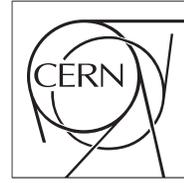


The Compact Muon Solenoid Experiment
Analysis Note



The content of this note is intended for CMS internal use and distribution only

2

3

April 12, 2010

4

Abstract

5

1 Determination of the jet energy resolution in photon-jet events by using E_T^{miss} Projection Fraction Method

Elif

1.1 Introduction

As mentioned in Section ??(this should be the introduction section of QCD background), the E_T^{miss} in QCD arises from either from the instrumental effects i.e. imperfect response of the detector and detector geometric acceptance, or from the semi leptonic decays of heavy quarks. The semi-leptonic contribution will be suppressed by the indirect lepton veto proposed to be applied in all-hadronic search. At start of the run, the detector simulation may not describe the details of the calorimeter response to jets and one may not be able to predict the E_T^{miss} distribution from QCD multijet production accurately. Thus it is essential to determine this background directly from data with minimal information from the Monte Carlo simulated events. To determine QCD background, many different data driven methods have been studied. Here we describe a procedure where the detector response to jets is measured directly from data using photon-jets events. If enough photon-jets events are not available, one may have to use dijet/multi-jet events to evaluate the tails in the detector response. The E_T^{miss} in QCD events can be predicted by smearing the “seed” events using the resolution functions. The seed events can be obtained either by selecting low E_T^{miss} multi-jet events or by scaling the p_T of jets such that the E_T^{miss} in an event is zero.

Given enough statistics, the detector response can be measured using photon-jet events. In these events, in $2 \rightarrow 2$ scattering approximation, the jet and the photon have the same p_T . Thus detector response/resolution can be measured by balancing the jet p_T with the photon p_T which is, thanks to excellent resolution of CMS ECAL, very accurately measured. Traditionally, two different methods have been used to measure the detector response to jets using photon-jet data: (a) photon-jet p_T balancing method and (b) E_T^{miss} projection fraction MPF method. MPF method uses the whole event and is thus less sensitive to additional activity in the event. On the other hand, it requires a good understanding of the E_T^{miss} which is measured using whole calorimeter and thus is more sensitive to detector noise. In addition, MPF method does not depend on the jet clustering algorithms and corrections must be made to determine the response to a jet from MPF response function. The p_T -balance method directly measures the calorimeter response to a jet but is sensitive to extra activity in the event and thus requires strong event selection cuts.

The detector response to jets mostly follows a Gaussian distribution. There is, however, a non-Gaussian tail which corresponds to severe effects of the detector. The Gaussian part of the response can be estimated by using a γ -jet sample where the p_T^γ is well measured. The non-Gaussian part can be either estimated using the same photon-jet sample if enough statistics are available, otherwise one has to resort to the dijet or multijet events. In these events, the relatively high E_T^{miss} is aligned with one of the jets as the probability of two or more jets fluctuating to a very low response in the same event is very small.

In this section, we describe the measurement of detector response function using MPF method. The results are compared to the *true* response determined using Monte Carlo information. The response is also compared to the response obtained using p_T -balance method. Procedure of measuring the tails in the response functions will be discussed later.

1.2 Missing E_T^{miss} Projection Method

A non-zero E_T^{miss} in an event with well-balanced transverse momentum indicates a different response to one or more the objects in the event than their true value. The derivation of R_{jet} in Refs. [1, 2] can be generalized as follows. If we denote the true transverse momentum of the two objects in the events as p_T^a and p_T^b , then, due to the conservation of the transverse p_T , we have

$$\mathbf{p}_{T,a} + \mathbf{p}_{T,b} = 0. \quad (1)$$

The quantities p_T^a and p_T^b are measured in the calorimeter with different values than their true values, which is the result of calorimeter responses ($R_{a,b}$) to objects a and b . As a result, E_T^{miss} is measured in the form of

$$\mathbf{p}_{T,a}^{\text{meas}} + \mathbf{p}_{T,b}^{\text{meas}} = -\mathbf{E}_T^{\text{miss}}, \quad (2)$$

where $\mathbf{p}_{T,a}^{\text{meas}} = R_a \mathbf{p}_{T,a}$ and $\mathbf{p}_{T,b}^{\text{meas}} = R_b \mathbf{p}_{T,b}$. If we assume that one of the objects, e.g. a , has a well measured p_T ¹⁾, then $\mathbf{p}_{T,a}^{\text{meas}} = \mathbf{p}_{T,a}$. Therefore, Eq.(2) becomes

$$\mathbf{p}_{T,a} + \mathbf{p}_{T,b}^{\text{meas}} = -\mathbf{E}_T^{\text{miss}}. \quad (3)$$

Table 1: Photon ID cuts, [Add photon selection cuts here](#)

Number of tracks
TrackIsolation
Calorimeter Isolation

After multiplying Eq.(3) with $\mathbf{p}_{T,a}$ and normalizing it with $|\mathbf{p}_{T,a}|^2$, we find the following expression for the estimated response of the calorimeter for the object b with the MPF method

$$R_b = 1 + \frac{\mathbf{E}_T^{\text{miss}} \cdot \mathbf{p}_{T,a}^{\text{meas}}}{|\mathbf{p}_{T,a}^{\text{meas}}|^2}. \quad (4)$$

1 For photon-jet events, the photon is well measured i.e. object a is identified with photon. The calorimeter response
 2 to hadronic recoil is given by R_b . In following the R_b is measured in bins of reconstructed photon p_T^γ . Thus the
 3 jet response and resolution are normalized to p_T^γ . The measured resolution is multiplied by $\langle p_T^{\text{Jet}} \rangle / \langle p_T^\gamma \rangle$
 4 for each bin to convert it to the fraction resolution of the jets in that bin. The measured response is parametrized
 5 as a function of the mean p_T^{jet} contributing to that bin. It is important as these response functions will be used to
 6 smear the jets and not the photons.

7 1.3 p_T -Balance Method

8 In this method, after event selection, response function is determining by ratio $p_T^{\text{Jet}}/p_T^\gamma$.

9 1.4 Data Set and Event Selection

10 We used Summer09-MC-31X-V3-v1/GEN-SIM-RECO γ -jet sample which was constructed by using CMSSW_3.1.2.
 11 The γ -jet events were generated with Pythia6 in different \hat{p}_T bins (0-15, 15-20, 20-30, 30-50, 50-80, 80-120, 120-
 12 170, 170-300, 300-500, > 500). The flat ntuples were made using **PAT version xxx**. The jets are corrected using
 13 **xxxx** corrections. The E_T^{miss} is corrected using **Type-1 corrections**. These jet corrections are determined from
 14 QCD dijet sample where the flavor composition of jets is different than the photon-jet sample. Thus at low p_T the
 15 average response p_T^{jet}/ptg is slightly higher than 1 and it approached 1 at large jet p_T .

16 The jets corresponding to the photons are removed from the jet collection. The γ -jet events were required to pass
 17 HLT_Photon25_L1R trigger, and required to have at least one photon with $p_T > 25$ GeV and $|\eta| < 2.4$. The events
 18 passing with the following selection cuts are used in further analysis.

- 19 • No **RECONSTRUCTED** lepton (e, μ) with $p_T > 10$ GeV
- 20 • Only 1 tight photons with $p_T > 50$ GeV and $|\eta| < 2.4$
- 21 • At least 1 jet with $p_T > 30$ GeV and $|\eta| < 2.4$
- 22 • Photon and leading jet are back to back with $|\Delta\phi(\gamma, j)| > 3.0$ radian
- 23 • No additional jet in the event $p_T^{\text{jet2}} < 0.1 \times p_T^\gamma$ or $p_T^{\text{jet2}} < 10$ GeV

24 1.5 Gaussian Response from γ -jet Events using MPF Method

We measured R_G defined by equation 5 in different p_T bins as given in Table 2.

$$R_G = 1 + \frac{\mathbf{E}_T^{\text{miss}} \cdot \mathbf{p}_{T,\gamma}^{\text{meas}}}{|\mathbf{p}_{T,\gamma}^{\text{meas}}|^2}. \quad (5)$$

25 In addition to R_G , following histograms were made for each photon bin.

¹⁾ In the case of a being a photon, it will have a very good energy scale due the resolution of the CMS ECAL and well established methods, such as $Z \rightarrow ee$. Similarly, if a is a barrel jet, the jet energy scale in the barrel will be much precise.

- 1 • p_T^γ
- 2 • p_T^{GenJet}
- 3 • p_T^{CaloJet}
- 4 • $p_T^{\text{CaloJet}}/p_T^{\text{rmGenJet}}$
- 5 • $p_T^{\text{CaloJet}}/p_T^\gamma$

6 The Monte Carlo events from each \hat{p}_T bin are appropriately weighted and combined. Each bin had at least 50

7 entries.

8 The event topology cuts used to suppress events with initial state or final state distributions are shown in Fig. 1.

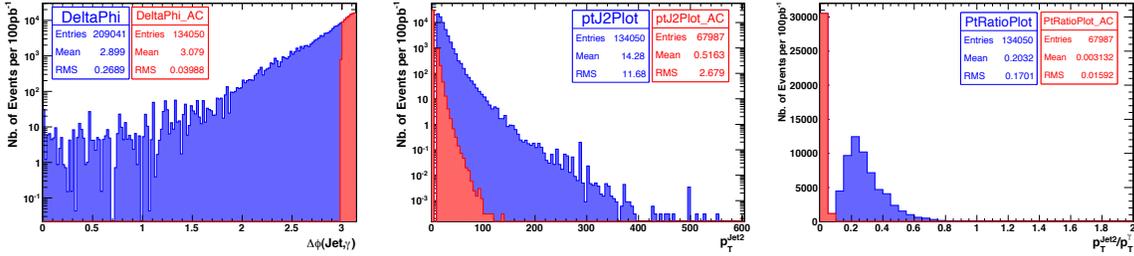


Figure 1: $|\Delta\phi(\gamma, j)|$ (left), $p_T^{\text{jet}2}$ (middle), and $p_T^\gamma/p_T^{\text{jet}2}$ (right). Blue distributions are before the cut and red distributions are after the event selection cuts.

9

Fig. 2 shows these six distributions from a $100\text{GeV} < p_T^\gamma < 105\text{GeV}$ bin used. The mean and the width of the R_G distribution was determined by fitting a Gaussian function for each p_T^γ bin. The mean of p_T^{CaloJet} and mean of p_T^γ was determined for each bin as given by ROOT. Using this procedure, the fractional resolution is defined with respect to p_T^γ where as the true resolution is measured with respect to the GenJet p_T . Moreover the jet p_T is smeared to predict the QCD E_T^{miss} spectrum. The resolution is redefined with respect to CaloJet p_T by scaling the resolution by ratio $\langle p_T^\gamma \rangle / \langle p_T^{\text{rmCaloJet}} \rangle$ where $\langle p_T^\gamma \rangle$ and $\langle p_T^{\text{rmCaloJet}} \rangle$ are the p_T of photon and p_T of calorimeter jet contributing to a given bin. The measured resolutions are parametrized as a function of p_T^{jet} using equation 6.

$$(\sigma_{R_G})^2 = p_0^2 + \left(\frac{P_1}{\sqrt{p_T}}\right)^2 + \left(\frac{P_2}{p_T}\right)^2 \quad (6)$$

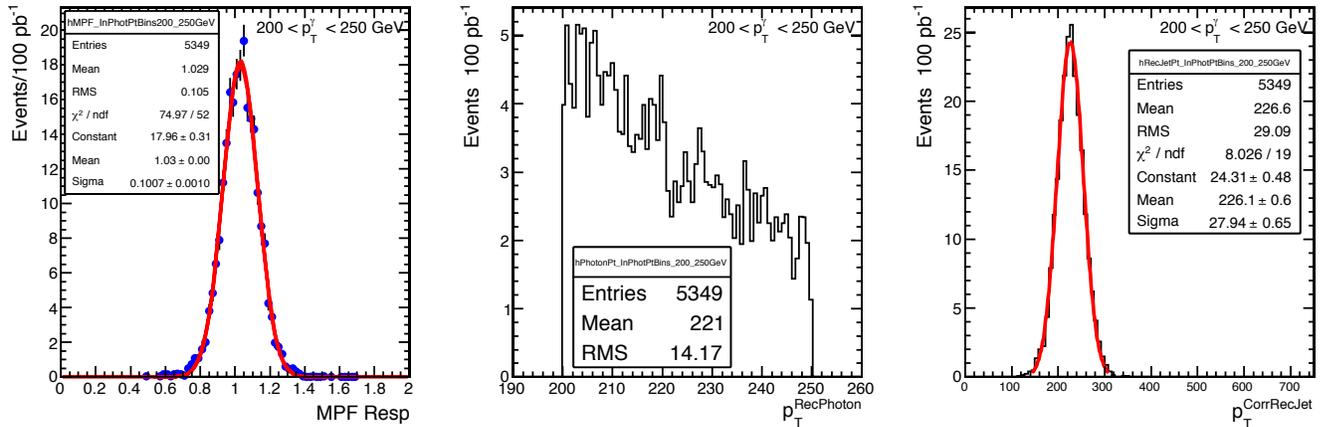


Figure 2: Measured R_G by MPF method (left), p_T of RecPhoton (middle) and p_T^{jet} (right) for a $100\text{ GeV} < p_T^\gamma < 105\text{ GeV}$ p_T^γ bin.

Table 2: Calorimeter to hadronic recoil using Missing p_T Projection Fraction (MPF) method **Please check the numbers. You have to fix the last row and last column.**

p_T^γ range	$\langle p_T^\gamma \rangle$	$\langle p_T^{\text{GenJet}} \rangle$	$\langle p_T^{\text{CaloJet(JES-cor)}} \rangle$	Response					
				CaloJet(JES-cor)/GenJet		MPF		$p_T^\gamma/p_T^{\text{CaloJet}}$	
				mean	$\sigma(\text{G fit})$	mean	$\sigma(\text{G fit})$	mean	$\sigma(\text{G fit})$
75-80	77.4	75.32	82.01	1.09	0.155	1.074	0.169		
80-85	82.3	79.81	86.7	1.09	0.147	1.074	0.164		
85-90	87.4	84.6	91.8	1.08	0.144	1.073	0.170		
90-95	92.4	90.37	97.6	1.08	0.135	1.074	0.135		
95-100	97.4	95.86	103.2	1.08	0.136	1.076	0.161		
100-105	102.4	100.2	107.9	1.08	0.141	1.070	0.138		
105-110	107.4	106.3	113.9	1.07	0.134	1.074	0.134		
110-120	114.9	113.2	119.8	1.06	0.131	1.055	0.137		
120-130	124.8	123.2	130.2	1.06	0.118	1.058	0.132		
130-140	134.9	133.7	140.7	1.05	0.113	1.054	0.125		
150-160	154.9	154.0	161.6	1.05	0.105	1.052	0.119		
160-170	164.7	163.5	171.2	1.04	0.103	1.043	0.112		
170-200	183.4	182.0	188.3	1.04	0.105	1.037	0.111		
200-250	221.0	220.3	226.1	1.03	0.092	1.030	0.101		
250-300	271.2	270.4	276.2	1.02	0.083	1.023	0.086		
300-400	337.1	336.1	3441.3	1.02	0.078	1.013	0.841		
400-600	458.8	457.6	450.3	1.01	0.070	1.006	0.073		
600-2000	709.9	710.4	712.5	1.00	0.063	1.002	0.060		

- 1 The influence of different cuts on $\Delta\phi(\gamma, jet)$ and $p_T^{\text{jet}2}$ to the MPF resolution can be seen in Fig. 4. The effect is
2 shown for two different photon p_T bins. Lower $\Delta\phi(\gamma, jet)$ cut and higher $p_T^{\text{jet}2}$ cuts give a larger value of σ and
3 the effect is more visible for the lower p_T range. And also fit parameters for different $\Delta\phi(\gamma, jet)$ and $p_T^{\text{jet}2}$ cuts
4 can be seen in Table 3 and Table 4

	$\Delta\phi(\gamma, jet) > 2.8$	$\Delta\phi(\gamma, jet) > 2.9$	$\Delta\phi(\gamma, jet) > 3.0$
p_0	0.2984 ± 0.0035	0.03002 ± 0.0034	0.02933 ± 0.0036
p_1	1.352 ± 0.041	1.350 ± 0.041	1.364 ± 0.042
p_2	7.621 ± 0.798	7.552 ± 0.807	6.763 ± 0.932

Table 3: Fit parameters for different $\Delta\phi(\gamma, jet)$ cut. The individual sigma values for two photon p_T bins are given in Fig. 4 (left).

	$p_T^{\text{jet}2} < 10 \text{ GeV}$	$p_T^{\text{jet}2} < 12 \text{ GeV}$	$p_T^{\text{jet}2} < 15 \text{ GeV}$	$p_T^{\text{jet}2} < 18 \text{ GeV}$	$p_T^{\text{jet}2} < 20 \text{ GeV}$
p_0	0.0293 ± 0.004	0.03103 ± 0.0033	0.3214 ± 0.0031	0.03319 ± 0.0030	0.03337 ± 0.029
p_1	1.364 ± 0.042	1.338 ± 0.042	1.316 ± 0.041	1.284 ± 0.041	1.275 ± 0.040
p_2	6.763 ± 0.932	7.4 ± 0.807	8.076 ± 0.689	9.28 ± 0.584	9.637 ± 0.541

Table 4: Fit parameters for different $p_T^{\text{jet}2}$ cut. The individual sigma values for two photon p_T bins are given in Fig. 4 (right).

5 1.6 Comparison of MPF Resolution with MC true Response

- 6 MC response ($p_T^{\text{recoJet}}/p_T^{\text{particleJet}}$) is measured in different particle jet (photon) p_T bins and mapped to the JEC
7 p_T^{jet} . The events selections listed in Section 1.5 are followed and particle jet (photon) p_T is used as an estimator.
8 For the events passed the selections criteria particle jets matched to the leading reconstructed jets **within R=0.3**.
9 The MC response was also measured in photon-jet sample without any event topology cuts. Comparison of two
10 measurements is shown in Fig. 5. It shows MC resolutions measured in particle jet (photon) p_T bins and mapped
11 to the JEC p_T^{jet} are in good agreement with the direct measurement without any event topology cuts.

- 12 The data driven MPF response is compared to true response in Fig. 6. Both the mean and resolution of two
13 measurements are close but not exactly same. It shows that any corrections needed to map the response measured

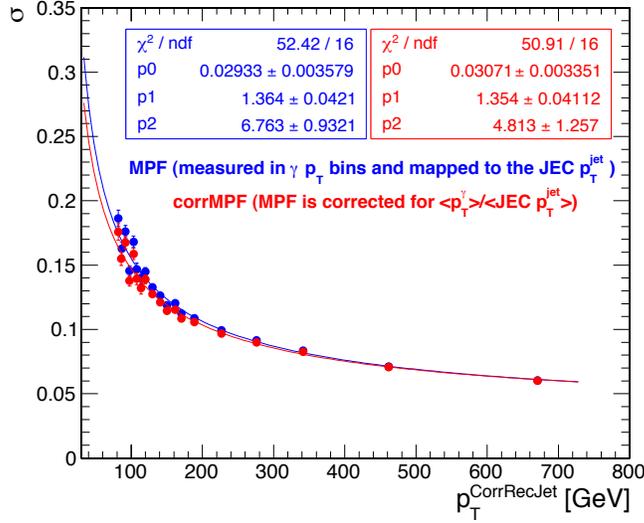


Figure 3: The resolution of MPF (blue) and corrected MPF (red) Gaussian responses as a function of JEC p_T^{jet} .

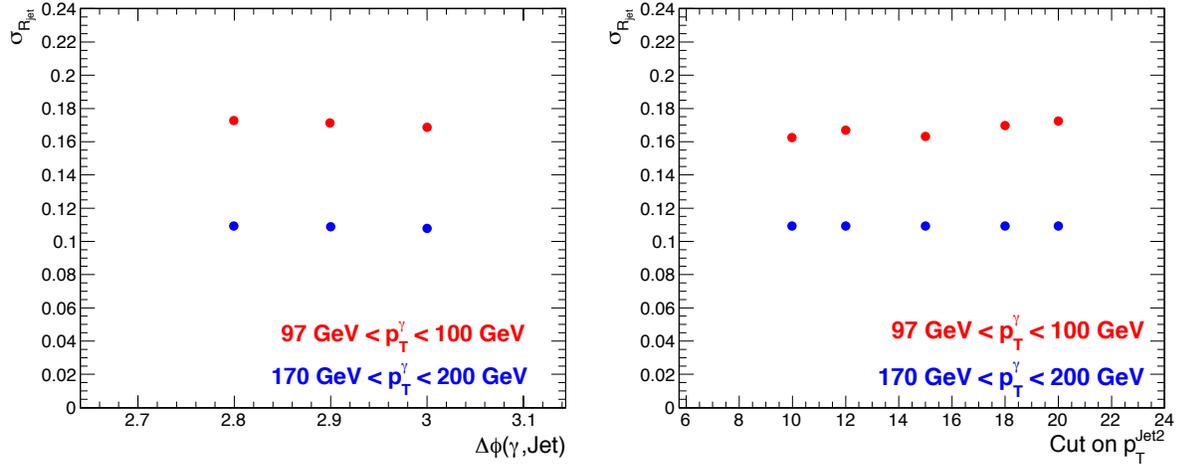


Figure 4: The effect of $\Delta\phi(\gamma, jet)$ (left) and $p_T^{\text{jet}2}$ (right) cuts on two different (90-95 GeV, 300-400 GeV) photon p_T bins.

- 1 using MPF to the “true” response are small. We are investigating the reason for these small differences and
- 2 finalizing the results.

3 1.7 Selection of Seed Events for Fake E_T^{miss} Calculation

4 Reason: To find smeared E_T^{miss} and prepare event group for final step.

5 We used QCD MadGraph (100-250, 250-500, 500-1000, 1000-inf H_T bins) ntuples produced with NT7 tags from
6 V5 PAT Layer1 samples [5]. Following criteria were used to select events to form the sample in the analysis:

- 7 • HLT_Jet110 trigger,
- 8 • no lepton (e, μ) with $p_T > 10$ GeV,
- 9 • no photons with $p_T > 25$ GeV,
- 10 • at least three jets $p_T > 50$ GeV and $|\eta| < 2.4$,
- 11 • $E_T^{\text{miss}} < 20$ GeV,

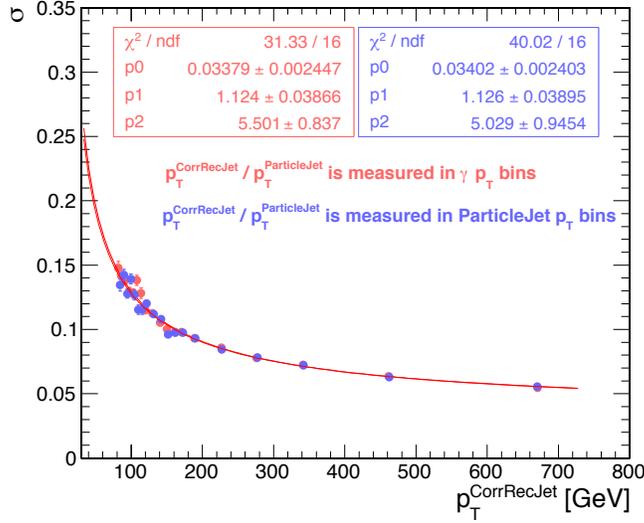


Figure 5: The resolution of MC Gaussian response function as a function of JEC p_T^{jet} .

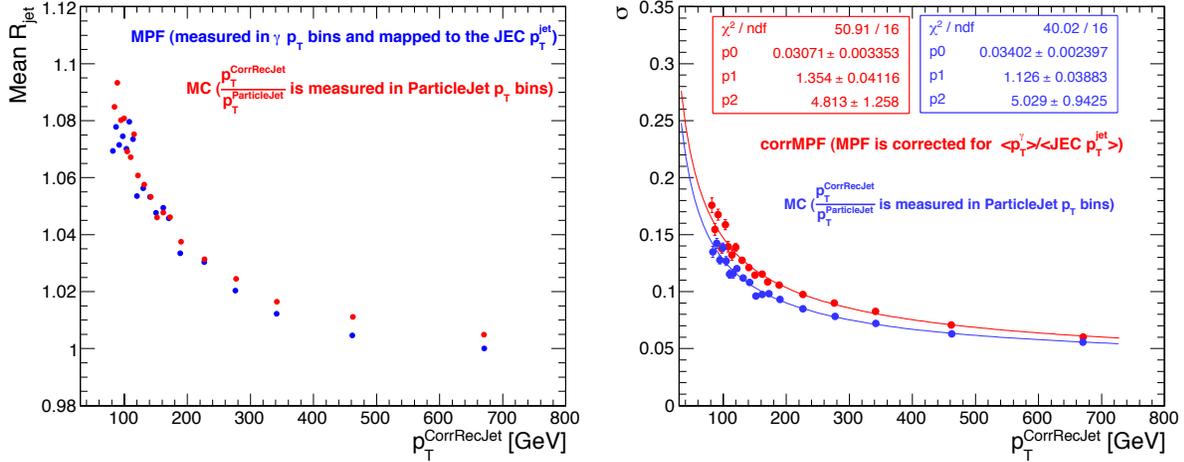


Figure 6: Mean R_{jet} vs JEC p_T^{jet} (left) and $\sigma_{R_{jet}}$ vs JEC p_T^{jet} (right) for MPF and MC measurements.

As a next step for a given jet, σ is calculated by using fit parameters from Section 1.5 and the jet p_T . Then a random number (smearing constant) from a Gaussian distribution with this σ is found. Smearing constant is applied to the p_T of jets in seed events and E_T^{miss} was changed accordingly. The jets are smeared with the formula

$$p^{\text{smeared}}(j) = c \times p(j) \quad (7)$$

where p is the four momentum vector of the jet and c is the smearing constant which was obtained randomly from the smearing function for each jet. The E_T^{miss} was recalculated with the formula

$$E_T^{\text{Estimated}} = E_T^{\text{miss}} + \sum_{i=0}^{N_{jets}} (p_T^{\text{smeared}}(j_i) - p_T(j_i)) \quad (8)$$

- 1 Then smeared seed events were normalized to QCD data. E_T^{miss} distribution for data and estimation can be seen in
- 2 Fig. 7. Estimated E_T^{miss} spectrum reaches up to 240 GeV. There is 10% overestimation and underestimation up to
- 3 140 GeV E_T^{miss} and above that the underestimation goes up to 50% and higher with increasing E_T^{miss} value.
- 4 The influence of selecting seed events can be seen in Fig. 8. Three different seed event groups are created with
- 5 three different E_T^{miss} cut (25, 20 and 15 GeV) and those seed events are smeared. Estimated E_T^{miss} is normalized
- 6 to observed E_T^{miss} distribution.

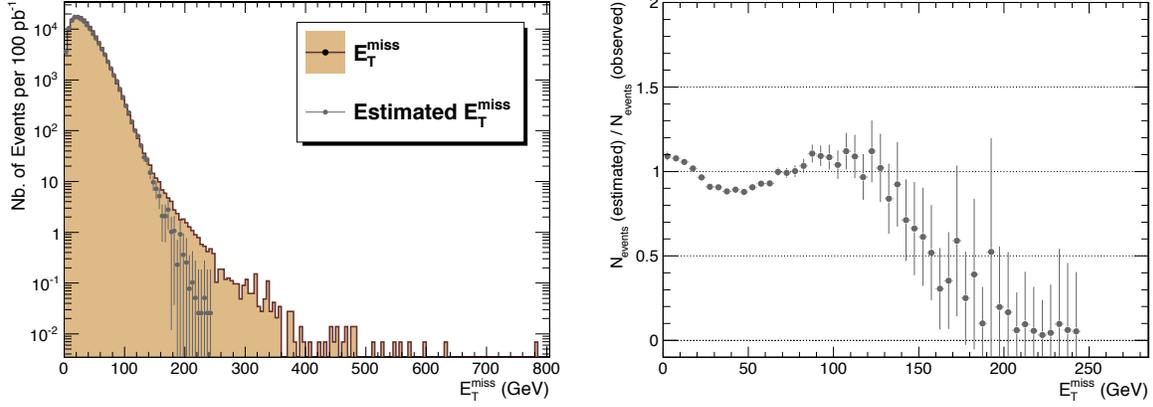


Figure 7: (Top) Estimated and observed E_T^{miss} distribution for QCD events with $N_{\text{jets}} \geq 3$, after SUSY jet selections ($p_T > 180, 150$ and 50 GeV). (Bottom) Ratio of estimated and observed E_T^{miss} .

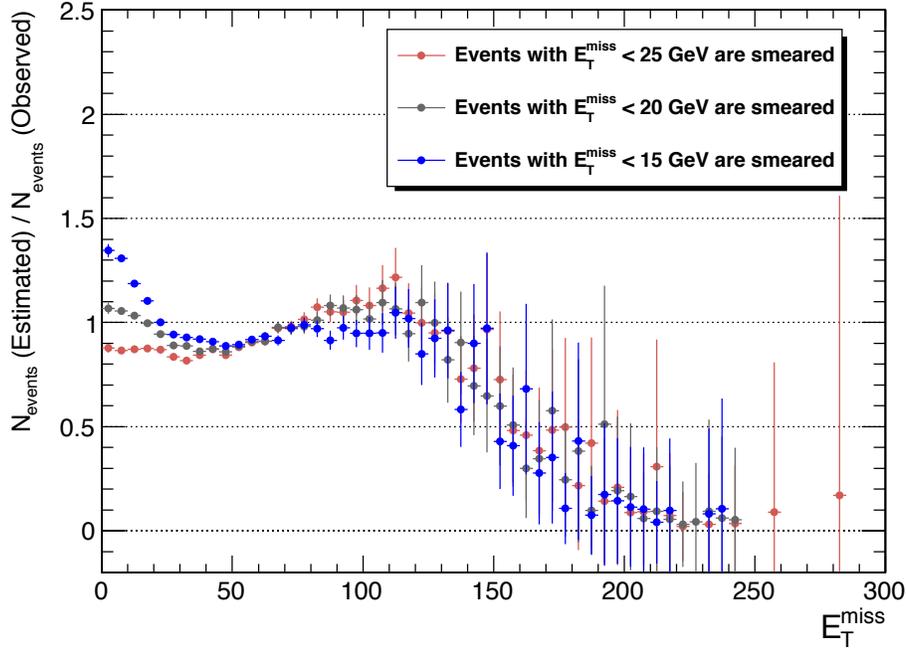


Figure 8: The effect of seed event selection on estimated E_T^{miss} distribution.

- 1 The result of this study shows that to be able to make an accurate estimation for E_T^{miss} tail we need to include a
- 2 non-Gaussian tail into our response measurements.

3 References

- 4 [1] F. Abe et al., Phys. Rev. Lett. 69 (1992) 2896.
- 5 [2] B. Abbott et al., Nucl. Instrum. Meth. A 424 (1999) 352394.
- 6 [3] G. Aad et al. (The ATLAS Collaboration), "Expected Performance of the ATLAS Experiment, Detector, Trig-
- 7 ger and Physics", arXiv:0901.0512.
- 8 [4] The official SUSY analysis package, SusyAna, is accessible at:
- 9 <https://twiki.cern.ch/twiki/bin/view/CMS/SusyPat>

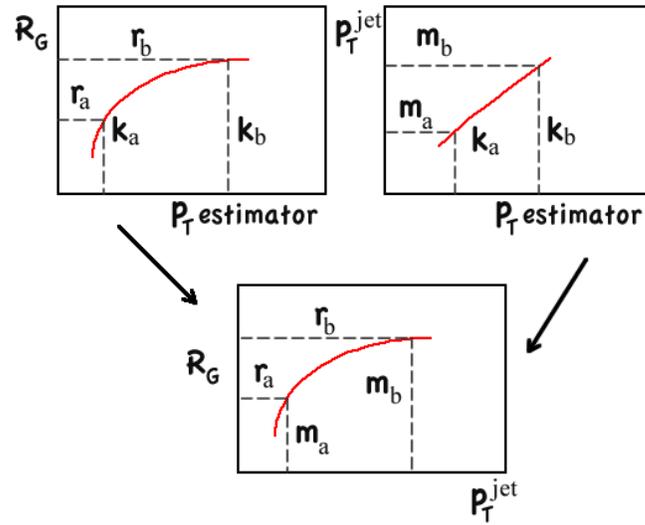


Figure 9: A sketch of R_{jet} derivation by using p_T estimator bins. First R_{jet} (top left) and p_T^{jet} (top right) are measured in p_T estimator bins. Then, R_{jet} is plotted against p_T^{jet} by using previous plots(bottom).

1 [5] SusyPatLayer1 Production:

2 <https://twiki.cern.ch/twiki/bin/view/CMS/SusyPatLayer1>