Evgeny Andronov

Analysis wagon for the pt-n correlations analysis

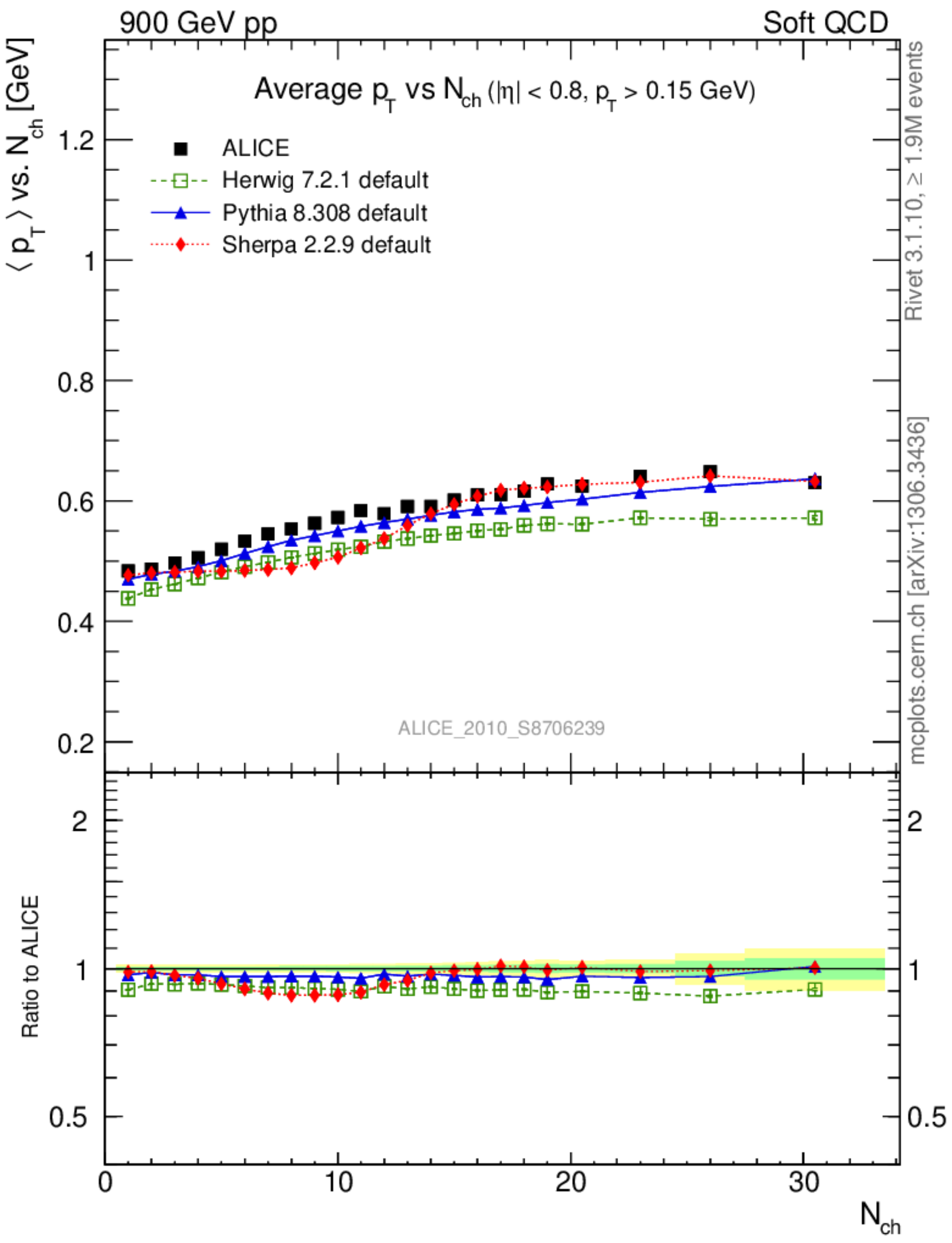
MPD CROSS-PWG, 8 OCTOBER 2024

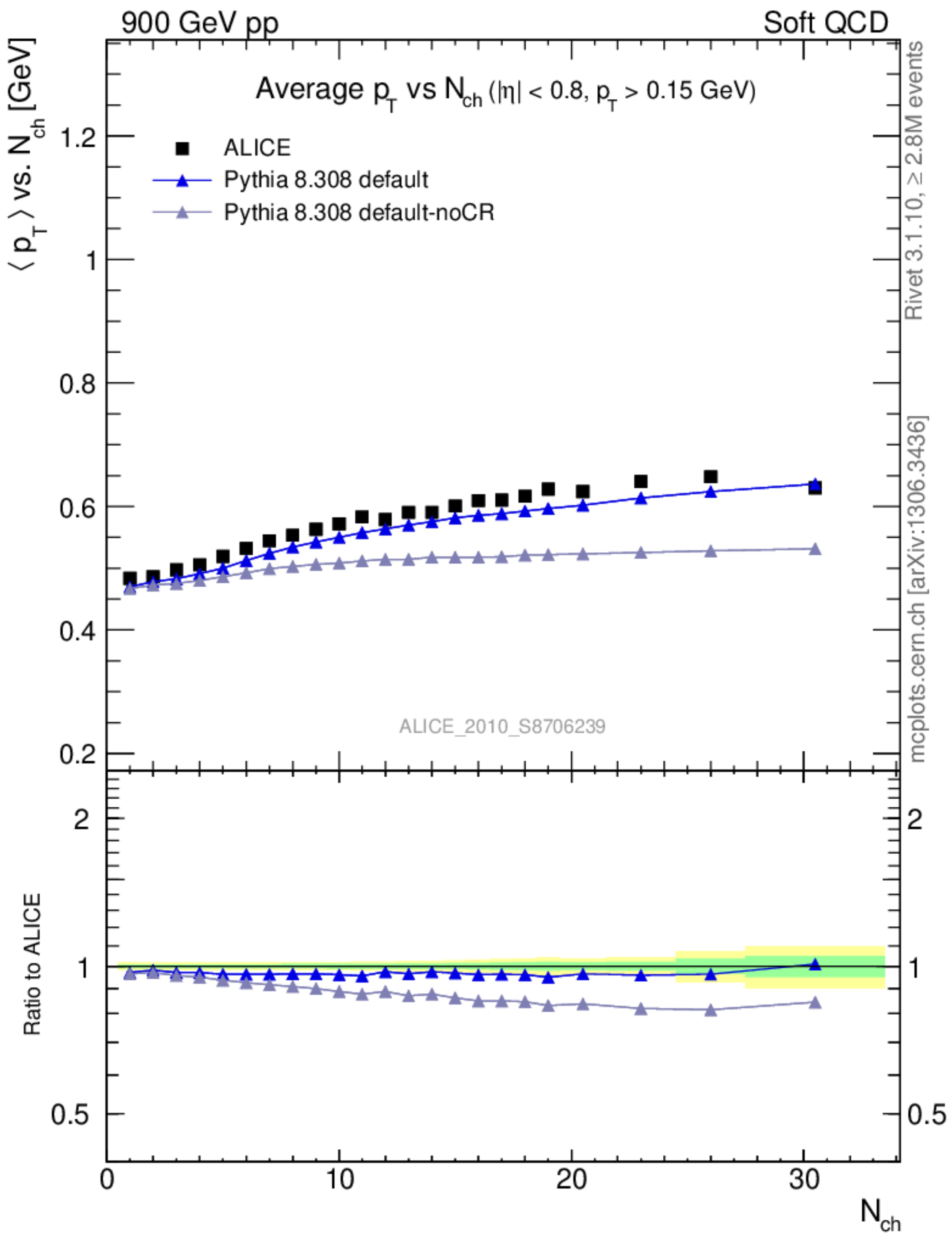
Goals of the project

- to prepare analyses wagon for
  - unidentified pT spectrum as a function of multiplicity (N) + moments of pT spectrum vs. N (i.e. pT vs N correlation function etc.)
  - analogous correlations but with observables taken in separated subevents (including fluctuation studies in terms of strongly intensive observables and pT cumulants)
  - studies of the mentioned observables for different centrality classes (i.e. the wagon should follow the pCentr wagon)

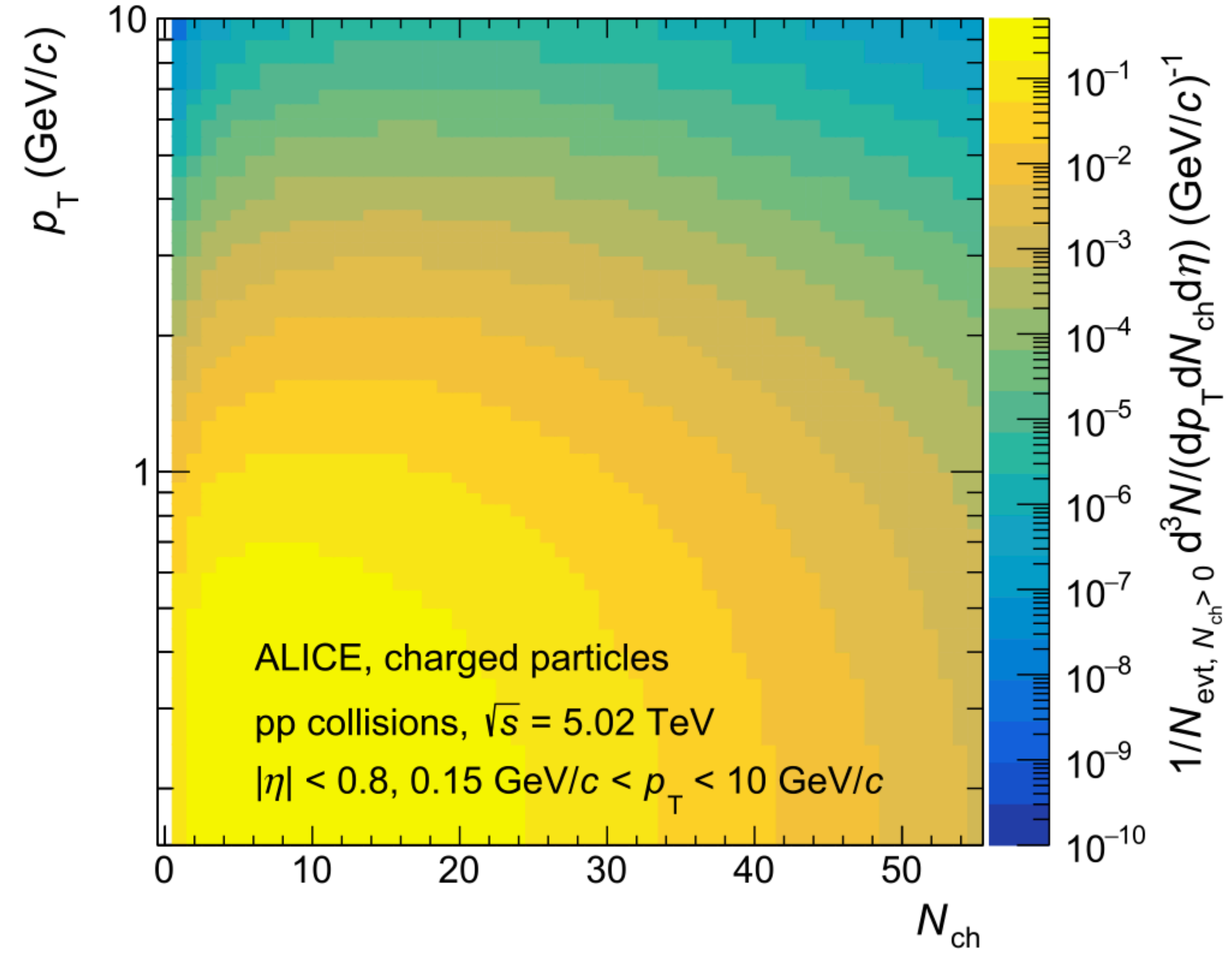
pT vs N correlations

- long history of measurements for broad energy range and for different colliding systems
- help to constrain models (famous example is introduction of color reconnection to the PYTHIA model)





pT vs N correlations



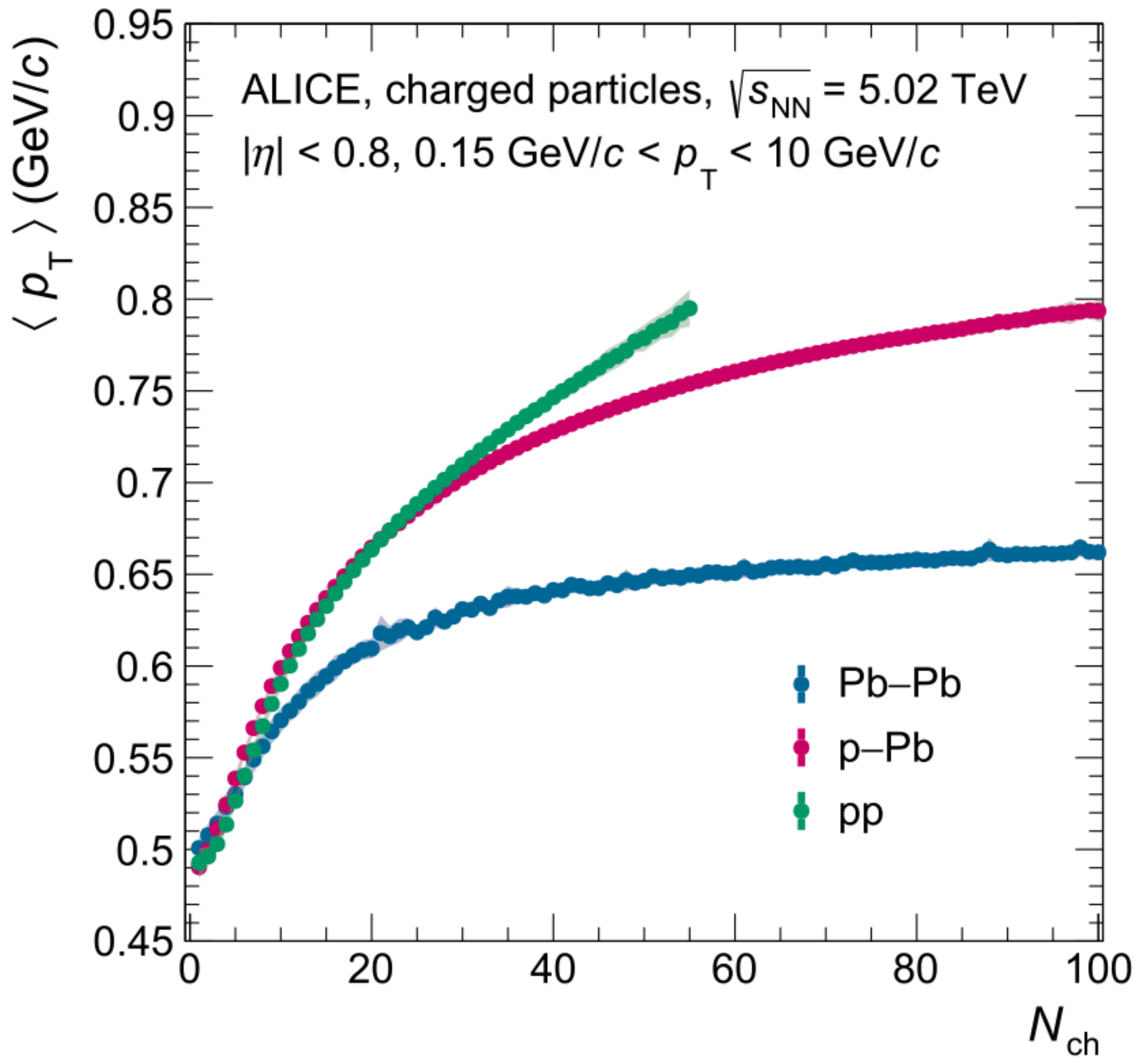
- only recently HI experiments started to publish full pT spectrum as a function of multiplicity
- therefore, one can have a look not only on the mean pT but on the higher moments as well

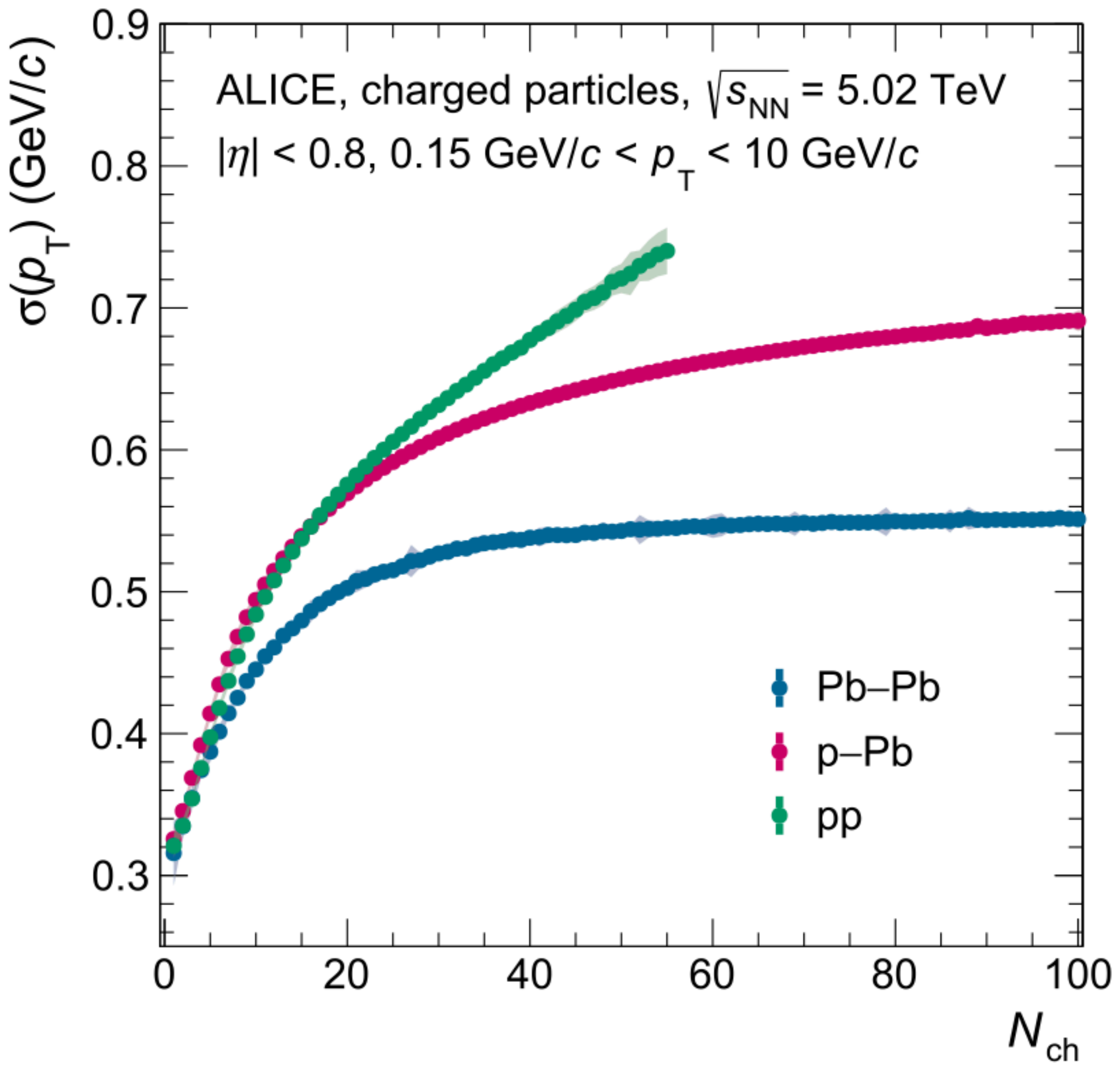
Data from «Multiplicity dependence of charged-particle production

in pp, p–Pb, Xe–Xe and Pb–Pb collisions at the LHC»

by ALICE Coll., Phys. Lett. B845,138110 (2023)

see also discussion https://inspirehep.net/literature/2767694 G.Biro et al.





Wagon

The structure and basic principles of the wagon were adopted from the available wagons:

```
[evandron@ncx104 fluctPt]$ ls -1
CMakeLists.txt
macros
MpdFluctPt.cxx
MpdFluctPt.h
MpdFluctPtLinkDef.h
MpdFluctPtParams.cxx
MpdFluctPtParams.h
```

Analysis code structure is standard:

```
void RunAnalyses(int nEvents = -1){
  gSystem->Load("libZdc.so");
  gSystem->Load("libMpdPhysics.so");


  MpdAnalysisManager man("ManagerAnal", nEvents);
  man.InputFileList("listShort.txt");
  man.ReadBranches("*");
  //man.ReadBranches("MCTrack,TpcKalmanTrack,ZdcDigi,Vertex,MPDEvent,TOFMatching");
  man.SetOutput("histos.root");

  MpdCentralityAll pCentr("pCentr", "pCentr");
  man.AddTask(&pCentr);

  MpdFluctPt taskFluctPt("pFluctPt", "pFluctPt");
  man.AddTask(&taskFluctPt);

  man.Process();
}
```

Datasets

The code was tested on the Request 25 production: Bi+Bi@9.2 AGeV (UrQMD)
The Request 26 production (Bi+Bi@9.2 AGeV (DCM–QGSM–SMM)) will be used for cross-validation

Event cuts

MpdAnalysisEvent contains info both on **MC** and **Reconstructed** versions of a given event

At the moment no effects concerning wrongful event selection were studied, i.e. loops over reconstructed and over pure MC tracks were ran on the same events that passed the following cuts:

- event should have at least two pure MC tracks (with **GetMotherId() =** -1) within experimental TPC acceptance
- event should have reconstructed vertex
- |VertexZ|<50cm

How to choose vertex position?

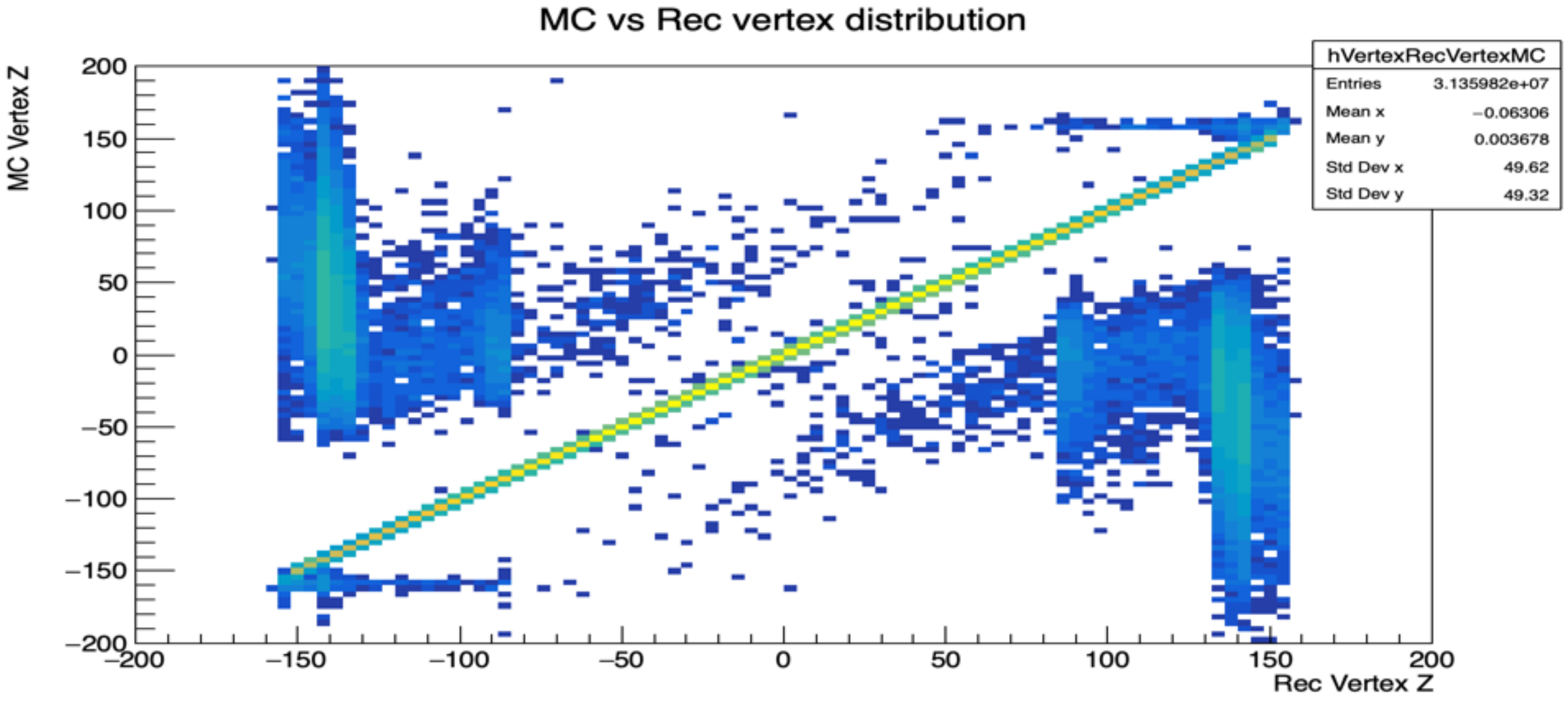
mMCHeader = event.fMCEventHeader;

mhMCVertex->Fill(mMCHeader->GetZ());

MpdVertex *vertex = (MpdVertex *)event.fVertex->First();

vertex->Position(mPrimaryVertex);

mPrimaryVertex.Z()



+/-50cm seems to be reasonable

Naming of tracks

pure MC track - primary, if

its motherid=-1

or

its ancestors are shortliving resonances (one has to check for all cascades of short-lived resonances that decay via EM or strong interaction, products of these decays should be treated as primary tracks)

pure MC track - good primary, if

it is ‘primary’ and it passed MC track cuts

Reconstructed track - good, if ← proxy for experimentally measured tracks

it passed Rec track cuts

Reconstructed track - good selected, if

it is ‘good’ and its matched pure MC track is good primary

Naming of tracks

pure MC track - primary, if

its motherid=-1

or

its ancestors are shortliving resonances (one has to check for all cascades of short-lived resonances that decay via EM or strong interaction, products of these decays should be treated as primary tracks)

pure MC track - good primary, if

it is ‘primary’ and it passed MC track cuts

Comparing RecTrack good и RecTrack good selected - one can estimate contamination

Reconstructed track - good, if

it passed Rec track cuts

Reconstructed track - good selected, if

it is ‘good’ and its matched pure MC track is good primary

Comparing RecTrack good selected и MC Track good primary - one can see effects of resolution and efficiency

Pure MC track cuts (mMCTracks = event.fMCTrack;)

- **kinematic cuts**
  - |eta|<0.8
  - 0.15<pT<2.0 GeV/c
- **distance of closest approach (to which vertex?)**
  - if(abs(mctrack->GetStartX() - mMCHeader->GetX()) > mParams.mDcaCut) return false;
  - if(abs(mctrack->GetStartY() - mMCHeader->GetY()) > mParams.mDcaCut) return false;
  - if(abs(mctrack->GetStartZ() - mMCHeader->GetZ()) > mParams.mDcaCut) return false;
  - mDcaCut =1 cm
- **electric charge = +/- 1** (it is strange that MpdMCTrack does not have GetCharge method and one has to play with pdgid)
- **particle species cut**
  - accept only pions, kaons, protons, muons and electrons

Rec track cuts (mMpdGlobalTracks = event.fMPDEvent->GetGlobalTracks();)

- **kinematic cuts**
  - |eta|<0.8
  - 0.15<pT<2.0 GeV/c
- **distance of closest approach (clearly to the reconstructed primary vertex)**
  - if (fabs(mpdtrack->GetDCAX()) > mParams.mDcaCut) return false;
  - if (fabs(mpdtrack->GetDCAY()) > mParams.mDcaCut) return false;
  - if (fabs(mpdtrack->GetDCAZ()) > mParams.mDcaCut) return false;
  - mDcaCut =1 cm
- **minimal number of TPC hits**
  - if (mpdtrack->GetNofHits() < mParams.mNofHitsCut) return false;
  - mNofHitsCut = 16

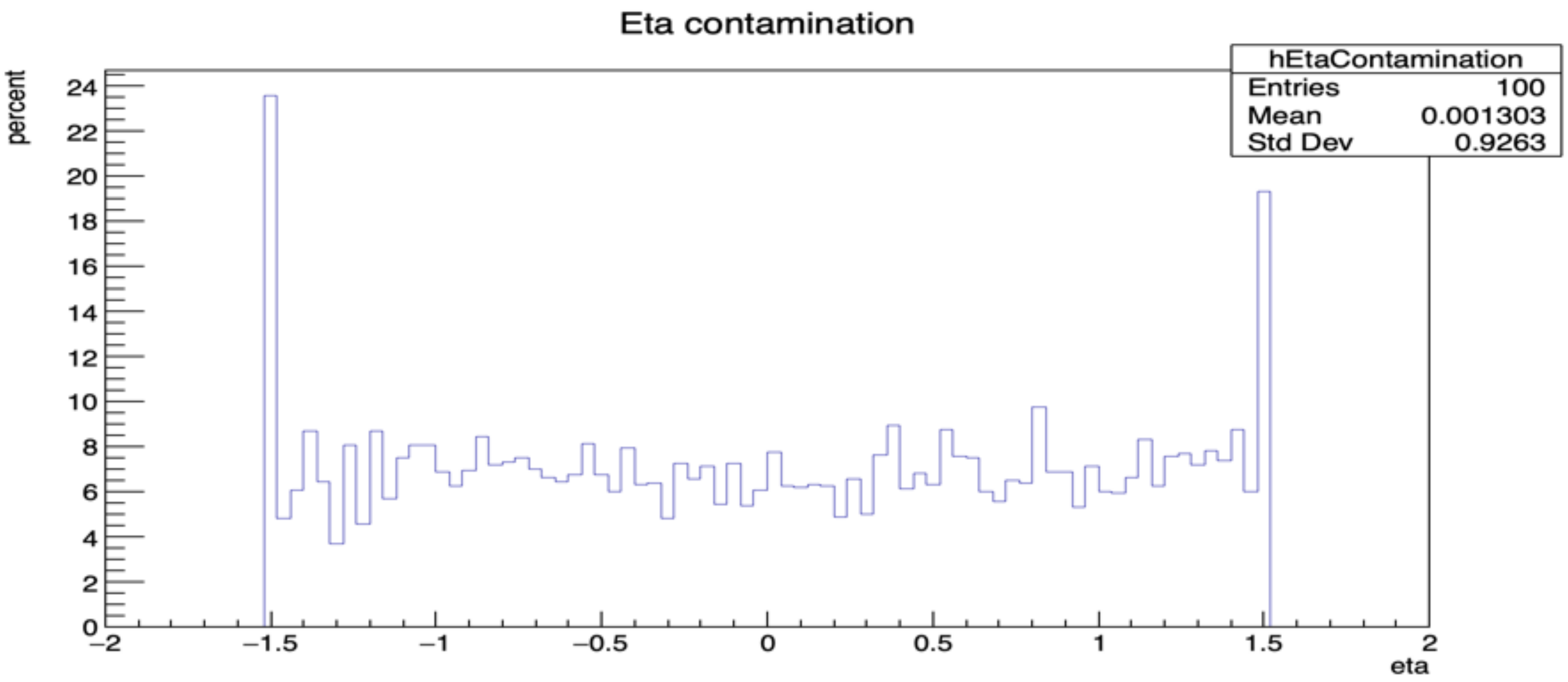
Contamination on pseudorapidity

from two eta spectra

Rec track good and Rec track good selected

percent of tracks that disappear after matching with good primary pure MC track

$$100 * \frac{RecTrackGood - RecTrackGoodSelected}{RecTrackGood}$$



|eta|<0.8 condition is lifted in order to plot this figure

Efficiency on pseudorapidity

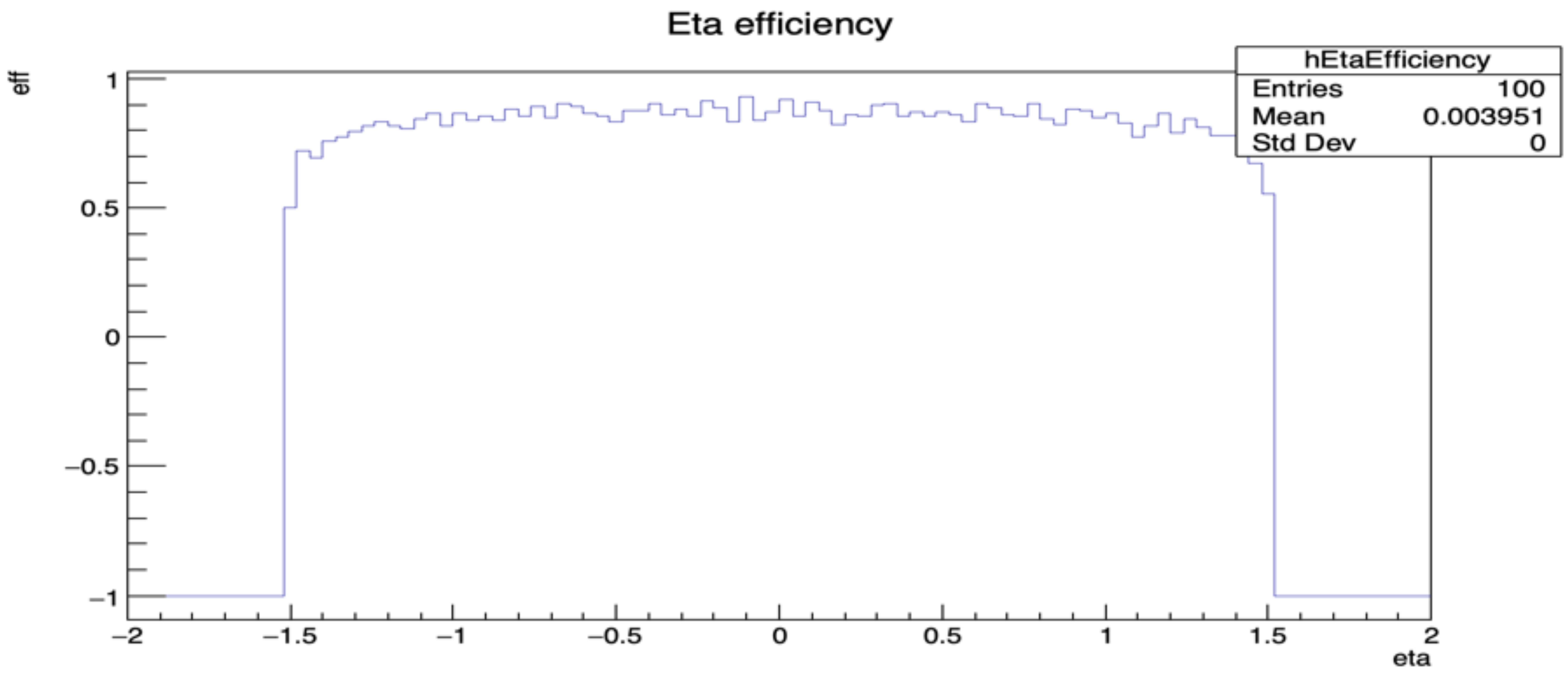
from two eta spectra:

Rec track good selected

and MC good primary

we draw the ratio:

$$\epsilon = \frac{RecTrackGoodSelected}{MCTrackGoodPrimary}$$



|eta|<0.8 condition is lifted in order to plot this figure

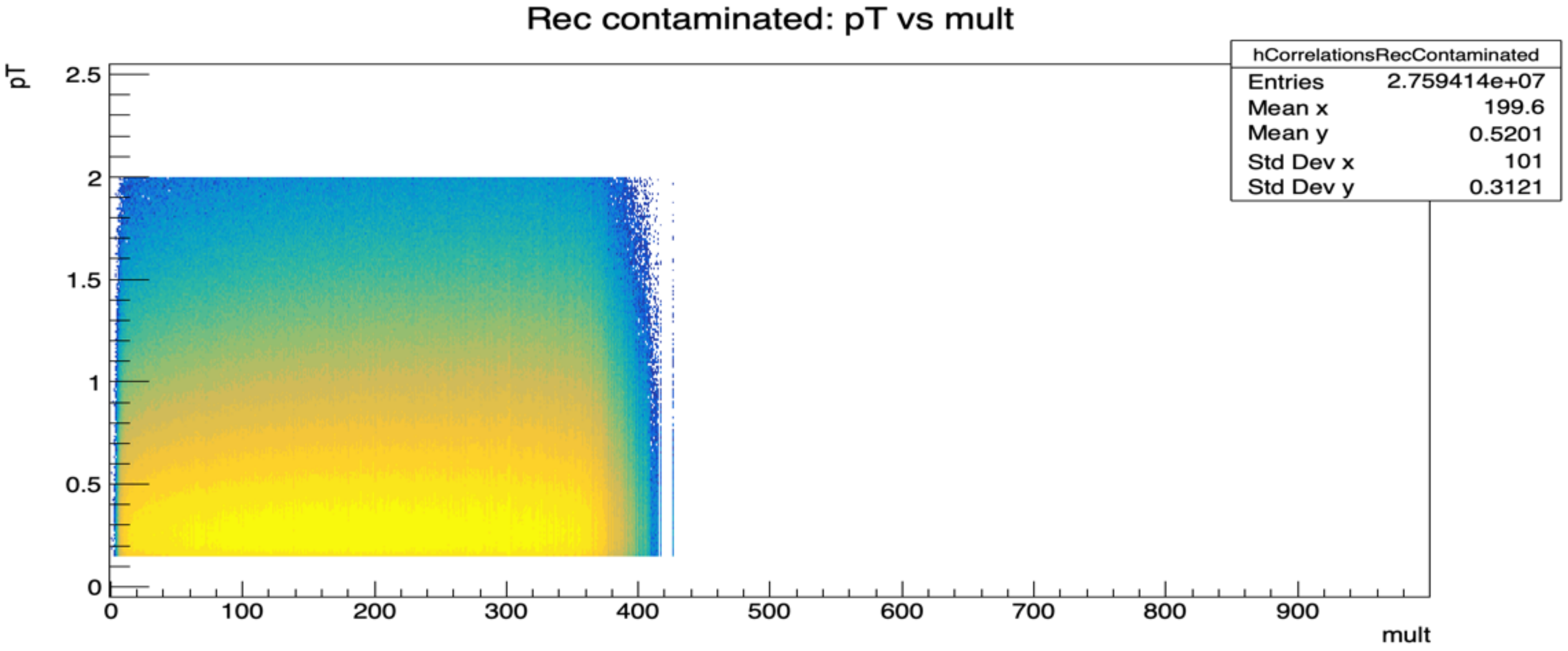
Algorithm of corrections
(a.k.a. sequentional unfolding)

Pics in this section are for low statitstics
'self-corrections'
(just to illustrate the idea)

What do we want to measure?

pT spectrum as a function of event multiplicity

experimentally we will be (hopefully) close to the ‘Rec contaminated’ level

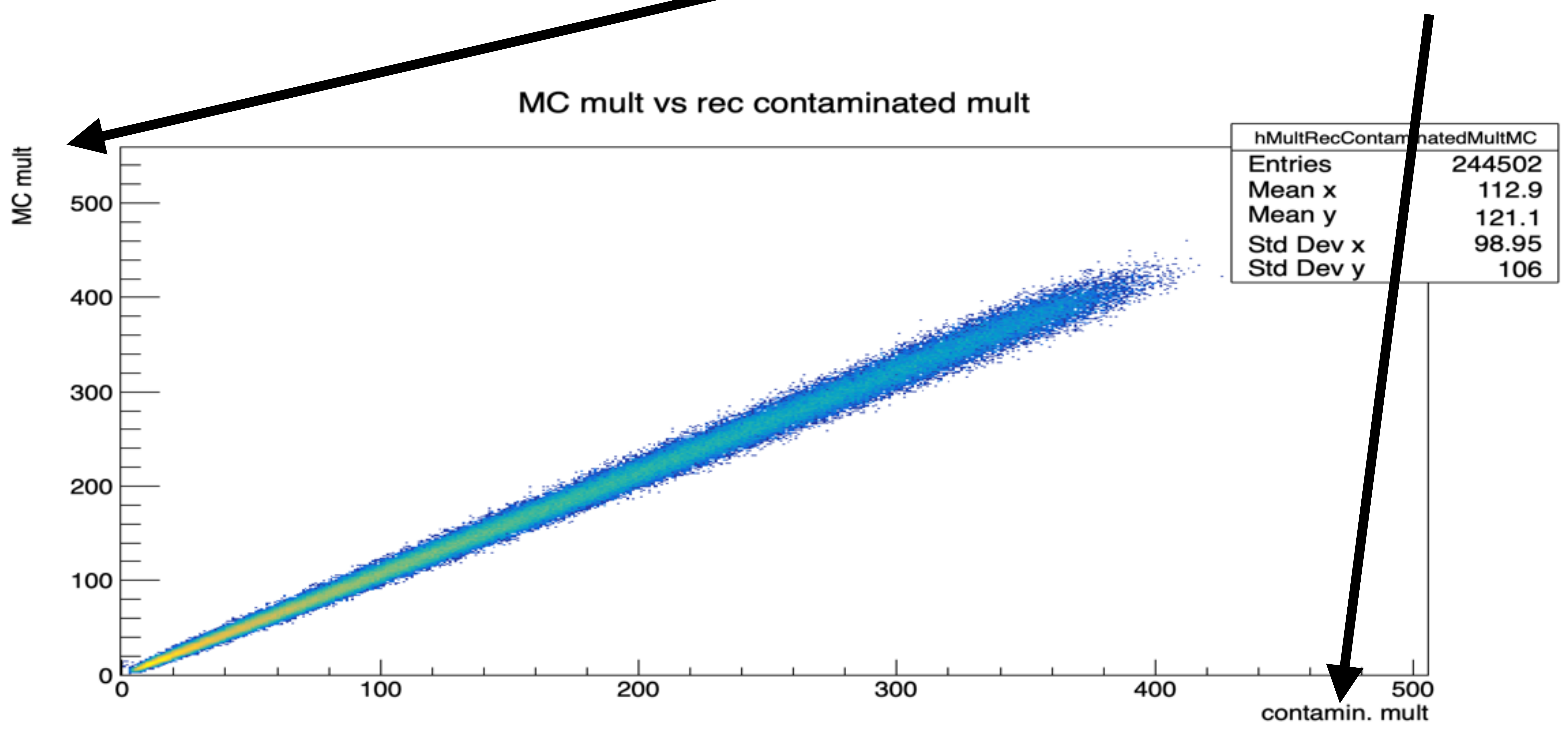


both multiplicity distribution and pT spectrum are distorted

suggestion - repeat ALICE procedure Phys. Lett. B845,138110 (2023)

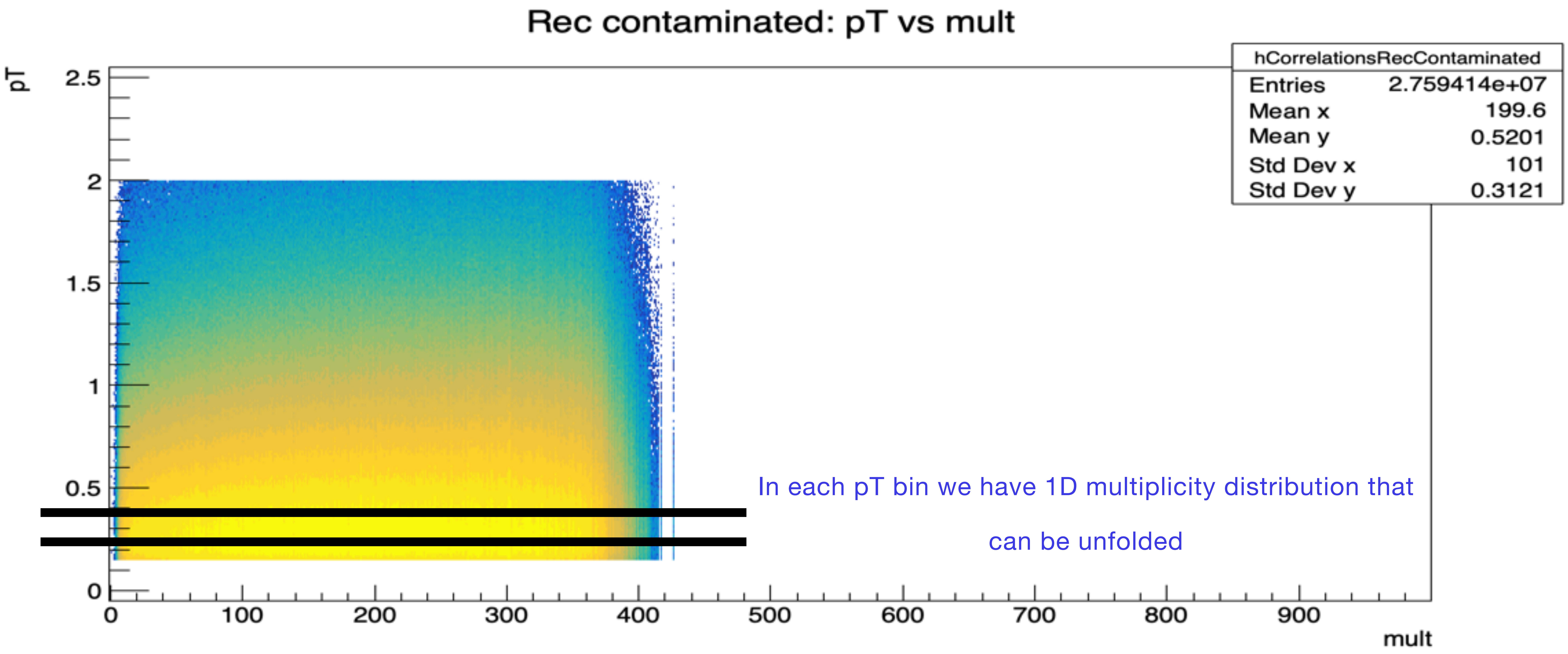
Curing the multiplicity

1) prepare response matrix via RooUnfold that shows correlations between multiplicities of good primary MC tracks and good Rec tracks



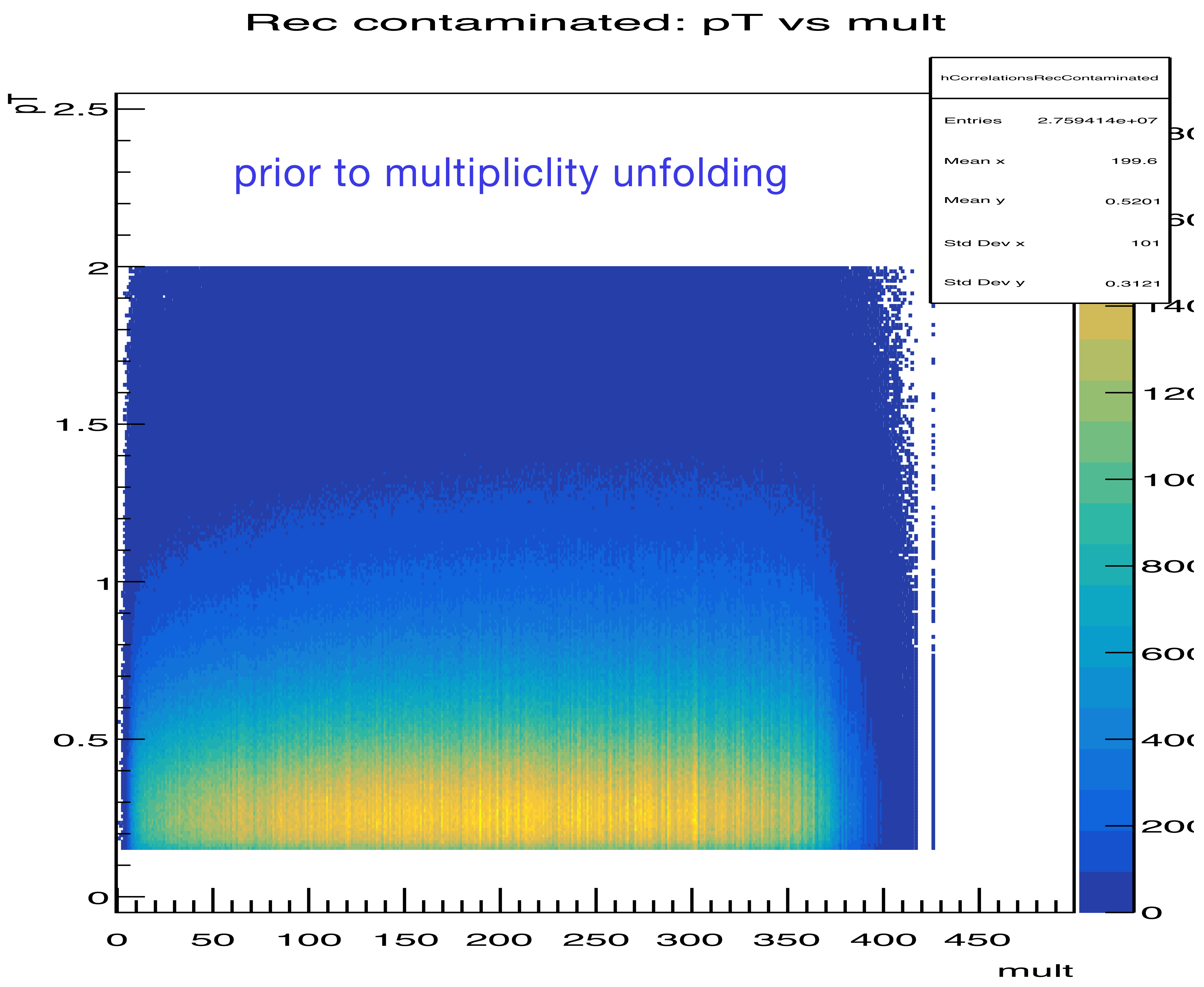
Curing the multiplicity

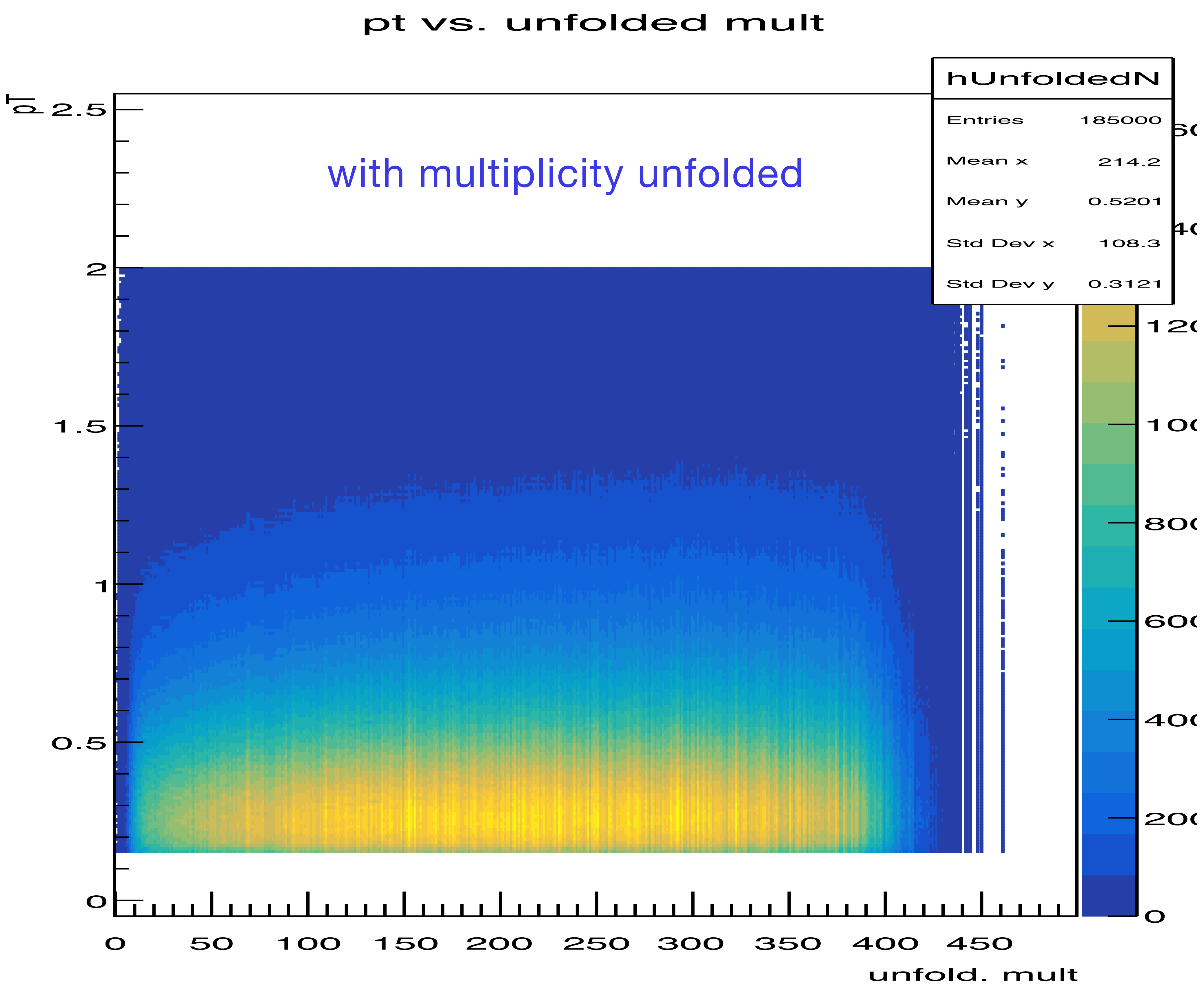
2) apply unfolding in each horizontal slice, i.e. for eact pT bin separately



Curing the multiplicity

3) so we get distorted pt vs corrected multiplicity histogram

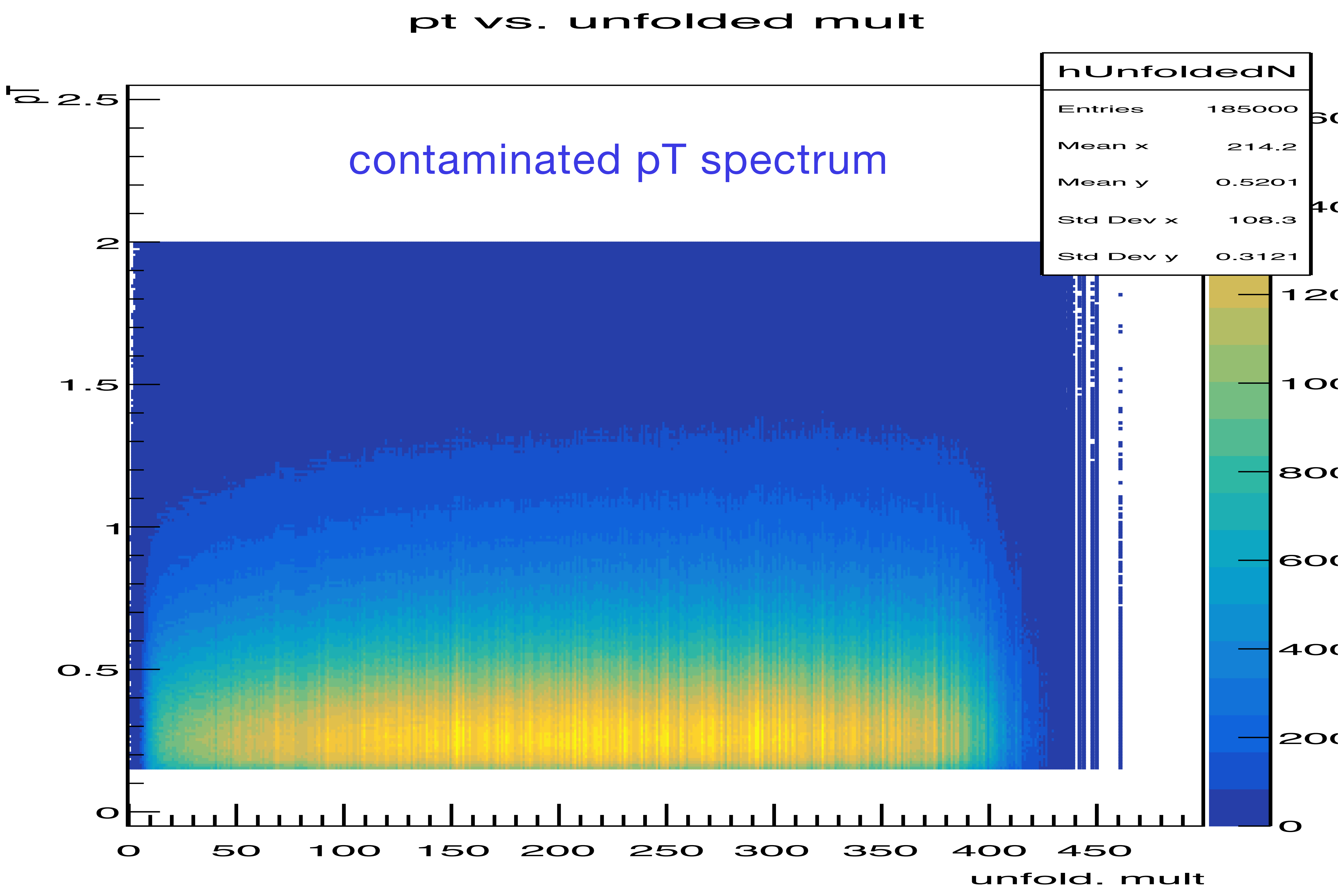


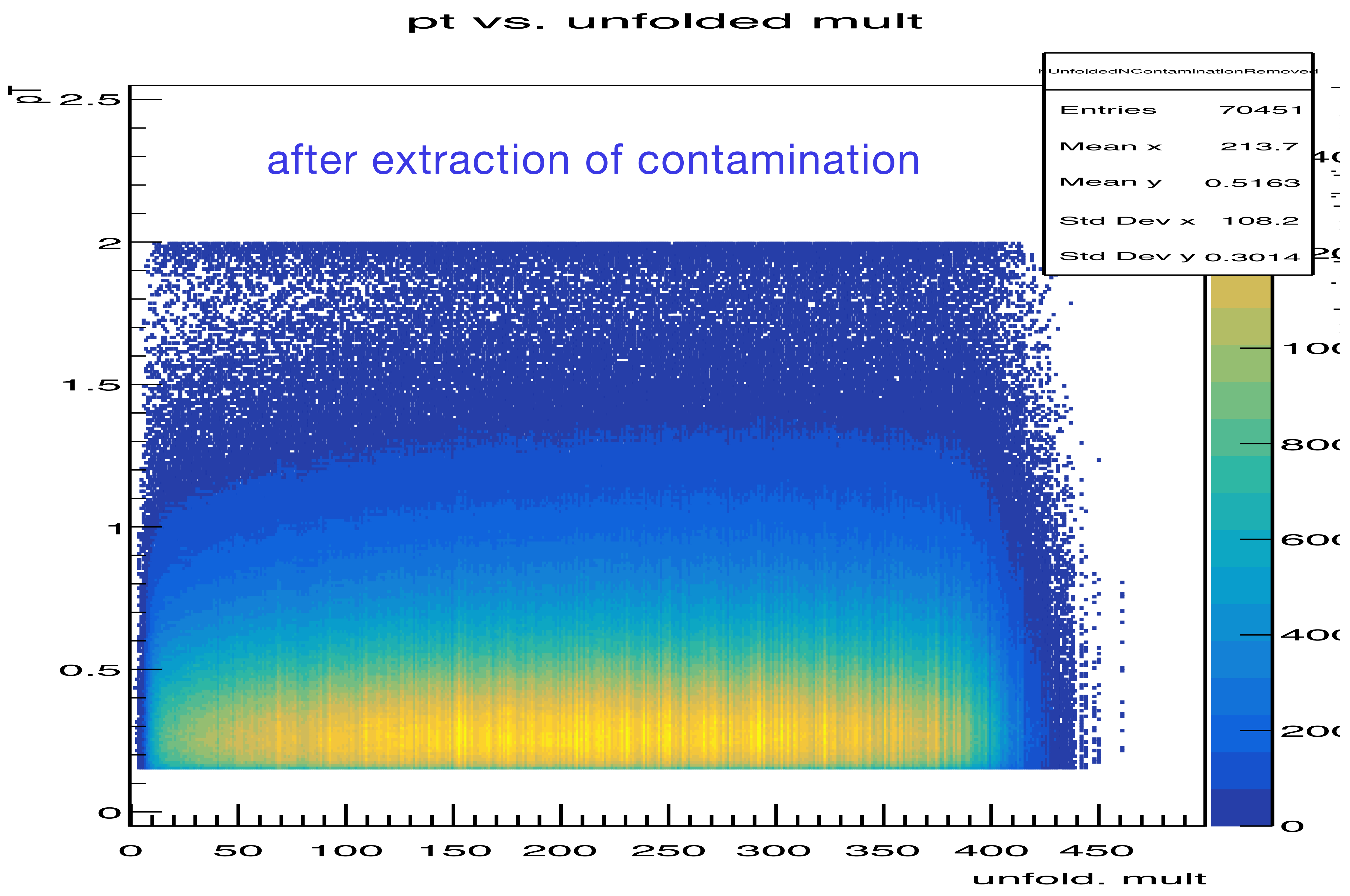


in short - we rescaled along the X axis in 'smart' way

Curing the contamination (brute force)

4) good Rec tracks contain fraction of fakes and secondaries, these fractions are calculated in each bin of reconstructed pT and good primary MC tracks' multiplicity. These fractions are removed in a multiplicative way bin-by-bin.



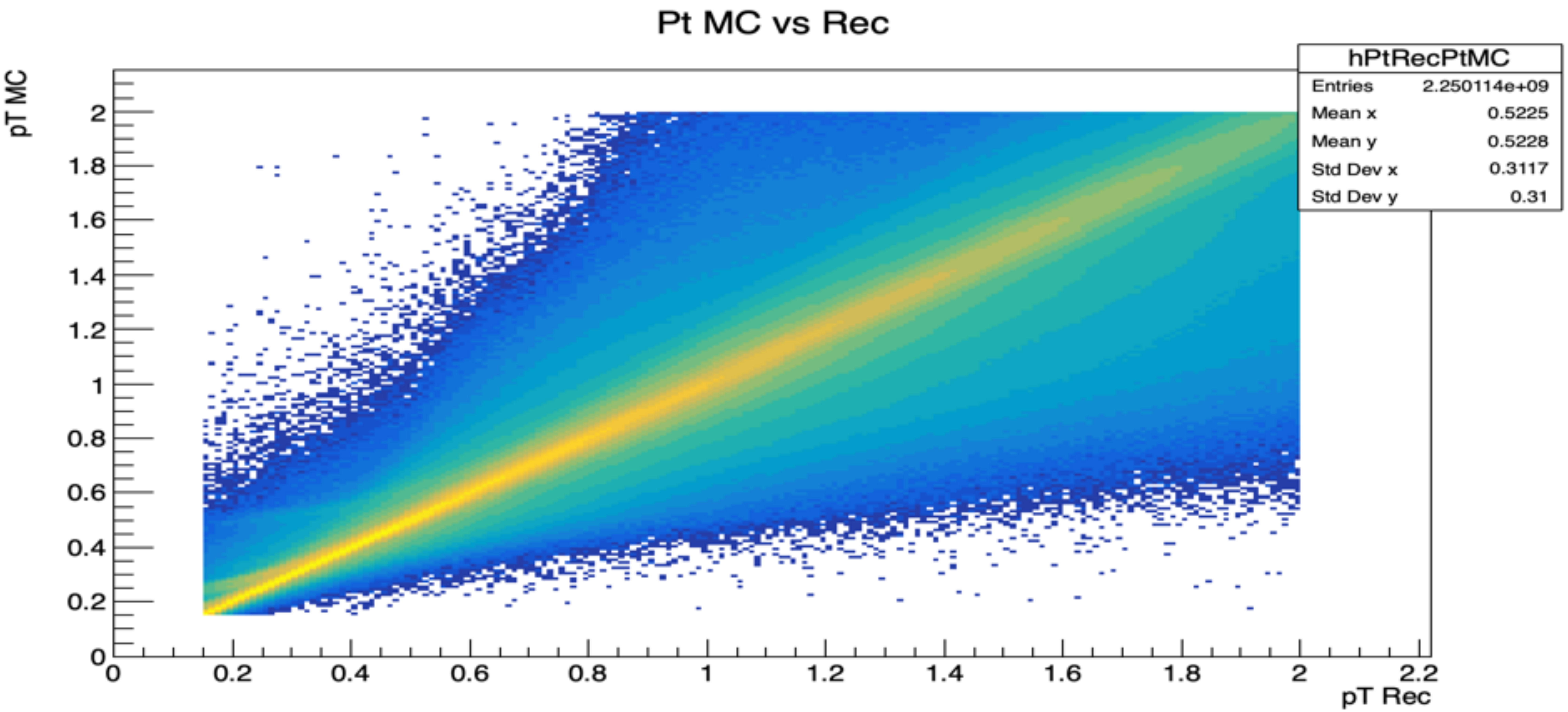


due to low stats here, when fraction of contamination was unknown

I filled the corrected hist with 0

Unfolding pT spectra

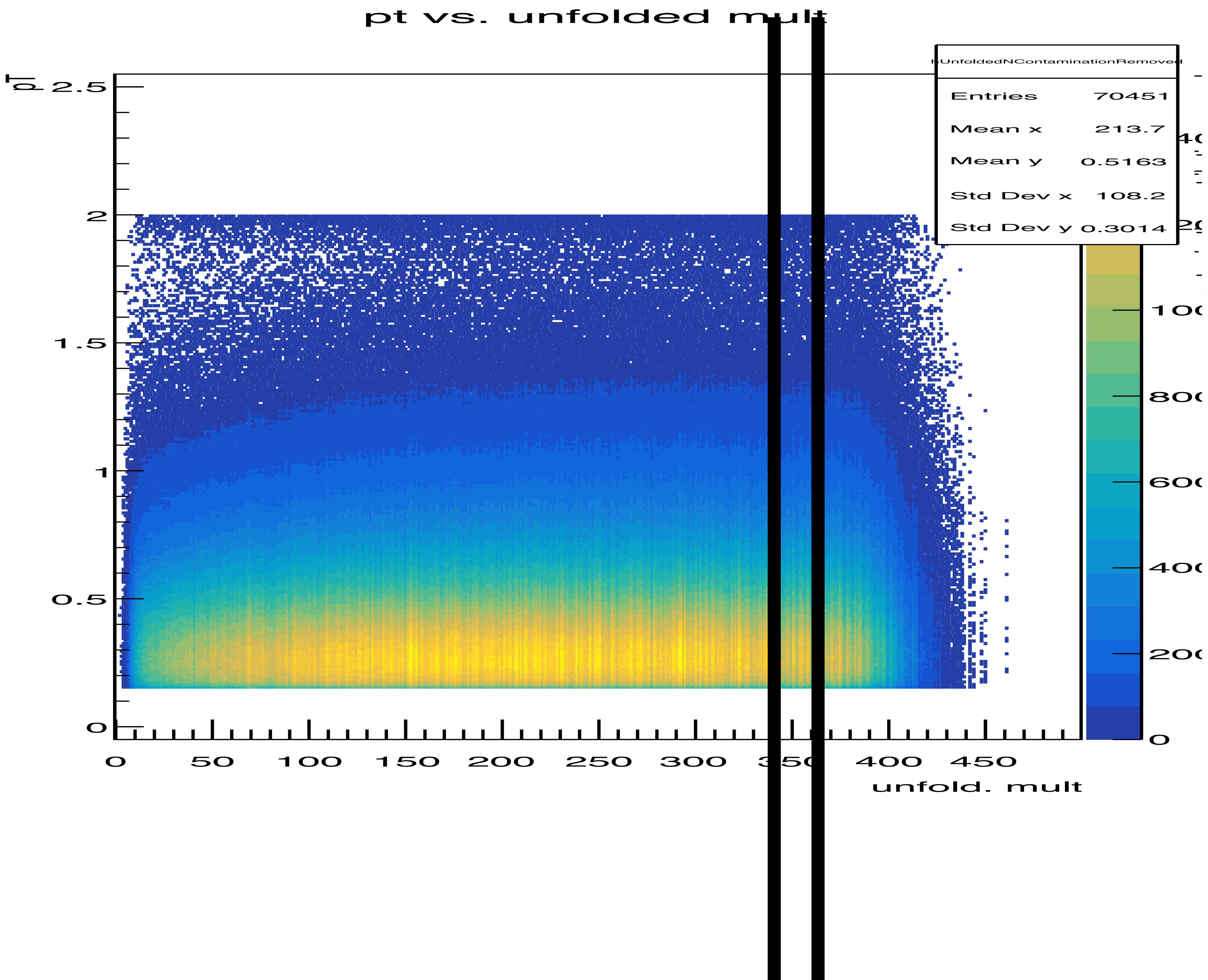
5) measured pT has non-zero resolution, one can prepare response matric between pT of good selected Rec track and pt of its matched good primary MC track



Note for future studies of systematics: some non-diagonal structures are visible - we should vary track cuts (e.g. min. #ofTPC hits)

Unfolding pT spectra

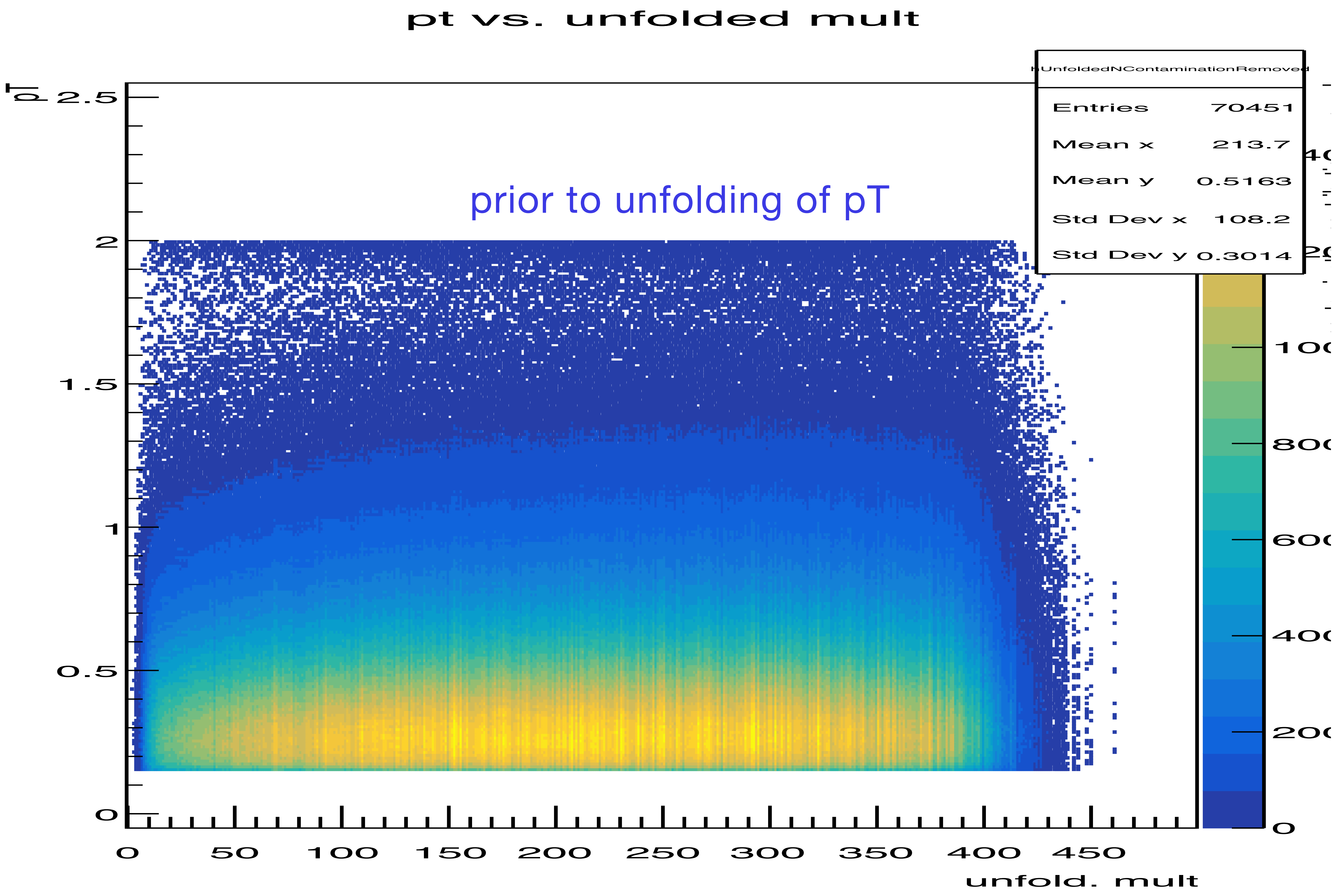
6) pT unfolding is applied for each multiplicity bin separately

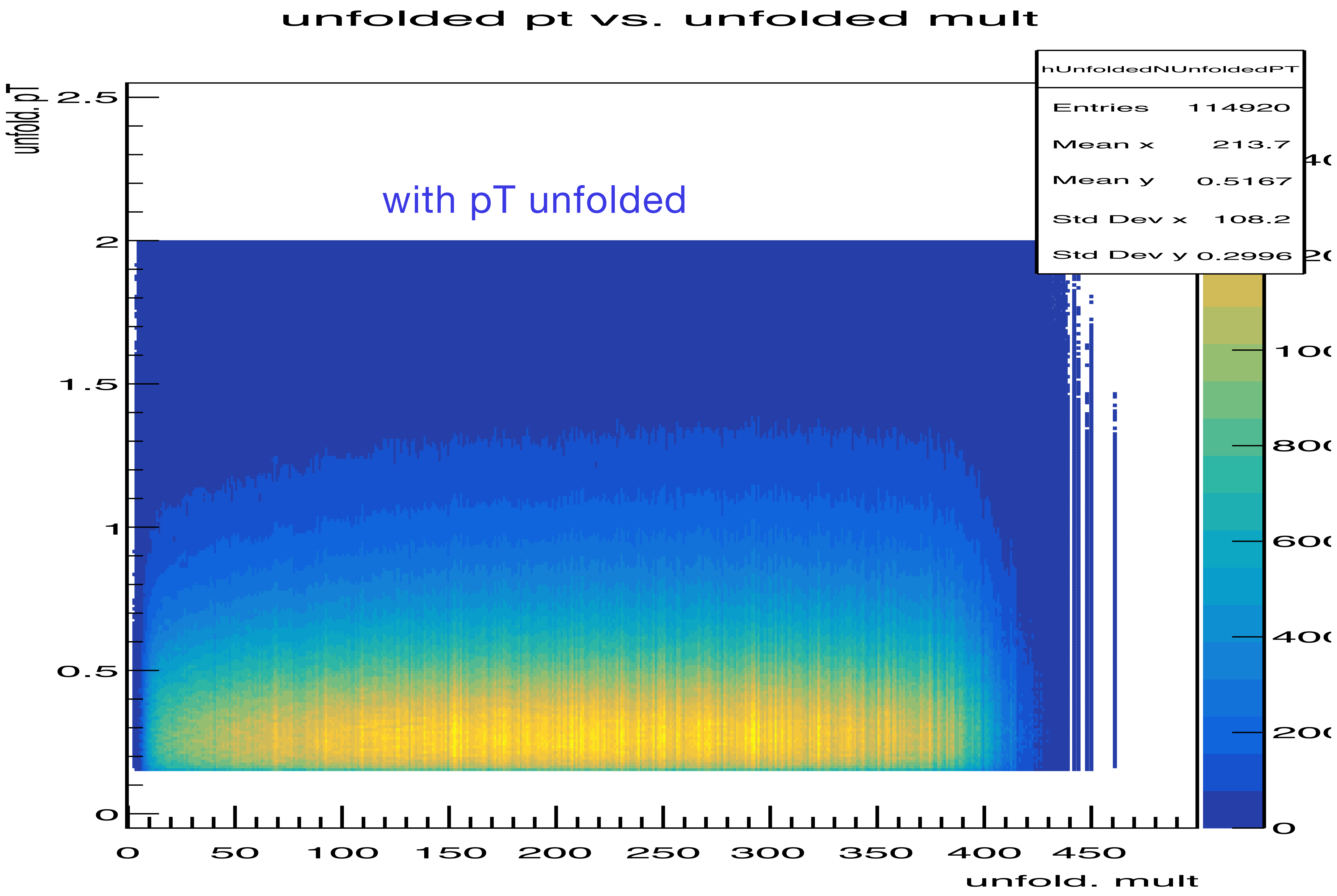


in each multiplicity bin we have 1D pT distribution that can be unfolded with 2d response matrix

Unfolding pT spectra

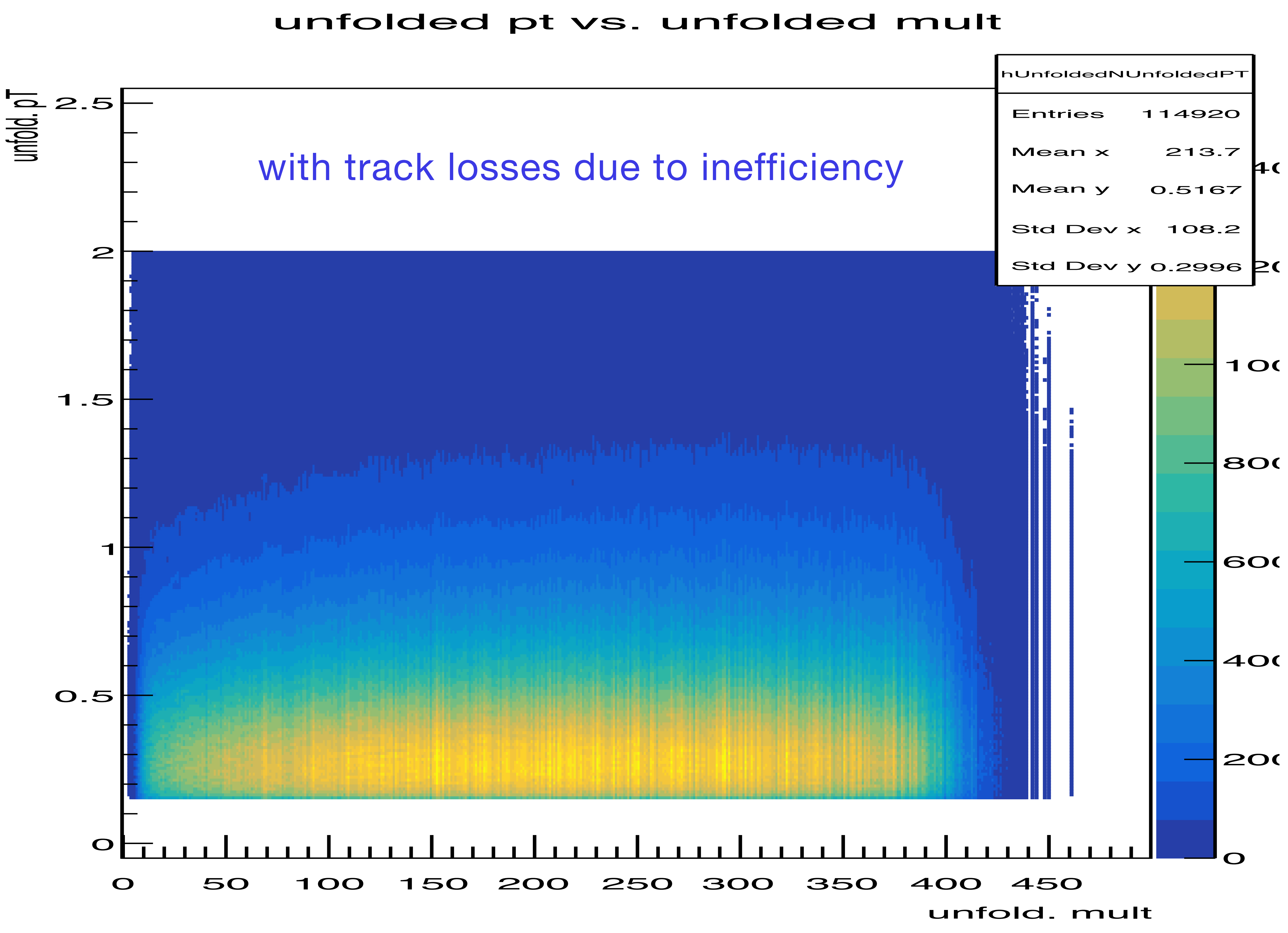
7) so we get corrected on cont. and resolution pt vs corrected multiplicity histogram

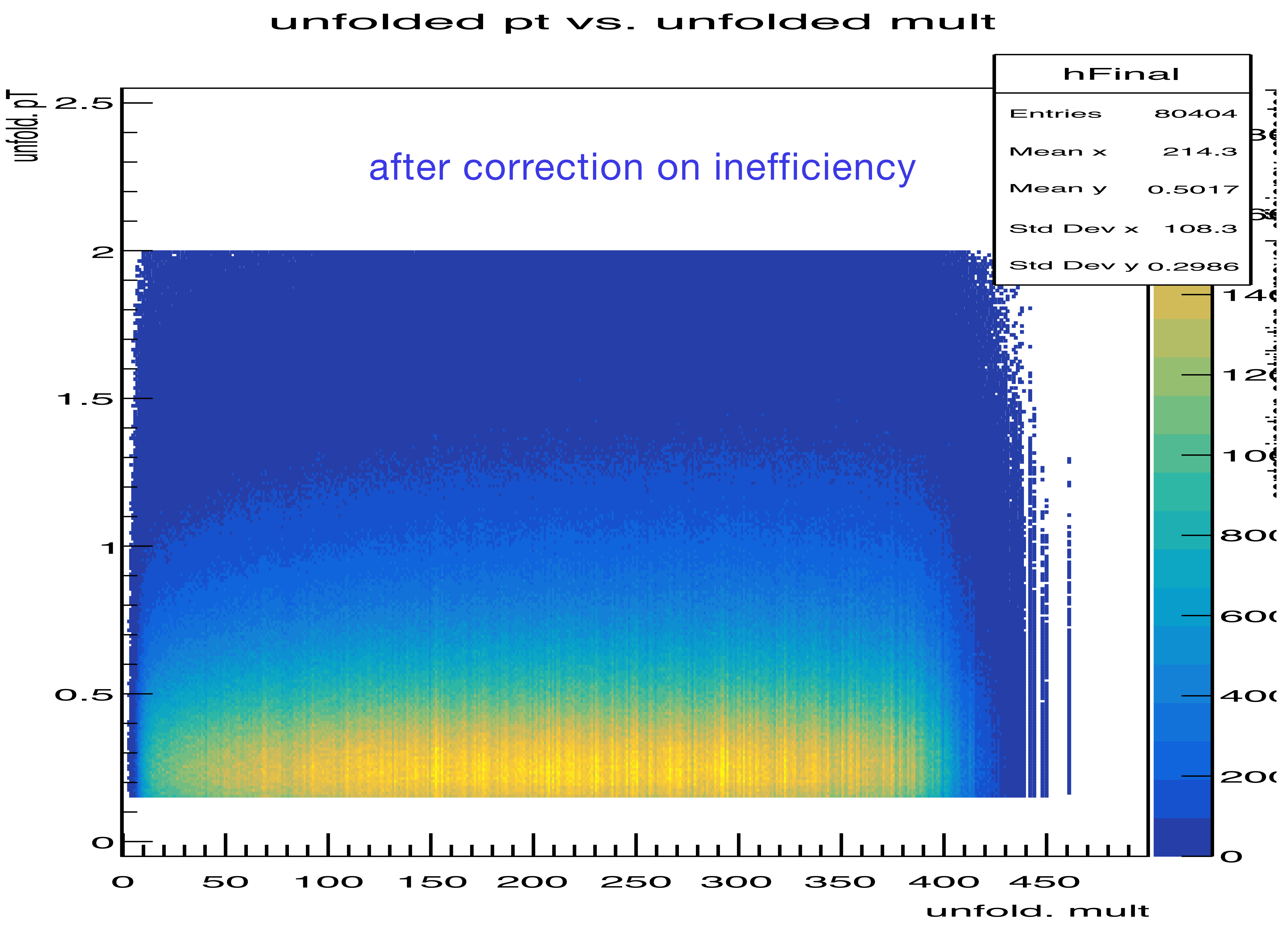




Corrections on inefficiencies (brute force)

8) similar to contamination step we prepare in advance 2d histogram (pT vs multiplicity) of efficiencies of reconstruction of a track. We restore up to 100% efficiency in a multiplicative way bin-by-bin.



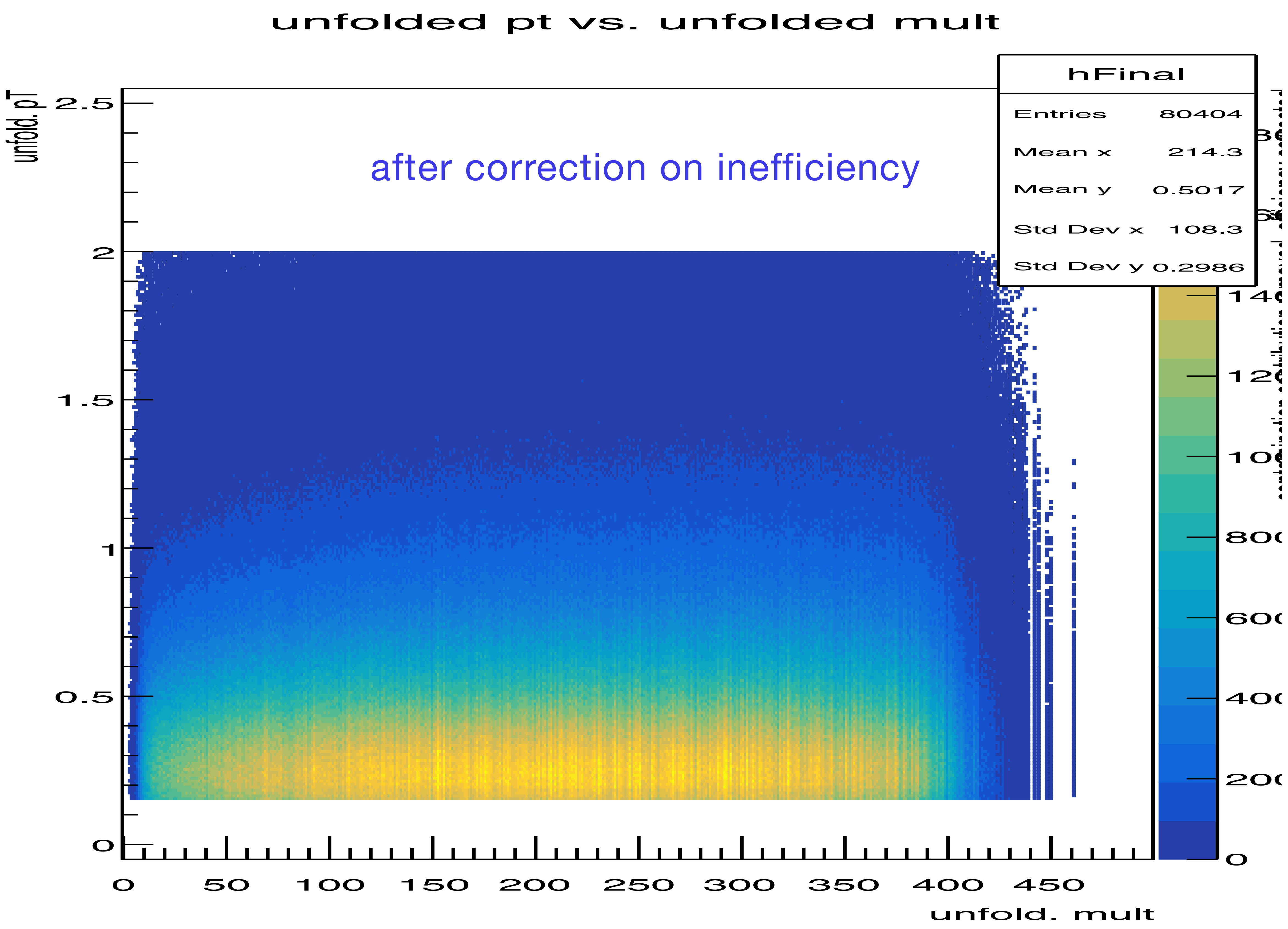


What’s next?

We should extract moments of pT spectrum as a function of multiplicity (so-called correlation functions)

We should compare with the results obtained for undistorted data (i.e. for good primary MC tracks)

By doing so we should conclude if the suggested procedure is viable



Tests of corrections

Correct Request25 by Request25

split dataset into train and test parts, create all the matrices based on train subset, apply corrections to test subset

Correct Request26 by Request25

cross-validation

For both points:

Corrections on artificial inefficiencies

reject randomly selected fraction of good reconstructed tracks (up to 50%) and check whether corrections will restore true level

Correct Request25 by Request25

1) Construct all ‘response matrices’ on train subset

2) Apply it to train subset

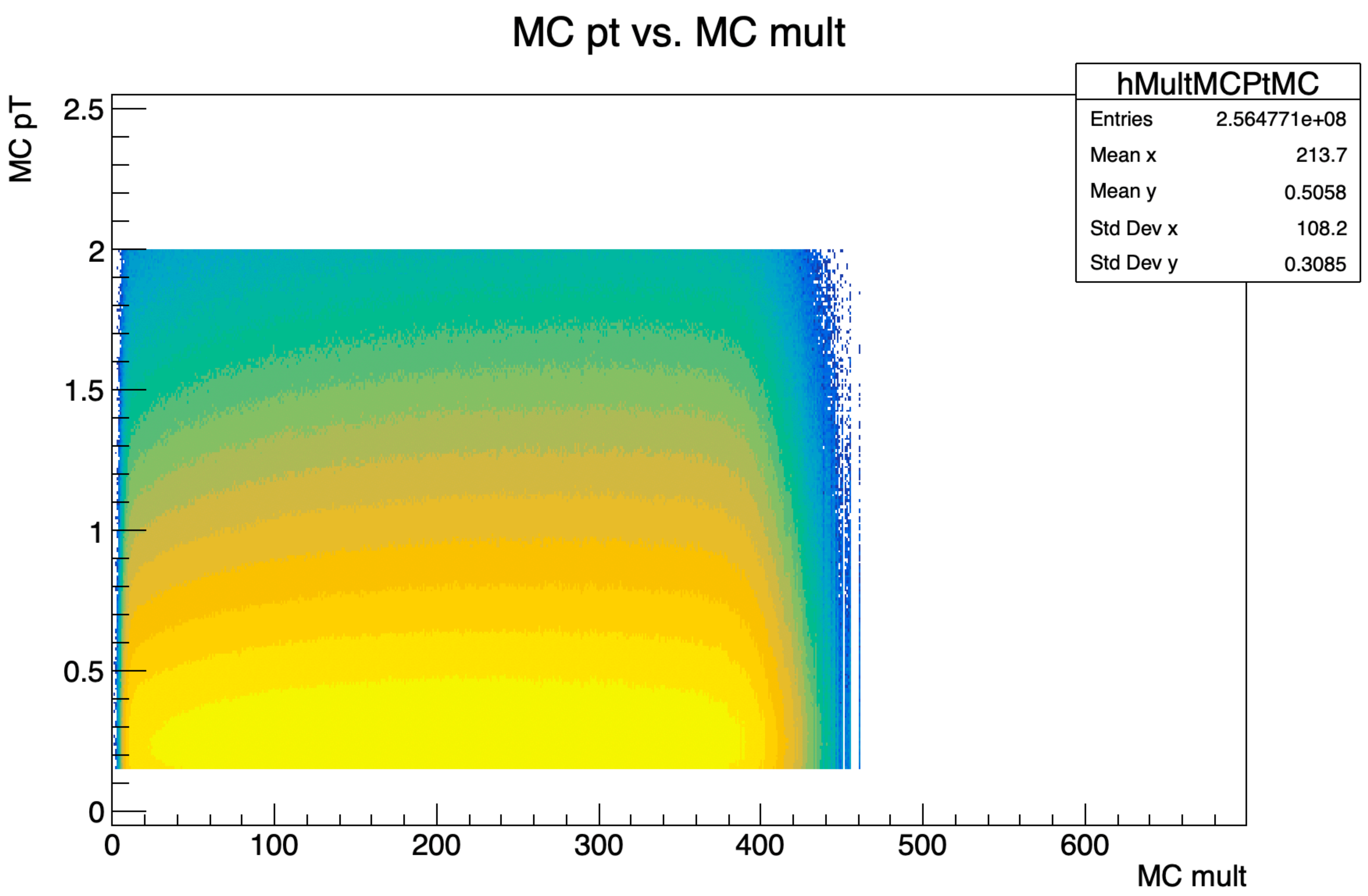
obtain values of correlation functions

3) In order to estimate statistical uncertainties train subset is split further to subsubsamples

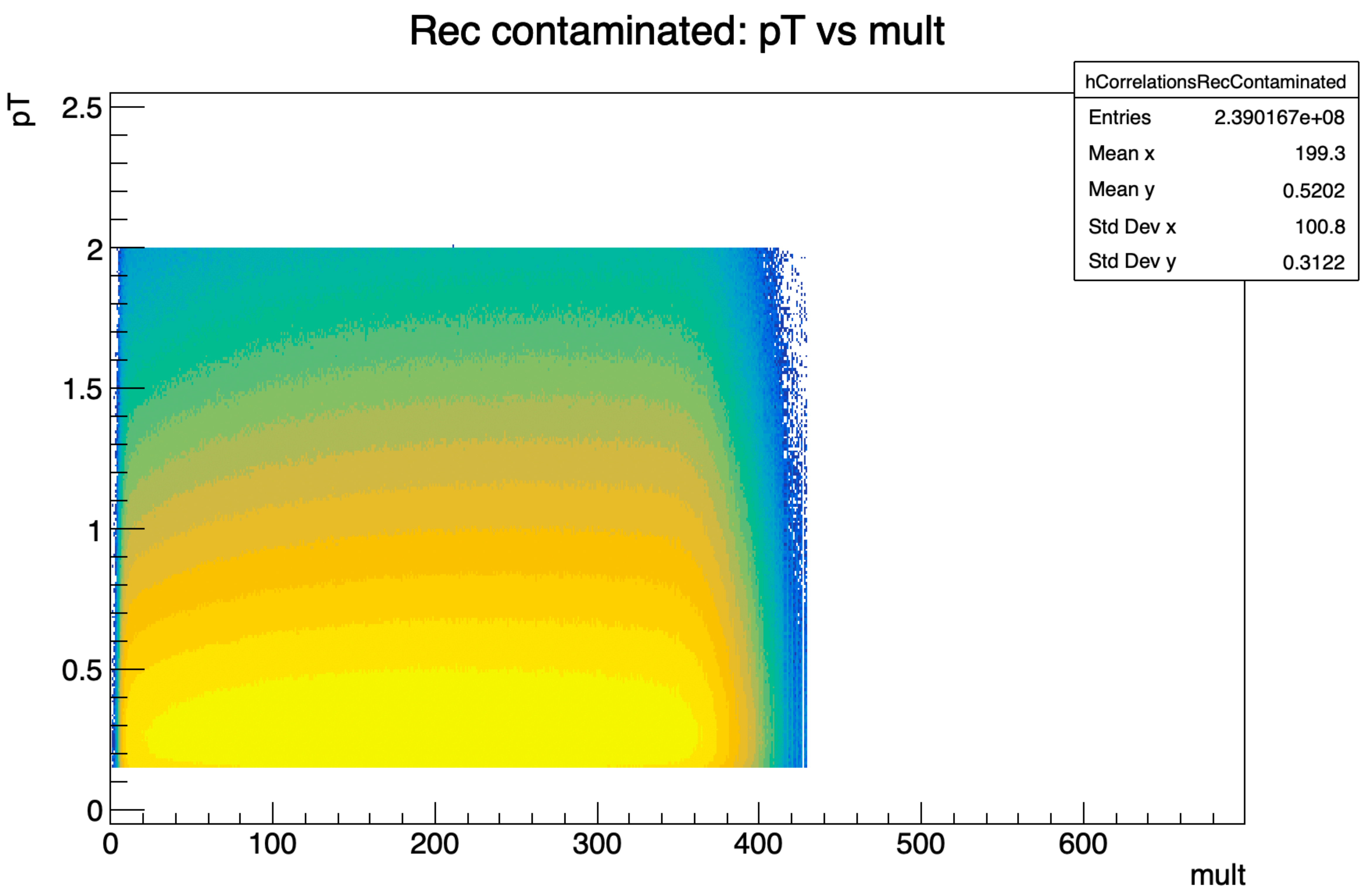
each subsubsample is corrected separately, we get a set of correlation functions, their variances provide uncertainty

Test dataset

what we want to achieve

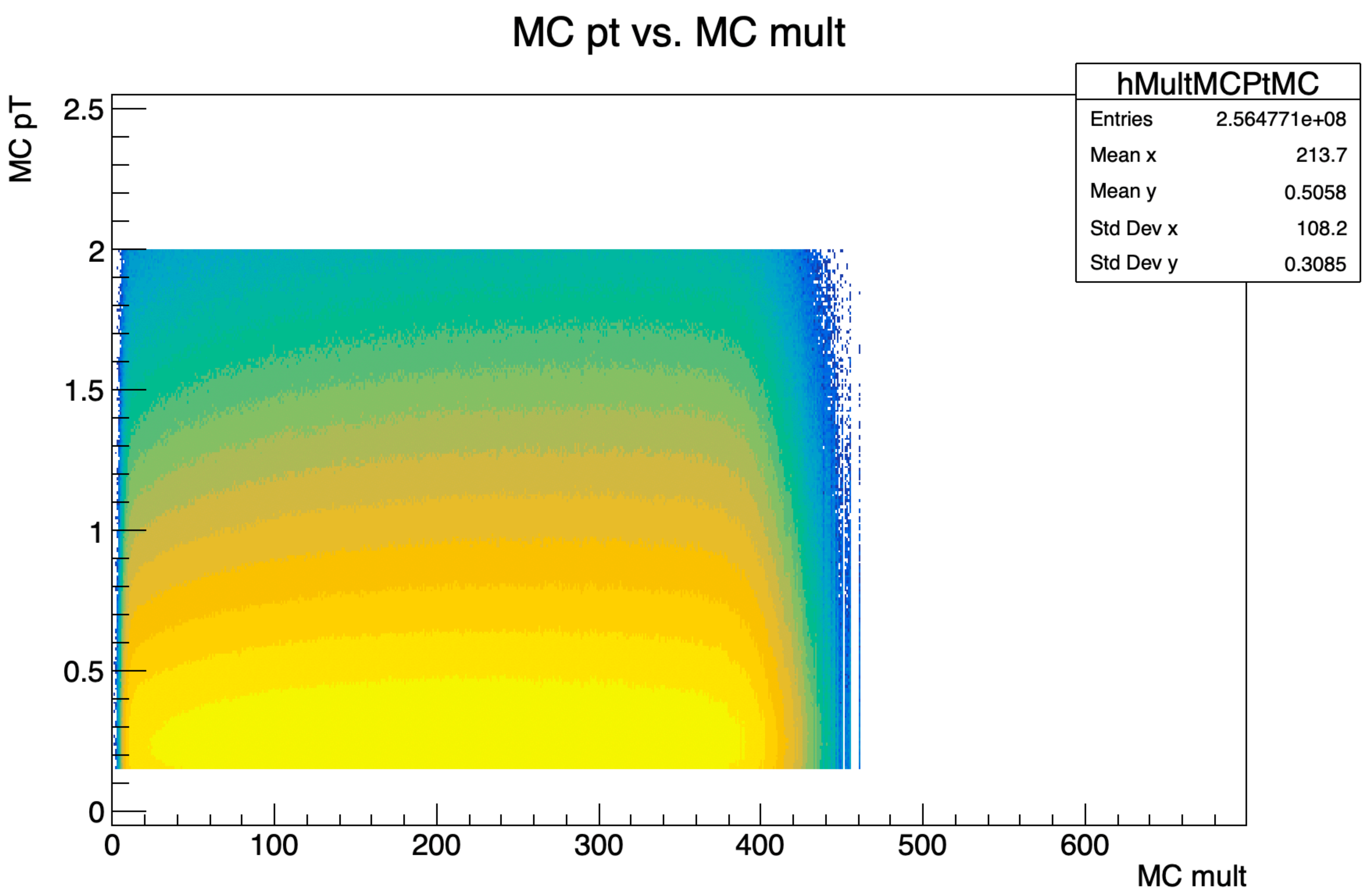


what we have on Rec level

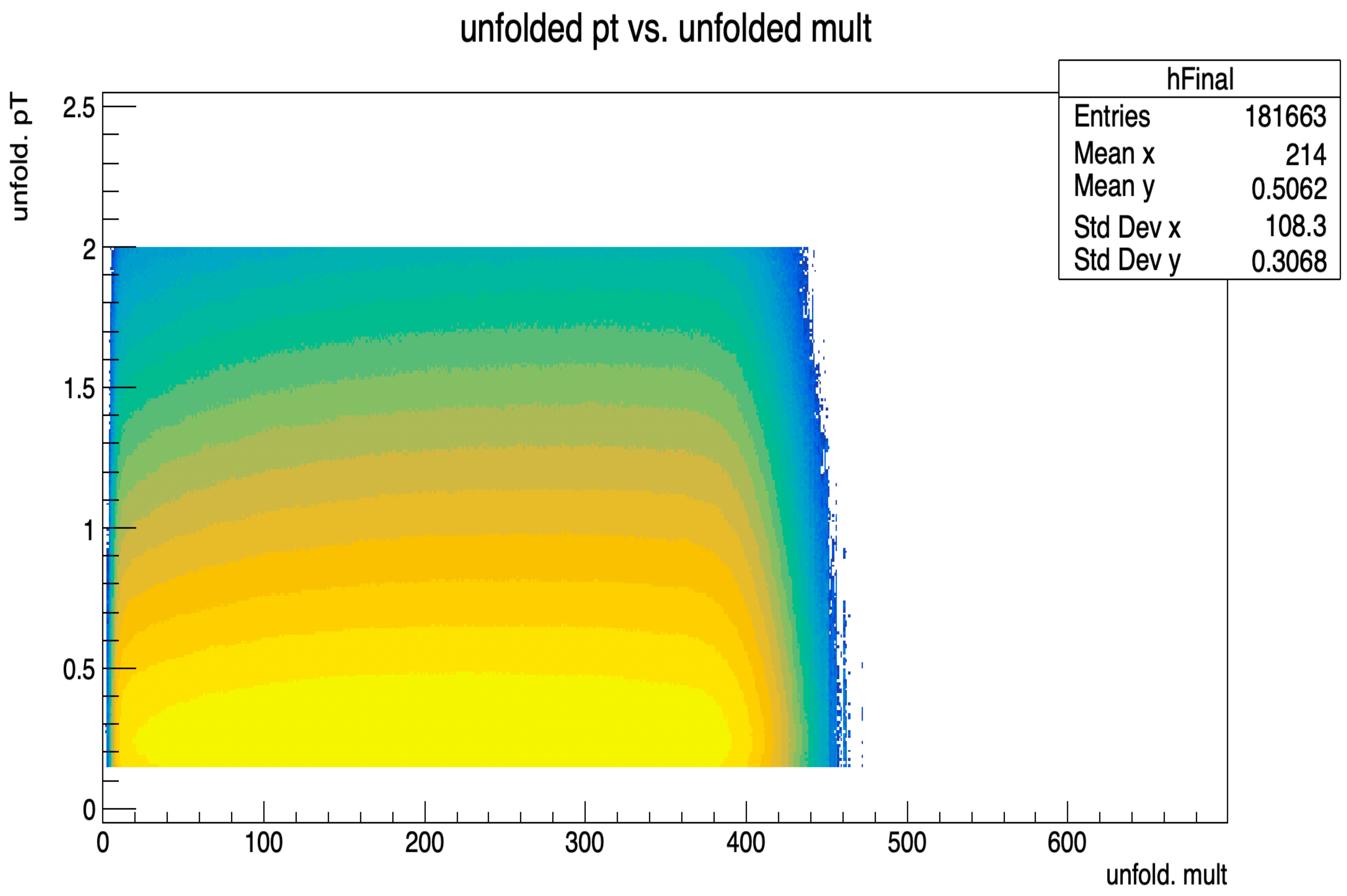


Test dataset

what we want to achieve

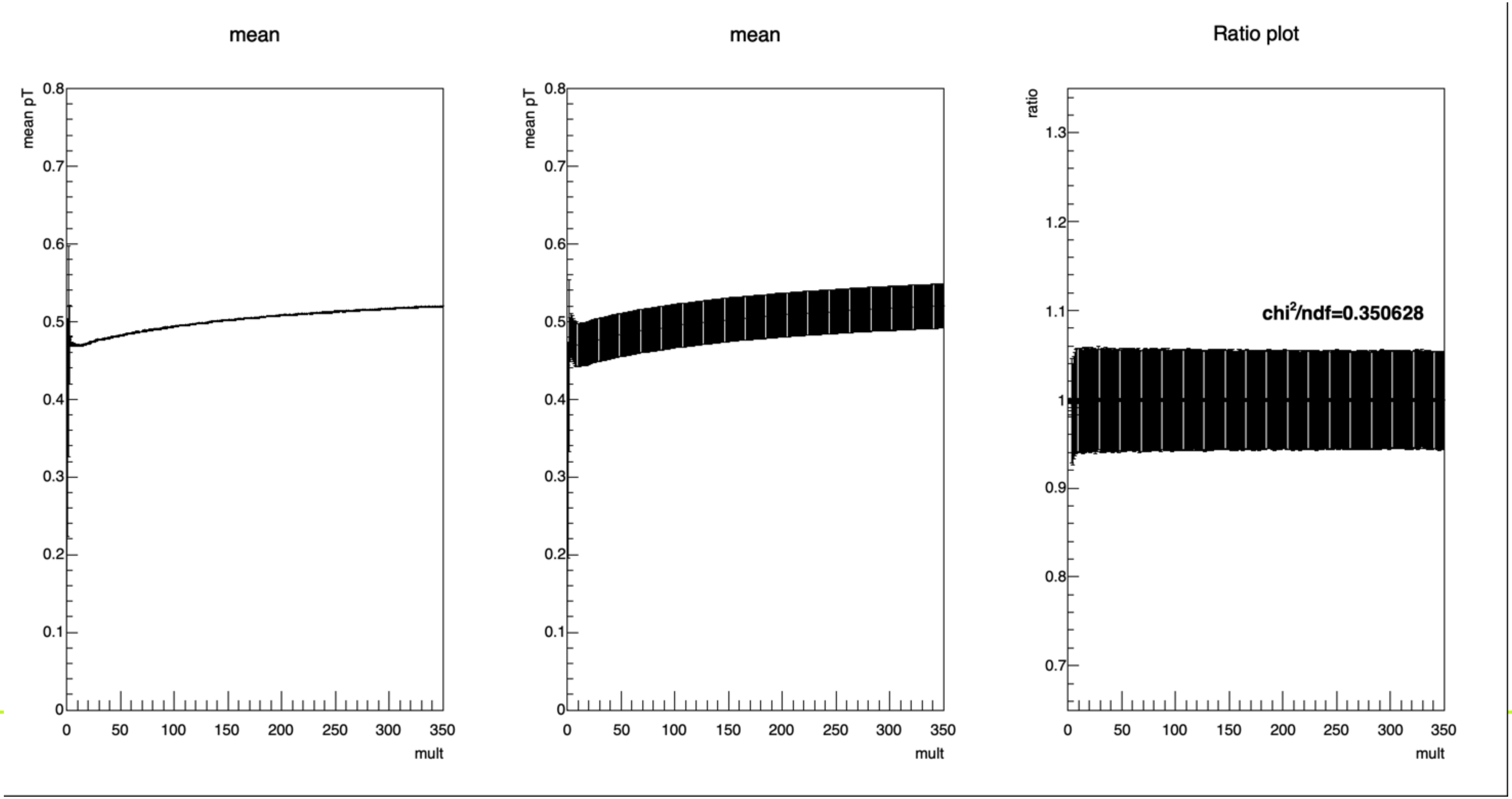


what we get after corrections



Test dataset: first moments

what we want to achieve | what we get after corrections | ratio

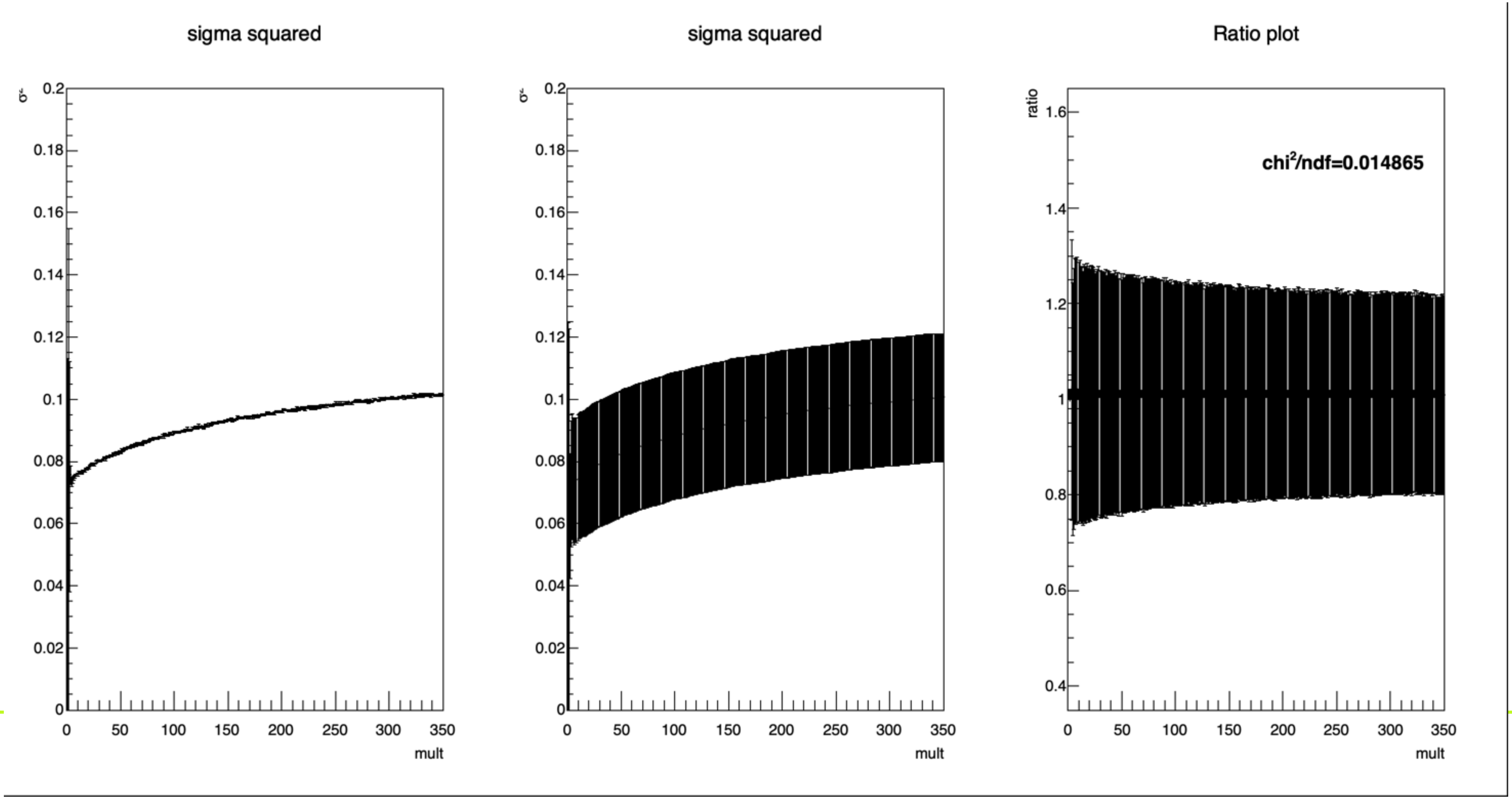


Test dataset: second moments

what we want to achieve

what we get after corrections

ratio



Correct Request26 by Request25

1) Construct all ‘response matrices’ on request25 set

2) Apply it to request26 set

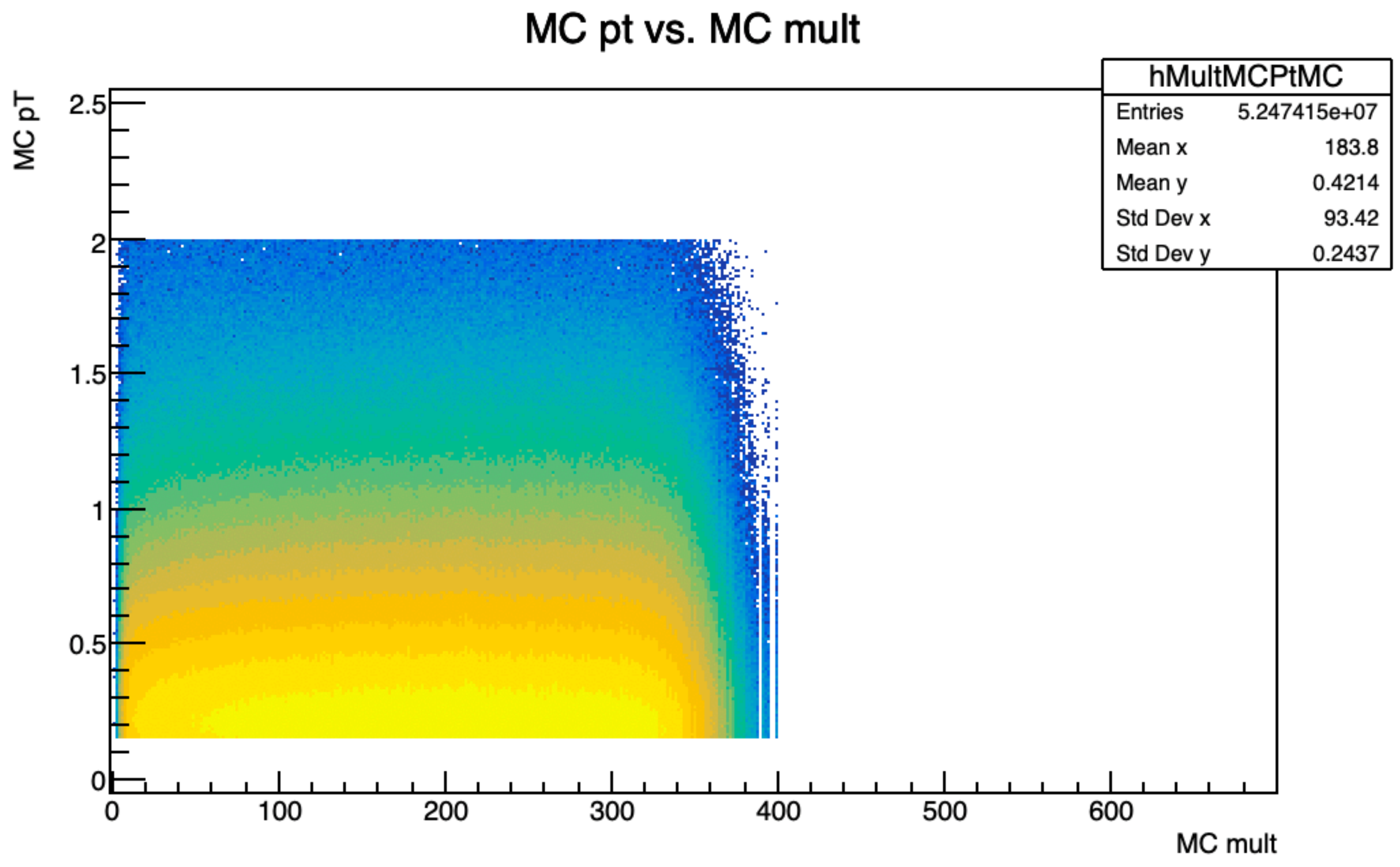
obtain values of correlation functions

3) In order to estimate statistical uncertainties request26 set is split further to subsamples

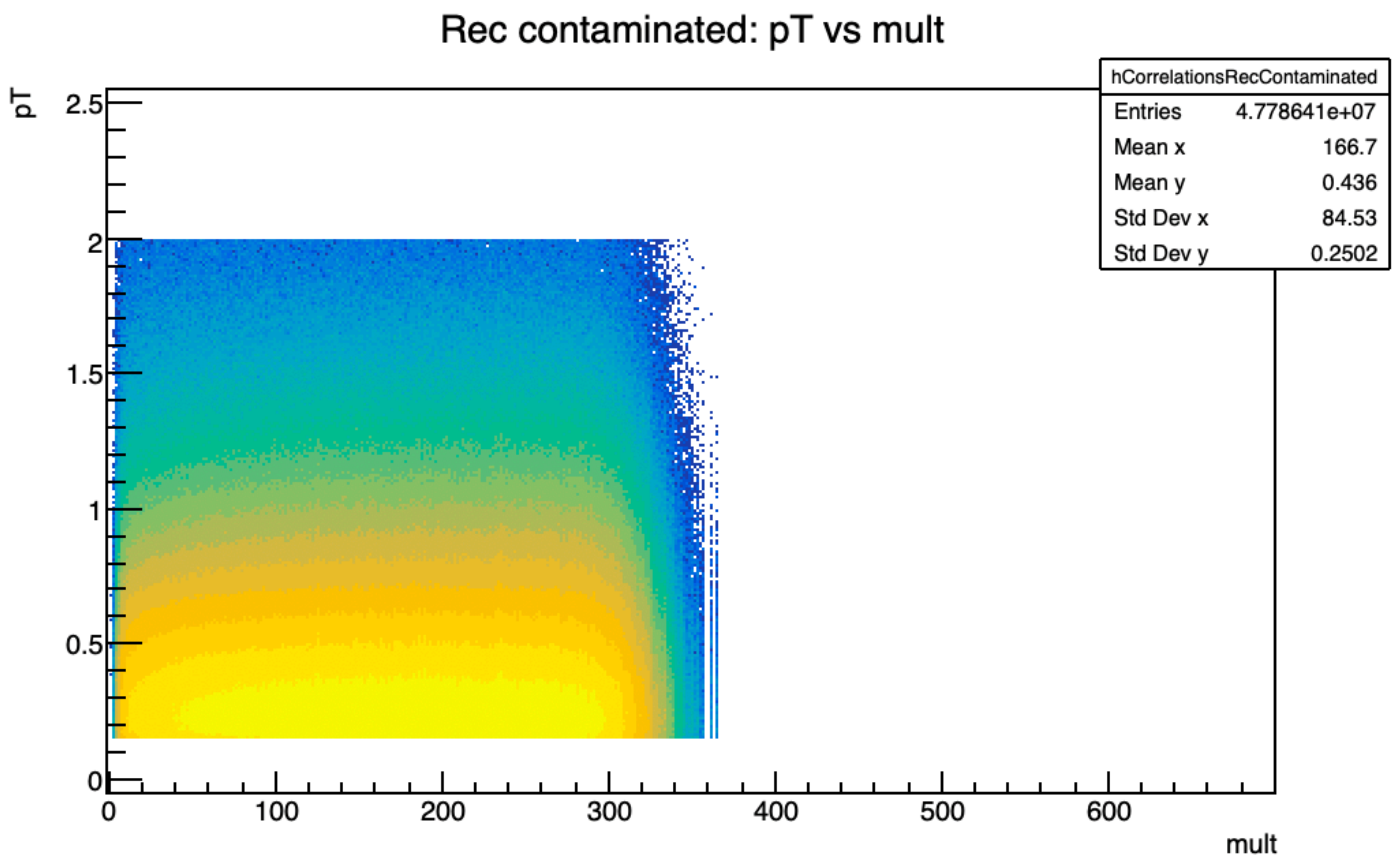
each subsample is corrected separately, we get a set of correlation functions, their variances provide uncertainty

Request26 dataset

what we want to achieve

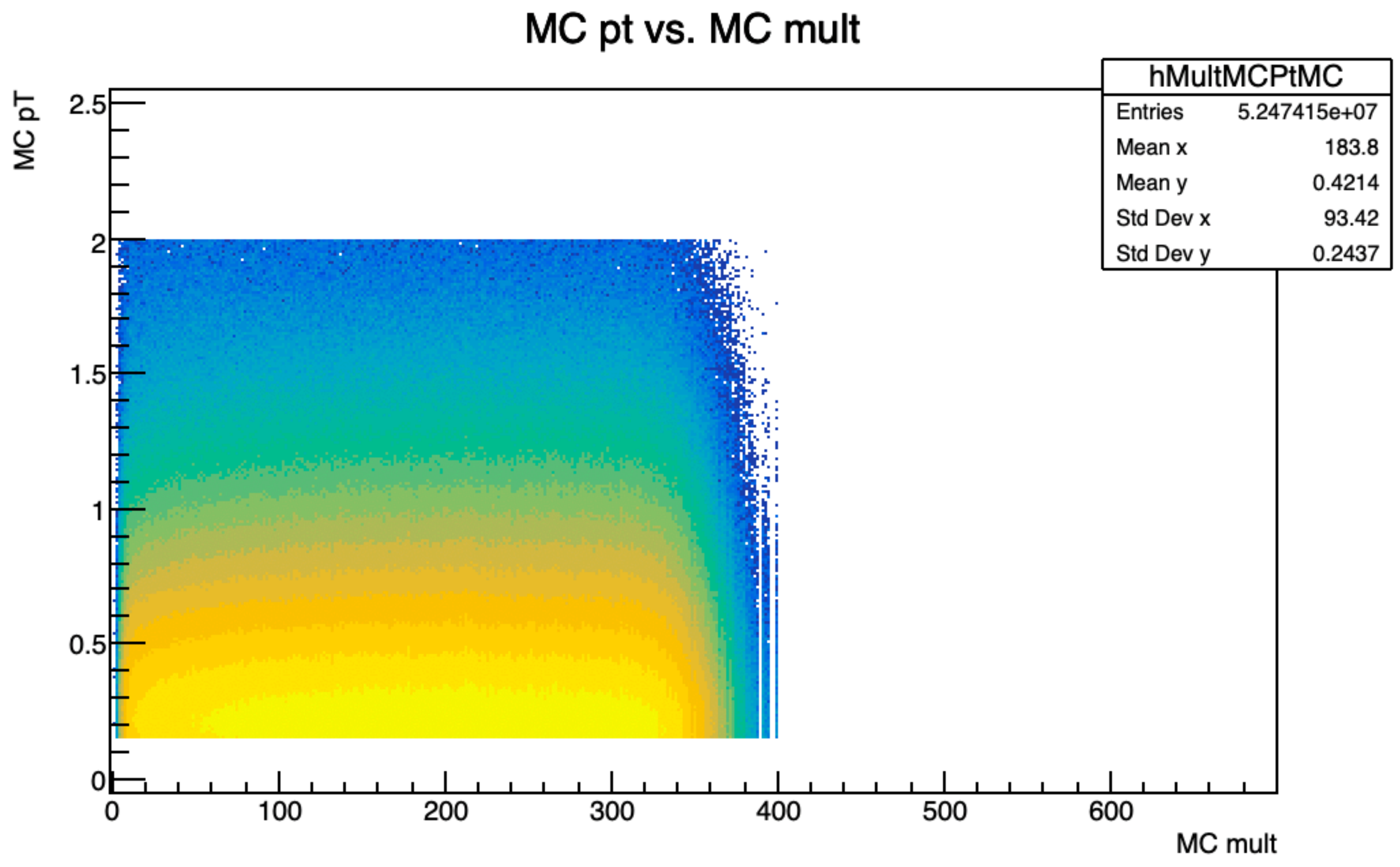


what we have on Rec level

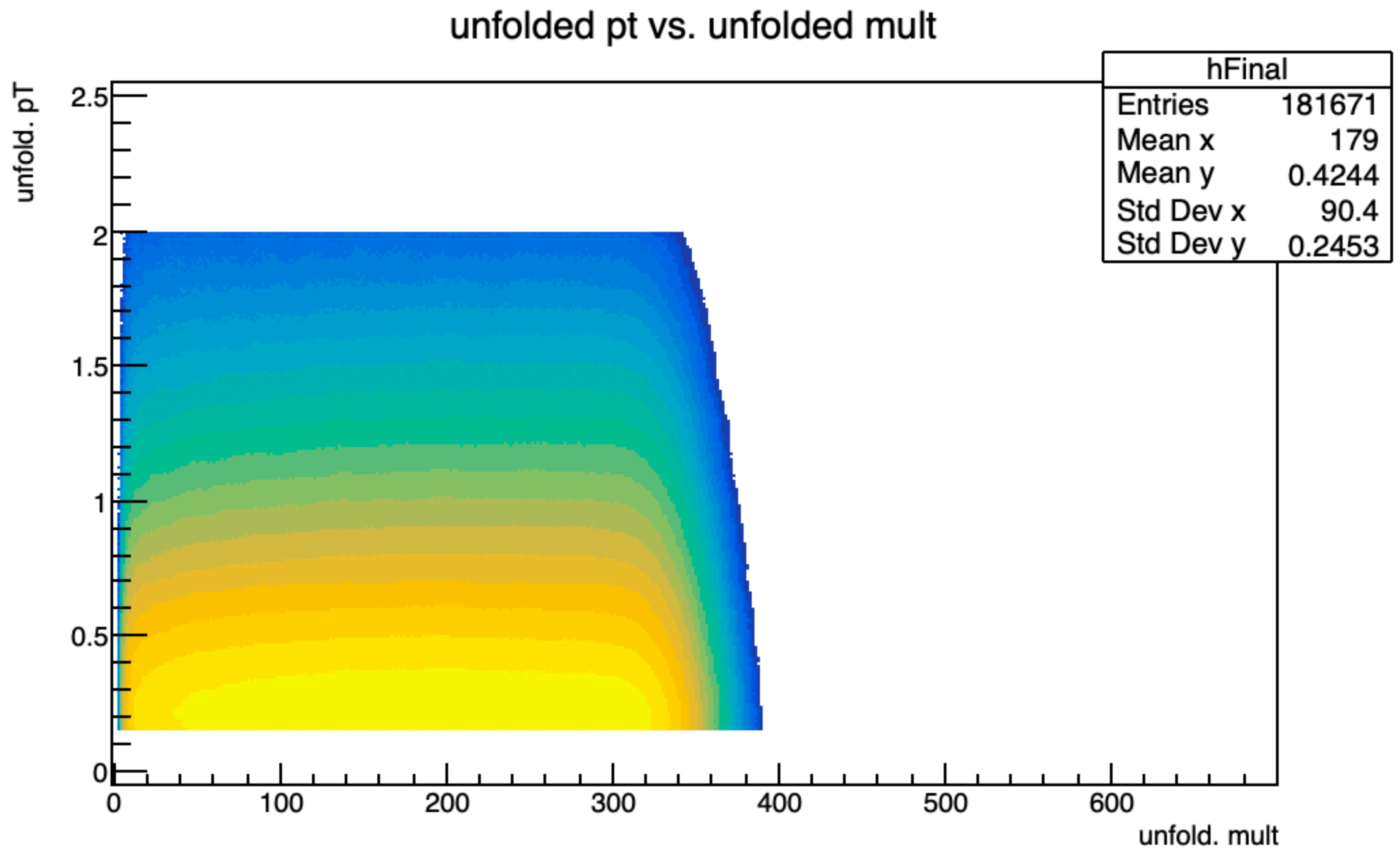


Request26 dataset

what we want to achieve

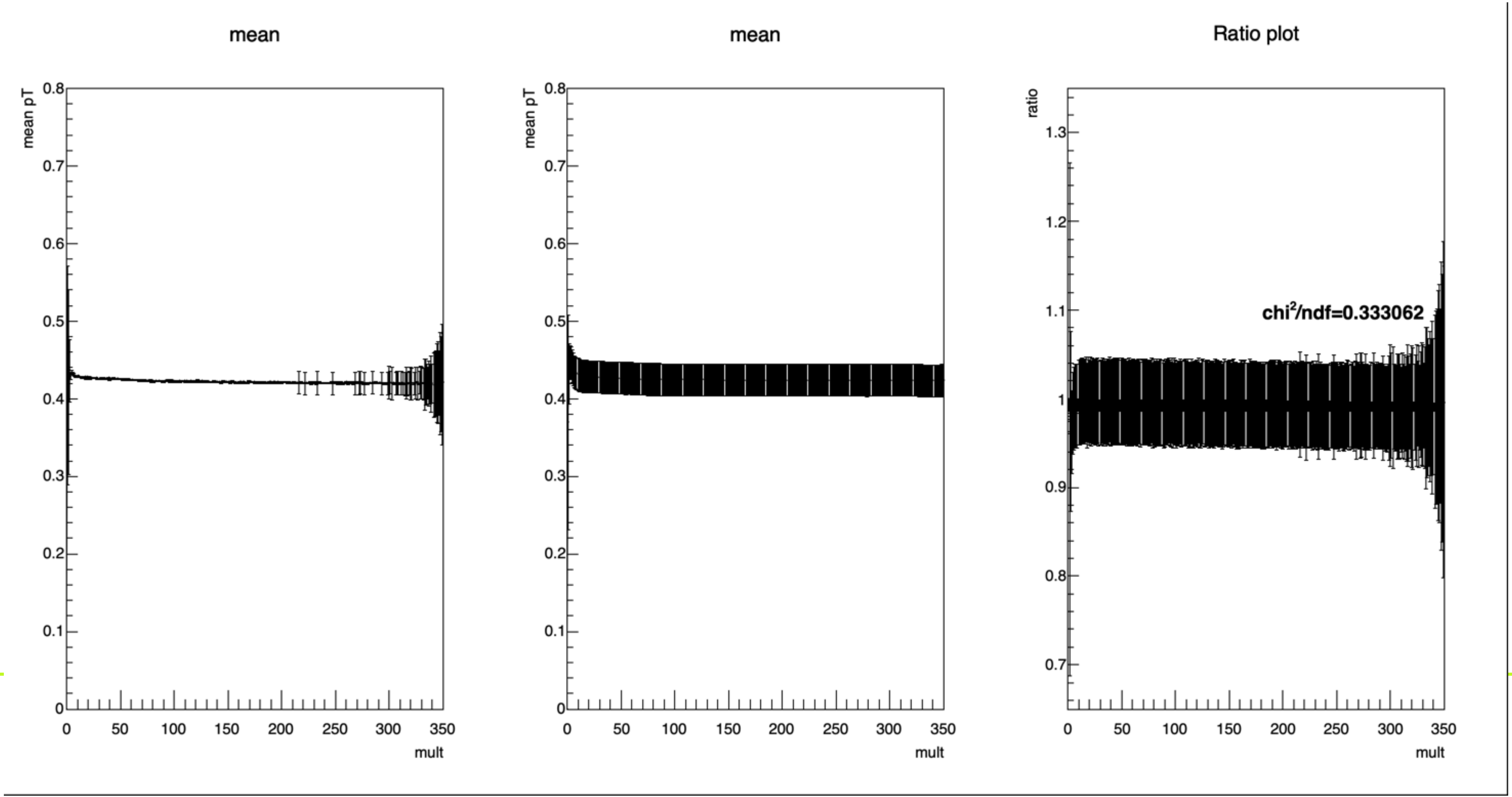


what we get after corrections



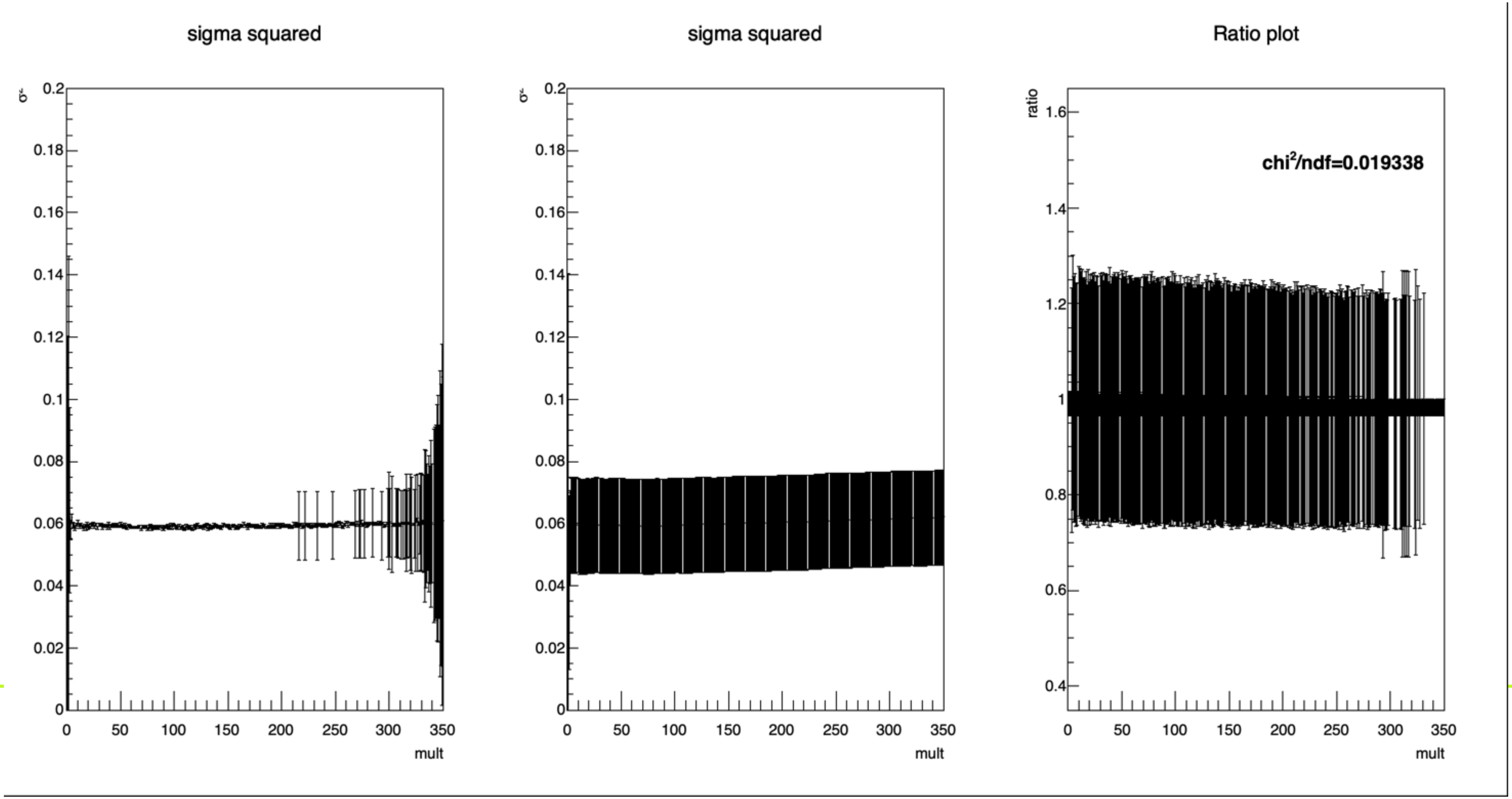
Request26 dataset: first moments

what we want to achieve | what we get after corrections | ratio



Request26 dataset: second moments

what we want to achieve | what we get after corrections | ratio



Artificial inefficiencies

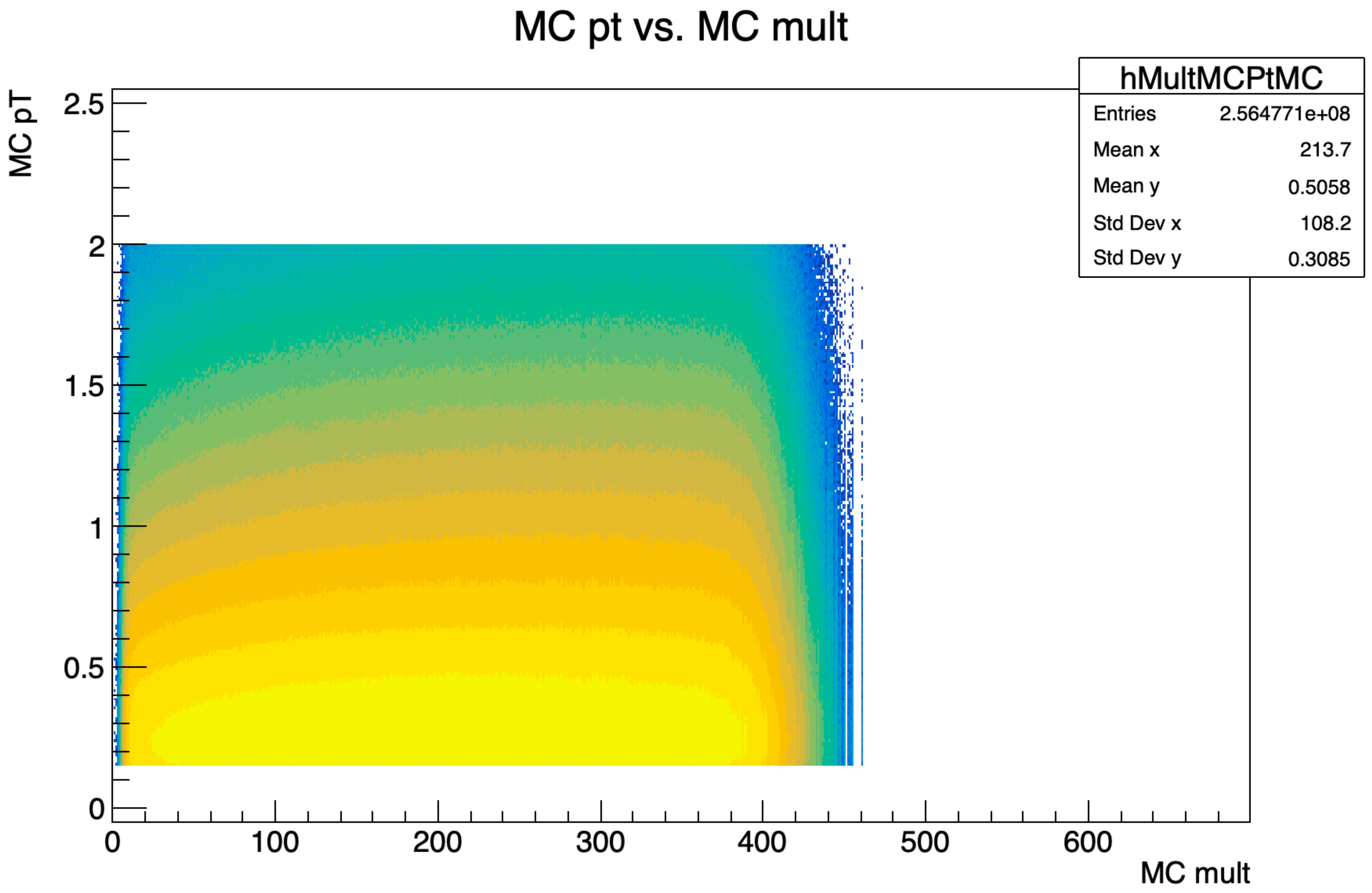
1) At the level of preparation of all ‘response matrices’ by chance remove some fraction of good Rec tracks

2) Remove the same fraction of good Rec tracks from the dataset to be corrected

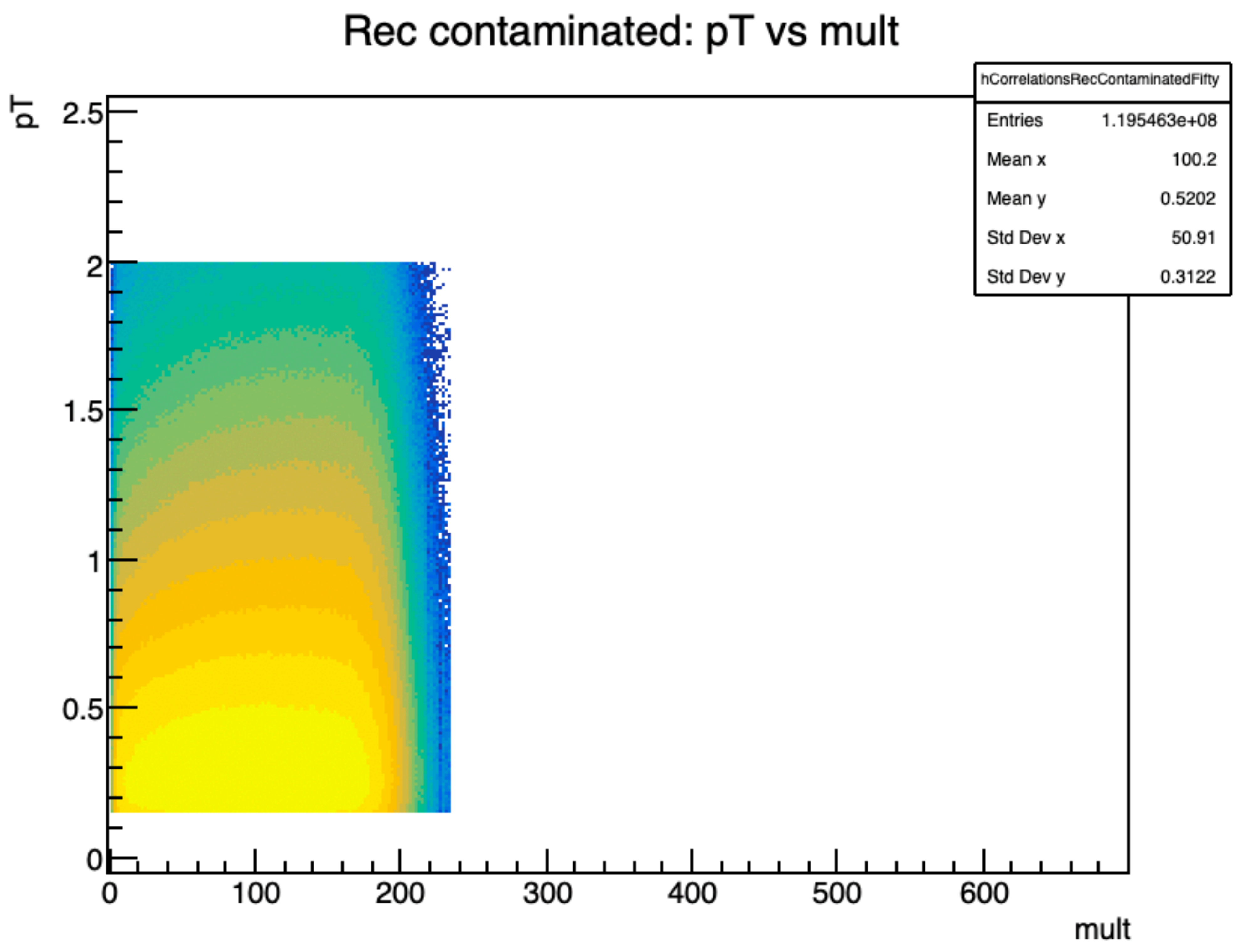
3) Apply corrections, compare results with unviolated case

Test dataset (+50% inefficiency)

what we want to achieve

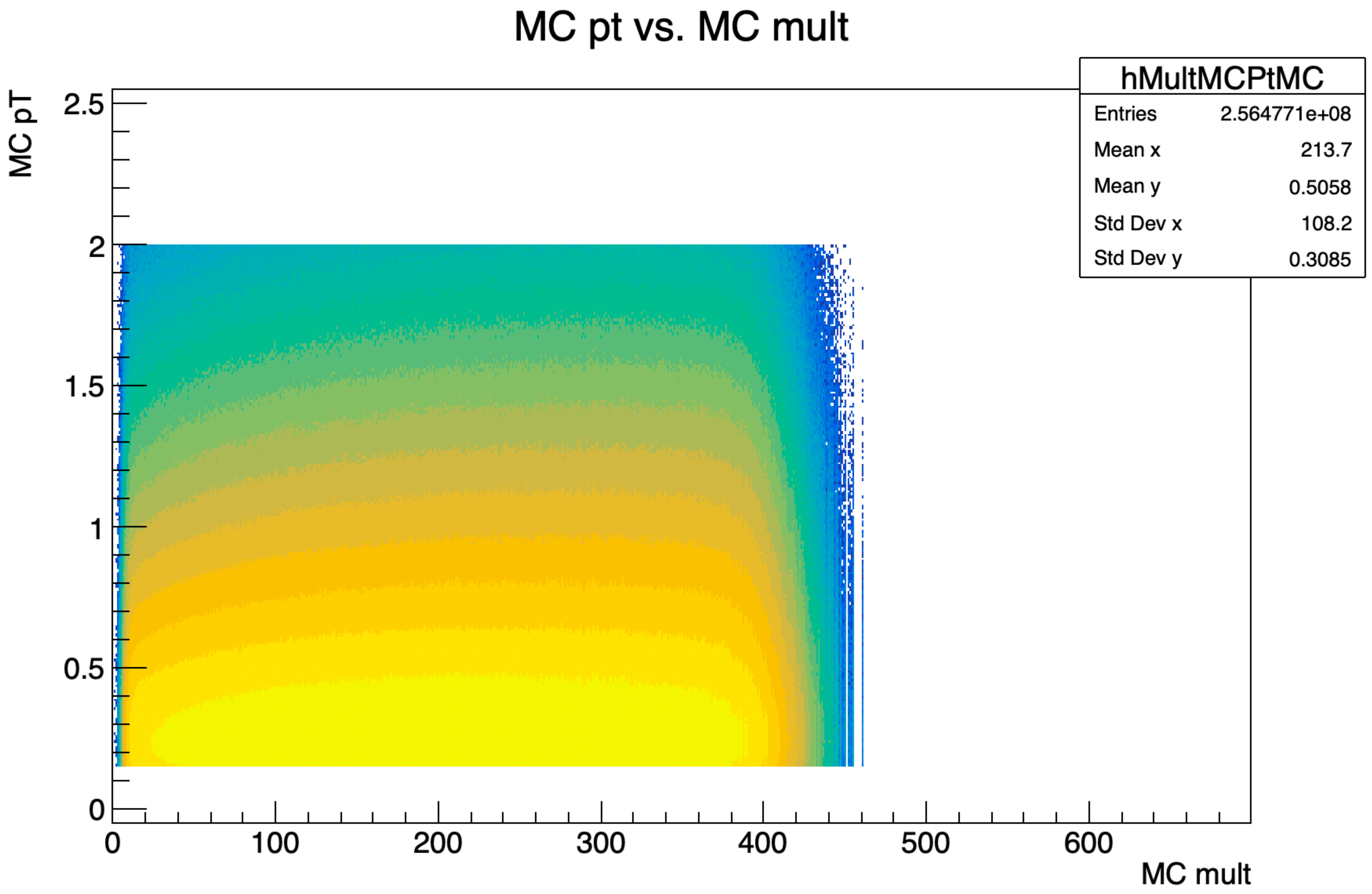


what we have on Rec level

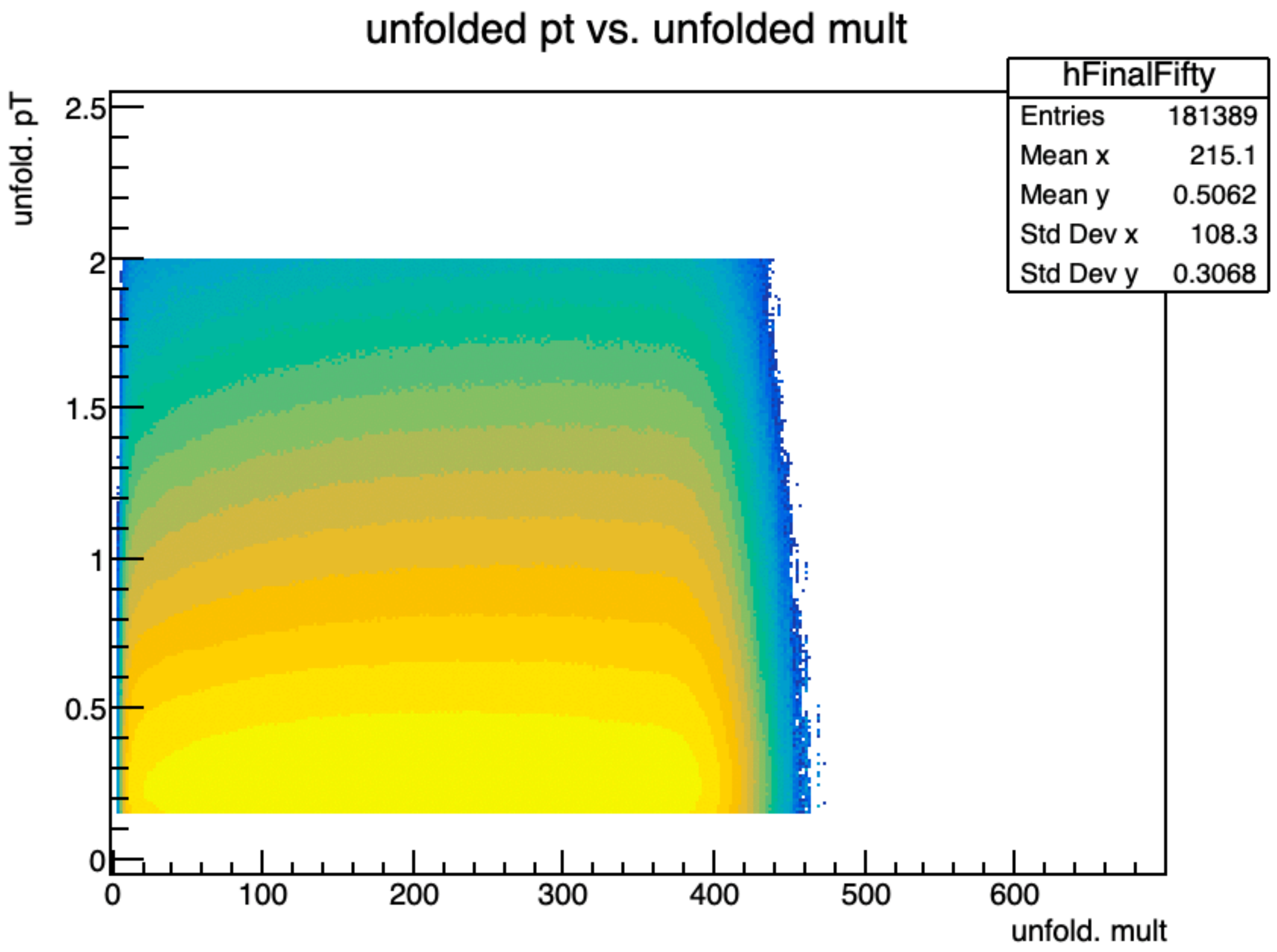


Test dataset (+50% inefficiency)

what we want to achieve

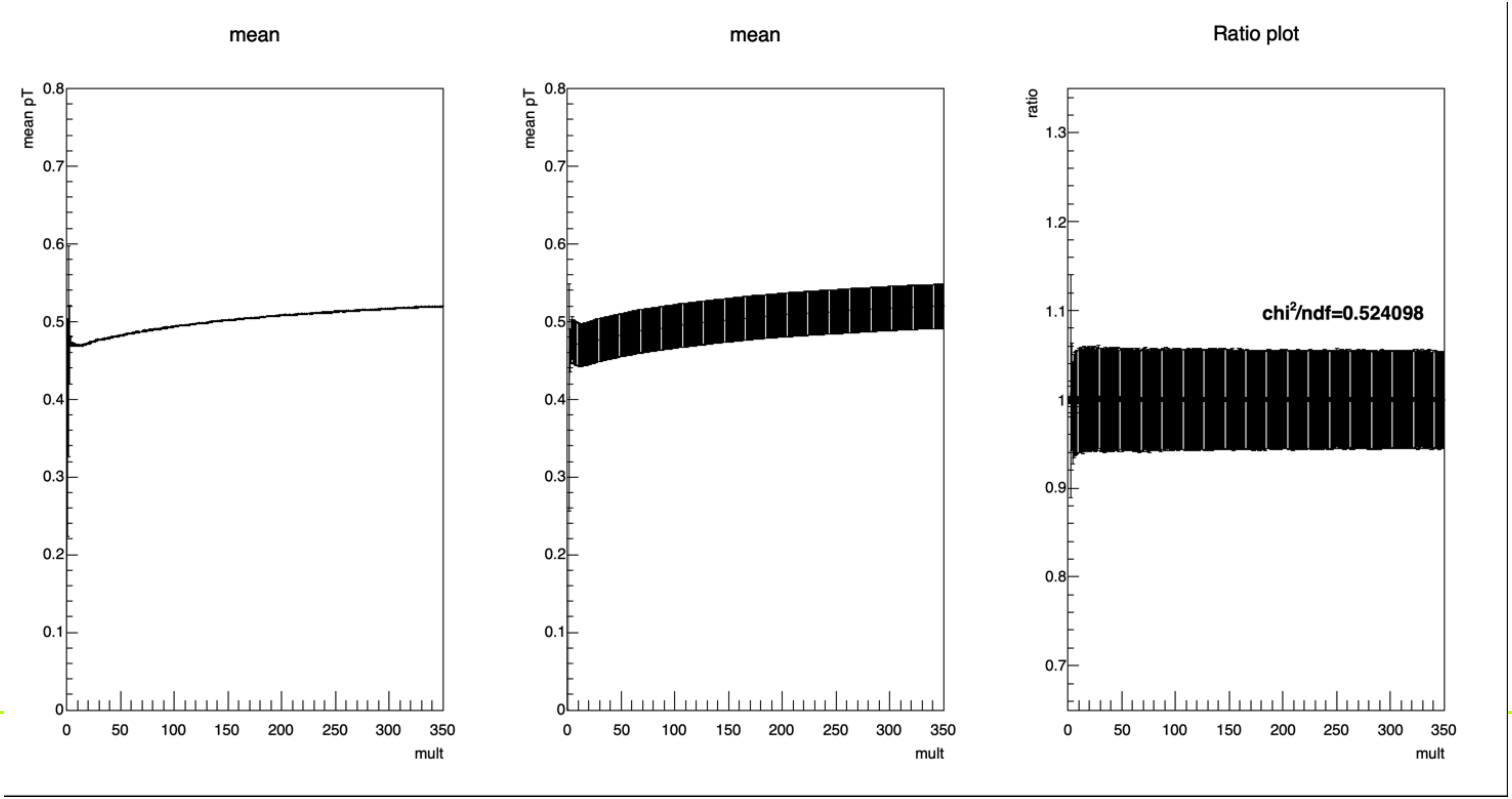


what we get after corrections



Test dataset (+50% inefficiency): first moments

what we want to achieve | what we get after corrections | ratio

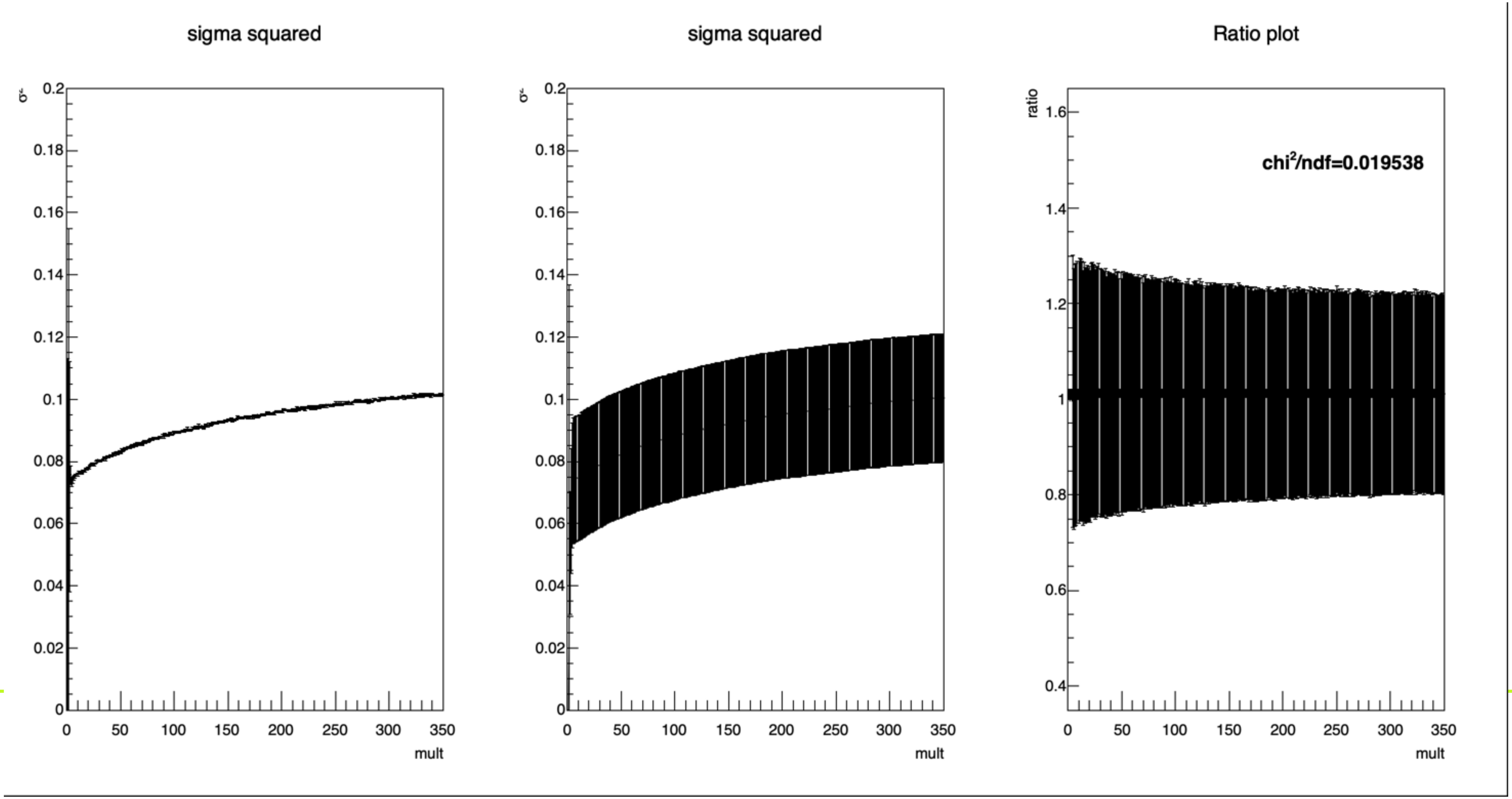


Test dataset (+50% inefficiency): second moments

what we want to achieve

what we get after corrections

ratio



Next steps

Check for other possible sources of systematic uncertainties

vary cuts

vary binning of pT axis and of multiplicity axis

Check other large productions

PHQMD and vHHLE+UrQMD

Thank you for your attention!

e.v.andronov@spbu.ru