Les Houches flavour studies

Giovanni Stagnitto
(on behalf of the flavour gang at the PhysTeV 2023 workshop)

Flavoured Jets at the LHC, IPPP Durham, 11-12.06.2024

LES HOUCHES 2023: PHYSICS AT TEV COLLIDERS

STANDARD MODEL WORKING GROUP REPORT

Conveners

Higgs Physics:
M. Donegà (CMS), S. Jones (Theory), K. Köneke (ATLAS), R. Röntsch (Theory)

Standard Model Phenomenology:
P. Azzurri (CMS), A. Hinzmann (CMS, Jet-Substructure subgroup), A. Huss (Theory), J. Huston (ATLAS), S. Marzani (Theory, Jet-Substructure subgroup), M. Pellen (Theory)

Monte Carlo generators, Tools, Machine-Learning:
S. Höche (Theory), J. McFayden (ATLAS), V. Mikuni (ML contact), S. Plätzer (Theory)

Authors and Participants

J. Andersen[1], B. Assi[2], K. Asteriadis[3], P. Azzurri[4a], G. Barone[5], A. Behring[6], A. Benecke[7], S. Bhattacharya[8], E. Bothmann[9], S. Caletti[10], X. Chen[11], M. Chiesa[12], A. Cooper-Sarkar[13], T. Cridge[14], A. Cueto Gomez[15], S. Datta[16], P. K. Dhani[17], M. Donegà[18], T. Engel[19], S. Ferrario Ravasio[6], S. Forte[20a,20b], P. Francavilla[4a,4b], M.V. Garzelli[21], A. Ghira[22a,22b], A. Ghosh[23,24], F. Giuli[25a,25b], L. Gouskos[5], P. Gras[26], C. Gütschow[27], Y. Haddad[28], L. Harland-Lang[27], F. Hekhorn[29,20a,20b], I. Helenius[29], A. Hinzmann[14], S. Höche[2], J. Holguin[30,31], A. Huss[6], J. Huston[32], T. Ježo[33], S. Jones[1], S. Kiebacher[34], M. Knobbe[9], R. Kogler[14], K. Köneke[19], L. Kunz[34], M. LeBlanc[5,30], P. Loch[35], G. Loeschcke Centeno[36], M. Löschner[14], A. Maas[37], G. Magni[38], A. Maier[14], M. Marcoli[1], S. Marzani[22a,22b], J. McFayden[36], P. Meinzinger[1], V. Mikuni[39], S. Moch[21], P. Nadolsky[40], D. Napoletano[41a,41b], M. Pellen[19], S. Plätzer[37,42], R. Poncelet[43], C. Preuss[44], H. Qu[45], K. Rabbertz[46], D. Reichelt[1], A. Rescia[14,22a], J. Roloff[5], R. Röntsch[20a,20b], S. Sanchez Cruz[45], T. Sarkar[4a], L. Scyboz[47,48], F. Sforza[22a,22b], A. Siódmok[49], G. Stagnitto[41a,41b], A. Tarek[32], R.S. Thorne[27], A. Valassi[50], J. Whitehead[43], J. Winter[51]

+ others (E. Lesser, G. Salam, …)

Contents

Flavoured jets hot topic in LH2023!
A dedicated paper in preparation

A lot of discussions (and headaches) around flavoured jets…

```
PART 1
Les Houches, there's a place you can go
I said, Les Houches, when you're short on ideas
You can stay there, and I'm sure you will find
Many ways to have a good time!

PART 2
Les Houches, where we all can discuss
I said, Les Houches, in the beautiful Alps
You can stay there, and then try to combine
truth and reco, then go drink wine!

CHORUS A
We want jet's flavour to be I. R. C. Safe
We want jet's flavour to be I. R. C. Safe
Cannot just count the b's, of an anti-kt
because even a soft gluon splits

PART 1
Les Houches, there's a place you can go
I said, Les Houches, when you're short on ideas
You can stay there, and I'm sure you will find
Many ways to have a good time!

PART 2
Les Houches, where we all can discuss
I said, Les Houches, in the beautiful Alps
You can stay there, and then try to combine
truth and reco, then go drink wine!

CHORUS B
We should reduce the ne-ga-tive weights
and compute processes at N3LO
This is our wishlist, to make Joey in peace
and then make the best of stats

CHORUS A x2
We want jet's flavour to be I. R. C. Safe
We want jet's flavour to be I. R. C. Safe
Cannot just count the b's, of an anti-kt
because even a soft gluon splits
```

…and even the LH2023 song “I.R.C. safe” (to the tune of YMCA) was dedicated to flavoured jets

Ongoing studies started in Les Houches

Compare new generation of theory-friendly flavour-dependent jet algorithms:
infrared safe to all orders (or up to high order)
and with exact (or close to exact) anti-k_t kinematics.

[Caletti, Larkoski, Marzani, Reichelt (2205.01109)] **SDF**

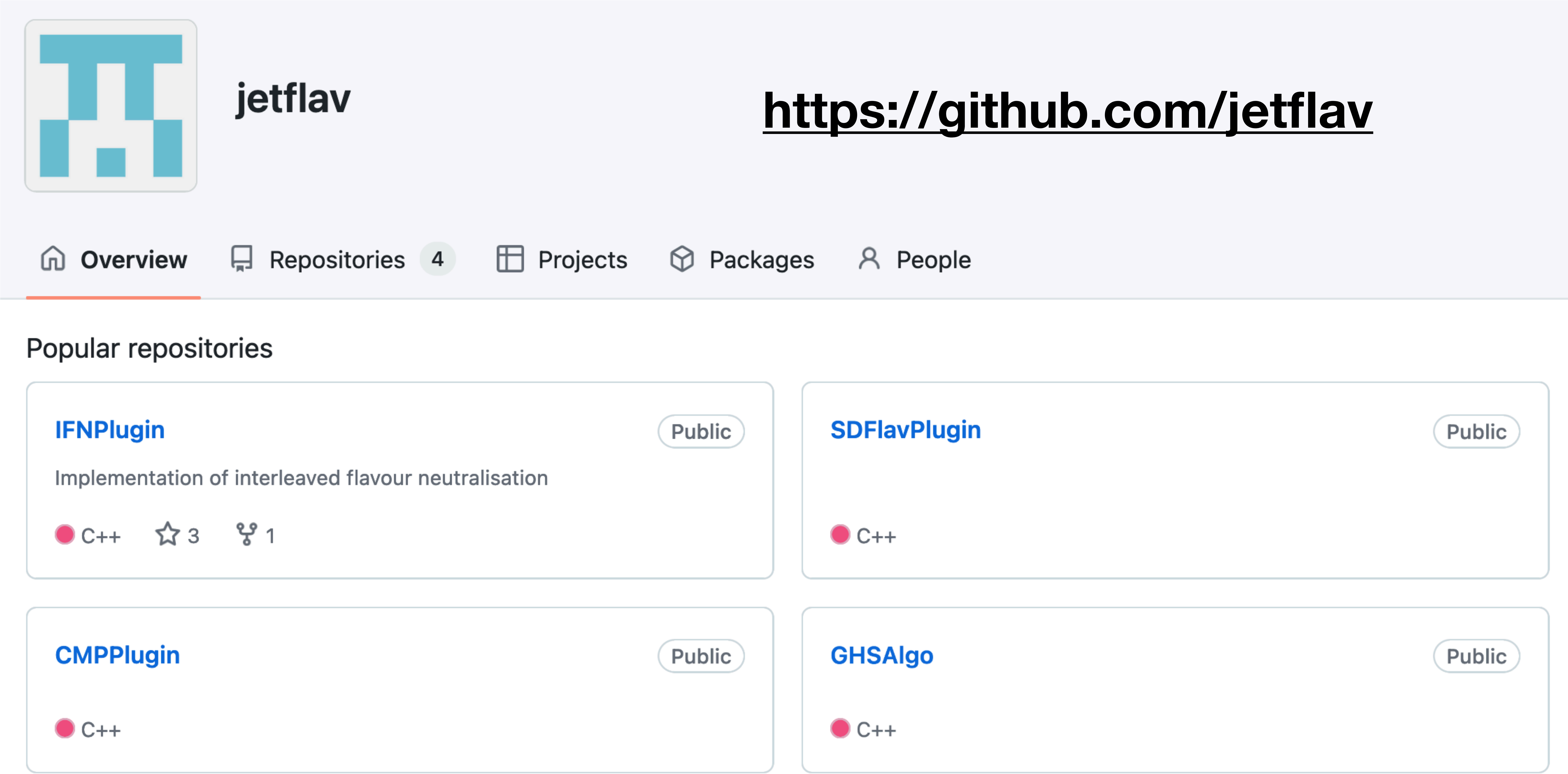
[Czakon, Mitov, Poncelet (2205.11879)] **CMP**

[Gauld, Huss, Stagnitto (2208.11138)] **GHS**

[Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler (2306.07314)] **IFN**

Phenomenological studies both at fixed-order and with (N)LO+PS simulations
$\rightarrow$ **in this talk I will focus on (N)LO+PS simulations**
see talk by Arnd tomorrow about fixed-order calculations

First step:
common framework with FastJet implementation of the four algorithms



A FastJet contribution (fjcontrib) is in preparation

Z+b-jet in the central region: detail of the analysis

MC@NLO sample of $Z(\to \mu^+\mu^-)$ + jet events with SHERPA (100M events)
LHC @ 13 TeV, PDF4LHC21, scale choice $E_T(Z)$

Generation cuts:
$p_T(\mu^\pm) > 20$ GeV; $|y(\mu^\pm)| < 2.4$; 71 GeV $< m(\mu^+\mu^-) <$ 111 GeV;
$p_T(\mu^+\mu^-) > 20$ GeV (avoid jet requirement at generation level)

Samples both at parton and stable heavy-hadron level

Rivet analysis FlavAlgAnalysis (https://github.com/DReichelt/LH23FlavAlgs)
Based on https://rivet.hepforge.org/analyses/CMS_2017_I1499471

Jet reconstructed with anti-k_t and then tagged (e.g. GHS, SDF)
or directly reconstructed and tagged with a flavoured algorithm (e.g. CMP, IFN)
$R = 0.5$, $p_T(j) > 30$ GeV, $|\eta(j)| < 2.4$

Default parameter choice for the algorithms

[Caletti, Larkoski, Marzani, Reichelt (2205.01109)] **SDF**

$\beta = 1, z_{\text{cut}} = 0.1$ (soft-drop parameters)

[Czakon, Mitov, Poncelet (2205.11879)] **CMP**

$a = 0.1$ (anti-k_t-like distance)

[Gauld, Huss, Stagnitto (2208.11138)] **GHS**

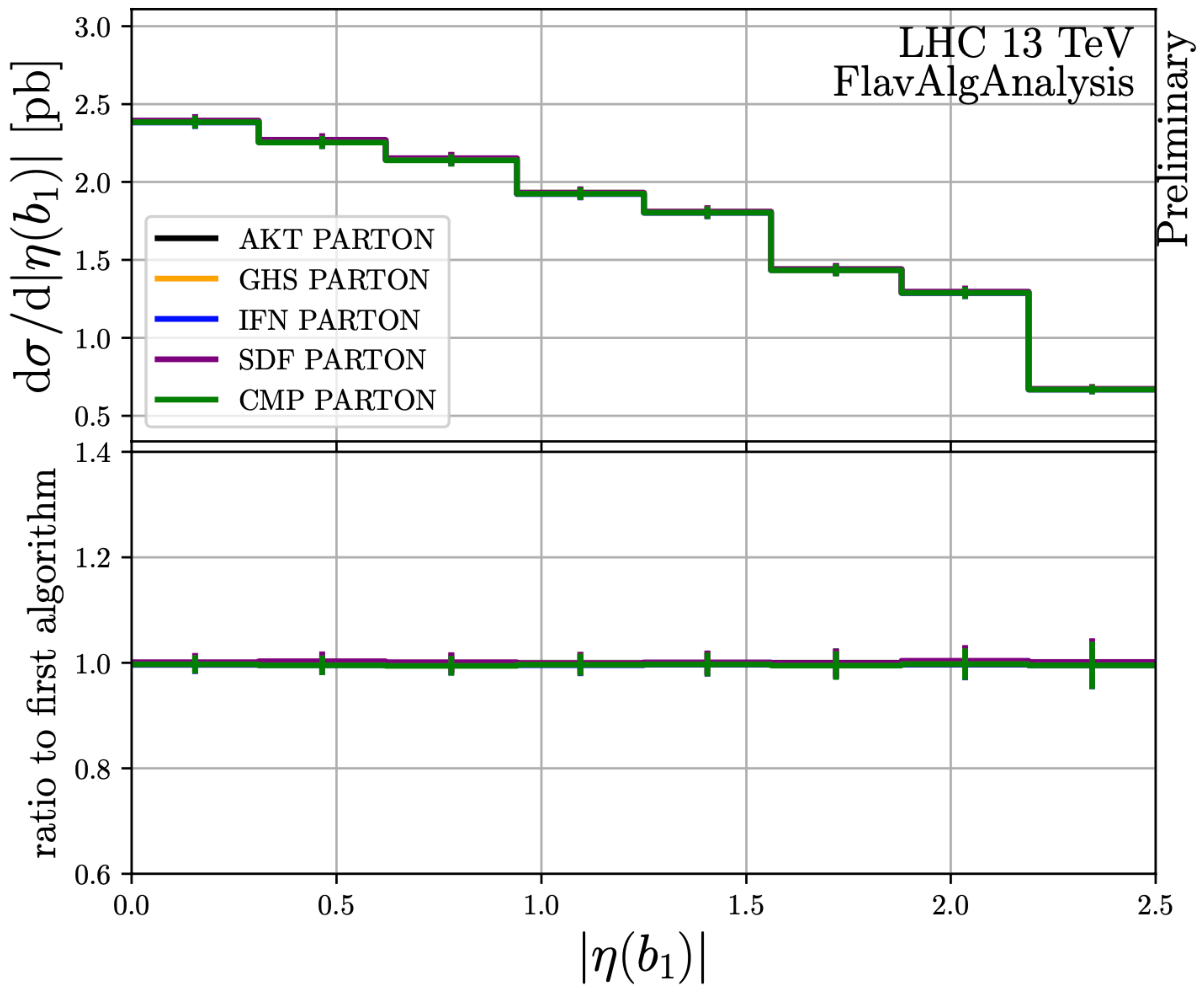
$\alpha = 1, \omega = 2$ (flavour-k_t-like and beam distance)

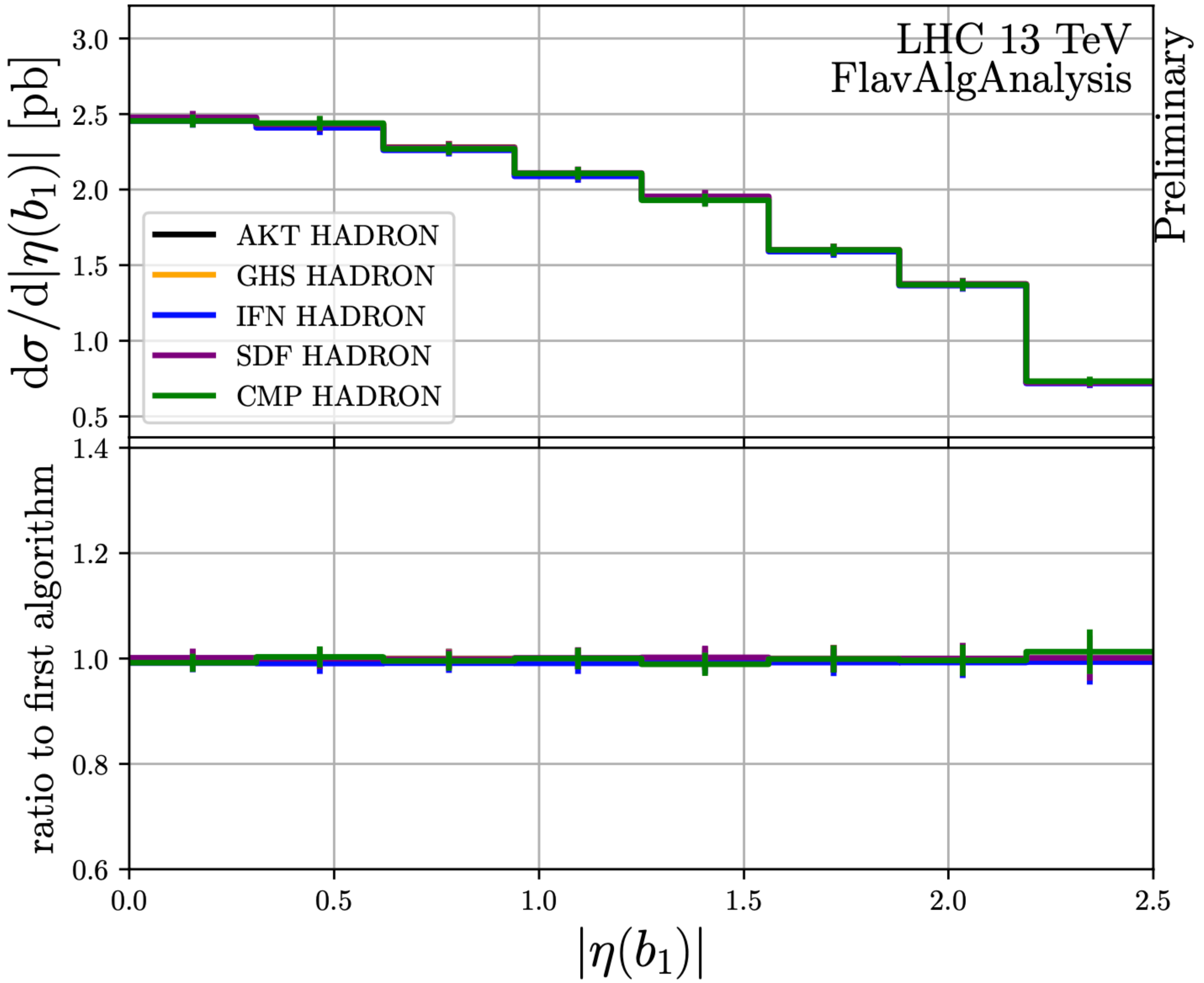
[Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler (2306.07314)] **IFN**

$\alpha = 1, \omega = 2$ (flavour-k_t-like distance)

In plots, **AKT** is the naive IRC-unsafe flavour tagging of anti-k_t jets
(does the jet contain a flavoured parton/hadron?)

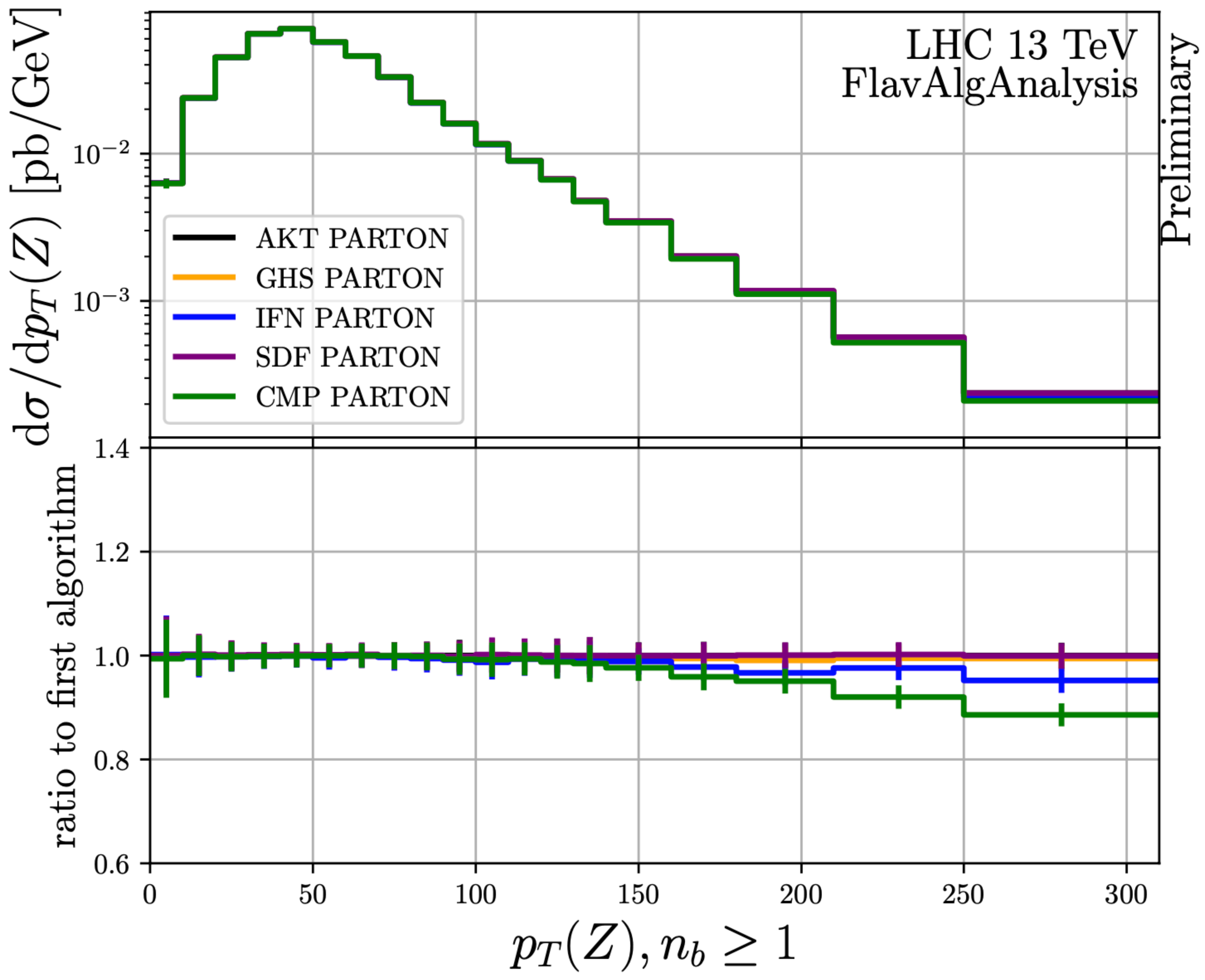
Z+b-jet in the central region: comparison of algorithms

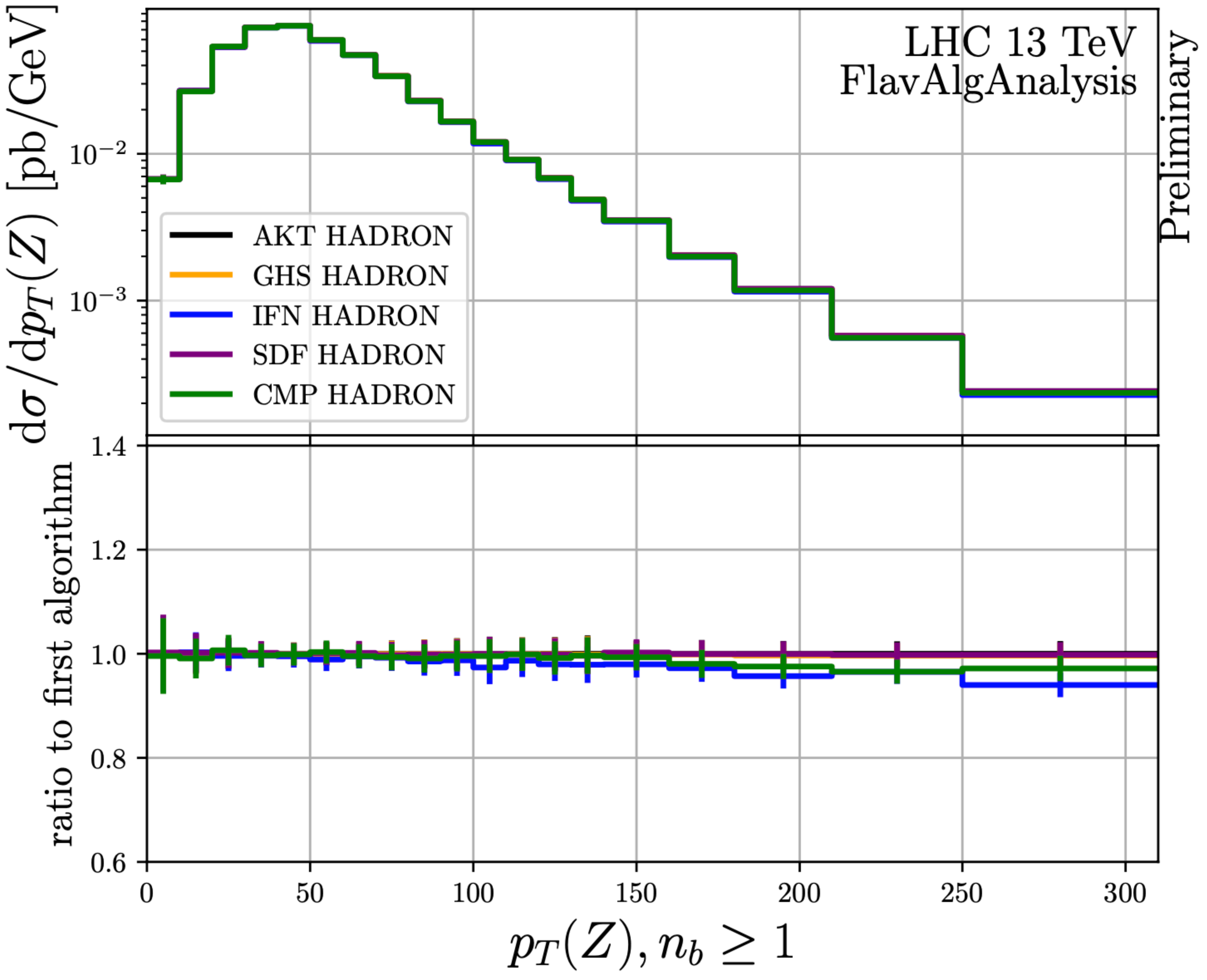




For many distributions, difference between algorithms negligible, both at parton and hadron levels: once you require IRC safety, constrained behaviour in most of the phase space?

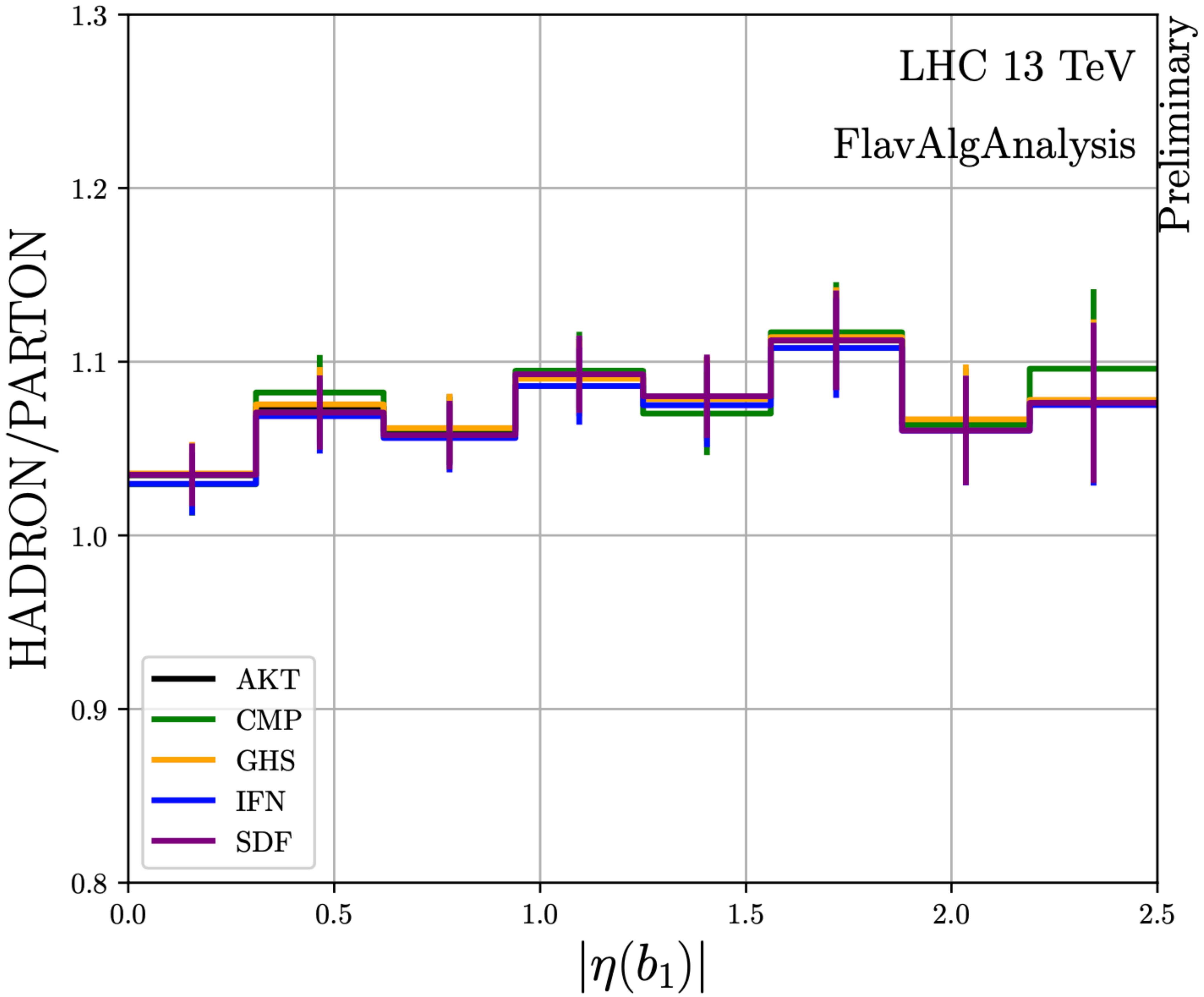
Z+b-jet in the central region: comparison of algorithms

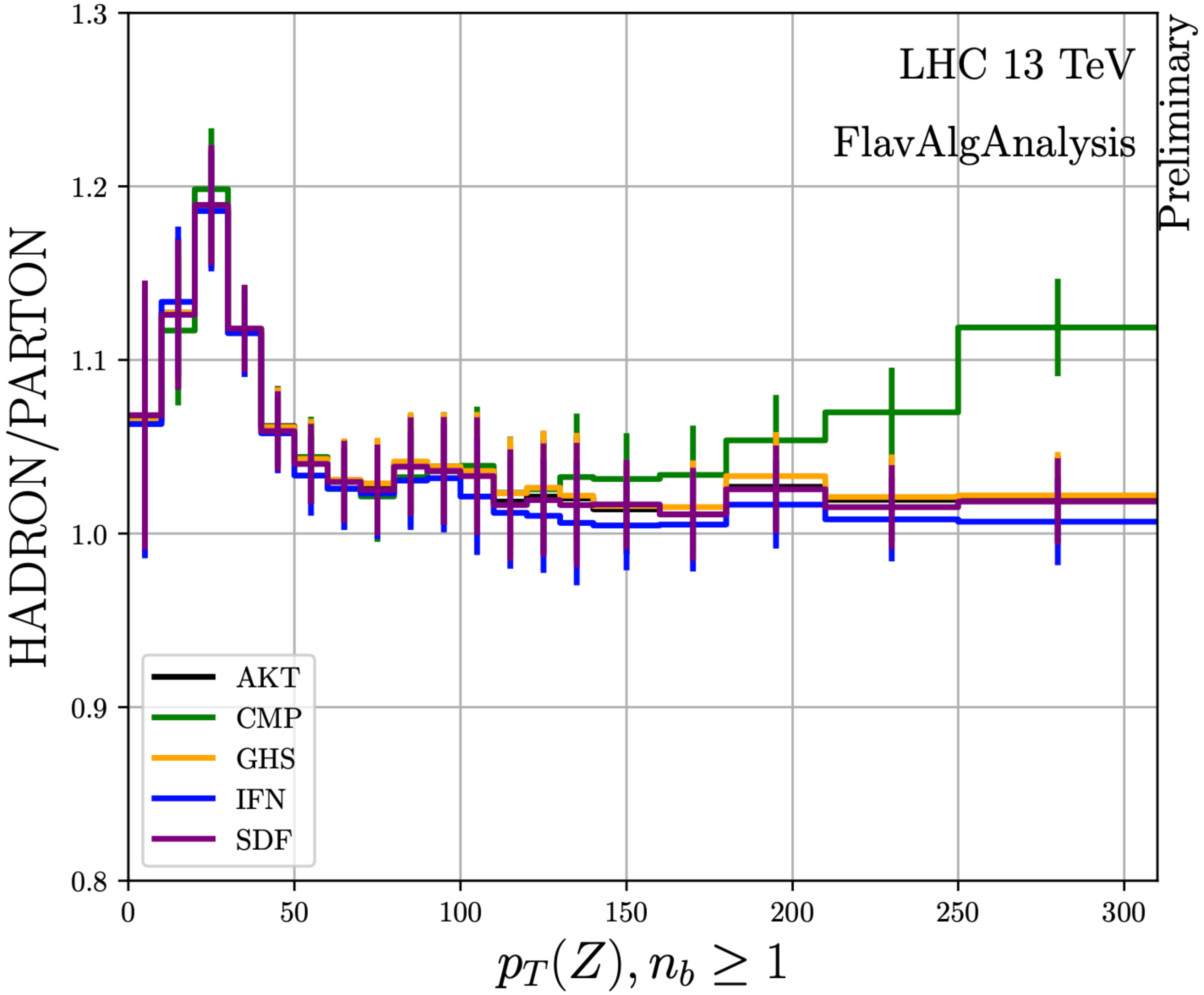




Visible difference at high-p_T: CMP and IFN outliers compared to AKT/GHS/SDF

Z+b-jet in the central region: comparison of algorithms





Unfolding correction as simple K-factor: similar to AKT for most distributions,
CMP behaviour at high $p_T(Z)$ to be understood

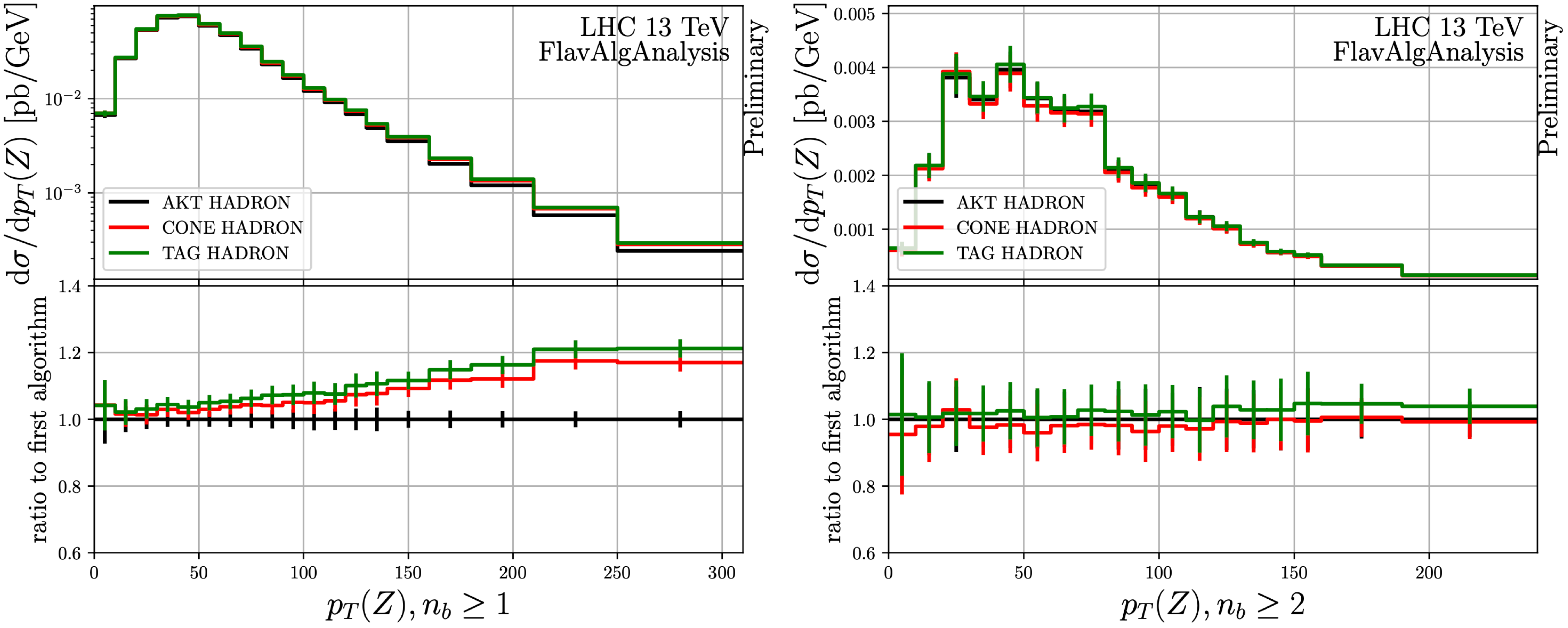
Z+b-jet in the central region: comparison of (unsafe) tagging strategies

AKT:
check if flavoured particle / hadron is inside the anti-k_t jet

CONE:
ATLAS-style tagging
(heavy hadron with $\Delta R(j,h) < 0.3$ and with $p_T > 5$ GeV)

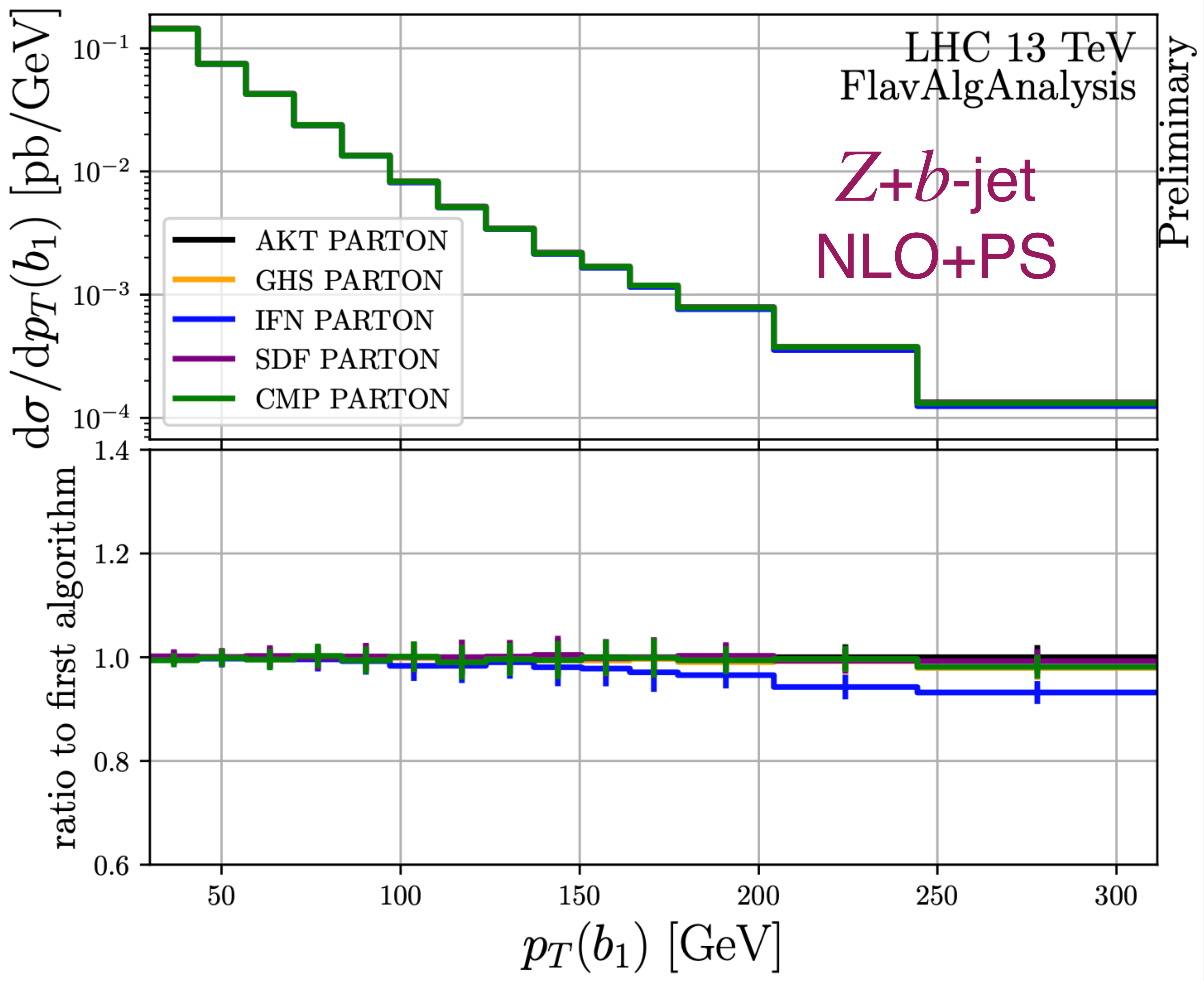
TAG:
CMS-style tagging
(bTagged method with ghost-tagging in Rivet)

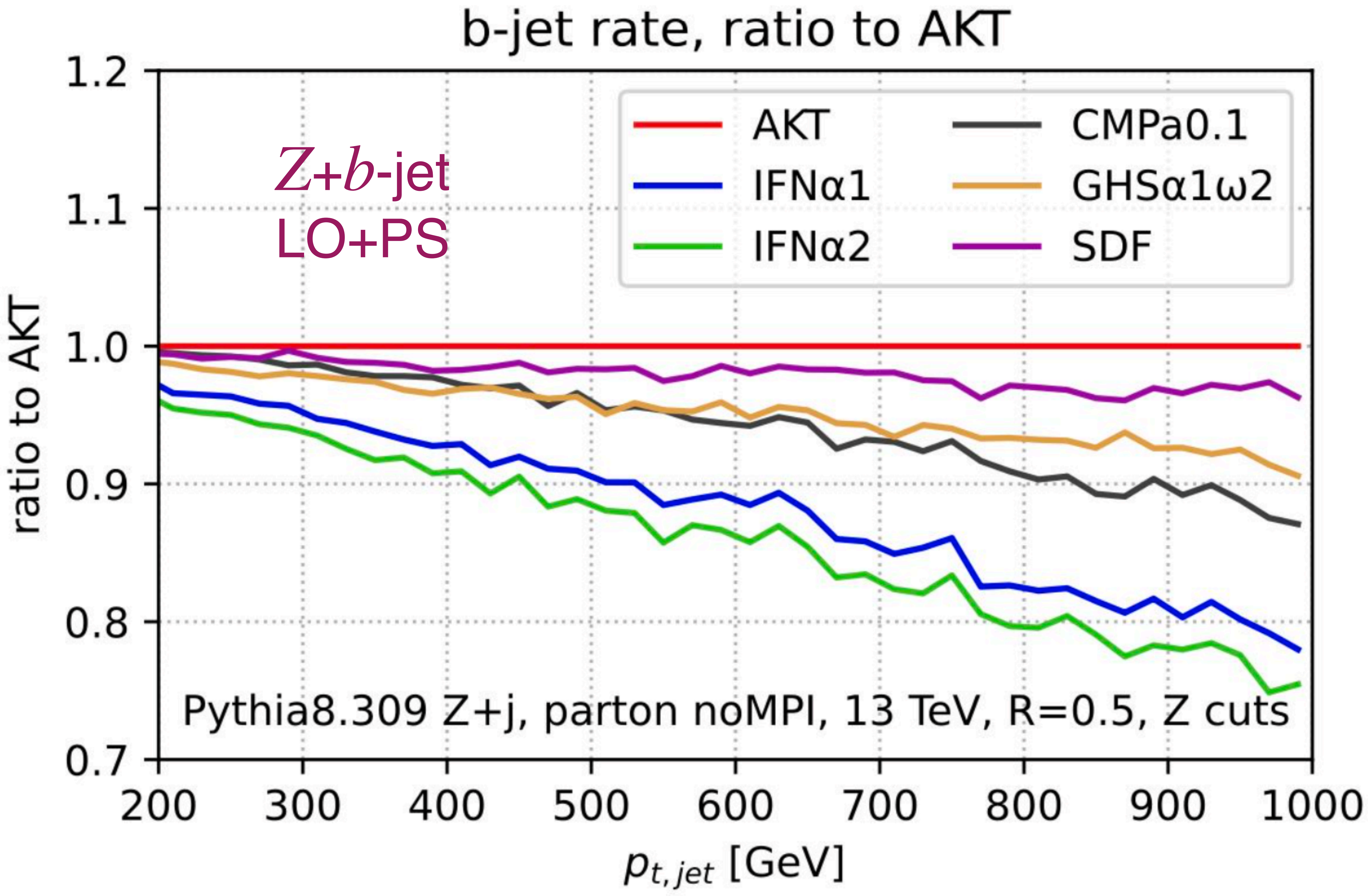
Z+b-jet in the central region: comparison of (unsafe) tagging strategies



Effect of multi-b tags important at high-p_T: by requiring presence of 2 b-jets, we have a reduced probability of double tags, mostly coming from first splitting

Z+b-jet in the central region: understanding the high-p_T behaviour





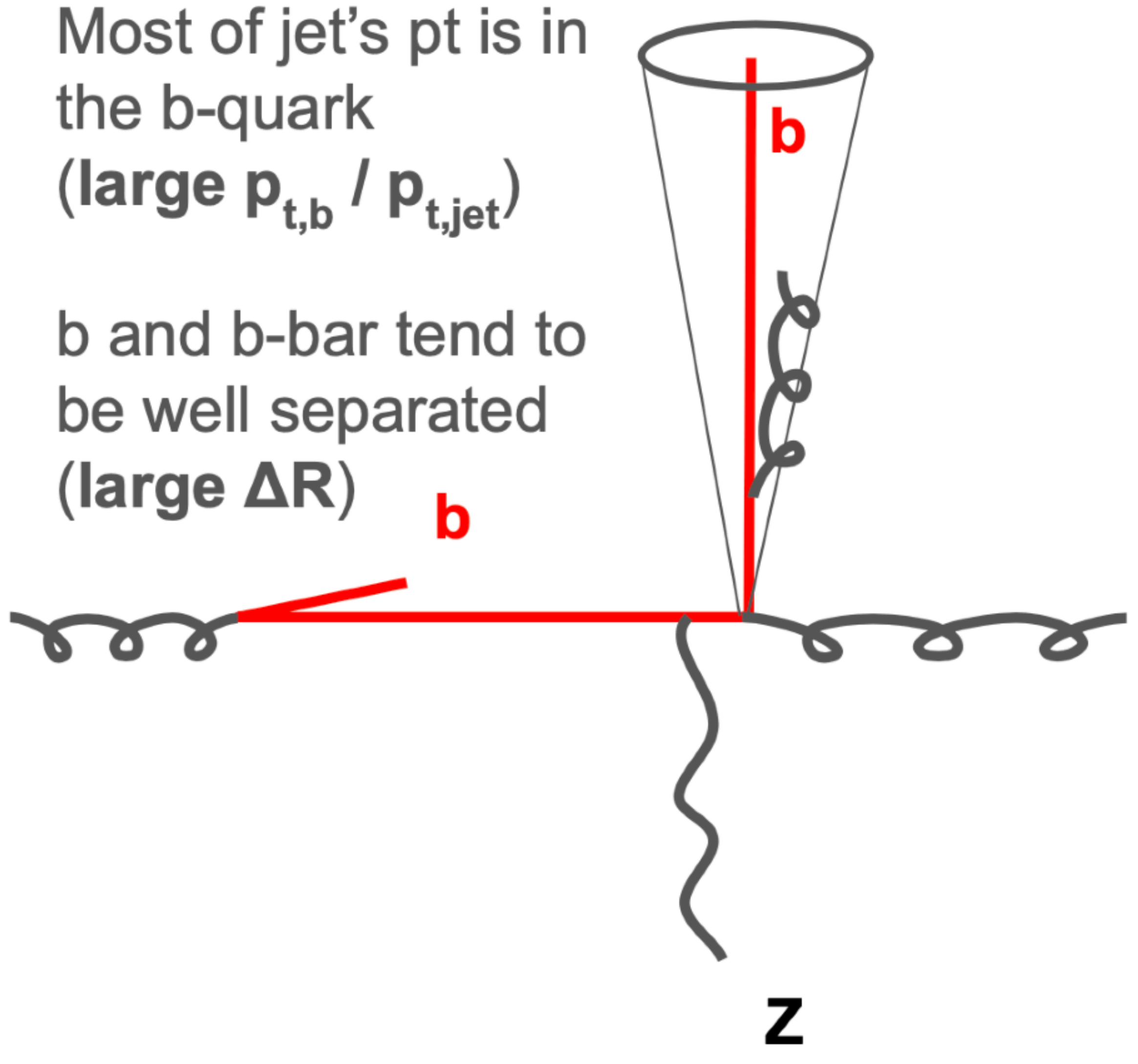
As we increase the p_T, the difference between algorithms is enhanced

Z+b-jet in the central region: understanding the high-p_T behaviour

Genuine b-jet

Most of jet's pt is in the b-quark (**large $p_{t,b}$ / $p_{t,jet}$**)

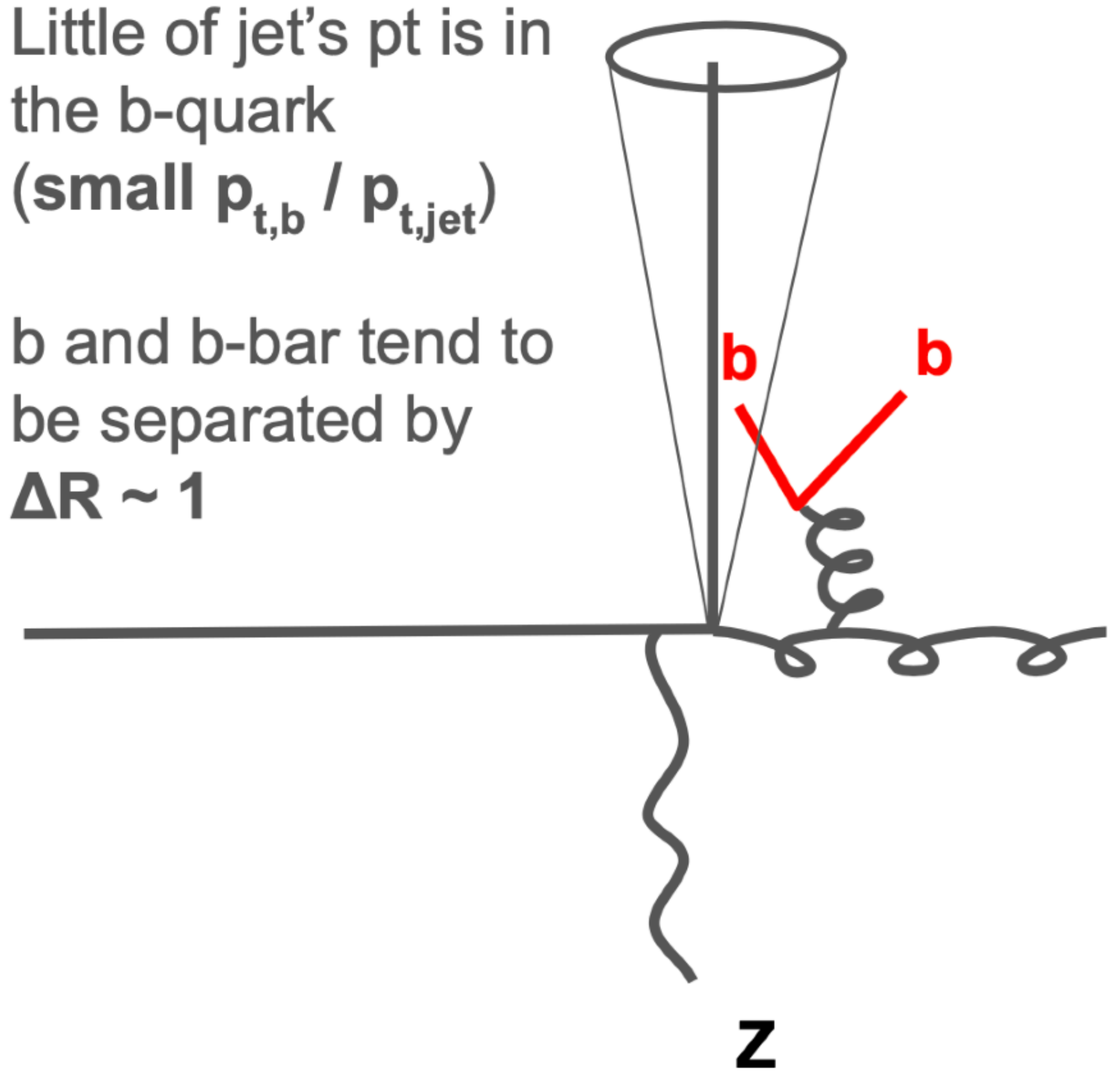
b and b-bar tend to be well separated (**large ΔR**)



Fake b-jet

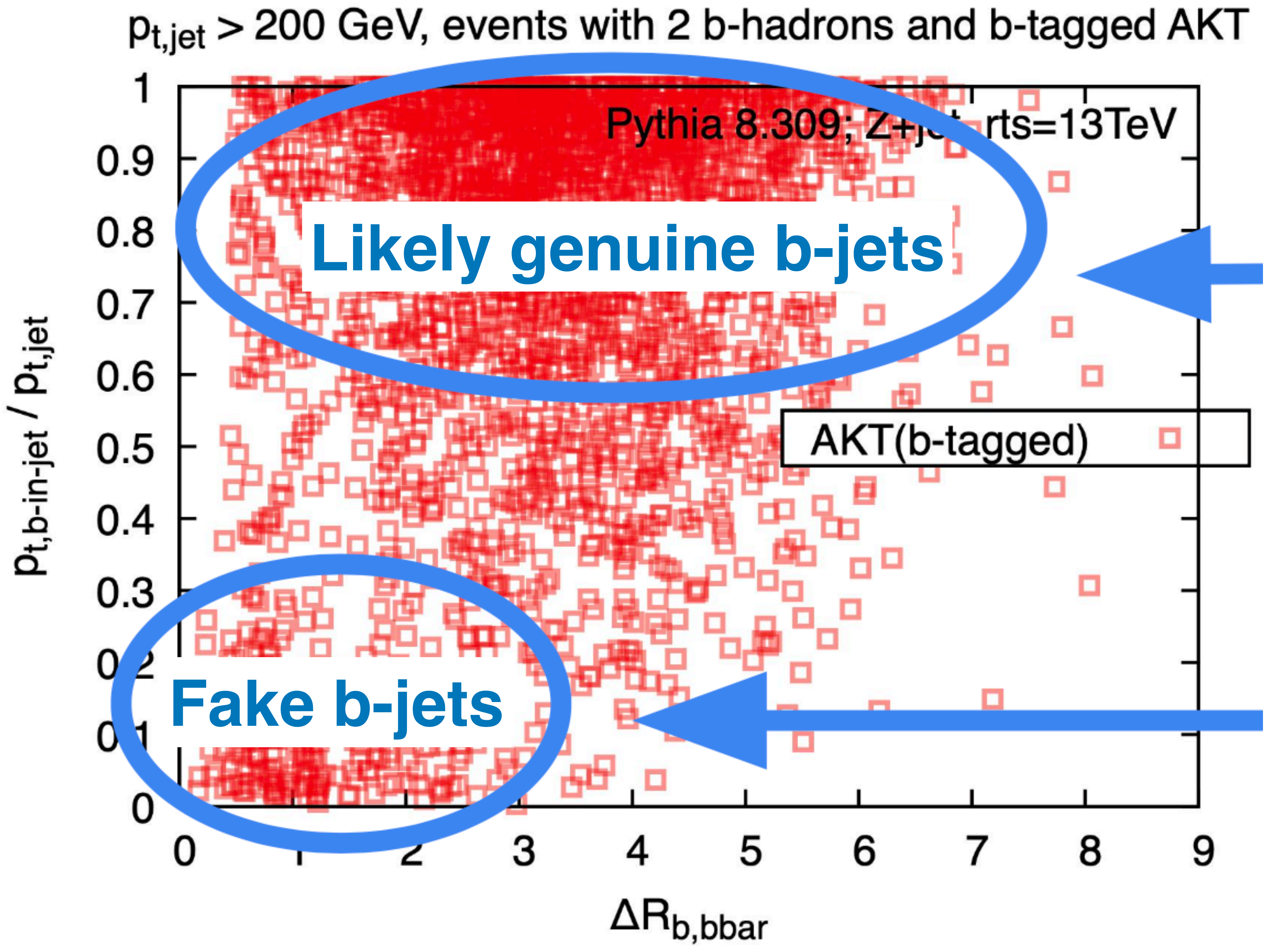
Little of jet's pt is in the b-quark (**small $p_{t,b}$ / $p_{t,jet}$**)

b and b-bar tend to be separated by **ΔR ~ 1**

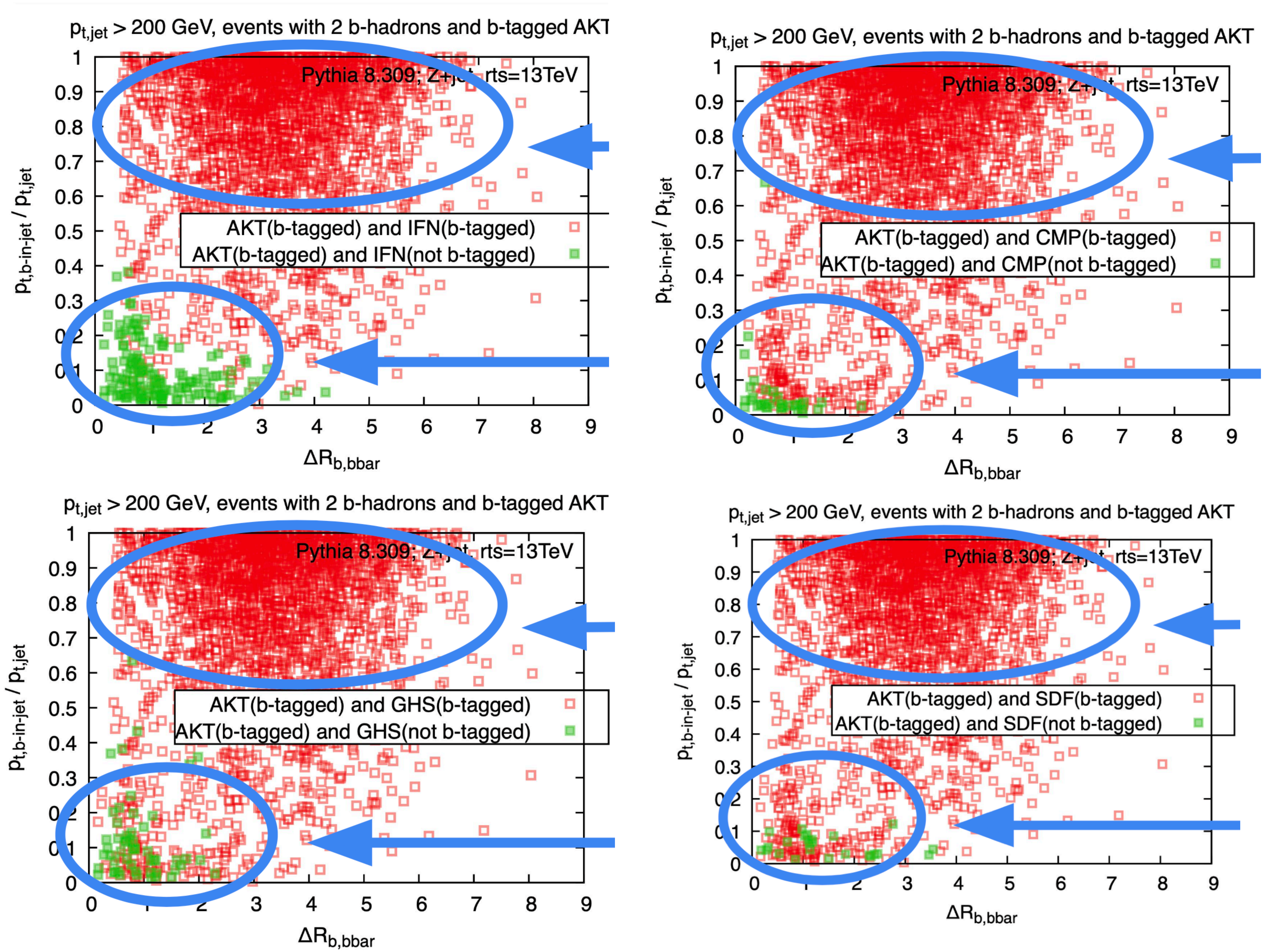


"Genuine" and "fake" b-jets can be disentangled in a $(p_{T,b}/p_{T,j},\ \Delta R_{b,\bar{b}})$ plane

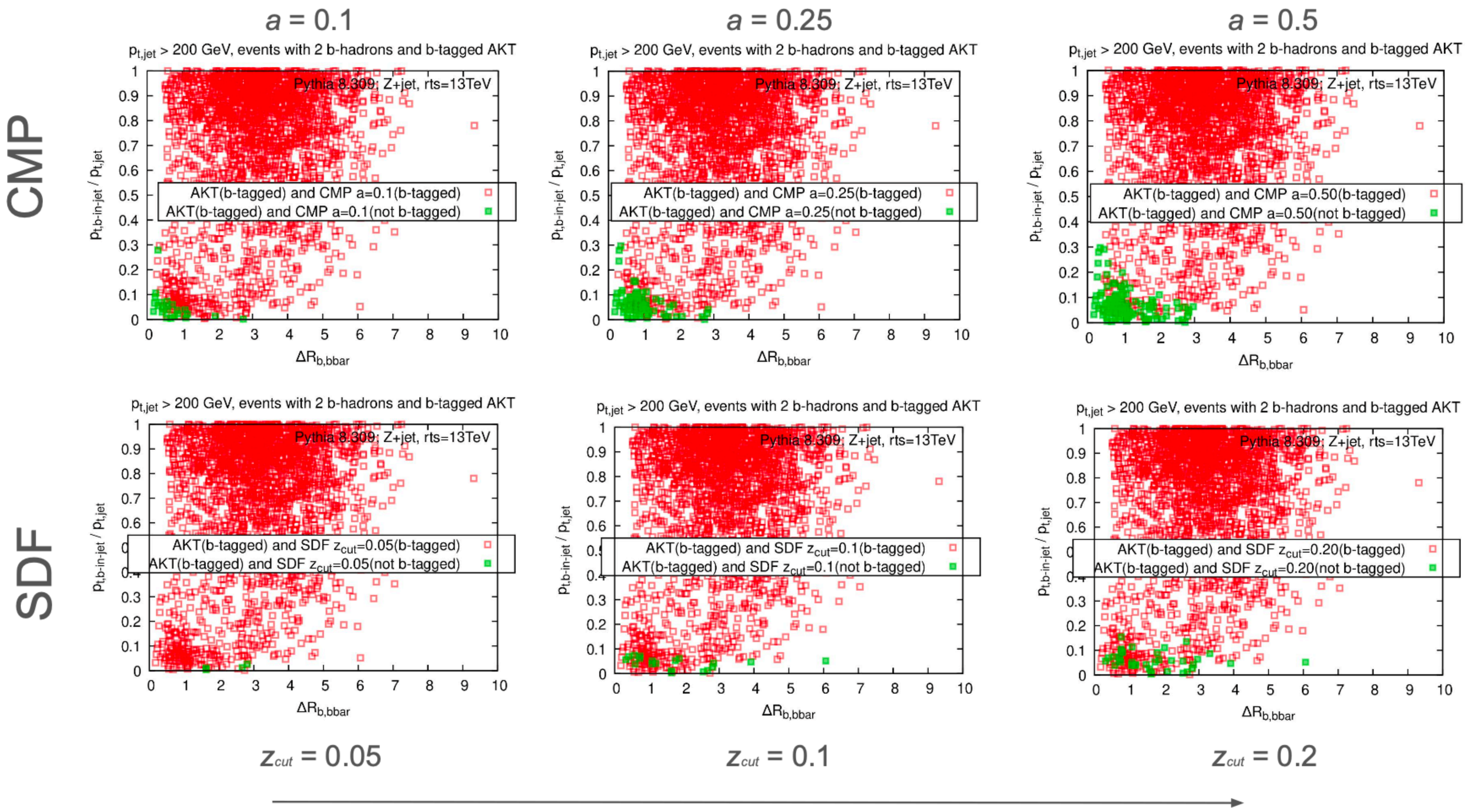
Naive AKT will tag both "genuine" and "fake" b-jets



Z+b-jet in the central region: understanding the high-p_T behaviour



Z+b-jet in the central region: understanding the high-p_T behaviour



Fraction of retained $b\bar{b}$ pairs depends on the value of the parameter entering the algorithm

As it become larger, more $b\bar{b}$ pairs are neutralised

Z+c-jet in the forward region: details

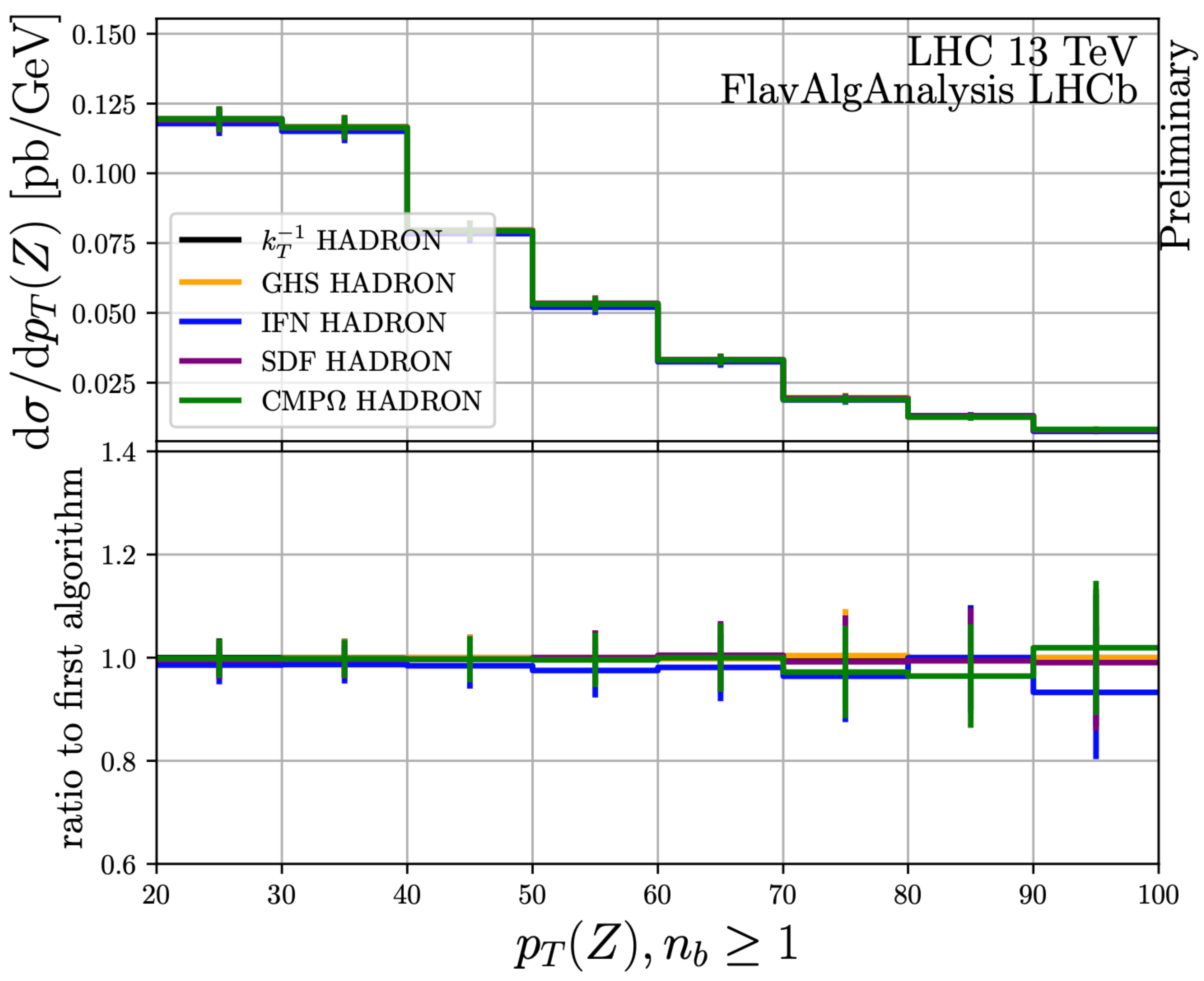
We first create LHEF with MadGraph5_aMC@NLO
[LHC @ 13 TeV, PDF4LHC21, scale choice $E_T(Z)$]
with very loose generation cuts (no requirement on rapidity of particles)

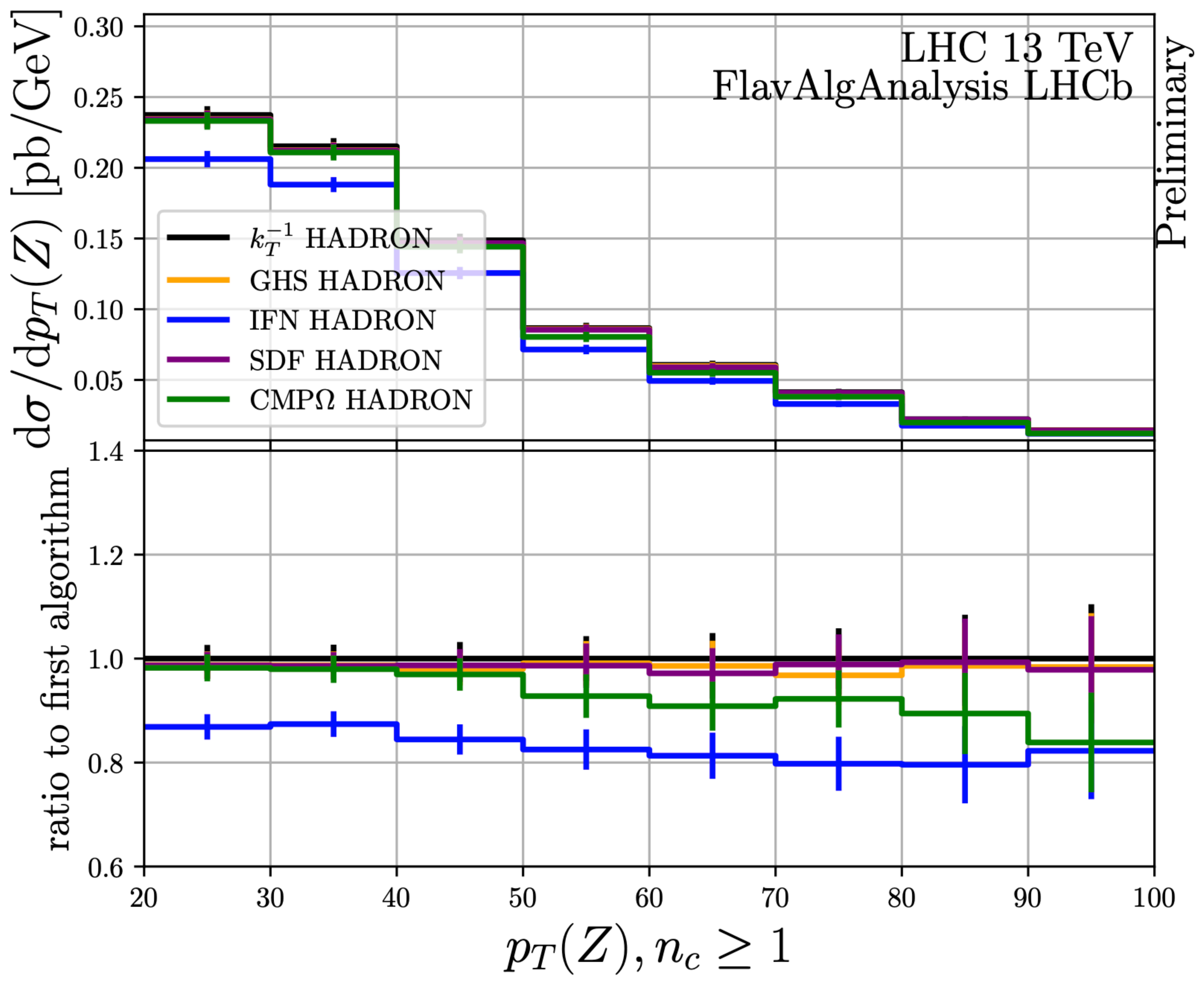
We then shower with Pythia8
and we apply analysis cuts similar to LHCb 2109.08084:
$p_T(\mu^\pm) > 20$ GeV; $2.0 < \eta(\mu^\pm) < 4.5$; 60 GeV $< m(\mu^+\mu^-) <$ 120 GeV

Jets are clustered with $R = 0.5$ and are required to have:
20 GeV $< p_T(j) <$ 100 GeV; $2.2 < \eta(j) < 4.2$; $\Delta R(\mu^\pm, j) > 0.5$

Same algorithms with default parameters as before,
but we now consider both charm and bottom as flavoured particles

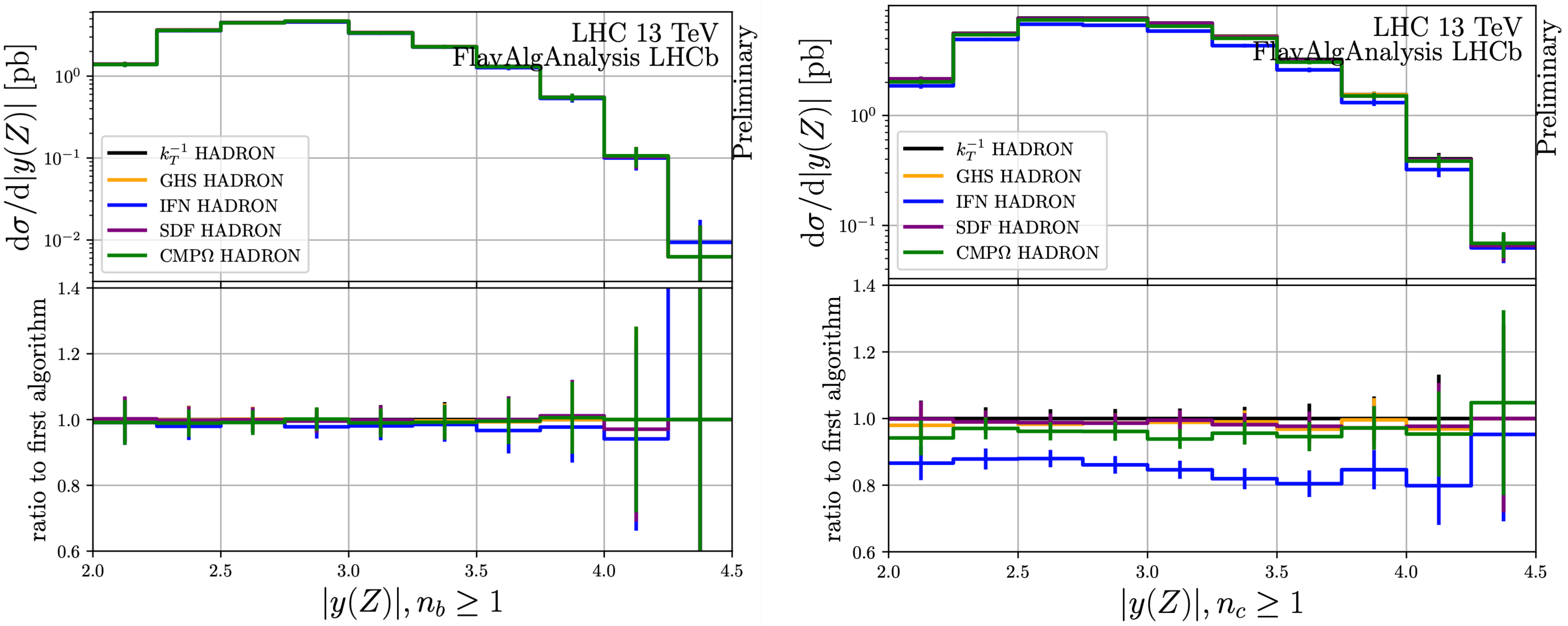
Z+c-jet in the forward region: comparison of algorithms





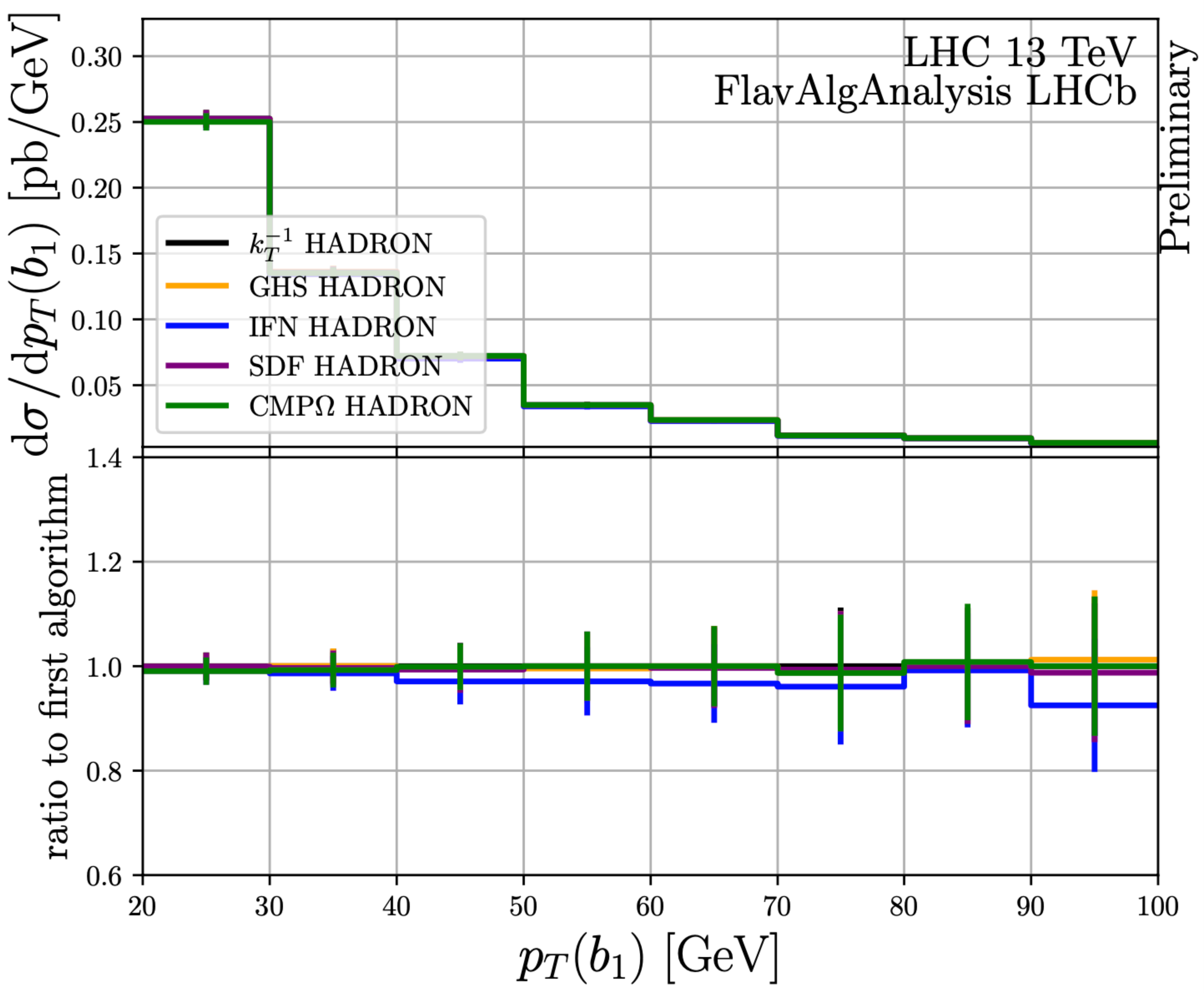
We note good agreement between algorithms in case of bottom,
and significant differences in the whole phase space in case of charm

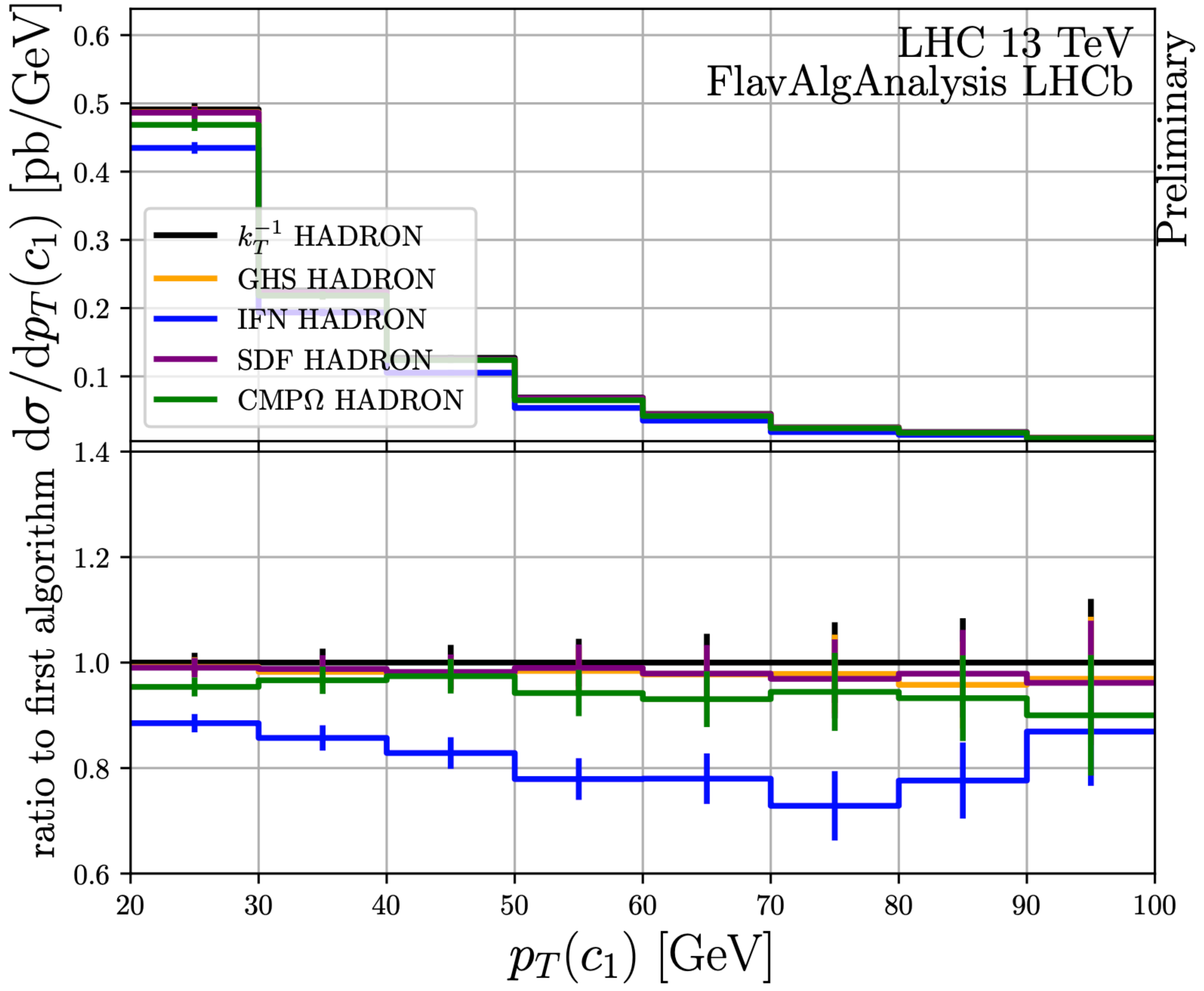
Z+c-jet in the forward region: comparison of algorithms



We note good agreement between algorithms in case of bottom,
and significant differences in the whole phase space in case of charm

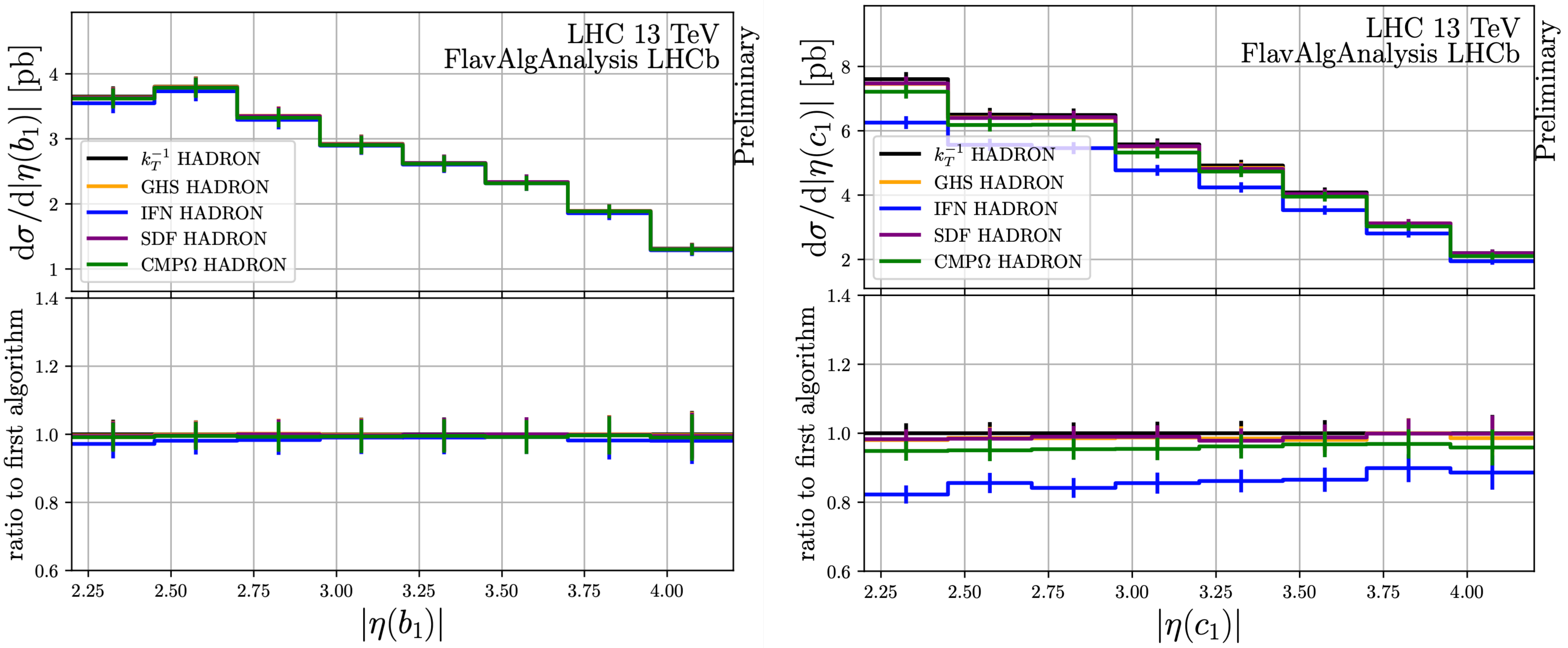
Z+c-jet in the forward region: comparison of algorithms





We note good agreement between algorithms in case of bottom, and significant differences in the whole phase space in case of charm

Z+c-jet in the forward region: comparison of algorithms



We note good agreement between algorithms in case of bottom,
and significant differences in the whole phase space in case of charm

Z+c-jet in the forward region: comparison of algorithms

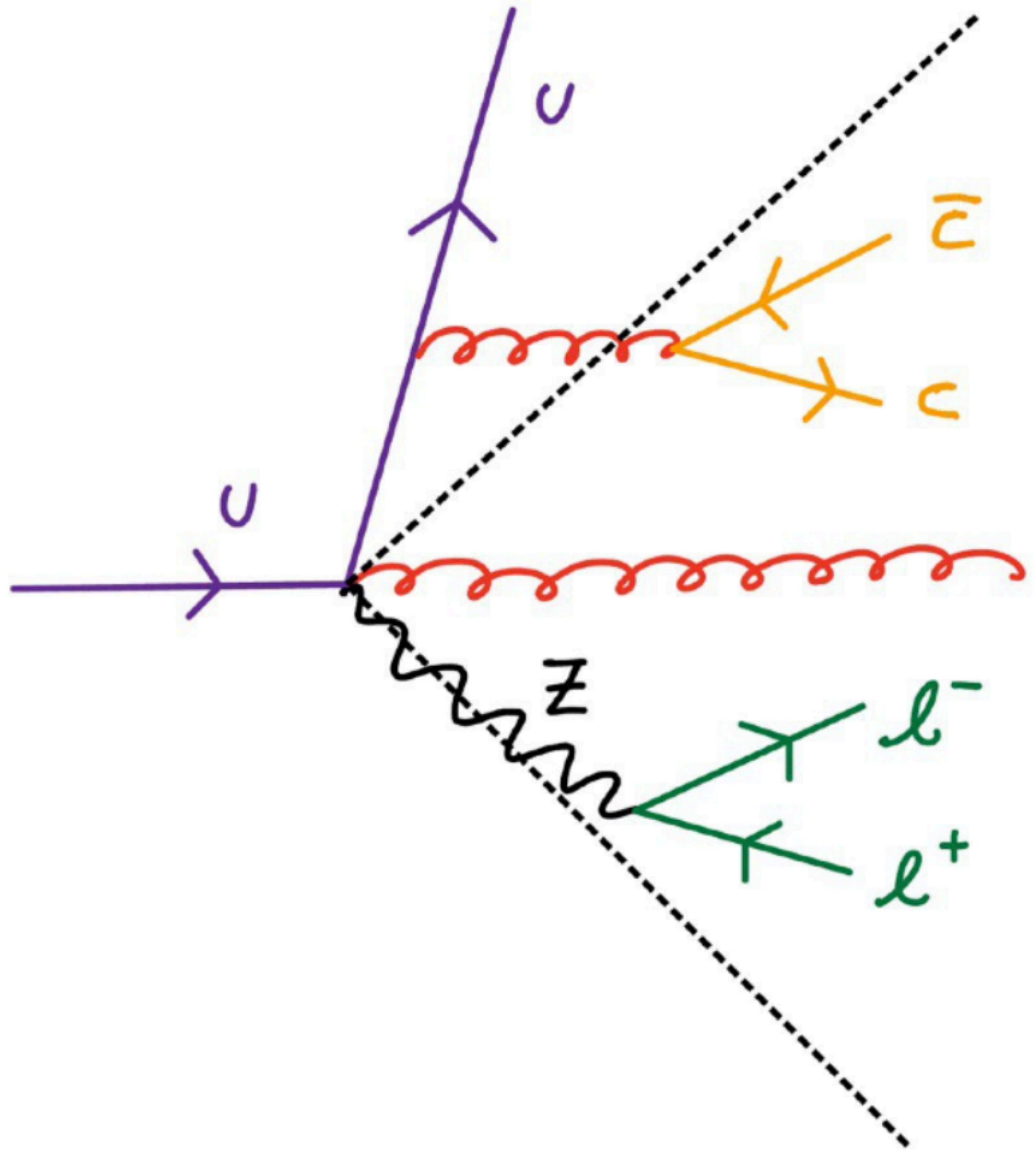
The difference we observe in case of charm needs to be understood, but one can make some considerations:

- more $g \to c\bar{c}$ splittings compared to $g \to b\bar{b}$ $\to$ stressing algorithms

- charm in the LHCb fiducial region likely to come from gluon splitting

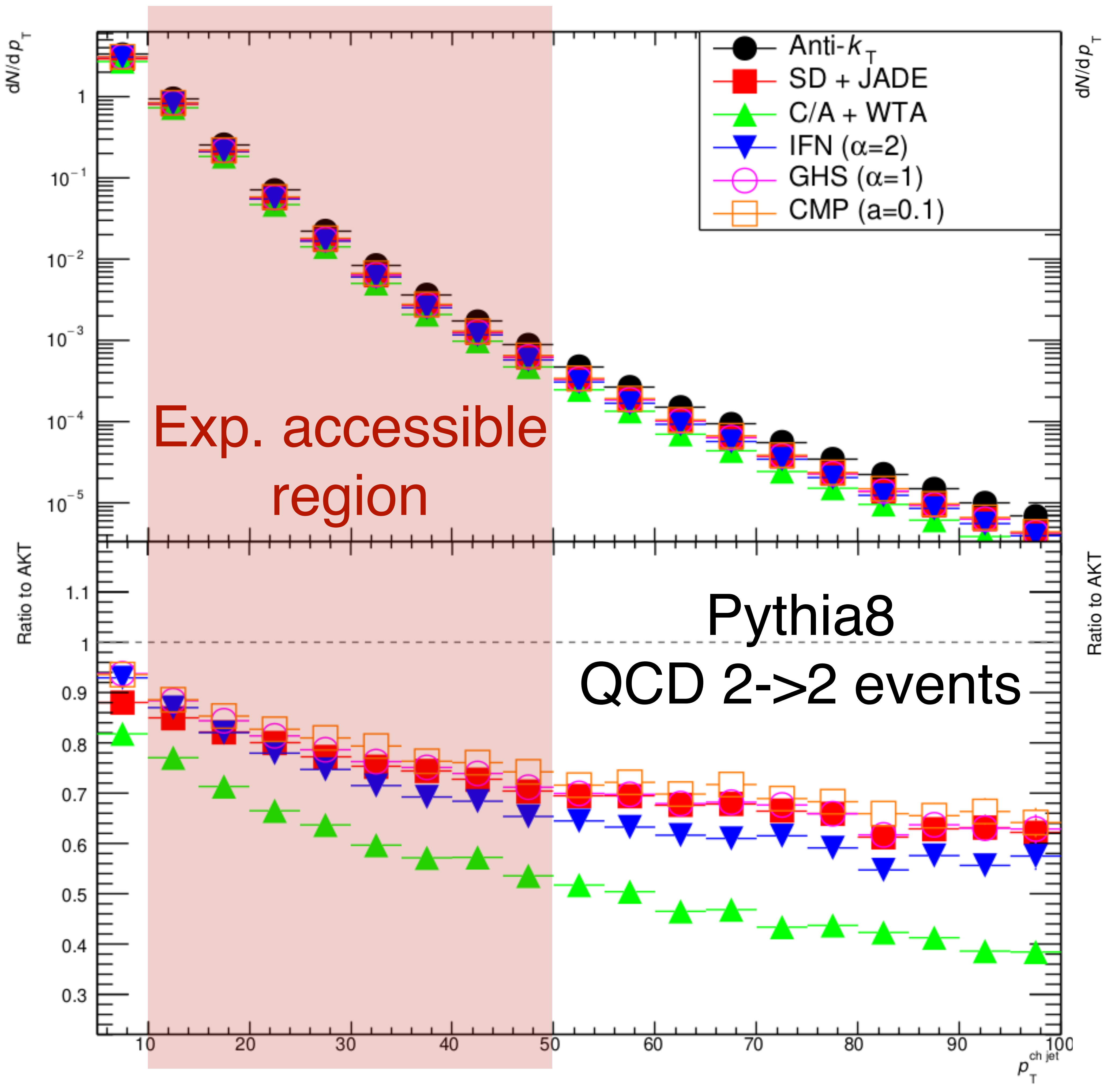
- IFN undershooting other algorithms $\to$ it tends to neutralise $c\bar{c}$ pairs more

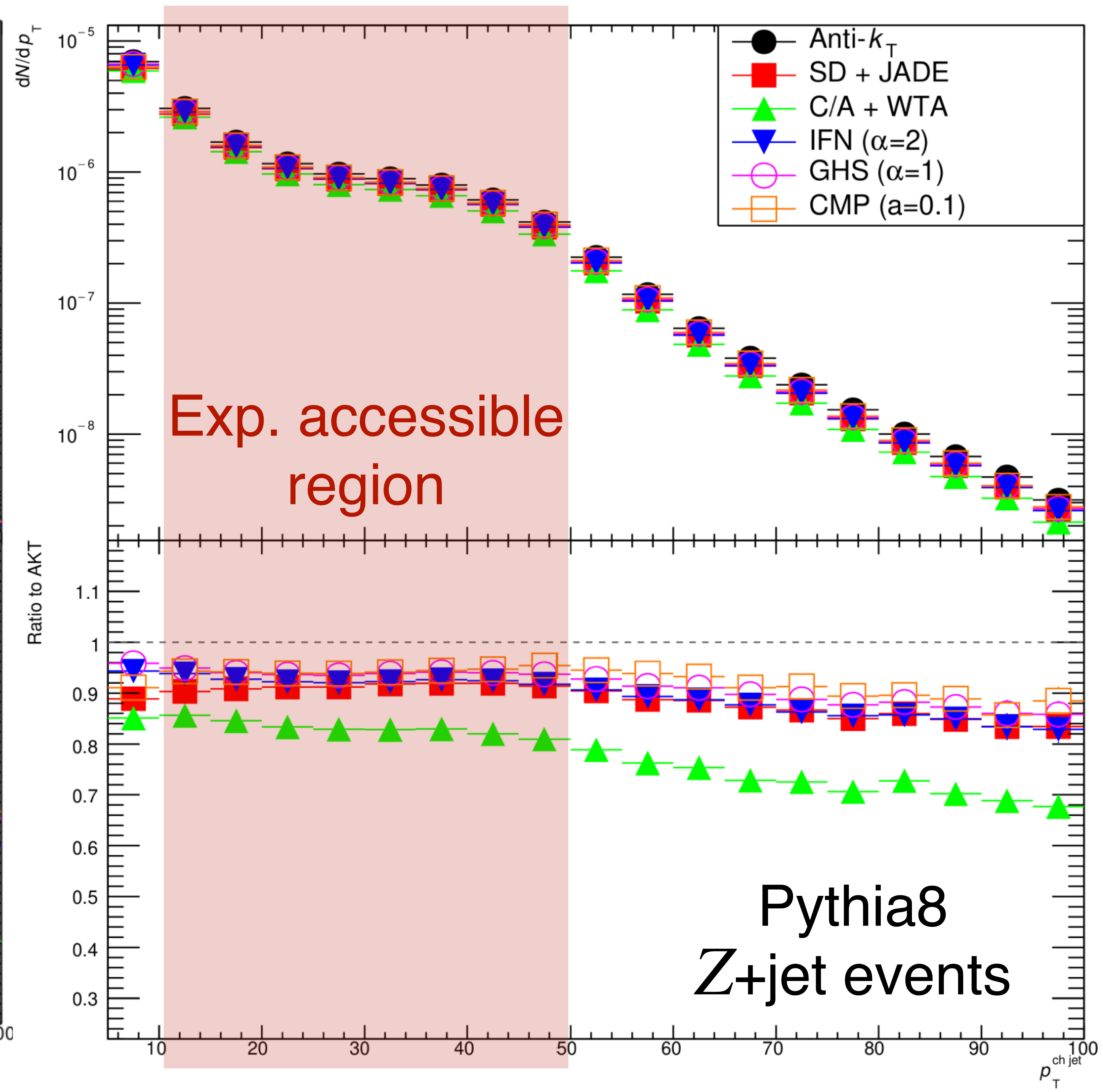
- interesting to vary parameters and perform a study similar to the Z+b-jet one



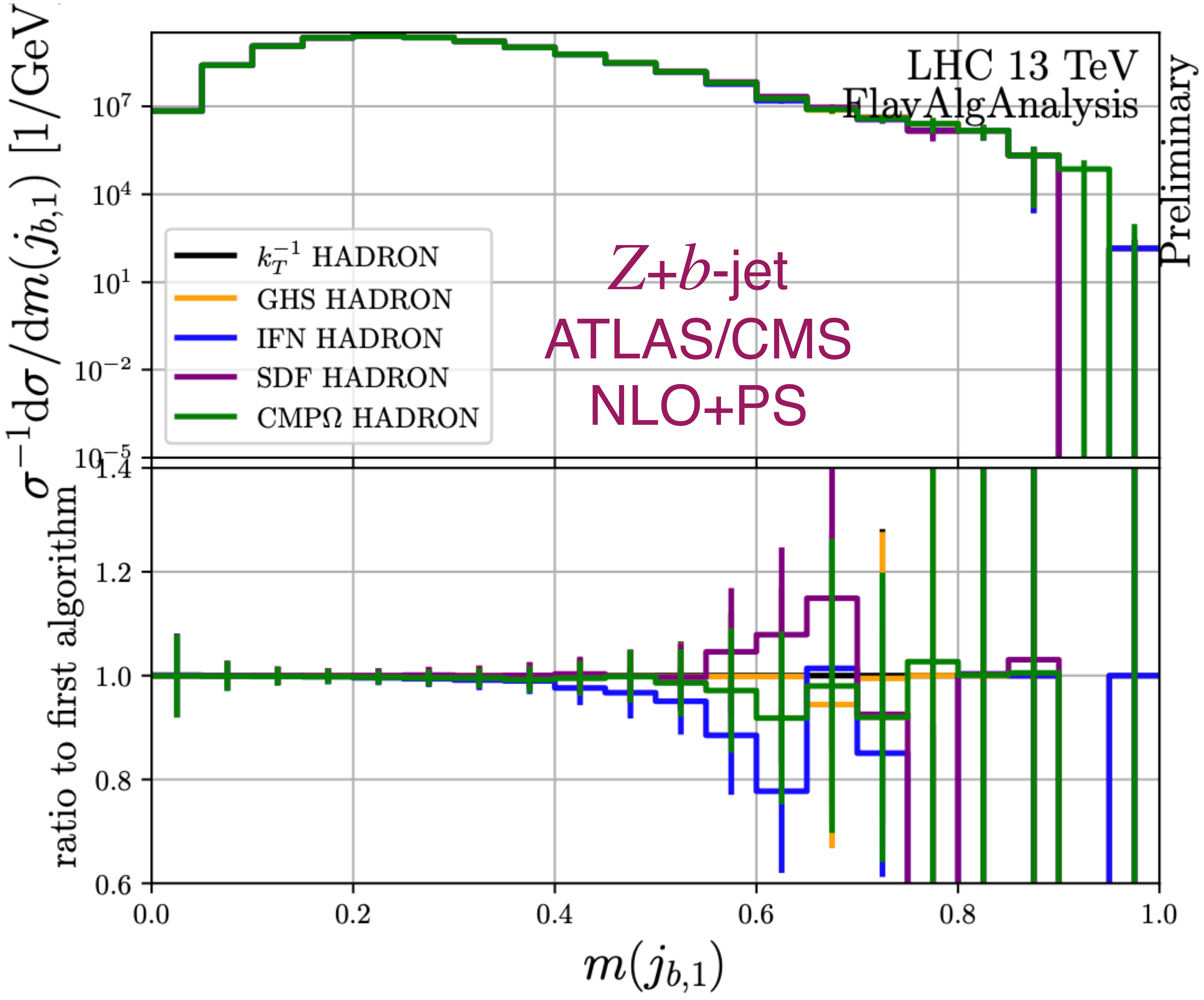
Tagging efficiency at LHCb

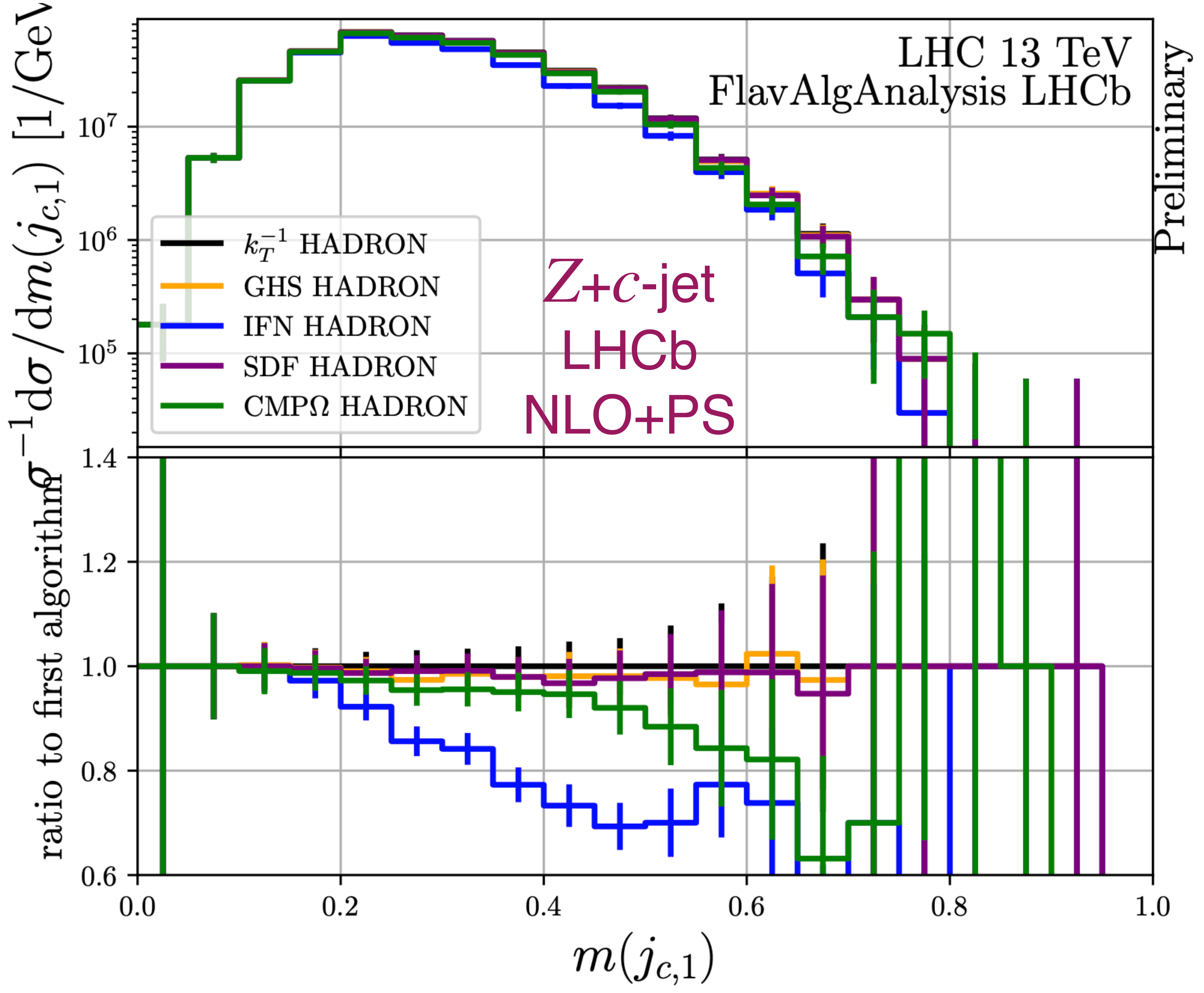
Rate of c-jet events in events with a D^0 meson ($R = 0.5$, $2 < \eta(D^0) < 5$, $2.5 < \eta(j) < 4.5$)



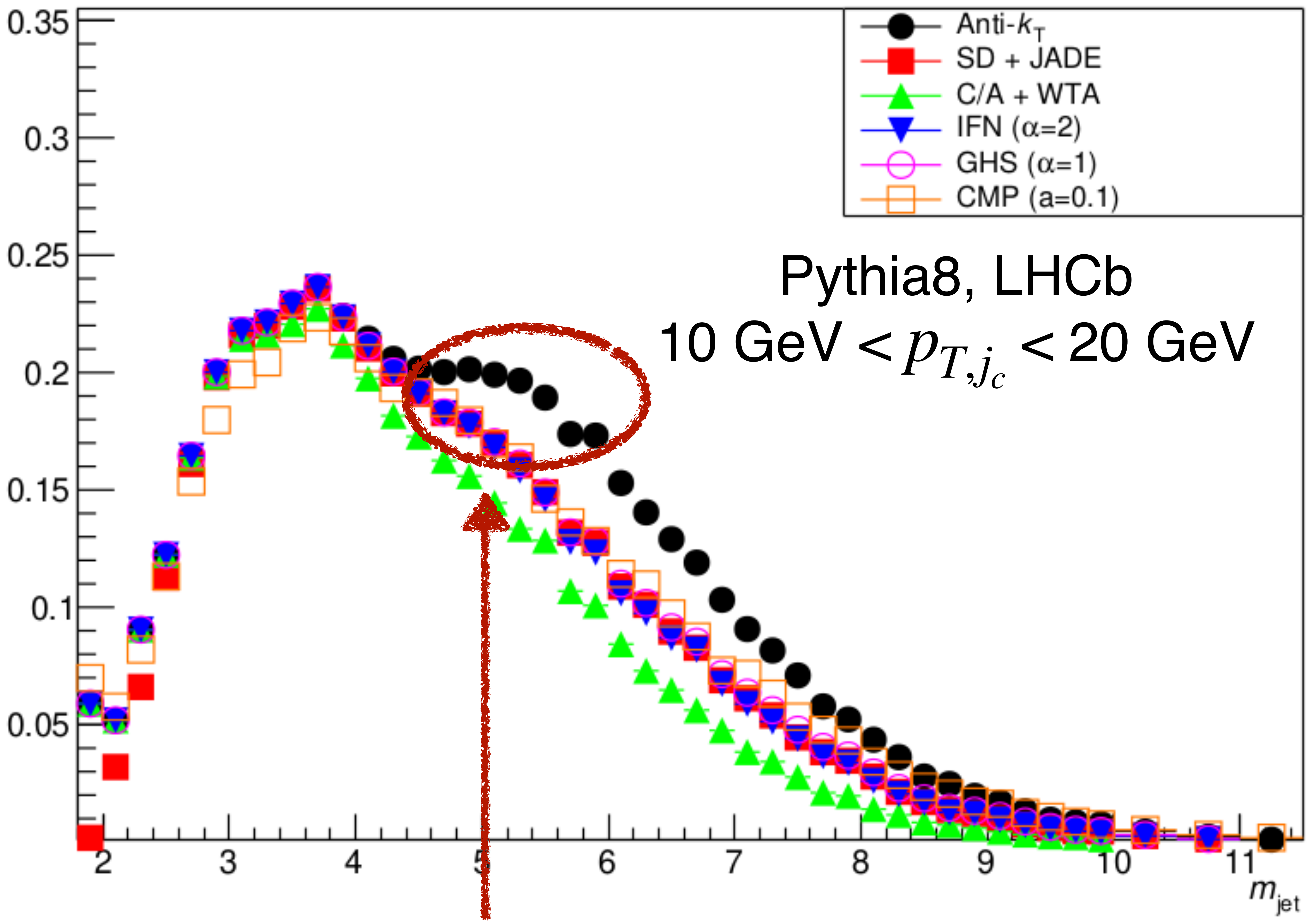


Jet substructure observables: rescaled jet mass $m(j) \equiv m/p_T$





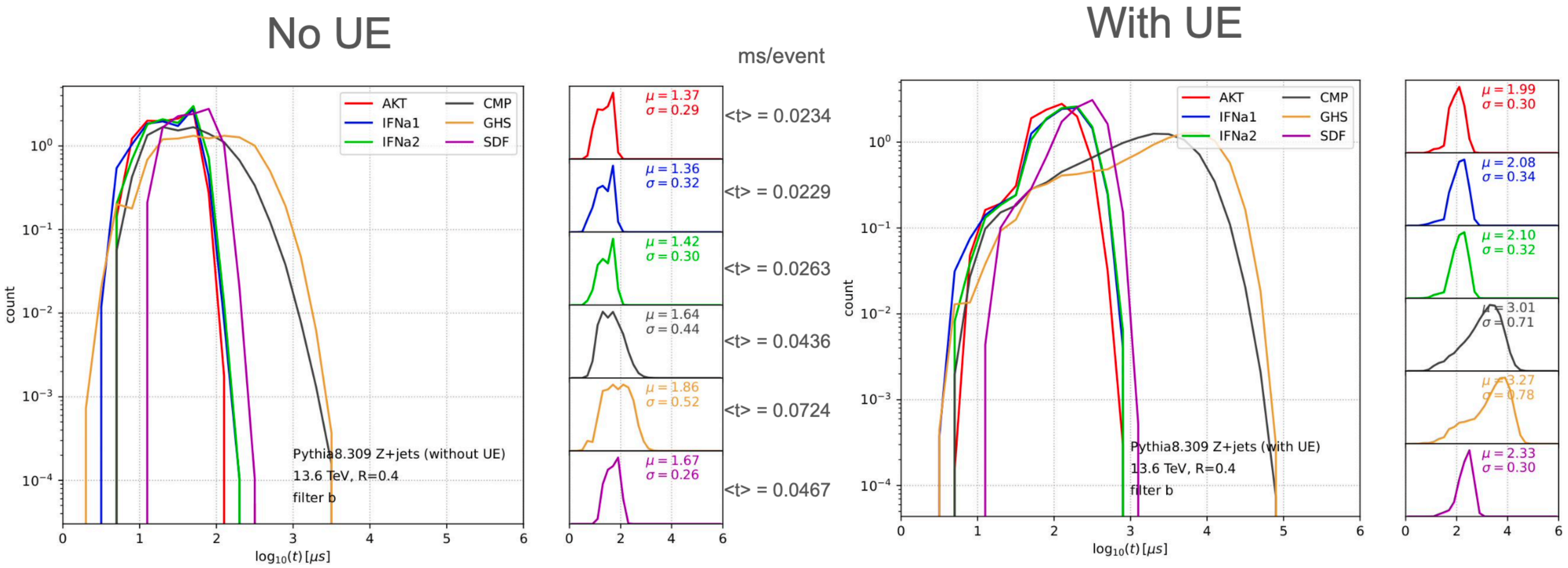
Jet substructure observables: jet mass



All algorithms remove the second peak
(probably due to $g \to c\bar{c}$ within the jet)

Performance studies of current FastJet implementations

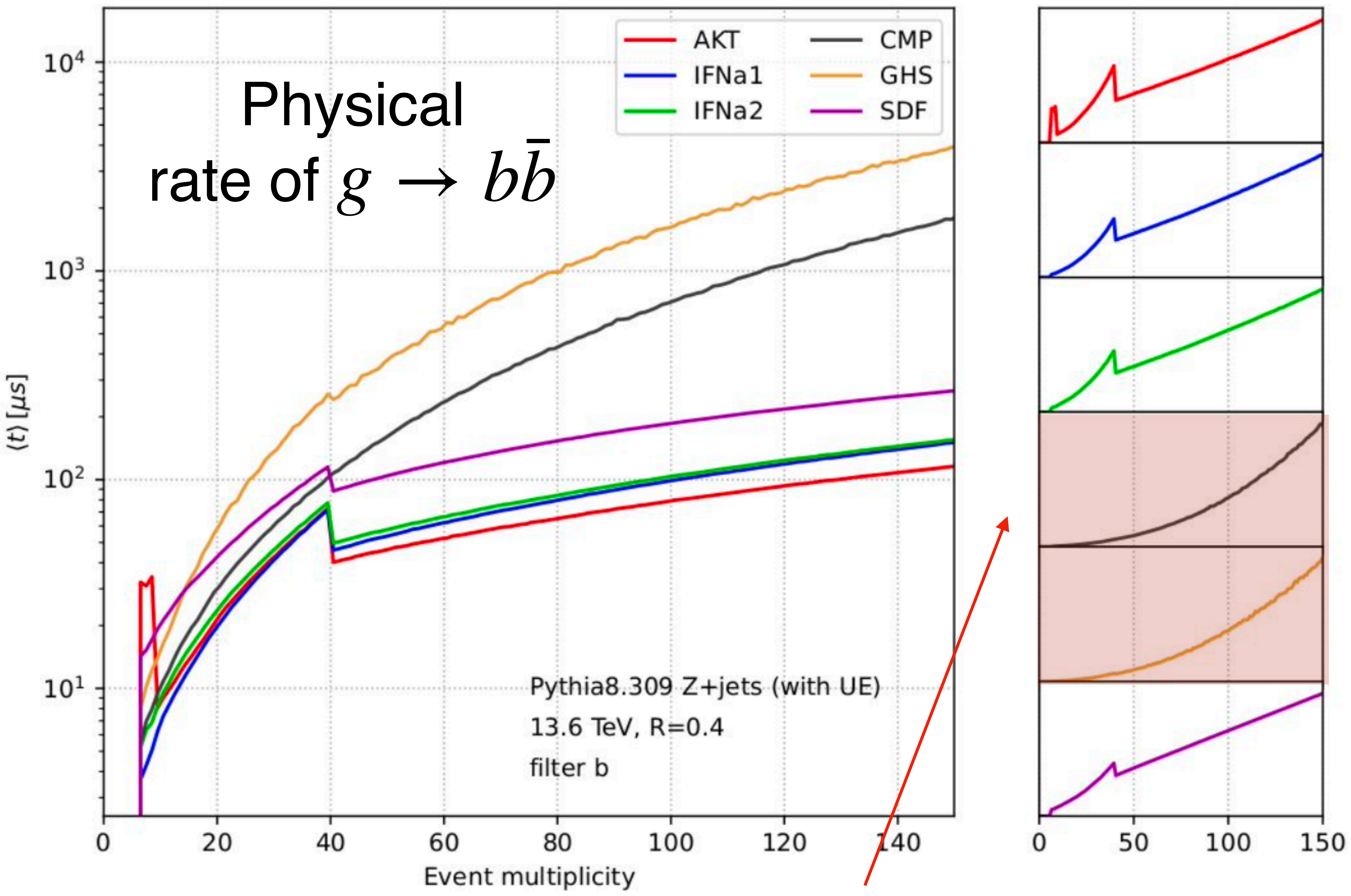
Clustering time of Z+jet Pythia8 events

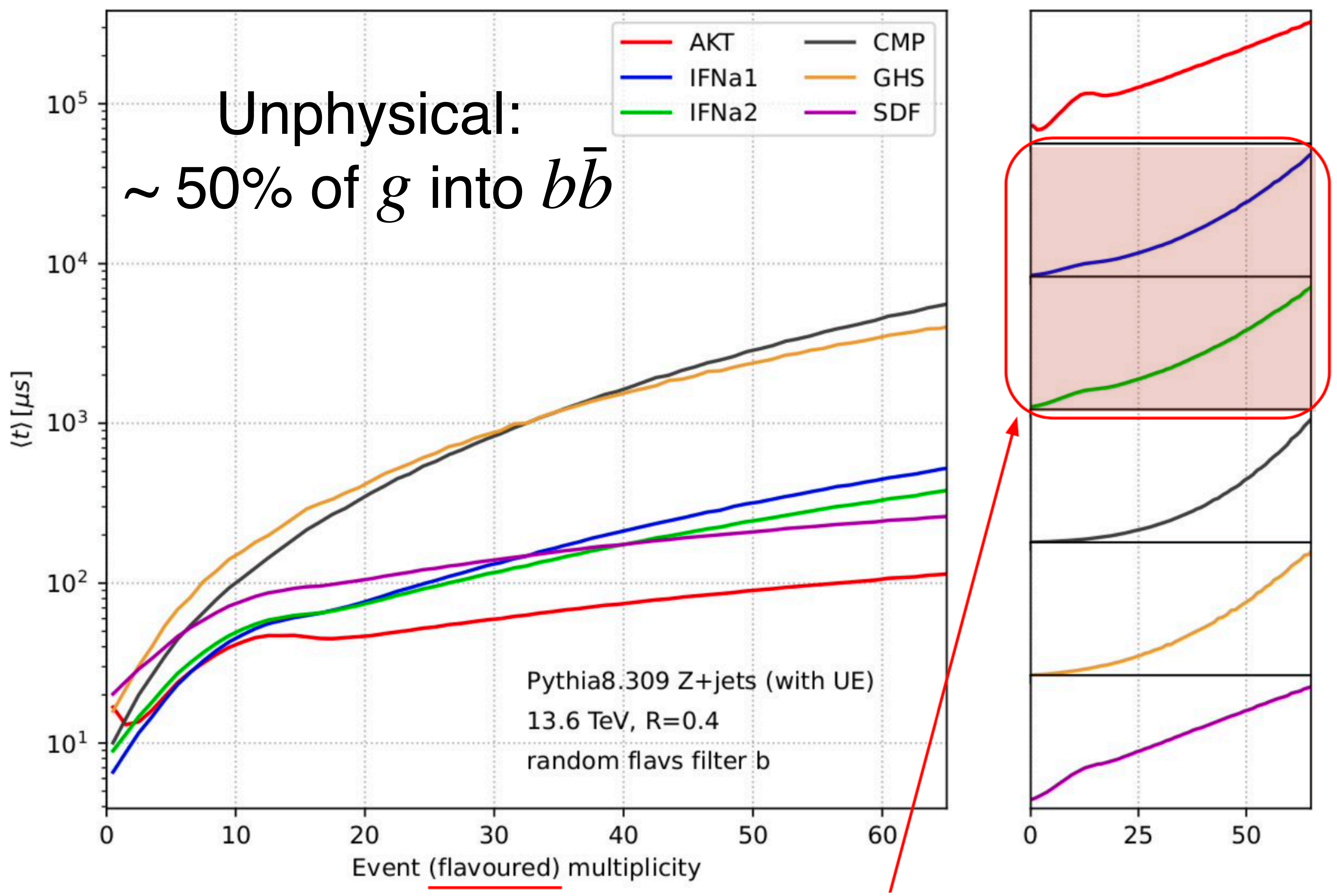


GHS/CMP slower with higher multiplicities
(some optimisation done in IFN, implementation of GHS/CMP to be improved)

Performance studies of current FastJet implementations

Clustering time of Z+jet Pythia8 events





Current implementations of GHS/CMP scale like N^2

Also IFN scaling is worsening with many flavoured particles

Conclusions (?)

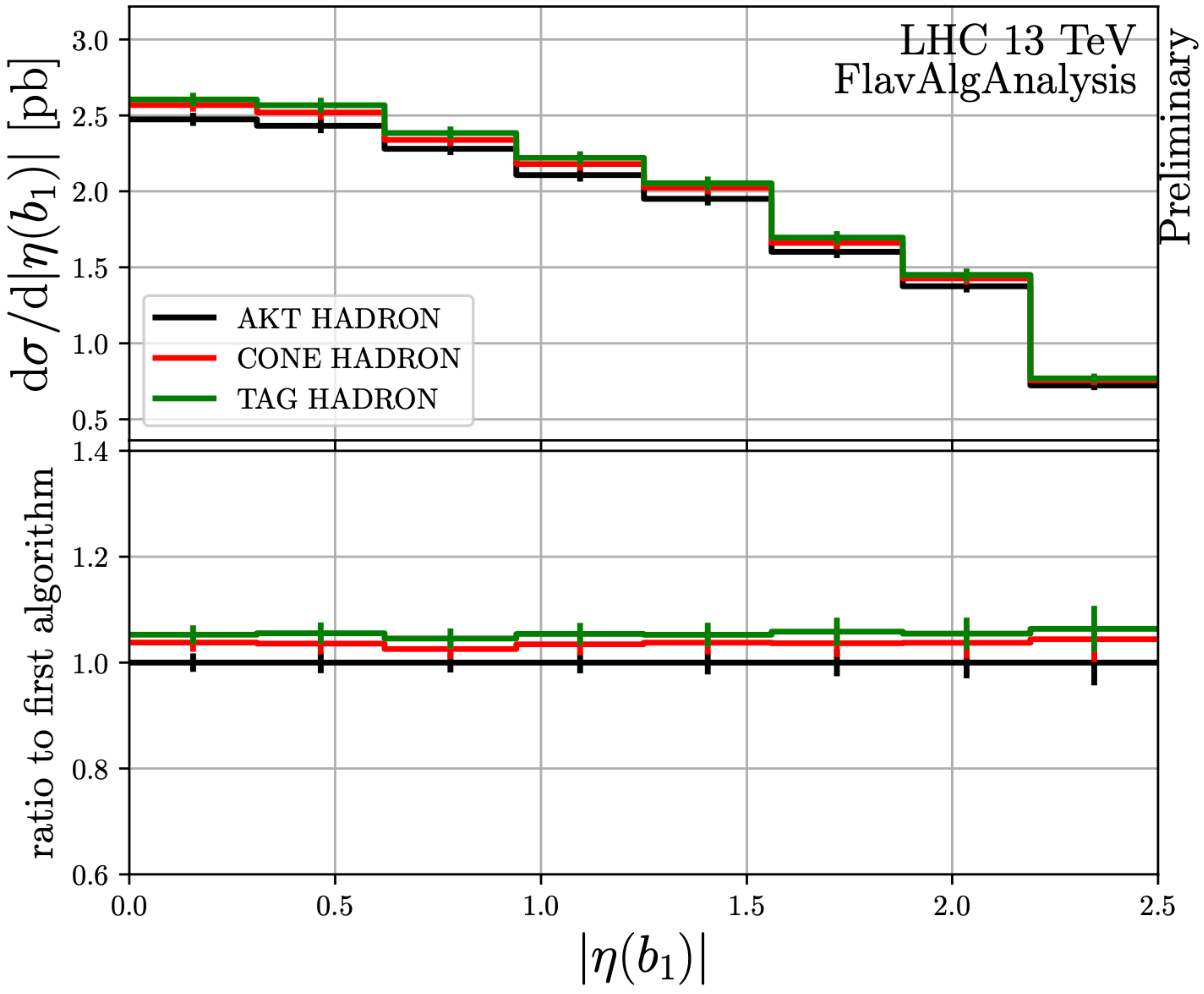
Many studies started in Les Houches, preliminary results available,
to be finalised and collected in a dedicated paper

Comparison between algorithms crucial
to assess size of MC corrections needed for theory-data comparison
and to explore adoption of “improved” labels for flavour tagging

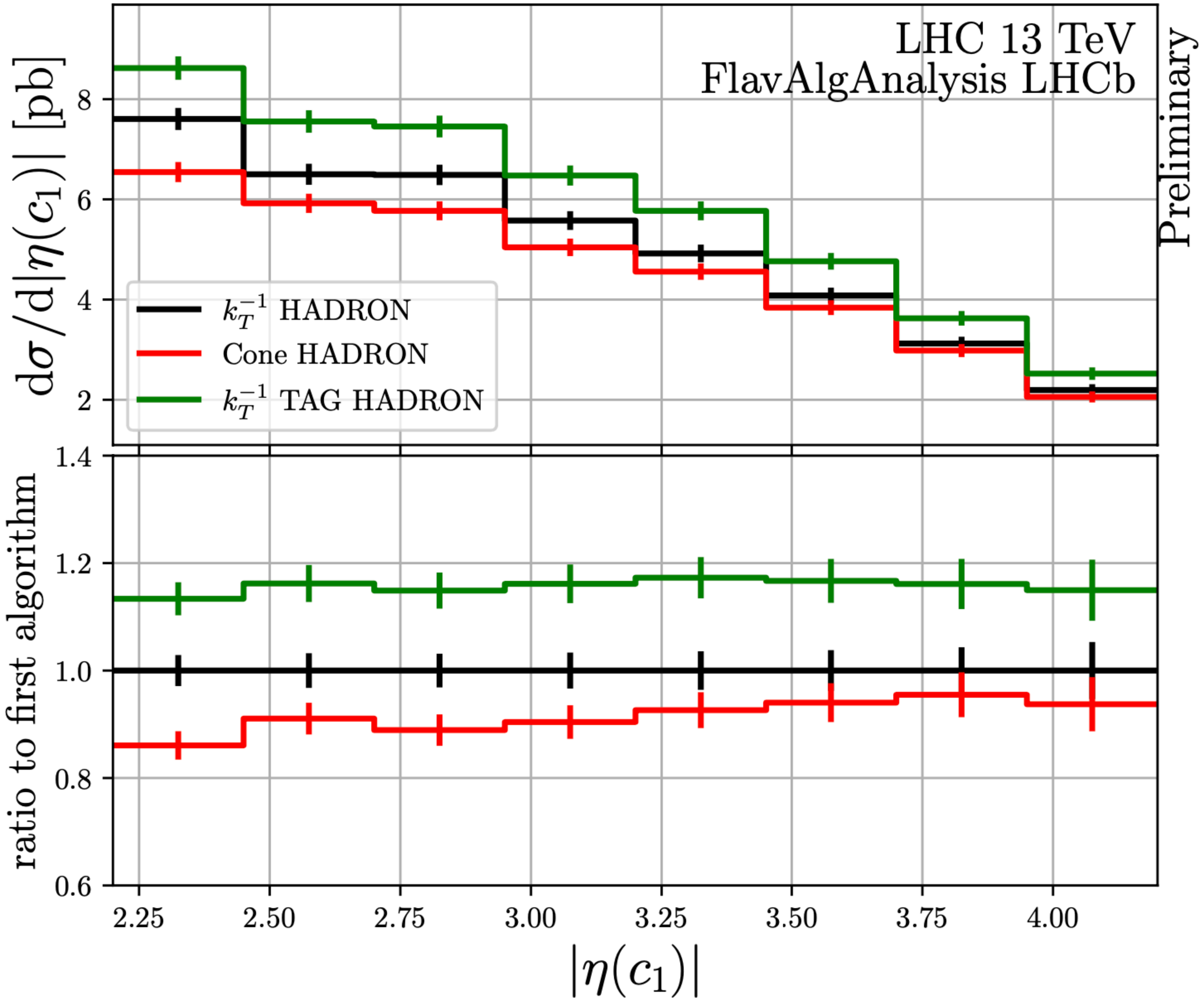
Next steps:
- Compare fixed-order predictions with NLO+PS ones
- Explore NLO+PS samples with other showers e.g. Herwig
- …

Many thanks to Ezra, Gavin, Ludo, Rene for providing me with material for this talk!

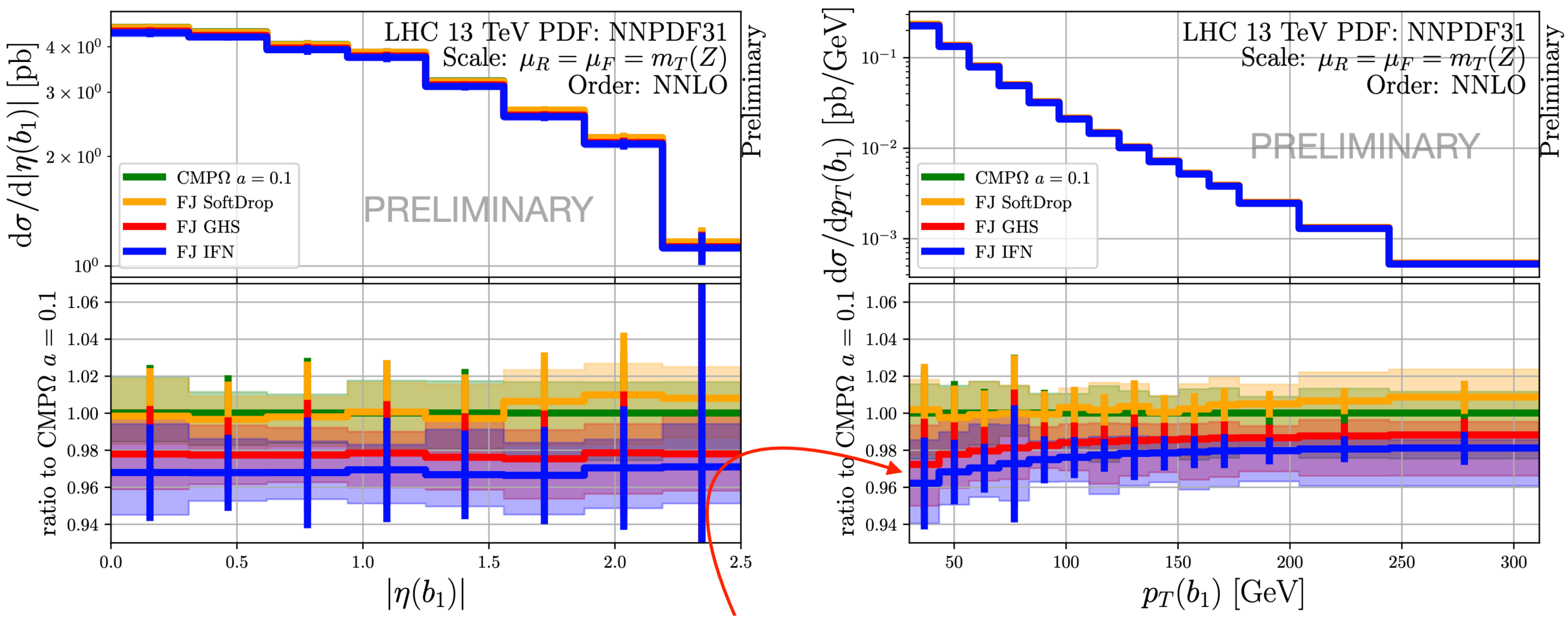
BACKUP



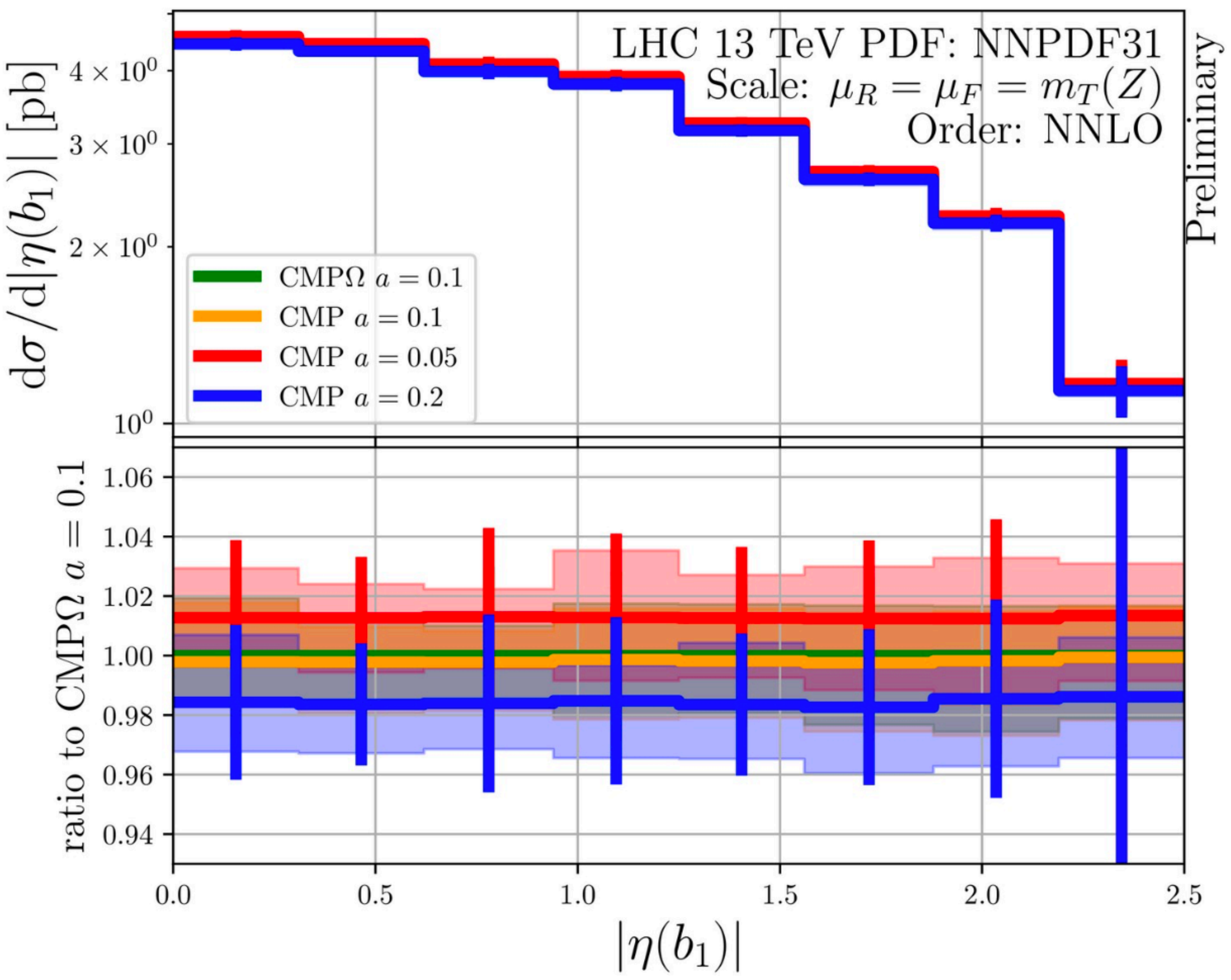
LHC 13 TeV
FlavAlgAnalysis
Preliminary
AKT HADRON
CONE HADRON
TAG HADRON
$d\sigma/d|\eta(b_1)|$ [pb]
ratio to first algorithm
$|\eta(b_1)|$



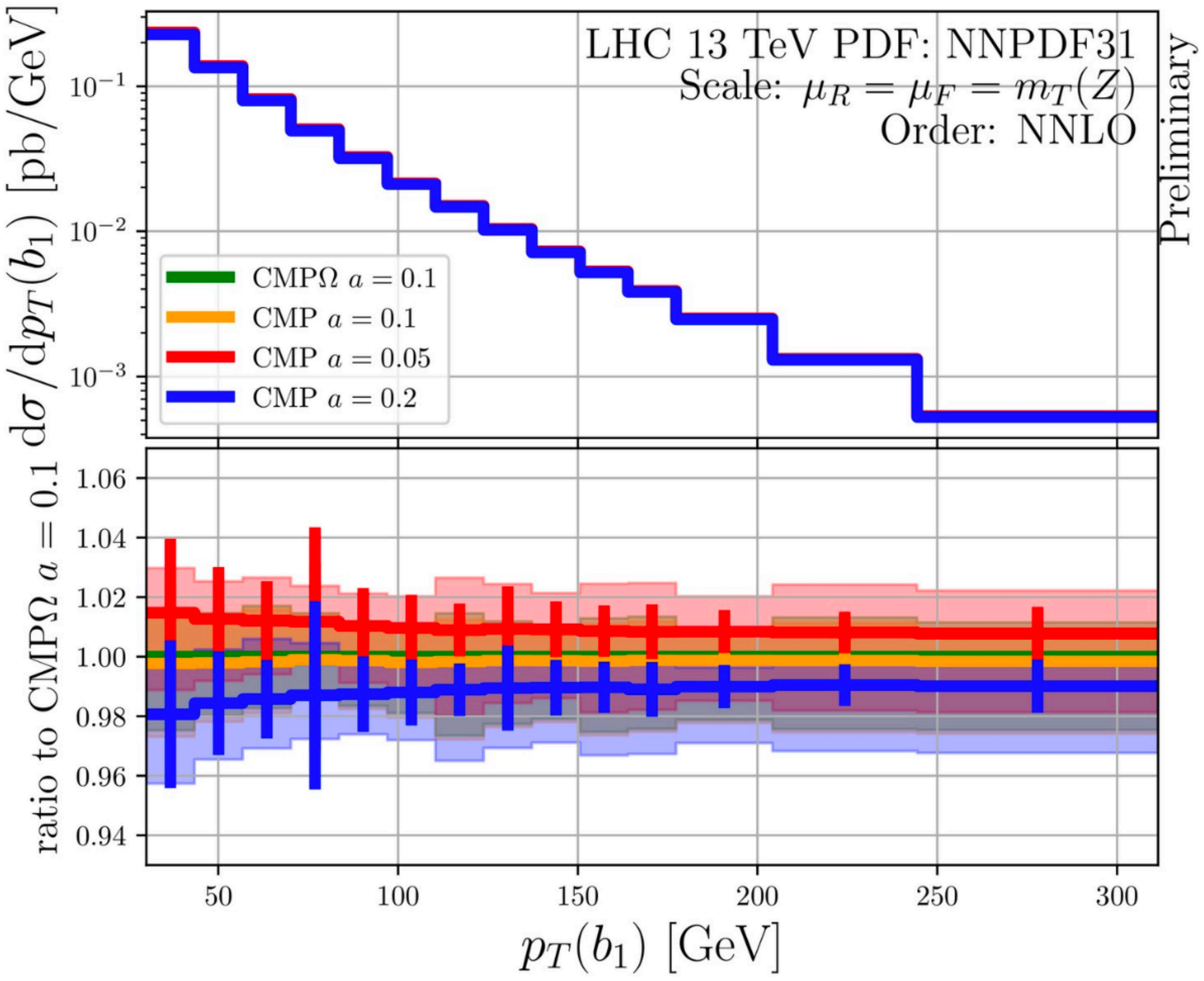
LHC 13 TeV
FlavAlgAnalysis LHCb
Preliminary
k_T^{-1} HADRON
Cone HADRON
k_T^{-1} TAG HADRON
$d\sigma/d|\eta(c_1)|$ [pb]
ratio to first algorithm
$|\eta(c_1)|$



LHC 13 TeV PDF: NNPDF31
Scale: $\mu_R = \mu_F = m_T(Z)$
Order: NNLO
Preliminary
PRELIMINARY
CMPΩ $a = 0.1$
FJ SoftDrop
FJ GHS
FJ IFN
$\mathrm{d}\sigma/\mathrm{d}|\eta(b_1)|$ [pb]
ratio to CMPΩ $a = 0.1$
$|\eta(b_1)|$
$\mathrm{d}\sigma/\mathrm{d}p_T(b_1)$ [pb/GeV]
$p_T(b_1)$ [GeV]



LHC 13 TeV PDF: NNPDF31
Scale: $\mu_R = \mu_F = m_T(Z)$
Order: NNLO
Preliminary
$d\sigma/d|\eta(b_1)|$ [pb]
CMPΩ $a = 0.1$
CMP $a = 0.1$
CMP $a = 0.05$
CMP $a = 0.2$
ratio to CMPΩ $a = 0.1$
$|\eta(b_1)|$



LHC 13 TeV PDF: NNPDF31
Scale: $\mu_R = \mu_F = m_T(Z)$
Order: NNLO
Preliminary
$d\sigma/dp_T(b_1)$ [pb/GeV]
CMPΩ $a = 0.1$
CMP $a = 0.1$
CMP $a = 0.05$
CMP $a = 0.2$
ratio to CMPΩ $a = 0.1$
$p_T(b_1)$ [GeV]