Top background to Supersymmetry searches in dilepton events

Timothy Brooks (brooks@cern.ch)
Royal Holloway, University of London

Abstract

Searches for Supersymmetry with two lepton final states are limited by backgrounds from events with top quark pairs. Top events mimic many of the characteristics of events involving supersymmetric particles including leptons, jets and missing energy. The discovery significance of a search, then, must account for the top contribution passing its selection. In addition, uncertainty in the size of this contribution degrades the significance of a search. Presented here is a method to measure the top background from data with a smaller uncertainty than estimates from Monte Carlo simulations.

Selection variables

In searching for events characteristic of Supersymmetry, variables that strongly descriminate against standard model proccesses are defined. The event variables used in this analysis are as follows:

- Missing transverse energy, $E_T^{miss} \equiv \left|\sum_i^N -\vec{E}_{T,i}\right|$
- Effective mass, $M_{eff} \equiv \sum_{i=1}^{4} P_T^{jet,i} + \sum_{i=1} P_T^{lep,i} + E_T^{miss}$
- Transverse sphericity, $S_T \equiv \frac{2\lambda_2}{(\lambda_1+\lambda_2)}$ where λ_1 & λ_2 are eigenvalues of the sphericity tensor (see Ref.[2]).

Events involving supersymmetric particles are expected to yeild high effective masses, significant missing energy and have uniform sphericities up to 1. In addition to these variables, multiplicities and transverse momenta (P^T) for individual particles detected in events are part of selection. Below is a selection criteria for events with two leptons.

Sample selection

Fig. 3 is an example final state that this search would be sensitive to. The final state leaves 2 leptons, 2 light quarks and missing energy in the form of neutrinos and the lightest supersymmetric particle (LSP), in this model, the neutralino. Reference [1] demonstrates searches for events including 0, 1 and 2 leptons in the final state. The search in this analysis, with 2 oppositely signed leptons, is further broken down into 3 selections for events with at least 2, 3 or 4 jets. These selections are defined as follows:

1. Exactly two leptons with $P_T > 10$ GeV and opposite charge.
2. $\geq [2, 3, 4]$ jets with $P_T > [50, 40, 40]$ GeV and a leading jet $P_T > [180, 100, 100]$ GeV
3. Missing transverse energy $E_T^{miss} > 80$ GeV.
4. $\Delta\phi(jet_i, E_T^{miss}) > 0.2$ for the $[2, 3, 3]$ leading jets.
5. $E_T^{miss} > [0.3, 0.25, 0.2] \times M_{eff}$.
6. Transverse sphericity, $S_T > 0.2$.

The distribution of transverse sphericity before the last cut is shown in Fig. 1 and Fig. 2. This shows the higher purity in the selected region to the right. Using these cuts, we can predict the expected number of signal and background events from Monte Carlo.

Monte Carlo results

Table 1 shows signal and background expectation values from Monte Carlo analysis normalised to one inverse femtobarn. These values are consistent with the findings of [1]. Samples analysed at 7 TeV with full GEANT4 ATLAS detector simulation and reconstruction:

- SUSY benchmark point 4 (JIMMY) 50,000 events.
- $t\bar{t}$ produced with (POWHEG + PITHIA) 200,000 events.
- WZ + jets produced with (ALPGEN + JIMMY) 35,000 events.

	2jet	3jet	4jet
Signal (SU4)	38.82	112.74	90.76
Background ($t\bar{t}$)	3.48	37.89	20.90
ATLAS work in progress			

Table 1: Signal and background values for 2, 3 & 4 jet analyses.

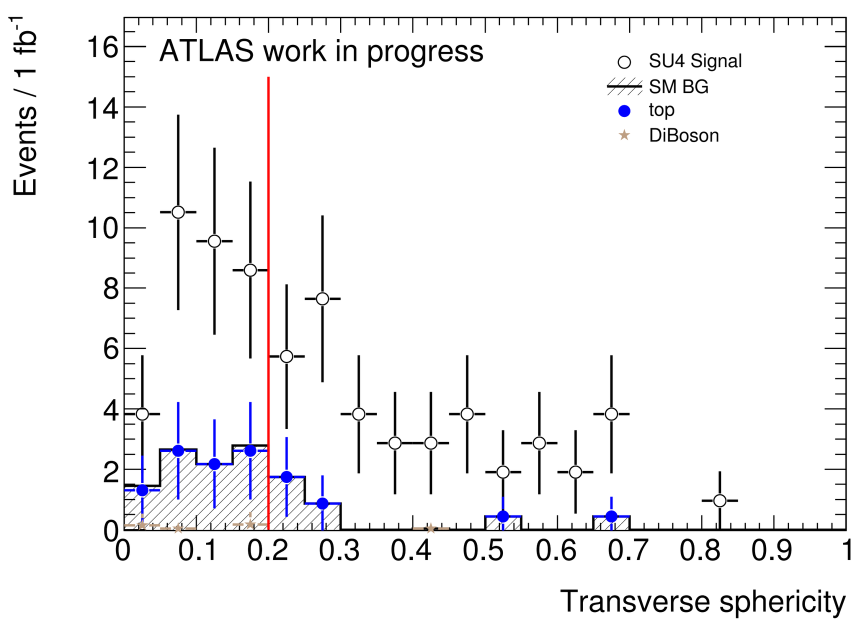

Figure 1: Transverse sphericity distribution after 2 jet selection.

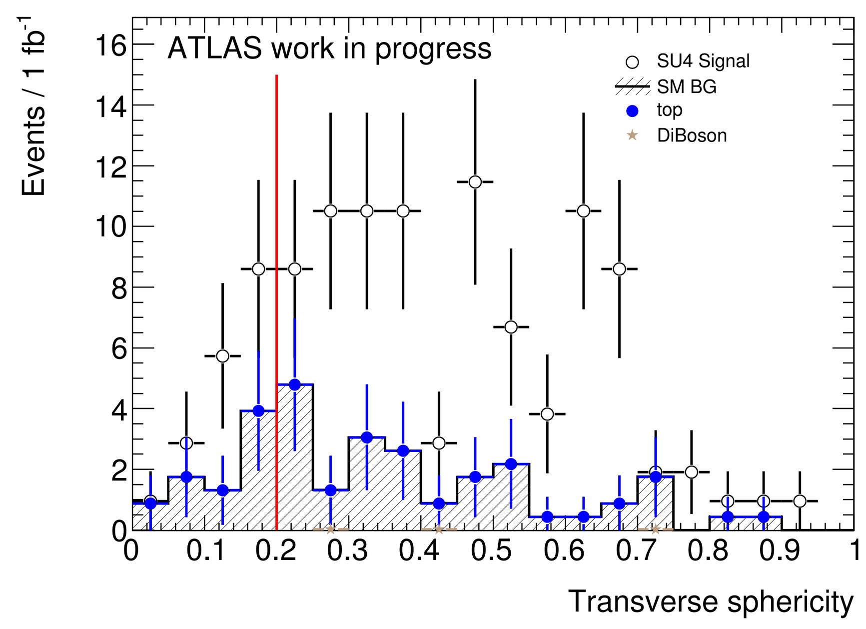

Figure 2: Transverse sphericity distribution after 4 jet selection.

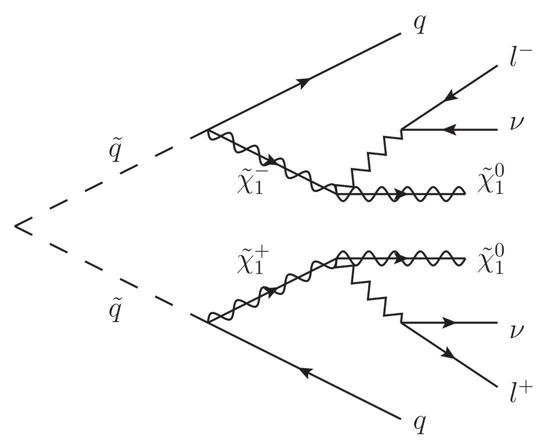

Figure 3: A system of supersymmetric decays with 2 leptons and 2 jets in the final state.

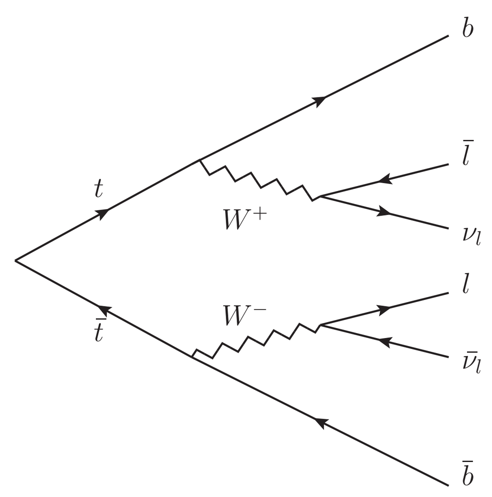

Figure 4: A di-leptonic $t\bar{t}$ decay with 2 leptons and 2 jets in the final state.

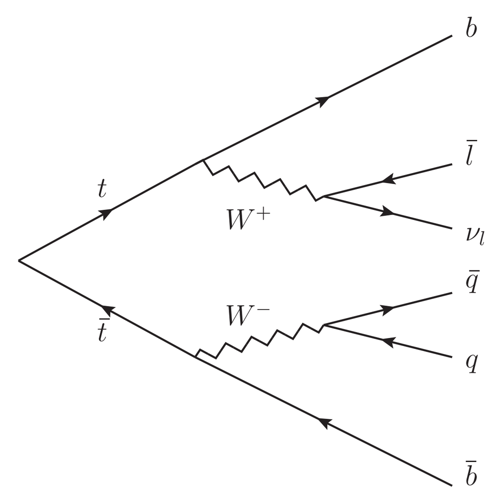

Figure 5: A semi-leptonic $t\bar{t}$ decay with 1 lepton and 4 jets in the final state.

Control sample

The absolute number of events carries a large uncertainty due to our models of the processes. This section outlines how we can make these selections, but use data to provide background estimates resulting in lower uncertainties.

The example Feynmann diagram in Fig. 3 has a final state of 2 leptons, 2 light quarks and missing energy in the form of neutrinos and the lightest supersymmetric particle (LSP), in this model, the neutralino. For a selection optimised for events like these, we can estimate the top background by forming a control sample from a region where the background is strongly enhanced. For a search channel of N jets and M leptons, we can look into a channel with $N+2$ jets and $M-1$ leptons. In top decay systems, this selects events where a W boson from a top decays leptonically in the search channel, and hadronically in a corresponding event in the control region.

In the $N = 2, M = 2$ search channel, $t\bar{t}$ with a final state such as Fig. 4 is the main background to the search. The control sample can be formed with a selection for events where $N = 4, M = 1$. This region would be composed mostly of events with the topology of Fig. 5.

For the events in the control sample, we can simulate the kinematic variables for top events in the signal region by attempting to identify a pair of jets consistent with coming from a W. Replacing these with a simulated lepton and neutrino, we should reproduce a sample consistent with the background of the signal region.

Analysis strategy

After preparing a search and a control sample, we must relate the number of events in the control sample to the number expected under a background only hypothesis. A ratio can be formed between the samples as

$$\tau = \frac{n\left(t\bar{t} \text{ accepted in } N+2 \text{ jet, } M+1 \text{ lepton selection}\right)}{n\left(t\bar{t} \text{ accepted in } N \text{ jet, } M \text{ lepton selection}\right)}. \quad (1)$$

Having estimated τ from data, the observation of experimental data can be modelled with Poisson statistics. The control region m has expectation value,

$$E[m] = \tau b, \quad (2)$$

where b is the background to the search selection and τ is per Eq.1. Using this value of b, the signal expectation is;

$$E[n] = \mu s + b. \quad (3)$$

Observing μ with a positive value indicates a discovery of beyond the standard model physics. Uncertainty from the absolute normalisation of the Monte Carlo is removed, only the relative uncertainty and the statistical error impact the measurement. It is expected that the normalisation of the Monte Carlo has a larger uncertainty than the ratio formed in Eq.1 such that this analysis will achieve a higher significance, or make an earlier discovery.

Multivariate techniques

In addition to improving the estimate of backgrounds to the search, the selection criteria can be optimised to increase signal to background ratios in the selections. Multivariate techniques combine variables from events and output one test statistic indicating to what degree the event appears like signal or background. This can create an selection criteria with greater significance than cuts on individual variables.

By comparing the output distribution of a multivariate classifier under some top enhancement cuts and some Supersymmetry enhancing cuts, the model dependency the classifier may have formed in training on Monte Carlo events can be limited. The application of boosted decision trees for this role is under investigation to enhance this analysis.

References

[1] The ATLAS collaboration, *Prospects for Supersymmetry discovery based on inclusive searches at a 7 TeV centre-of-mass energy with the ATLAS detector*, ATL-COM-PHYS-2010-336.

[2] The ATLAS collaboration, *Expected Performance of the ATLAS Experiment*, CERN-OPEN-2008-020 (pg1514+).