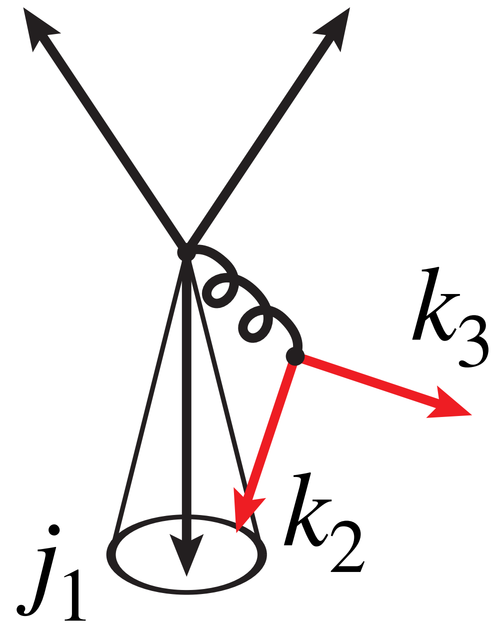



# Flavoured jets: trials and tribulations

Rhorry Gauld

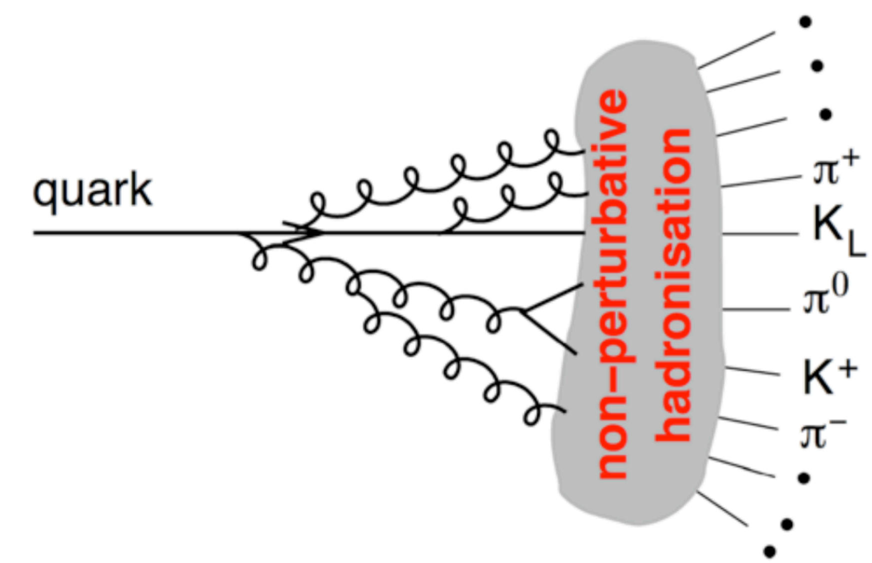
**Heavy** Flavours at High pT (30/11/23)

Higgs Centre for Theoretical Physics

MAX-PLANCK-INSTITUT FÜR PHYSIK

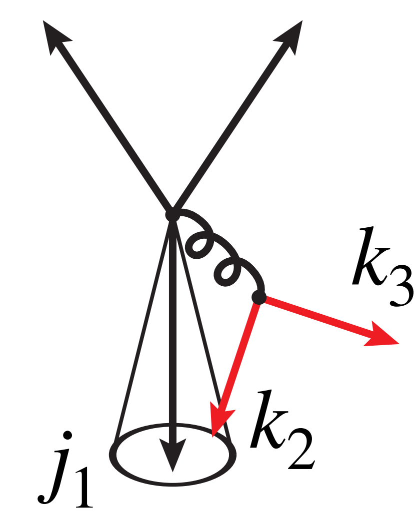
# 1) Introduction

- Jet reconstruction basics



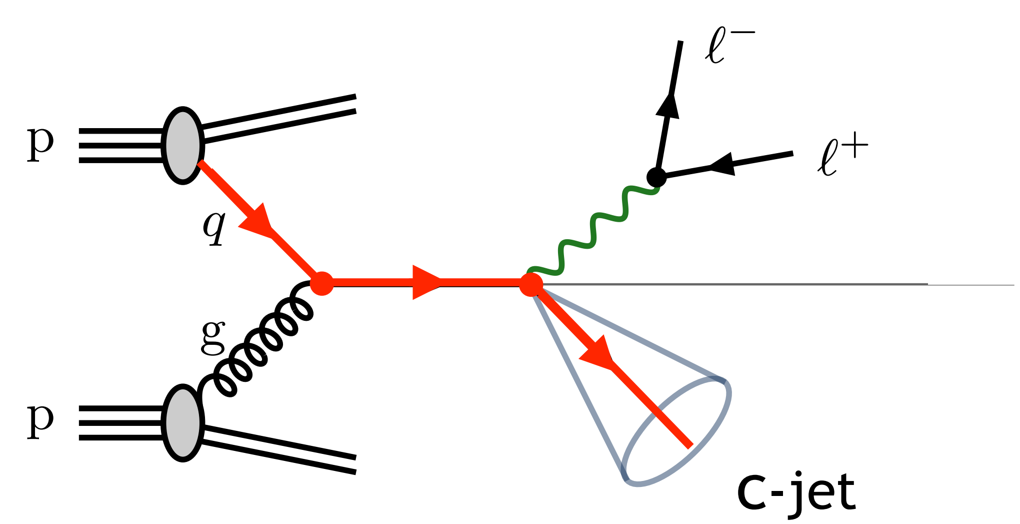

# 2) Jet flavour

- Selected histories
- Recent progress
- IRC tests




# 3) Implications/Applications

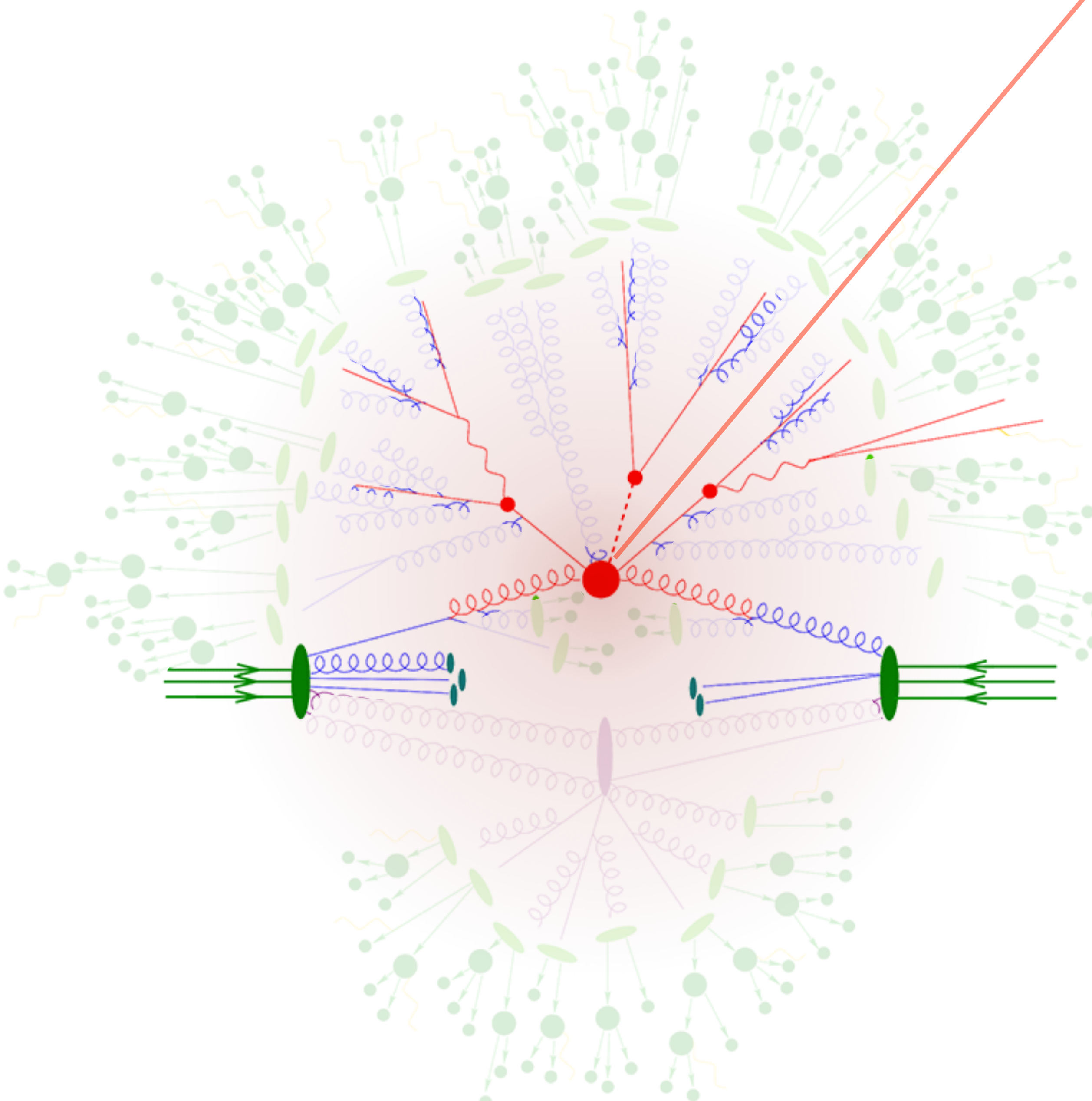
- Where does it (not) matter
- Z + c-jets



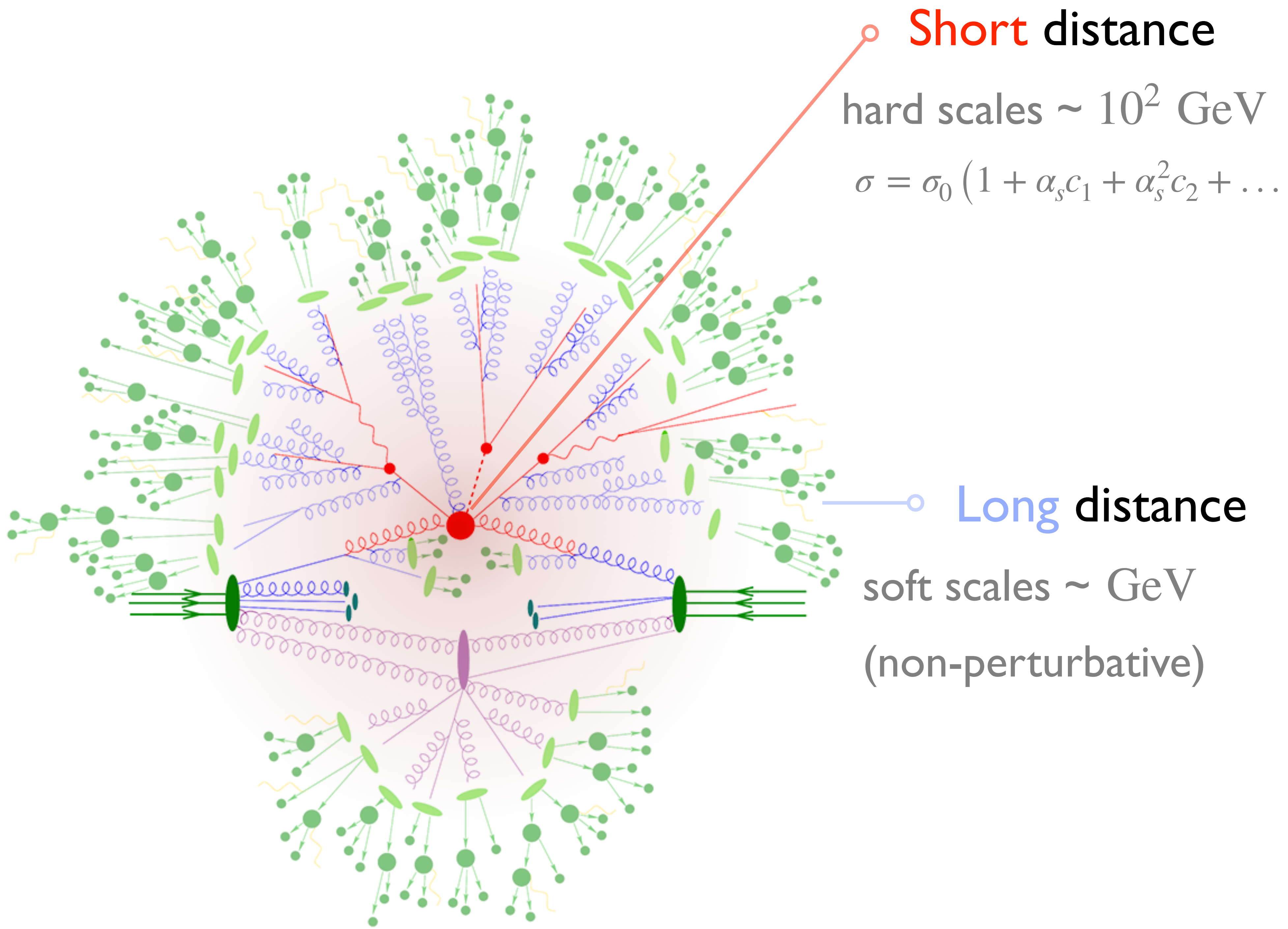
# **Short** distance

hard scales ~ $10^2$ GeV

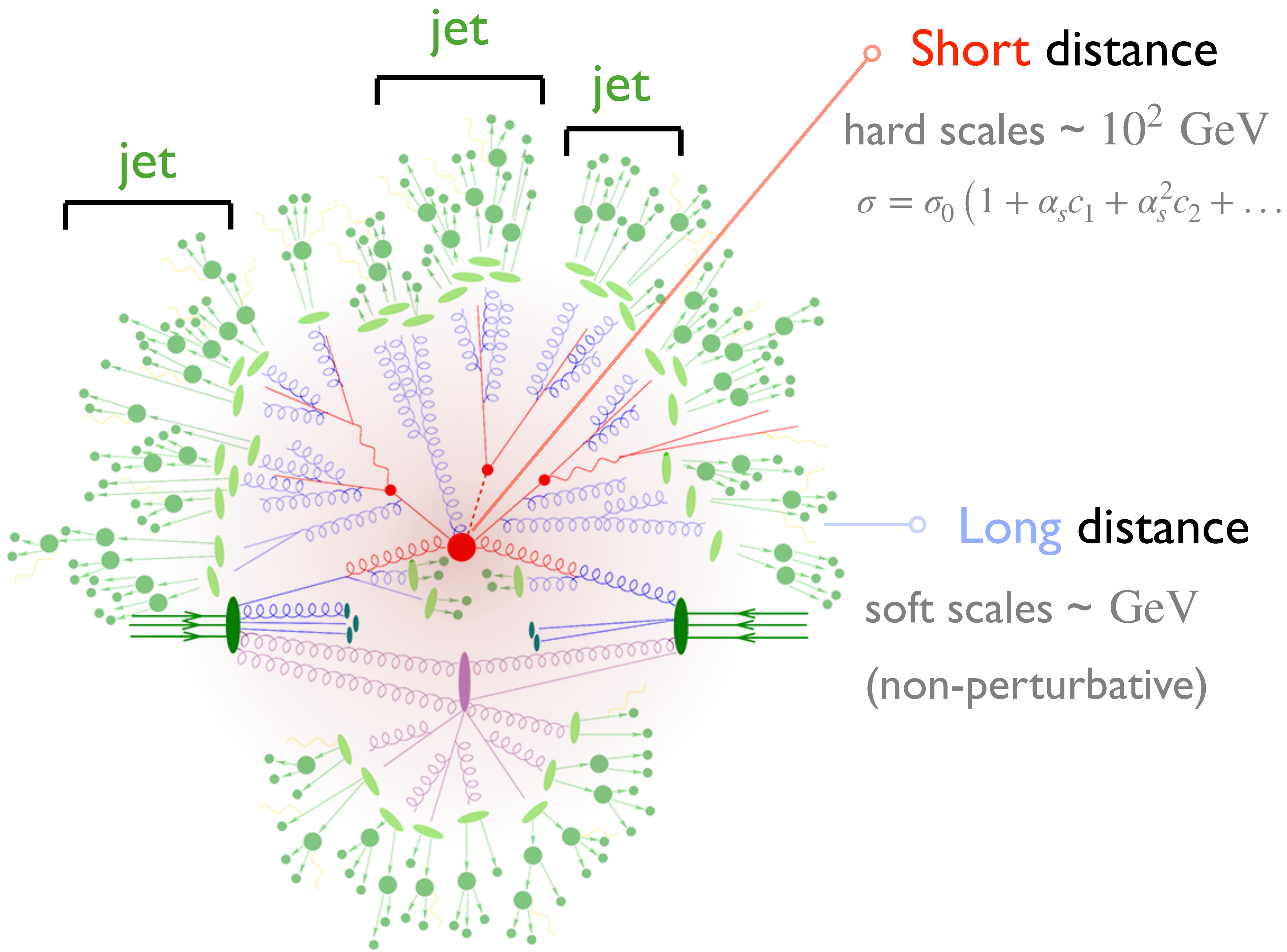
$$\sigma = \sigma_0 \left(1 + \alpha_s c_1 + \alpha_s^2 c_2 + \ldots\right)$$



$$PP \to f + X$$


Short distance
hard scales ~ $10^2$ GeV
$\sigma = \sigma_0 \left(1 + \alpha_s c_1 + \alpha_s^2 c_2 + \dots\right)$
Long distance
soft scales ~ GeV
(non-perturbative)

$$PP \to f + X$$




$$PP \to f + X$$

How to describe the QCD radiation in events?

a jet algorithm which is insensitive to Long distance physics

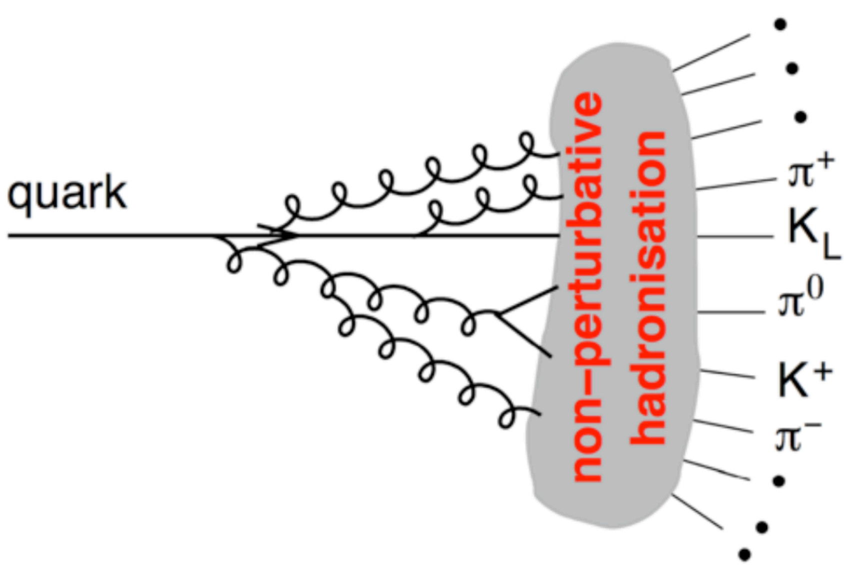
# Jets at the LHC

Experimentally: apply an algorithm to particle flow objects (Kaons, Pions,…)
(e.g. ATLAS arXiv:1703.10485, CMS arXiv:1706.04965, LHCb arXiv:1310.8197)

The anti-$k_T$ algorithm (Cacciari, Salam, Soyez arXiv:0802.1189) applied to these objects

Simple version

➡ Reconstruct hadronic jets (~collimations of hadronic radiation)




anti-$k_T$ algorithm → set of jets $\{j_1, \ldots, j_m\}$

particle flow objects

# Jets at the LHC

Experimentally: apply an algorithm to particle flow objects (Kaons, Pions,…)
(e.g. ATLAS arXiv:1703.10485, CMS arXiv:1706.04965, LHCb arXiv:1310.8197)

The anti-$k_T$ algorithm (Cacciari, Salam, Soyez arXiv:0802.1189) applied to these objects

Or… initialise a list of particles (pseudo jets) from these objects

Introduce distance measures between particles (pseudo jets) and a Beam:

$$d_{ij} = \min\left(k_{Ti}^{2p}, k_{Tj}^{2p}\right) \frac{\Delta R_{ij}^2}{R^2} \qquad \Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

$$d_{iB} = k_{Ti}^{2p}$$

anti-$k_T$ (p=-1)

(Inclusive) clustering proceeds by identifying the min. distance:

- If it is $d_{ij}$ combine particles ij (update list to contain combined particle)
- If it is $d_{iB}$ , identify i as a jet and remove from list

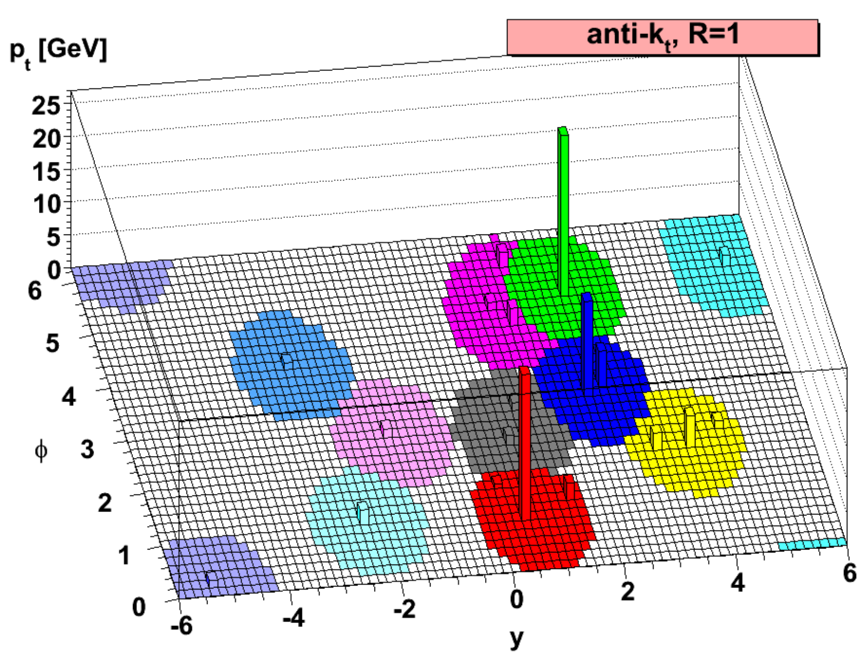
[repeat until list is empty]

# Jets at the LHC

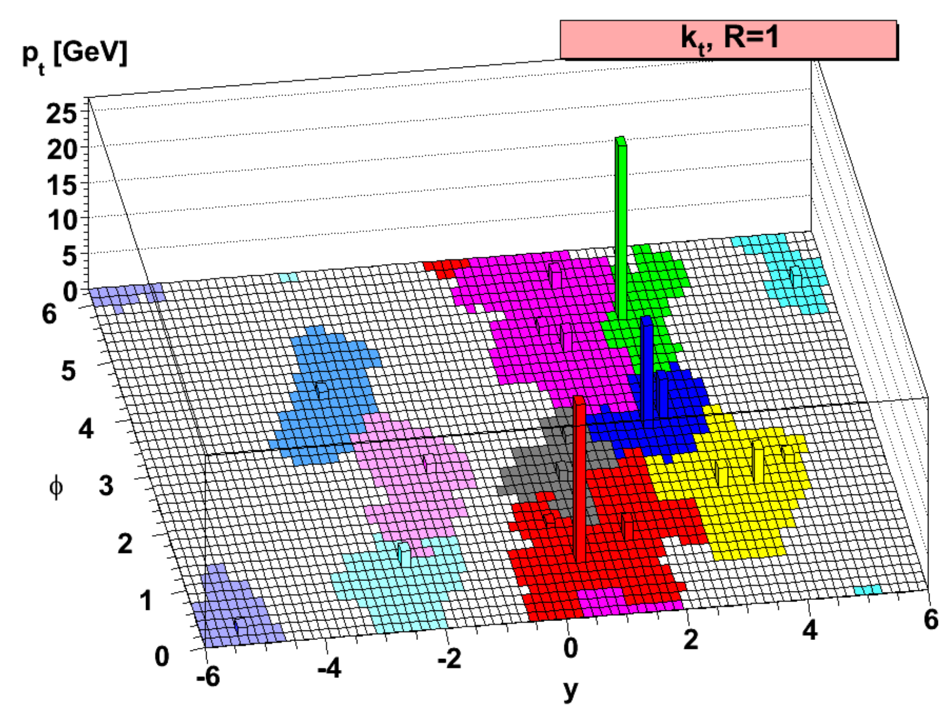
Experimentally: apply an algorithm to particle flow objects (Kaons, Pions,…)
(e.g. ATLAS arXiv:1703.10485, CMS arXiv:1706.04965, LHCb arXiv:1310.8197)

The anti-$k_T$ algorithm (Cacciari, Salam, Soyez arXiv:0802.1189) applied to these objects

anti-$k_T$ has nice geometrical properties (used in all LHC analyses)




anti-$k_T$ (p=-1)




$k_T$ (p=1)

## InfraRed and Collinear safe observables

- Those not impacted by collinear splitting(s) or emission(s) of soft particles

- Those calculated in terms of quarks and gluons where the $m_q \to 0$ limit does not introduce singularities (Stermann, Weinberg '77)

➡ Can (reliably) use fixed-order perturbation theory

$$d\sigma^{\text{data (meas.)}}_{\text{PP}\to\text{f}+\text{X}} \quad \text{vs} \quad d\sigma^{\text{fixed-order}}_{\text{PP}\to\text{f}+\text{X}}$$

## KLN theorem (Kinoshita '62, Lee & Nauenberg '64)

- For such observables, a cancellation of IRC divergences between virtual and real emissions is ensured (order-by-order)

- IRC unsafe observables can be defined, all-order-resummation/factorisation theorems typically required (PDF evolution, obs. dependent resummation)

## InfraRed and Collinear safe observables

- Those not impacted by collinear splitting(s) or emission(s) of soft particles
- Those calculated in terms of quarks and gluons where the $m_q \to 0$ limit does not introduce singularities (Stermann, Weinberg '77)

➡ Can (reliably) use fixed-order perturbation theory

$$d\sigma^{\mathrm{data\,(meas.)}}_{\mathrm{PP}\to \mathrm{f}+\mathrm{X}} \quad \text{vs} \quad d\sigma^{\mathrm{fixed-order}}_{\mathrm{PP}\to \mathrm{f}+\mathrm{X}}$$

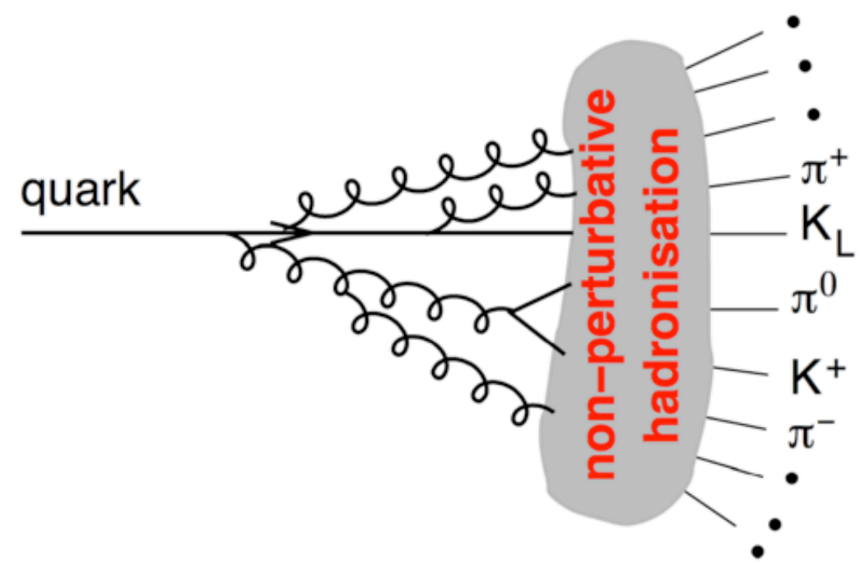
Clearly, one can use a massive calculation to compute unsafe observables

Some costs are:

- $\ln\left[m/Q\right]$ corrections at each order in perturbation theory
- Resummation of some universal corrections absent (heavy flav. PDFs)

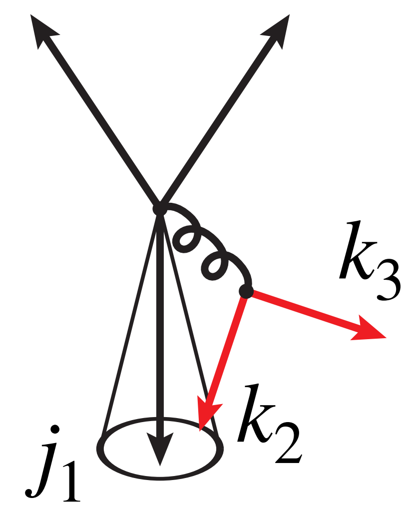
# 1) Introduction ✓

- Jet reconstruction basics



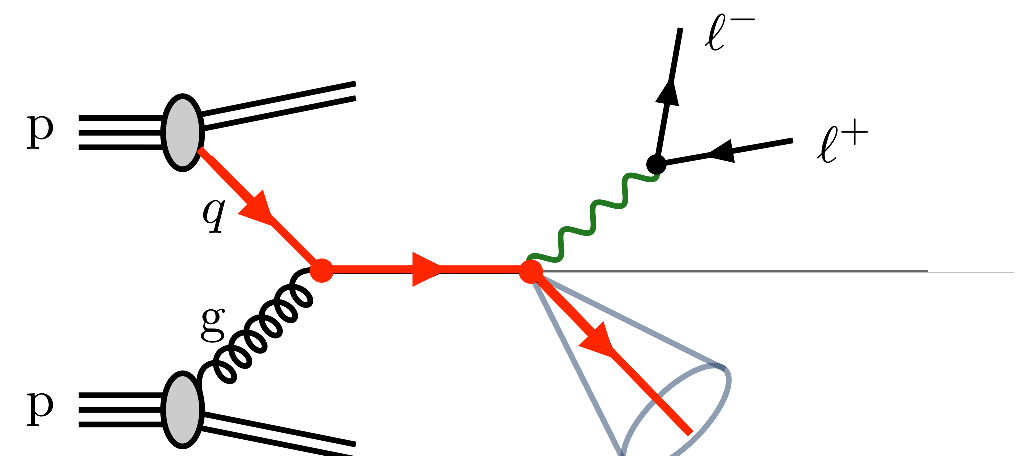

# 2) Jet flavour

- Selected histories
- Recent progress
- IRC tests




# 3) Implications/Applications

- Where does it (not) matter
- Z + c-jets



# Heavy-flavour jets at the LHC

Typical experimental approaches of defining jet flavour (truth/data level):
(ATLAS arXiv:1504.07670, CMS arXiv:1712.07158, LHCb arXiv:1504.07670)

i) First identify flavour-blind anti-$k_T$ jets in a fiducial region

ii) Tag these jets with flavour by the presence of 1 or more D/B hadrons

$$\Delta R(j, D/B) < 0.5$$

iii) [ATLAS/LHCb] Apply $p_T$ requirement to D/B hadron ~ $p_T^{D/B} > 5$ GeV

# Heavy-flavour jets at the LHC

Typical experimental approaches of defining jet flavour (truth/data level):

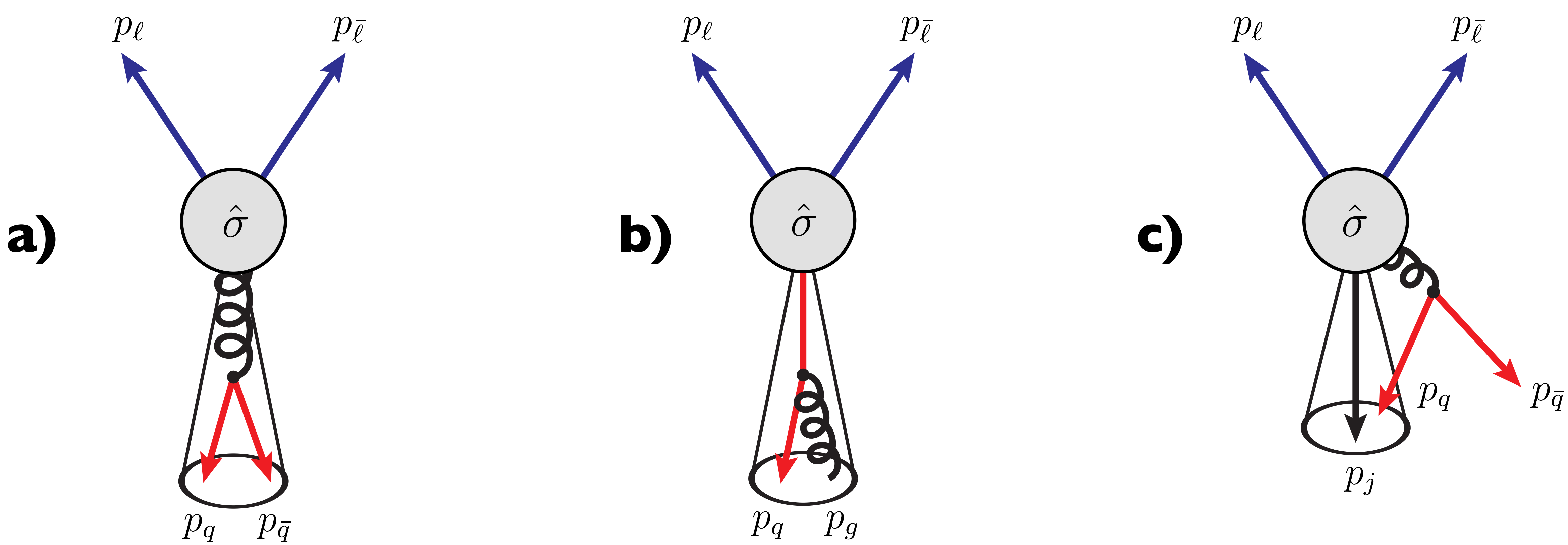
(ATLAS arXiv:1504.07670, CMS arXiv:1712.07158, LHCb arXiv:1504.07670)

i) First identify flavour-blind anti-$k_T$ jets in a fiducial region

ii) Tag these jets with flavour by the presence of 1 or more D/B hadrons

$$\Delta R(j, D/B) < 0.5$$

iii) [ATLAS/LHCb] Apply $p_T$ requirement to D/B hadron ~ $p_T^{D/B} > 5$ GeV



# Jet flavour is not a new problem …

<u>Comments</u>

flavoured jet algorithm
$k_T$ jets

"The Flavour-$k_T$ algorithm"
(Banfi, Salam, Zanderighi **BSZ**: hep-ph/0601139)

Theoretical Physics | Published: 19 May 2006

## Infrared-safe definition of jet flavour

A. Banfi, G.P. Salam & G. Zanderighi

*The European Physical Journal C - Particles and Fields* **47**, 113–124(2006) | Cite this article

**109** Accesses | **71** Citations | Metrics

(1) Quantum flavour assignment (or modulo even 2):

$$b = +1, \ \bar{b} = -1$$

(2) Flavour specific clustering (a flavoured jet algorithm)

$$d_{ij} = \frac{\Delta y_{ij}^2 + \Delta\phi_{ij}^2}{R^2} \begin{cases} \max(k_{ti}, k_{tj})^\alpha \min(k_{ti}, k_{tj})^{2-\alpha} & \text{softer of } i, j \text{ is flavoured,} \\ \min(k_{ti}, k_{tj})^\alpha & \text{softer of } i, j \text{ is unflavoured} \end{cases}$$

# … but it is an active one

Comments

"The Flavour-$k_T$ algorithm"
(Banfi, Salam, Zanderighi **BSZ**: hep-ph/0601139)

flavoured jet algorithm
$k_T$ jets

… selected recent work with anti-$k_T$ jets

Practical jet flavour through NNLO
(Caletti, Larkoski, Marzani, Reichelt **CLMR**: arXiv:2205.01109)

substructure

Infrared-safe flavoured anti-$k_T$ jets
(Czakon, Mitov, Poncelet **CMP**: arXiv:2205.11879)

flavoured jet algorithm
damping function

A dress of flavour to suit any jet
(Gauld, Huss, Stagnitto **GHS**: arXiv:2208.11138)

flavour assignment algorithm
(reconstructed jets are input)

Flavoured jets with exact anti-$k_T$ kinematics and tests of infrared and collinear safety
(Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler **IFN**: arXiv:2306.07314)

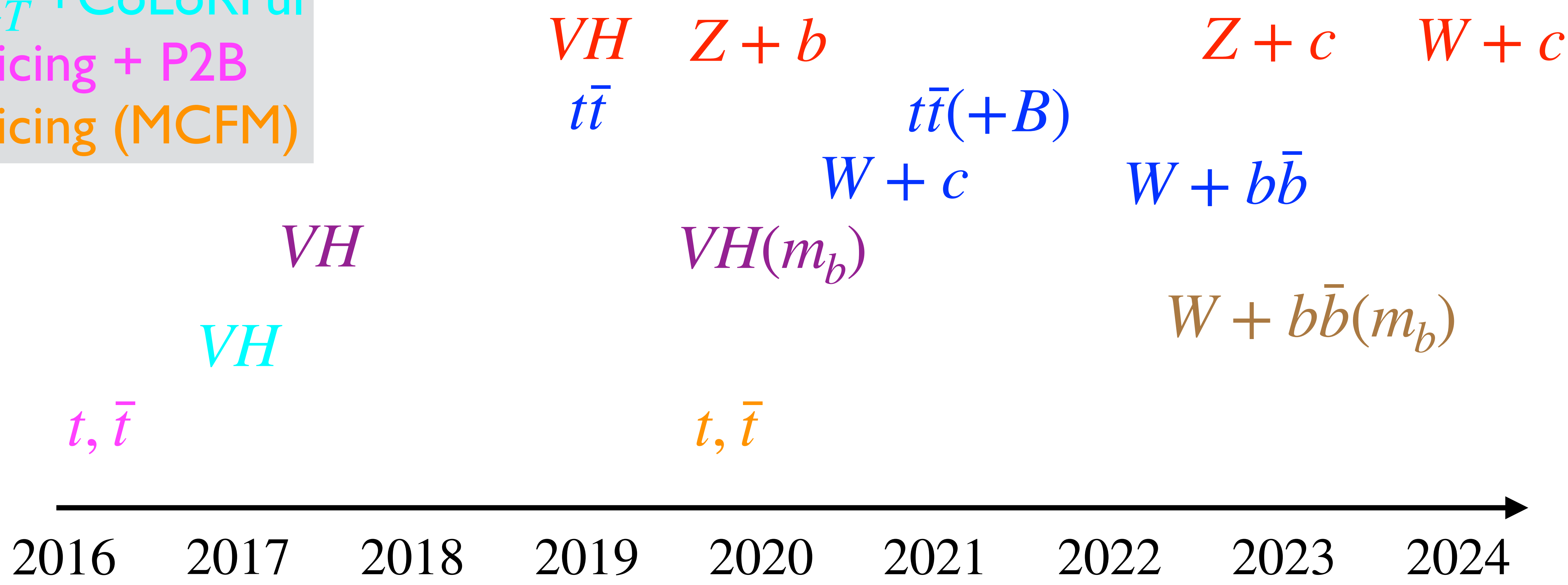
substructure
IRC tests up to $\mathcal{O}(\alpha_s^6)$

# Recent NNLO progress with flavoured jets

**Methods**

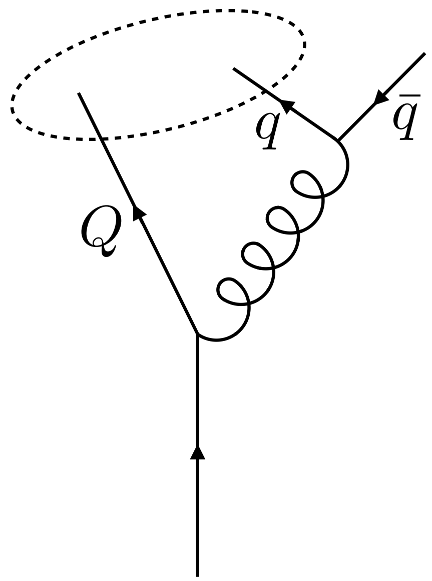
- Antenna
- Stripper
- Nested SC
- $Q_T$ subtraction
- $Q_T$ +CoLoRFul
- Slicing + P2B
- Slicing (MCFM)

All calculations consider flavoured jets
(see Rene's talk)




apologies if some missing

## Practical jet flavour through NNLO
(Caletti, Larkoski, Marzani, Reichelt **CLMR**: arXiv:2205.01109)




Use of SoftDrop:

Uses Jade and $\beta > 0$

## Infrared-safe flavoured anti-$k_T$ jets
(Czakon, Mitov, Poncelet **CMP**: arXiv:2205.11879)

$$d_{ij}^{(F)} \equiv d_{ij} \times \begin{cases} \mathcal{S}_{ij}\,, & \text{if both } i \text{ and } j \text{ have non-zero flavour of opposite sign,} \\ 1\,, & \text{otherwise.} \end{cases} \tag{2.8}$$

$$\mathcal{S}_{ij} \equiv 1 - \theta\left(1-\kappa_{ij}\right)\cos\left(\frac{\pi}{2}\kappa_{ij}\right) \quad \text{with} \quad \kappa_{ij} \equiv \frac{1}{a}\,\frac{k_{T,i}^2 + k_{T,j}^2}{2k_{T,\max}^2}\,. \tag{2.9}$$

damping function, vanishes in double soft limit (overcomes $E^2$ of $d_{ij}$ )

# A dress of flavour to suit any jet

(Gauld, Huss, Stagnitto **GHS**: arXiv:2208.11138)

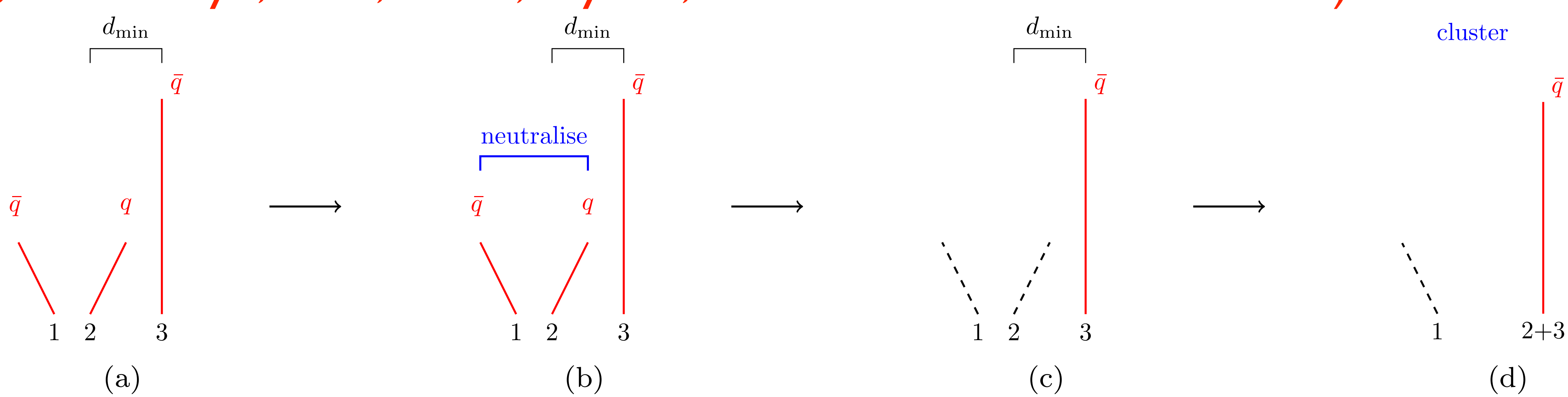
algorithm assigns $\hat{f}_i$ to $j_k$

set of jets $\{j_1, \ldots, j_m\}$ ⟵ set of flavoured objects $\{\hat{f}_1, \ldots, \hat{f}_n\}$

(flavoured particles not required to be final state!)

# Flavoured jets with exact anti-k$_T$ kinematics and tests of infrared and collinear safety

(Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler **IFN**: arXiv:2306.07314)




**Interleaved Flavour Neutralisation**

$$u_{ik} \equiv [\max(p_{ti}, p_{tk})]^{\alpha}[\min(p_{ti}, p_{tk})]^{2-\alpha} \times \Omega^2_{ik}\,, \quad (7a)$$

$$\Omega^2_{ik} \equiv 2\left[\frac{1}{\omega^2}(\cosh(\omega \Delta y_{ik}) - 1) - (\cos\Delta\phi_{ik} - 1)\right], \quad (7b)$$

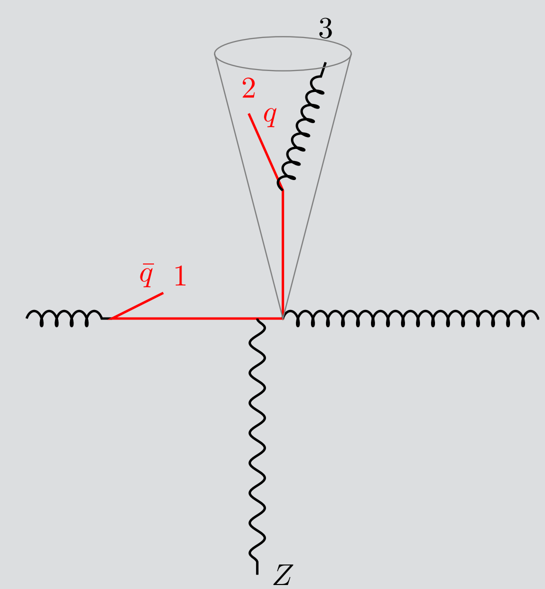
$$u_{ik} \equiv [\max(p_{ti}, p_{tk})]^{\alpha} [\min(p_{ti}, p_{tk})]^{2-\alpha} \times \Omega_{ik}^2 \,, \quad \text{(7a)}$$

$$\Omega_{ik}^2 \equiv 2 \left[ \frac{1}{\omega^2} \left( \cosh(\omega \Delta y_{ik}) - 1 \right) - \left( \cos \Delta\phi_{ik} - 1 \right) \right], \quad \text{(7b)}$$

Take $\omega = 1$, for small $\Delta y_{ik}$ and $\Delta\phi_{ik}$ this behaves as $\Delta R^2 = \Delta y^2 + \Delta\phi^2$

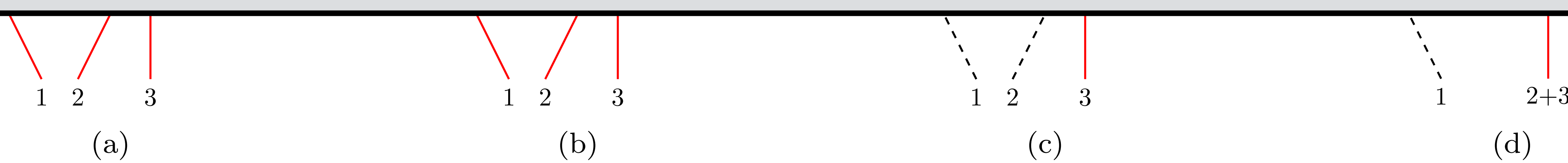
However, for larger $\Delta y$ the exponent leads to a larger distance:

e.g. the [12] clustering probability damped




Analytic/numerical tests suggest $\omega = 3 - \alpha$ is safe

Such a distance also critical for original CMP and GHS algorithms (next slides)



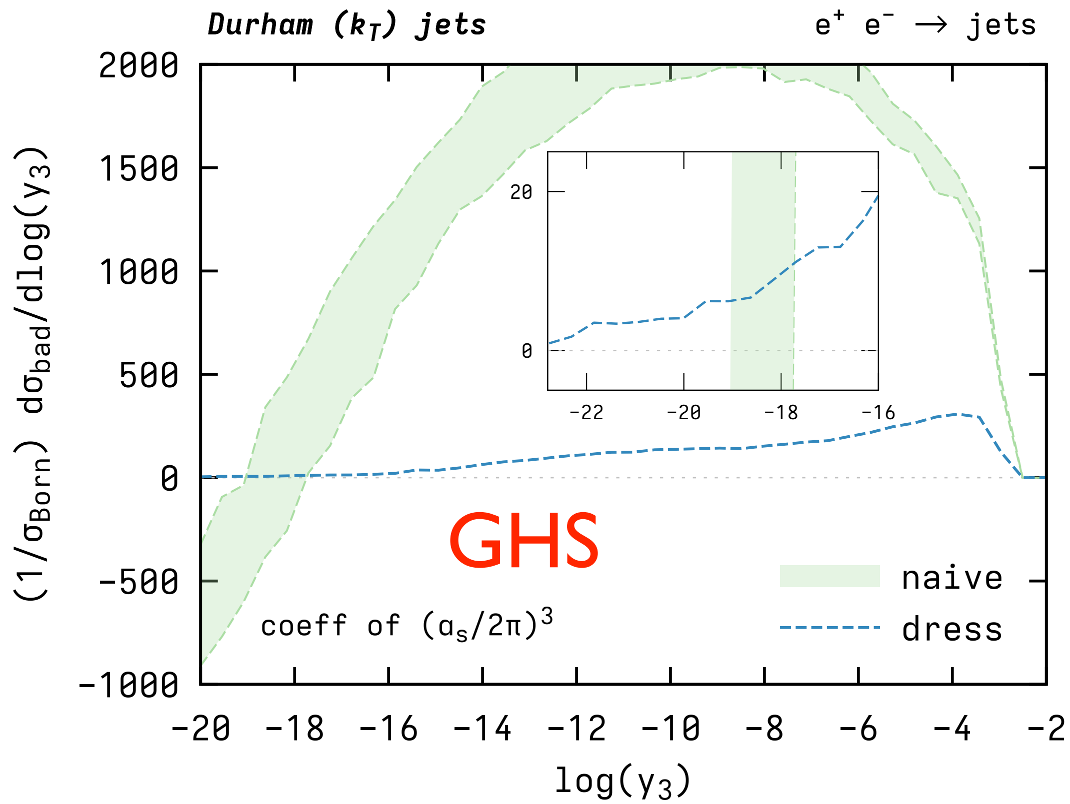

Interleaved Flavour Neutralisation

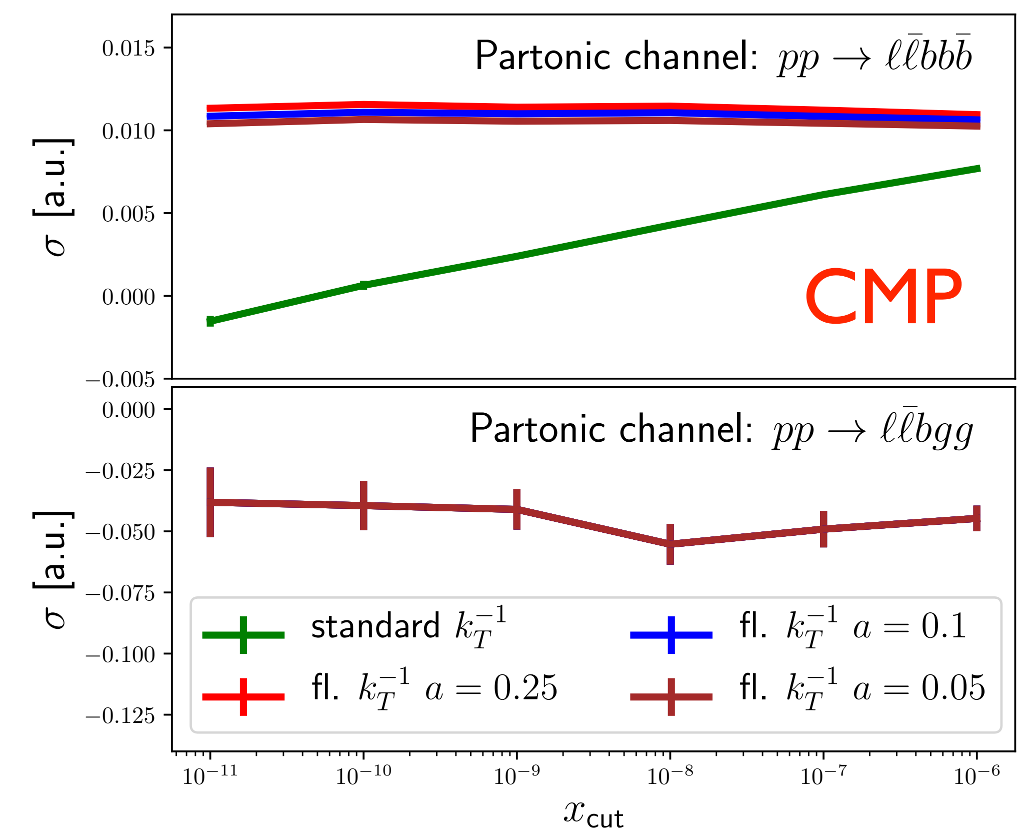

$$u_{ik} \equiv [\max(p_{ti}, p_{tk})]^{\alpha} [\min(p_{ti}, p_{tk})]^{2-\alpha} \times \Omega_{ik}^2 \,, \quad \text{(7a)}$$

$$\Omega_{ik}^2 \equiv 2 \left[ \frac{1}{\omega^2} \left( \cosh(\omega \Delta y_{ik}) - 1 \right) - \left( \cos \Delta\phi_{ik} - 1 \right) \right], \quad \text{(7b)}$$

# Tests of infrared and collinear safety

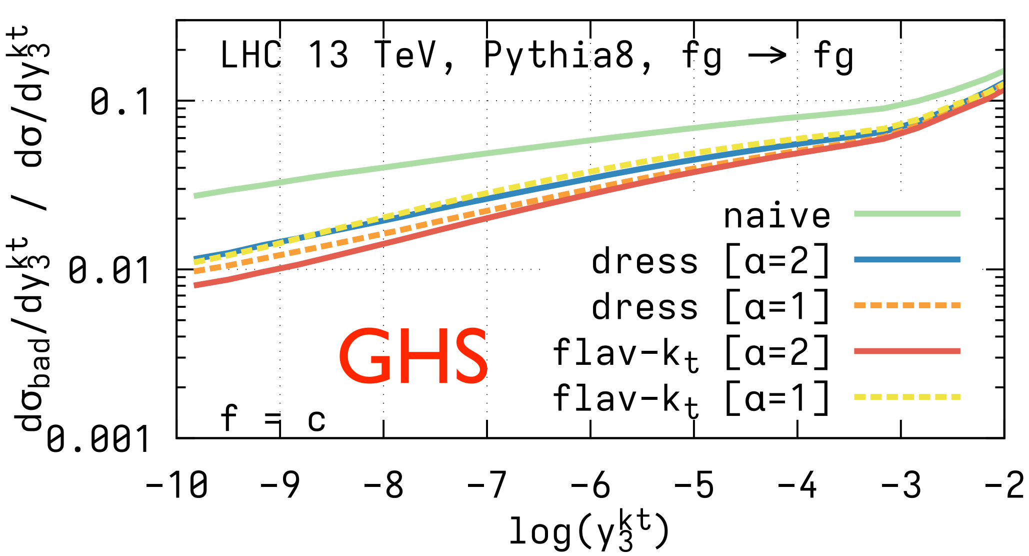
Use of massless fixed-order calculations:

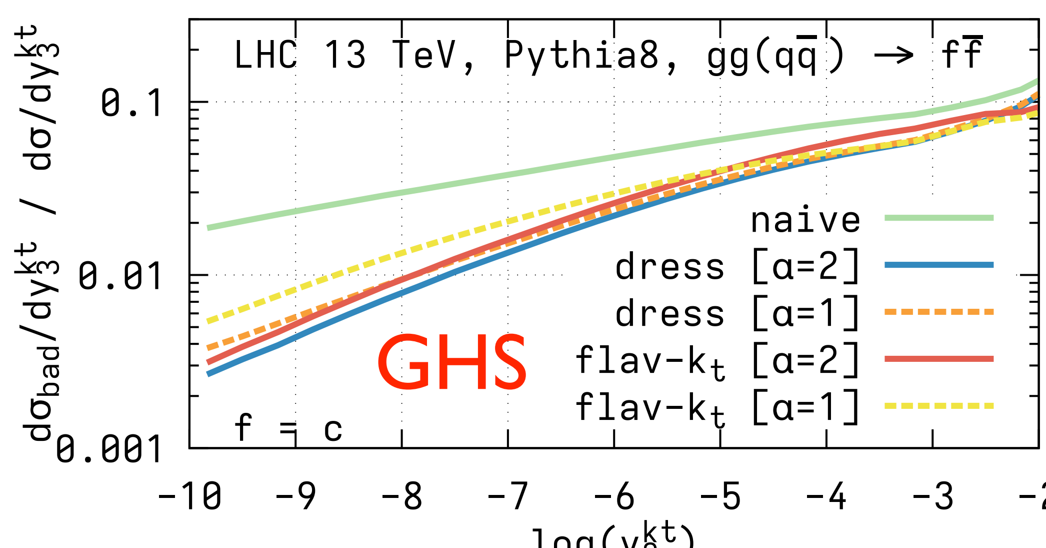



Test `bad' assignments in IRC regime

Insensitivity to technical parameter



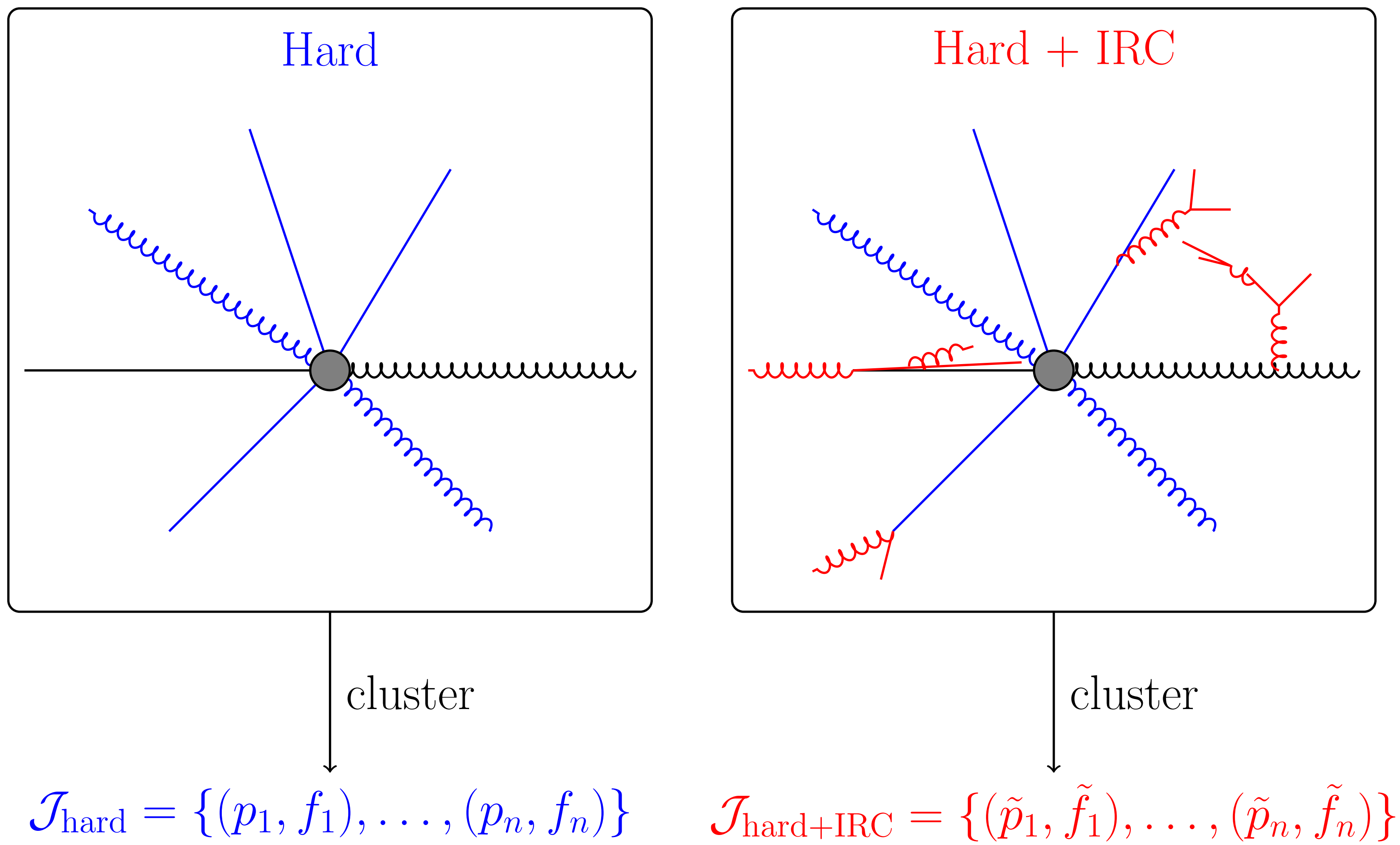




All order sensitivity tests via PS (qualitative, potentially misleading)

# Tests of infrared and collinear safety

Use of systematic framework, numerically test up to $\mathcal{O}(\alpha_s^6)$

(Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler **IFN**: arXiv:2306.07314)




Add a selected set of IRC emissions to test the jet algorithms safety (insensitivity):

Final-state hard collinear **(FHC)**
Final-state double soft **(FDS)**

Initial-state hard collinear **(IHC)**
Initial-state double soft **(IDS)**

# Tests of infrared and collinear safety

Use of systematic framework, numerically test up to $\mathcal{O}(\alpha_s^6)$

(Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler **IFN**: arXiv:2306.07314)

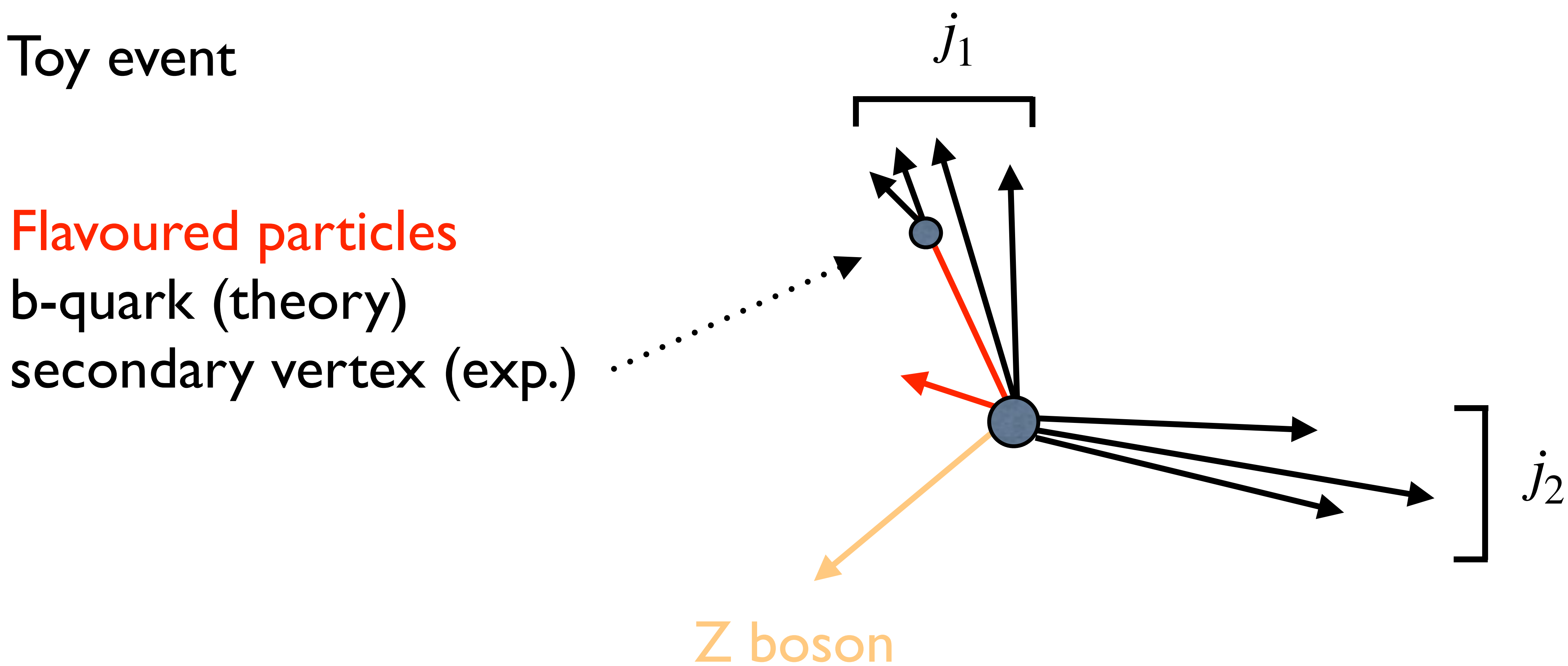
| order relative to Born | | anti-$k_t$ | flav-$k_t$ $(\alpha = 2)$ | CMP | GHS$_{\alpha,\beta}$ $(2,2)$ | anti-$k_t$+IFN$_\alpha$ | C/A+IFN$_\alpha$ |
|---|---|---|---|---|---|---|---|
| $\alpha_s$ | FHC | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | IHC | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $\alpha_s^2$ | FDS | $\times_{\text{II B}}$ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | IDS | $\times_{\text{II B}}$ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | FHC×IHC | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| | IHC$^2$ | ✓ | ✓ | $\times_{\text{C 2}}$ | ✓ | ✓ | ✓ |
| | FHC$^2$ | ✓ | ✓ | ✓ | $\times_{\text{C 4}}$ | ✓ | ✓ |
| $\alpha_s^3$ | IHC×IDS | | $\sim_{\text{C 1}}$ | $\times_{\text{C 3}}$ | $\sim_{\text{C 1}}$ | ✓ | ✓ |
| | rest | | | | | ✓ | ✓ |
| $\alpha_s^4$ | IDS×FDS | | | | $\times_{\text{C 5}}$ | ✓ | ✓ |
| | rest | | | | | ✓ | ✓ |
| $\alpha_s^5$ | | | | | | ✓ | ✓ |
| $\alpha_s^6$ | | | | | | ✓ | ✓ |

Summary table from arXiv:2306.07314 of IRC tests for CMP, GHS, IFN

note: for GHS the IHC² configuration does not appear for $e^+e^- \to$ jets at $\mathcal{O}(\alpha_s^3)$

# An example of a failure (GHS)

Toy event




algorithm assigns $\hat{f}_i$ to $j_k$

set of jets $\{j_1, \ldots, j_m\}$ ← set of flavoured objects $\{\hat{f}_1, \ldots, \hat{f}_n\}$

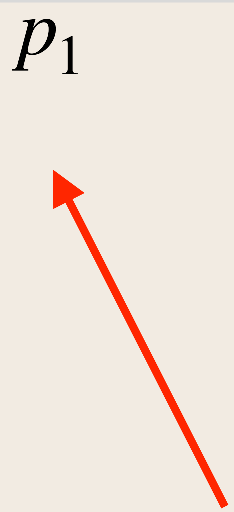
(flavoured particles not required to be final state!)

Whole procedure has a few stages:

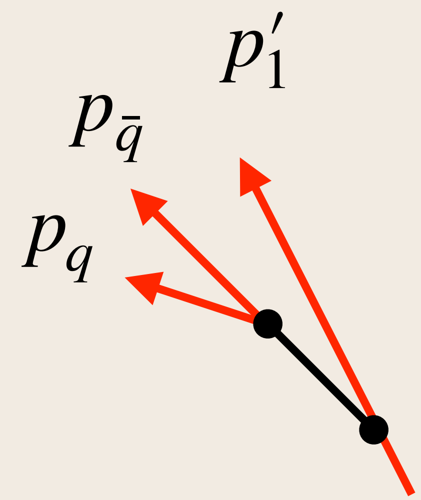
1. Prepare a set of flavoured objects (use soft-drop for a collinear dressing)
2. Use the flavour-$k_T$ distance (and an association criterion) to assign $\hat{f}_i$ to $j_k$
3. Sum up/count flavours per jet

(in part 2 momenta of objects un-changed)

# An example of a failure (GHS)




Scenario A: $p_1$ enters stage 2




add DCollinear

Scenario B: $p_1, p_q, p_{\bar{q}}$ enters stage 2

$d_{q\bar{q}} \sim 0$ [removed] $\quad p_1' \sim z p_1$

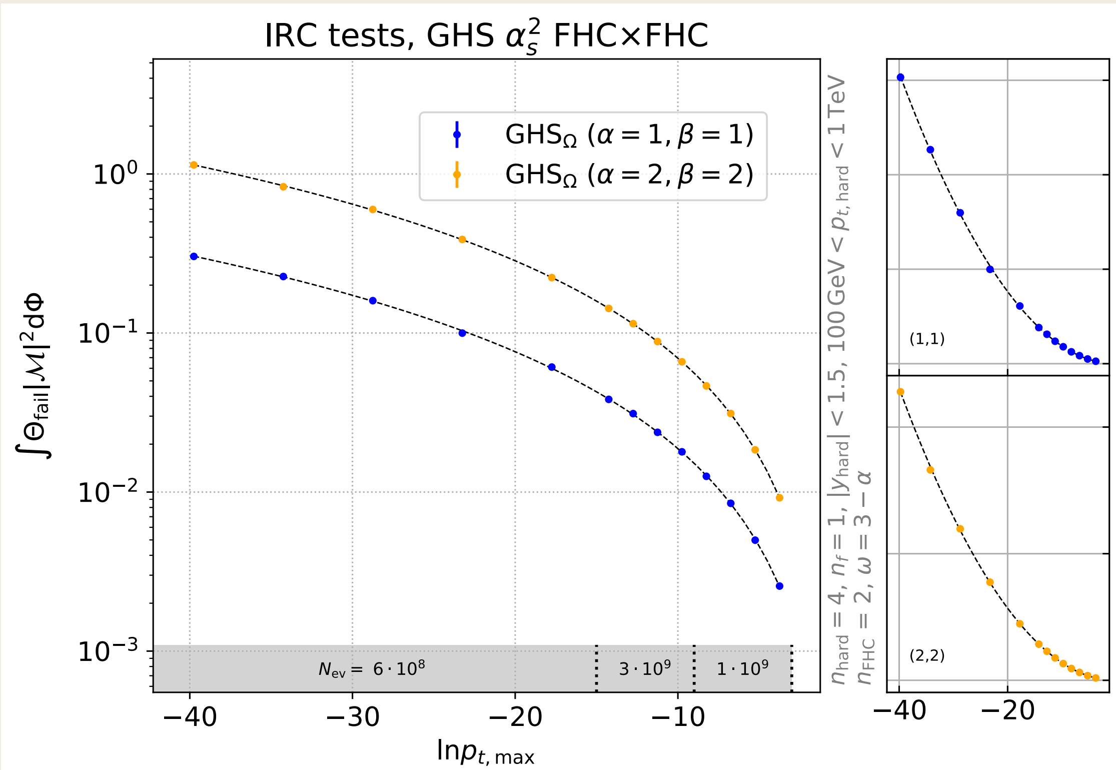
depending on event, can alter result

Whole procedure has a few stages:

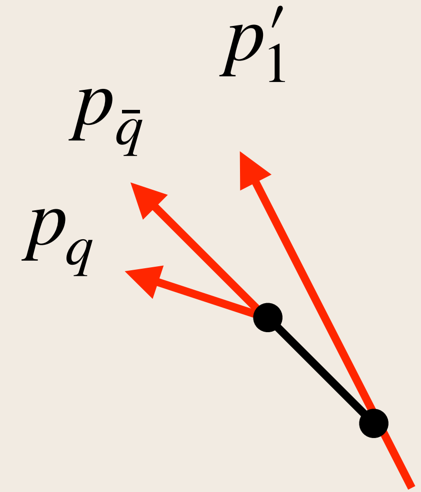
1. Prepare a set of flavoured objects (use soft-drop for a collinear dressing)
2. Use the flavour-k$_T$ distance (and an association criterion) to assign $\hat{f}_i$ to $j_k$
3. Sum up/count flavours per jet

(in part 2 momenta of objects un-changed)

# An example of a failure (GHS)




GHS Failure rate (from IFN paper)




add DCollinear

Scenario B: $p_1, p_q, p_{\bar{q}}$ enters stage 2

$d_{q\bar{q}} \sim 0$ [removed] $\quad p'_1 \sim z p_1$

depending on event, can alter result

Whole procedure has a few stages:

1. Prepare a set of flavoured objects (use soft-drop for a collinear dressing)
2. Use the flavour-k$_T$ distance (and an association criterion) to assign $\hat{f}_i$ to $j_k$
3. Sum up/count flavours per jet

(in part 2 momenta of objects un-changed)

# Changes required for GHS

(i) To combine stages 1. and 2. (without the need for SoftDrop)
(ii) Modify the flavour-kT distance similar to the IFN approach

algorithm assigns $f_i$ to $j_k$

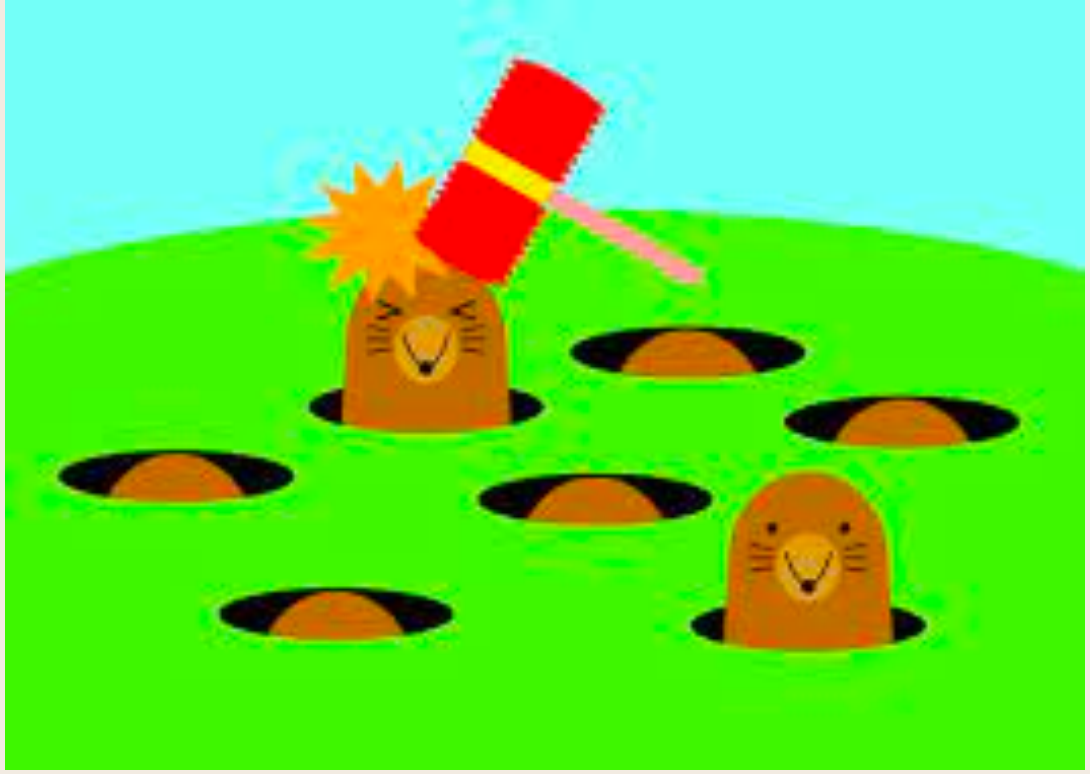
set of jets $\{j_1, \ldots, j_m\}$ ← set of flavoured & flavourless objects $\{f_1, \ldots, f_n, p_1, \ldots, p_m\}$

(flavoured particles not required to be final state!)

Revision passes the tests up to $\mathcal{O}(\alpha_s^6)$:

Tested for gen-$k_T$ algorithm (-1,0,1)

An erratum to be submitted for GHS



Many thanks to the IFN authors for providing the test suite from
(Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler **IFN**: arXiv:2306.07314)

# Summary of part 2

Several anti-$k_T$ jet-flavour algs. with varying features, inputs, parameters

**CMP**: Modified clustering w.r.t. standard anti-$k_T$

**IFN:** Interleaved with the clustering history (gives exact anti-$k_T$ kinematics)

**GHS**: Does not require flavours to be present in initial jet reconstruction

…

Some commonalties between algorithms (**BSZ**, **CLMR**, **CMP**, **GHS**, **IFN**, ..):

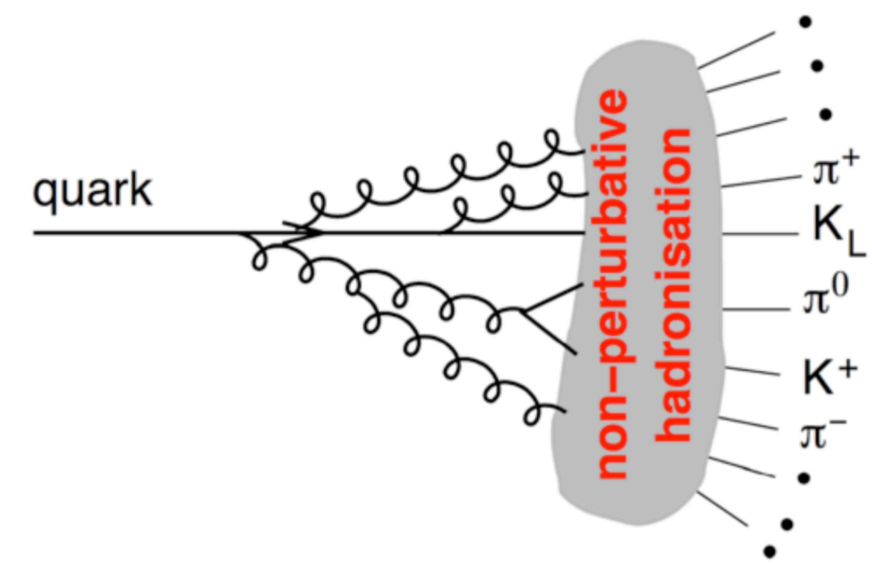
Flavour counting (or modulo even 2): $q = +1, \ \bar{q} = -1$

Require knowledge of flavoured particles in event (full phase space, no $p_T^q$ cuts) (except CLMR?)

**All in conflict with experimental approaches**

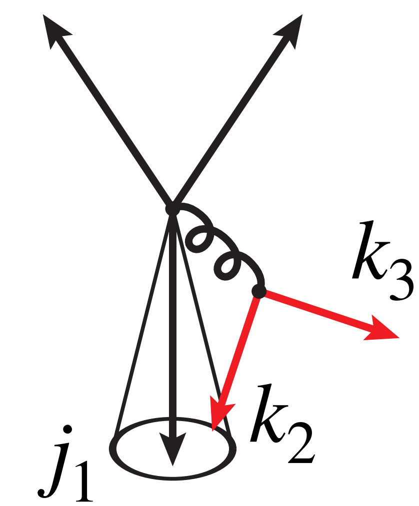
## 1) Introduction ✓

- Jet reconstruction basics



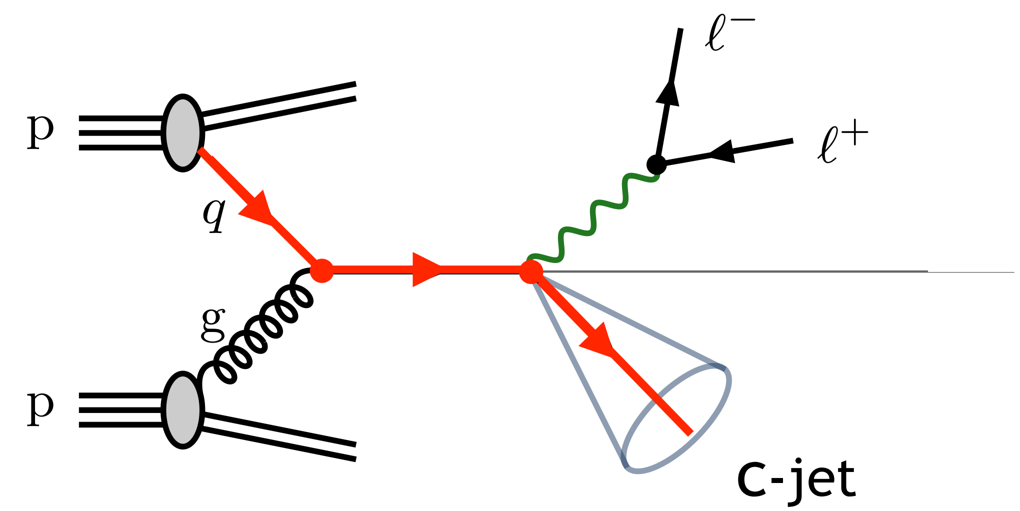

## 2) Jet flavour ✓

- Selected histories
- Recent progress
- IRC tests




## 3) Implications/Applications

- Where does it (not) matter
- Z + c-jets



# (To what extent) does it IRC safety matter?

Are massive computations sufficiently precise to enable (unsafe) comparisons

Essentially, how large is the unsafe component numerically

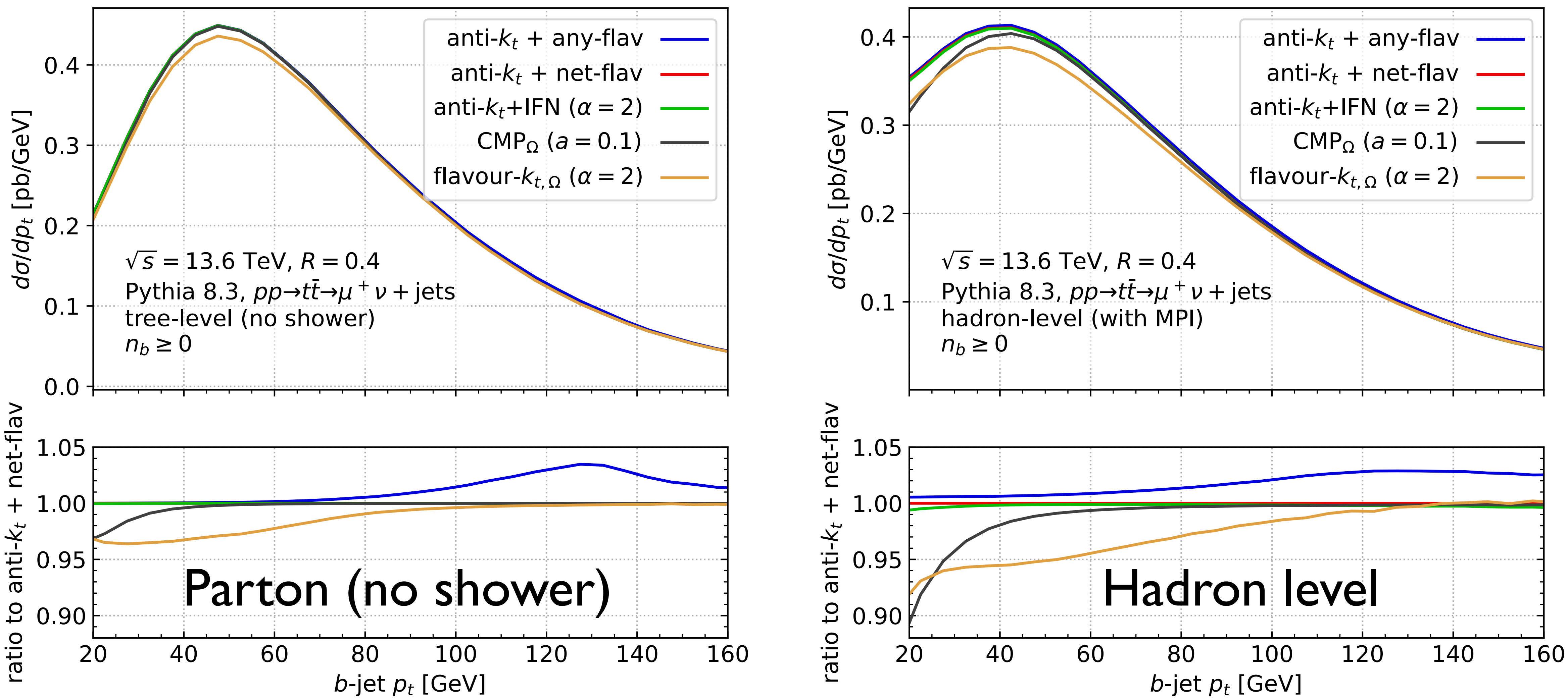
$$\mathrm{d}\sigma^{\mathrm{data\,(meas.)}}_{\mathrm{PP}\to\mathrm{f}+\mathrm{X}} \quad \text{vs} \quad \mathrm{d}\sigma^{\mathrm{theory}}_{\mathrm{PP}\to\mathrm{f}+\mathrm{X}}$$

Some cases to consider: $pp \to t\bar{t}\,,\ pp \to Z + \mathrm{f} - \mathrm{jet}$

# (To what extent) does it IRC safety matter?

## Semi-leptonic $t\bar{t} + X$ (NLO+PS)

(Caola, Grabarczyk, Hutt, Salam, Scyboz, Thaler **IFN**: arXiv:2306.07314)




(a) (b)

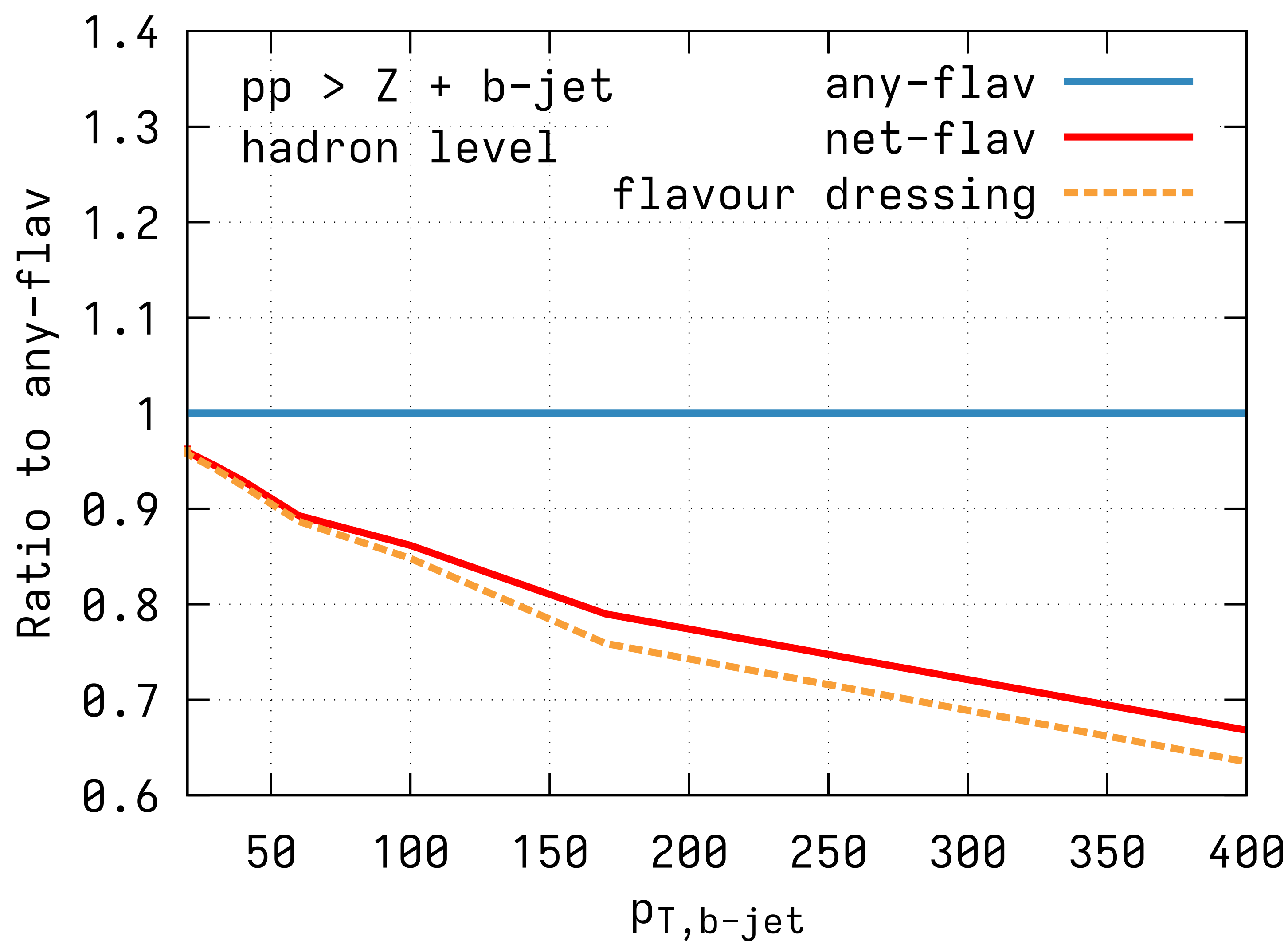
"any-flav": includes double-tagged b-jets (like "Exp")

"net-flav": $b = 1\,, \bar{b} = -1$ (removes $g \to b\bar{b}$ collinear component)

"net-flav" and **IFN** almost identical, differ from "Exp" style by **[0-3]%**

# (To what extent) does it IRC safety matter?

Z+b-jet (NLO+PS)




"any-flav": includes double-tagged b-jets (like "Exp")

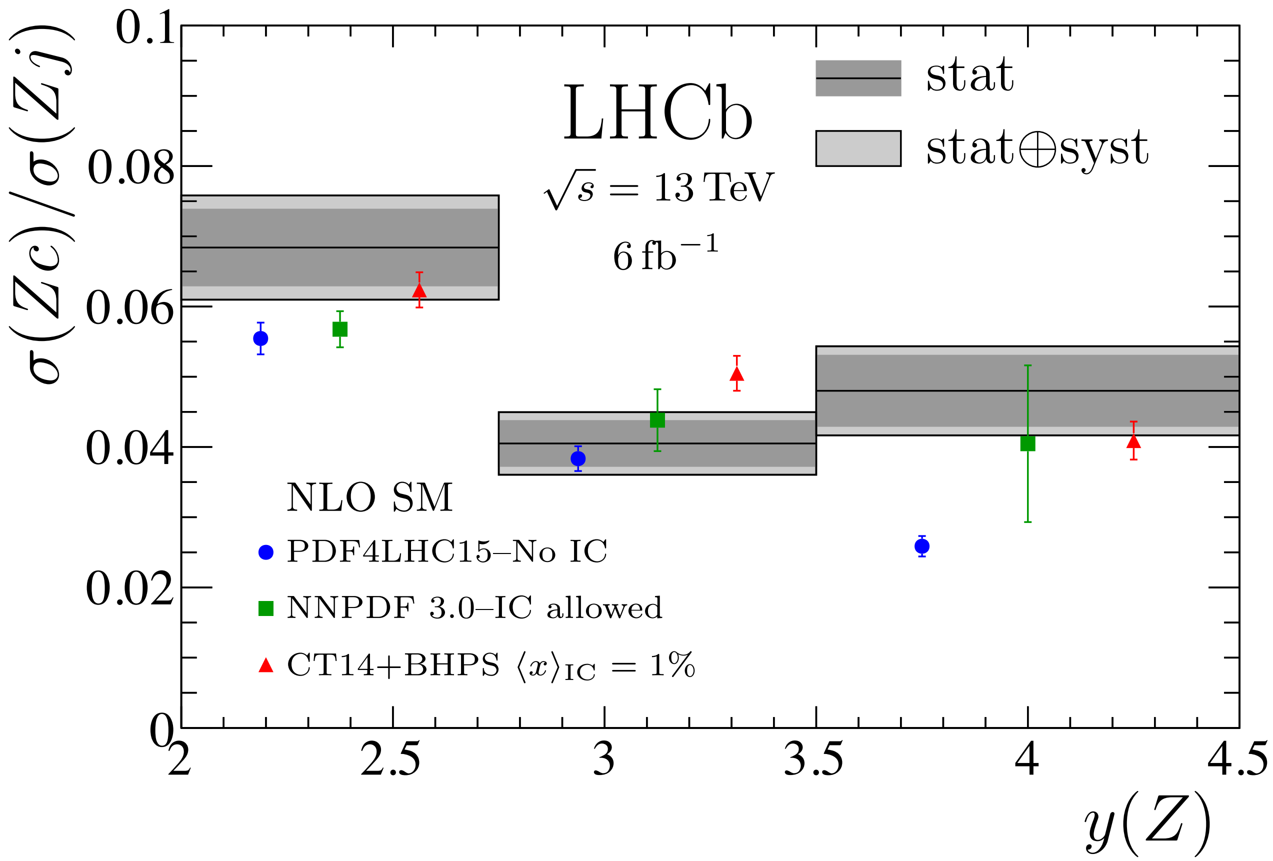
"net-flav": $b = 1\,, \bar{b} = -1$ (removes $g \to b\bar{b}$ collinear component)

"net-flav" and **GHS** few percent differences at high pT

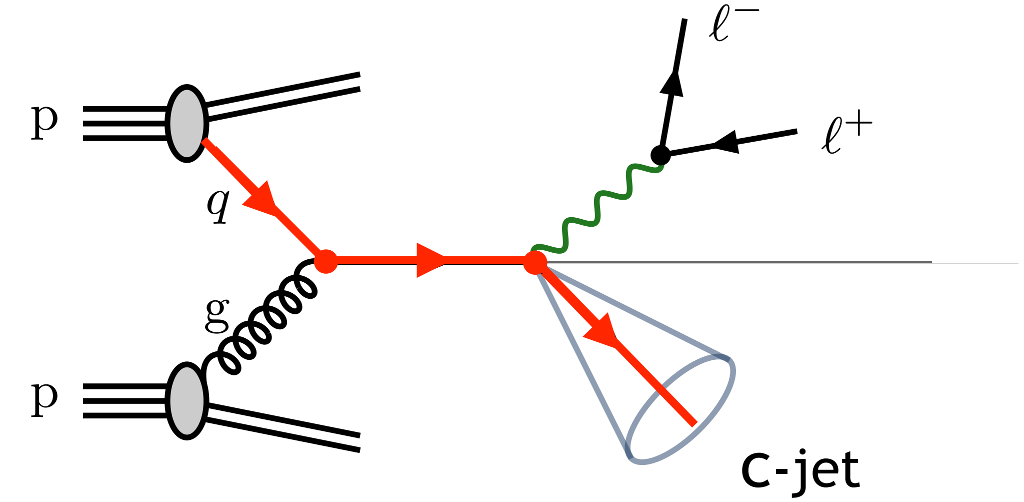
Two important points: $pp \to Z + (g \to b\bar{b})$ at NLO, and the b-quark PDF

# (To what extent) does it IRC safety matter?

## Z+c-jet at LHCb: IRC safety and MPI




LHCb measurement




(Stefano's talk earlier)

Experimentally, "any-flav" with a $p_T^D > 5$ GeV requirement on the c-jet

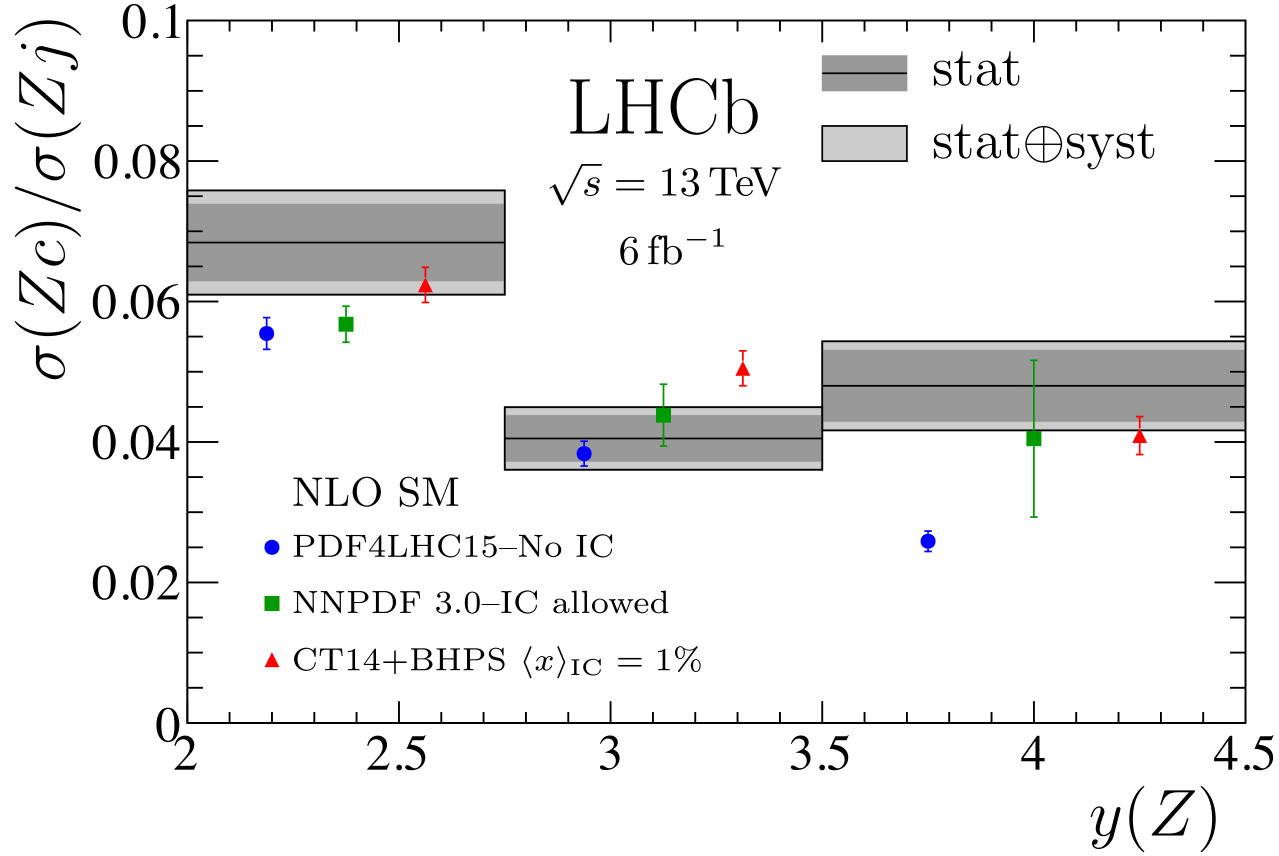
Aim is to interpret the measurement in a collinear PDF fit (intr. charm)

The massless scheme (requiring IRC safe jet at F.O.) allows this

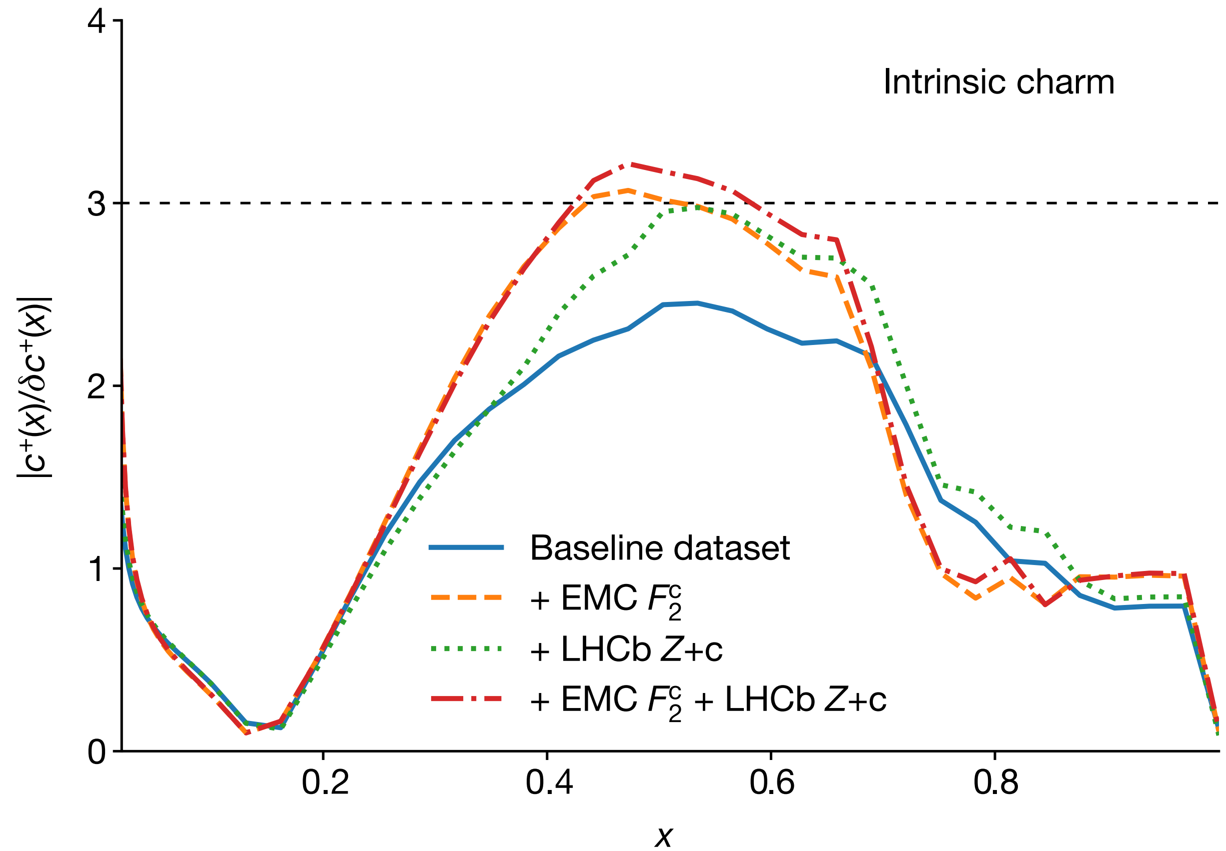
(see Fabrizio's talk tomorrow RE: massive initial-states)

# (To what extent) does it IRC safety matter?

## Z+c-jet at LHCb: IRC safety and MPI




LHCb measurement




(Stefano's talk yesterday)

Experimentally, "any-flav" with a $p_T^D > 5$ GeV requirement on the c-jet

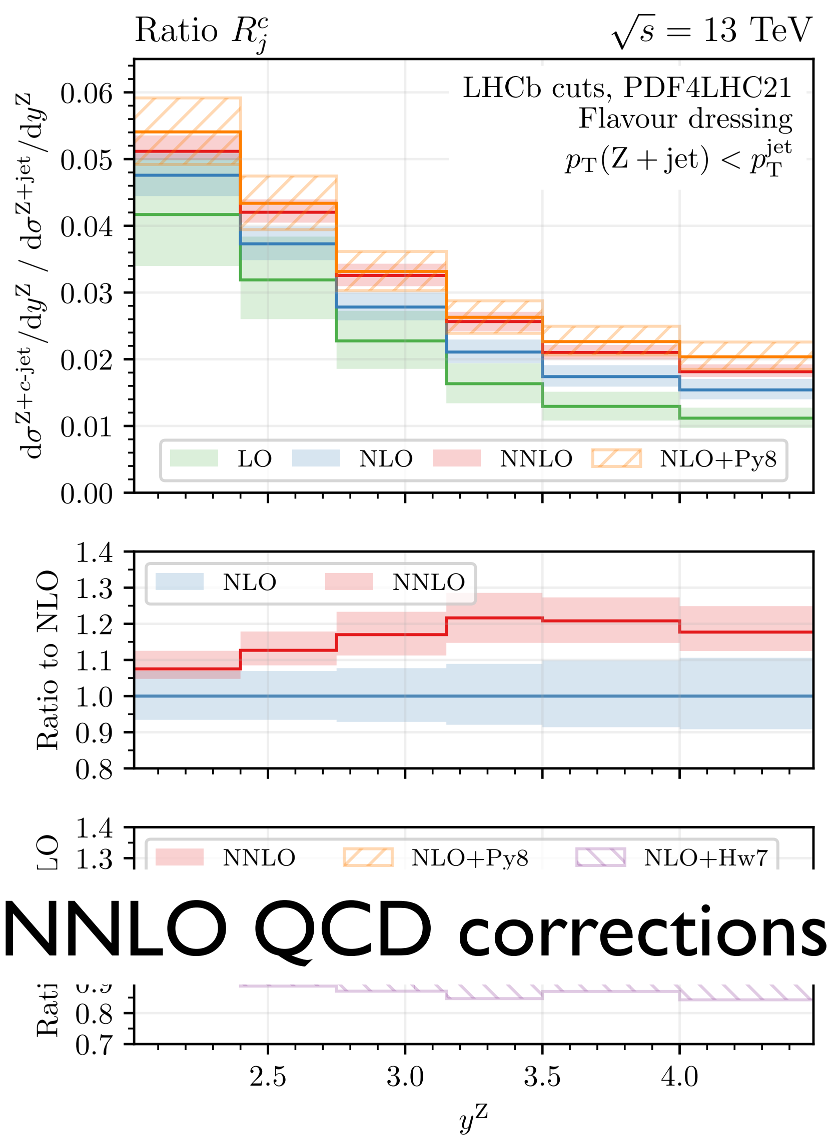
Aim is to interpret the measurement in a collinear PDF fit (intr. charm)

The massless scheme (requiring IRC safe jet at F.O.) allows this

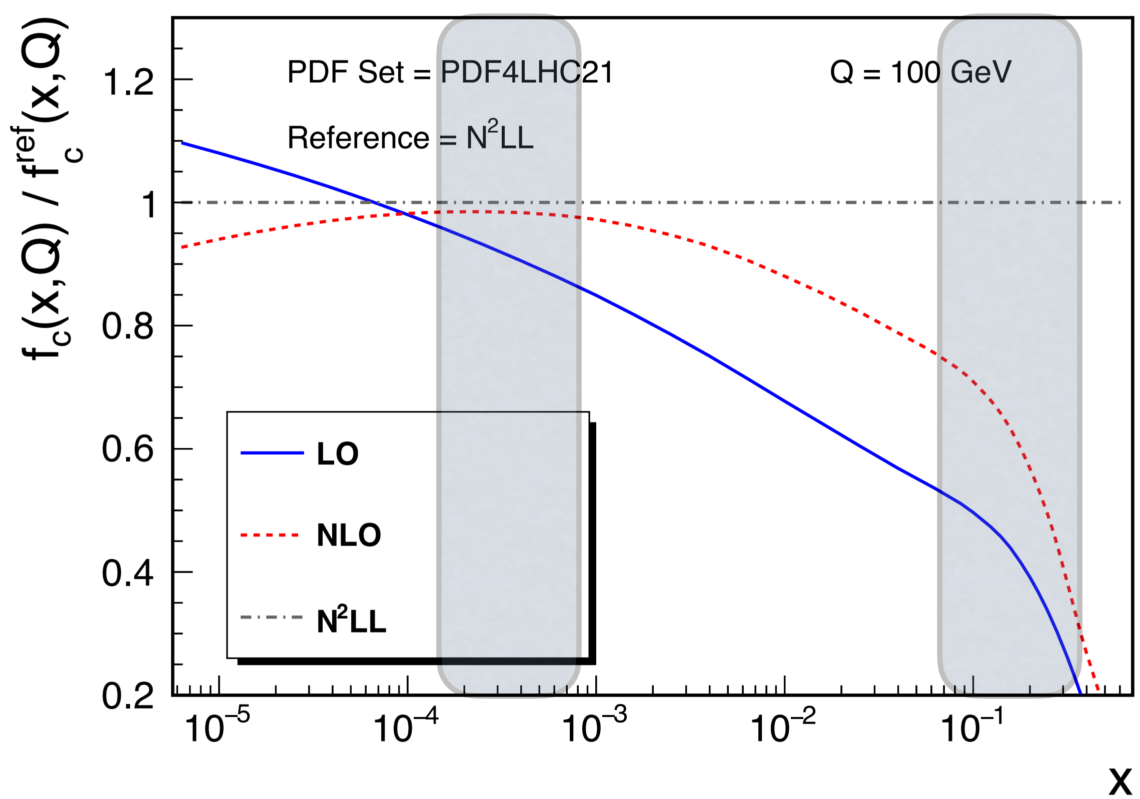
(see Fabrizio's talk tomorrow RE: massive initial-states)

# (To what extent) does it IRC safety matter?

(RG et al., arXiv:2302.12844)




NNLO QCD corrections




c-quark PDF
massive NNLO probs. insufficient

Experimentally, "any-flav" with a $p_T^D > 5$ GeV requirement on the c-jet

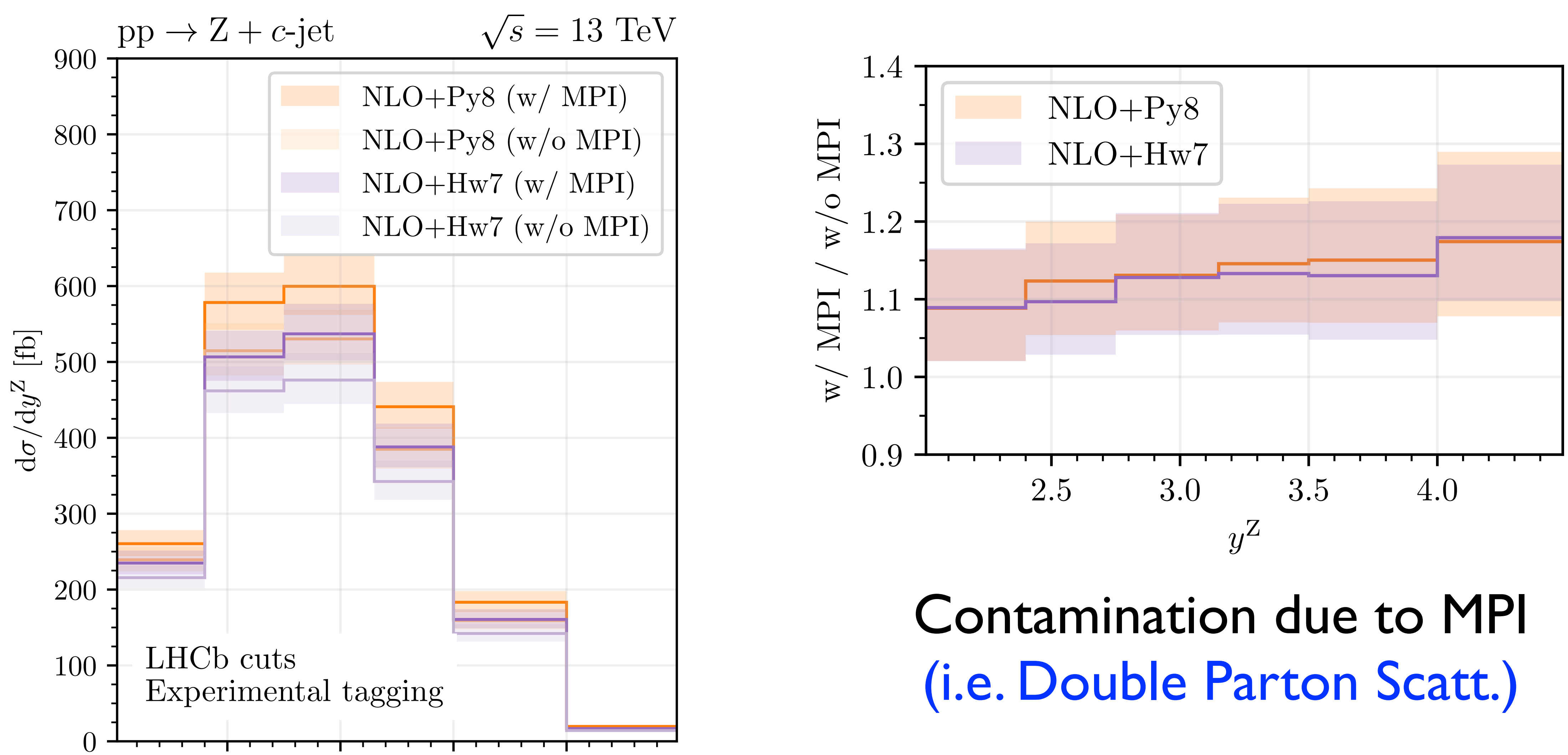
Aim is to interpret the measurement in a collinear PDF fit (intr. charm)

The massless scheme (requiring IRC safe jet at F.O.) allows this

(see Fabrizio's talk tomorrow RE: massive initial-states)

# Final comment about MPI effect

(RG et al., arXiv:2302.12844)




Contamination due to MPI
(i.e. Double Parton Scatt.)

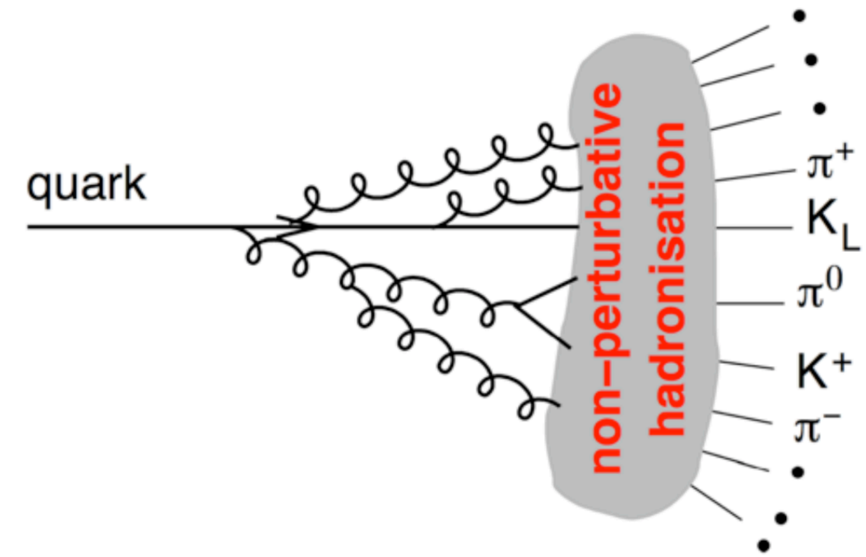
Experimentally, "any-flav" with a $p_T^D > 5$ GeV requirement on the c-jet

Aim is to interpret the measurement in a collinear PDF fit (intr. charm)

The massless scheme (requiring IRC safe jet at F.O.) allows this

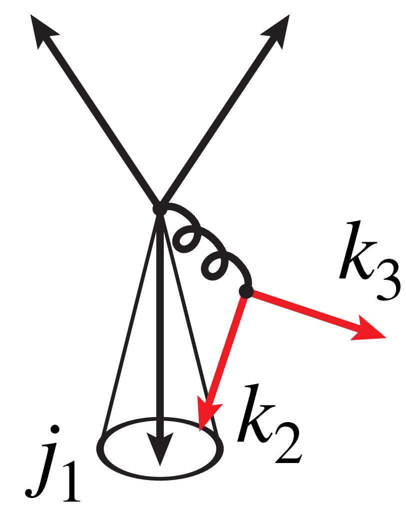
(see Fabrizio's talk tomorrow RE: massive initial-states)

# 1) Introduction



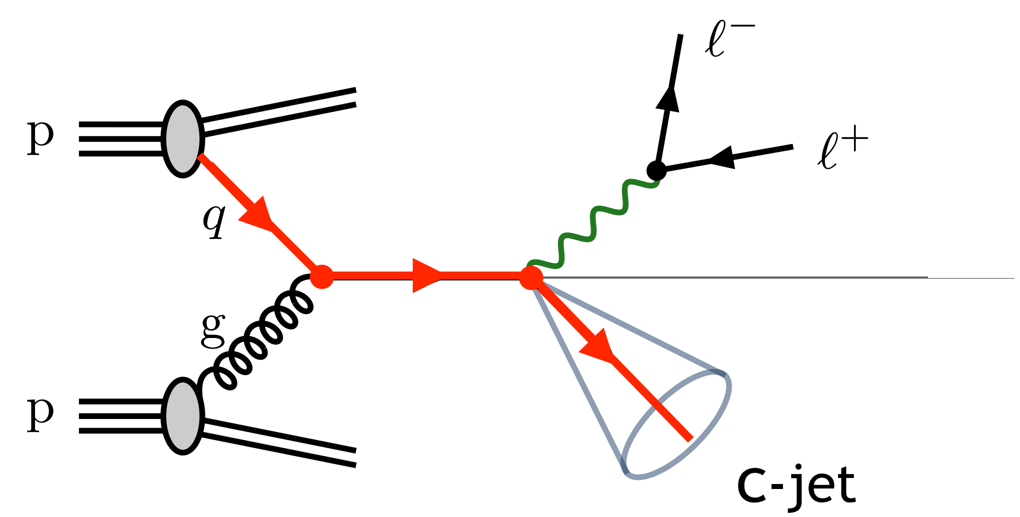

- Jet reconstruction basics

# 2) Jet flavour




- Selected histories
- Recent progress
- IRC tests

# 3) Implications/Applications




- Where does it (not) matter
- Z + c-jets

# Summary of main points

i) Several IRC safe jet definitions of anti-$k_T$ jet flavour now exist

* Their experimental feasibility varies
* If used at 'truth' or unfolded level that does not really matter

ii) For some processes, using massive NNLO w/ unsafe is probably fine

* E.g. $t\bar{t}$ or VH (for which NNLO+PS is anyway available for)

(Behring et al. arXiv:2003.08321)

iii) For others, an IRC safe approach is important

* E.g. Z+c-jets (particularly from point of view of PDF fits)

iv) Alternatively, improve theory to better deal with collinear unsafety?

* Fragmentation approaches to jet flavour or exclusively with hadrons

(WTA aproach, Caletti et al., arXiv:2205.01117, Larkoski et al. , arXiv: 2310.01486)

# Whiteboard

# Jet flavour (accepting collinear unsafety)

## A Fragmentation Approach to Jet Flavor

**Simone Caletti,**[1] **Andrew J. Larkoski,**[2] **Simone Marzani,**[1] **and Daniel Reichelt**[3]

[1] *Dipartimento di Fisica, Università di Genova and INFN, Sezione di Genova, Via Dodecaneso 33, 16146, Italy*

[2] *SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA*

[3] *Institute for Particle Physics Phenomenology, Department of Physics, Durham University, Durham DH1 3LE United Kingdom*

*E-mail:* simone.caletti@ge.infn.it, larkoski@slac.stanford.edu, simone.marzani@ge.infn.it, daniel.reichelt@durham.ac.uk

Abstract: An intuitive definition of the partonic flavor of a jet in quantum chromodynamics is often only well-defined in the deep ultraviolet, where the strong force becomes a free theory and a jet consists of a single parton. However, measurements are performed in the infrared, where a jet consists of numerous particles and requires an algorithmic procedure to define their phase space boundaries. To connect these two regimes, we introduce a novel and simple partonic jet flavor definition in the infrared. We define the jet flavor to be the net flavor of the partons that lie exactly along the direction of the Winner-Take-All recombination scheme axis of the jet, which is safe to all orders under emissions of soft particles, but is not collinear safe. Collinear divergences can be absorbed into a perturbative fragmentation function that describes the evolution of the jet flavor from the ultraviolet to the infrared. The evolution equations are linear and a small modification to traditional DGLAP and we solve them to leading-logarithmic accuracy. The evolution equations exhibit fixed points in the deep infrared, we demonstrate quantitative agreement with parton shower simulations, and we present various infrared and collinear safe observables that are sensitive to this flavor definition.

# Jet flavour (accepting collinear unsafety)

## Flavor Fragmentation Function Factorization

**Andrew J. Larkoski**[a] **and Duff Neill**[b]

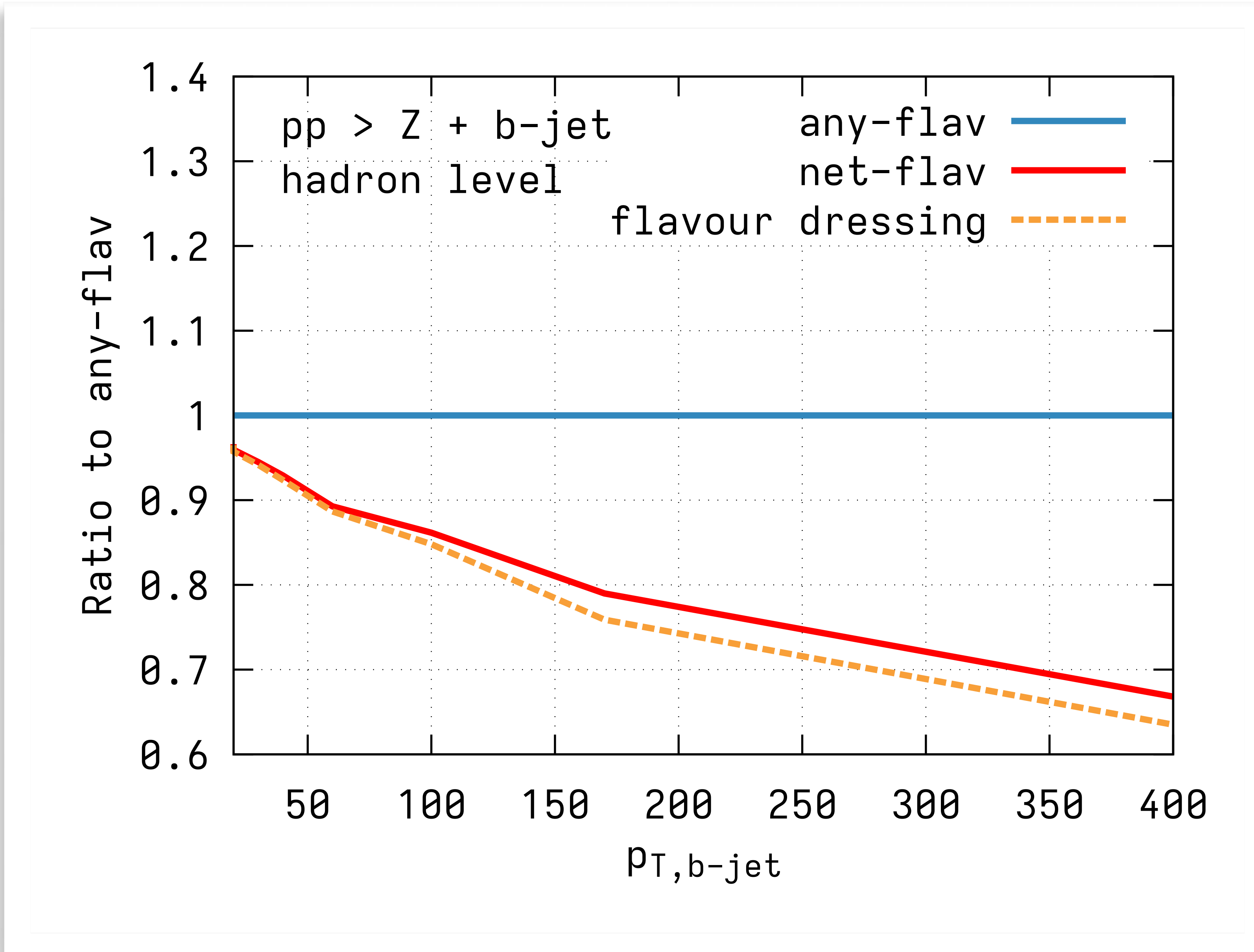
[a] *Department of Physics and Astronomy, University of California, Los Angeles, CA 90095, USA*
*Mani L. Bhaumik Institute for Theoretical Physics, University of California, Los Angeles, CA 90095, USA*

[b] *Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico, 87545, USA*

*E-mail:* larkoski@ucla.edu, dneill@lanl.gov

Abstract: A definition of partonic jet flavor that is both theoretically well-defined and experimentally robust would have profound implications for measurements and predictions especially for heavy flavor applications. Recently, a definition of jet flavor was introduced as the net flavor flowing along the direction of the Winner-Take-All axis of a jet which is soft safe to all orders, but not collinear safe. Here, we exploit the lack of collinear safety and propose a factorization theorem of perturbative flavor fragmentation functions that resum collinear divergences and describe the evolution of flavor from the short distance of jet production to the long distance at which hadronization occurs. Collinear flavor evolution is governed by a small modification of the DGLAP equations. We present a detailed all-orders analysis and identify exact relations that must hold amongst the various anomalous dimensions by probability conservation and the existence of fixed points of the renormalization group flow. We explicitly validate the factorization theorem at one-loop order, and demonstrate its consistency at two loops in particular flavor channels. Starting at two-loops, constraints on phase space imposed by flavor measurements potentially allow for non-trivial soft contributions, but we demonstrate that they are scaleless and so explicitly vanish, ensuring that soft particles are summed inclusively and all divergences are exclusively collinear in nature. This factorization theorem opens the door to precision calculations with identified flavor in the infrared.

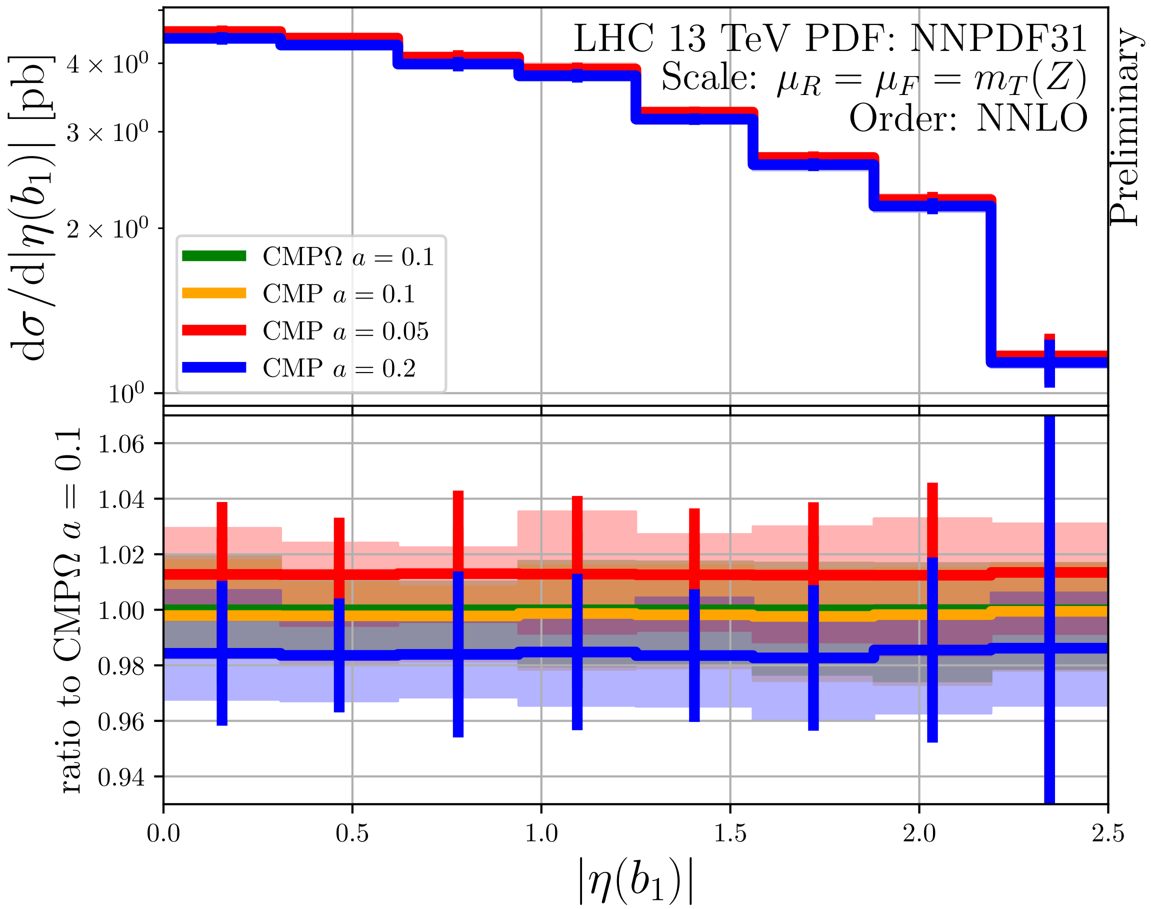
# Jet flavour (accepting collinear unsafety)

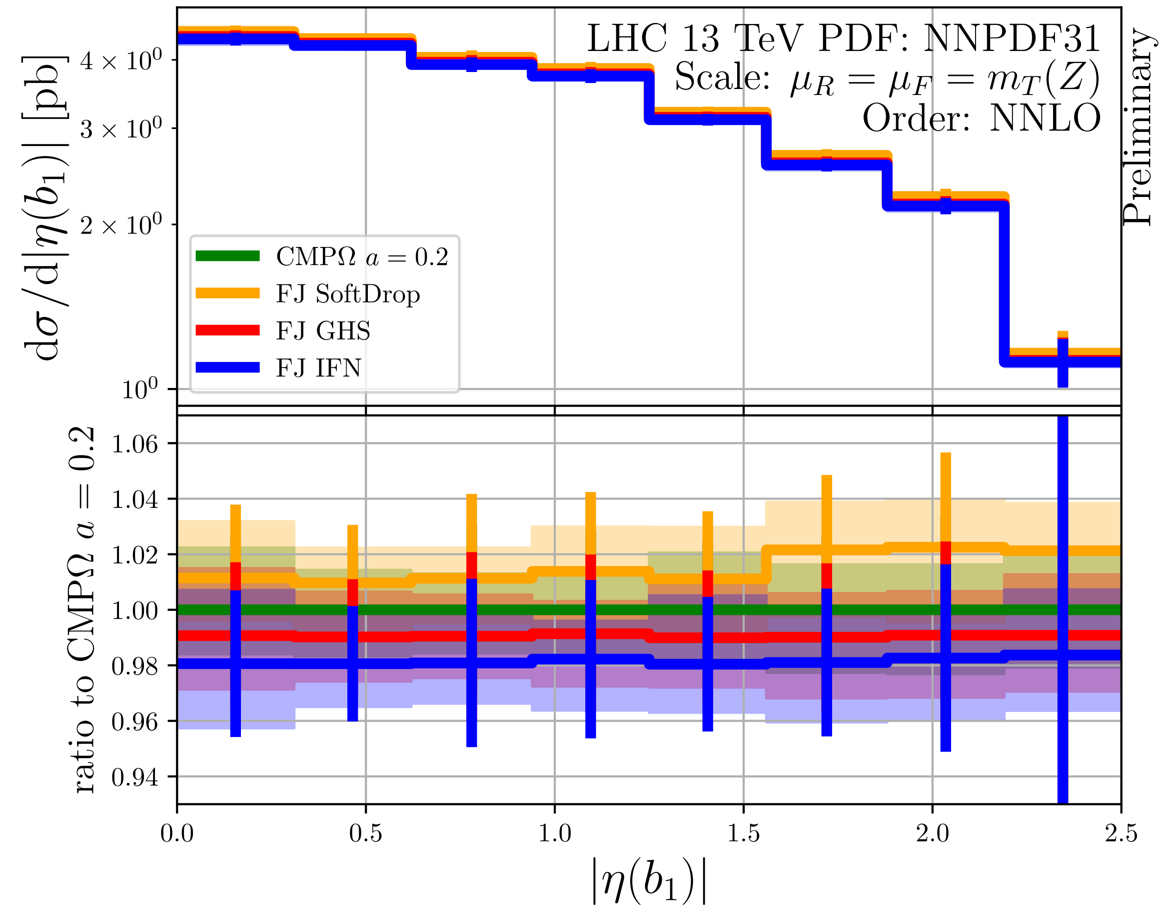



A better description of this collinear unsafe region clearly desired

Note: there is also the $p_T^D > \ 5 \text{ GeV}$ issue (would require resummed pp>Z+B)

# Z+b-jet comparison (CMP, and others)



# (Revised)$^n$ flavour dressing

1. Initialise empty sets $\text{tag}_k = \varnothing$ for each jet $j_k$ to accumulate all flavoured particles assigned to it.

2. Populate a set $\mathcal{D}$ of distance measures based on all possible pairings:

   (a) For each unordered pair of particles $k_i$ and $k_j$, add the distance measure $d_{k_i k_j}$, except in the following case: if one (or both) of the two particles is flavourless and they do not share the same jet association (in order words, we allow for distances involving flavourless particles only with other particles in the same jet). When the two

3. While the set $\mathcal{D}$ is non-empty, select the pairing with the smallest distance measure:

   (a) $d_{k_i k_j}$ is the smallest: the two particles merge into a new particle $k_{ij}$ with the sum of 4-momenta and sum of flavours. Update all entries in $\mathcal{D}$ that involve $k_i$ or $k_j$ with new distances involving $k_{ij}$.

   (b) $d_{k_i j_k}$ is the smallest: assign the particle $k_i$ to the jet $j_k$, $\text{tag}_k \to \text{tag}_k \cup \{k_i\}$, and remove all entries in $\mathcal{D}$ that involve $k_i$.

   (c) $d_{k_i B_\pm}$ is the smallest: discard particle $k_i$ and remove all entries in $\mathcal{D}$ that involve $k_i$.

4. The flavour assignment for jet $j_k$ is determined according to the accumulated flavours in $\text{tag}_k$.

# (Revised)$^n$ flavour dressing

For the distance measure between two final-state objects $a$ and $b$ (particles or jets) we use

$$d_{ab} = \Omega_{ab}^2 \;\max\big(p_{\mathrm{T},a}^{\alpha}, p_{\mathrm{T},b}^{\alpha}\big)\;\min\big(p_{\mathrm{T},a}^{2-\alpha}, p_{\mathrm{T},b}^{2-\alpha}\big)\,, \tag{1a}$$

with

$$\Omega_{ab}^2 = 2\left(\frac{1}{\omega^2}(\cosh(\omega\Delta y_{ab}) - 1) - (\cos\Delta\phi - 1)\right)\,. \tag{1b}$$

The distance between a particle and a hadron beam in the direction of positive (+) or negative (−) rapidity is

$$\begin{aligned} d_{iB_\pm} &= \max\big(p_{\mathrm{T},i}^{\alpha}, p_{\mathrm{T},B_\pm}^{\alpha}(y_i)\big)\;\min\big(p_{\mathrm{T},i}^{2-\alpha}, p_{\mathrm{T},B_\pm}^{2-\alpha}(y_i)\big)\,,\\ p_{\mathrm{T},B_\pm}(y) &= \sum_{k=1}^{m} p_{\mathrm{T},j_k}\Big[\Theta(\pm\Delta y_{j_k}) + \Theta(\mp\Delta y_{j_k})\;\mathrm{e}^{\pm\Delta y_{j_k}}\Big]\,, \end{aligned} \tag{1c}$$

with the rapidity separation $\Delta y_{j_k} = y_{j_k} - y$ and $\Theta(0) = \frac{1}{2}$.

# Accumulating flavour quantum numbers

**BSZ** hep-ph/0601139 and 0704.2999

**How to count flavour quantum numbers:**

- With charge info. ($q$ vs $\bar{q}$), then $q = +1$ and $\bar{q} = -1$

  (net flavour is sum of the $q_i$ and $\bar{q}_j$ assigned to jet $j_n$)

- If one cannot (e.g. experiment), $|q| = |\bar{q}| = 1$

  (net flavour is sum [modulo 2] of the $q_i$ and $\bar{q}_j$ assigned to jet $j_n$)

  [i.e. even tagged jets are NOT flavoured]

# (1/3 OLD) collinear-safe flavoured objects

(**RG**, Huss, Stagnitto arXiv:2208.11138)

flavoured particles (quarks, hadrons) not collinear safe. Define new objects:

i) Initialise a list of all particles

ii) Add to the list all flavoured particles, removing any overlap

iii) Calculate the distances $d_{ij} = \Delta R_{ij}^2$ between all particles

iv) If $d_{ij}^{\min} > \Delta R_{\mathrm{cut}}^2$ terminate the clustering. Otherwise:

1. (i & j flavourless) replace i & j in the list with combined object ij
2. (i or j flavoured) combine i and j if:

$$\frac{\min(p_{T,i}, p_{T,j})}{p_{T,i} + p_{T,j}} > z_{\mathrm{cut}} \left( \frac{\Delta R_{ij}}{R_{\mathrm{cut}}} \right)^{\beta}$$

[Soft-drop]
(Larkoski et al. arXiv:1402.2657)

Otherwise:

(i & j flavoured) remove both from list

(i or j flavourless) remove only flavourless object

[Repeat until list empty, or no flavoured particles left]

# (2/3 OLD) Association criterion and counting

(**RG**, Huss, Stagnitto arXiv:2208.11138)

We now have have $\{j_1, \ldots, j_m\}, \left\{\hat{f}_1, \ldots, \hat{f}_n\right\}$

We introduce an **Association criterion** for $\hat{f}_a$ with $j_b$ (some possibilities):

- the flavoured particle $f_a$ is a constituent of jet $j_b$
- or $\Delta R(\hat{f}_a, j_b) < R_{\rm tag}$
- or Ghost association of $\hat{f}_a$ (include direction of $\hat{f}_a$ in anti-$k_T$ clustering)

(association criterion required as not assumed that $f_a$ is a stable particle)

Introduce a **Counting** or **Accumulation** for flavour:

- with charge info. ($q$ vs $\bar{q}$), then $q = +1$ and $\bar{q} = -1$ (net flavour is sum)
- if one cannot (e.g. experiment), $q = \bar{q} = 1$ (net flavour is sum modulo 2)

  [i.e. jets with even number of $q_i + \bar{q}_j$ are NOT flavoured]

# (3/3 OLD) The flavour dressing algorithm

(**RG**, Huss, Stagnitto arXiv:2208.11138)

We now have have $\{j_1, \ldots, j_m\}, \left\{\hat{f}_1, \ldots, \hat{f}_n\right\}$, association, and counting rules

Dressing algorithm:

- Calculate a set of distances between the flavoured objects, jets and beam:
  - [ff] $d_{ab}$ between all all flavoured objects $\hat{f}_a$ and $\hat{f}_b$
    - [fj] $d_{ab}$ between $\hat{f}_a$ and $j_b$ ONLY if there is an association
    - [fB] $d_{aB}$ for all $\hat{f}_a$ without a jet association
- Find the minimum distance of all entries in the list
  - if it is an [fj] assign $\hat{f}_a$ to $j_b$ (removing entries involving $\hat{f}_a$ from list)
  - otherwise just remove $\hat{f}_a$ [fB] or $\hat{f}_a$ and $\hat{f}_b$ [ff] from the list

[repeat until list empty]

- The flavour of each jet is then just the accumulation of its flavour

# (4/3 OLD) The distance measure

(**RG**, Huss, Stagnitto arXiv:2208.11138)

Stage 2: Calculate $d_{ab}$ for all $\{j_1, \ldots, j_m\}, \left\{\hat{f}_1, \ldots, \hat{f}_n\right\}$

$$d_{ab} = \Delta R_{ab}^2 \max\left(p_{T,a}^{\alpha}, p_{T,b}^{\alpha}\right) \min\left(p_{T,a}^{2-\alpha}, p_{T,b}^{2-\alpha}\right)$$

(this is the original flavour-$k_T$ distance)

$$d_{aB\pm} = \max\left(p_{T,a}^{\alpha}, p_{T,B_\pm}^{\alpha}(y_{\hat{f}_a})\right) \ \min\left(p_{T,a}^{2-\alpha}, p_{T,B_\pm}^{2-\alpha}(y_{\hat{f}_a})\right)$$

(rapidity dependent measure of the Beam)

- If min. is $d_{fj}$ : assign $f$ quantum number to j, [remove $f$ from list]
- If min. is: $d_{ff}$ or $d_{fB}$ , [remove $f$ from list]

[repeat until list is empty]

Note: only evaluate $d_{fj}$ if the $f$ is associated to the jet (e.g. a constituent)

[complete details in back-up slides]

# Application of the algorithm (pp)

(**RG**, Huss, Stagnitto arXiv:2208.11138)

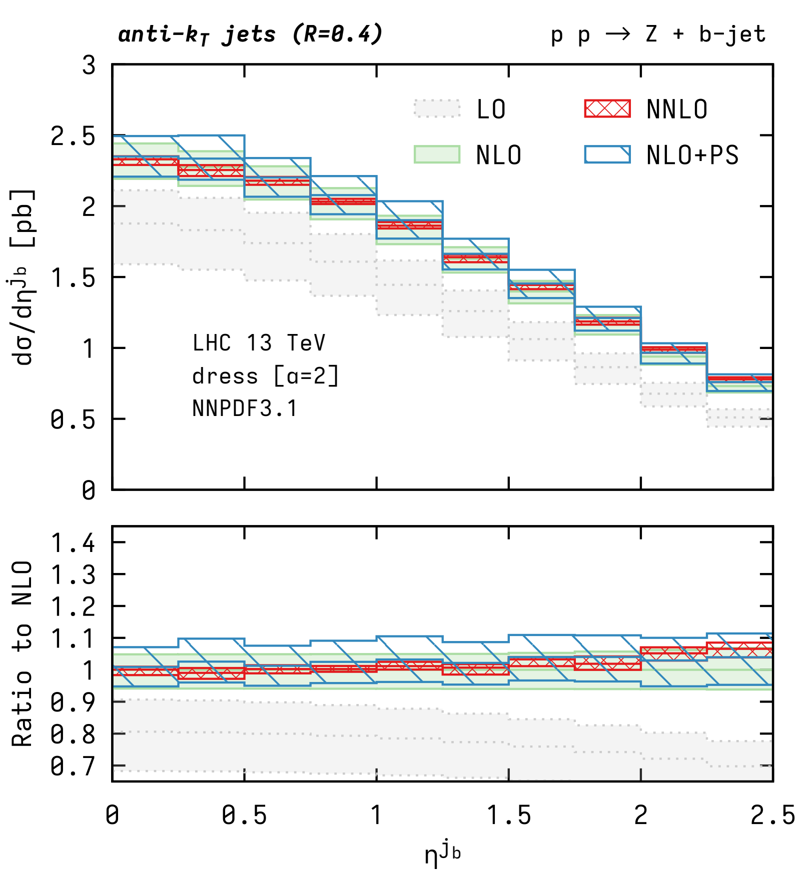
Now consider the process $pp \to Z + b - \text{jet}$ in Fiducial region (13 TeV, CMS-like)

(N)NLO at fixed-order w/ NNLOJET, **RG** et al. arXiv:2005.03016

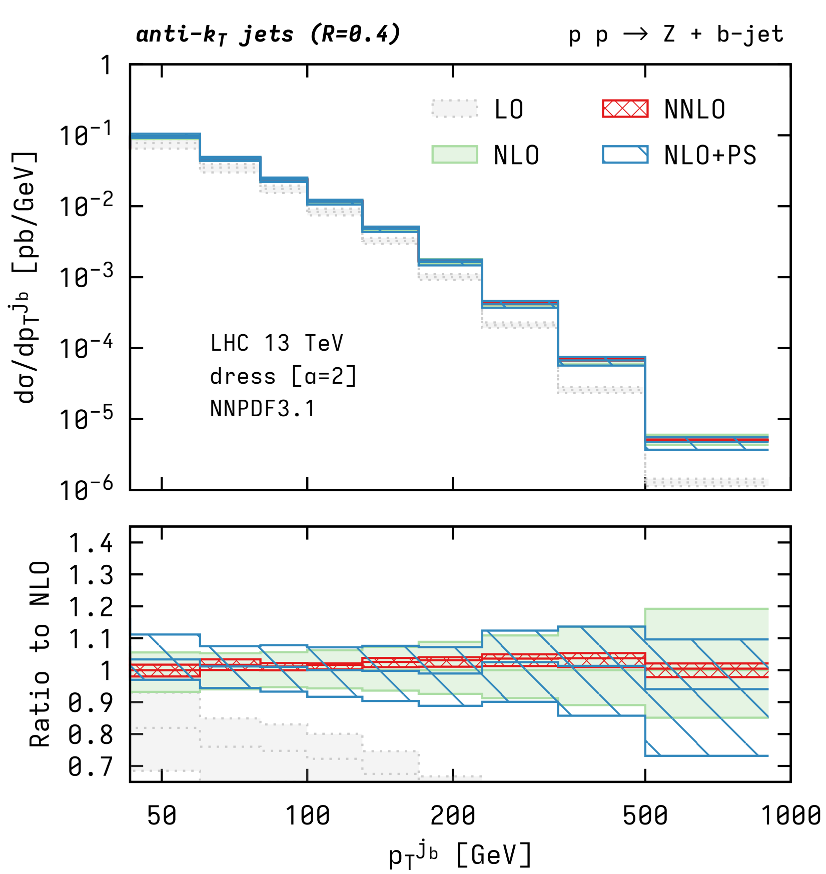
NLO+PS Hadron-level with aMC@NLO interfaced to Pythia8

Tests sensitivity to: all-order effects, hadronisation (also FO IRC safety in pp)

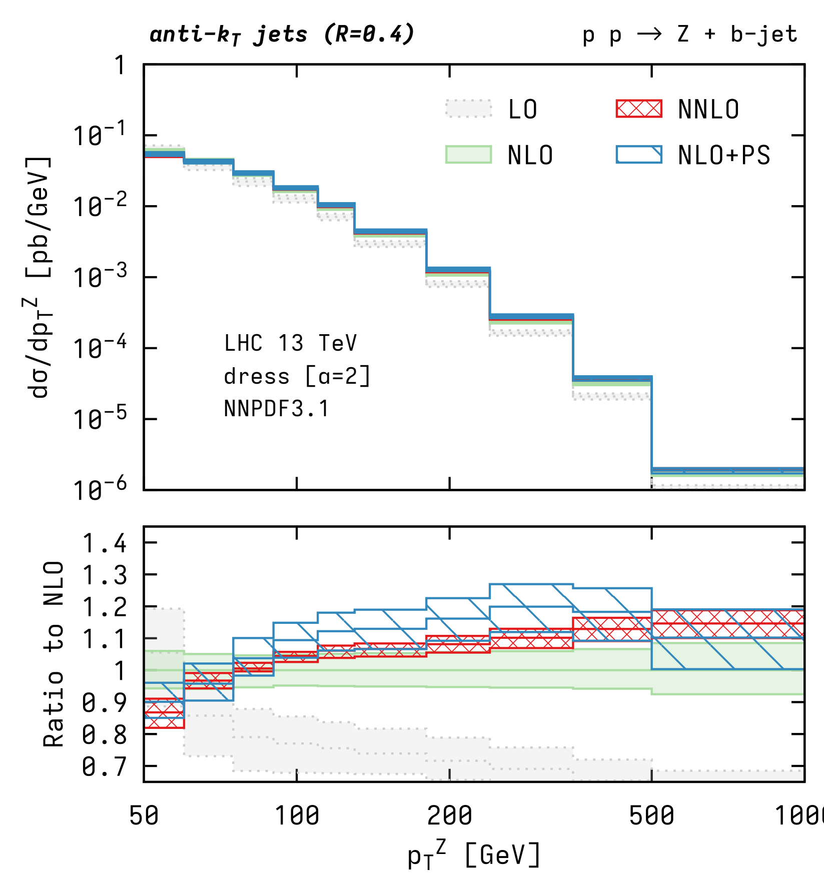
$\eta_{\text{b-jet}}$




$p_{T,\text{b-jet}}$




$p_{T,Z}$



# Applications: Z+c-jet at LHCb

LHCb measurement (13 TeV), arXiv: 2109.08084

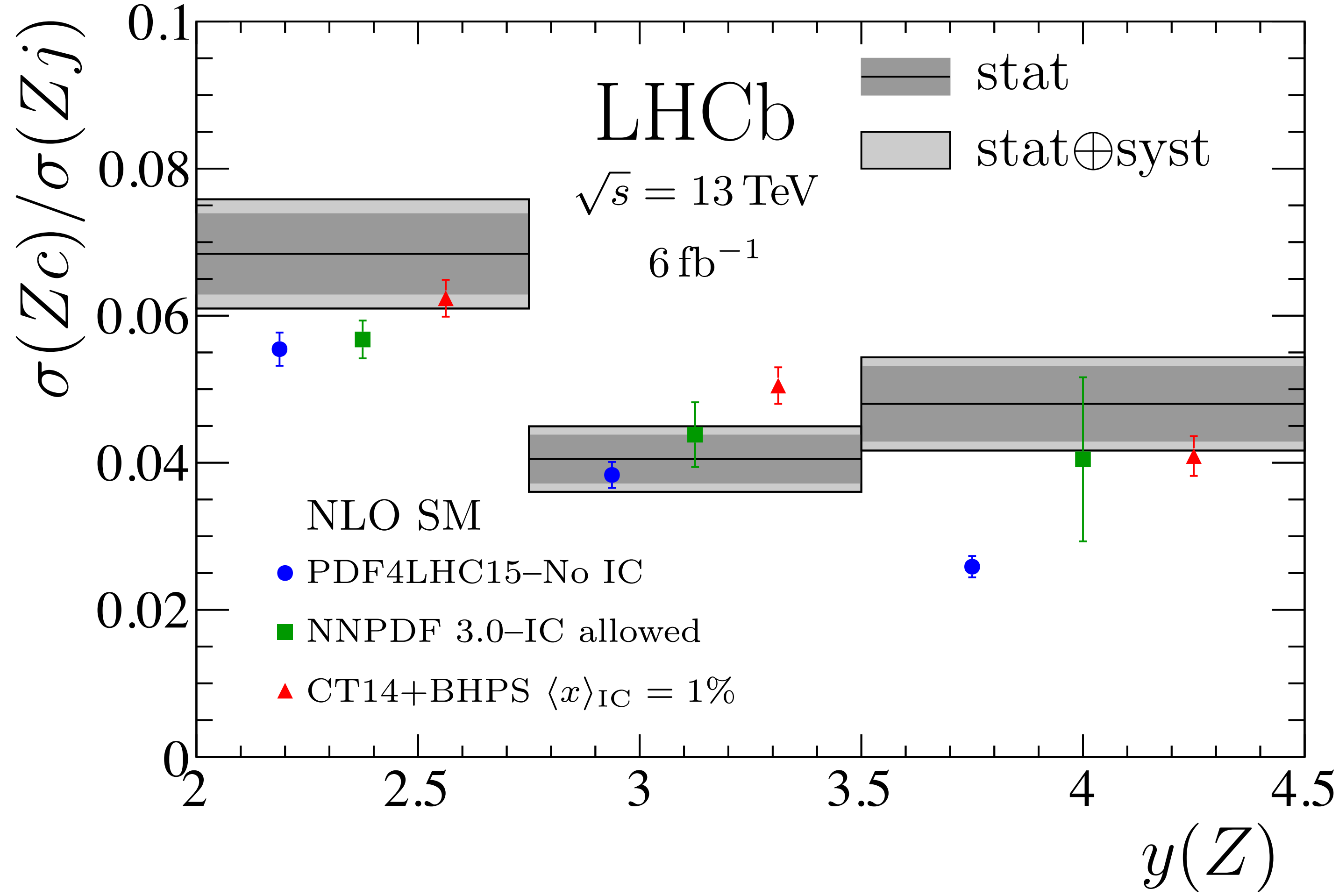
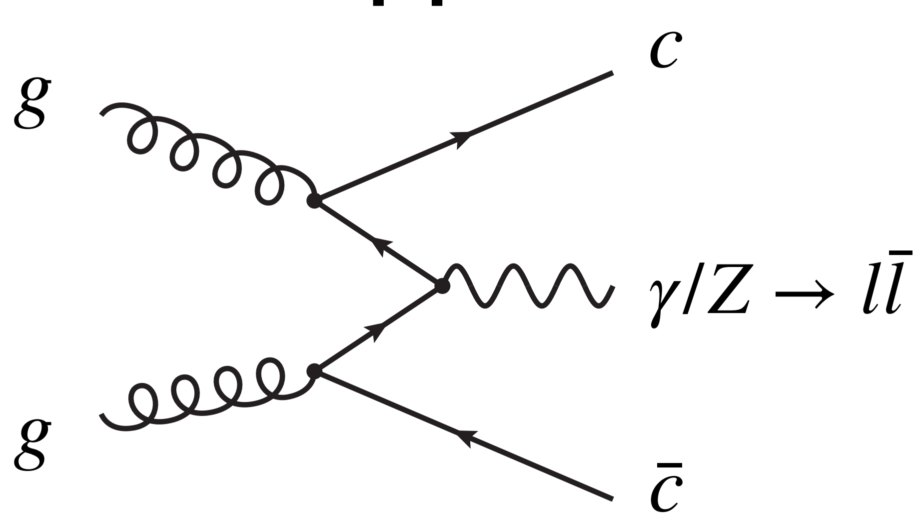
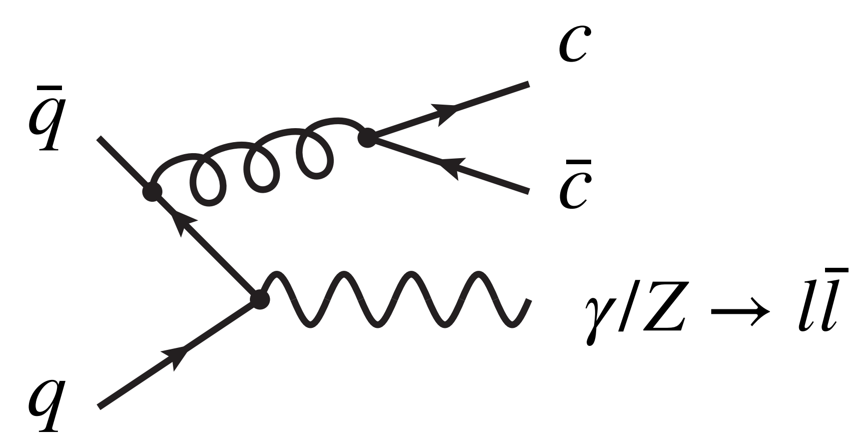



Table 1: Definition of the fiducial region.

| | |
|---|---|
| $Z$ bosons | $p_{\rm T}(\mu) > 20$ GeV, $2.0 < \eta(\mu) < 4.5$, $60 < m(\mu^+\mu^-) < 120$ GeV |
| Jets | $20 < p_{\rm T}(j) < 100$ GeV, $2.2 < \eta(j) < 4.2$ |
| Charm jets | $p_{\rm T}(c \text{ hadron}) > 5$ GeV, $\Delta R(j, c \text{ hadron}) < 0.5$ **IRC unsafe** |
| Events | $\Delta R(\mu, j) > 0.5$ |

# Applications: Z+c-jet at LHCb

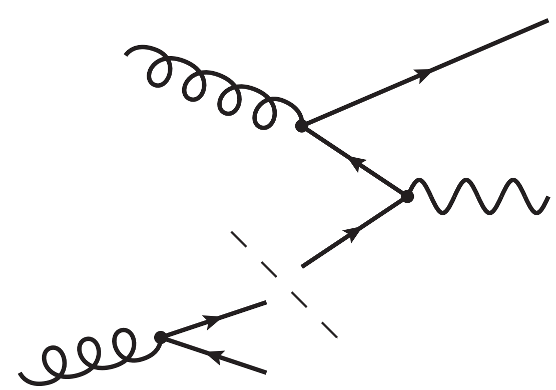



Calculated in the 3fs scheme (i.e. $n_f^{\max} = 3$ in PDFs, and $\alpha_s$ evolution)

$$d\sigma^{3fs} = d\sigma^{m_c=0} \quad + \quad d\sigma^{\ln[m_c]} \quad + \quad d\sigma^{m_c}$$

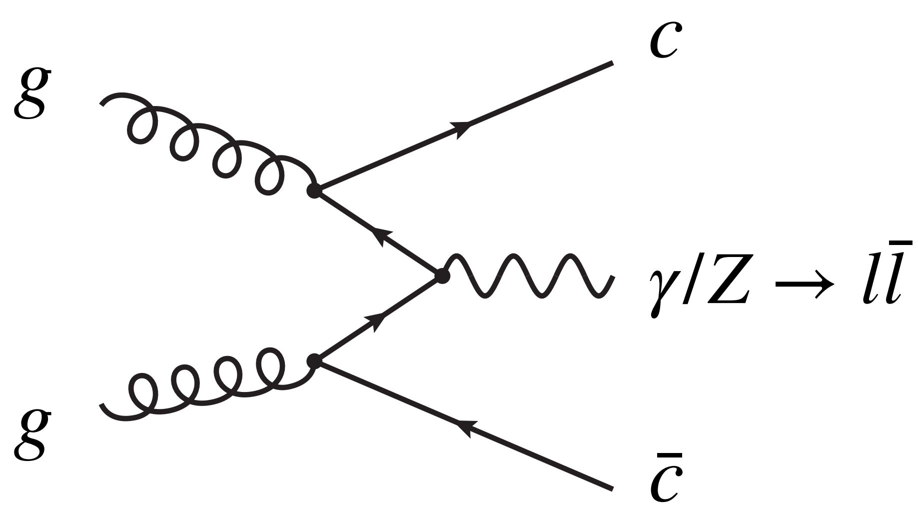
Massless component
$\mathcal{O}(\alpha_s^2 n_f)$ in 4fs

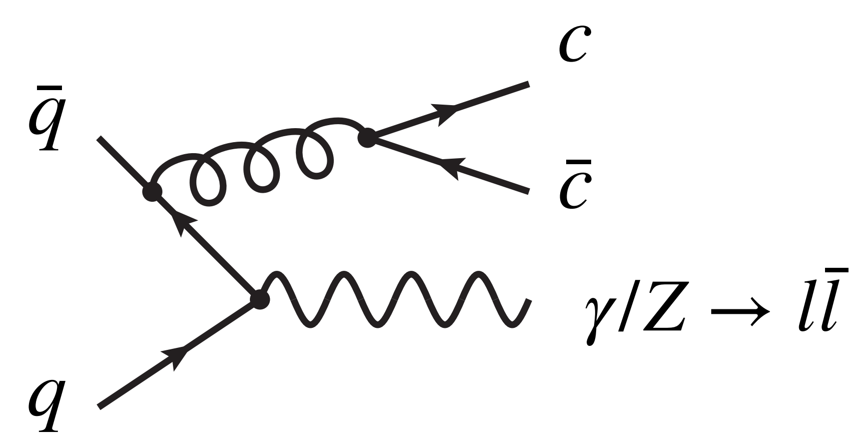


$\mathcal{O}(m_c^2)$ effects
(exact kinematics)

Note, initial-state mass singularities still there (even with IRC safe jet alg.)

# Applications: Z+c-jet at LHCb




Calculated in the 3fs scheme (i.e. $n_f^{\max} = 3$ in PDFs, and $\alpha_s$ evolution)

$$\mathrm{d}\sigma^{3fs} = \mathrm{d}\sigma^{m_c=0} \quad + \quad \mathrm{d}\sigma^{\ln[m_c]} \quad + \quad \mathrm{d}\sigma^{m_c}$$

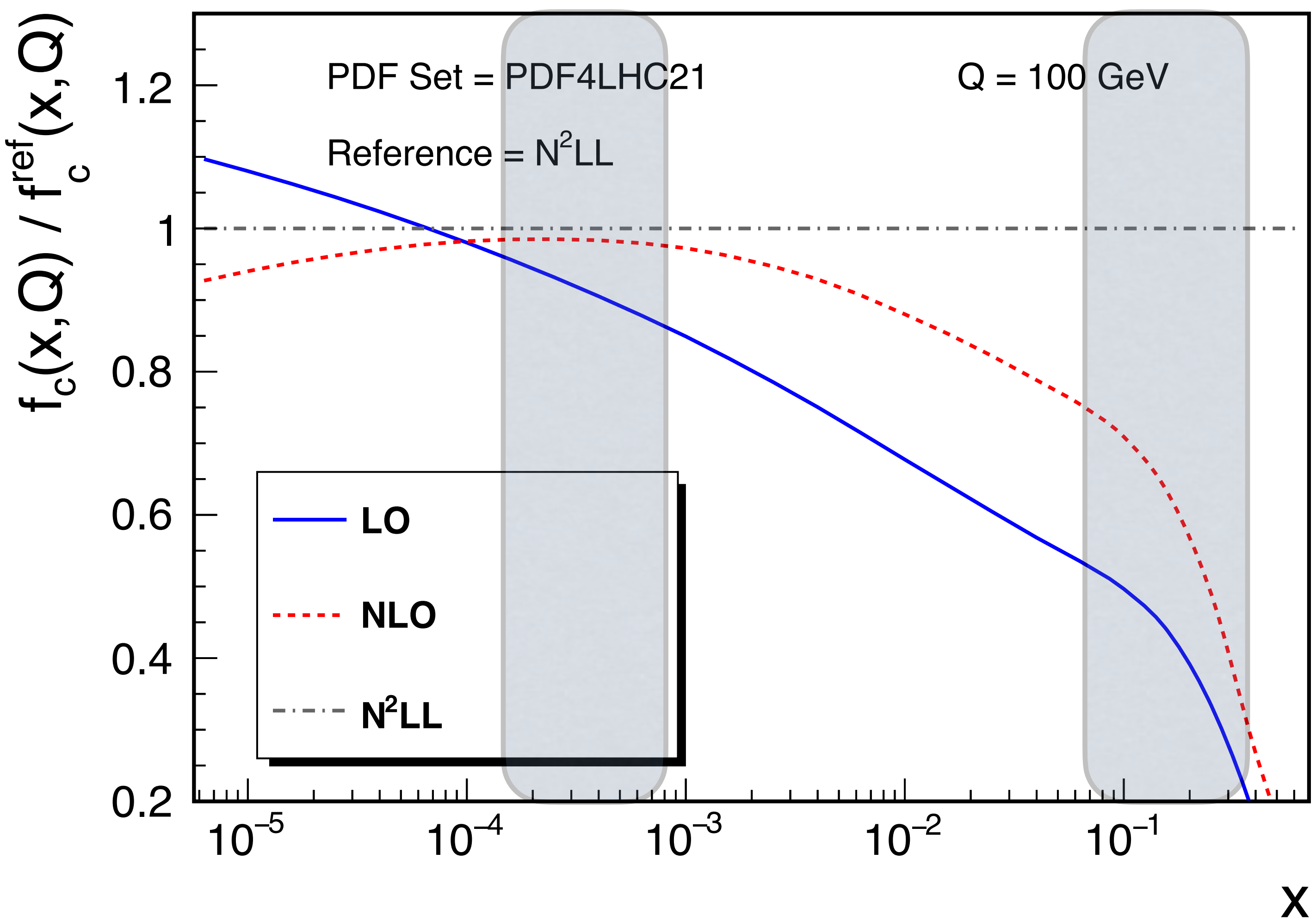
**0.220 = +0.0364 +0.203 -0.019 [pb]**

**100% = +16% +92% -8%**

Note, initial-state mass singularities still there (even with IRC safe jet alg.)

# Applications: Z+c-jet at LHCb

The perturbative corrections are enormous: resummation critical
(this class of logarithm resummed by PDF evolution)




$< x_1 > \sim 0.2$

$< x_2 > \sim 6 \times 10^{-4}$

LHCb cross-section: Leading Log (1st order) = 0.203pb, Leading Log (resumed) = 0.332pb

I am showing fixed-order pdf versus a resummed one (PDF evolution)

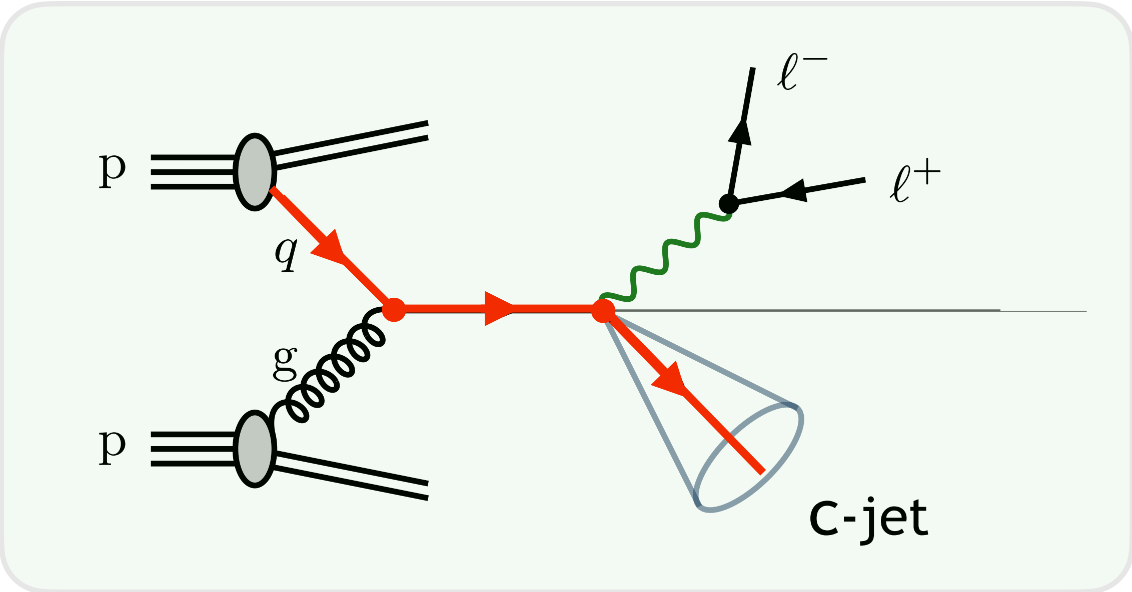
$\alpha_s^m \ln^n[\mu_F^2/m_c^2], \quad m \geq n$

Note! $\alpha_s \ln[m_Z^2/m_c^2] \approx 1.0$

# Applications: Z+c-jet at LHCb

**RG**, Gehrmann-De Ridder, Glover, Huss, Rodriguez Garcia, Stagnitto, arXiv:2302.12844

- Theory study based on SPS predictions (no MPI corrections)
- Consider a fiducial region matching that of the LHCb experiment




Introduce the constraint

$$p_T(Zj_c) \leq p_T(j_c)$$

Predictions are provided in a Massive - Variable Flavour Number Scheme

**RG**, Gehrmann-De Ridder, Glover, Huss, Maier, arXiv:2005.03016, **RG**, arXiv:2107.01226

$$\mathrm{d}\sigma^{M-VFNS} = \mathrm{d}\sigma^{ZM-VFNS} + \mathrm{d}\sigma^{pc}$$

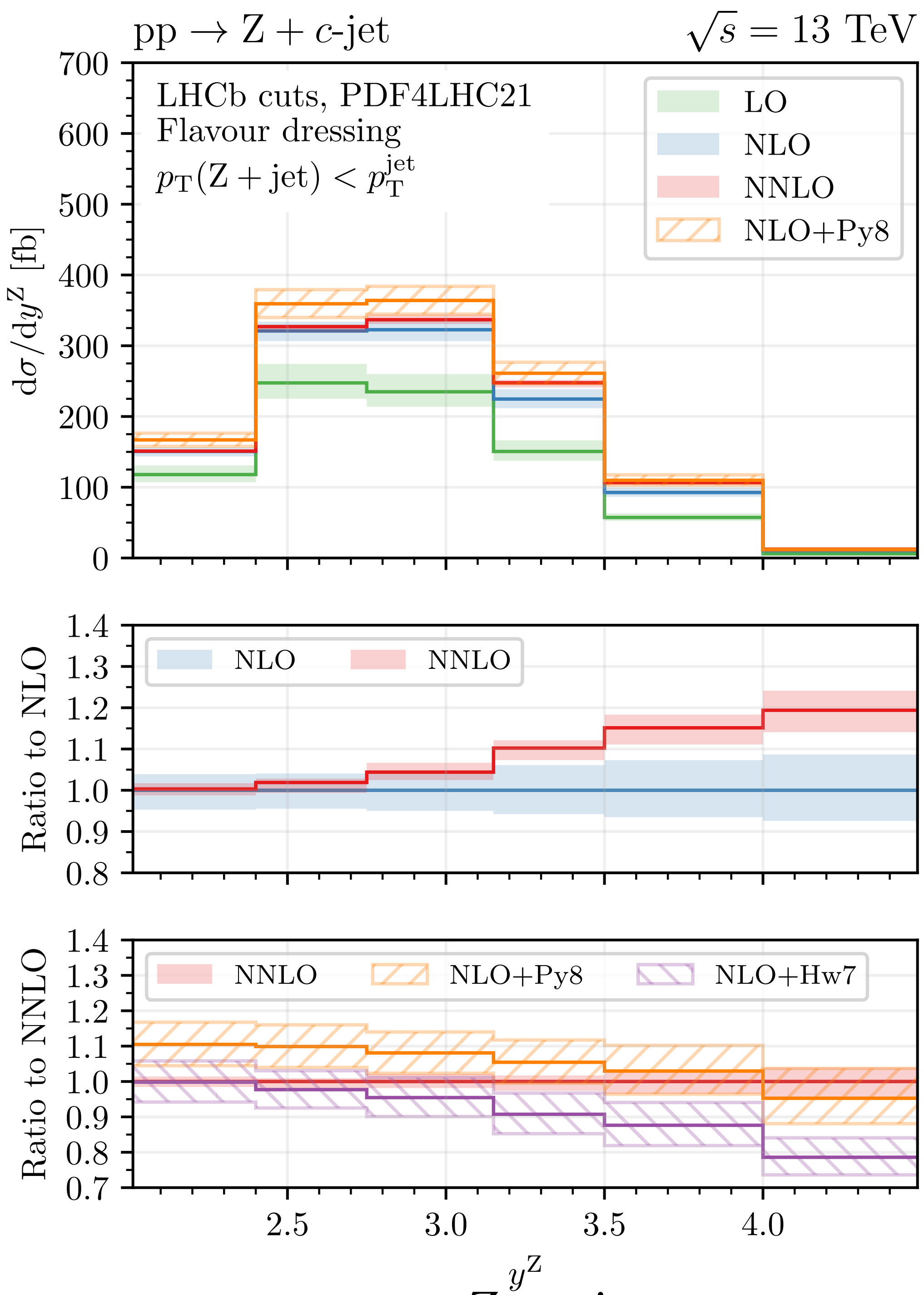
NNLO QCD predictions via the Z+jet antenna subtraction calculation

Gehrmann-De Ridder, Gehrmann, Glover, Huss, Morgan, arXiv:1507.02850

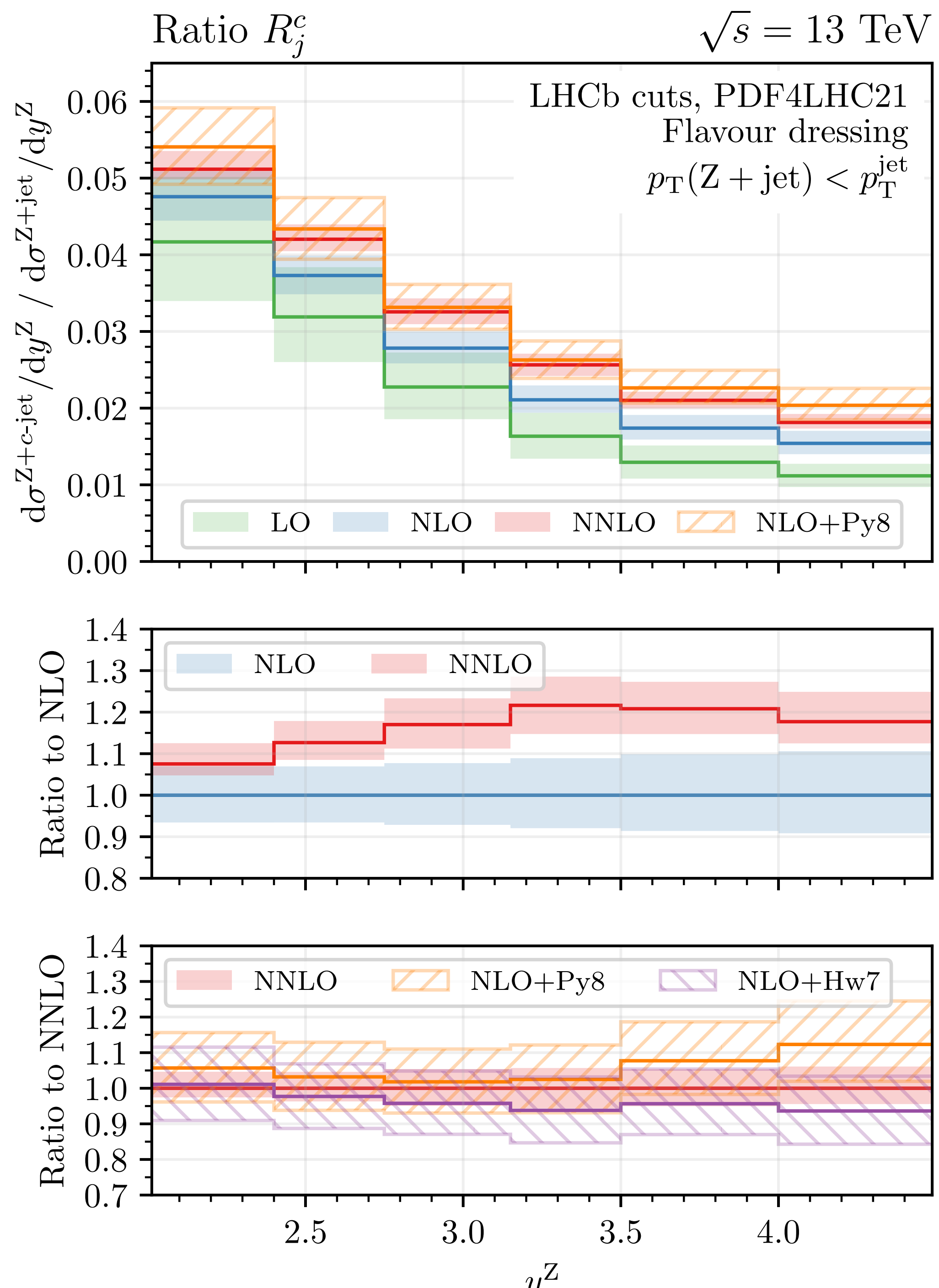
$\alpha_{G_\mu}$ scheme, 7-point scale variation around $E_{T,Z}$ , and the PDF4LHC21 set

arXiv:2203.05506

# Applications: Z+c-jet at LHCb




$$\frac{d\sigma^{Z+c-jet}}{dy_Z}$$




$$\frac{d\sigma^{Z+c-jet}}{dy_Z} \bigg/ \frac{d\sigma^{jet}}{dy_Z}$$

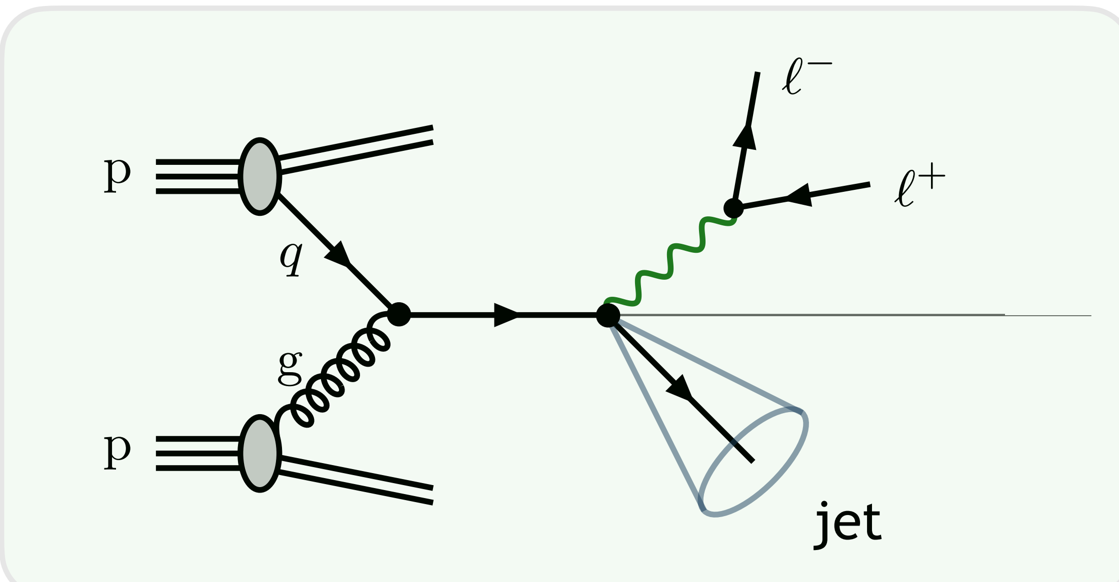
NNLO QCD corrections positive and grow with $y_Z$

# Applications: Z+c-jet at LHCb … MPI

Possibility for multiple hard interactions in a single pp-collision

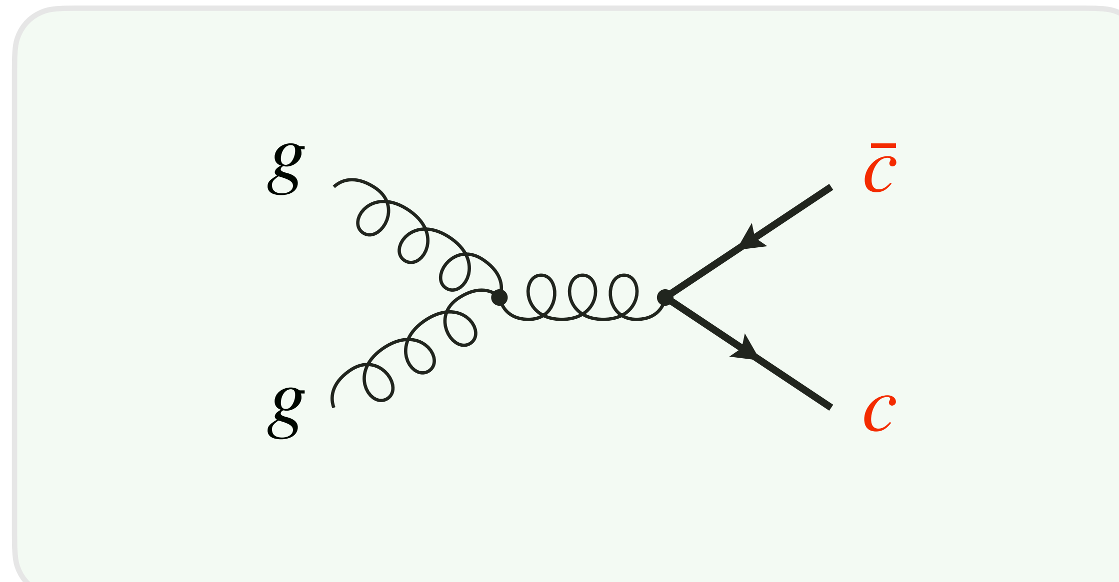
e.g. single-parton-scattering (SPS), double-parton-scattering (DPS), …

Hard Process 1 (HP1) = Z+jet




The jet is flavour inclusive

Hard Process 2 (HP2) = $c\bar{c}$




Large cross-section at LHCb

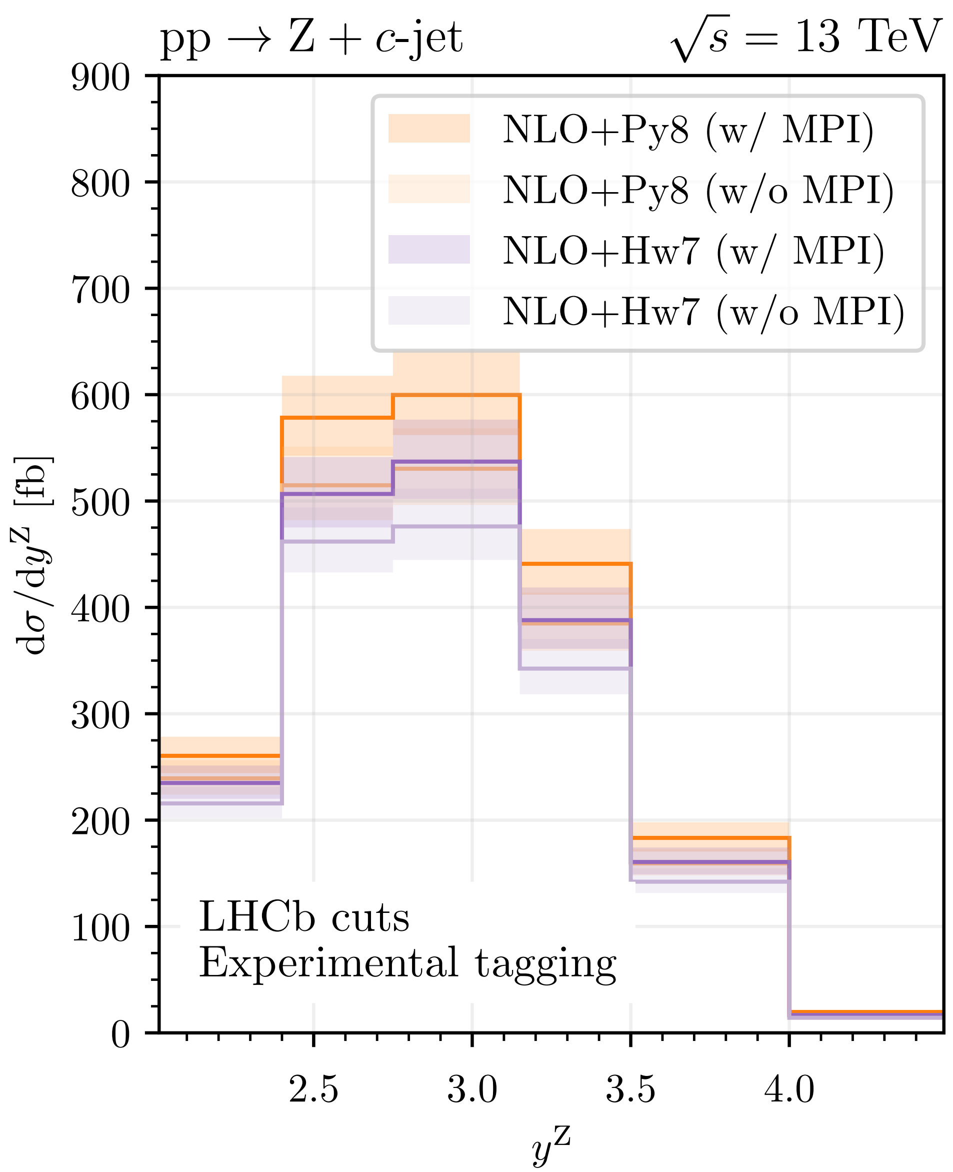
Probability that $\Delta R(j_{HP1}, c_{HP2}) \leq 0.5$ leading to a charm tagged jet

(small phase-space compensated by large $c\bar{c}$ cross-section)

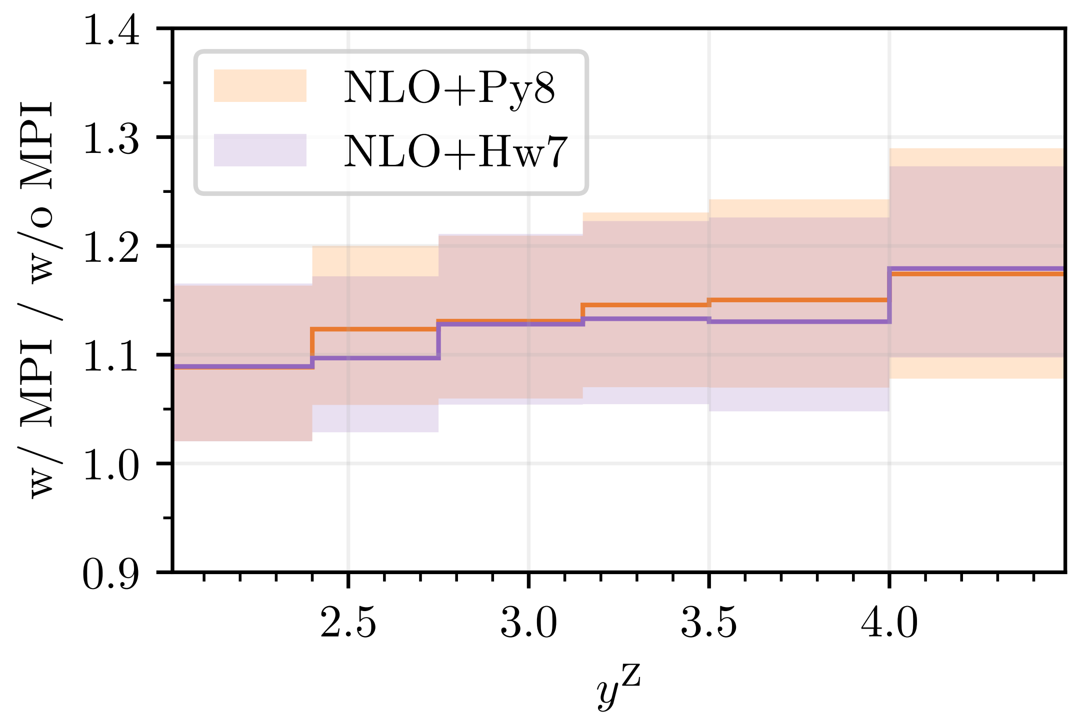
# Applications: Z+c-jet at LHCb … MPI

Possibility for multiple hard interactions in a single pp-collision

e.g. single-parton-scattering (SPS), double-parton-scattering (DPS), …



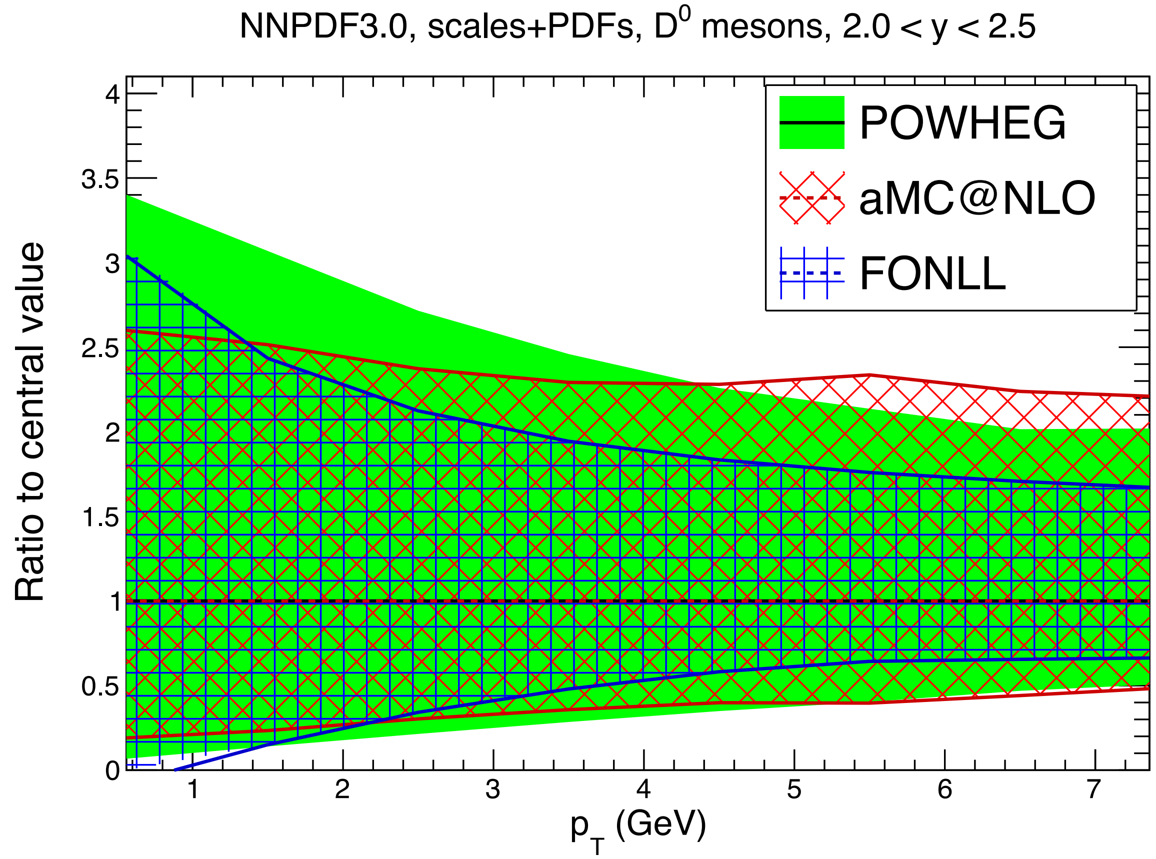

Z+c-jet (MPI on / MPI off)

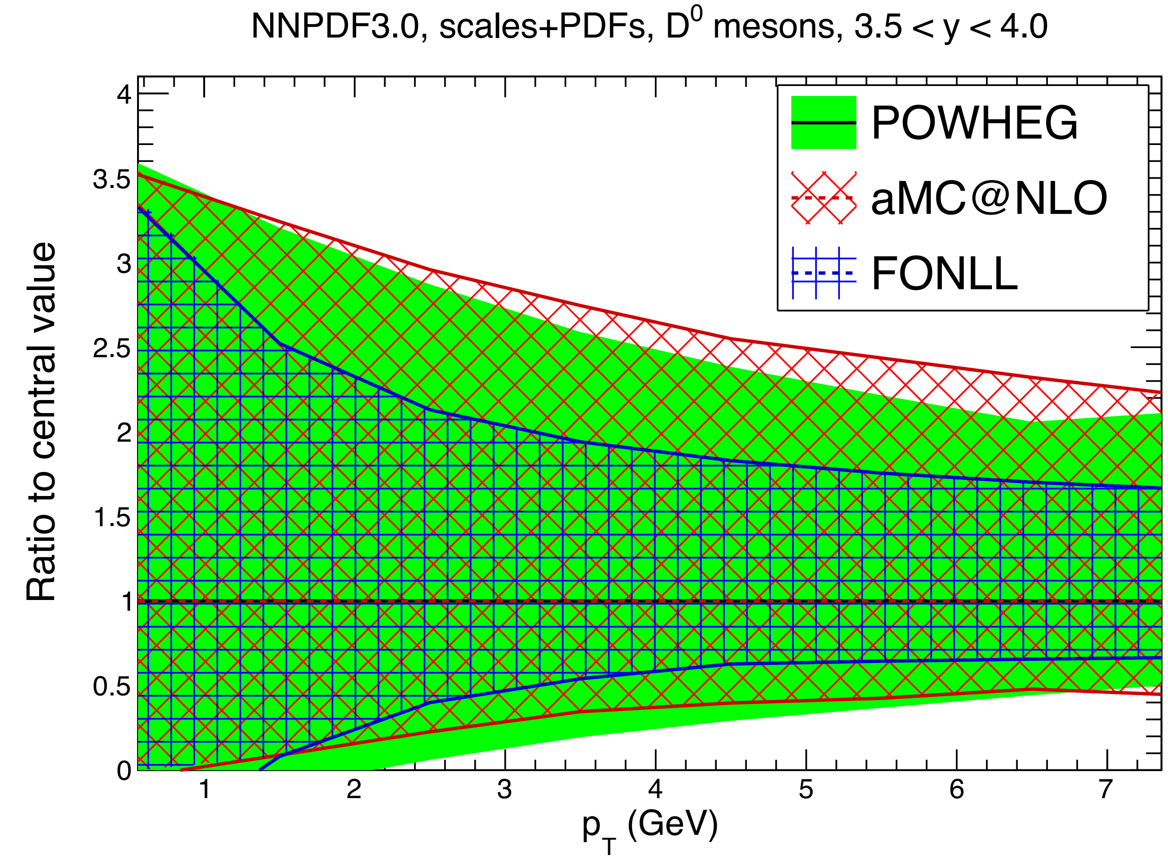



MPI correction required when the considered observable is sensitive to the <u>combination</u> of H1 and H2 (a genuine physics effect not described by SPS)

# Heavy-quark pair production

(RG et al., arXiv:1506.08025)




These are the theory uncertainties (PDF+scales) for D-cross section at LHCb

With a requirement of $P_{T,c} > 5$ GeV QCD uncertainties >> 50% (at best)

The charm MPI component generates a ~15% contribution to LHCb Z+c-jet $\sigma$

Extracting the SPS component will lead to increased uncertainties (>>7.5%)

# Z+b-jet and unfolding

How to account for theory-experiment mismatch?

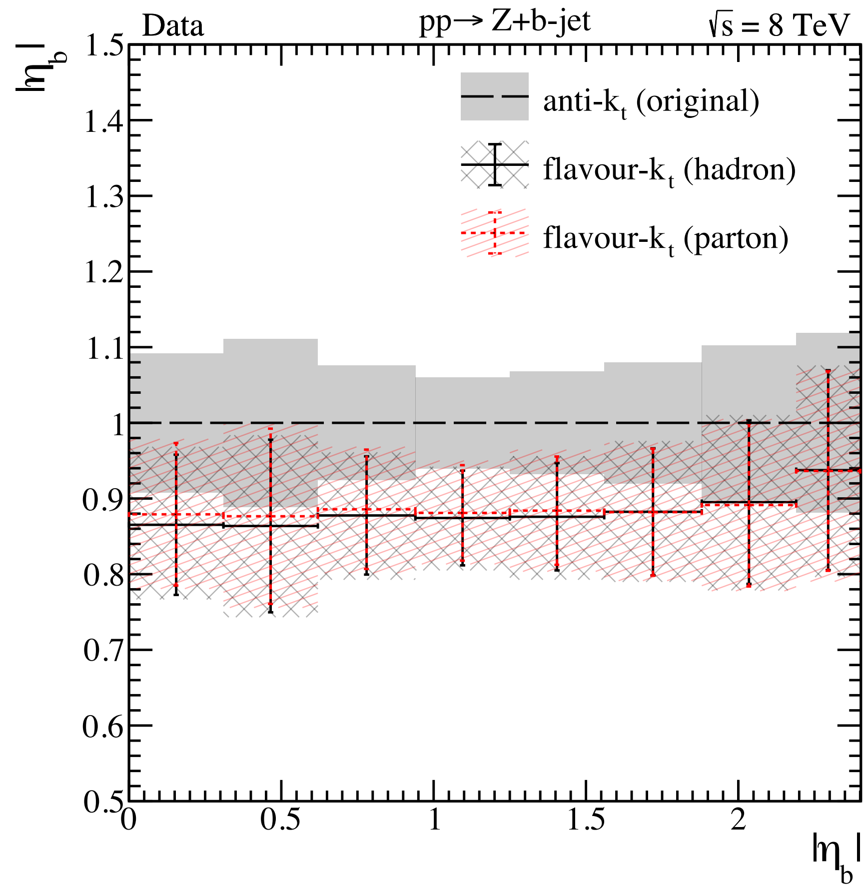
[Gauld, Gehrmann-De Ridder, Glover, Huss, Majer] *PRL* 125 (2020) 22, 222002

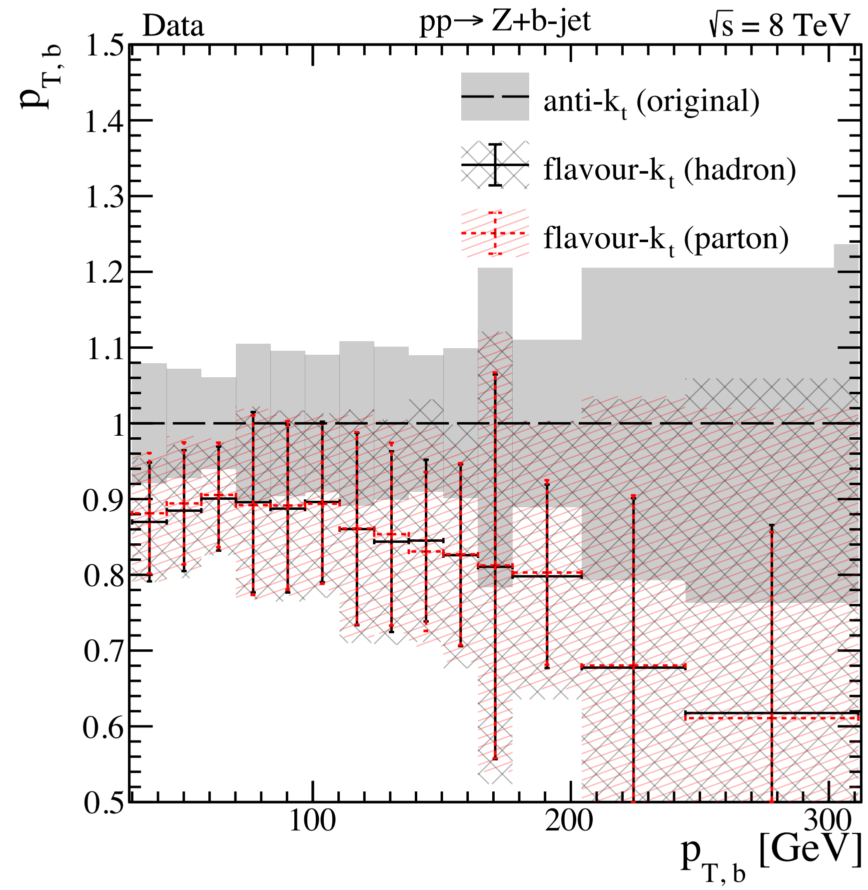
Use an NLO + Parton Shower prediction (which can evaluate both)

1) Prediction at parton-level, flavour-$k_T$ algorithm **(Theory)**

2) Prediction at hadron-level, anti-$k_T$ algorithm **(Experiment)**

Calculate an "Unfolding" correction from 2) Experiment → 1) Theory




We use RooUnfold (following the procedure used in the exp. analyses)