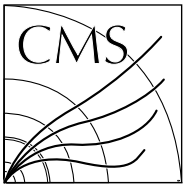
Available on CMS information server

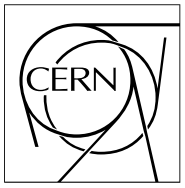
CMS NOTE 2010/version0



The Compact Muon Solenoid Experiment

CMS Note

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



October 23, 2010

Jet-veto Efficiency for the WW production cross section in pp collisions at $\sqrt{s} = 7$ TeV

D. Barge, C. Campagnari, P. Kalavase, D. Kovalskyi, V. Krutelyov, J. Ribnik

University of California, Santa Barbara, Santa Barbara, USA

W. Andrews, D. Evans, F. Golf, J. Mulmenstadt, S. Padhi, Y. Tu, F. Wurthwein, A. Yagil

University of California, San Diego, San Diego, USA

L. Bauerdick, I. Bloch, K. Burkett, I. Fis, Y. Gao, O. Gutsche, B. Hooberman

Fermi National Accelerator Laboratory, Batavia, USA

Abstract

This note describes a data-driven method to estimate the jet-veto signal efficiency in data and its systematic uncertaintes for the WW analysis. In this method, the WW jet-veto efficiency in data is parametrized as the product of the Z jet-veto efficiency in the data and the WW/Z jet-veto efficiency in the MC. We measure the Z jet-veto efficiency using the 3.1/pb data and compared to various MC samples. The dominant systematic uncertainty in the WW jet-veto efficiency is from the WW/Z jet-veto efficiency ratio estimation. The WW jet-veto efficiency is estimated as $61.1 \pm 0.9 \pm 4.5$ % using the corrected PFJet at the jet-veto thresold of 25 GeV.

Contents

1 Introduction

For the cross section measurement of $pp \rightarrow WW \rightarrow 2l2\nu_l$ [1], one of the dominant backgrounds is the top-background. The main difference between the top-background and WW signal is the presence of one or two extra b-jets in the final states. Thus, one efficienct way to suppress the top-background is to apply jet-veto, vetoing the events with leading jet p_T above a thresold at a given η range. The specific p_T thresolds and the η ranges are discussed in this note.

The jet-veto efficiency of WW can not be measured directly from the data. We have to use the MC to certain extent. In this note, we describe a partially data-driven approach to measure WW jet-veto signal effcicieny in data. In this approach, we use the well-defined process $pp \rightarrow Z \rightarrow 2l$ as a control region to test the agreement in the jet spectrum between data and MC. We then estimate the jet-veto signal efficiency on WW as the jet-veto efficiency of $pp \rightarrow Z \rightarrow 2l$ on data multiplied by the WW/Z jet-veto efficiency ratio estimated on the Monte Carlo, shown in Equation 1. Table 1 summarizes the data and MC samples used in this note.

The uncertainties of jet-veto signal efficiency come from both the matrix element calculations and the jet reconstruction. In the leading order matrix element calculation, the $pp \rightarrow WW \rightarrow 2l2\nu$ process does not contain jet. Extra jets are primarily due to the initial state radition. It makes the jet energy spectrum sensitive to the corrections beyond the LO. We could estimate this systematic error by comparing the predictions from various MC generators, such as, the Pythia, Madgraph and the MC@NLO. On the jet reconstruction side, the effects from the jet energy correction and its uncertainties can propagate towards the jet-veto efficiency estimation as well. Validation of the jet response between the data and MC is thus required.

The note is organized as follows. In section 2, we discuss the jet-veto efficiency measurement on Z data, in which the choice of the jet-veto η region is examined. In section 3, we describe the WW/Z jet-veto efficiency ratio estimate and its uncertainties. In section 4, we discuss the effect of the jet energy response to the results. Finally, in section 5, we present the jet-veto efficiency measurement for WW in data.

$$\epsilon_{WW}^{data} \approx \epsilon_{Z}^{data} \times (\epsilon_{WW}^{MC}/\epsilon_{Z}^{MC}) = \epsilon_{Z}^{data} \times R_{WW/Z}^{MC} \tag{1}$$

2 Jet-veto Efficiency Measurement on Z data

In this section, we describe the jet-veto efficiency measurement on the $pp \rightarrow Z \rightarrow ll$, comparing the data with various Monte Carlo samples (Table 1).

We use the same lepton selections as in the WW analysis to select the two leptons in the $Z \rightarrow ll$ decays. Additional requirements to select the Z events are listed as follows,

- dilepton mass inside the Z mass window: $|m_{ll} - m_Z| < 15$GeV;
- if more than one di-lepton candidates are selected in an event, the one with di-lepton mass closest to the Z mass is chosen

Table 1: List of datasets

Collision Data	
Luminosity	Dataset
3.1	/EG_Run2010A-PromptReco-v4_RECO/ /EG_Run2010A-PromptReco-v4_RECO/ /EG_Run2010A-Jul16thReReco-v2_RECO/ /EG_Run2010A-Jun14thReReco_v1_RECO/ /MinimumBias_Commissioning10-SD_EG-Jun14thSkim_v1_RECO/
Monte Carlo	
Pythia	/Zee_Spring10-START3X_V26_S09-v1/ /Zmumu_Spring10-START3X_V26_S09-v1/ /WW_Spring10-START3X_V26_S09-v1
Madgraph	/ZJets-madgraph_Spring10-START3X_V26_S09-v1 /VVJets-madgraph_Spring10-START3X_V26_S09-v1/
MC@NLO	/Zgamma_ee_M20-mcatnlo_Spring10-START3X_V26_S09-v1/ /Zgamma_mumu_M20-mcatnlo_Spring10-START3X_V26_S09-v1/ /WWtoEE-mcatnlo_Spring10-START3X_V26_S09-v1/ /WWtoEPlusMuMinus-mcatnlo_Spring10-START3X_V26_S09-v1/ /WWtoEPlusTauMinus-mcatnlo_Spring10-START3X_V26_S09-v1/ /WWtoMuMu-mcatnlo_Spring10-START3X_V26_S09-v1/ /WWtoMuPlusEMinus-mcatnlo_Spring10-START3X_V26_S09-v1/ /WWtoMuPlusTauMinus-mcatnlo_Spring10-START3X_V26_S09-v1/ /WWtoTauTau-mcatnlo_Spring10-START3X_V26_S09-v1/ /WWtoTauPlusEMinus-mcatnlo_Spring10-START3X_V26_S09-v1/ /WWtoTauPlusMuMinus-mcatnlo_Spring10-START3X_V26_S09-v1/

2.1 Jet-veto η region

Aside from the maximum leading jet p_T , the choice of η range where the jet-veto is applied is also important. In the summer08 analysis of WW [1], the jet-veto region is $|\eta| < 3.0$. However, applying the jet-veto in the full η range ($|\eta| < 5$) gives the best signal over background performance. It also helps in estimating the remaining top-background after the full selection, as it is harder to estimate the top background in the region $3 < |\eta| < 5$ outside the tracking acceptance. As the jets in the region $3 < |\eta| < 5$ have worse resolution than the central jets with $|\eta| < 3$, we check the data/MC performance.

Figure 1 shows the leading jet p_T in the region of $3 < |\eta| < 5$, and Figure 2 shows the respective jet-veto efficiencies and the MC/data ratios. We find the MC/data jet-veto efficiency ratios are close to unity for all the jet algorithms in all the MC samples considered. Therefore, we choose to apply the jet veto in the full η range $|\eta| < 5.0$.

2.2 Jet-veto efficiency estimate on the full $|\eta| < 5$ region

Figure 3 shows the jet-veto efficiencies and the MC/data ratios, using the uncorrected JPT jets and PFJet [2]. We find similar performance between these two types of jets even without the jet energy correction. Figure 4 shows the jet-veto efficiencies using the L2L3 corrected CaloJet, TrkJet, PFJet and JPT Jets. The MC@NLO sample is found to have softer jet energy spectrum compared to data, yielding a larger jet-veto efficiency. The jet-veto efficiencies in the Pythia/Madgraph MC samples agree well with the data, with a discrepancy level of within 5%. The Z jet-veto efficiencies using the uncorrected and corrected PFJets at three different energy cut (20, 25, 30) GeV are tabulated in Table 2-3.

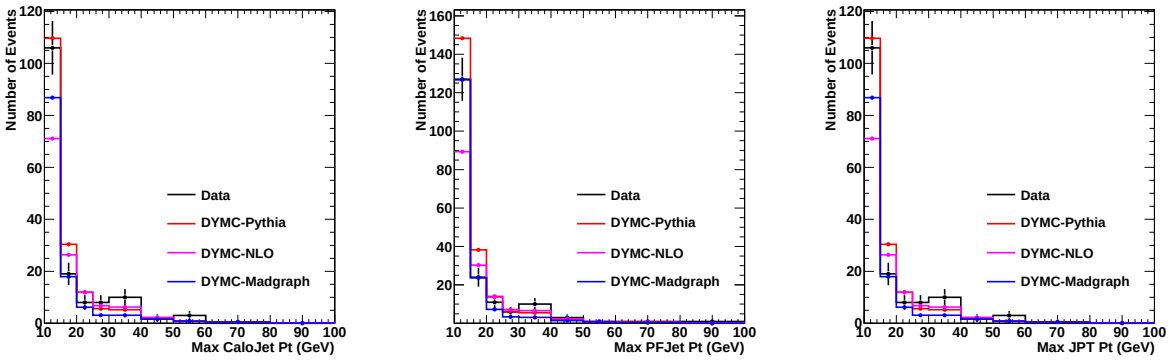


Figure 1: The leading jet p_T using uncorrected CaloJets (left), PFJets (middle) and JPTs (right) within $3 < |\eta| < 5$. The MC samples are normalized to the data by the number of events.

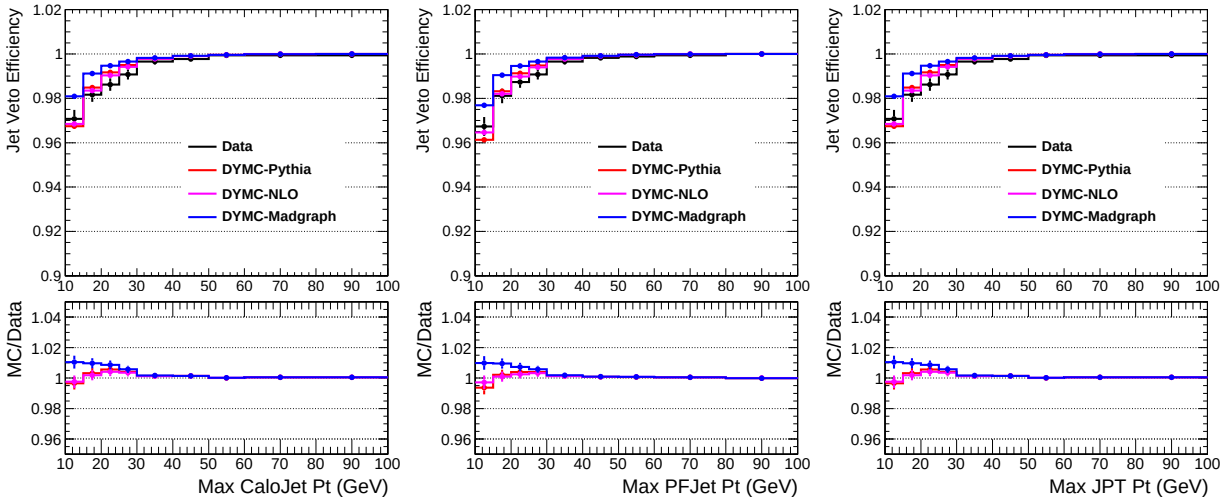


Figure 2: The jet-veto efficiency and MC/data ratios using uncorrected CaloJets (left), PFJets (middle) and JPTs (right) within $3 < |\eta| < 5$.

Table 2: The jet-veto efficiency and the MC/data ratios on the Z events, using the uncorrected PFJet.

Jet-veto Cut (GeV)	Pythia/Data (%)	Madgraph/Data (%)	MC@NLO/Data (%)	$\epsilon_Z^{data}(\%)$
20	100.4± 1.3	103.4± 1.4	107.3± 1.4	76.3± 1.0
25	100.7± 1.1	102.1± 1.1	105.1± 1.2	82.6± 0.9
30	100.2± 0.9	100.9± 0.9	103.5± 1.0	87.2± 0.8

Table 3: The jet-veto efficiency and the MC/data ratios on the Z events, using the corrected PFJet with L2L3, plus the residual corrections for data.

Jet-veto Cut (GeV)	Pythia/Data (%)	Madgraph/Data (%)	MC@NLO/Data (%)	$\epsilon_Z^{data}(\%)$
20	100.3± 1.5	104.8± 1.6	109.6± 1.7	71.3± 1.1
25	101.4± 1.3	103.7± 1.3	106.8± 1.3	78.8± 1.0
30	100.5± 1.0	101.8± 1.0	104.4± 1.1	84.5± 0.9

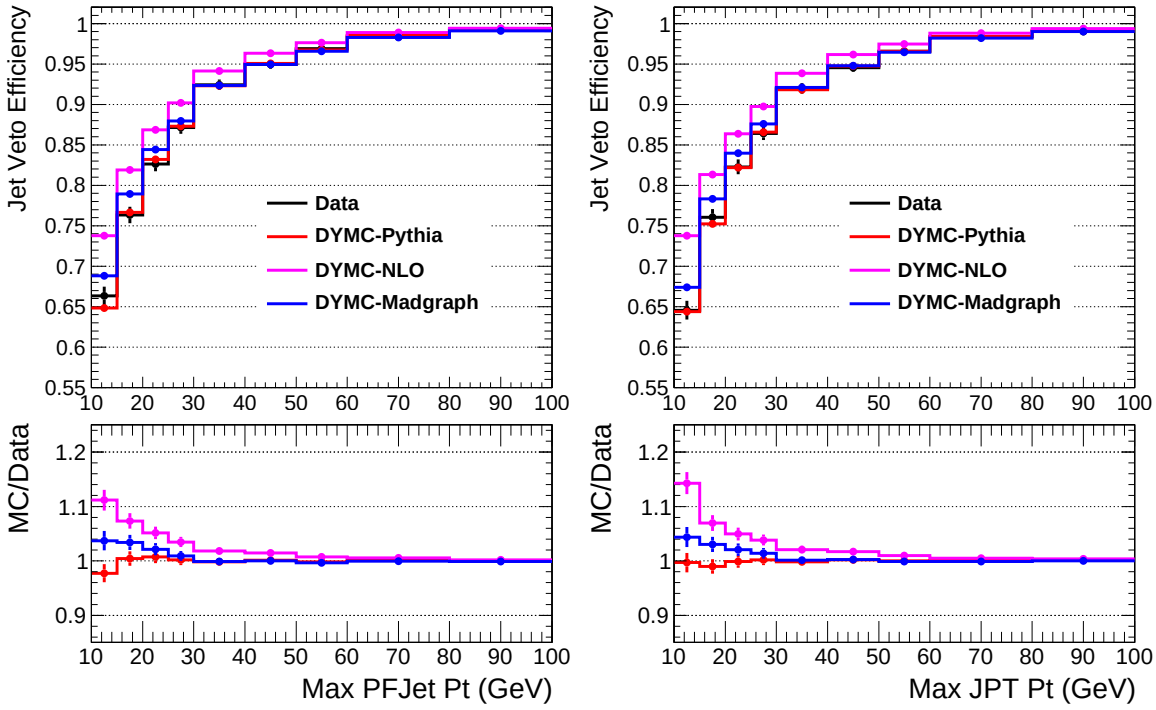


Figure 3: The jet-veto efficiency and MC/data ratios using uncorrected PFJets (left) and JPTs (right) within $|\eta| < 5$.

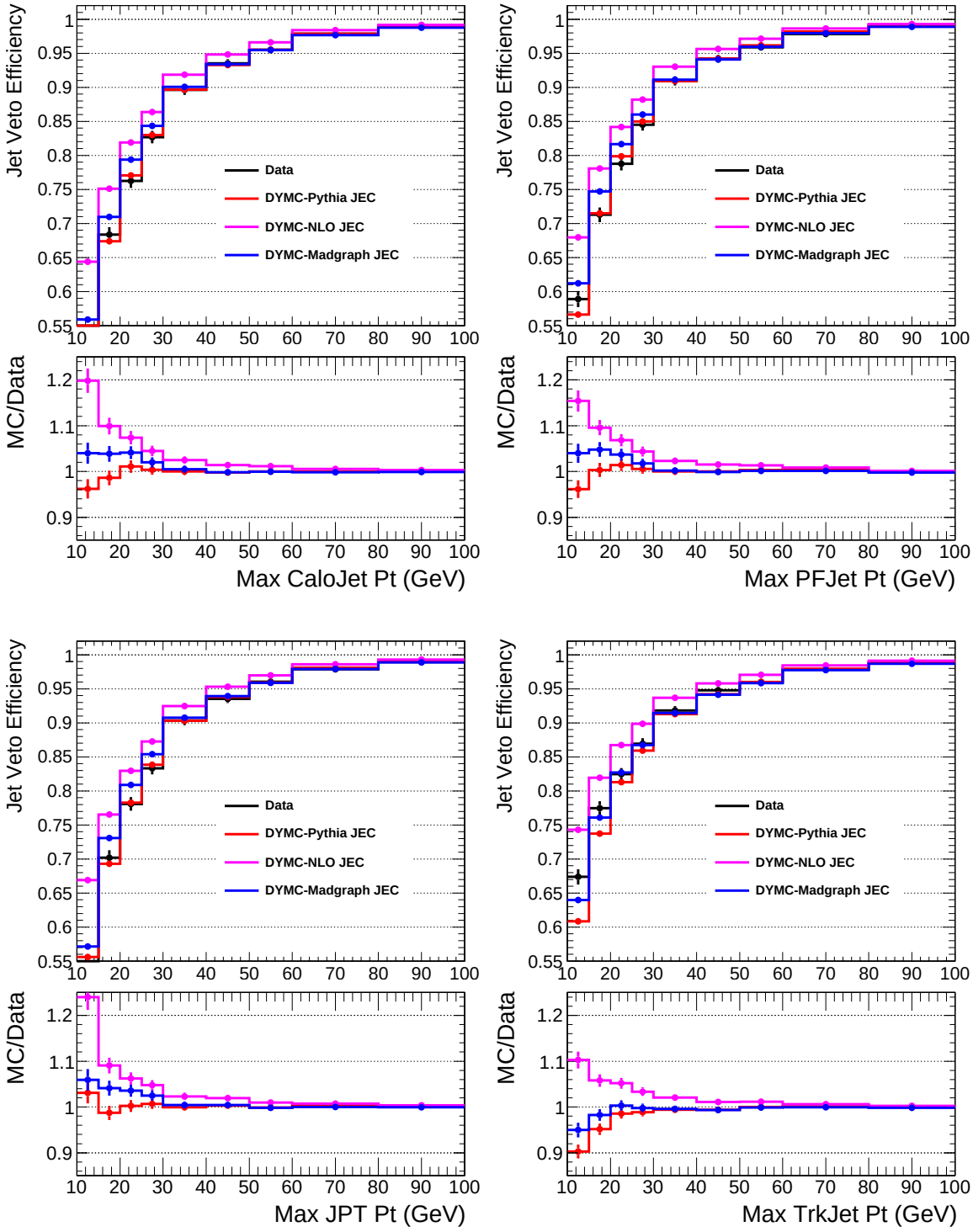


Figure 4: The jet-veto efficiency and MC/data ratios using corrected CaloJets (top left), PFJets (top right), JPTs (bottom left) and TrkJets (bottom right) within $|\eta| < 5$.

3 WW/Z Jet-veto Efficiency Ratio Estimate

In this section, we describe the WW/Z jet-veto efficiency ratio $R_{WW/Z}$ (Equation 1) estimation. WW production has different kinematics from the Z, as the energy scale in the former is much larger than the latter (see Figure 5, left). This indicates that the incoming partons in WW production have a larger energy on average than Z production, which results in a harder ISR parton energy spectrum, shown in Figure 5 (right).

Figure 6-7 show the WW/Z jet-veto efficiency ratio dependence on the leading jet p_T , using the uncorrected and corrected jets respectively. The predictions from Madgraph agree with the MC@NLO within 1-2%. The predictions from Pythia MC differ from the other two in the order of 10% on average. Note that the jet-veto efficiencies in the Z control region from Pythia and Madgraph MC samples agree within 3% (section 2). The large Pythia/Madgraph difference in the $R_{WW/Z}$ is due to the different leading jet p_T spectrum in the WW productions. More specifically, Pythia predicts a much softer jet energy spectrum than Madgraph in the WW production.

In Pythia, the matrix element calculation for WW production includes only the leading order contributions. The ISR is modeled through the parton showering in the soft-collinear limit. On the other hand, both Madgraph and MC@NLO MCs take into account upto 1 parton ISR in the matrix element calculation. The large Pythia/Madgraph difference could be due to the imperfect modelling of ISR in Pythia parton showering. The good agreement between Pythia and Madgraph in the Z jet-veto efficiency is likely because Pythia parton showering is tuned well on Z region using data.

To disentangle the jet reconstruction differences, we study the WW/Z jet-veto efficiency ratio dependence on both the generator level jet energy and the parton energy, shown in Figure 8. The large difference of Pythia predictions from Madgraph and MC@NLO persists at the generator level. We use the MCFM[**?**] MC generator to probe the WW/Z jet-veto efficency ration dependence on the matrix element level parton energy. The events are generated with up to 1 extra parton in the final states. The parton p_T is taken as the magnitude of the recoiling four or two leptons $p_T$in the $WW \rightarrow 2l2\nu$ or $pp \rightarrow Z \rightarrow 2l$ decays. The jet-veto efficiency curve obtained using the parton p_T is between Pythia and Madgraph (MC@NLO). We change the normalization (μ_R) and factorization (μ_F) scales to $1/2$ and 2 times the conventional values, and find the differences in the jet-veto efficiency smaller than the Pythia/Madgraph differences.

We choose Madgraph as the MC sample to calculate the WW/Z jet-veto efficiency ratio $R^{MC}_{WW/Z}$. This is because Madgraph MC predicts well the jet-veto efficiency in Z events, and it includes extra one-parton radiation in the matrix element calculation of WW production. We assign half the Pythia/Madgraph difference as the systematic error for $R^{MC}_{WW/Z}$. The results are tabulated Table 4-5 using the uncorrected and corrected PFJets for three different jet energy cut (20, 25, 30) GeV.

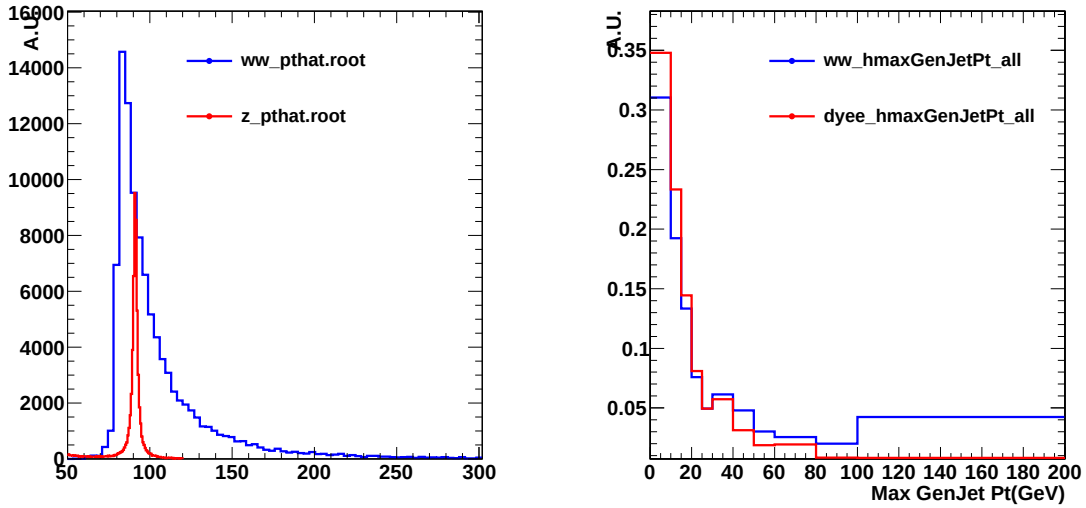


Figure 5: The $\hat{s}$ (left) and the leading GenJets for $pp \rightarrow WW$ (Blue) and $pp \rightarrow Z$ (Red).

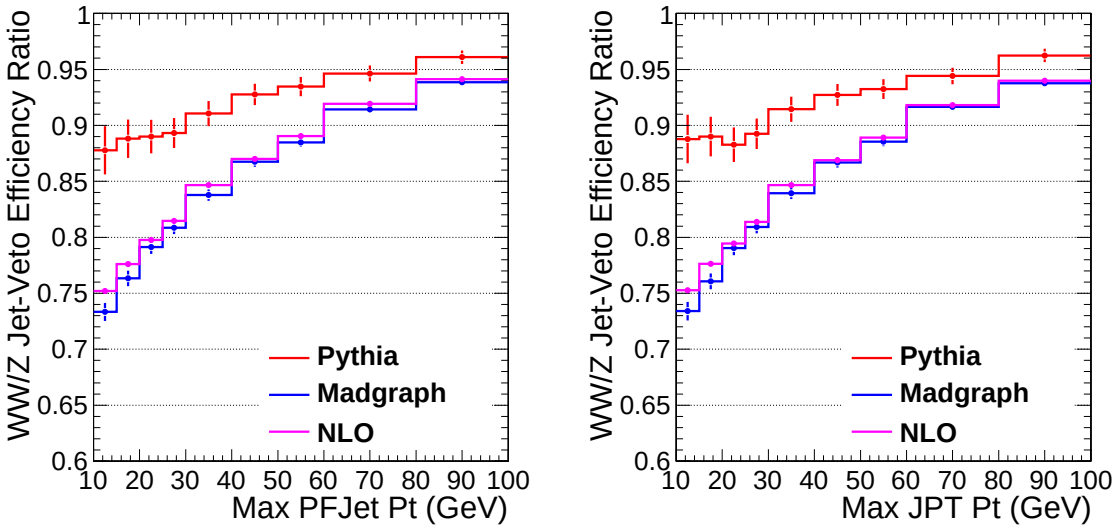


Figure 6: The WW/Z jet-veto efficiency ratio $R_{WW/Z}$ using uncorrected PFJets (left) and JPTs (right).

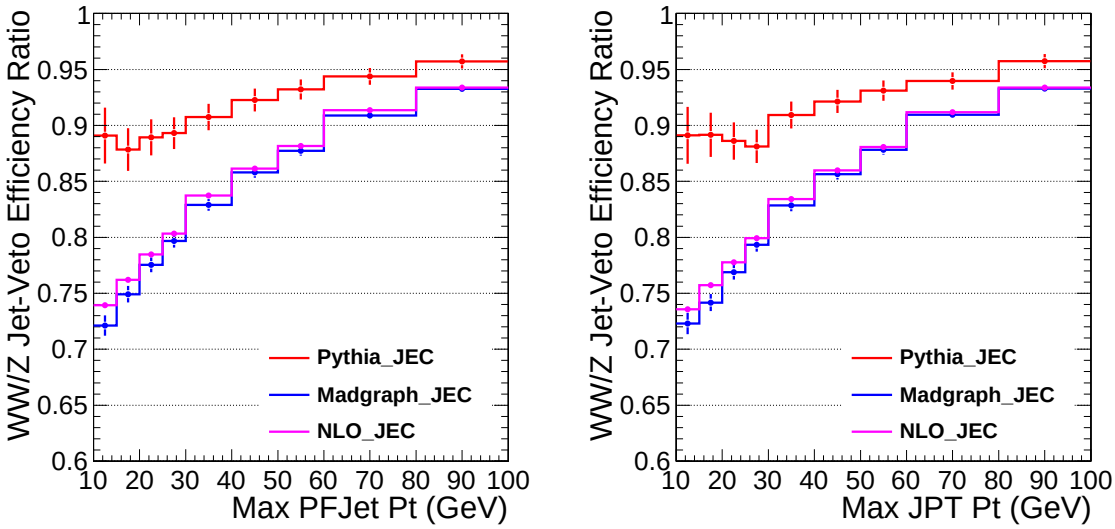


Figure 7: The WW/Z jet-veto efficiency ratio $R_{WW/Z}$ using corrected PFJets (left) and corrected JPTs (right).

4 Jet Response Validation Using $Z + 1$ Jet Events

The jet energy correction [3] uncertainties give rise to addtional uncertainties in evaluating both the Z jet-veto efficiency and the WW/Z jet-veto efficiency ratios.

The bulk of the jet energy correction are the L2L3 corrections, see Figure 9 (left). The same L2L3 corrections are applied to both data and MC. For the data, additional residual corrections (Figure 9, right) are applied to improve the data/MC agreement [4]. The jet energy correction in the p_T range of (20-30) GeV are important for the jet-veto signal efficiency. As the jet energy corrections are primarily derived for the high p_T jets, we perform a simple cross check on the data/MC comparisons of the jet response, using the Z+jet events, similar to the method described in [5].

The Z+jet events are selected with the following two additional requirements beyond the selections described in section 2.

- The leading jet and the Z are back-to-back: $|\Delta\phi(\vec{p}_{leadingjet} - \vec{p}_Z)| < 0.2$;
- The maximum p_T of other jets in the event is $0.1 \times p_T(leadingjet)$.

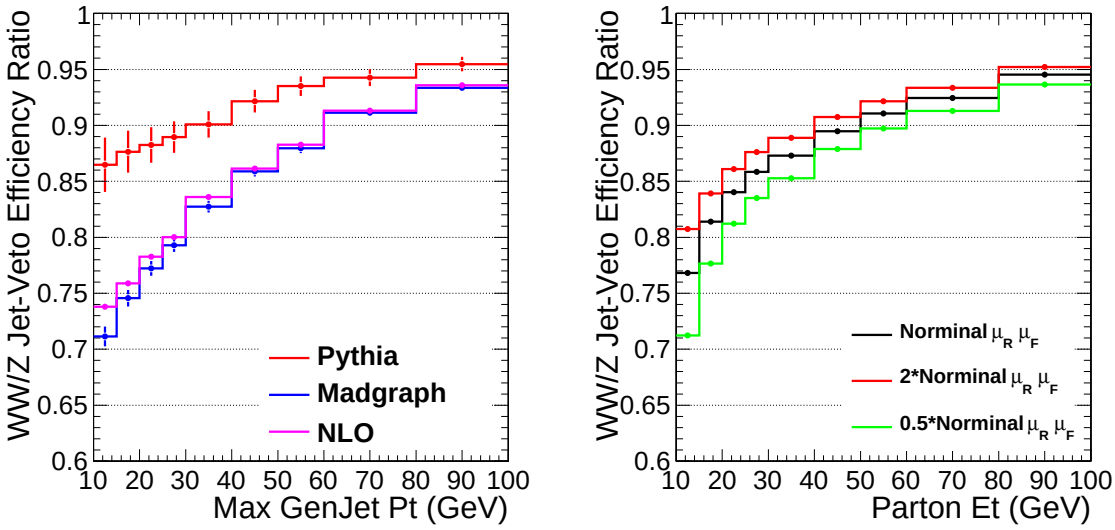


Figure 8: The WW/Z jet-veto efficiency ratio $R_{WW/Z}$ using GenJets (left) and the parton p_T(right). The parton four-momentum is calculated as the sum of the four (two) leptons in the case of $WW \to 2l2\nu_l$ ($Z \to 2l$).

Table 4: The WW/Z jet-veto signal efficiency ratio using uncorrected PFJets. The efficiency ratio is quoted as the value using PFcJets based on the Madgraph MC sample. The uncertainty on $R^M_{WW/Z}$ is taken as half of the difference between Pythia and Madgraph.

Jet-veto Cut (GeV)	$\sigma R^{MC}_{WW/Z}$ (%)
20	76.3± 0.7± 6.2
25	79.1± 0.6± 4.9
30	80.9± 0.6± 4.2

Table 5: The WW/Z jet-veto signal efficiency ratio using L2L3 corrected PFJets. The efficiency ratio is quoted as the value using PFJets based on the Madgraph MC sample. The uncertainty on $R^{MC}_{WW/Z}$ is taken as half of the difference between Pythia and Madgraph.

Jet-veto Cut (GeV)	$\sigma R^{MC}_{WW/Z}$ (%)
20	74.9± 0.7± 6.5
25	77.5± 0.7± 5.7
30	79.7± 0.6± 4.8

Figure 10 shows the jet energy scale derived from the Z+jet events for PFJets. The jet energy scale is defined as the $p_T(leadingjet)/p_T(Z)$. We see a small systematic shift of about 5% for the jets in the p_T range of (20-30) GeV. However, the data is statistically limited. Note that the efficiency difference for a correction around 11% (Figure 9, left) yields a difference about 4% in the Z jet-veto efficency uncertainty and 1% in the $R_{WW/Z}$. We estimate that the jet energy correction induced systematic uncertainty is less than 2%, assuming the jet energy correction uncertainty is 5%. For the time being, the contributions of this difference is ignored in the jet-veto efficiency estimation. We will repeat the study with more data and apply appropriate corrections based on more precise measurement of this data/MC shift.

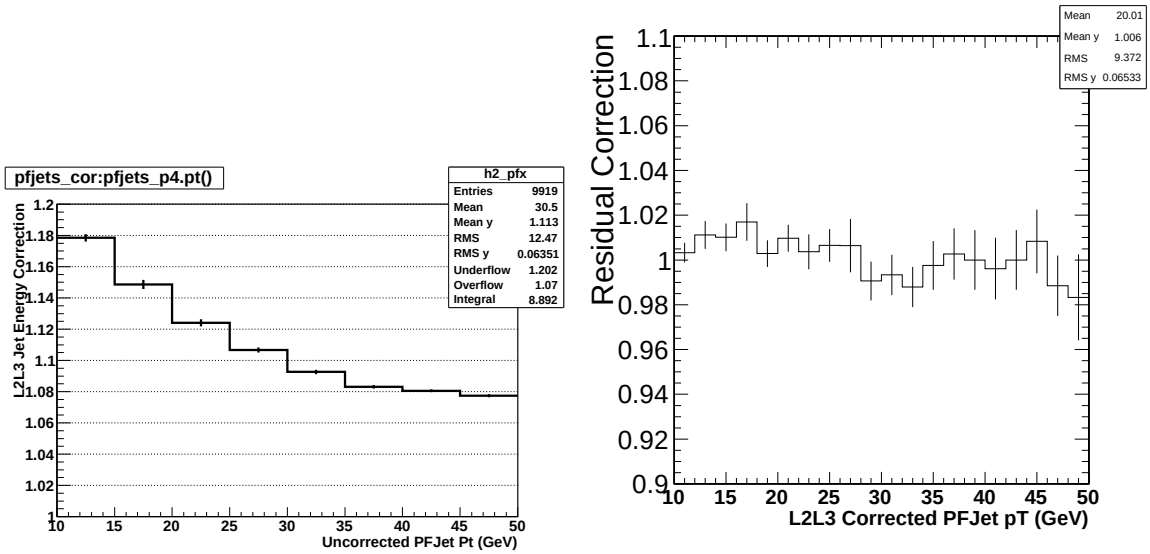


Figure 9: The L2L3 jet energy correction (left) and the residual corrections in data (right) for PFJets.

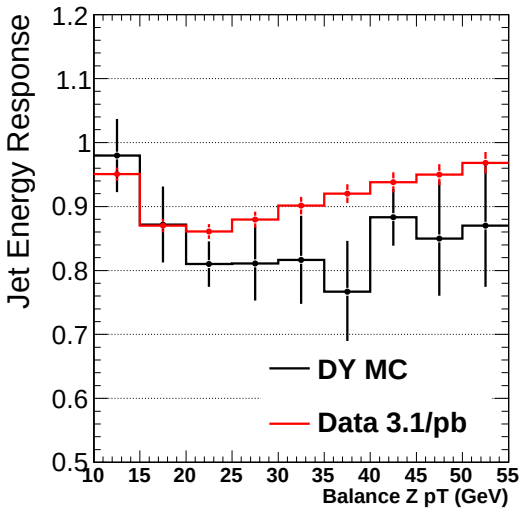


Figure 10: The jet energy scale derived from the Z+jet events for PFJets.

5 Summary and Conclusion

In summary, we have described a partially data-driven method to estimate the WW jet-veto efficiency and its systematic uncertainties in data. The WW jet-veto efficiency in data is parametrized as the product of the Z jet-veto efficiency in data and the WW/Z jet-veto efficiency in MC.

The Z jet-veto efficiency is measured in data and compared to various MC samples, shown in Table 2-3. We have observed good agreement between data and the Pythia and Madgraph MC samples. The jet-veto efficiency predicted by MC@NLO is found to be larger than the measurement in data.

On the other hand, we have compared the WW/Z jet-veto efficiency ratio using the Pythia, Madgraph and MC@NLO MC samples. The predictions from the Madgraph and MC@NLO MC samples agree with each other within 1%. The WW jet energy spectrum in Pythia is significantly softer than the other two, yielding a difference of more than 10% in the WW/Z jet-veto efficiency ratio at the jet p_T cut of 20 GeV. Based on the above findings, we choose Madgraph as the nominal MC to calculate the WW/Z jet-veto efficiency ratio and assign half of the difference between the Pythia and Madgraph predictions as the $R^{MC}_{WW/Z}$ uncertainties.

The WW jet-veto efficiencies using uncorrected and corrected PFJets at p_T cut of (20, 25, 30) GeV are tabulated in Table 6-7.

Table 6: The WW jet-veto signal efficiencies using uncorrected PFJets. The WW/Z jet-veto efficiency ratios ($R^{MC}_{WW/Z}$) use the Madgraph MC sample.

Jet-veto Cut (GeV)	ϵ_Z^{data} (%)	$R^{MC}_{WW/Z}$ (%)	ϵ_{WW} (%)
20	76.3± 1.0	76.3± 0.7± 6.2	58.3± 0.9± 4.8
25	82.6± 0.9	79.1± 0.6± 4.9	65.4± 0.9± 4.1
30	87.2± 0.8	80.9± 0.6± 4.2	70.5± 0.8± 3.7

Table 7: The WW jet-veto signal efficiencies using corrected PFJets. The WW/Z jet-veto efficiency ratios ($R^{MC}_{WW/Z}$) use the Madgraph MC sample.

Jet-veto Cut (GeV)	ϵ_Z^{data} (%)	$R^{MC}_{WW/Z}$ (%)	ϵ_{WW} (%)
20	71.3± 1.1	74.9± 0.7± 6.5	53.4± 1.0± 4.6
25	78.8± 1.0	77.5± 0.7± 5.7	61.1± 0.9± 4.5
30	84.5± 0.9	79.7± 0.6± 4.8	67.3± 0.9± 4.1

References

[1] CMS AN -2009/042, "Prospects for measuring the WW production cross section in pp collisions at $\sqrt{s} =$ 10 TeV"

[2] CMS PAS JME-07-003, "Performance of Jet Algorithms in CMS"

[3] CMS PAS JME-10-003, "Jet Performance in pp Collisions at $\sqrt{s} = 7$ TeV"

[4] https://hypernews.cern.ch/HyperNews/CMS/get/JetMET/1017.html

[5] CMS PAS JME-09-005, "Determination of the jet energy scale using $Z \rightarrow e^+e^-$ + jet $p_T$balance and a procedure for combining data driven corrections", CMS PAS JME-09-009, "Calibration of the absolute jet energy scale with $Z(\rightarrow \mu^+\mu^-)$ + jet events at CMS"

[6] http://mcfm.fnal.gov/