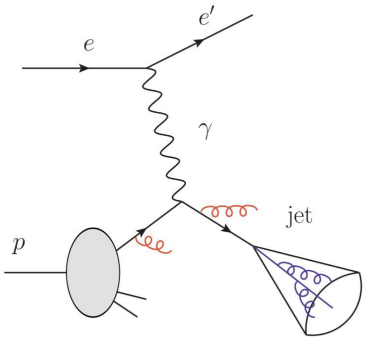
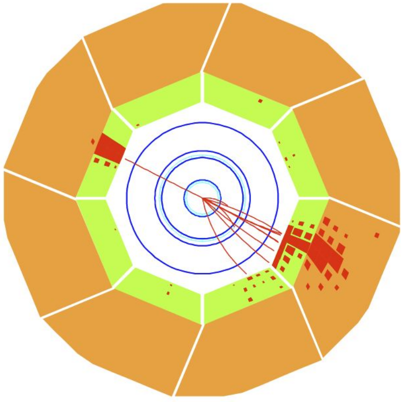
Multi-differential Jet Substructure Measurement in High Q^2 DIS Events with HERA-II Data

Vinicius M. Mikuni, Ben Nachman

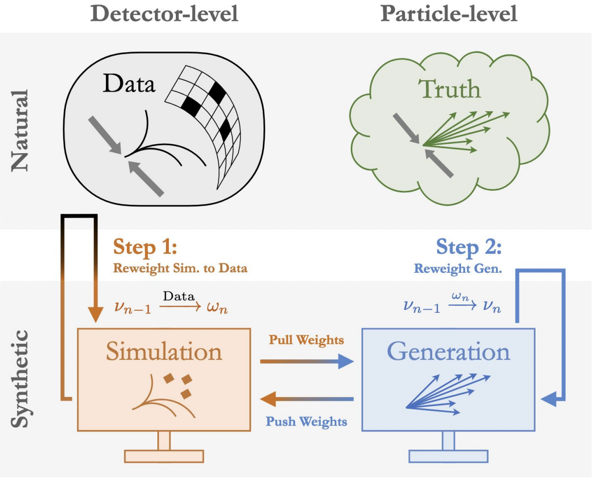
Review

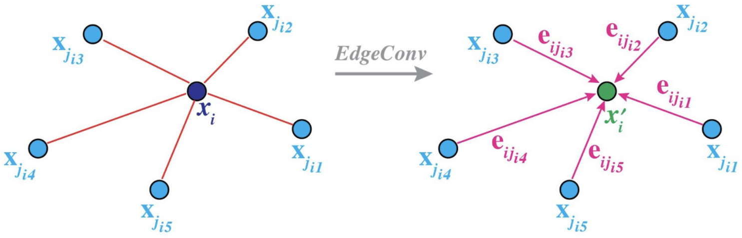
1: Definition of measure observables



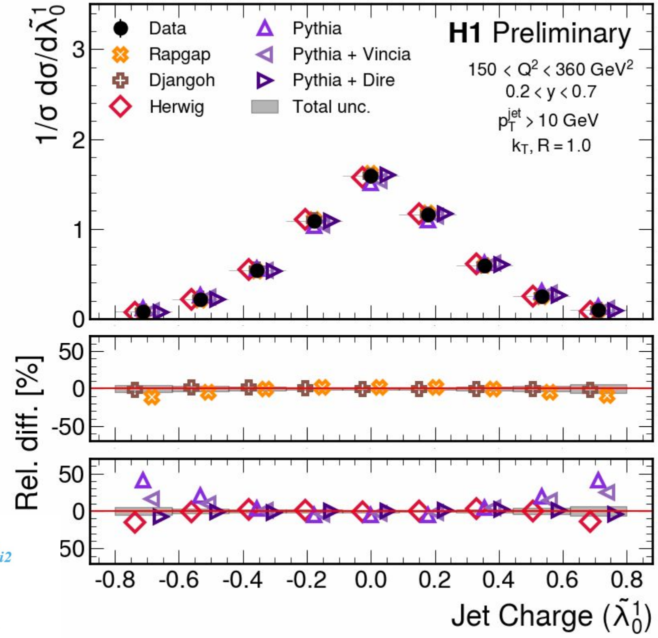


2: Unfolding methodology





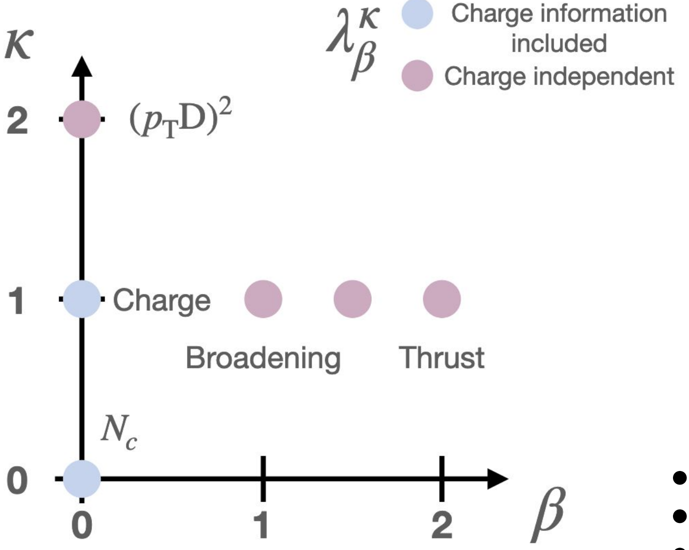
3: Multi-differential cross section results

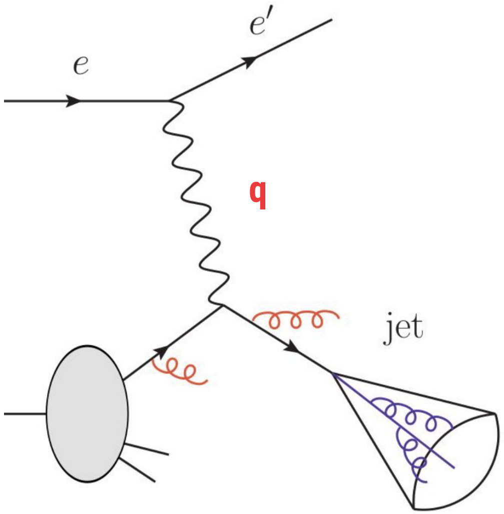


Jet angularities

Use jet observables to study different aspects of QCD physics:

- IRC safe λ^1_a, a = [0,0.5,1] and unsafe p_T**D** angularities
- Charge dependent observables: Q_j and N_c
- Study the evolution of the observables with energy scale $Q^2 = -q^2$





- z_i: longitudinal momentum fraction
- q_i: charge
- R_i distance from jet axis in (eta,phi)

$$\lambda^\kappa_\beta = \sum_{i\in \text{jet}} z_i^\kappa \left(\frac{R_i}{R_0}\right)^\beta$$

$$\tilde{\lambda}^\kappa_0 = Q_\kappa = \sum_{i\in \text{jet}} q_i \times z_i^\kappa .$$

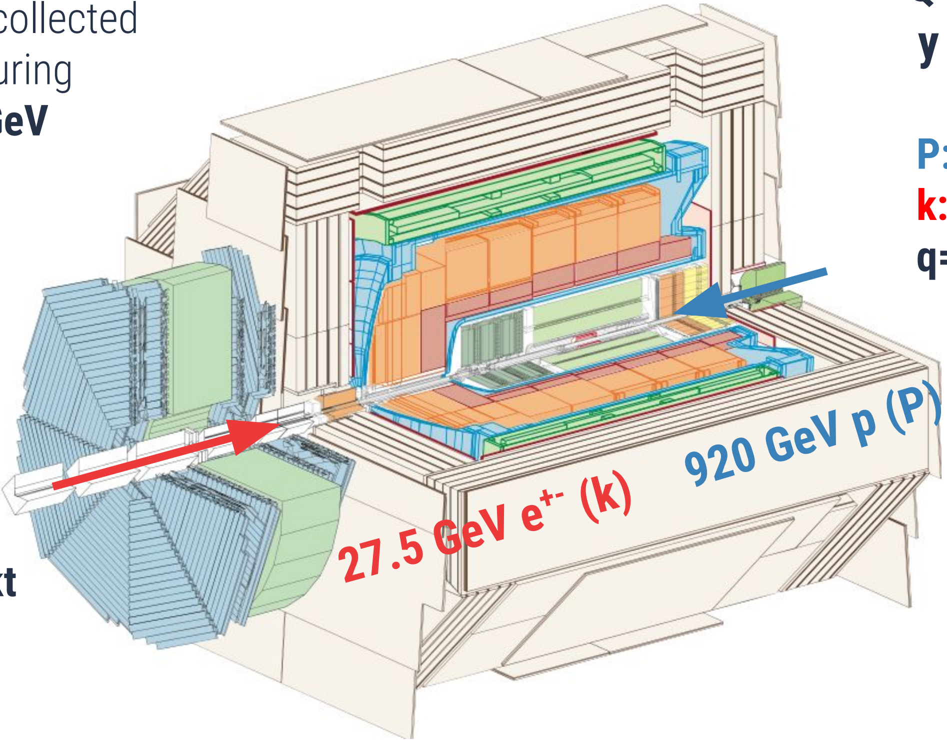
Experimental setup

Using **228 pb^{-1}** of data collected by the **H1 Experiment** during **2006** and **2007** at **318 GeV center-of-mass energy**

Phase space definition:

- $0.2 < y < 0.7$
- $Q^2 > 150\ \mathrm{GeV}^2$
- Jet $p_T > 10$ GeV
- $-1 < \eta_{lab} < 2.5$

Jets are clustered with **kt** algorithm with **R=1.0**



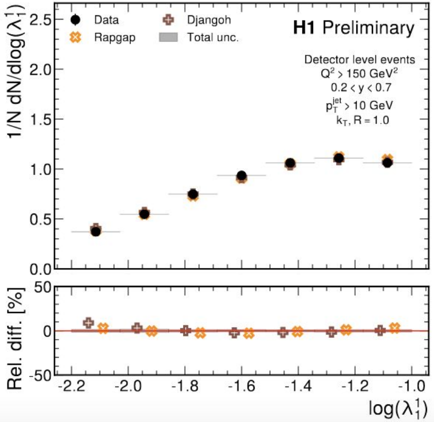
$$Q^2 = -q^2$$

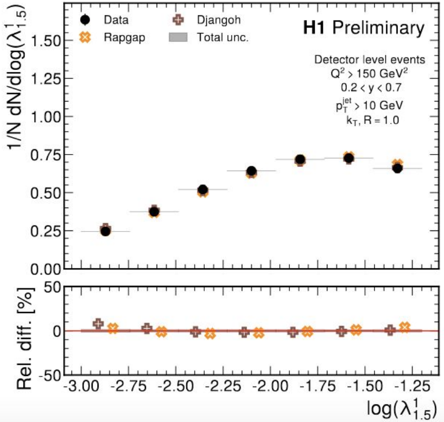
$$y = Pq / pk$$

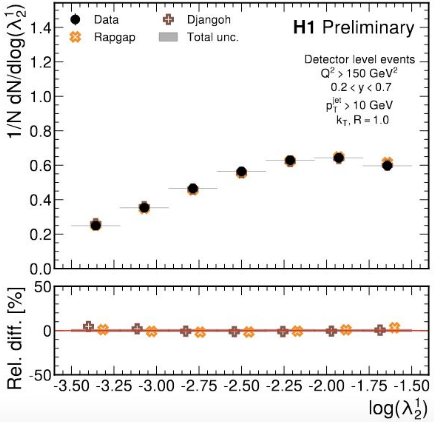
P: incoming proton 4-vector
k: incoming electron 4-vector
q=k-k' : 4-momentum transfer

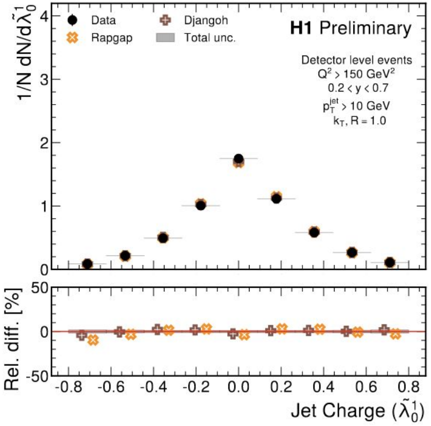
Reconstructed hadrons using combined detector information: **energy flow algorithm**

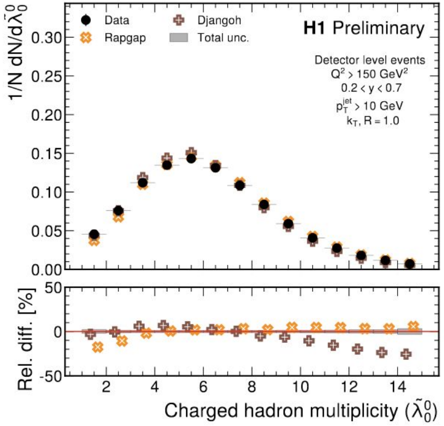
Total experimental uncertainty at reconstruction level at the **% level**!

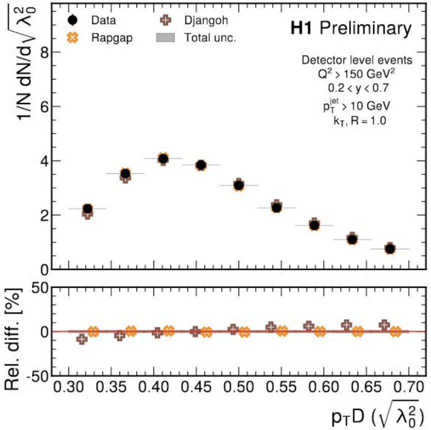








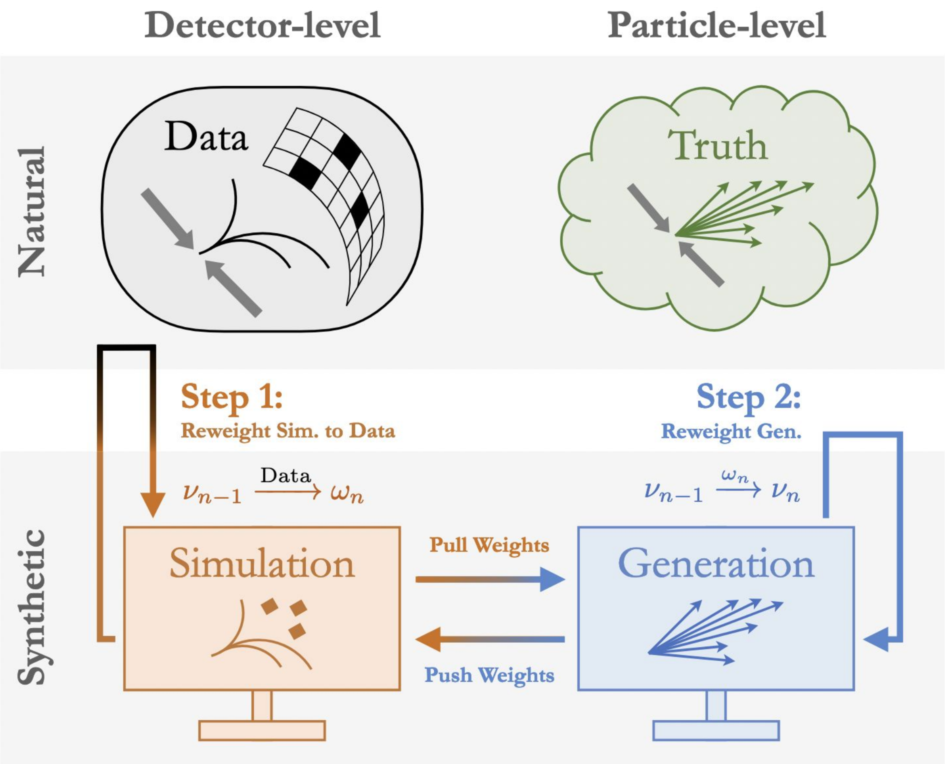




Part 2

Unfolding strategy

Omnifold*



2 step iterative approach

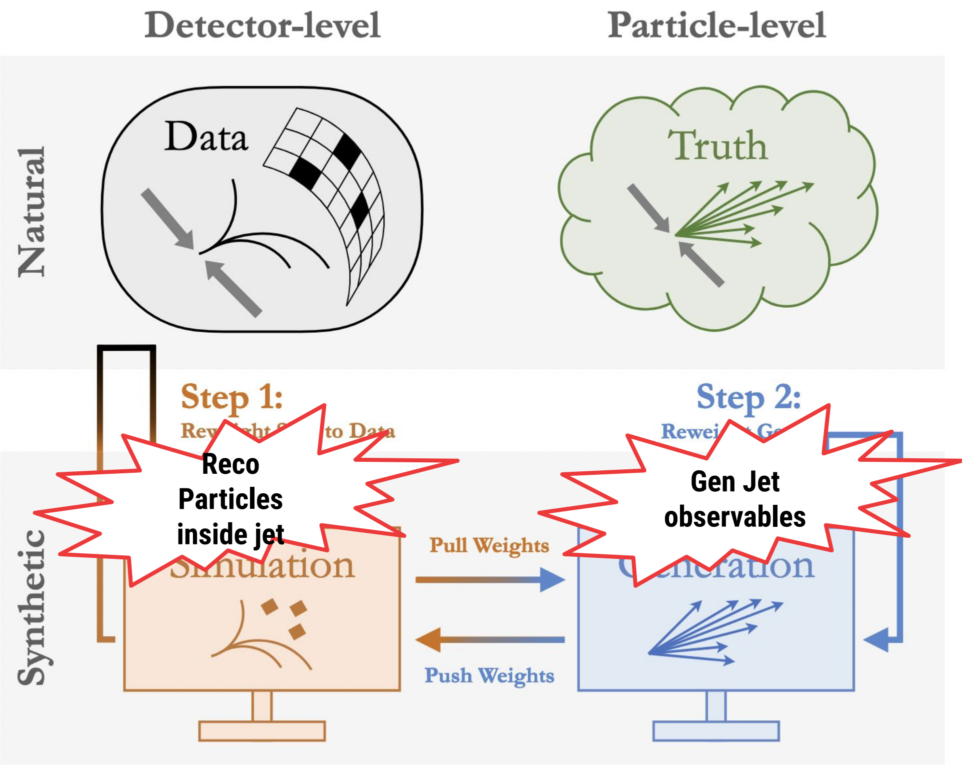
- Simulated events after detector interaction are reweighted to match the data
- Create a “new simulation” by transforming weights to a proper function of the generated events

Machine learning is used to approximate **2** likelihood functions:

- **reco MC to Data** reweighting
- **Previous** and **new Gen** reweighting

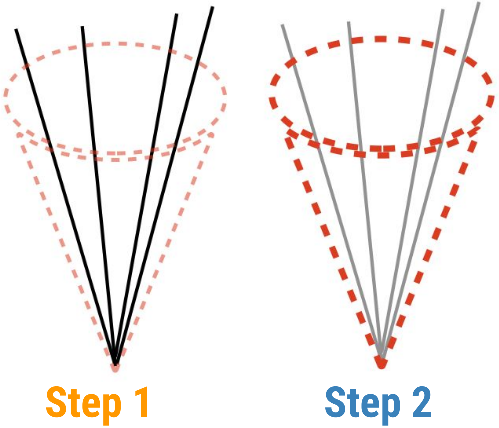
* Andreassen et al. PRL 124, 182001 (2020)

Omnifold



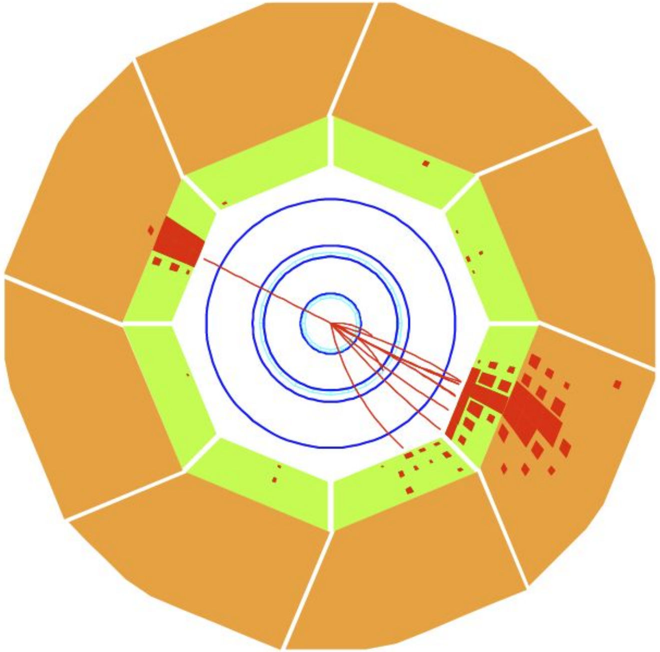
Different input levels for each step

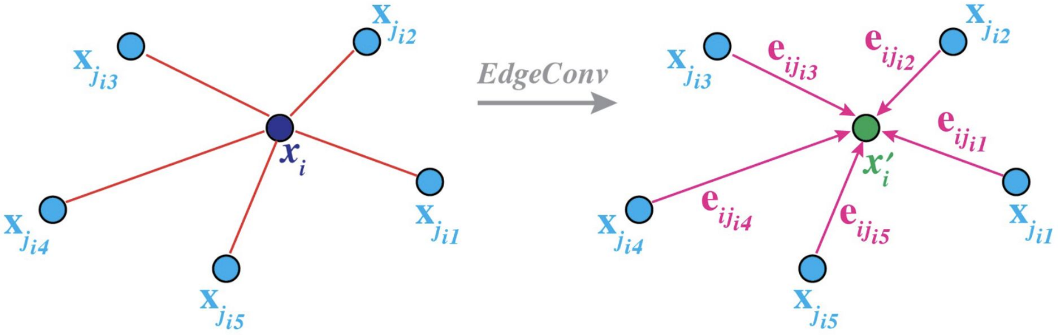
- **Step 1** particles are used as inputs
- **Step 2** uses the set of observables planned to unfold



Extracting particle information

- Particle information is extracted using a **Point cloud transformer*** model
- Model takes **kinematic properties** of particles as inputs
- Connect **k=10** nearest neighbors in η-φ to learn the relationship between particles.
- Built in symmetries: **permutation invariance**
- Consider up to **30** particles per jet

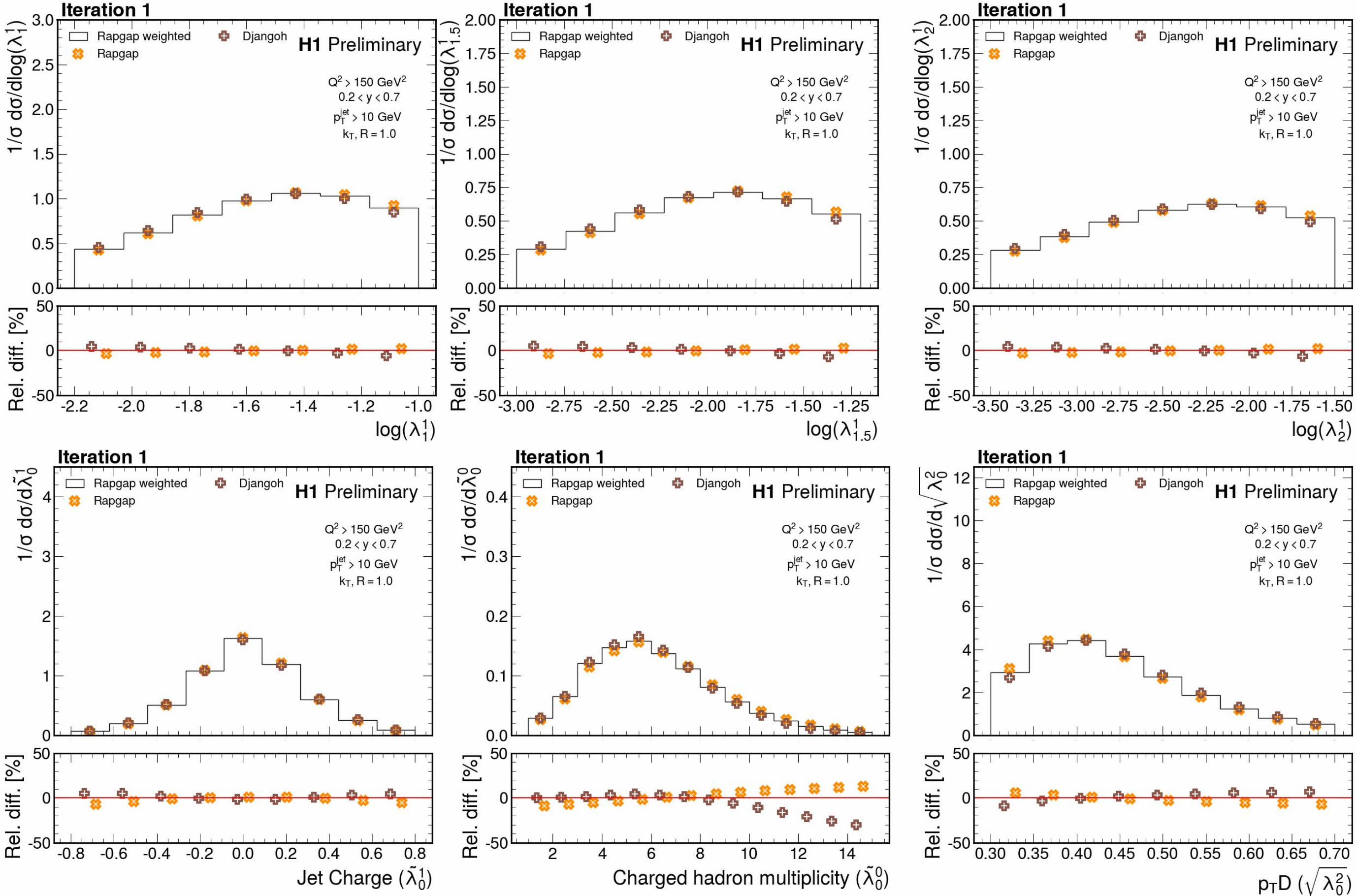




* V. Mikuni and F. Canelli 2021 *Mach. Learn.: Sci. Technol.* **2** 035027

Closure test

All distributions are unfolded **simultaneously without binning and without jet substructure information used at reco level!**



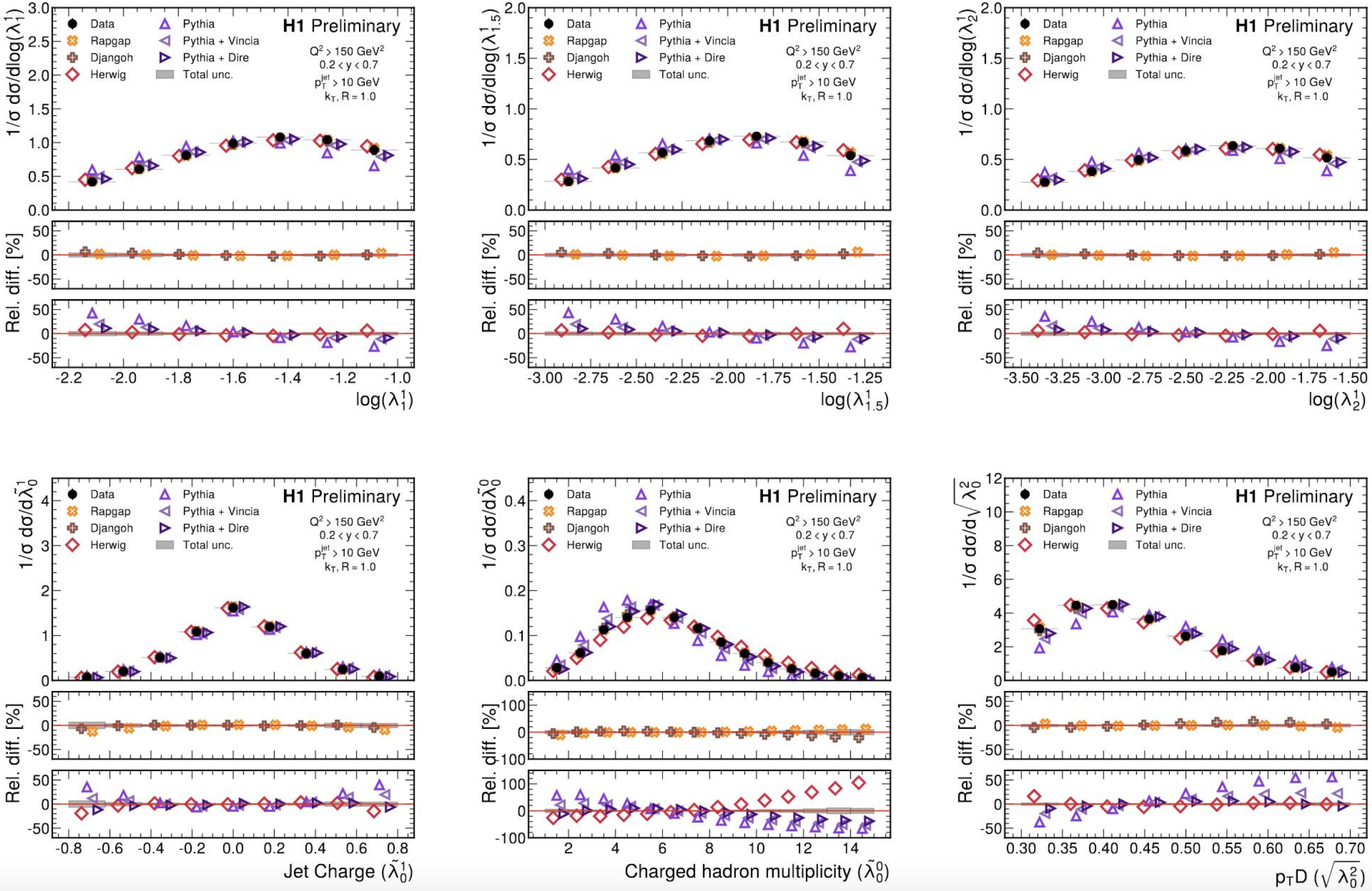
Verify the model **consistency**: start from the **Rapgap** simulation and unfold the response based on the **Djangoh** simulation

Total of **6 iterations** used to derive the main results

Part 3

Unfolded results

Inclusive

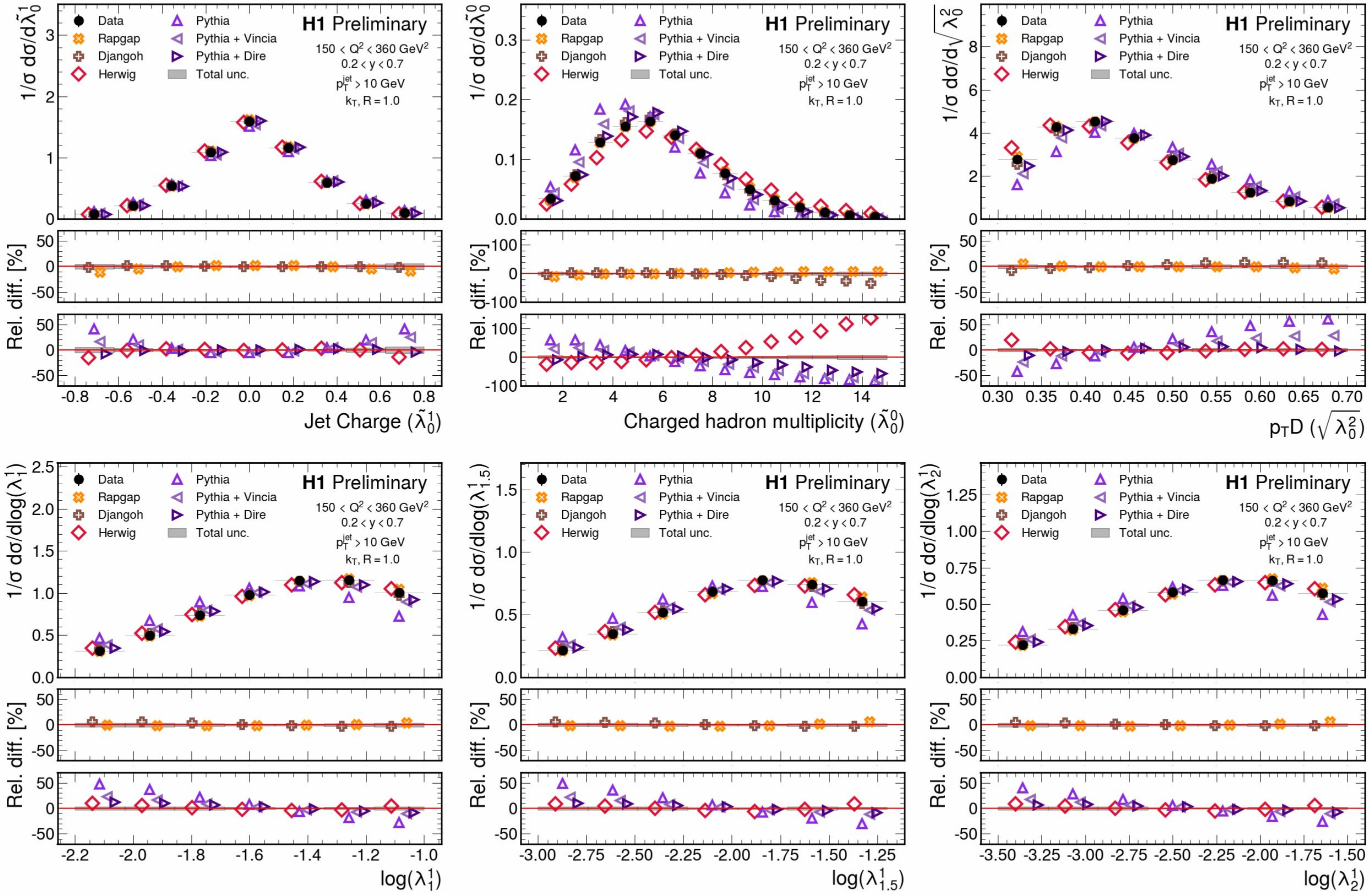


Dedicated DIS generators do a good job **everywhere**, especially **Rapgap**

Herwig does a good job for all distributions besides charge particle multiplicity

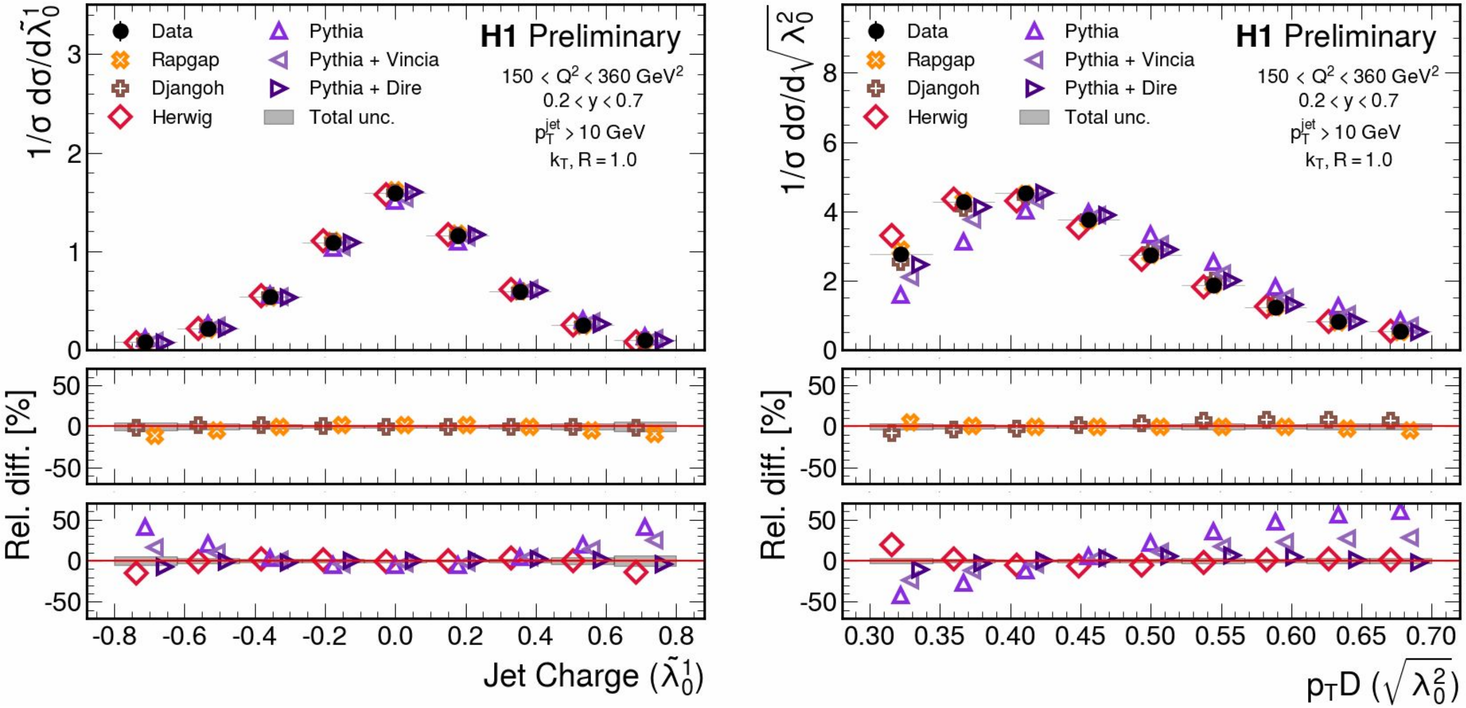
Alternative parton showers for **Pythia** do better than nominal, specially **Dire**

Multi-differential

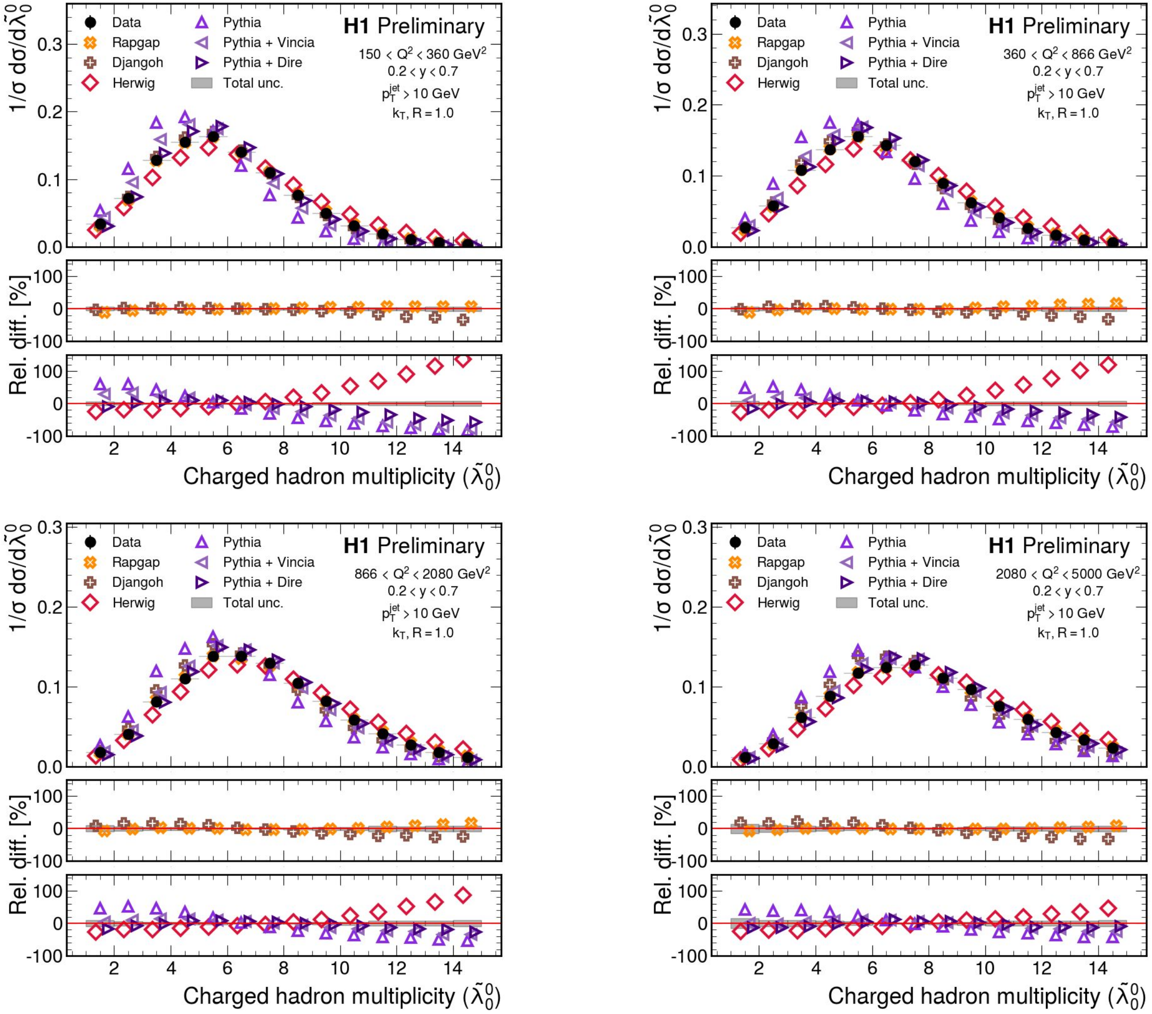


Unfolded results are also presented for **4 Q^2 intervals**

Multi-differential



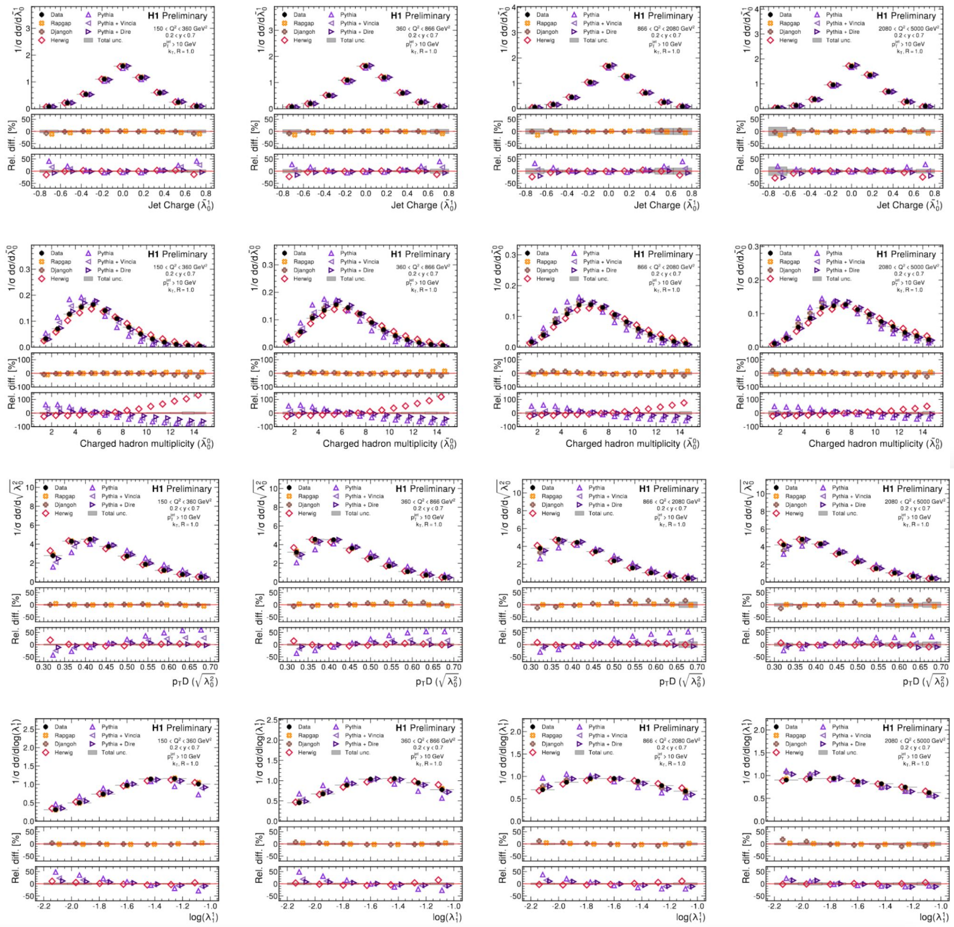
More quark like distributions as Q^2 increases

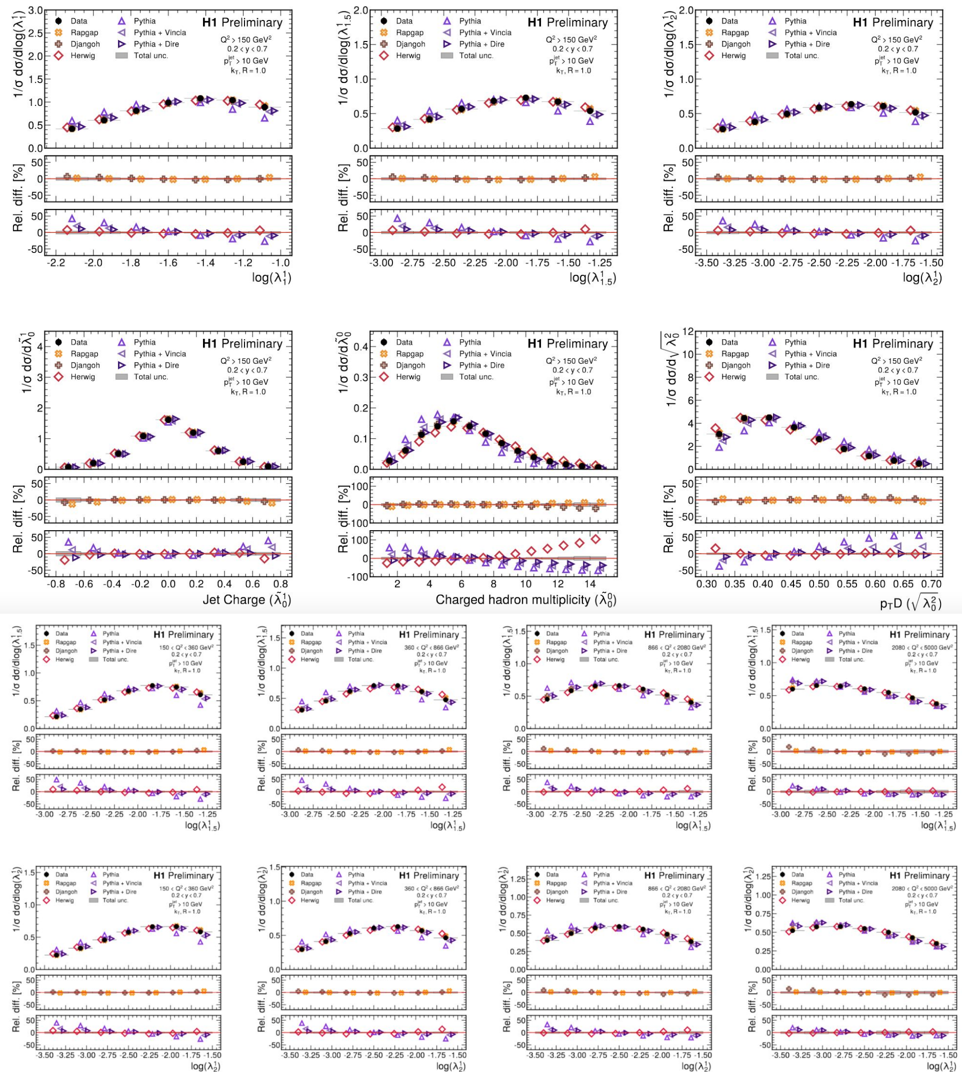


Agreement between general purpose generators **improves** at higher Q^2

Multi-differential

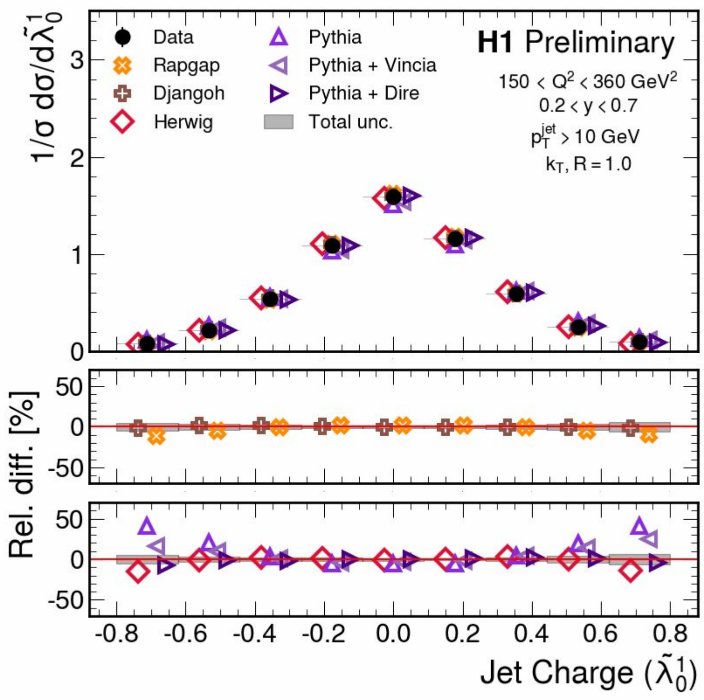
All Q^2 intervals and **distributions** are unfolded **simultaneously!**





Conclusions and prospects

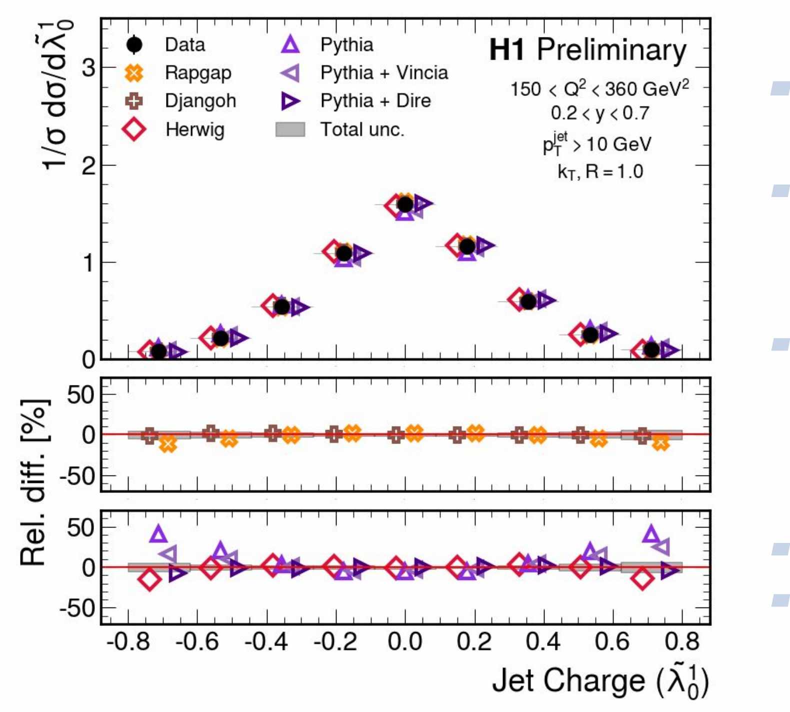
Unfolding the EIC



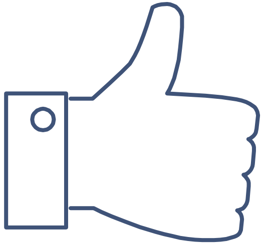
- Future step to be investigated is to use **only particles** for the unfolding, allowing the simultaneous unfolding of **any** jet observable
- To achieve this goal is rely on:
  - Precise object reconstruction and calibration
  - Manageable number of objects: O(100)
- Should also study the feasibility of a **full event unfolding**
- Unfolded results could be shared by the **whole collaboration:**
  - Accelerate new measurements
  - Analysis design based on unfolded information

$$\lambda_\beta^\kappa = \sum_{i\in\text{jet}} z_i^\kappa \left(\frac{R_i}{R_0}\right)^\beta$$

Conclusions and prospects



- Jet observables are an unique laboratory to study **QCD properties**
- **Energy scale** evolution for each jet observable measured in multiple **Q^2 intervals from 150 to 5000 GeV^2**
- Detector effects are corrected using the **Omnifold method** with particles as inputs using **graph neural networks**
  - Unbinned and simultaneous unfolding
- Good agreement for dedicated DIS generators, **Herwig** described all distributions besides track multiplicity while **Dire** parton shower has the best agreement for the **Pythia** implementations
- First step towards unfolding **any** jet observable in one go
- Preliminary results available at: H1prelim-22-034



THANKS!

Any questions?

Backup

Systematic uncertainties

Systematic uncertainties currently considered

- **HFS energy scale:** +- 1%
- **HFS azimuthal angle:** +- 20 mrad
- **Lepton energy:** +- 0.5% (mainly affects Q^2)
- **Lepton azimuthal angle:** +- 1 mrad (mainly affects Q^2)
- **Model uncertainty:** differences in unfolded results between Djangoh and Rapgap
- **Non-closure uncertainty:** Differences between the expected and obtained values of the closure test
- **Statistical uncertainty:** Standard deviation of 100 bootstrap samples with replacement

MC Generators

Lund string hadronization model and **CTEQ6L** PDF set

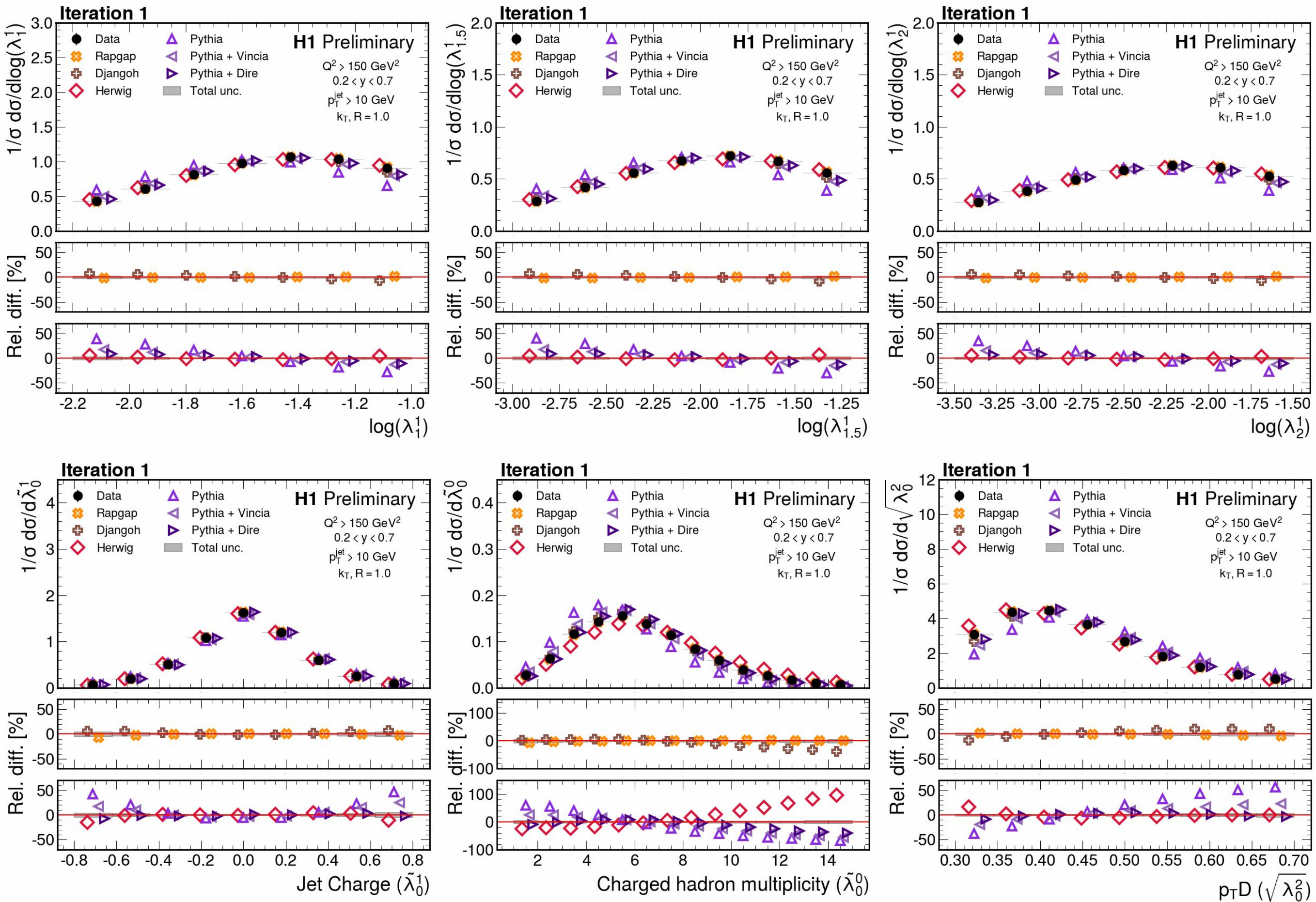
- **Djangoh:** Dipole model from Ariadne
- **Rapgap**: PS from leading log approximation

Pythia 8.3: default NNPDF3.1 PDF

- **Vincia**: p_T ordered antenna and NNPDF3.1 PDF
- **Dire**: dipole model, similar to Ariadne and MMHT14nlo68cl PDF

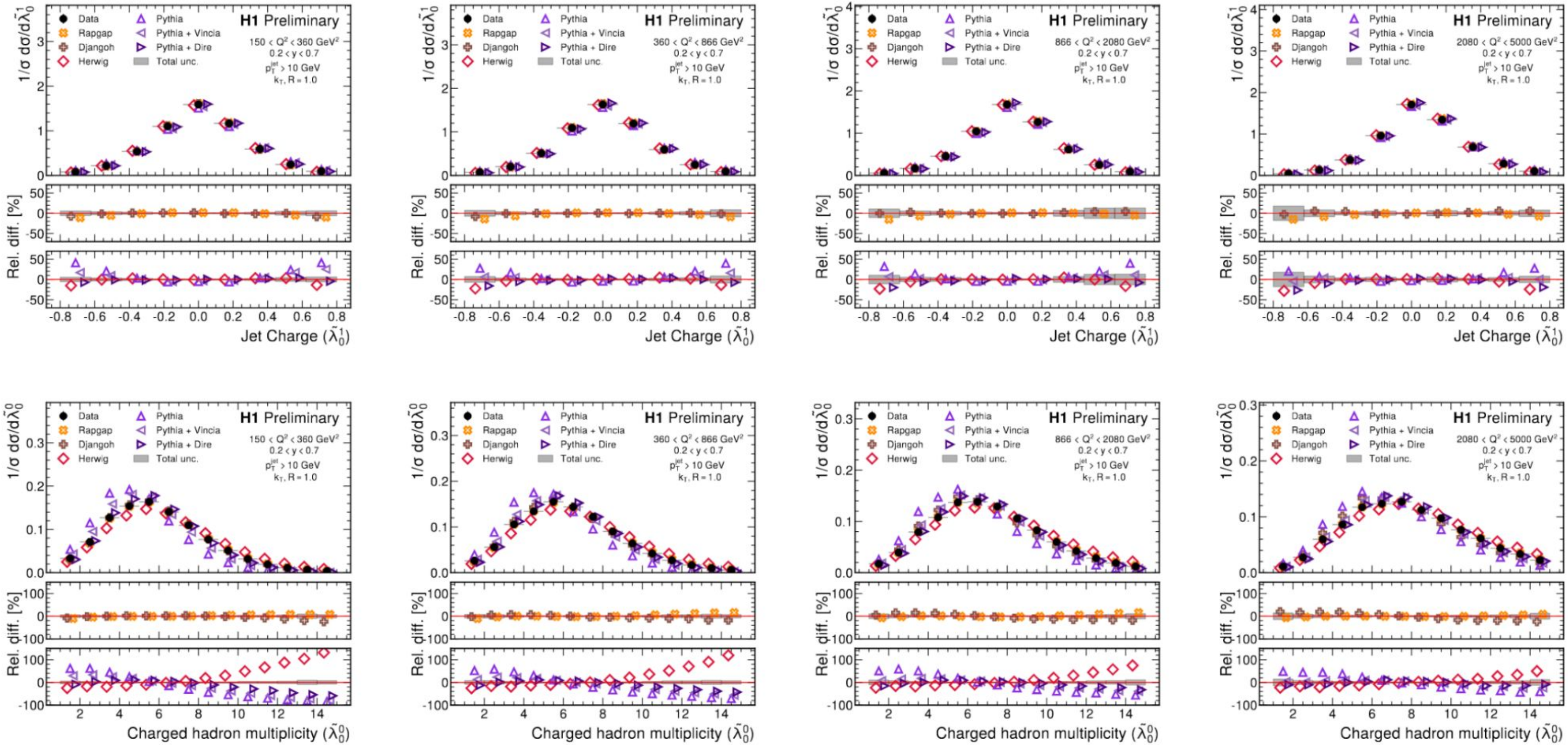
Herwig 7.2: Cluster hadronization and CT14 PDF set

Iteration dependence

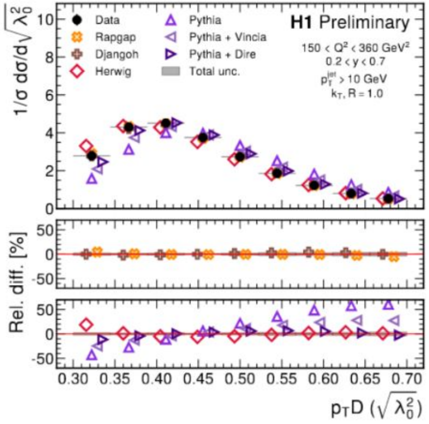


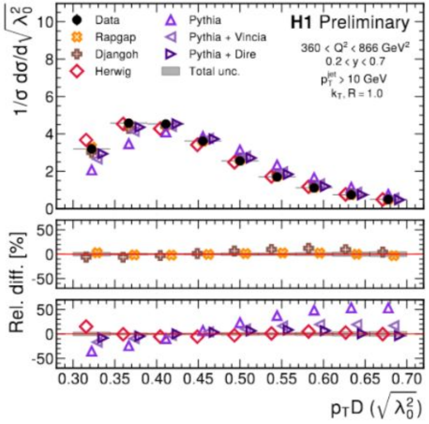
Results are stable even at high omnifold iteration numbers. Total number of iterations chosen based on the total systematic uncertainty

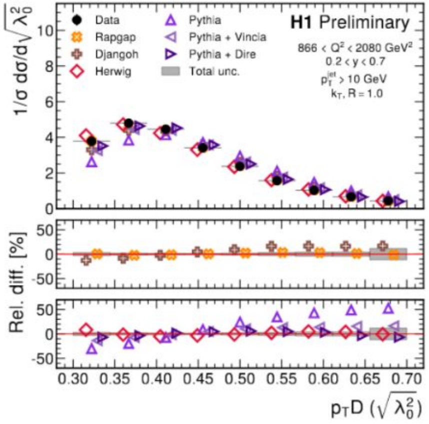
Multi Differential

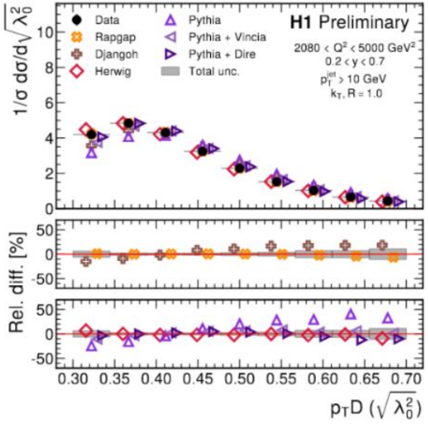


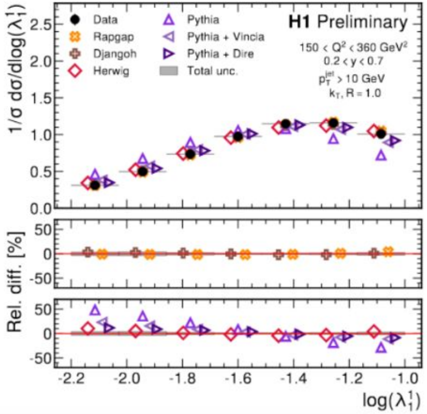
Multi Differential

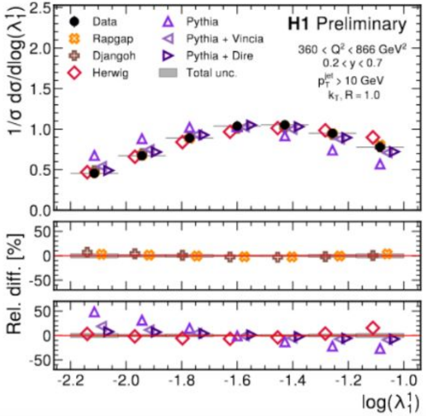


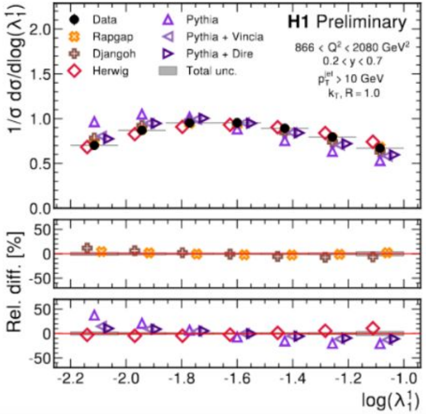


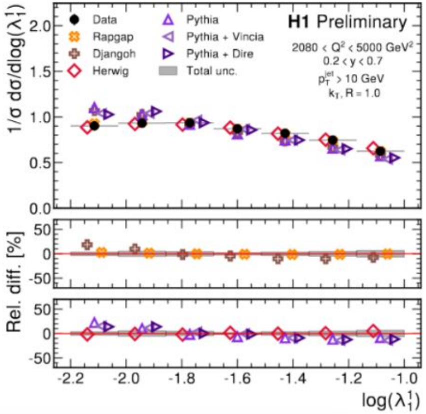












Multi Differential