University of Zurich[UZH]
Physik-Institut

Improving M_{bb} resolution in ZH→llbb

Two staged approach : b-specific jet corrections followed by jet corrections with MET

Jennifer Ngadiuba
Clemens Lange
Benjamin Kilminster

VHbb meeting
19th April 2013
CERN

Motivation

- The Higgs dijet invariant mass is the most effective discriminant to distinguish the Higgs signal from backgrounds
  - Higgs-signal → Gaussian-like resonance
  - Backgrounds → exponential-like smoothly falling

- Improving the jet energy resolution improves the Higgs dijet invariant mass resolution

- Jet energy resolution = RMS[($p_{T,genB}$ - $p_{T,recoJet}$)/$p_{T,recoJet}$]
  - $p_{T,genB}$ = p_T of the generator level b-quark
  - $p_{T,recoJet}$ = p_T of the reconstructed jet

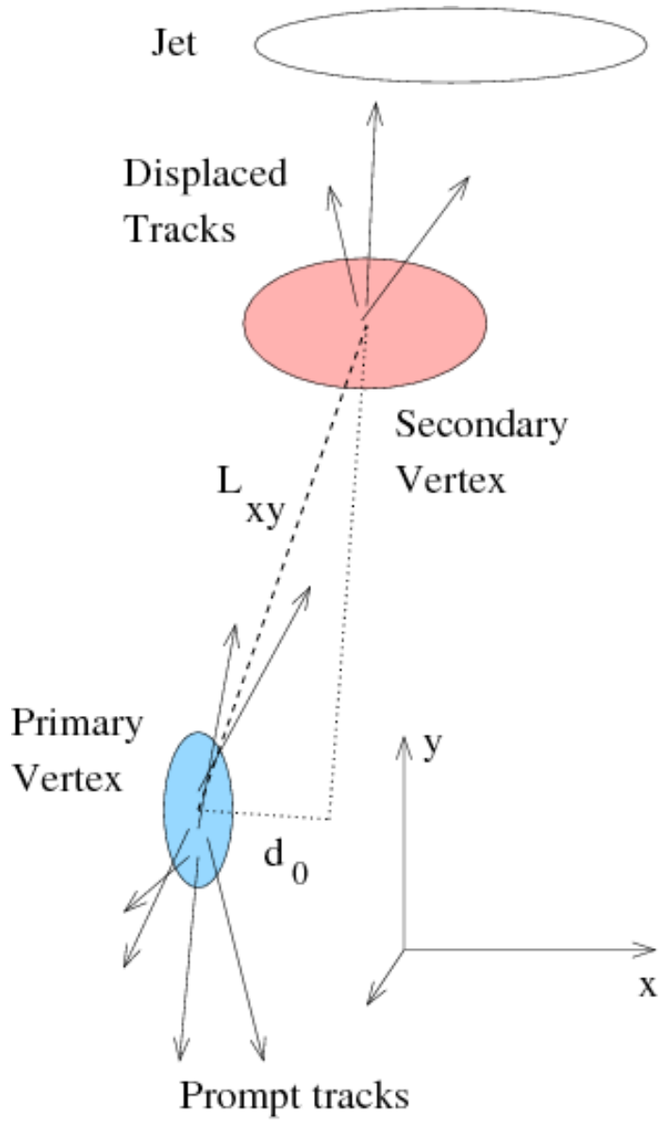
- **Goal:** provide an estimate of the true b-quark energy and correct the reconstructed jet energy

b-jet energy correction

- CDF implemented a method based on Neural Network (ROOT) (arxiv.org/pdf/1107.3026.pdf)
  - a correction function is computed using specific properties of b-jets as input variables
  - target the p_T of the generator level b-quark
  - the Higgs dijet invariant mass resolution improved from ~15% to ~11% (~27%) in WlnHbb
- Similar regression implemented for CMS (Niklas Mohr talk for Hbb meeting)
  - BDT regression implemented in TMVA
  - input variables: b-specific properties, jet kinematic and MET-related information
  - target generator level jet-p_T
  - the two jets are kept uncorrelated in the training
  - Higgs dijet invariant mass resolution improvement of ~20% in ZllHbb

b-jet properties

- Large lifetime of B-hadrons (~1.5ps)
  - observable flight distance
  - secondary vertex displaced from the primary vertex
- Large multiplicity of charged particles in the final state
- Large mass
  - charged particle tracks incompatible with the primary vertex
  - high impact parameter



Goal:

improving individual b-jet energy measurement and resolution exploiting the correlation between the b-specific properties and the b-jet energies

Missing Transverse Energy

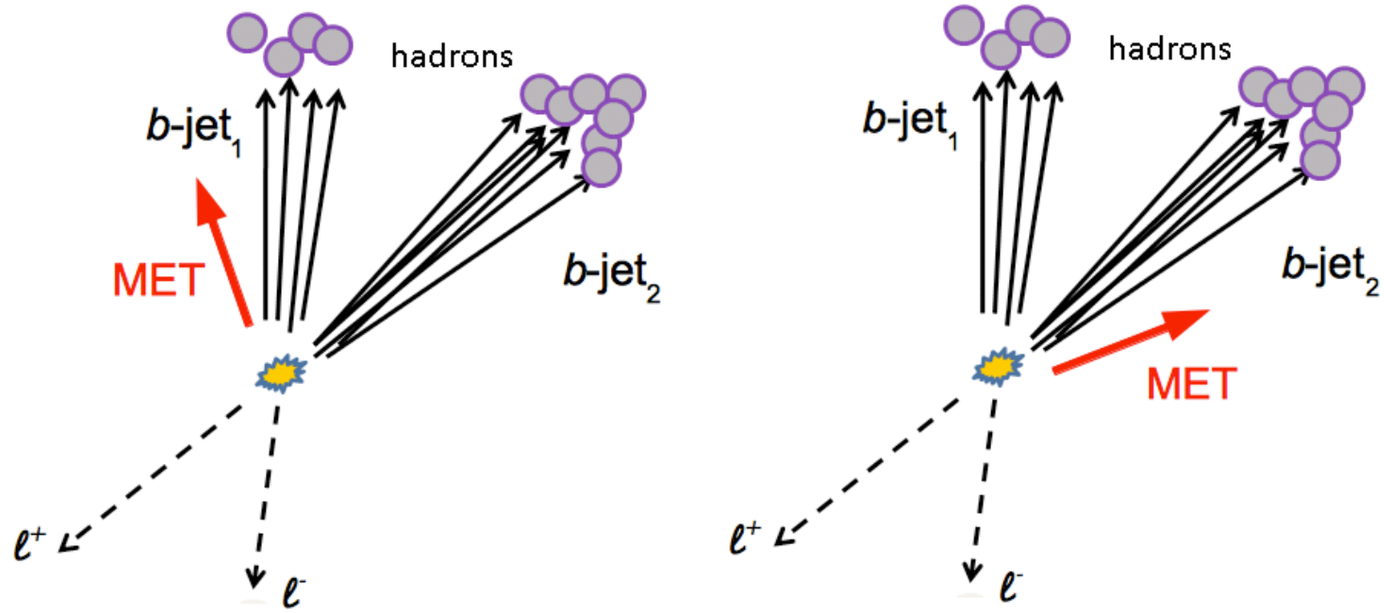
- Presence of Missing Transverse Energy
  - calorimeter resolution
  - presence of neutrino in the jet

→ mis-measurement of the b-jet energy

- The MET direction points in the direction of the jet that is more likely mis-measured

$ZH \rightarrow llbb$ events

$\sum \vec{p}_{T,lep} + \sum \vec{p}_{T,jet} \neq \vec{0}$



Goal:
exploit the MET value and direction compared to the jet to correct the reconstructed jet energies

Technical details

> A correction function is computed in order to approximate the reconstructed b-jet energy to the MC generated b-quark energy

> The method exploits Multi-Layer Perceptron Neural Networks (implemented in ROOT) in two steps:

1.

- A first NN is trained using

  - b-specific input variables (Secondary Vertex)

  # tracks in the SV of the $jet_{1,2}$, p_T of the lead track in the $jet_{1,2}$

  vtx-mass, vtx-pT - mass and pT of the SV of the $jet_{1,2}$

  vtx-3dL, vtx-3deL - 3D flight lenght and error of the SV of the $jet_{1,2}$

  - jet kinematic input variables:

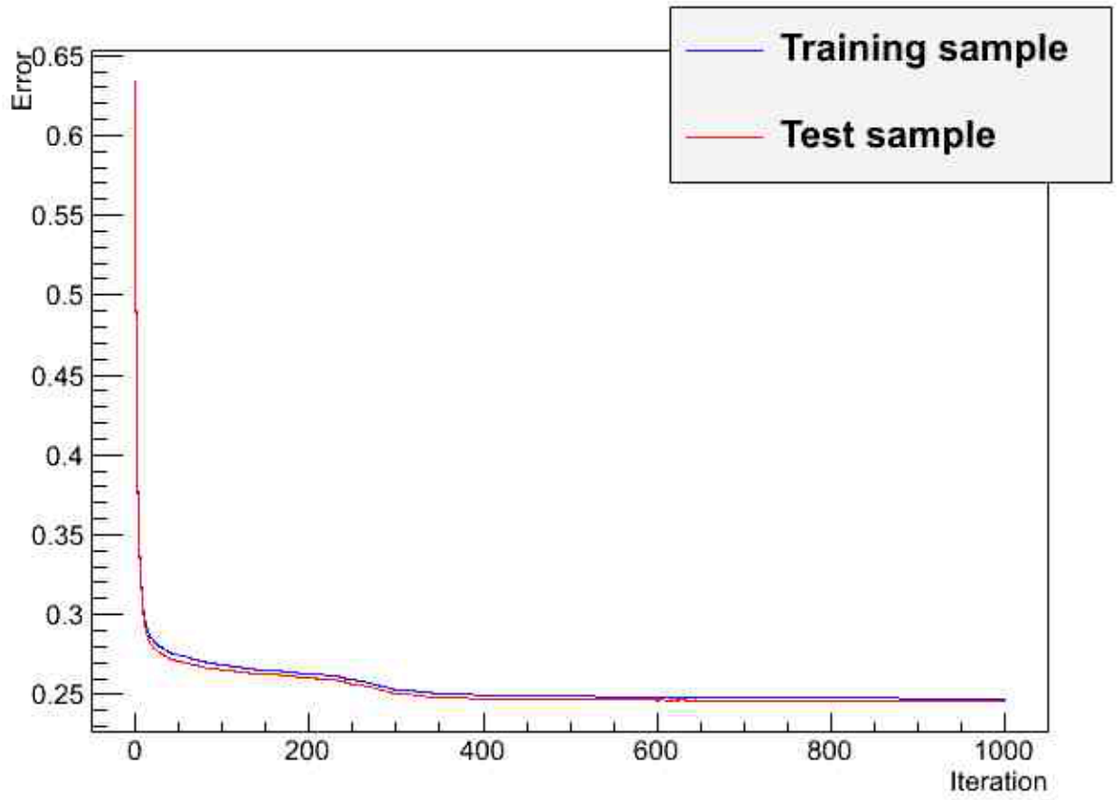
  standard corrected $jet_{1,2}$ energy and p_T

16 input variables

pT of the generator level b-quark: NEW in CMS!

- Target two scale factors (one for each jet): **$SF = p_{T,genB}/p_{T,recoJet}$**
- Outputs: two correction factors which are applied to the jets
  → NNb-corrected jets energy, jets pT, MET

NN training



- BFGS method with 1 hidden layer of 32 neurons for 1000 epochs
- On a sample of b-tagged jets from Higgs candidate matching the generator level b-quarks, chosen from a MC of ZllHbb events @ M_H = 125GeV
- jet-quark match → $dR = \sqrt{(\phi_{genB} - \phi_{jet})^2 + (\eta_{genB} - \eta_{jet})^2} < 0.5$

NN training

2.

- A second NN is trained using

  - **MET input variables**

    b-corrected MET, MET phi, MET-jet$_{1,2}$ projection

  - **jet kinematic input variables:**

    NNb corrected jet$_{1,2}$ energy and p_T, jet$_{1,2}$ eta, jet$_{1,2}$ phi

  - **PU correction:**

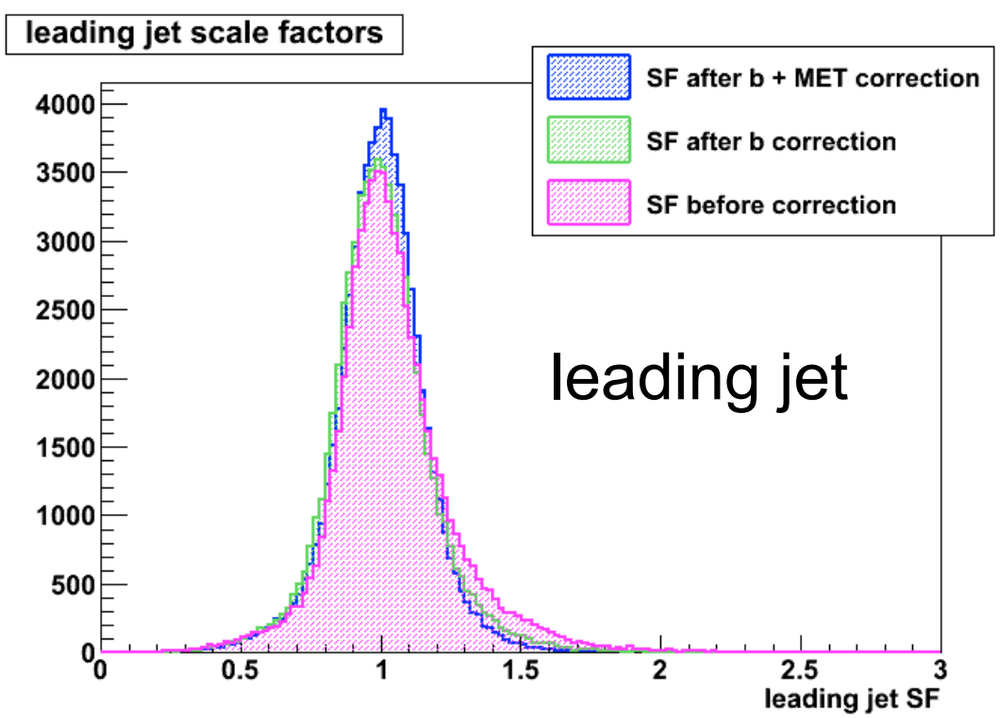
    # Primary vertices

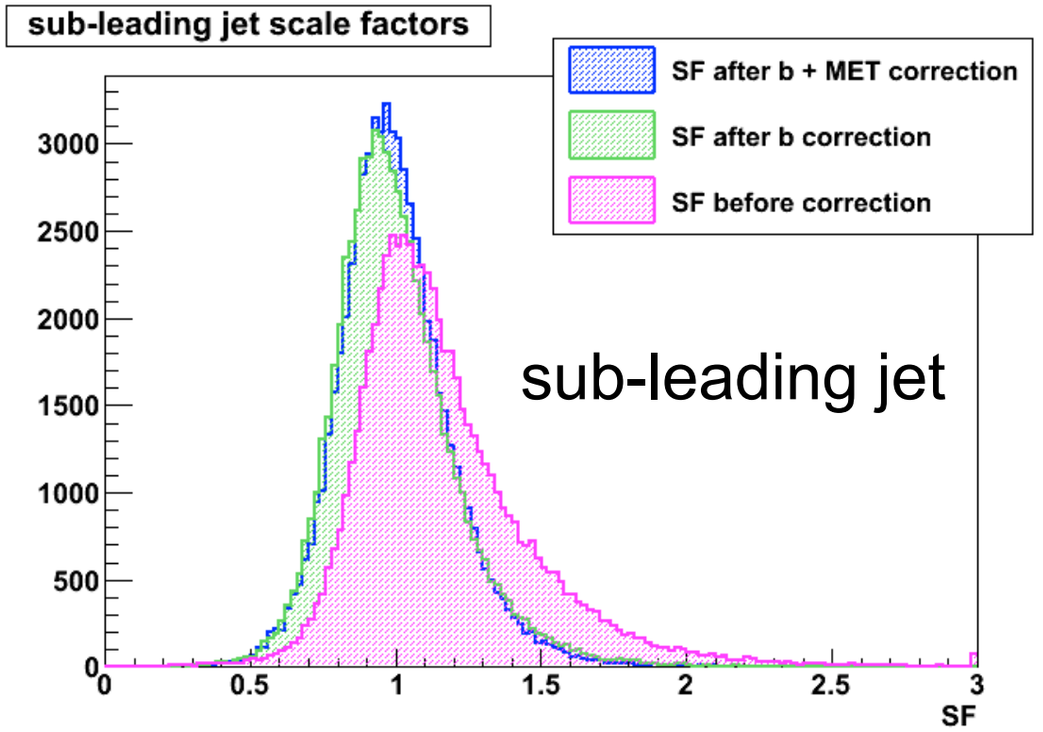
11 input variables

- Target two scale factors (one for each jet): **SF = $p_{T,genB}/p_{T,recoJet}$ (NNb-corr)**
- Outputs: 2correction factors which are applied to the jets
  → NN_b+NN_{MET} - corrected jets energy and jets p_T
- BFGS method with 1 hidden layer of 22 neurons for 1000 epochs
- On a sample of b-tagged jets matching with generator level b-quarks chosen from a MC of ZllHbb events @ M_H = 125GeV

Scale factors

$SF = p_{T,genB}/p_{T,recoJet}$ —NNb correction→ $SF_b = p_{T,genB}/p_{T,recoJet}$ (NNb-corr) —NN_{MET} correction→ $SF_{MET} = p_{T,genB}/p_{T,recoJet}$ (NNMET-corr)





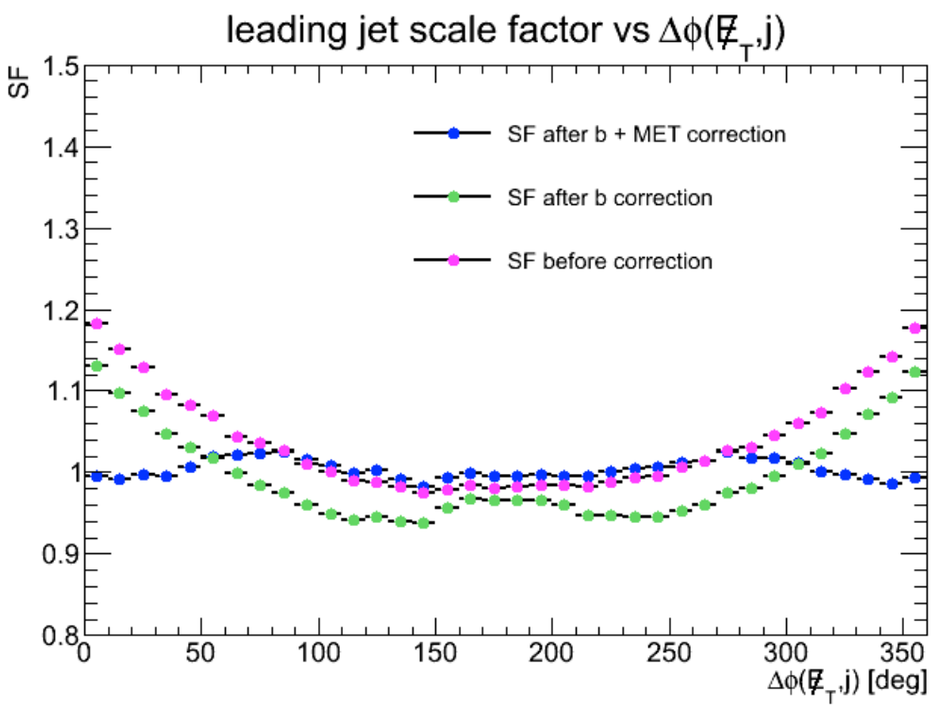
- Jets selection:
  - 2 b-tagged jets from the Higgs candidate with pT > 20 GeV, dR < 0.5
  - no selection applied on additional jets

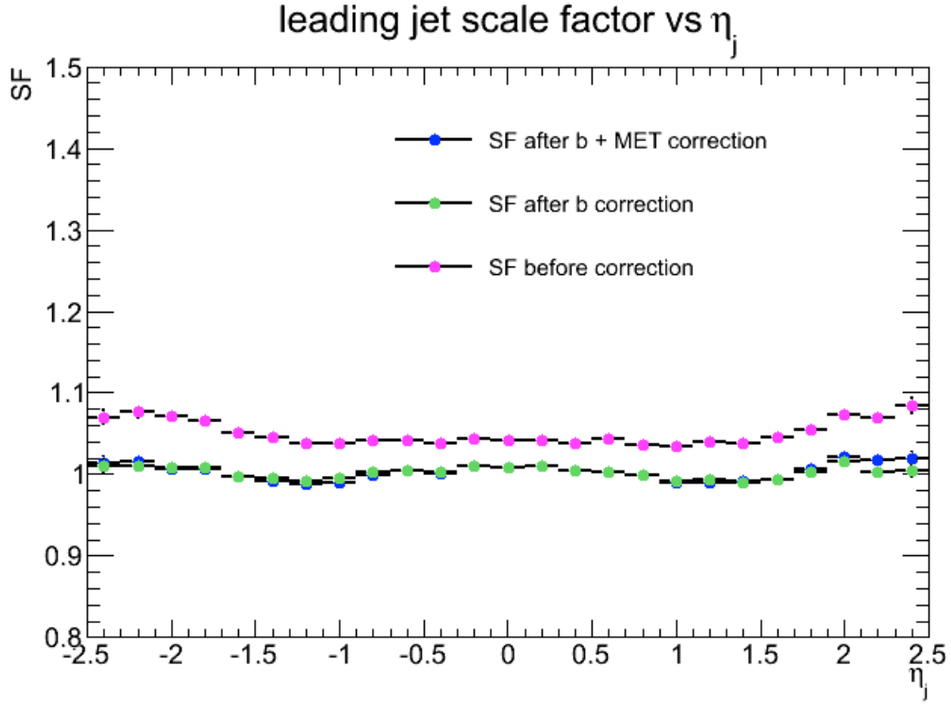
Scale factors

Scale factors vs DPhi(MET,j)

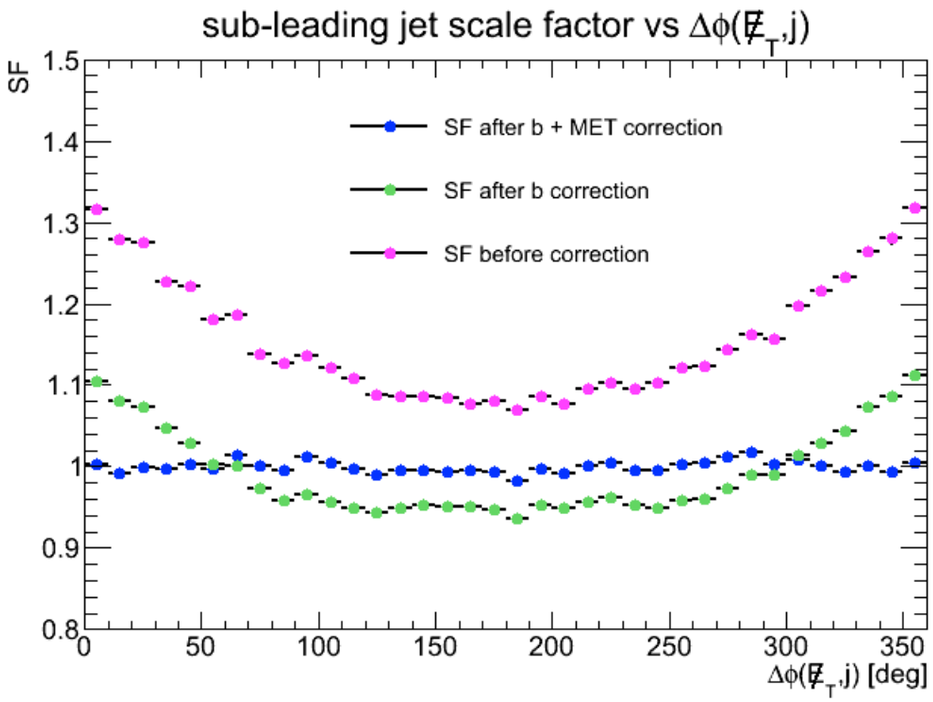
Scale factors vs jet eta

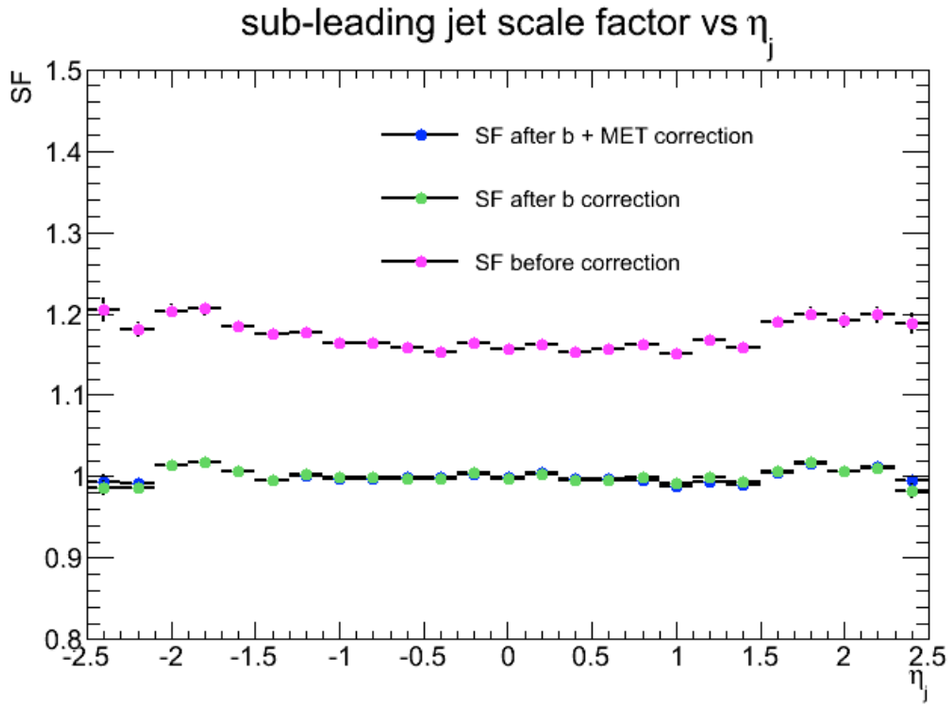
leading jet





sub-leading jet

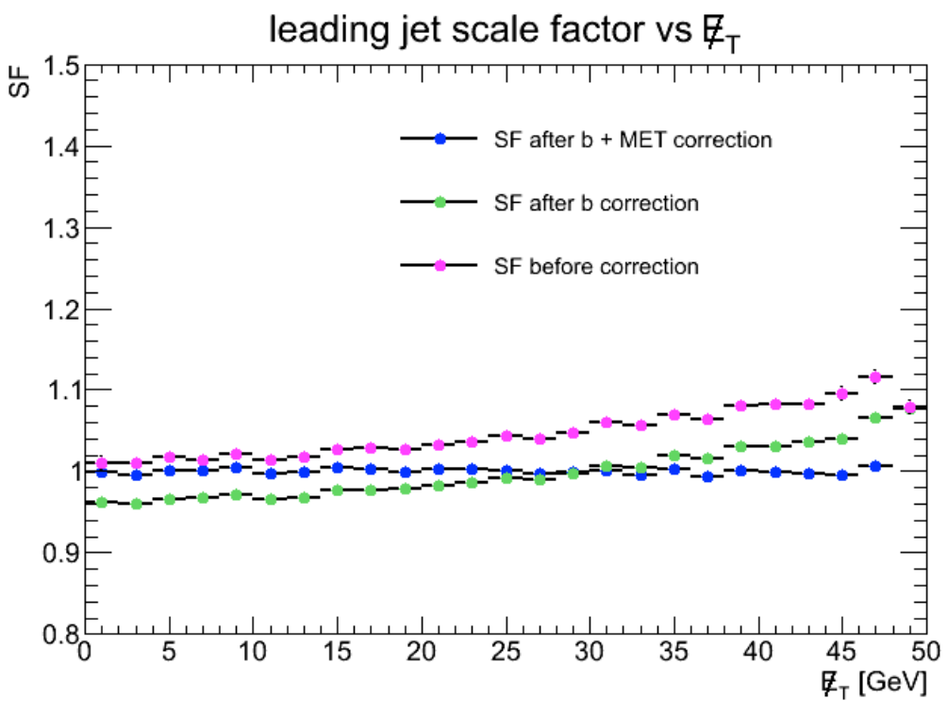




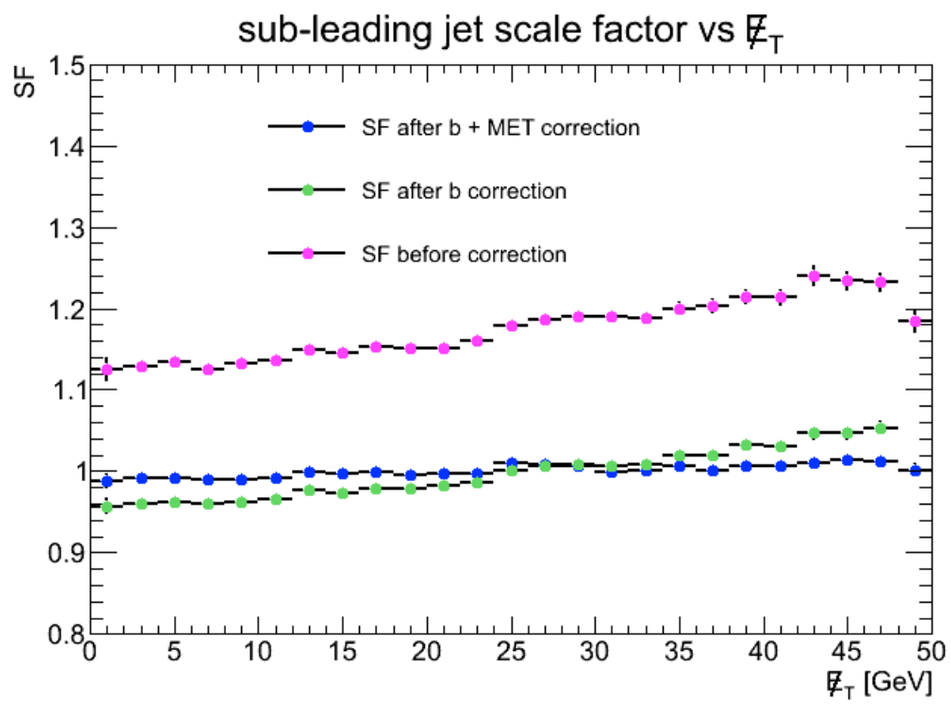
Scale factors

Scale factors vs MET

leading jet



sub-leading jet

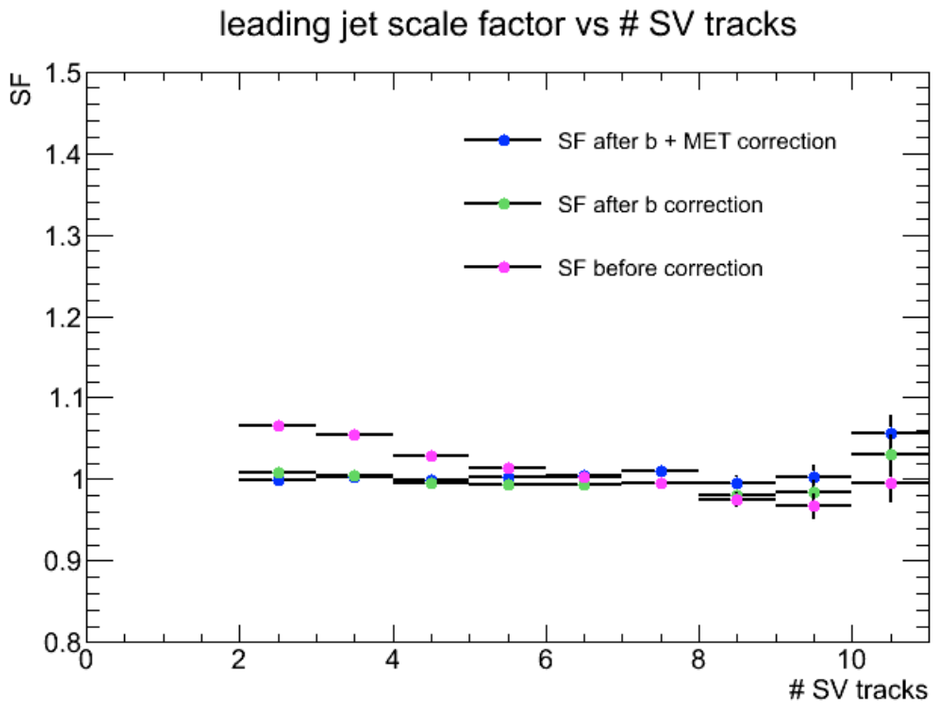


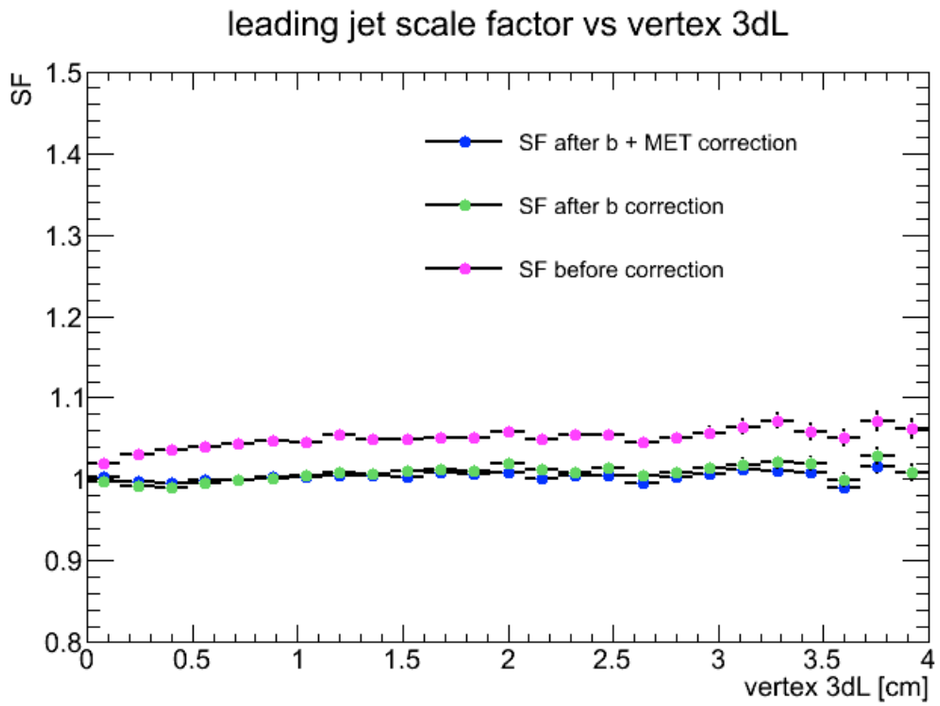
Scale factors

Scale factors vs # SV tracks

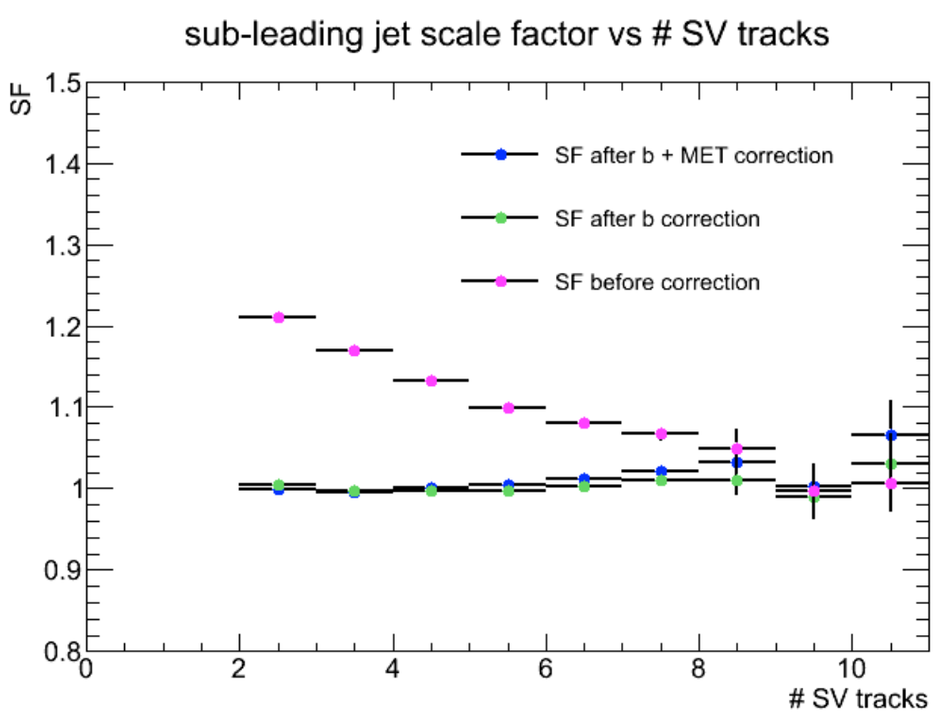
Scale factors vs vertex 3dL

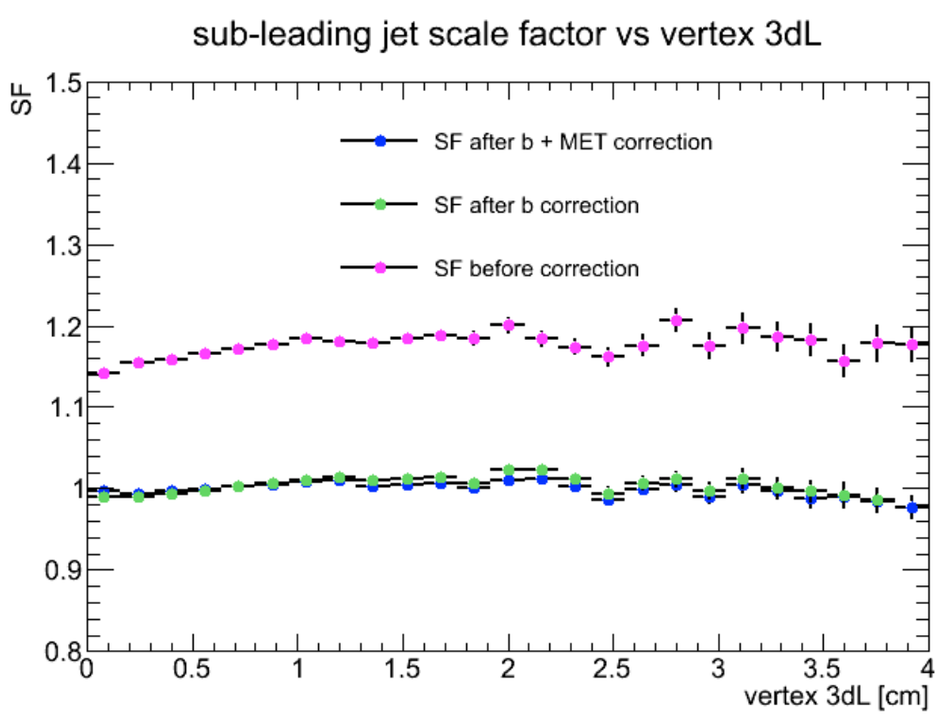
leading jet





sub-leading jet



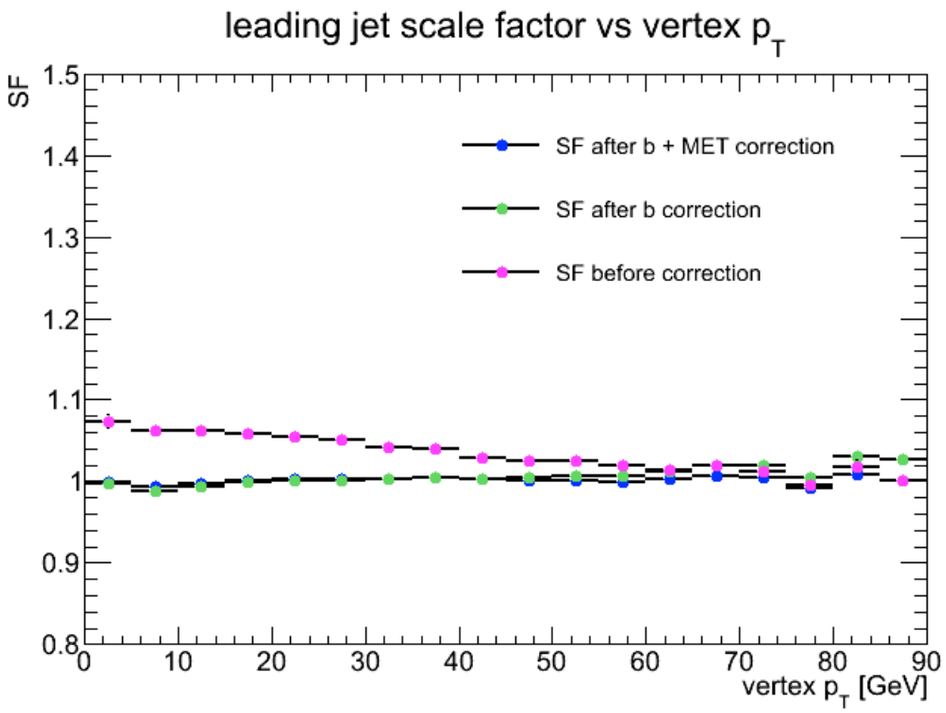


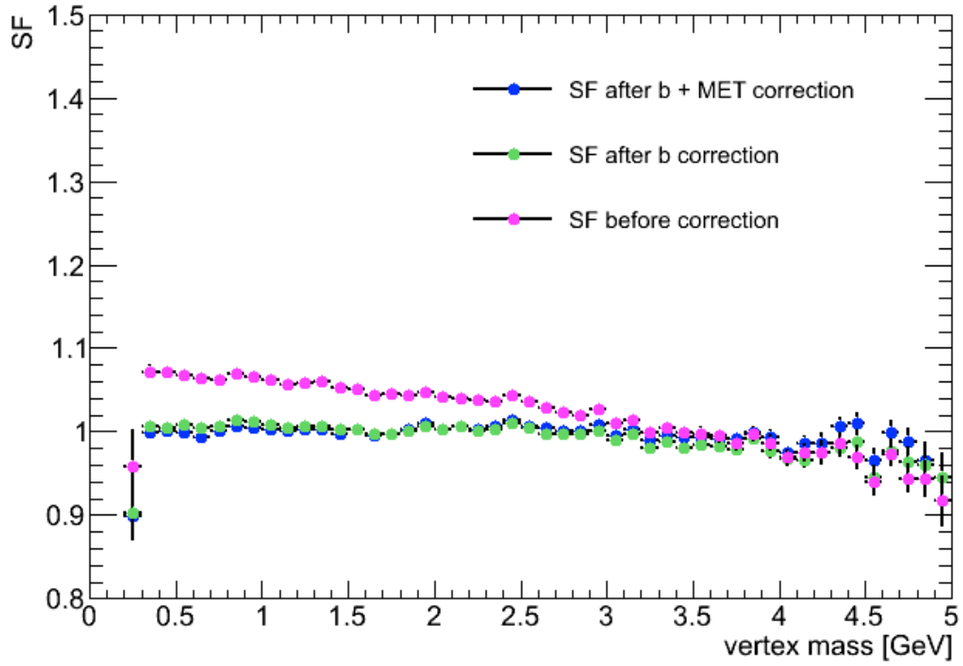
Scale factors

Scale factors vs vertex pT

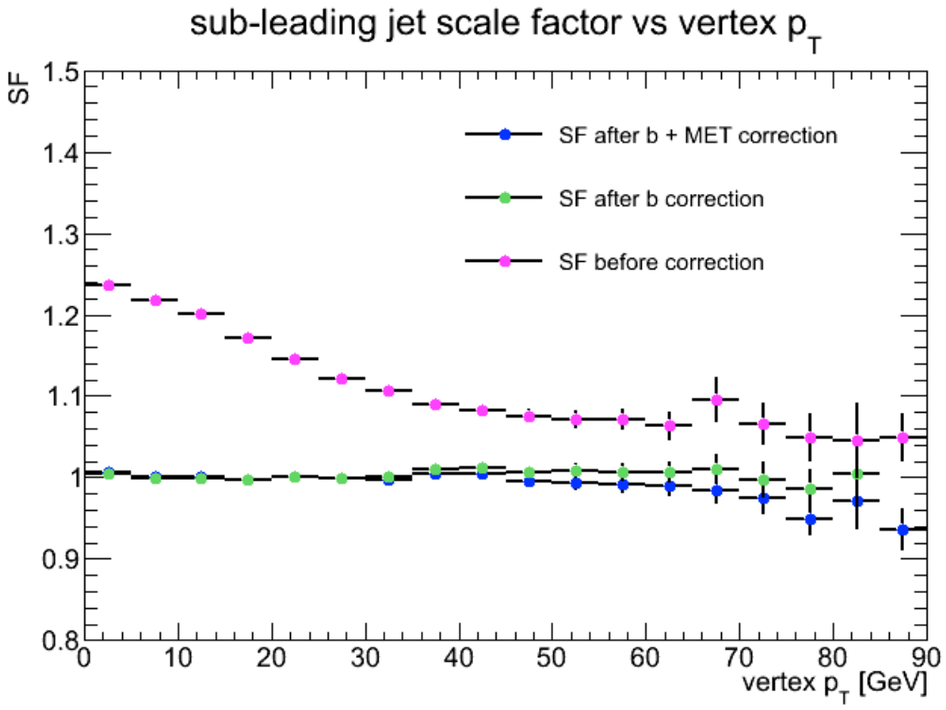
Scale factors vs vertex mass

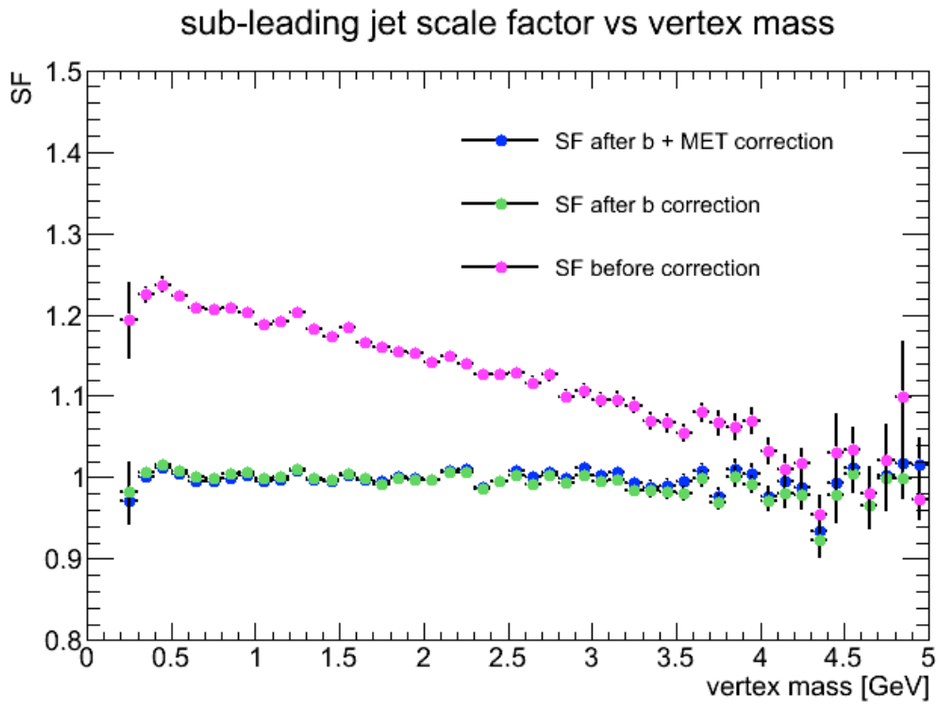
leading jet



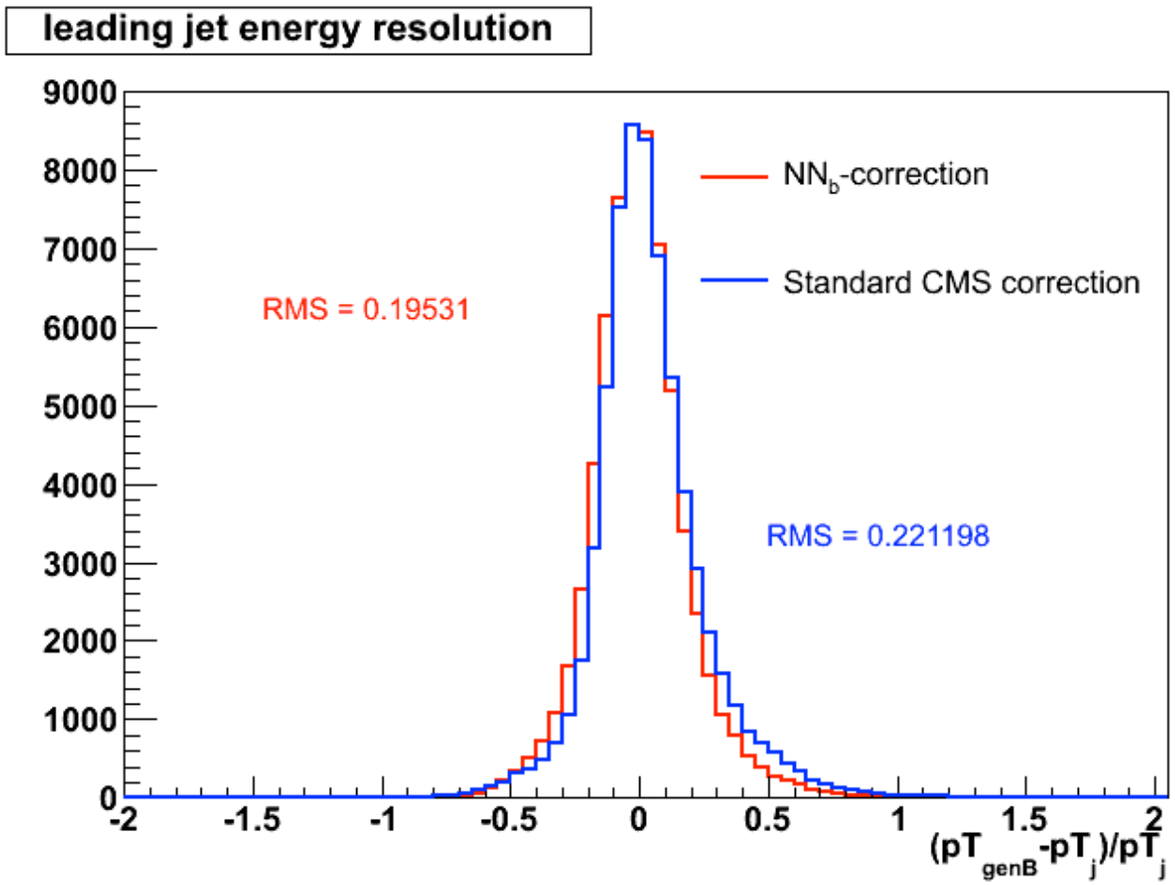


sub-leading jet

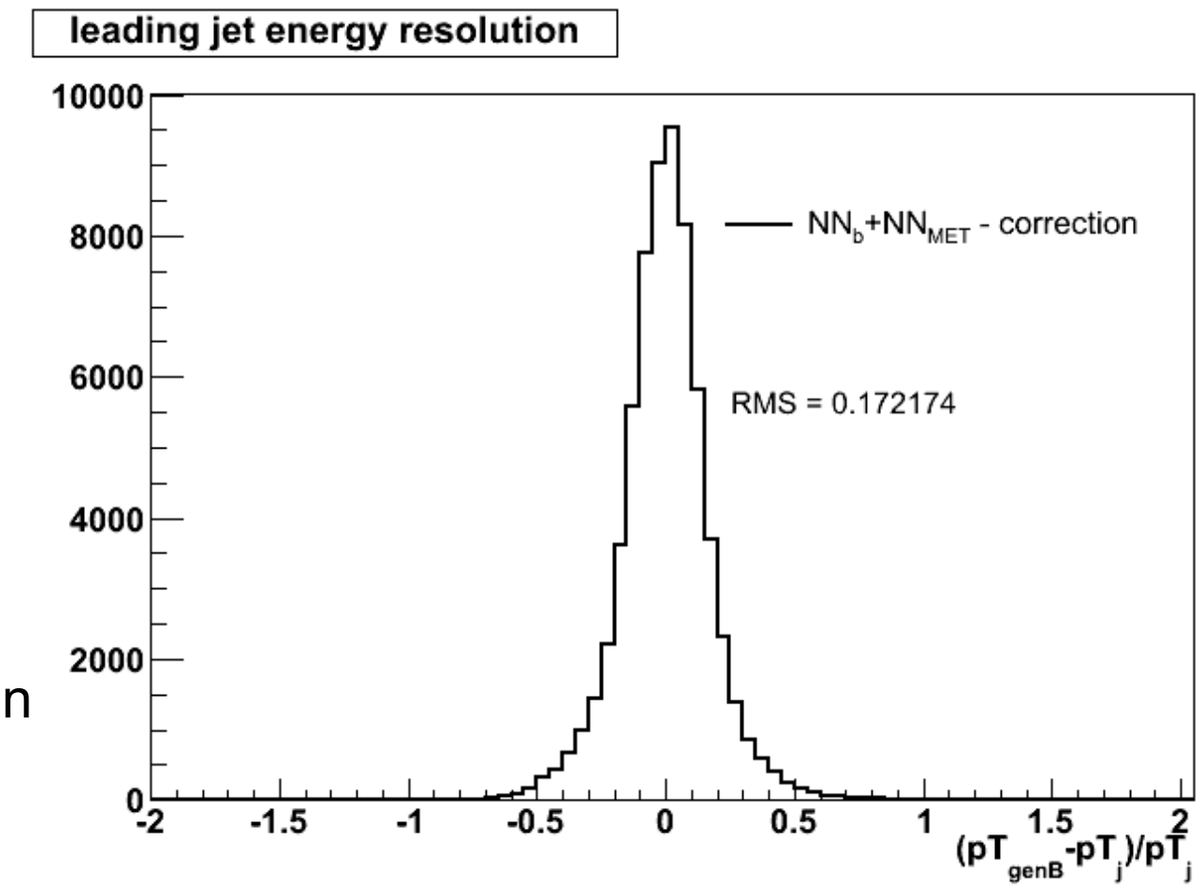




jet energy resolution



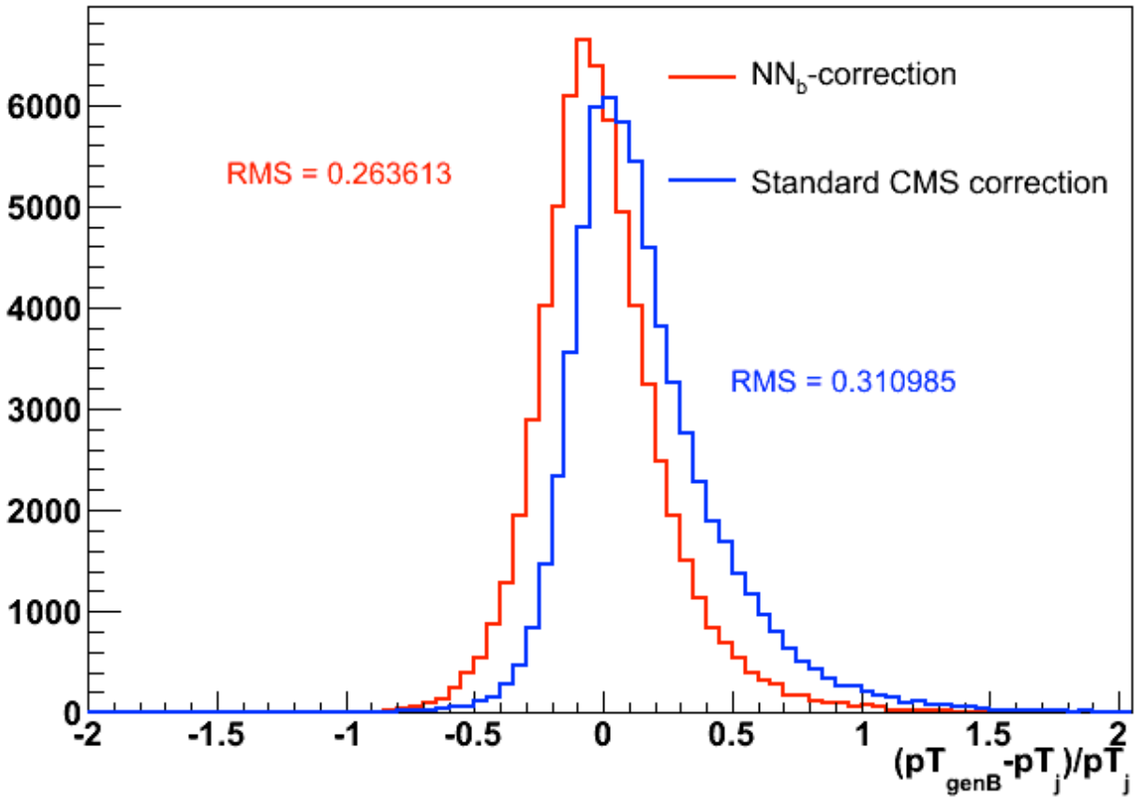
- Jets selection:
  - 2 b-tagged jets from the Higgs candidate with pT > 20 GeV, dR < 0.5
  - no selection applied on additional jets



- **Leading jet energy resolution improved**
  - from ~22% to ~20% after NN_b correction
  - from ~20% to ~17% after NN_{MET} correction

jet energy resolution

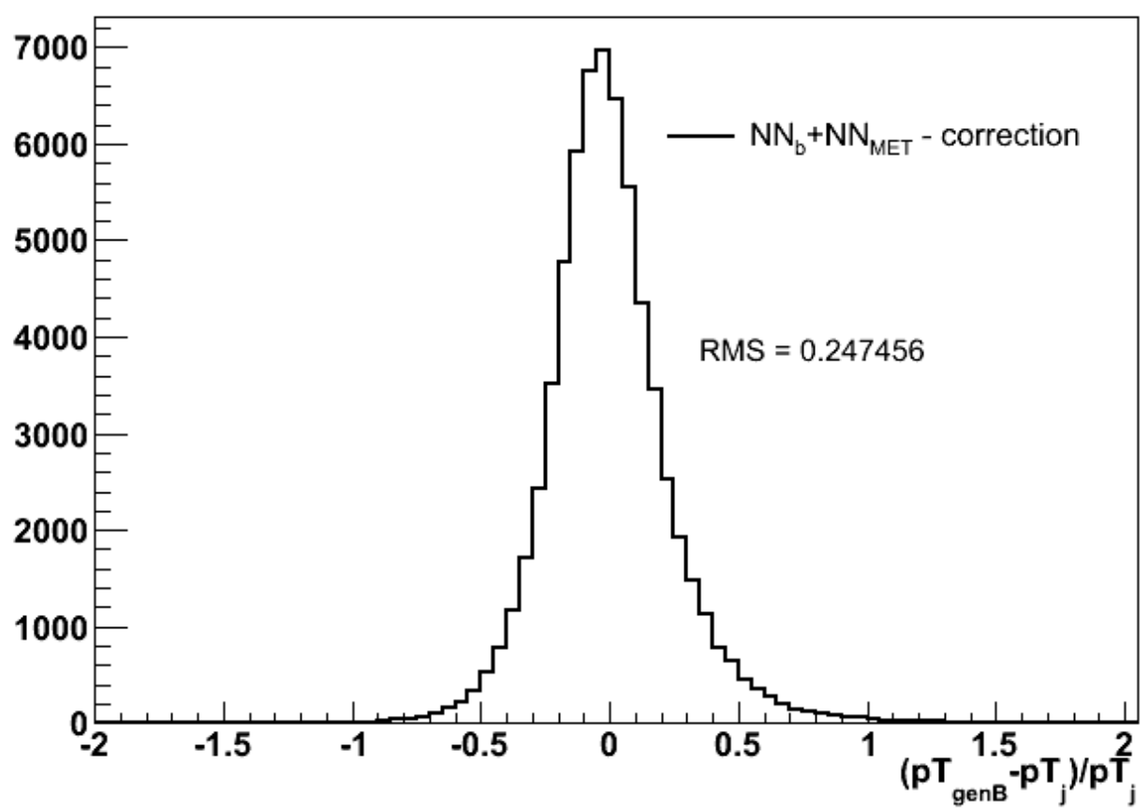
soft jet energy resolution



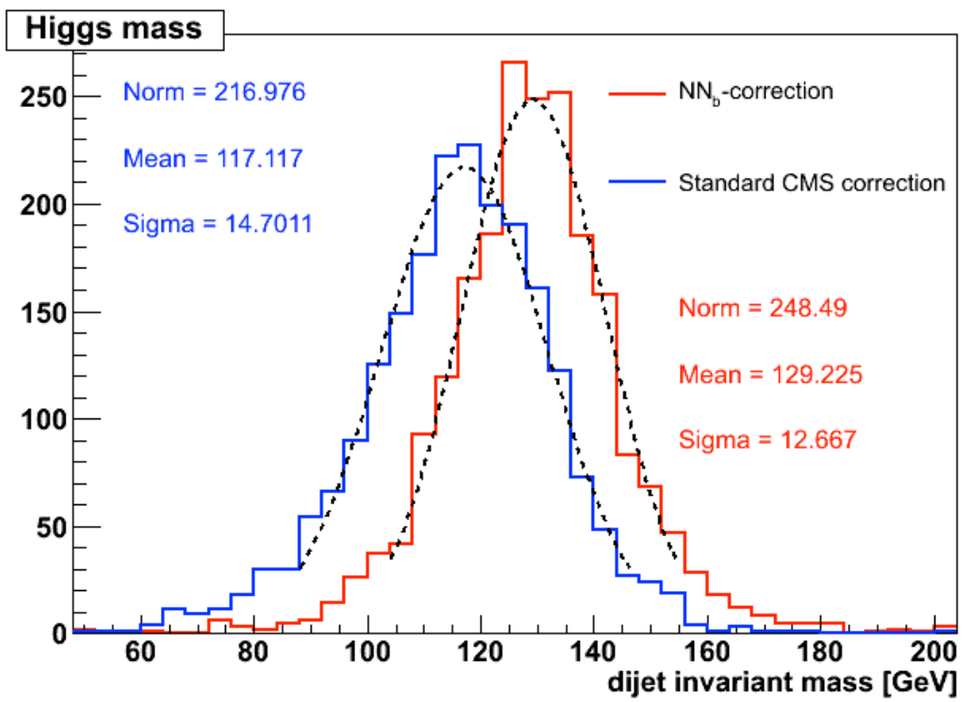
- Jets selection:
  - 2 b-tagged jets from the Higgs candidate with pT > 20 GeV, dR < 0.5
  - no selection applied on additional jets

- **Sub-leading jet energy resolution improved**
  - from ~31% to ~26% after NN$_b$ correction
  - from ~26% to ~25% after NN$_{MET}$ correctior

sub-leading jet energy resolution



Reconstructed dijet invariant mass

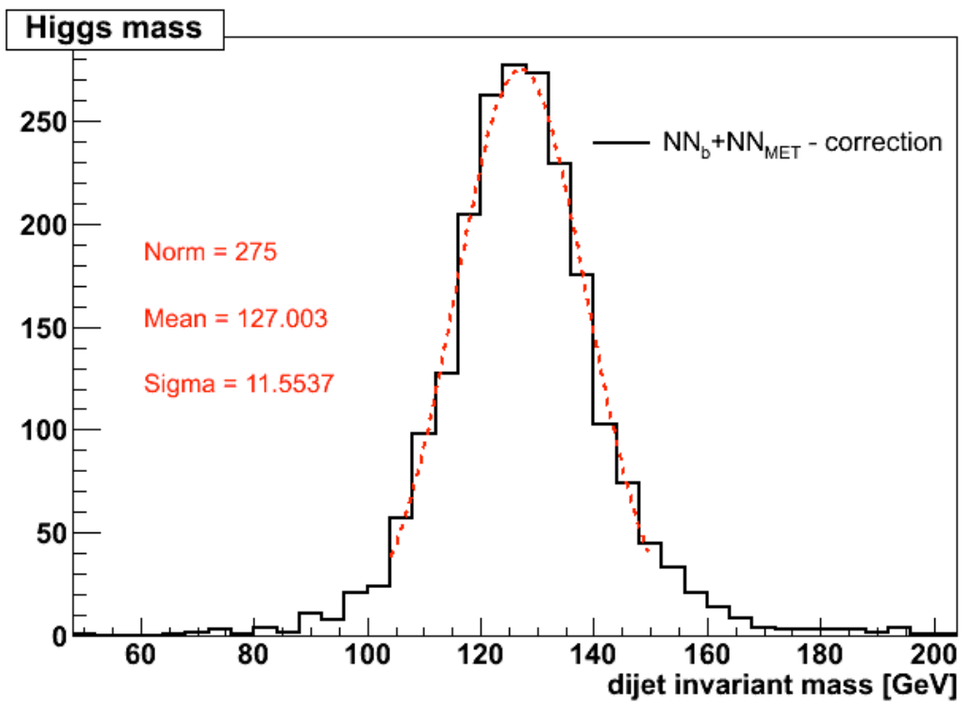


Leptons selection

- pT > 20 GeV
- |eta| < 2.5
- # additional leps = 0

Jets selection

- pT > 20 GeV
- |eta| < 2.5
- CSV l-j > 0.9
- CSV sl-j > 0.5
- # additional jets = 0
- # fat jets = 0

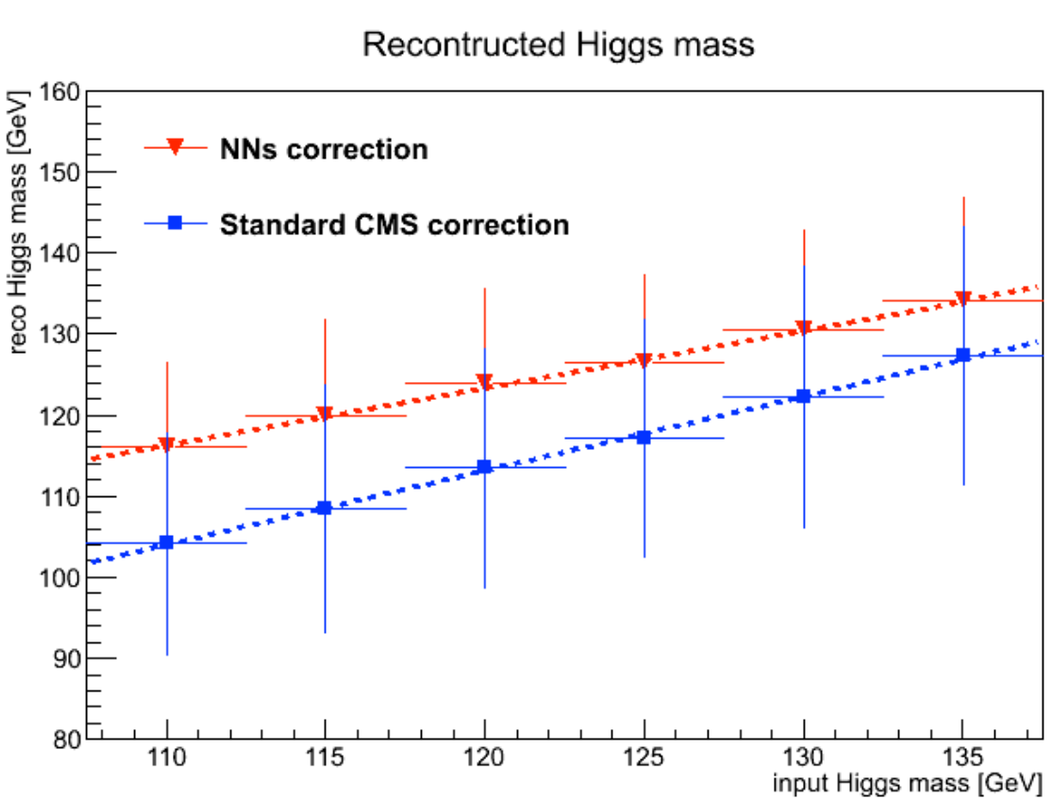


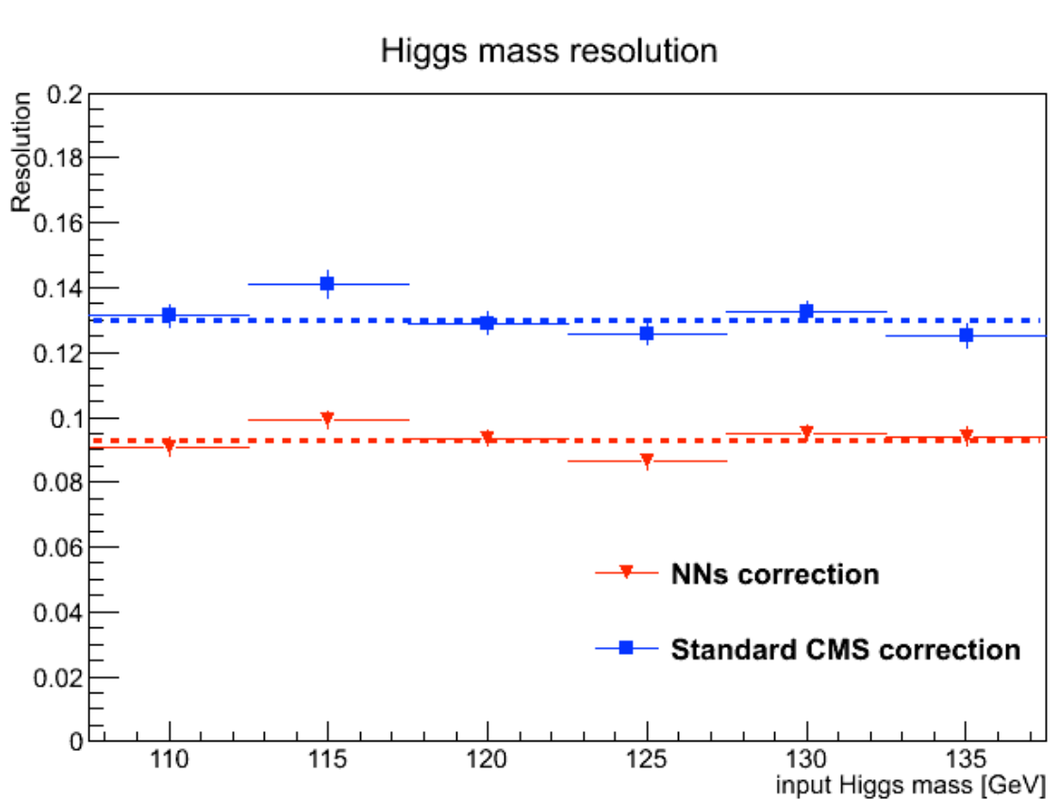
Higgs dijet invariant mass resolution improved

- from ~13% to ~10% after NN_b correction
- from ~10% to ~9% after NN_{MET} correction

Reconstructed dijet invariant mass

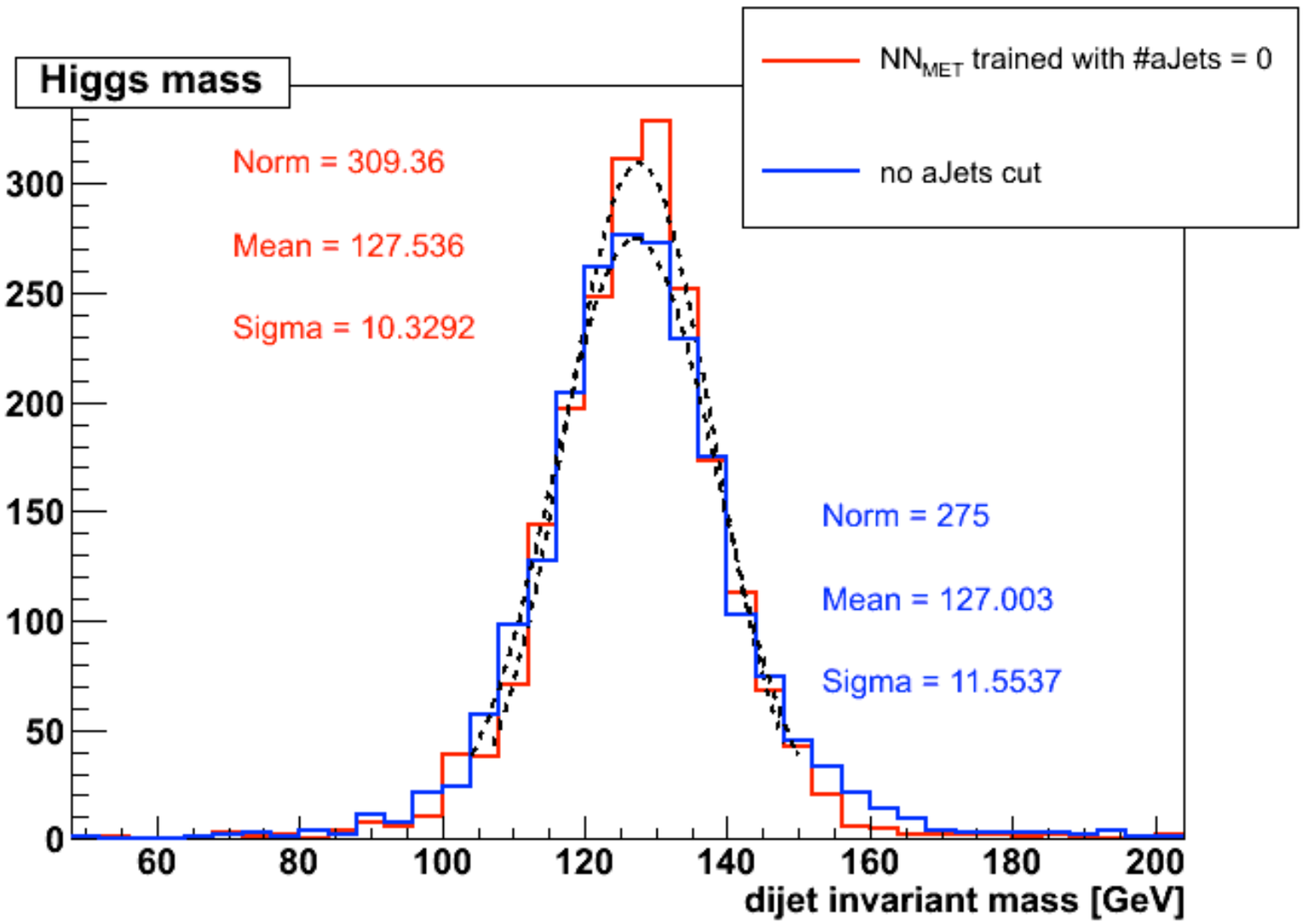
- A bias of the jet-energy correction with respect to a particular Higgs mass in the training is avoided training the NNs @M_H = 110-135 GeV (5 GeV steps)
- Then the NN correction function is evaluated on each of the Higgs boson masses in the range
- Resolution improved from the standard-corrected jets result of ~13% to the NNs-corrected jets result of ~9% (~30%)





Reconstructed dijet invariant mass

- NNMET trained again on a sample of events where no additional jets are found apart from the Higgs candidate jets
  - further improvement of the Higgs mass resolution from ~9% to ~8%



Conclusions

- Used new approach to perform two-stage jet energy correction
  - b-specific corrections on jet-by-jet basis
  - MET-specific corrections on event-by-event basis
  - Maintains correlations between MET and both jets

- Obtained improvement in M_{jj} resolution to 8-9%

- Some additional studies to do
  - Tests of background sculpting

- Look forward to contributing to Hbb group !