Missing Transverse Energy Significance with CMS

Jim Alexander, Lawrence Gibbons, Aleko Khukhunaishvili
Cornell University
Presented by Dayong Wang, University of Florida

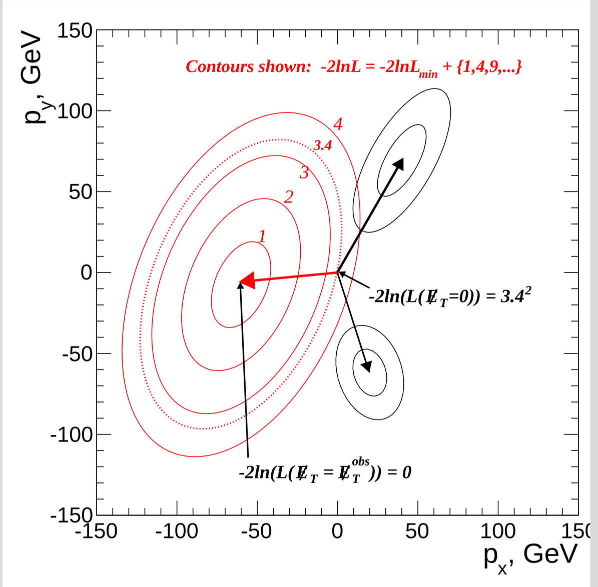
Abstract

The missing transverse energy significance assesses, on an event-by-event basis, the likelihood that the observed Missing Transverse Energy (MET) is consistent with a fluctuation from zero because of detector-related limitations, like measurement resolution. We present the formal definition of the significance and show performance studies of particle-flow MET significance in dijet and $W \to e\nu$ data samples collected with the CMS detector.

Definition

The determination of the significance requires evaluation of the uncertainty on the total measured transverse energy: $\vec{E}_T^{total} = \sum \vec{E}_{T_i} = -\vec{\not{E}}_T$.

Given the true transverse momentum $\vec{e}_{T_i}$, the reconstructed transverse momentum $\vec{E}_{T_i}$ and their difference $\vec{\varepsilon}_i = \vec{E}_{T_i} - \vec{e}_{T_i}$ for each reconstructed object, the likelihood of observing a total transverse momentum $\vec{\varepsilon}$ under the null hypothesis, for two momentum vectors, is given by:



$$\mathcal{L}(\vec{\varepsilon}) = \int p_1(\vec{\varepsilon}_1|\vec{e}_{T_1}) p_2(\vec{\varepsilon}_2|\vec{e}_{T_2}) \delta(\vec{\varepsilon} - (\vec{\varepsilon}_1 + \vec{\varepsilon}_2))\, d\vec{\varepsilon}_1\, d\vec{\varepsilon}_2.$$

For an arbitrary number of momentum vectors, the likelihood is obtained by applying this equation recursively. The significance is then defined to be the log-likelihood ratio:

$$\mathcal{S} \equiv 2\ln\left(\frac{\mathcal{L}(\vec{\varepsilon} = \sum \vec{\varepsilon}_i)}{\mathcal{L}(\vec{\varepsilon} = 0)}\right)$$

This definition is completely general and will accommodate any probability density function.

The Gaussian Case

The Gaussian probability density function is given by:

$$p_i(\vec{\varepsilon}_i|\vec{e}_{T_i}) \sim \exp\left(-\frac{1}{2}\vec{\varepsilon}_i^{\,T} \mathbf{V}_i^{-1} \vec{\varepsilon}_i\right),$$

where $\mathbf{V}_i$ is the 2×2 covariance matrix associated with i-th object.

For many contributing vectors, the integration yields the following expression for the overall likelihood:

$$\mathcal{L}(\vec{\varepsilon}) \sim \exp\left(-\frac{1}{2}(\vec{\varepsilon})^T \left(\sum_i \mathbf{V}_i\right)^{-1} (\vec{\varepsilon})\right)$$

The covariance matrix $\mathbf{U}_i$ for each reconstructed momentum vector is initially specified in a natural coordinate system having one axis aligned with the measured $\vec{E}_{T_i}$ vector:

$$\mathbf{U}_i = \begin{pmatrix} \sigma^2_{E_{T_i}} & 0 \\ 0 & E^2_{T_i}\sigma^2_{\phi_i} \end{pmatrix}.$$

We adopt the simplifying assumption that the E_T and ϕ measurements are uncorrelated. This matrix is rotated into the CMS $x - y$ reference frame with the rotation matrix $R(\phi_i)$ to give $\mathbf{V}_i$.

Combining all equations yields the final working expression for the significance:

$$\mathcal{S} = \left(\sum_{i \in X} \vec{E}_{T_i}\right)^T \left(\sum_{i \in X} R(\phi_i)\mathbf{U}_i R^{-1}(\phi_i)\right)^{-1} \left(\sum_{i \in X} \vec{E}_{T_i}\right)$$

Implementation

We focus on the calculation of particle-flow MET significance, $\mathcal{S}_{\rm PF}$. The $\vec{\varepsilon}$ sums are taken over the following objects:

- isolated muons (using tracking resolutions)
- isolated electrons (using electromagnetic calorimeter resolutions)
- jets (using jet resolutions)
- remaining particles (using particle resolutions according to particle type)

Resolutions for jets are obtained from Monte Carlo simulations by matching generated and reconstructed jets.

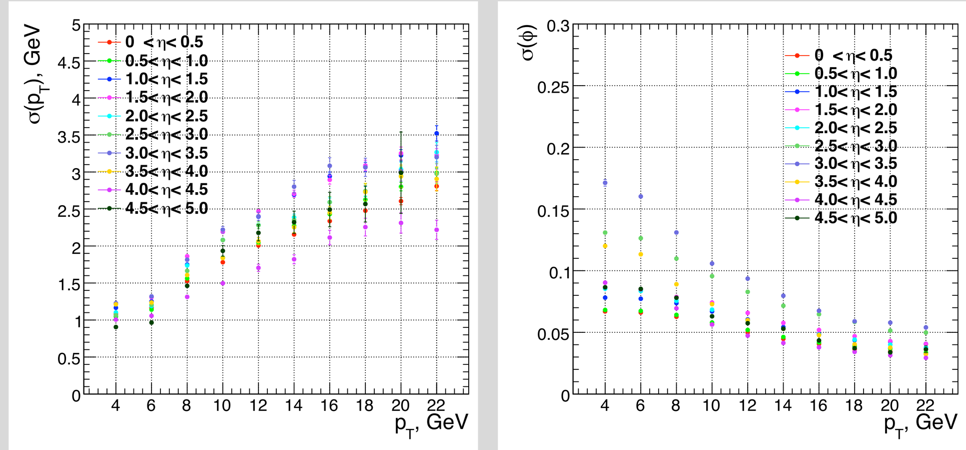


Figure: Low transverse momentum jet p_T and ϕ resolutions.

Jet resolutions naturally take into account fluctuations in the fraction of the total particle content that has been reconstructed, as well as resolution effects due to spuriously reconstructed particles, both of which contribute to the particle-flow $\not{E}_T$ uncertainty.

$\mathcal{S}_{\rm PF}$ Performance in Dijet Events

In the Gaussian case discussed above, $\mathcal{S}_{\rm PF}$ is simply a χ^2 with two degrees of freedom. Therefore, it should exhibit a flat probability of χ^2, $\mathcal{P}(\chi^2)$, in a sample containing events that have a true MET of zero. We select dijet events by requiring a primary vertex near the center of the detector ($|z| < 24$ cm, $\rho < 2$ cm), a single jet trigger with a threshold of 30 GeV and at least two jets with $p_T > 60$ GeV and $|\eta| < 2.3$.

Performance Study

The $\mathcal{S}_{\rm PF}$ distribution closely follows the ideal exponential behaviour, and the $\mathcal{P}(\chi^2)$ distribution is close to uniform.

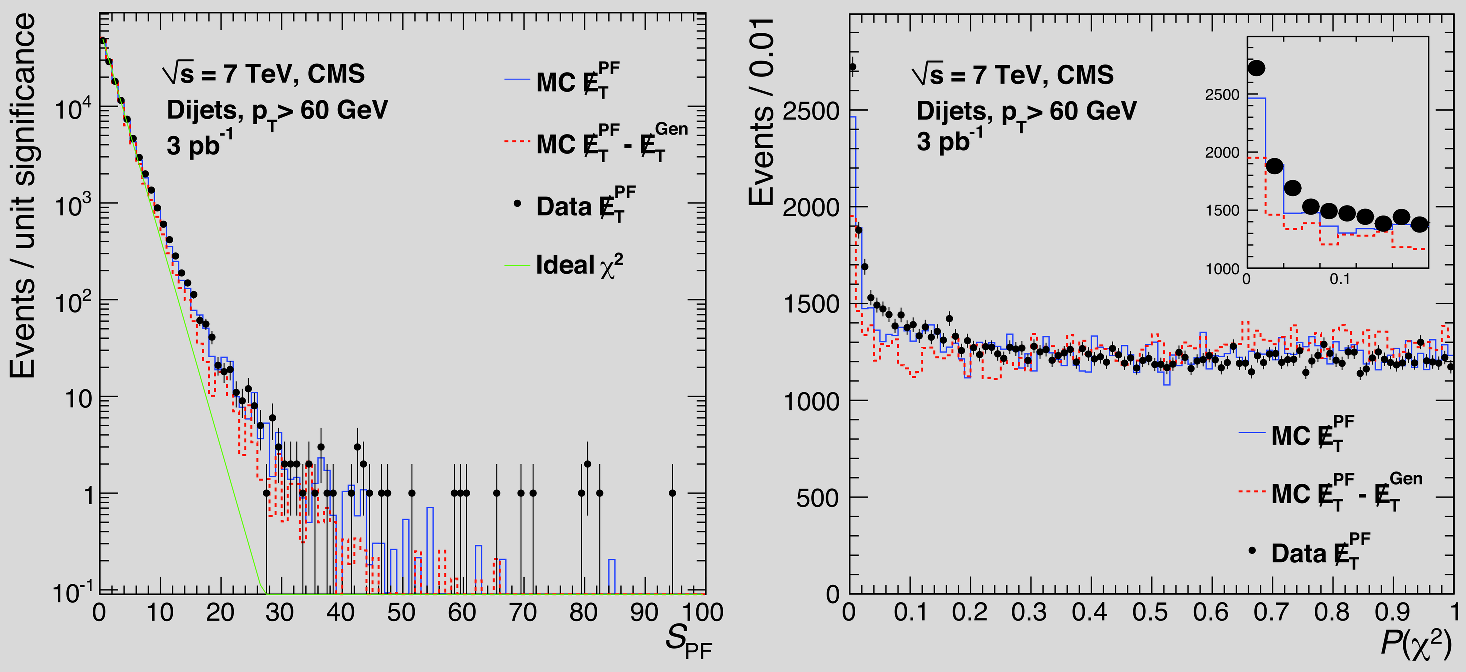


Figure: $\mathcal{S}_{\rm PF}$ and $\mathcal{P}(\chi^2)$ distributions in di-jet events.

About half of the peak at zero in $\mathcal{P}(\chi^2)$ arises from the events containing true MET, either from physics or from the finite detector acceptance. The remainder of the excess at low probability are events that have at least one high-p_T jet whose response is in the non-Gaussian tail.

$\mathcal{S}_{\rm PF}$ in Pile-Up Events

A powerful feature of the $\not{E}_T$ significance is that its distribution is quite insensitive to pileup for events with no true MET.

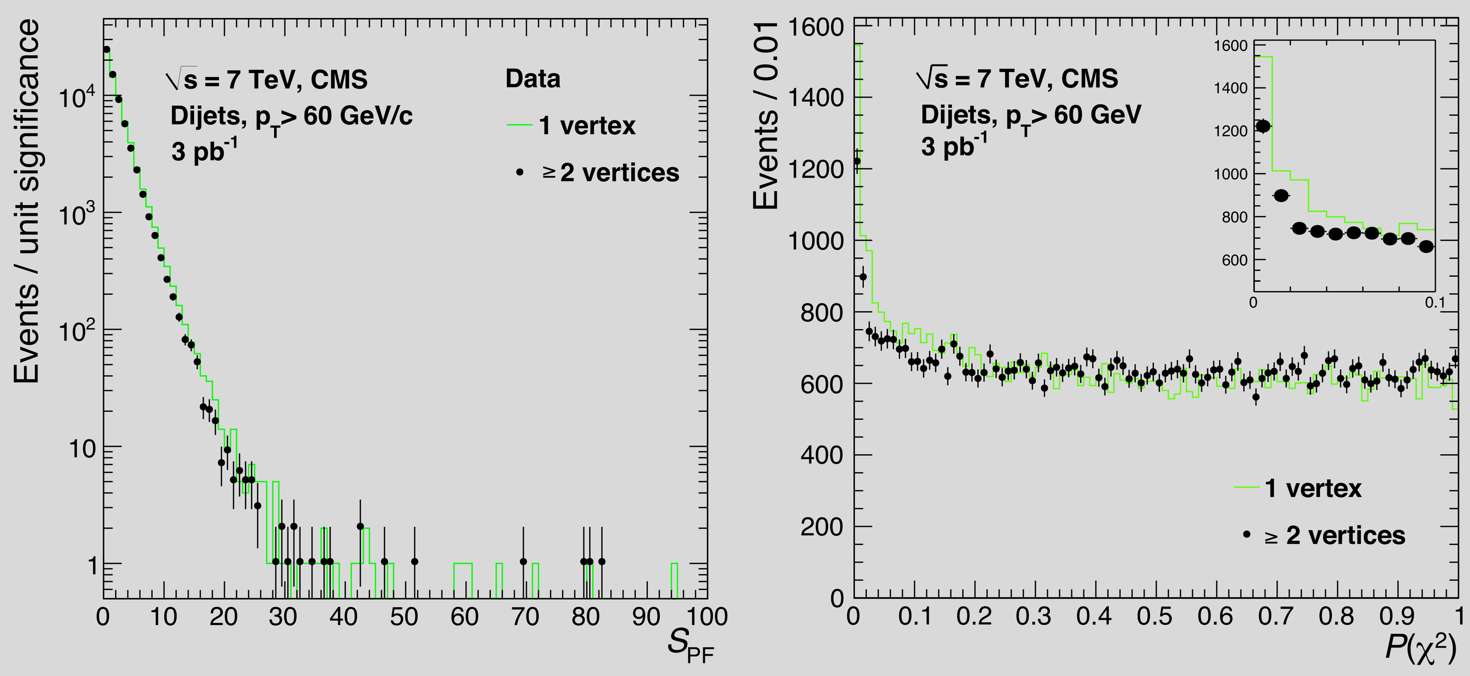


Figure: $\mathcal{S}_{\rm PF}$ and $\mathcal{P}(\chi^2)$ in data for the events with and without pile-up.

As long as the correct resolutions are input, the significance should still have a pure exponential behaviour with a uniformly distributed $\mathcal{P}(\chi^2)$.

Application to $W \to e\nu$ Events

As a case study, we examine the potential gain from using the significance variable as part of the selection criteria for a $W \to e\nu$ analyses. We use criteria similar to those used in the W cross section measurement, which uses an electron selection criteria which is 80% efficient. One analysis option would be to relax the electron isolation to a 95%-efficient criterion and introduce the $\not{E}_T$ significance to help control backgrounds.

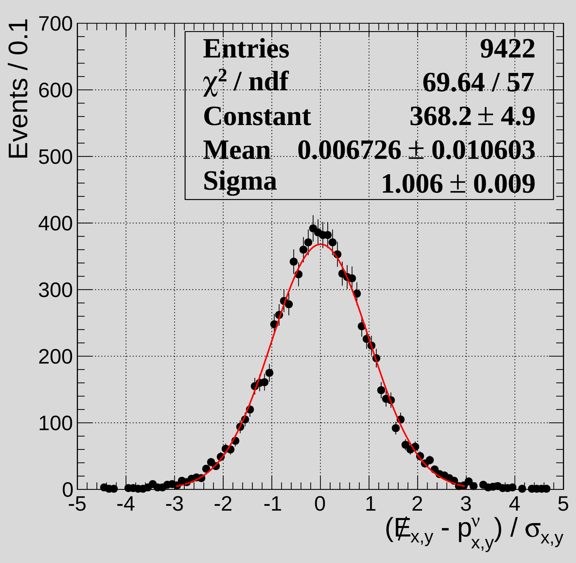
Pull Distribution

As a check of the covariance matrix elements, we look at the distributions of the pull variable:

$$\frac{\not{E}_{x,y} - p^{\nu}_{x,y}}{\sigma_{x,y}}$$

in the Monte Carlo $W \to e\nu$ sample.

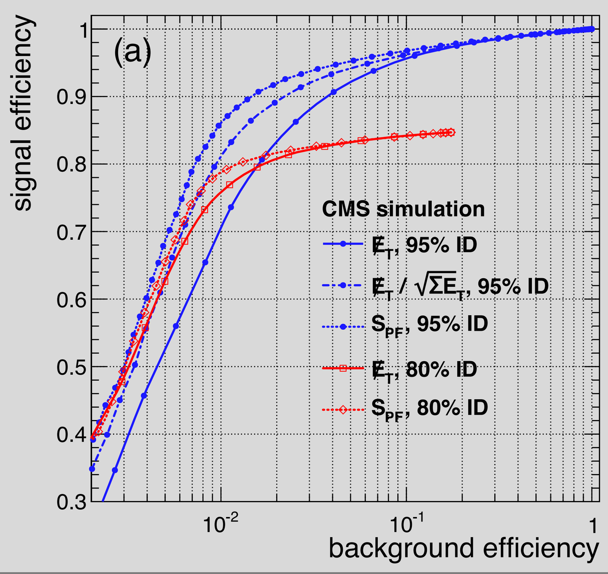
The width of the pull distribution is close to unity.



Efficiency vs Rejection

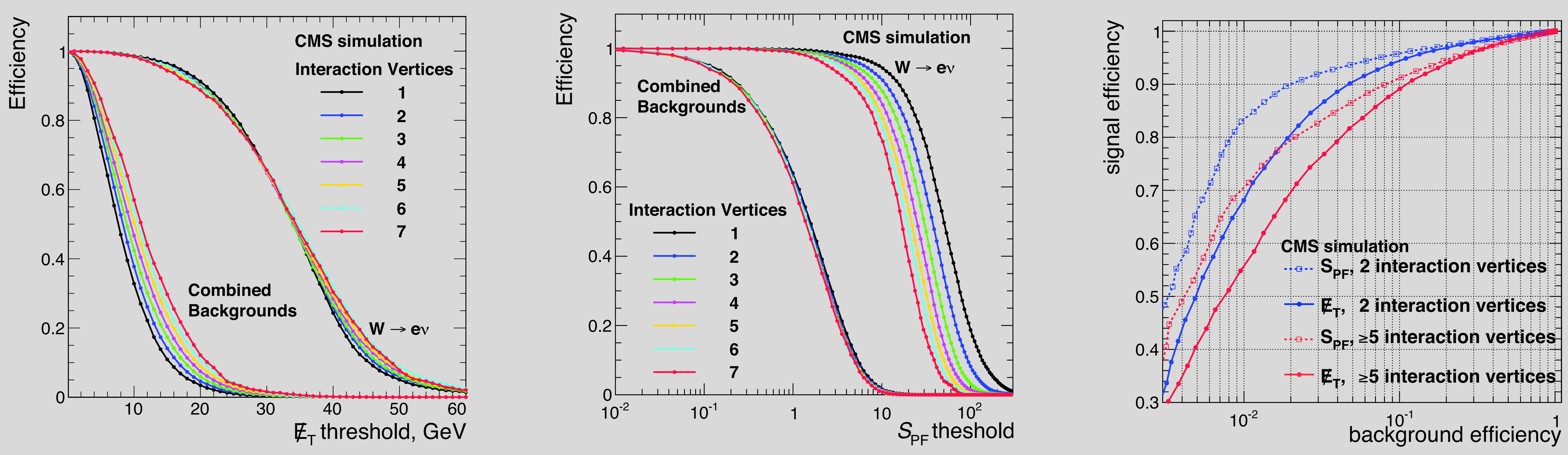
Efficiency for signal versus background for increasing thresholds on $\not{E}_T$, $\mathcal{S}_{\rm PF}$, and $\not{E}_T/\sqrt{\sum E_{T_i}}$ with either tight or loose electron selection criteria.

Use of $\mathcal{S}_{\rm PF}$ with looser electron selection outperforms all other combinations for background rejection at a given signal efficiency.



Pile-Up in $W \to e\nu$ Events

We study the behaviour of signal and total background efficiencies for minimum $\not{E}_T$ or $\mathcal{S}_{\rm PF}$ thresholds and for different numbers of interaction vertices (pile-up) in simulation. The multijet and γ+jets backgrounds, which have no genuine $\not{E}_T$, dominate. The background contribution at higher $\not{E}_T$ grows as pile-up increases, while the $\mathcal{S}_{\rm PF}$ levels remain quite stable.



Differentiation of signal from background degrades for both $\not{E}_T$ and $\mathcal{S}_{\rm PF}$ as pile-up increases. Regardless of the amount of pile-up, however, $\mathcal{S}_{\rm PF}$ always provides a superior signal-to-background ratio compared to $\not{E}_T$.

Data-MC Comparison

Finally, we compare the significance distributions for data and Monte Carlo and see that the agreement is good.

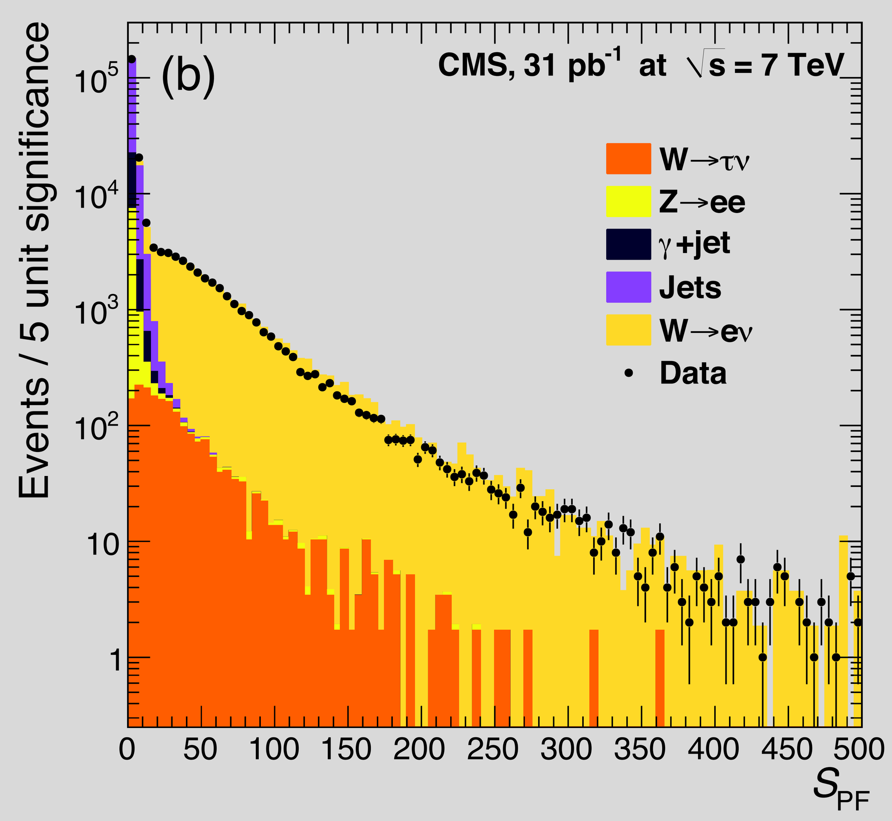


Figure: $\mathcal{S}_{\rm PF}$ distribution in data and MC.

As expected, the backgrounds without genuine $\not{E}_T$ tend to have low values of $\mathcal{S}_{\rm PF}$ while signal events, which have real MET, extend to high values of $\mathcal{S}_{\rm PF}$.

References

[1] CMS Collaboration, *MET Performance in pp Collisions at $\sqrt{s} = 7$ TeV*, arXiv:0271575 (2011).

[2] CMS Collaboration, *Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV*, JHEP 01 (2011) 80. doi:10.1007/JHEP01(2011)080.

 E-mail:ak427@cornell.edu