Model Independent Characterization of New Physics with Early Data

Philip Schuster (SLAC)
U.C. Davis HEFTI MET Workshop, April 1, 2009

work with Johan Alwall and Natalia Toro (arXiv:0810.3921)
Model-Independence?

For searches, model-independent means “recyclable”: Results should allow multiple model comparisons to broadly applicable exclusions.

If a signal is observed, then what?
Characterizing New Physics

With a signal, the pretense behind “model-independence” is absent

There’s only one model of nature -- we want to identify it!

The point should be to describe the data, then draw and test inferences
To learn what model describes nature, I want to check consistency of the data with a wide variety of guesses

I’m not an experimentalist, not a detector expert, and not particularly experienced doing careful exp. analysis

Is there a form of the data that I can study?

- I get raw distributions, and a detector simulator
  - I analyze

- First-stage analysis
  - Experimentalists analyze
  - and/or

I analyze
One Theorist’s Perspective

To learn what model describes nature, I want to check consistency of the data with a wide variety of guesses.

I’m not an experimentalist, not a detector expert, and not particularly experienced doing careful exp. analysis.

Is there a form of the data that I can study?

- I get raw distributions, and a detector simulator
- and/or

I analyze

Experimentalists analyze

First-stage analysis

What and why?
(Who does what is an open question)
Simplify, Simplify, Simplify...

Establish approximate mass scale, quantum number, and decay chain determined by the data...
(This is what a model-independent characterization means)

Simplify the model space to sift relevant from irrelevant and resolvable from un-knowable details
(appropriate for early data, low statistics)

Disregard structure that’s “hard” to measure

Simplify until description is typically over-constrained by data

How?
Production and Decay Approximation

\[ |\mathcal{M}|^2 = f(s, \xi) \]
\[ \xi = \beta_{34} \cos \theta = \frac{\hat{t} - \hat{u}}{\hat{s}} = \frac{p_z}{\sqrt{s}} \]

\[ |\mathcal{M}|^2 = \frac{\pi \alpha_s^2}{s^2} \left( \frac{4 m_q^2 - t}{9} \right) + \frac{[m_q^2 - t]s + 2m_q^2(m_q^2 - t)}{(t - m_q^2)^2} + u + st + su + tu \quad \text{(channels)} \]

\[ \frac{d\sigma}{d\hat{t}} = \int \quad \text{Parton Luminosity} \quad \times \quad \text{Phase Space (Threshold)} \quad \times |\mathcal{M}|^2 \]

Homogeneous function of energy $\rightarrow$ uniform shape indep. of $|\mathcal{M}|^2$
\[ |\mathcal{M}|^2 \propto x^m \xi^n \]

This implies
\[ \frac{d\sigma}{dx_T} \propto x_T(1 + x_T)^{m+q} \]

which is independent of \( n \)!

Similarly, rapidity distribution is independent of \( m \)

Information is lost at hadron colliders!
Going On-Shell...

The Basic Idea:

Example: Top Quark
Masses, Rates, and Topology
vs. Amplitudes

Dominant Top Properties:

\[ \sigma(gg \rightarrow t\bar{t}) \]
\[ \text{Br}(t \rightarrow bW) \]
\[ m_t, m_W, m_b \]

Detailed Top Properties:

\[ \frac{d\sigma}{d\hat{t}} \]
\[ W \text{ helicity} \]
\[ t \text{ charge} \]
On-Shell-Effective-Theory

Production:

\[ 2 \rightarrow 1 \]
Use Breit Wigner

\[ 2 \rightarrow 2 \]
\[ |\mathcal{M}|^2 = A + B \left( 1 - \frac{s_{\text{thresh}}}{s} \right) \]

or

\[ |\mathcal{M}|^2 = A + B \left( \frac{s}{s_{\text{thresh}}} - 1 \right) \]

“Normal” Behavior

“Contact” Operator Behavior

“Dominant \( \xi \) correction can be included (not usually necessary)

Decay:

- Polynomial in \( \cos \theta \): rank determined by spins, coefficients by masses. Spin correlations can be included...use a more powerful tool (i.e. MadGraph for example)

- Single-object lab-frame distributions, and many correlations, well approximated by phase space decays.

See: hep-ph/0703088 for detail...
PDFs Simplify Further...

Good physics reasons for simplicity of description

Cross sections fall like

\[ \sim \frac{1}{E^{5-6}} \]

Lowest mass process heavily dominates!

Single production hypothesis not bad

For a given **model**, the observable OSET is much simpler than the complete one.
Simplifying SUSY-like Physics

Common signal of SUSY-like models:

Jets + Missing energy + (leptons?)

If jets+MET+leptons excess(es) are seen, it’s reasonable to assume SUSY-like physics interpretation!
The First Three Questions

Start building evidence for structure with questions that are (relatively) easy and of high theoretical interest.

1) Which colored particles dominate production?
2) What color-singlet decay channels are present, and in what fractions?
3) How b-rich are the events?

Easiest to frame quantitative questions in terms of sharply specified models – what models should we choose, to have a good chance of fitting any jets+MET+leptons signal from SUSY-like physics?
Four Simplified Models

1) Which colored particles dominate production?

Either Gluon partner \( G \) or Quark partner \( Q \)

2) What color-singlet decay channels are present, and in what fractions?

Models with **one** produced species, **one**-stage cascade decay (produced species either \( G \) or \( Q \)).

3) How b-rich are the events?

\( G \): Produce gluon partners that decay to \( q\bar{q}, b\bar{b}, \) or \( t\bar{t} +\text{LSP} \)

\( Q \): Pair-produce partners of \( q_{12}, b, \) and \( t \)

Total of four models

GOAL: As simple as possible to answer these three questions + fit ANY new physics in SUSY-like class well
Simplified Models of Lepton Cascades

From gluon partner:

From quark partner:

Masses

$M_G$
$M_I$
$(M_L)$
$M_{LSP}$

*on or off-shell
Constraining Masses

$H_T = \sum p_T$ (GeV)

(sum over up to 4 jets + leptons + missing ET)

$M_G - M_{LSP}$:

$M_{I} - M_{LSP}$:

Plots from PGS study:
- data=SUSY model
- 500 pb-1
- details in 0810.3921
Constraining $\sigma$ and BR’s

Branching ratios are a detector-independent translation of the lepton counts!

<table>
<thead>
<tr>
<th></th>
<th>$\sigma$ (pb)</th>
<th>$B_{LSP}$</th>
<th>$B_{W}$</th>
<th>$B_{Z}$</th>
<th>$B_{ll}$</th>
<th>$B_{lv}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>11.3</td>
<td>0.0</td>
<td>0.914</td>
<td>0.02</td>
<td>0.063</td>
<td>—</td>
</tr>
<tr>
<td>Green</td>
<td>13.1</td>
<td>0.613</td>
<td>—</td>
<td>0.03</td>
<td>0.052</td>
<td>0.30</td>
</tr>
<tr>
<td>± (***)</td>
<td>0.1</td>
<td>0.04</td>
<td>0.05</td>
<td>0.02</td>
<td>0.005</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Plots from PGS study
- data=SUSY model
- 500 pb$^{-1}$
- details in 0810.3921
Claim:

For a wide variety of signatures, and MSSM parameter regions, these simplified models work remarkably well!

Suggests that applicability will extend beyond the MSSM.

Designed for answering early new-physics questions and establishing the correct range of topologies and rates.

see: arXiv:0810.3921
Building Models from Simplified Models

Experimental comparison:
Simplified Model (Leptons) vs. Data
(shown over ttbar background)

Theorist’s comparison
Simplified Model (PGS) vs. 3 SUSY models (PGS)

(not PGS vs. CMS/ATLAS!)

Many systematic errors factor out for a PGS vs. PGS comparison...
Comparing Gluon and Squark Partners

Two ways to get jet & lepton counts in simplified models:
- quark partner decays to 1 jet with W’s in cascades
- gluon partner decays to 2 jets with no hadronic W/Z in cascades

Real physics can interpolate between the two!

Models look different, but only distinguishable with more statistics!

*Can’t even distinguish 100% gluino from 100% squark, let alone mixture*
Over-constrained Models are Useful

Identify Distributions that cannot be explained without adding structure beyond simplified models

Softer lepton source in signal than simplified models: can’t match while keeping invariant mass distribution agreement – indicative of e.g. multiple cascades, but refined two-cascade model would be under-constrained
(Study heavy flavor separately from leptons)

From gluon partner:

From quark partner:

Different structures / different patterns of b-tag multiplicity
Using Simplified Model Fits

Important to see several kinds of results

• Simplified model best fits

• Parameter uncertainties, particularly careful treatment of weakly constrained parameters

• Comparisons of the data to expectations for best-fit simplified model — both for distributions used in the fit and for diagnostics

Back-of-the-envelope analysis

• “Good fit” suggests what regions of parameter space to study in model-building

• “Bad fit” suggestive of additional structure (multiple species production, multiple cascades in decays, etc...)

Quantitative comparison

• Can compare predictions of any model to simplified model predictions (e.g. in PGS) to gauge consistency with data.
Find a signal

Compare data to models

Simplified models

Qualitative discrepancies

Refined models

Synthesis & Speculation
Find a signal

Compare data to models

Qualitative discrepancies

Refined models

Synthesis & Speculation

Experiment

Theory
Discussion...
Backup
Preliminary Interpretation

When we do get distributions, there will be a lot we can do

Easy Cases:

- Self-calibrating signal, like a mass peak
- HT observable
  - peak $\sim 1.7 \times$ Mass difference (depending on decay chain) is roughly encoded

Di-object mass can have distinctive phase space cutoff, giving a constraint on decay chain mass difference

$$m_{\text{edge}} = \frac{\sqrt{(m_2^2 - m_L^2) (m_L^2 - m_{\text{LSP}}^2)}}{m_L},$$

or

$$m_{\text{end}} = m_2 - m_{\text{LSP}}.$$
Preliminary Interpretation

What about less kinematically sharp distributions?

Easy to compare to well-simulated guesses...much harder to turn out physical quantities (masses, branching ratios, cross sections...or even “detector-corrected” distributions)

even in principle, distributions not narrow

further smeared by detector

Jet Count

B Count

Jet ET

Lepton ET
Goals for Early Characterization

We want to find consistent & predictive explanations of all the data...then discriminate options, measure parameters...etc

Obstacles:
- distributions with no sharp features do not map clearly onto a set of particles, masses and decays
- many regions of parameter space to consider in each model
Constraining $\sigma$ and BR’s

Signatures quite distinctive (dilepton pairs on Z peak, opposite-flavor leptons, ...) except $B_W$ looks like $B_{lv} \times 0.32 + B_{LSP} \times 0.68$.

Study extreme limits, e.g. $B_W=0$, or $B_{lv}=0$
Additional constraints

Exchanging $W \leftrightarrow (l\nu + \text{direct})$ changes jet multiplicities, and correlation with lepton counts.

Choosing gluon/squark partner also changes jet multiplicities.

Varying particle masses changes kinematic distributions.

$H_T = \sum p_T$ (GeV) (sum over up to 4 jets + leptons + missing ET)
Building Models from Simplified Models

Experimental comparison:
Simplified Model (Leptons) vs. Data
(shown over ttbar background)

Theorist’s comparison
Simplified Model (PGS) vs. 3 SUSY models (PGS)

(not PGS vs. CMS/ATLAS!)
Building Models from Simplified Models

Experimental comparison:

Simplified Model (Leptons) vs. Data
(shown over ttbar background)

Theorist’s comparison

Simplified Model (PGS) vs. 3 SUSY models (PGS)

(not PGS vs. CMS/ATLAS!)
Building Models from Simplified Models

Experimental comparison:

Simplified Model (Leptons) vs. Data
(shown over ttbar background)

Theorist’s comparison

Simplified Model vs. 3 SUSY models

---

Experimental comparison:

# Evts/Bin

- 20
- 40
- 60
- 80
- 100
- 120
- 140
- 160

G-LCM best fit

On-Shell C

Off-Shell B

Off-Shell A

---

Theorist’s comparison

# Evts/Bin

- 50
- 100
- 150
- 200
- 250

pseudoData

2Lep2Jet_counts

G-LCM best fit

On-Shell C

Off-Shell B

Off-Shell A
Building Models from Simplified Models

Experimental comparison:
Simplified Model (Heavy flavor) vs. Data
(shown over ttbar background)

Theorist’s comparison
Simplified Model vs. 3 SUSY models

Number of B Jets (pT>30 GeV) (in lepton-inclusive region)
Constraining $\sigma$ and BR’s

Branching ratios well constrained by these counts (aside from the W/Lnu ambiguity):

<table>
<thead>
<tr>
<th></th>
<th>$\sigma$ (pb)</th>
<th>$B_{LSP}$</th>
<th>$B_W$</th>
<th>$B_Z$</th>
<th>$B_{ll}$</th>
<th>$B_{l\nu}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>11.3</td>
<td>0.0</td>
<td>0.914</td>
<td>0.02</td>
<td>0.063</td>
<td>—</td>
</tr>
<tr>
<td>Green</td>
<td>13.1</td>
<td>0.613</td>
<td>—</td>
<td>0.03</td>
<td>0.052</td>
<td>0.30</td>
</tr>
<tr>
<td>± (**)</td>
<td>0.1</td>
<td>0.04</td>
<td>0.05</td>
<td>0.02</td>
<td>0.005</td>
<td>0.01</td>
</tr>
</tbody>
</table>

** Don’t take these errors too seriously!! No backgrounds, etc.

Masses: Best fit to kinematics, with LSP fixed at 100 GeV

...can try to find models that reproduce it
W vs Inu Modes

Within each of the two models (quark-partner or gluon-partner initiated), $W \leftrightarrow (l\nu+\text{direct})$ changes jet multiplicities, and correlation with lepton counts.

Lepton-veto region

2-Lepton region (different cuts)

(in some cases, lepton kinematics also constrains these fractions)
Comparing Gluon and Squark Partners

Two ways to get jet & lepton counts in simplified models:
- quark partner decays to 1 jet with W’s in cascades
- gluon partner decays to 2 jets with no hadronic W/Z in cascades
Real physics can interpolate between the two!

Lepton-veto region

Models look different, but not distinguishable without more statistics!
Better observables also help.
Constraining $\sigma$ and BR’s

<table>
<thead>
<tr>
<th></th>
<th>$\sigma$ (pb)</th>
<th>$B_{qq}$</th>
<th>$B_{bb}$</th>
<th>$B_{tt}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>11.4</td>
<td>0.44</td>
<td>0.56</td>
<td>—</td>
</tr>
<tr>
<td>Red</td>
<td>11.4</td>
<td>0.33</td>
<td>0.03</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Counts appear consistent with one pair-produced particle decaying to $bb$ or $q's$ (high heavy-flavor fraction)

$b$ kinematics most consistent with top pairs

# of b-tagged jets ($p_T > 30$ GeV)

$p_T$ of leading b-tagged jet

qq, bb, and tt

qq and bb
Constraining $\sigma$ and BR’s

<table>
<thead>
<tr>
<th></th>
<th>$\sigma$ (pb)</th>
<th>$B_{qq}$</th>
<th>$B_{bb}$</th>
<th>$B_{tt}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>11.4</td>
<td>0.44</td>
<td>0.56</td>
<td>—</td>
</tr>
<tr>
<td>Red</td>
<td>11.4</td>
<td>0.33</td>
<td>0.03</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Counts appear consistent with one pair-produced particle decaying to $bb$ or $q$’s (high heavy-flavor fraction)

$b$ kinematics most consistent with top pairs

Number of $B$ Jets ($p_T > 30$ GeV) (in lepton-inclusive region)

$qq$, $bb$, and $tt$

$qq$ and $bb$

$# $Evts/Bin$

$# $Evts/Bin$

$\#$ of $b$-tagged jets ($p_T > 30$ GeV)

$p_T$ of leading $b$-tagged jet
Constraining $\sigma$ and BR’s

<table>
<thead>
<tr>
<th></th>
<th>$\sigma$ (pb)</th>
<th>$B_{qq}$</th>
<th>$B_{bb}$</th>
<th>$B_{tt}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>11.4</td>
<td>0.44</td>
<td>0.56</td>
<td>—</td>
</tr>
<tr>
<td>Red</td>
<td>11.4</td>
<td>0.33</td>
<td>0.03</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Counts appear consistent with one pair-produced particle decaying to $bb$ or $q's$ (high heavy-flavor fraction)

b kinematics most consistent with top pairs

Weak deviation suggestive of additional 2b source that does not also imply 4b (e.g. in SUSY – top squark direct production, gluino-squark assoc. production)