OutliersRejection Impact and Tracking Performance Studies on Cosmic Data

Kevin Burkett, Yanyan Gao*

Fermi National Accelerator Laboratory

15 May 2009 / CERN

CMS Tracking POG Working Meeting

Outline

- CMSSW environment and datasets
- OutliersRejection impact on the cosmic Data
  - Track parameters comparision
  - Rejected hits population based on 1-track events
  - Pull distribution comparision on MC
- Tracking performance study using MultiTrackValidator
  - Track efficiency
  - Track fake rate
  - Track parameter resolutions
  - Track parameter pull distributions
- Summary and Plans

CMSSW Environment and Datasets

- Release: CMSSW_2_2_5

- CRAFT Data: Run 69382

/data/Commissioning08/Cosmics/RAW-RECO/CRAFT_ALL_V11_225_ReReco_FromTrackerPointing_v1/

- Cosmic Monte Carlo

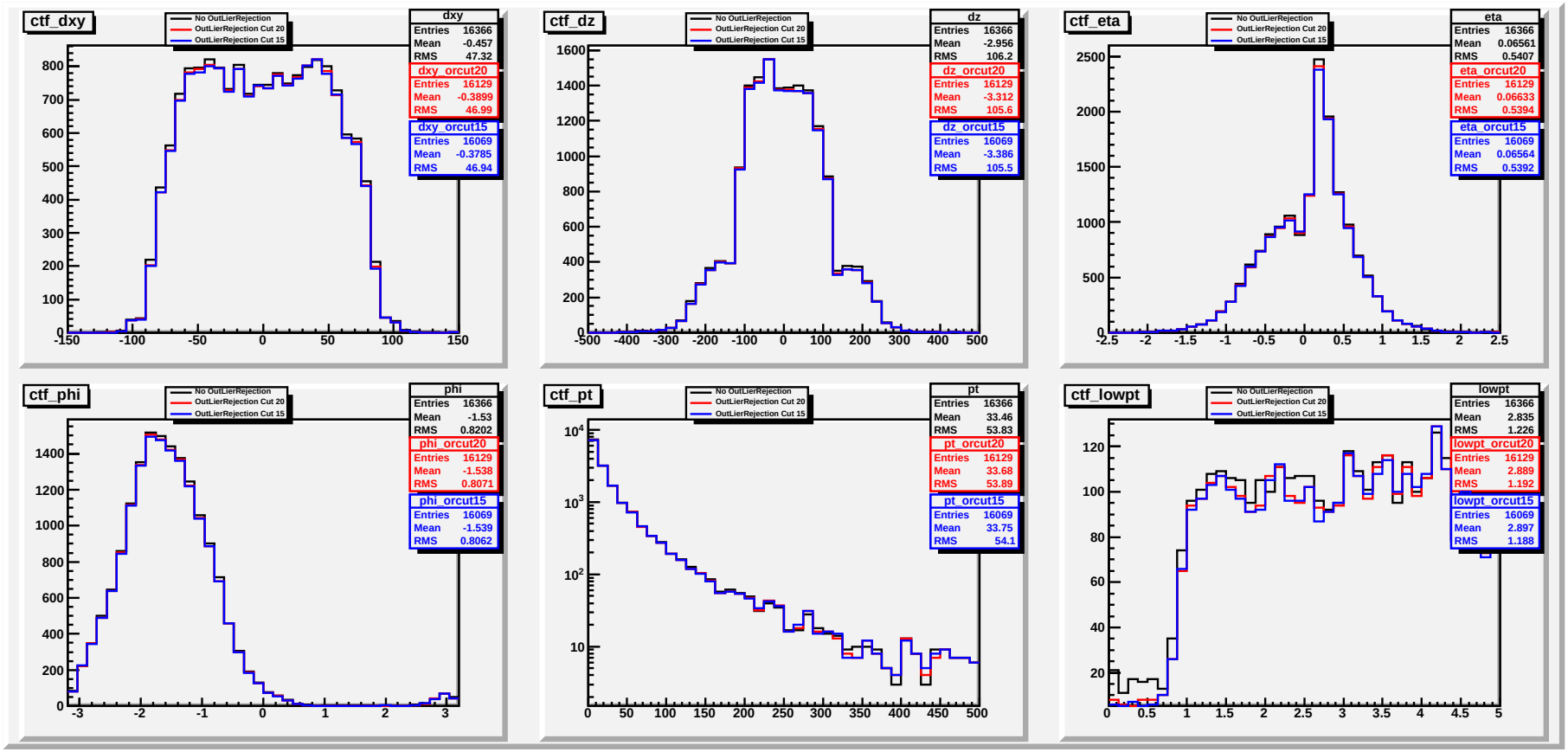
/mc/Summer08/TkAlCosmics4T/GEN-SIM-RECO/COSMMC_22X_TK_225-v2/

Brief Reminder of OutliersRejection

- OutliersRejection is implemented in the final fit
  - Reject hits with large χ^2 over a thresold cut and refit the track
  - The rejected hits could be noisy, from another track, or generated from complex nuclear processes.
- The χ^2 cut is optimized to satisfy the following:
  - Not to worsen the tracking efficiency
  - Improve fake rate
  - Improve as much as possible the quality of the tracks
  - Keep the CPU time at a reasonable level
- Test for OutliersRejection cuts $\chi^2 > 15, 20$
- BreakTrajWith2ConsecutiveMissing is turned off (ON by default)
- Twiki Page: https://twiki.cern.ch/twiki//bin/view/CMS/OutliersRejection

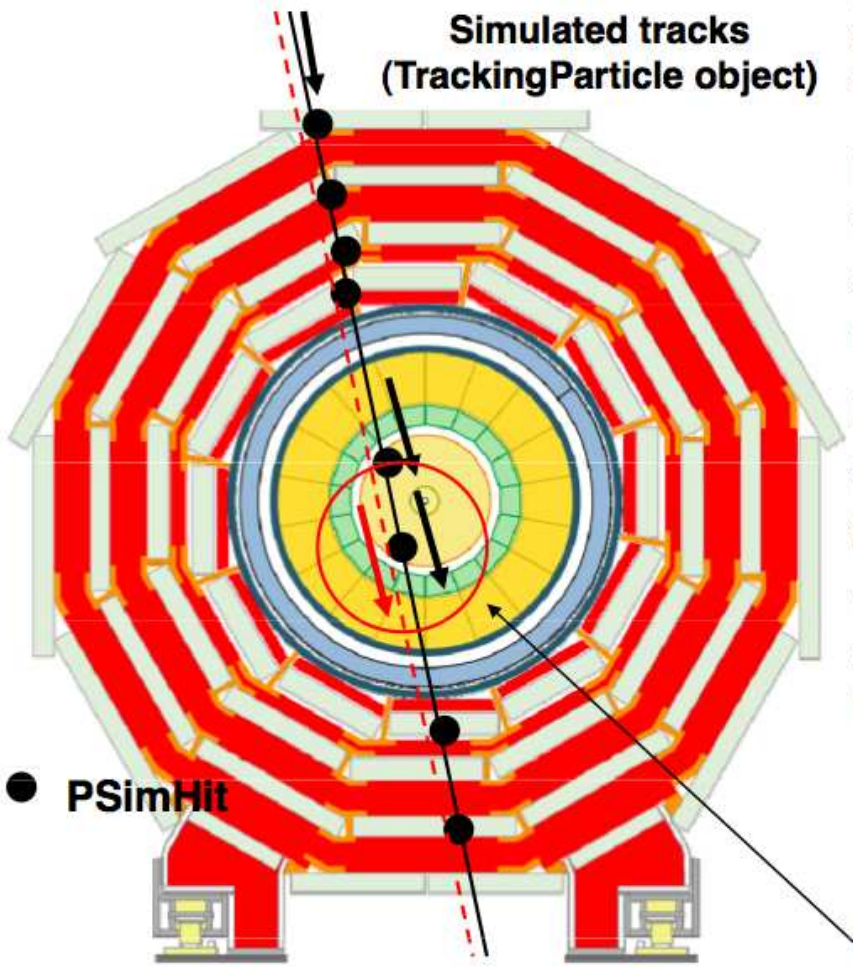
OutliersRejection Impact on Track Parameters

- Selection efficiency: $\chi^2 > 20 \rightarrow 98.6\%$ $\chi^2 > 15 \rightarrow 98.2\%$
  - Black: $\chi^2 > -1$ – Red: $\chi^2 > 20$ – Blue: $\chi^2 > 15$



- OutliersRejection removes low p_T ($<\sim 1$ GeV)
- Track parameters are consistent with/without the OutliersRejection

Use of Tracking Particles (from Boris Mangano)



Each PSimHit of the TrackingParticle contains the state of the particle on the surfaces of sensitive detector volumes.

Looping over all the PSimHits of a TrackingParticle, the state on the closest sensor to the impact point can be determined.

Using the track propagator service, the state at the impact point can be evaluated too.

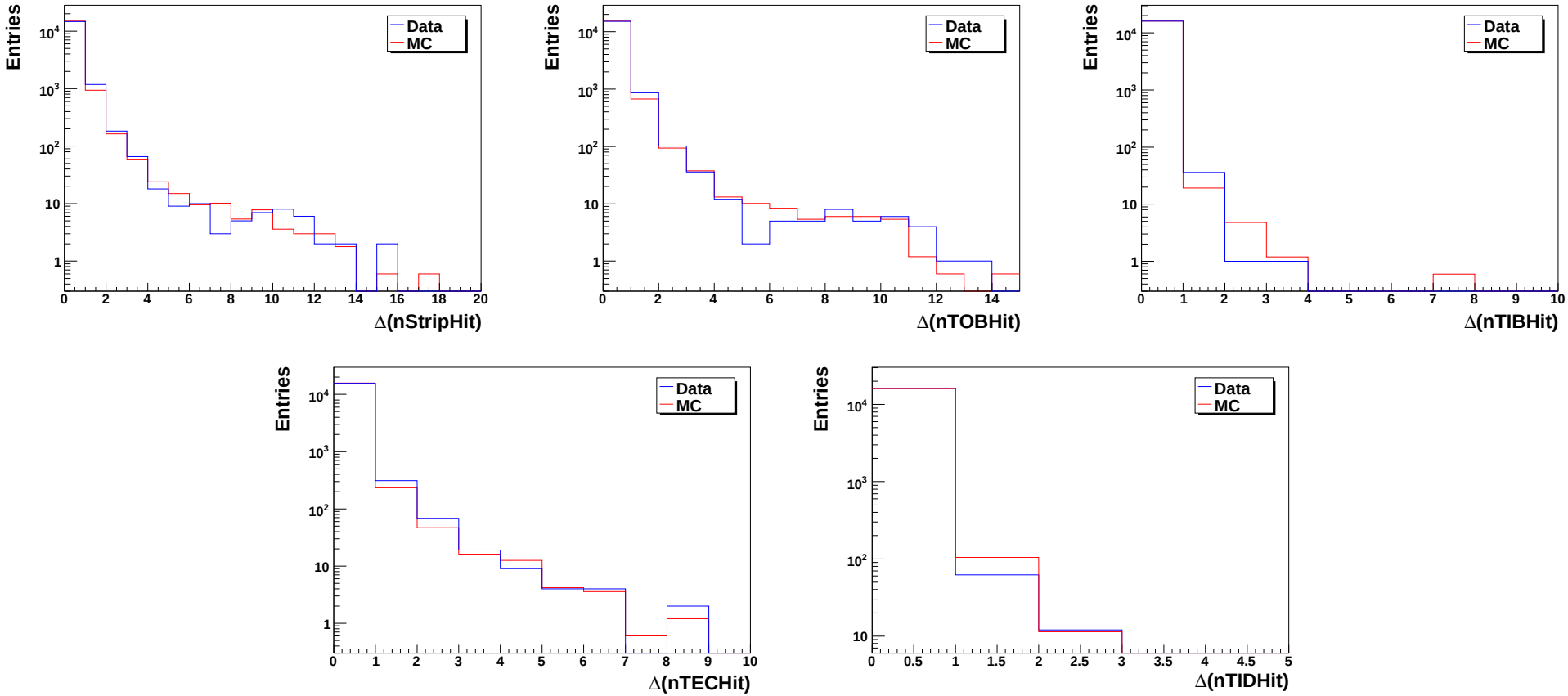
The accuracy of the extrapolation is exactly the same as the one for reconstructed tracks

Simulated and reconstructed state are now consistent

http://indico.cern.ch/materialDisplay.py?contribId=6&materialId=slides&confId=53547

Rejected Hits Population in Tracker

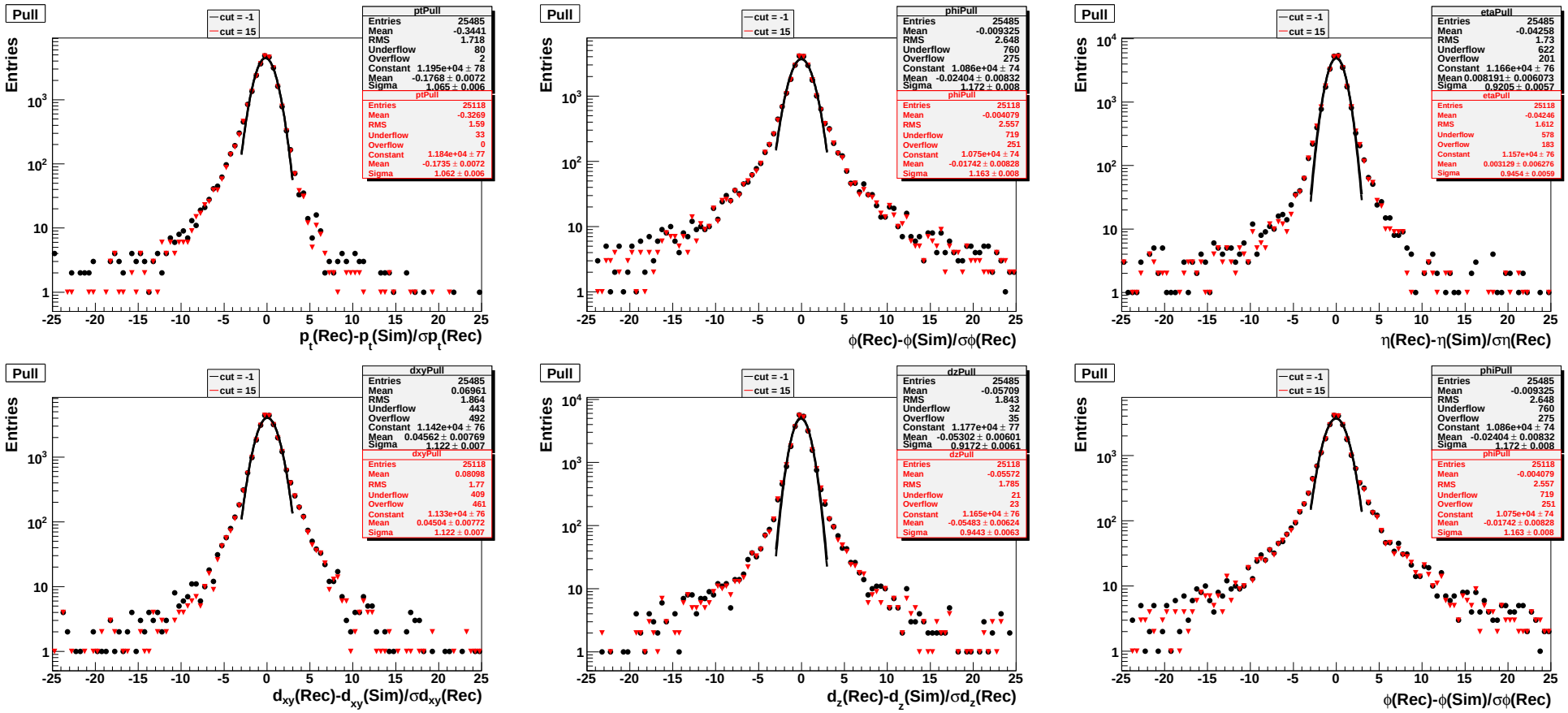
- Compare the data and MC with $\chi^2 > 15$
- Only 1-track events are used, and data is normalized to MC



- No Pixel hits are rejected
- Most rejected hits are located in TOB and TEC

OutliersRejection Impact on Track Parameter Pulls

- SimtoRec: AssociatorByHits with shared hits fraction > 50%



- Slight improvement in cleaning the tails of the pull distributions
  - Smaller impact than the collision data, see Giuseppe Cerati's talk:

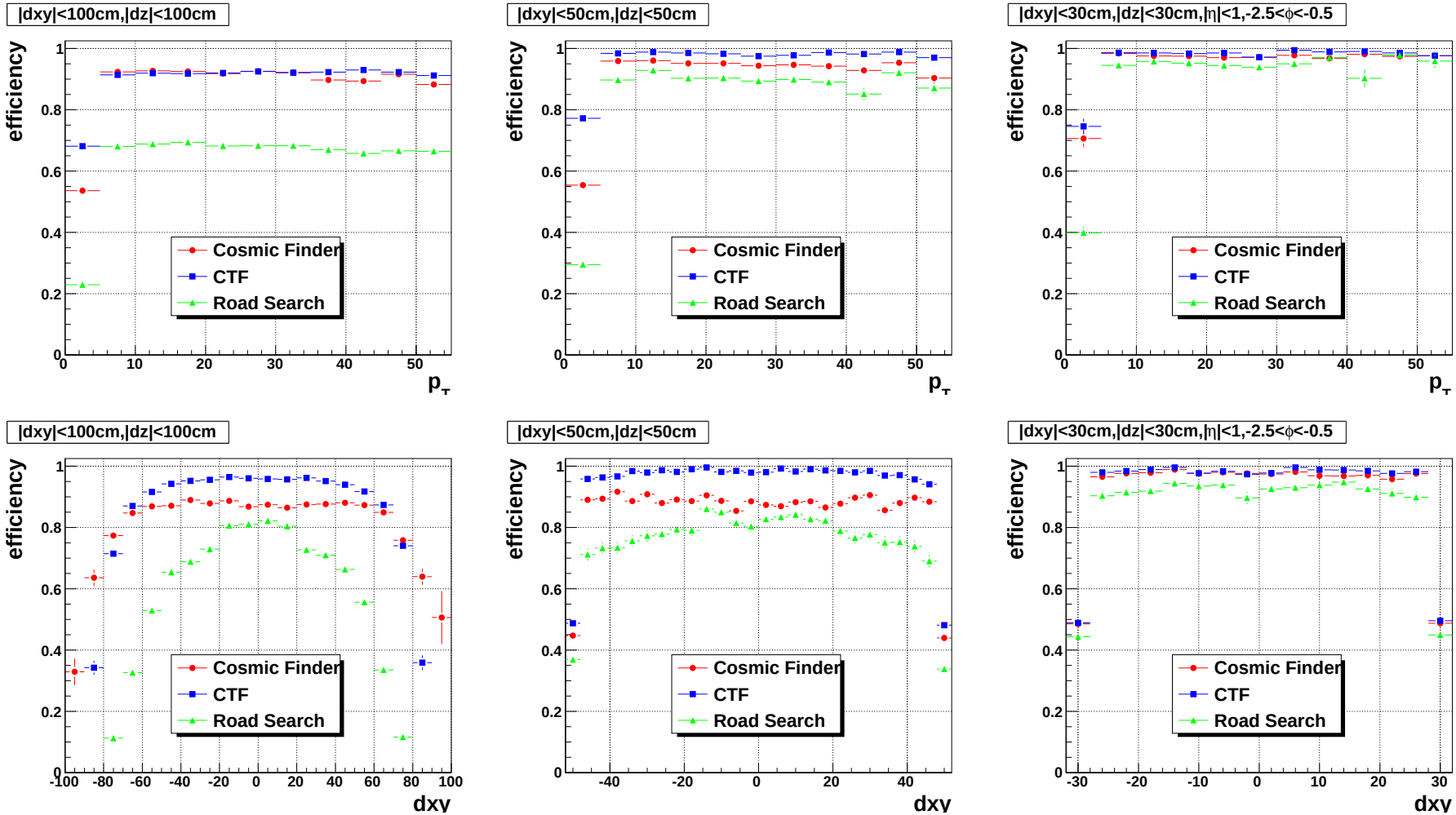
    http://indico.cern.ch/getFile.py/access?contribId=4&sessionId=1&resId=1&materialId=slides&confId=19586
  - The fake hits and Pixel hits are much less in cosmic data

Tracking Performance Using MultiTrackValidator

- Package: CMSSW/Validation/RecoTrack/
- Tracking Efficiency $\epsilon = \frac{N_{\texttt{assoc.}}}{N_{\texttt{simul.}}}$, with 3 cuts on simulated particle
  - nHits> 2, $p_T > 1$ GeV, $|dxy| < 100\texttt{cm}$, $|dz| < 100\texttt{cm}$
  - nHits> 2, $p_T > 1$ GeV, $|dxy| < 50\texttt{cm}$, $|dz| < 50\texttt{cm}$
  - "lhc-like" nHits> 2, $p_T > 1$ GeV, $|dxy| < 30\texttt{cm}$, $|dz| < 30\texttt{cm}$, $|\eta| < 1$, $-2.5 < \phi < -0.5$
- Fake rate $= 1 - \frac{N_{\texttt{assoc.}}}{N_{\texttt{reco.}}}$ with 2 cuts on reconstructed tracks
  - nHits> 2, $p_T > 1$ GeV, $|dxy| < 100\texttt{cm}$, $|dz| < 100\texttt{cm}$
  - "lhc-like" nHits> 2, $p_T > 1$ GeV, $|dxy| < 30\texttt{cm}$, $|dz| < 30\texttt{cm}$, $|\eta| < 1$, $-2.5 < \phi < -0.5$
- MC Associator: AssociatorByHits (shared hits fraction $>$ 50%)
- Twiki Page: https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMultiTrackValidator

Tracking Efficiency

- Common Cuts: nHits> 2, $p_T > 1$ GeV,
  - $|dxy| < 100$cm, $|dz| < 100$cm
  - $|dxy| < 50$cm, $|dz| < 50$cm
  - "lhc-like" cuts (previous page)



- "lhc-like" cuts (right) results agree with Jeremy Andrea's findings

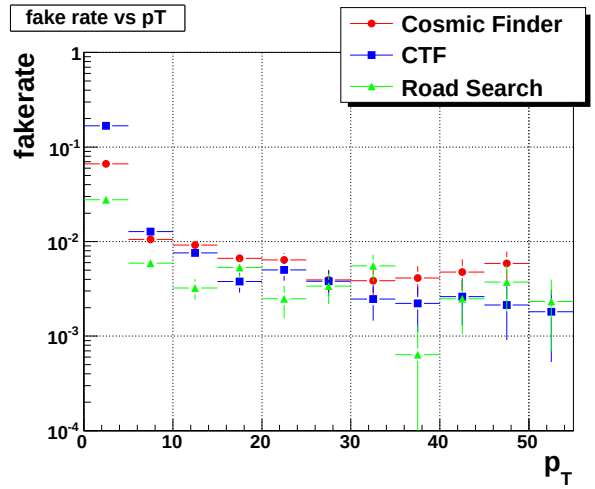
http://indico.cern.ch/materialDisplay.py?contribId=0&materialId=slides&confId=50704

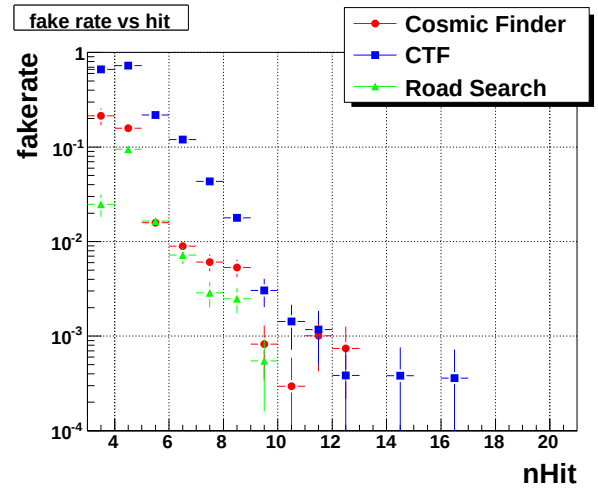
Tracking Fake Rate (I)

- Fake rate is dominated by low p_T and small nRecHit tracks

— No quality cut on tracks

100,000 MC Events



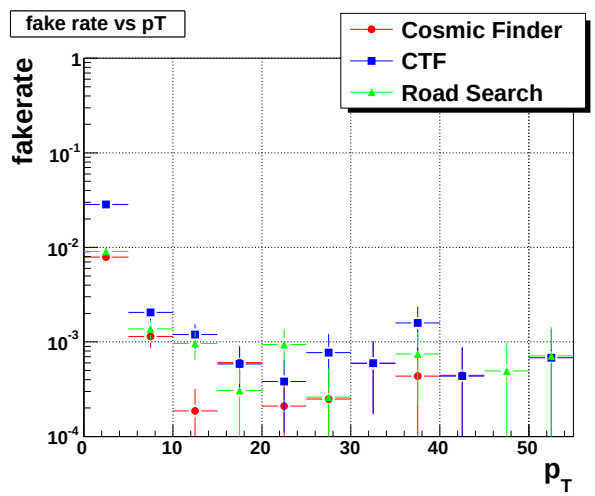


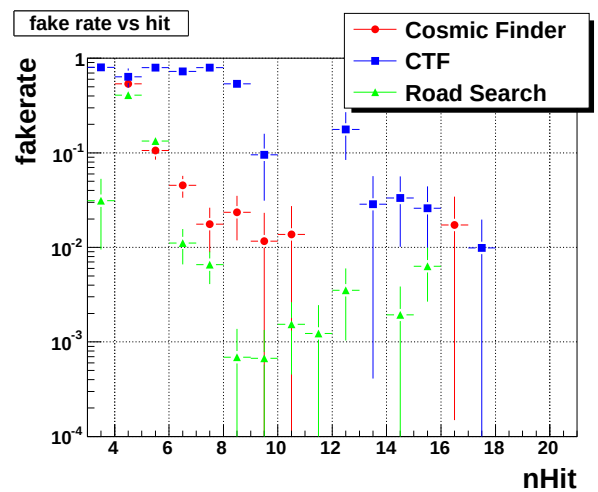
— "lhc-like" cuts on tracks

$|dxy| < 30\mathtt{cm}$, $|dz| < 30\mathtt{cm}$

$|\eta| < 1$, $-2.5 < \phi < -0.5$

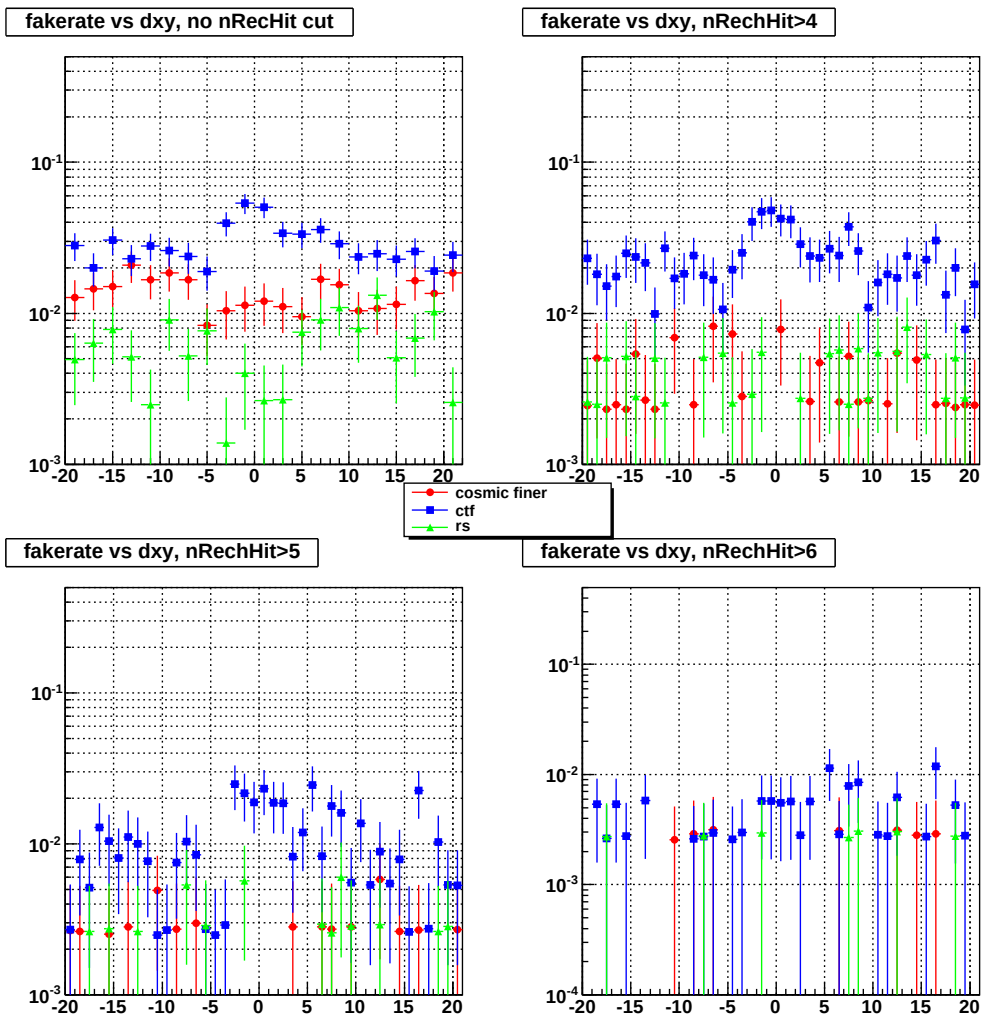
2,000,000 MC Events



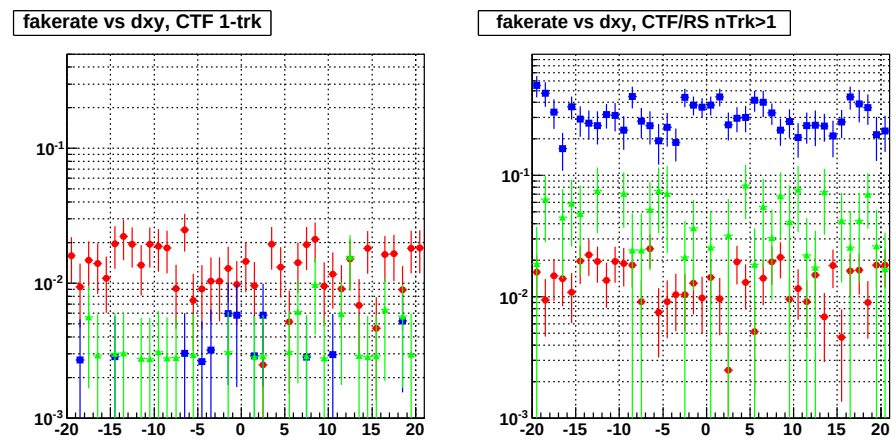


Tracking Fake Rate (II)

- See excess of fake rate $|dxy| < 5\mathtt{cm}$ for CTF
  - The excess goes away with more RecHits



- The CTF fake rate is dominated by multi-tracks events



- Need improvement in the CTF?

The p_T Resolution

- The p_T resolution is defined as $\frac{p_T(\texttt{rec})-p_T(\texttt{sim})}{p_T(\texttt{rec})}$

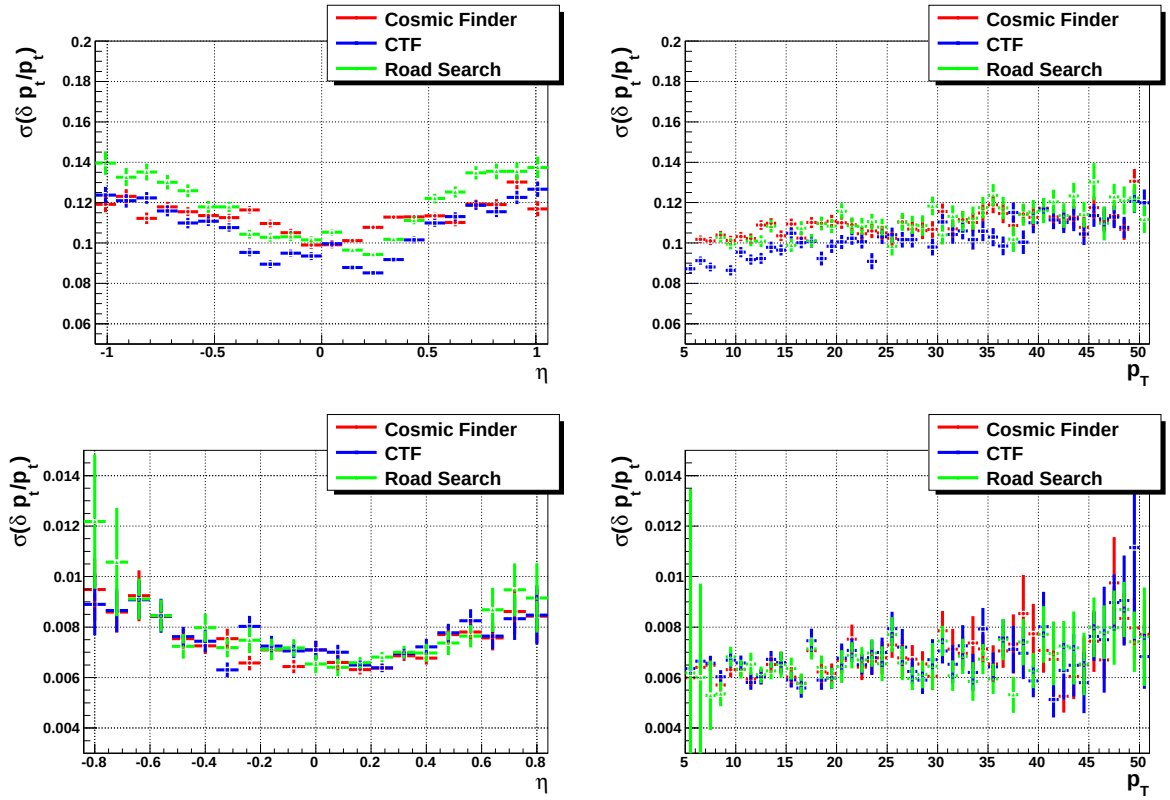
  - No quality cut on tracks

  - Alignment cuts on tracks

nRecHit> 9, nPXBHit> 1

$|dz| < 30cm$, $p_T > 5$GeV

Resolution reduces 10X

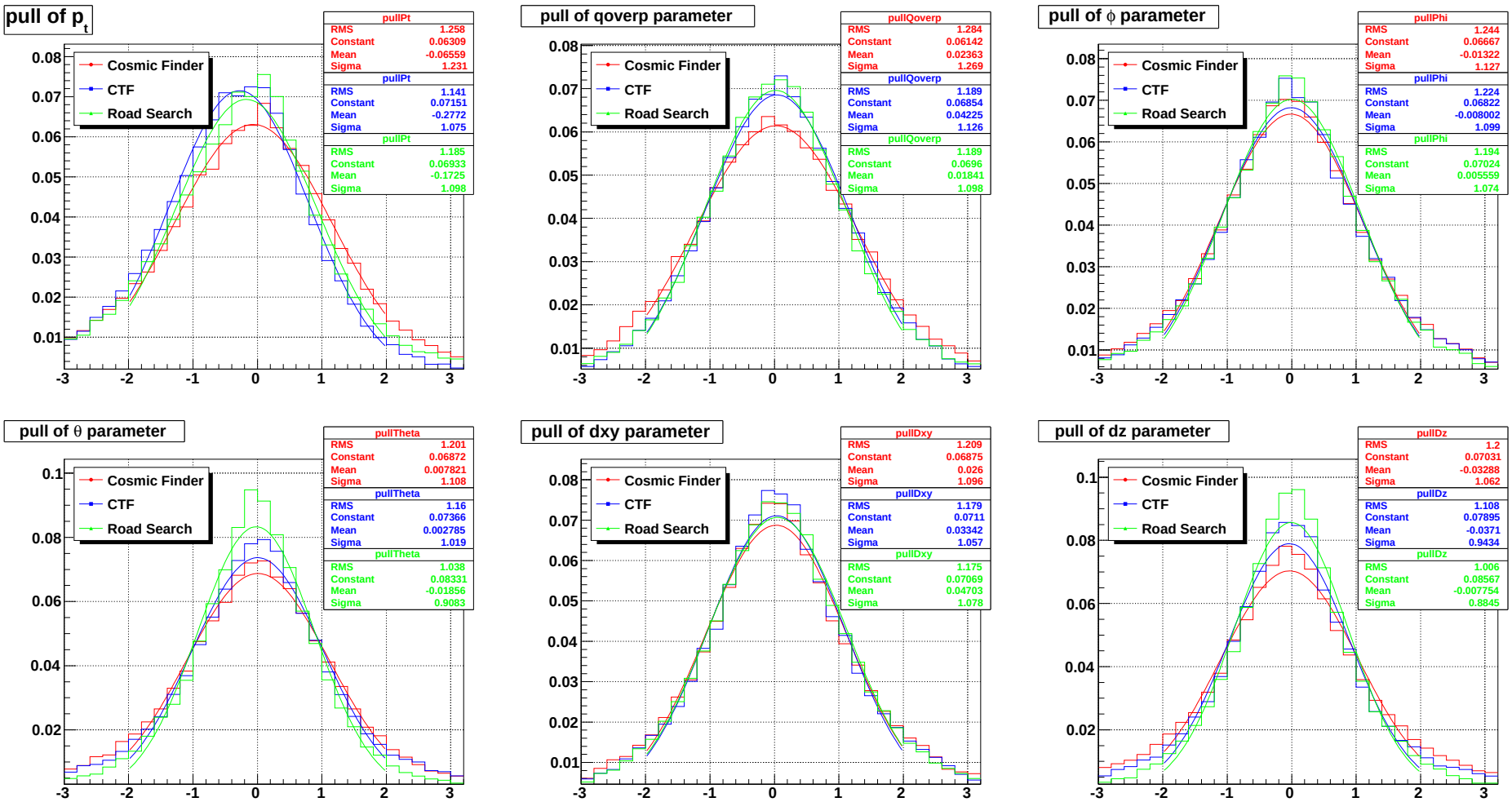


- Track parameter resolutions agree with Nhan Tran's findings:

http://indico.cern.ch/materialDisplay.py?contribId=9&materialId=slides&confId=54735

Track Parameter Pull Distribution

- No quality cut on the reconstructed tracks



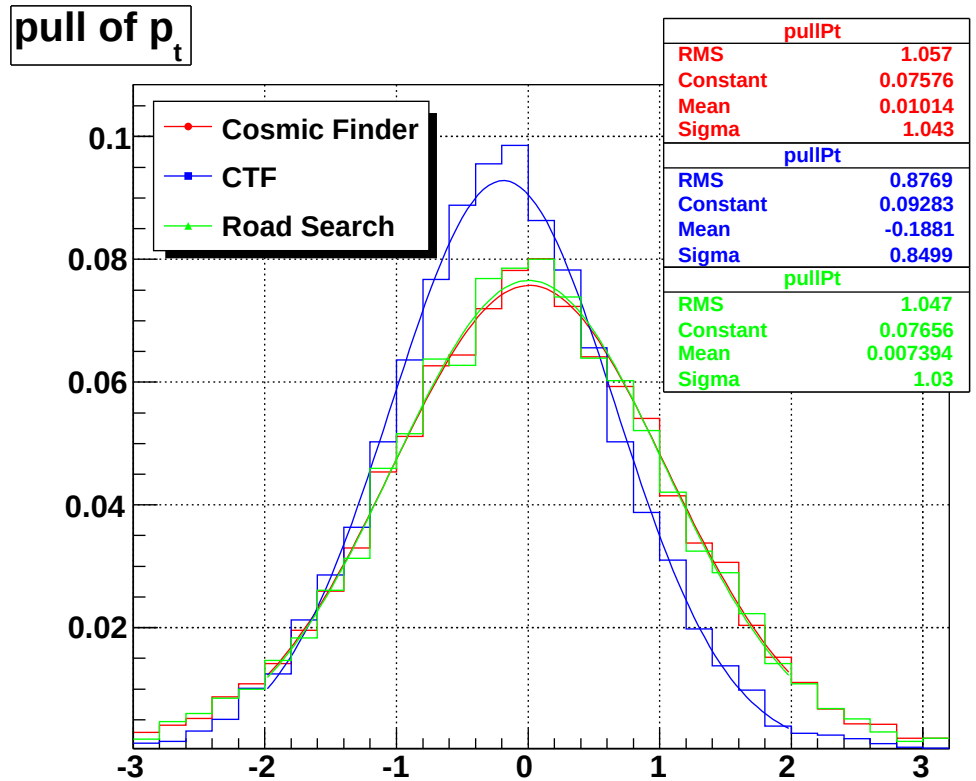
- See large shift in the mean of p_T pull for CTF (!), but not in Q/p ?

The CTF p_T Pull Issue (I)

- Use the "Standard Alignment Cut" as shown by Nhan

http://indico.cern.ch/materialDisplay.py?contribId=9&materialId=slides&confId=54735

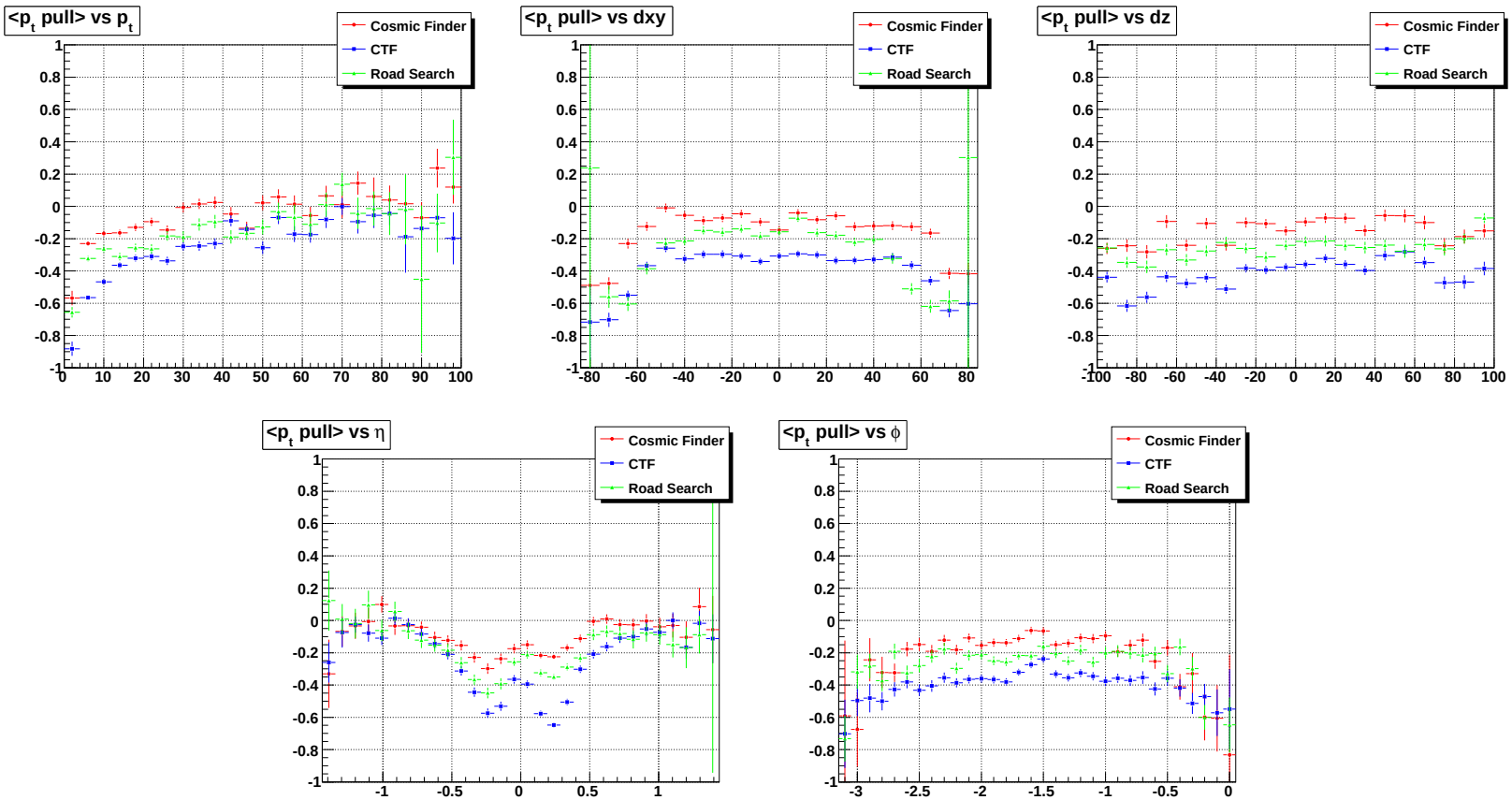
- nRecHit> 9
- nPXBHit> 1
- $|dz| < 30\mathrm{cm}$
- $p_T > 5$ GeV



- The shift in $<p_T$ pull$>$ persists for CTF (not COS/RS !) tracks
- The p_T pull agrees with Nhan's finding with Cosmic Finder tracks
- Thanks to Nhan for the useful discussion

The CTF p_T Pull Issue (II)

- Plot the mean of p_T pull versus track parameters



- There is no quality cut on the tracks
- CTF tracks have larger shift compared to Cosmic Finder and RS

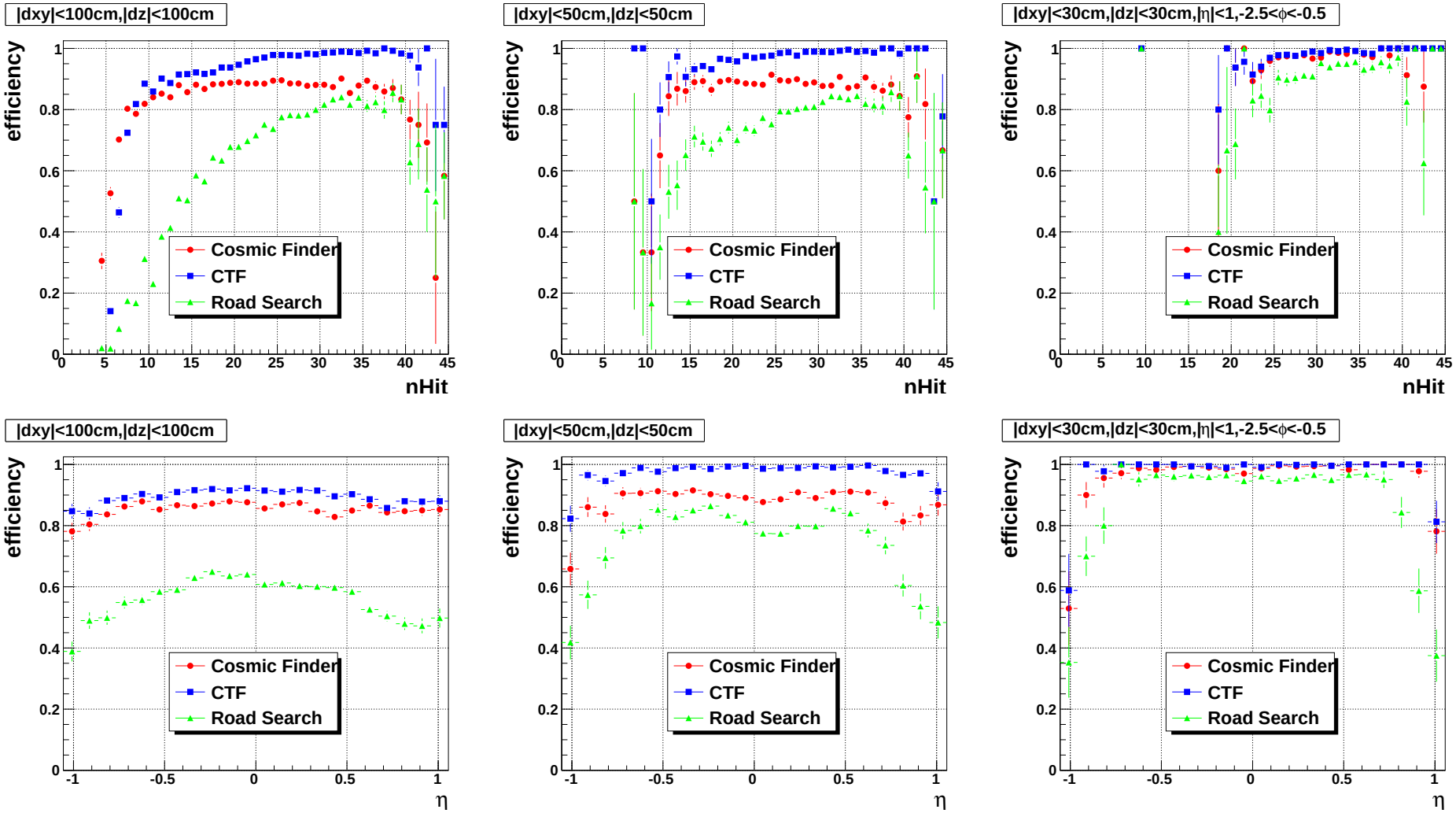
Summary and Plans

- OutliersRejection shows improvement on the cosmic Data
  - Rejection rate is 1.8% with $\chi^2 > 15$
  - Rejected hits mainly locate in TOB/TEC
  - The tails of track paramter pull distributions improve slightly
- Tracking performance study using MultiTrackValidator
  - Track efficiency ~98.5% for "lhc-like" tracks with p_T<50 GeV
  - Track fake rate ~ 0.1% for "lhc-like" tracks with p_T<50 GeV
  - The p_T resolution <1% for $p_T < 50$GeV and $|\eta| < 1.0$
  - See large shift (> 0.2!) in <p_T pull> distribution for CTF tracks
- Plans
  - Follow up on the asymmetric p_T pull distribution for CTF tracks

BACKUP SLIDES

Tracking Efficiency

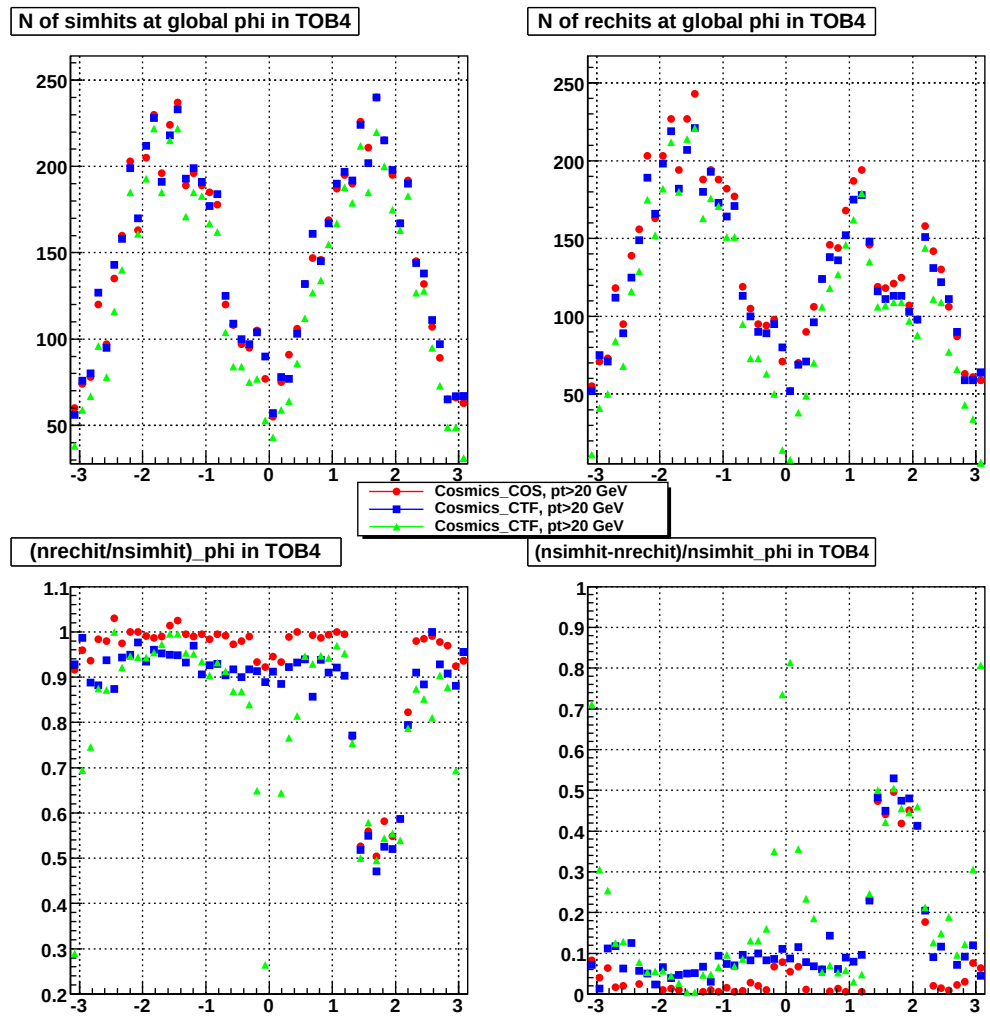
- $\epsilon = \frac{N_{\texttt{assoc.}}}{N_{\texttt{simul.}}}$ AssociatorByHits (shared hits fraction $>$ 50%)



Hits ϕ top-bottom asymmetry

- There are some dead modules in TOB4 top ($\phi = \pi/2$).

– $p_T > 20$ GeV



- The dead modules are simulated in MC, but masked in reconstruction.