Next Generation Perturbative QCD for Hadron Structure: Preparing for the Electron-Ion Collider

Jets and their substructure at electron-proton colliders

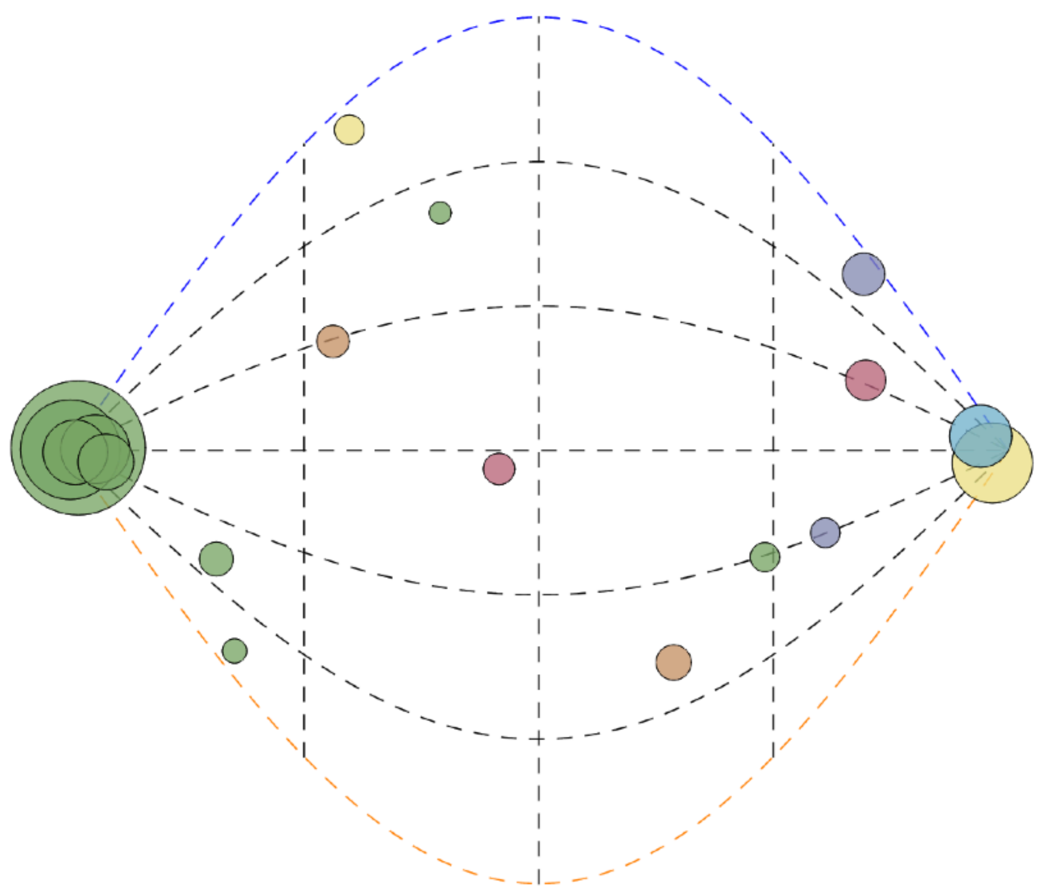
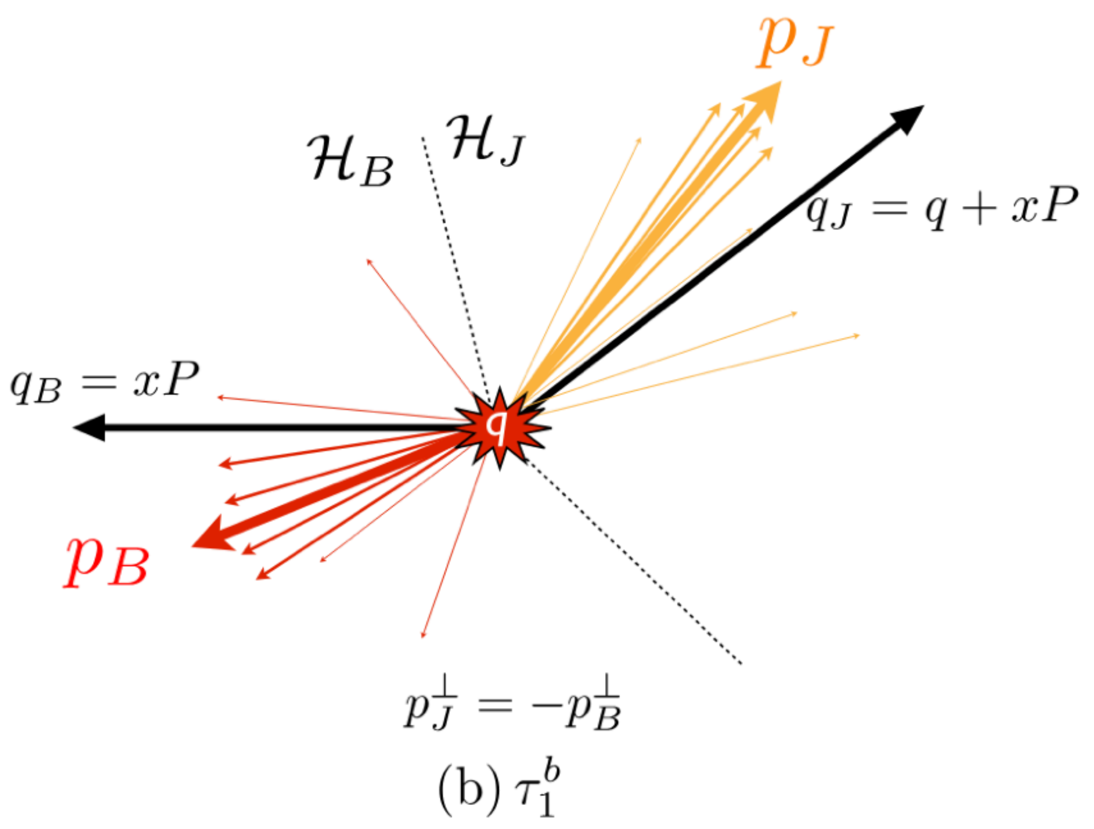
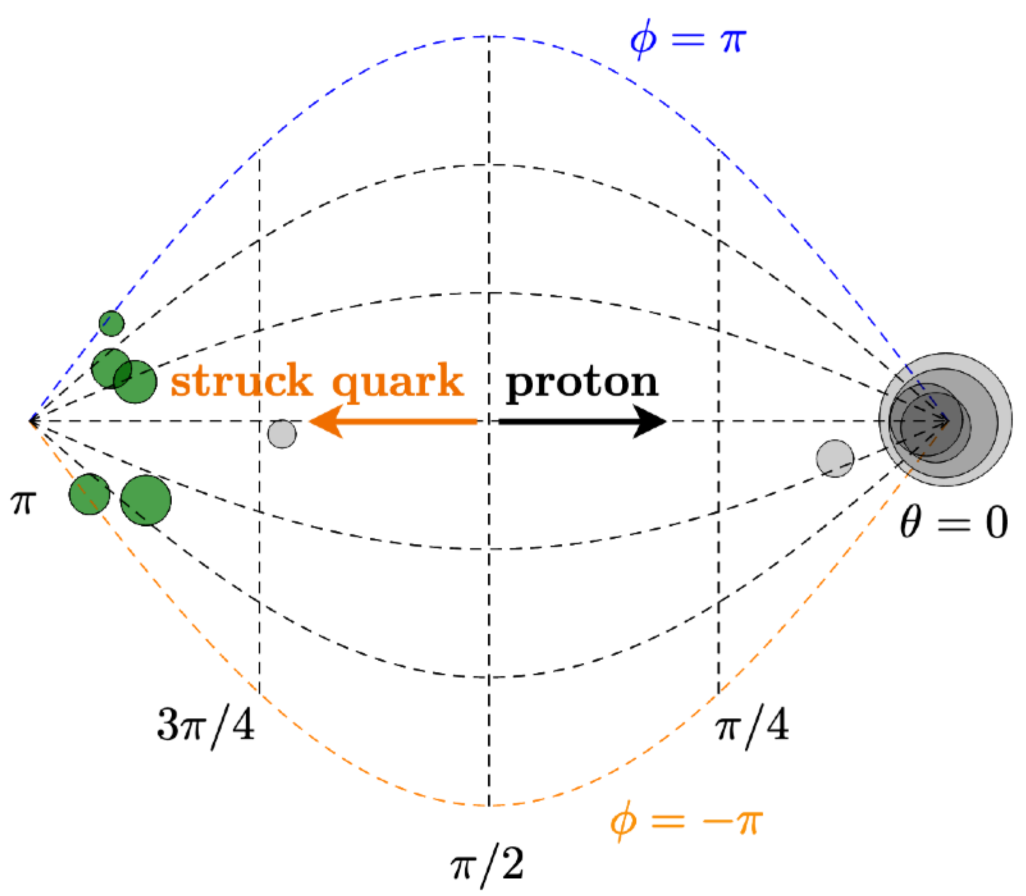
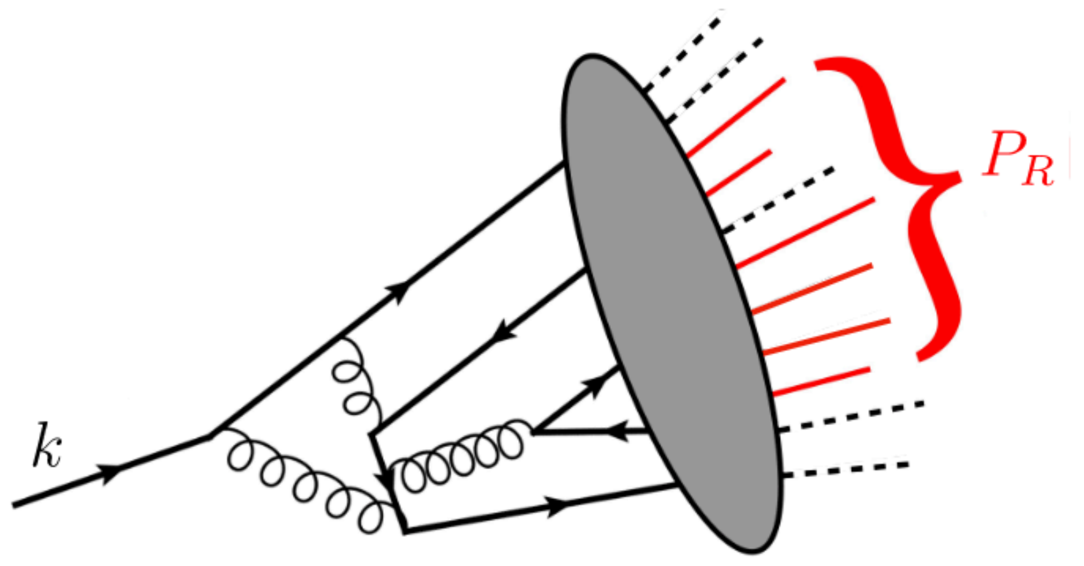
Giovanni Stagnitto

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

INFN

MITP (Mainz), 21.10.2024

Outlook

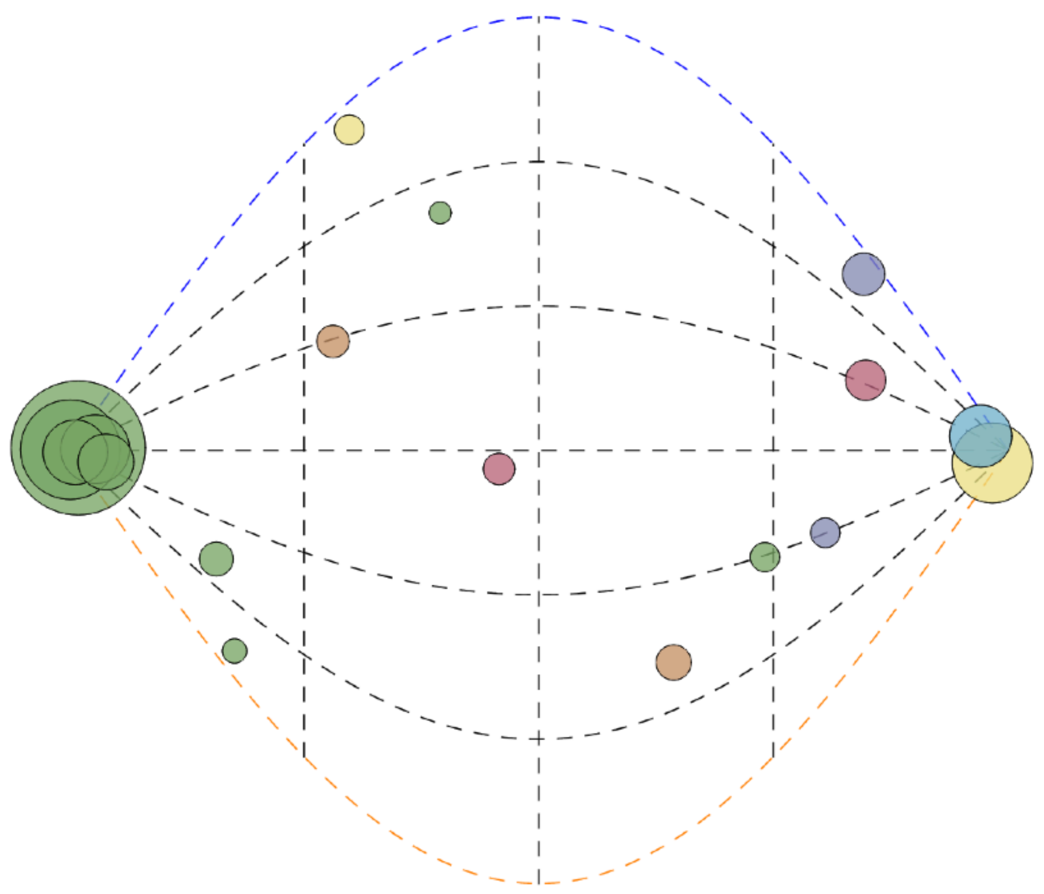
Jet algorithms	Event shapes	Grooming	Track functions
			

How to exploit the expertise gained from the LHC in DIS context?

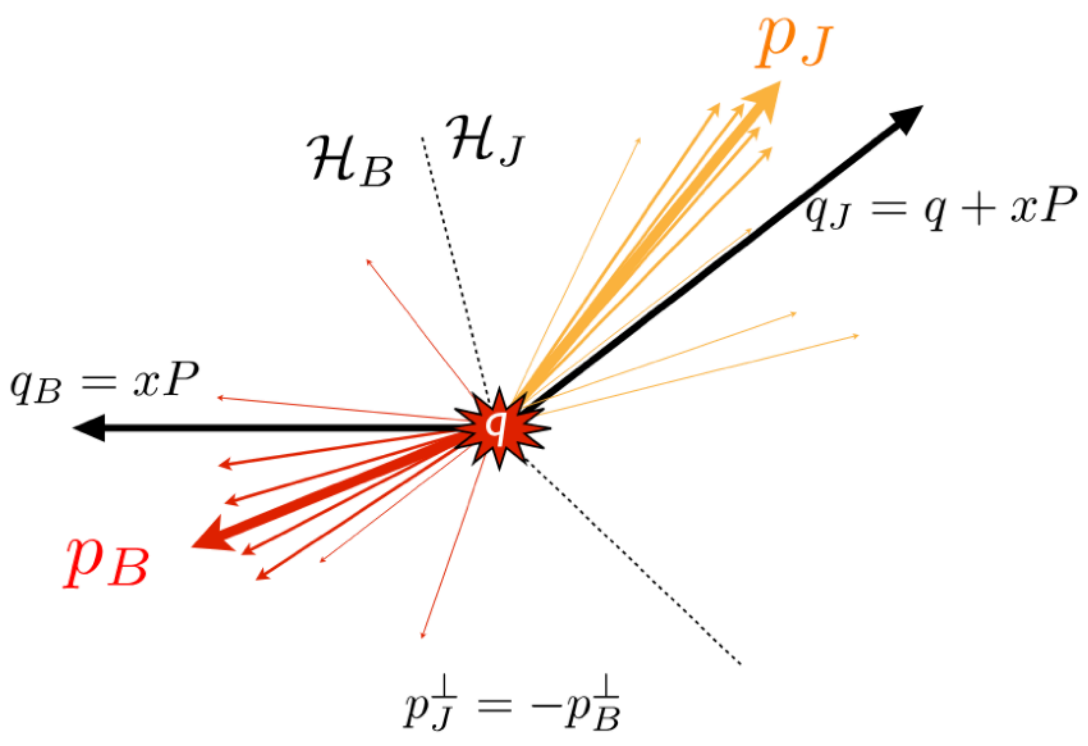
Personal selection of results that are representative for on-going progress.
Apologies for any relevant omission of references.

Outlook

Jet algorithms

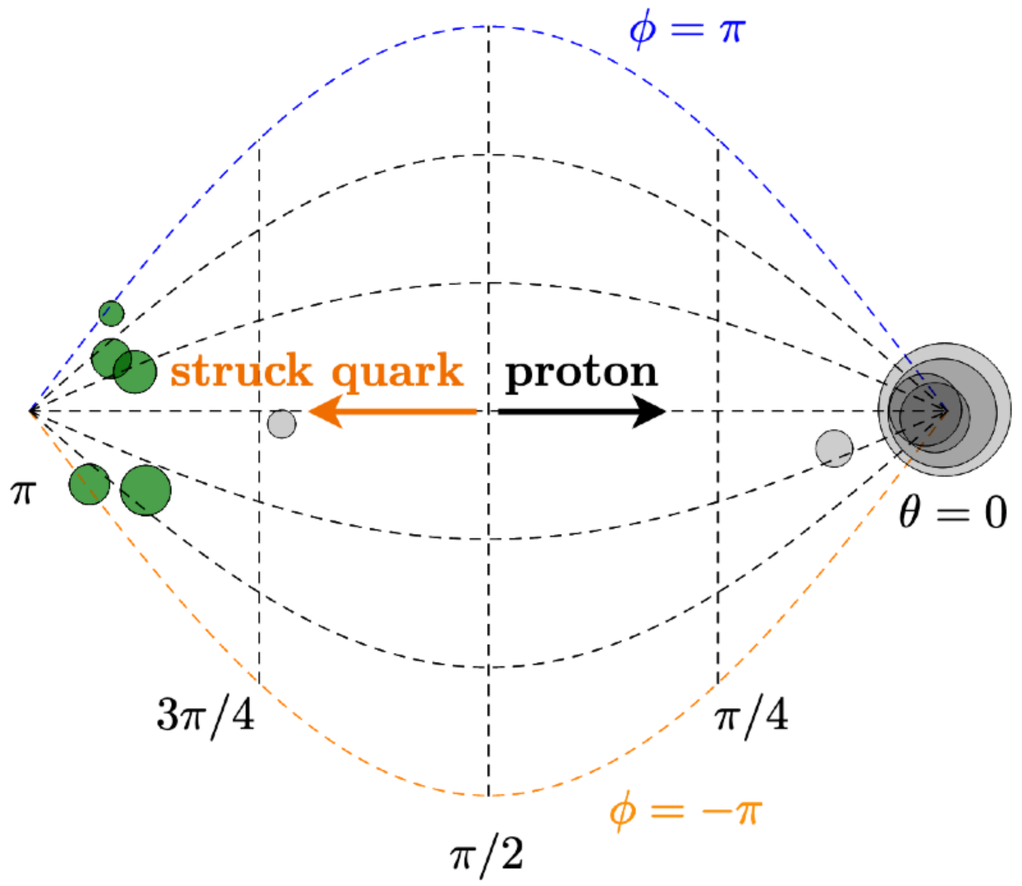


Event shapes

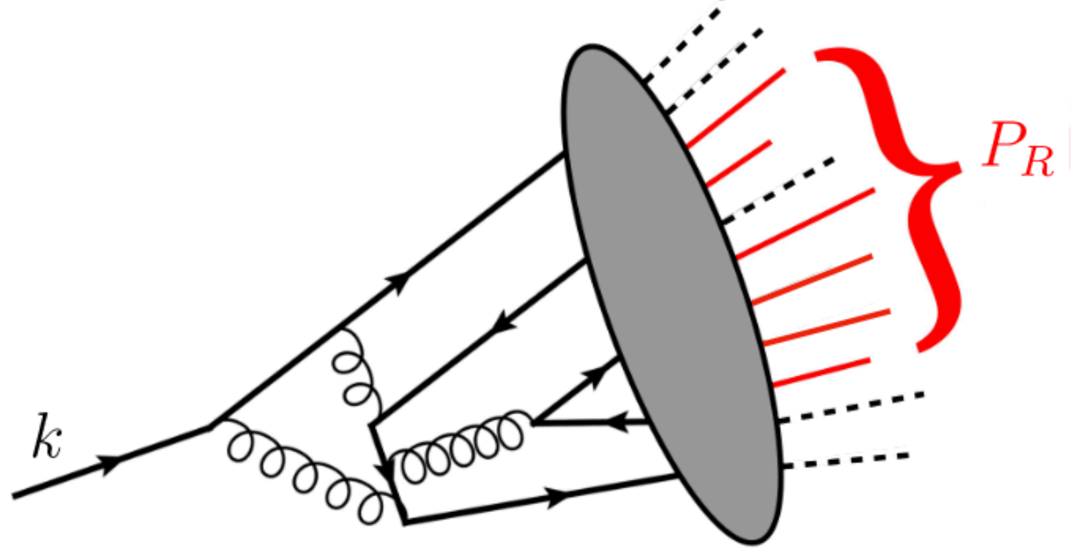


(b) τ_1^b

Grooming



Track functions

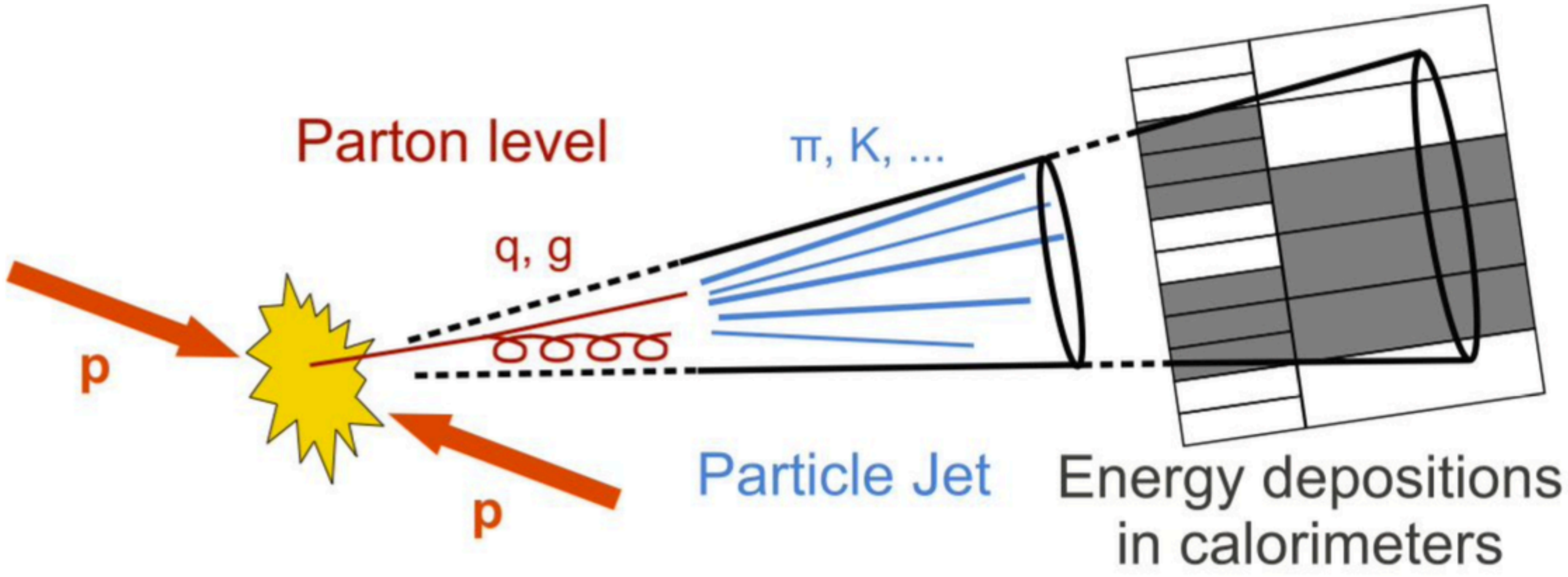


How to exploit the expertise gained from the LHC in DIS context?

Personal selection of results that are representative for on-going progress.
Apologies for any relevant omission of references.

What is a jet?

A jet is the macroscopic manifestation of QCD dynamics at high energies
i.e. most of the particles are soft or tend to be emitted at small angles



Naive definition: collimated bunch of hadrons flying roughly in the same direction
Proper definition: a collection of hadrons defined by means of a jet algorithm
At the LHC we usually adopt *sequential recombination clustering algorithms*
that can applied to objects at parton, particle or detector level.

Example: the gen-k_t family of clustering algorithms

Given distances d_{ij} and beam distances d_{iB} defined as:

See e.g. [Salam '09]

$$d_{ij} = \min\left(p_{ti}^{2p}, p_{tj}^{2p}\right)\frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = p_{ti}^{2p}$$

with transverse momentum p_t and angular distance $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ with y rapidity y and azimuthal angle ϕ, apply the following algorithm:

1. identify all initial objects as *pseudo-jets*
2. find the minimum distance:
   - d_{ij}: recombine the pseudo-jet (i, j) into a new pseudo-jet k by summing the 4-momenta and update the distances
   - d_{iB}: declare the pseudo-jet i as a final jet and remove all distances involving i.
3. iterate until there are no pseudo-jets left.

The parameter R is called *jet radius* (usually taken between 0.4 and 1)

Example: the gen-k_t family of clustering algorithms

The value of the parameter p defines the algorithm:

See e.g. [Salam '09]

$$d_{ij} = \min\left(p_{ti}^{2p}, p_{tj}^{2p}\right) \frac{\Delta R_{ij}^2}{R^2}\,, \quad d_{iB} = p_{ti}^{2p}$$

- $p = 1$: k_t algorithm
- $p = 0$: Cambridge/Aachen (C/A) algorithm
- $p = -1$: anti-k_t algorithm

Example: the gen-k_t family of clustering algorithms

The value of the parameter p defines the algorithm:

See e.g. [Salam '09]

- $p = 1$: k_t algorithm $\rightarrow$ mass/virtuality ordering

$$d_{ij} = \min\left(p_{ti}^2, p_{tj}^2\right) \frac{\Delta R_{ij}^2}{R^2}\,, \quad d_{iB} = p_{ti}^2$$

Distance measure reflects the splitting probability $P_{k\to ij}$:

$$P_{k\to ij} \sim \frac{\alpha_s}{\min\left(p_{ti}^2, p_{tj}^2\right)\, \Delta R_{ij}^2} \sim \frac{1}{d_{ij}}$$

(at the LHC we use variables invariant under longitudinal boosts, such as p_t and Δ_R; energies and angles are not invariant)

- $p = 0$: Cambridge/Aachen (C/A) algorithm
- $p = -1$: anti-k_t algorithm

Example: the gen-k_t family of clustering algorithms

See e.g. [Salam '09]

The value of the parameter p defines the algorithm:

- $p = 1$: k_t algorithm $\rightarrow$ mass/virtuality ordering
- $p = 0$: Cambridge/Aachen (C/A) algorithm $\rightarrow$ pure angular ordering

$$d_{ij} = \frac{\Delta R_{ij}^2}{R^2}, \quad d_{iB} = 1$$

- $p = -1$: anti-k_t algorithm

Example: the gen-k_t family of clustering algorithms

See e.g. [Salam '09]

The value of the parameter p defines the algorithm:

- $p = 1$: k_t algorithm $\rightarrow$ mass/virtuality ordering
- $p = 0$: Cambridge/Aachen (C/A) algorithm $\rightarrow$ pure angular ordering
- $p = -1$: anti-k_t algorithm $\rightarrow$ unphysical?

$$d_{ij} = \min\left(\frac{1}{p_{ti}^2}, \frac{1}{p_{tj}^2}\right)\frac{\Delta R_{ij}^2}{R^2}\,, \qquad d_{iB} = \frac{1}{p_{ti}^2}$$

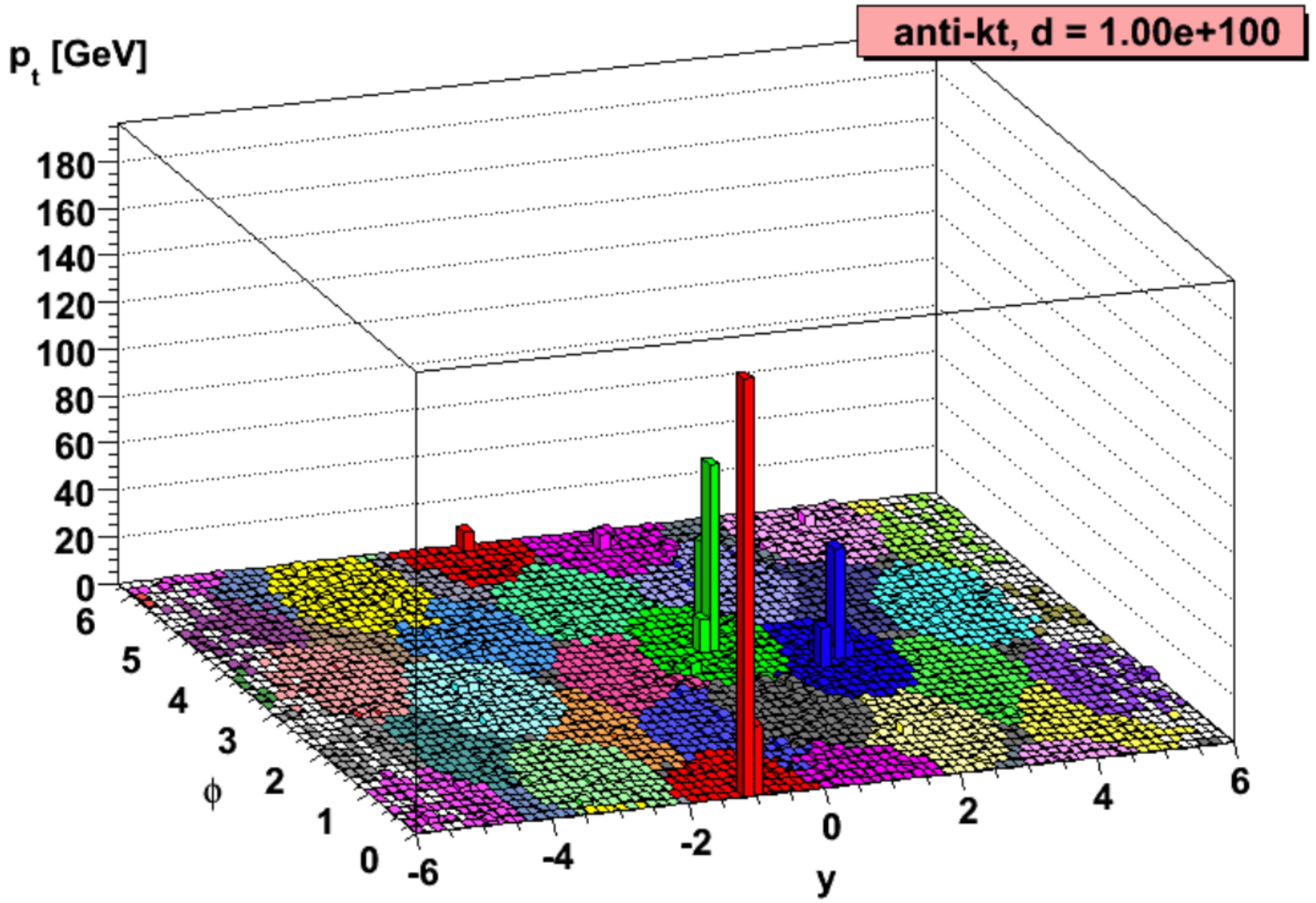
It tends to favour clustering involving hard particles rather than first recombining soft particles or particles close in angle.

Example: the gen-k_t family of clustering algorithms

The value of the parameter p defines the algorithm:

See e.g. [Salam '09]

- $p = 1$: k_t algorithm $\rightarrow$ mass/virtuality ordering
- $p = 0$: Cambridge/Aachen (C/A) algorithm $\rightarrow$ pure angular ordering
- $p = -1$: anti-k_t algorithm $\rightarrow$ unphysical?



Circular jets in a theory-friendly way

A couple of comments: inclusive vs. exclusive

We presented the *inclusive* version of the algorithms:
- all particles are included in final-state jets
- what gets called a jet is determined by the radius R

Also *exclusive* version is possible
e.g. exclusive k_t algorithm:

$$d_{ij} = \min(p_{ti}^2, p_{tj}^2)\Delta R_{ij}^2\,, \qquad \Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$
$$d_{iB} = p_{ti}^2\,,$$

- every particle is assigned either to a beam-jet or a final-state jet.
- the termination condition requires a d_{cut}: $d_{ij}, d_{iB} > d_{\text{cut}}$.

A couple of comments: longitudinally vs. spherically invariant

We presented the *longitudinally* invariant version of the algorithms
i.e. distances are manifestly invariant under longitudinal boosts
as they depend on $(y_i - y_j)$, ϕ_i and p_{ti}

Also a *spherically* invariant version is possible
e.g. gen-k_t in spherical coordinates

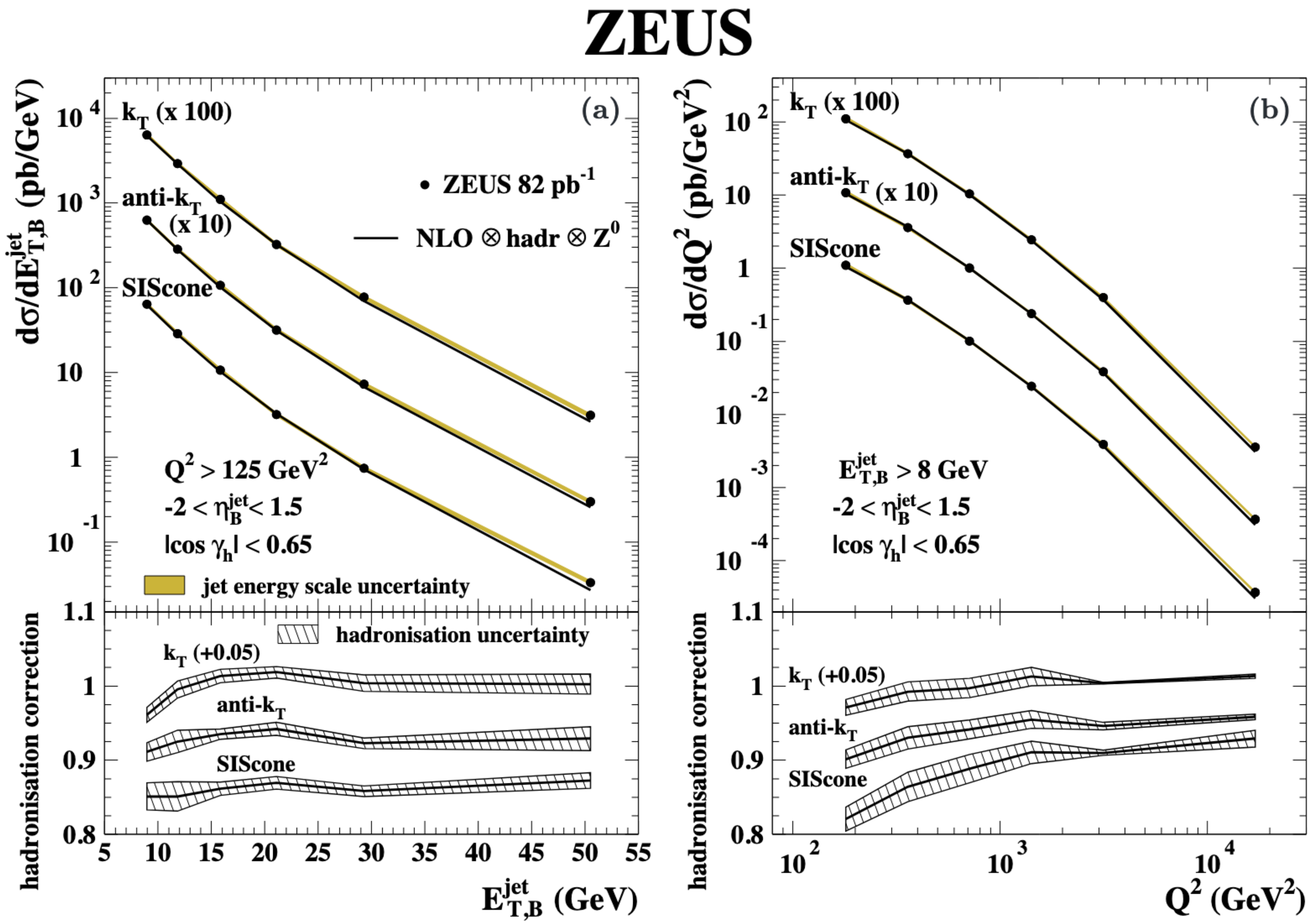
$$d_{ij} = \min(E_i^{2p}, E_j^{2p}) \frac{(1 - \cos\theta_{ij})}{(1 - \cos R)}$$

$$d_{iB} = E_i^{2p} \,,$$

Suited for jet reconstruction in e^+e^- annihilation

What about DIS?

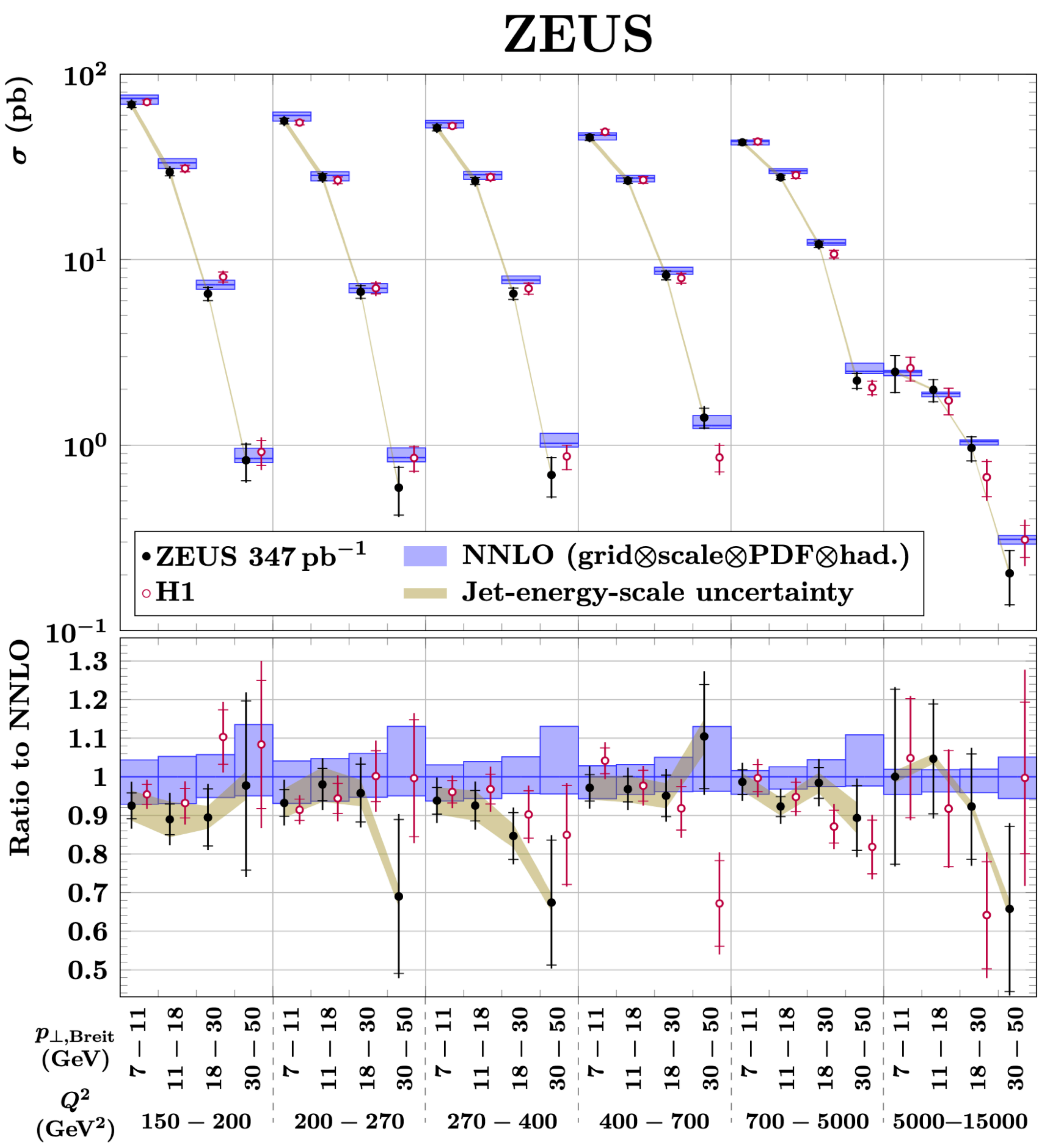
Preferred algorithm at HERA was inclusive longitudinally invariant k_t algorithm



However, for high-energy jets at HERA, it was established that the k_t and anti-k_t algorithms have very similar performances.

[ZEUS ’10]

The measured cross sections for the three jet algorithms have similar shapes and normalisations.



[ZEUS ’23]

Recent re-analysis of HERA data.

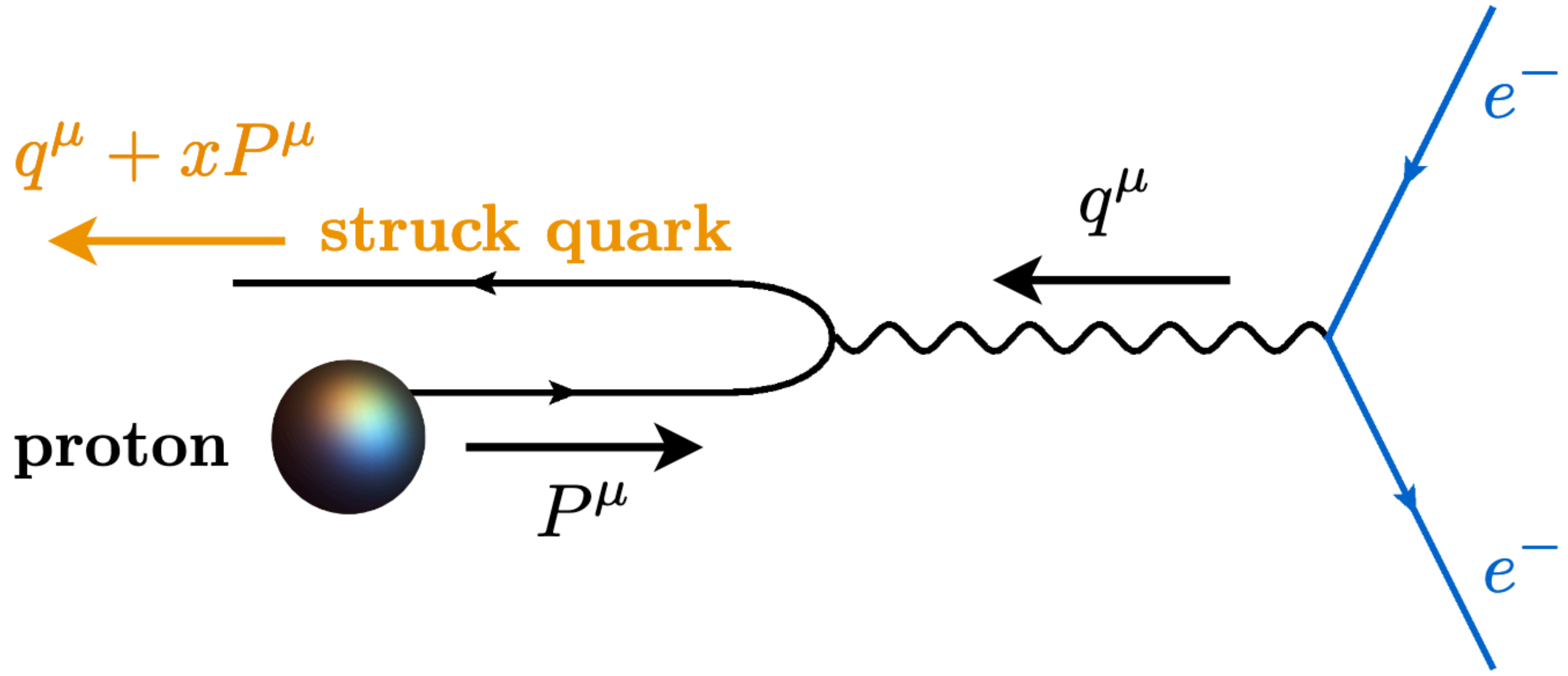
Inclusive jet cross section with more than 70% of the entire luminosity.

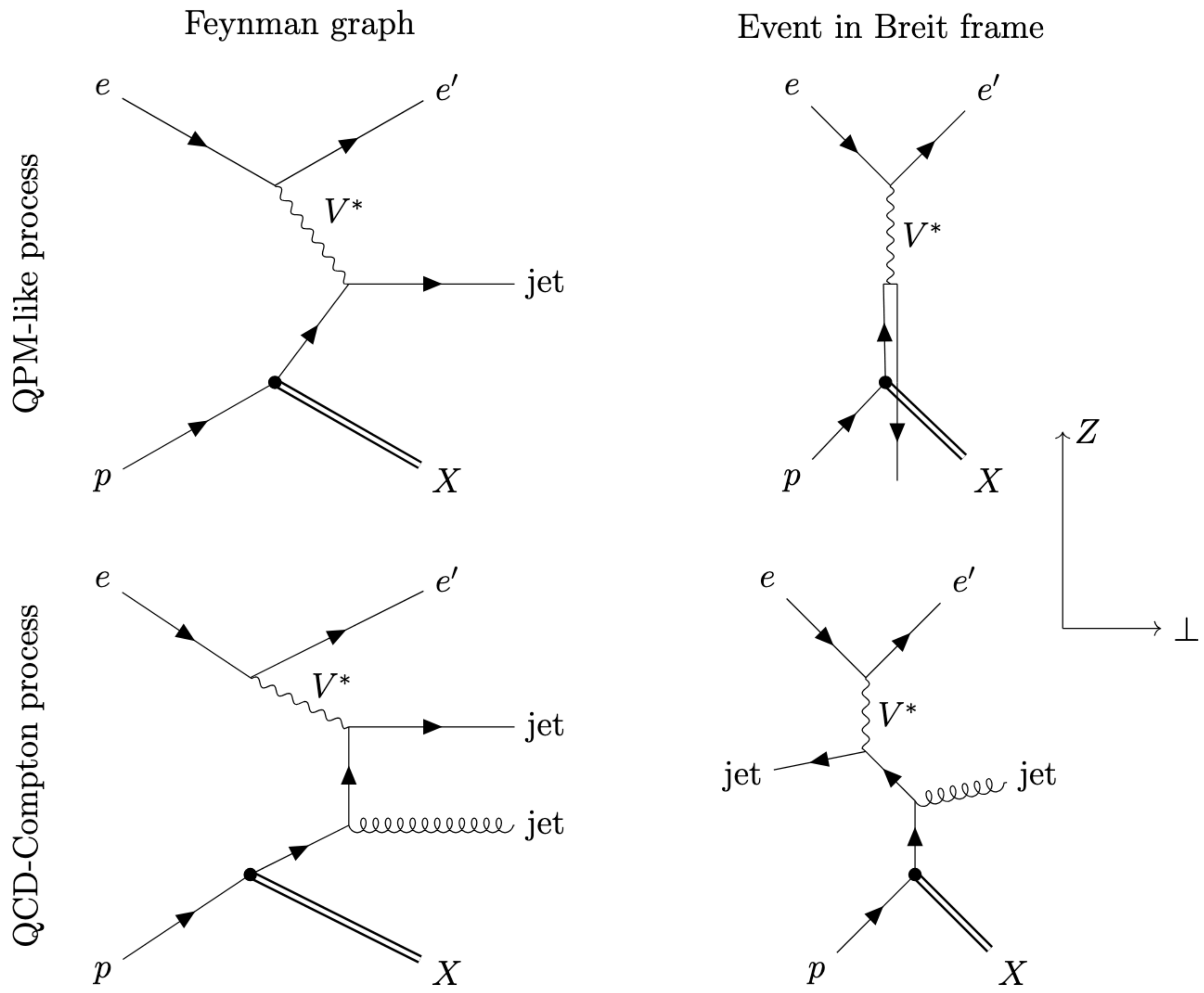
Measurement compatible with H1 and nice agreement with NNLO theory calculations from NNLOJET
[Currie, Gehrmann, Huss, Niehues ’17]

Also new extraction of strong coupling:

$$\alpha_s(M_Z^2) = 0.1142 \pm 0.0019$$

HERA analysis identifies jets in the Breit Frame: parton and boson collide head-on





Analysis phase space

$$150\,\mathrm{GeV}^2 < Q^2 < 15\,000\,\mathrm{GeV}^2$$
$$0.2 < y < 0.7$$
$$7\,\mathrm{GeV} < p_{\perp,\mathrm{Breit}} < 50\,\mathrm{GeV}$$
$$-1 < \eta_{\mathrm{lab}} < 2.5$$

In particular by requiring a minimum transverse momentum in this frame, one suppresses single-jet production of zeroth order in α_s.

Interesting observables with jets: lepton-jet azimuthal correlations

Proposed in [Liu, Ringer, Vogelsang, Yuan ’18] to study TMDs, both in lepton-nucleon and lepton-nucleus collisions, as a complementary approach to SIDIS TMD.

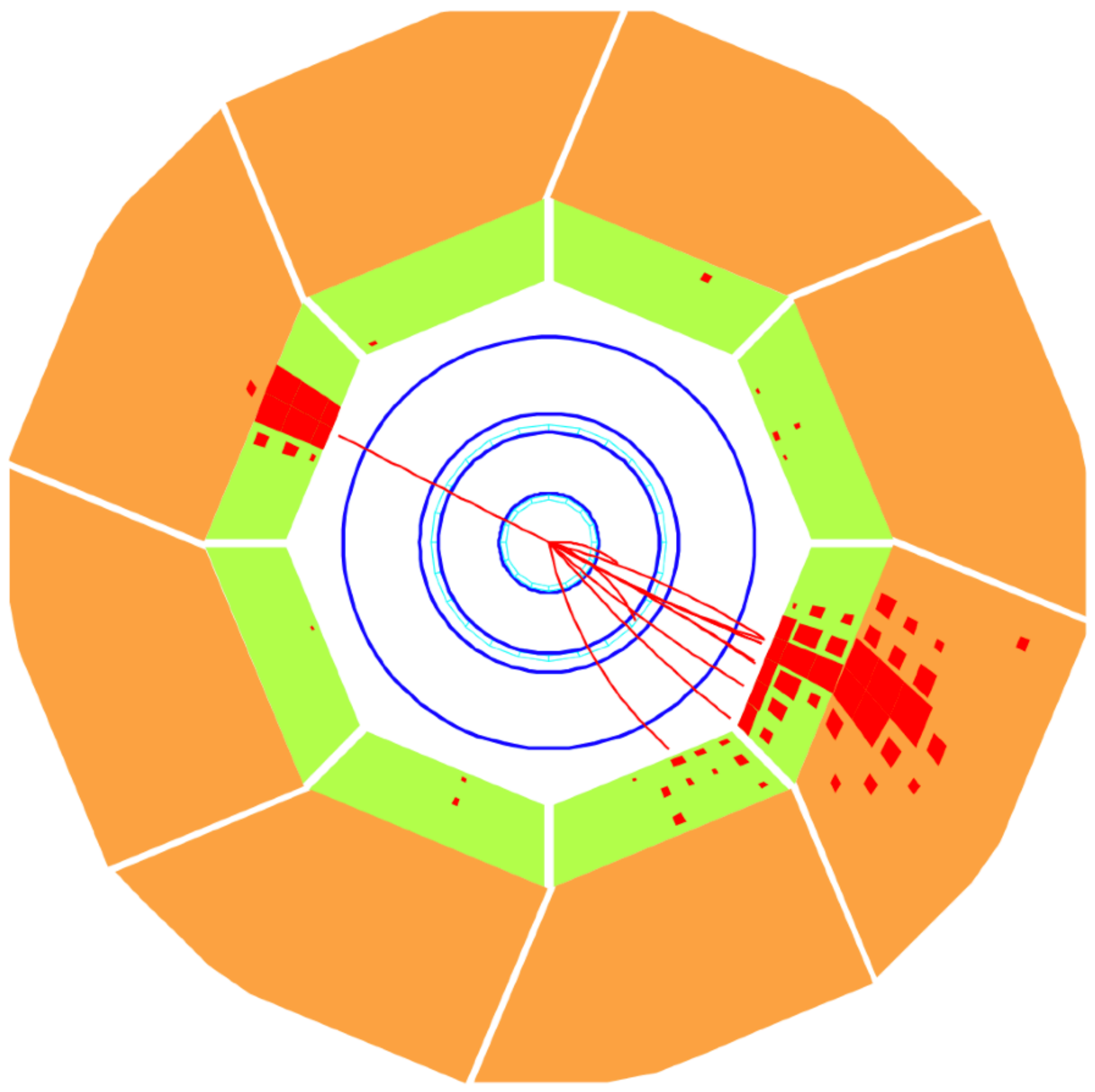
At Born level, lepton and jet are produced back-to-back in the laboratory frame.

$$e + p \to e + \text{jet} + X$$

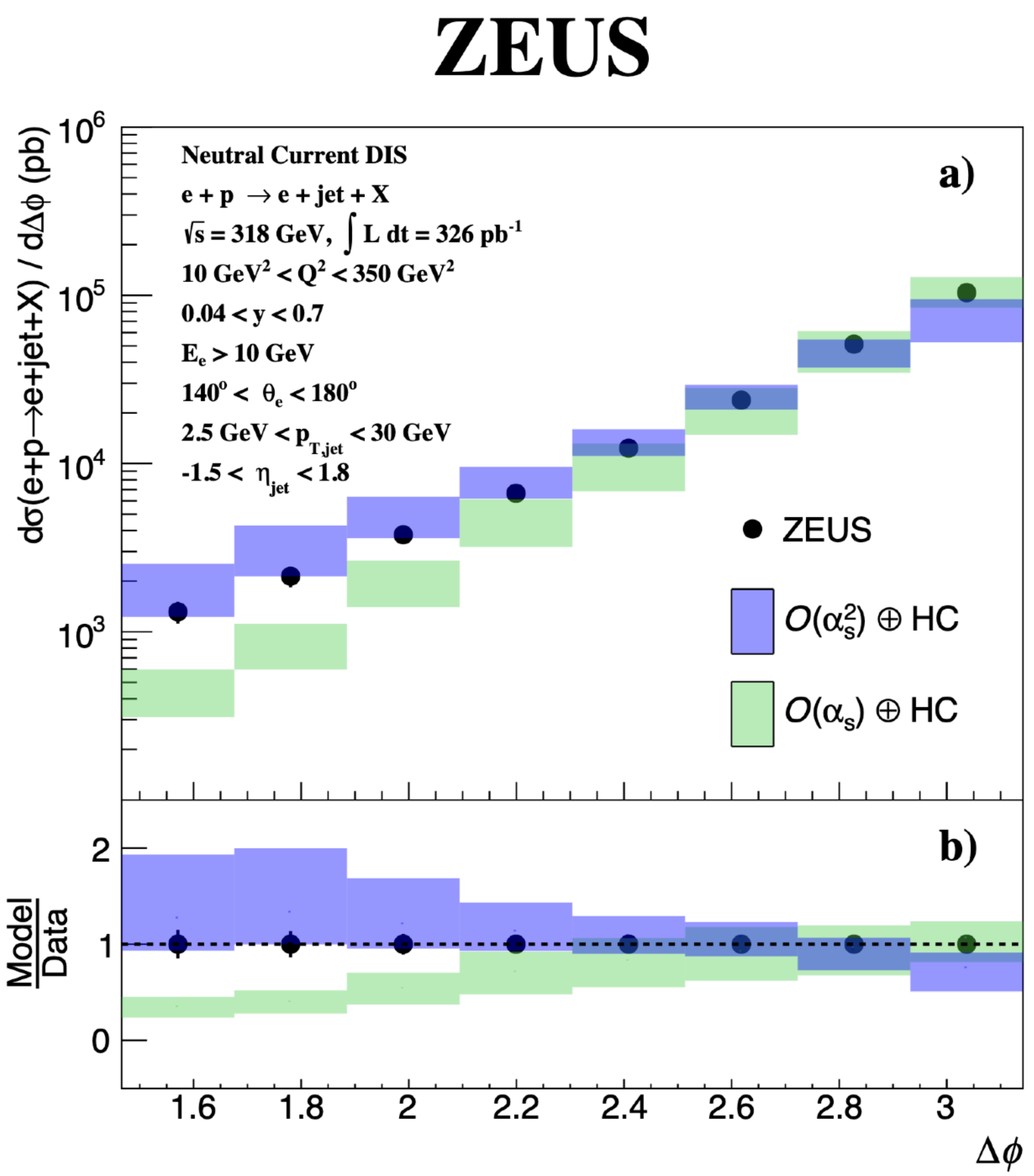
Observables:

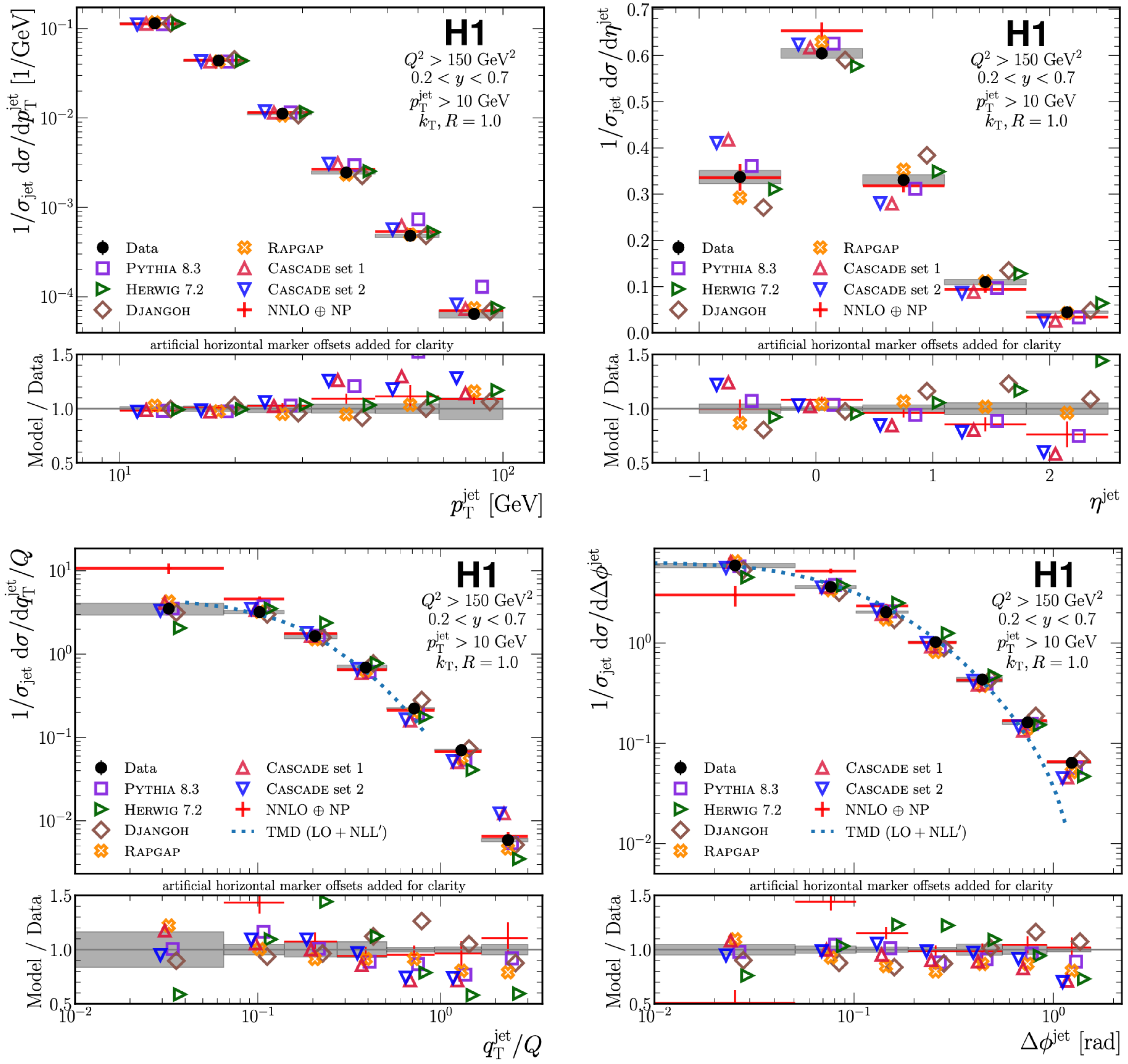
- azimuthal correlation $\Delta\phi = |\phi_e - \phi_{\text{jet}}|$
- momentum imbalance $q_T = |\vec{p}_T^{\,e} + \vec{p}_T^{\,\text{jet}}|$

When $\Delta\phi \to \pi$ or $q_T \to 0$, soft gluons and/or non-zero transverse momentum of the struck parton.



Recent measurements by [ZEUS '24] and [H1 '21]

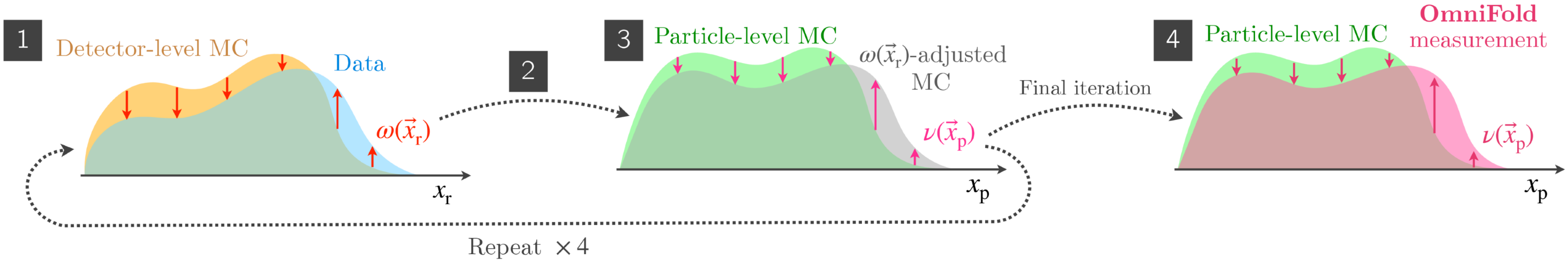




NNLO fixed-order predictions with projection-to-Born method [Borsa, De Florian, Pedron '21], supplemented with hadronization corrections from MC

Excursus: machine learning-based unfolding

The H1 measurement adopts *OmniFold* [Andreassen, Komiske, Metodiev, Nachman, Thaler '21].
Unfolding performed simultaneously for eight observables: $(\vec{p}_T^{\,e}, p_z^e, p_T^{\rm jet}, \eta^{\rm jet}, \phi^{\rm jet}, q_T^{\rm jet}/Q, \Delta\phi^{\rm jet})$



Two classifiers (neural networks) trained to distinguish the:

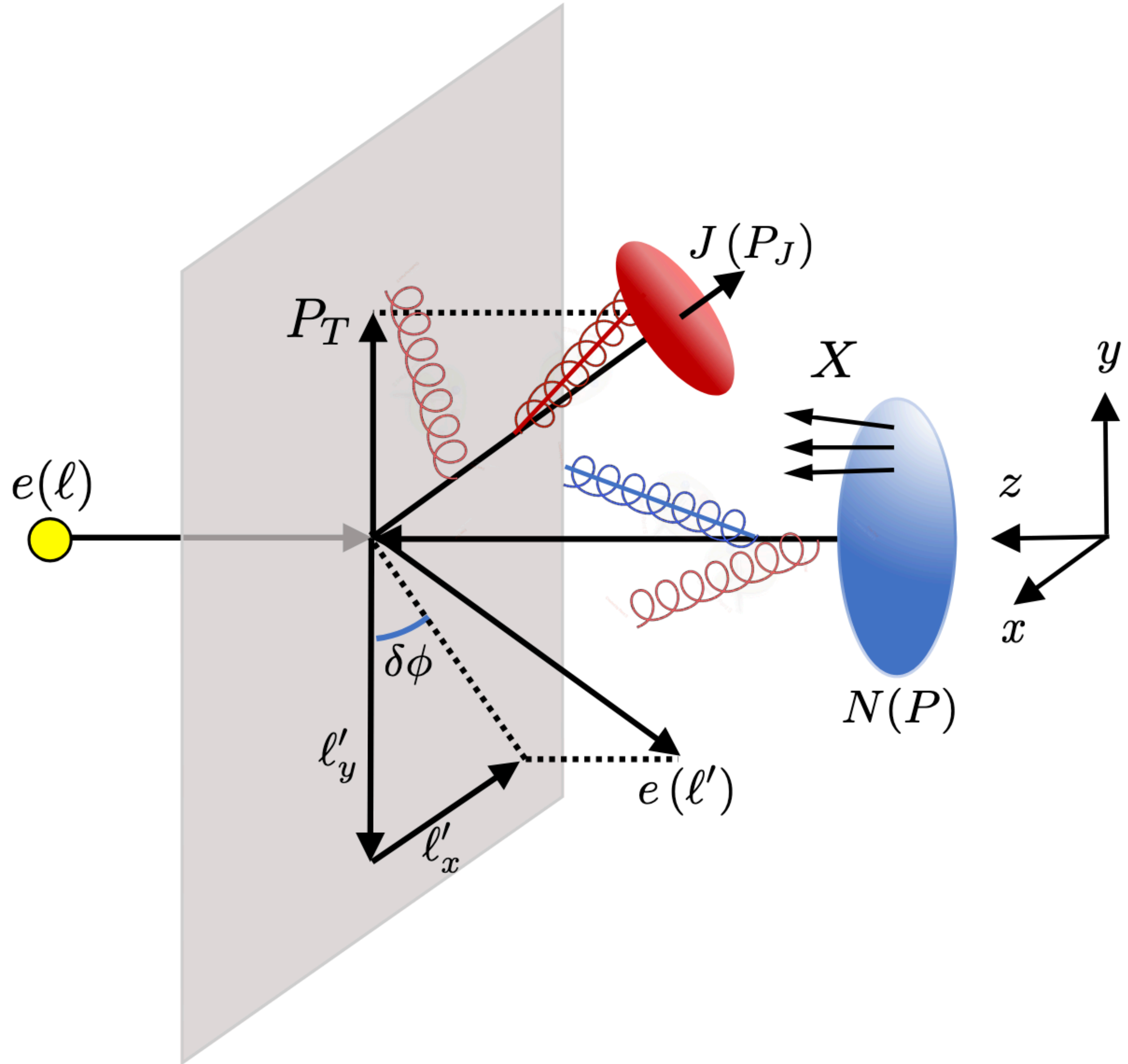
- detector-level MC (with ν weight initialised to 1) from data -> ω weight
- particle-level MC from ω-adjusted particle-level MC -> ν weight

The final output is a set of "truth" events that encode the measurement (histograms for any observable can be produced by binning these events)

Recent theory developments for $\delta\phi$

[Fang, Ke, Shao, Terry '23]

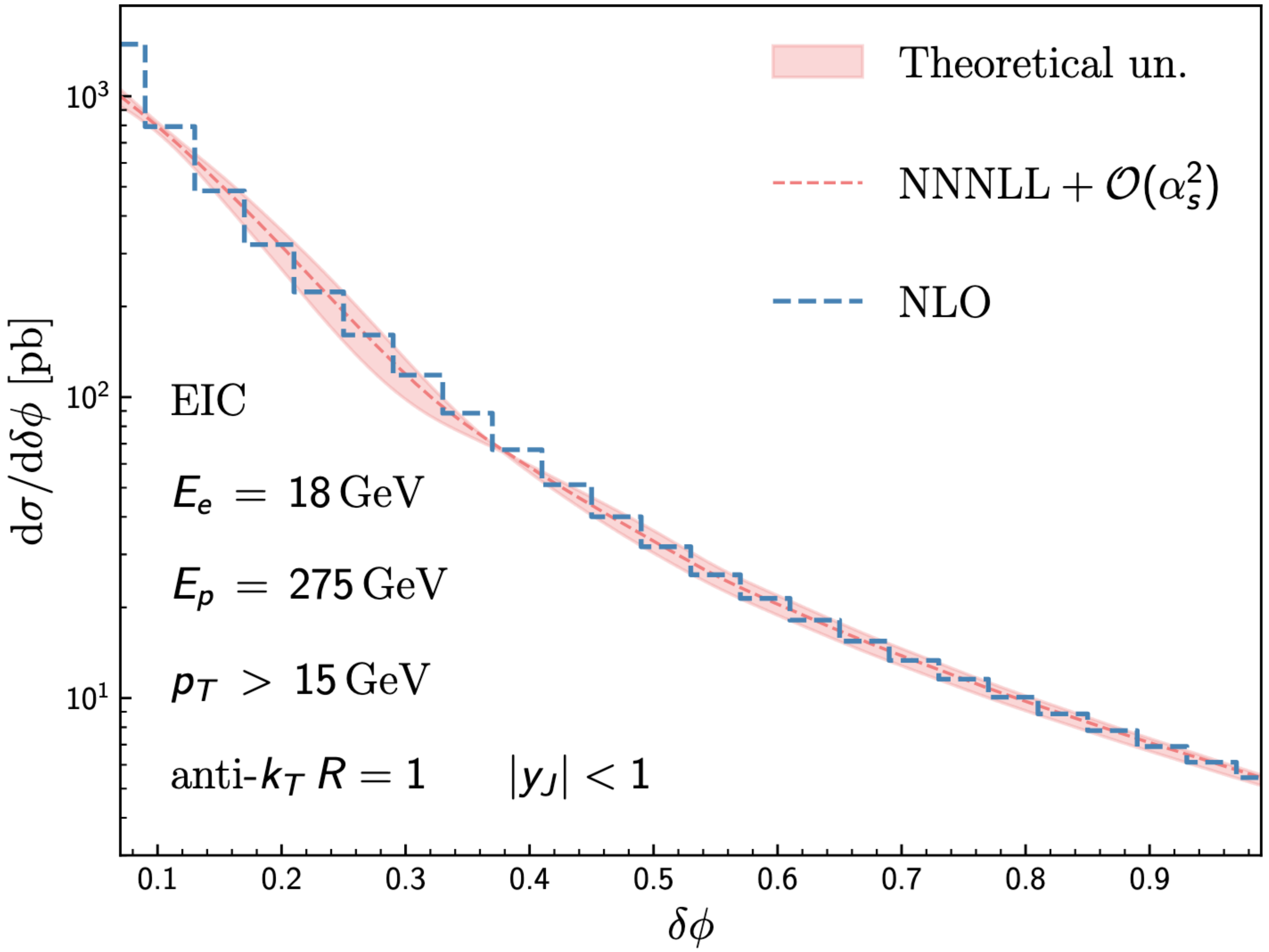
Factorisation in SCET and NNLL resummation in the $\delta\phi \to 0$ region



Use of winner-take-all (WTA) axis to remove non-global logarithms

[Fang, Gao, Li, Shao '24]

Resummation pushed to N3LL and matching to fixed-order result with NLOJET++



New jet algorithms for DIS

- Is possible to capture jets from the fragmentation of the struck-quark? In the Breit frame, they point to the beam direction i.e. infinite rapidity.

- Using a spherically-invariant algorithm solves the problem, but at the price of breaking longitudinal invariance.

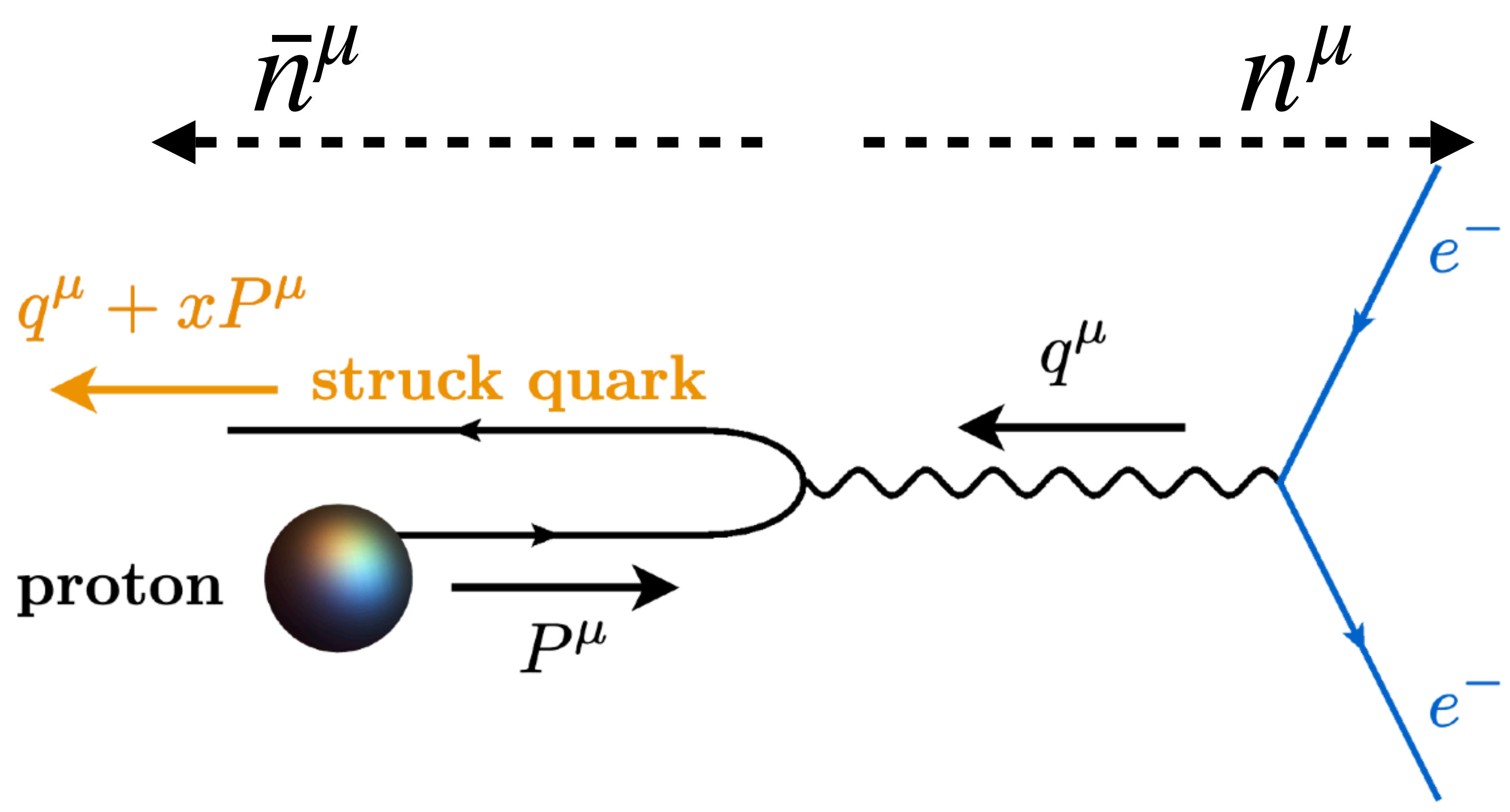
Possible solution: Centauro
[Arratia, Makris, Neill, Ringer, Sato '20]

$$d_{ij} = \left[(\Delta f_{ij})^2 + 2 f_i f_j (1 - \cos\Delta\phi_{ij})\right]/R^2 \,, \quad d_{iB} = 1$$

$$f_i = \frac{2p_i^\perp}{n \cdot p_i}$$

Forward region ($n \cdot p_i \to 0$): beam distance always smaller

Backward region ($n \cdot p_i \sim 1$): f_i is angle w.r.t. the photon axis



$$q^\mu = \frac{Q}{2}(\bar{n}^\mu - n^\mu) = Q(0,0,0,-1)$$

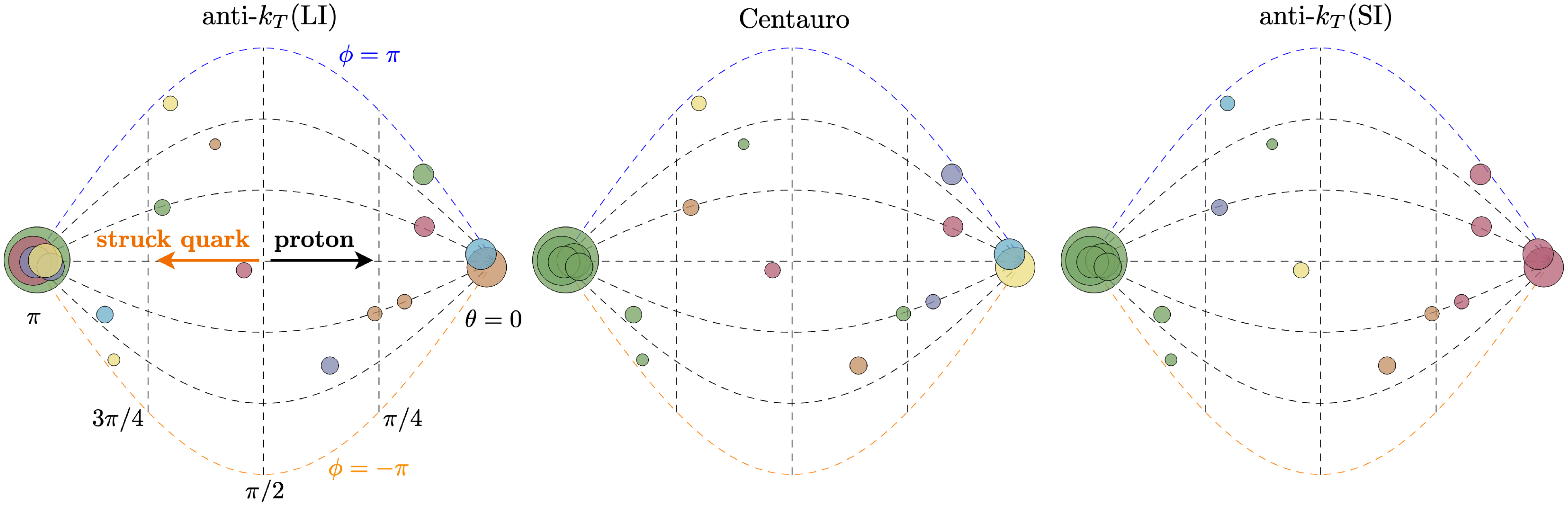
$$P^\mu \simeq Q/(2x_B)n^\mu = Q/(2x_B)(1,0,0,+1)$$

$$p_q^\mu = xP^\mu + q^\mu \simeq (Q/2)\bar{n}^\mu$$

New jet algorithms for DIS: Centauro

In the backward region similar to spherically-invariant (SI) anti-k_t

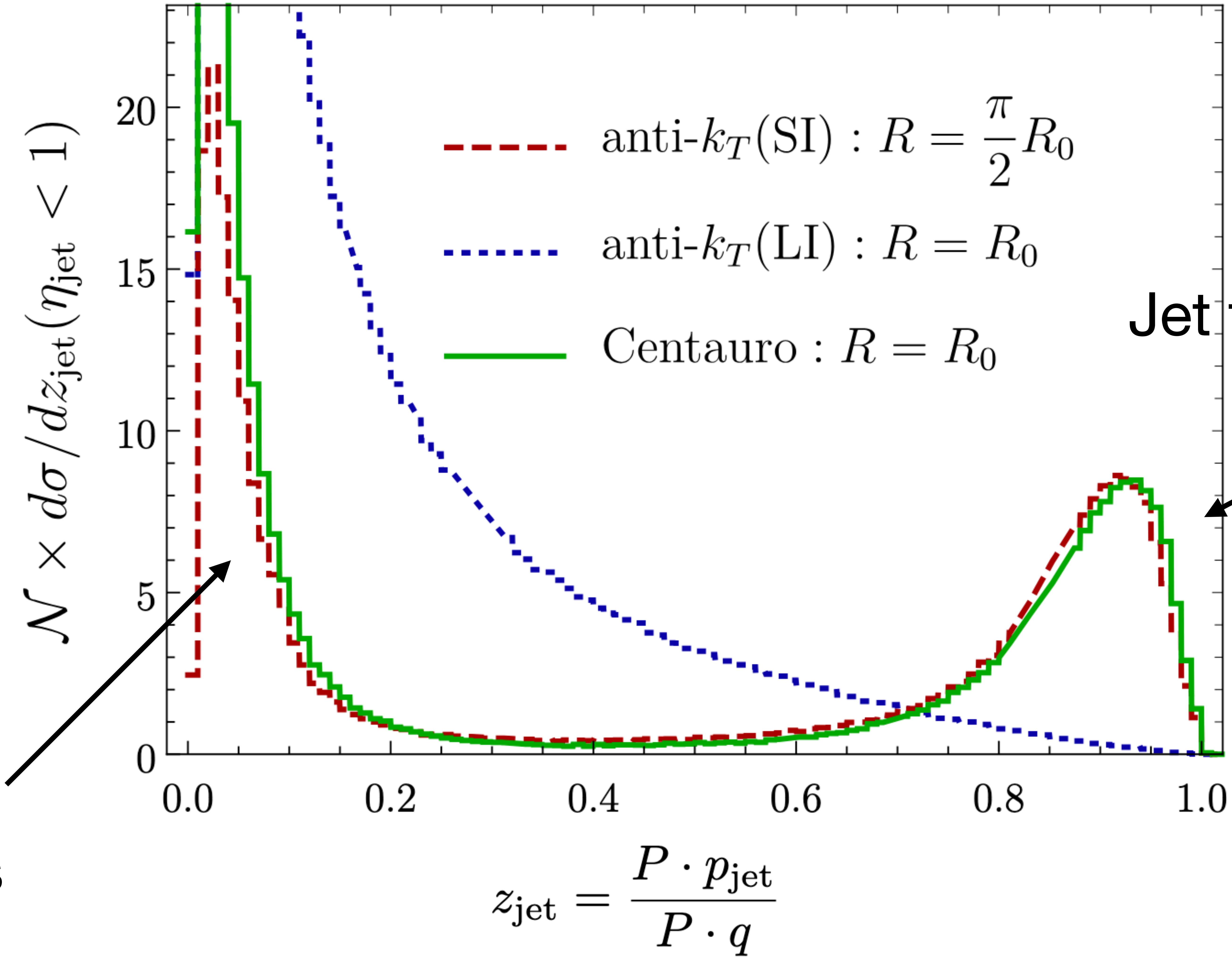
In the forward region similar to longitudinally-invariant (LI) anti-k_t



New jet algorithms for DIS: Centauro

Backward region

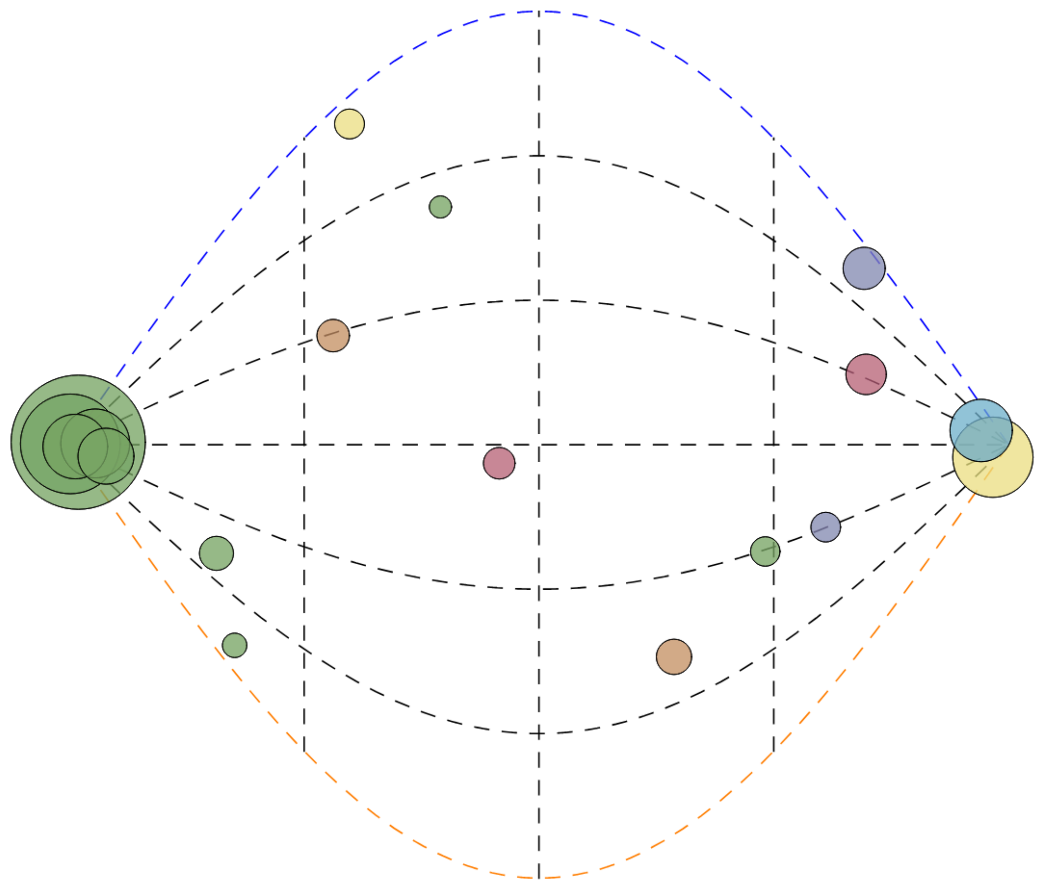
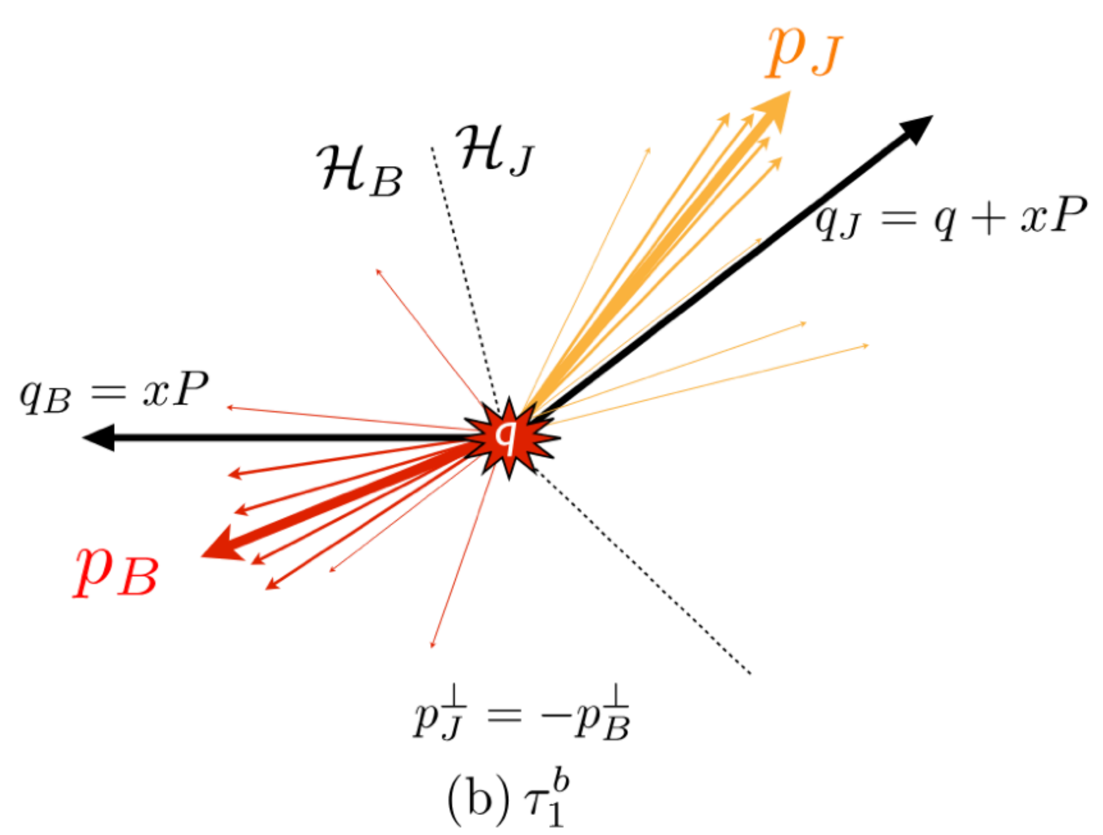
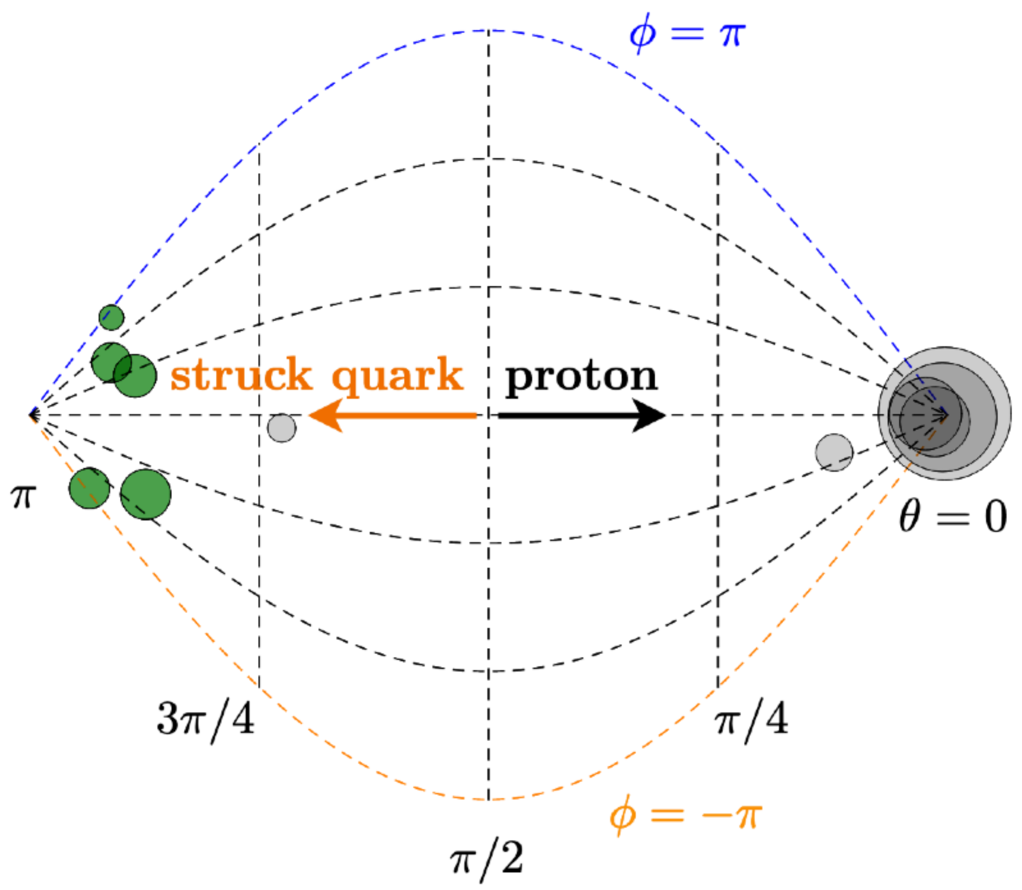
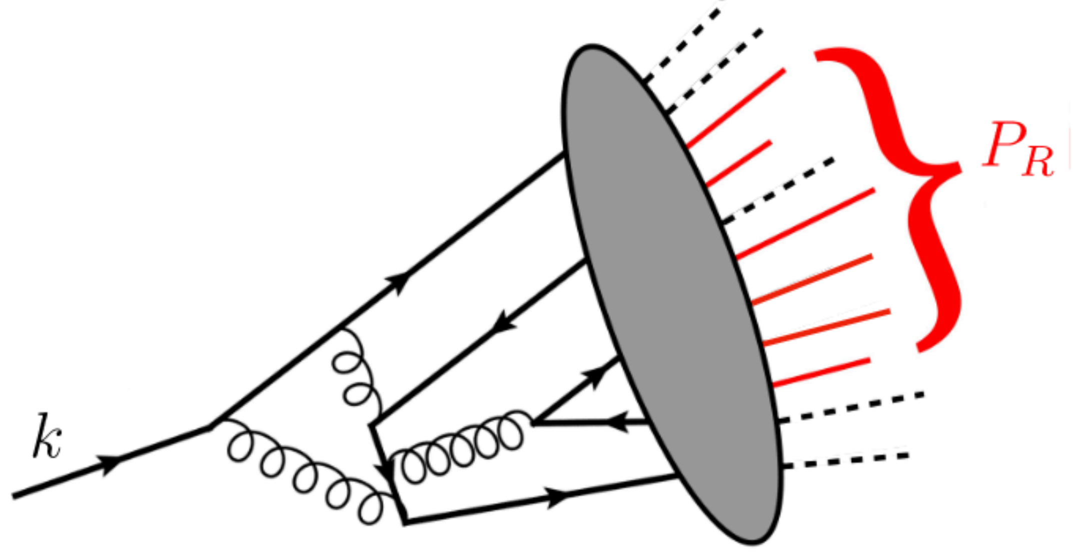
$(\eta_{\rm jet} < 0)$



Jet from struck-quark fragmentation

Mid-rapidity low-energy jets

Outlook

Jet algorithms	Event shapes	Grooming	Track functions
			

How to exploit the expertise gained from the LHC in DIS context?

Personal selection of results that are representative for on-going progress.
Apologies for any relevant omission of references.

Global event shapes at hadron colliders

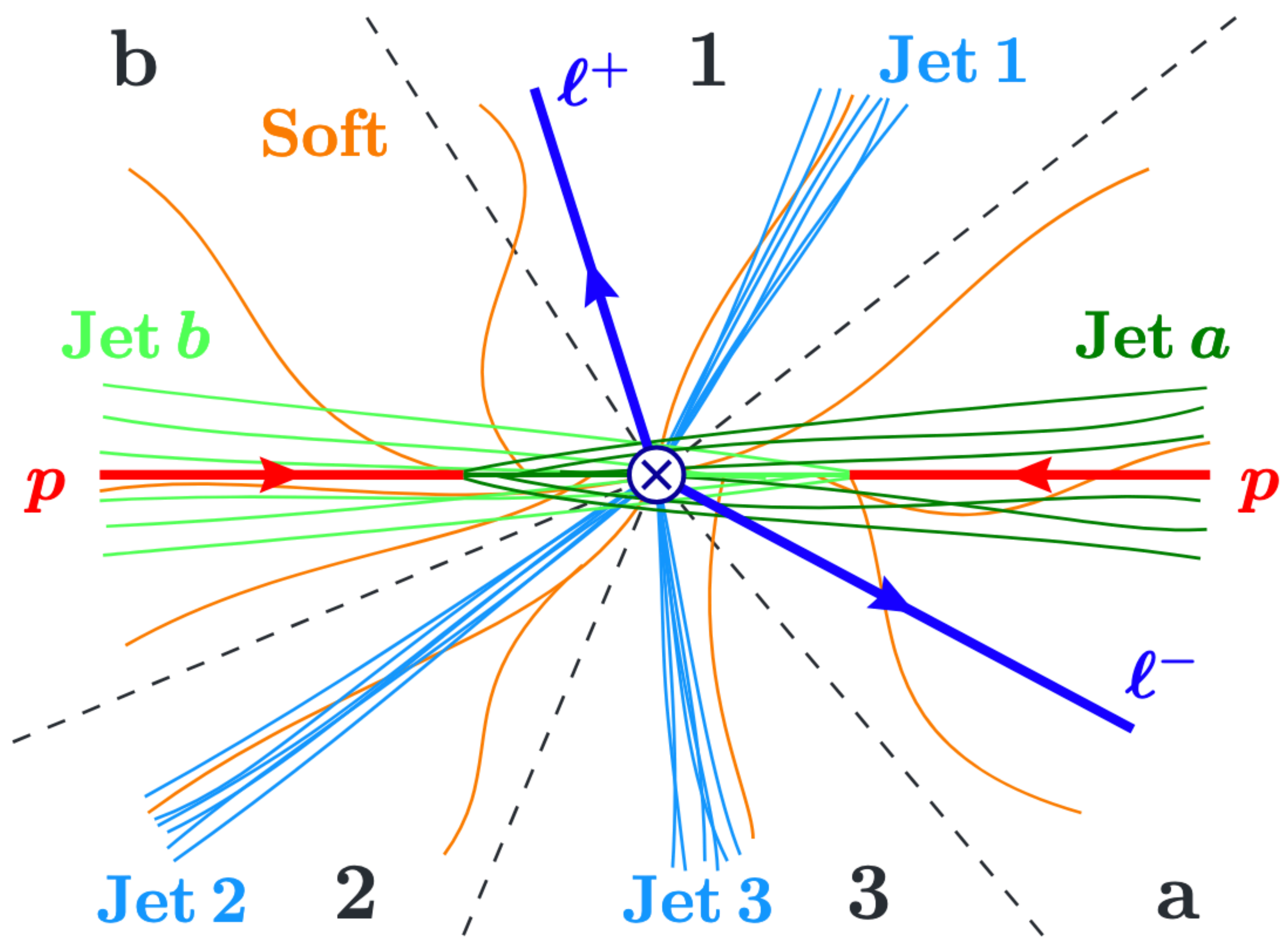
Will focus on N-jettiness
[Stewart, Tackmann, Waalewijn '10]

$$\tau_N = \frac{2}{Q^2} \sum_k \min\{q_a \cdot p_k,\, q_b \cdot p_k,\, q_1 \cdot p_k,\, \ldots,\, q_N \cdot p_k\}$$

q_a, q_b and $q_1, \ldots, q_N$ are massless reference momenta

It vanishes for exactly N infinitely narrow jets

Factorization formula with jet and beam functions



(c) $pp \to$ leptons plus jets.

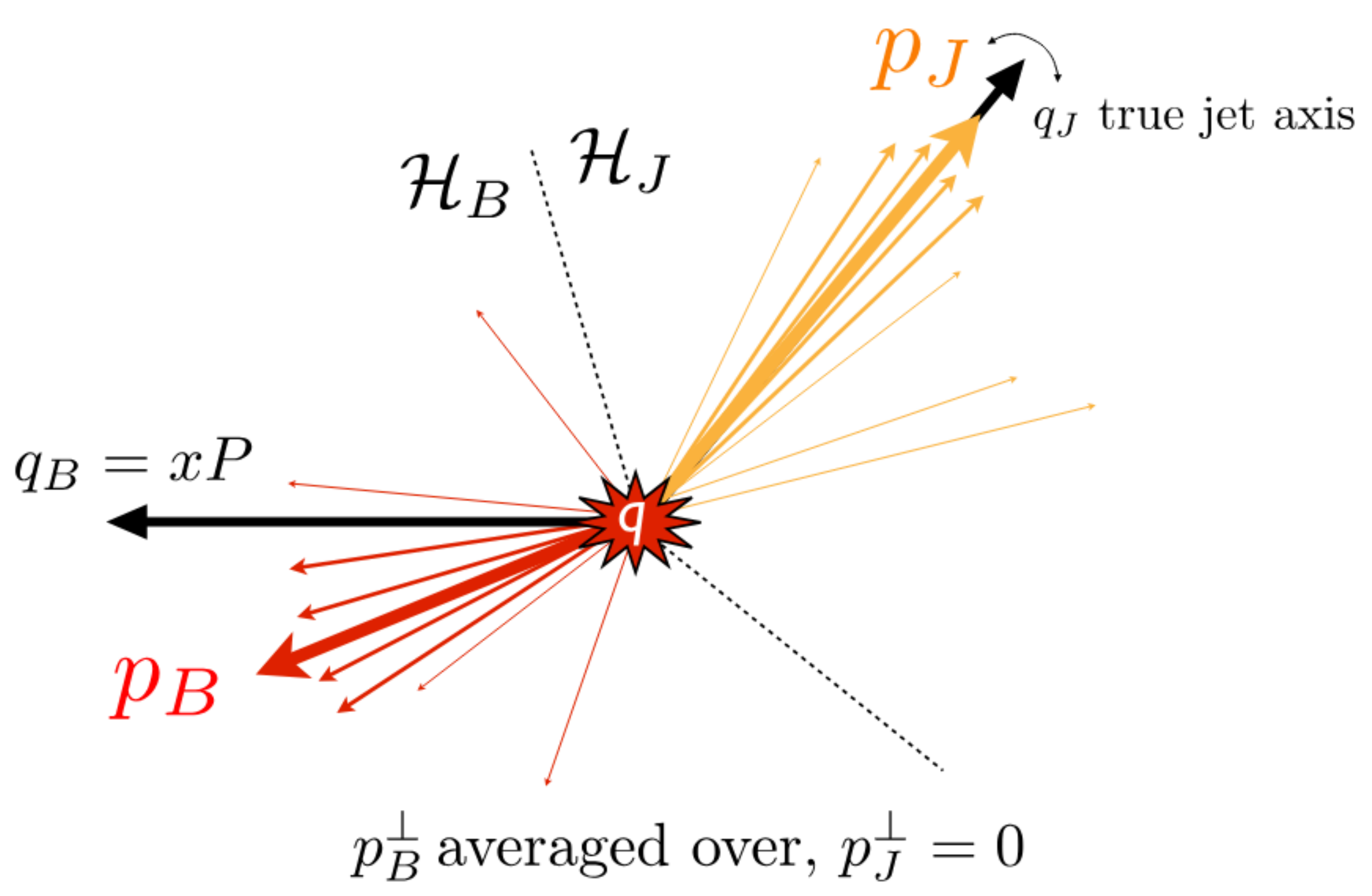
Used in a variety of contexts: as resolution parameter for slicing calculations (0/1-jettiness subtraction); in matching fixed-order with parton shower (GENEVA framework); as discriminating variable between signal and background events; …

One-jettiness for DIS

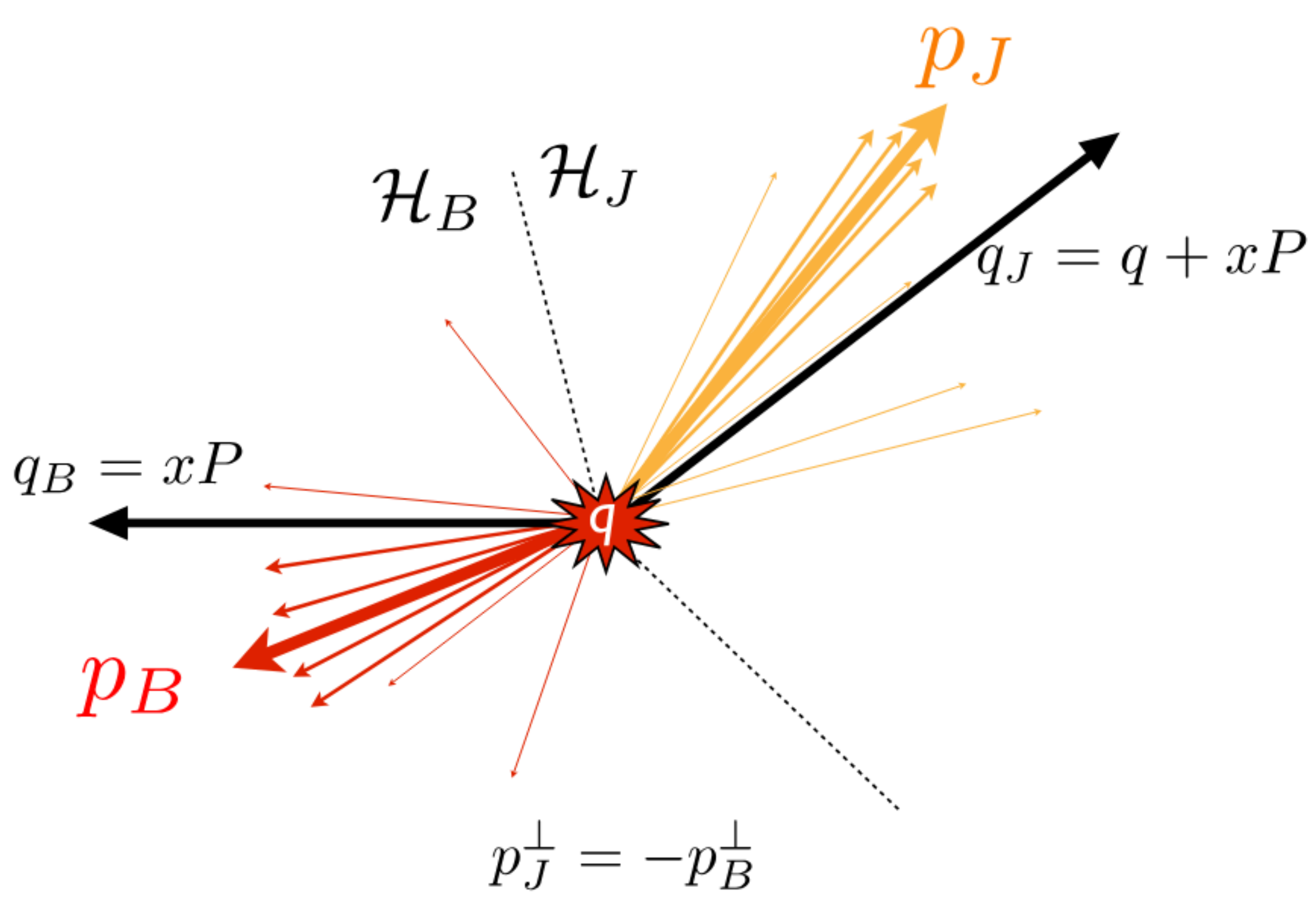
3 possible definitions explored in [Kang, Lee, Stewart '13]

$$\tau_1 = \frac{2}{Q^2} \sum_{i \in X} \min\{q_B \cdot p_i, q_J \cdot p_i\}$$

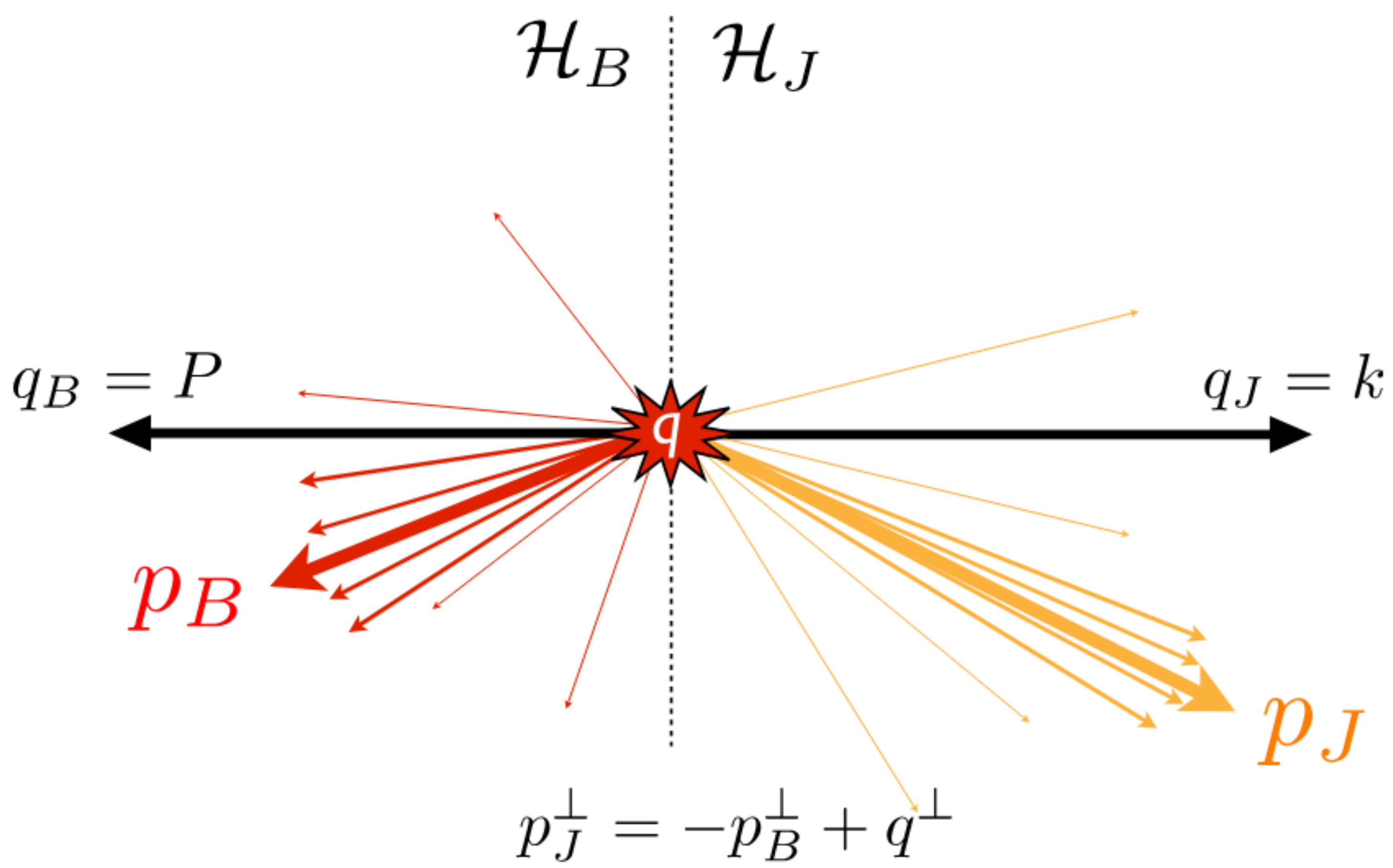
$$\begin{aligned} \tau_1^a &: \quad q_B^a = xP\,, \quad && q_J^a = \text{jet axis} \\ \tau_1^b &: \quad q_B^b = xP\,, \quad && q_J^b = q + xP \\ \tau_1^c &: \quad q_B^c = P\,, \quad && q_J^c = k\,, \end{aligned}$$



(a) τ_1^a



(b) τ_1^b



(c) τ_1^c

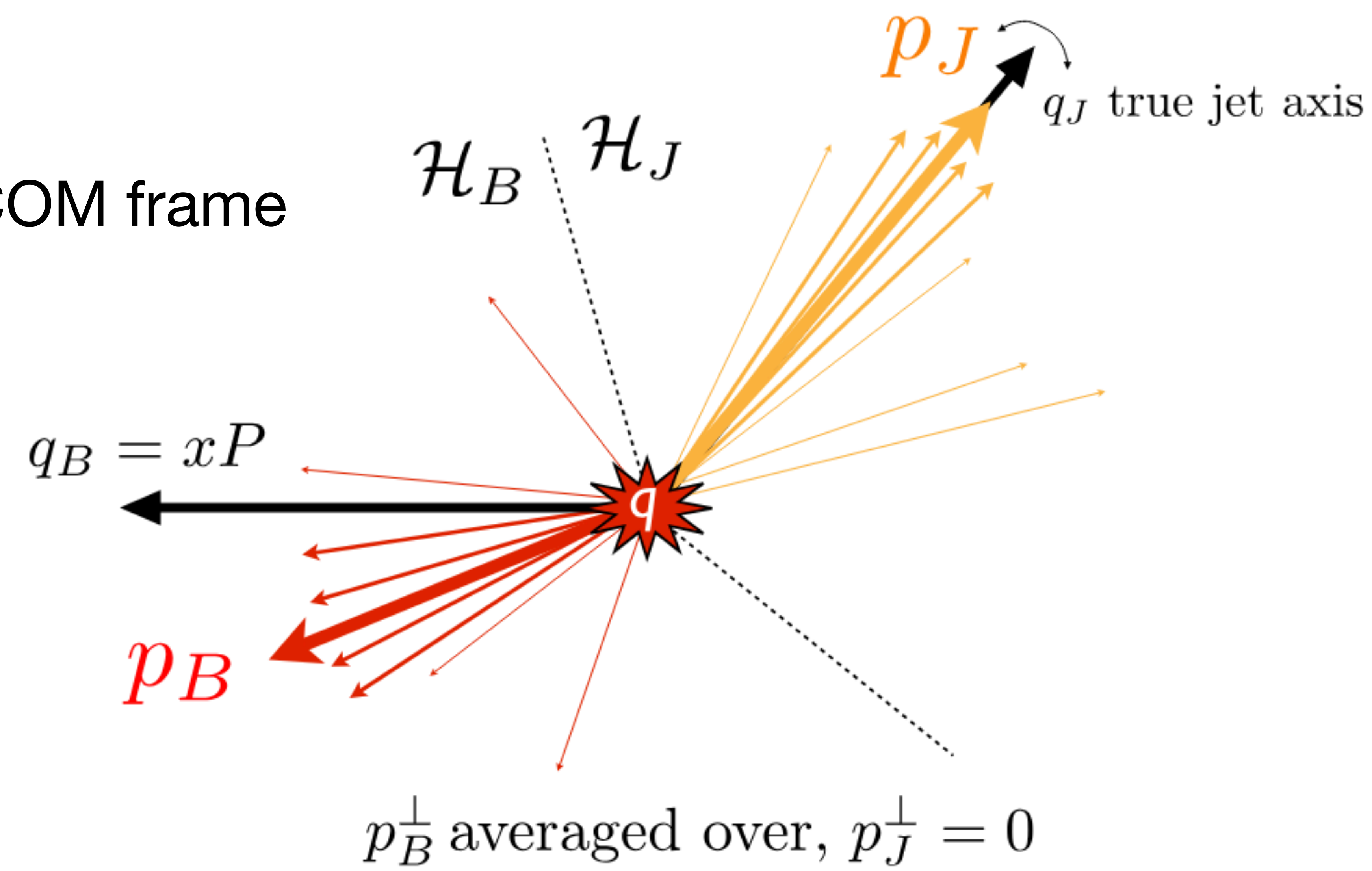
One-jettiness for DIS

3 possible definitions explored in [Kang, Lee, Stewart '13]

$$\tau_1 = \frac{2}{Q^2} \sum_{i \in X} \min\{q_B \cdot p_i, q_J \cdot p_i\}$$

$$\begin{array}{lll}
\tau_1^a : & \boxed{q_B^a = xP\,, \qquad q_J^a = \text{jet axis}} \\
\tau_1^b : & q_B^b = xP\,, & q_J^b = q + xP \\
\tau_1^c : & q_B^c = P\,, & q_J^c = k\,,
\end{array}$$

eP COM frame



(a) τ_1^a

Simplest factorisation formula, as q_J coincides with the jet axis
-> ordinary beam and jet functions

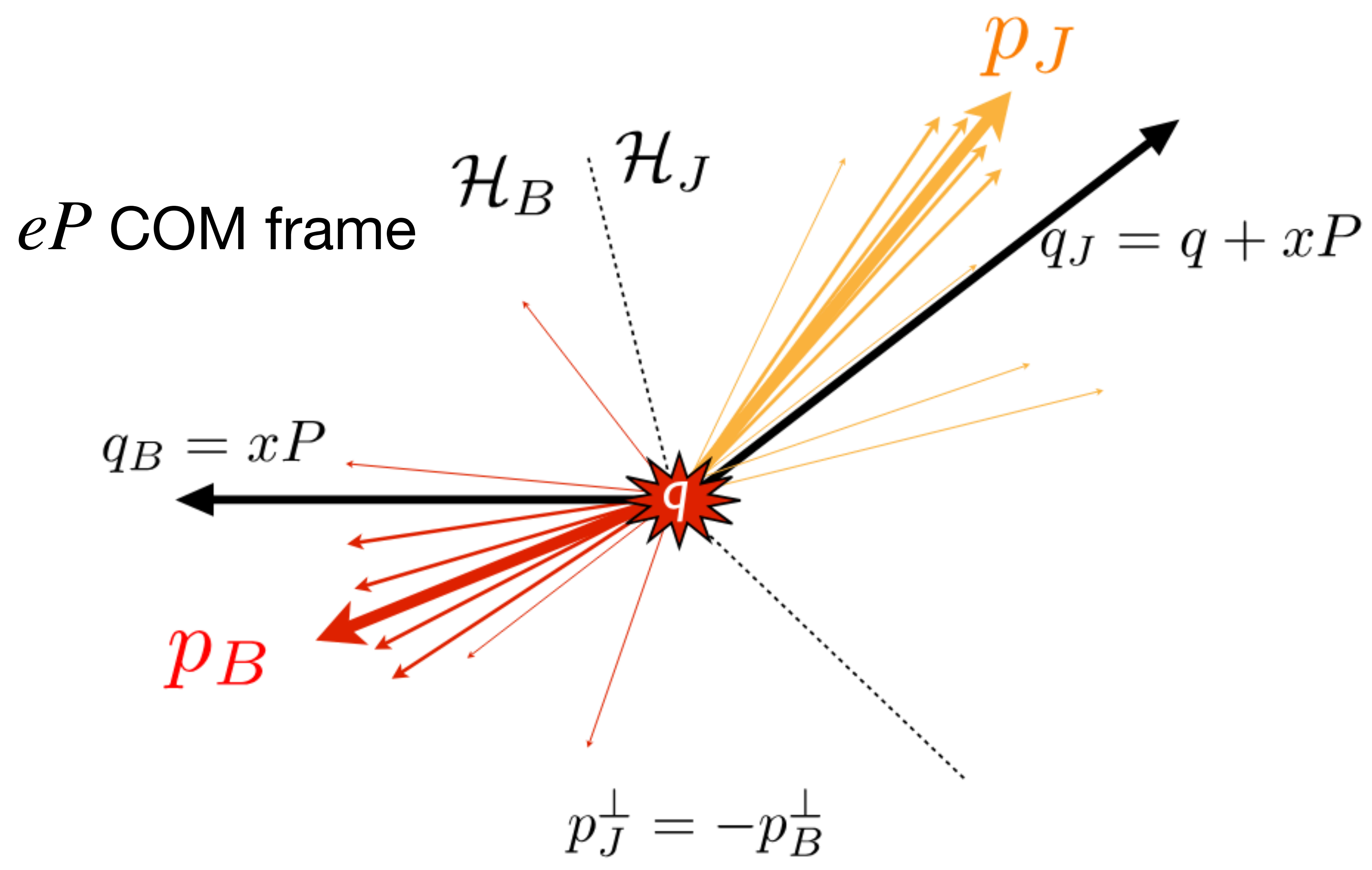
Predictions pushed to N3LL+$\mathcal{O}(\alpha_s^2)$
[Cao, Kang, Liu, Mantry '24]

One-jettiness for DIS

3 possible definitions explored in [Kang, Lee, Stewart '13]

$$\tau_1 = \frac{2}{Q^2} \sum_{i \in X} \min\{q_B \cdot p_i, q_J \cdot p_i\}$$

$$\begin{aligned} \tau_1^a &: \quad q_B^a = xP\,, \quad && q_J^a = \text{jet axis} \\ \tau_1^b &: \quad \boxed{q_B^b = xP\,, \quad q_J^b = q + xP} \\ \tau_1^c &: \quad q_B^c = P\,, \quad && q_J^c = k\,, \end{aligned}$$



(b) τ_1^b

q_J is not the physical jet axis

-> sensitive to p_T of ISR

-> generalised p_T-dependent beam function

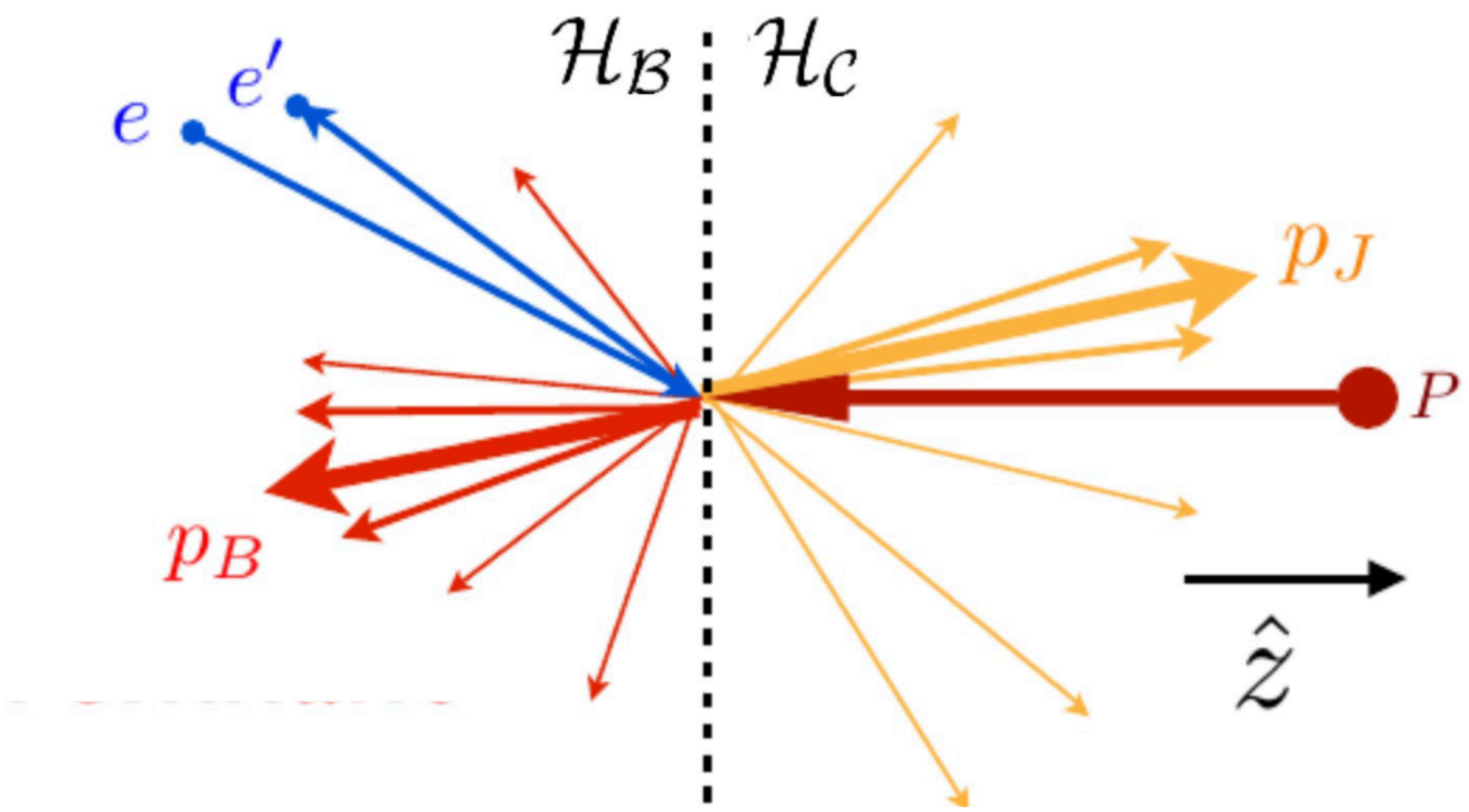
One-jettiness for DIS

3 possible definitions explored in [Kang, Lee, Stewart '13]

$$\tau_1 = \frac{2}{Q^2} \sum_{i \in X} \min\{q_B \cdot p_i, q_J \cdot p_i\}$$

$$\begin{array}{llll}
\tau_1^a : & q_B^a = xP\,, & q_J^a = \text{jet axis} \\
\tau_1^b : & \boxed{q_B^b = xP\,, & q_J^b = q + xP} \\
\tau_1^c : & q_B^c = P\,, & q_J^c = k\,,
\end{array}$$

Breit frame



beam and jet regions are back-to-back hemispheres in the Breit frame

$$\tau_1^b \overset{\text{Breit}}{=} 1 - \frac{2}{Q} \sum_{i \in \mathcal{H}_J^b} p_{z\,i} \equiv \tau_Q$$

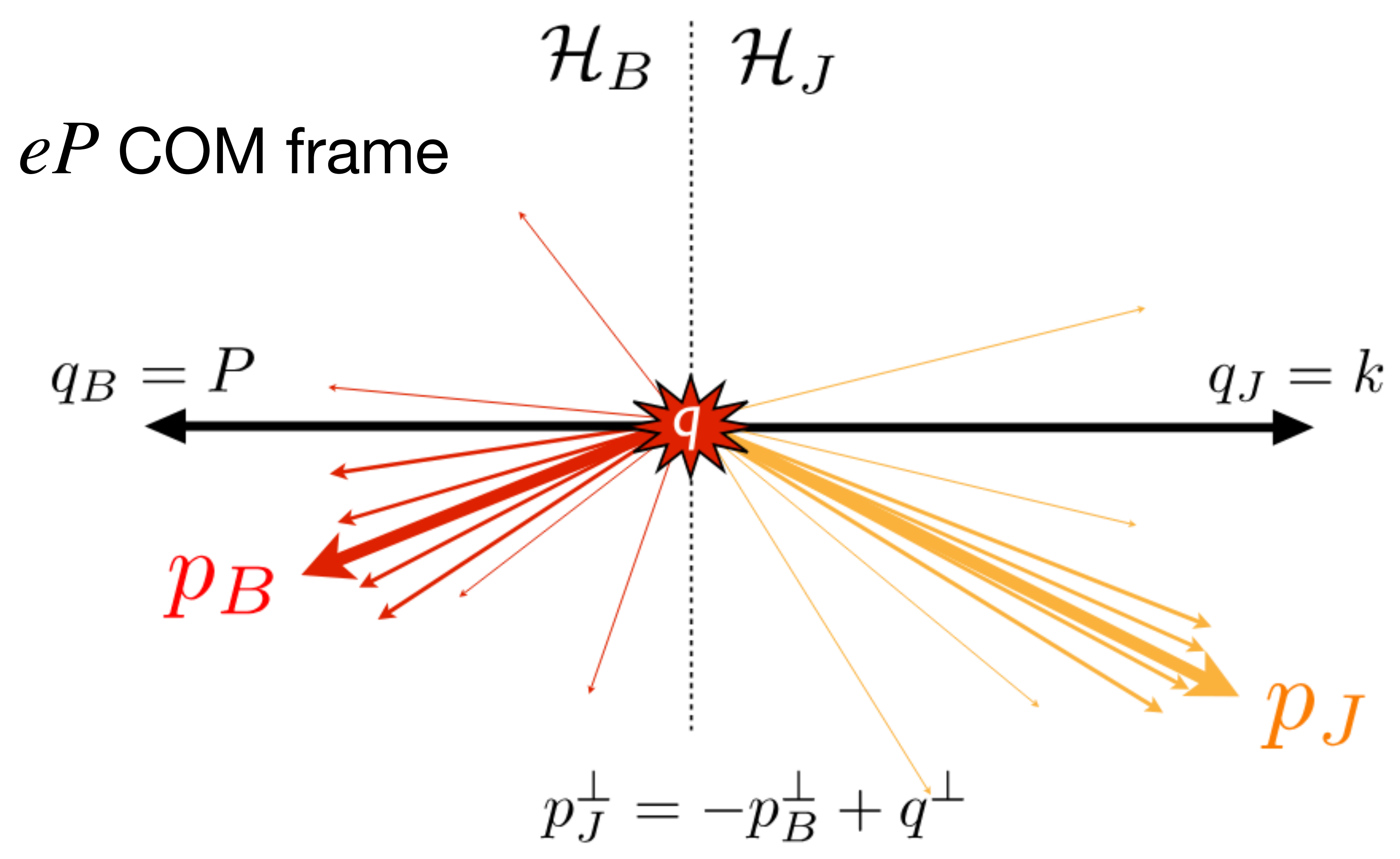
equivalent to DIS thrust [Antonelli, Dasgupta, Salam '99]

One-jettiness for DIS

3 possible definitions explored in [Kang, Lee, Stewart ’13]

$$\tau_1 = \frac{2}{Q^2} \sum_{i \in X} \min\{q_B \cdot p_i, q_J \cdot p_i\}$$

$$\begin{aligned}
\tau_1^a &: \quad q_B^a = xP\,, \quad && q_J^a = \text{jet axis} \\
\tau_1^b &: \quad q_B^b = xP\,, \quad && q_J^b = q + xP \\
\tau_1^c &: \quad \boxed{q_B^c = P\,, \quad q_J^c = k\,,}
\end{aligned}$$



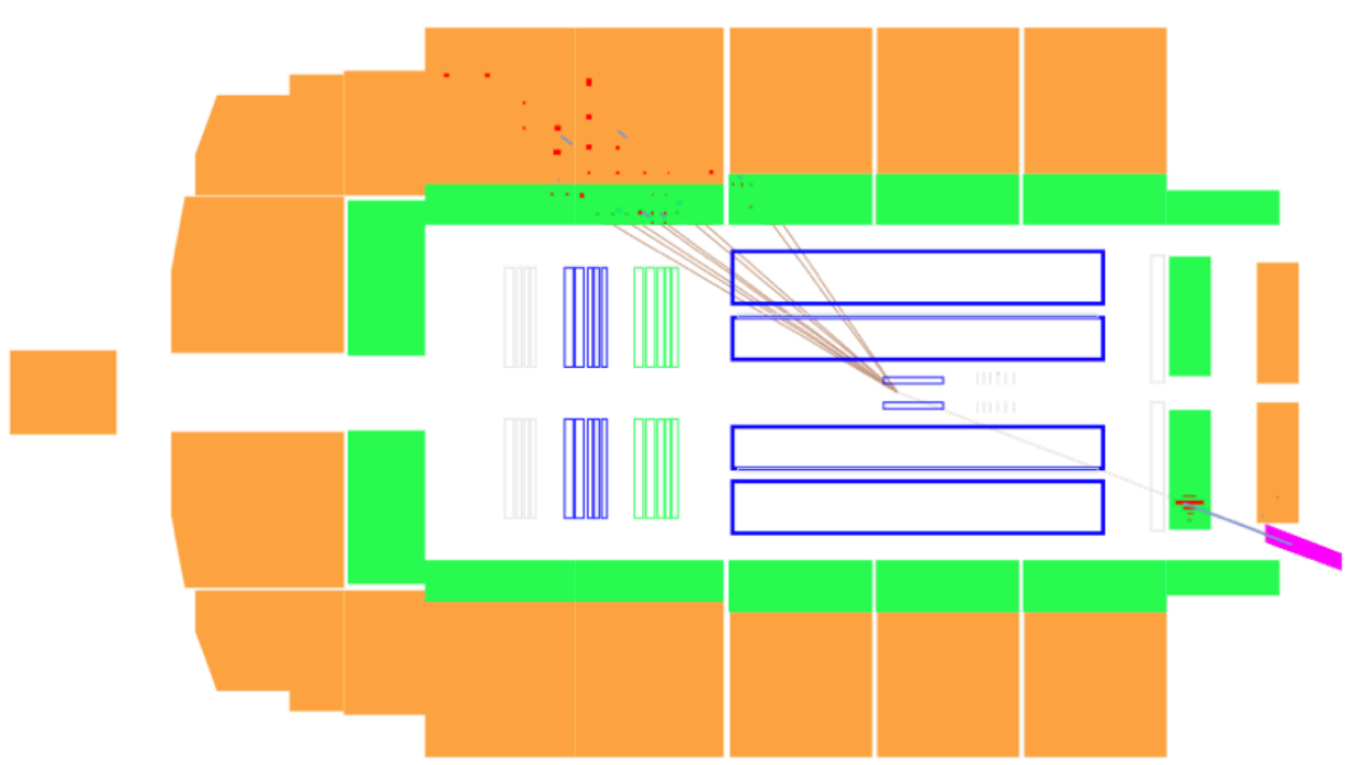
(c) τ_1^c

back-to-back hemispheres in COM frame

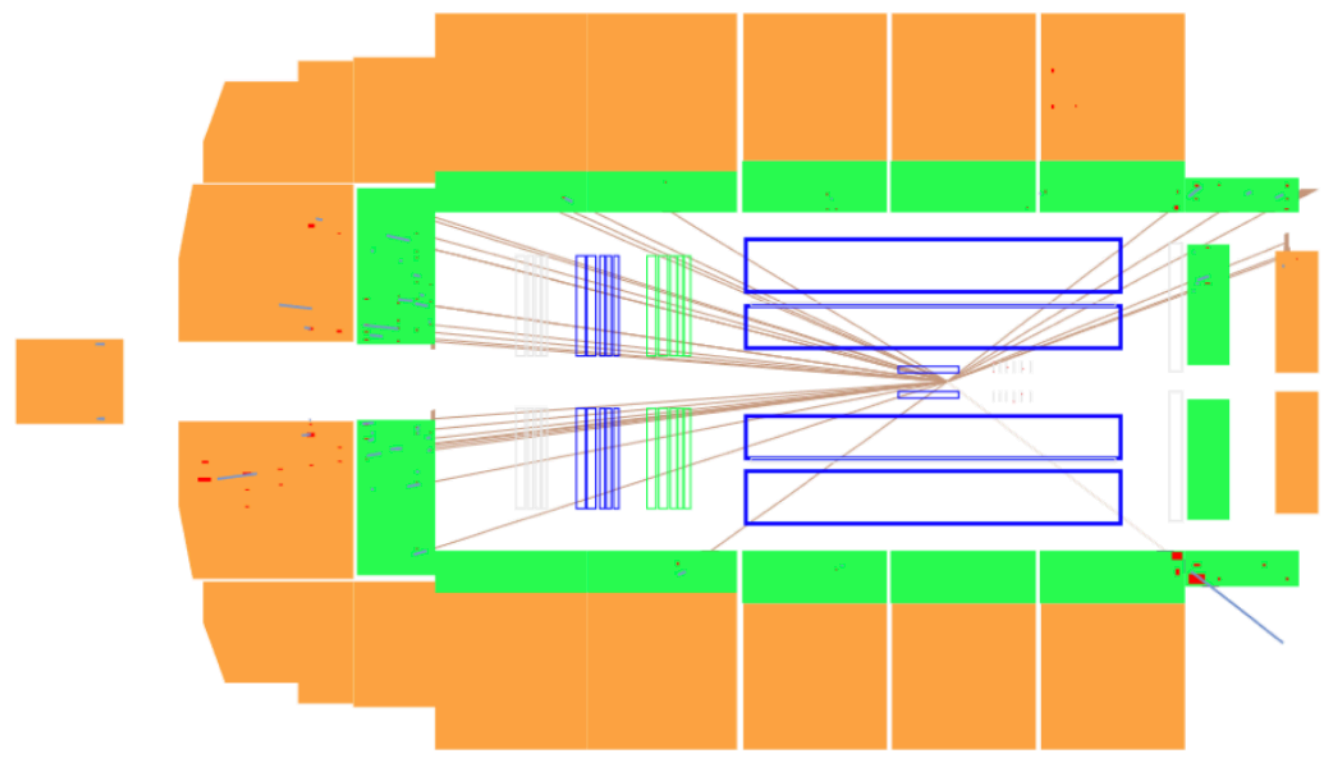
However, factorisation theorem requires a jet in a direction fairly close to the initial electron direction (= q_J axis) e.g. inelasticity $y \sim 1$

τ_1^b version recently measured at H1 [H1 '24]

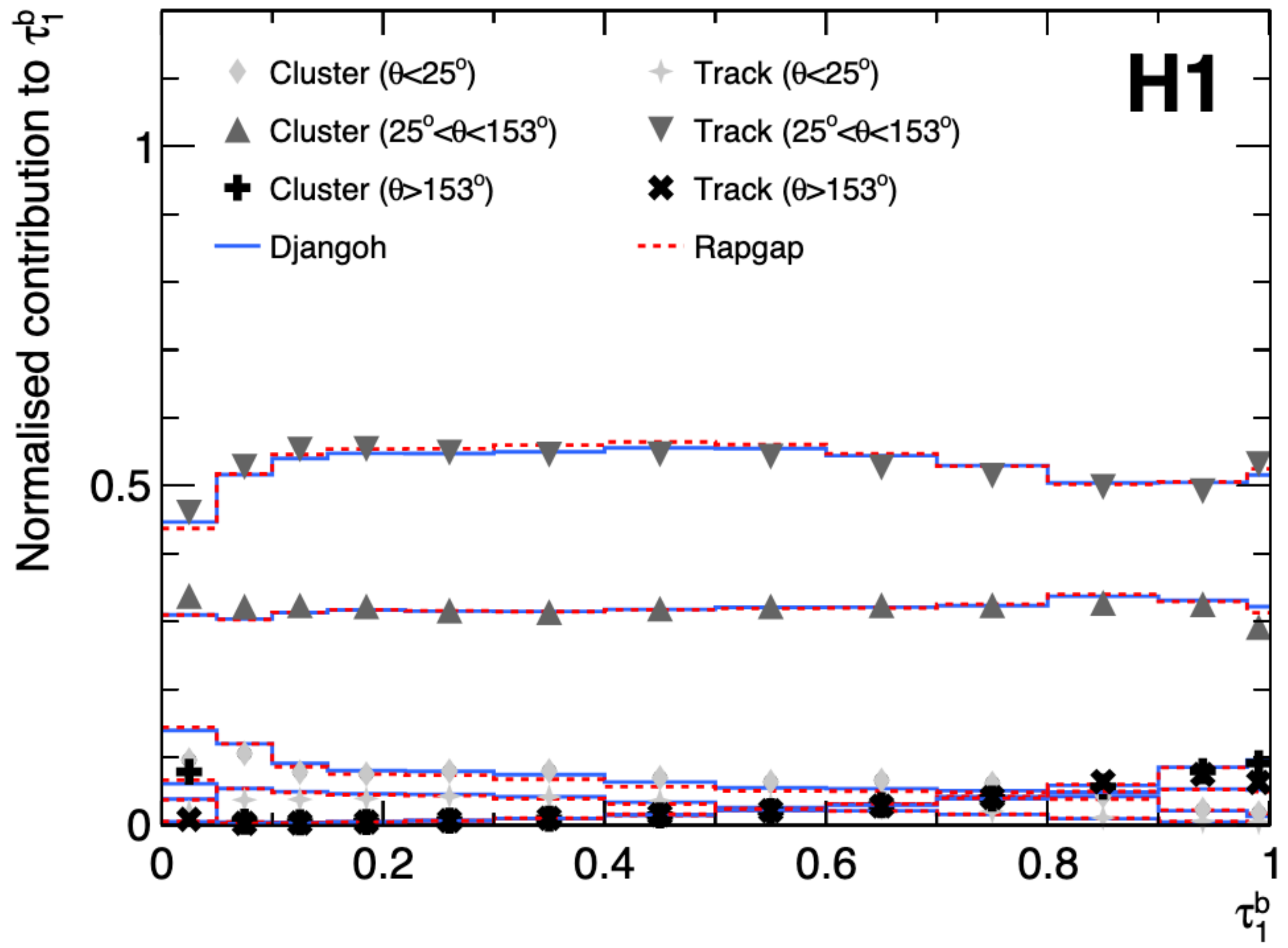
Visualisation of the 1-jettiness with event displays

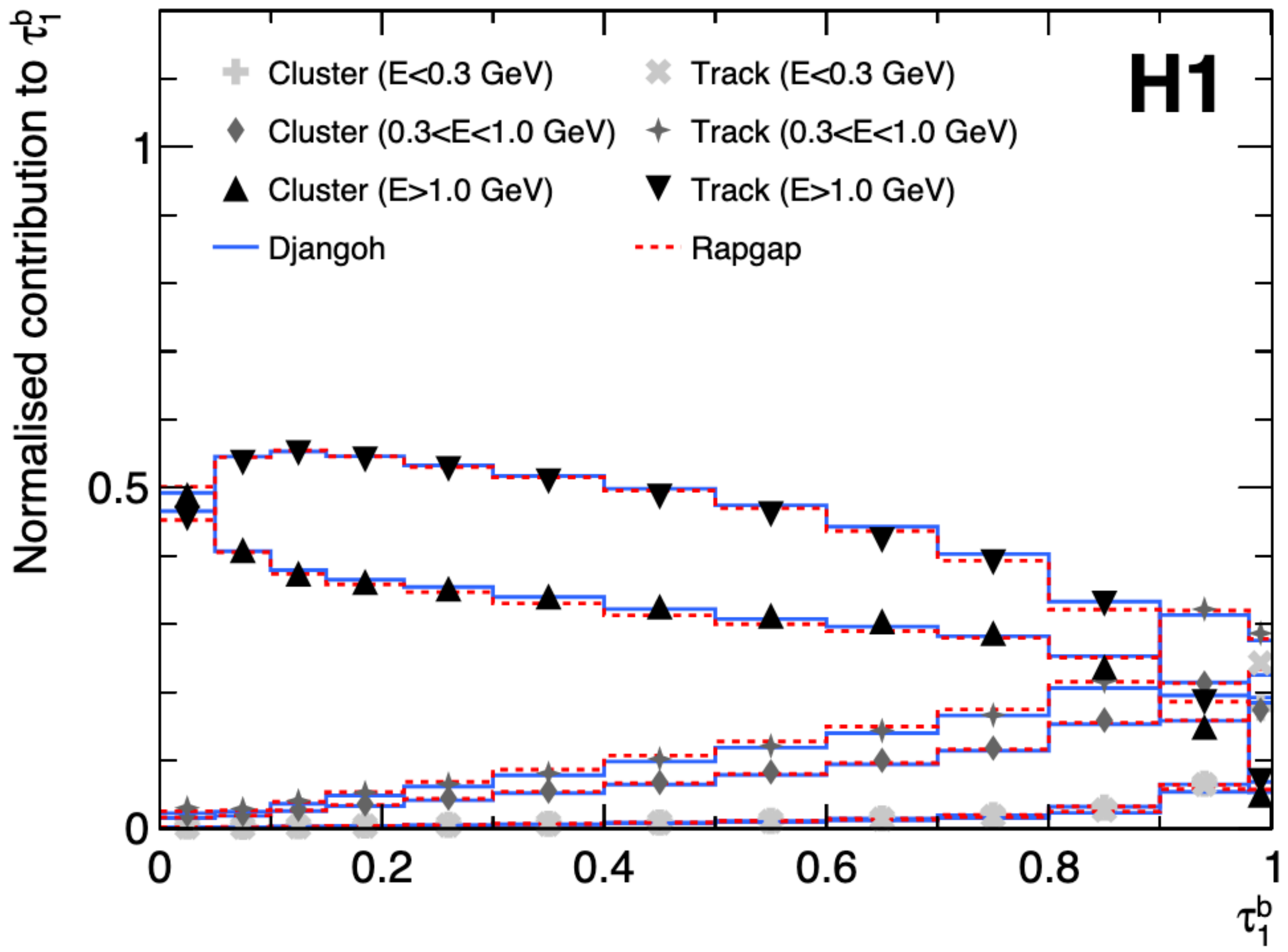


- DIS 1-jet configuration
- Most HFS particles collinear to scattered parton → Small τ_1^b

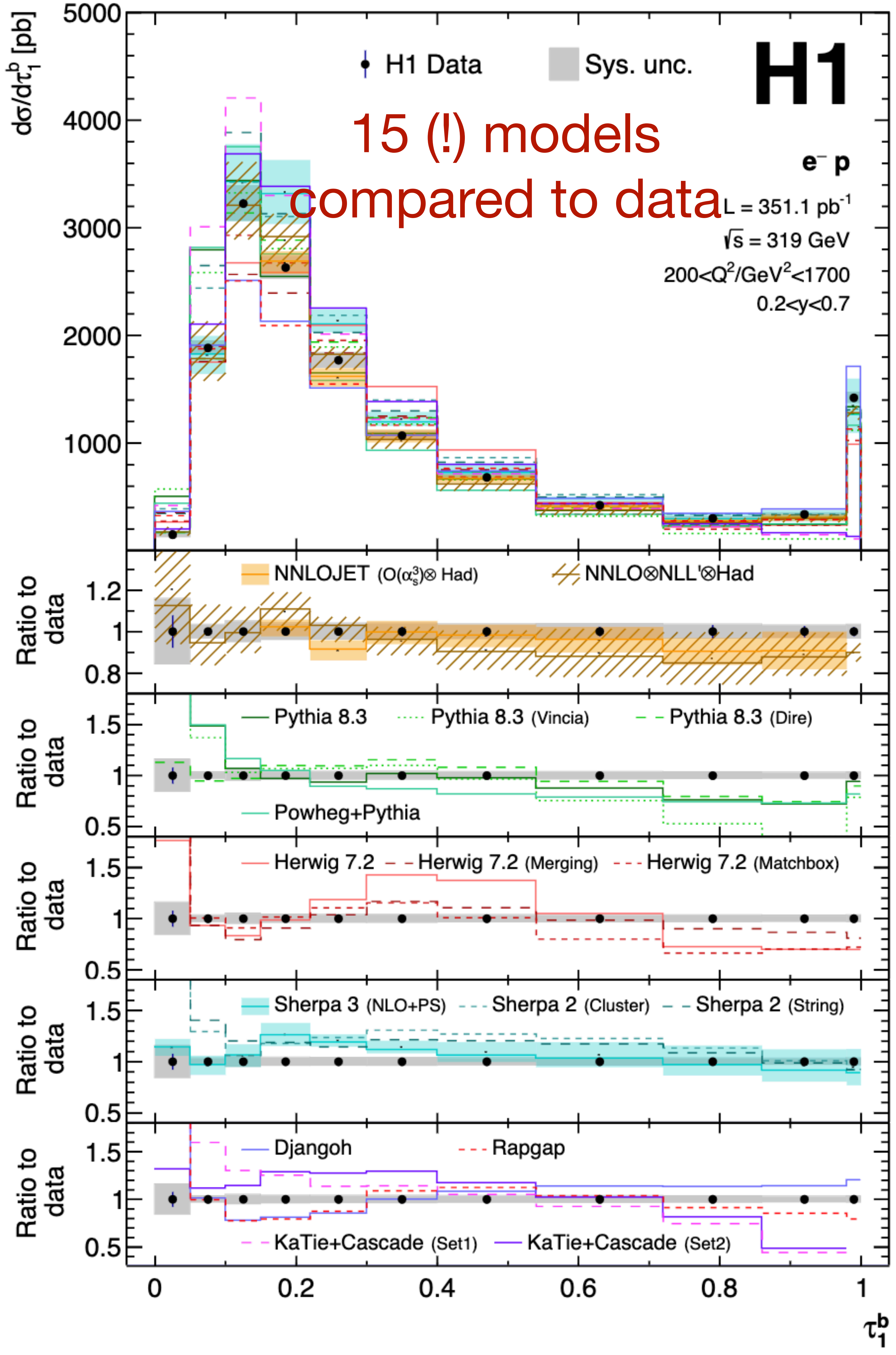


- Dijet event
- More and larger contributions to the sum over the HFS → Large τ_1^b



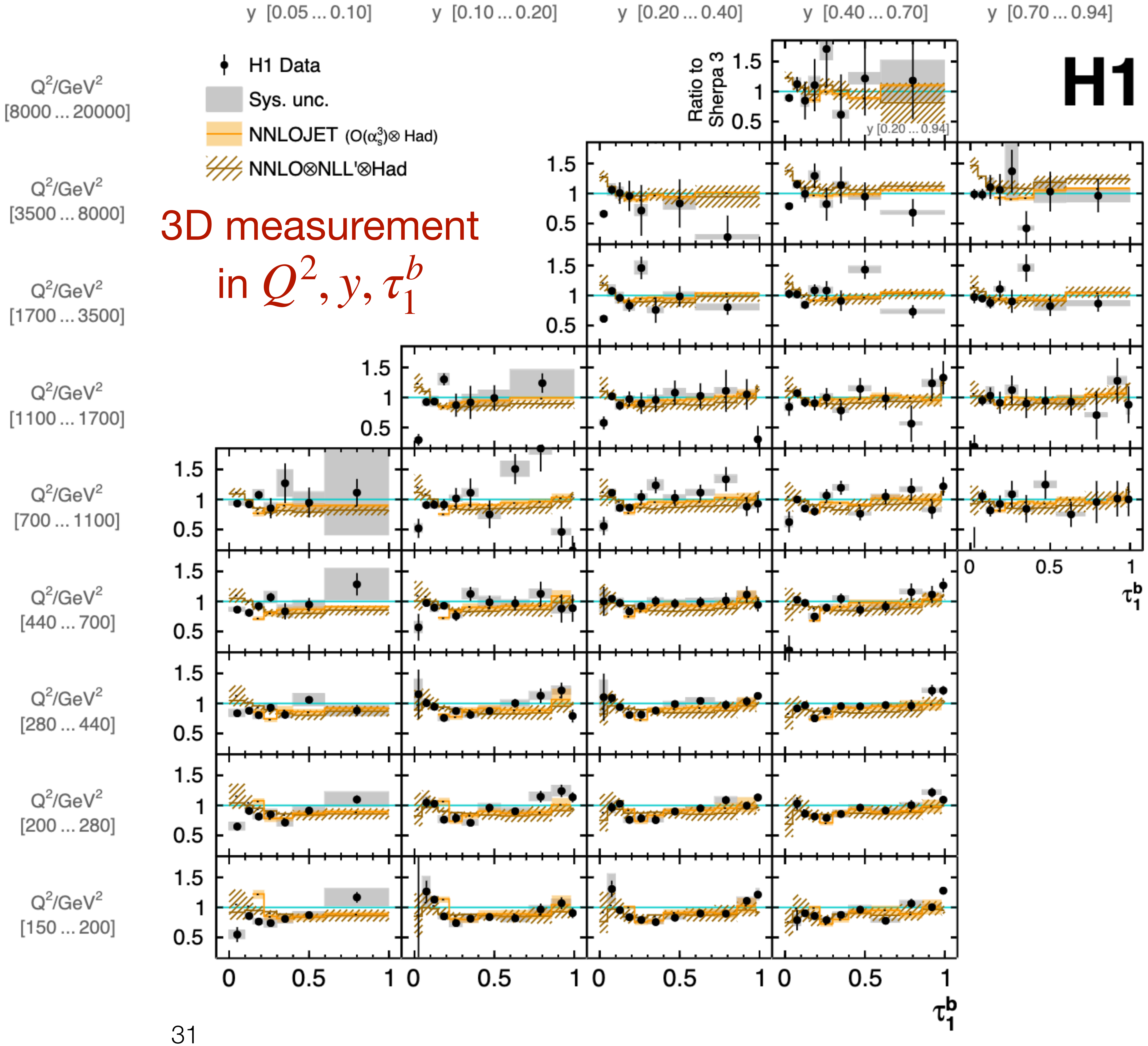


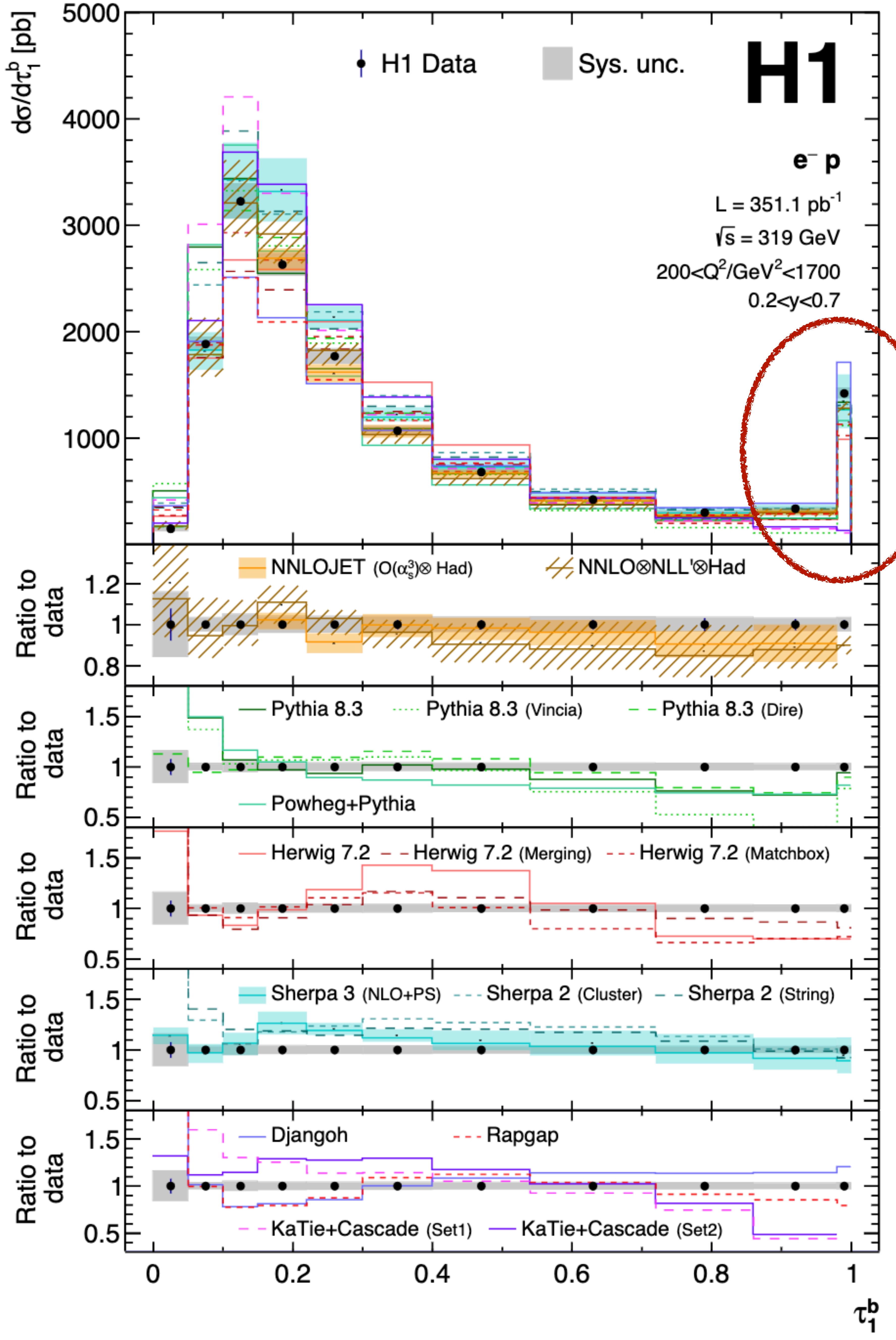
τ_1^b is a global observable, but well-measured particles provide the dominant contribution



15 (!) models compared to data

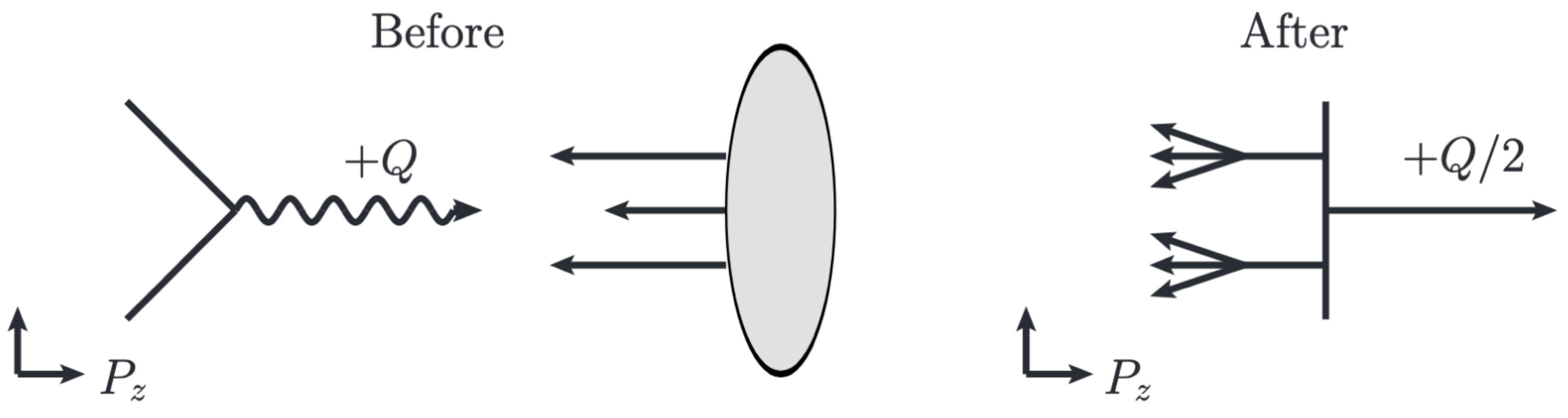
3D measurement in Q^2, y, τ_1^b



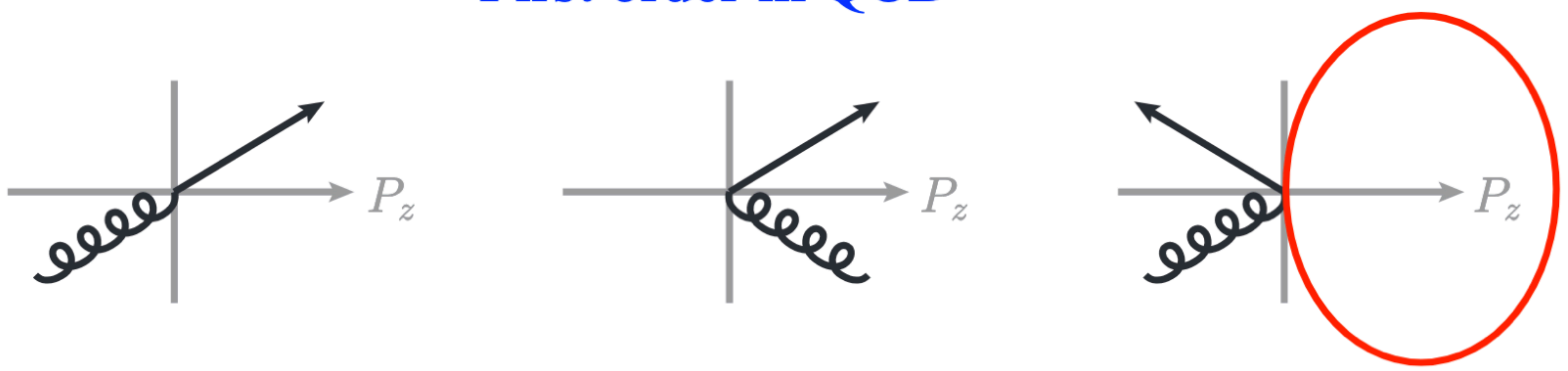


Delta-distribution at $\tau_1^b = 1$: events with empty current hemisphere

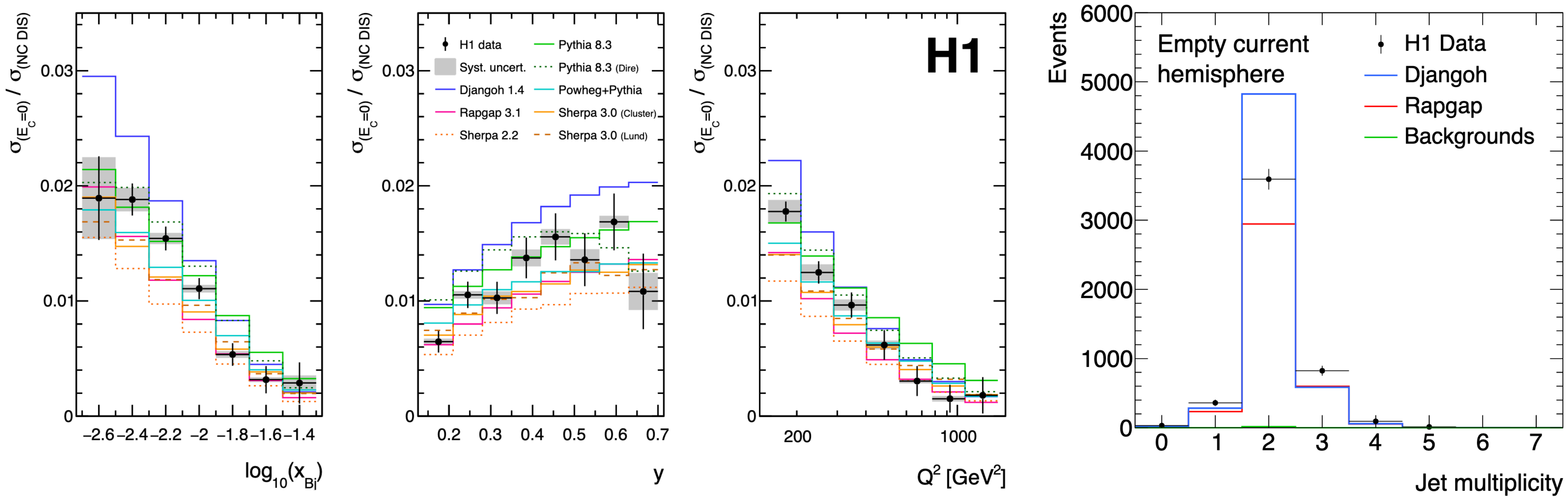
Leading order: parton model



First order in QCD



Observation of NC DIS events with an empty hemisphere in the Breit frame [H1 ’24]

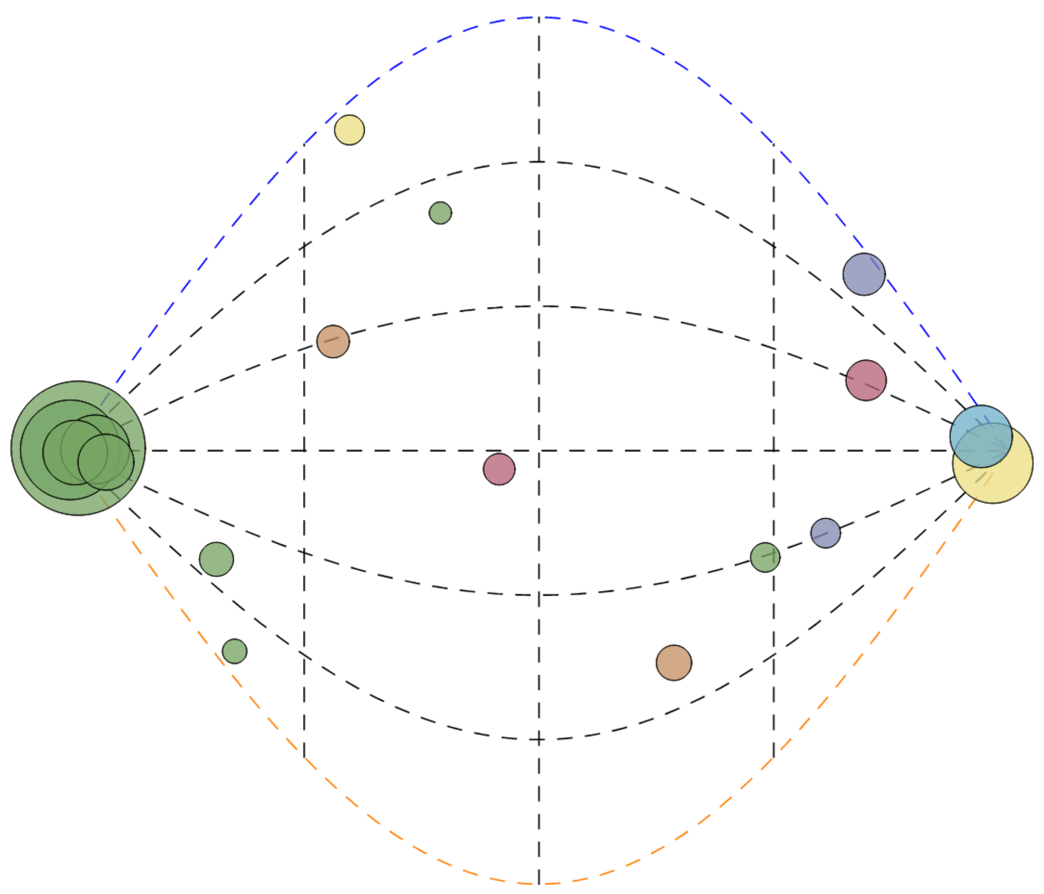


Fraction of empty current hemisphere: around 1% of the events, with a dependence on x_B, y, Q^2,

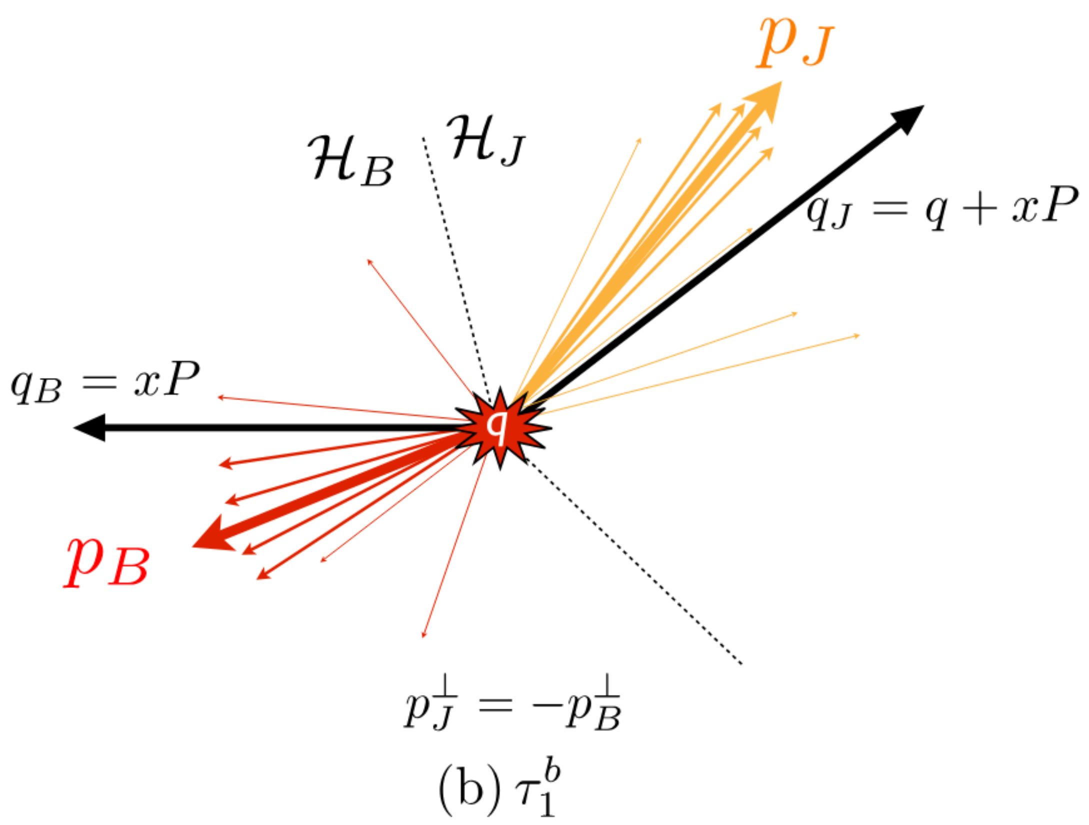
Predominantly 2-jet events (k_t algorithm with $R = 1$)

Outlook

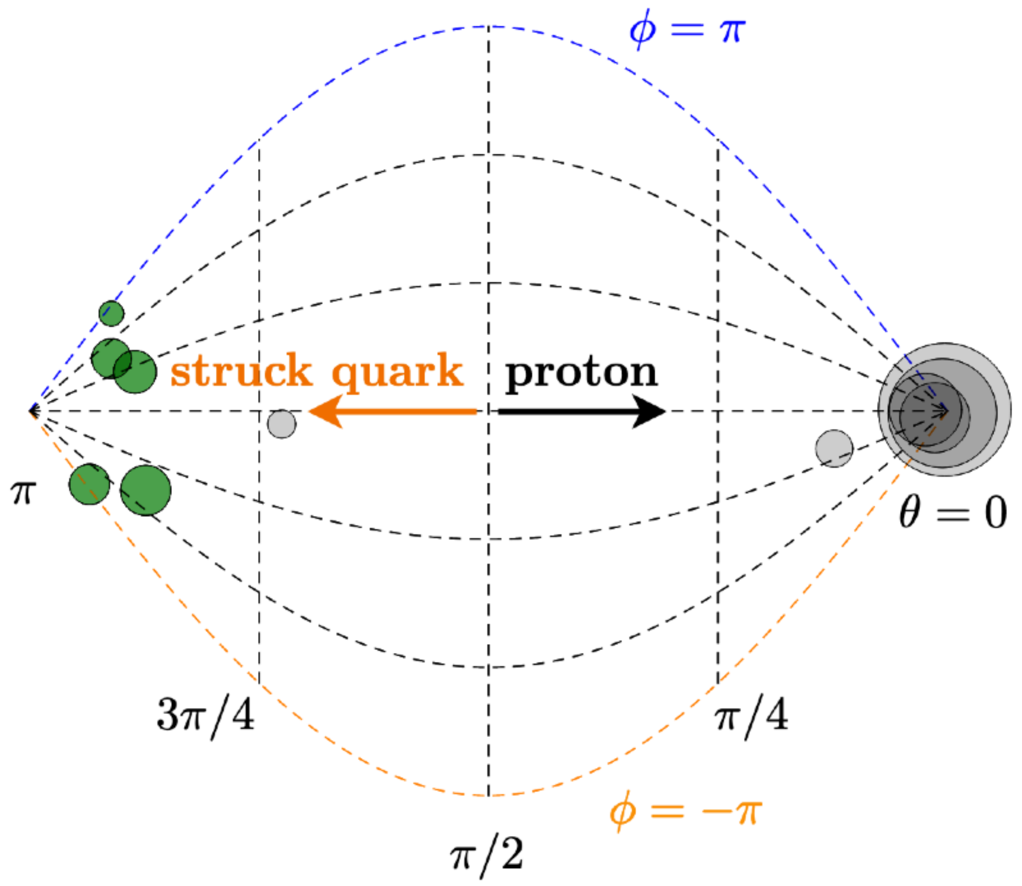
Jet algorithms



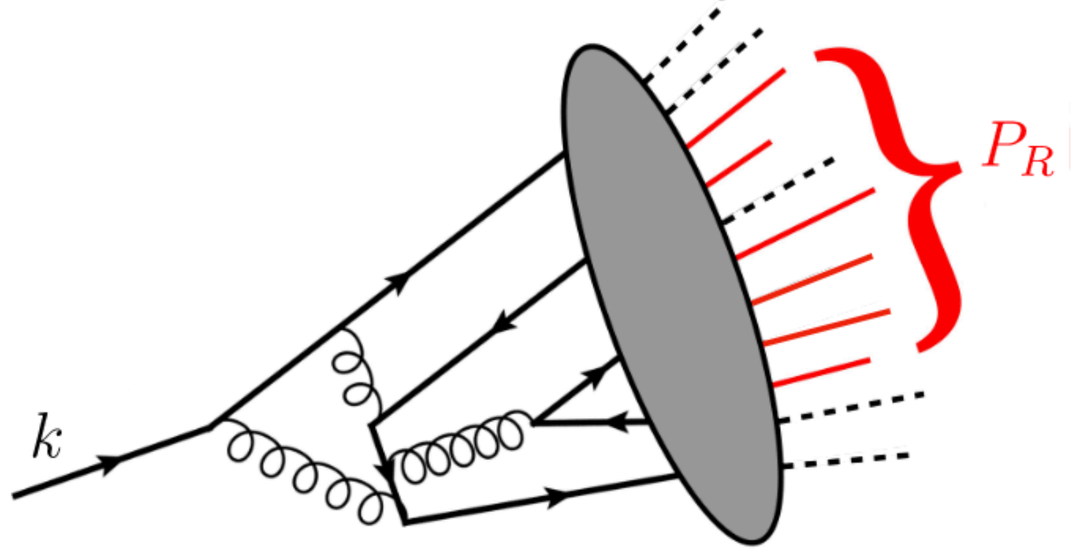
Event shapes



Grooming



Track functions



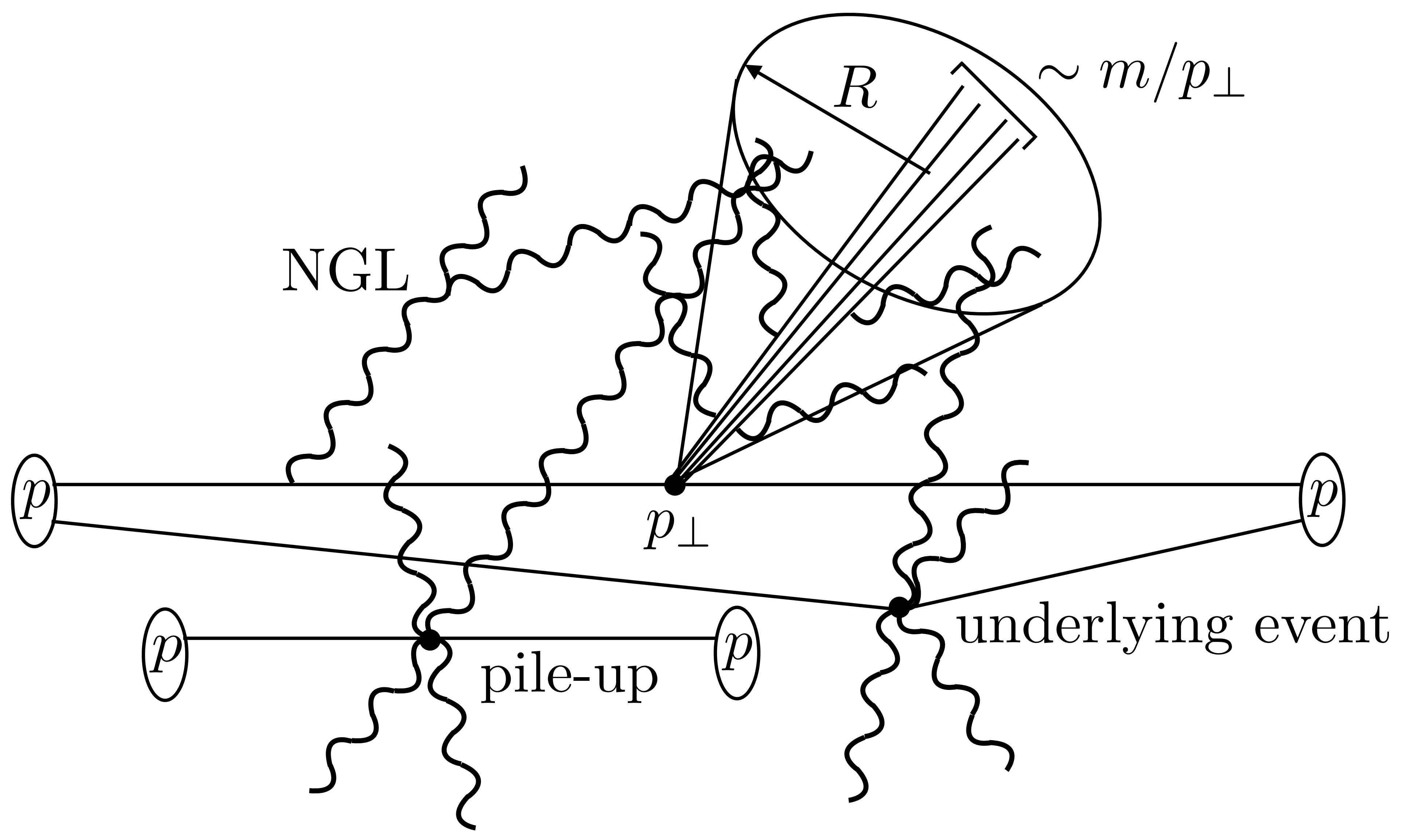
How to exploit the expertise gained from the LHC in DIS context?

Personal selection of results that are representative for on-going progress.
Apologies for any relevant omission of references.

Jet substructure: grooming

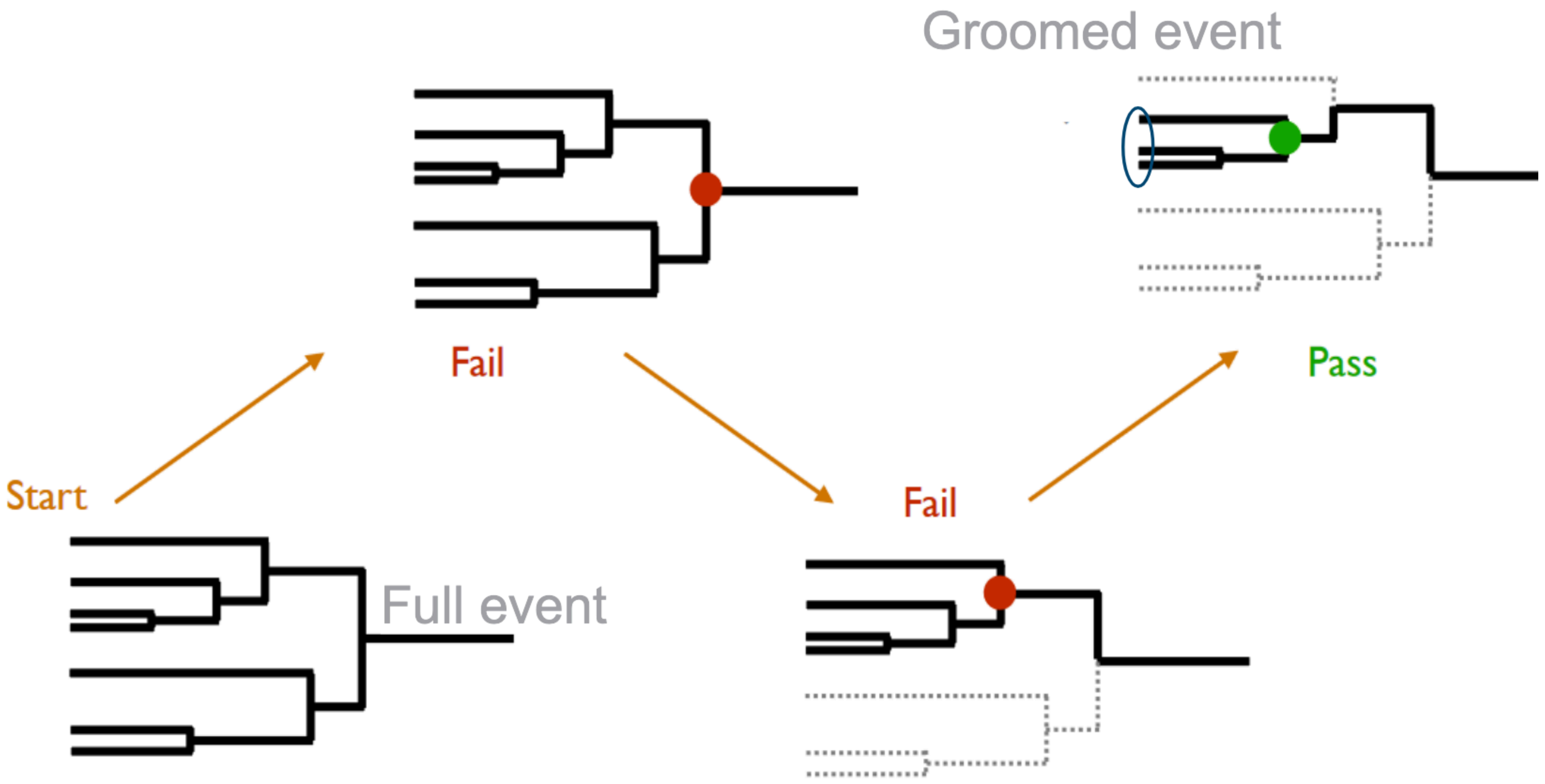
Grooming in LHC context originally designed to remove background contamination in jet, while keeping the bulk of perturbative radiation

Non-global logarithms (NGLs), underlying events (UE), pile-up (PU) mostly appear as <u>soft wide-angle radiation</u> inside the jet



Grooming with Soft Drop

[Larkoski, Marzani, Soyez, Thaler ’14]



When $\beta \geq 0$, soft (large-angle) emissions are removed from the jet

- Recluster the jet with C/A algorithm i.e. angular-ordered history

- At each declustering step $(i+j) \to i+j$, check the condition:

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

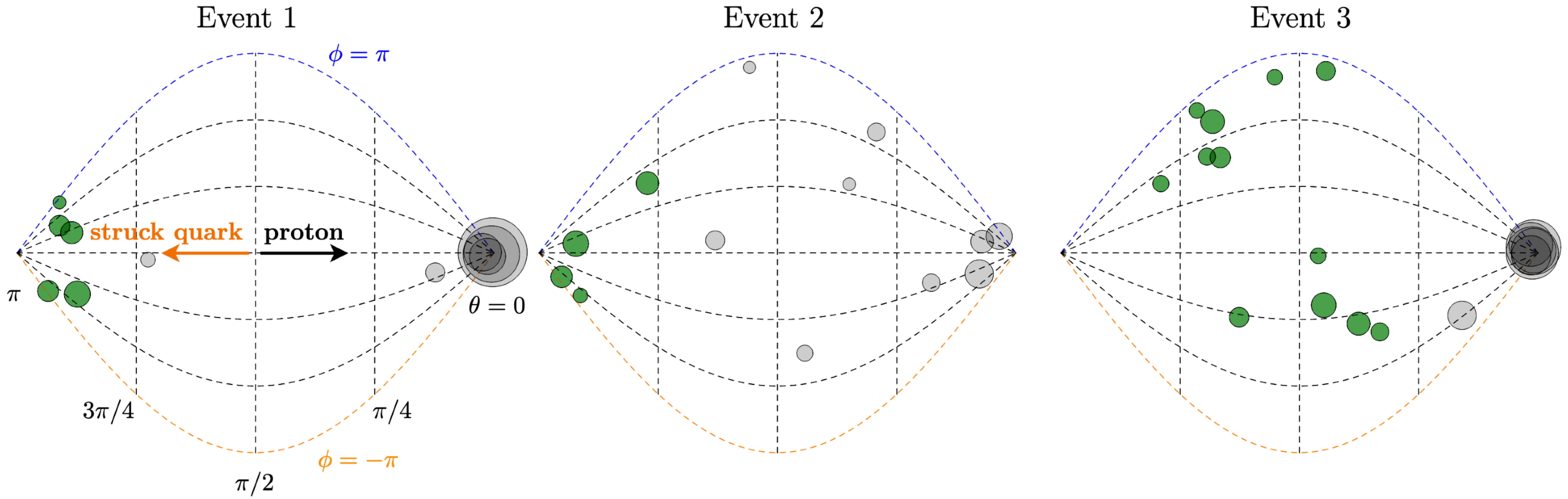
- If satisfied, declare $(i+j)$ as the groomed jet. Otherwise, iterate on the subjet with the largest p_t

Adapt Soft Drop to DIS scenario

[Makris '21]

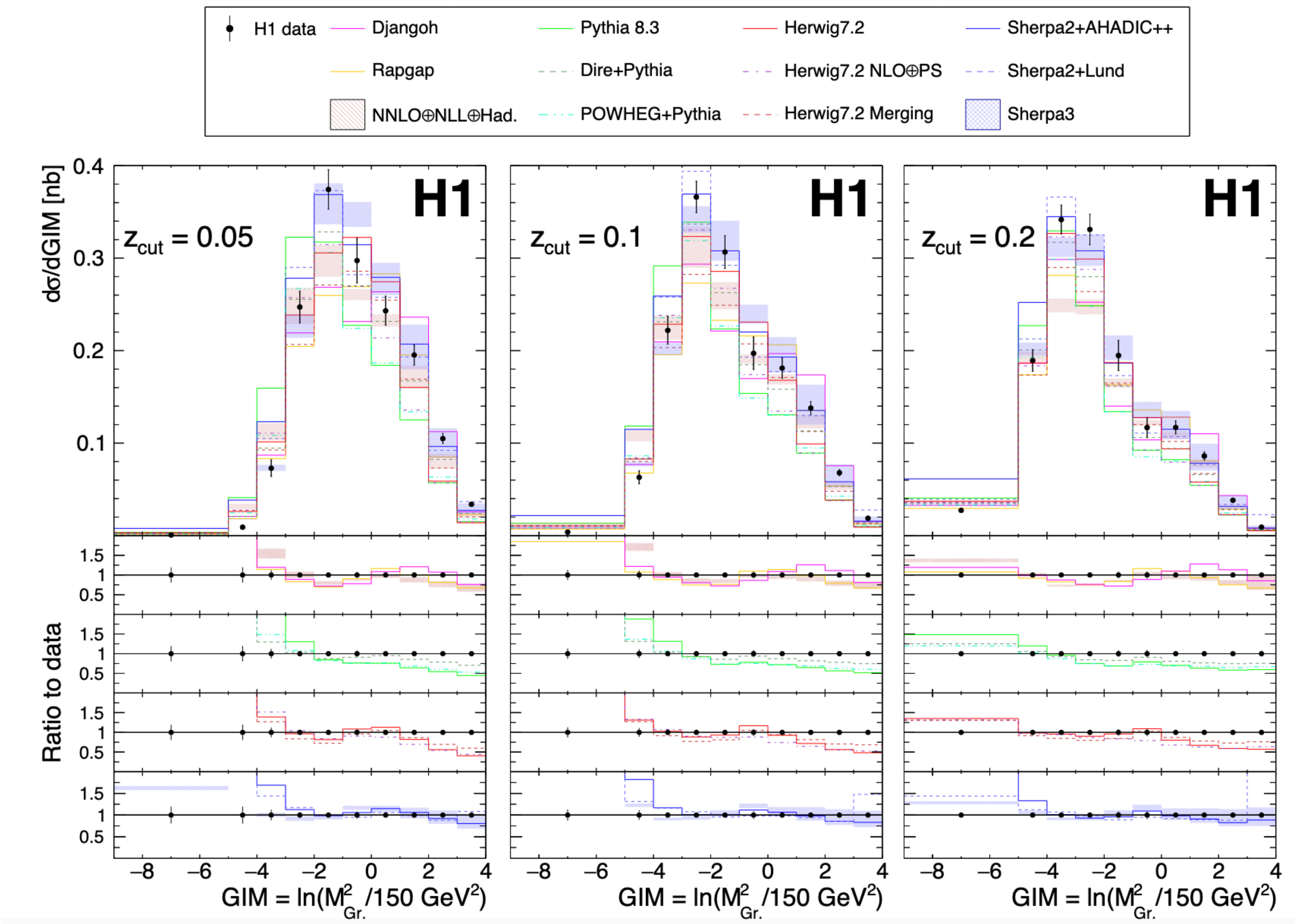
Recluster with Centauro distance and modify Soft Drop condition:

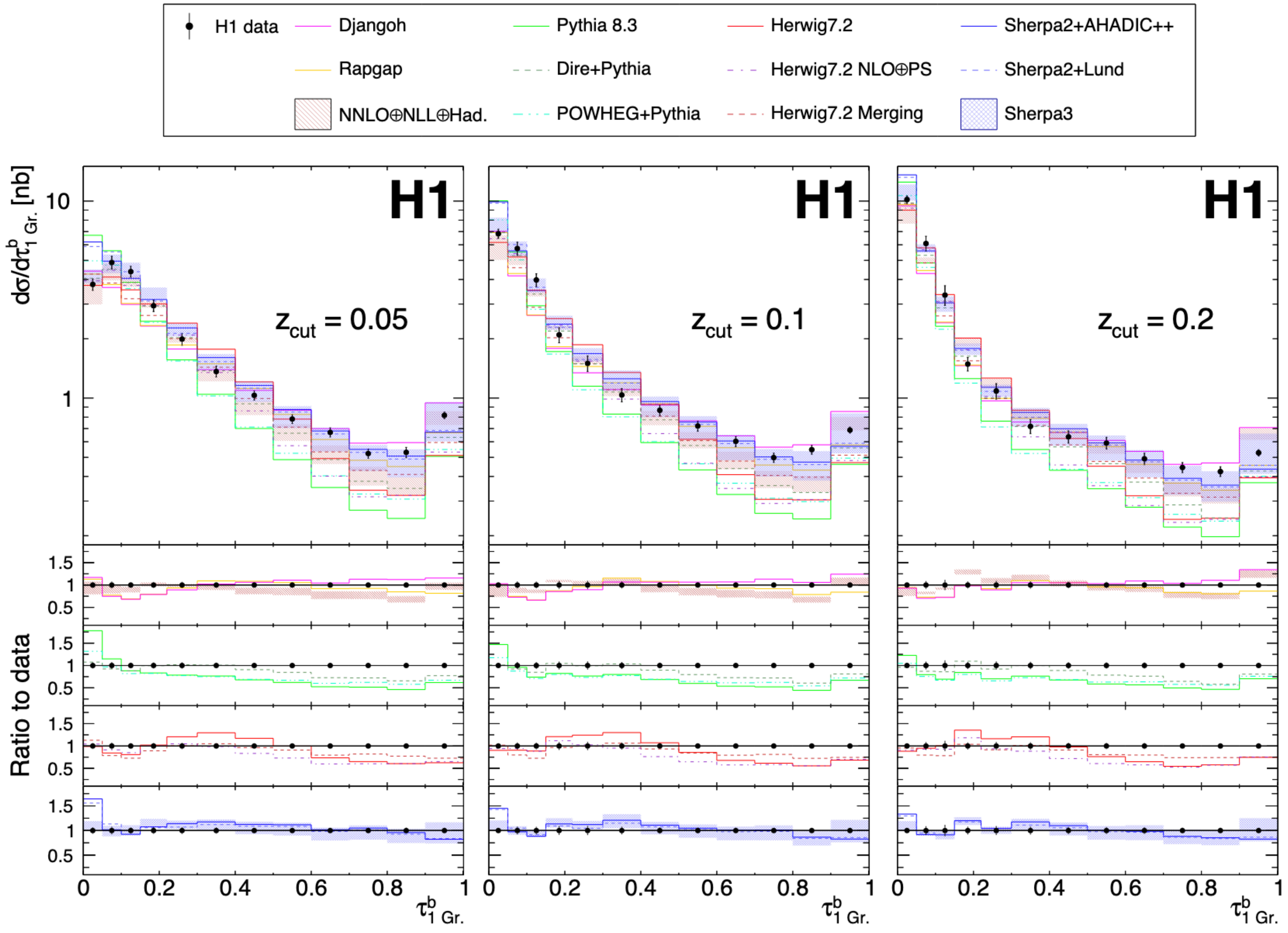
$$\frac{\min(p_{t,i}, p_{t,j})}{p_{t,i} + p_{t,j}} > z_{\text{cut}} \left(\frac{\Delta R_{ij}}{R} \right)^{\beta} \longrightarrow \frac{\min(z_i, z_j)}{z_i + z_j} > z_{\text{cut}} \qquad z_i = \frac{P \cdot p_i}{P \cdot q}$$



Only particles either hard or collinear to the direction of the current are kept

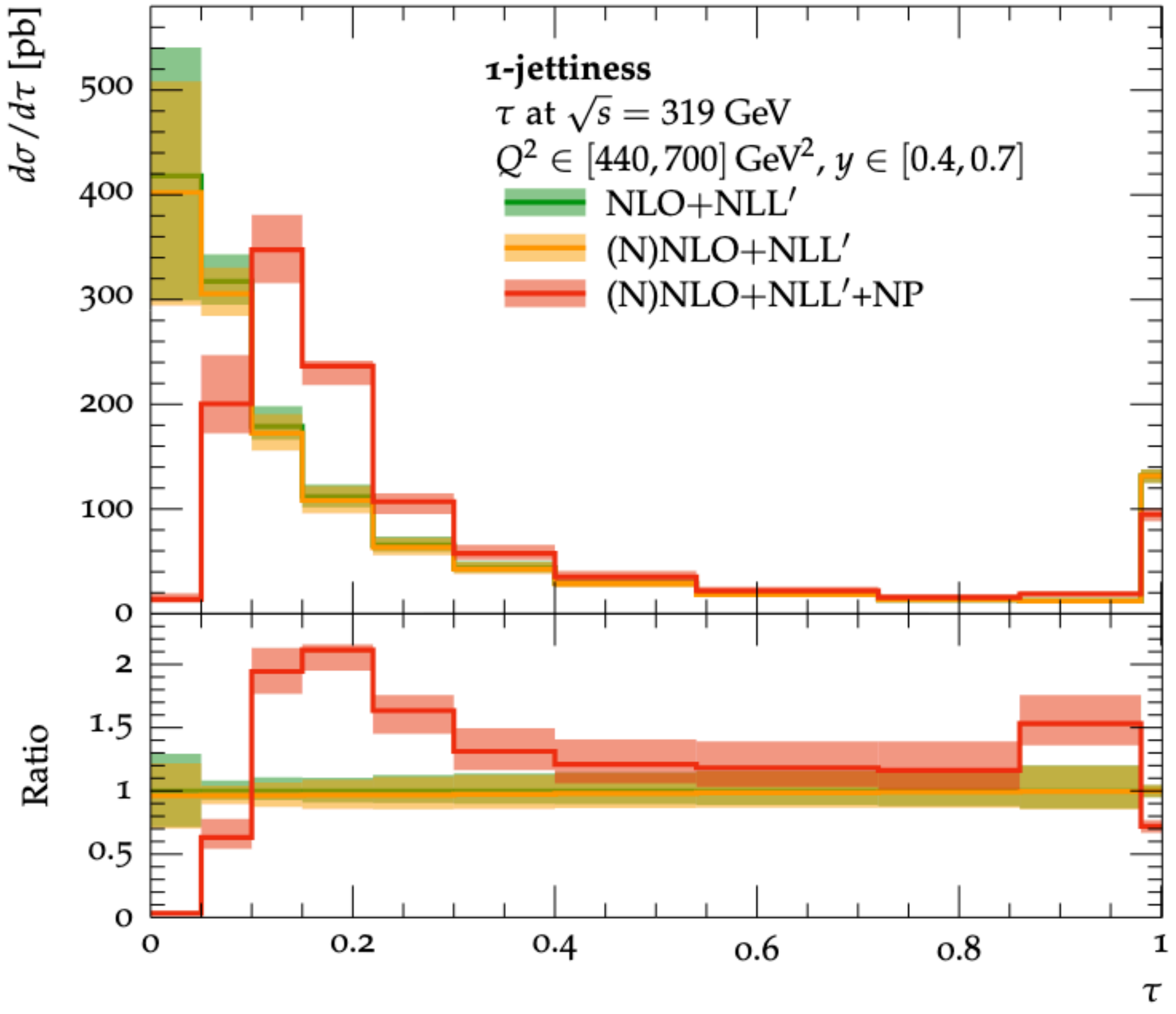
Measurements of groomed invariant mass and groomed one-jettiness [H1 ’24]

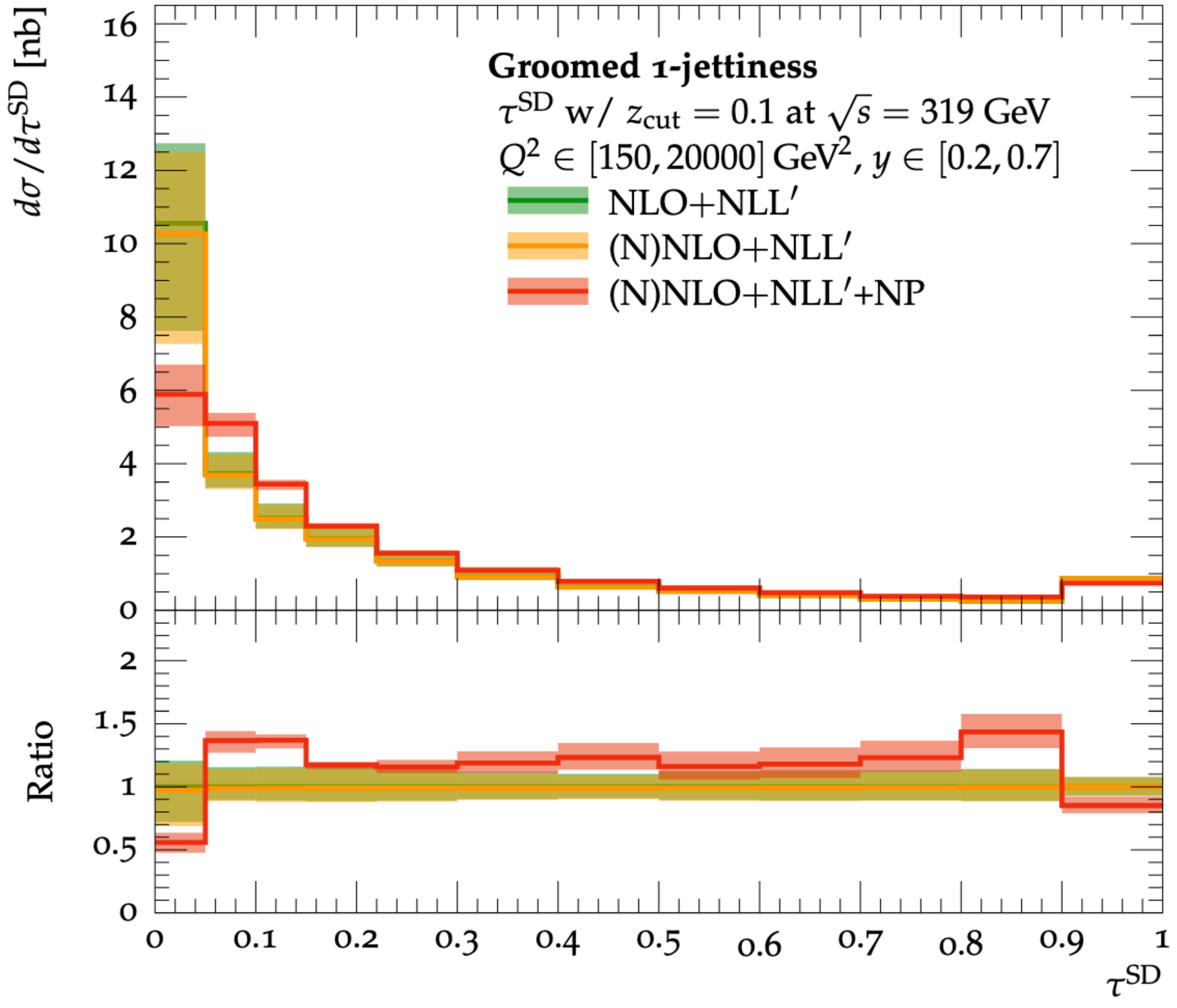




New calculations: [Knobbe, Reichelt, Schumann ’23]

- NNLO+NLL+Had, resummation with CAESAR formalism
- “SHERPA3”: MEPS@NLO with $ep \to e + 1{,}2j$ @ NLO + $3{,}4j$ @ LO





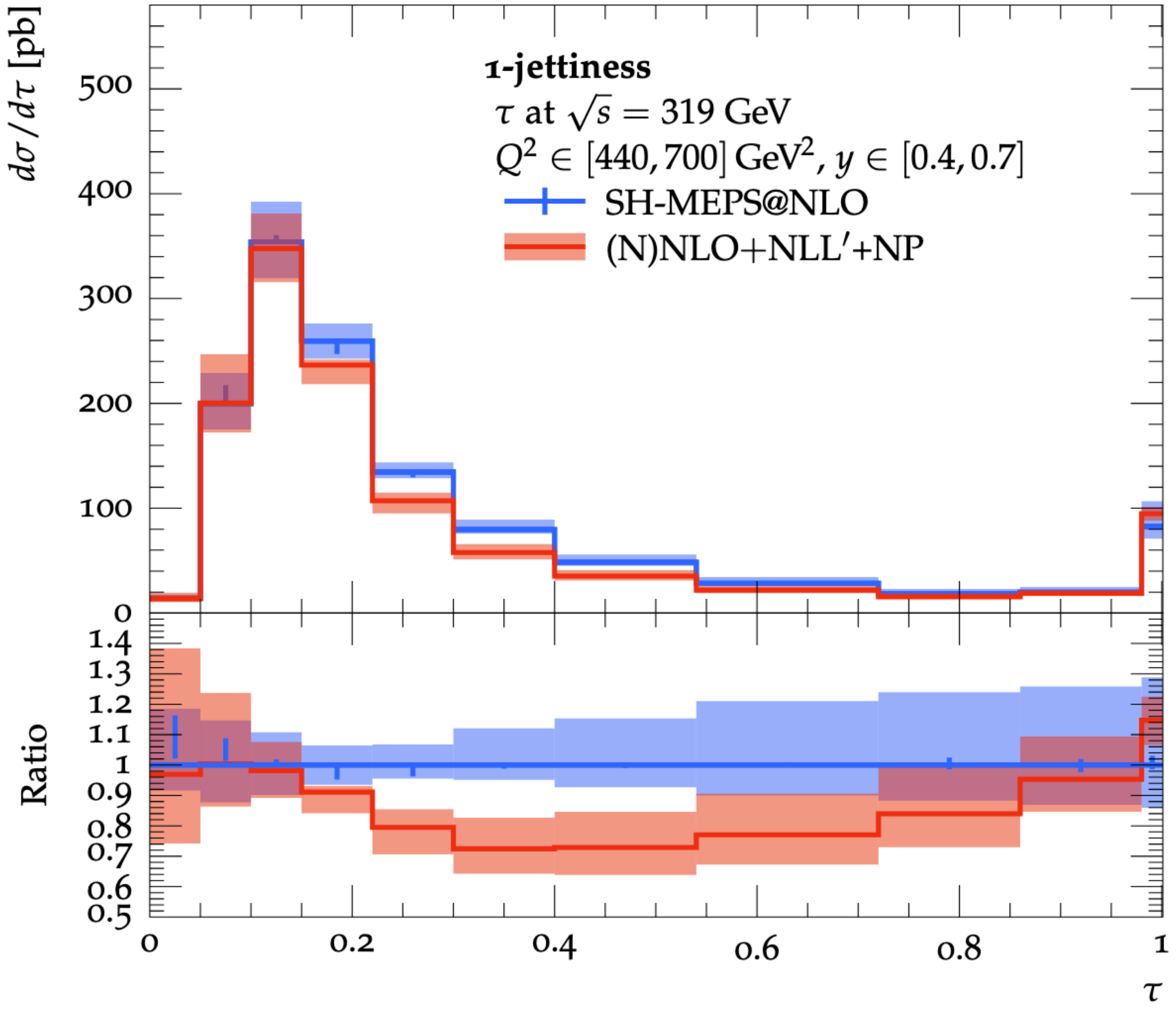
Benefits of grooming

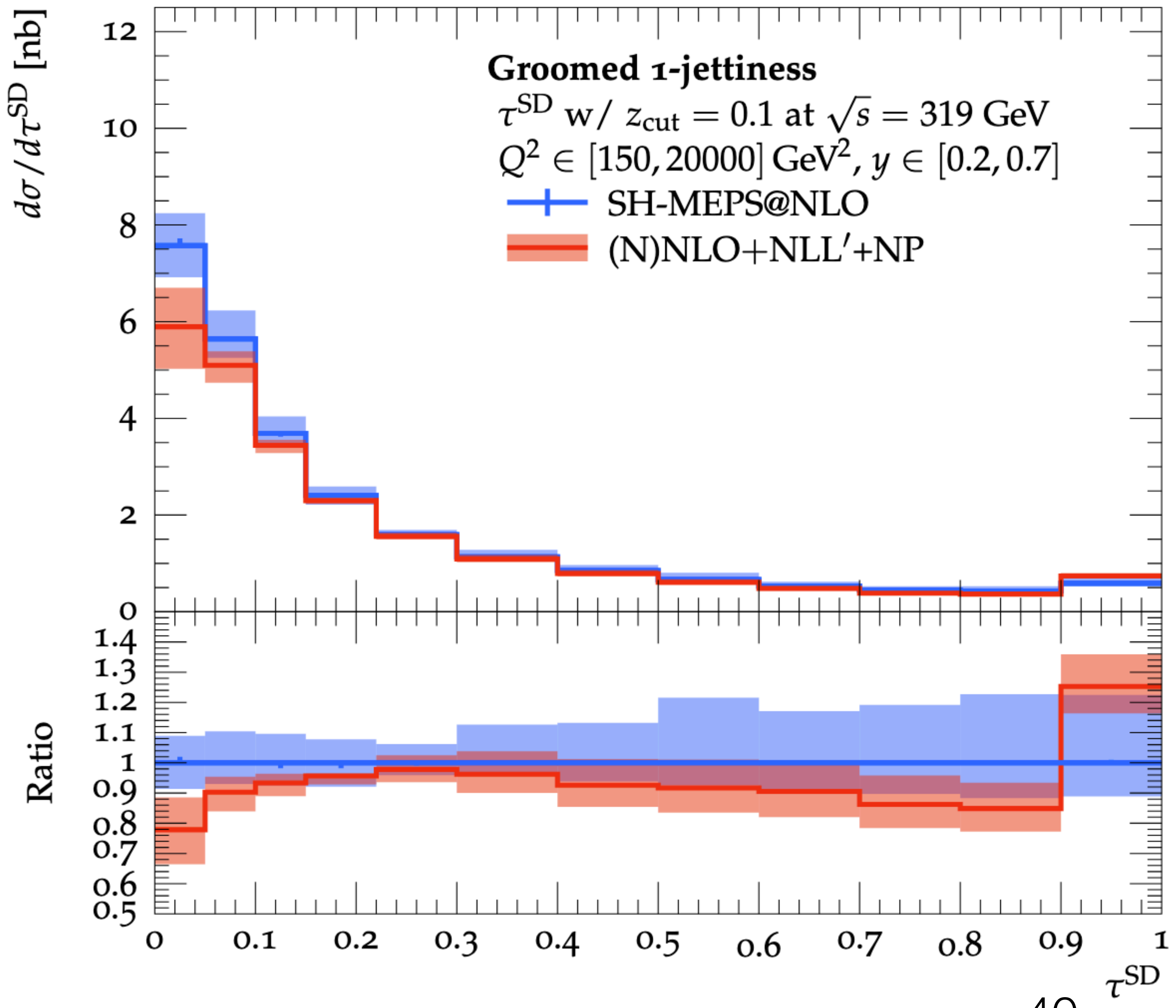
Reduction of non-perturbative corrections

Two calculations compatible within uncertainties

Better agreement between parton-level and detector-level

Less dependence on acceptance cuts





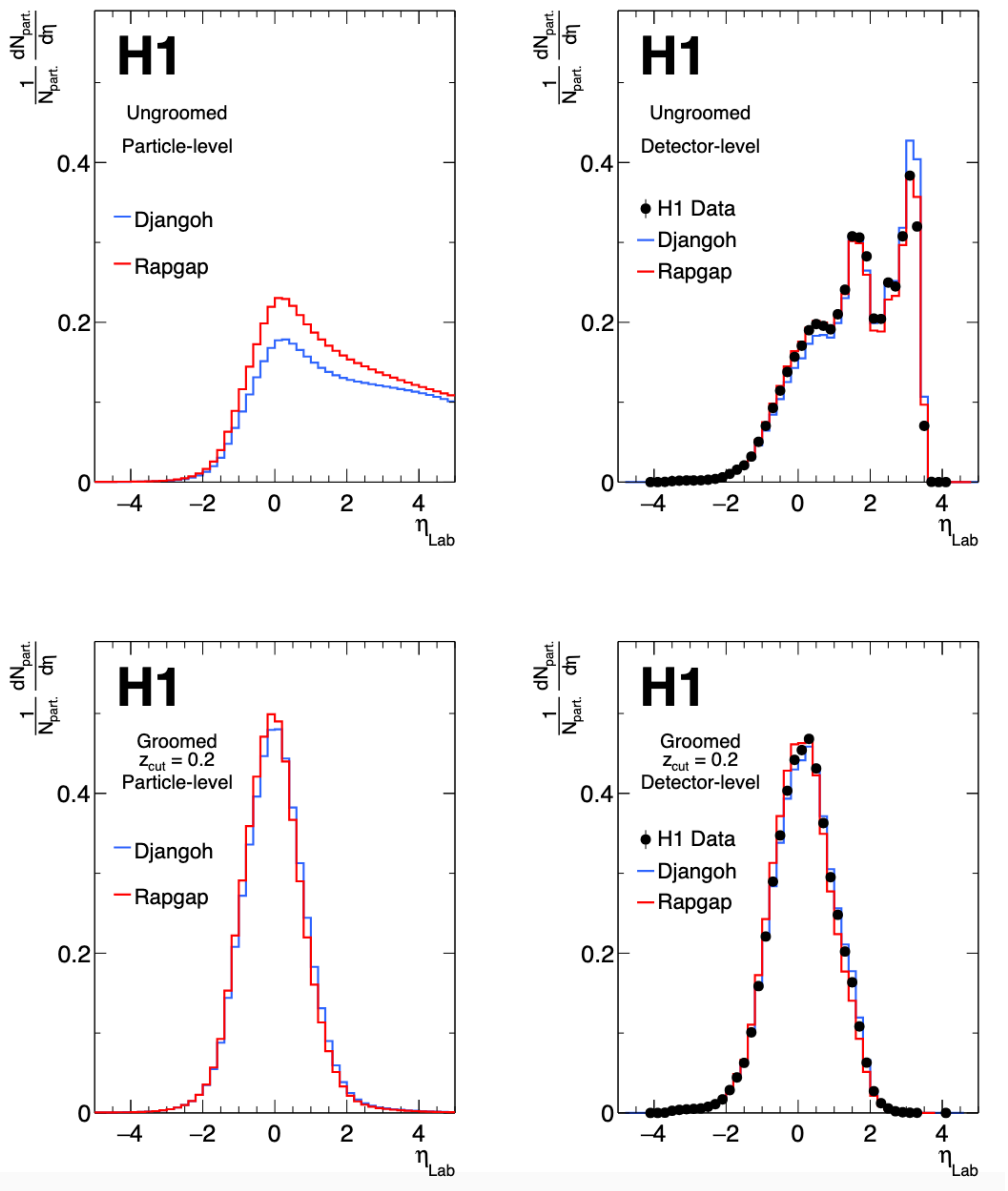
Benefits of grooming

Reduction of
non-perturbative corrections

Two calculations compatible
within uncertainties

Better agreement between
parton-level and detector-level

Less dependence
on acceptance cuts



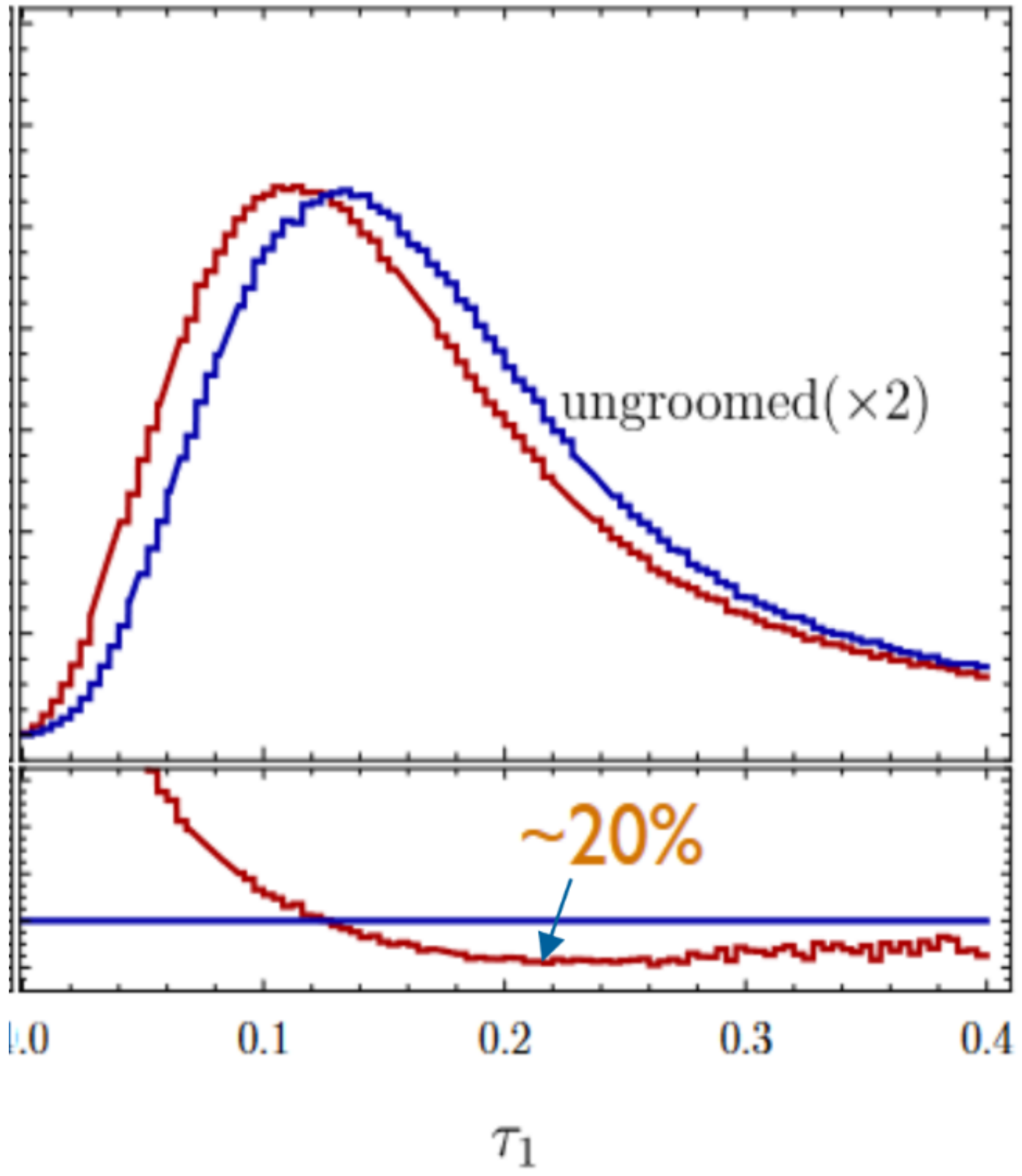
Benefits of grooming

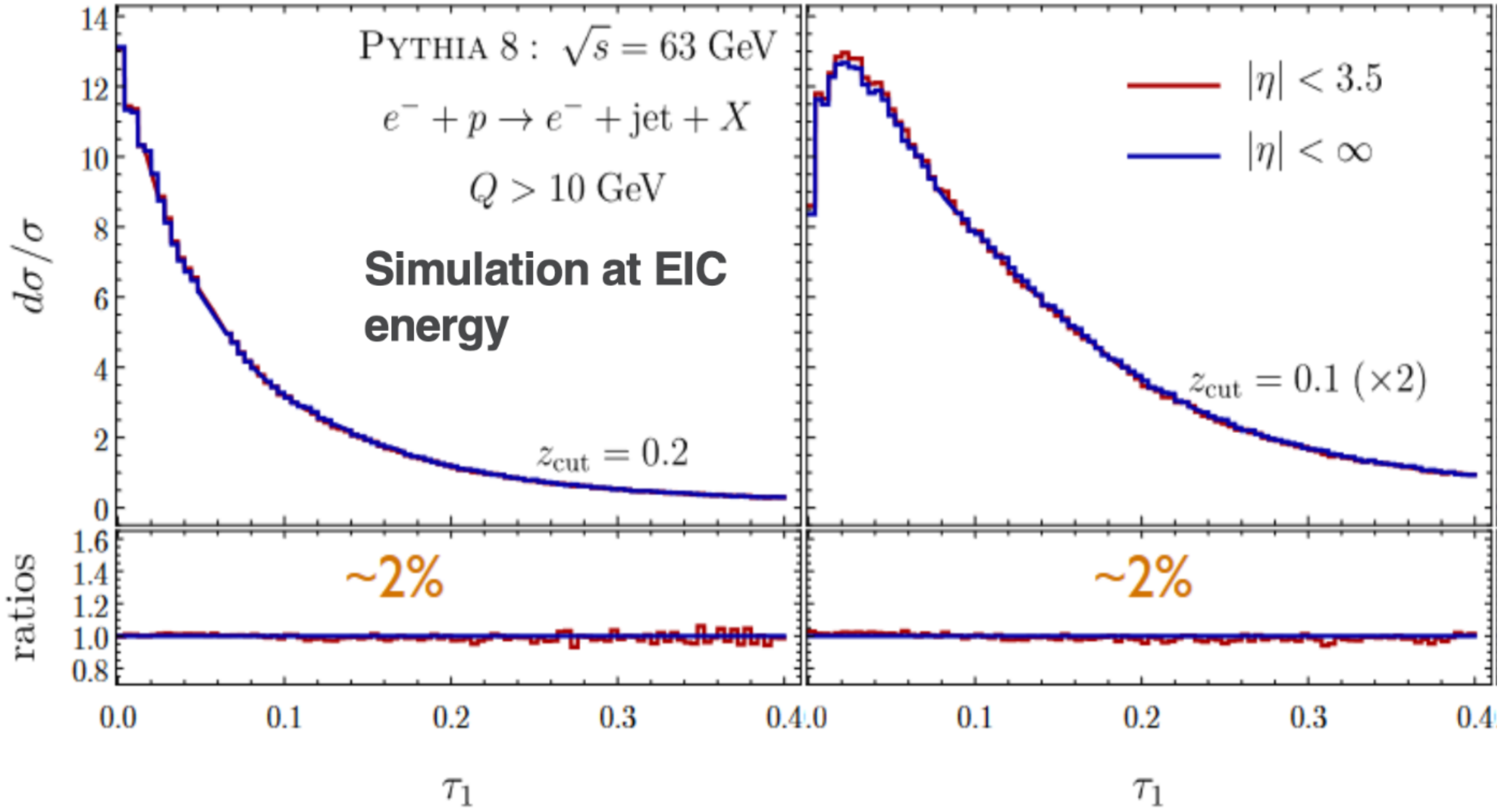
Reduction of non-perturbative corrections

Two calculations compatible within uncertainties

Better agreement between parton-level and detector-level

Less dependence on acceptance cuts





Benefits of grooming

Reduction of
non-perturbative corrections

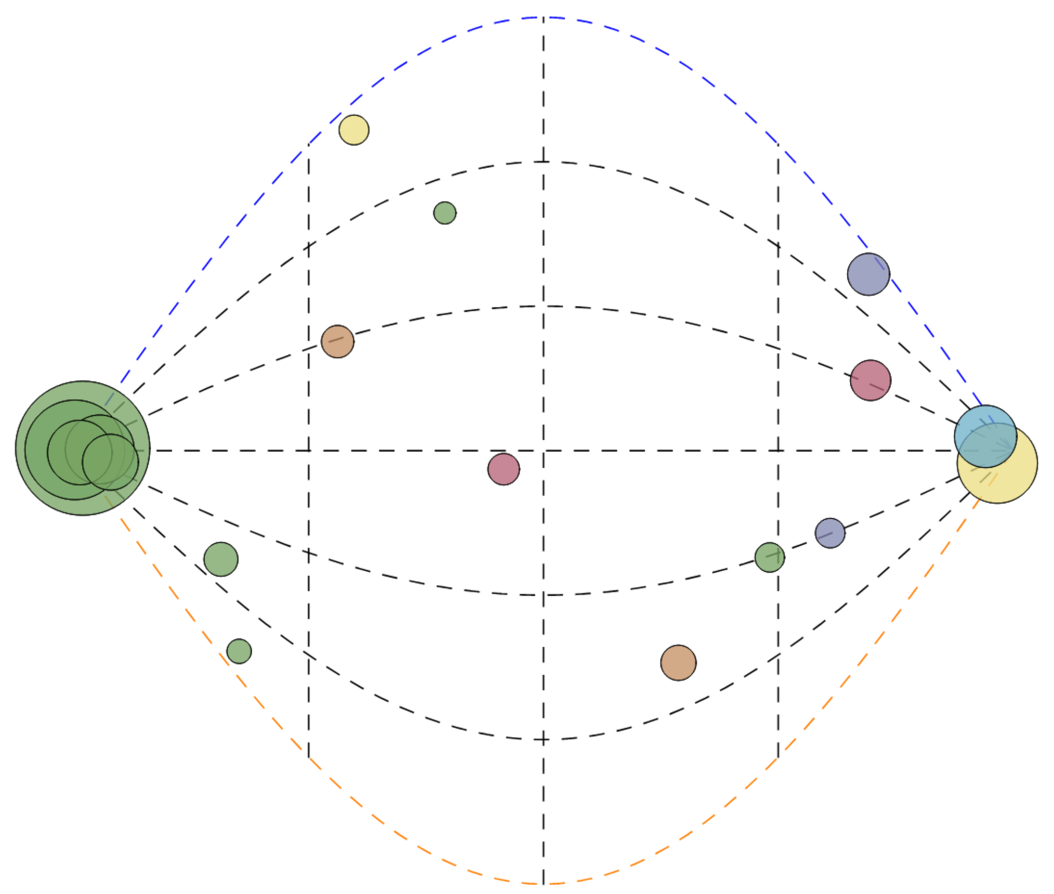
Two calculations compatible
within uncertainties

Better agreement between
parton-level and detector-level

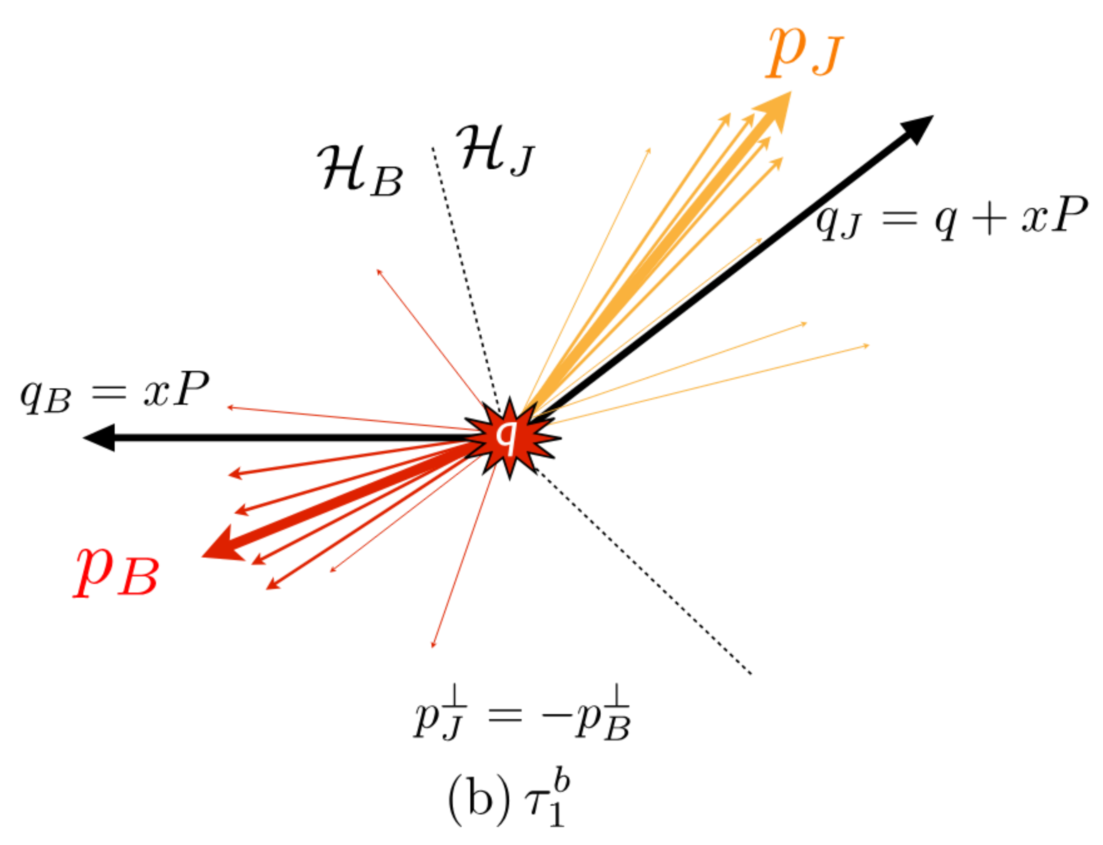
Less dependence
on acceptance cuts

Outlook

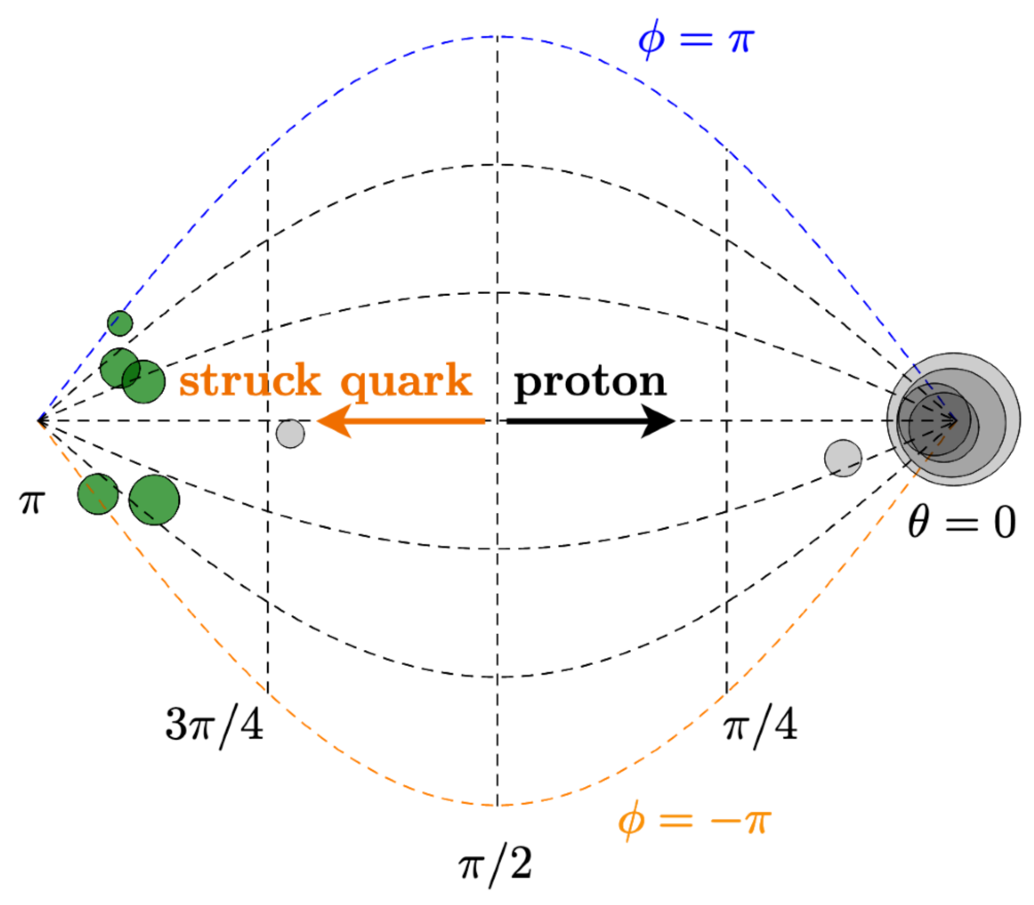
Jet algorithms



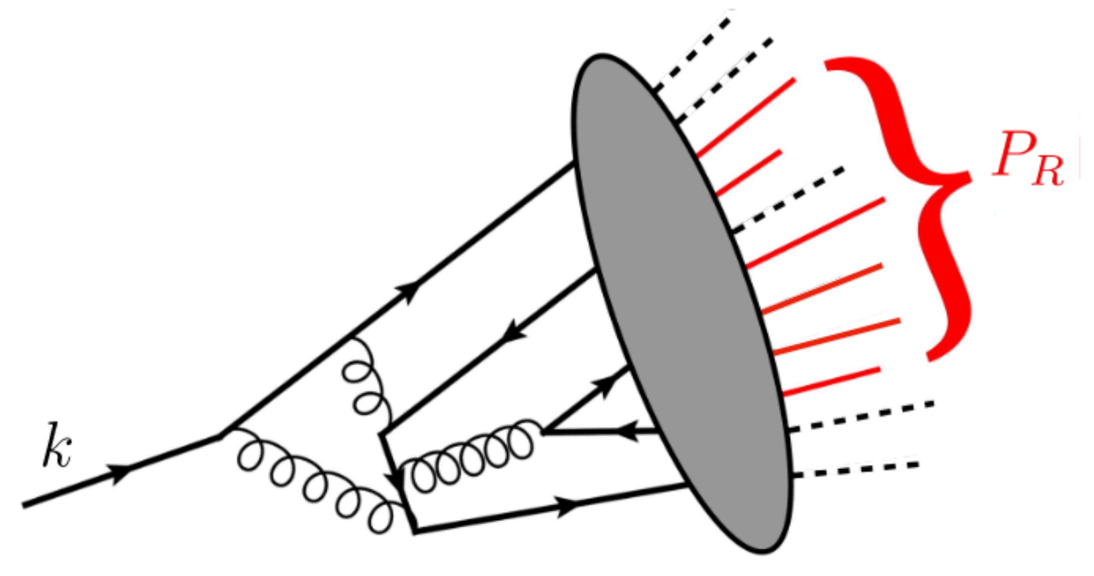
Event shapes



Grooming



Track functions



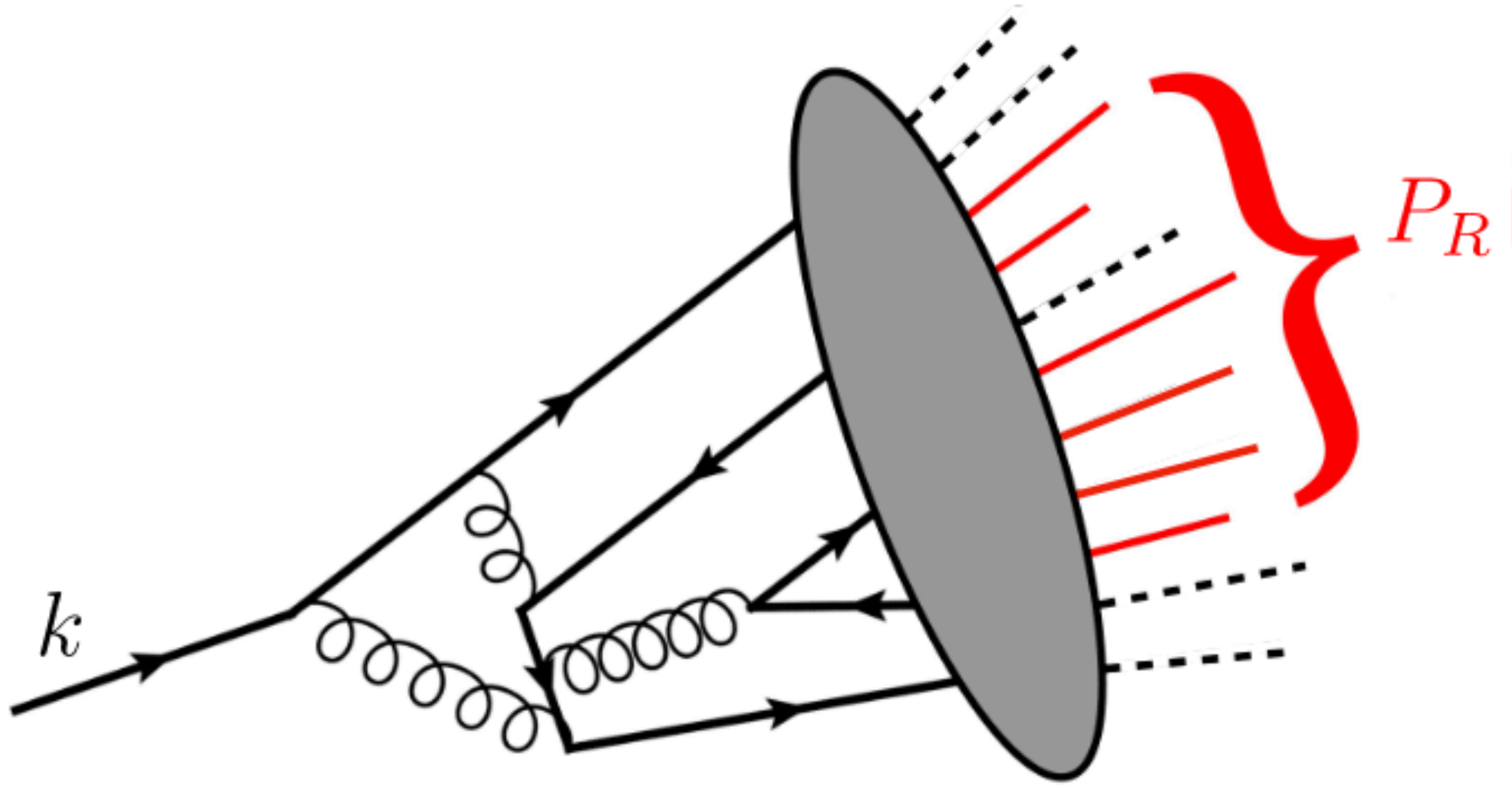
How to exploit the expertise gained from the LHC in DIS context?

Personal selection of results that are representative for on-going progress.
Apologies for any relevant omission of references.

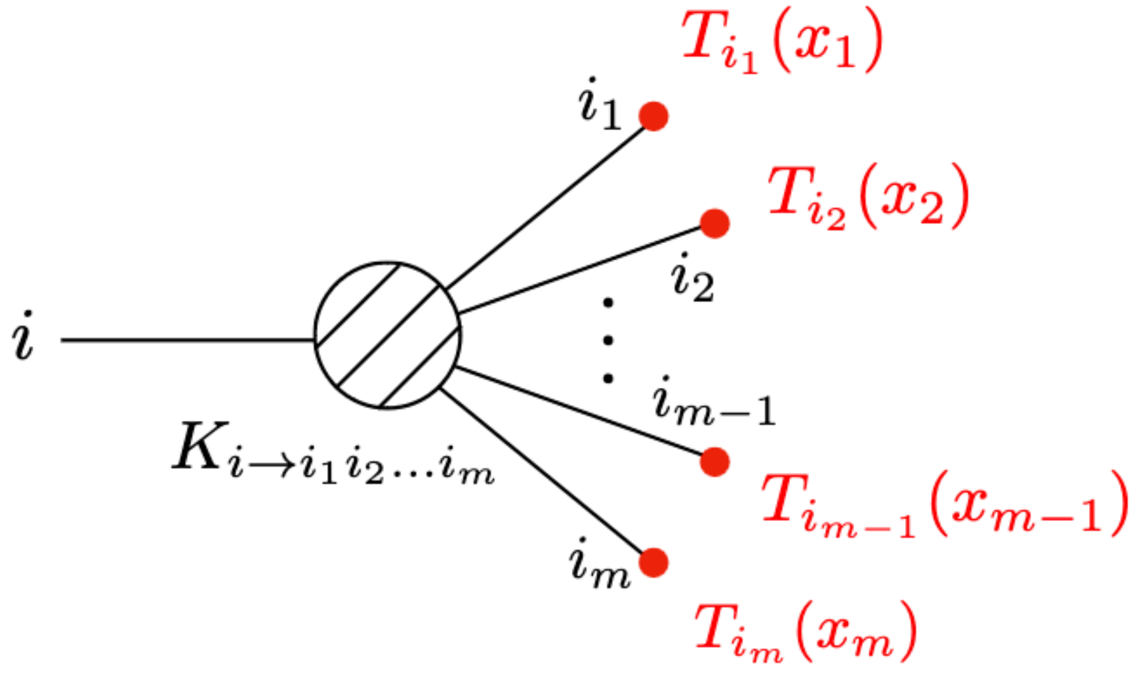
Track functions

[Chang, Procura, Thaler, Waalewijn ’13]

Non-pertubative objects describing fragmentation of a parton into a <u>subset of hadrons</u> (e.g. charged hadrons). They naturally encode multi-hadron fragmentation.



Evolution determined by a *non-linear* equation with perturbatively calculable kernels from multi-collinear splitting functions [Chen, Jaarsma, Li, Moult, Waalewijn, Zhu ’22], with a non-perturbative initial condition extracted from simulations/data



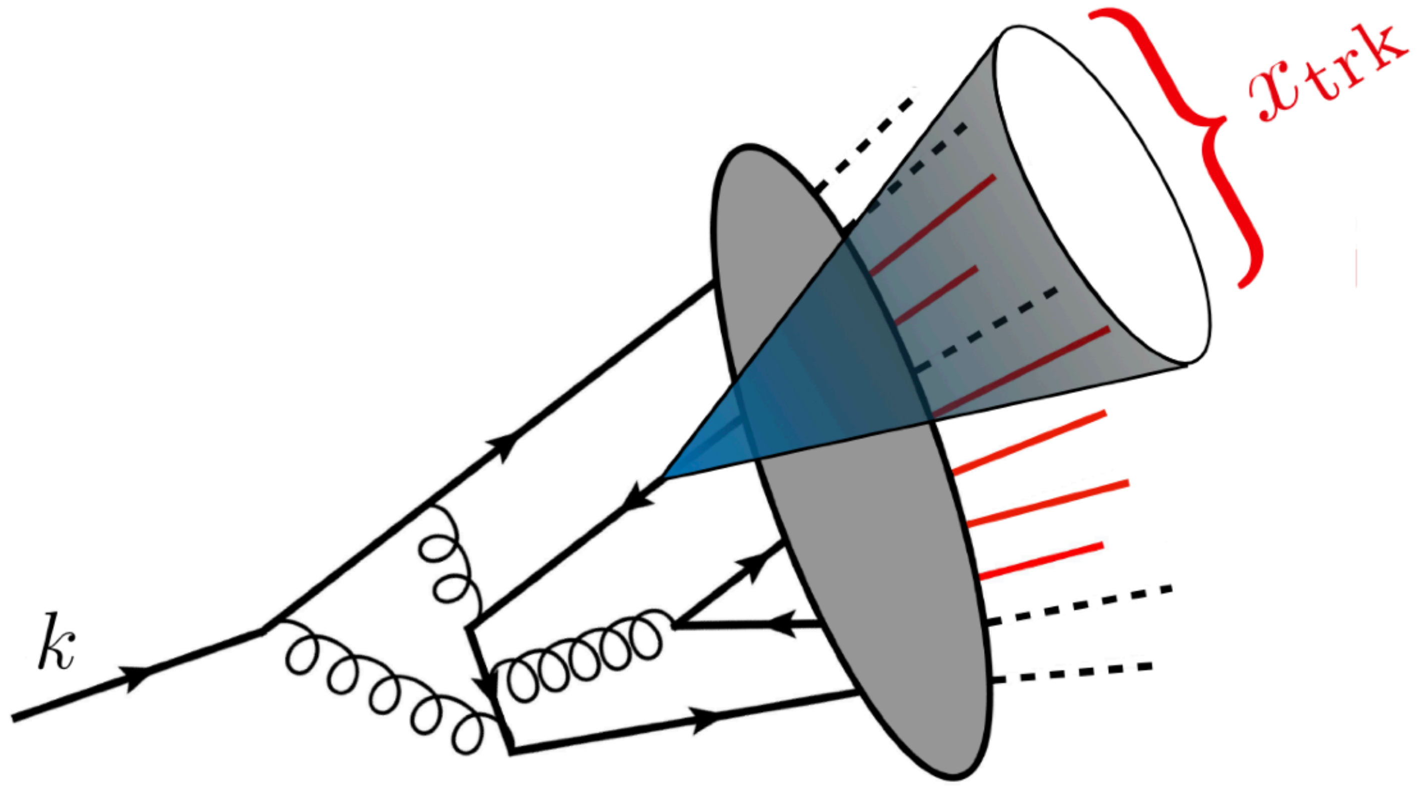
$$\frac{\mathrm{d}}{\mathrm{d}\ln\mu^2} T_i(x) = K_{i\to i} T(x) + \sum_{\{i_1,i_2\}} K_{i\to i_1 i_2} \otimes T_{i_1}(x_1) T_{i_2}(x_2) + \sum_{\{i_1,i_2,i_3\}} K_{i\to i_1 i_2 i_3} \otimes T_{i_1}(x_1) T_{i_2}(x_2) T_{i_3}(x_3)$$

$$K_{i\to i_1 i_2} \otimes T_{i_1} T_{i_2}(x) = \int_0^1 \mathrm{d}x_1 \mathrm{d}x_2 \; T_{i_1}(x_1) T_{i_2}(x_2) \int_0^1 \mathrm{d}z_1 \mathrm{d}z_2 \; \delta(1-z_1-z_2)\delta(x-z_1x_1-z_2x_2) K_{i\to i_1 i_2}(z_1,z_2)$$

Extracting track functions at the LHC (and HERA/EIC)

[Lee, Moult, Ringer, Waalewijn '23]

Presence of jets require *semi-inclusive* track jet functions, depending on x_{trk} = energy fraction of charged hadrons inside a jet.

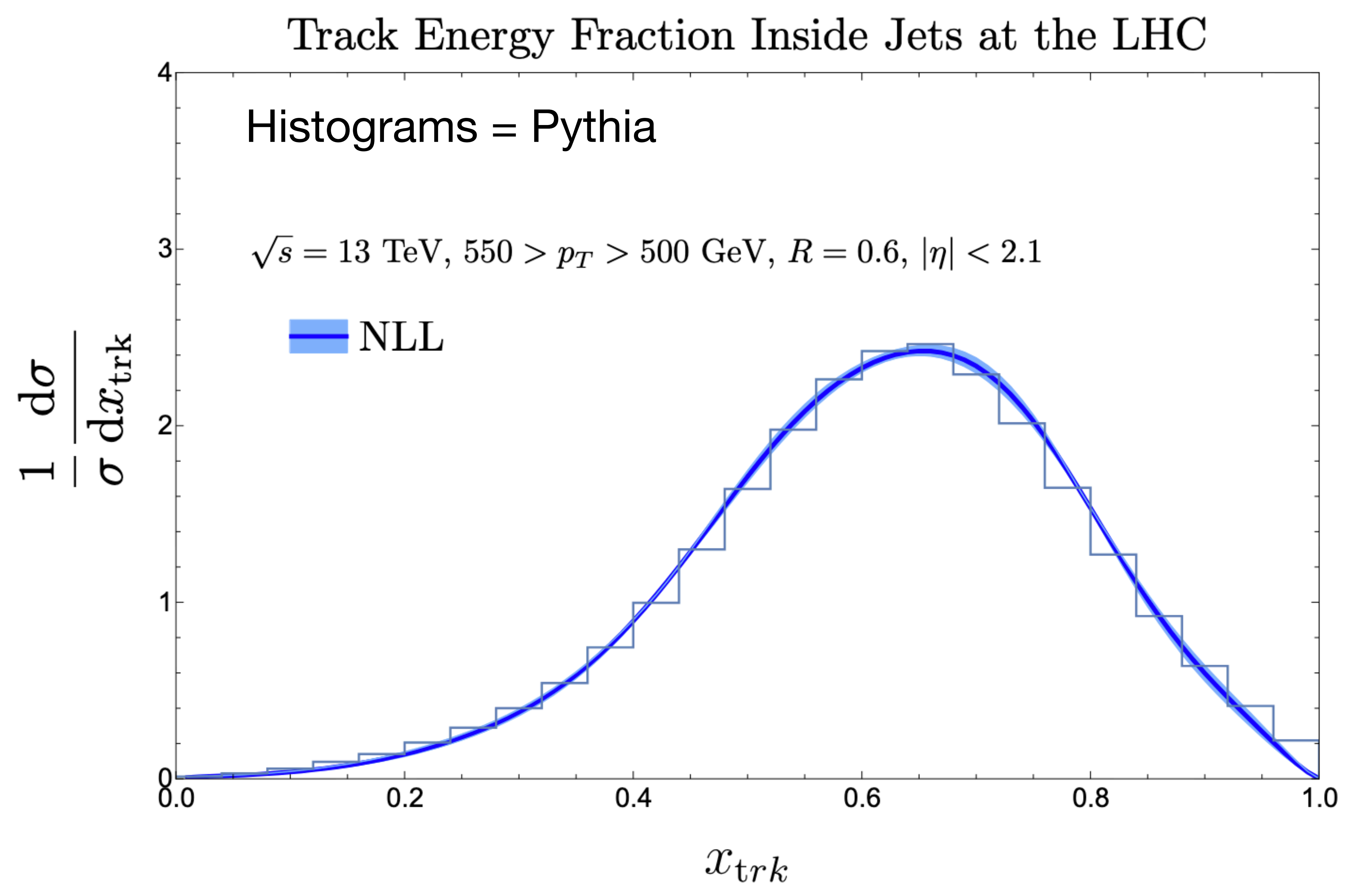


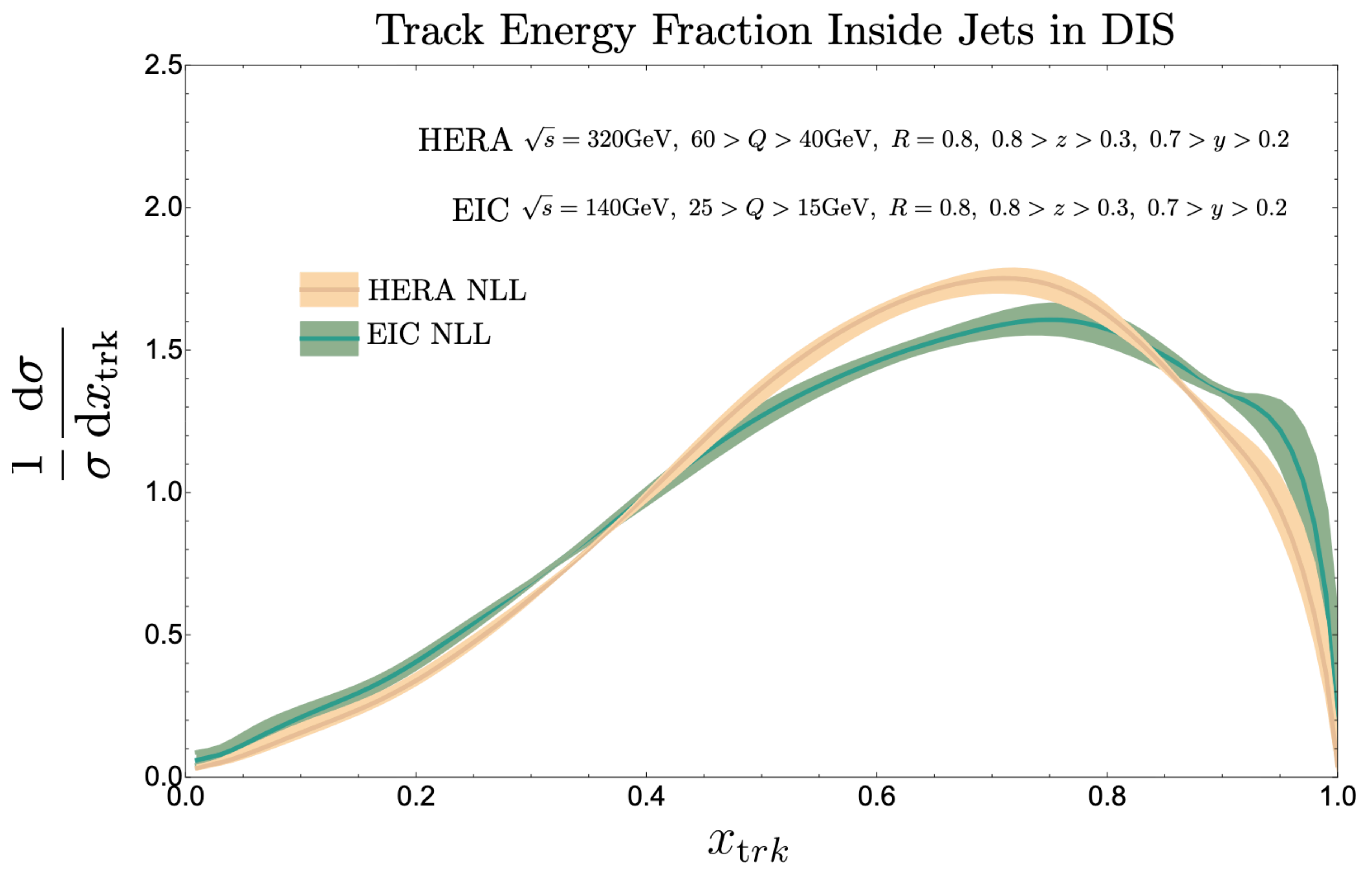
$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_T\,\mathrm{d}\eta\,\mathrm{d}x_{\mathrm{trk}}} = \mathcal{H}_i(p_T/z,\eta,\mu) \otimes_z \mathcal{G}_{i\to\mathrm{trk}}(z, x_{\mathrm{trk}}, p_T R, \mu)$$

Matched onto standard track functions with perturbatively (jet-algorithm dependent) matching coefficient

$$\mathcal{G}_{i\to\mathrm{trk}}(z, x_{\mathrm{trk}}, p_T R, \mu) = \sum_{m=1} \sum_{i_1,.,i_m} \mathcal{J}_{i\to[i_1,\dots,i_m]} \otimes \prod_{k=1}^{m} T_{i_k}(x_{\mathrm{trk}}) + \mathcal{O}\left(\frac{\Lambda_{\mathrm{QCD}}}{p_T R}\right)$$

Extracting track functions at the LHC (and HERA/EIC)



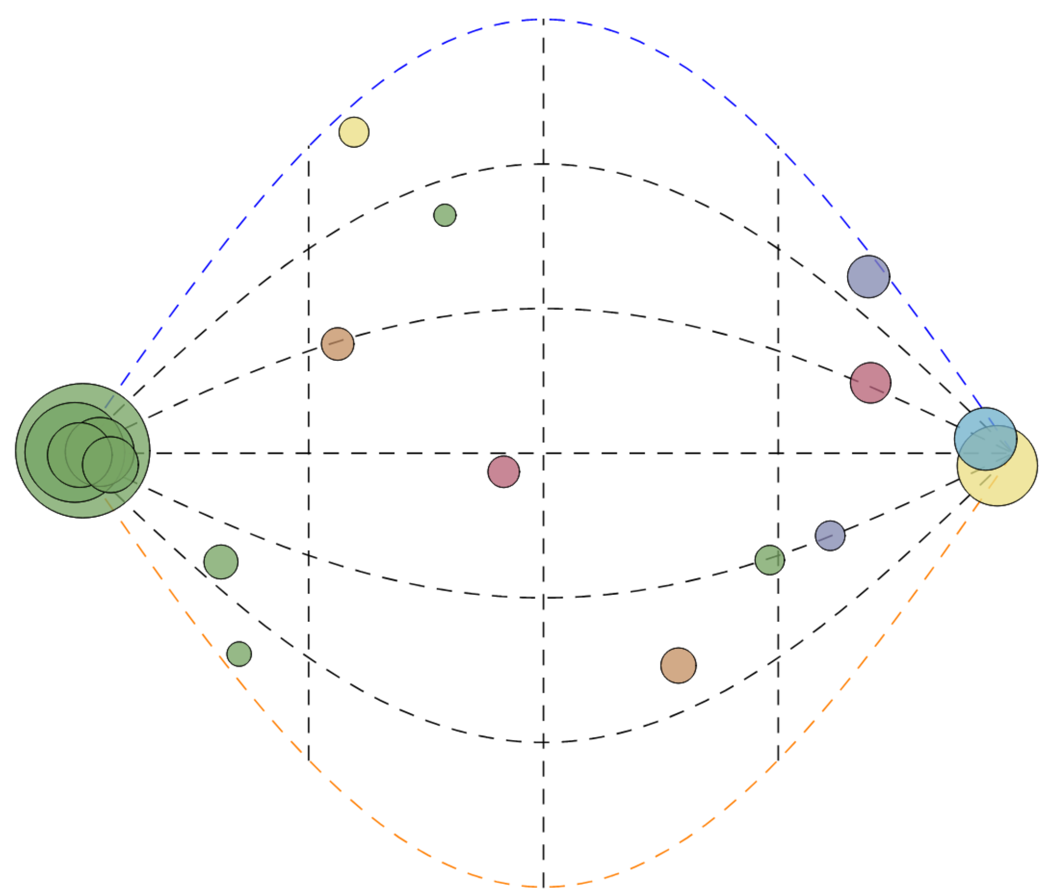


Track function with simpler behaviour at small momentum fractions w.r.t. FF
(at the price of a more complicated RG evolution)

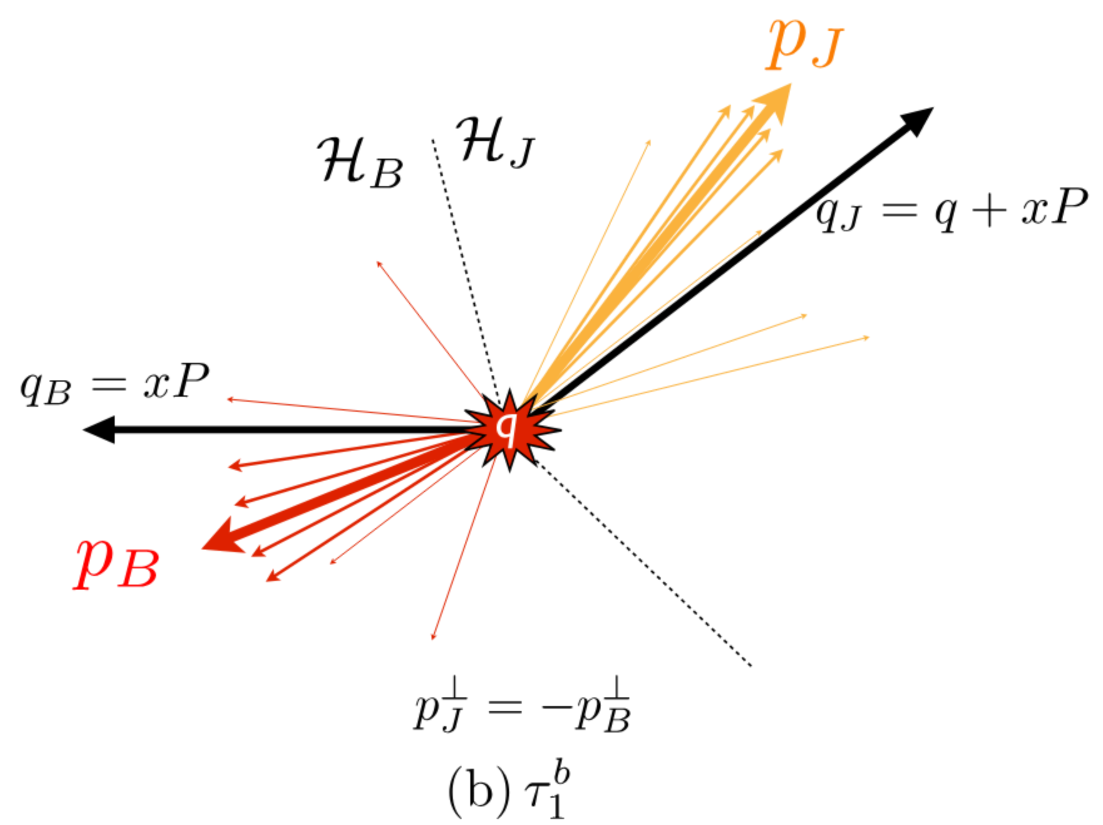
Summary: towards a second youth for DIS

We have seen many examples of ideas, techniques, tools, …
born at the LHC in the context of jet physics
that are being (successfully) ported to ep colliders

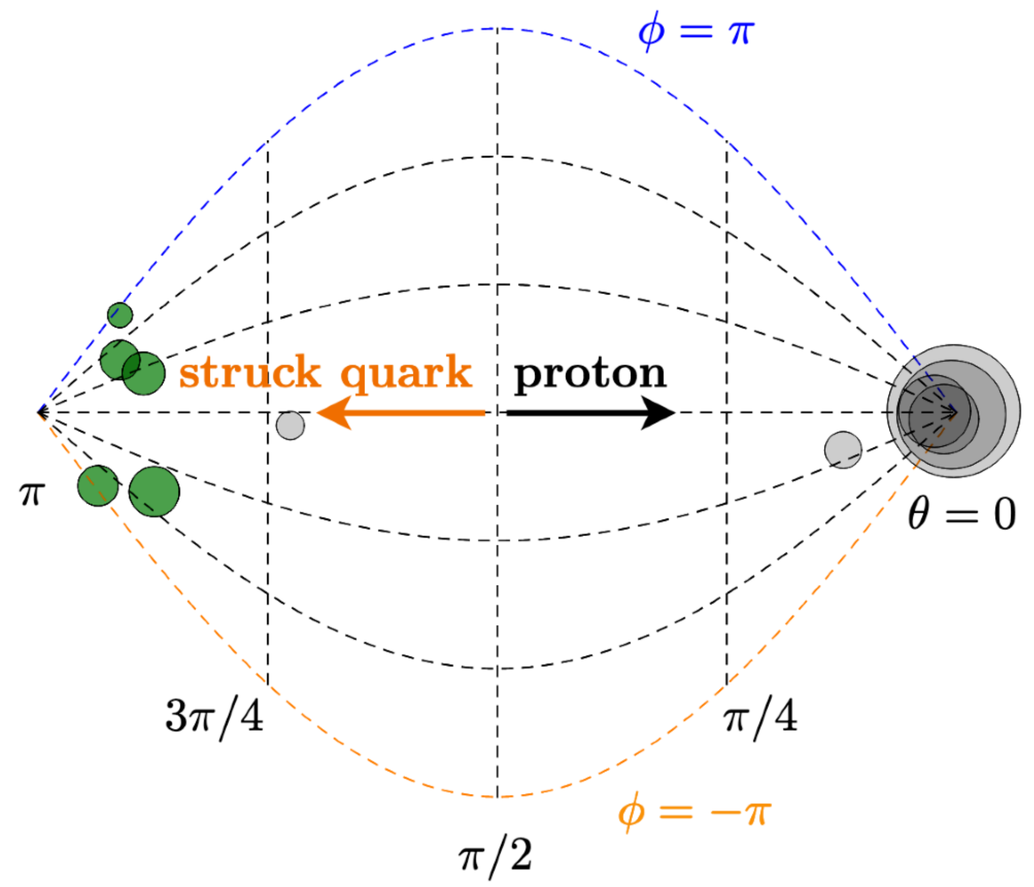
Jet algorithms



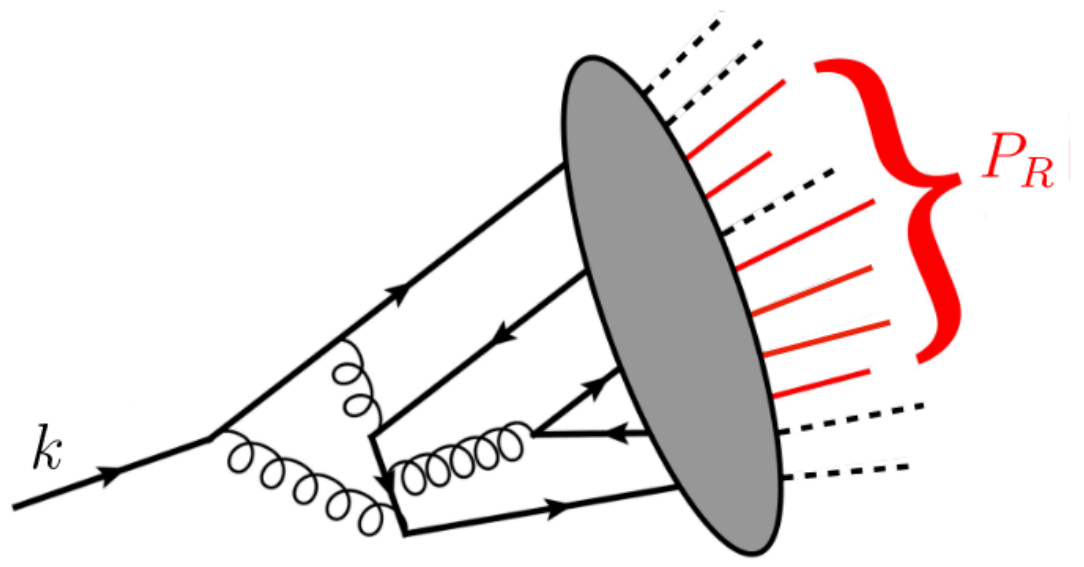
Event shapes



Grooming

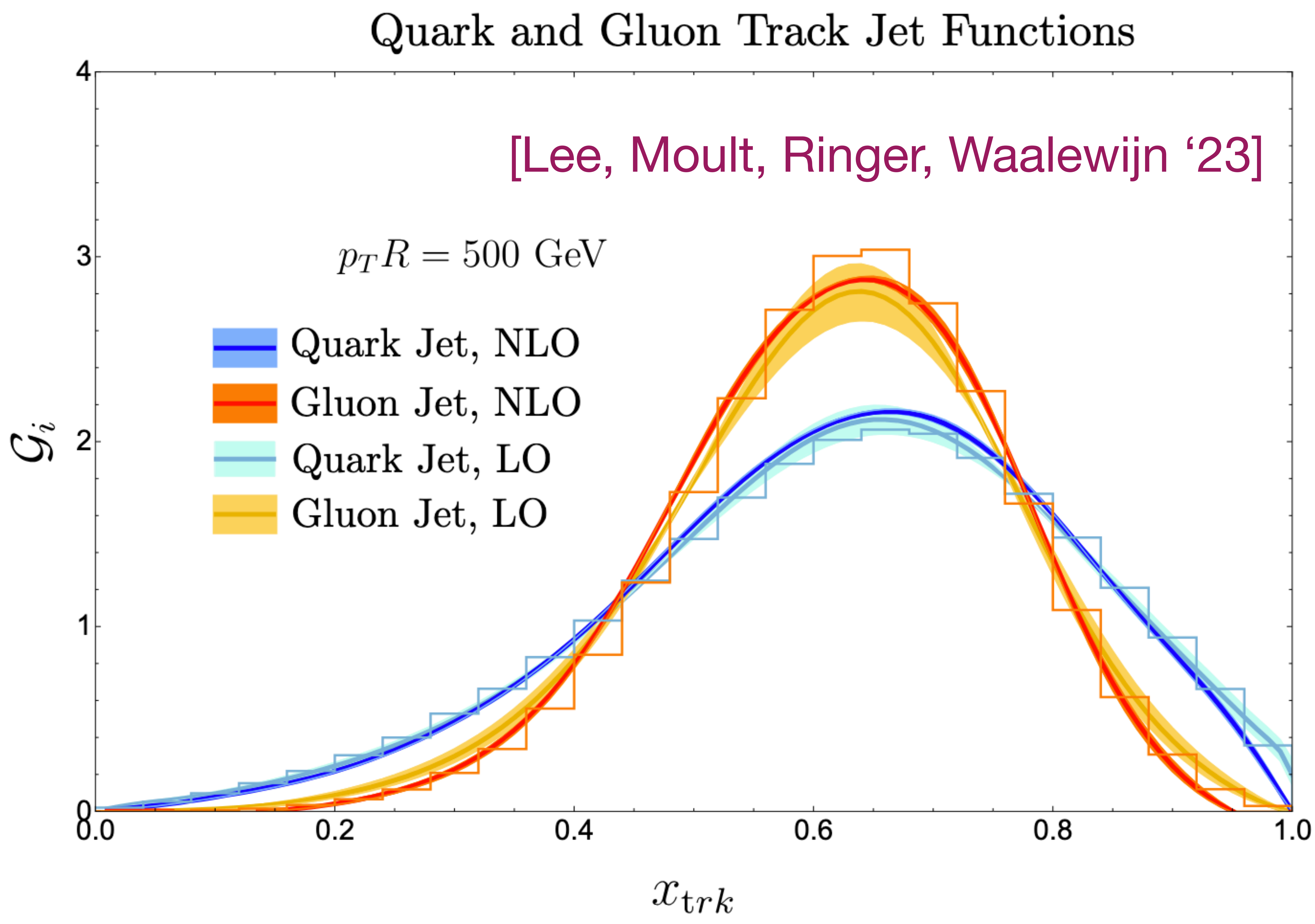


Track functions



Many more to come in the next years!

BACKUP



Quark and Gluon Track Jet Functions
[Lee, Moult, Ringer, Waalewijn '23]
$p_T R = 500$ GeV
Quark Jet, NLO
Gluon Jet, NLO
Quark Jet, LO
Gluon Jet, LO
$\mathcal{G}_i$
$x_{\rm trk}$
0
1
2
3
4
0.0
0.2
0.4
0.6
0.8
1.0

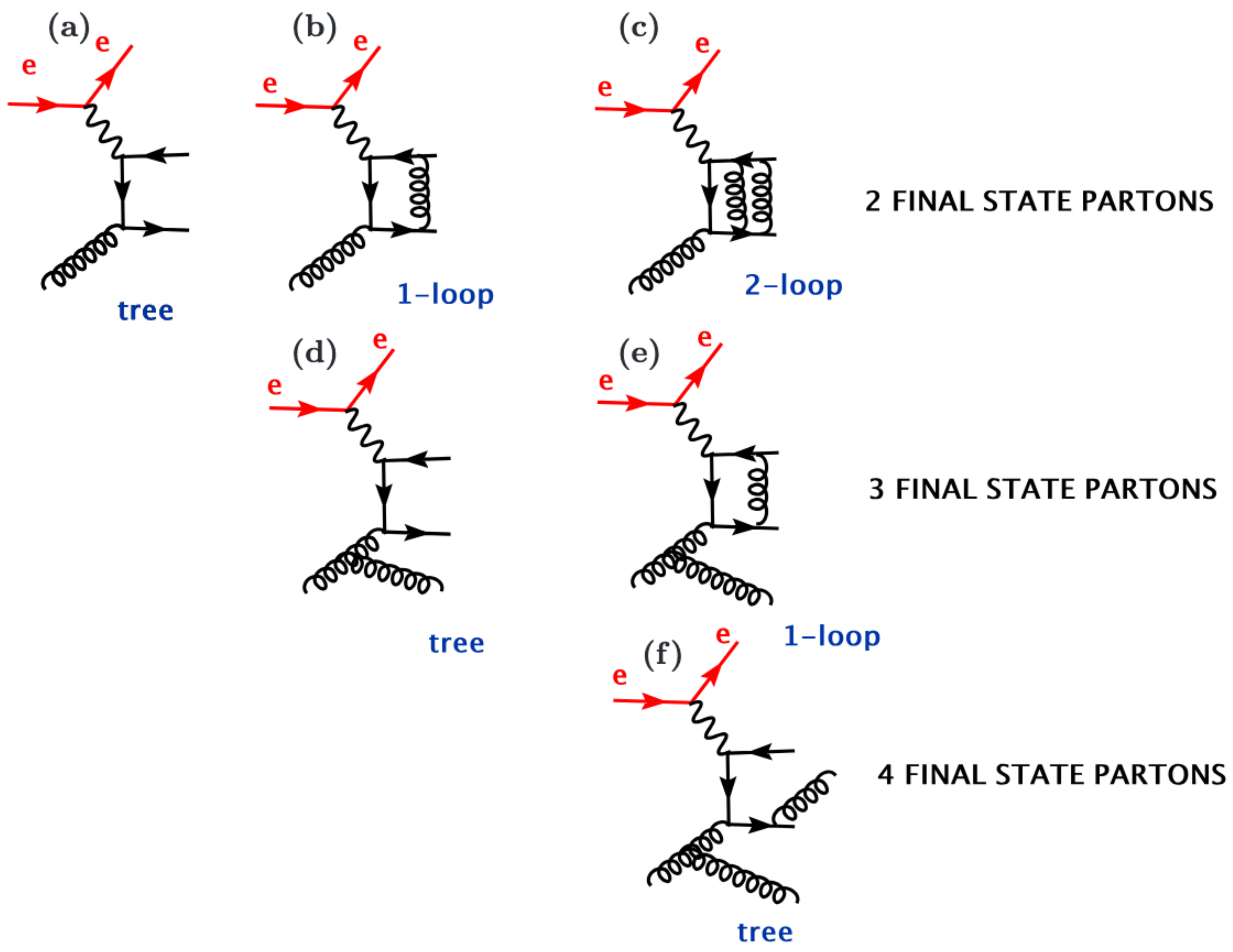


Figure 1: *Examples of Feynman diagrams contributing to inclusive-jet production in the Breit frame up to $\mathcal{O}(\alpha_s^3)$. The calculations of inclusive-jet cross sections up to $\mathcal{O}(\alpha_s^2)$ include the (a) leading-order diagrams and (b) virtual and (d) real corrections. The lowest-order diagrams that contribute to the cross-section difference between the anti-k_T and k_T algorithms are of type (f). The calculations of the cross-section difference between the SIScone and k_T algorithms include diagrams of type (d), (e) and (f).*

This section briefly reviews the MultiFold technique introduced in Ref. [65, 66]. Let $\vec{x} = (\vec{p}_{\mathrm{T}}^{e}, p_z^e, p_{\mathrm{T}}^{\mathrm{jet}}, \eta^{\mathrm{jet}}, \phi^{\mathrm{jet}}, q_{\mathrm{T}}^{\mathrm{jet}}/Q, \Delta\phi^{\mathrm{jet}})$. MultiFold is an iterative, two-step procedure. Let $X_{\mathrm{data}} = \{\vec{x}_i\}$ be the set of events in data and $X_{\mathrm{MC,truth}} = \{\vec{x}_{\mathrm{truth},i}\}$ and $X_{\mathrm{MC,reco}} = \{\vec{x}_{\mathrm{reco},i}\}$ be sets of events in simulation with a correspondance between the two sets. In simulation, we have a set of observables at particle-level ('truth') and detector-level ('reco') for each event. If an event does not pass the particle-level or detector-level event selection, then the corresponding set of observables are assigned a dummy value $\vec{x} = \emptyset$. Each event i in simulation is also associated with a weight w_i. MultiFold then proceeds iteratively by repeating the following two steps to iteratively adjust a set of event weights ν_i:

1. Train a classifier f to distinguish the weighted simulation at detector-level from the data. The loss function is the binary cross entropy:

$$L_1[f] = -\sum_{\vec{x}_i \in X_{\mathrm{data}}} \log(f(\vec{x}_i)) - \sum_{\vec{x}_i \in X_{\mathrm{MC,reco}}} \nu_i\, w_i\, \log(1 - f(\vec{x}_i))\,, \tag{A1}$$

   where both sums only include events that pass the detector-level selection. For events that pass the detector-level selection, define $\lambda_i = \nu_i \times f(\vec{x}_i)/(1 - f(\vec{x}_i))$ for $\vec{x}_i \in X_{\mathrm{MC,reco}}$. For events that do not pass the detector-level selection, $\lambda_i = \nu_i$.

2. Train a classifier g to distinguish the particle-level simulation weighted by ν from the particle-level simulation weighted by λ. The loss function is once again the binary cross entropy:

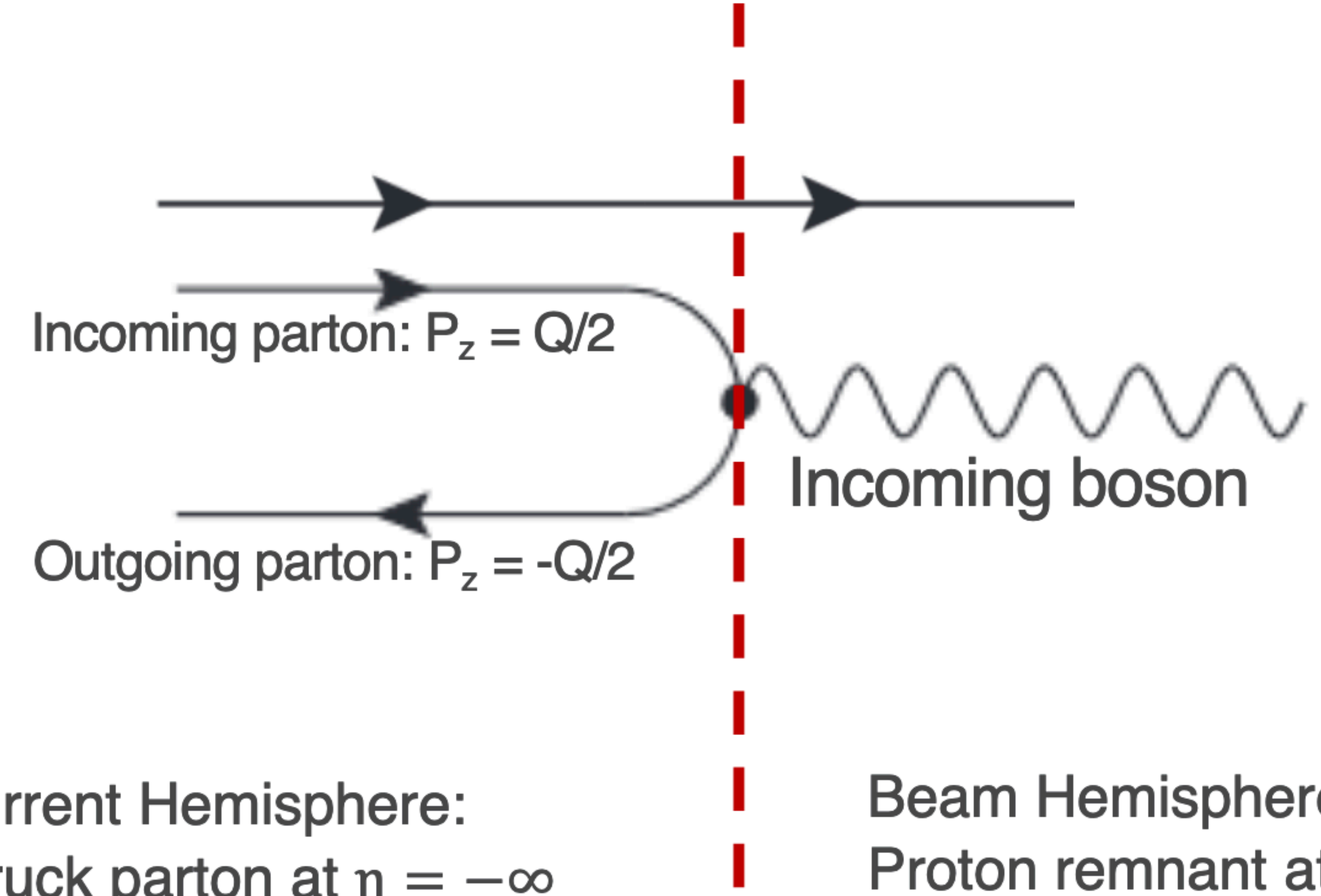
$$L_2[g] = -\sum_{\vec{x}_i \in X_{\mathrm{MC,truth}}} \lambda_i\, w_i\, \log(g(\vec{x}_i)) - \sum_{\vec{x}_i \in X_{\mathrm{MC,truth}}} \nu_i\, w_i\, \log(1 - g(\vec{x}_i))\,, \tag{A2}$$

   where both sums only include events that pass the particle-level selection. For events that pass the particle-level selection, define $\nu_i = \nu_i \times g(\vec{x}_i)/(1 - g(\vec{x}_i))$ for $\vec{x}_i \in X_{\mathrm{MC,truth}}$. For events that do not pass the particle-level selection, ν_i is left unchanged from its previous value.

The process is initialized by $\nu_i = 1$ for all events. The $f/(1-f)$ or $g/(1-g)$ form for the weights is a well-known (see e.g. Ref. [111, 112]) approximation for the likelihood ratio of the two samples in the left and right sums in each equation. After iterating the above procedure some number of times, the final result is constructed by making histograms with the truth events using the final $\{\nu_i w_i\}$ weights.

Born-level DIS in the Breit Frame

Event $p_T = 0$



Current Hemisphere:
Struck parton at $\eta = -\infty$

Beam Hemisphere:
Proton remnant at $\eta = \infty$