Pileup mitigation techniques for MissingET at CMS

Matthieu Marionneau

PU Mitigation workshop @ CERN
May, 18th 2014

Introduction

The missing transverse energy (MET) is the most complicated object we can reconstruct in CMS or ATLAS

- all detector signals are combined
- sensitive to noise, object mis-reconstruction
- sensitive to the pileup, by adding extra-energy fluctuations

What is done in CMS to mitigate the pileup impact on MET ?

- definition
- detector level improvement and general corrections
- pileup mitigation algorithms
- the close and far future

MissingET in CMS : definition

Use of a global event description technique (also called "particle-flow") :

- unique interpretation of each event, with the reconstruction of an exhaustive list of particles using all the informations collected in the detector

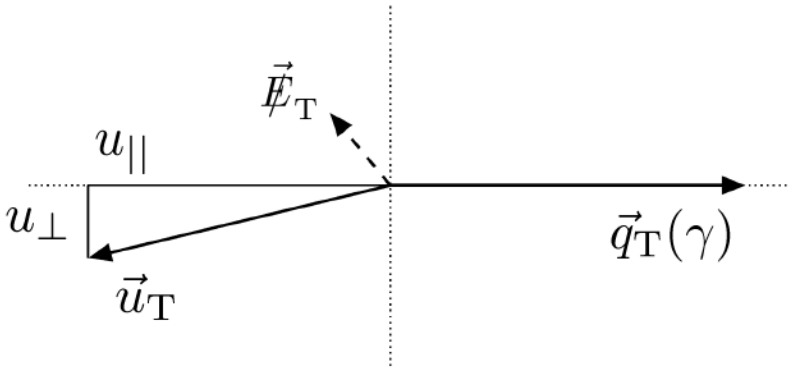
Definition : $\vec{\not{E}}_T = -\sum_{\text{Reconstructed particles}} \vec{p_T}$ + jet energy scale corrections propagated to MET

MET scale and resolution computed in Z/γ events using the hadronic recoil : $\vec{u}_T = -\vec{q}_T - \vec{\not{E}}_T$

- q_T = transverse momentum of Z or γ
- q_T known with accuracy → MET resolution and energy scape dominated by the precision of the recoil measurement

Decomposition of the recoil gives :

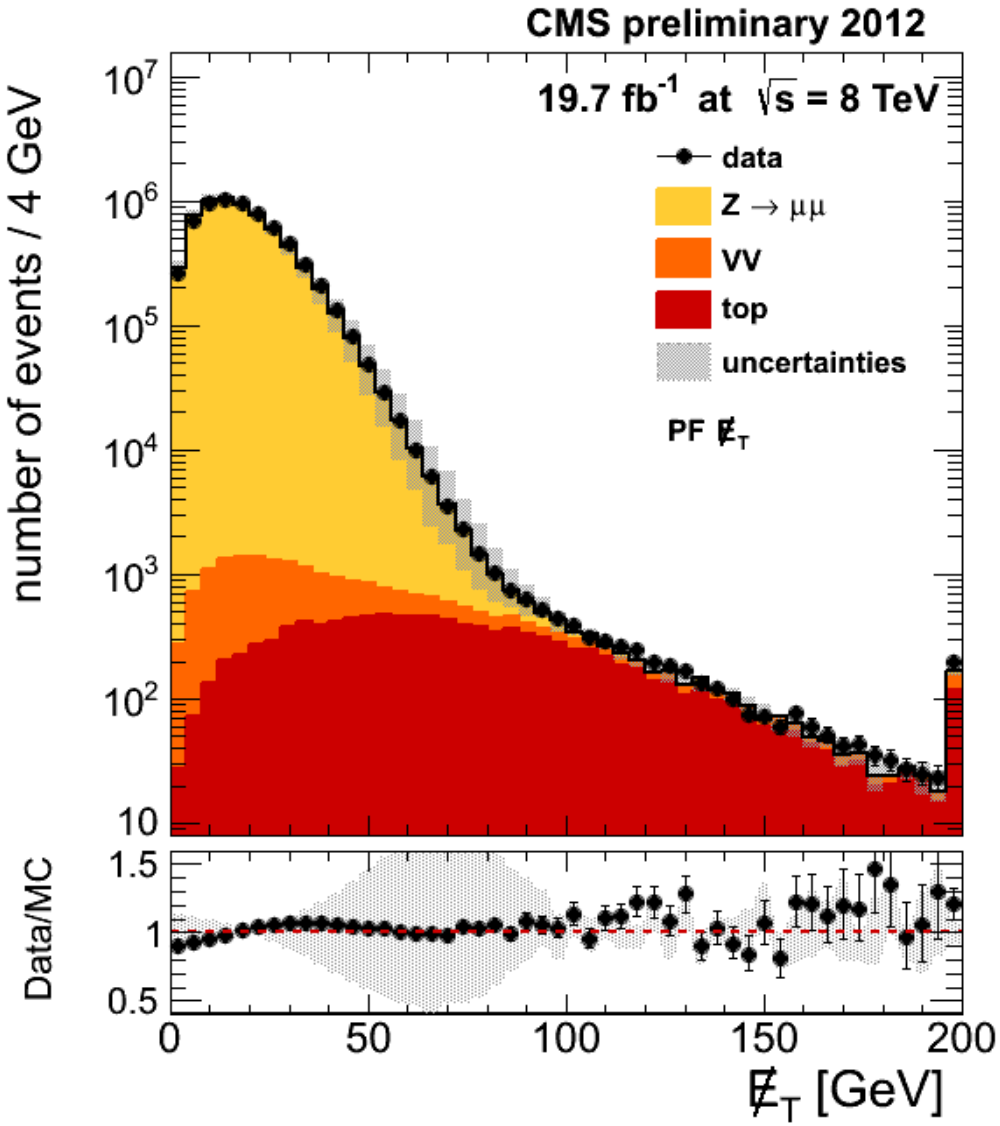
- $u_{\parallel}$ dominated by jets
- $u_{\perp}$ dominated by the underlying event, noise and pileup



MissingET in CMS : modeling in simulation

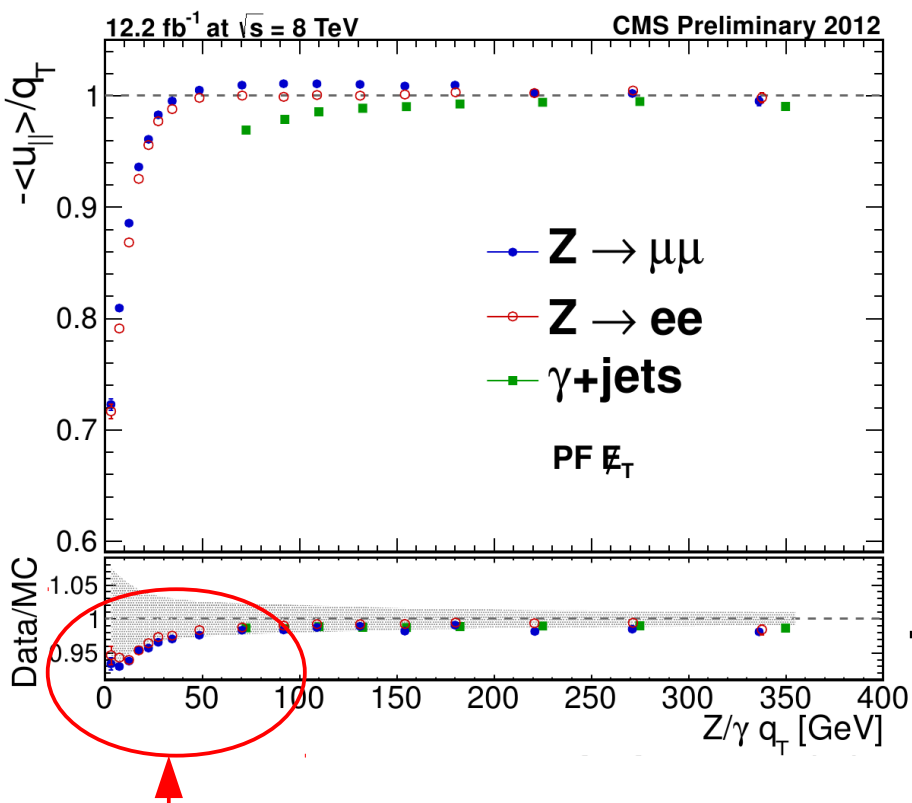
Good overall modeling of the MET in the simulation

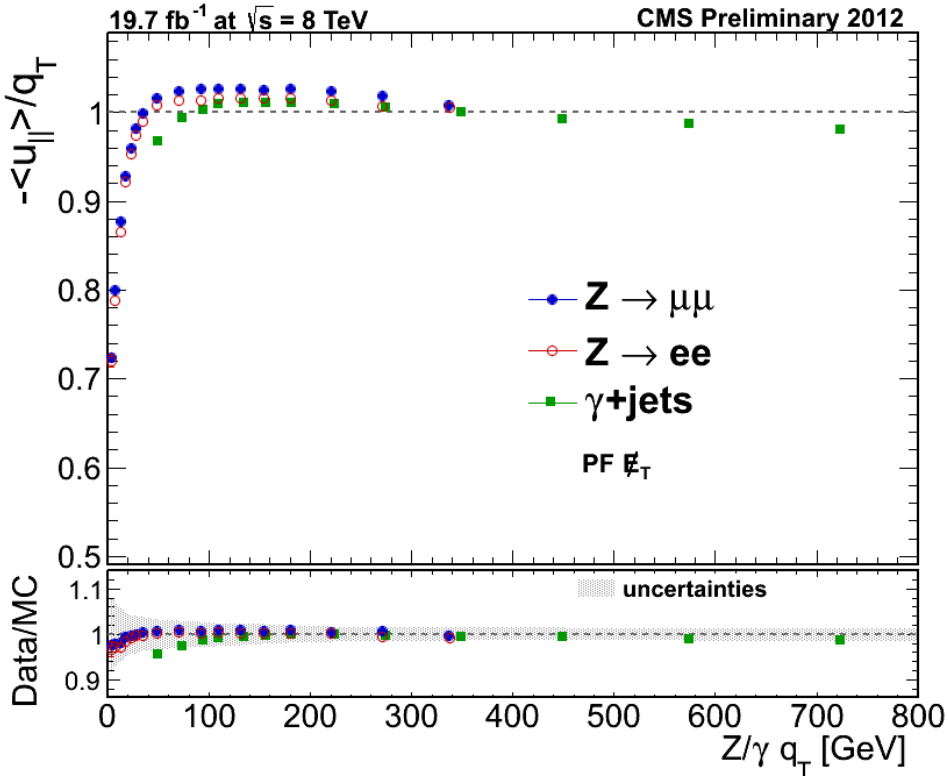
Uncertainties dominated by the jet energy resolution and unclustered energy scale uncertainties



MissingET scale and pileup

In-time pileup is uniformly distributed, and since 2010 no real impact has been seen on the MET scale



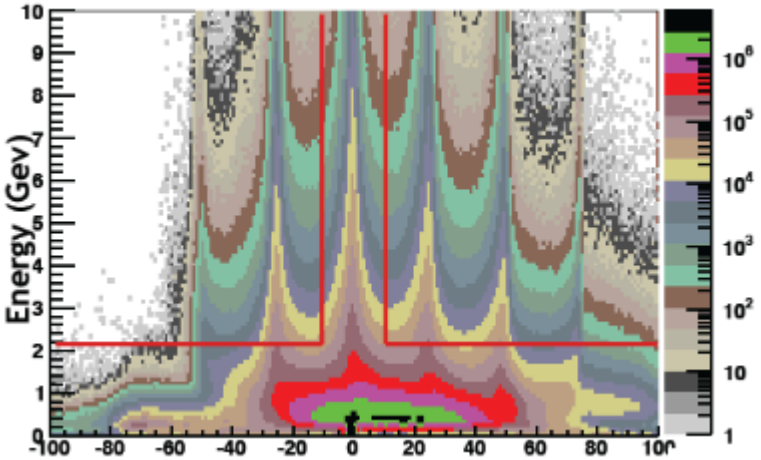


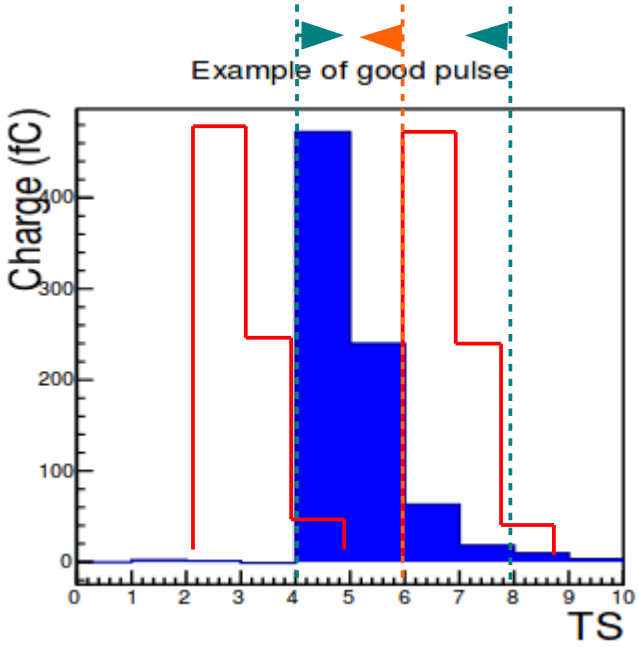
However, a lack of out-of-time pileup in simulation has been identified as the source of a small data/MC discrepancy → small impact on the reconstructed signals in the event in "low energy processes", leading to a small event distortion and biased MET scale

How to mitigate pileup effects ?

Detector signal reconstruction improvements during the run-I :

- ECAL timing information → reject out-of-time pileup
- HCAL reconstruction window changed from 100 ns to 50 ns





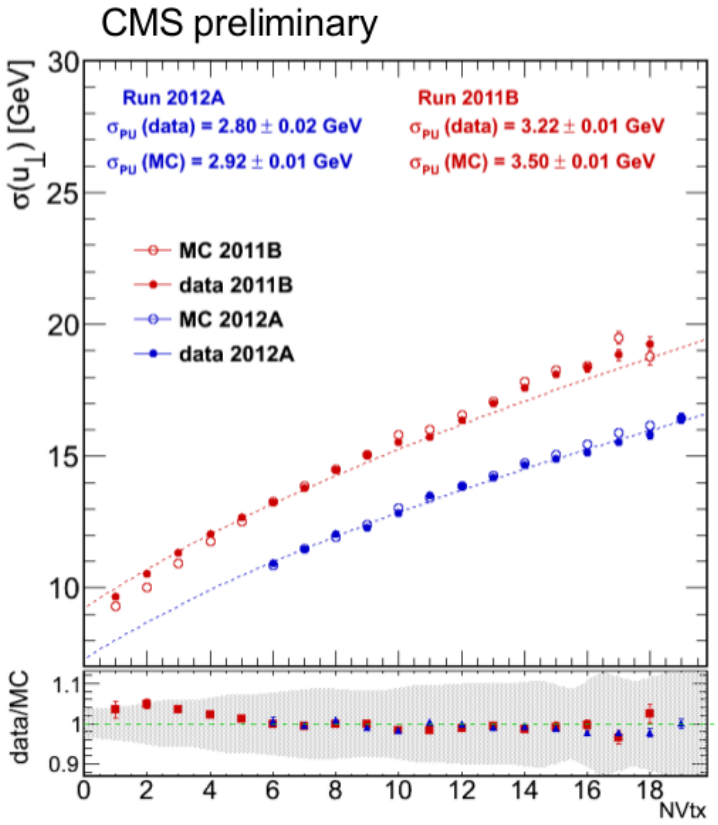
Pileup correction applied to the MET :

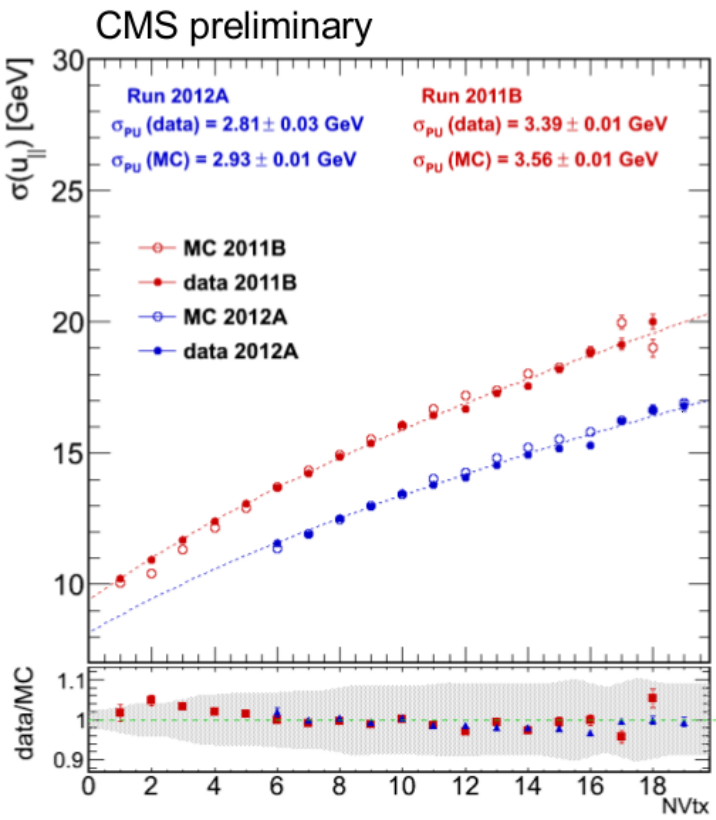
- Artificial balance of the neutral pileup activity using the charged information

$$\vec{\not{E}}_{\mathrm{T}}^{\,\mathrm{corr}} = \vec{\not{E}}_{\mathrm{T}} - \vec{\Delta}_{\mathrm{PU}} = \vec{\not{E}}_{\mathrm{T}} - \sum_{\text{PU-vertices}} f(v)\frac{\vec{v}}{v}$$

with $\vec{v} = \sum_{\text{charged}} \vec{p}_{\mathrm{T}}$ for a given vertex, and f(v) a function reflecting the average mismeasured/missing neutral activity as a function of the charged activity

Resolution performances



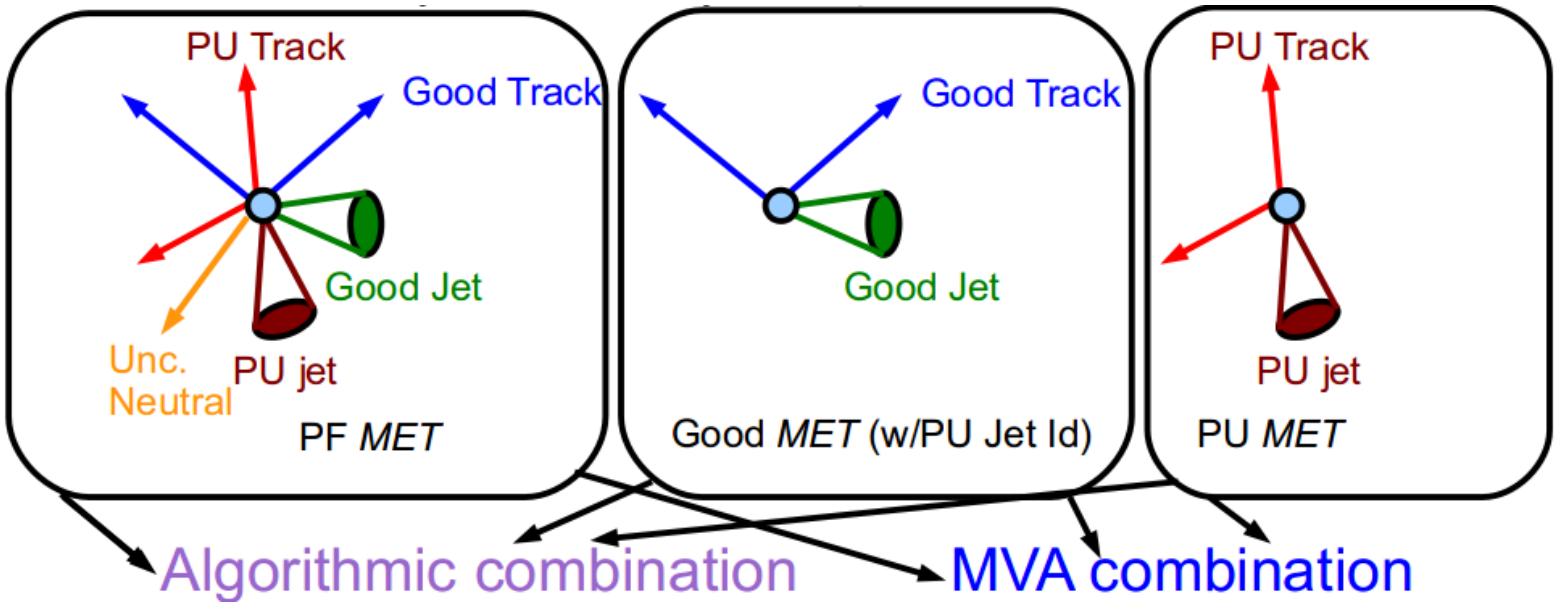


Calo-timing & pileup correction

Use of calorimeter timing and the general neutral pileup activity correction :
→ 15% improvement of the missingET resolution (mostly from the correction)

Algorithms improvements

Additional handle : change the definition of the missing transverse energy by using algorithms able to mitigate the pileup effects

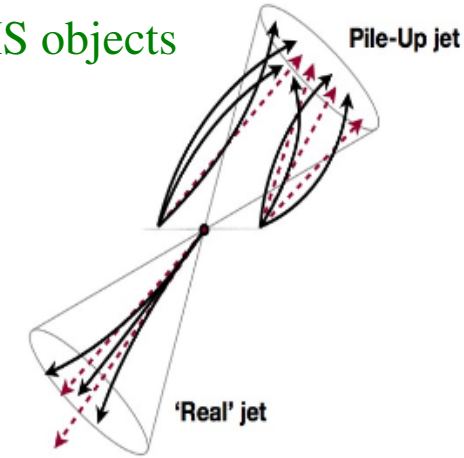


Aim : separate HS objects from pileup objects and recompute the MET with HS objects

Track identification : links to vertices

Neutrals : cannot do anything outside of jets

Jets : multivariate discriminator filled with jet shape and vertexing variables

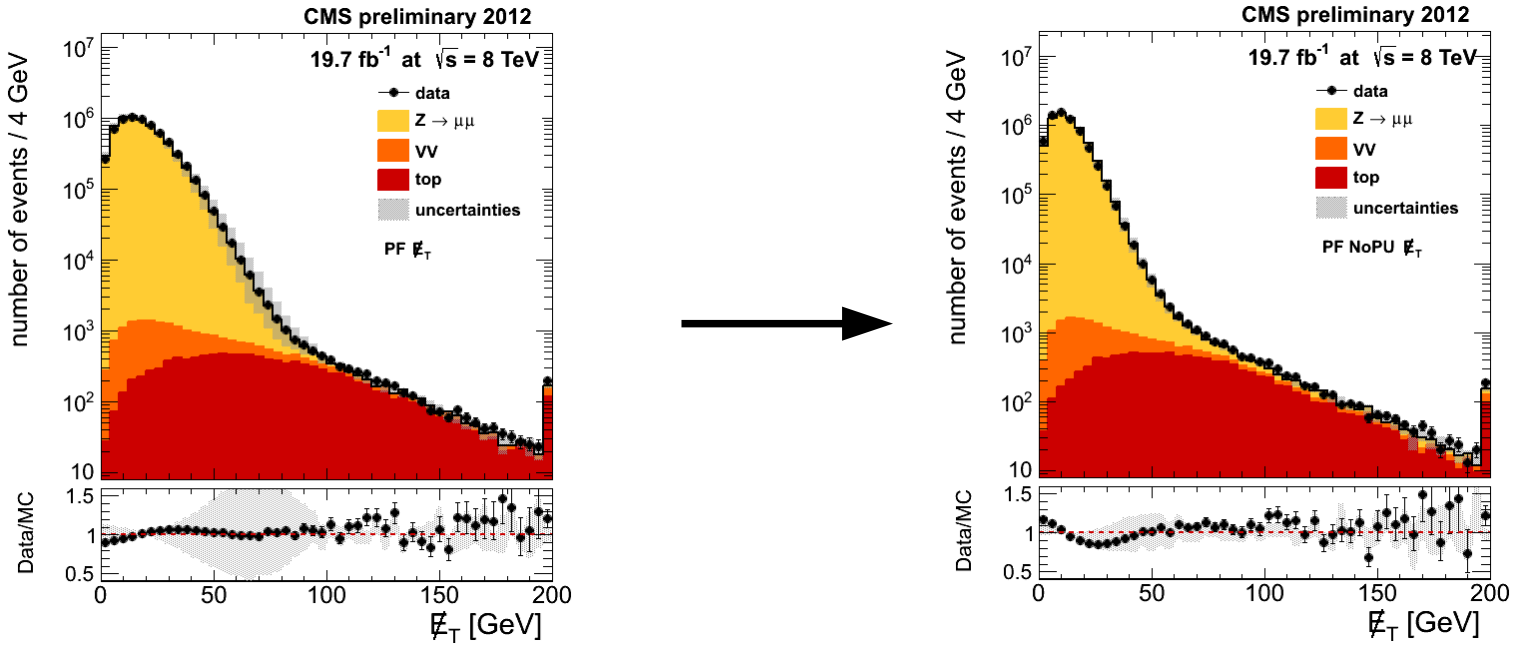


The NoPileup PF MET algorithm

Redefinition of the missingET reconstruction algorithm with a weighted linear sum of reconstructed particles in the event :

$$\vec{\not{E}}_{\mathrm{T}} = -\left[\sum_{\text{leptons}} \vec{p}_{\mathrm{T}} + \sum_{\text{HS-jets}} \vec{p}_{\mathrm{T}} + \sum_{\text{HS-charged}} \vec{p}_{\mathrm{T}} + S_{\mathrm{F}} \cdot \left(\alpha \cdot \sum_{\text{PU-charged}} \vec{p}_{\mathrm{T}} + \beta \cdot \sum_{\text{neutrals}} \vec{p}_{\mathrm{T}} + \gamma \cdot \sum_{\text{PU-jets}} \vec{p}_{\mathrm{T}} + \delta \cdot \vec{\Delta}_{\mathrm{PU}}\right)\right]$$

with the tuned parameters $\alpha = \gamma = \delta = 1$, $\beta = 0.6$ and $S_{\mathrm{F}} = \dfrac{\sum_{\text{HS-charged}} p_{\mathrm{T}}}{\sum_{\text{HS-charged}} p_{\mathrm{T}} + \sum_{\text{PU-charged}} p_{\mathrm{T}}}$

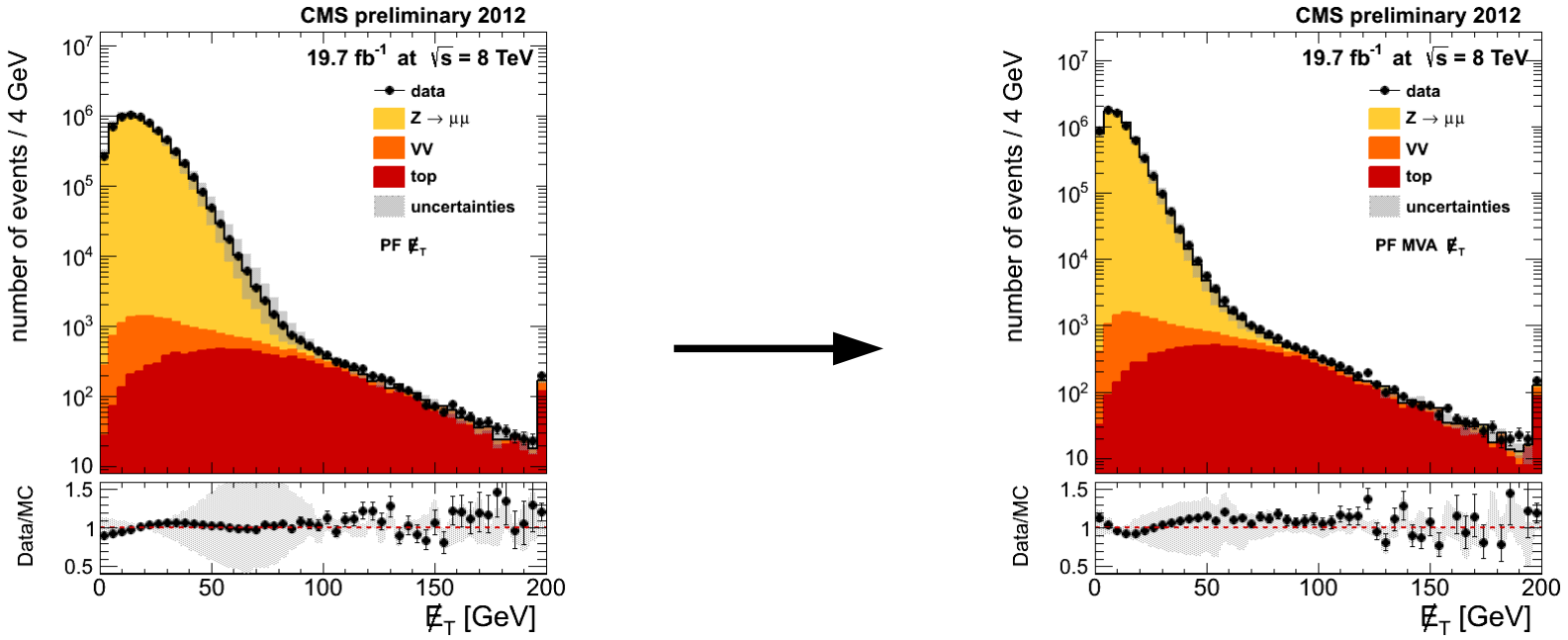


The MVA PF MET algorithm

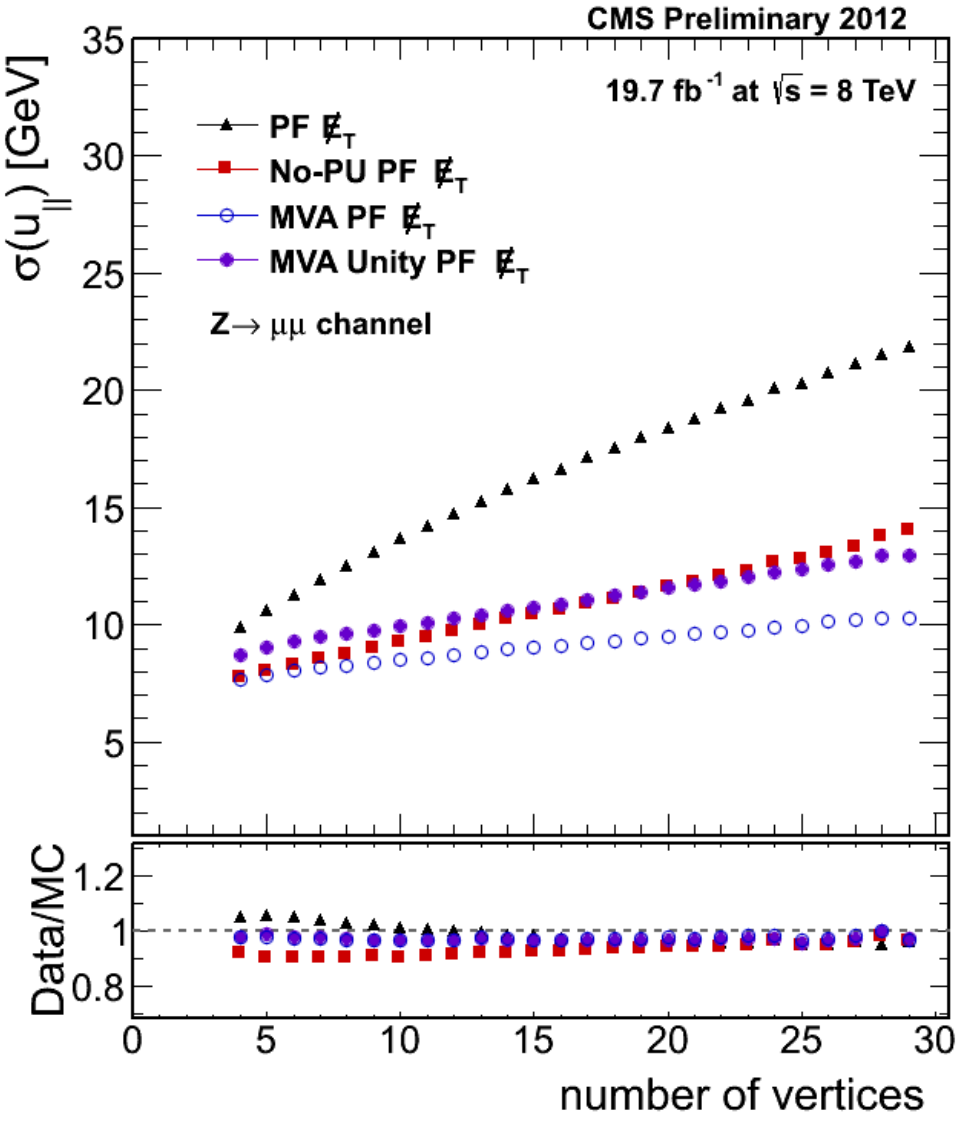
Defined with a MVA regression : trained on Z events → targeted MET is zero

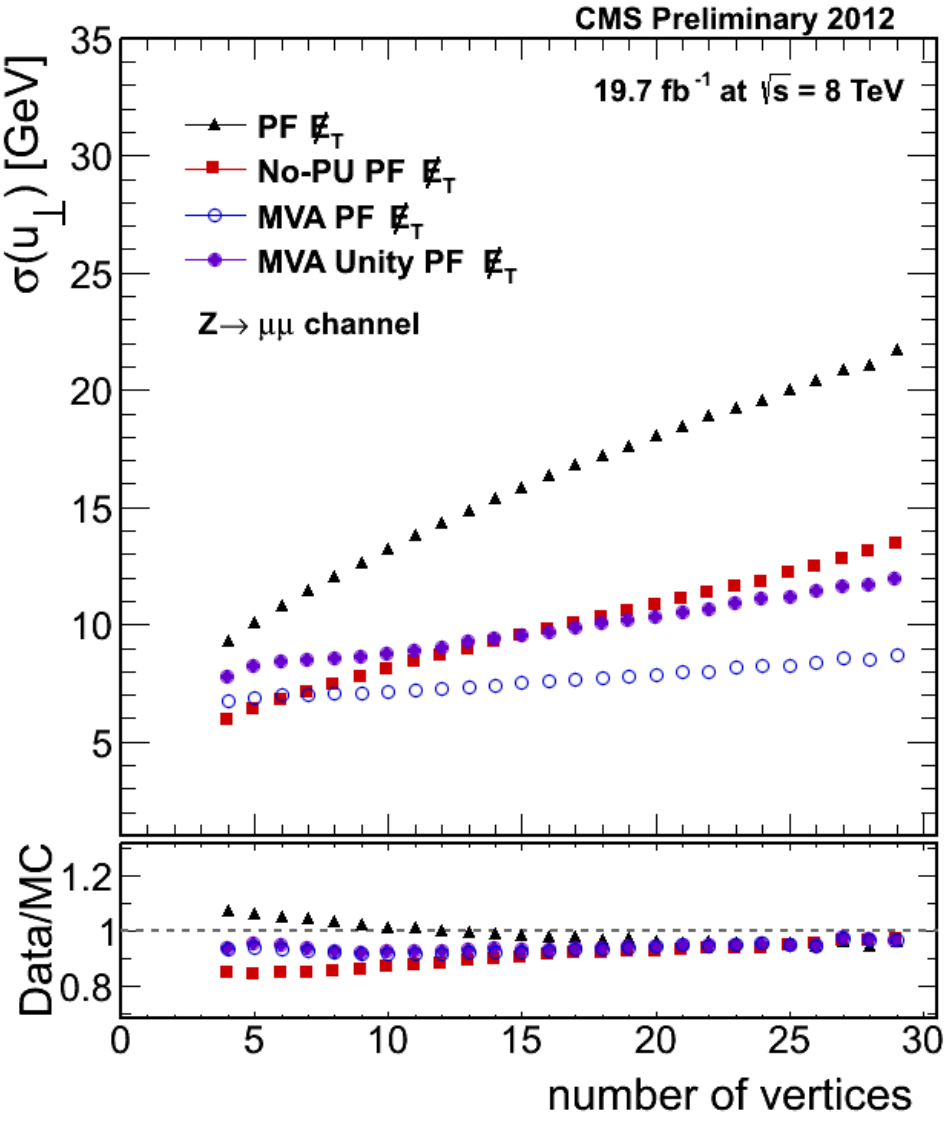
Use of 5 different "MET-like" variables constructed with :

- all particles reconstructed in the event (PF MET)
- HS charged particles
- HS charged particles + neutral particles within HS jets
- PU charged particles + neutral particles within PU jets
- HS charged particles + neutral particles within HS jets + isolated neutral particles



MissingET resolution





Performances in numbers

MET resolution can be parametrized as a quadratic sum of a hard scatter component and a pileup component identical per pileup interaction :

$$f(N_{vtx}) = \sqrt{\sigma_c^2 + \frac{N_{vtx}}{0.7} \times \sigma_{PU}^2}$$

	σ_c (GeV)	$R = \sigma_c(\text{data})/\sigma_c(\text{MC})$	σ_{PU} (GeV)	$R = \sigma_{PU}(\text{data})/\sigma_{PU}(\text{MC})$
PF MET	6.04 ± 0.03	$0.75 \pm 0.12 \pm 0.22$	3.25 ± 0.00	$1.05 \pm 0.03 \pm 0.11$
noPU PF MET	6.10 ± 0.02	$1.17 \pm 0.02 \pm 0.07$	1.83 ± 0.00	$1.04 \pm 0.02 \pm 0.05$
MVA PF MET	7.30 ± 0.01	$1.01 \pm 0.05 \pm 0.07$	1.14 ± 0.00	$1.04 \pm 0.09 \pm 0.12$

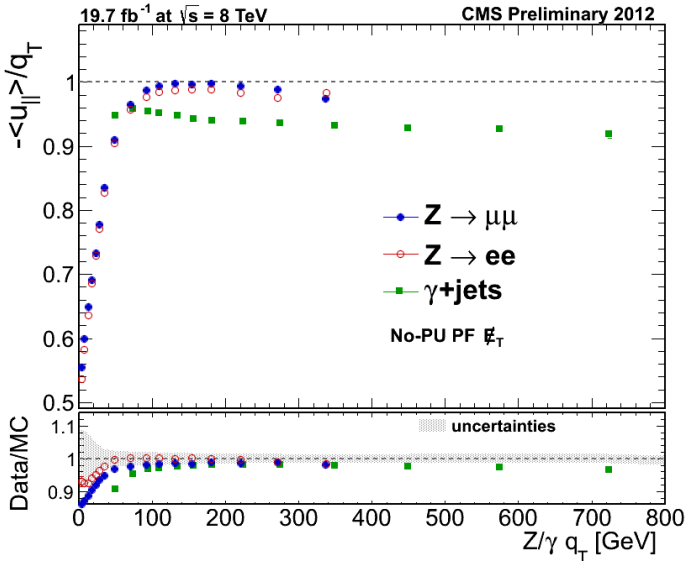
Improvement of the MET resolution by a factor 1.8 to 2.8
using alternate algorithms

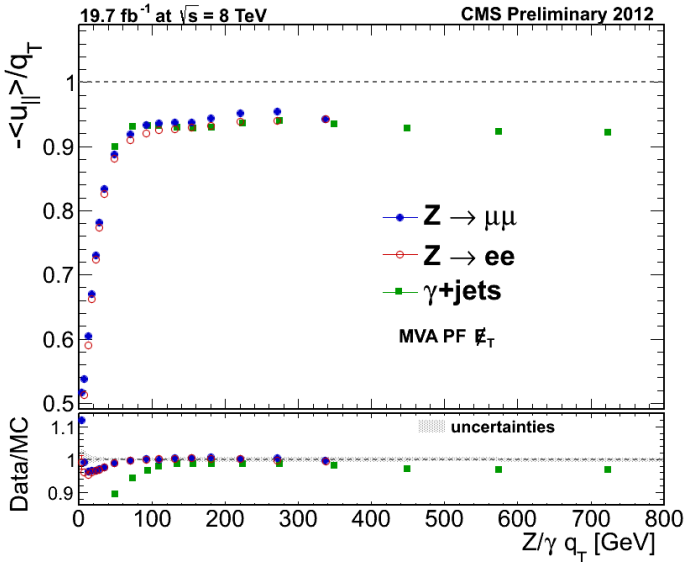
Drawbacks

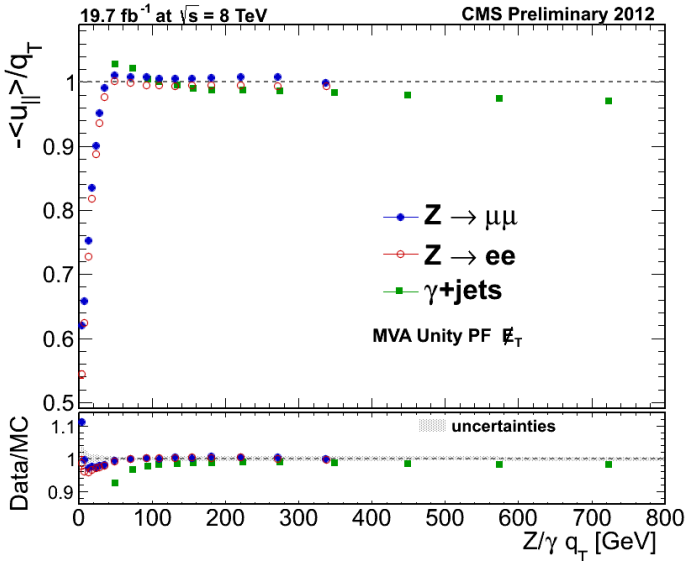
Scaling down the contribution of particles in an event is just removing energy

If only pileup is removed everything is fine, but this is never the case

→ impact on the MET energy scale







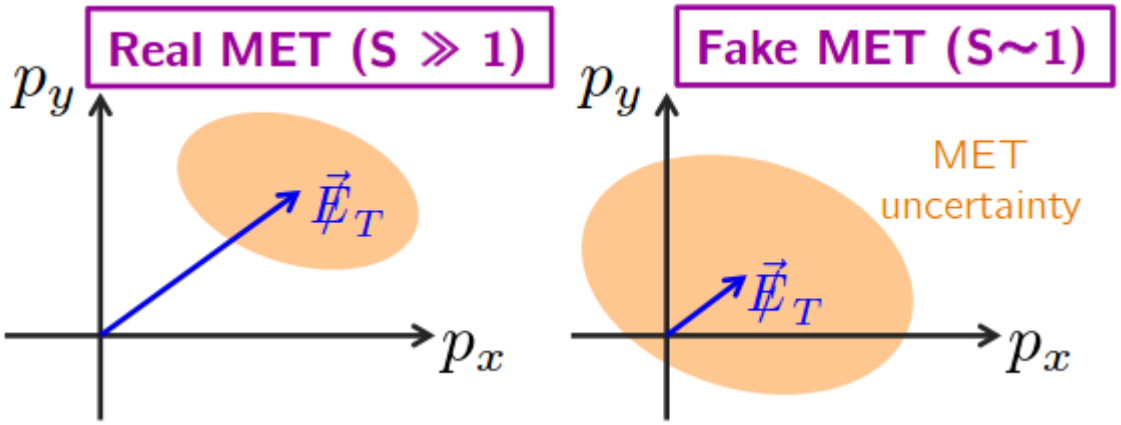
Depending of what you want to do, it can be a problem or not :

- W mass, top mass, h → ττ need an accurate MET energy scale
- selection purposes (i.e separating Z → 2l from WW → 2l2ν) less sensitive

Special MVA training can be designed to get a unity response

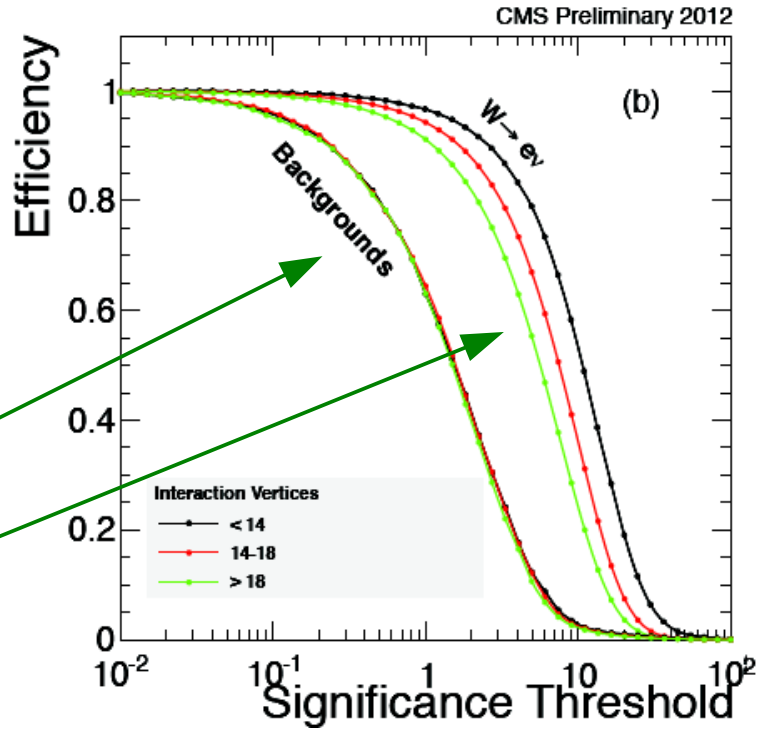
MissingET significance

$$\mathcal{S} \sim \frac{\not\!\!E_T^2}{\sigma^2_{\not\!\!E_T}}$$



The idea : using all the event content, build a variable reflecting the likelihood to obtain genuine missingET in such topology

- Algorithm insensitive to the pileup for events with no genuine MET
- Sensitive when presence of genuine MET



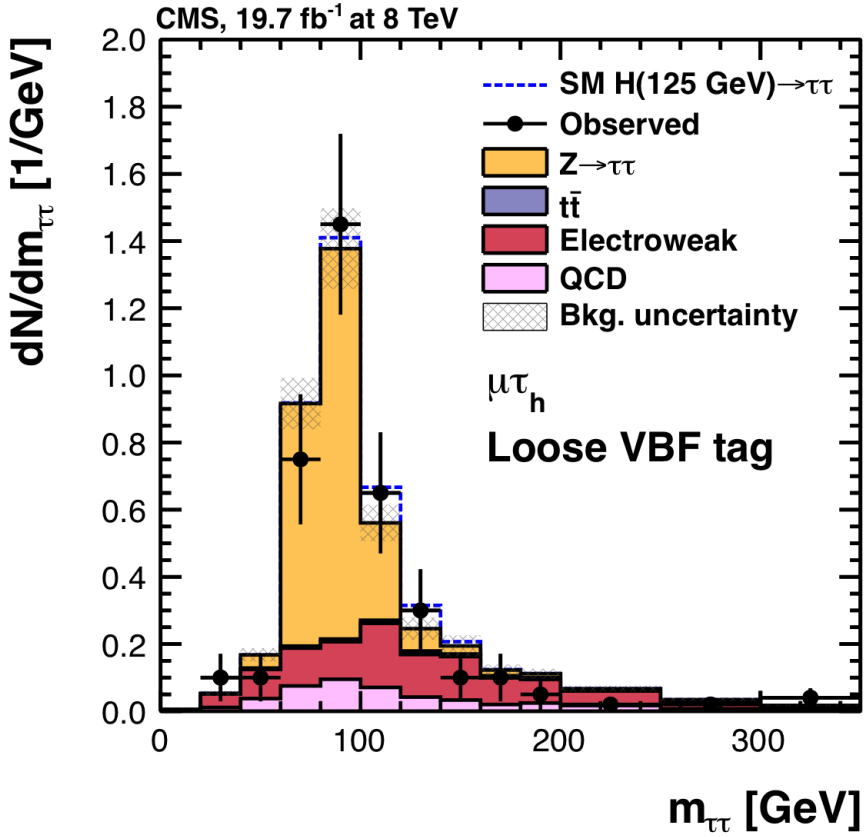
Situations of interest

Separation of events with no MET from signal :

- dark matter vs Z → 2l / photon+jet
- ttbar vs Z → 2l
- etc.

Precision measurements :

- W mass measurement
- H → ττ, MVA MET used by CMS, gain of ~30% on the resolution with no loss of response



The future : to infinity and not so beyond

The runII and the 25ns BX running scenarii :

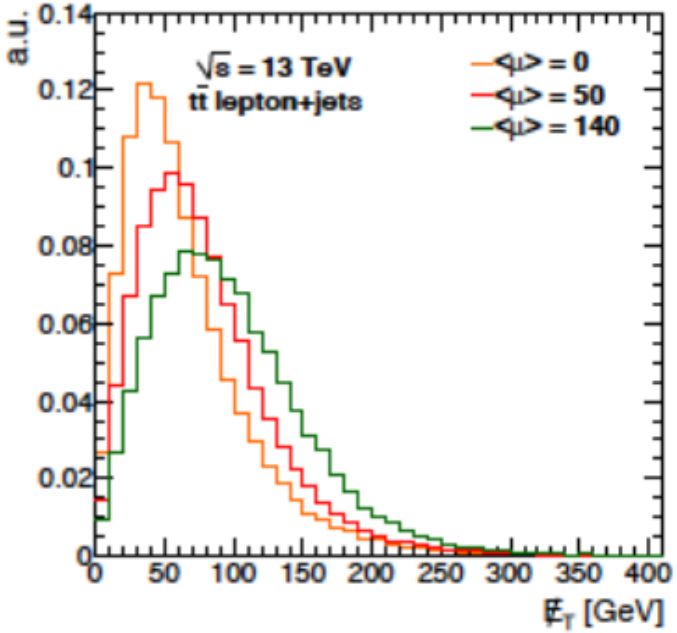
OOT pileup, first estimations gives an extra ~2 GeV contribution (in quadrature) on resolution for BX=-1 and independently for BX=+1

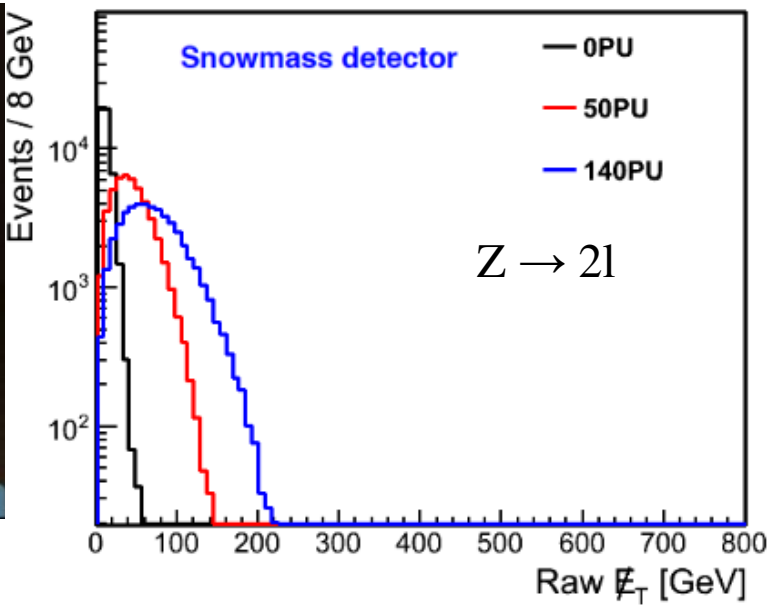
→ calorimeter/tracking timing information is crucial to mitigate those effects, currently in development in CMS

In-time pileup : same degradation expected per pileup interaction so far, but dedicated mitigation algorithms may become less and less sensitive

New more general pileup mitigation algorithms under development (soft killer, PUPPI, etc)

And after 2018?





Summary

The missingET reconstruction has been working very well during the run I in CMS

- good modeling of the observed distribution
- development of techniques to mitigate the increasing pileup effects
- detector signal reconstruction
- alternate algorithms robust against pileup

Run II and farther

- the 25 ns BX scenarii are currently under studies in CMS
  → deterioration of the resolution estimated at ~2 GeV / out of time pileup interaction
  → optimization of alternate algorithms may be needed

- in time pileup : alternate algorithms to be tested, some re-optimization may be needed

- new algorithms looks promising (e.g. PUPPI)

- at high PU level (> 50-80) ? the great unknown

Backup