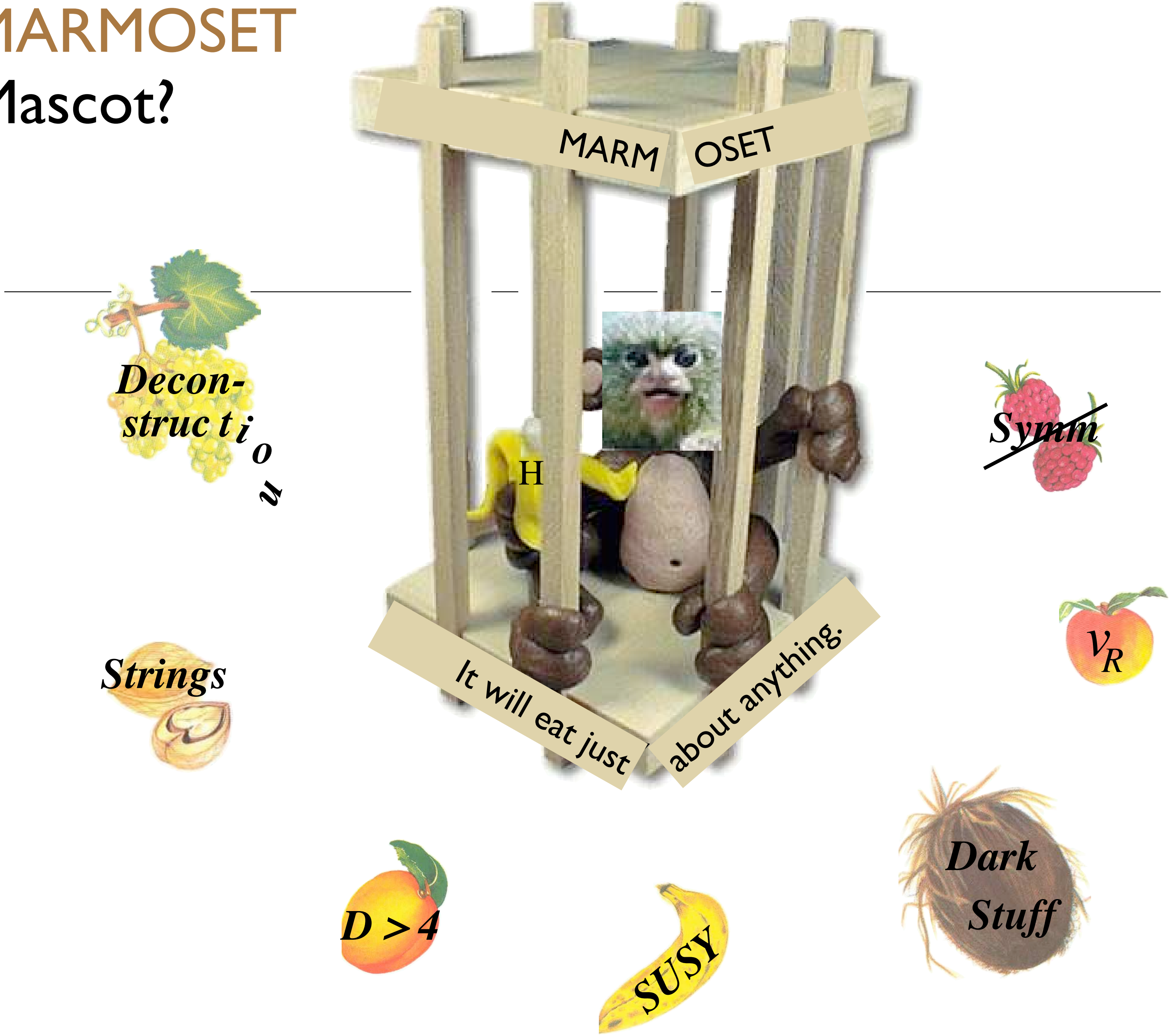
- MARMOSET Needs Debuggers...

  ```
  cvs checkout Marmoset1
  ```

TASI 2002 T-Shirt
(Björn Lange)



The MARMOSET Mascot?



Backup Slides

Theory and the LHC

N years until LHC data

$$N < 3$$

Flavor? Dark Matter? Little Hierarchy Problem?
Little M-theory? Continue Model Building? Landscape?
Higher Dimension Operators? LHC-thropics? ILC?

Theory and the LHC

Two Important Monte Carlo-esque Issues

Standard Model Background Estimation

- Jets/Jet Definitions
- Parton Shower / Matrix Element Merging
- Low Multiplicity NLO Monte Carlo
- High Multiplicity NLO Calculations

Beyond my expertise...

Signal Monte Carlo for Exclusions/Discovery

- Human Time to Code Specific Models in Tree Level MC
- Computer Time to Efficiently Scan Large Class of Models
- Assigning Error Bars
- Comparing Data to MC if Model is Unknown

Enter MARMOSET...

Qualitative Success

Mocking Up Gluino Pairs

$$m_{\text{eff}} = \sum_i p_T^i$$

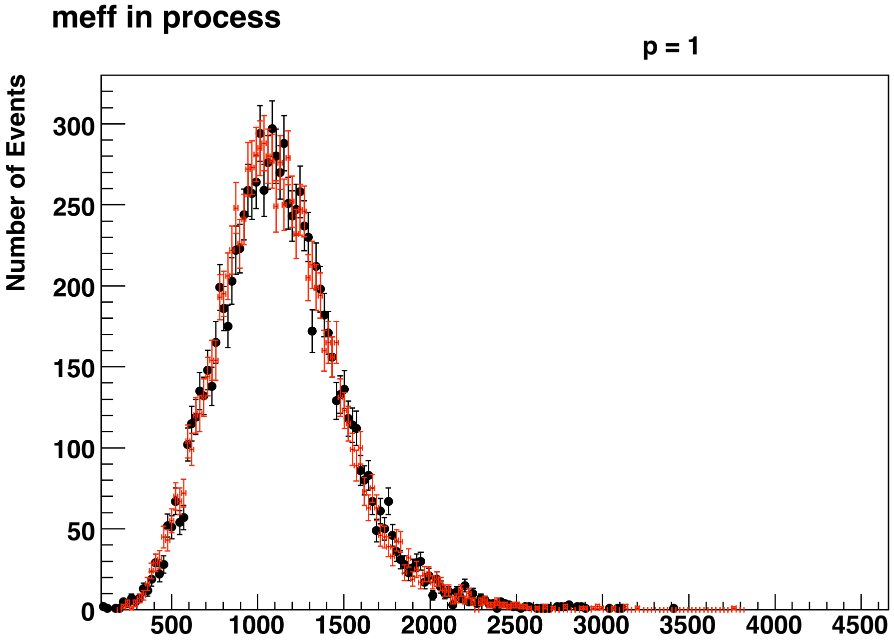


Figure 3: Meff distribution for $|M|^2 = const$

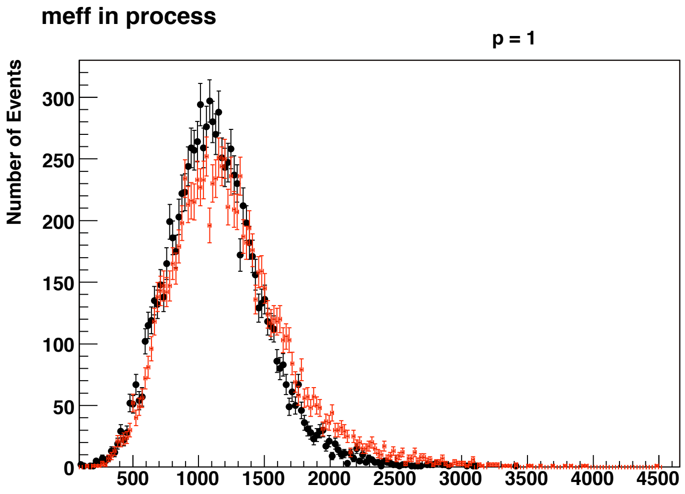


Figure 5: Meff distribution for a $f\bar{f} \to f\bar{f}$ type matrix element.

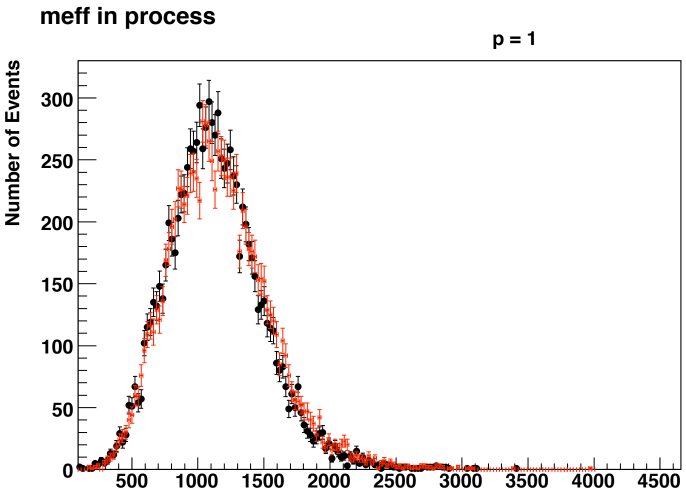


Figure 4: Meff distribution for a $gg \to f\bar{f}$ type matrix element.

Worst Case Scenario

Gluino-Neutralino (i.e. Heavy-Light) Associated Production

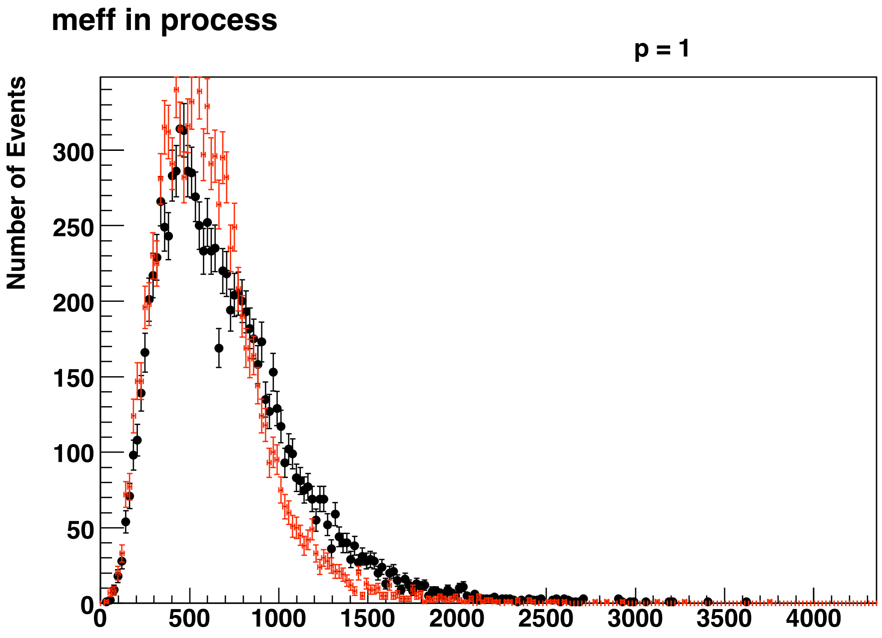


Figure 6: Meff distribution for $|M|^2 = const$

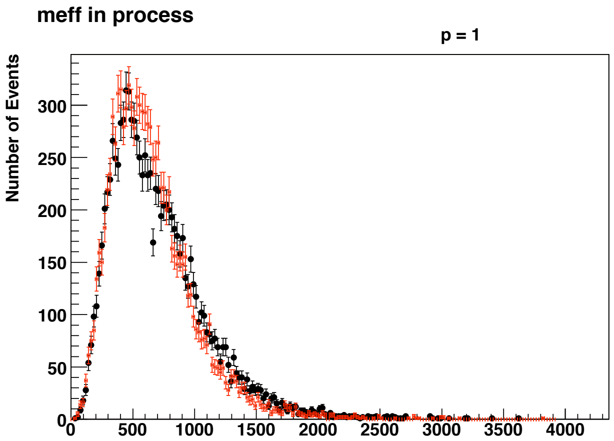


Figure 7: Meff distribution for a t-channel $f\bar{f} \to f\bar{f}$ type matrix element. $m_I = 900$ GeV

Flat amplitudes fail if produced particles explore phase space or if amplitude has singular structure. Is error just in tail?

OSET MC Organization

- Every tree is a separate MC file.
- Cross Sections and Branching Ratios are selected after MC generation.
- (Not enough MC for the desired rate? You can dynamically make more.)
- Reusable signal MC is ideal for experiments that have detailed detector simulations.
- Bonus for inclusive data analysis...

Trileptons in Action...

MARMOSET Demonstration

“Unmotivated” Searches?

Consider this crazy scenario...

- As an experimentalist, you’ve worked really hard to understand the effect of anomalous missing energy on di-jet invariant mass distributions. (Missing E_T dependent Jet Energy Scales?)
- Can you put this knowledge to use in exotic searches?
- How about looking for di-jet resonances in events with one lepton and missing energy?

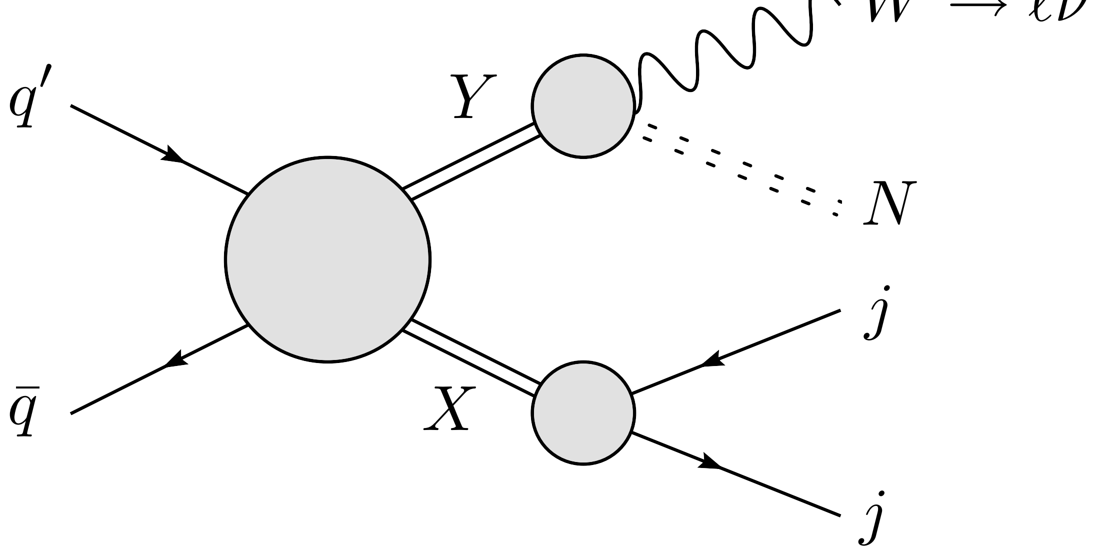
$$p\bar{p} \to (X \to jj)(W \to \ell\nu)\not{E}_T$$

(I’m not advocating this approach, only mentioning how OSETs suggest different analyses.)

“Unmotivated” Searches?

X to 2 Jets, Leptonic W, Large Missing Energy

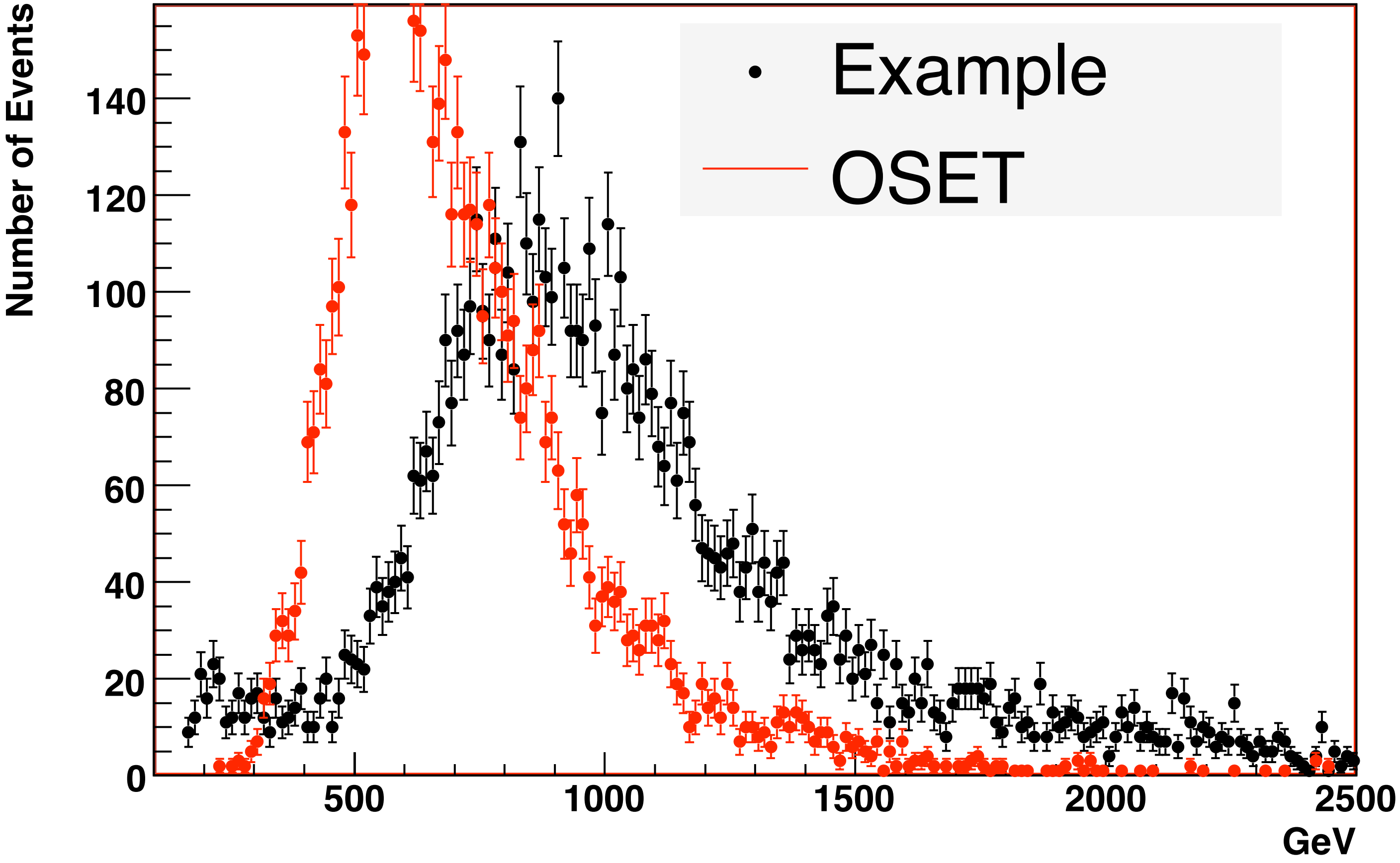
- Is there a good model that gives this final state?
- All you need is something to estimate kinematics of this final state.
- How about...



- Use data or interesting experimental techniques to motivate searches instead of models.

Michigan v1 vs. v2

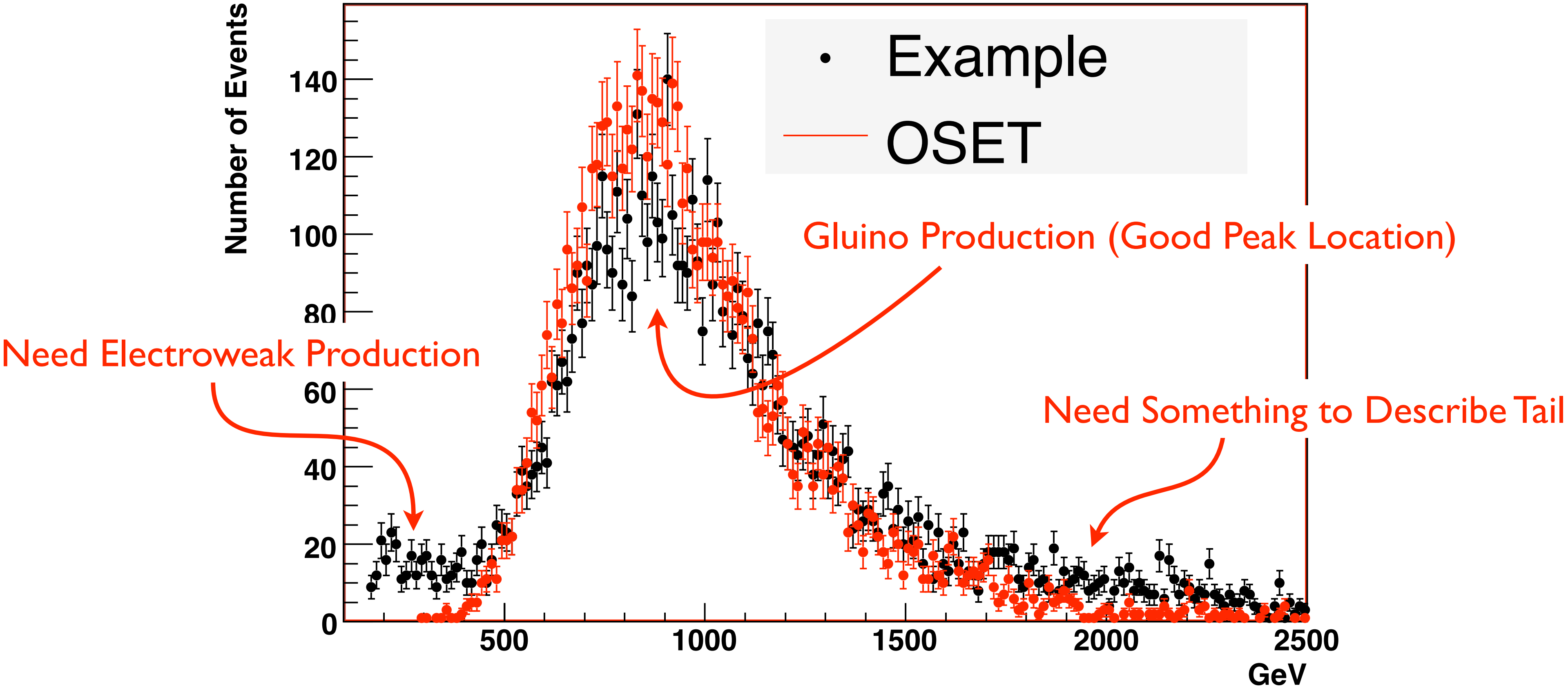
$$m_{\text{eff}} = \sum_i p_T^i$$



Results of a Global Fit

An OSET with Just Gluino Production

Masses are Fixed at Correct Values for Simplicity



Results of a Global Fit

An OSET with Just Gluino Production

```
                                  Target   Best   Error   +*****|*****+
l=0  b=0  j=0  (    0<pT<  500)    101.0    0.0    10.1   +*****|
l=0  b=0  j=0  (  500<pT< 1300)      5.0    0.0     2.6       **|
l=0  b=0  j=2  (    0<pT<  500)    156.0    2.8    12.6   +*****|
l=0  b=0  j=2  ( 1300<pT<14000)      8.0    1.4     3.2       **|
l=0  b=0  j=4  (    0<pT<  500)     43.0   14.9     7.4     ****|
l=0  b=0  j=4  ( 1300<pT<14000)     42.0   18.5     7.5      ***|
l=0  b=0  j=6  (    0<pT<  500)      9.0   14.2     4.5         |*
l=0  b=0  j=6  (  500<pT< 1300)    291.0  337.4    23.1         |**
l=0  b=0  j=6  ( 1300<pT<14000)    106.0   43.3    11.8    *****|
l=0  b=0  j=8  ( 1300<pT<14000)     86.0   24.9    10.3    *****|
l=0  b=1  j=0  (    0<pT<  500)      3.0    0.0     2.1        *|
l=0  b=1  j=2  (    0<pT<  500)     10.0    4.3     3.8       **|
l=0  b=1  j=4  (  500<pT< 1300)    295.0  338.1    23.2         |**
l=0  b=1  j=6  (    0<pT<  500)     10.0   17.8     4.9         |**
l=0  b=1  j=6  (  500<pT< 1300)    622.0  669.8    33.2         |*
l=0  b=1  j=6  ( 1300<pT<14000)    164.0   91.6    15.2    *****|
l=0  b=1  j=8  (  500<pT< 1300)    324.0  352.3    24.0         |*
l=0  b=1  j=8  ( 1300<pT<14000)    156.0   74.6    14.5   +*****|
```

Results of a Global Fit

An OSET with All Three Production Modes