A general search with the ATLAS detector in *pp* collisions at 8 TeV

Fabio Cardillo

Helmholtz Alliance meeting „Physics at the Terascale“

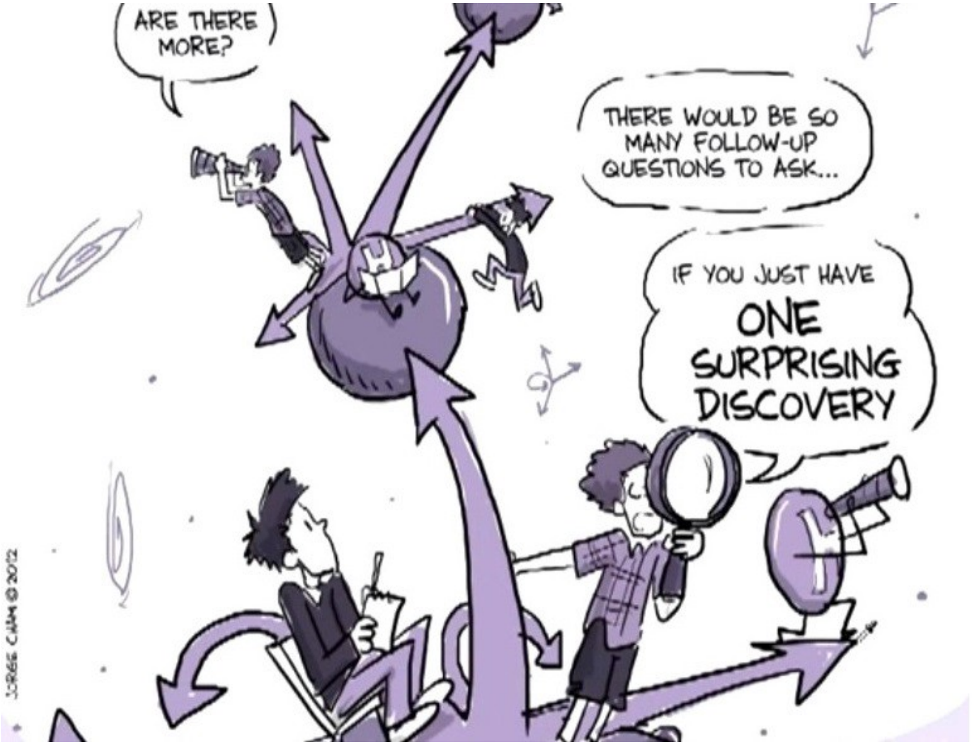
December 2, 2014

Introduction

Why are we doing Model Independent searches ?

- Is there new Physics at the Terascale ?
- Where to search ?

  Too many different theories to test.
- A model independent discovery tool is needed.
- We want to look everywhere.

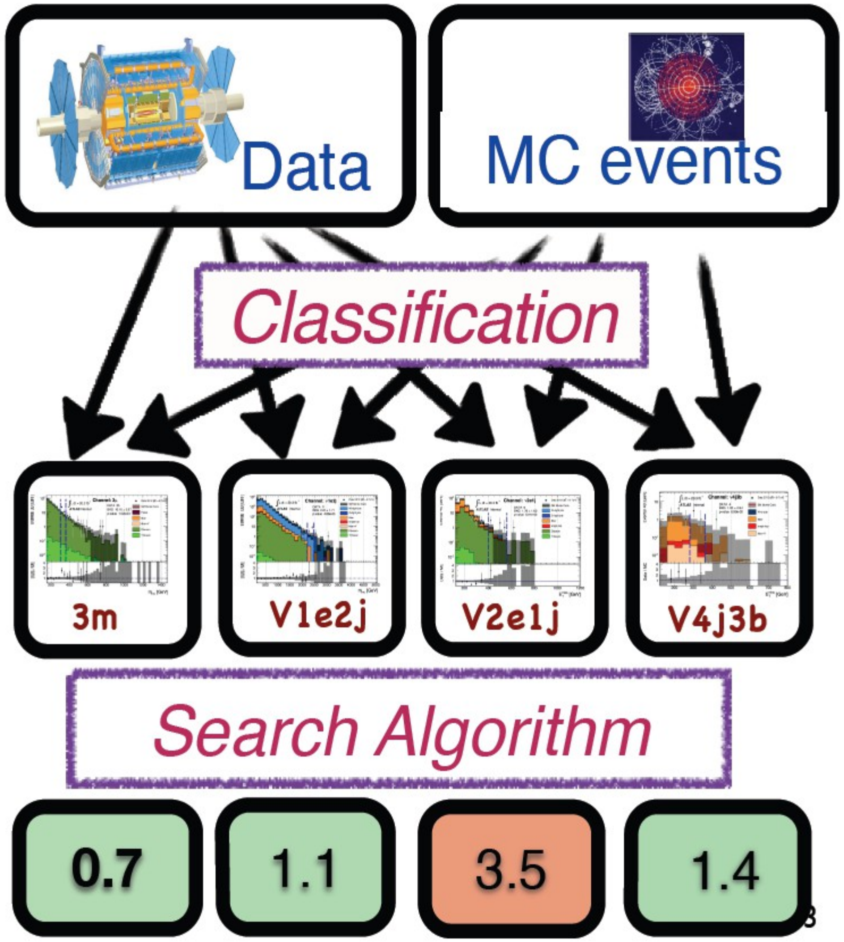


The aim is to reveal unexpected discrepancies between data and the SM expectation by looking in as many event topologies as possible.

Analysis overview

How does a Model Independent search work ?

- A search for deviations from the Standard Model prediction in as many as possible final states in a detector.
- Data can be subdivided in exclusive classes depending on the number and type of physics objects.
- A statistical search algorithm is used to find deviation between data and the SM prediction in kinematic distributions.
- Distributions of m_{eff}, m_{inv} and E_T^{miss} can be investigated.



Analysis overview

Classification:

- Data events are selected by asking a chain of several triggers.
- Events are assigned to exclusive classes depending on the number and type of high level objects reconstructed.
- This subdivision can be regarded as a classification according to the most important features in data.

μ	a tight isolated muon $p_T > 25$ GeV	b	a b-tagged jets with MV1 tag (70% OP)
e	a tight isolated electron $p_T > 25$ GeV	γ	a tight isolated photon $p_T > 40$ GeV
j	a jet with $p_T > 50$ GeV	ν	if $E_T^{miss} > 150$ GeV

Search algorithm:

- To quantify the level of agreement between the data and the SM expectation and to identify regions of deviations, we use a statistical search algorithm.
- The algorithm loops over all possible connected bin regions in the histograms, calculating for each of them a p-value, taking into account both statistical fluctuations and systematic uncertainties.
- The region of smallest p-value is taken as region of interest

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

Background estimation

- The background estimation is mainly Monte Carlo driven.
- All processes producing high p_T objects are considered for the SM estimate. The contributing theoretical and experimental uncertainties are taken into account.
- Re-weighting procedures are applied to some of the MC samples to improve the modeling of the SM background
- For some topologies, data driven ways have to be used to enhance the background modeling (*ABCD* method for low p_T leptons)
- The ex. uncertainties are always fully correlated. The theoretical uncertainty is composed of an uncorrelated and a fully correlated fraction to cover normalization uncertainties due to cross section errors.

Processes considered for the SM estimate
$\pm$ sys. uncertainty $(\sigma_{uncorr.} + \sigma_{corr.})$

- Z + jets $\pm 14\%_{uncorr}$ $\pm 5\%_{corr}$
- W + jets $\pm 14\%_{uncorr}$ $\pm 5\%_{corr}$
- Di-Boson $\pm 30\%_{uncorr}$ $\pm 10\%_{corr}$
- $t\bar{t}$ $\pm 19\%_{uncorr}$ $\pm 6\%_{corr}$
- $t\bar{t}$ + W/Z $\pm 30\%_{uncorr}$ $\pm 10\%_{corr}$
- SingleTop $\pm 19\%_{uncorr}$ $\pm 6\%_{corr}$
- SingleTop + Z $\pm 30\%_{uncorr}$ $\pm 10\%_{corr}$
- Multijet $\pm 30\%_{uncorr}$ $\pm 0\%_{corr}$
- γ + jets $\pm 30\%_{uncorr}$ $\pm 10\%_{corr}$
- Z + γ $\pm 50\%_{uncorr}$ $\pm 20\%_{corr}$
- W + γ $\pm 50\%_{uncorr}$ $\pm 20\%_{corr}$
- Tri-Boson $\pm 50\%_{uncorr}$ $\pm 20\%_{corr}$
- Higgs@125GeV $\pm 10\%_{uncorr}$ $\pm 10\%_{corr}$

8 TeV Results

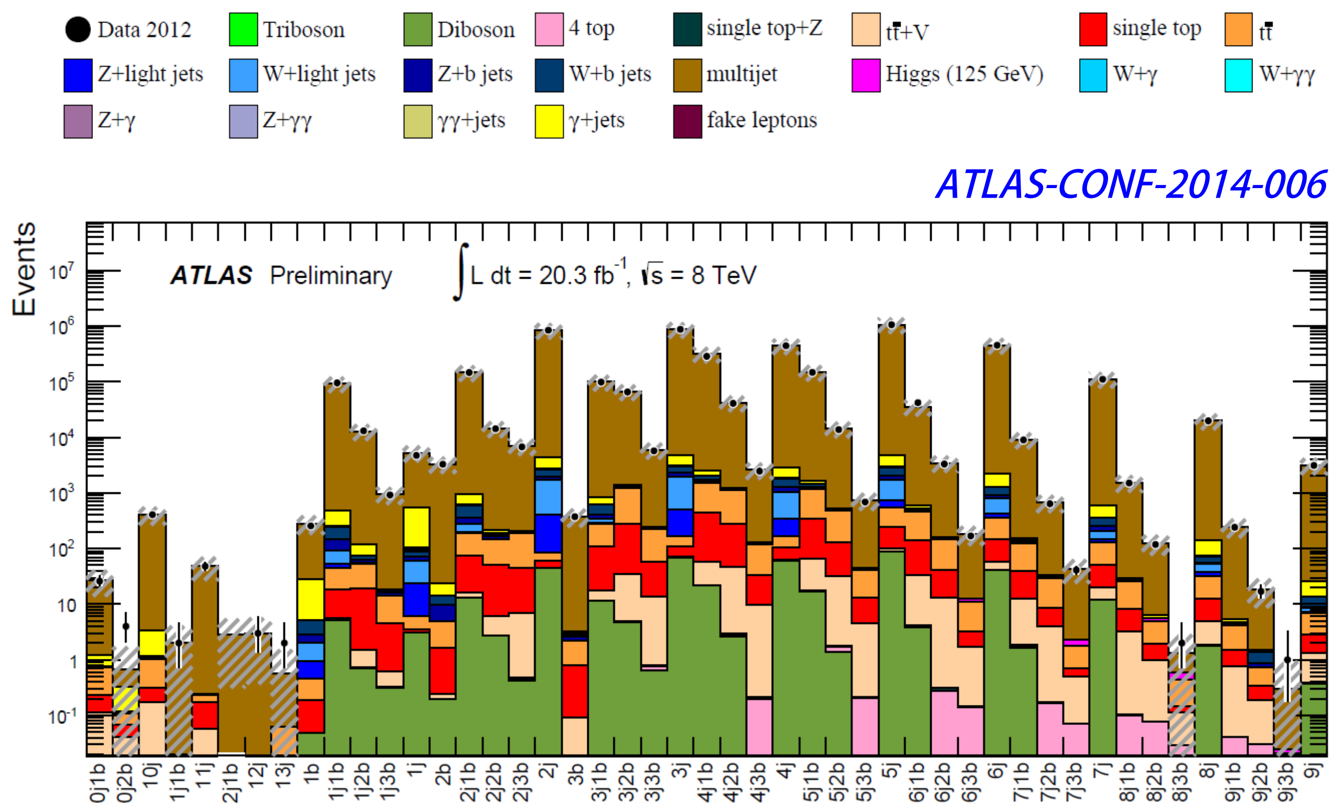
The results for 8 TeV of this analysis were published 2013 in ATLAS-CONF-2014-006
http://cds.cern.ch/record/1666536

- 713 different detector topologies have been found using this classification scheme.
- Data events are found in 573 event classes. The number of classes with a total SM expectation > 0.1 is 697.
- A total of 16 event classes have SM < 0.1 events, but at least one data event. These channels are investigated by eye but not considered for the scan.
- Some channels close to signal regions from other searches haven't been considered for the scan in order to prevent unblinding them.
- Removed channels with more than 3 *b*-tagged jets and channels with more than two photons because we were not able to model them correctly with Monte Carlo.

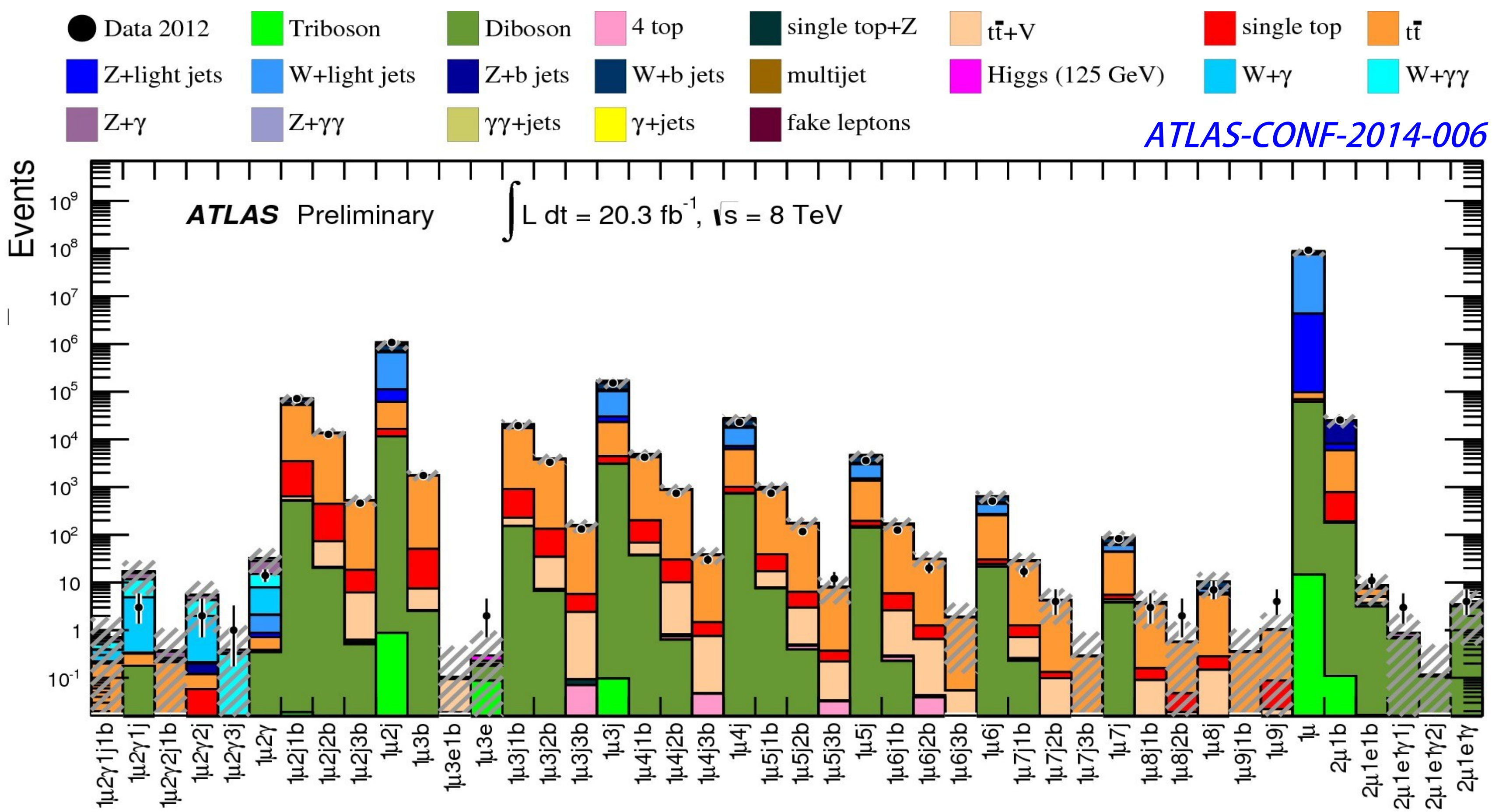
8 TeV Results: Global event yields

The global event yields of all channels can be compared to the background expectation.

The total number of data events for all exclusive classes which contain at least one event or a SM background > 0.1 are shown in the global plots presented.

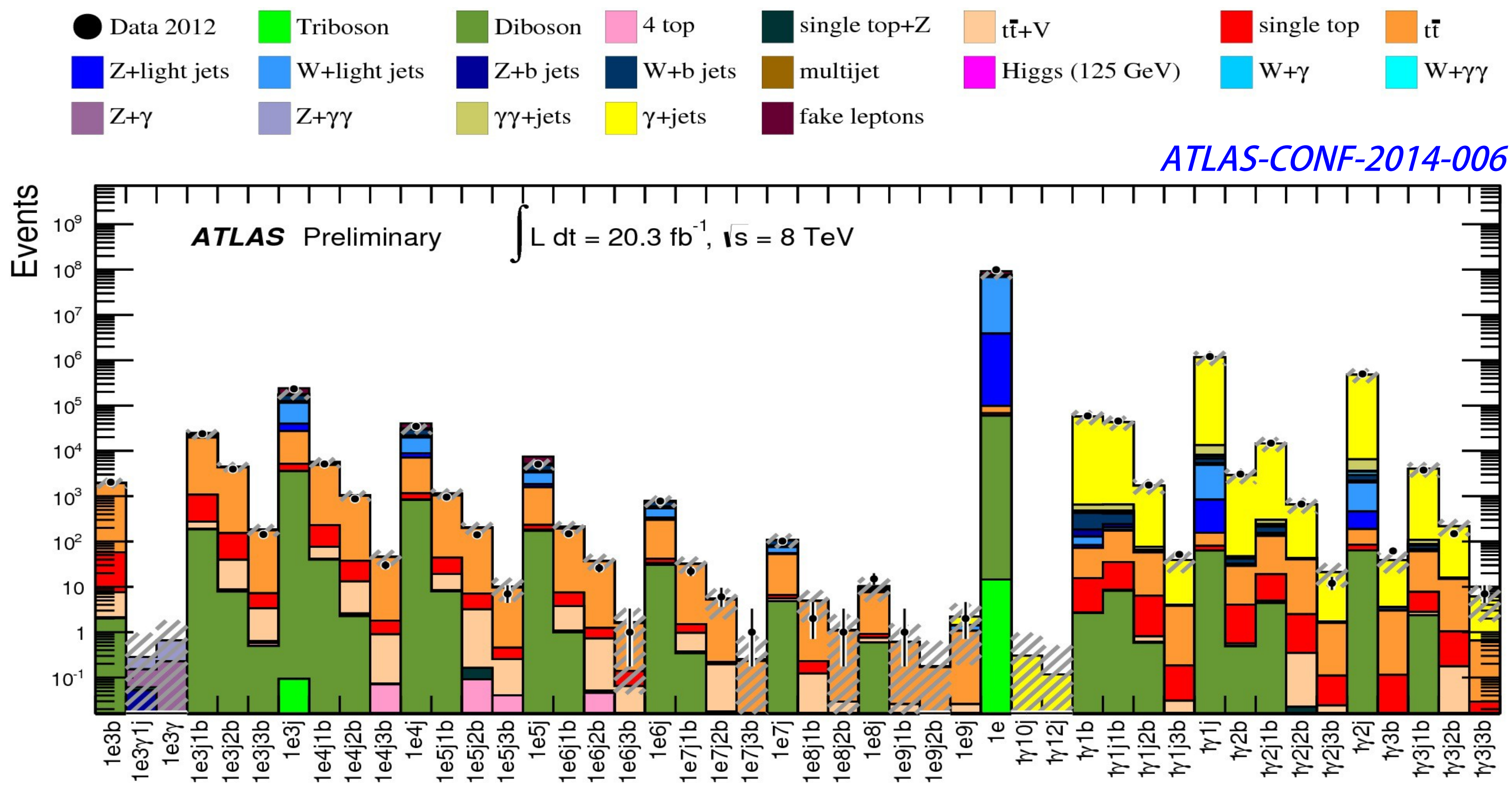


8 TeV Results: Global event yields



Total number of events for some channels containing muons

8 TeV Results: Global event yields



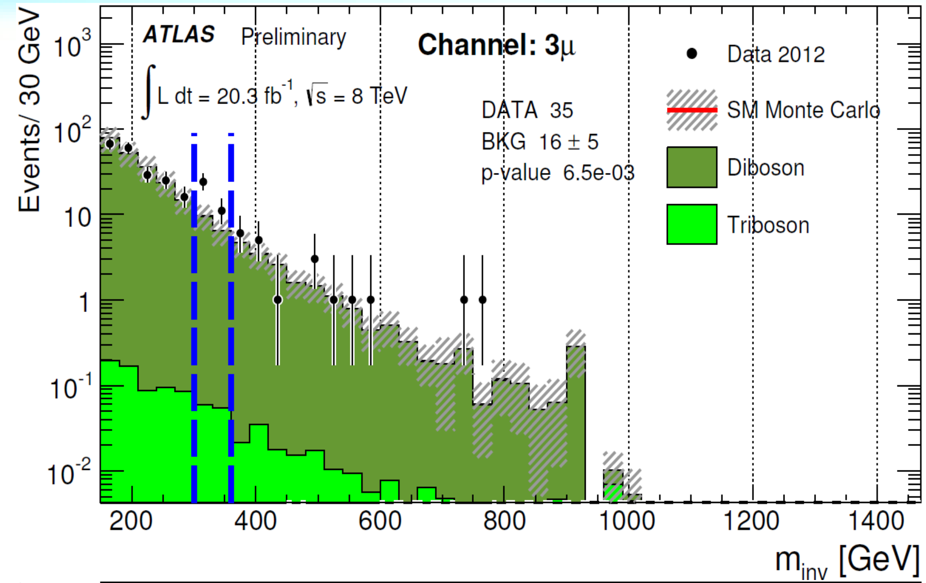
Total number of events for some channels containing electrons or photons

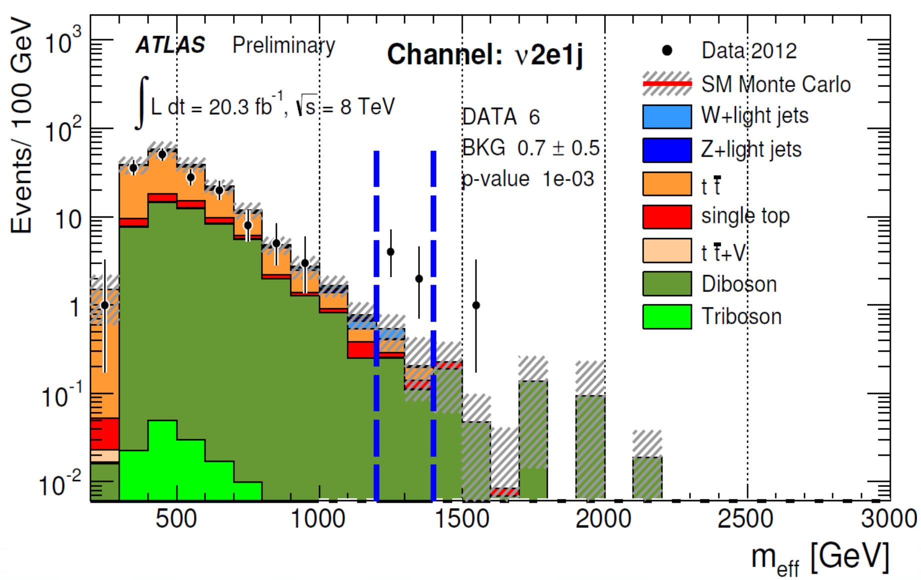
Scan results

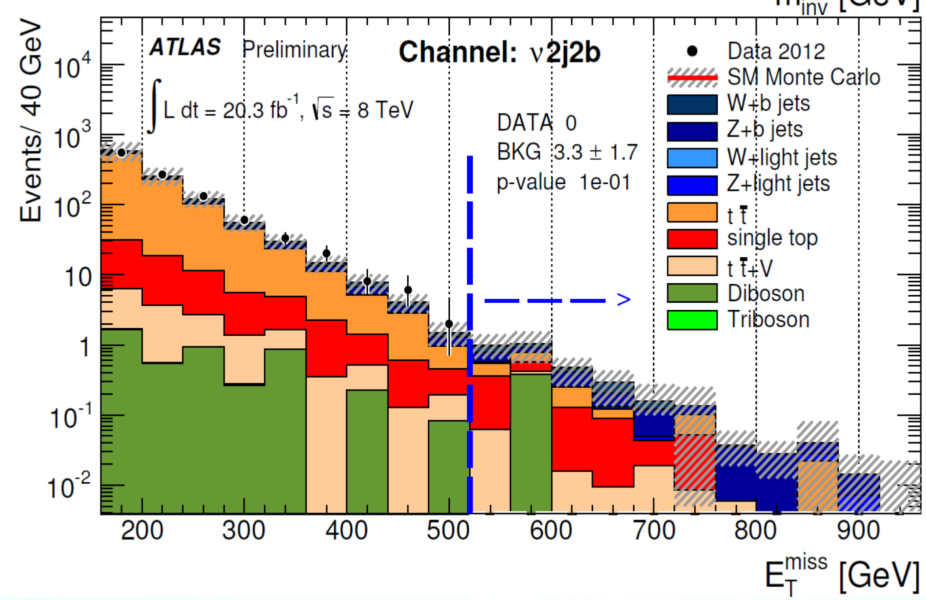
ATLAS-CONF-2014-006

Results from the scan algorithm:

- The blue dashed lines label the region of interest (smallest *p*-value) selected by the scan for each distribution in each class
- The algorithm is sensitive to excesses and deficits in data and was used on the m_{eff}, m_{inv} and E_T^{miss} distributions for all channels.



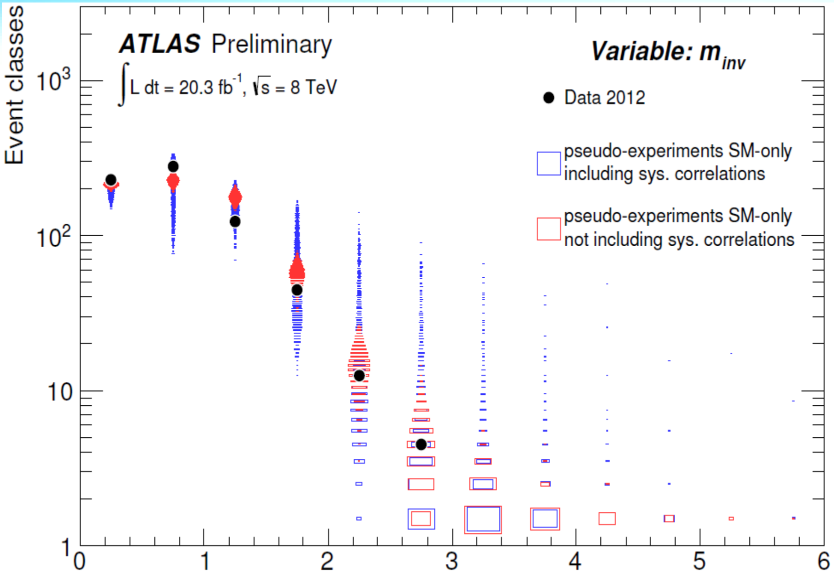


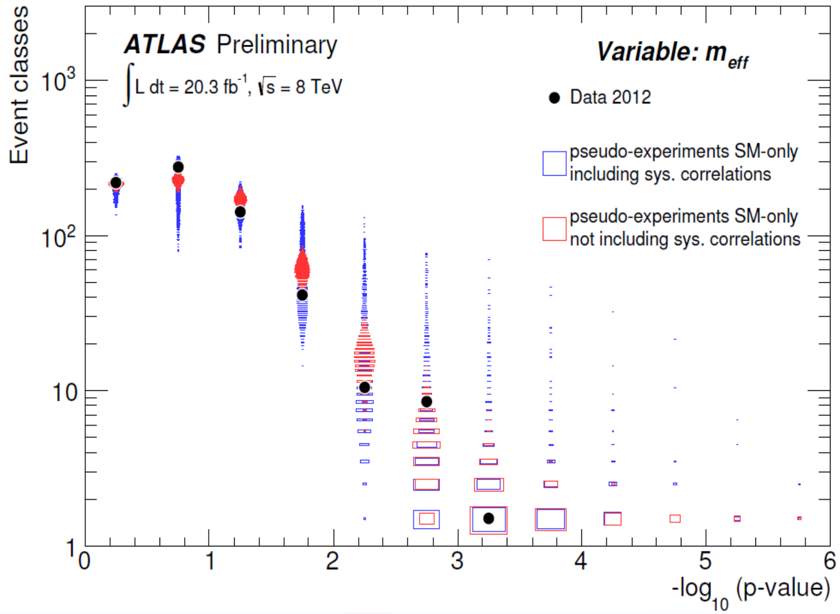


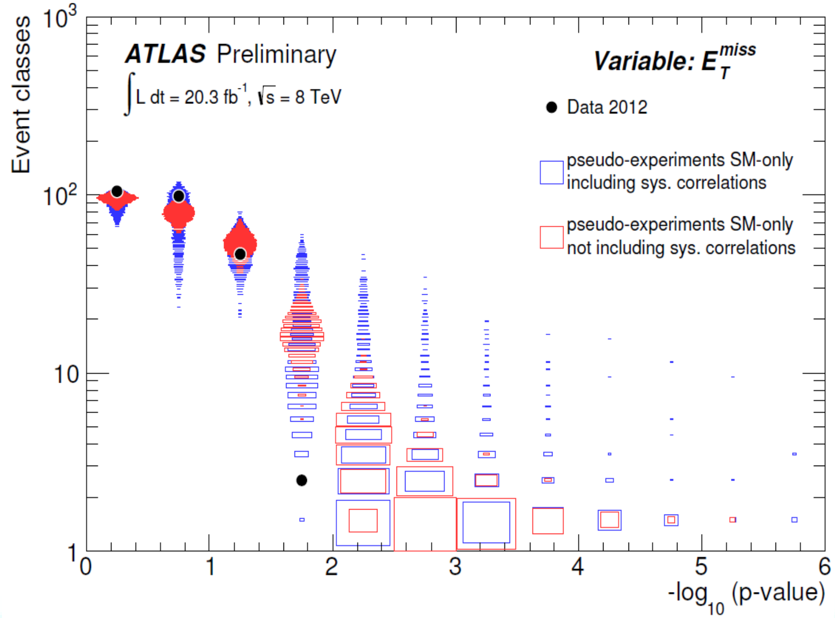
Scan results: *p*-Values

ATLAS-CONF-2014-006

- The distribution of the local p-values observed in data compared to the expectation.
- The expected distribution is obtained from pseudo-experiments assuming the „*SM-only*“ hypothesis.
- Observed significances are predicted by the theoretical distribution and are therefore explainable by statistical fluctuations within the uncertainties of the model.





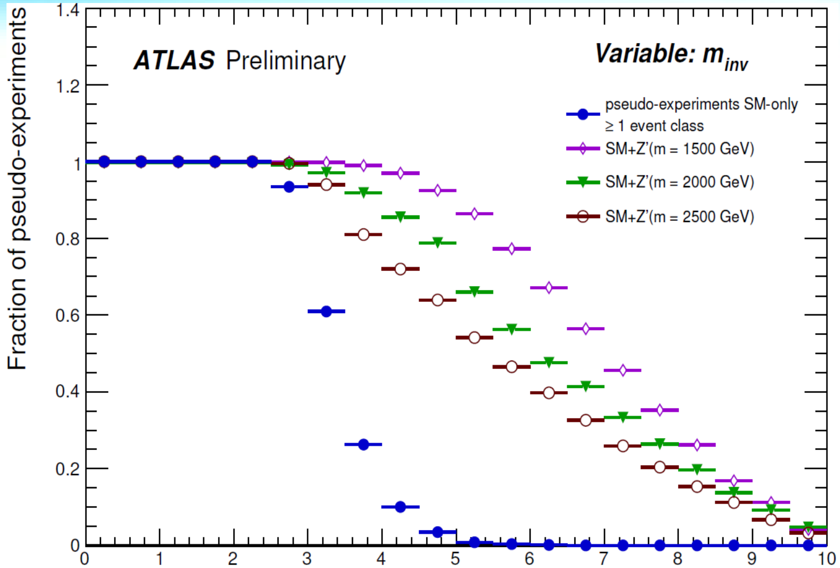


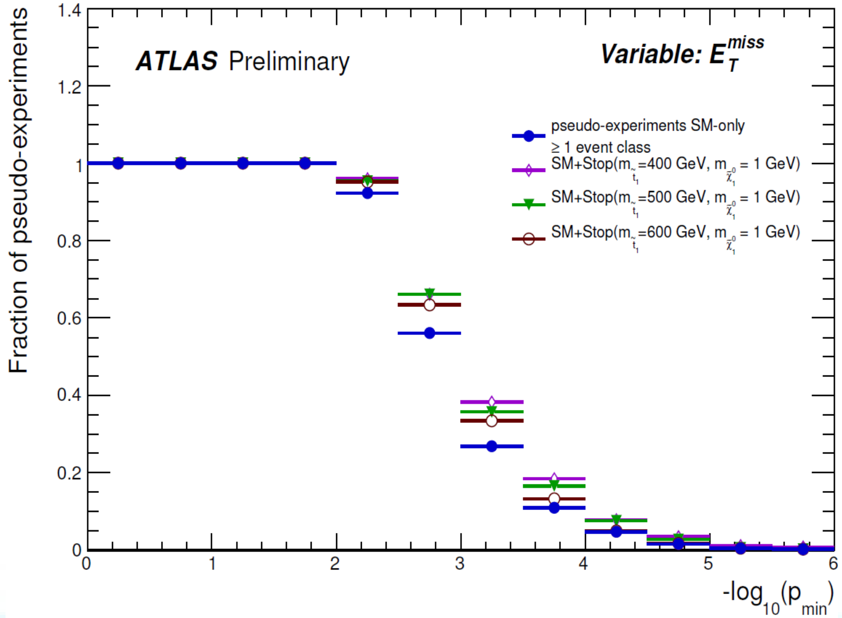
Scan results: Sensitivity

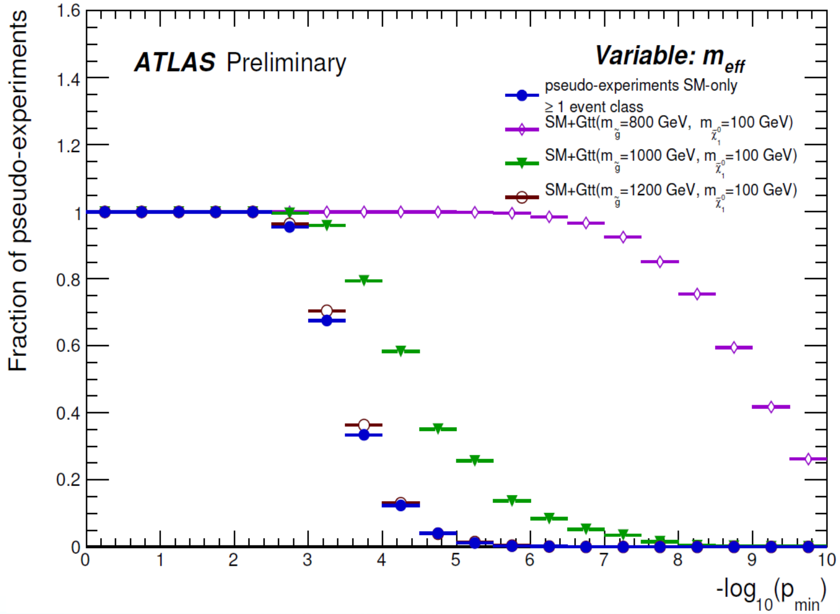
ATLAS-CONF-2014-006

Sensitivity of the analysis method to some benchmark models:

- Fraction of pseudo-experiments having at least one event class with a *-log$_{10}$(p)* greater than the one shown on the axis for pseudo-experiments assuming the *„SM-only“* & *„SM+Signal“* hypothesis.
- Less sensitive in E_T^{miss} (probably due to overestimation of JES uncertainty).







Conclusion & Outlook

Conclusions :

- A model independent general search provides an complementary approach to the dedicated analyses on ATLAS data.
- The analysis has been performed on the complete dataset collected during the year 2012 corresponding to an integrated luminosity of 20.3 fb^{-1}.
- No significant deviation from the Standard Model expectation has been observed. Actually, a remarkable good agreement between the observed deviations in data and the fluctuations expected from the SM was found.

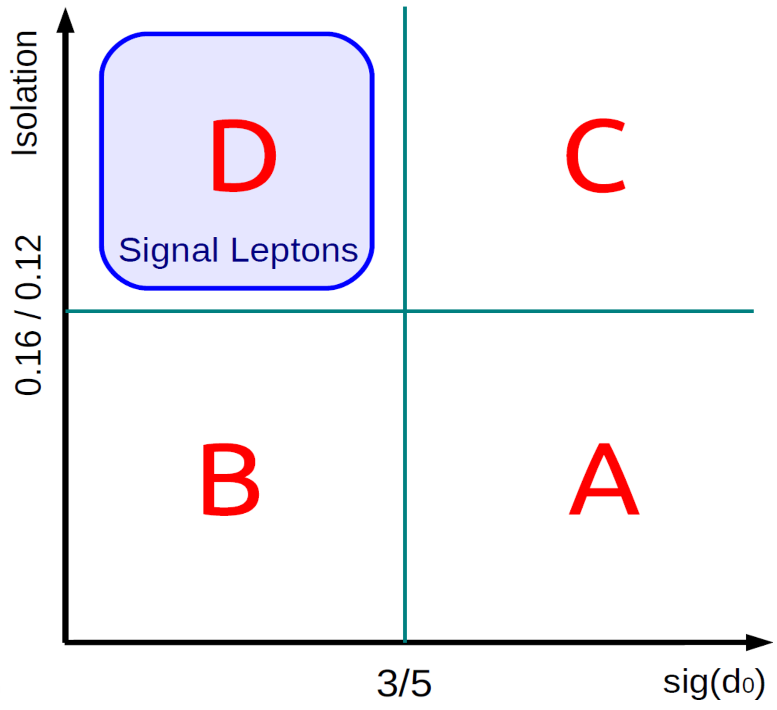
Outlook :

- This analysis can be extremely interesting considering the discovery potential of Run II. It can be used to give indications of unexpected signals in early Run II stages (first few fb^{-1}).
- In case of an observation of a significant deviation, the interesting region can be disclosed for the relevant analysis group to examine it with dedicated analysis techniques.
- Work will be continued during 2015 to enhance the strategy of background estimation, the treatment of analysis objects, as well as the statistical interpretation of the collected data.

Backup

Fake lepton estimation

- Single lepton + X channels have a big contribution of fake leptons in the low p_T region.
- These fakes cannot be modeled by Monte Carlo. The fake contribution is estimated using a data-driven approach.



ABCD method using the p_T^{cone} isolation and the transverse impact parameter significance $d_0/\sigma(d_0)$ of leptons.

- Fake enriched region A with no isolation and no $d_0/\sigma(d_0)$ requirement.
- Classify leptons according to position in ABCD plane
- Extract fake rate and extrapolate to signal region using

$$D = B \cdot C/A$$

Muons			Electrons		
Category	$d_0/\sigma(d_0)$	Isolation $p_T cone30/p_T$	Category	$d_0/\sigma(d_0)$	Isolation $p_T cone/p_T$
A	≥ 3	≥ 0.12	A	≥ 5	≥ 0.16
B	< 3	≥ 0.12	B	< 5	≥ 0.16
C	≥ 3	< 0.12	C	≥ 5	< 0.16
D	< 3	< 0.12	D	< 5	< 0.16

Fake lepton estimation

Estimation of fake contribution in low p_T region of lepton channels.

- ABCD method is able to model fake rates in these regions sufficiently
- The uncertainty of the fake background is determined with a Closure Test on Monte Carlo, where the method is applied on a pseudo-data sample composed of the usual MC and simulated fakes.
- According to the test results, the fake background is assigned with a 50% uncertainty.