Stony Brook University

Center for Frontiers in Nuclear Science

Measurement of Groomed Event Shapes in e+p DIS

Henry Klest for the H1 Collaboration

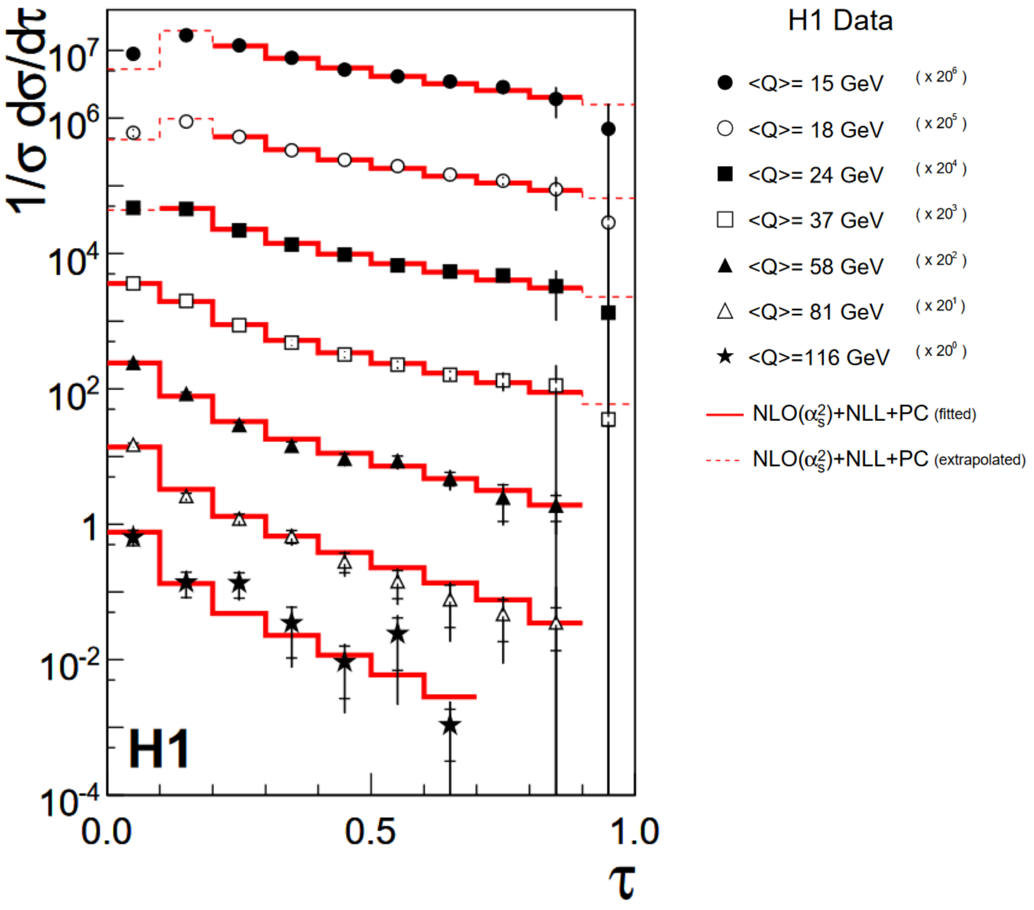
DIS2022, May 3

FAR BEYOND

Event Shapes

$$\tau = 1 - T \quad \text{with} \quad T = \frac{\sum_h |\vec{p}_{z,h}|}{\sum_h |\vec{p}_h|}$$

- Inclusive observables where all particles contribute
  - E.g. Thrust – measures degree of collimation along an axis
- **Sensitive to QCD across scales**
- Calculable to high precision in perturbation theory
  - Fixed-order QCD → tail of thrust distribution
  - Soft-collinear effective theory (SCET) calculations → peak of thrust distribution
- Used extensively in e^+e^- and Breit frame e+p collisions



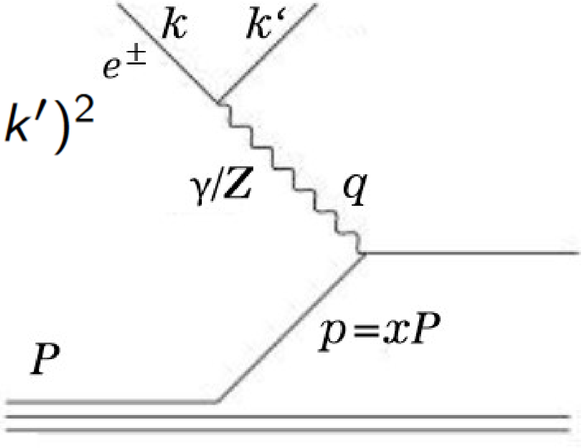
arXiv/0512014v1

Inclusive DIS & Breit Frame

$$Q^2 = -q^2 = -(k - k')^2$$

$$y = \frac{p \cdot q}{p \cdot k}$$

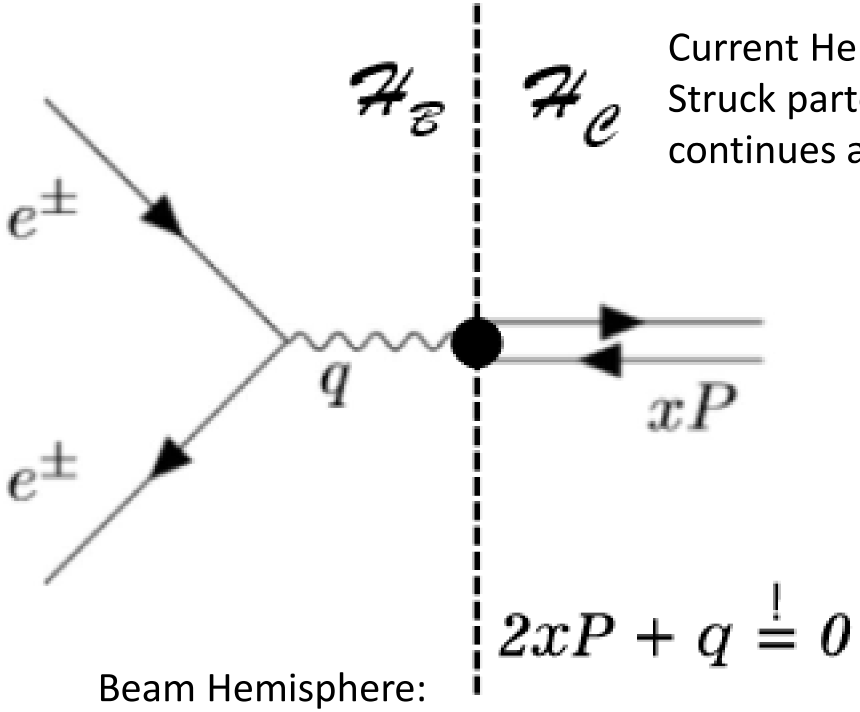
$$x = -\frac{q^2}{2P \cdot q} = \frac{Q^2}{2P \cdot q}$$



- HERA-II data
  - $Q^2 > 150$ GeV², $0.2 < y < 0.7$
  - No direct $x_{Bj.}$ cut applied
  - 352 pb^{-1} collected

- Breit Frame
  - Defined as the frame where $2x_{Bj.}P + q = 0$
    - Divides event into two hemispheres: “beam”/”remnant”/”target” hemisphere and “current”/”struck parton” hemisphere
  - Exchanged boson reverses struck parton’s momentum
    - Parton has $\overrightarrow{xP}$ incoming, -$\overrightarrow{xP}$ outgoing

Current Hemisphere:
Struck parton
continues at η = -∞



Beam Hemisphere:
Proton remnant
continues at η = ∞

arXiv:2006.10751

Centauro

- Jet algorithm with asymmetric clustering distance measure
  - Suited for clustering Born-level DIS in the Breit frame

- Here Centauro is used to produce a clustering tree for the full event

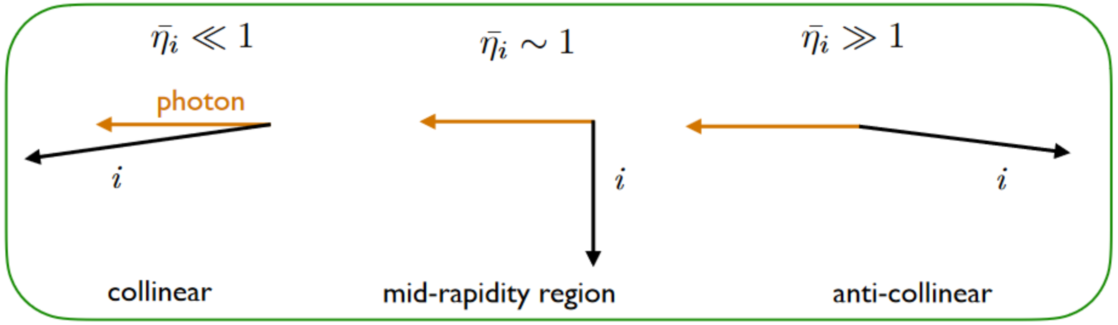
Definition:

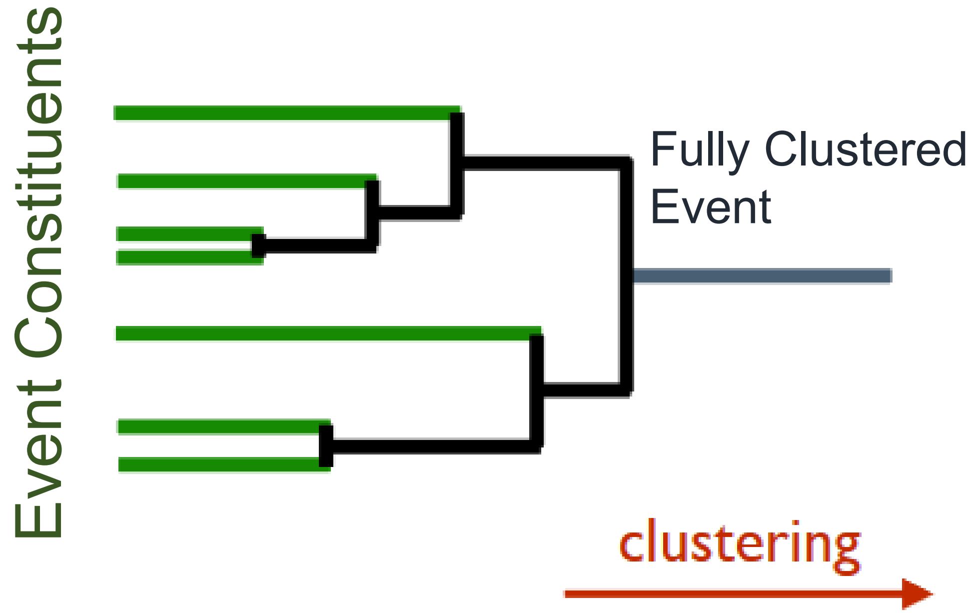
$$d_{ij} = (\bar{\eta}_i - \bar{\eta}_j)^2 + 2\bar{\eta}_i\bar{\eta}_j(1 - \cos(\phi_i - \phi_j))$$

Breit frame

$$\bar{\eta}_i = 2\sqrt{1 + \frac{q \cdot p_i}{x_B P \cdot p_i}} \xrightarrow[\text{frame}]{\text{Breit}} \frac{2p_i^{\perp}}{p_i^{+}}$$

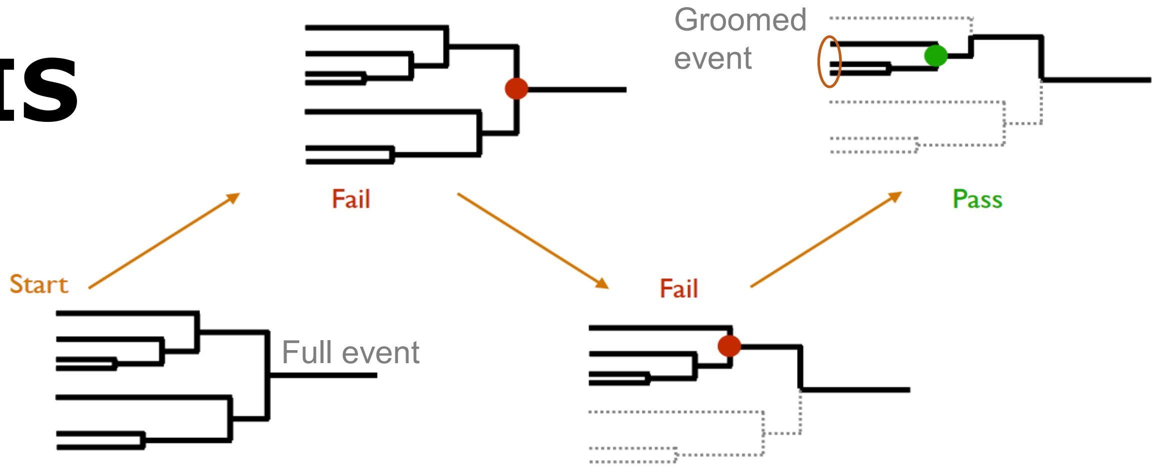
Limits:





Figures courtesy of Y. Makris [2]

Event Grooming in DIS



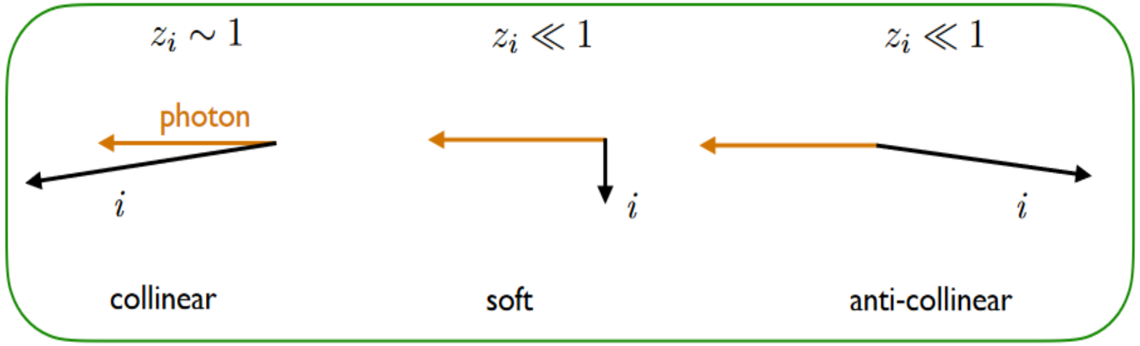
- Whole event is clustered into one "jet"

- Iteratively de-cluster until grooming condition is passed
  - Analogous to Soft Drop in p+p

- Groomed events are similar to groomed jets!

$$z_i = \frac{P \cdot p_i}{P \cdot q} \xrightarrow[\text{frame}]{\text{Breit}} z_i = n \cdot p_i / Q = p_i^+ / Q \,.$$

Limits (geometric interpretation in the Breit frame)

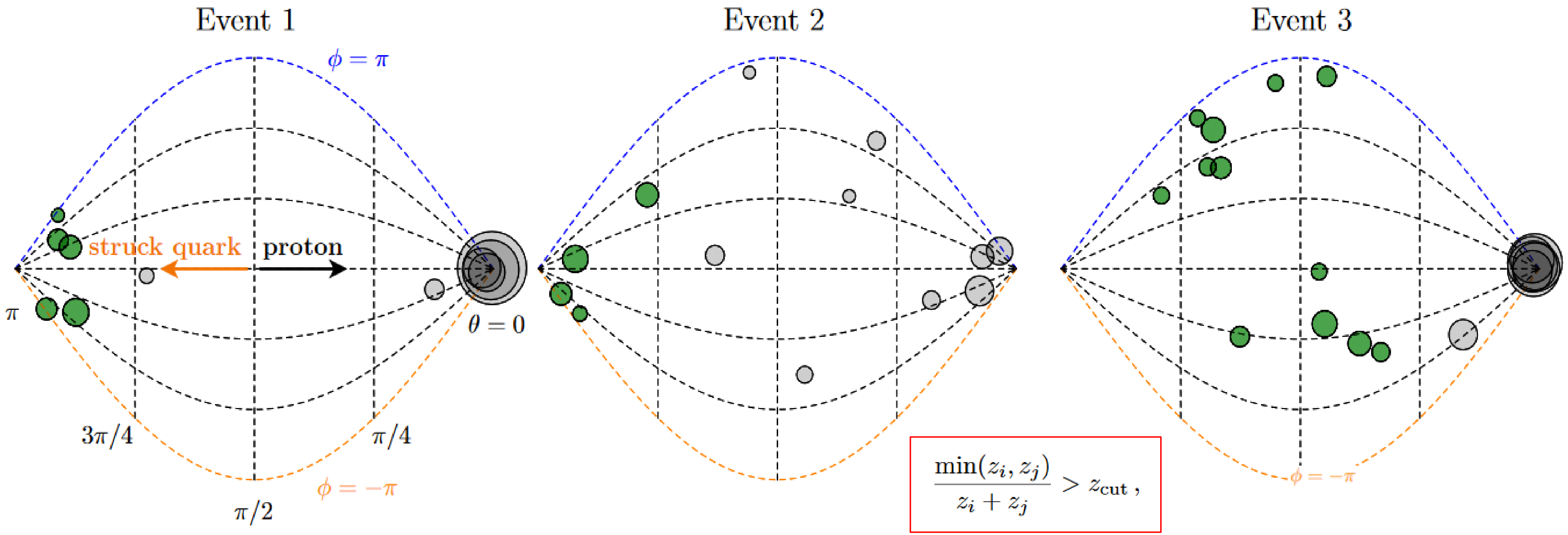


$$\frac{\min(p_{t1}, p_{t2})}{p_{t1} + p_{t2}} > z_{\text{cut}} \longrightarrow \frac{\min(z_i, z_j)}{z_i + z_j} > z_{\text{cut}}$$

p+p Soft Drop condition

DIS grooming condition

Breit Frame Event Displays



$$\frac{\min(z_i, z_j)}{z_i + z_j} > z_{\text{cut}} \,,$$

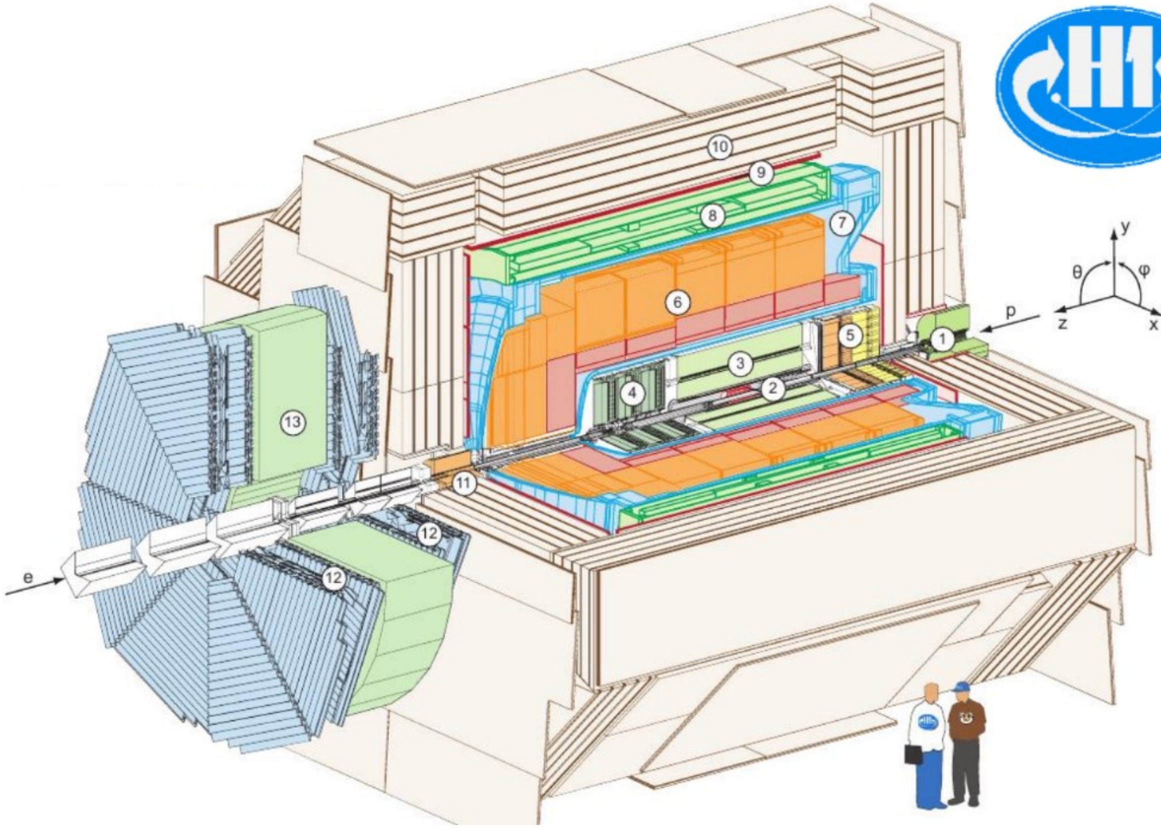
Figure 2. Visualization of three PYTHIA 8 events at $\sqrt{s} = 63$ GeV and $Q \sim 10$ GeV before and after grooming. The particles in this events are represented by disks on the unfolded sphere. Green disks represent particles that pass grooming where grayed-out particles are removed from the event by the grooming procedure. For the grooming parameter we use here $z_{\text{cut}} = 0.1$

H1 Detector

- HERA
  - World's only high energy electron-proton collider

$$E_e = 27.6 \text{ GeV}, E_p = 920 \text{ GeV}$$
$$\rightarrow \sqrt{s} = 319 \text{ GeV}$$

  - 352 pb^{-1} collected in HERA-II run period from 2003-2007
- H1 Experiment
  - Hermetic detector with asymmetric design
    - Drift chamber + silicon tracking
    - High-resolution LAr calorimeter
  - Trigger on high-energy hadronic or EM LAr cluster
    - > 99% efficient for $y < 0.7$

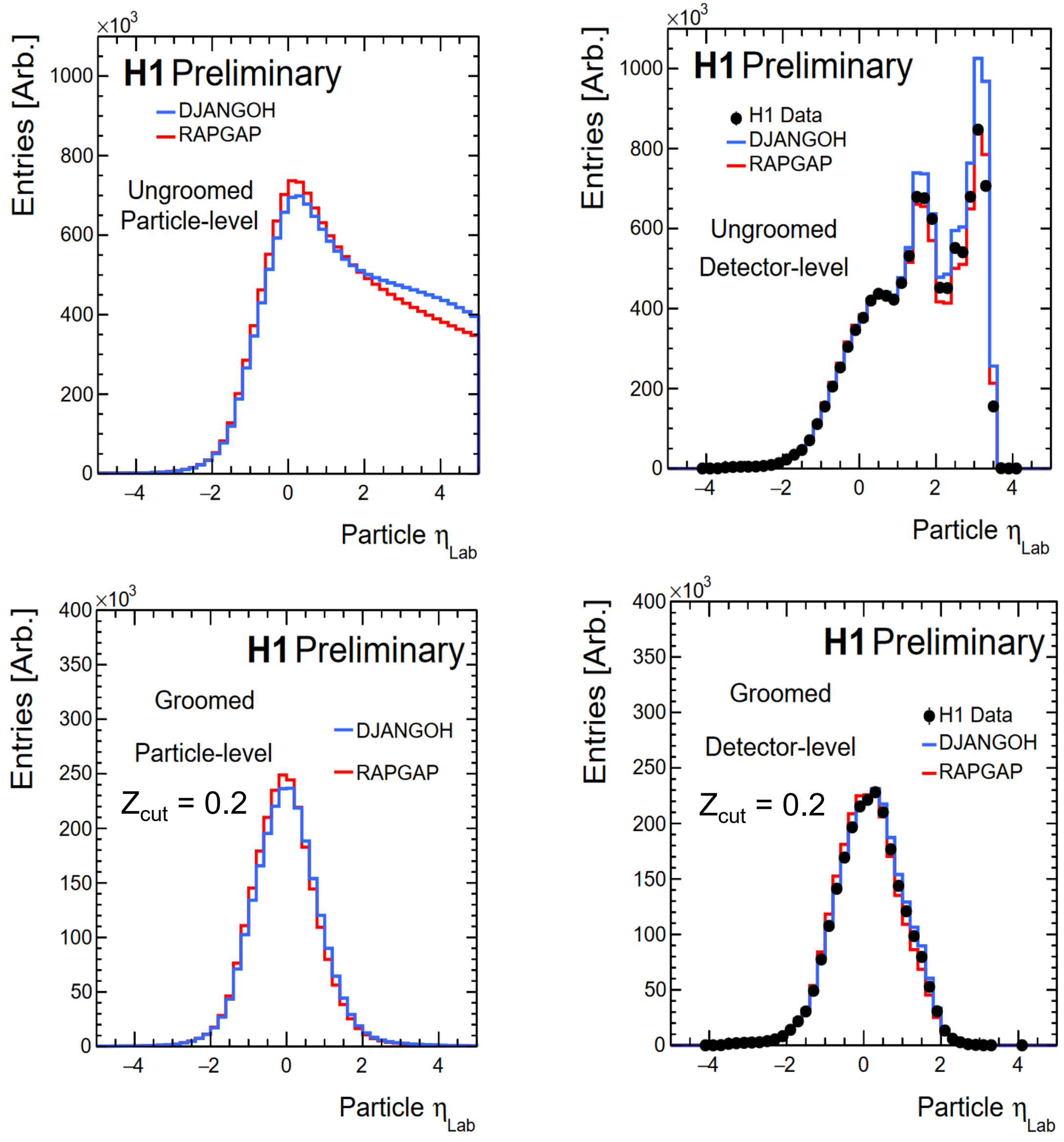


H1 LAr Calorimeter Specifications

Electromagnetic part	Hadronic part
10 to 100 cm^2	50 to 2000 cm^2
20 to 30 X_0 (30784)	4.7 to 7 λ_{abs} (13568)
$\approx 11\%/\sqrt{E_e} \oplus 1\%$	$\approx 50\%/\sqrt{E_h} \oplus 2\%$

Grooming Benefits

- No underlying event, why groom?
  - Less affected by lab-frame detector acceptance
  - Mitigate QCD remnant, ISR
  - No theoretically challenging non-global logarithms
- Ungroomed detector-level shows significant difference from particle-level
  - Detector acceptance, efficiencies
- Grooming events brings particle-level and detector-level distributions into much better agreement!



Observables

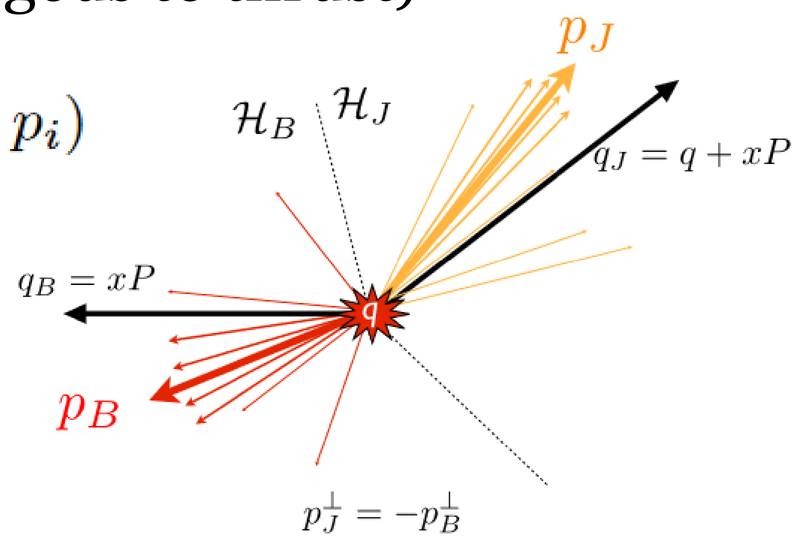
- After grooming procedure, a subset of particles survives
  - Event shape is calculated with these particles
  - Two event shapes studied here

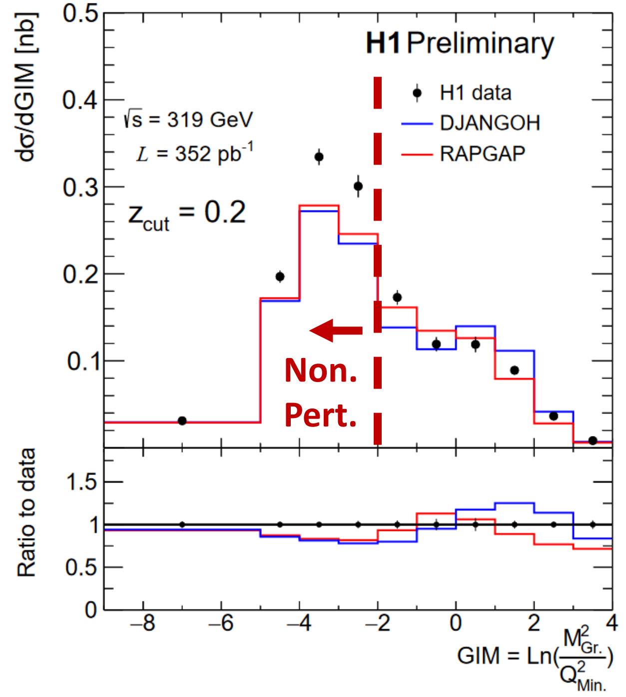
- Groomed Invariant Mass (GIM) $M^2_{Gr.} = (\sum_i p_i^\mu)^2$

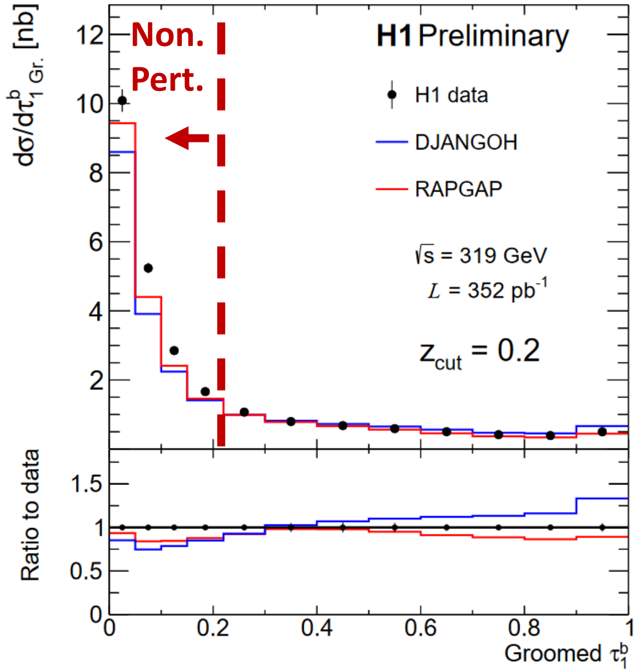
- Groomed 1-Jettiness τ_1^b (analogous to thrust)

$$\tau_1 = \frac{2}{Q^2} \sum_{i \in \text{gr. ent.}} \min(q_B \cdot p_i,\, q_J \cdot p_i)$$

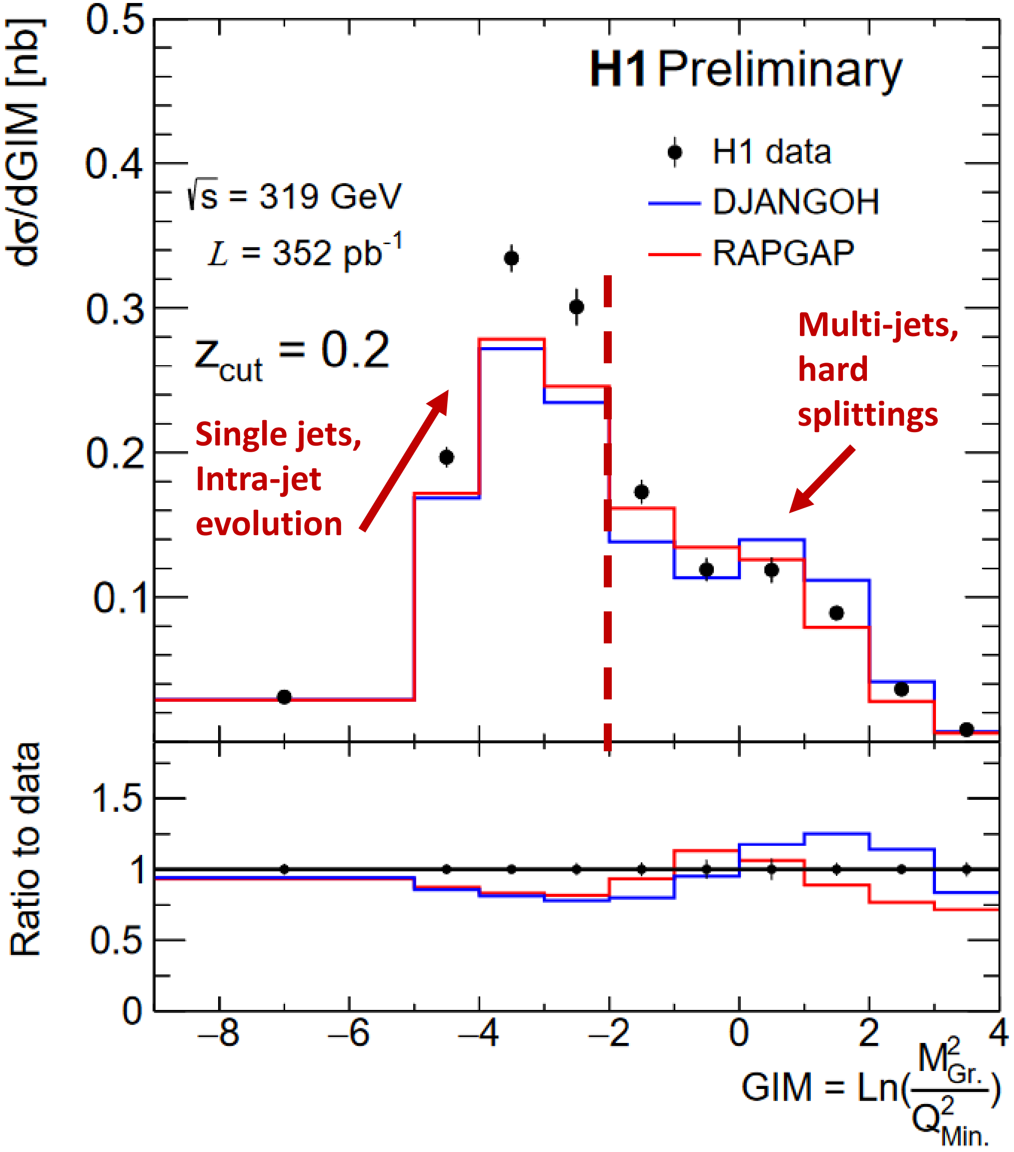
$$\tau_1^b \rightarrow \quad q_J = q + xP, \quad q_B = xP$$

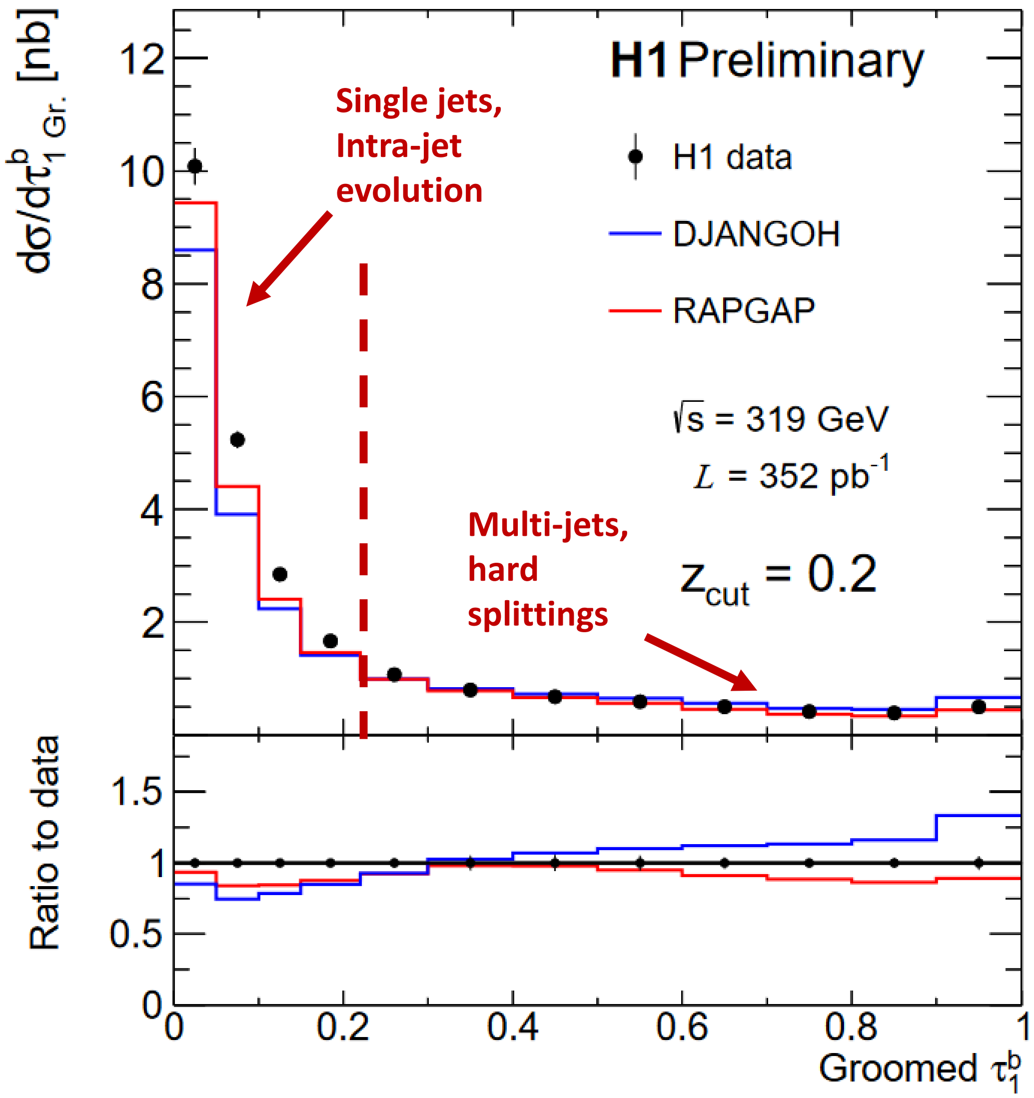




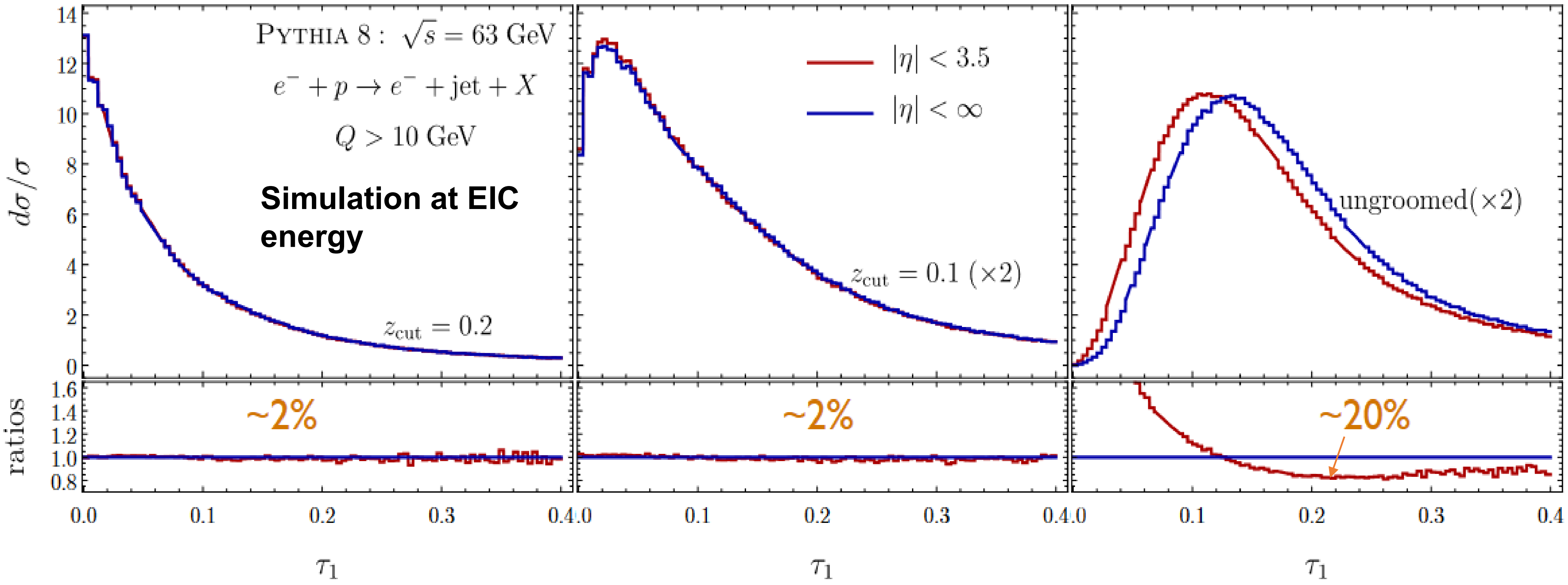


Observables



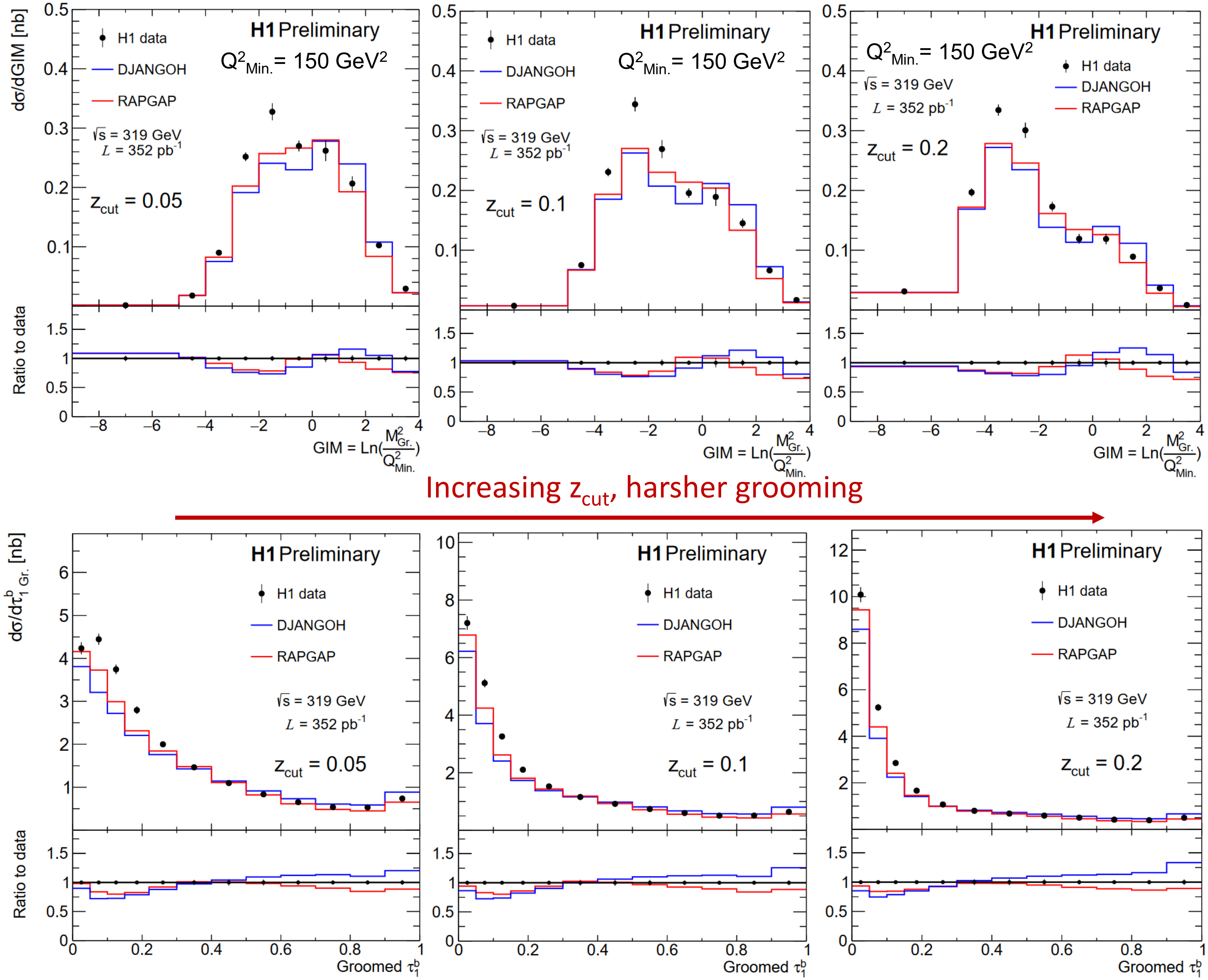


Groomed vs. Ungroomed 1-jettiness



Grooming enables precision measurement!

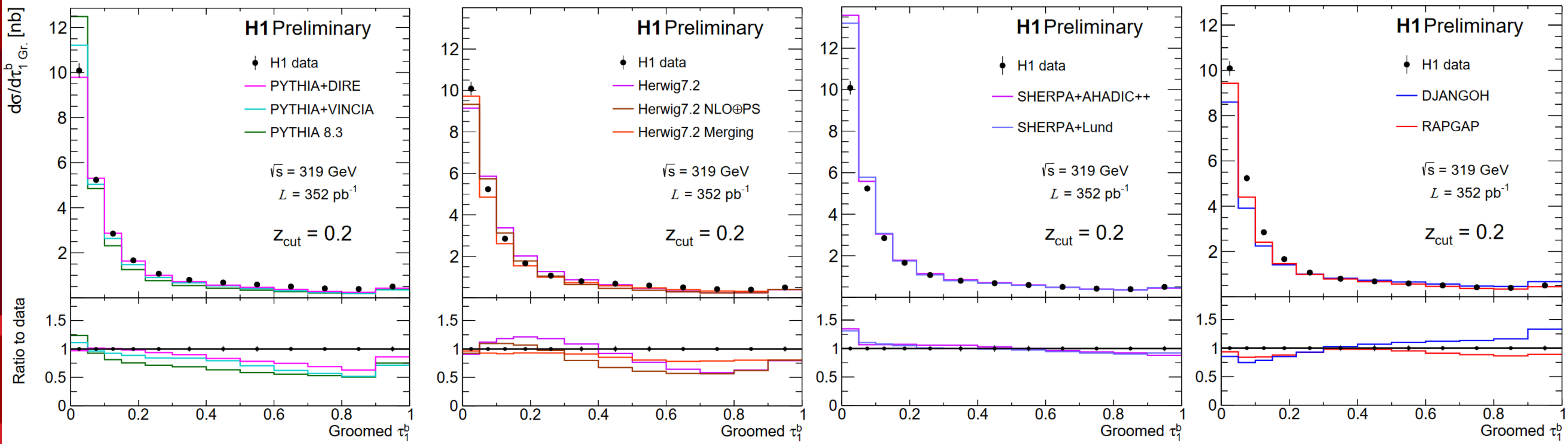
Results



- Data is corrected for real QED ISR and FSR
- Uncertainty on data is statistical ⊕ systematic
  - Dominated by model uncertainty from bin-by-bin correction
- RAPGAP and DJANGOH
  - Standard H1 MCs
  - Both use LEPTO for matrix elements $O(\alpha_S)$
- DJANGOH:
  - Color dipole model PS + string fragmentation
- RAPGAP:
  - DGLAP PS + string fragmentation

Results – Groomed 1 Jettiness

$$\tau_1 = \frac{2}{Q^2} \sum_{i \in \mathrm{gr.\ ent.}} \min(q_B \cdot p_i, q_J \cdot p_i)$$

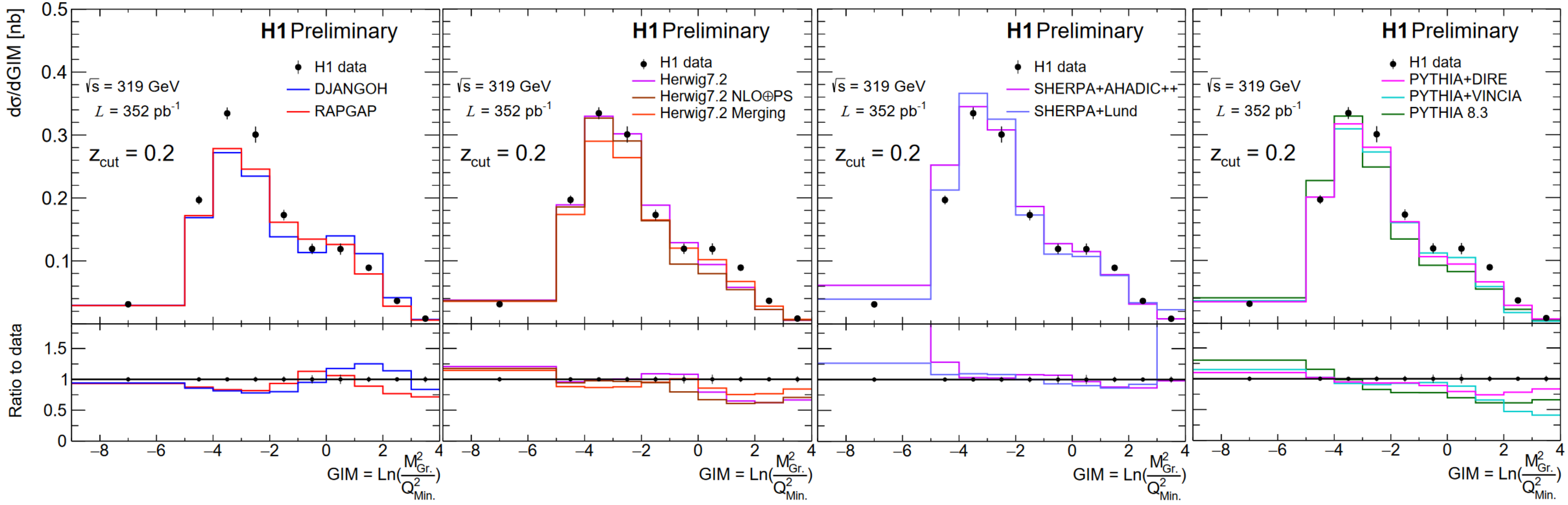


- PYTHIA – Version 8.3
  - VINCIA – Antenna Shower
  - DIRE - Dipole shower + multijet merging
- Herwig – Version 7.2 (Angular-ordered)
  - NLO ⊕PS – AO Shower, subtractive matching
  - Merging - Dipole shower + multijet merging
- SHERPA – Version 2.2.12 (MEPS@NLO)
  - AHADIC++ - Cluster Fragmentation
  - Lund – String Fragmentation

- Best tail region from SHERPA, RAPGAP
  - Fixed-order
- Best peak region from DIRE, Herwig Merging
  - Resummation, parton shower, hadronization

Results – Groomed Invariant Mass

$$M^2_{Gr.} = (\sum_i p_i^{\mu})^2$$

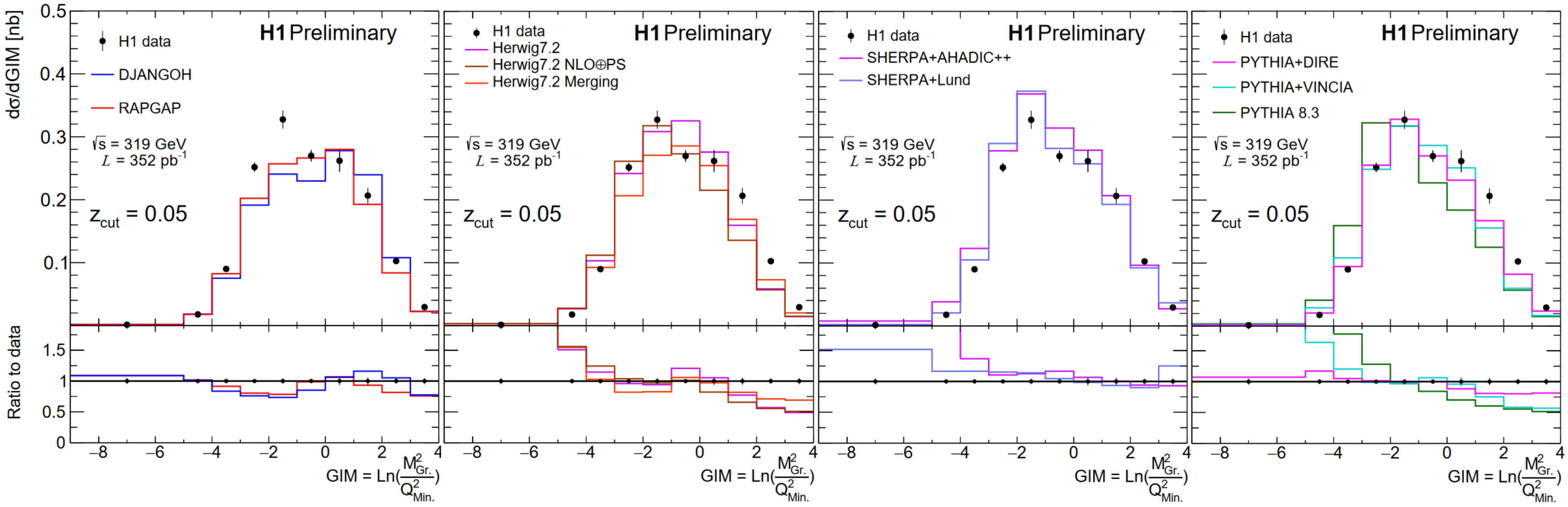


- PYTHIA – Version 8.3
  - VINCIA – Antenna Shower
  - DIRE - Dipole shower + multijet merging
- Herwig – Version 7.2 (Angular-ordered)
  - NLO ⊕PS – AO Shower, subtractive matching
  - Merging - Dipole shower + multijet merging
- SHERPA – Version 2.2.12 (MEPS@NLO)
  - AHADIC++ - Cluster Fragmentation
  - Lund – String Fragmentation

- $Q^2_{Min.}$ = 150 GeV²
- Best high mass region from SHERPA
  - Fixed-order
- Best low mass region from Herwig, DIRE
  - Resummation, hadronization

Results – Groomed Invariant Mass

$$M^2_{Gr.} = (\sum_i p^\mu_i)^2$$



- PYTHIA – Version 8.3
  - VINCIA – Antenna Shower
  - DIRE - Dipole shower + multijet merging
- Herwig – Version 7.2 (Angular-ordered)
  - NLO ⊕PS – AO Shower, subtractive matching
  - Merging - Dipole shower + multijet merging
- SHERPA – Version 2.2.12 (MEPS@NLO)
  - AHADIC++ - Cluster Fragmentation
  - Lund – String Fragmentation

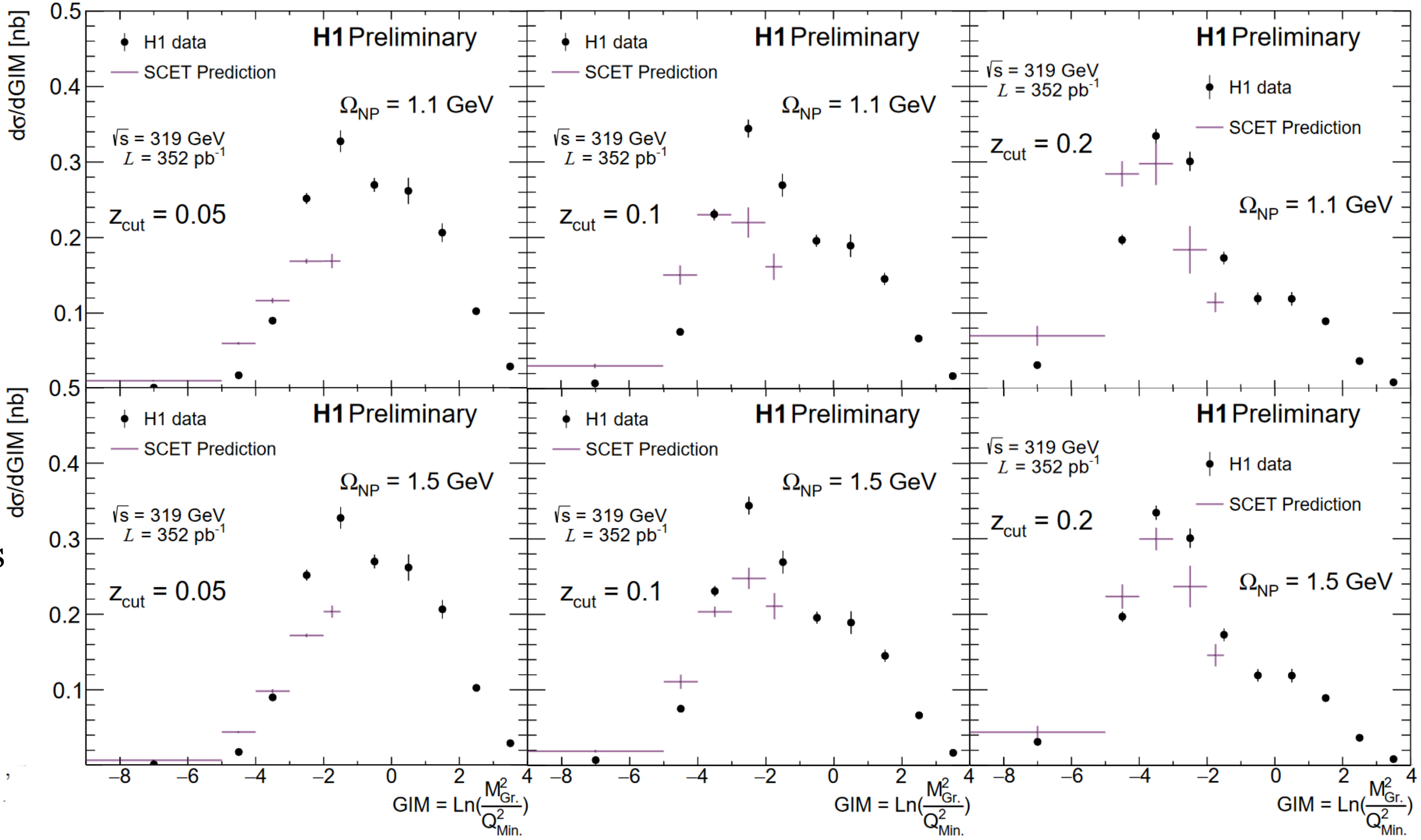
- Generally, predictions become less accurate at lower z_{cut}
  - Less grooming → Less removal of remnant hemisphere radiation
  - Remnant hemisphere is typically less well described by MC models

Results

- Analytic - SCET
  - From Y. Makris [1]
  - Evaluated at two values of Ω_{NP}
    - Shape function mean
  - No fixed-order calculation yet incorporated
- Agreement improves with increasing z_{cut}, Ω_{NP}
  - Non-perturbative effects are significant!
  - Factorization validity improves to higher z_{cut}

$$\frac{d\sigma_{\text{had.}}}{dx dQ^2 dm^2_{\text{gr.}}} = \int d\epsilon \frac{d\sigma}{dx dQ^2 dm^2_{\text{gr.}}}\left(m^2_{\text{gr.}} - \frac{\epsilon^2}{z_{\text{cut}}}\right) f_{\text{mod.}}(\epsilon) \,,$$

$$f_{\text{mod.}}(\epsilon) = N_{\text{mod.}} \frac{4\epsilon}{\Omega^2} \exp\left(\frac{2\epsilon}{\Omega}\right)$$

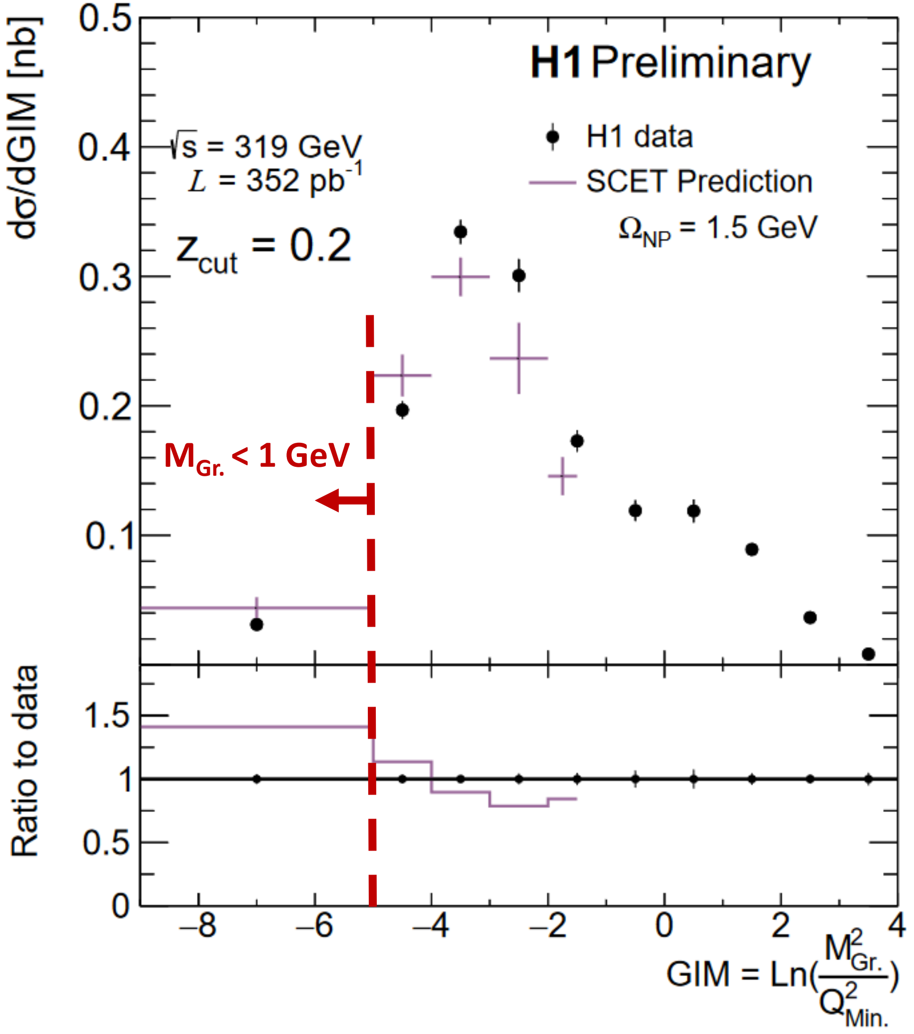


Groomed Invariant Mass - Future

- At small invariant masses, individual hadron masses play a large role
- Analytic predictions most accurate at small masses, in the region defined by:

$$1 \gg z_{\text{cut}} \gg m^2_{\text{gr.}}/Q^2$$

- EIC will have significant advantages in this region
  - Hadron ID, high statistics
  - More differential measurement possible
  - New theory tools+data for high-precision studies of NP sector
- LHeC will access larger region where factorization holds → higher Q^2



Conclusion

- H1 has performed the first measurement of groomed event shapes in DIS
  - H1prelim-22-033
  - See also the ungroomed 1-jettiness preliminary: H1prelim-21-032

- Data has been compared to a variety of MC predictions from SHERPA, PYTHIA, HERWIG, DJANGOH, RAPGAP, as well as analytic predictions from SCET
  - None of the models studied here agree completely with data within uncertainties
  - Signifies that DIS MC models could use improvement, especially in light of upcoming EIC
    - Archived HERA data will necessarily play an important role in this initiative!

Check out other H1 Talks!

Multi-differential Jet Substructure Measurement in High Q^2 Deep-Inelastic Scattering with the H1 Detector
Machine learning-assisted measurement of multi-differential lepton-jet correlations in deep-inelastic scattering with the H1 detector
Probing hadronization and jet substructure with leading particles in jet at H1
Measurement of the 1-jettiness event shape observable in deep-inelastic electron-proton scattering at HERA

Bibliography

[1] - Revisiting the role of grooming in DIS, Y. Makris - arXiv:2101.02708

[2] – Groomed and energy-energy-correlation event shapes at EIC, Y. Makris - LBL Seminar: Oct. 2020

[3] – PYTHIA 8.3 Manual - arXiv:2203.11601

[4] – SHERPA 2.2.12 Manual - sherpa.hepforge.org/doc/SHERPA-MC-2.2.12.html

[5] – herwig.hepforge.org

Backup

region 1: $1 \gg z_{\rm cut} \gg m^2_{\rm gr.}/Q^2$

$$\lambda = \frac{m^2_{\rm gr.}}{Q^2} \,,$$

$$p = (p^+, p^-, p^\perp)$$

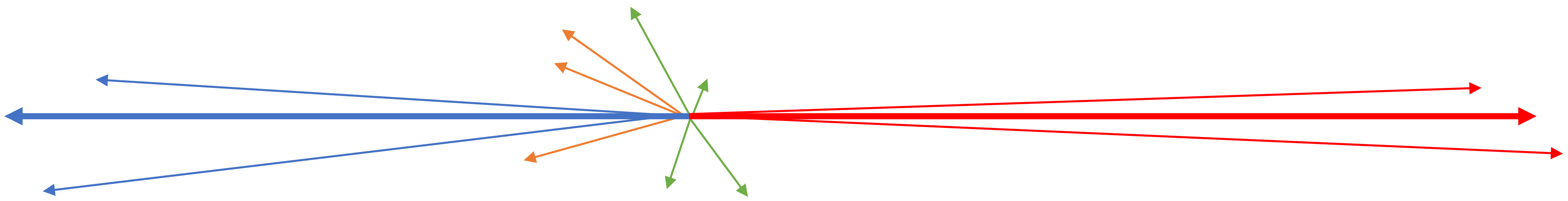
$$\begin{aligned}
n\text{-collinear:} \quad & p_n \sim Q(z_{\rm cut}, 1, \sqrt{z_{\rm cut}}) \,, \\
\text{soft:} \quad & p_s \sim Q z_{\rm cut}(1,1,1) \,, \\
\text{collinear-soft:} \quad & p_{cs} \sim Q(\lambda, z_{\rm cut}, \sqrt{z_{\rm cut}\lambda}) \,, \\
\bar{n}\text{-collinear:} \quad & p_{\bar{n}} \sim Q(1, \lambda, \sqrt{\lambda}) \,,
\end{aligned}$$

Jet Direction (Breit η=-∞) = $\bar{n}$-collinear direction

Beam Direction (Breit η=-∞) = n-collinear direction

Soft Radiation (Isotropic)

Collinear-soft radiation, wide-angle soft radiation mostly along jet direction



$$\frac{d\sigma}{dx\,dQ^2\,dm^2_{\rm gr.}} = H(Q,y,\mu) S(Q z_{\rm cut}, \mu) \sum_f \mathcal{B}_{f/P}(x, Q^2 z_{\rm cut}, \mu) \int de_{\bar{n}}\, de_{cs}\, \delta(m^2_{\rm gr.} - e_{\bar{n}} - e_{cs}) J(e_{\bar{n}}, \mu^2) \mathcal{C}(e_{cs} z_{\rm cut}, \mu^2) \times \left[1 + \mathcal{O}\left(z_{\rm cut}, \frac{m^2_{\rm gr.}}{z_{\rm cut} Q^2}\right)\right] , \quad (15)$$

In Region 1, shape of distribution depends only on jet and collinear-soft functions, which are independent of Q region 1: $1 \gg z_{\rm cut} \gg m^2_{\rm gr.}/Q^2$

$$\lambda = \frac{m^2_{\rm gr.}}{Q^2}\,,$$

$$p = (p^+, p^-, p^\perp)$$

$$n\text{-collinear:}\quad p_n \sim Q(z_{\rm cut}, 1, \sqrt{z_{\rm cut}})\,,$$

$$\text{soft:}\quad p_s \sim Q z_{\rm cut}(1,1,1)\,,$$

$$\text{collinear-soft:}\quad p_{cs} \sim Q(\lambda, z_{\rm cut}, \sqrt{z_{\rm cut}\lambda})\,,$$

$$\bar{n}\text{-collinear:}\quad p_{\bar{n}} \sim Q(1, \lambda, \sqrt{\lambda})\,,$$

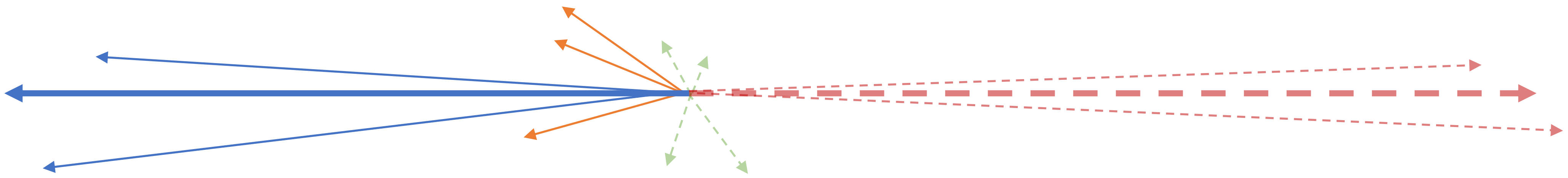
Jet Direction (Breit η=-∞) = $\bar{n}$-collinear direction

Beam Direction (Breit η=-∞) = n-collinear direction

Soft Radiation (Isotropic)

Collinear-soft radiation, wide-angle soft radiation mostly along jet direction

Grooming causes only Jet and Collinear-soft radiation to contribute to shape of distribution in the single-jet (low invariant mass) limit

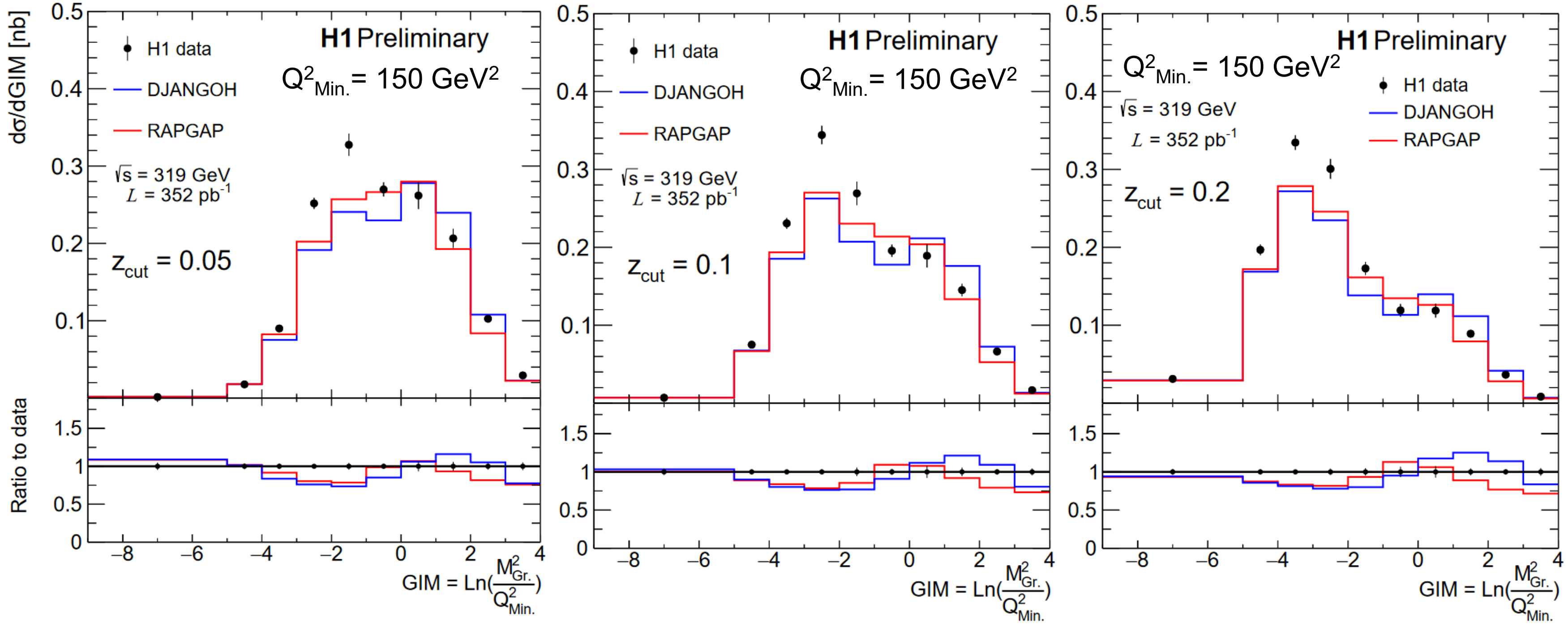


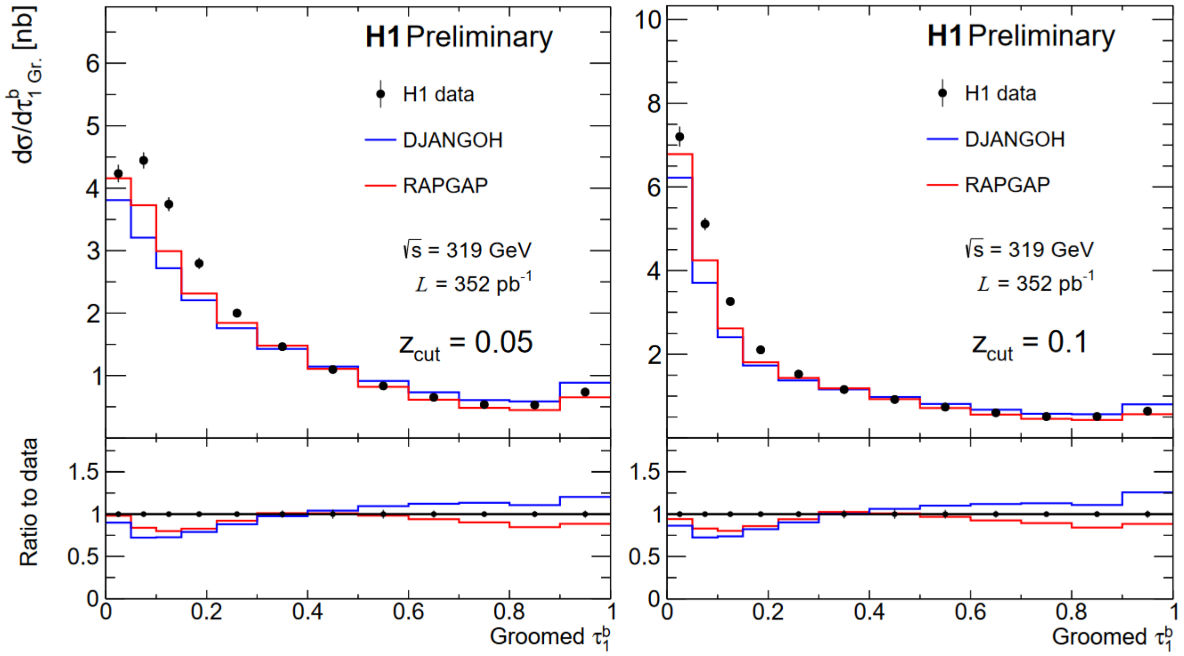
$$\frac{d\sigma}{dx\,dQ^2\,dm^2_{\rm gr.}} = H(Q,y,\mu)\,S(Qz_{\rm cut},\mu)\sum_f \mathcal{B}_{f/P}(x, Q^2 z_{\rm cut}, \mu)\int de_{\bar{n}}\,de_{cs}\;\delta(m^2_{\rm gr.} - e_{\bar{n}} - e_{cs})\,J(e_{\bar{n}},\mu^2)\,\mathcal{C}(e_{cs} z_{\rm cut}, \mu^2)$$
$$\times\left[1 + \mathcal{O}\left(z_{\rm cut}, \frac{m^2_{\rm gr.}}{z_{\rm cut}Q^2}\right)\right], \quad (15)$$

In Region 1, shape of distribution depends only on jet and collinear-soft functions, which are independent of Q

Results

- Data is corrected for real QED ISR and FSR
- Uncertainty on data is statistical ⊕ systematic
- RAPGAP and DJANGOH
  - Standard H1 MCs
  - Both use LEPTO for matrix elements $O(\alpha_S)$
- DJANGOH:
  - Color dipole model for parton shower + string fragmentation
- RAPGAP:
  - DGLAP parton shower + string fragmentation

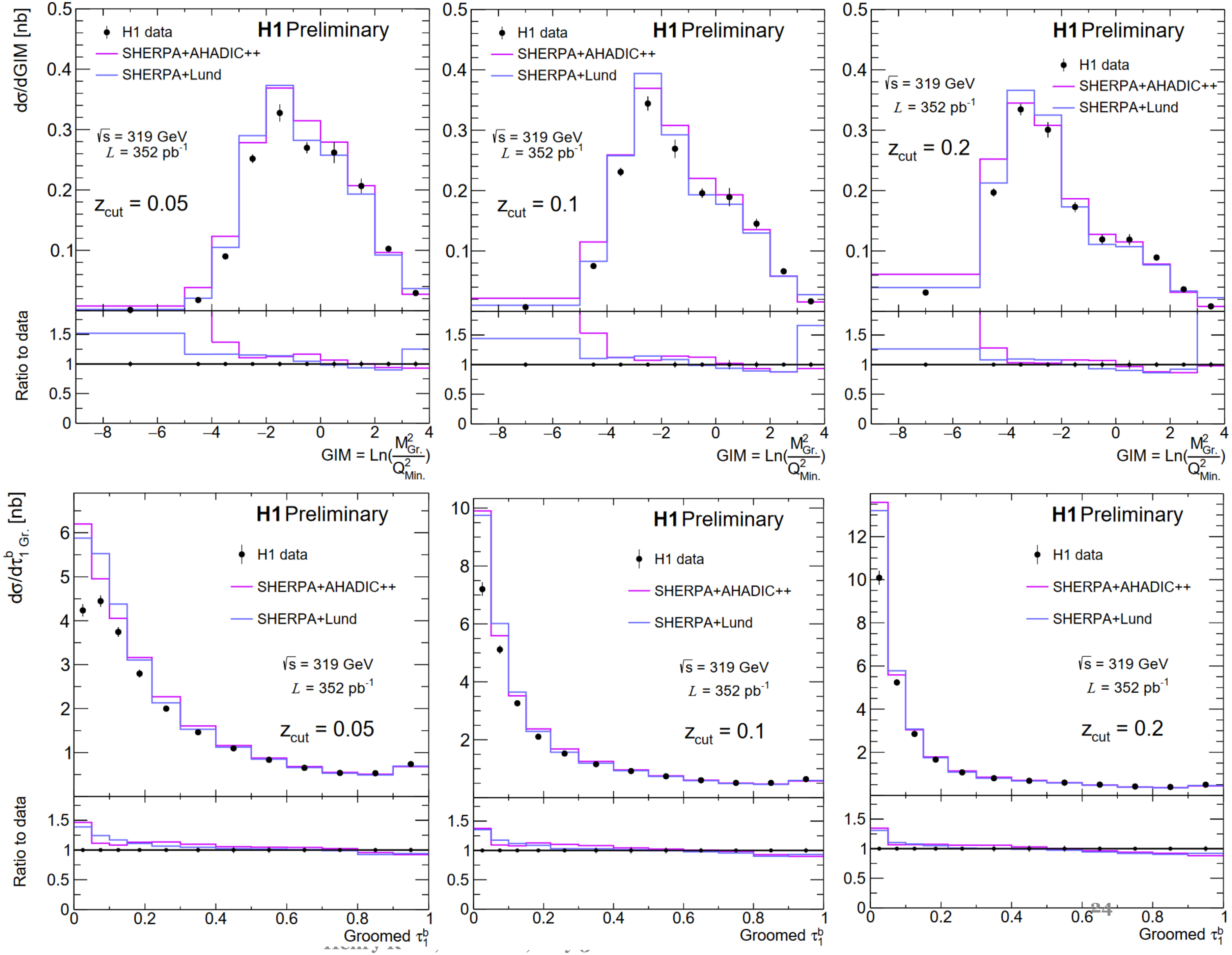




Data is corrected for real QED ISR and FSR
Uncertainty on data is statistical ⊕ systematic
RAPGAP and DJANGOH
- Standard H1 MCs
- Both use LEPTO for matrix elements $O(\alpha_S)$
DJANGOH:
- Color dipole model for parton shower + string fragmentation
RAPGAP:
- DGLAP parton shower + string fragmentation

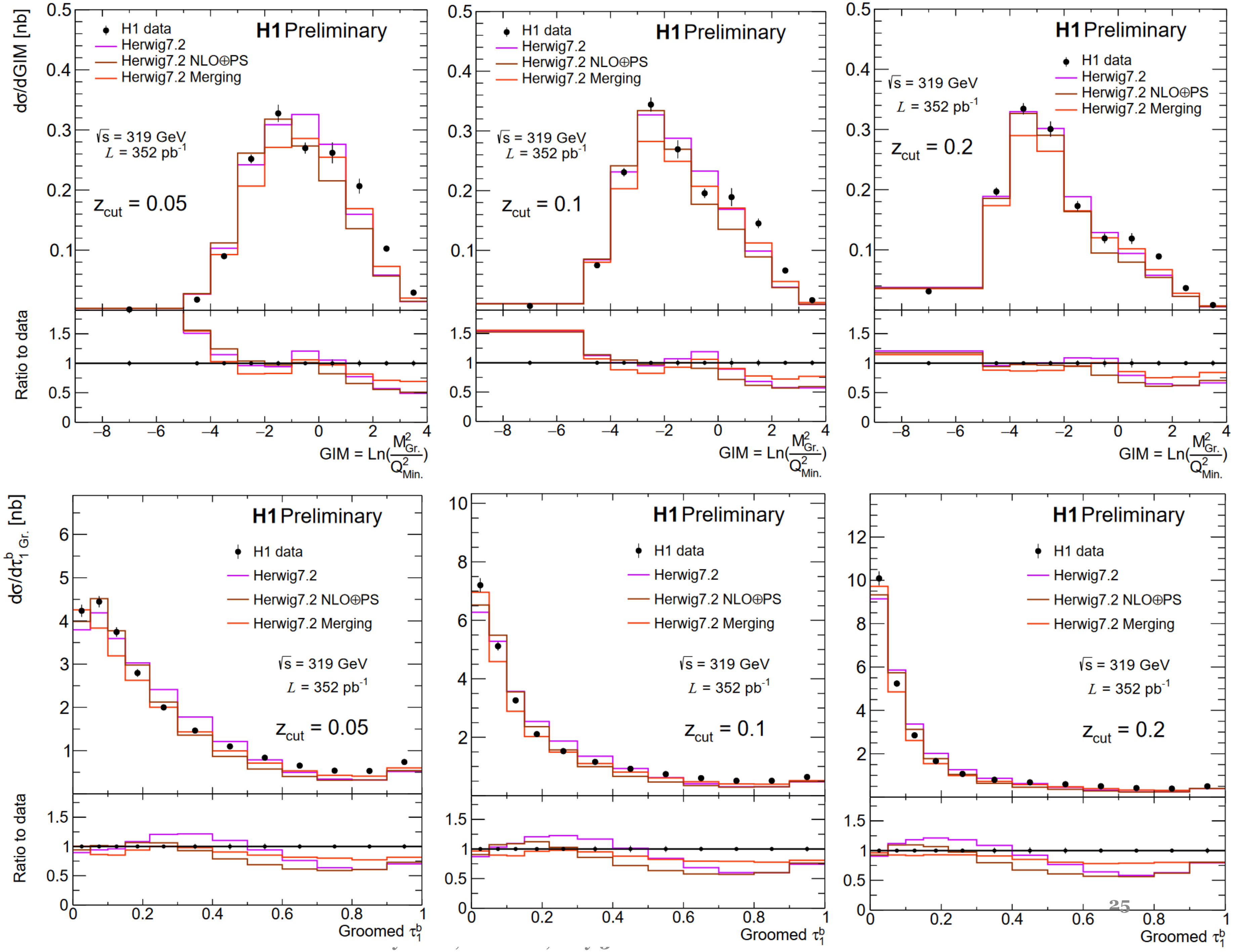
Results

- SHERPA
  - Version 2.2.12
  - MEPS@NLO
- AHADIC++:
  - SHERPA native cluster hadronization model
- Lund:
  - Lund string model from PYTHIA
- Both models provide good description of fixed-order region



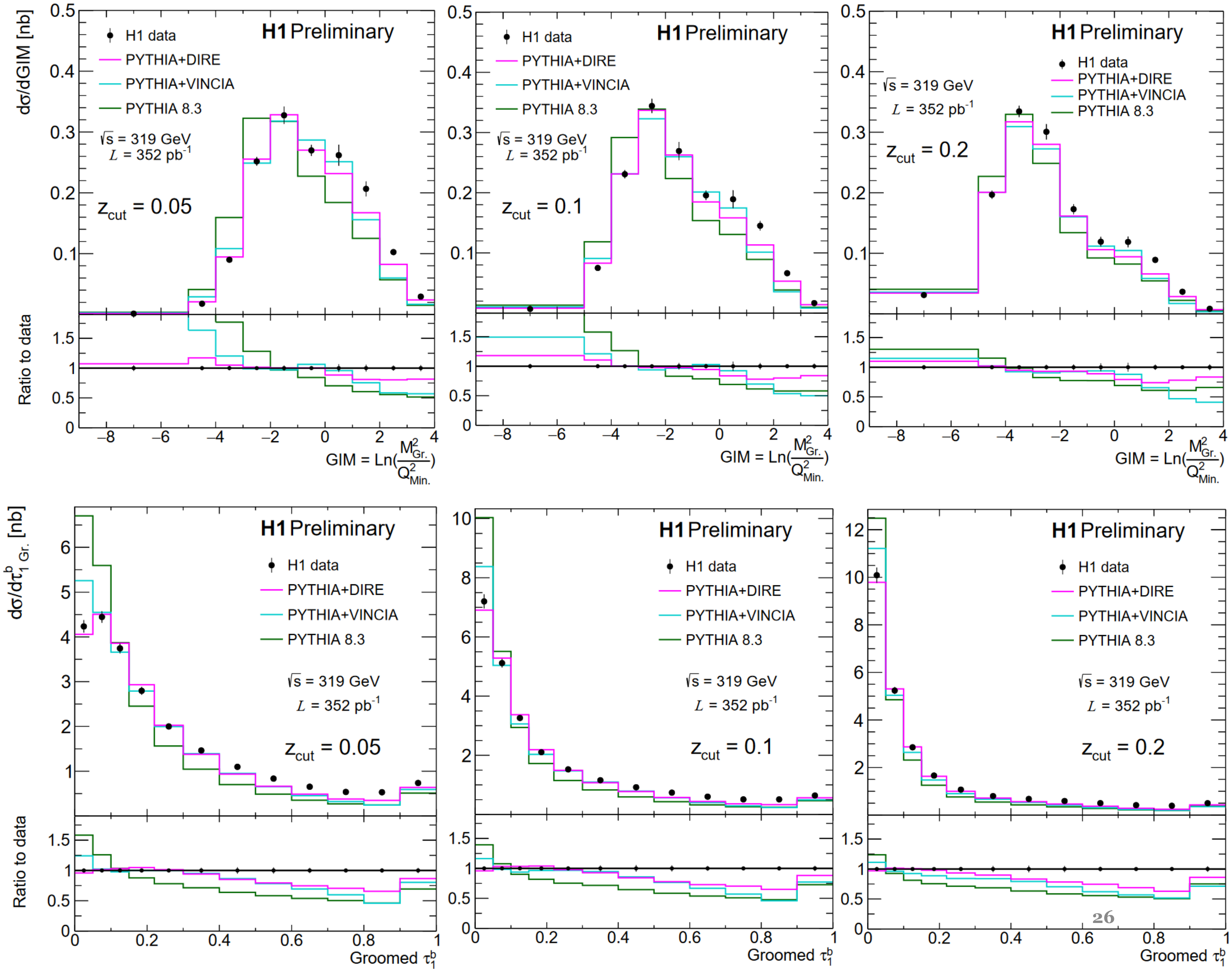
Results

- Herwig
  - Version 7.2.2
- NLO⊕PS:
  - Herwig internal implementation of MC@NLO via Matchbox
- Merging:
  - Dipole shower with multijet merging



Results

- PYTHIA
  - Version 8.3
  - No external matrix elements
- DIRE:
  - Dipole resummation
  - Excellent description in resummation region
- VINCIA:
  - Antenna shower

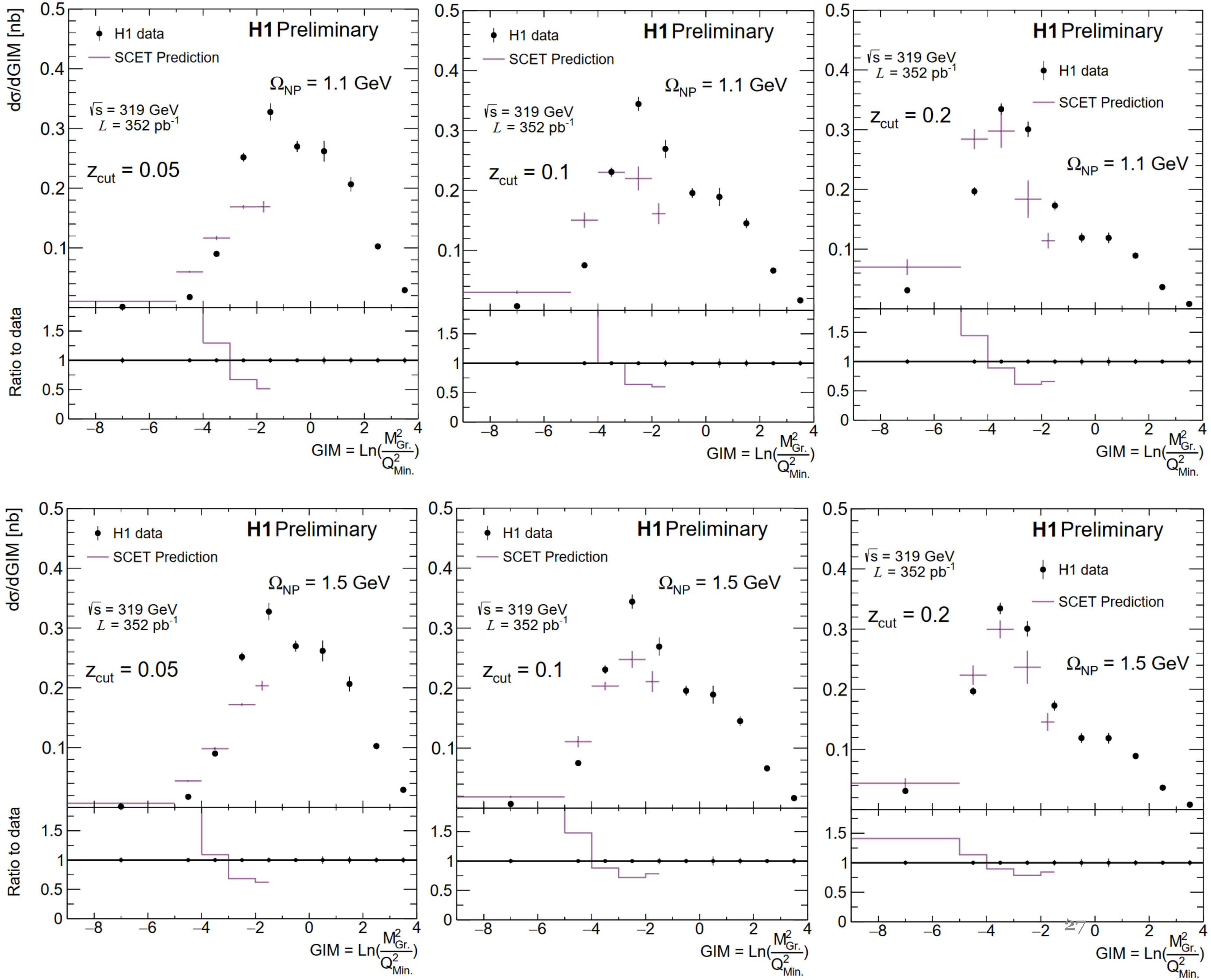


Results

- Analytic - SCET
  - From Y. Makris [1]
  - Evaluated at two values of Ω_{NP}
    - Shape function mean
  - No fixed-order calculation yet incorporated
- Agreement improves with increasing z_{cut}, Ω_{NP}
  - Non-perturbative effects are significant!
  - Factorization validity improves to higher z_{cut}

$$\frac{d\sigma_{\text{had.}}}{dx dQ^2 dm^2_{\text{gr.}}} = \int d\epsilon \frac{d\sigma}{dx dQ^2 dm^2_{\text{gr.}}}\left(m^2_{\text{gr.}} - \frac{\epsilon^2}{z_{\text{cut}}}\right) f_{\text{mod.}}(\epsilon) \, ,$$

$$f_{\text{mod.}}(\epsilon) = N_{\text{mod.}} \frac{4\epsilon}{\Omega^2} \exp\left(\frac{2\epsilon}{\Omega}\right)$$



Centauro

$$d_{ij} = \min(p_{Ti}^{2p}, p_{Tj}^{2p})\Delta R_{ij}^2/R^2\,, \qquad d_{iB} = p_{Ti}^{2p}$$

vs.

$$d_{ij} = (\Delta\bar{\eta}_{ij})^2 + 2\bar{\eta}_i\bar{\eta}_j(1-\cos\Delta\phi_{ij})\,,$$

where

$$\bar{\eta}_i \equiv 2\sqrt{1+\frac{q\cdot p_i}{x_B P\cdot p_i}} \xrightarrow[\text{frame}]{\text{Breit}} \frac{2p_i^{\perp}}{p_i^{+}}\,,$$

vs.

$$d_{ij} = \min(E_i^{2p}, E_j^{2p})\frac{1-c_{ij}}{1-c_R}\,, \qquad d_{iB} = E_i^{2p}$$

where $c_{ij} = \cos\theta_{ij}$ and $c_R = \cos R$.

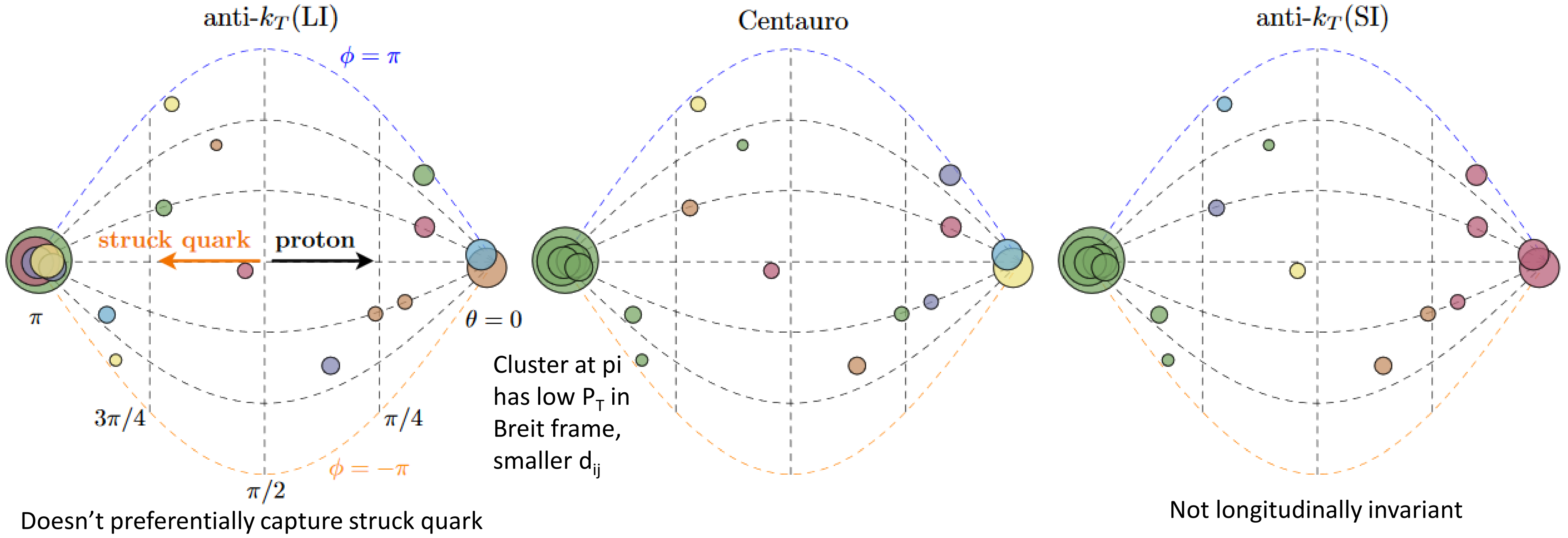
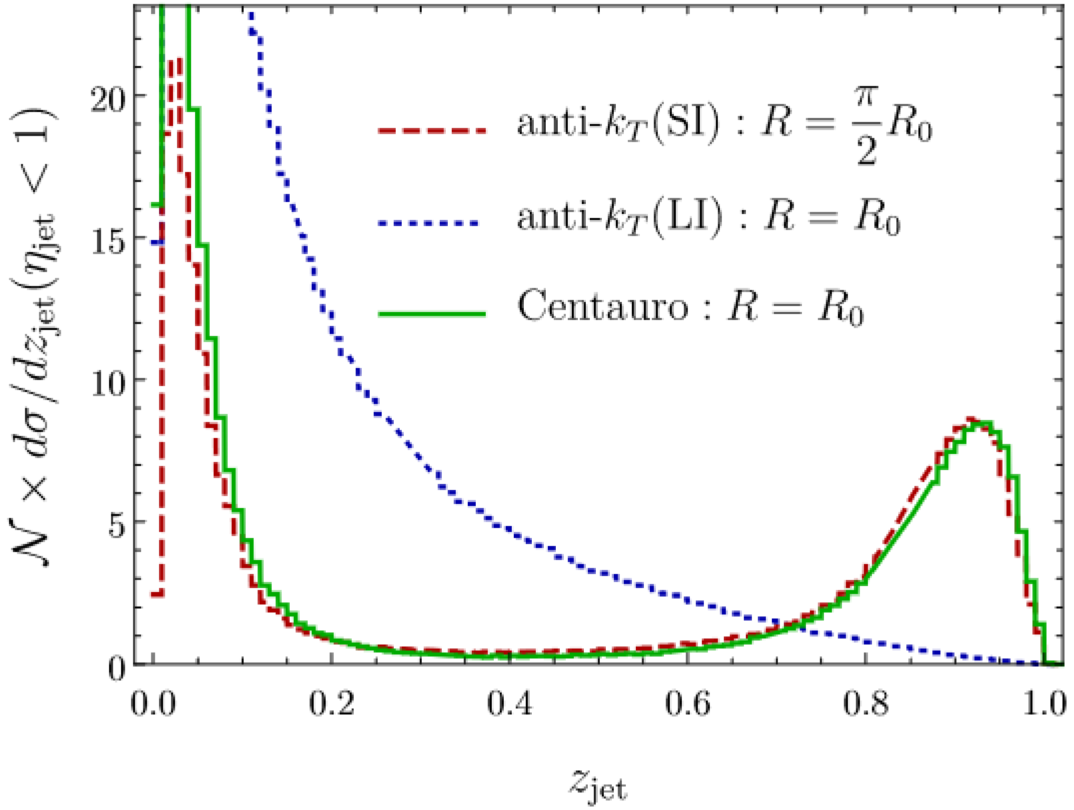


Figure 2. Jet clustering in the Breit frame using the longitudinally-invariant anti-k_T(LI), Centauro, and spherically-invariant anti-k_T(SI) algorithms in a DIS event simulated with PYTHIA 8. Each particle is illustrated as a disk with area proportional to its energy and the position corresponds to the direction of its momentum projected onto the unfolded sphere about the hard-scattering vertex. The vertical dashed lines correspond to constant θ and curved lines to constant ϕ. All the particles clustered into a given jet are colored the same.



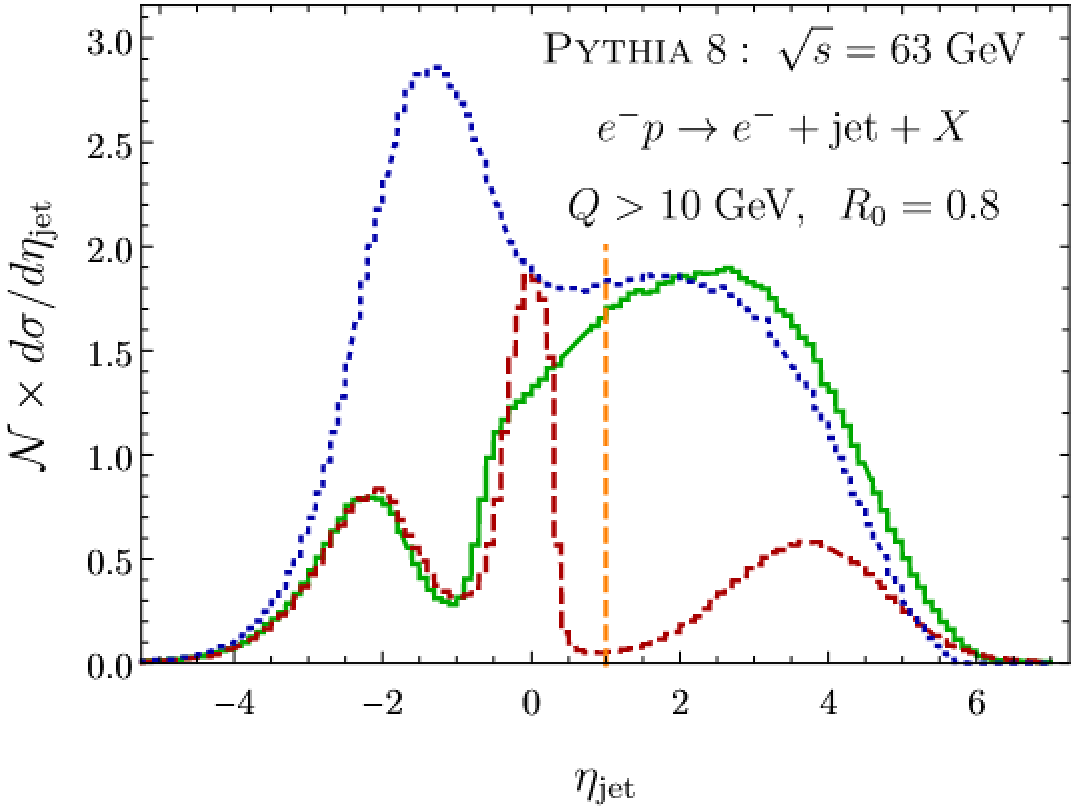


Figure 4. Pseudorapidity (top panel) and momentum fraction $z_{\rm jet}$ (bottom panel) of jets clustered with anti-k_T(LI), anti-k_T(LI) and Centauro algorithms in the Breit frame. Here $\mathcal{N}$ is an overall normalization constant chosen to improve readability and is the same for all curves in a graph.

Centauro

$$d_{ij} = \min(p_{Ti}^{2p}, p_{Tj}^{2p})\Delta R_{ij}^2/R^2 \, , \qquad d_{iB} = p_{Ti}^{2p}$$

vs.

$$d_{ij} = (\Delta\bar{\eta}_{ij})^2 + 2\bar{\eta}_i\bar{\eta}_j(1-\cos\Delta\phi_{ij}) \, ,$$

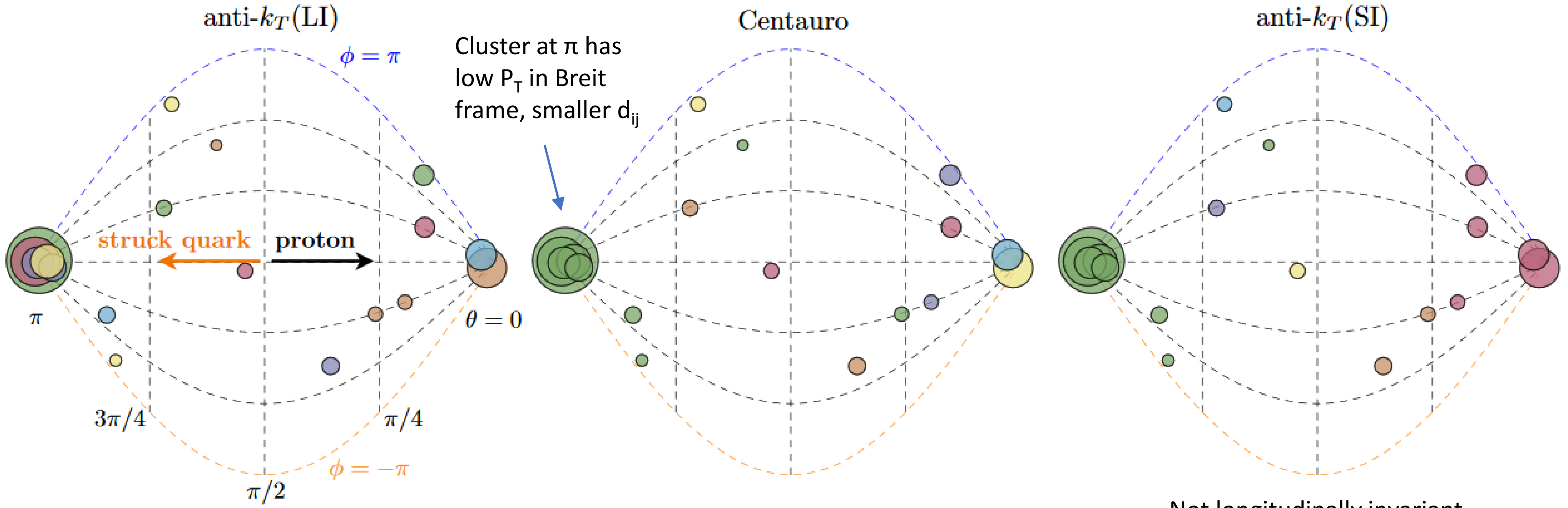
where

$$\bar{\eta}_i \equiv 2\sqrt{1+\frac{q\cdot p_i}{x_B P\cdot p_i}} \xrightarrow[\text{frame}]{\text{Breit}} \frac{2p_i^\perp}{p_i^+} \, ,$$

vs.

$$d_{ij} = \min(E_i^{2p}, E_j^{2p})\frac{1-c_{ij}}{1-c_R} \, , \qquad d_{iB} = E_i^{2p}$$

where $c_{ij} = \cos\theta_{ij}$ and $c_R = \cos R$.



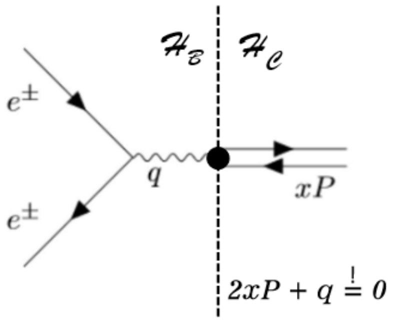
Doesn't preferentially capture struck quark

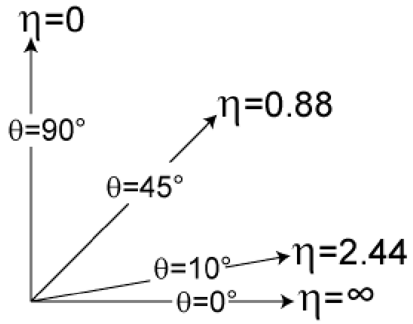
Not longitudinally invariant

Figure 2. Jet clustering in the Breit frame using the longitudinally-invariant anti-k_T(LI), Centauro, and spherically-invariant anti-k_T(SI) algorithms in a DIS event simulated with PYTHIA 8. Each particle is illustrated as a disk with area proportional to its energy and the position corresponds to the direction of its momentum projected onto the unfolded sphere about the hard-scattering vertex. The vertical dashed lines correspond to constant θ and curved lines to constant ϕ. All the particles clustered into a given jet are colored the same.

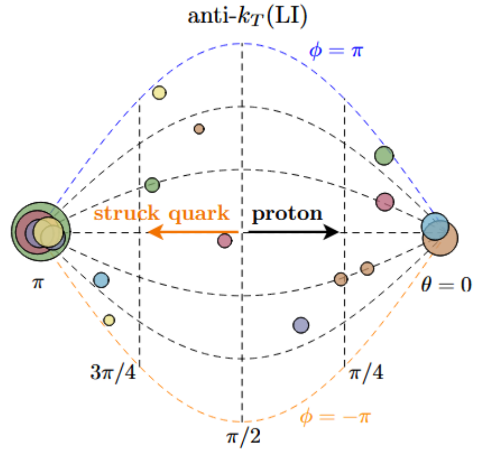
Stony Brook University

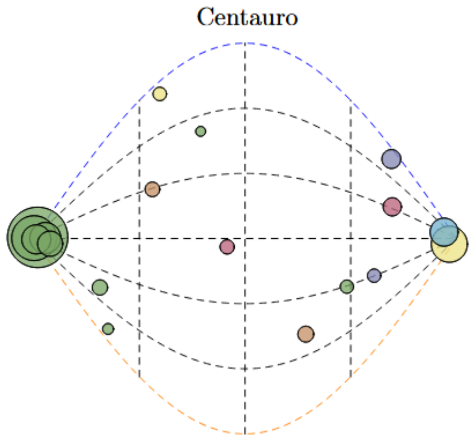
Centauro

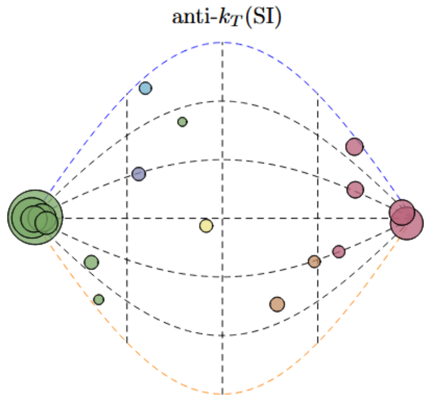




- Jet algorithm with asymmetric clustering measure
  - Treat current hemisphere and beam hemisphere differently
- Typical longitudinally-invariant jet algorithms cluster in (rapidity, azimuthal angle) space and fail to capture the born-level configuration in the Breit frame
  - Particles close to struck-parton direction have divergent rapidity, and therefore divergent distance between each other!
    - Makes study of single-jet Born level configuration impossible!
- Use spherically invariant clustering (polar angle, azimuthal angle) in the struck-parton direction and longitudinally invariant in beam direction
- Tends to create one hard jet in struck-parton direction and many weak single particle jets in beam direction, which can easily be filtered away







z_{cut}	0.05	0.1	0.2
Pythia8.3	0.31%	1.3%	6.3%
Pythia+Vincia	0.52%	1.7%	6.9%
Pythia+DIRE	0.47%	1.2%	5.6%
SHERPA+AHADIC++	0.03%	0.31%	3.6%
SHERPA+LUND	0.09%	0.59%	4.9%
HERWIG	0.038%	0.36%	3.6%
HERWIG+Merging	0.04%	0.39%	3.6%
HERWIG+MC@NLO	0.04%	0.39%	3.8%
DJANGO (Gen.)	0.09%	0.5%	4.0%
RAPGAP (Gen.)	0.07%	0.4%	3.7%
DJANGO (Det.)	1.0%	2.3%	7.8%
RAPGAP (Det.)	0.9%	2.2%	7.6%
H1 Data	1.0%	2.3%	7.7%

Table 1: Percentage of events that fail grooming. Rapgap and Djangoh are listed for both detector and generator level.