Generic Signatures of New Physics at the LHC

M. E. Peskin
August 2006
Certain new physics signatures at the LHC are unmistakable and straightforward to interpret:

\[ Z' \rightarrow \mu^+ \mu^- \]

black hole production

But other -- maybe more probable -- types of new physics are more subtle to observe.
SUSY is canonically associated with the signature of

missing ET + multijets

Actually, in my SSI lectures, I argued that (almost) any model of electroweak symmetry breaking that also leads to cosmic dark matter predicts a large cross section for complex events with missing ET and multiple jets.

If the colored particles in the new sector have masses below 1 TeV, the cross sections are large -- 10s of pb!
\[ \sigma_{\text{tot}}[\text{pb}] \]

PROSPINO

\[ \sqrt{S} = 14 \text{ TeV} \]

\[ m [\text{GeV}] \]
Unfortunately, as is well appreciated, missing ET is not a good first indicator of new physics. Many detector imperfections can fake missing ET, and some Standard Model physics processes make this easier to do.

**Bob Cousins** has a more precise version of this comment:

We would like a $5\sigma$ anomaly to claim a discovery. But, the tails of distributions at the LHC are difficult to understand, and are generated in complex ways. Will we be able to claim any knowledge at $5\sigma \sim Prob < 10^{-4}$?
So, I think that this is an important research problem (even for theorists):

Assume that we really have conventional SUSY with squarks and gluons at $M \sim 600$ GeV, giving a substantial missing energy signal.

With 1 fb-1 of data, can we really prove that the Standard Model is violated, and that we not are just screwing up?

The question of whether the model is actually SUSY is a secondary one. We will have many years to figure this out (and it may take another accelerator to answer it definitively).
What makes it possible to solve Cousin’s problem is that SUSY and similar models lead to many distinct signatures:

<table>
<thead>
<tr>
<th>observable</th>
<th>irreducible bkgd</th>
<th>mismeasurement bkgd</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma$(multi-jet)</td>
<td>QCD multijets</td>
<td>QCD 3-jets, $t\bar{t}$</td>
</tr>
<tr>
<td>$E_T$ w. &gt; 2 jets</td>
<td>$Z \rightarrow \nu\bar{\nu}$</td>
<td>QCD multijets, $t\bar{t}$, $W +$ jets</td>
</tr>
<tr>
<td>$\ell + E_T$</td>
<td>$W +$ jets, $t\bar{t}$, $WZ$</td>
<td>QCD multijets, $b\bar{b}$+ jets, multijets, $b\bar{b}$+jets, $t\bar{t}$</td>
</tr>
<tr>
<td>$\ell^+\ell^+$</td>
<td>$W^+W^+$</td>
<td></td>
</tr>
<tr>
<td>$\ell^+\ell^- + E$</td>
<td>$ZZ$</td>
<td>$Z +$ jets, $W +$ jets, $t\bar{t}$</td>
</tr>
<tr>
<td>$Z^0 + E$</td>
<td>$ZZ$</td>
<td>$Z +$ jets,</td>
</tr>
</tbody>
</table>

The first new physics discovery at the LHC will likely be in the form of 3 $\sigma$ anomalies in many or all of these channels.
This does mean that the backgrounds to each of these signatures must be understood, in a way that is based on the data.

The research problem is to have a plausible strategy for how this can be done.

The rest of this talk will have some not-very-coherent remarks on this problem. Hopefully, we can work toward some real solutions over the next year.
There are two types of backgrounds that must be considered:

Irreducible background sources that mimic the signature.

a much-discussed example is

\[ q\bar{q} \rightarrow Z^0 + n \text{ jets} , \quad Z^0 \rightarrow \nu\bar{\nu} \]

Reducible backgrounds that fake signatures through mismeasurement or in tails of distributions:

- \(10^{-4}\) probability for jets to fake \(E_T\)
- \(10^{-2}\) probability for leptons to find cracks in the detector
- \(10^{-2}\) probability for light quarks to fake b vertices
Of the irreducible backgrounds, only
\[ q\bar{q} \rightarrow Z^0 + n \text{ jets} , \quad Z^0 \rightarrow \nu\bar{\nu} \]
has a cross section large enough to mimic the SUSY signal.

For single leptons, and especially for dileptons, we are in a much better situation:
\[ \sigma(pp \rightarrow WW + X) \sim 30 \text{ pb} \]
decreasing to 100 fb if we require \( p_T > 100 \) GeV. To fake a SUSY signal, we also need multiple jets.

Like-sign dileptons can be a particularly powerful probe: These are generated in typical \( pp \rightarrow \tilde{g}\tilde{g} + X \) events, but the first irreducible SM background comes from \( pp \rightarrow W^+W^+ + g + X \)
Here are Padhi’s results on backgrounds to the missing ET and missing ET + single lepton signatures. Note the large effect of W (missing lepton) and $tt$ events compared to the irreducible backgrounds.
In principle, the background from

\[ q\bar{q} \rightarrow Z^0 + n \text{ jets} , \quad Z^0 \rightarrow \nu\bar{\nu} \]

can be completely controlled experimentally by measuring

\[ q\bar{q} \rightarrow Z^0 + n \text{ jets} , \quad Z^0 \rightarrow \mu^+\mu^- , e^+e^- \]

with overall pT balance. SUSY cannot fake these events. This reaction is also a good source of data to understand the reaction

\[ q\bar{q} \rightarrow W + n \text{ jets} \]

It should be noted that the Z BR to neutrinos is 6 times that to leptons, so it will be important to study samples with fewer jets or smaller jet ET to get higher statistics. Then the extrapolation to large ET, visible hadronic mass must be understood. The cross sections are roughly:

\[ \sigma(pp \rightarrow Z^0 + n \; j, \; Z^0 \rightarrow \ell^+\ell^-) \sim 10,000 \; \text{pb} \cdot \left( \frac{1}{30} \right)^n \]
Experimental measurements can be compared to LO event generators (esp. ALPGEN) and (if Lance and Carola work hard) to NLO multiparton calculations. Precision comparison requires matching of pertubative QCD and parton shower calculations.

Here are some issues that one would like to study:

What is the **normalization** of the data/true QCD relative to LO as a function of the number of jets? Is there a smooth way to extrapolate to large numbers of jets using LO as a basis? How does one choose the scales for the $n$ factors of $\alpha_s$?

For large (TeV) values of the visible hadronic mass, how many **hard partons** are needed to make events that contribute to the observed $n$-jet signal?

How many **b-tags** are there in these events? Are we correctly estimating b production from gluon splitting?
$\frac{d\sigma}{dE_T}[\text{pb/GeV}]$

$W \rightarrow e\nu + \geq n \text{ jets}$

CDF Run II Preliminary

CDF Data $\int dL = 320 \text{ pb}^{-1}$

- $W$ kin: $E_T^e \geq 20 [\text{GeV}]$; $|\eta|^\ell \leq 1.1$
- $M_T^W \geq 20 [\text{GeV/c}^2]$; $E_T^\nu \geq 30 [\text{GeV}]$

Jets: JetClu $R=0.4$; $|\eta|<2.0$

hadron level; no UE correction

LO Alpgen + PYTHIA
Total $\sigma$ normalized to Data

$1\text{st jet}$

$4\text{th jet}$
Turn now to reducible backgrounds, for which tails of distributions combined with mismeasurement can fake the SUSY signatures.

In Padhi’s study, ordinary multi-jet reactions are relatively unimportant, but $t\bar{t}$ makes important contributions to the background.

This is not surprising. The cross section for $t\bar{t}$ production at the LHC is huge - 840 pb - and $t\bar{t}$ events begin with large visible hadronic energy.

I remind you that b-tagging is not necessary to find top quarks at the LHC. Look at events with isolated lepton + 4 jets, pick the 3 jets with highest pT, already top quarks should appear.
\( \sigma_{\text{jet}}(E^\text{jet}_T > \sqrt{s}/20) \)

\( \sigma_{\text{jet}}(E^\text{jet}_T > 100 \text{ GeV}) \)

\( \sigma_{\text{Higgs}}(M_H = 150 \text{ GeV}) \)

\( \sigma_{\text{Higgs}}(M_H = 500 \text{ GeV}) \)
The problem is now to understand the tails of the production distribution of top quarks, especially toward high visible hadronic mass. The physics contribution is from events with hard gluon radiation in the process of $t\bar{t}$ production.

It is a very interesting question whether one can define a sample of $t\bar{t} + g$ events for detailed study that is not potentially polluted by SUSY background.

Studying QCD event samples with gluon emitting brings back the problems that we saw in $Z + \text{jets}$ of matching between different orders of QCD calculation.
So let me present a little more material on this issue:

DO has measured the azimuthal correlation of the two hardest jets in a multijet event.

For $\Delta\phi > 2\pi/3$, these configurations can be produced by $2 \rightarrow 3$ parton QCD reactions. For smaller angles, they require $2 \rightarrow 4$ reactions, i.e., reactions at the next order.

Here is the comparison to leading order, and to a parton-level NLO calculation by Zoltan Nagy.
HERWIG and PYTHIA generate parton showers, so in principle they can generate an arbitrarily large number of partons. However, the basic algorithms of these programs are correct only when the extra partons are approximately collinear with the primary core partons.

The most recent versions of HERWIG and PYTHIA try to fix this problem by boosting the rate of partons produced at wide angles. For PYTHIA, Peter Skands has introduced the terminology:

‘wimpy showers’: generate extra partons using the collinear approximation, and stop at a pT comparable to the highest pT in the core process.

‘power showers’: generate extra partons using the collinear approximation, and stop only at a pT comparable to the $\sqrt{s}$ of the full hadron-hadron system.

The first is more technically correct; the second is a better model of the extra jets.
One can use these various methods to estimate the rates for producing 1, 2 extra jets in association with $t\bar{t}$ at the LHC. Finding agreement between theory and experiment for these rates will give us some confidence that we can predict backgrounds to new physics due to the tails of SM top production.

It would be even better to have a detailed method for merging QCD calculations at different levels; the measurements would give a nice test of that theory.
Here are estimates by Plehn, Rainwater, and Skands of extra gluon radiation in supersymmetry events. The effect is much smaller than for top quarks. This is similar to the case of top quarks at the Tevatron, where the available energy limits extra gluon radiation.

600 GeV gluino pair production:

600 GeV squark pair production:
Finally, I would like to note two classes of signatures that are not usually discussed as generic signatures of SUSY but should be considered in a list of anomalies.

The first is total jet activity or visible mass, measured by

\[ H_T = \cancel{E}_T + E_{T\ell} + \sum_i E_{Ti} \]

CDF has succeeded in identifying top quark events at the Tevatron without b tagging, based on the large $H_T$ together with other kinematic features of the events.

In the same way, SUSY events at the LHC are expected to populate higher $H_T$ regions than QCD, W + jets, top + jets, as we have seen in Padhi’s plots.
W+3 jets

CDF Run2, 194 pb$^{-1}$

\[ \ttbar: 65.8 \pm^{22.1}_{21.5} \]

\[ \text{multijet: } 32.7 \pm^{0.0}_{0.0} \]

\[ \text{W-like: } 419.6 \pm^{28.9}_{28.9} \]

W + 4 jets

CDF Run2, 194 pb$^{-1}$

\[ \ttbar: 57.2 \pm^{15.6}_{15.7} \]

\[ \text{multijet: } 7.4 \pm^{0.0}_{0.0} \]

\[ \text{W-like: } 52.8 \pm^{16.8}_{14.9} \]
The second signature is an anomalous class of events with multiple b tags, together with high $H_T$ or $E_T$ or both.

In models of electroweak symmetry breaking, there can be special effects on the spectrum for the partners of b and t. In SUSY, some b and t squarks are typically lighter than the other squarks, especially in models with large $\tan \beta$. At the SUSY point LCC3, for example, most of the SUSY production is pair-production of b and t squarks. Then every SUSY event has a pair of b jets, with $m(bb) > \text{TeV}$ and accompanied by additional jets and/or leptons.

A subtlety here is that we will need to separate $t\bar{t}$ events with one missing $\nu$ from this multiply-b-tagged sample. Again, we see the need for a good discriminator between the top and SUSY event samples.
Summary:

SUSY and other models of electroweak symmetry breaking and dark matter imply a variety of signatures of new physics that should produce significant anomalies over the Standard Model expectation in the first 1 fb-1 of data.

The discovery of new physics at the LHC will most likely involve observation of 3 $\sigma$ anomalies in at least 5 of these channels.

There are no apparent obstacles to discovering these anomalies, but we need to be prepared. It will be time well spent to develop a systematic approach to these observations in preparation for the first LHC data.