Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

CMS FastSim for LHC Run II

Lukas Vanelderen
University of Hamburg
on behalf of the CMS collaboration

LPCC workshop Geneva
March 19, 2014

Outline

- CMS FastSim crash course
- Performance overview
- Recent/ongoing/future developments

CMS FastSim crash course

simulated hits from energy deposits in active and dead material

- ideal cylindrical geometry
- material concentrated in infinitely thin layers

reconstructed hits with FullSim “digitizers"

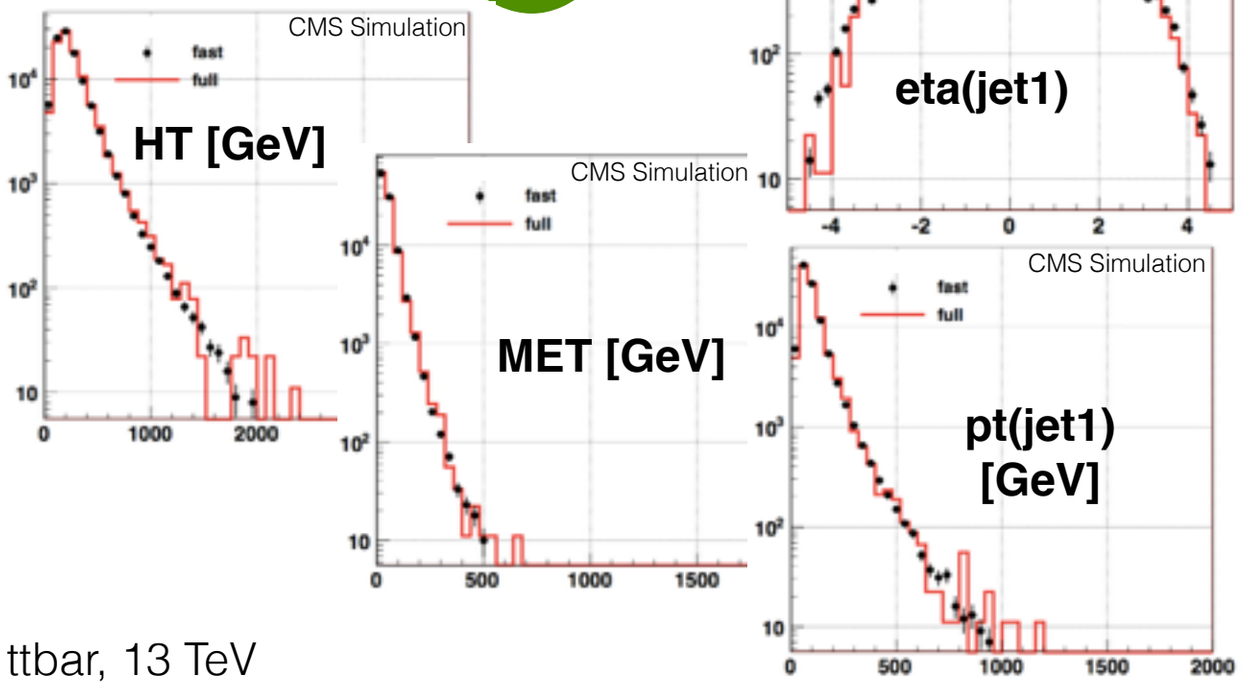
- => e.g. full emulation of electronics (noise)
- exception: simple spatial smearing for tracker hits

full **event reconstruction**

- exception: track finding based on truth information

~5s per ttbar event
(8 TeV, summer 2012 pu profile)

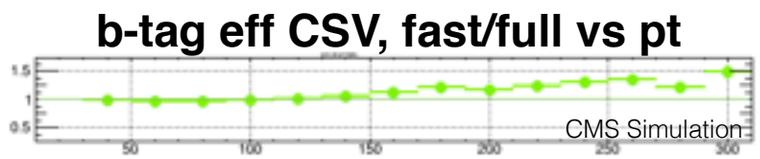
fast vs **full** overview

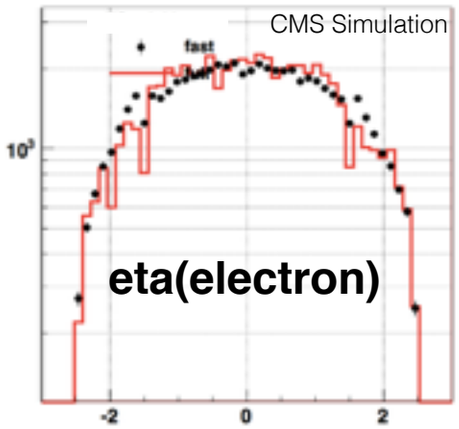


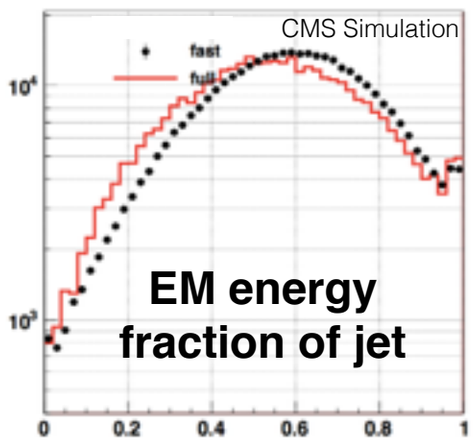
ttbar, 13 TeV

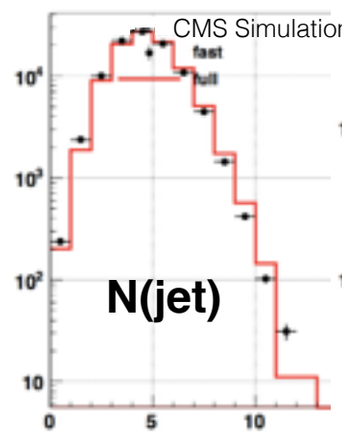
- modelling of hadronic variables is mostly excellent
- e.g. HT, MET, jet eta and pt

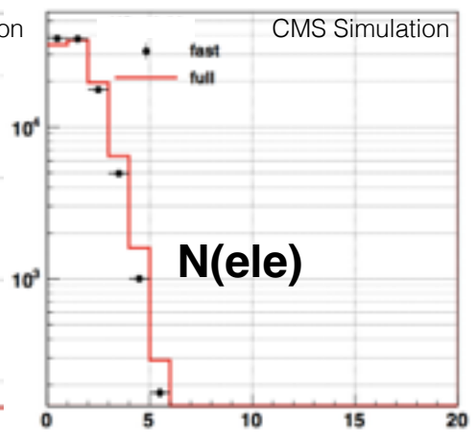
fast vs **full** overview











ttbar, 13 TeV

not so well by FastSim:

- upper left: ratio of
  - (btag efficiency in fastsim vs jet pt) /
  - (btag efficiency in fullsim vs jet pt)
- different jet energy fractions (EM, neutral hadronic, charged hadronic)
- high tail of jet multiplicities
- lepton multiplicities
- impact of detector structure on e.g. electron efficiency not perfectly well described.(see top plot)

3 ways to improve

core developments:
interaction with material,
hit conversion:
simulated -> reconstructed

generally applicable

responsibility of fastsim group

low-level bias corrections:
correcting low level reconstruction

analysis-level bias corrections:
for analysis-level reconstruction

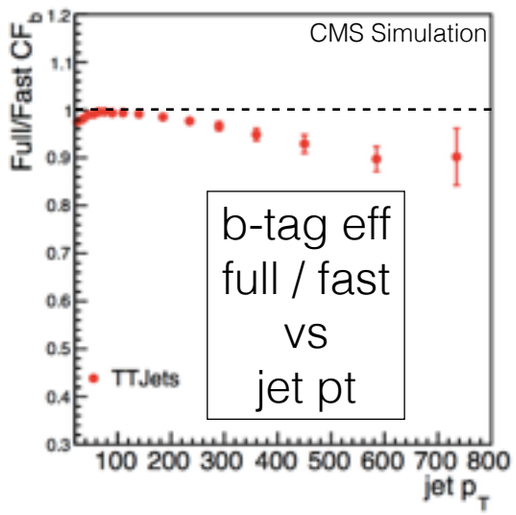
~analysis specific

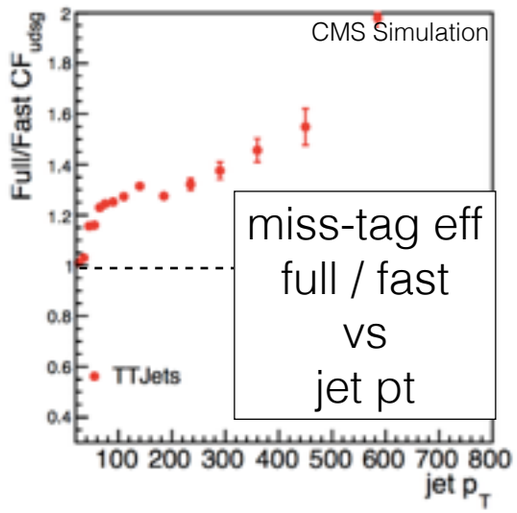
responsibility of analysis groups

tracker developments

analysis-level bias corrections

Run I situation





improving this behaviour is one of the 2014 priorities for CMS FastSim

plots show ratio of

(b-tag efficiency vs jet pt in **fullsim**) /

(b-tag efficiency vs jet pt in **fastsim**)

left: for b-jets

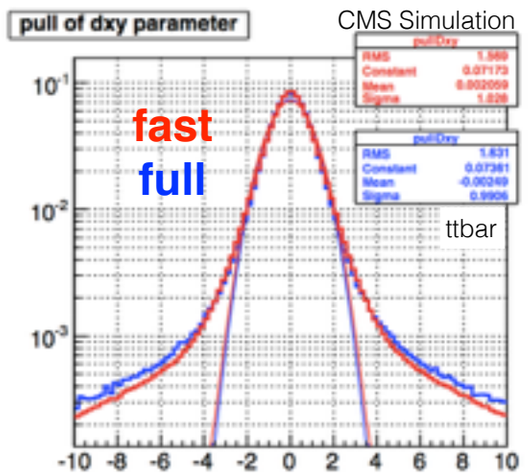
right: for light jets

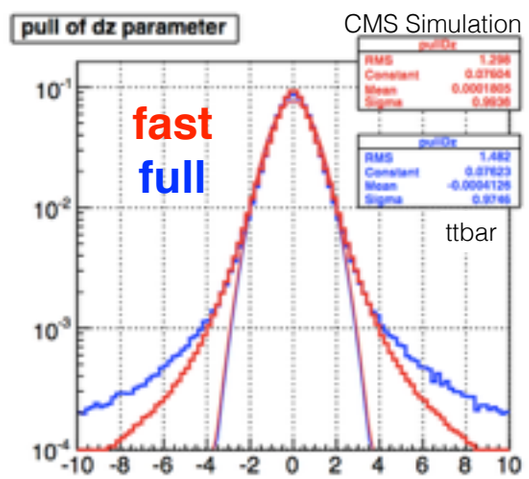
these ratios are used to correct biasses in fastsim

b-tagger used is CSVL : Combined Secondary Vertex with Loose working point

core developments I

- does gaussian hit smearing describe the tails of track vertex resolution?
- study **planned** for 2014
- expect very limited impact: vertex pull described well up to 4 sigma



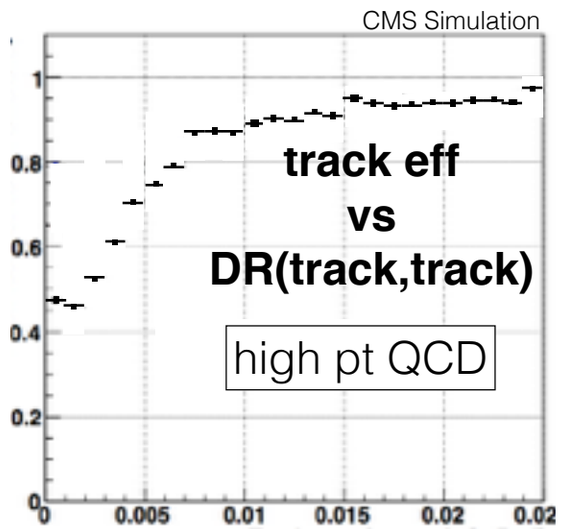


top plot: pull of dxy: signed impact parameter

bottom plot: pull of dz: distance between primary vertex and z-vertex

core developments II

- in high pt jets, tracker hits from different tracks often merge
- CMS fastsim does not model merging of tracker hits
- introduction of tracker hit merging **planned** for 2014

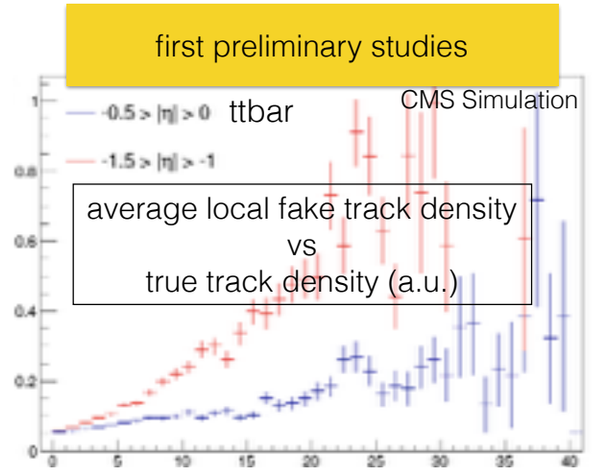


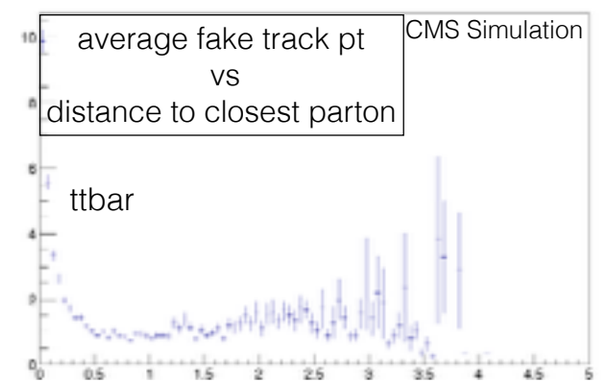
plot shows typical effect of tracker hit merging:

in high pt jets, track efficiency drops when tracks are too close

low level bias corrections

- mimic track finding effects
- **ongoing**: shoot well-parameterised fake tracks
- **idea:** combine this with a parameterisation of track inefficiencies





- FastSim bases track finding on truth information.
- As a consequence, track finding effects are not modelled in FastSim.
- Here: first attempts to parameterise these effects
  - average local fake track density is correlated with local true track density
  - fake track pt is correlated with angular distance between track and closest parton
- New idea: fakes and track inefficiencies often occur simultaneously.
  - => a simultaneous parameterisation of fakes and track inefficiencies might be desirable.
- Definition of local track density: event is chopped up in small rectangular bins in eta-phi space and the number of tracks in the bin are counted

core developments III

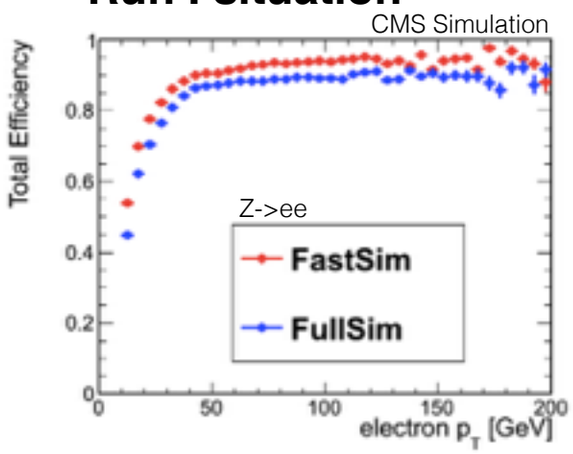
- parameterisation as on previous slides can be extremely difficult: not all fakes / inefficiencies can be parameterised
- **idea**: apply full tracking / full seeding for a (very) tight selection of generator particles
- selection based on e.g. deltaR(track,track) track pt, distance between sim hits,….

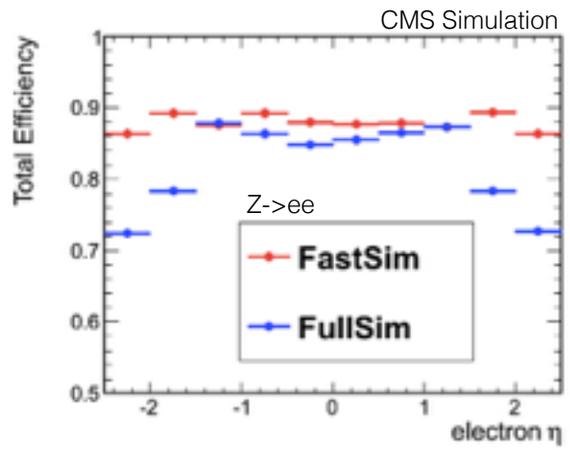
the selection criteria for tracks to undergo full reconstruction should be very tight, since track finding requires a lot of CPU

ECAL developments

analysis-level bias corrections for e and mu efficiency

Run I situation



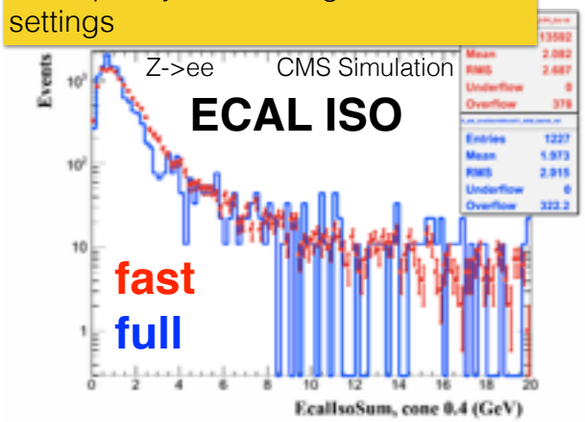


- bias in electron efficiency depends on electron pt and eta
- electron efficiency correction factors were derived as a function of electron pt and eta

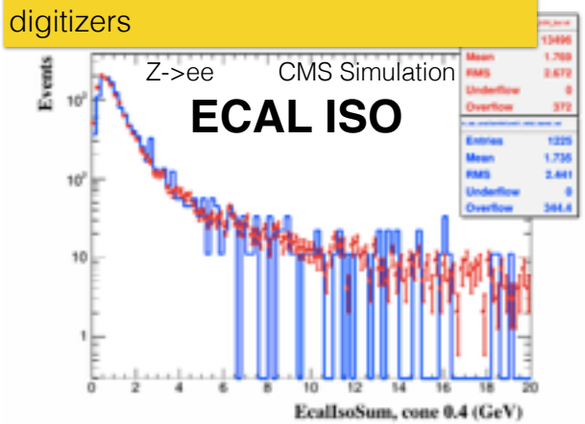
core developments I

- **done:** use of FullSim digitizers in FastSim
- => new settings of electronics settings propagate to FastSim
- => corresponding noise models propagate to FastSim
- => easier maintenance

discrepancy after change in readout settings



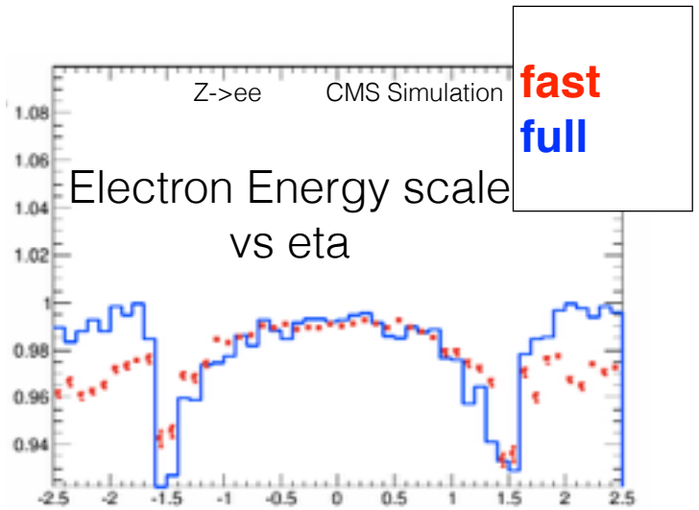
fixed after introduction of full sim digitizers



ECAL ISO:

energy deposited in ECAL in cone with radius R=0.4 around electron

core developments II

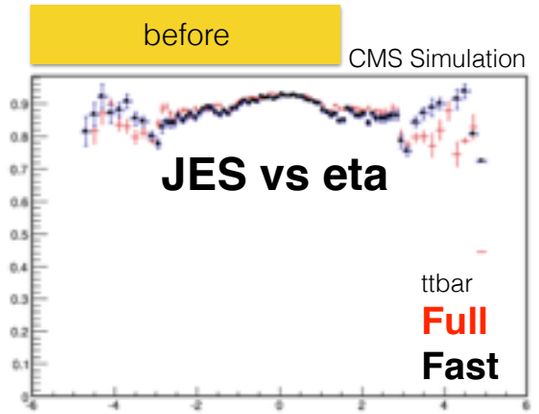


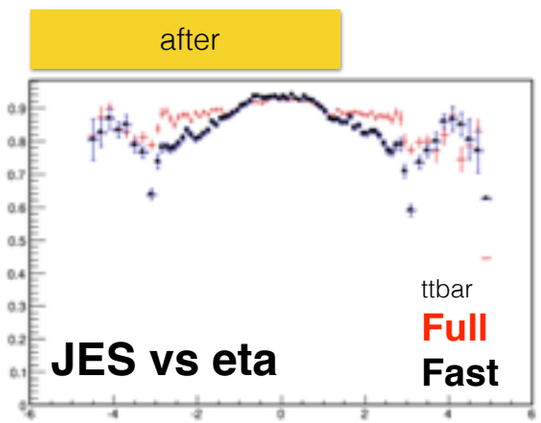
- impact of inaccurate modelling of EM showers increases with increasing precision of reconstruction
- time to review/update the shower model

HCAL developments

core developments I

- **done:** use of FullSim digitizers in FastSim
- **done:** retune required after introduction digitisers
- First retune not fully satisfactory. Further fine tuning ongoing see e.g. next slide

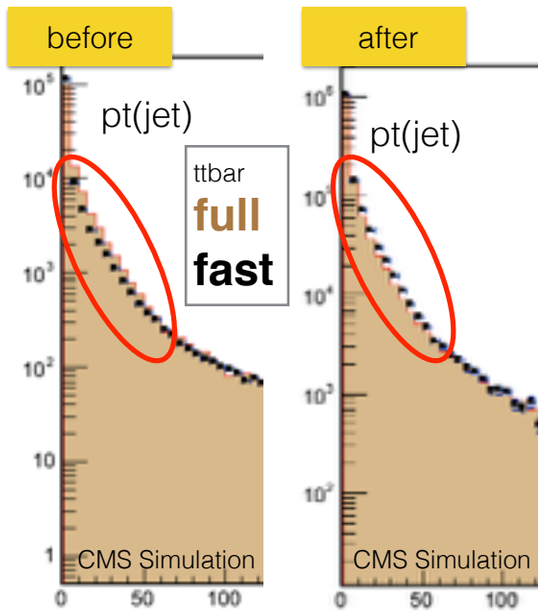




JES: Jet Energy Scale = jet reco pt / jet true pt

low-level bias corrections

- neutral hadronic component of jets underestimated
- **done:** introduced random parameterised hadronic clusters
- more fundamental treatment **planned**: fine tuning of hadronic shower model



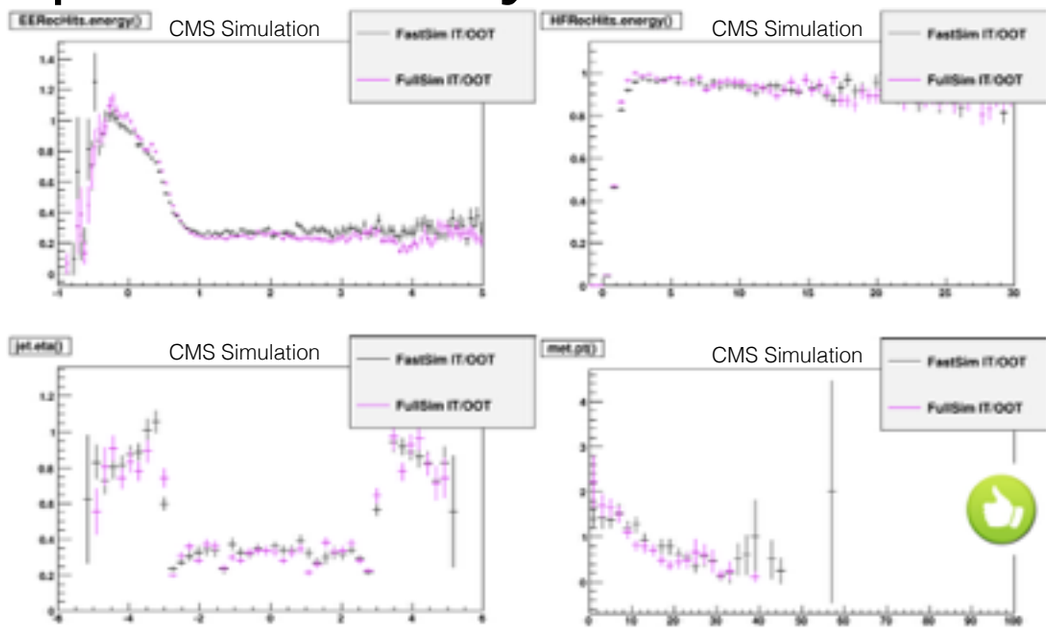
However: these kind of low-level bias corrections are often suboptimal:
they are often sensitive to small changes in reconstruction algorithms,
and while they fix the modelling of one variable,
they might brake the modelling of other variables.
Still, this fix is know to work properly for most variables of interest.

Pile up developments

25ns, more pile up
=> new challenges

- importance of out of time pile up will increases
  => cpu and memory intensive
- **old mechanism**:
  mixing at genlevel, only in time pu
- **new mechanism**:
  - mixing at simlevel via full sim digitizers
  - exception: tracks mixed at reco level,
    in time pu only

preliminary validation



Relation between In Time (IT) and Out Of Time (OOT) pu well modeled in FastSim

Conclusions

- CMS FastSim performance for Run II mostly ranges between acceptable and good

- Many recent/ongoing developments
  core, low-level, analysis-level

- CMS FastSim can even do MUCH better

  - studies towards improved high pt b-tagging
    definitely on the 2014 menu

  - hopefully we can put hadronic and EM showers
    on the menu too

Many thanks to the CMS FastSim developers and validators!