Investigation of the MC jet bias to our analysis

Please note that due to personal reasons, I have moved back to the US. Consequently my contract with Uni Heidelberg will end very soon because it does not allow me to work from abroad.

The content of this presentation represents a suggestion for a compromise between the items requested from the last pre-EB meeting and the very limited time possible to finish the publication.

Dhevan Gangadharan
22-2-2021

Main properties of the light-flavor jet PHP MC samples

- Pythia 6.220 in PHP mode with Multiparton interactions switched on
- $\hat{p}_{T,min}$ = 1.9 GeV (regulates IR divergence with a hard cutoff).
- Jet requirements
  - Previously used the “SlowJet” package that comes with PYTHIA 8 to reconstruct massive jets.
  - However, from Sebastian Mergelmeyer's talk 10 years ago, it’s clear that the light-flavor jet generator preselection employed **massless jets**.
  - Now I use the FastJet package (v 3.3.4) to reconstruct massless jets.
    - **Light-flavor jet emulation:**
      At least 1 massless jet with ET > 3 GeV and |y| < 3.0.
    - **Light-flavor dijet emulation:**
      At least 2 massless jets with ET > 4 GeV and |y| < 3.0.

Notes about generator level ZEUS MC Jet reconstruction

- Documentation is quite sparse but some information was found in the original fortran code: /afs/desy.de/group/zeus.zsmsm/ZEUSSysSoft/Released/zeus/Programs/orange/v2013a/src/montecarlo/mcjets.fpp & findjets.fpp

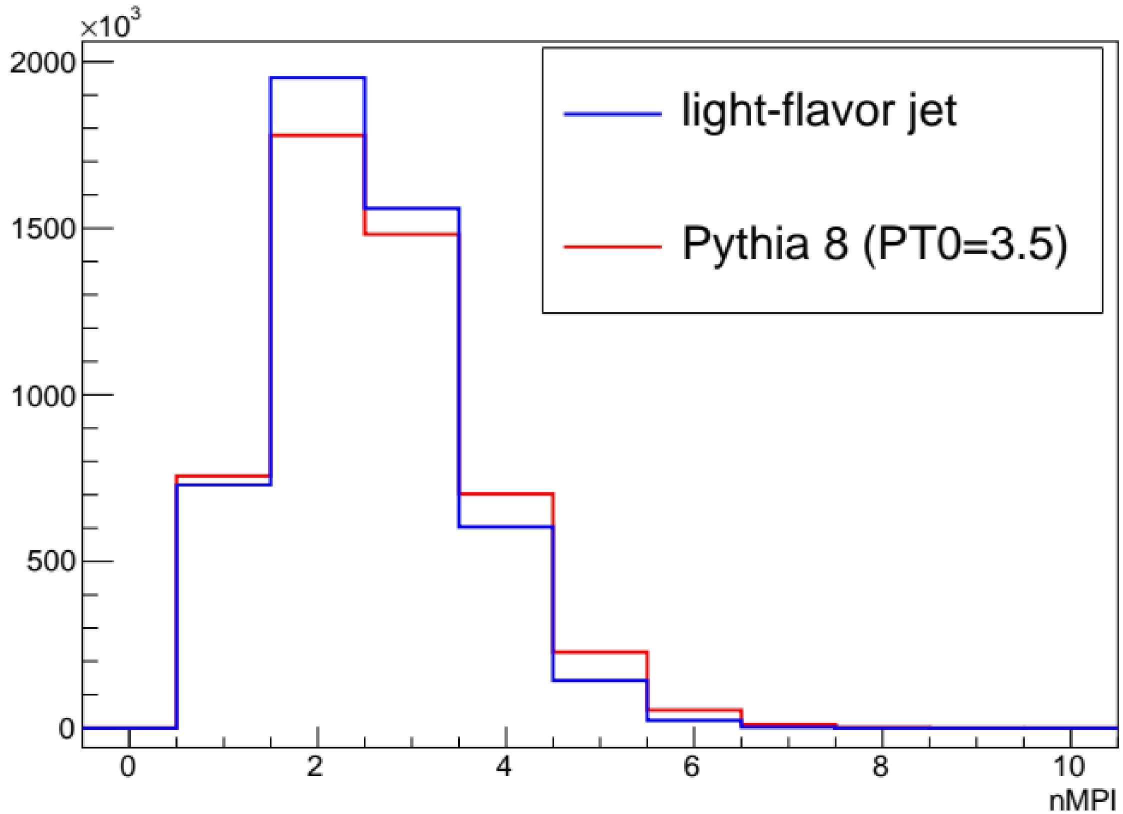
- MC massless hadron jets are stored in orange block MCHJETS (MC Hadron Jets).

- Jets are sorted by their ET in the lab frame and at most 10 jets are saved.

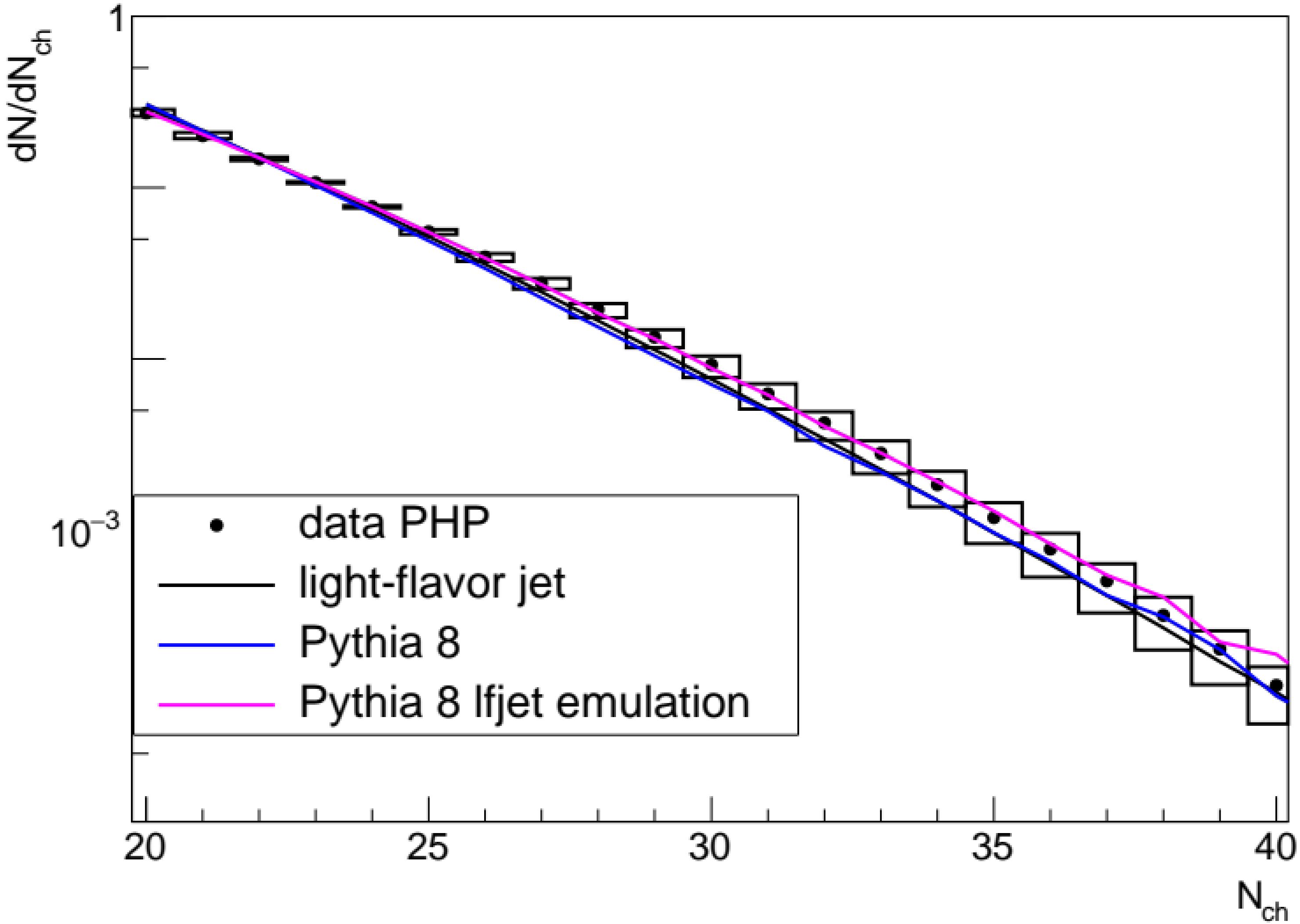
	<u>Jets used for the lfjet emulation</u>	<u>Jets stored in MCHJETS for comparison of distributions</u>
Clustering	Final-state visible particles. Also tried final-state hadrons only	Final-state visible particles. Also tried final-state hadrons only
Recombination scheme	pT scheme (Massless jets)	pT scheme (Massless jets)
|eta| cut	< 3	< 3
ET cut	**> 3 GeV**	**> 1.5 GeV**

MPI setting

- light-flavor jet (Pythia 6) has MPI turned ON but the Pythia 6 MPI framework is quite different from Pythia 8.
- This difference is not fully understood but comparisons of the nMPI distributions lead us to choose PT0 = 3.5 as a reasonable correspondence.

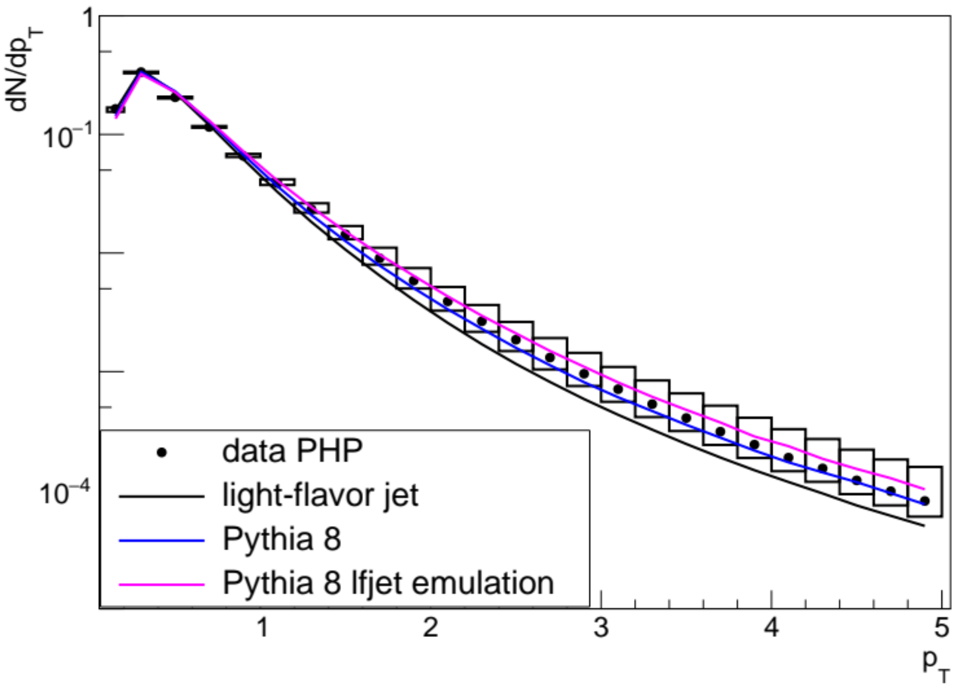


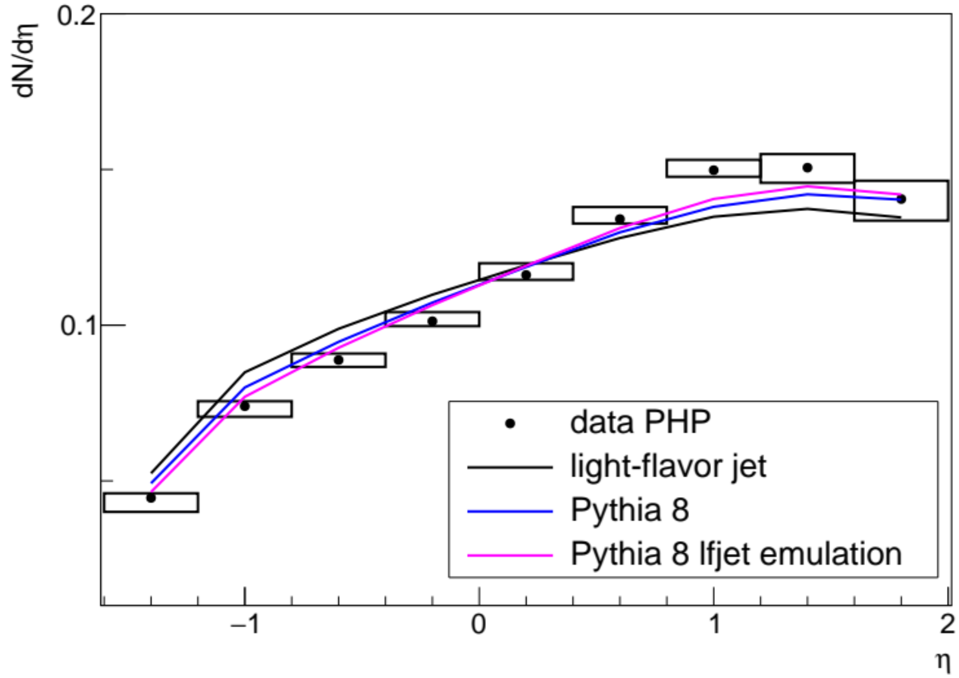
Light-flavor jet compared to Pythia 8 and ZEUS data



- Full systematic uncertainties attached to black points except for the one pertaining to the jet bias of lfjet (to be discussed at end of talk).

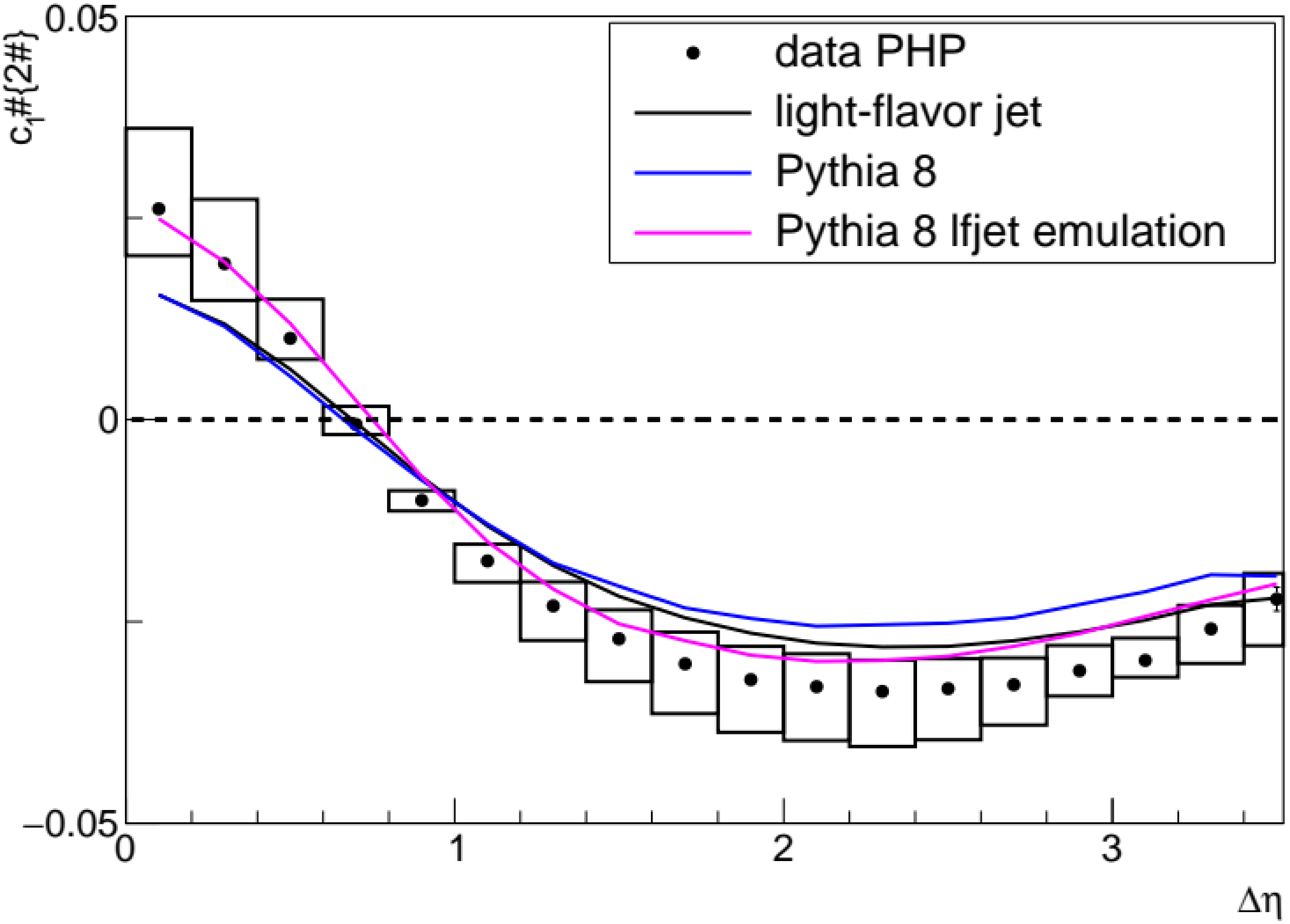
Light-flavor jet compared to Pythia 8 and ZEUS data





- Full systematic uncertainties attached to black points except for the one pertaining to the jet bias of lfjet (to be discussed at end of talk).

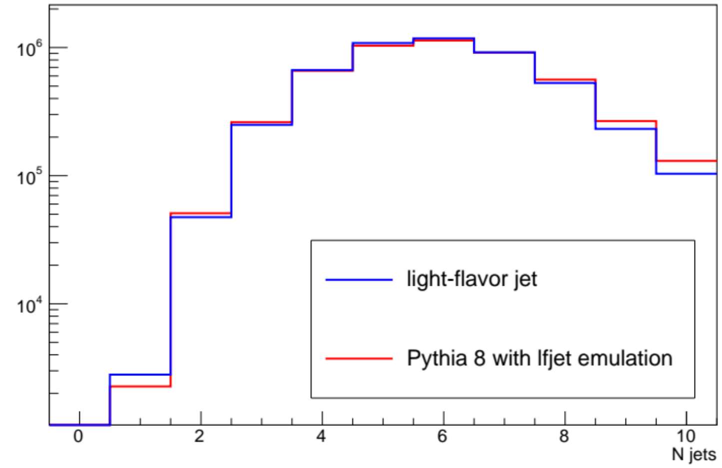
Light-flavor jet compared to Pythia 8 and ZEUS data



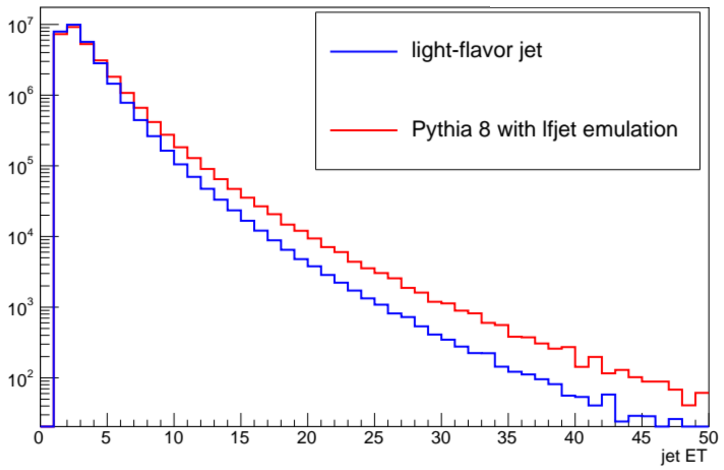
- Full systematic uncertainties attached to black points except for the one pertaining to the jet bias of lfjet (to be discussed at end of talk).

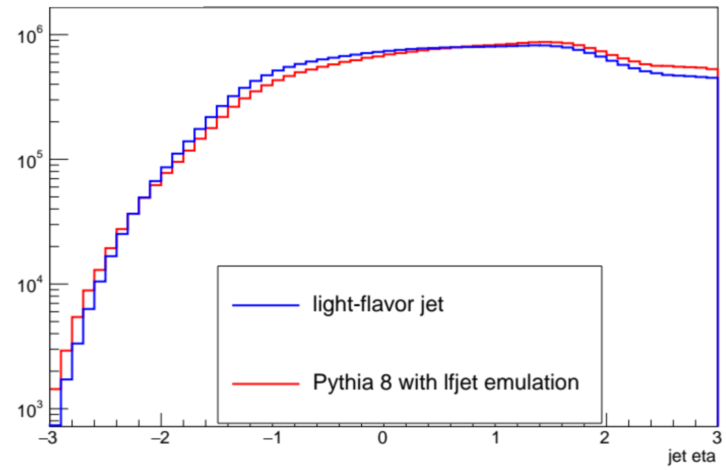
Jet distributions

- Final-state **visible** particles used for jet clustering.



Pythia 8 distribution scaled to match the integral of lfjet

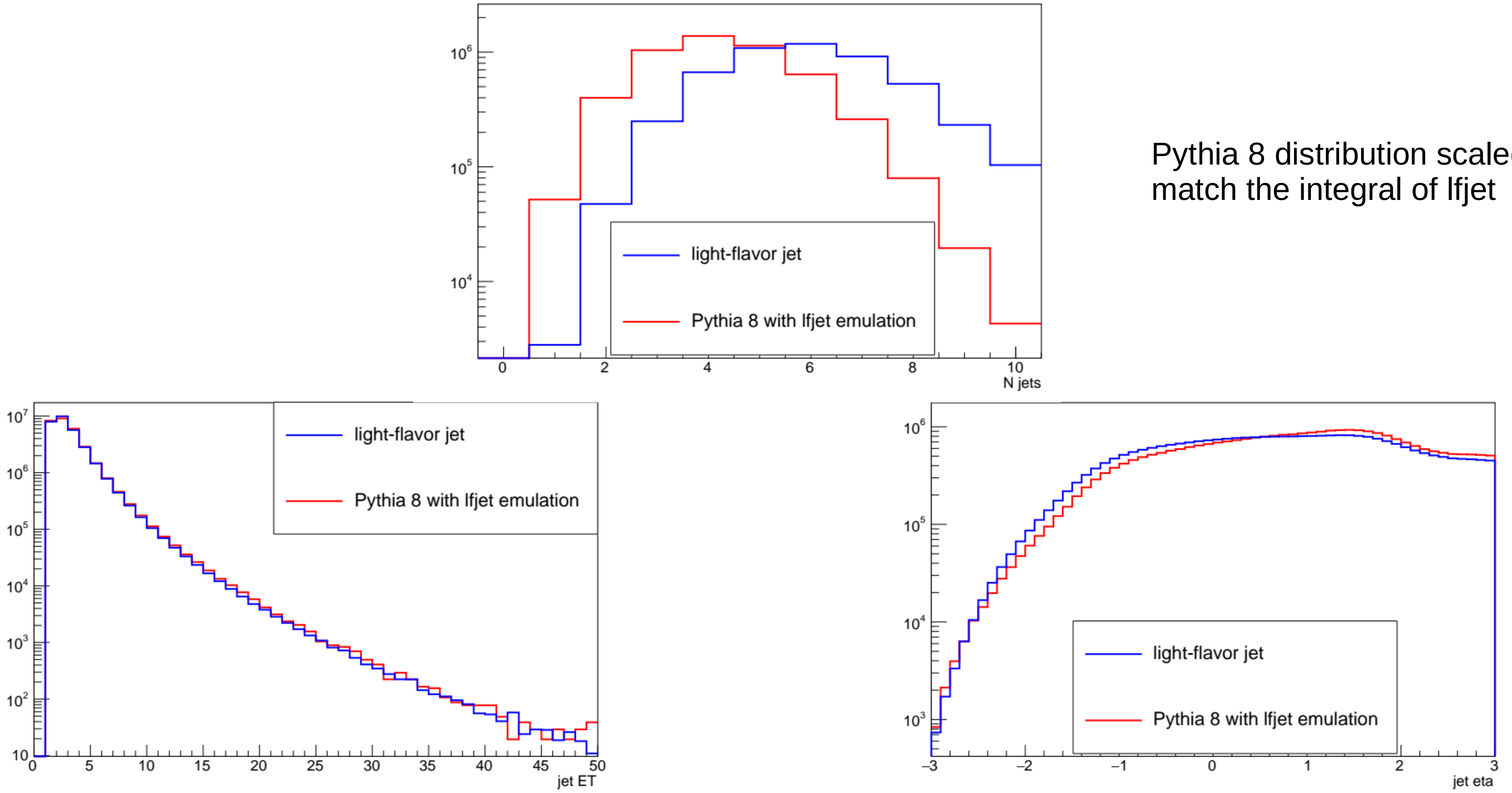




- Light-flavor jet emulation using **visible** particles works well for "N jets" and "jet eta" distributions but not so well for jet ET.
- Maybe there's some feature of the orange ntuple block MCHJETS that I missed or that is undocumented...

Jet distributions

- Final-state **hadrons** used for jet clustering.



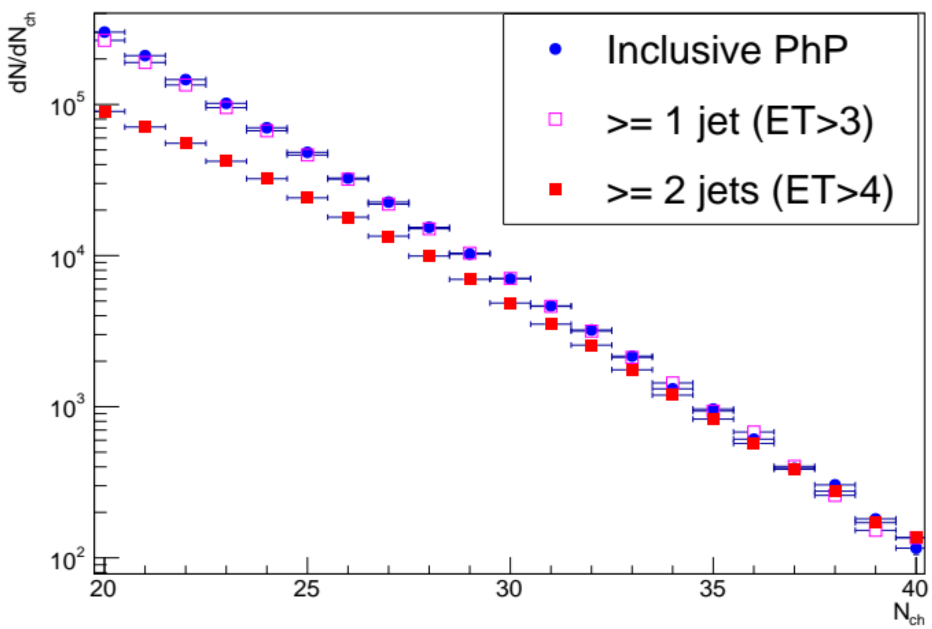
Pythia 8 distribution scaled to match the integral of lfjet

- Light-flavor jet emulation using **hadrons** works well for “jet ET” and “jet eta” distributions but not so well for “N jets”.

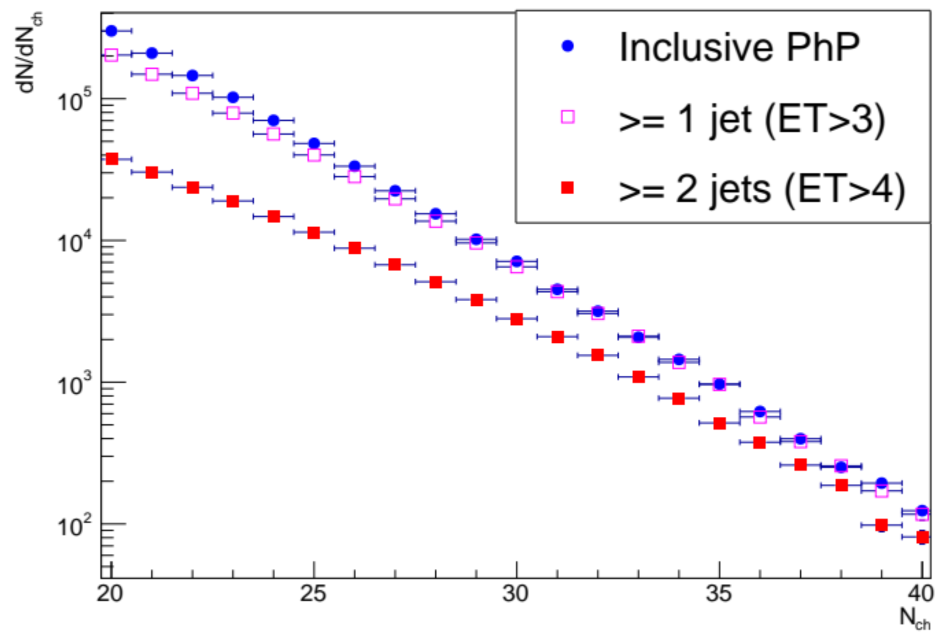
Sequential reduction/alteration of MC phase-space with the jet requirements.

Pythia 8.303
PHP MC

Clustering: visible particles



Clustering: hadrons



1 million inclusive PHP pythia events with Nch >= 20 generated.

8% of events rejected with 1 jet cut (emulates lfjet).

61% of events rejected with 2 jet cut (emulates lfdijet).

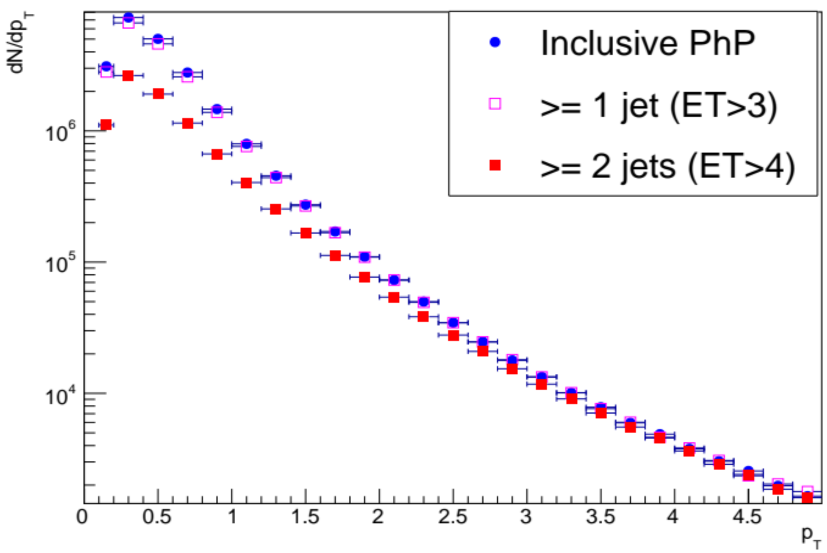
26% of events rejected with 1 jet cut (emulates lfjet).

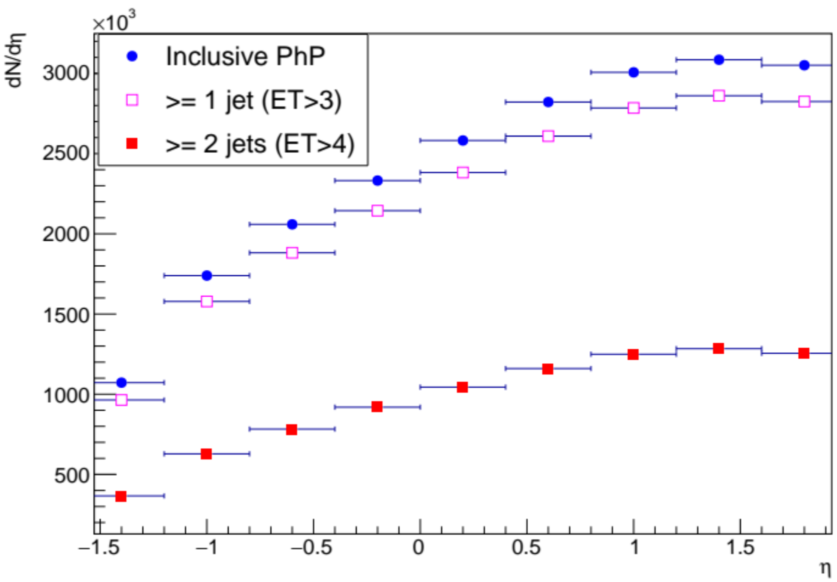
82% of events rejected with 2 jet cut (emulates lfdijet).

Sequential reduction/alteration of MC phase-space with the jet requirements.

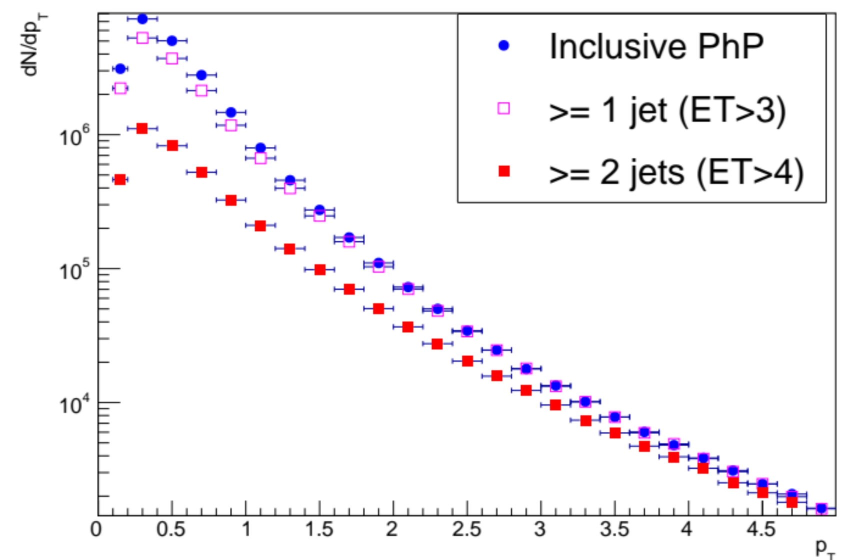
Pythia 8.303
PHP MC

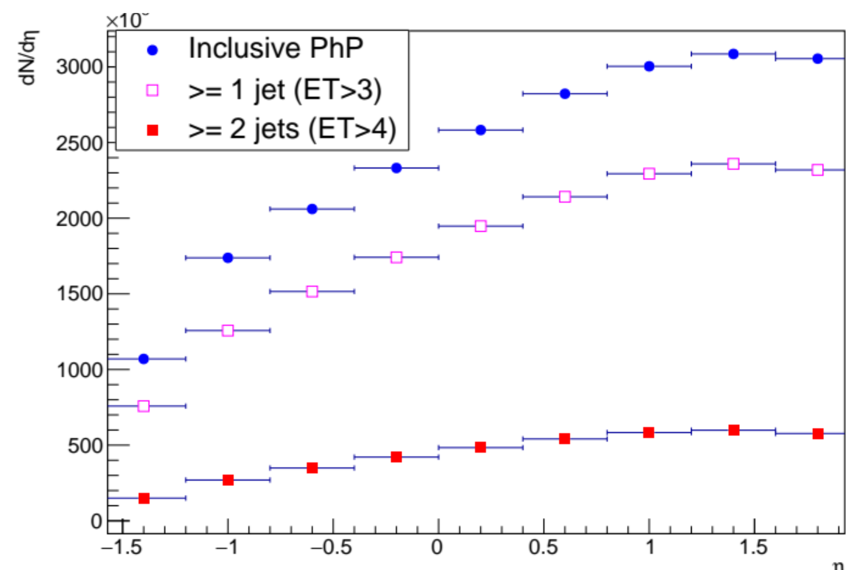
Clustering: visible particles





Clustering: hadrons

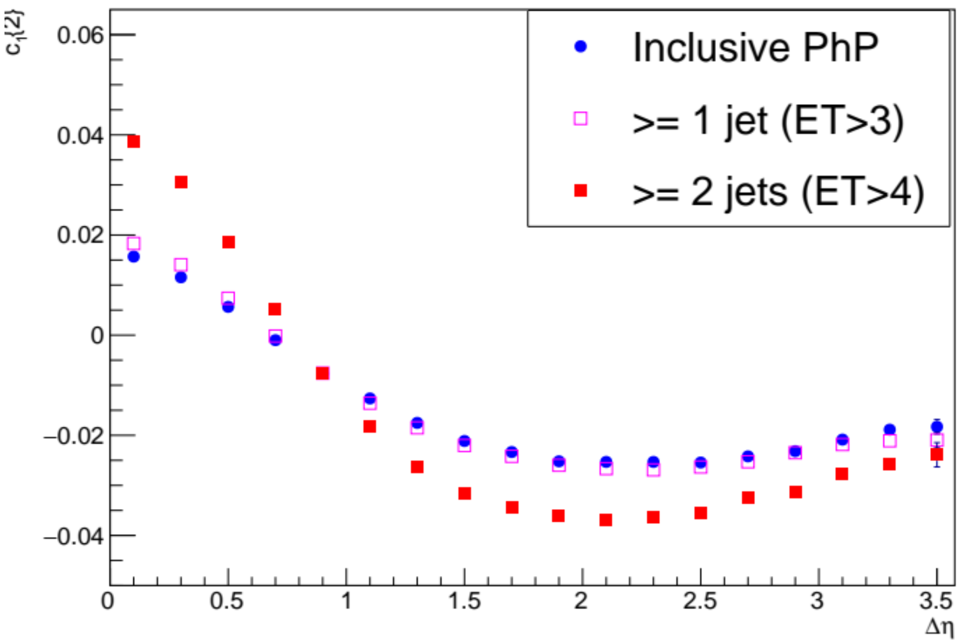




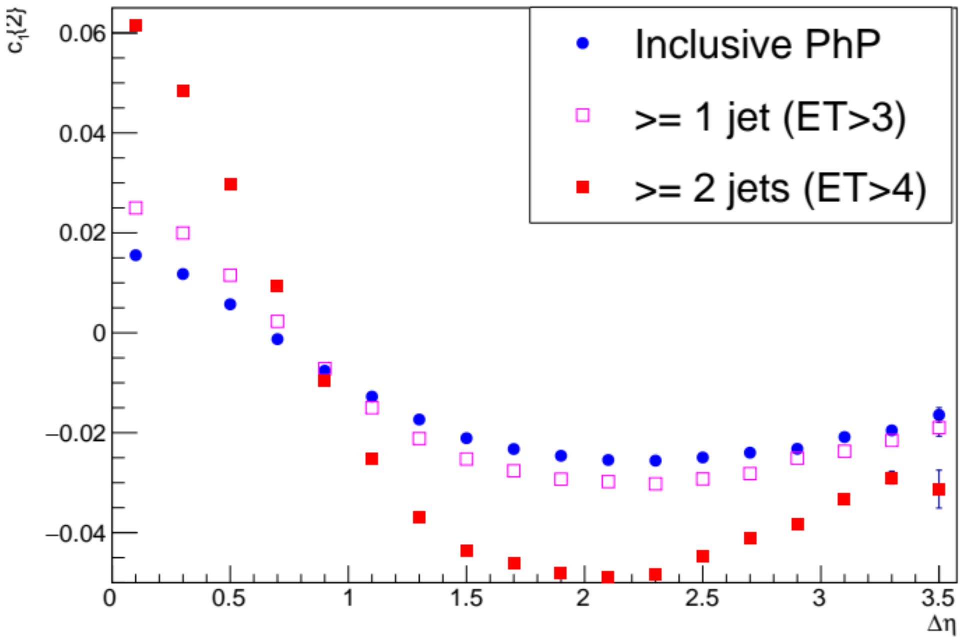
Sequential reduction/alteration of MC phase-space with the jet requirements.

Pythia 8.303
PHP MC

Clustering: visible particles



Clustering: hadrons



Estimating the bias caused by the missing phase-space in our MC

Although the reduction/alteration of MC phase-space with the jet requirements is significant, its effect on our measurements may not be large.

We rely on MC simulations of the ZEUS detector for the following corrections:

1) Tracking efficiency corrections
2) Trigger-bias corrections

We can get a reasonable estimate of the impact of the MC jet bias by comparing the ZEUS results obtained using:

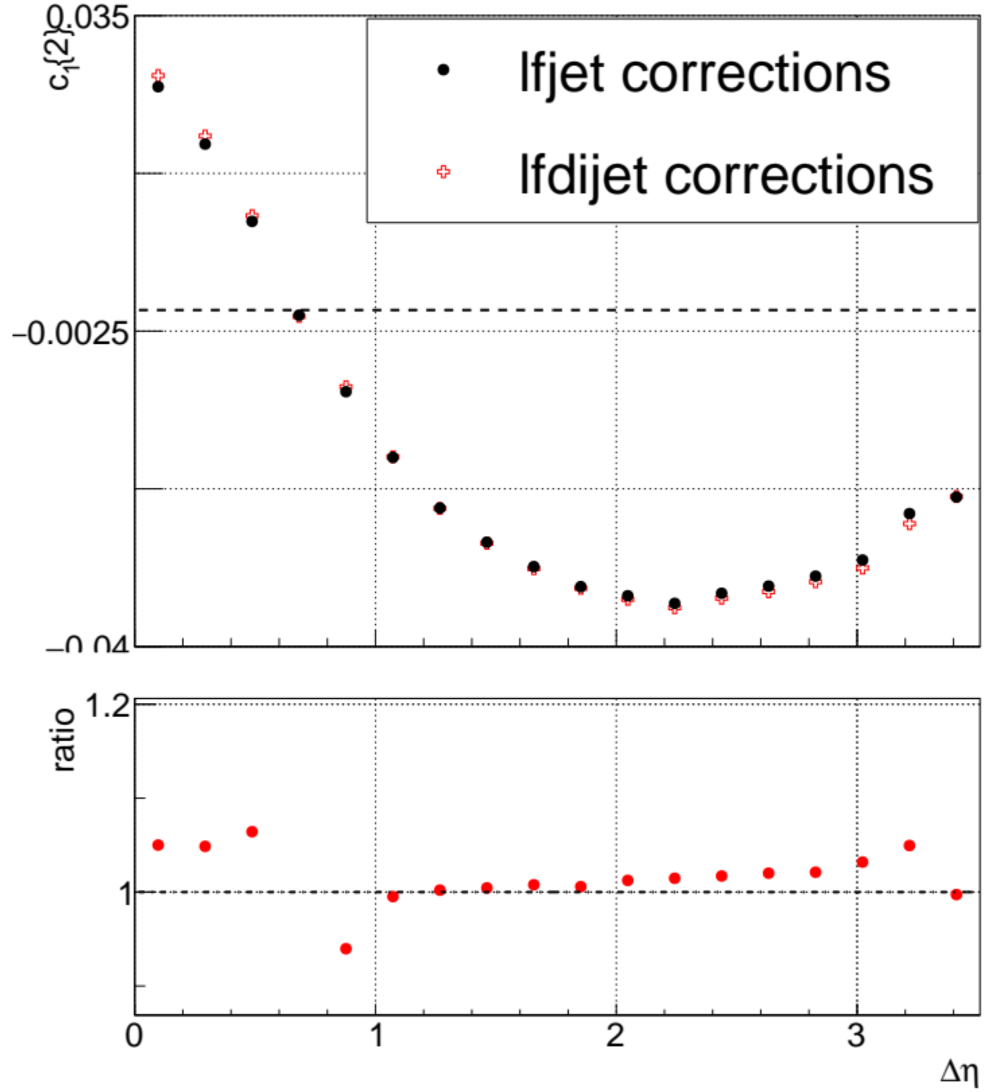
- light-flavor jet correction factors

to those using

- light-flavor **dijet** correction factors

Light-flavor jet vs dijet MC corrections to ZEUS data

ZEUS



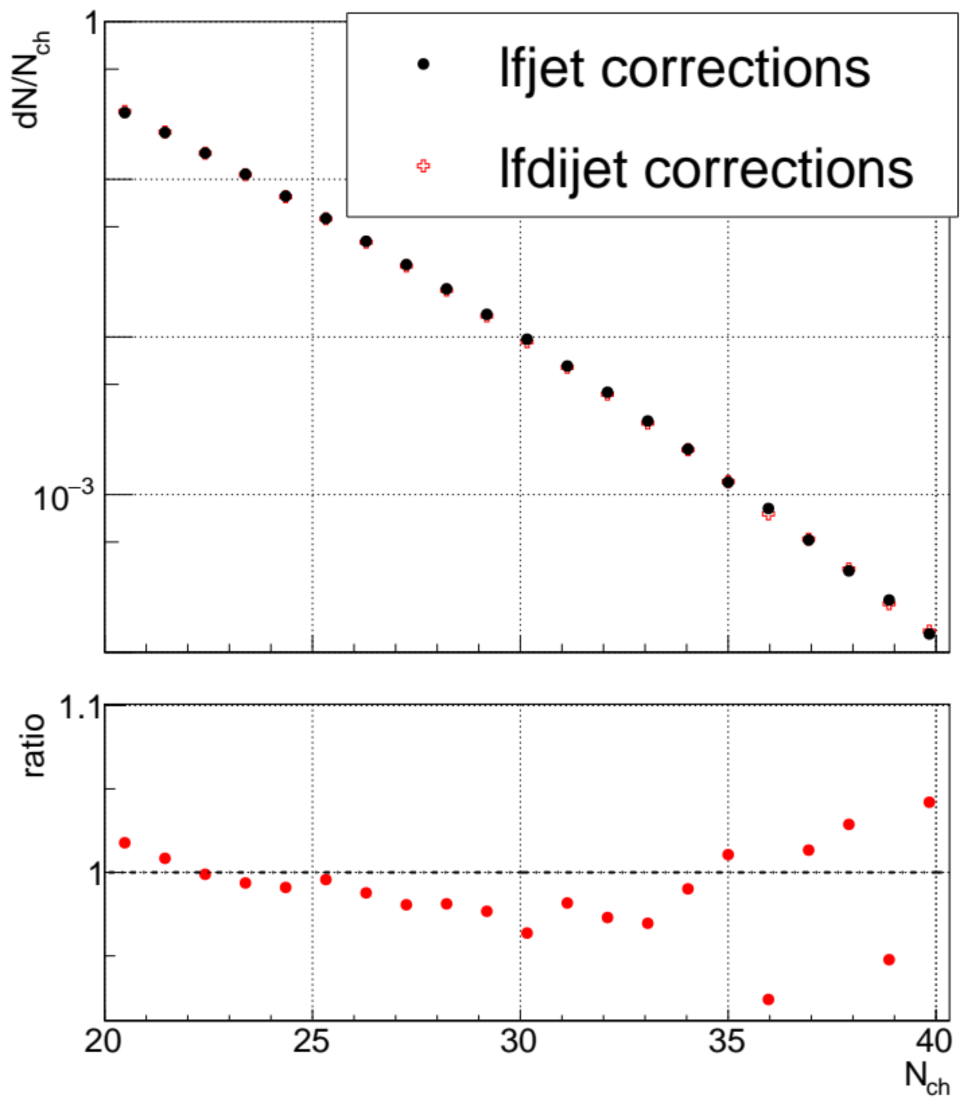
Assesses the effect of missing phase-space between lfjet and lfdijet.

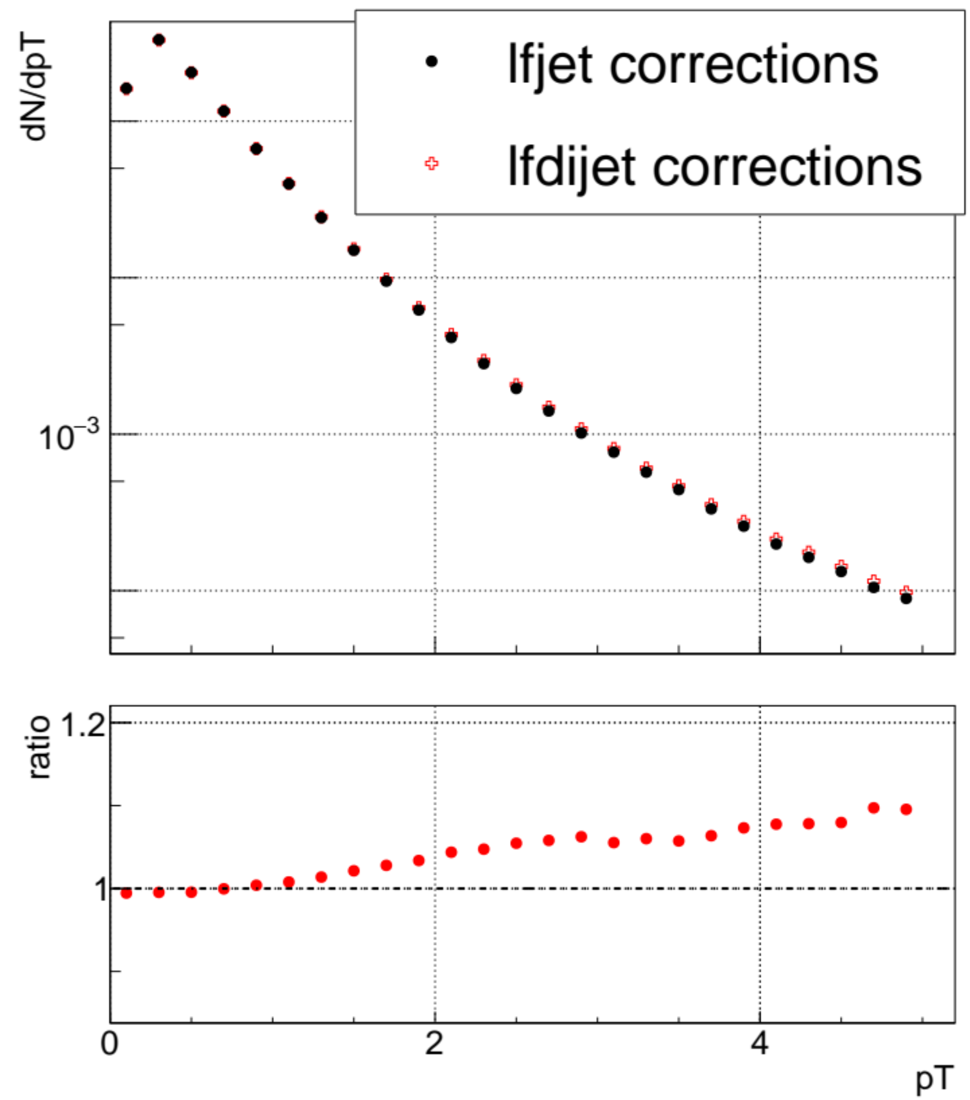
Two-particle correlations vary by <~ 5%.

ZEUS data with light-flavor jet vs dijet MC corrections

ZEUS

N_{ch} and p_T distributions vary by <~ 10%.





Conclusion about the missing phase-space in our MC

The impact of the missing MC phase-space to our measurements is not large (~10%).

Proposal:

- Remove the previous "**black factor**"

$$D_{corrected} = \frac{D_{PhP}}{D_{PhP}^{\geq 1jet}} \frac{D_{lfjet}}{D_{lfjet}^{\mathrm{Trigger}}} w^{(n)} D_{rec}$$

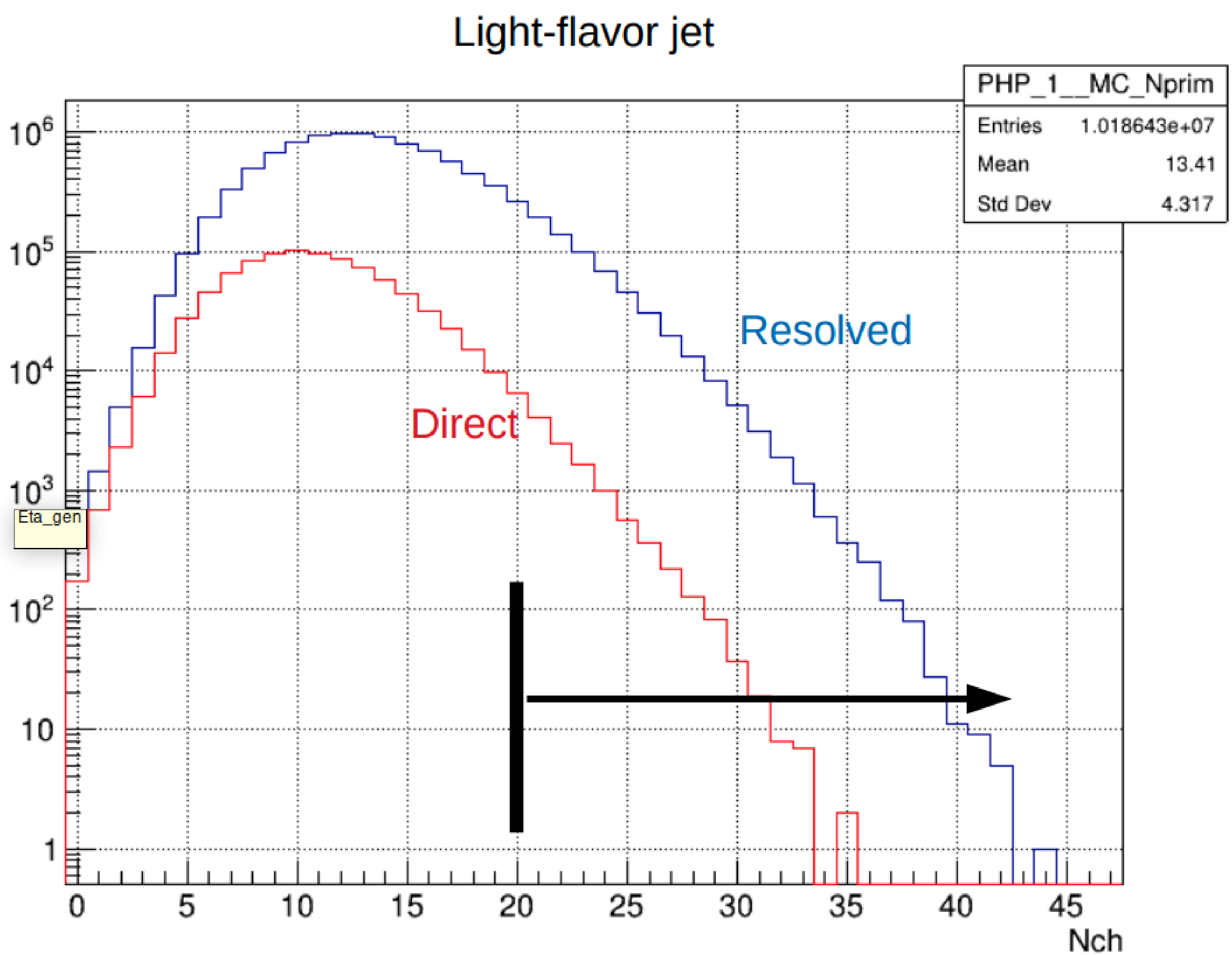
Typically a ~10% correction
I compute this from the Pythia that I generate

Typically a ~30% correction
Extracted from light-flavor jet MC

Typically a ~10% correction
Extracted from light-flavor jet MC

- And assign a systematic uncertainty based on the difference of lfjet vs lfdijet corrected results. Should it be symmetrized?
- The previous systematic assigned to the black factor was the factor itself, ~10%, and the new systematic is also ~10%.

Fraction of direct PhP events at high multiplicity



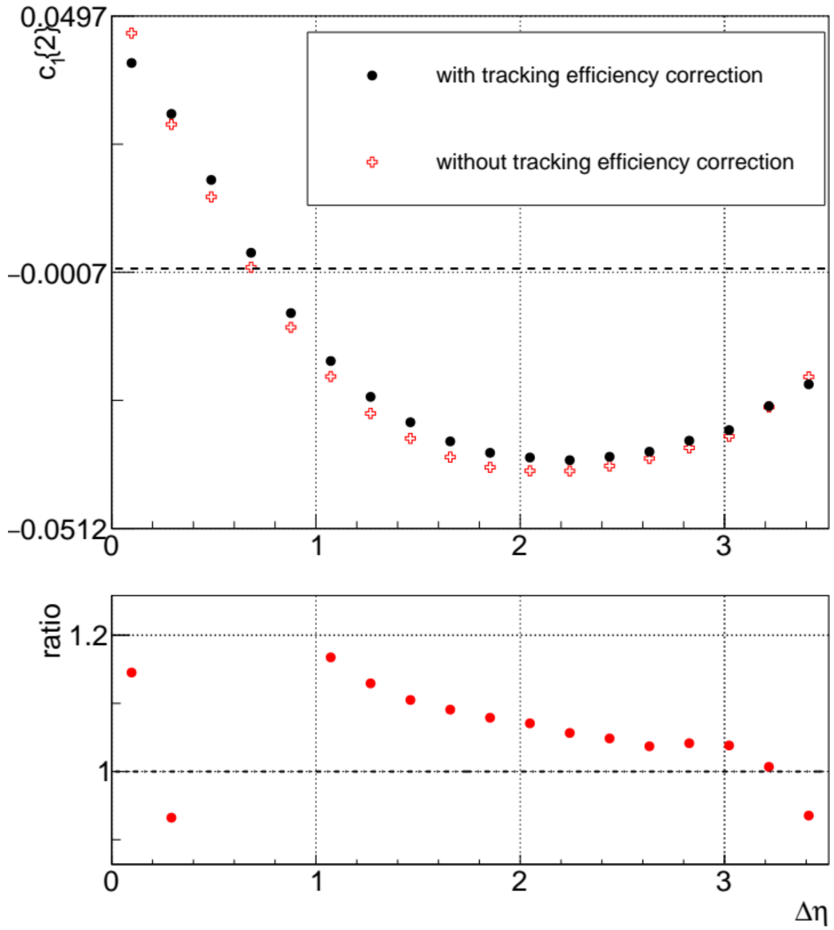
- The calculated fraction of direct events generated for N_{ch} >= 20 is 1.5%

Backup

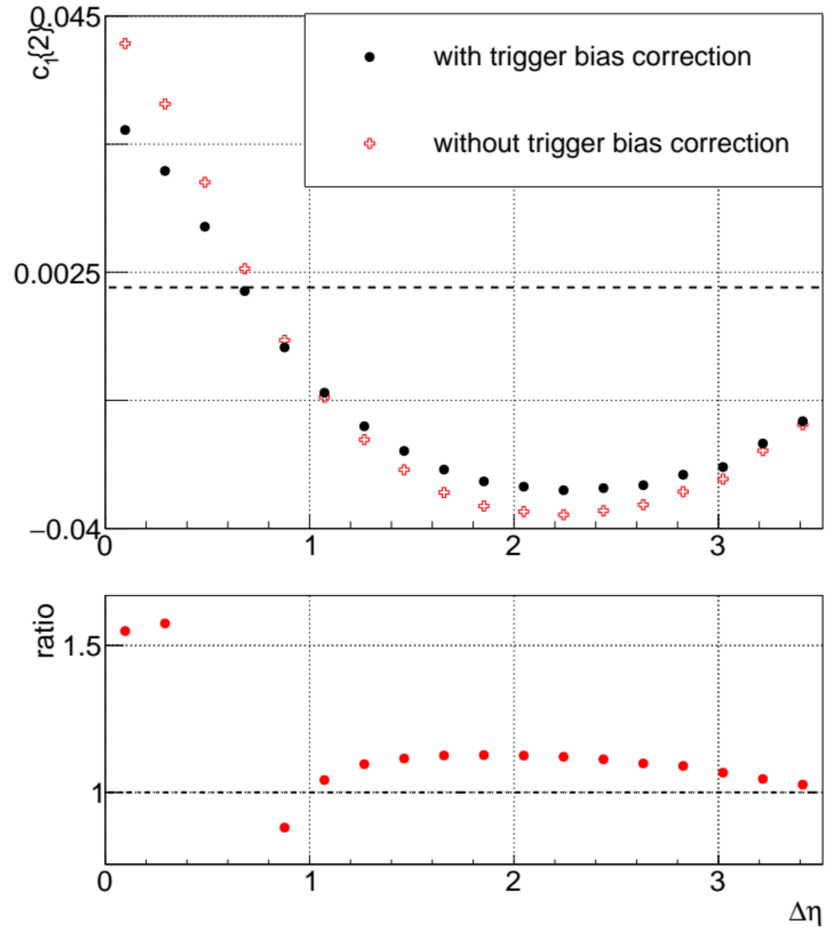
Effect of MC correction factors to the ZEUS data

ZEUS

1) Tracking efficiency corrections



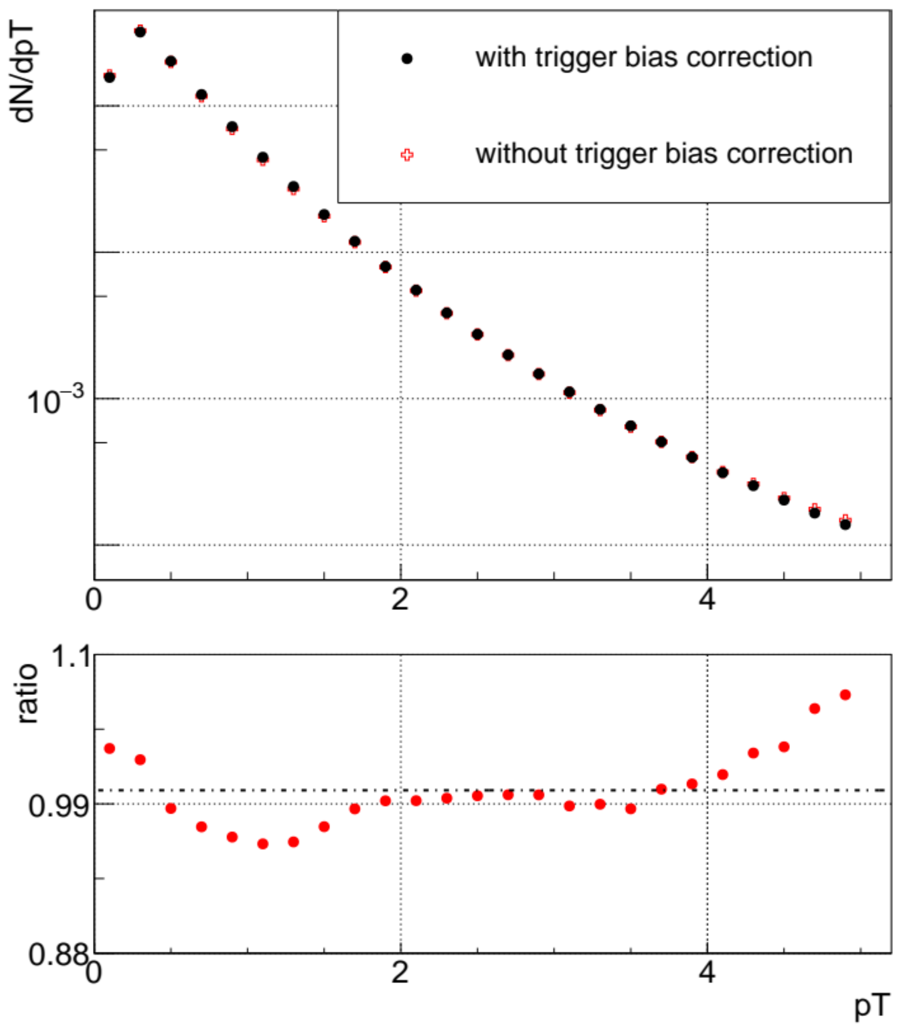
2) Trigger bias corrections



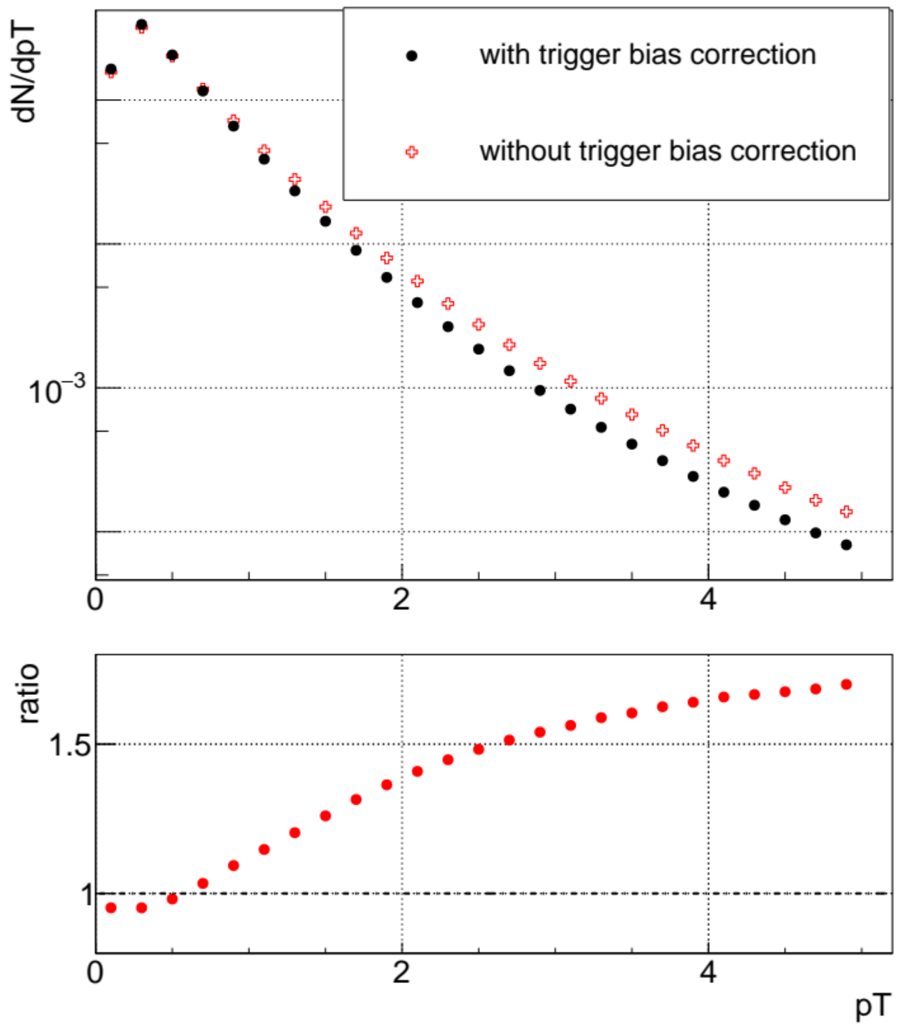
Effect of MC correction factors to the ZEUS data

ZEUS

1) Tracking efficiency corrections



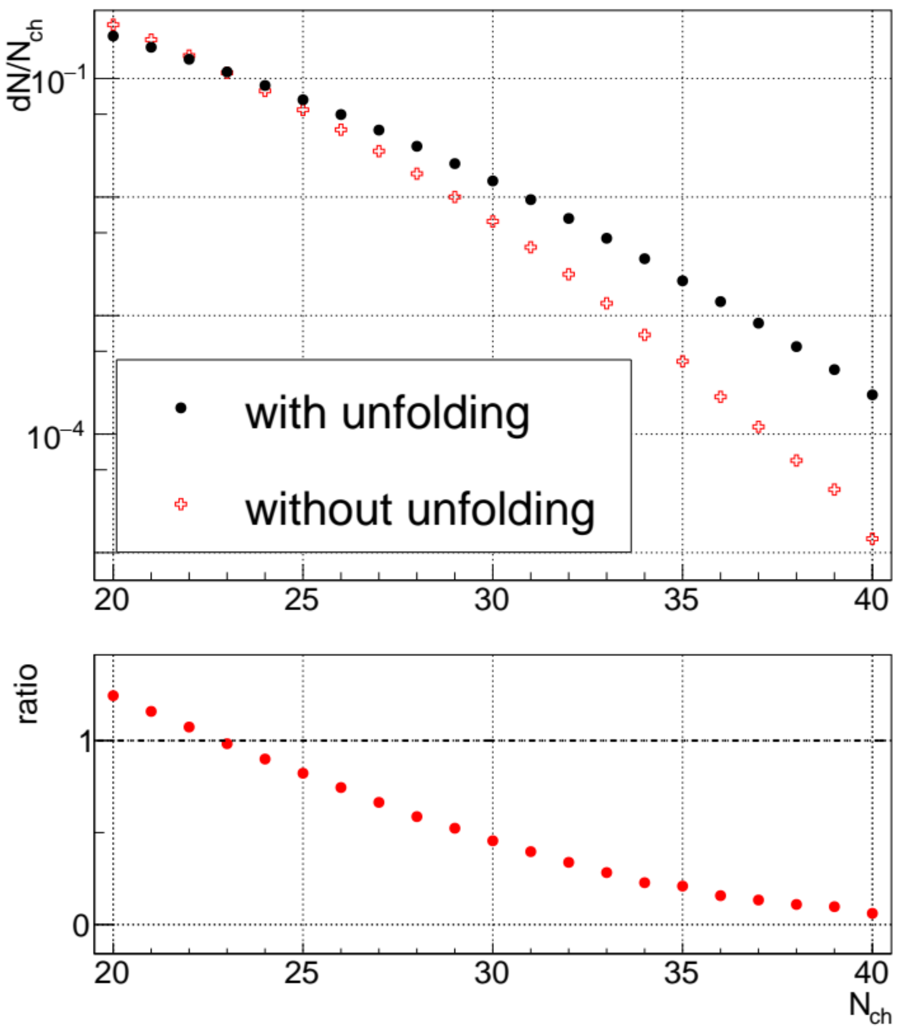
2) Trigger bias corrections



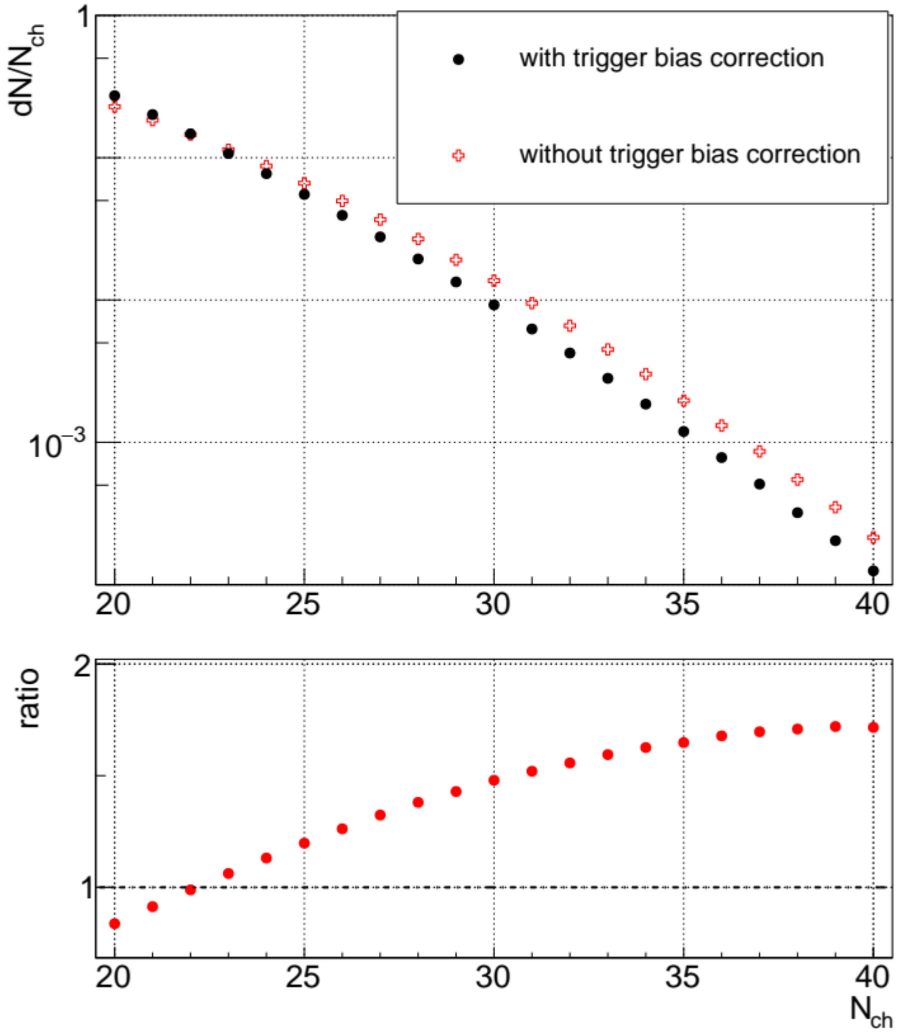
Effect of MC correction factors to the ZEUS data

ZEUS

1) Tracking efficiency corrections



2) Trigger bias corrections



About the Inclusive Dijet PhP MC

Important MC not only for HFL, but also for prompt photon in PhP.

Generator-level cuts in the HFL group's inclusive dijet MC:

- *Preselection in AMADEUS*
  2 jets with $|\eta| < 3$, $E_{\mathrm{T}} > 4$ GeV in hadronic final state (massless)
  indicated by `.JJ.ET4.` in the funnel name!
- *Phase space cut in PYTHIA*
  $\hat{P}_{\mathrm{T}} > 2$ GeV

In May I showed their limited efficiencies and how to compute them.

In June prompt photon people reported a large amount of events was "missing" in the MC, especially with their very low 1-jet E_{T}-cut.

Expected Sizes

New preselection

- $\hat{P}_T > 1.9$ GeV
- 1 jet with $|\eta| < 3$ and $E_T > 3$ GeV

- 1008 mio. resolved events
- 85 mio. Direct events