ATLAS INTERNAL NOTE

ATL-PHYS-PUB-2006-000

August 10, 2007

The ATLAS SUSY Search Strategy with $100\ \mathrm{pb}^{-1} - 1\ \mathrm{fb}^{-1}$ of Data

Tobias Golling, Sven Vahsen, Joseph Virzi
Lawrence Berkeley Laboratory

Abstract

If Supersymmetry (SUSY) exists, what is the optimum way to search for it with the ATLAS experiment at the LHC? The answer depends on a number of factors we cannot know before the LHC turns on: how SUSY is broken, the SUSY mass scale, the in situ detector performance, and the cross section of standard model processes at LHC beam energies.

Focusing on mSUGRA models, we compare the sensitivity and robustness of different search strategies, for multiple physics and detector performance scenarios. We employ a full GEANT4 simulation of the ATLAS detector for all background processes and several mSUGRA points. We also examine a larger range of mSUGRA and more general SUSY models with fast simulation. We perform a detailed validation of the fast simulation. The optimum event selection criteria yield a typical 5-sigma sensitivity for mSUGRA-like SUSY up to $m_{\mathrm{squark}} = xxx$ and $m_{\mathrm{gluino}} = yyy$ for $100\ \mathrm{pb}^{-1}$, and up to $m_{\mathrm{squark}} = xxx$ and $m_{\mathrm{gluino}} = yyy$ for $1\ \mathrm{fb}^{-1}$.

This note documents the authors' contribution to ATLAS SUSY CSC note #5.

1 Introduction

2 Generation of Monte Carlo Samples

2.1 Standard Model Background: Pythia

Searches for Supersymmetry are characterized by events with large missing transverse energy and jets and leptons with high p_T , given by the lightest supersymmetric particle escaping the detector undetected and the large SUSY mass scale.

Matrix element generators such as ALPGEN are believed to best model the dominant background contributions to SUSY signals, such as W+jets, Z+jets and QCD multijet production[1)]. However, the ALPGEN generation is very involved for several reasons: To get an inclusive sample of $\geq N$ jets samples have to be generated exclusively with a fixed number of jets by applying the so-called MLM matching and the samples are added afterwards. This generation mechanism is very inefficient, time-consuming and error-prone.

Alternatively, the SUSY group decided to pursue a second more conservative way to use PYTHIA for the background description, which can also serve as a means to cross-check or as a back-up solution.

Due to limited resources only 1 fb^{-1} can be fully simulated and only with tight generation filters[2)].

2.1.1 Signal region

The generation filters for the signal region reflect the minimum cuts for a SUSY search: Missing transverse energy above 100 GeV, respectively 80 GeV, and at least two jets with high p_T , i.e. the leading jet above 80 GeV, and the second-leading jet above 40 GeV:

- For the multijet samples J4-J8 the JetMETEstimator with following settings is used:
  JetMETEstimator.TruthJetCollection = "Cone4TruthJets"
  JetMETEstimator.doTrgCuts = False
  JetMETEstimator.TotalMET = 100

- For the vector boson samples the MissingEtFilter is used:
  MissingEtFilter.MEtCut = 80.*GeV

- For all samples the TruthJetFilter is used with following settings:
  TruthJetFilter.Njet=2;
  TruthJetFilter.NjetMinPt=40.*GeV;
  TruthJetFilter.NjetMaxEta=5;
  TruthJetFilter.jet_pt1=80.*GeV;
  TruthJetFilter.TruthJetContainer="Cone4TruthJets";

The properties of the PYTHIA samples representing the SUSY signal region are summarized in Table 1.

The QCD multijet samples are generated in the usual bins of p_T , using "pysubs ckin 3 140." and "pysubs ckin 4 280." e.g. for J4. Events with ckin below 140 GeV are not considered, since they do not contribute significantly to the SUSY signal region.

The vector boson samples are generated in association with a quark or a gluon. This leads to a more efficient generation, since the filter removes events with less than two jets:

- "pysubs msub 15 1" and "pysubs msub 30 1" for Z,

---

1) MC@NLO is used to generate top quark pair production.

2) J4, J5 and J6 samples are produced with a lower luminosity given the large cross section.

Table 1: Properties of PYTHIA samples in the signal region, including the production cross section, the filter efficiency, the integrated luminosity of the fully simulated sample and the number of events passing the filter. No filters are applied to J7 and J8, since the filter efficiencies are close to 100 %.

process and (CSC ID)	cross section [pb]	$\varepsilon_{\text{filter}}$ [%]	L [pb^{-1}]	N_{events} [k]
J4 (8090)	3.16 10^5	0.29	10.9	10
J5 (8091)	1.25 10^4	2.85	140	50
J6 (8092)	344	19.60	148	10
J7 (8093)	5.3	(63.00)	943	5
J8 (8094)	2.21 10^{-2}	(89.00)	2.3 10^5	5
$W \to e\nu$ (8270)	343	14.30	1020	50
$W \to \mu\nu$ (8271)	343	8.35	1050	30
$W \to \tau\nu$ (8272)	343	16.30	990	55
$Z \to \nu\nu$ (8190)	246	16.80	968	40
$Z \to \tau\tau$ (8191)	46.3	9.72	1112	5

Table 2: Properties of PYTHIA samples in the background region, including the production cross section, the filter efficiency, the integrated luminosity of the fully simulated sample and the number of events passing the filter.

process and (CSC ID)	cross section [pb]	$\varepsilon_{\text{filter}}$ [%]	L [pb^{-1}]	N_{events} [k]
$W \to e\nu$ (8273)	3.73 10^4	0.79	86	25
$W \to \mu\nu$ (8274)	3.73 10^4	0.22	122	10
$W \to \tau\nu$ (8275)	3.73 10^4	0.41	99	15
$Z \to \nu\nu$ (8192)	2.16 10^4	0.27	87	5
$Z \to \tau\tau$ (8193)	9.88 10^3	0.35	143	5

- “pysubs msub 16 1” and “pysubs msub 31 1” for W

The missing transverse energy in the vector boson samples originates directly or indirectly from the p_T of the boson, which justifies the following generation cut:

- “pysubs ckin 3 80.0”

2.1.2 Data-driven background estimation

In order to estimate the SUSY background from data, one typically identifies signal-free regions, topologically similar to the signal region and possibly close in phase space and extrapolates the background to the signal region. The region of low missing transverse energy can be used for this. To study these background estimation methods the vector boson samples are generated without the missing transverse energy filter and without the “pysubs ckin 3 80.0” cut, see Table 2.

QCD multijet samples without a cut on the missing transverse energy are produced centrally in the CSC effort, see CSC ID 5013-5017.

The background originating from $Z \to \nu\nu$ can also be estimated from $Z \to ee$ or $Z \to \mu\mu$ events by replacing the charged leptons by neutrinos. $Z \to ee$ or $Z \to \mu\mu$ samples are generated without the missing transverse energy filter but with the “pysubs ckin 3 80.0” cut, see Table 3.

Table 3: Properties of PYTHIA samples used to estimate the background from $Z \to \nu\nu$, including the production cross section, the filter efficiency, the integrated luminosity of the fully simulated sample and the number of events passing the filter. No filter is applied to $Z \to ee$, since the filter efficiency is close to 100 %.

process and (CSC ID)	cross section [pb]	$\varepsilon_{\text{filter}}$ [%]	L [pb^{-1}]	N_{events} [k]
$Z \to ee$ (8194)	46.2	(92.20)	108	5
$Z \to \mu\mu$ (8195)	46.4	20.70	521	5

2.2 Standard Model Background: Alpgen

2.3 mSUGRA Signal Monte Carlo

In later sections of this note we use a large mSUGRA Monte Carlo event sample to investigate the ATLAS discovery reach in terms of the mSUGRA parameters m_0 and $m_{1/2}$ (the universal scalar and gaugino mass at the GUT scale), as well as in terms of the weak-scale squark and gaugino masses. The mSUGRA Monte Carlo sample was produced privately by the authors, but using publically available and validated scripts for each stage of production.

We started with 315 (fixme: actually we made 1500 points for 5 tan beta values - comment on that) sets of mSUGRA input parameters:

- $m_0 = \{0, 100, ...1900, 2000 \text{ GeV}/c^2\}$,
- $m_{1/2} = \{0, 100, ..., 1300, 1400 \text{ GeV}/c^2\}$,
- $\tan\beta = 30$, $sgn(\mu) = +1, A_0 = 0$,

where $\tan\beta$ is related to the Higgs sector via $\tan\beta =< H_2 > / < H_1 >$ ($< H_2 >$ gives mass to the up quark and H_1 gives mass to the down quark), $sgn(\mu)$ is a CP violating phase in the Higgs mixing term $\mu_0 H_1 H_2$ of the Lagrangian, and A_0 is the universal trilinear coupling at the GUT scale.

For each set of input parameters, we used ISAJET 7.74 to run the Renormalization Group Equations (RGEs) down to the electro-weak scale, to obtain SUSY particle masses and couplings there. ISAJET reported success and a valid weak-scale solution for 287 of the 315 input parameter sets. See caption of Figure 1 for further details. For the 287 “good mSUGRA points” we saved the SUSY Les Houches Accord (SLHA) file produced by ISAJET (fixme: mention what’s in it. Spectrum + Decay table. Are we using the decay table, or is Pythia re-generating it?). (fixme: comment on the hack we had to make to ISAJET to get convergence for some of the points.)

For each of the 287 good mSUGRA model we generated 5000 events with Pythia 6.4 under Athena 12.0.31. The low energy spectrum (and decay table???) for each model was specified by reading in the SLHA file from ISAJET with the Pythia option IMSS(1)=11. (We have received numerous inquieries as to how we did this. The ability to specify the SLHA input file was added to the ATLAS Pythia wrapper Pythia_i at the author’s request: the relevant parameter of Pythia_i is “SusyInputFile”.) We subsequently ran the generated events through the fast simulation package ATLFAST (fixme: version, reference), using the official production transform csc_atlfast_trf.py from AtlasProduction 12.0.6.1. We then ran the resulting atlfast AOD through HPTV/SUSYView (version of athena, version of group area, reference) to make standardized ntuples for analysis in root.

All files from this production are available to ATLAS collaboration members. They can be found on the lxplus linux cluster and on the castor storage system at CERN. For details, see our production status webpage at http://phyweb.lbl.gov/atlaswiki/index.php?title=LBLPrivateProduction.

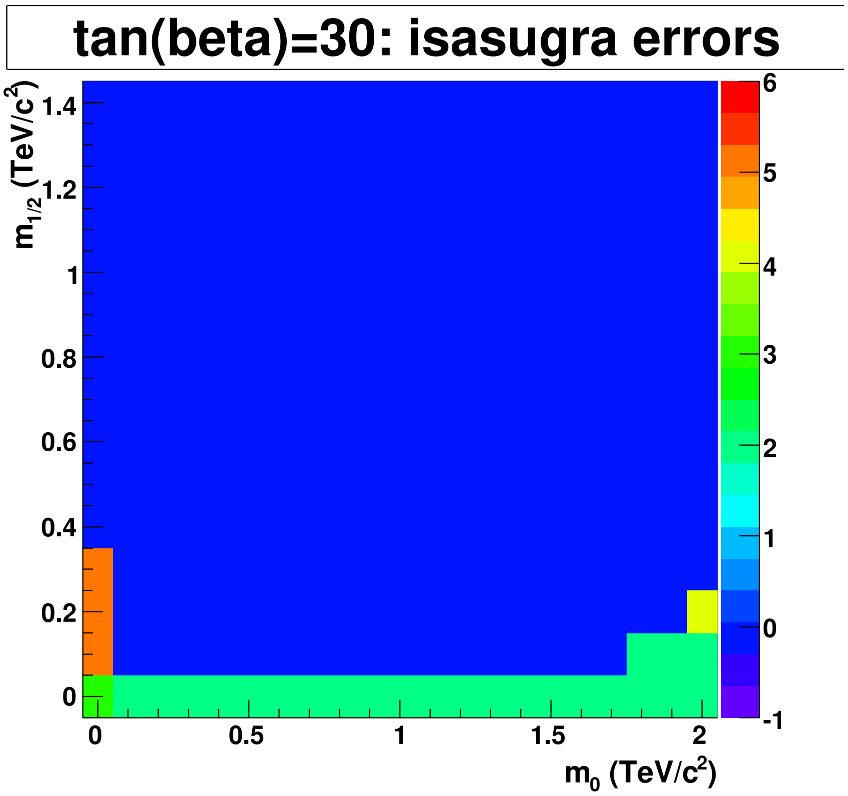

Figure 1: ISAJET status versus m_0 and $m_{1/2}$ for 315 mSUGRA models with $\tan\beta = 30, sgn(\mu) = +1, A0 = 0$. For the blue points in the center of the plot ISAJET converged on a valid low energy sparticle spectrum. For the remaining 28 points near the edge of the plot a valid low energy spectrum could not be obtained. The color/number shown for those points encode the error message returned by ISAJET: 2 (light green) = "NO EW SYM. BREAKING!", 3 (green) = "$M(H_P)^2 < 0$", 4 (yellow) = "NO RGE SOLUTION FOUND", 5 (red) = "TACHYONIC PARTICLES".

Table 4: Triggers considered for SUSY analyses, unprescaled at 10^{31} cm^{-2}s^{-1}, where PT stands for Passthrough.

Level 1		Level 2		Event Filter	
Name	Rate [Hz]	Name	Rate [Hz]	Name	Rate [Hz]
L1_J100 (thehold?)	8.9	PT	8.9	PT	8.9
L1_4J23 (thehold?)	17	PT	17	PT	17
L1_TE380	7.2	PT	7.2	PT	7.2
L1_JE256	10	–	10	–	10
L1_XE52	10.25	–	10.25	–	10.25
L1_XE72	small	–	small	–	small
L1_J70_XE30	50?	–	50?	–	50?
L1_2J42_XE30	10	–	10	–	10

3 Detector performance in SUSY events

3.1 jets, met, electrons, muons

3.2 comparison of fast and full simulation

3.3 fast simulation fudge factors

3.4 trigger efficiencies, trigger strategy

The first step in a SUSY analysis, just like in any other analysis, is to make sure that the events one is interested in are all recorded. In other words a set of triggers needs to be defined which efficienctly selects signal events while keeping the rate at a manageable level. In particular, SUSY triggers need to be unprescaled.

As mentioned earlier, SUSY events are characterised by:

- jets with high transverse energy,
- many jets,
- large transverse energy deposition in the calorimeter,
- large missing transverse energy,
- optionally one or more leptons.

For each of these characteristics or any combination a corresponding trigger can be defined. The focus is here on an initial instantaneous luminosity of 10^{31} cm^{-2}s^{-1}. Preliminary trigger menus exist for this instantaneous luminosity. Table 4 shows a list of the most promising unprescaled triggers for SUSY events.

A trigger emulation exists for the triggers defined in Table 4 and fully simulated events. However, trigger information does not exist for ATLfast simulation. In the following a procedure is defined that is used to parameterise trigger efficiencies on fully simulated samples and apply them to the ATLfast samples. This allows to simulate the effect of the chosen triggers on the grid of mSUGRA points, see Section 2.3.

Table 5: Available and self-made triggers in the CSC Monte Carlo samples.

Level 1	
Name	Threshold [GeV]
L1_J170	80
L1_J300	150
L1_4J45	20
L1_XE50	50
L1_XE100	100
L1_J_115	115
L1_4J_25	25
L1_XE70	70
L1_TE380	380
L1_TE500	500
L1_TE600	600
L1_JE256	256
L1_JE350	350

3.5 Parameterisation of trigger efficiencies

Single object efficiencies are measured for individual trigger conditions at Level 1, Level 2 and the Event Filter using fully simulated Monte Carlo events. The efficiencies are parameterised as a function of the relevant kinematic variables. These parameterisations can then be folded into the ATLfast simulated events.

In principle the trigger efficienies can be derived by a (data-driven) method that could be also applied to data by using reference triggers which allow to use events that do not pass the trigger under consideration. However, it is not the focus of this study to derive trigger efficiecnies from the data. Instead the method is based on a Monta Carlo sample where no trigger selection has been applied, i.e. a zero-bias trigger:

- Find the reconstructed physics object, e.g. a jet (this is the denominator).
- Find a trigger object with the corresponding trigger requirement (e.g. L1 jet p_T cut) that matches the reconstructed object (numerator). A match for a jet object is defined as $\Delta R < 0.3$, see Fig. **??**.
- The ratio (numerator / denominator) is an estimate of the trigger efficiency.
- Parameterise the ratio as a function of the the relevant kinematic variables, typically: p_T , η, ϕ.
- For higher level triggers it is required that the all lower level trigger requirements are fulfilled.

The trigger is emulated by assigning to each physics object a probability to pass the trigger as given by the corresponding value of the parameterisations. For each object a random number is generated. The object trigger requirement is fulfilled if the number is smaller than the assigned probablity. Event triggers are built from the object triggers.

3.6 Sanity checks

The method described in the previous section is validated by comparing event trigger efficiencies from the method with the ones given by the Monte Carlo trigger emulation. Fig. **??**.

A possible dependence of the trigger efficiency on the Monte Carlo sample is studied. Fig. **??**.

The trigger efficiency is parameterised as a function of kinematic quantities of fully simulated events. However, the parameterisations are applied to ATLfast simulated events. Any deficiency of the ATLfast simulation to describe the full simulation (e.g. in the energy scale or the resolution) might lead to an improper calculation of the trigger efficiency. It is, however, desired to separate issues related to the trigger and the simulation. If necesary, corrections will be applied to the ATLfast objects to better describe the full simulation.

4 Optimizing the event selection

4.1 different figures of merit

4.2 final states: 0 lepton, 1 lepton, lepton blind, (2 lepton later)

4.3 combining final states / multiple cut sets

5 Systematic Uncertainties

5.1 Standard Model Cross Sections

5.2 Detector Performance

5.3 Other Systematics

6 Summary and conclusion