

Extracting backgrounds to SUSY searches from LHC data

Vadym Zhuravlov* on behalf of the ATLAS and CMS collaborations

Max-Planck-Institut für Physik

Föhringer Ring 6, 80805 München, Germany

E-mail: Vadym.Zhuravlov@cern.ch

Any discovery of new physics relies on detailed understanding of the Standard Model background. At the LHC, we expect to extract the backgrounds from the data itself, with minimum reliance on Monte Carlo simulations. We describe new developments in ATLAS and CMS on such data-driven techniques, and prospects for their application on first data.

*The 2009 Europhysics Conference on High Energy Physics,
July 16 - 22 2009
Krakow, Poland*

*Speaker.

1. Introduction

High energy jets, missing transverse energy (E_T^{miss}) and possibly leptons are the typical signature of R-parity conserved SUSY events at the LHC. The observation of an excess of events with these signatures with respect to the Standard Model prediction may manifest the presence of SUSY. Due to the poor knowledge of Standard Model cross-sections, parton distribution functions, underlying event and parton showering at the LHC energy scale, as well as insufficient knowledge of the detector itself, a reliable prediction of the Standard Model backgrounds should be derived mainly from the experimental data with reduced reliance on Monte Carlo simulations. In ATLAS and CMS the searches for a SUSY signal in events with a multijet and E_T^{miss} signature are classified according to the number of reconstructed leptons (electrons or muons). In this note we present several data-driven methods recently developed by the ATLAS and CMS collaborations to estimate the Standard Model background in SUSY searches with veto on leptons (no-lepton mode) and with one lepton (one-lepton mode).

2. Invisible Z boson decay

The process of Z boson production followed by the decay $Z \rightarrow \nu\nu$ gives an essential contribution to the Standard Model background in the no-lepton mode. The conventional way to estimate it is based on a sample of $Z \rightarrow \ell\ell$ events ($\ell = e, \mu$) [1], [2]. Both reconstructed leptons in such events are removed, and their momenta are accounted for by the E_T^{miss} . This method suffers from the low statistics of $Z \rightarrow \ell\ell$ sample, mainly due to the low branching ratio of the $Z \rightarrow \ell\ell$ decay with respect to the invisible decay.

Another possibility is to estimate the contribution of $Z \rightarrow \nu\nu$ process from $W \rightarrow \mu\nu$ control sample [3]. The process of W boson production is rather similar to Z boson production, however the correction factor must be estimated from the theory and applied to the $W \rightarrow \mu\nu$ estimate. The $W \rightarrow \mu\nu$ sample is selected requiring one isolated muon in addition to the no-lepton mode selection cuts. The resulting sample has substantial contributions of $t\bar{t}$ and QCD events, in addition to $W \rightarrow \mu\nu$. The $t\bar{t}$ contribution can be estimated with a data-driven technique exploiting the multiplicity of b-tagged jets. QCD contribution is also estimable from the data by reversing the muon isolation cut. The statistical (systematic) uncertainty of the method is estimated to be 29% (23%) for an integrated luminosity of 100 pb^{-1} . The additional theoretical uncertainty is 8%.

Similar to Z and W bosons one can also use the prompt photon production in association with jets [3]. As in the previous technique, this method also requires a correction due to the difference in the structure of photon and Z boson couplings to quarks. It has been shown, that for photons with large transverse momenta ($p_T > 200 \text{ GeV}$) this correction is small. The photons with transverse momentum $p_T > 100 \text{ GeV}$ are selected by requiring an electron veto and combined tracker and calorimetric isolation. The main backgrounds in the resulting sample are the QCD events with secondary photons and the events with an electron faking a photon. Both contributions do not exceed 3% each and can be estimated by data-driven techniques. Figure 1 (left) shows the result of the estimation before the theory correction compared to the distribution of $Z \rightarrow \nu\nu$ events. The statistical uncertainty of the method is estimated to be 10%, the systematics is estimated to be 20% (for a luminosity of 100 pb^{-1}).

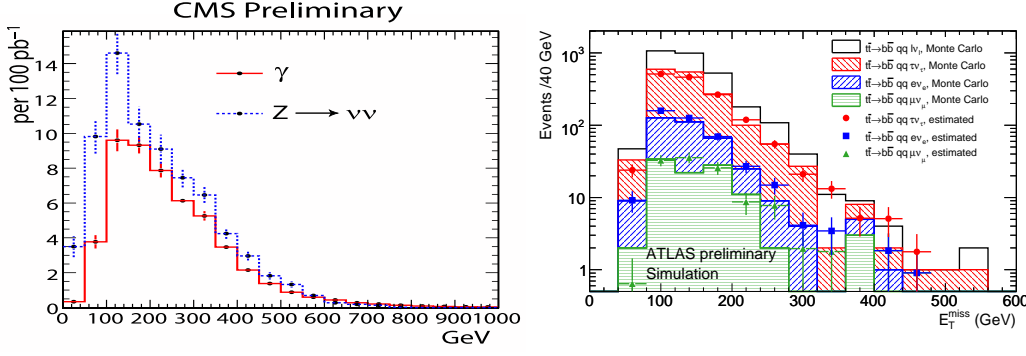


Figure 1: Left: distribution of E_T^{miss} in $Z \rightarrow \nu\nu$ events compared to an estimate derived from the prompt photon sample by correcting for the branching ratios and for the efficiency of the selection. The residual disagreement is mainly due to the difference in the photon and Z boson couplings to quarks and can be estimated from the theory. Right: distribution of E_T^{miss} of $t\bar{t} \rightarrow b\bar{b}l\nu_l q\bar{q}$ events (corresponding to an integrated luminosity of 1.6 fb^{-1}) passing the no-lepton mode selection (open histogram) and fractions of events with a tau lepton (left hatching), a lost electron (right hatching) and a lost muon (horizontal hatching), compared to the estimate of these fractions from the data-like sample.

3. $t\bar{t}$ background

Monte Carlo simulations show that the $t\bar{t}$ gives an important contribution to the Standard Model background in the no-lepton mode and the dominating contribution is the $t\bar{t} \rightarrow b\bar{b}l\nu_l q\bar{q}$ process mainly coming from the events with a tau lepton and events with a non-identified electron or muon. The method of estimation of these two contributions is based on the “seed sample” enriched by the $t\bar{t} \rightarrow b\bar{b}l\nu_l q\bar{q}$ events extracted from the data. The $t\bar{t} \rightarrow b\bar{b}\tau\nu_\tau q\bar{q}$ component is reproduced by replacing the electron or muon of the seed event with a tau lepton and simulating the tau lepton decay. The contribution of $t\bar{t} \rightarrow b\bar{b}l\nu_l q\bar{q}$ events with a non-identified electron is estimated by replacing the electron of the seed event by a jet reconstructed along with the electron. For the estimation of the $t\bar{t} \rightarrow b\bar{b}\mu\nu_\mu q\bar{q}$ component the muon in the seed event is considered as non-identified, but the probability of its contribution to the estimation of E_T^{miss} in the event is taken into account. Fig. 1 (right) shows the E_T^{miss} distribution of the various components of $t\bar{t}$ background compared to the data-driven estimates. The statistical (systematic) uncertainty of the method is estimated to be 8% (36%) for an integrated luminosity of 200 pb^{-1} .

4. Tiles method

The Tiles Method [4] is designed to estimate the total Standard Model background in the one-lepton mode and is based on the analysis of the two main discriminant variables in this mode, effective mass¹ m_{eff} and transverse mass² m_T . The $m_{\text{eff}}-m_T$ plane is divided in quadrants, called tiles. Each tile has contributions from the SUSY signal and from the background. The fractions

¹Effective mass m_{eff} is defined as a scalar sum of transverse momenta of the four leading jets, transverse momenta of all identified leptons and E_T^{miss} .

²The transverse mass m_T is defined by $m_T^2 = 2(|\mathbf{p}_T|E_T^{\text{miss}} - \mathbf{p}_T \cdot \mathbf{p}_T^{\text{miss}})$, where \mathbf{p}_T is the transverse momentum of the lepton and $\mathbf{p}_T^{\text{miss}}$ is the missing-transverse-energy two-vector.

of the Standard Model events entering each tile are estimated from the Monte Carlo. The yields of SUSY and background events are estimated by solving the system of equations

$$\begin{aligned}
 n_A &= f_A \cdot n_{SM} + SUSY_A \\
 n_B &= f_B \cdot n_{SM} + SUSY_B \\
 n_C &= f_C \cdot n_{SM} + SUSY_C \\
 n_D &= f_D \cdot n_{SM} + SUSY_D \\
 SUSY_A \cdot SUSY_D &= SUSY_B \cdot SUSY_C
 \end{aligned}
 \tag{4.1}$$

Here A, B, C and D denote four tiles, n_X is the measured number of events in the tile X , f_X is the fraction of Standard Model events in the tile X , n_{SM} is the unknown number of Standard Model events in all tiles and $SUSY_X$ is the unknown number of SUSY events in the tile X . The last equation comes from the assumption of no correlation between m_{eff} and m_T for signal events. The correlation of the variables for the Standard Model contribution is fully taken into account by the values of f_X . The method can be generalized to a $N \times N$ tiles configuration. In such a configuration the system of equations (4.1) is over-constrained and a solution can be found by minimizing the extended log-likelihood estimator constructed from the measured number of events in the tile and the expected number of events calculated as the sum of Standard Model and SUSY contributions. The statistical error of the method depends on the tiles configuration and can be optimized for a given luminosity. The optimal number of tiles for a luminosity of 1 fb^{-1} is found to be 8×8 . A configuration with many tiles has the advantage of the possibility to add new parameters to the fit, for example, a linear correlation factor, or separation of the various background contributions.

5. Conclusion

The $Z \rightarrow \nu\nu$ background in the no-lepton mode can be estimated from three different control samples, $Z \rightarrow \mu\mu$, $W \rightarrow \mu\nu$ and prompt photons. Consistency check between W, Z and photon samples provides an important element in the understanding of any new signal. The data-driven method of $t\bar{t}$ background estimation in the no-lepton mode based on the replacement technique is discussed. The Tiles method allows the explicit accounting of variable correlation and presence of SUSY in the one-lepton mode.

References

- [1] G. Aad, et al., *Data-Driven Determinations of W, Z and Top Backgrounds to Supersymmetry*, in *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, arXiv:0901.0512, CERN-OPEN-2008-020, 2009, pp. 1525–1561.
- [2] G. L. Bayatian, et al., *CMS physics: Technical design report*, CERN-LHCC-2006-001, 2006.
- [3] The CMS Collaboration, *Data-Driven Estimation of the Invisible Z Background to the SUSY MET plus Jets Search*, CMS PAS SUS-08-002, 2008.
- [4] The ATLAS Collaboration, *Background Estimation for Inclusive SUSY Searches - The Tiles Method*, ATLAS note ATL-PHYS-PUB-2009-077, 2009.
- [5] The ATLAS Collaboration, *Data-Driven Determination of $t\bar{t}$ Background to Supersymmetry Searches in ATLAS*, ATLAS note ATL-PHYS-PUB-2009-083, 2009.