Simplified Models for a First Characterization of New Physics at the LHC

Philip Schuster (SLAC)
West Coast LHC Theory Meeting, UCLA, Nov. 21, 2008

work with Johan Alwall and Natalia Toro (arXiv:0810.3921)
(also thanks to JA and NT for use of graphics/slides in this talk)
Outline

Assumption: A robust jets+MET excess has been seen at the LHC

Why a model-independent characterization of new physics is valuable.

The first three questions to ask about the new physics

Four “simplified models” to frame and answer these questions

How the simplified models are constrained, and how to use them
Why new physics?

- **Hierarchy problem**: What cancels top contribution to Higgs mass?

- Symmetry $\Rightarrow$ “partners” with same 3-2-1 quantum numbers as Standard Model particles.

- Minimal: top, SU(2)$\times$U(1) gauge boson partners – top partner is colored, will be produced (if light enough)

- SUSY, Randall-Sundrum, or Universal Extra Dimensions: Spacetime symmetry $\Rightarrow$ partners for **all** SM particles
Why new physics?

- Hierarchy Problem → partners
- TeV-scale Dark matter? → parity
  - If partner states are odd under a new parity, lightest parity-odd particle is stable and a DM candidate (also helps guarantee proton stability)

- Two consequences for LHC searches
  - new particles produced in pairs
  - some collision energy in lost to invisible particles (2 of them)

- We’ll call models with partners and parity “SUSY-like”: e.g. weak-scale supersymmetry, universal extra dimensions, and “theory space” models (e.g. Little Higgs) w/ T-parity
Well-Motivated Signature

If jets+MET+leptons excess(es) are seen, it’s reasonable to assume SUSY-like physics interpretation!
Jets+MET excess is evidence for SUSY-like new physics.

.....how do we learn more about it?

Focus on questions that (within SUSY-like framework) are **almost guaranteed** to be relevant and **accessible in the first few years of the LHC** (really 1-10x discovery luminosities).
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~1.7*Mass difference (depending on decay chain) is roughly encoded

\[ m_{edge} = \sqrt{(m_2^2 - m_L^2) (m_L^2 - m_{LSP}^2)} \]

\[ m_{end} = m_2 - m_{LSP}. \]

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

What about less kinematically sharp distributions?

Jet Count

B Count

even in principle, distributions not narrow

further smeared by detector

Jet ET

Lepton ET

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

lepton-inclusive signal region (3 jets, $p_T>75$ GeV, $HT>350$, $MET>100$)

Assume experimentalists have understood and subtract backgrounds (we’ll make life easy: ignore them completely!)

Data with expected background (tt only)

# of b-tagged jets (transverse momentum > 30 GeV)
Example

Suppose a theorist were to guess the correct model, and simulate it in PGS ....

Data vs. correct model (wrong simulator)

Limited by accuracy of model-builder’s description of detector

(I have used one version of PGS as my “detector” and “experimentalist’s simulator”, and a different version as my “theorist’s simulator”)

# of b-tags (backgrounds subtracted)

b-counts off by almost 50% in some bins!
Unrealistic Game Plan

lepton-inclusive signal region (3 jets, pT>75 GeV, HT> 350, MET > 100)

“Inverting” efficiencies based on # of jets, pT’s, and etas in each event? (Very hard, not standard)
**Alternative Approach**

**Experimenter’s comparison**

- Red line: model X with gluon-partner pair production, $32\% \rightarrow qq \ LSP, 68\% \rightarrow tt \ LSP$

(limited by experimental collaboration’s detector modeling)

**Data vs. some model “X” (well simulated)**

**Instead of comparing their model to data, any theorist can simulate their model and model X in PGS**

- **model X vs. my favorite model**

(There are systematic errors in this procedure — e.g. if model X mis-models kinematics and PGS mis-models efficiencies but they are partially corrected)
We’re kind of prepared

Let’s look at what CMS and ATLAS are prepared to use:

**Benchmarks** : Good for designing searches – most production & decay topologies of interest are in a benchmark, but most **combinations** of topologies are not.

**mSUGRA (or similar constrained frameworks)** : Fixed mass ratios still prohibit qualitative changes to spectrum; need many different frameworks to cover phenomenology.

**MSSM (and other many-parameter models)** : Technically challenging to optimize, very hard to present and interpret globally (e.g. if >1 region of parameter space is consistent with data)
A Need For Something Else

Broader problem:
Suppose a good MSSM fit is found. MSSM parameter variations change *multiple* observables – if I am studying a non-MSSM model, how do I know whether it’s consistent?

Want:
A small (manageable) parameter space
Coverage of large range of possible phenomenology
Physical parameters → broadly useful error bars
Motivating A Compromise

It’s easy to propose a simple model

If we’re going to pick them ahead of time, we should have specific “high value” physics questions in mind

Our models should be the simplest possible for the given questions

- Once there’s robust evidence for new physics:

  Signal distributions → Simplified Model parameter fits & comparison → Model hypothesis-testing (inside & outside collaborations)
  → More sophisticated discrimination among models
  → Precise parameter determination
  → Searches for sub-leading processes predicted by models
  → ...
The First Three Questions

Very simple questions for a broad-brush characterization of SUSY-like data

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?

Answers to these questions motivate consistent spectra!
heavy-flavor enhanced: 
( $\gg 1/3$ but $< 100\%$ )

 EW

(3rd gen. enhanced by phase space)

 EW

(3rd gen. enhanced by $y_t$)

 EW

(squark prod. dominates, produces light flavors)

 Large 1,2 lepton decay modes (consistent w/on-shell slepton):

 $\tilde{\chi}_1^0$ (one EW state—no cascades)

 strong

 $\tilde{l}$ $\tilde{B}/\tilde{W}$ (skips slepton)

 $\tilde{l}$ $\tilde{W}$ and/or $\tilde{h}$

 Qualitative features motivate basins of attraction...
The First Three Questions

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?
The Difficult Part

What makes the choice of simplified models not so obvious is:

Models need to answer our earlier physics questions without introducing too many hard to constrain “flat” directions that complicate experimental analysis.

Models should have broad coverage of well-motivated physics topologies included.

Models need to be sufficiently detailed to describe the kinematics of the objects that define the signal regions (i.e. trigger objects).

Detector effects can be encapsulated in the model fits provided we properly describe:

- leading (3-4) jets
- missing energy profile
- leading (1-3) leptons

Need some way of assessing systematic errors that might arise from oversimplification.
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 \textit{one} produced species, \textit{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 \( t\bar{t} \) +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:

\[ \sigma_G \quad B_{LSP} \quad B_{W/Z} \quad B_{\ell\ell} \quad B_{\ell\nu} \]

*on or off-shell

\[ M_G \quad M_I \quad (M_L) \quad M_{LSP} \]

From quark partner:

\[ \sigma_Q \quad B_{LSP} \quad B_{W/Z} \quad B_{\ell\ell} \quad B_{\ell\nu} \]

\[ M_Q \quad M_I \quad (M_L) \quad M_{LSP} \]
Simplified models don’t include all possible SUSY-like behavior in models:

- Quark-partner and gluon-partner production
- Different decay modes for LH and RH quark partners
- Multiple cascades
  etc...

But our goal is to answer the three questions – additional structure can often be guessed from comparisons of data distributions to those predicted by simplified models.

Trying to match too much at once → larger parameter space, less constrained and harder to present. No in-principle reason not to extend the models, but important to understand what we can do with simplest models.
Simplified Models of Lepton Cascades

From gluon partner:

Parameters:

- **One** total production cross-section
- **Five** branching fractions (sum to 1); three easy
- **Three** masses (four if slepton on-shell)

Can be constrained in data (2 parameters harder)
For a detailed analysis of case studies confined to well-motivated “signal regions”, as well as omissions and caveats, please see: arXiv:0810.3921
Constraining $\sigma$ and BR’s

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

Study extreme limits, e.g. $B_W=0$, or $B_{\text{IV}}=0$
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.
Why do Simplified Models Work Well?

Good physics reasons for apparent simplicity

Cross sections fall like $\sim \frac{1}{E^{5-6}}$ Lowest mass process heavily dominates!

Single production hypothesis not bad for well-spaced spectrum.

Competing processes require similar mass scale Easier to model jets as coming from single source in this case
detailed multiplicity is tricky...
Subtle Substructure

What about the modeling of the production and decays?

Cross Sections dominated near thresholds:

$$\frac{d\sigma}{dt} = \int \text{Parton Luminosity} \times \int \text{Phase Space (Threshold)} \times |M|^2$$

$$\rightarrow |M|^2$$ well approximated by constant!

Homogeneity of PDF in $E_{cm}$ and $y_{cm}$

(systematic & universal corrections necessary for highly asymmetric kinematics

formally correct for simple pT, eta observables; useful much more broadly)

See: hep-ph/0703088 for detail...

Most important structure captured by an “OSET” scheme...
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 $t\bar{t}$ +LSP

$Q$: Pair-produce partners of $q_{12}$, $b$, and $t$

Total of four models
Heavy Flavor Models

From gluon partner:

\( \sigma_G \)

\( B_{qq} \)

\( B_{bb} \)

\( B_{tt} \)

Masses

\( M_G \)

\( M_{LSP} \)

From quark partner:

\( \sigma_Q \)

\( \sigma_B \)

\( \sigma_T \)

Masses

\( M_{Q/T/B} \)

\( M_{LSP} \)

Different structures / different patterns of b-tag multiplicity
Gluon Heavy Flavor Model: Top/Bottom Fractions

(1) fit using only b-jet counts, and no tt mode:
- detector-independent characterization of b-jet fraction
- check consistency with one source of b-jet pairs

(2) include tt, lepton counts
- is it consistent for all leptons come from tops?
- check kinematics, too

Note the omission of lepton cascades here!
Another kind of information

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 multiple cascades
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.

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

see: arXiv:0810.3921
Using Simplified Model Fits

Important to see several kinds of results

• Simplified model best fit
• 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.
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!)

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!)
Conclusions

Many kinds of new physics could be visible at the LHC in jets+missing energy searches

- How to characterize them to maximize theory returns?
- And in a detector-independent way?

Simplified models are a concrete proposal:

- 2 models for leptonic cascades, 2 models for heavy flavors
- Few, simple parameters – easy to fit and easy to interpret

Fits of simplified models to data facilitate qualitative and quantitative comparisons to data by theorists outside collaborations
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
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)

# Evts/Bin

0
100
200
300
400
500
600
700
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>$B_{\text{LSP}}$</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>

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

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

Low-significance discrepancy...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)
**Constraining Masses**

\[ \sigma_G \rightarrow G \rightarrow q \bar{q} \]

- **Masses**
  - \( M_G \)
  - \( M_I \) (\( M_L \))
  - \( M_{LSP} \)

**\( H_T = \sum p_T \) (GeV)**

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

**OSSF (e+e- and \( \mu^+\mu^- \)) events:**

di-lepton invariant mass

(note: this “data” is different from the other slides)
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 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

qq, bb, and tt
qq and bb

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

$p_T$ of leading b-tagged jet

pseudoData
G-HFM (exc lep)
G-HFM (inc lep)
Constraining $\sigma$ and BR’s

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

<table>
<thead>
<tr>
<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>
</tr>
<tr>
<td>Red</td>
<td>11.4</td>
<td>0.33</td>
<td>0.03</td>
</tr>
</tbody>
</table>

# 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)