Available online at www.sciencedirect.com
ScienceDirect
Nuclear Physics B Proceedings Supplement 00 (2014) 1–3
Nuclear Physics B Proceedings Supplement

A Model Independent General Search for new physics in ATLAS

S.Amoroso (on behalf of the ATLAS collaboration)[a]

[a] *Albert-Ludwigs Universität Freiburg, Hermann-Herdern Str.3, D-79102 Germany*

Abstract

We present results of a model-independent general search for new phenomena in proton-proton collisions at a centre-of-mass energy of 8 TeV with the ATLAS detector at the LHC. The data set corresponds to a total integrated luminosity of 20.3 fb^{-1}. Event topologies involving isolated electrons, photons and muons, as well as jets, including those identified as originating from b-quarks (b-jets) and missing transverse momentum are investigated. The events are subdivided according to their final states into exclusive event classes. For the 697 classes with a Standard Model expectation greater than 0.1 events, a search algorithm tests the compatibility of data against the Monte Carlo simulated background in three kinematic variables sensitive to new physics effects. No significant deviation is found in data. The number and size of the observed deviations follow the Standard Model expectation obtained from simulated pseudo-experiments.

Keywords: ATLAS experiment, New Physics, model independent searches

1. Introduction

The LHC proton–proton collision data recorded by the ATLAS experiment have been thoroughly analyzed for specific signals of physics beyond the Standard Model (SM), and stringent limits have been set on new physics models. Although these searches cover a wide variety of possible event topologies, they are not exhaustive. Events produced by new interactions or new particles might still be hidden in the data. The analysis described here extends these specific searches with a model-independent approach, designed to be comprehensive for new physics (NP) signals appearing at high transverse momenta (p_T) [1].

2. Event selection and Classification

The analysis uses multiple triggers: single lepton and photon, single or multijet and E_T^{miss}. To avoid double counting each event is taken from the trigger in which the loosest selection is applied.

Events are then subdivided into exclusive classes based on the number and types of objects reconstructed: electrons (e), muons (μ), photons (γ), jets (j), b-jets (b) and E_T^{miss} (ν) are considered. The subdivision can be regarded as a classification according to the most important features of the events. The p_T cuts applied on top of the trigger selection and the labels used for each object are summarized in Table 1.

Object	Label	Lower p_T cut
jet	j	50 GeV
b-jet	b	50 GeV
electron	e	25 GeV
muon	μ	25 GeV
photon	γ	40 GeV
E_T^{miss}	ν	150 GeV

Table 1: List of objects used for the event classification with their label and lower p_T requirement.

3. Background Estimation

The background estimate is obtained mostly from Monte Carlo simulation. All processes producing high

p_T objects with non negligible cross-section are considered for the SM estimate: inclusive jets production, $W/Z/\gamma$ production in association with jets, single top and top pair production including their association with jets and vector bosons, diboson, triboson and Higgs production.

Background contributions with exactly one fake lepton are determined from data with an ABCD method, using as variables the relative track isolation of the lepton and the track impact parameter significance, and reverting the requirements applied to signal leptons.

In classes containing only j and b the multijet MC samples are scaled to data with normalization factors, ranging from 0.4 to 1, derived separately in each exclusive jet multiplicity class (e.g. for the $1j2b$ the same normalization factor as for the $3j$ category is used). The simulated W/Z+jets and top pair samples are also reweighed to improve the modeling in classes containing ν, by reweighing the truth-level p_T distribution of the boson or top quark pair.

4. Systematic Uncertainties

The systematic uncertainties have both uncorrelated components and components which are correlated between event classes and within different bins of kinematic distributions in the same event class. In the search algorithm, all uncertainties are treated as uncorrelated. Correlations according to the correlated uncertainty components are, however, used for the generation of the pseudo-experiments. Experimental uncertainties are treated as correlated and the uncertainty due to the limited number of MC events is treated as uncorrelated in the pseudo-experiment generation. Theoretical uncertainties are assigned to represent typical uncertainties of the models and generators used. They are assigned per subprocess and subdivided into an uncorrelated and a correlated component. All theory uncertainties are treated as constant over the studied parameter space.

5. Search Algorithm

To quantify the level of agreement between the data and the SM expectation and to identify regions of deviations, we make use of a search algorithm to scan distributions in form of histograms. The algorithm has been applied to the effective mass m_{eff} (the scalar sum of the p_T of the objects defining the class, including the E_T^{miss}), the visible invariant mass and the E_T^{miss} distributions. The visible invariant mass (m_{inv}) is defined for each event class as the invariant mass calculated with all objects defining the class besides the E_T^{miss}. The m_{eff} scan is applied to all 697 event classes with at least 0.1 expected events, while event classes with a single object have been removed from the scan of m_{inv}. The E_T^{miss} scan is applied only to categories containing ν (250 in total). The bin size of the scanned distributions is chosen to reflect the expected resolution of each variable in a given class, with values ranging from 20 to 500 GeV.

In every histogram the number of data events N_{obs} and the expectation N_{SM} with its total systematic uncertainty δN_{SM} (obtained by summing linearly the absolute value of the uncertainty of each bin) are determined for each possible connected bin region made with a minimum of two bins. A statistical estimator p is used to judge which region is of most interest. It is derived from the convolution of a Poisson probability density function (pdf), to account for statistical uncertainties, with a Gaussian pdf, $G(b; N_{\text{SM}}, \delta N_{\text{SM}})$, with mean N_{SM} and width δN_{SM}, to include systematic uncertainties. It is defined as:

$$p = \begin{cases} A \int_0^\infty db\, G(b; N_{\text{SM}}, \delta N_{\text{SM}}) \sum_{i=N_{\text{obs}}}^{\infty} \frac{e^{-b} b^i}{i!} & \text{if } N_{\text{obs}} \geq N_{\text{SM}} \\ A \int_0^\infty db\, G(b; N_{\text{SM}}, \delta N_{\text{SM}}) \sum_{i=0}^{N_{\text{obs}}} \frac{e^{-b} b^i}{i!} & \text{if } N_{\text{obs}} < N_{\text{SM}}. \end{cases} \quad (1)$$

The factor A ensures that the pdf is normalized to unity. The value of p gives an estimate of the probability that the SM expectation fluctuates upwards or downwards with respect to the data in a given region. Such a method is able to find narrow resonances and single outstanding events as well as signals spread over large regions of phase space in distributions of any shape.

To avoid being sensitive to the effect of poor MC statistics, regions where the total background prediction has an uncertainty greater than 100% are discarded by the algorithm. If all regions in an event class have an uncertainty larger than 100% no region is selected, and a p-value of 1 is assigned to the class. Following this procedure we find 63 such event classes for the scan of m_{eff}, 76 for the m_{inv} scan and 42 for the E_T^{miss}scan.

6. Pseudo-experiments

The probability that a statistical fluctuation occurs somewhere in the event class distributions is modelled by pseudo-experiments. In this procedure, the data are replaced by pseudo-data which are generated according to the SM expectation.

We generated 2000 "pseudo ATLAS experiments", each consisting of the same event classes and distributions as considered in the data. The search algorithm is applied to each of these in the same way as for data.

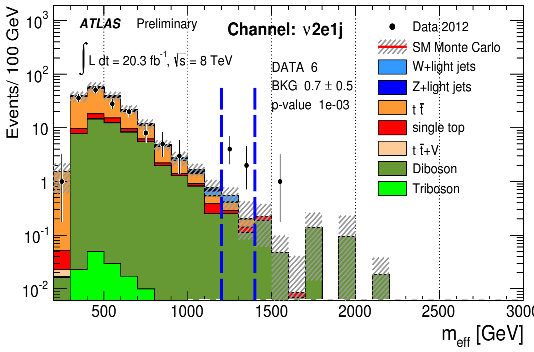

Figure 1: $m_{\rm eff}$ distribution for the event class with two electrons, one jet and $E_{\rm T}^{\rm miss}$ ($\nu 2e1j$). The dashed vertical lines indicate the region of interest which has the smallest p-value (0.0013) for this event class.

The p-value distributions of the "pseudo ATLAS experiment" and their statistical properties can be compared with the p-value distributions obtained from data. The effect of bin-by-bin correlations is taken into account in the generation of pseudo-experiments.

7. Results

After classification data events are found in 573 event classes. Agreement between data and the SM prediction is observed for most event classes. The number of classes with an SM expectation larger than 0.1 is 697. These classes are further considered for the statistical analysis. A total of 16 event classes have an SM expectation of less than 0.1 events, but at least one data event; two data events are found only in the $\nu 2e1j$ event class. To illustrate how the algorithm works Figure 1 shows the $m_{\rm eff}$ distribution for the event class with two electrons, one jet and $E_{\rm T}^{\rm miss}$; where the region of greatest deviation found by the search algorithm in this distribution is indicated with vertical lines. In Figure 2 we show the fraction of pseudo-experiments in which the scan finds at least one, two, or three event classes with deviations below a given p-value ($p_{\rm min}$). No event class is found in data with a local p-value below 10^{-4}, consistent with the expectation from pseudo-experiments. The largest deviation has a local p-value of $7 \cdot 10^{-4}$ and is found in the $m_{\rm eff}$ distribution of a class with one electron, one muon, one photon and two jets.

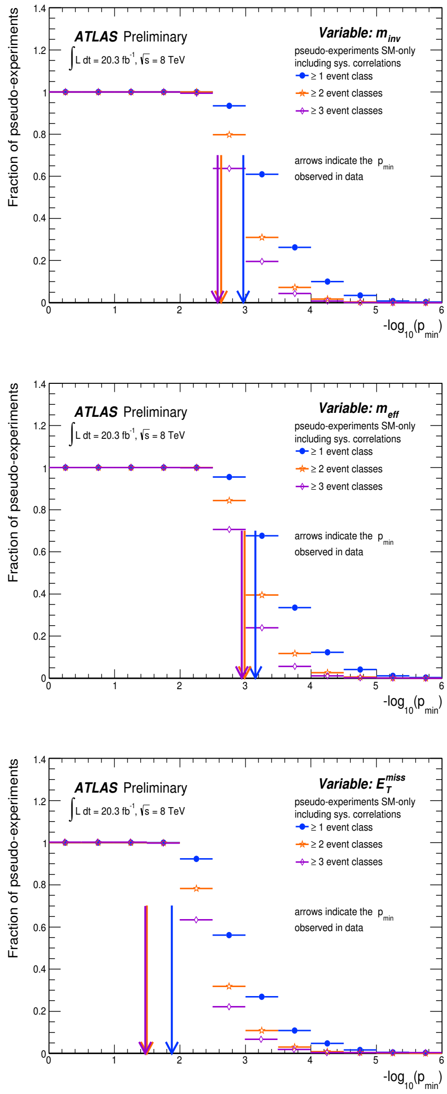

Figure 2: The fraction of pseudo-experiments which have at least one, two and three event classes in which the scan finds a deviation with a p-value below a given threshold ($p_{\rm min}$) in the scan of the visible invariant mass, effective mass and missing energy distributions (from top to bottom respectively). The values observed in data are indicated by the arrows. Pseudo-experiments are generated under the SM-only hypothesis.

References

[1] The ATLAS Collaboration, ATLAS-CONF-2014-006, A general search for new phenomena with the ATLAS detector in pp collisions at $\sqrt{s}$ = 8 TeV.