Dressing jets with flavour in an infrared safe way

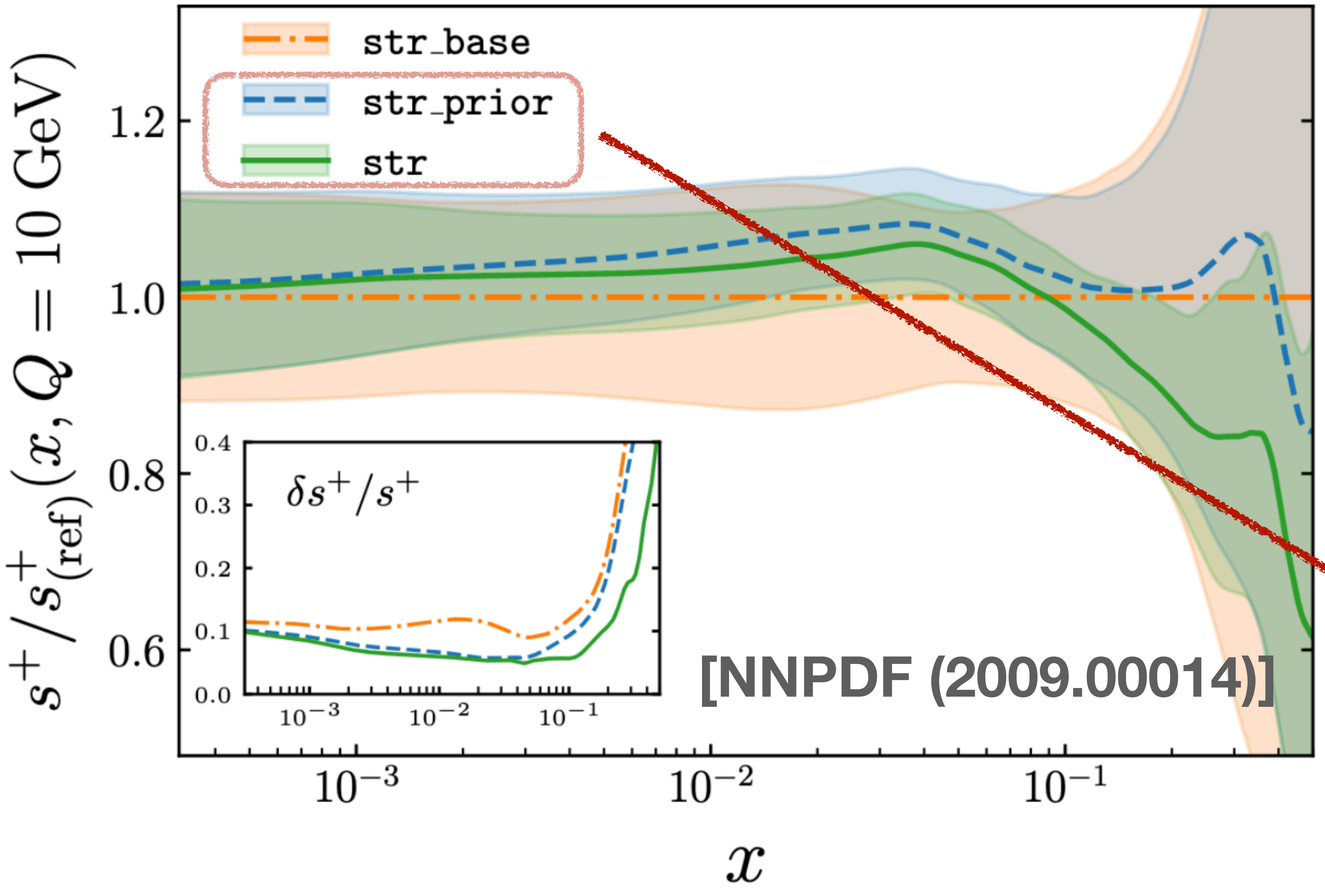
Giovanni Stagnitto

University of Zurich UZH

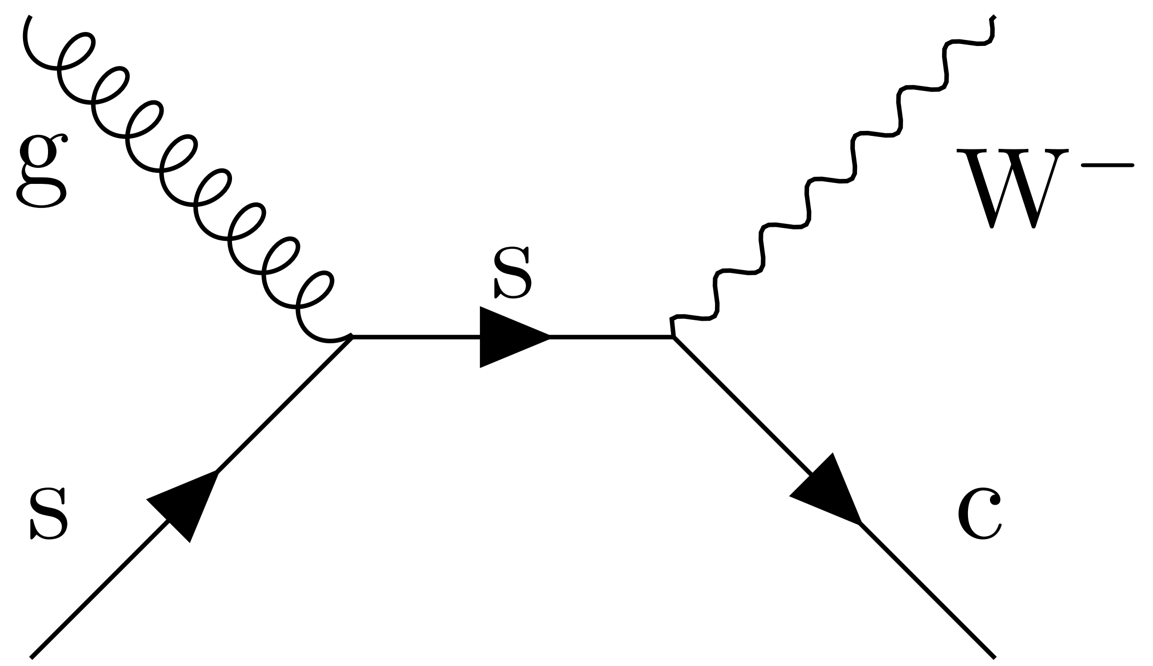
Swiss National Science Foundation

International Workshop on the High Energy CEPC, 26.10.2022

Why flavoured jets?



An example: W+c-jet
unique probe into the strange PDF

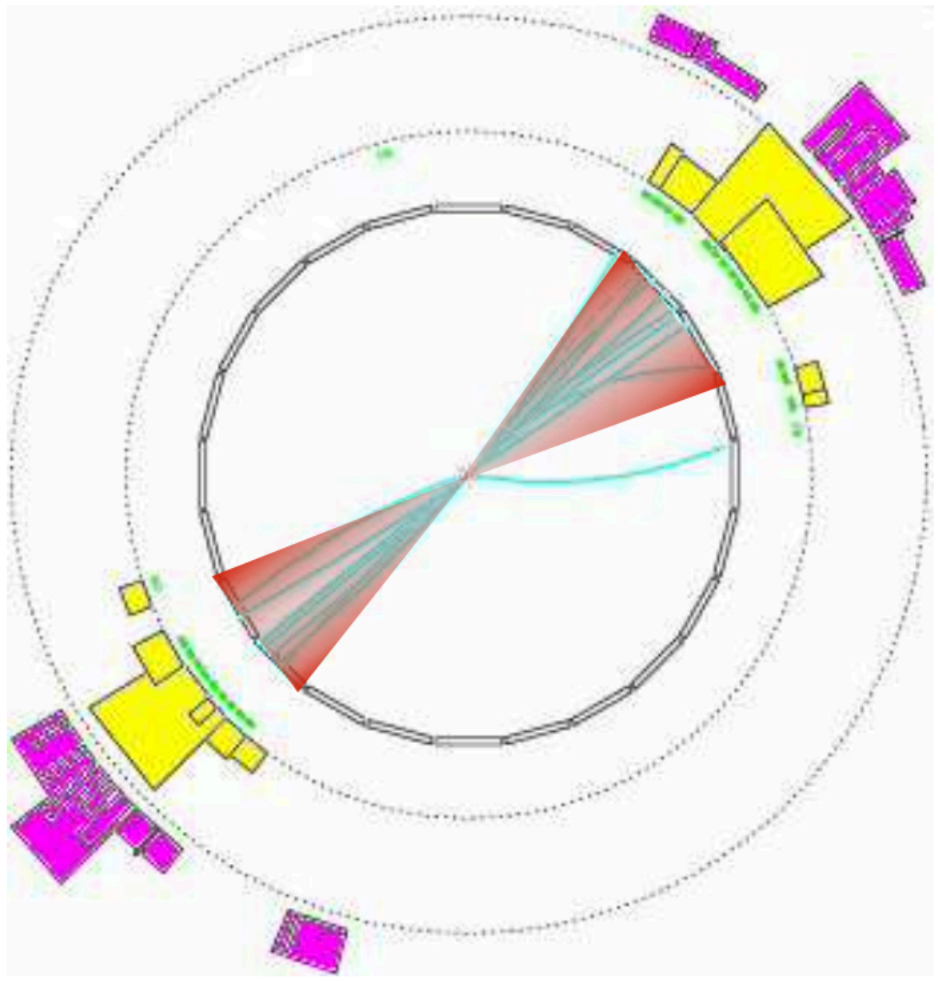


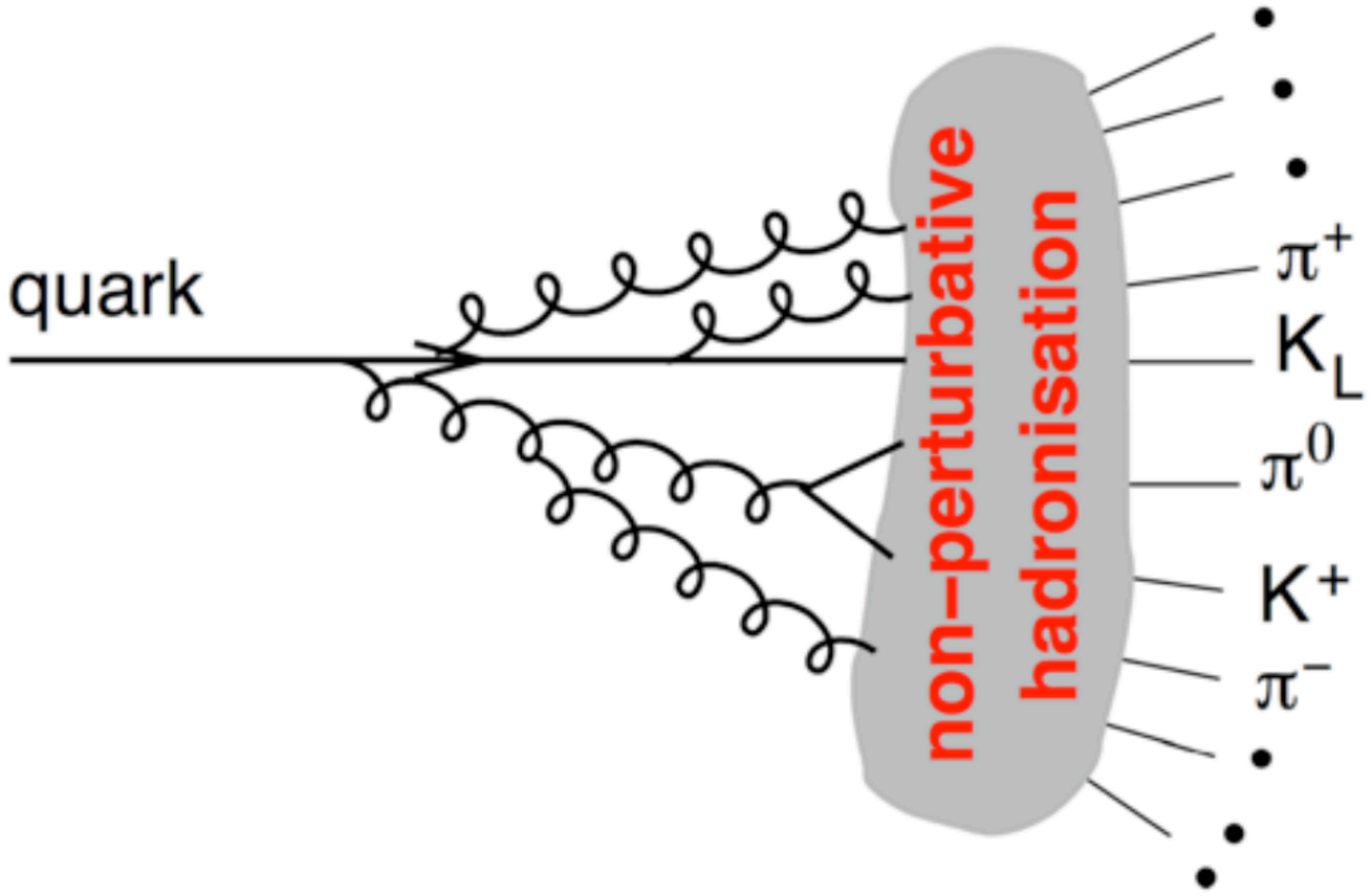
contain [ATLAS (1402.6263)] and [CMS (1310.1138)] 7 TeV data

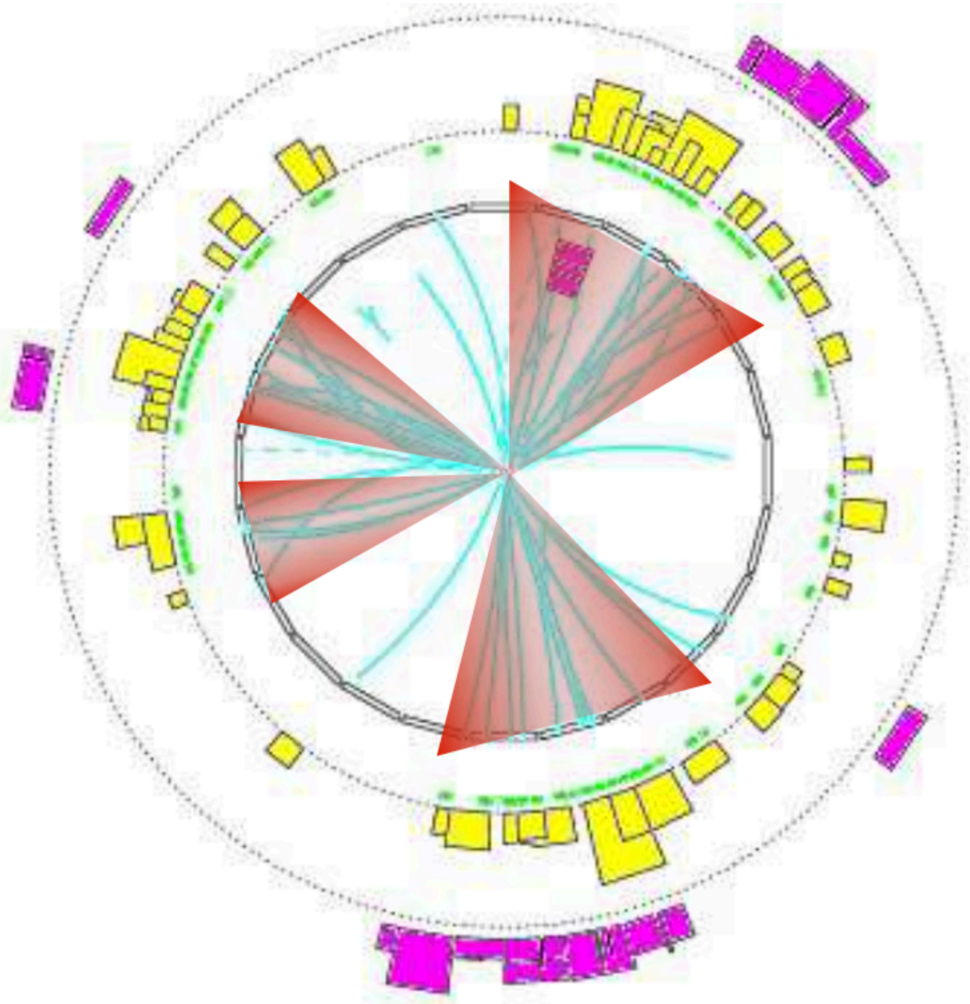
… but **flavoured jets appear everywhere**:
top physics, Higgs physics, new physics searches, …

What are jets?

Naive definition: **collimated bunch of hadrons** flying roughly in the same direction







2 clear jets

3 jets?
or 4 jets?

Proper definition: a collection of hadrons defined by means of a **jet algorithm**

"Jet [definitions] are legal contracts between theorists and experimentalists"
MJ Tannenbaum

IRC safety

An observable is **infrared and collinear safe** if, in the limit of a **collinear splitting**, or the **emission of an infinitely soft** particle, the observable remains **unchanged:**

$$O(X; p_1, \ldots, p_n, p_{n+1} \to 0) \to O(X; p_1, \ldots, p_n)$$
$$O(X; p_1, \ldots, p_n \parallel p_{n+1}) \to O(X; p_1, \ldots, p_n + p_{n+1})$$

This property ensures cancellation of **real** and **virtual** divergences in higher order calculations

If we wish to be able to calculate a jet rate in perturbative QCD the jet algorithm that we use must be IRC safe:
soft emissions and collinear splittings must not change the hard jets

slide stolen from Matteo Cacciari

Popular jet algorithms

e^+e^- colliders

$$d_{ij} = 2\min(E_i^2, E_j^2)\,(1 - \cos\theta_{ij})$$

k_t or Durham algorithm

pp colliders

$$d_{ij} = \min(p_{T,i}^{2p}, p_{T,j}^{2p})\frac{\Delta y^2 + \Delta\phi^2}{R^2}\,, \quad d_{iB} = p_{T,i}^{2p}$$

$p = 1$: k_t algorithm

$p = 0$: Cambridge/Aachen algorithm

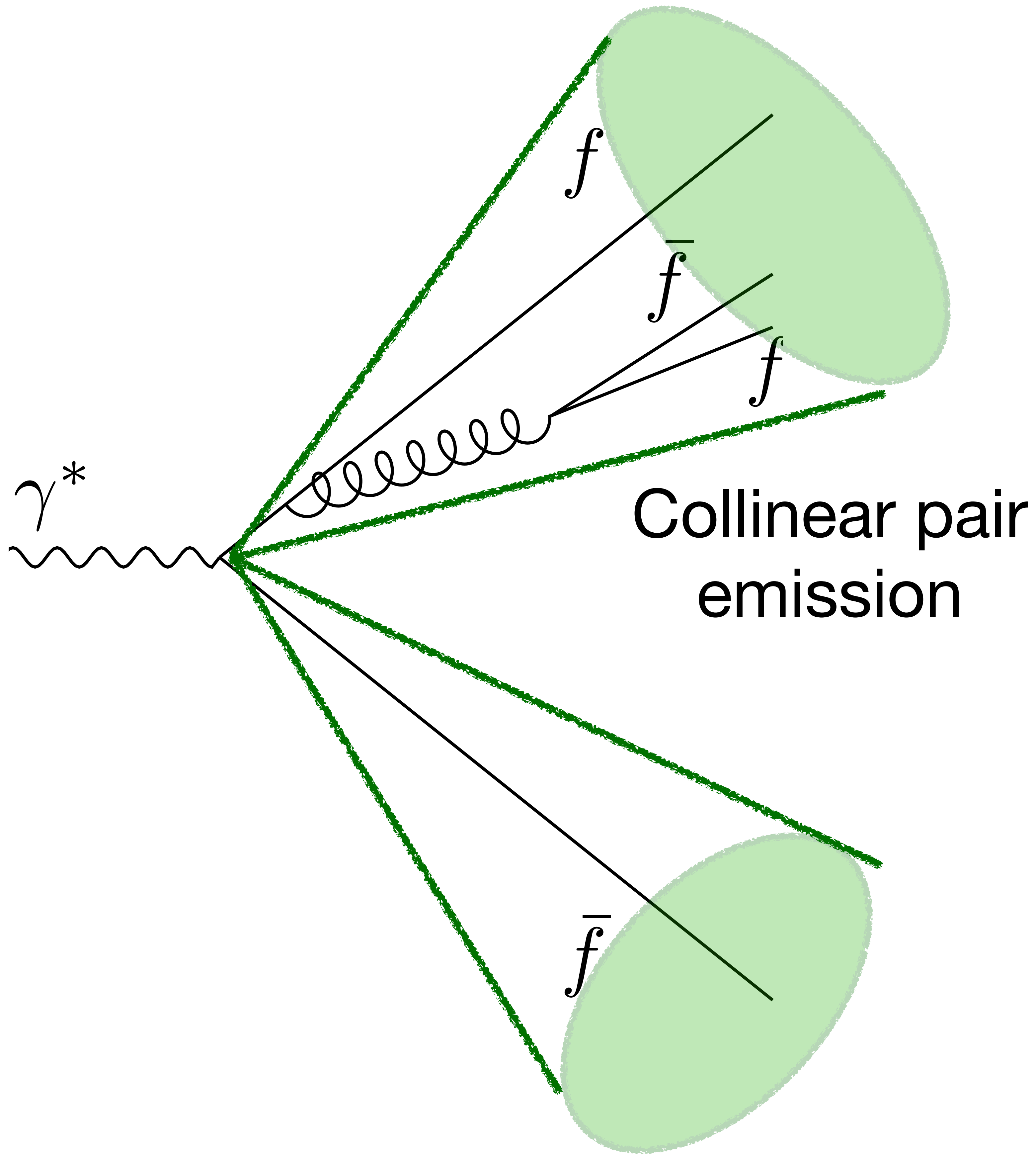
$p = -1$: anti-k_t algorithm

All IRC safe

IRC safe flavour definition of jets?

$e^+e^- \rightarrow$ jets with the k_t algorithm (flavour agnostic)
Issues at $\mathcal{O}(\alpha_s^2)$, when on top of $e^+e^- \rightarrow f\bar{f}$ we add a $f\bar{f}$ pair coming from gluon splitting.

When the $f\bar{f}$ pair from gluon is soft or collinear, the jet algorithm must return two flavoured jets.

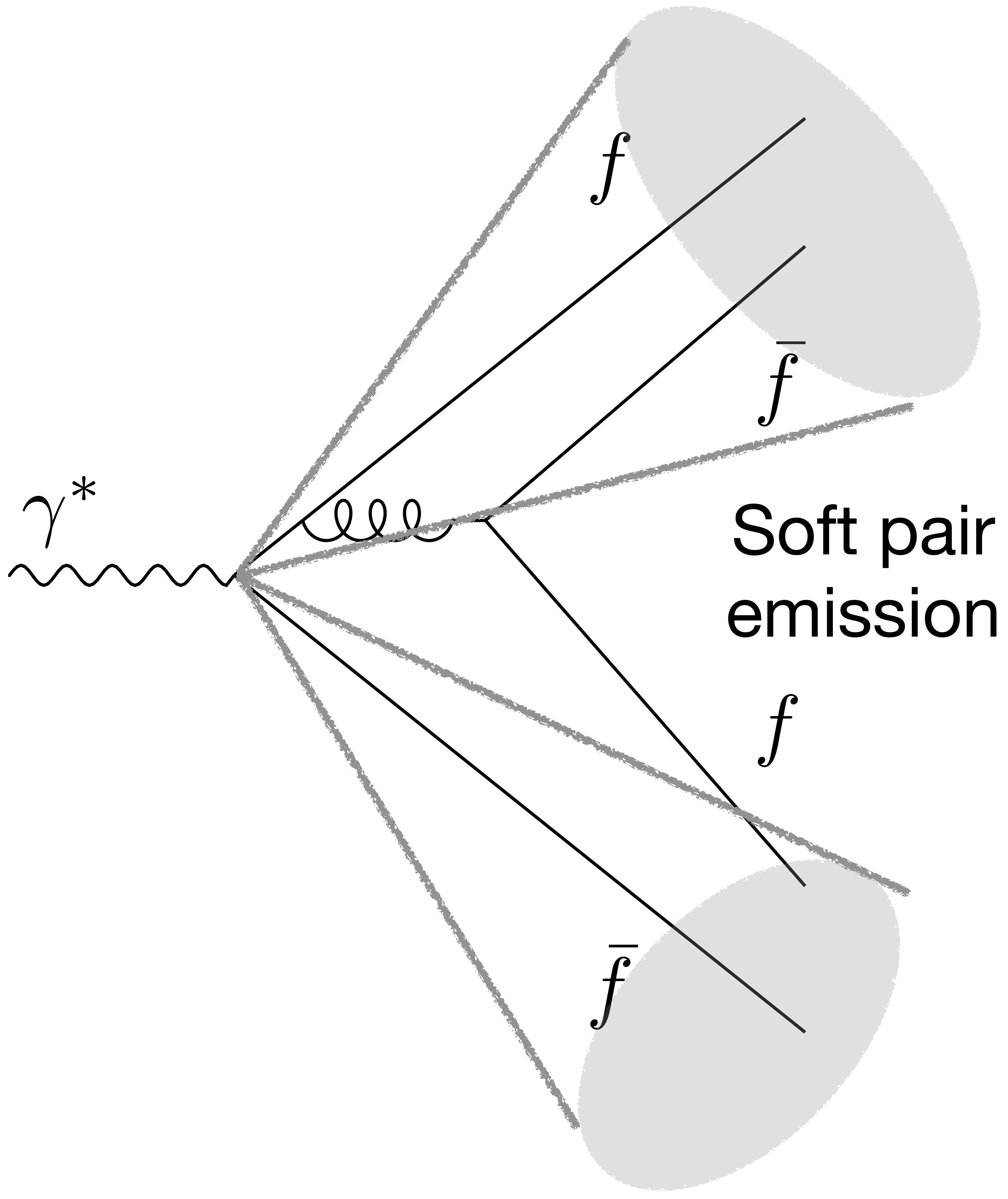


Collinear safe
(as any IRC safe flavour-agnostic algorithm)

IRC safe flavour definition of jets?

$e^+e^- \to$ jets with the k_t algorithm (flavour agnostic)
Issues at $\mathcal{O}(\alpha_s^2)$, when on top of $e^+e^- \to f\bar{f}$ we add a $f\bar{f}$ pair coming from gluon splitting.

When the $f\bar{f}$ pair from gluon is soft or collinear, the jet algorithm must return two flavoured jets.



Soft **unsafe**
due to polluting large-angle soft pair

The flavour-k_t algorithm

[Banfi, Salam, Zanderighi (hep-ph/0601139)]

1. Introduce a distance measure $d_{ij}^{(F)}$ between every pair of partons i, j:

$$d_{ij}^{(F,\alpha)} = (\Delta\eta_{ij}^2 + \Delta\phi_{ij}^2) \times \begin{cases} \max(k_{ti}, k_{tj})^\alpha \min(k_{ti}, k_{tj})^{2-\alpha}\,, & \text{softer of } i,j \text{ is flavoured,} \\ \min(k_{ti}^2, k_{tj}^2)\,, & \text{softer of } i,j \text{ is flavourless,} \end{cases} \tag{17}$$

as well as distances to the two beams,

$$d_{iB}^{(F,\alpha)} = \begin{cases} \max(k_{ti}, k_{tB}(\eta_i))^\alpha \min(k_{ti}, k_{tB}(\eta_i))^{2-\alpha}\,, & i \text{ is flavoured,} \\ \min(k_{ti}^2, k_{tB}^2(\eta_i))\,, & i \text{ is flavourless,} \end{cases} \tag{18}$$

and an analogous definition of $d_{i\bar{B}}^{(F,\alpha)}$ involving $k_{t\bar{B}}(\eta_i)$ instead of $k_{tB}(\eta_i)$ (both defined as in eqs. (15) and (16)).[9] As in section 2 we have introduced a class of measures, parametrised by $0 < \alpha \leq 2$.

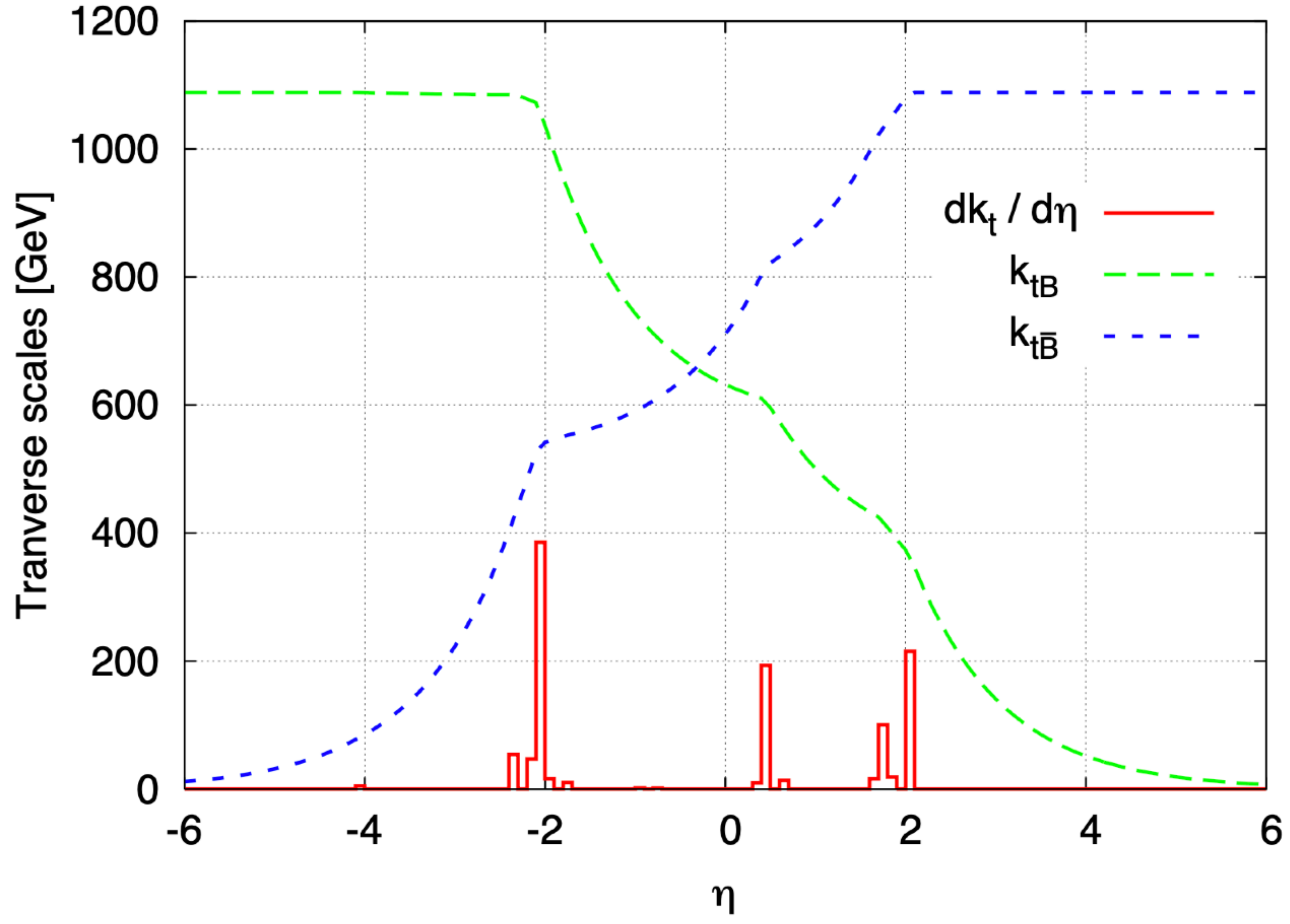
2. Identify the smallest of the distance measures. If it is a $d_{ij}^{(F,\alpha)}$, recombine i and j; if it is a $d_{iB}^{(F,\alpha)}$ ($d_{i\bar{B}}^{(F,\alpha)}$) declare i to be part of beam B ($\bar{B}$) and eliminate i; in the case where the $d_{iB}^{(F,\alpha)}$ and $d_{i\bar{B}}^{(F,\alpha)}$ are equal (which will occur if i is a gluon), recombine with the beam that has the smaller $k_{tB}(\eta_i)$, $k_{t\bar{B}}(\eta_i)$.

3. Repeat the procedure until all the distances are larger than some d_{cut}, or, alternatively, until one reaches a predetermined number of jets.[10,11]

Modified beam distance:

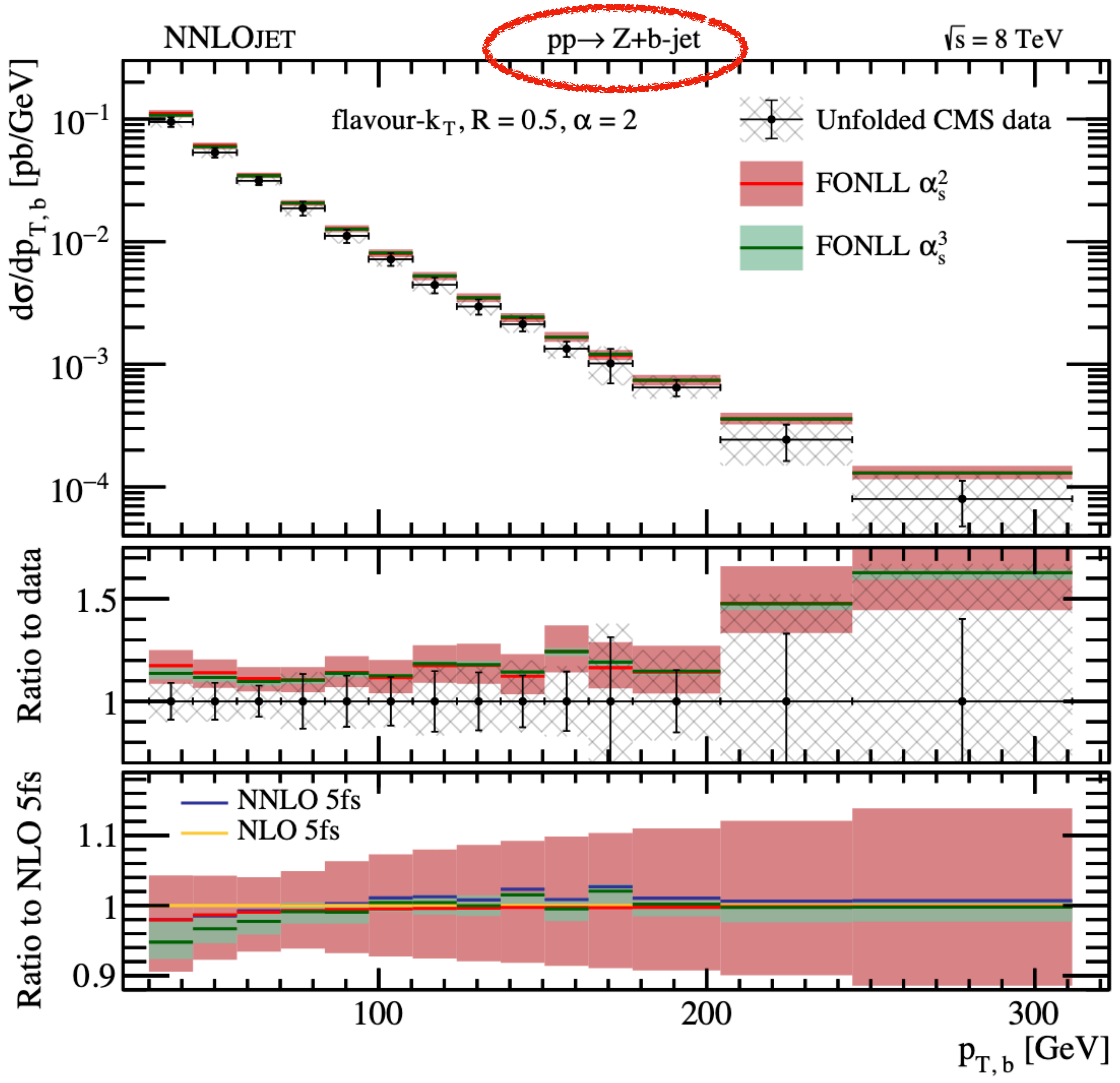
$$k_{tB}(\eta) = \sum_i k_{ti}\left(\Theta(\eta_i - \eta) + \Theta(\eta - \eta_i)e^{\eta_i - \eta}\right)$$

$$k_{t\bar{B}}(\eta) = \sum_i k_{ti}\left(\Theta(\eta - \eta_i) + \Theta(\eta_i - \eta)e^{\eta - \eta_i}\right)$$

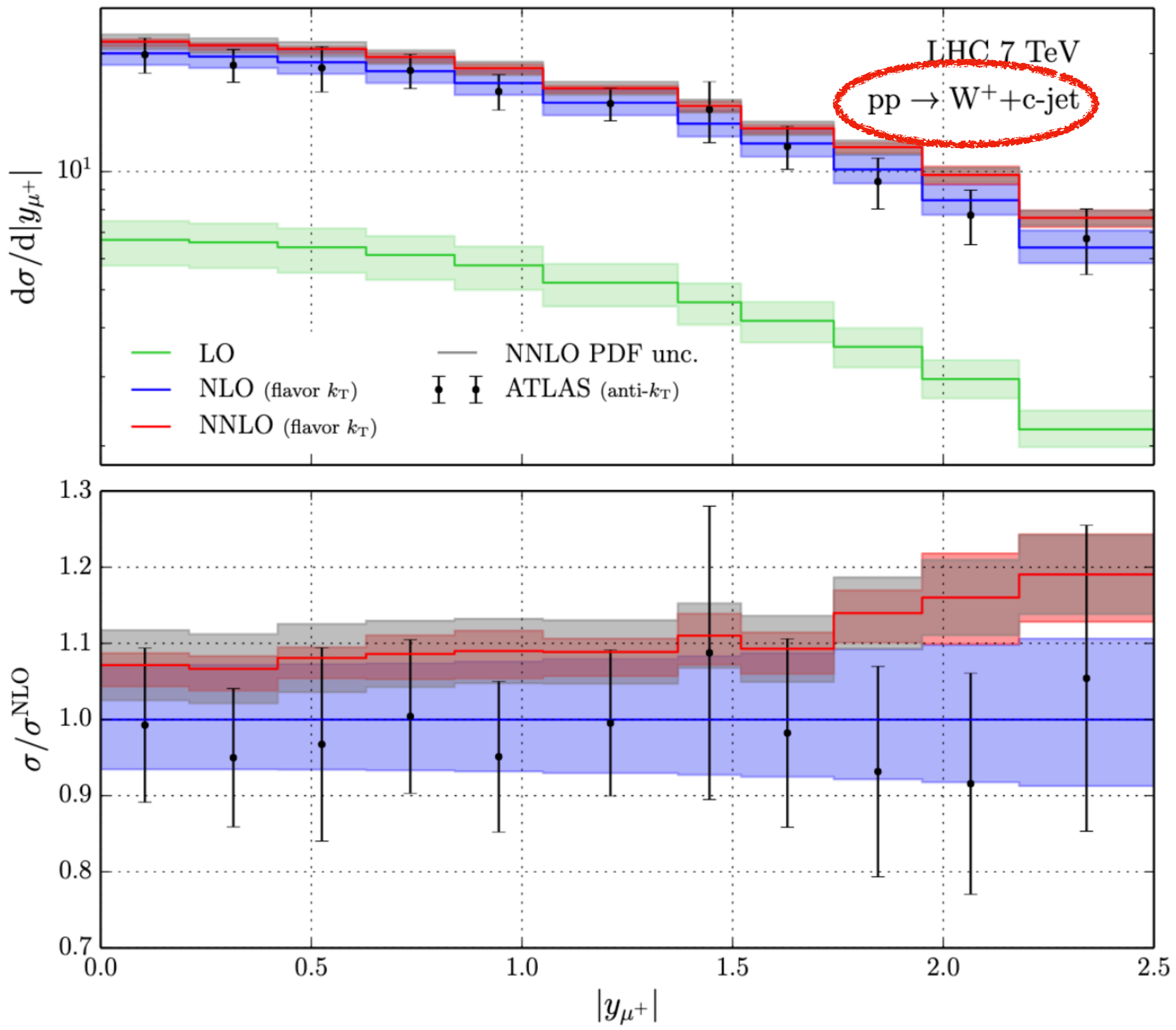


IRC flavour safe to all orders,
but different kinematics
(because new distance)

NNLO predictions with flavour-k_t



[Gauld et al. (2005.03016)]



[Czakon et al. (2011.01011)]

Comparison with experimental data **not straightforward**

Recent proposals

- based on Soft Drop grooming techniques
  [Caletti, Larkoski, Marzani, Reichelt (2205.01109)]
- through the alignment of flavoured particles along the Winner-Take-All axis
  [Caletti, Larkoski, Marzani, Reichelt (2205.01117)]
- through a modification of anti-k_t clustering distance
  [Czakon, Mitov, Poncelet (2205.11879)]
- with successive iterations of flavour-k_t and anti-k_t
  [Caletti, Fedkevych, Marzani, Reichelt (2108.10024)]
- using jet angularities and primary Lund jet plane as discriminants
  [Fedkevych, Khosa, Marzani, Sforza (2202.05082)]

However, none of the above reproduces the same jets as anti-k_t,
can be applied to a generic process with one or more jets
and it is IRC safe to all orders.

The flavour dressing algorithm

[Gauld, Huss, GS (2208.11138)]

Flavour assignment *factorised* from jet reconstruction:
we assign flavour to flavour-agnostic jets in an IRC safe way

Inputs:

- *Flavour agnostic jets* $\{j_k\}$
- *Flavoured clusters* $\{\hat{f}_i\}$
- *Association criterion*
- *Accumulation criterion*

The flavour dressing algorithm: inputs

- *Flavour agnostic jets* $\{j_k\}$*:*
  set of jets obtained with an IRC safe jet algorithm (e.g. gen-k_t family), possibly after a fiducial selection.

- *Flavoured clusters* $\{\hat{f}_i\}$

- *Association criterion*

- *Accumulation criterion*

The flavour dressing algorithm: inputs

- *Flavour agnostic jets* $\{j_k\}$

- *Flavoured clusters* $\{\hat{f}_i\}$*:*
  built out of quarks (e.g. c, b) or stable heavy-flavour hadrons (e.g. D, B), by dressing them with radiation close in angle (see below)
  “Naked” flavoured objects are collinear unsafe

- *Association criterion*

- *Accumulation criterion*

The flavour dressing algorithm: inputs

- *Flavour agnostic jets* $\{j_k\}$
- *Flavoured clusters* $\{\hat{f}_i\}$
- *Association criterion:* whether f_i is "associated" to j_k
  At parton-level simply if f_i is a constituent of j_k
  Other options: $\Delta R(f_i, j_k) < R_{\text{tag}}$, ghost association, …
  Flavour assignment based only on the association is not IRC safe
- *Accumulation criterion*

The flavour dressing algorithm: inputs

- *Flavour agnostic jets* $\{j_k\}$
- *Flavoured clusters* $\{\hat{f}_i\}$
- *Association criterion*
- *Accumulation criterion:* how to “sum” flavours
  - sum flavoured if unequal number of f and $\bar{\mathrm{f}}$ (need charge information)
  - sum flavoured if odd number of f or $\bar{\mathrm{f}}$ (if no charge information)

Definition of flavoured cluster $\hat{f}_i$

1. Initialise a set with all the flavourless objects p_i (particles used as input to jets) and all the flavoured objects f_i (bare flavours), avoiding double counting if necessary.

2. Find the pair with the smallest angular distance ΔR_{ab}:
   — flavourless p_a, p_b: combine p_a and p_b into a flavourless p_{ab};
   — flavoured f_a, f_b: remove both from the set;
   — flavoured f_a, unflavoured p_b: remove p_b from the set and check a Soft Drop criterion

$$\frac{\min(p_{t,a}, p_{t,b})}{(p_{t,a} + p_{t,b})} > z_{\text{cut}} \left(\frac{\Delta R_{ab}}{\delta R} \right)^{\beta}$$

   to **recombine collinear while preserving soft.** [default: $\delta R = 0.1$, $z_{\text{cut}} = 0.1$, $\beta = 2$]
   If satisfied, combine f_a and p_b into a flavoured f_{ab}.

3. Iterate while there are at least two objects in the set until $\Delta R_{ab} > \delta R$.
   The momentum of $\hat{f}_i$ is given by the accumulated momentum into f_i.

The flavour dressing algorithm

1. Define tag_k = flavoured clusters assigned to jet j_k (initialised as empty for all jets) and populate set of distances:
   - $d(\hat{f}_i, \hat{f}_j)$ between flavoured clusters;
   - $d(\hat{f}_i, \hat{j}_k)$ if flavoured cluster $\hat{f}_i$ *associated* to jet j_k
   - $d_B(\hat{f}_i)$ if $\hat{f}_i$ not associated to any jet.

   Distances (including beam) inherited from the flavour-k_t algorithm:

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

$$d_{B\pm}(f) = \max\left(p_{t,f}^{\alpha}, p_{t,B_\pm}^{\alpha}(y_f)\right)\ \min\left(p_{t,f}^{2-\alpha}, p_{t,B_\pm}^{2-\alpha}(y_f)\right)$$

The flavour dressing algorithm

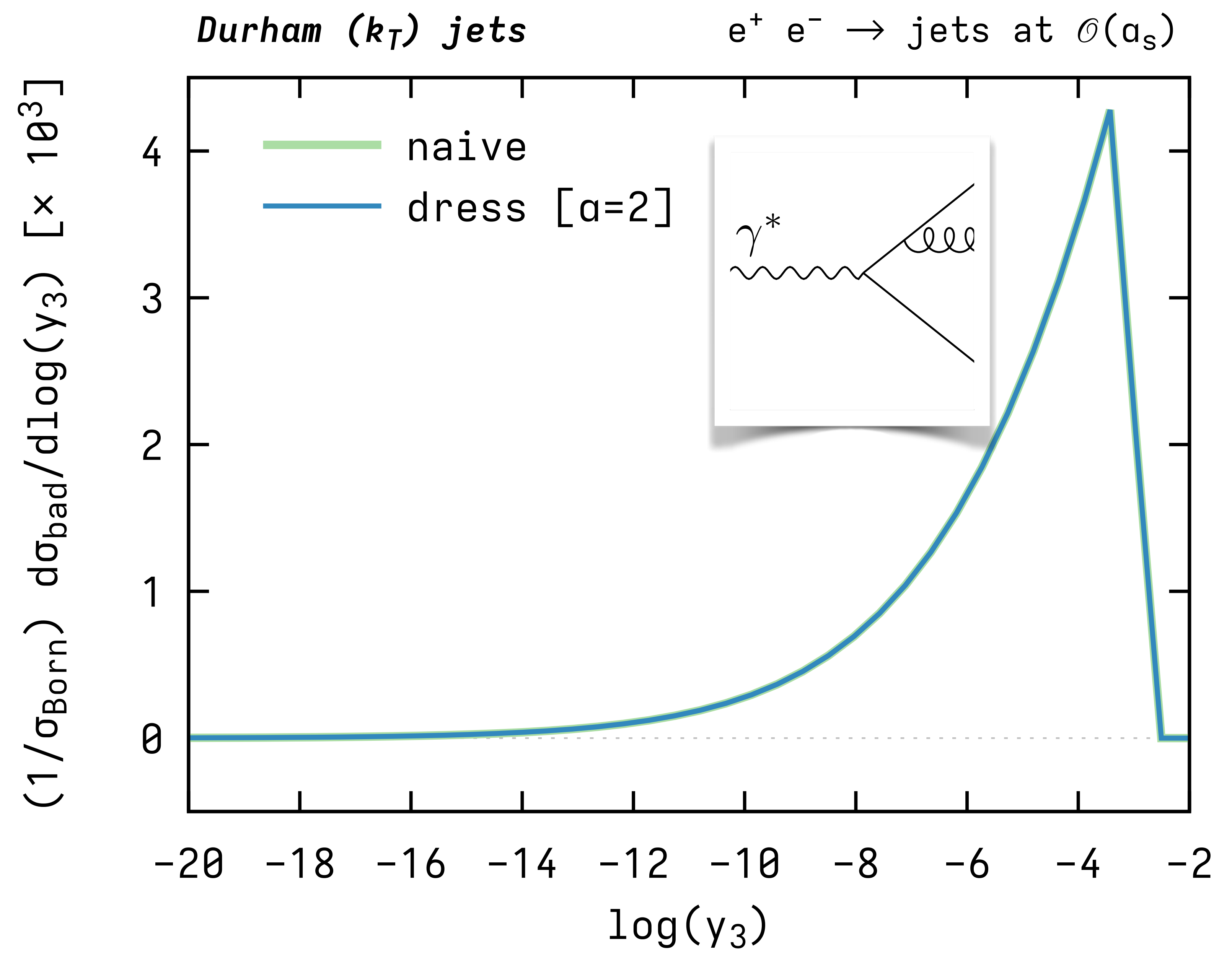
2. While the set of distances is not empty, select the smallest distance:
   $\rightarrow d(\hat{f}_i, \hat{f}_j)$:
   the two flavours "annihilate", hence remove distances that involve $\hat{f}_i$ or $\hat{f}_j$;
   $\rightarrow d(\hat{f}_i, \hat{j}_k)$:
   update $\mathrm{tag}_k = \mathrm{tag}_k \cup \{f_i\}$, then remove distances that involve $\hat{f}_i$.
   $\rightarrow d_B(\hat{f}_i)$:
   discard flavour $\hat{f}_i$ and remove all entries that involve $\hat{f}_i$.

3. Assign flavour to jet j_k according to tag_k and *accumulation* criterion.

IRC safety test in $e^+e^- \to$ jets

Vanishing mis-identification of flavours in the fully unresolved regime = IRC safety

only soft and/or collinear radiation

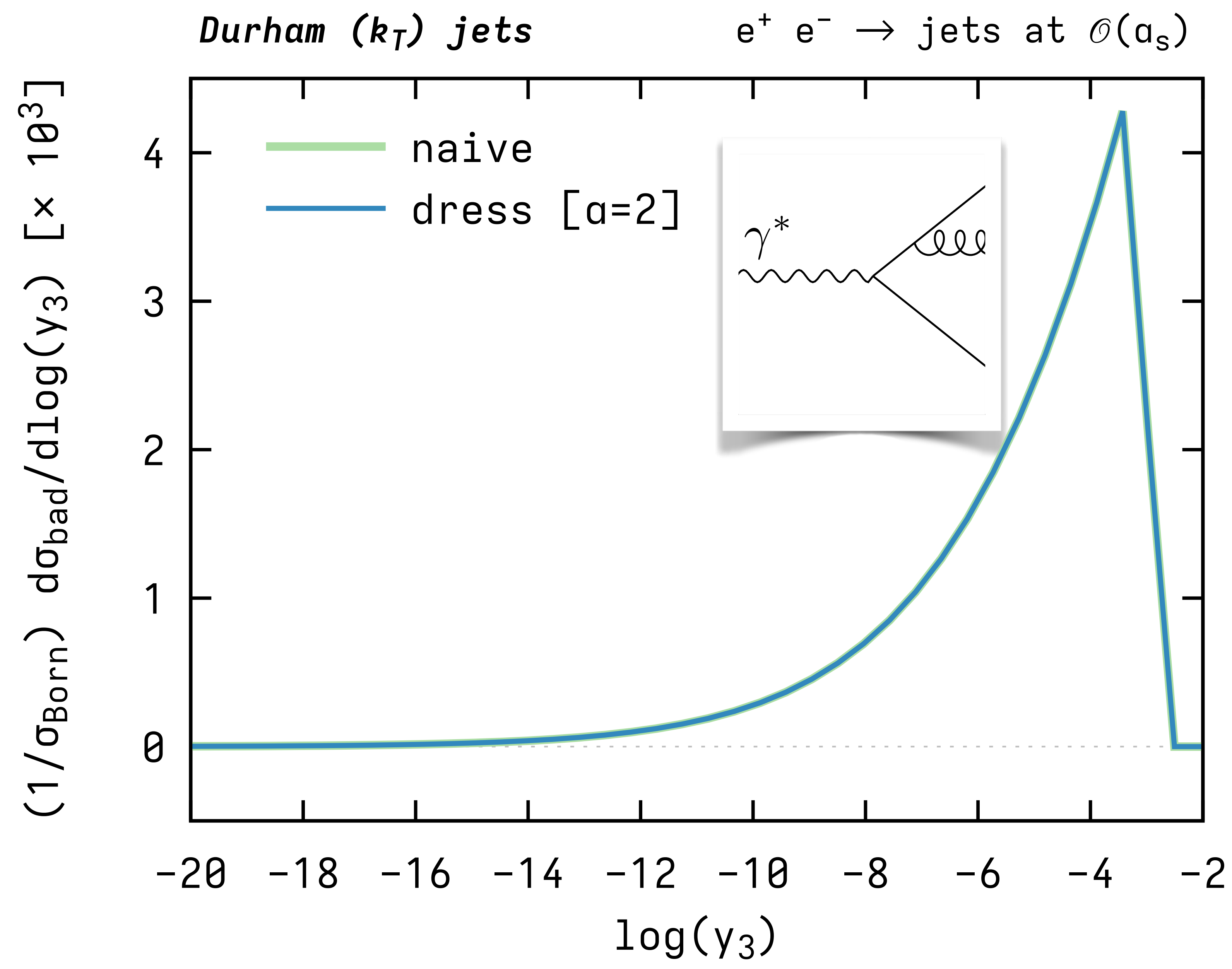


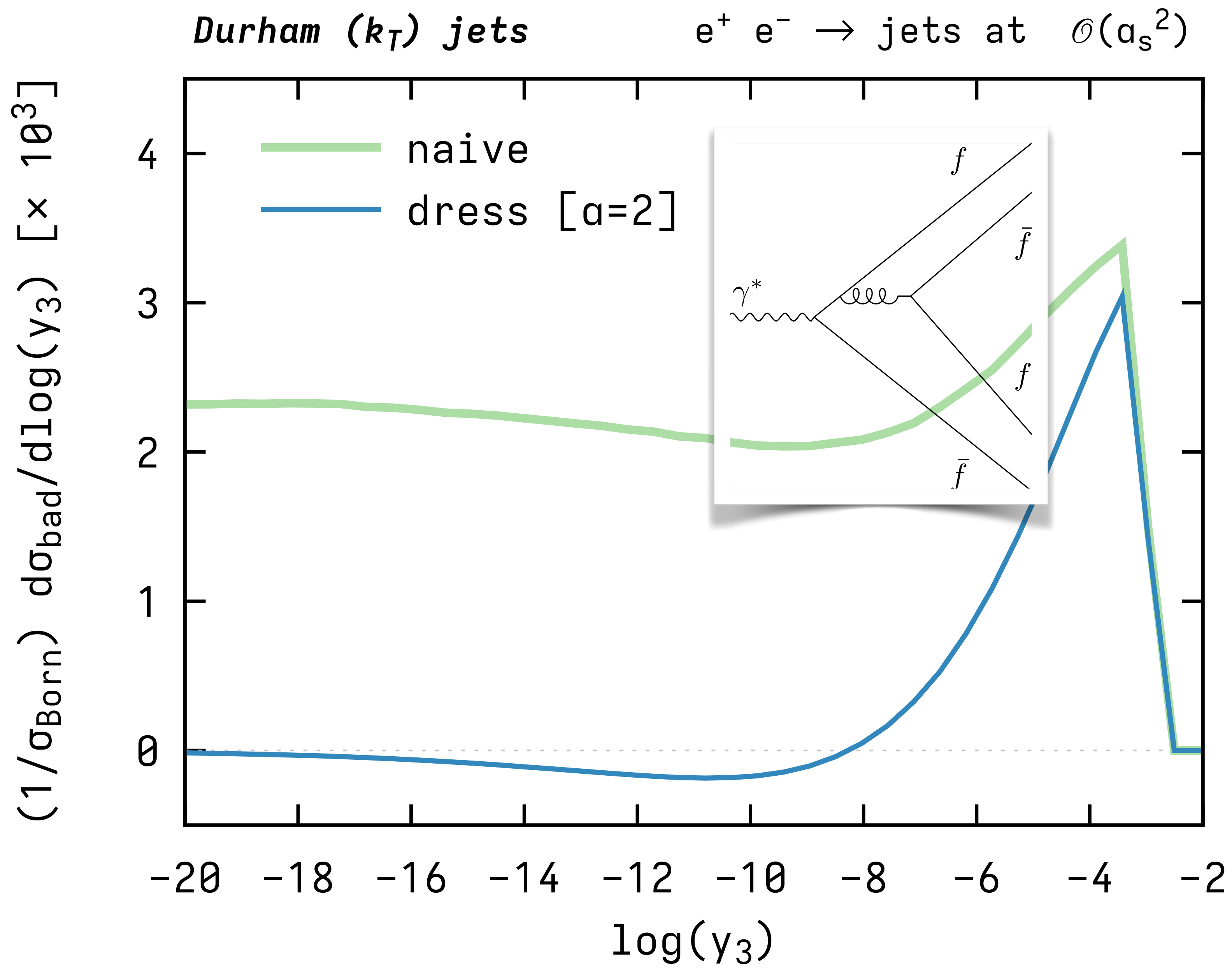
Any gen-k_t algo is safe!

IRC safety test in $e^+e^- \to$ jets

Vanishing mis-identification of flavours in the fully unresolved regime = IRC safety

only soft and/or collinear radiation





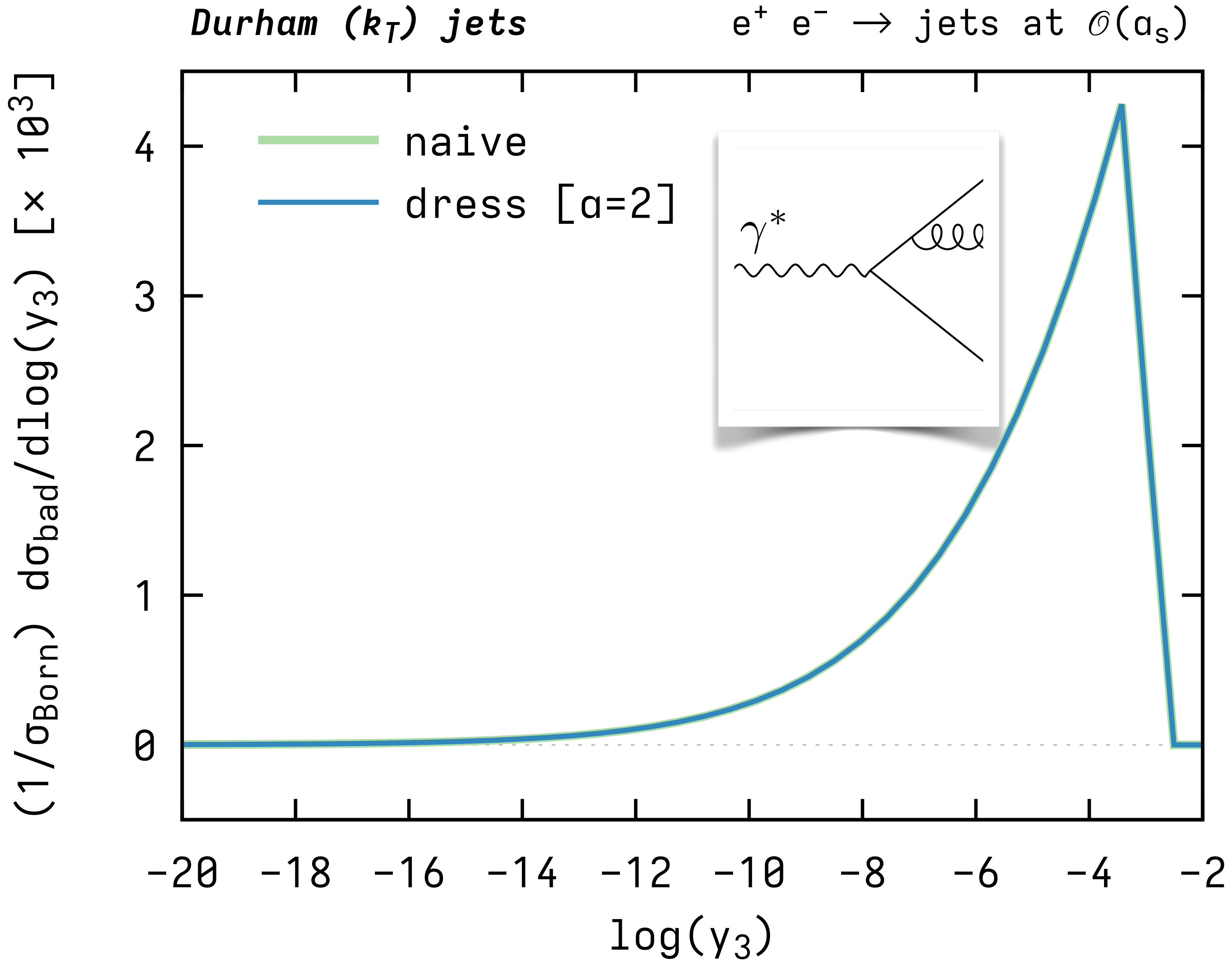
Any gen-k_t algo is safe!

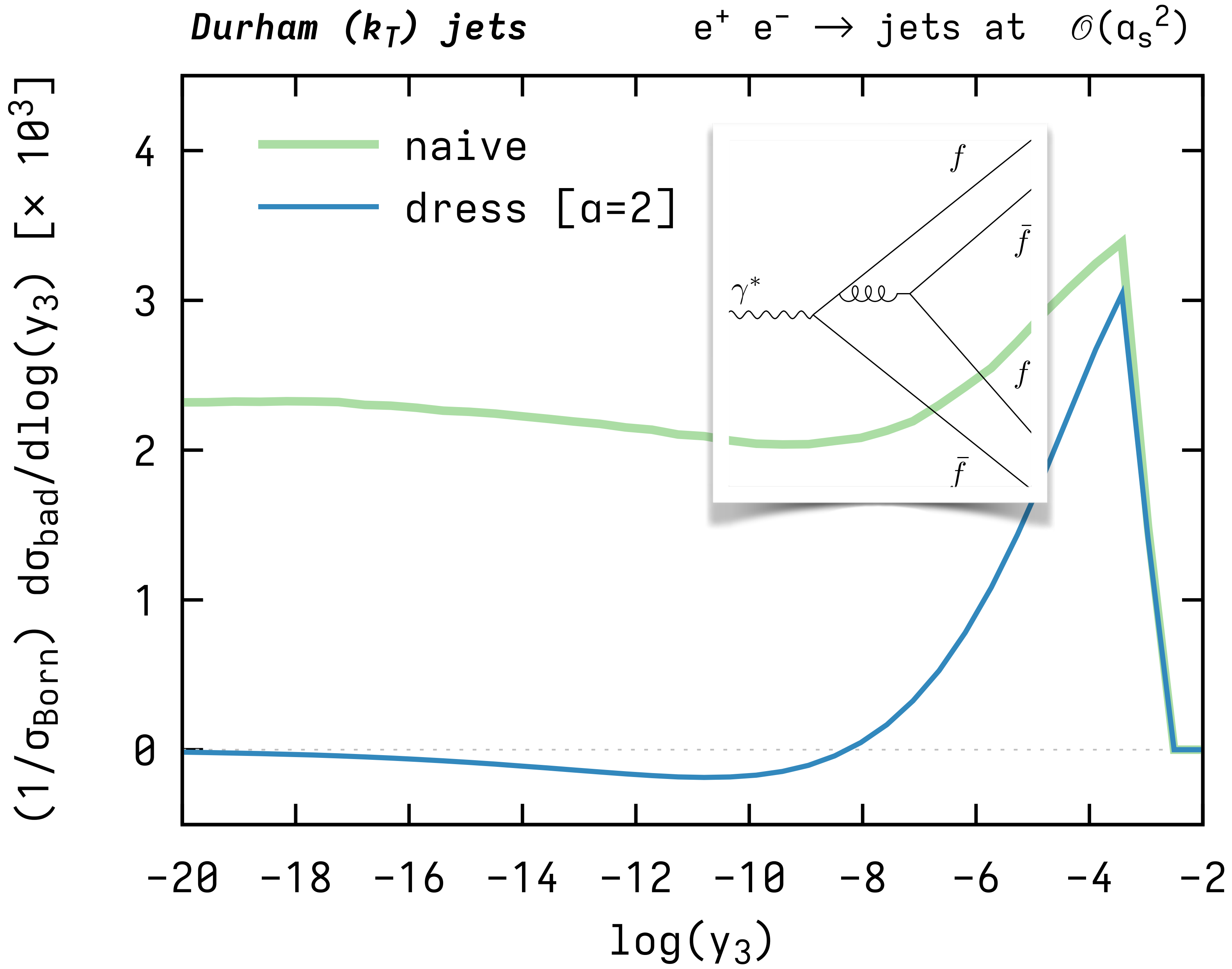
Naive dressing unsafe,
flavour dressing safe!

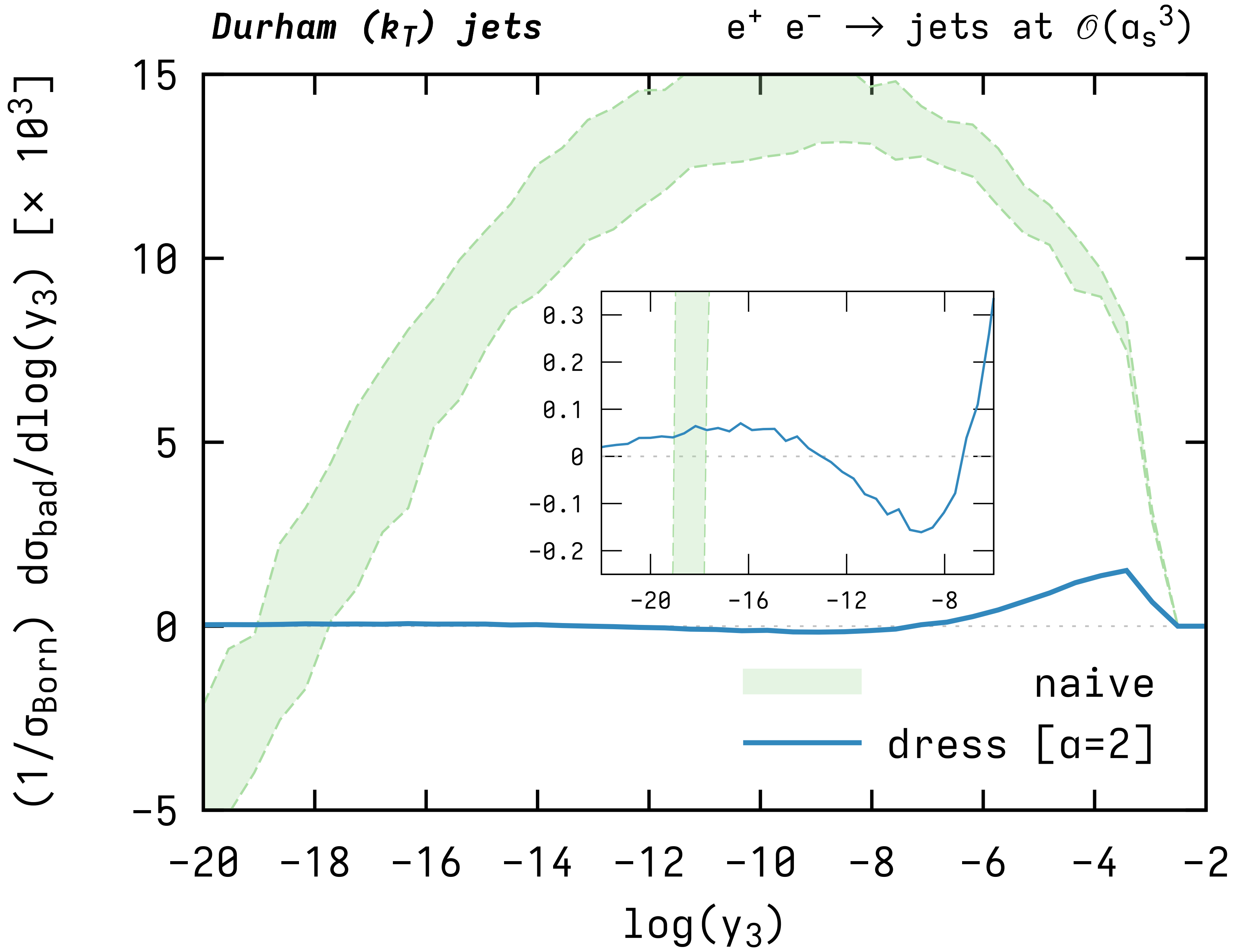
IRC safety test in $e^+e^- \to$ jets

Vanishing mis-identification of flavours in the fully unresolved regime = IRC safety

only soft and/or collinear radiation





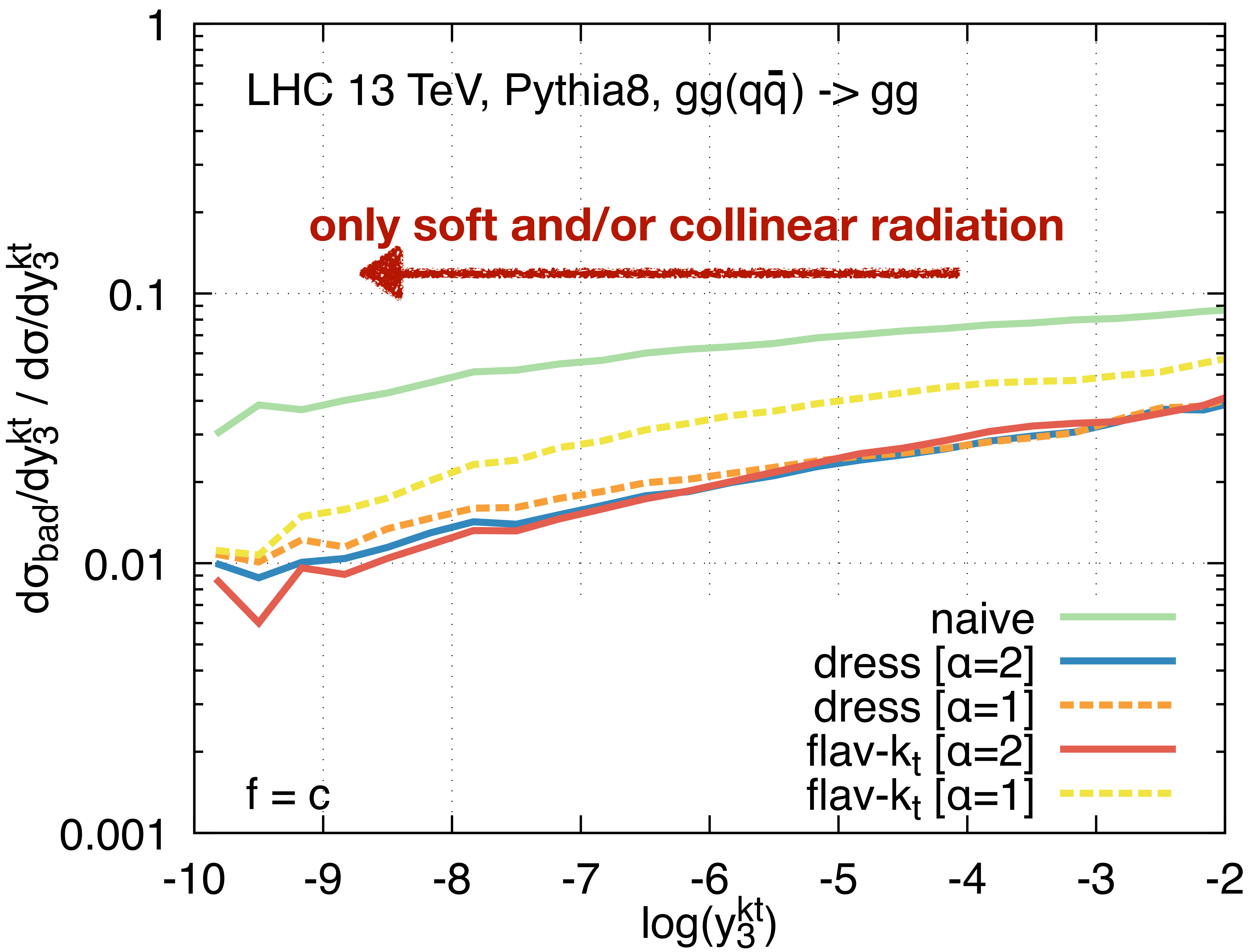


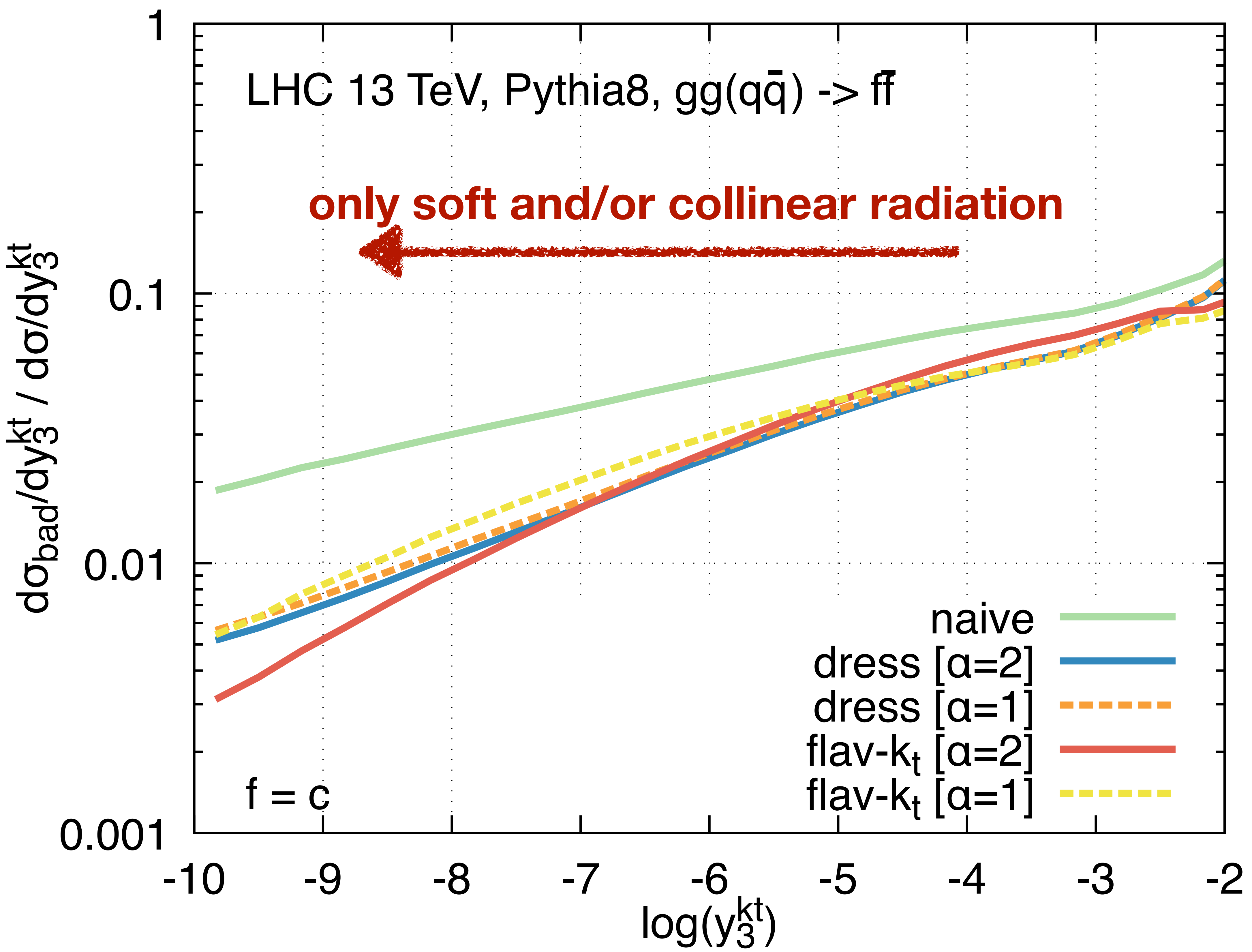
Any gen-k_t algo is safe!

Naive dressing unsafe, flavour dressing safe!

Naive dressing unsafer, flavour dressing still safe!

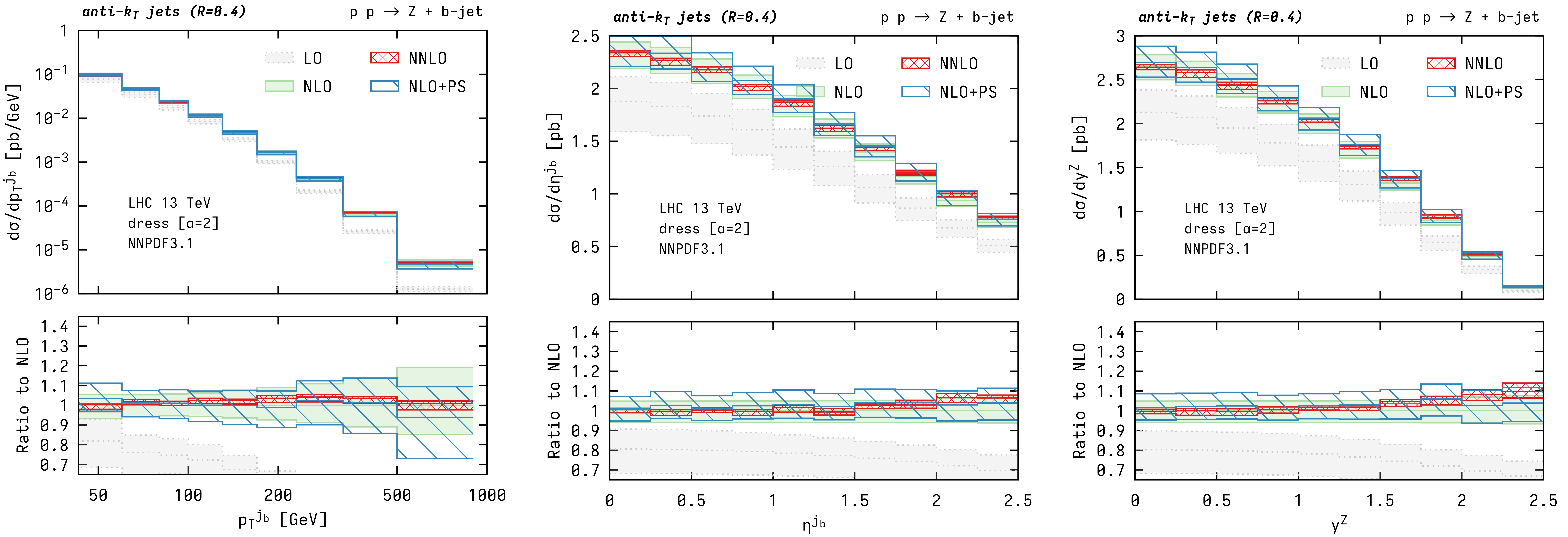
IRC sensitivity in $2 \to 2$ QCD events in pp





Flavour dressing approaches zero faster than a naive flavour tagging as $y_3^{k_t} \to 0$

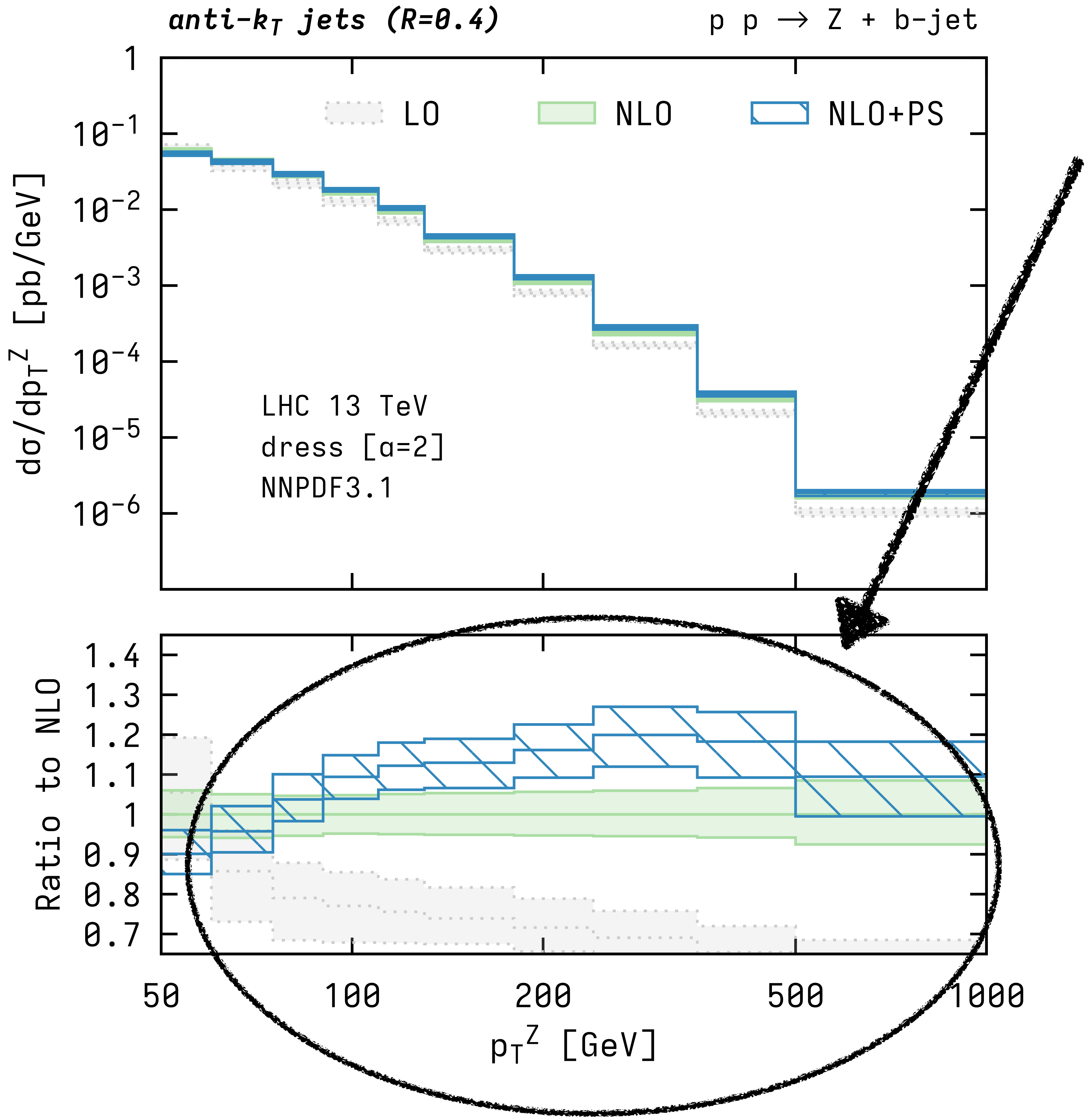
Test in a realist scenario: $Z + b$-jet



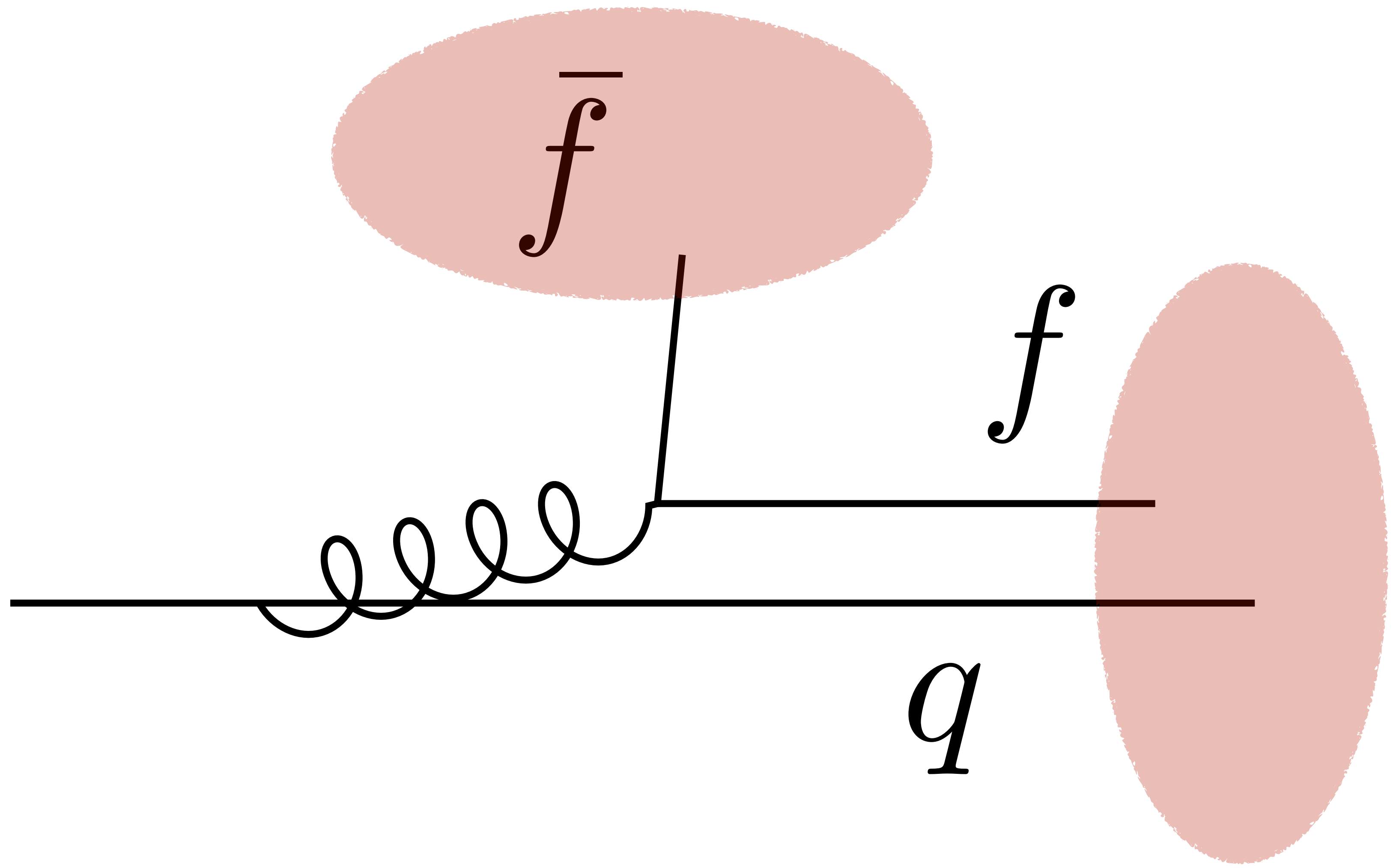
Remarkable agreement between (N)NLO and NLO+PS

→ for most distributions **largely insensitive to all-order corrections**

Test in a realist scenario: $Z + b$-jet

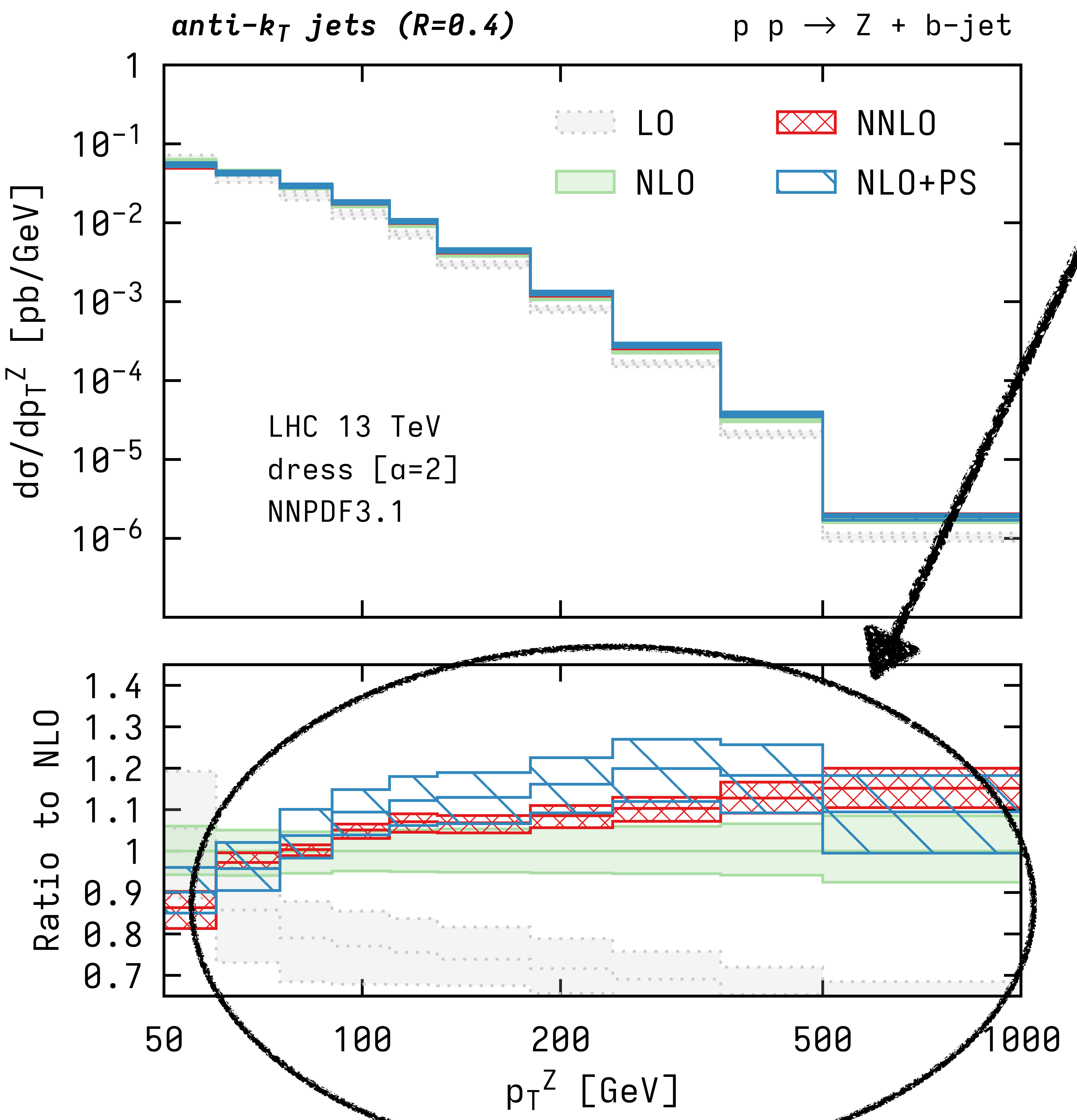


Some sensitivity observed in p_T^Z, likely due to:

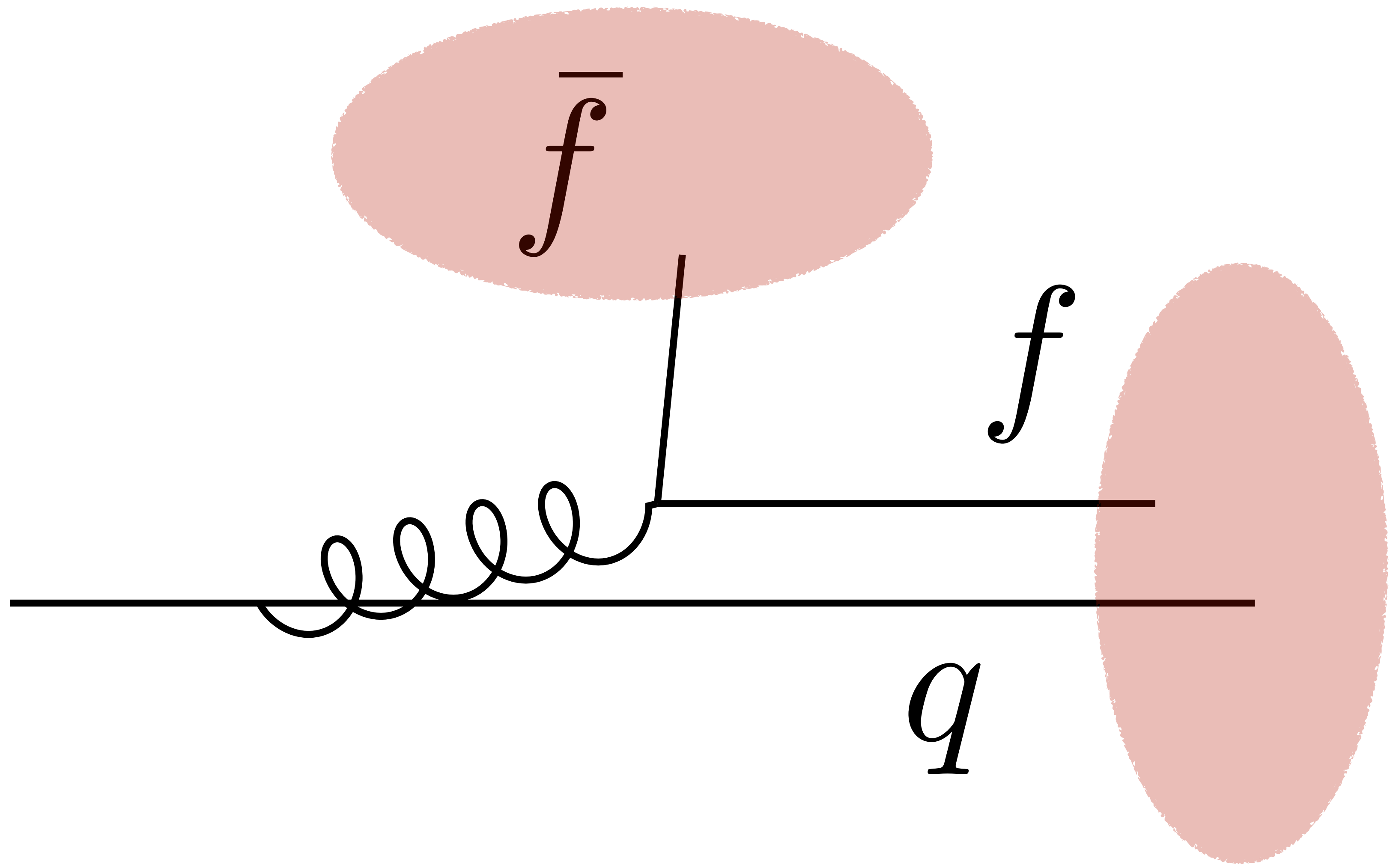


Even if IRC finite, it leads to large migration of (unflavoured)-jet into the b-jet sample.

Test in a realist scenario: $Z + b$-jet



Some sensitivity observed in p_T^Z, likely due to:



Effect **captured at NNLO**

Conclusions

With favour dressing, **flavour assignment *factorised* from the initial jet reconstruction**, hence it can be **combined with any IRC safe definition of a jet**

Thanks to an IRC safe flavour assignment to all orders in perturbation theory, we can compute massless fixed-order predictions, and in the case of massive calculations, we have a suppressed sensitivity on mass logarithms $\log(Q^2/m_f)$

Interesting to explore: experimental feasibility of the algorithm, how flavour dressing behaves for other processes and observables, and how it compares to the other approaches recently proposed.

BACKUP

Flavour anti-k_t algorithm

[Czakon, Mitov, Poncelet (2205.11879)]

Recent proposal: modify anti-k_t distance when flavoured particles involved

$$d_{ij} = R^2 \min(k_{T,i}^{-2}, k_{T,j}^{-2}) \cdot S_{ij}^{a}\,, \quad d_B = k_{T,i}^{-2}$$

where $S_{ij} \neq 1$ only when i and j are of opposite flavour

$$S_{ij}^{a} = 1 - \theta(1-\kappa)\cos\left(\frac{\pi}{2}\kappa\right)\,, \quad \kappa = \frac{1}{a}\frac{k_{T,i}^2 + k_{T,j}^2}{2\,k_{T,\max}^2}$$

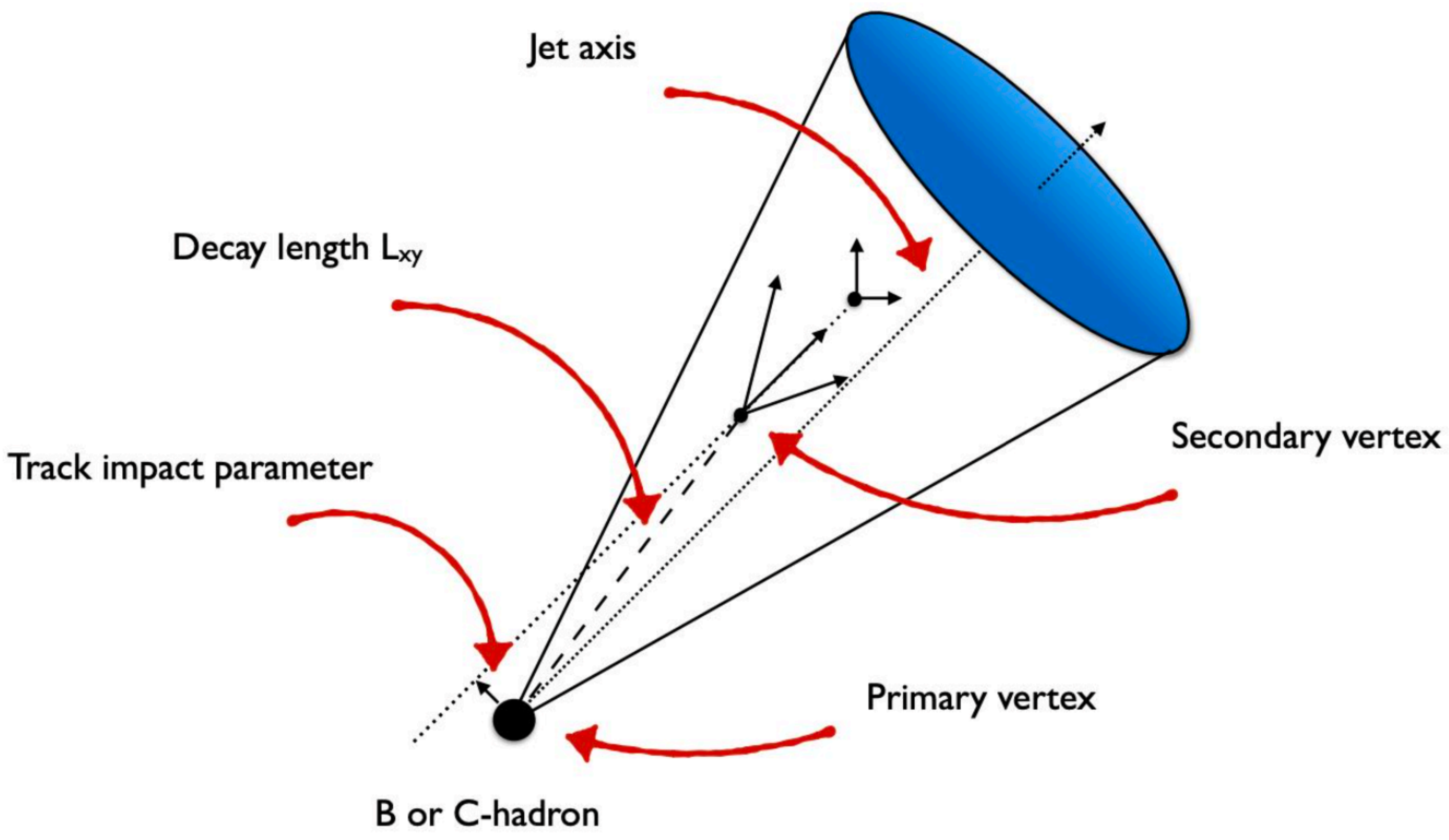
One recovers (IRC flavour unsafe) anti-k_t jets when $a \to 0$.

Quite significant dependence of the result on the parameter a.

Flavoured jets (experiments)

A jet defined as flavoured if it contains at least one heavy hadron within $\Delta R < R$ from the jet axis and with $p_T > p_{T,\mathrm{cut}}$ (*naive* tagging)

This is the "truth" labelling used in Monte Carlo samples, used to train a ML architecture ("High-level tagger") which adopts low-level variables as inputs



IRC flavour safety

The experimental definition is both **collinear and soft unsafe**

"A jet defined as flavoured if it contains
at least one heavy hadron
within $\Delta R < R$ from the jet axis
and with $p_T > p_{T,\mathrm{cut}}$"

Soft large angle $g \to b\bar{b}$
polluting different jets

$g \to b\bar{b}$ is always flavoured
even in the collinear limit
(an "even tag" removal is
enough to fix this)

$b \to bg$ collinear with the gluon
carrying most of the momentum
(would an identified particle, hence FF)