The Compact Muon Solenoid Experiment
Conference Report

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

24 September 2009

Vector Boson plus Jets background to the early MET plus Jets Search

Mariarosaria D'Alfonso for the CMS Collaboration

Abstract

A search in large missing transverse momentum (MET) plus jets final state topology has great potential for discovery of new physics involving dark matter candidates early in LHC running. Some Standard Model processes produce the same observable signature. One of the main background is expected to be from higher order QCD contributions to single vector boson, which is difficult to predict with Monte Carlo. A strategy to estimate this contribution from data for the early search with the CMS detector using different control samples is presented.

Presented at *SUSY09: 17th International Conference On Supersymmetry And The Unification Of Fundamental Interactions*

Vector Boson plus Jets background to the early MET plus jets search

Mariarosaria D'Alfonso
(on behalf of the CMS collaboration)

University of California, Santa Barbara

Abstract. Searches in large missing transverse momentum (MET) plus jets final state topology have great potential for discovery of new physics involving dark matter candidates during early LHC running. Standard Model processes that produce the same observable signature constitute the background to this type of searches. One of the main background is expected to come from higher-order QCD radiation is single vector boson production, which is difficult to model with Monte Carlo simulations. A strategy to estimate this background using data driven methods with various control samples is presented for the early search with the CMS detector.

Keywords: ???

PACS: ???

INTRODUCTION

Discovery of new physics could take the form of an excess of events with high Missing Transverse Energy (MET) and energetic jets. For example MET plus single jet channel can be used to probe models involving extra dimensions; MET plus multijet channels are presented as a smoking gun for Super-Symmetric extensions of the Standard Model (SM); MET + forward jets channels is a discovery topology for an invisible Higgs boson produced through vector boson fusion.

Some Standard Model processes produce the same MET+jets signature and more copiously. For example a vector boson plus jets can mimic the MET plus jets signature. High p_T Z decaying to neutrinos constitute such an irreducible background; another example is W + Jets events with the W decaying leptonically but fail the lepton veto.

In these Beyond Standard Model (BSM) searches, the challenge is to convince ourselves that the supposed excess is not due to SM processes. For this reason the BSM searches will benefit from having more than one method to estimate the same background as multiple cross checking.

In this review a few examples of data driven methods studied by the CMS collaboration to estimate the invisible Z + Jets background for the early LHC searches, are presented.

DATA DRIVEN BACKGROUND OF INVISIBLE Z: MOTIVATION

It is desirable to estimate backgrounds by taking them directly from data. This minimize the dependence on Monte Carlo simulations, which are less reliable especially for high multiplicity jets. Also the dependence on the calibration of the experimental apparatus can be large for the early data taking. In so-called "data-driven" methods, one typically predicts distributions in regions where new physics may contribute (the "search region"), by extrapolating from SM-dominated control regions. One of the techniques is to identify a set of events that are the same as the backgrounds in search region, up to differences that one knows how to correct for.

A straightforward strategy to estimate the invisible Z boson production is to extract from data the events where a Z boson is decaying into two muons and to replace the muons with neutrinos. But considering the different branching ratios, the dimuon control sample would have only one sixth of the cross sections of the $Z \rightarrow \nu\bar{\nu}$ background we want to estimate. An additional factor of two of the reduction in the event yield is expected from muon acceptance and reconstruction efficiency. This control sample should be chosen to have the same jet multiplicity and the p_T of the boson should be selected accordingly to the MET cut applied to the search (typically Z p_T is few hundred GeV/c). Unfortunately few $Z \rightarrow \mu\mu$ events satisfying all the search criteria are expected in data samples of less than 1/fb of integrated luminosity. Thus a different strategy is needed to be effective for an early searches when only $O(100) \text{ pb}^{-1}$ will be available.

For these reasons the use of $Z \rightarrow \mu\mu$ + jets and $Z \rightarrow ee$ + jets with lower multiplicity, the W + jets and Photon

+ jets is investigated in [1] [2] [3]. The number of jets in these events are not easily predicted, but a good approximation do not depend upon whether boson is a Z, W or a Photon.

All the above methods exploits the use of the ratios of production of similar processes. In this way uncertainties associated to the SM production rates and to artifact of the detector simulation cancel to first order. This approach has numerous advantages in modeling the effect of the energy from the underlying event, detector effects like dead and noisy channels and uncertainty from the jet energy scale since the spectra in p_T , the distribution in η , and the composition (e.g. quark versus gluon) of the jets in the two processes are similar to first order.

In addition, uncertainties from a New Physics contamination that can distort the control sample, can be reduced. If the BSM events are copiously produced in the muon channel, they will probably not be in the same way or not at all in the photon channel. Since it is impossible to anticipate the nature of the BSM model, multiple cross check are necessary.

In the next sections, the use of different control samples for modeling the invisible Z background is reviewed.

$$Z(\rightarrow \mu\mu) + \text{Jets}$$

The invisible Z background can be derived from the sample of events containing a Z boson produced in conjunction with jets, and decaying into two muons. The high-jet multiplicity background is extrapolated from the lower-jet multiplicity events using the phenomenological assumption that $\frac{Z+n \text{ jets}}{Z+(n+1) \text{ jets}}$ is constant as a function of jet multiplicity. At first approximation, for each additional jet the cross section falls by a factor proportional to the strong coupling. In figure 1 an illustration of the measurement of the number of events as a function of jet multiplicity in the inclusive Z + Jets candle data is presented. In practice, events with $Z \rightarrow \mu\mu + \geq 3$ jets (and Z boson $p_T > 200$ GeV) will be normalized relative to the observed $Z \rightarrow \mu\mu + \geq 2$ jets events counts via the slope which will be measured in data from counting the $Z \rightarrow \mu\mu + \geq 1$ jet and $Z \rightarrow \mu\mu + \geq 2$ jets events. The Z decaying into di-electron events will also be used for this normalization purpose. In figure 2 the shape of the MET in $Z \rightarrow \nu\bar{\nu}$ is compared to the met distribution in $Z \rightarrow \mu\mu$ when the muons are not included in the computation of the event momentum balance. This method benefit from the fact that the Z candle is an almost pure control sample and also the theoretical correspondence with the invisible Z sample is trivial.

Source: CMS-TDR-Volume II

FIGURE 1. Illustration of the measurement of the number of events as a function of the jet multiplicity in the Z + Jets data. No Z p_T requirement is used.

Source: CMS-TDR-Volume II

FIGURE 2. MET in the $Z \rightarrow \mu\mu + \geq 2$ jets candle sample and normalized MET in $Z \rightarrow \nu\bar{\nu} + \geq 2$ jets sample.

$$\gamma + \text{Jets}$$

In [2], it is shown that γ + Jets production can be used to model the invisible Z background.

This control sample result in an large statistics with respect to the $Z \rightarrow \mu\mu$ and $W \rightarrow \mu\nu$ since there is no suppression from a branching fraction. In addition, this channel profits from the precise energy resolution of the CMS electromagnetic calorimeter.

Once the photon p_T is measured for events passing the selection criteria for the MET+jets search, the vector sum of the photon E_T and the calorimeter MET of the event is computed. The latter MET-like quantity is corrected for the photon selection efficiency, for the difference in electroweak coupling between photon and Z to quarks, and

Source: CERN-CMS-PAS-SUS-08-002

FIGURE 3. Demonstration of the inclusion of theoretical corrections obtained from generation level MC sample, and applied to full simulation samples.

for the $Z \rightarrow \nu\bar{\nu}$ branching fraction. All the theoretical corrections will be calculated to next-to-leading order, such techniques are expected to be available in the near future [4]. In figure 3 a demonstration of the inclusion of theoretical corrections applied to MC samples with full detector simulation, is presented.

The relative normalization of photon+jets to Z+jets might be affected by the collinear photon production, but this is expected to be mitigated by the isolation requirement on the photon. The effect of the difference in mass between the two bosons is reduced by the high p_T selection. In general there is also a difference in the η distribution of the photons with respect to Z bosons due to different vector and axial couplings. However as shown in figure 4 at sufficiently high p_T the bosons tend to be produced in both cases in the central region.

The background to the photon control sample arise from the misidentification of jets as photons. Indeed, jets can fake photons when they fragment into an hard π^0 or η . This QCD background is kept low by the isolation cuts resulting in only 1 part in 50 remaining after the full event selection. Despite this, a limit from data is still needed for the maximum number of events of fakes photons that could contaminate the photon control sample. The excellent CMS tracker provides an handles to confirm this low background expectation. A prompt photon has 50% probability of converting in the CMS tracker and in this case the momentum of the conversion tracks adds up to the photon momentum. In the case of a π^0

Source: CERN-CMS-PAS-SUS-08-002

FIGURE 4. Correspondence between Z and $\gamma \eta$ after the high p_T selection.

Source: CERN-CMS-PAS-SUS-08-002

FIGURE 5. Distribution of the sum of the conversion tracks for prompt and fake photon.

the probability to have a conversion is 75% and will only account for one of the two photons from the $\pi^0 \rightarrow \gamma\gamma$ ("secondary photons"). In figure 5 the different behavior of the sum of the conversion tracks for prompt and secondary photons is shown.

Source: CERN CMS PAS SUS 08 002

FIGURE 6. Expected MET-like spectrum from the muon plus jets control sample.

$W(\rightarrow \mu\nu) + \text{Jets}$

For this control sample, events with exactly one isolated muon with $p_T > 20$ GeV/c are selected. All the other search cuts like jet counts, are then applied. The MET-like quantity is equivalent to the calorimeter MET and is not corrected for the small calorimeter deposit of the muon. The QCD multijets background is expected not to pass the relatively high muon p_T and the high MET selection criteria and can be estimated using an orthogonal muon sample obtained inverting the isolation requirement in the muon identification. The top background can be significant for the higher jet multiplicity events and can be estimated from the b quark content, i.e. by using the ratio of tagged/untagged events.

In figure 6 the MET-like quantity is plotted in the context of the muon plus 3 jets control sample. The other SM sources that contaminate this selection are also shown. In figure 7 a comparison between MET estimated from the $W \rightarrow \mu\nu$ events and the selected $Z \rightarrow \nu\bar{\nu}$ events plus jets is plotted in the context of met plus monojets search: the good agreement is observed between the two distributions.

A further complication is that BSM physics processes are typically characterized by multi-jet events, but when a real lepton is present in the decay chain, such events may themselves pass the muon+jets selection. The contamination can be high since the selection cuts are inherited from the MET+jets search and are designed to enhance the BSM presence. For this reason a comparison

Source: CERN-CMS-PAS-EXO-08-011

FIGURE 7. Selected $Z \rightarrow \nu\bar{\nu} + \text{jet}$ events and spectrum estimated from the $W(\rightarrow \mu\nu) + \text{Jets}$ procedure in the context of a MET + monojet search.

between all the vector boson plus jets estimations can be an important step in understanding the presence (or not) of contamination in the background predictions.

Conclusion

In this paper, methods studied by the CMS collaboration to predict the invisible Z background at high p_T produced with multiple jets have been reviewed in the context of early searches for new phenomena with the signature of MET plus jets. Ideas for predicting other vector boson plus jets backgrounds are being explored. Extensive cross-checking and multiple redundant methods are important since the challenge is to convince ourselves that the observed excess is not just an under-estimation of SM processes.

REFERENCES

1. CMS collaboration, CMS Physics TDR: Volume II, J. Phys. G: Nucl. Part. Phys. 34 995-1579, doi:10.1088/0954-3899/34/6/S01.
2. CMS collaboration, *Data-driven estimation of the invisible Z background to the SUSY MET plus jets search*, CERN-CMS-PAS-SUS-08-002.
3. CMS collaboration, *Search for mono-jet final states from ADD extra dimensions*, CERN-CMS-PAS-EXO-08-011.
4. David A. Kosower et al., *Next-to-Leading Order QCD Predictions for W+3-Jet Distributions at Hadron Colliders*, arXiv:0907.1984v1 [hep-ph]