Experimental aspects of jet flavour definition

LHC EW WG General Meeting – 10th July 2024

Federico Sforza (University and INFN Genoa)

INFN ATLAS EXPERIMENT Università di Genova

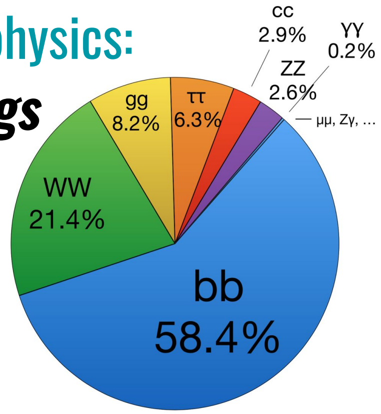
The importance of Beauty anche Charm (quarks)

Heavy-flavour (HF) quarks, *b* & *c*, are important for exploring fundamental physics:

⇒ *A window on the Higgs properties:*

- Dominating $H \to b\bar{b}$ decay, allow high stat. test of rare production modes
- Test of Yukawa coupling to 2nd generation using $H \to c\bar{c}$ decay

Higgs BR:



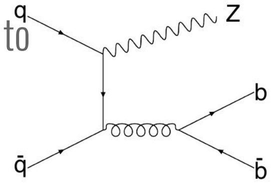
⇒ *Crucial for top-quark identification and present in many BSM models*

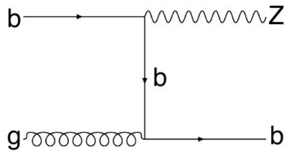
⇒ *Probe of complex QCD processes:*

- How to include HF-quark mass in pQCD predictions?
- Reliability of state-of-the-art MC+PS generators at LHC?
- Test the HF content and models of proton PDFs?

E.g. LO for Z+*b* in QCD Flavour Schemes (*FS*):

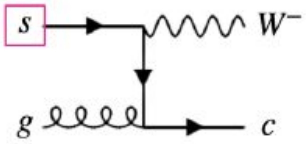
4FS: gluon splitting to massive *b*'s

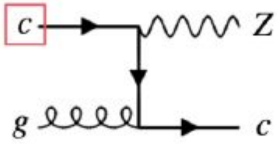




5FS: massless *b* in proton PDF

The proton strangeness





Intrinsic charm in proton PDF

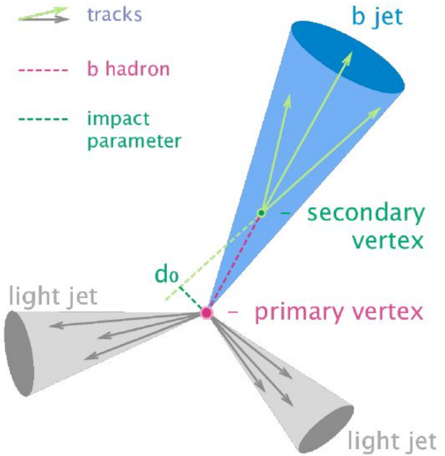
How LHC experiments deal with high-p_T HF-jet identification

HF-jet identification, i.e. tagging, relies on detectable HF-hadron characteristics inside reco-jets:

→ Long lifetime, *$O(10^{-12}$ s)*, and complex decay chains give secondary/tertiary decays displaced from primary vertex

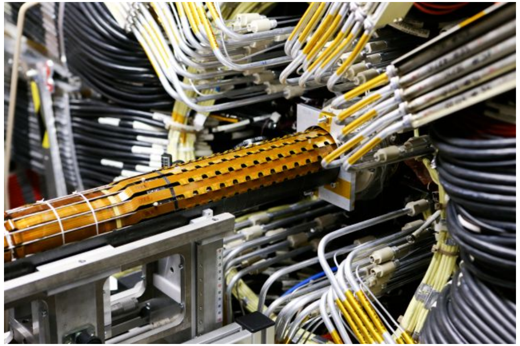
→ Inv. mass of B- & D-meson using charged particle tracks

→ Charged hadron trajectories with *impact parameter* > 0



→ Information recorded by state-of-the-art tracking detectors extremely close to beamline (E.g. ATLAS IBL or LHCb VeLo)

JINST 13 T05008 (2018)



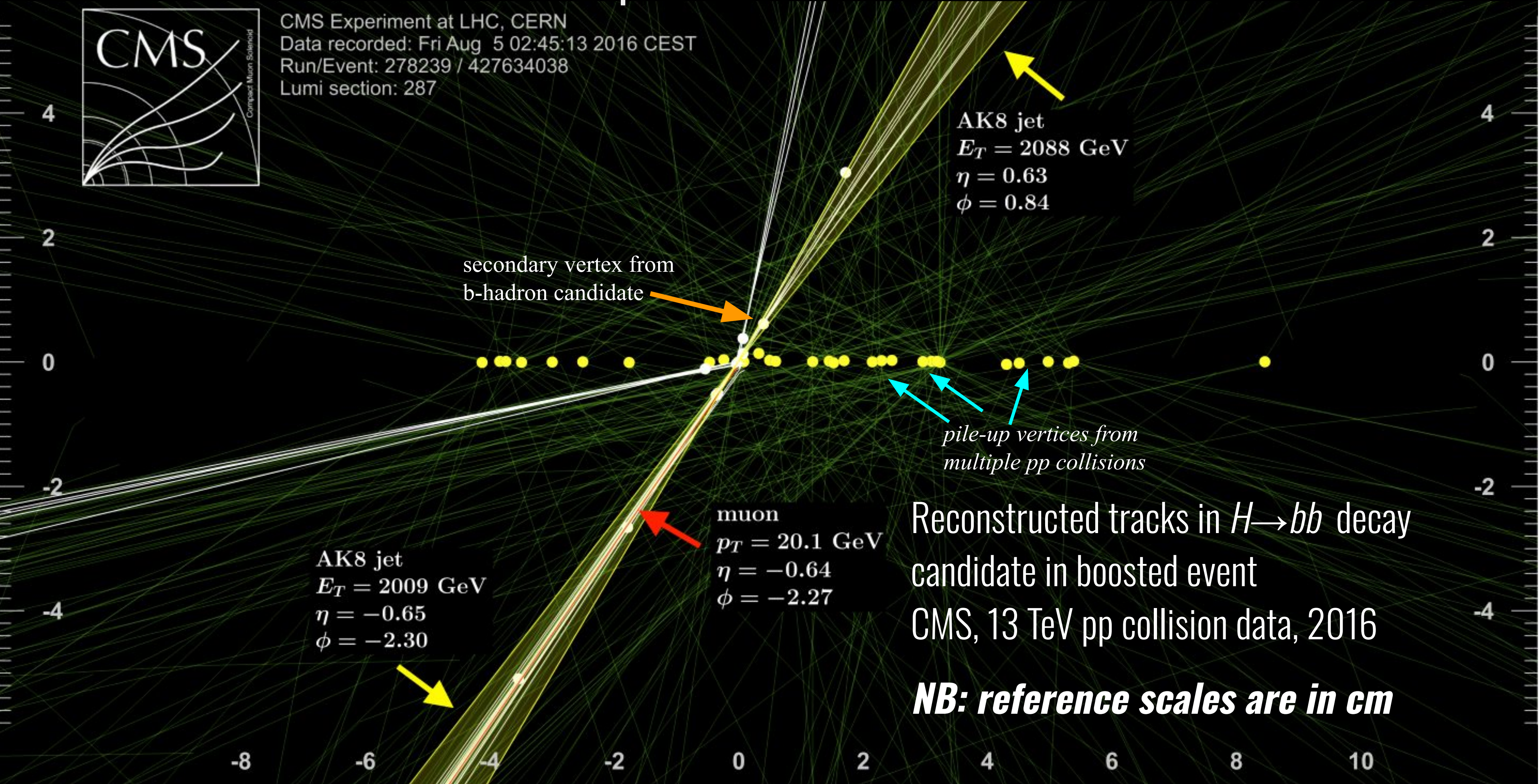
→ Often advanced machine-learning (ML) algorithms used to condense tracking & jet info for optimal separation ***b-* vs *c-* vs *light-jets***

*"Reco tag" added on top of **reconstructed & calibrated** anti-k_T jet*

"MC truth tag" added if ≥1 HF-hadron lies within anti-k_T jet

How does it look a high-p_T b-jet from the Higgs boson?

[CMS-DP-2017-032 (2017)]



Reconstructed tracks in *H→bb* decay candidate in boosted event

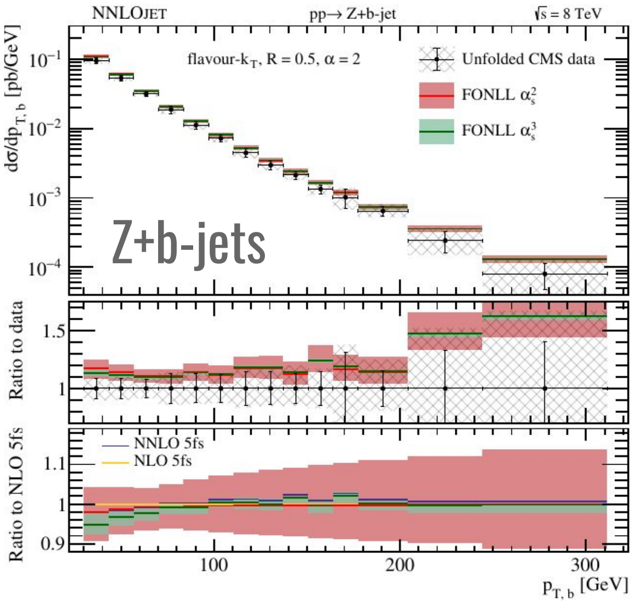
CMS, 13 TeV pp collision data, 2016

NB: reference scales are in cm

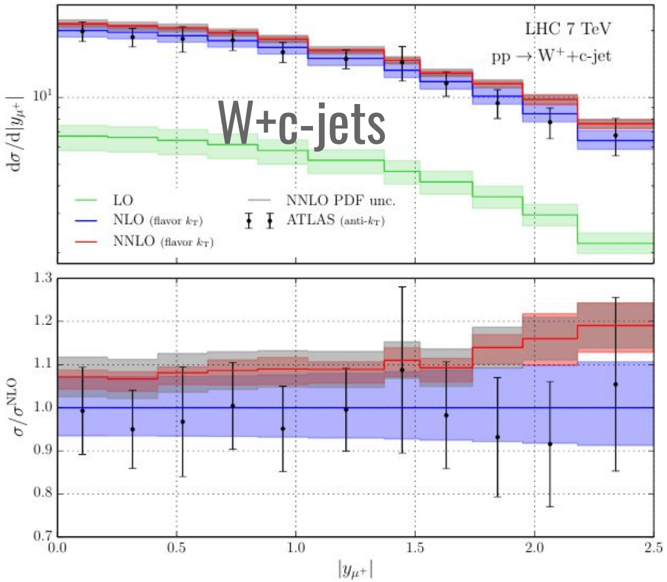
Flavour is now also part of QCD pheno NNLO revolution!

Beautiful NNLO predictions that started to become available since 2020!

Gauld et al. arXiv:2005.03016 [hep-ph] (2020)



Czakon et al. arXiv:2011.01011 [hep-ph] (2020)

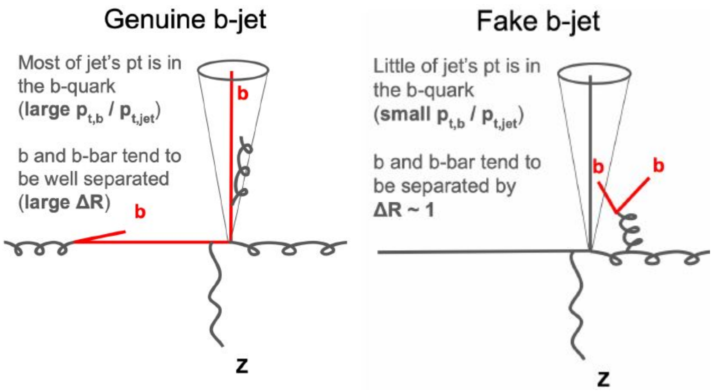


⇒ ***drastic reduction of scale uncertainties to ~% level going from NLO to NNLO !***

NB: *but misleading comparison against data*, which is unfolded using anti-k_T jets while predictions need ***IRC-safe flavour assignation…*** *here using* ***flavour-k_T jet algorithm*** [arXiv:hep-ph/0601139] (2006)

The (many) IRC safe (possible) solutions to jet flavour labelling

An algorithm is IRC safe if able to correctly label jet flavour in these and similar cases



Four IRC safe algorithms († including post-IRC safety test adaptations)

Flav-k_t hep-ph/0601139†	CMP 2205.11879†	Flav-Dressing 2208.11138† GHS	IFN 2306.07314
modified k_t-like distance when quark is softer	modified anti-k_t like distance for low-p_t quark pairs	after-burner on jets above p_t threshold	separates flavour-recomb. from kinematic recomb.
Flavoured jets have different effective radius & kinematics	Jets with flavour ≠ anti-k_t also have ≠ kinematics	Identical kinematics to reference alg.	Identical kinematics to reference alg.
replaces k_t alg	replaces anti-k_t alg	works with anti-k_t, C/A & k_t	works with anti-k_t, C/A (incl. substructure)
Banfi, GPS, Zanderighi	*Czakon, Mitov, Poncelet*	*Gauld, Huss, Stagnitto*	*Caola, Grabarczyk, Hutt, GPS, Scyboz, Thaler*

Flavour-k_T algorithm cannot be used for experimental jet definition ⇒ push from pheno community for a (maybe too wide) set of IRC-safe & (possibly) experimentally suitable jet flavour labelling algorithms!

Issues: *each jet algo has 1-3 free settings, may need anti-k_T jet changes, may need extensive truth-info, etc..*

A workshop to use and discuss new flavoured jets algorithms

Feasibility studies, algorithm comparisons, and best practices need to be developed in order to bridge experimental and theoretical communities ⇒ Workshop@IPPP organized last month!

Flavoured Jets at the LHC

Jun 11, 2024, 9:00 AM → Jun 12, 2024, 6:00 PM Europe/London

PH8 (Physics)

Joey Huston (Michigan State University), Michael Spannowsky (IPPP, Durham University), Simone Marzani (Università di Genova and INFN Sezione di Genova)

https://conference.ippp.dur.ac.uk/event/1301/

Description: Heavy-flavour jets play a central role in LHC phenomenology, ranging from Higgs boson physics, searches for new particles, as well as measurements of Standard Model properties.

Rapid progress on the experimental side has been matched by novel high-precision theoretical calculations, which include NNLO corrections in QCD. In order to exploit this progress, new jet algorithms that allow for jet flavour labelling in an infra-red and collinear (IRC) safe way have been proposed.

This workshop aims to bring together theorists and experimentalists working on this topic to assess the experimental feasibility of these new jet algorithms quantitatively.

Summarizing later a few findings of the many, still in progress, studies done by colleagues for the workshop:

⇒ **Test of jet flavour labelling effects in various contexts**

NB: results mostly taken from authors, no big work from my side

Analysis of jet-flav algorithms in practical terms

Studies by
R. Grabarczyk: talk

How to test the effect of new IRC safe flavour-labels in experiments?

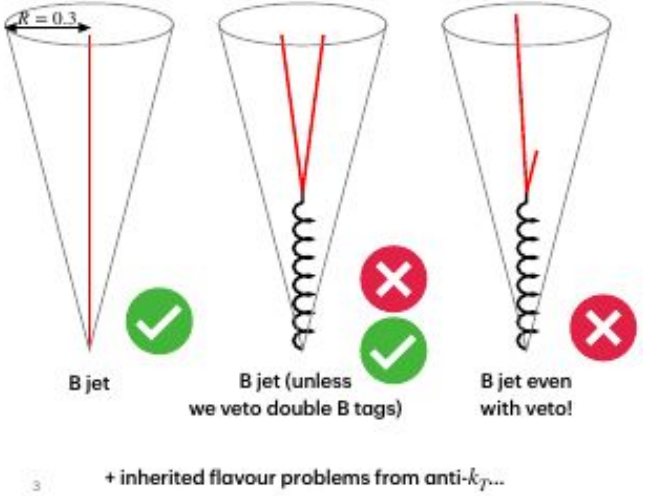
Compare vs default jet-label used by experiments in HF-tagging ML training in MC with b-/c-jets

Default ATLAS cone-label: HF-hadrons with $p_{T,HF}$ > 5 GeV and ΔR(jet,HF-had.) < 0.3.
Label anti-k_T jet as b-jet if ≥ 1 if b-had. is found, label anti-k_T jet as c-jet if ≥1 c-had. is found and no b-had

Default CMS ghost matching: HF-hadrons with $p_{T,HF}$ > 5 GeV.
Set p_T to 10^{-20} GeV, count HF-content in anti-k_T jet after new clustering

Parameters choice for IRC safe algorithms in FASTJET:

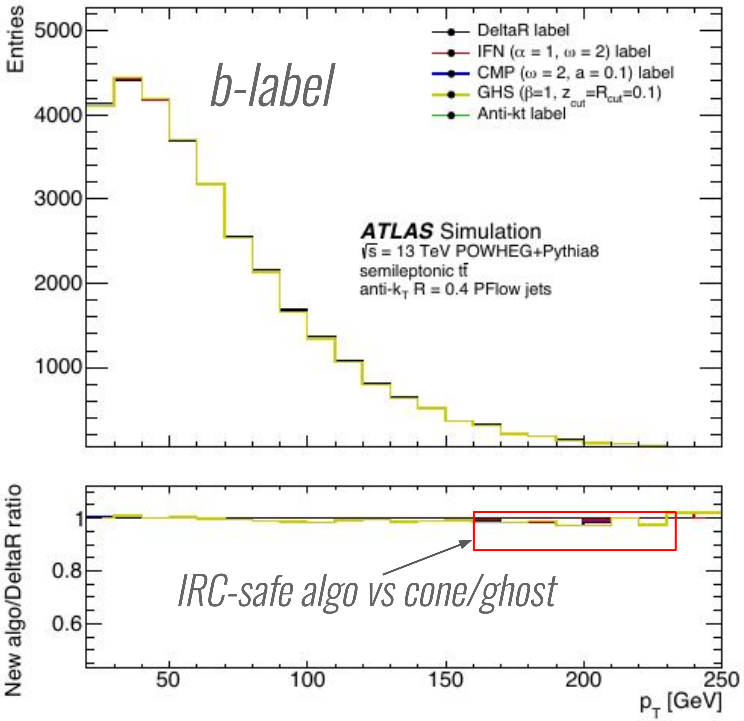
- Flavour anti-k_T (CMP): $\omega = 2, a = 0.1$
- Flavour dressing (GHS): $\beta = 1, R_{\mathrm{cut}} = z_{\mathrm{cut}} = 1$
- Interleaved Flavour Neutralisation (IFN), $\omega = 2, \alpha = 1$

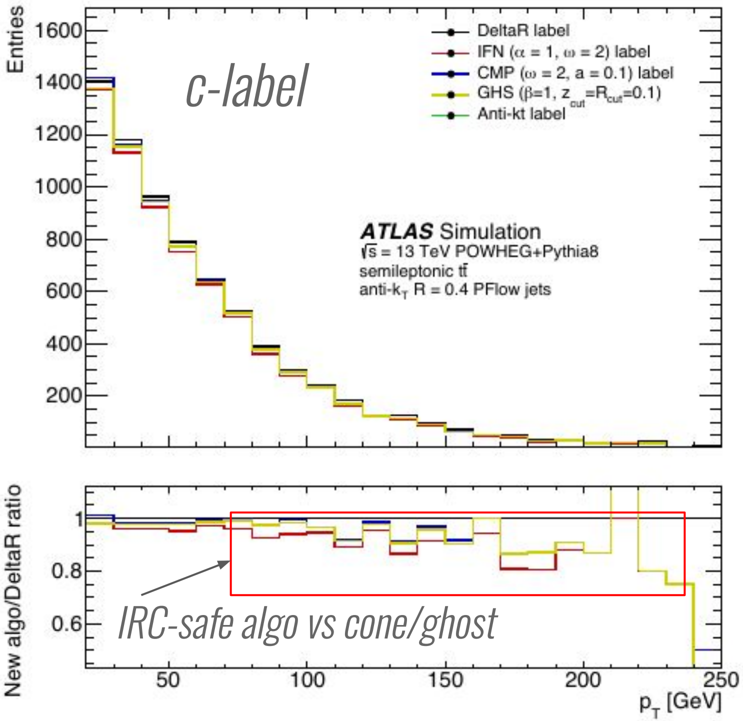


Analysis of jet-flav algorithms in Top-quark decay

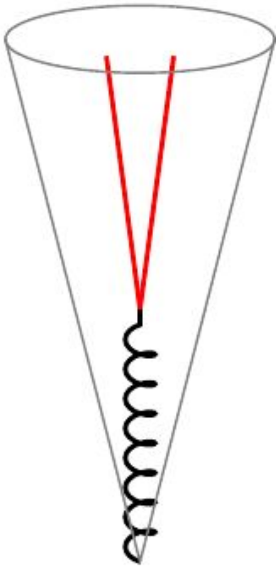
→ *Jets from top decay are well separated and do not origin from gluon*

→ *However small differences appear at medium p_T for b-jets, and earlier for c-jets*





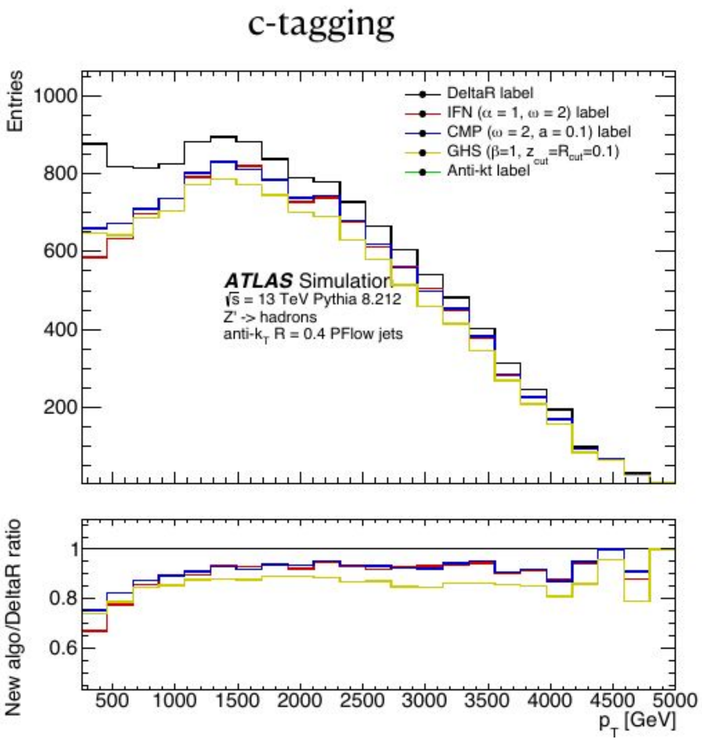
Difference comes from:

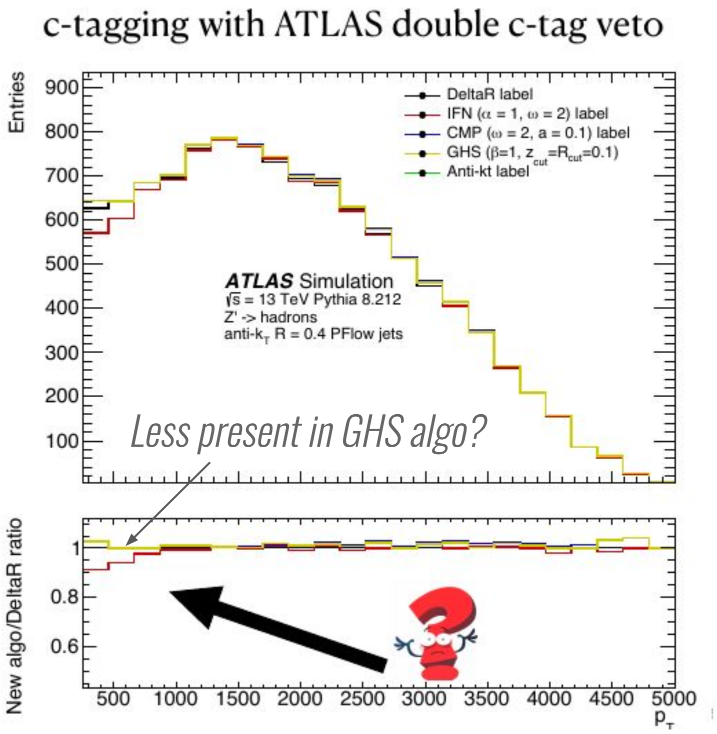


Analysis of jet-flav algorithms for high-mass Z'→bb or Z'→cc

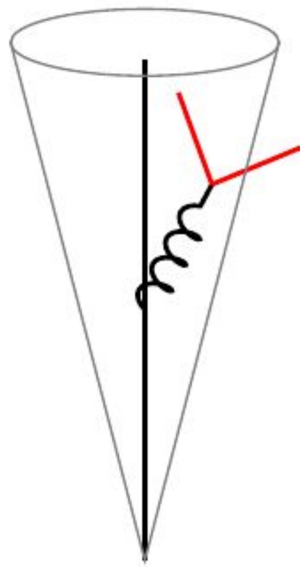
→ Similar pattern for Z' →bb or Z' →cc, *but additional effect at low p_T* (stronger for c-jets)

→ Reduced after veto of g→cc in c-jet (N=2 c-had.) ⇒ *Origin from g→cc with out-of-jet emission!?*





Difference from:

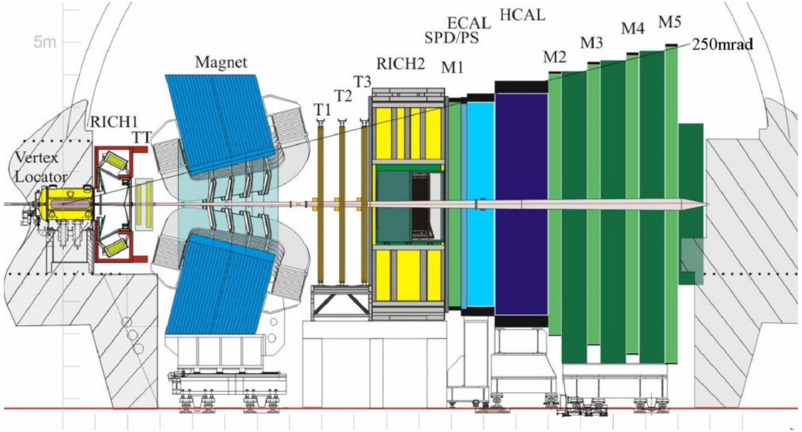


An other side of jet-flavour comparison: LHCb exclusive states

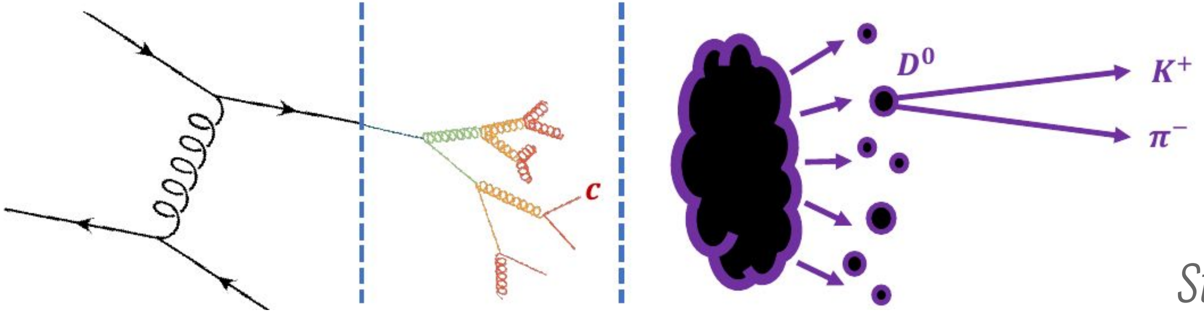
→ LHCb has special kinematic coverage: $2 < \eta < 5$

→ Excellent single-hadron reconstruction and ID

→ Common LHCb analyses select QCD multi-jet final states for exclusive HF-decay reconstruction



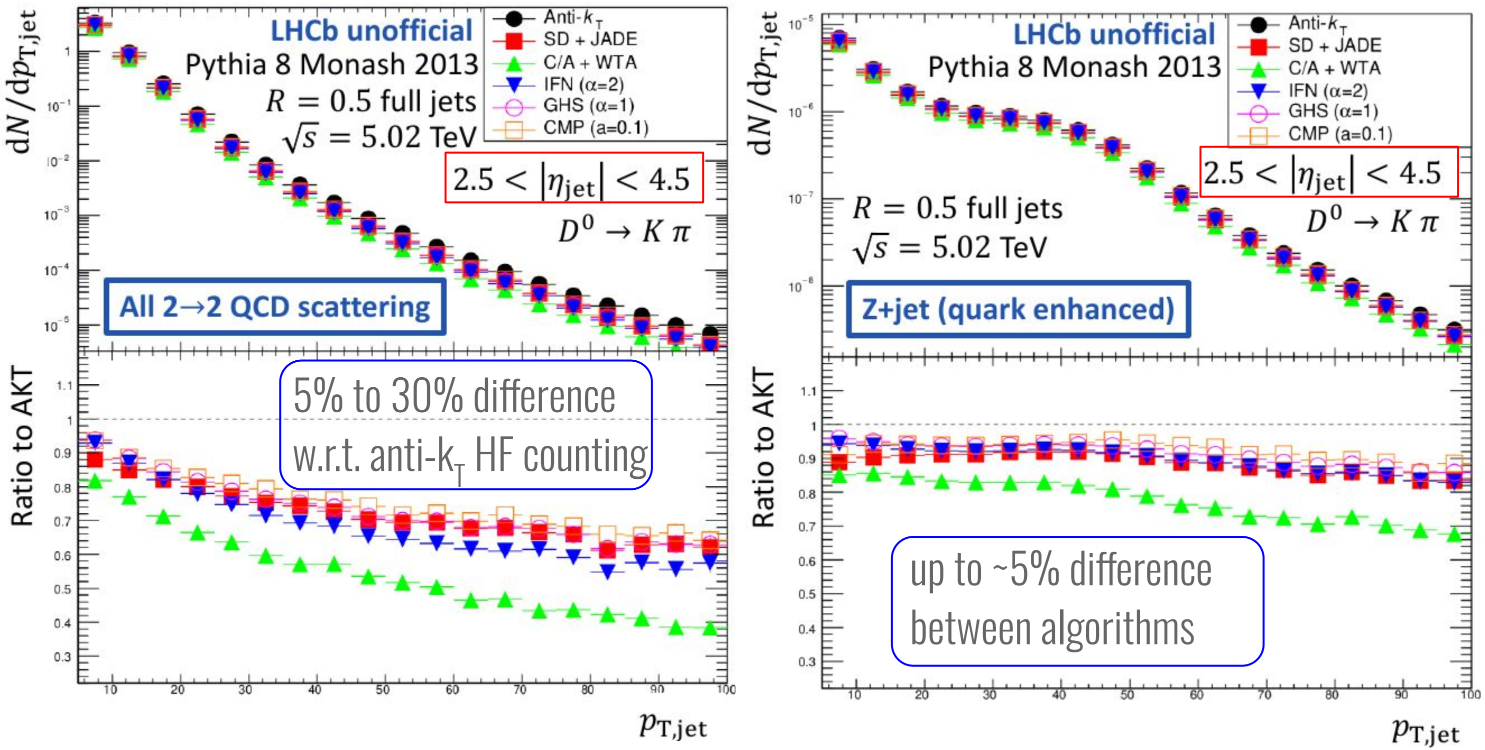
→ Testing IRC-safe jet-flavour algorithm in c-jets with exclusive decay: $D^0 \to K^\pm \pi^\mp$



Py8 used to generate LO 2→2 MC+PS events for di-jet and Z+jets production in $2 < \eta < 5$

Studies presented by E. Lesser [talk]

An other side of jet-flavour comparison: LHCb exclusive states



NB: two additional jet-flavour label algorithm shown:

→ SD+JADE [arXiv:2205.01117]

→ C/A WTA [arXiv:2205.01109]

not fully IRC-safe

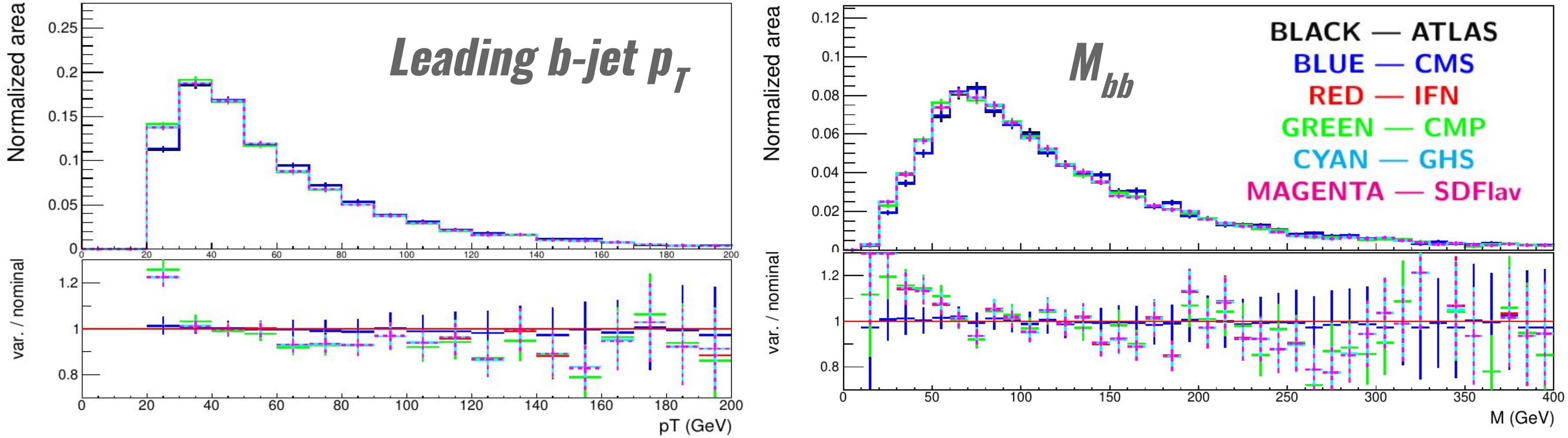
Striking difference depending on tested process ⇒ Consequence of quark/gluon content in MC!

Jet flavour test as close as possible to real measurements: Rivet

Testing jet-flav. algorithms in Z+bb "ATLAS-style" Rivet analysis on Z+jets MGaMC LO+PS:

1. Anti-k_t jets with CMS-style ghost flavour labelling
2. Anti-k_t jets with ATLAS-style ΔR flavour labelling
3. IFN algorithm
4. CMP algorithm
5. GHS algorithm
6. SDFlav algorithm

NB: work in progress, by A. Rescia [talk]



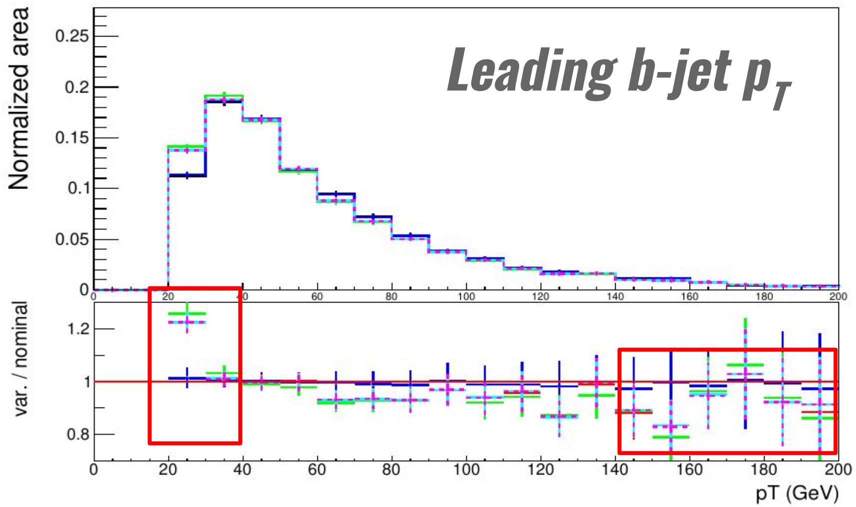
Jet flavour test as close as possible to real measurements: Rivet

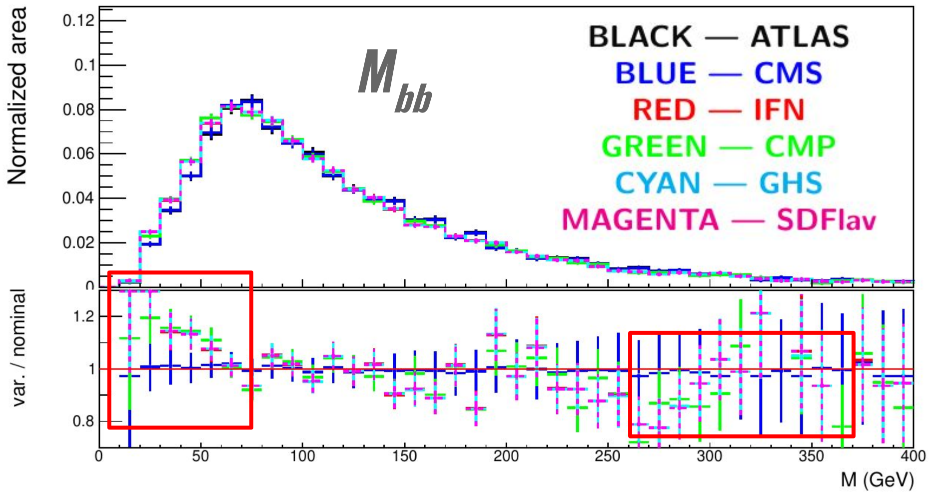
Difference observed at the level of clustered jet kinematic:

⇒ *Reason at low p_T* : anti-k_T jets include HF-hadron decay products (untouched by ghost/cone labelling). *IRC-safe jet-flav. labelling needs undecayed HF-hadrons*

⇒ *Reason at high p_T* : g⟶bb splitting assigned as light-flavour jet

NB: work in progress, by A. Rescia [talk]





Last example: Z+HF measurement compared to NNLO predictions

Z+HF measurement [arXiv:2403.15093 [hep-ex]] used IFN flavour-jet algorithm implemented in FASTJet for comparison vs NNLO predictions calculated with the same IRC-safe algorithm

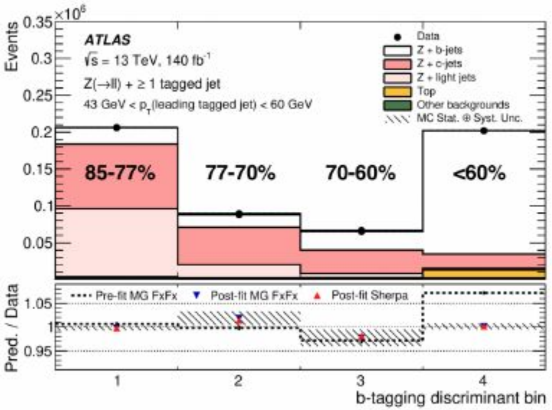
→ Using full Run 2 data and pseudo continuous b-tagging, for kinematic-dependent background fit

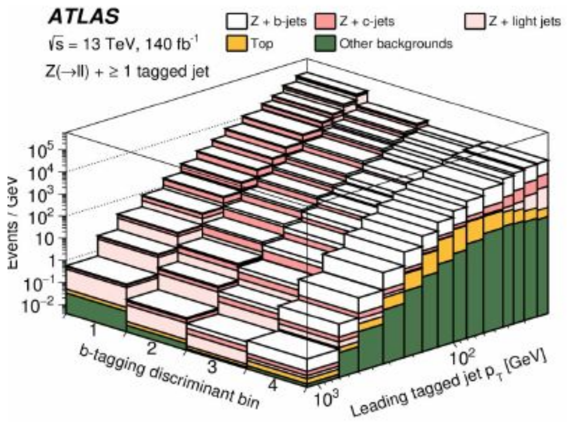
→ Multi-differential measurement, *3-10% precision*, of: ***Z + ≥ 1 b-jet***, ***Z + ≥ 2 b-jets***, ***Z + ≥ 1 c-jet***

→ Measurement shows mild sensitivity to Intrinsic Charm (IC)

→ Analysis presented at the Durham workshop by Y.Yu [talk]

NB: paper just re-submitted, HEPData and Rivet soon available





Effect on ATLAS measured Z+≥ 1 b-jet cross section

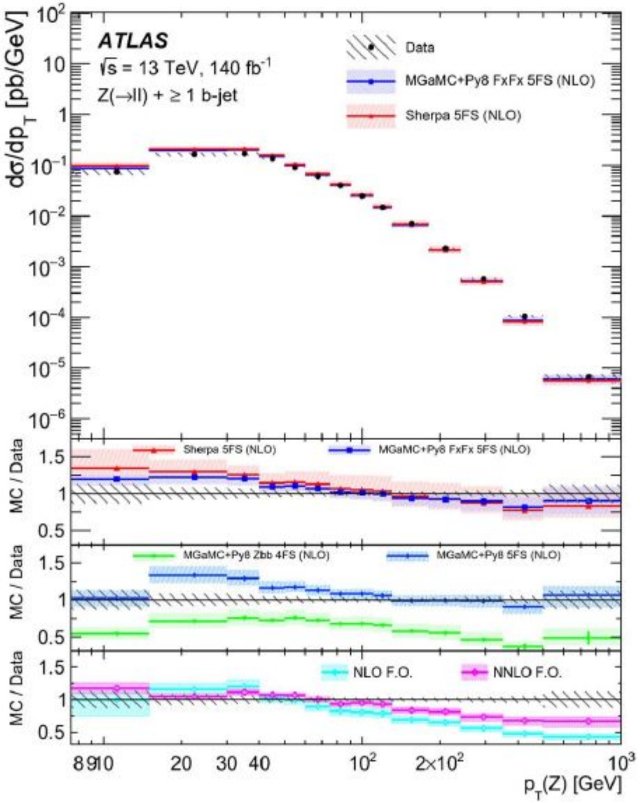
→ b-jet p_T: 20 GeV to ~1 TeV, compared to many predictions

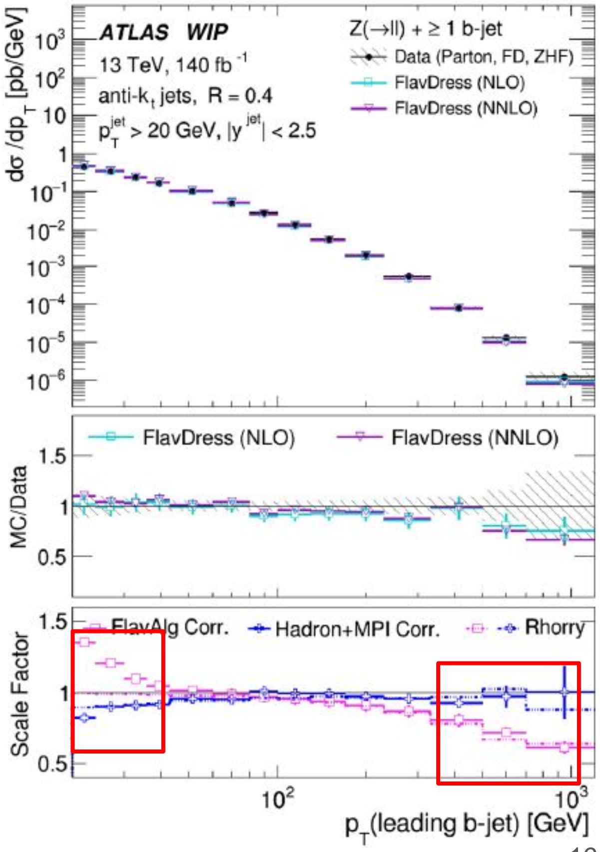
→ unfolding to anti k_T jets

→ 2 step correction of NNLO prediction to data unfolding:
1) parton IFN→hadron IFN
2) IFN hadron → cone-hadron

→ Sensitive to g→bb at high p_T

→ HF-decay clustering at low p_T (compared vs MC w/o HF-decay)





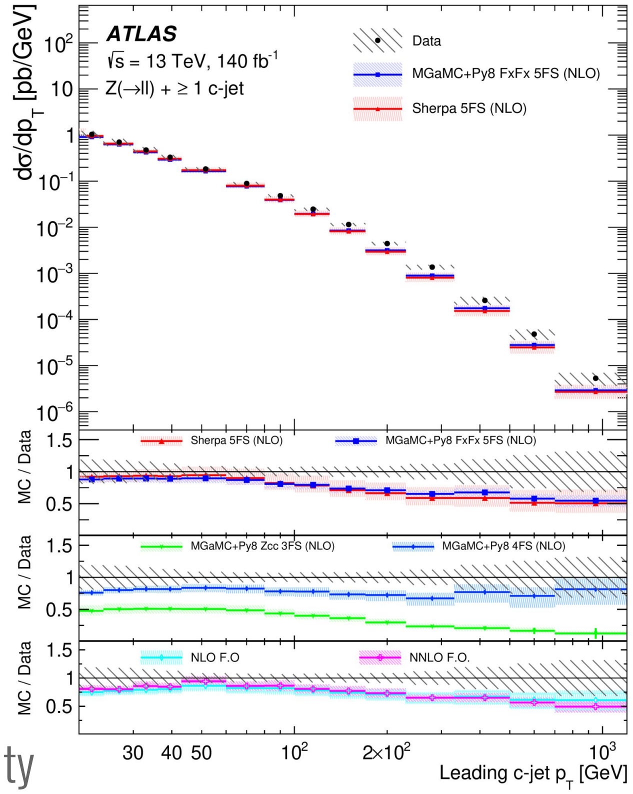
Effect on ATLAS measured Z+≥ 1 c-jet cross section

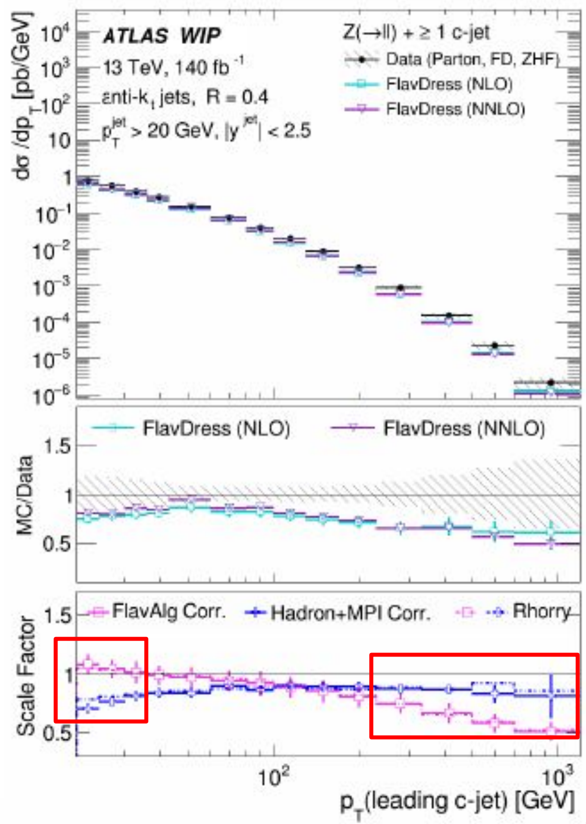
→ **c-jet p_T**: range from 20 GeV to ~1 TeV, compared to predictions after unfolding to anti k_T jets, NNLO comparison after jet-flav IFN algorithm 2-step correction

→ Minor effect of HF-decay clustering at low p_T (less energy in HF-cascade) but visible MPI effect

→ Sensitive to g→cc at high p_T sizable effect, up to 50%!

→ May be relevant for IC sensitivity





Summary and prospects

→ Jet flavour physics has rich and only partially explored landscape at the LHC

→ Experimental measurements and theoretical NNLO calculations are opening a new precisions era, but matching the two is not trivial...

→ The need of IRC-save jet flavour definition have lead to many options on the market, experiments started to use them

→ *Discussion just started:* choice of the "best" algorithm? Practical issues as computing time and parameter settings?

What's your favourite flavour?

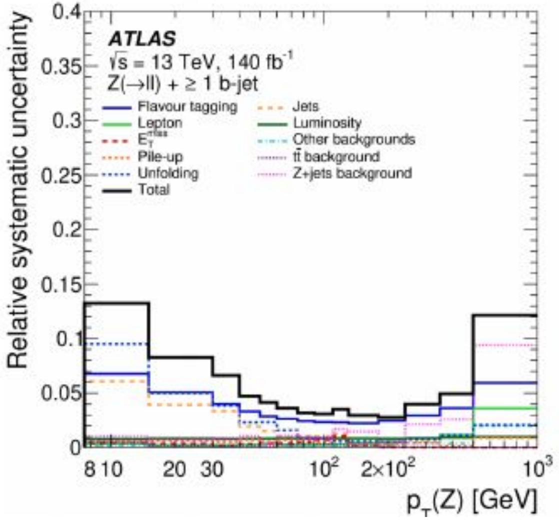
→ Test show sizable effect on observables depending on topology and (as expected) on g→qq contribution, however MC based studies needs to be compared to data: ***how?***

Extra slides & Backup Material

Uncertainties on the cross section measurements

- **x2 improved precision on Z + b-jets** measurements with respect to previous ATLAS results
- Dominant uncertainty contributions from
  flavour-tagging, jet energy scale and resolution and unfolding
- Statistical uncertainty on data <1%
  Differential distributions: total unc. **<5% in Z+≥1 b-jet, ~10-15% in Z+≥2 b-jets and Z+≥1 c-jet** for modest p_T

Source of uncertainty	$Z(\to \ell\ell) + \geq 1$ b-jet [%]	$Z(\to \ell\ell) + \geq 2$ b-jets [%]	$Z(\to \ell\ell) + \geq 1$ c-jet
Flavour tagging	3.6	5.7	10.3
Jet	2.4	4.3	6.5
Lepton	0.3	0.3	0.4
E_T^{miss}	0.4	0.5	0.3
Z+jets background	0.6	1.5	1.6
Top background	0.1	0.3	<0.1
Other backgrounds	<0.1	0.2	0.1
Pile-up	0.6	0.6	0.2
Unfolding	3.3	5.8	5.0
Luminosity	0.8	0.9	0.7
Total [%]	5.6	9.4	13.2



Inclusive cross-section results

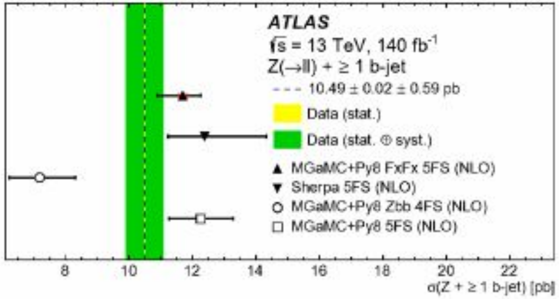
$\sigma(Z+\geq1$ b-jet$) = 10.49 \pm 0.02$ (stat.) ± 0.59 (syst.) pb

$\sigma(Z+\geq2$ b-jets$) = 1.39 \pm 0.01$ (stat.) ± 0.13 (syst.) pb

$\sigma(Z+\geq1$ c-jet$) = 20.89 \pm 0.07$ (stat.) ± 2.77 (syst.) pb

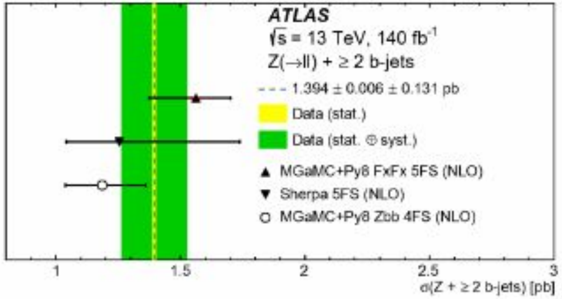
$Z+\geq 1\ b$-jet

- Good description from 5FS
- 4FS with large underestimation



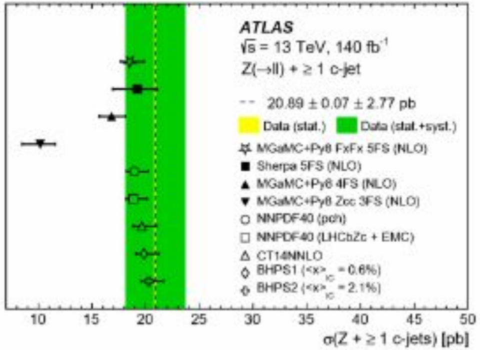
$Z+\geq 2\ b$-jet

- 4FS and 5FS agrees with data
- much sizable MHOU for Sherpa

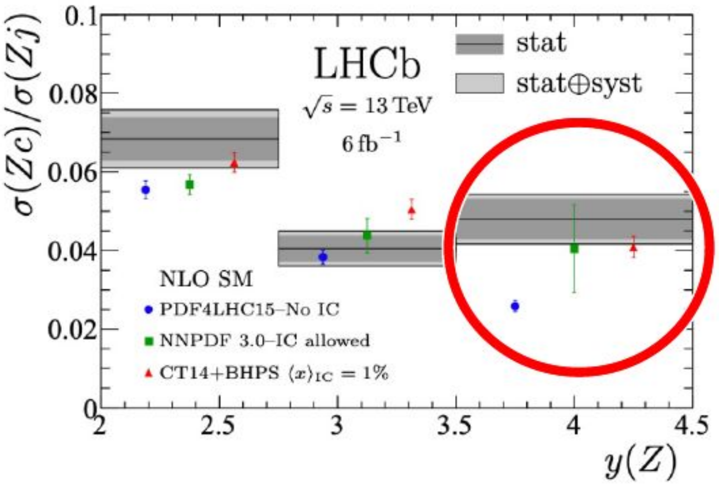


$Z+\geq 1\ c$-jet

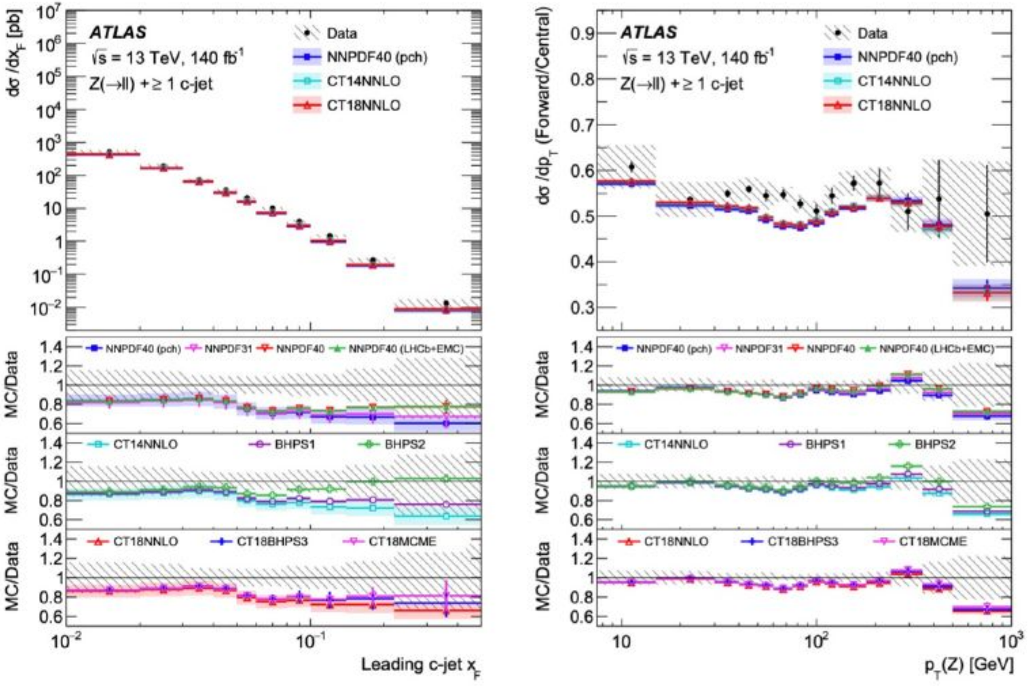
- 5FS in agreement with data
- 3FS with large underestimation



Probing HF content of the proton



LHCb 13 TeV, arXiv:2109.08084, PRL128 (2022)



(a) (b)

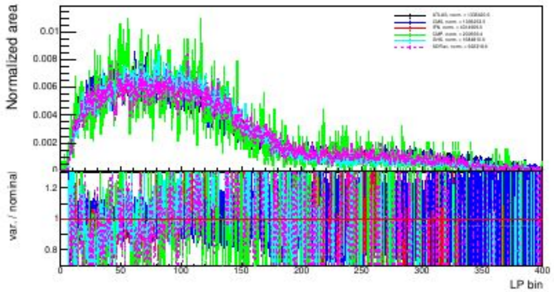
ATLAS13 TeV, Z+c-jet, 140 fb^{-1} arXiv:2403.15093

Is intrinsic charm present in the proton? May need a better understanding of the theory, and in particular the heavy flavor jet algorithm.

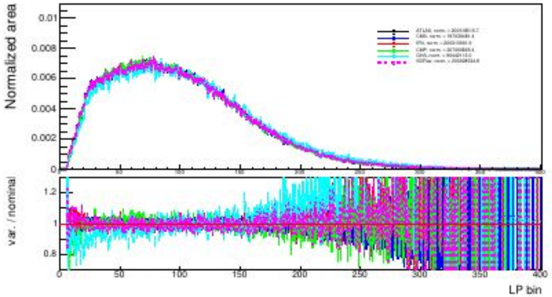
Testing JSS with different jet flav algorithms

Lund Jet Plane

$Z + b\bar{b}$ leading



$Z + jj$ leading



LJP seems to be mostly independent of choice of algorithm

GHS shows trend in $Z + jj$