Data/MC Discrepancy In Average Track Multiplicity and Its Impact on MET

Kevin Burkett, *Yanyan Gao* (FNAL)

Dmytro Kovalskyi (UCSB)

CMS TRK/B-Tag Prompt FeedBack Meeting

April 7, 2010

Outline

- Compare <nTracks> and pT between 7 TeV and 900 GeV
- Compare Data/MC <nTracks> Ratio at 7 TeV
- Compare Data/MC <nTracks> Ratio at 900 GeV
- <nTracks> Data/MC Difference Effect to MET

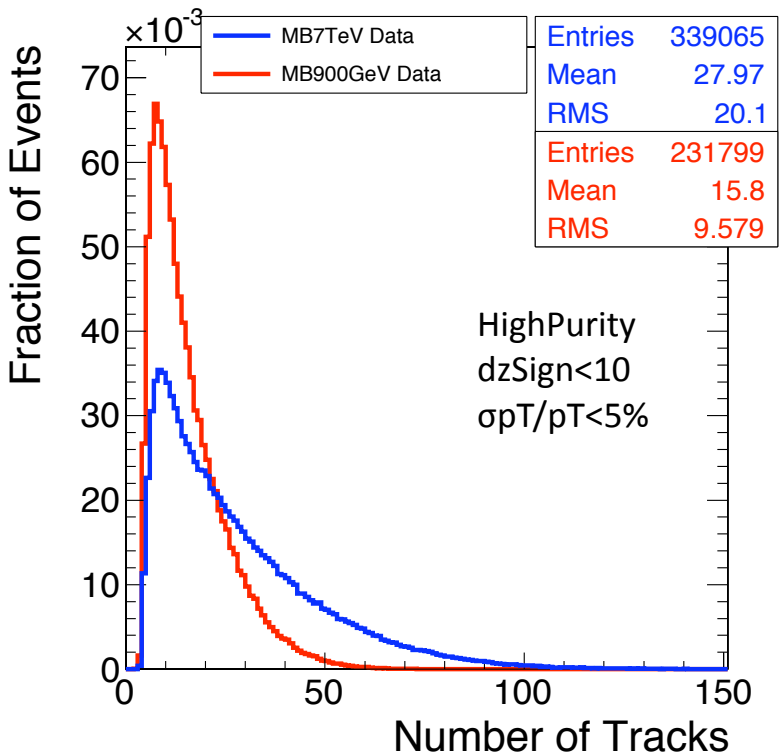
  * ReWeight Data according to Data/MC nTracks distribution

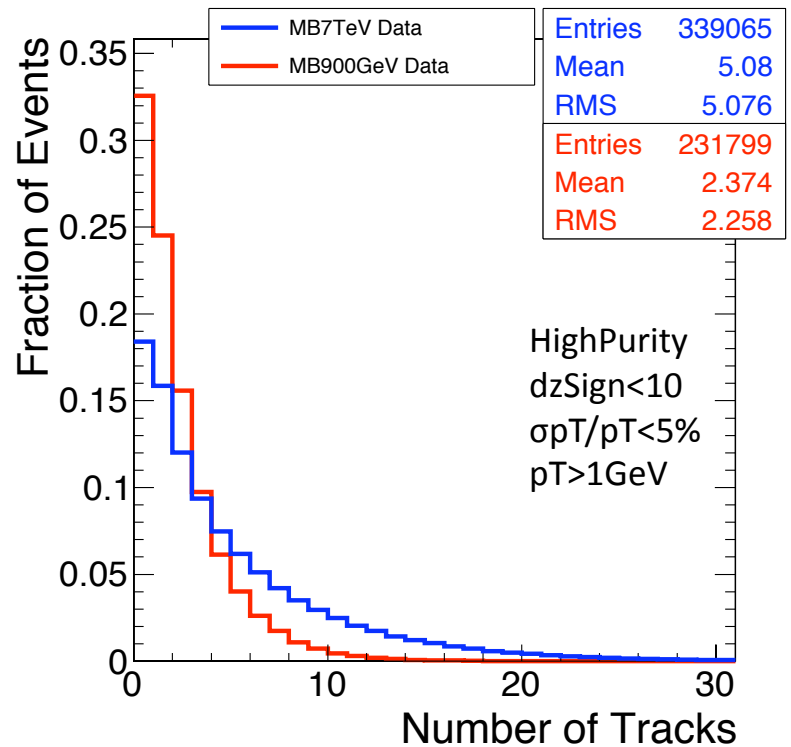
- Summary

1. Release: 356
2. Data (132440/477): /MinimumBias/Commissioning10-*Apr1Skim_GOODCOLL*-v1/RAW-RECO
3. MC: /MinBias/Spring10-START3X_V26A_356ReReco-v1/GEN-SIM-RECO
4. MET values are taking out of the box without additional cleaning

Compare 7TeV and 900 GeV nTracks

- We *expect* <nTracks> increase roughly *(80%-100%)* from 0.9->7 TeV, *observe 80%*





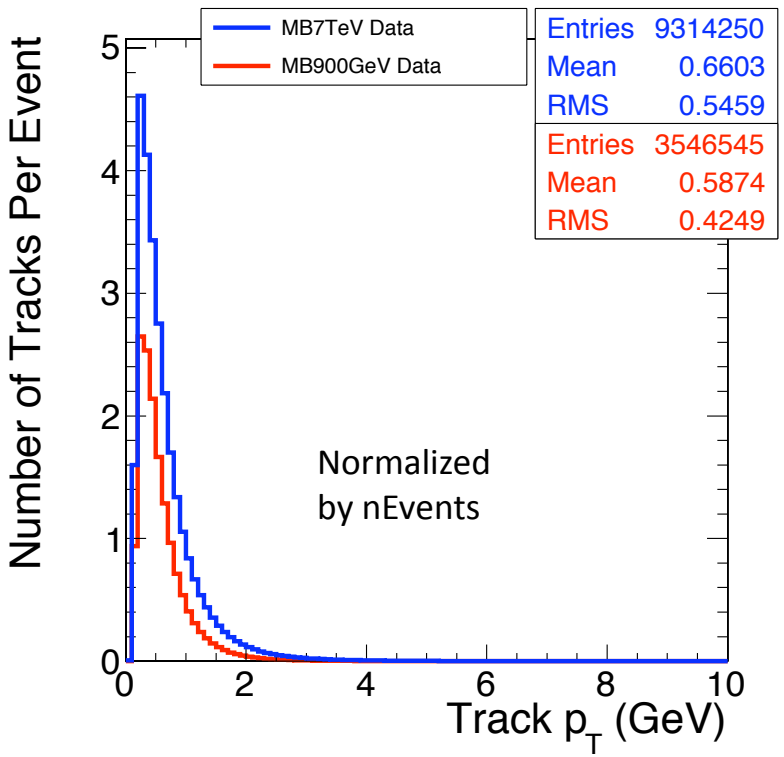
- *<nTracks>(√s) two good fit functions:* *arXiv:0912.0023[hep-ex]*

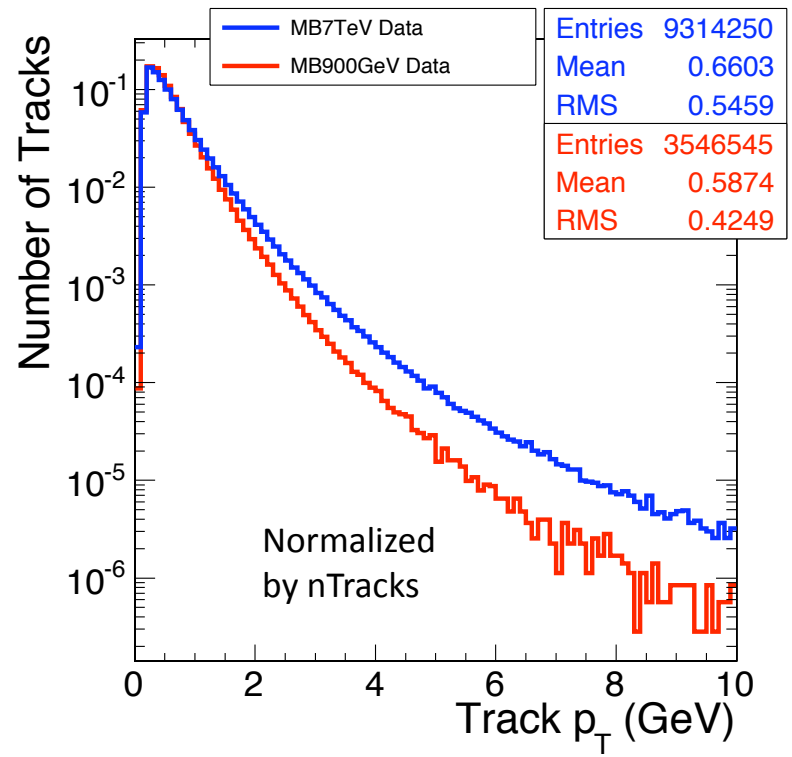
1) $a + 2b\ln(\sqrt{s}) + 4c\ln^2(\sqrt{s}),\ \ (a = 16.65, b = -3.147, c = 0.334)$ <nTracks>(7TeV)/<nTracks>(900GeV) = 1.8

2) $b\sqrt{s}^{2n},\ \ (b = 3.102, n = 0.178)$ <nTracks>(7TeV)/<nTracks>(900GeV) = 2.1

Compare 7TeV and 900 GeV nTracks

- From 0.9->7 TeV, we also expect more high pT tracks

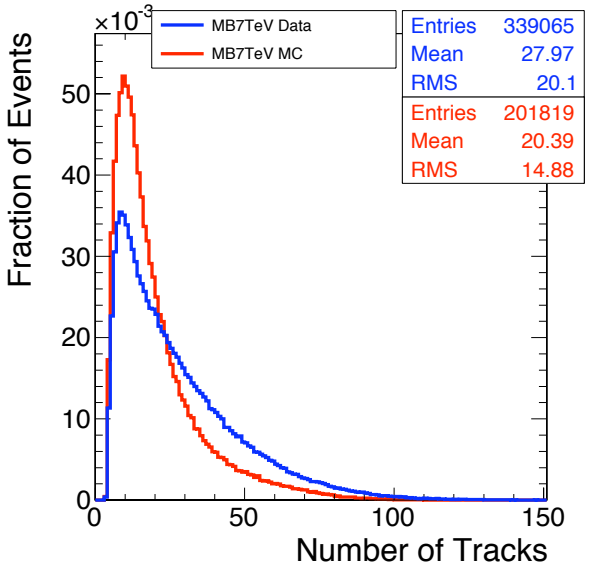


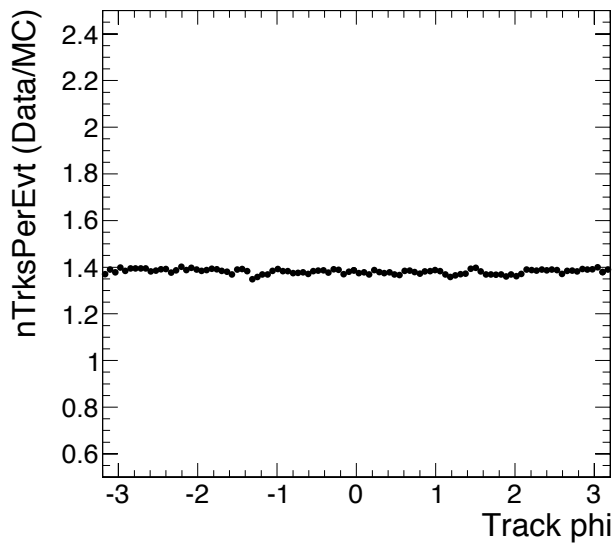


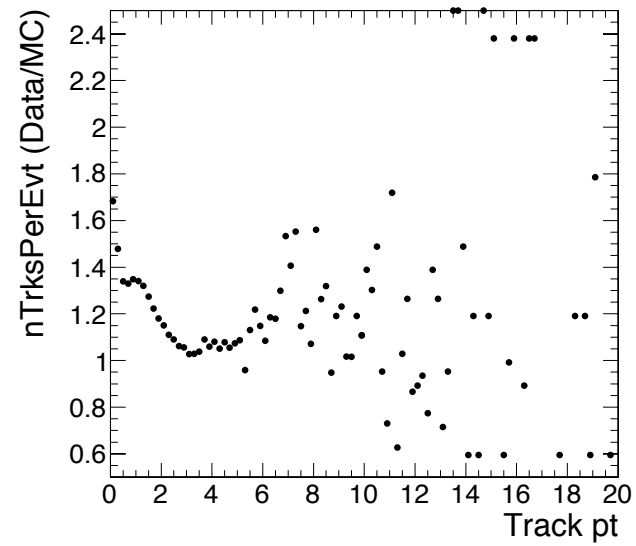
- Track multiplicity and pT at 7 TeV data looks reasonable compared to 900GeV

7TeV Data/MC nTracks Difference (1/3)

- Track selection "*HighPurity && |dzSign| < 10 && σpT/pT < 5%*"



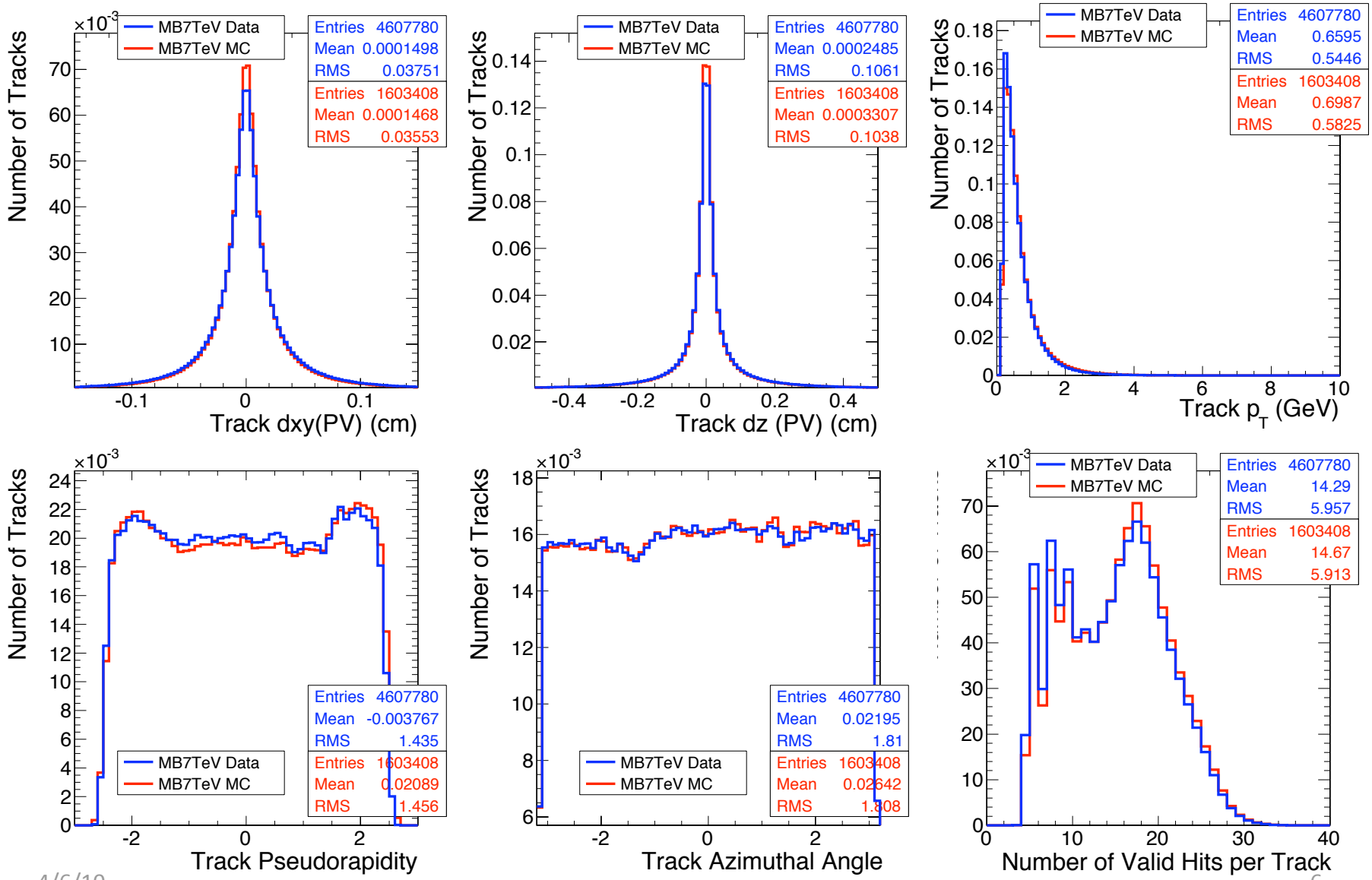




- The track distributions are consistent on track base
- <nTracks> in Data is ~ 37% (*~ 8 tracks*) more than MC, excess in pT < ~ 4GeV
- High pT (> 4 GeV) tracks are limited by statistics
- Both absolute and relative data/mc discrepancies are larger in 7 TeV than in 900GeV

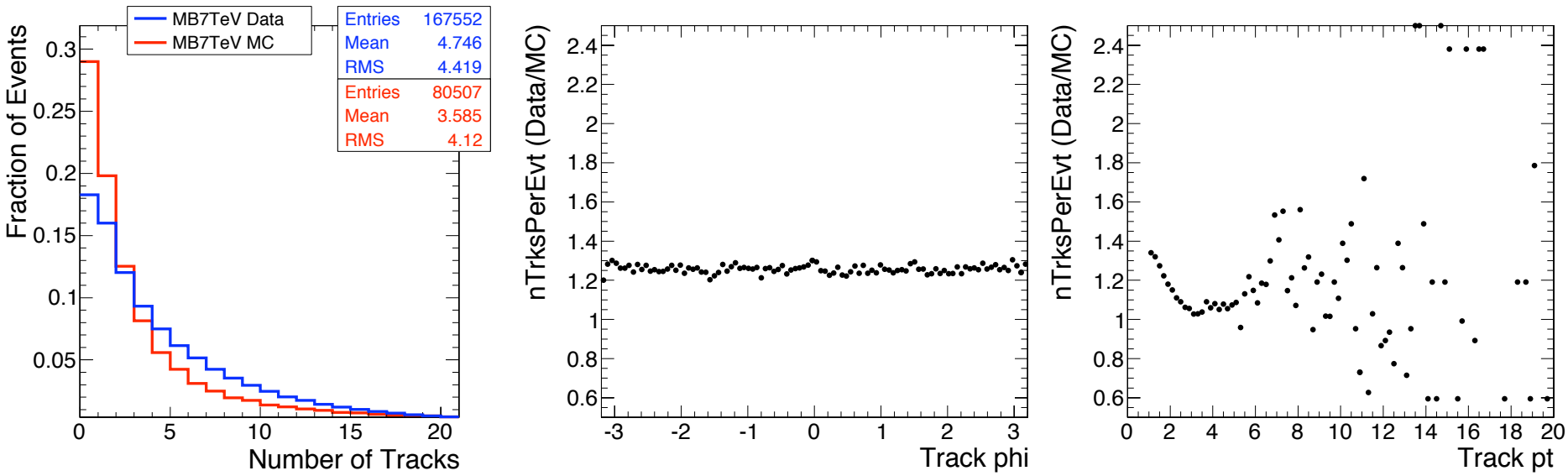
Track Distributions NormBynTracks (2/3)

- Track selection "*HighPurity && |dzSign| < 10 && σpT/pT < 5%*"



7TeV Data/MC nTracks Difference (3/3)

- Track selection "*HighPurity && |dzSign| < 10 && σpT/pT < 5% && pT > 1GeV*"



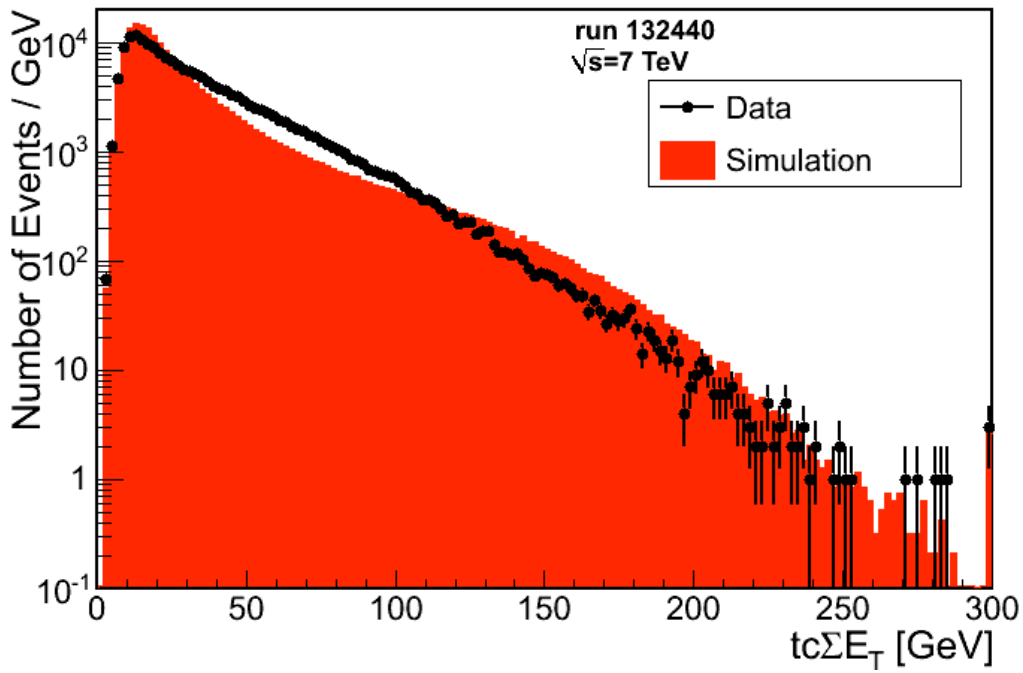
- The track distributions are consistent on track base
- <nTracks> in Data is ~ 25% (*~ 1-2 tracks*) more than MC, excess in pT < ~ 4GeV
- High pT (> 4 GeV) tracks are limited by statistics
- Relative data/mc discrepancy at 7 TeV is roughly the same as at 900GeV
- Absolute data/mc discrepancy is larger at 7 TeV than in 900GeV

Data/MC Discrepancy in MET SumEt

- nTracks Data/MC discrepancy indicates a particle level inefficiency in CMS MC simulation
- It should be considered when we compare data/mc, beyond Tracking scope
- Example: big data/mc discrepancy is observed in *SumEt but not in MET:*

1. *tcMET: Warren Andrews in MET working meeting*
2. *Similar discrepancy is observed in caloMET and pfMET*

http://indico.cern.ch/conferenceDisplay.py?confId=89806

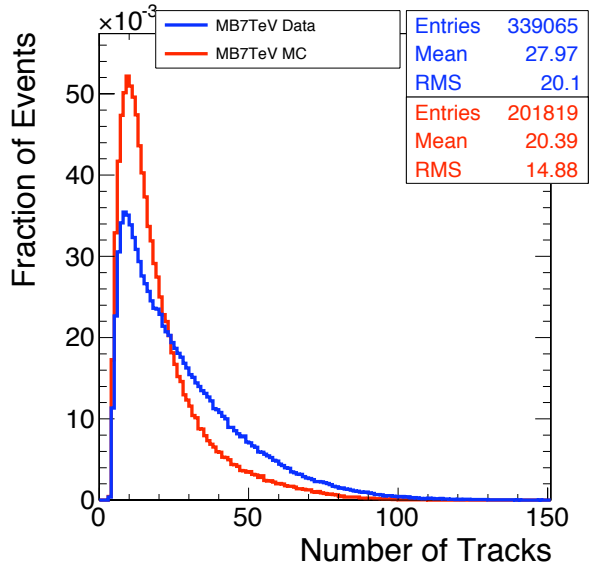


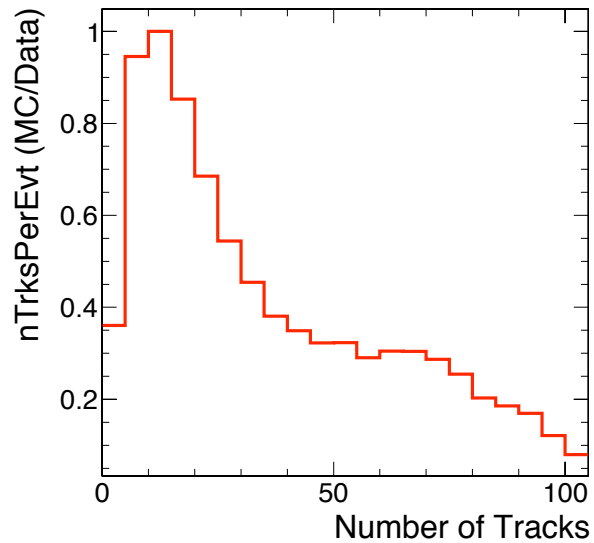
Hypothesis: If we get the nTracks distribution consistent, would it help matching sumEt?

Re-Weight Data nTracks Distribution (1/4)

- Divide MC/Data nTracks histograms, and scale the largest Ratio to 1

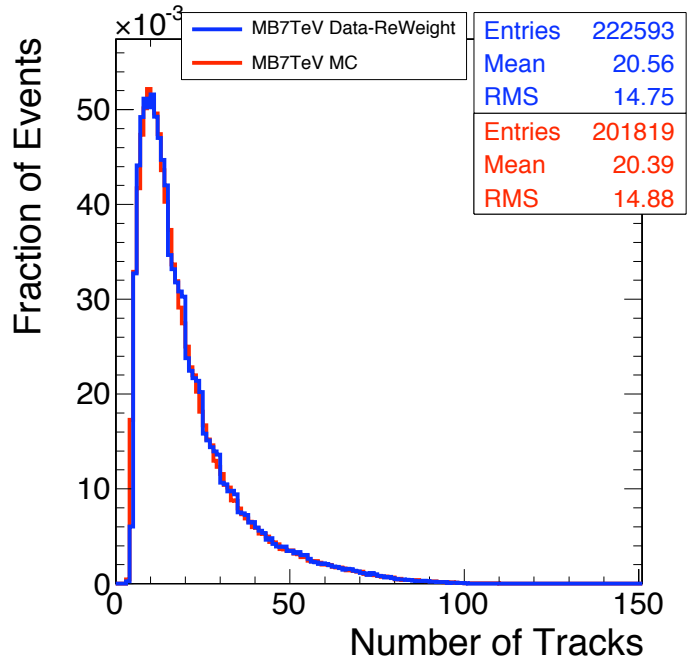
- Correct up to nTracks = 100





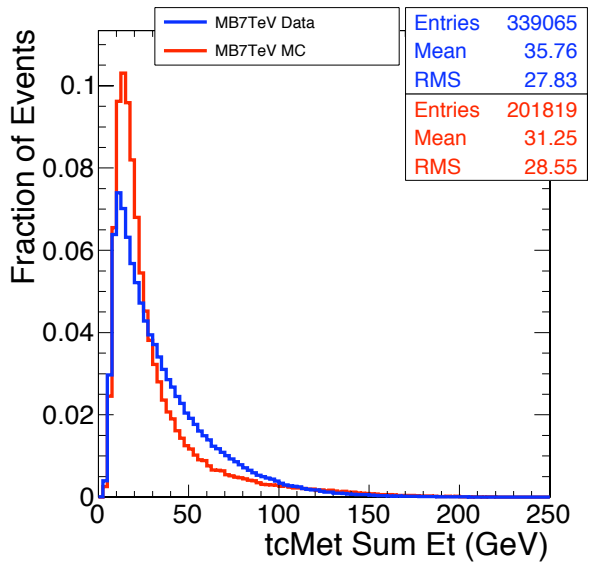
- After reweighting, obtained consistent nTracks distribution

- Track distributions are consistent between ReWeighted-Data and MC

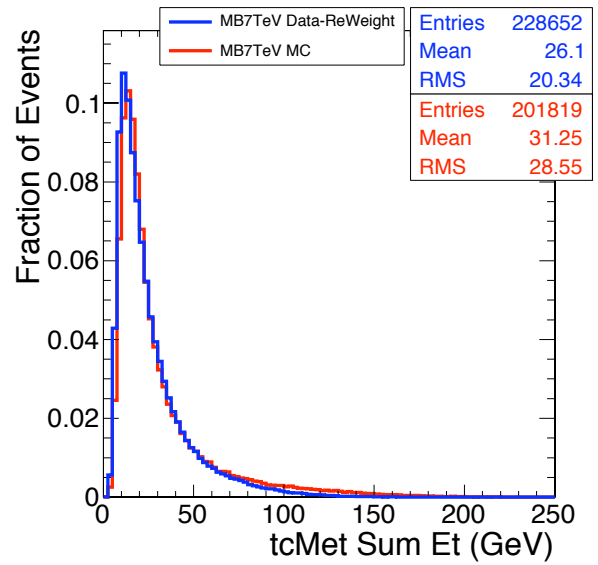


Re-Weight Data Effect to tcMET (2/4)

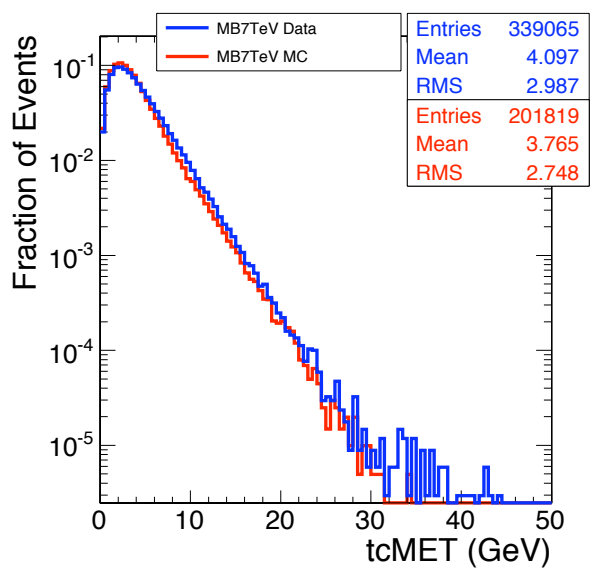
- Sum Et



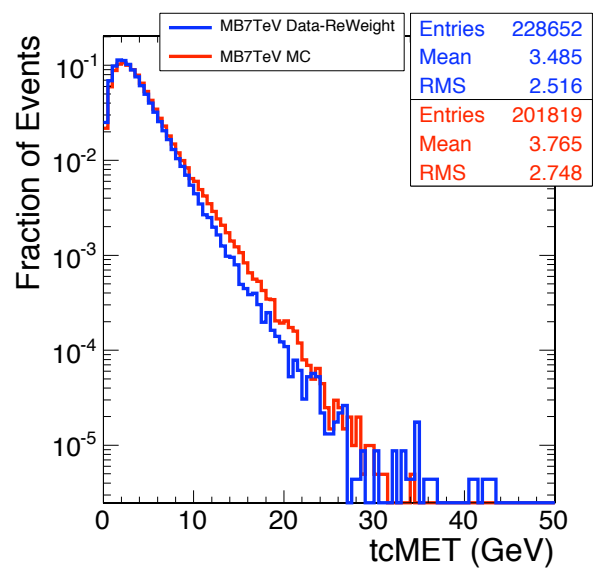
Reweight Data



- MET

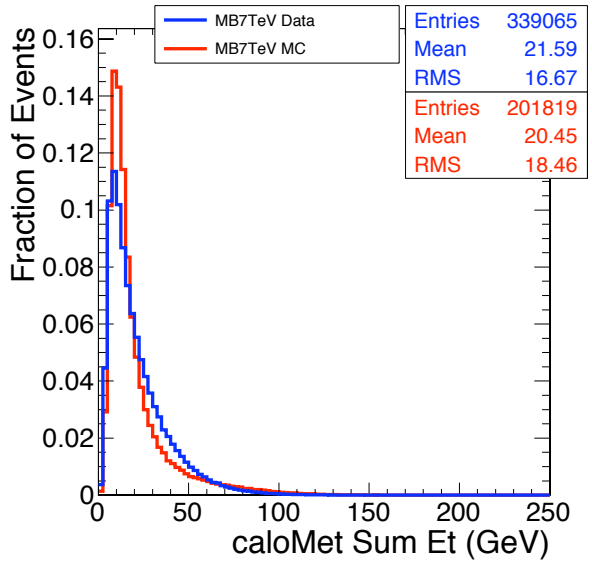


Reweight Data

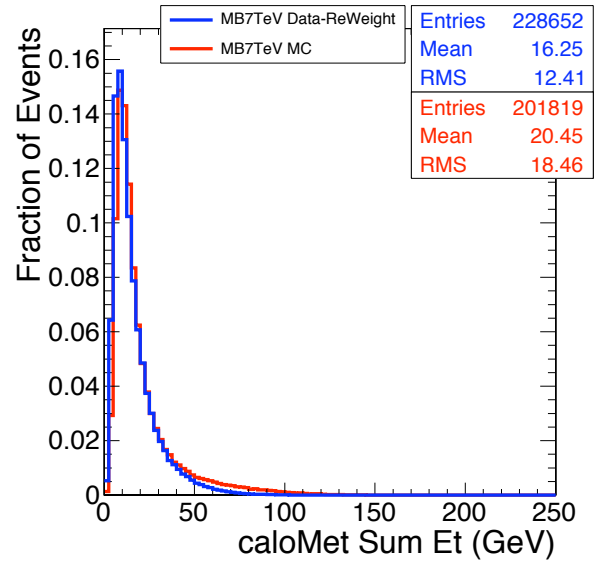


Re-Weight Data Effect to CaloMET (3/4)

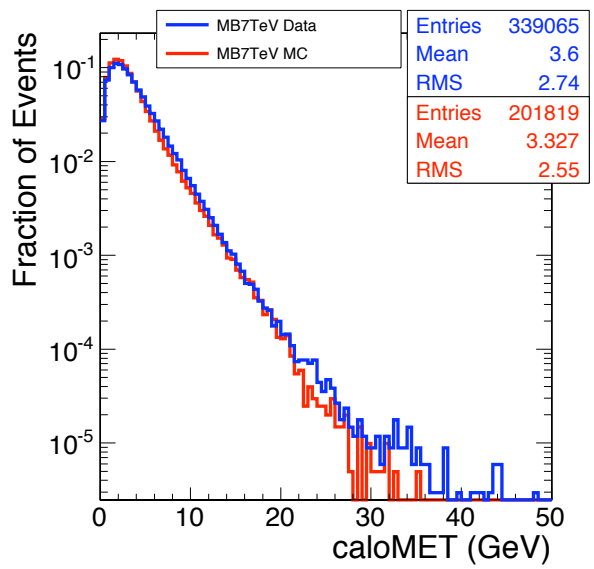
- Sum Et



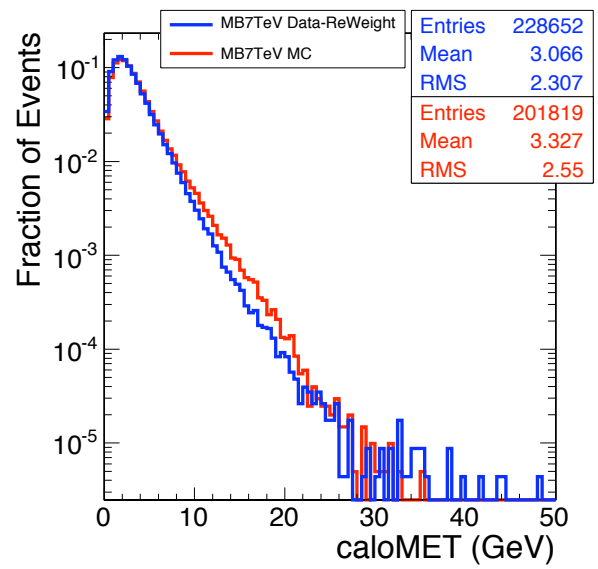
Reweight Data



- MET

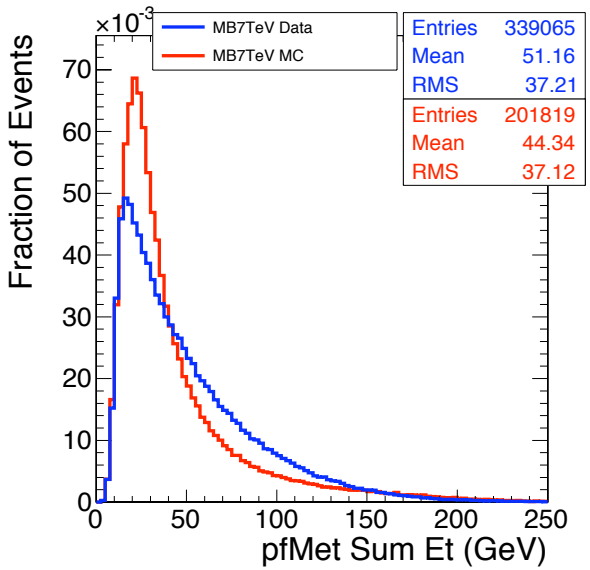


Reweight Data

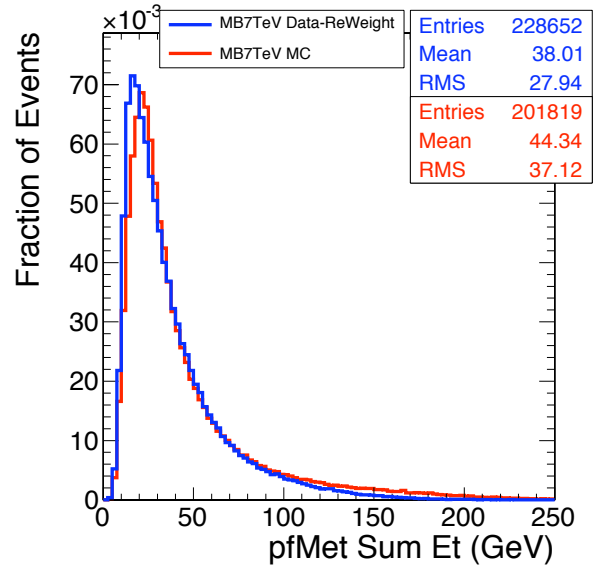


Re-Weight Data Effect to pfME (4/4)

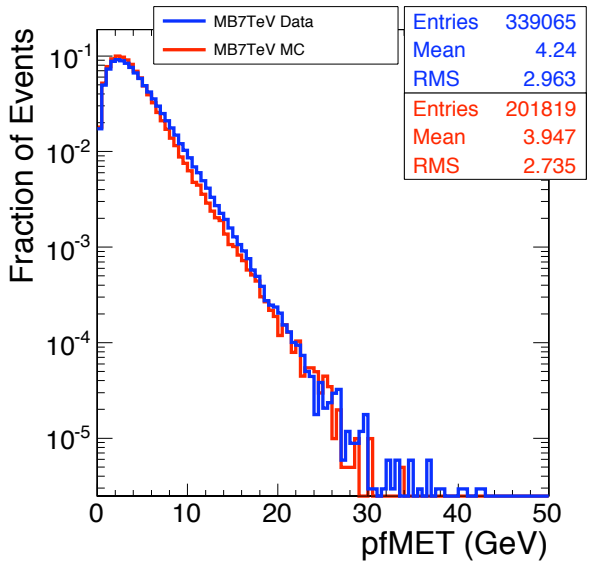
- Sum Et



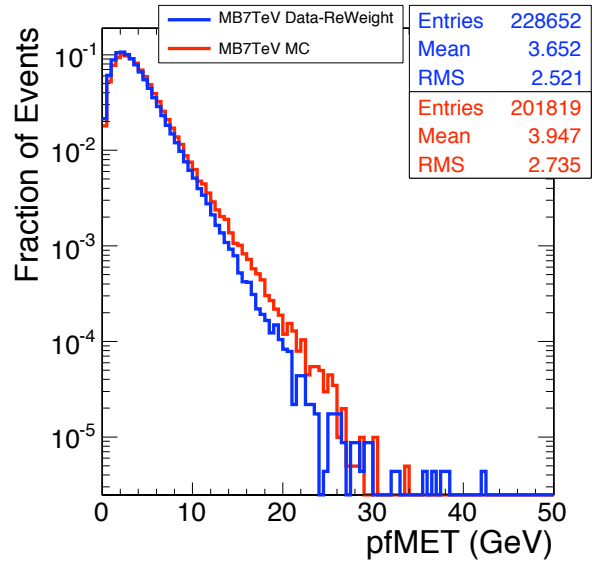
Reweight Data



- MET



Reweight Data



Summary

- The <nTracks> at 7 TeV is about 1.8 times of the value at 900 GeV

- Data/MC excess in <nTracks> at 7 TeV show larger discrepancy than 900GeV

  ** Data has ~ 37 % (~ 8 tracks) more tracks mainly in pT < 4 GeV*
  ** Data has ~ 25% (> 1 track) more tracks with pT > 1 GeV*

- Check the <nTracks> Data/MC discrepancy contribution to MET SumEt

  ** ReWeight Data according to Data/MC nTracks distribution*

  ** ReWeight MC according to MC/Data nTracks distribution (backup slides)*

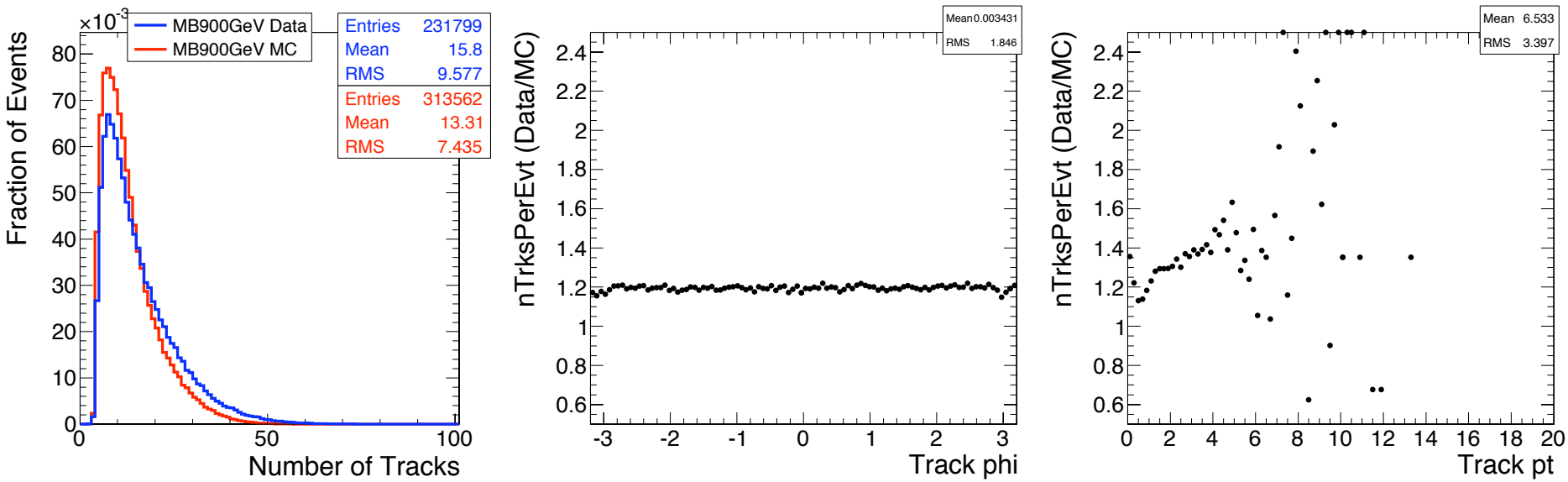
  ** In both methods, SumEt match better in lower end and the MET not affected*

- Pythia MC used in CMS simulates less soft particles, it need to be retuned!

BACKUP SLIDES

900GeV Data/MC nTracks Difference (1/2)

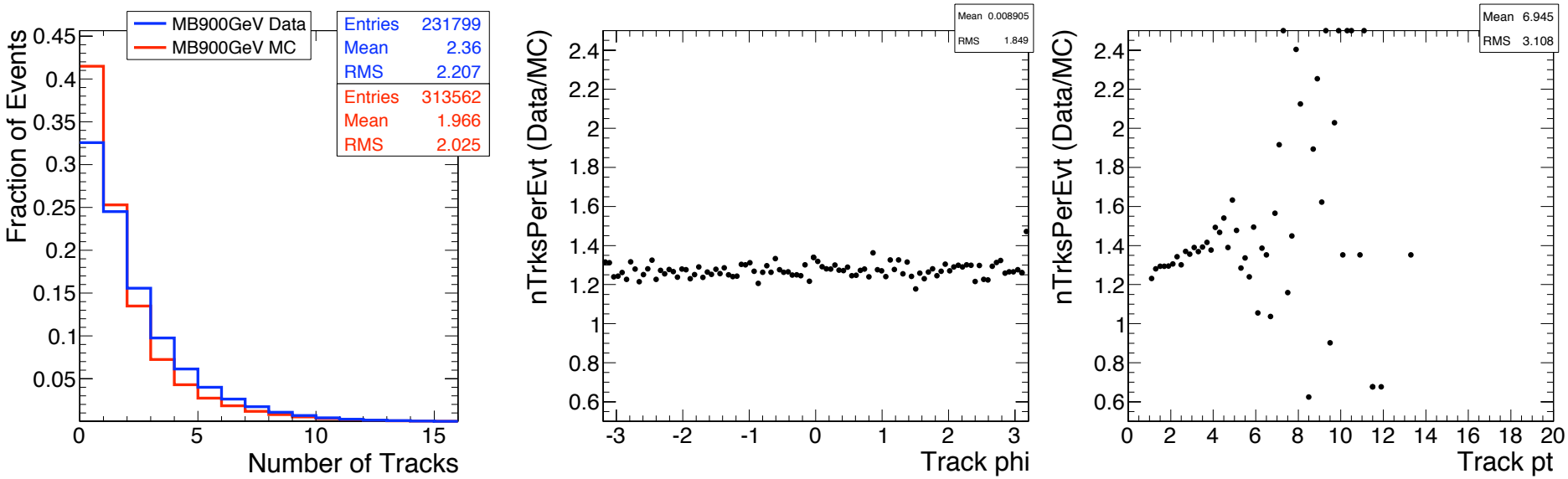
- It has been known that pythia MC used in CMS not tuned well, simulating less soft particles
- Track selection "*HighPurity && |dzSign| < 10 && σpT/pT < 5%*"



- The track distributions are consistent on track base (see TRK-10-001)
- <nTracks> in Data is ~ 20% (*~ 2-3 tracks*) more than MC, excess in pT < ~ 4GeV
- High pT (> 4 GeV) tracks are limited by statistics.

900GeV Data/MC nTracks Difference (2/2)

- Track selection "*HighPurity && |dzSign| < 10 && σpT/pT < 5% && pT > 1GeV*"

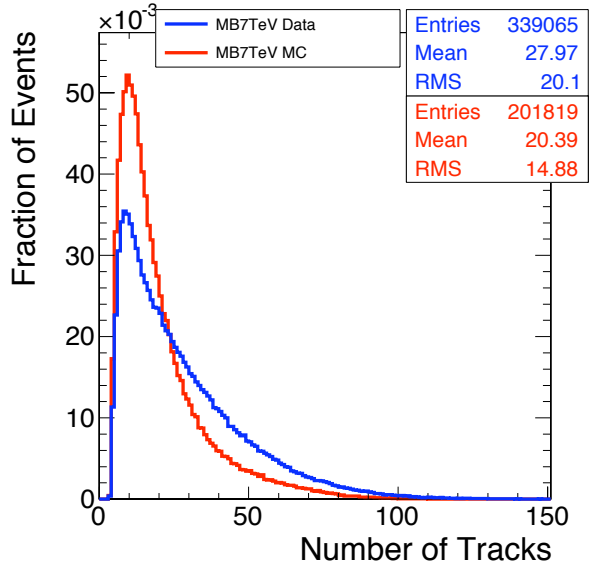


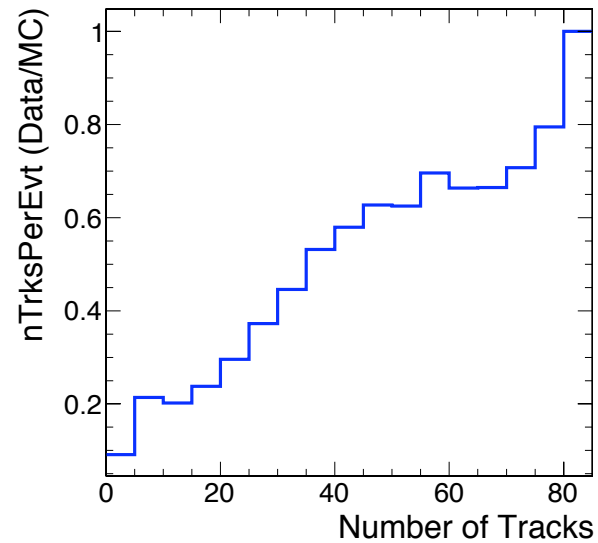
- The track distributions are consistent on track base
- <nTracks> in Data is ~ 25% (*< 1 track*) more than MC, excess in pT < ~ 4GeV
- High pT (> 4 GeV) tracks are limited by statistics.

Re-Weight MC nTracks Distribution (1/4)

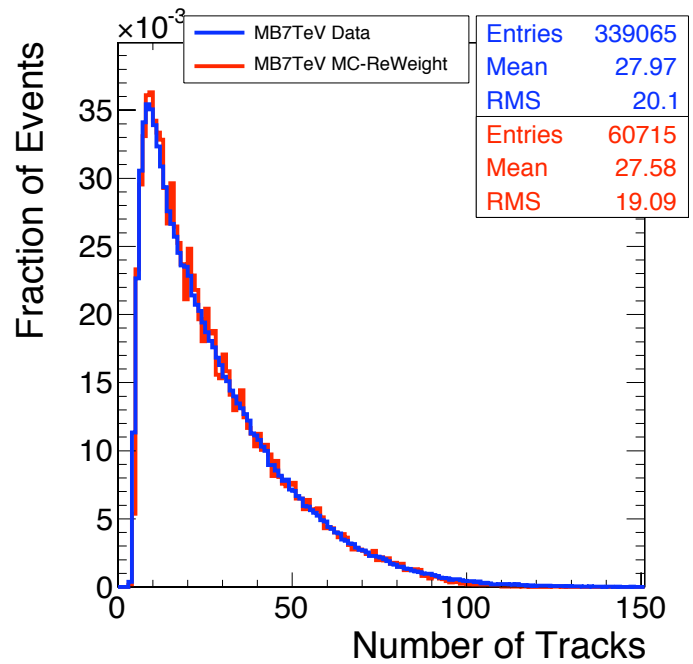
- Divide Data/MC nTracks histograms, and scale the largest Ratio to 1

- *Correct up to nTracks = 80*



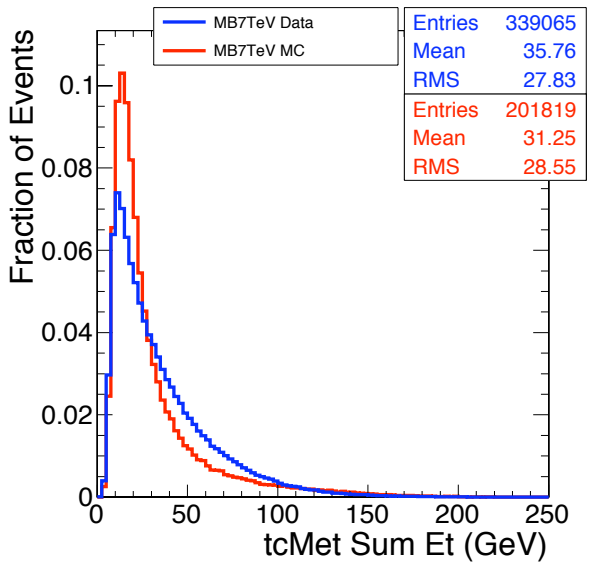


- After reweighting, obtained consistent nTracks distribution

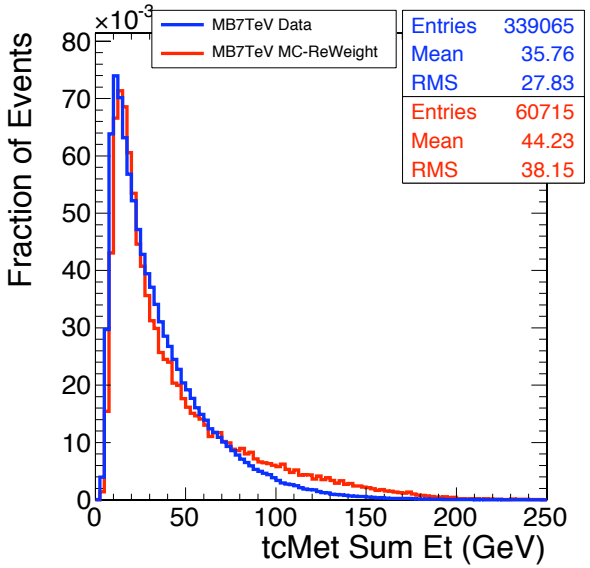


Re-Weight MC Effect to tcMET (2/4)

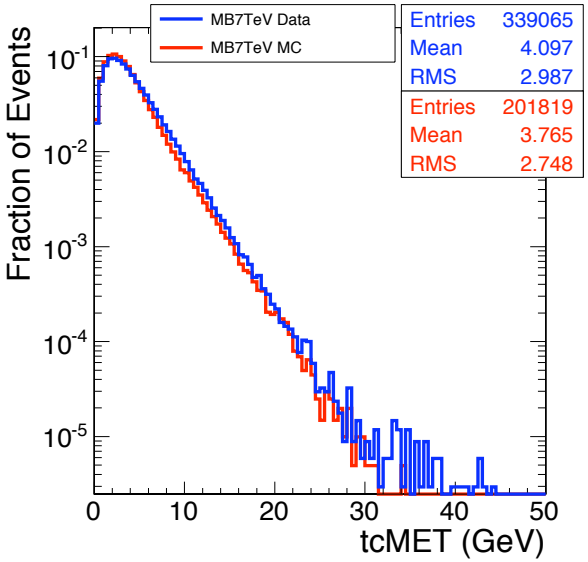
- Sum Et



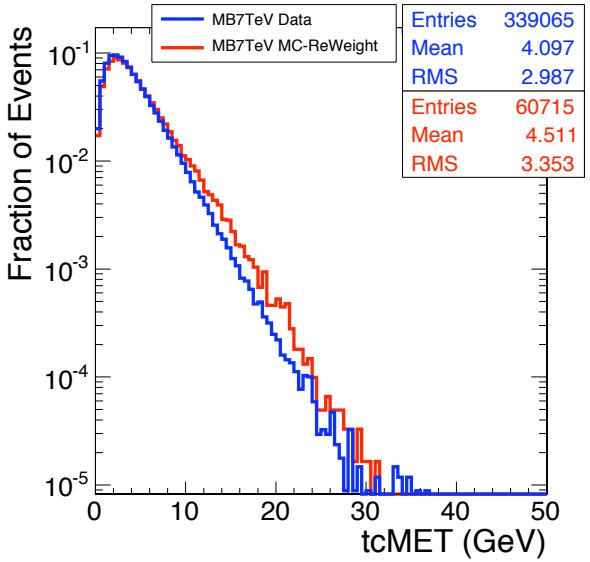
Reweight MC



- MET



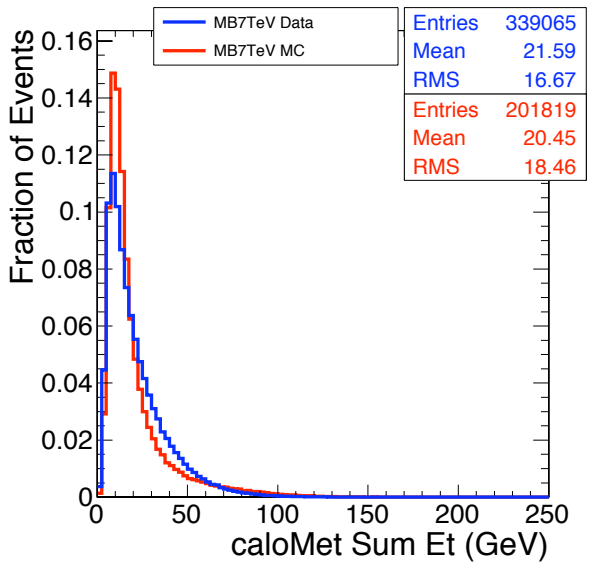
Reweight MC



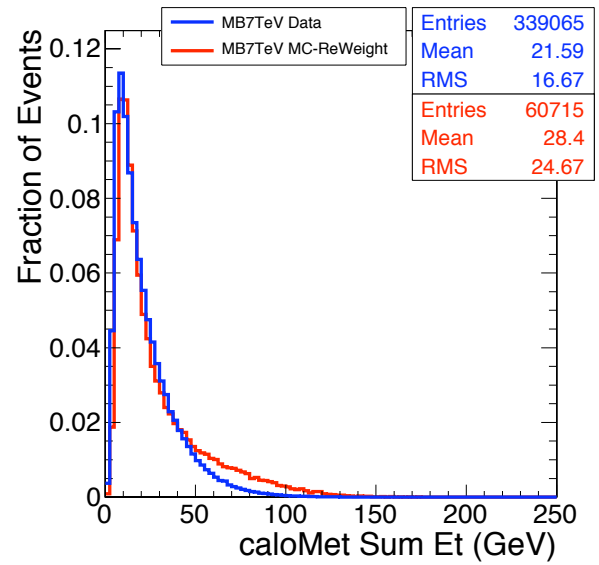
ReWeight up to nTracks = 80, creating artificial long tail in MC

Re-Weight MC Effect to CaloMET (3/4)

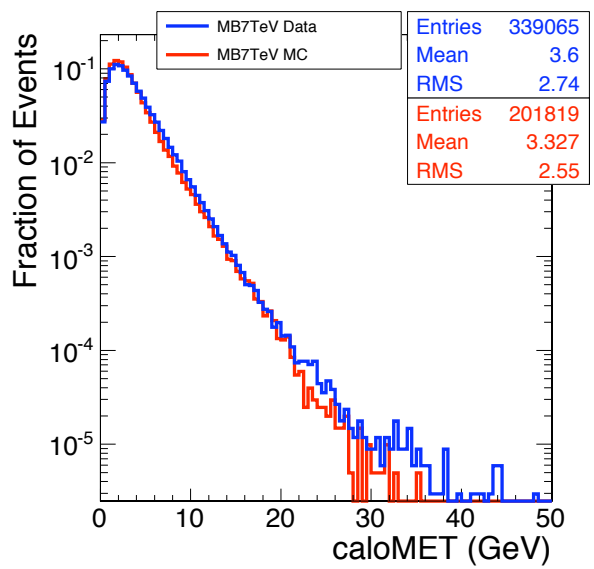
- Sum Et



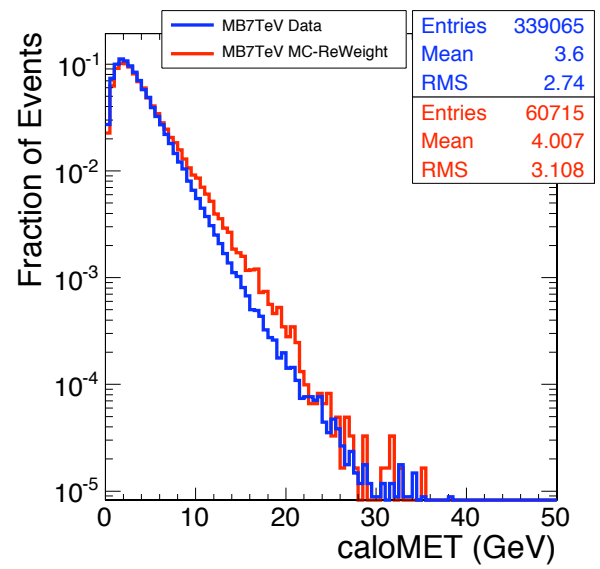
Reweight MC



- MET



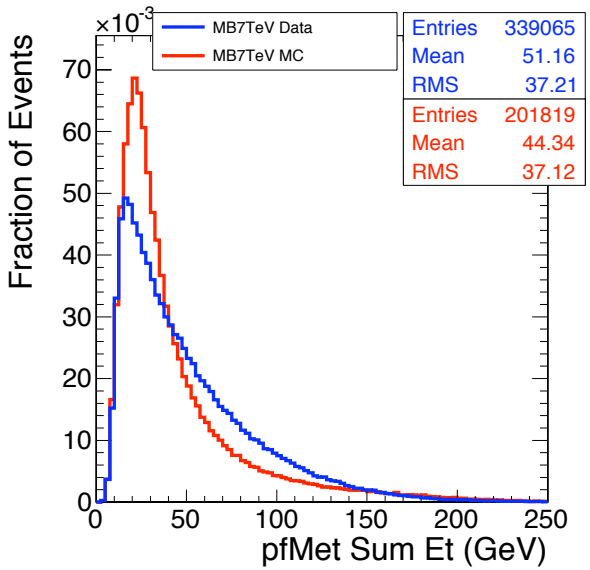
Reweight MC



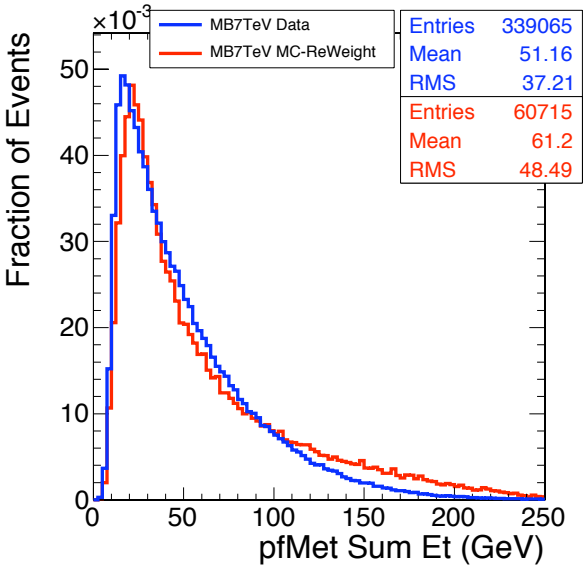
ReWeight up to nTracks = 80, creating artificial long tail in MC

Re-Weight MC Effect to pfMET (4/4)

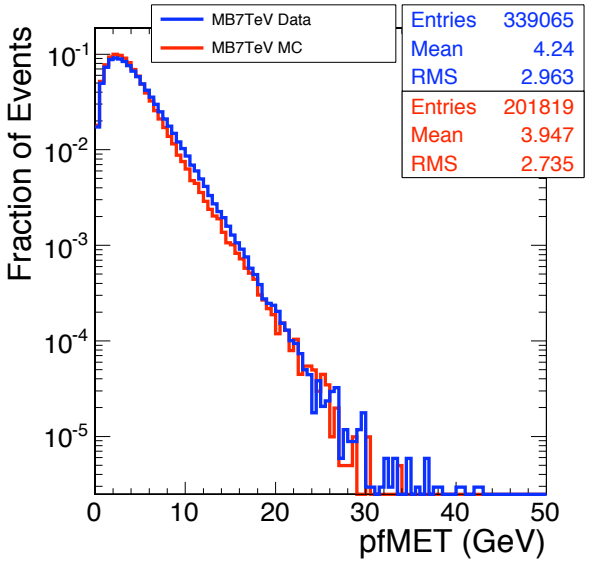
- Sum Et



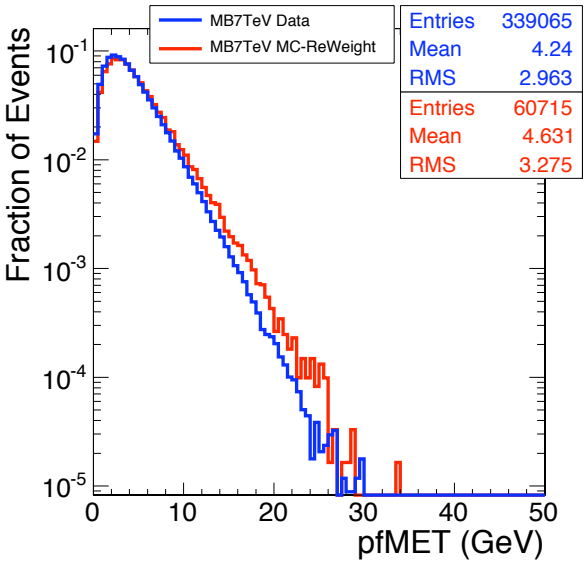
Reweight MC



- MET



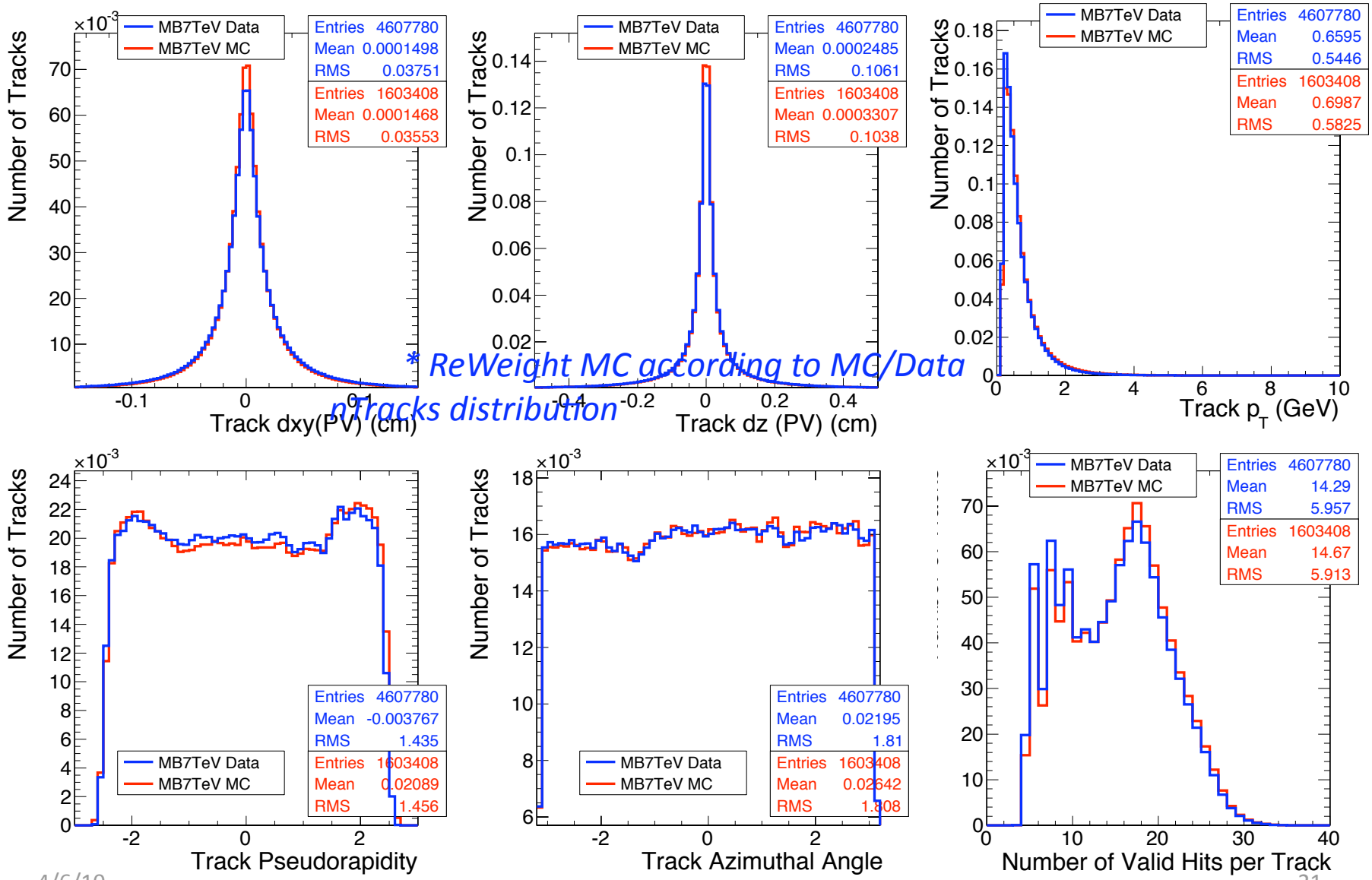
Reweight MC



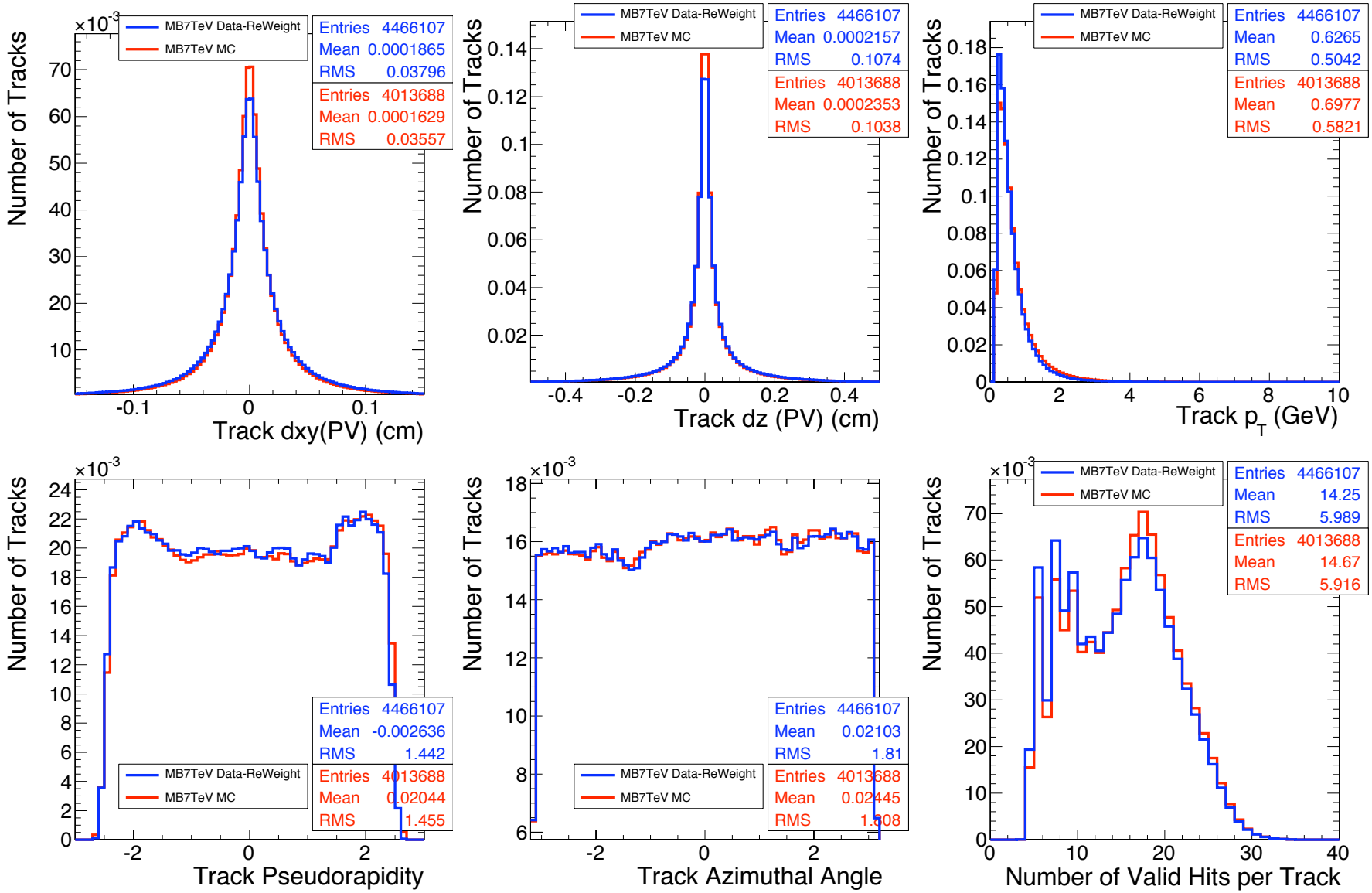
ReWeight up to nTracks = 80, creating artificial long tail in MC

Track Distributions NormBynTracks -MC ReWeighted

- Track selection "*HighPurity && |dzSign| < 10 && σpT/pT < 5%*", MC reweighted



** ReWeight MC according to MC/Data nTracks distribution*

Track Distributions NormBynTracks -Data ReWeighted

- Track selection "*HighPurity && |dzSign| < 10 && σpT/pT < 5%*", Data reweighted