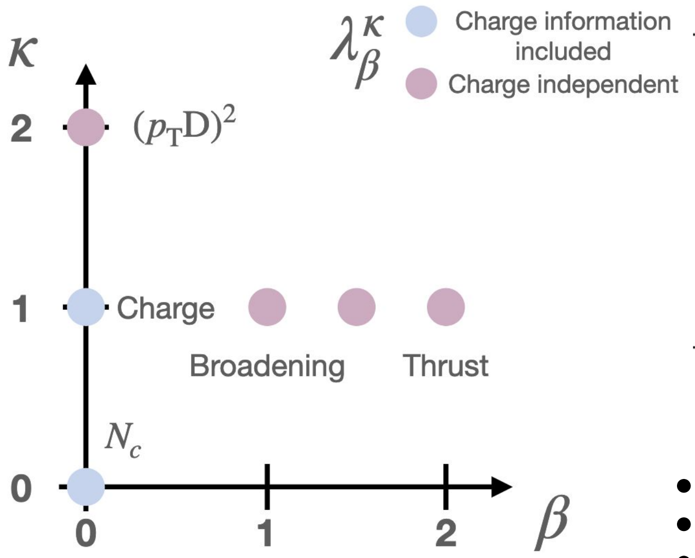
Multi-differential Jet Substructure Measurement in High Q^2 DIS Events with HERA-II Data

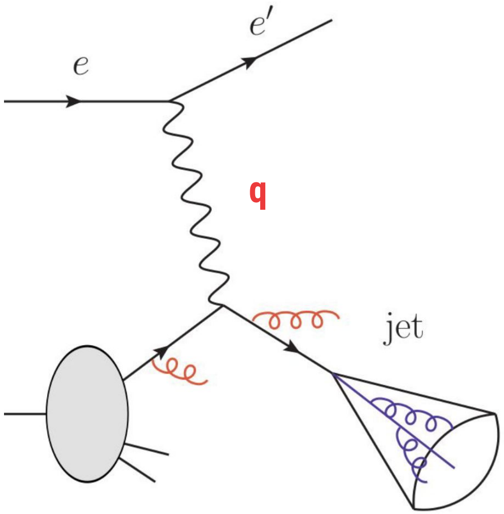
Vinicius Mikuni, on behalf of the **H1 Collaboration**

Jet angularities

Use jet observables to study different properties of QCD physics:

- Infrared and collinear (IRC) safe λ^1_a, a = [0,0.5,1] and unsafe **p_TD** 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)

$$\lambda_\beta^\kappa = \sum_{i\in\text{jet}} z_i^\kappa \left(\frac{R_i}{R_0}\right)^\beta \qquad \tilde{\lambda}_0^\kappa = Q_\kappa = \sum_{i\in\text{jet}} q_i \times z_i^\kappa .$$

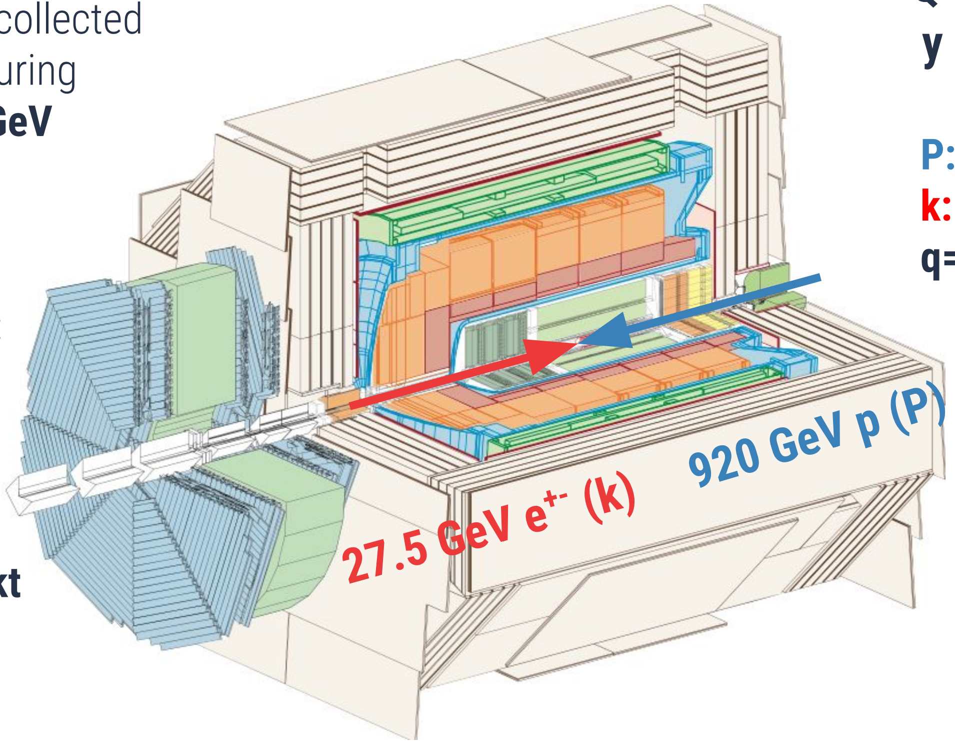
Experimental setup

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**



$Q^2 = -q^2$

$y = Pq / pk$

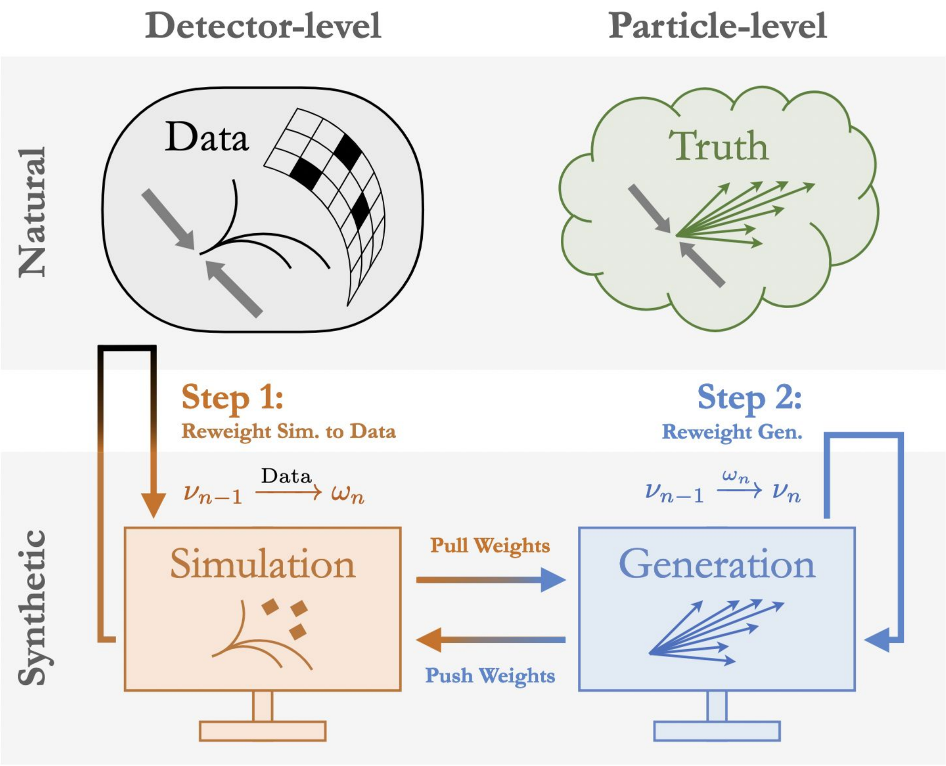
P: incoming proton 4-vector

k: incoming electron 4-vector

q=k-k' : 4-momentum transfer

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

Omnifold*



2 step iterative approach

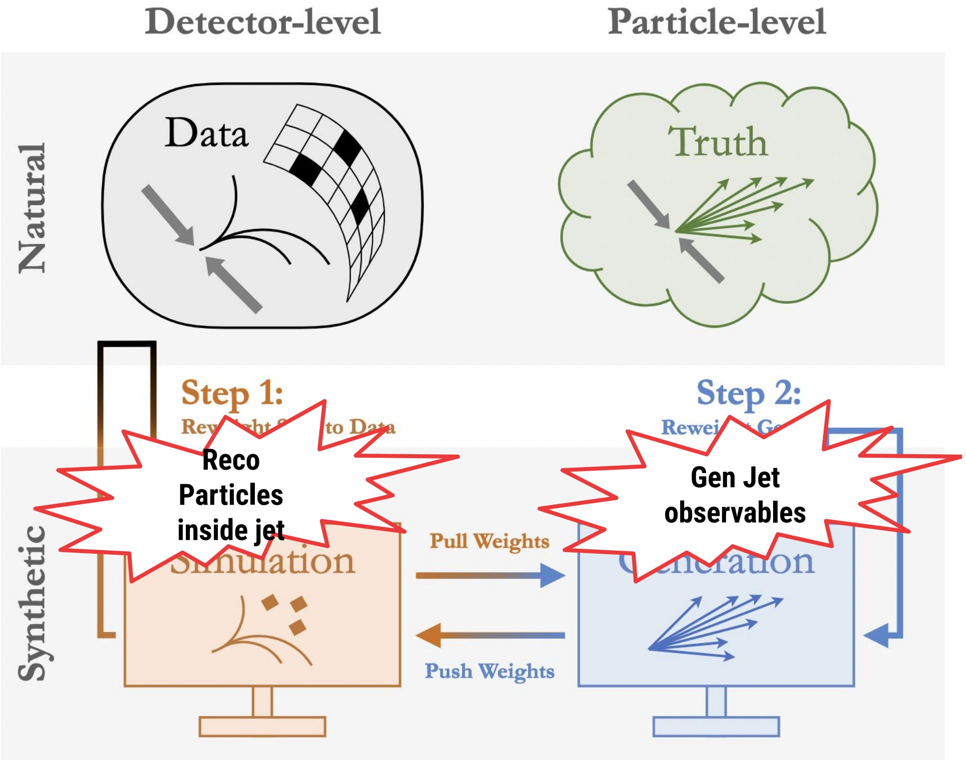
- Simulated events after detector interaction are reweighted to match the data
- Create a "new simulation" by transforming weights to a proper function of the generated events

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

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

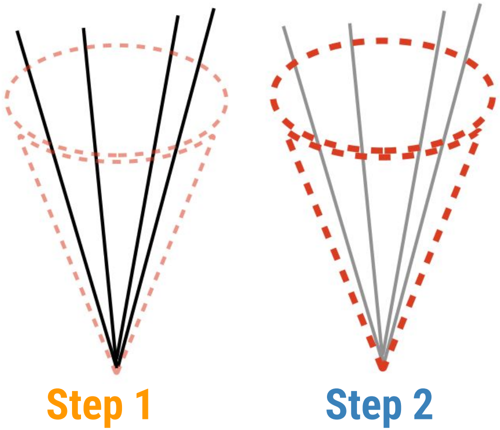
* Andreassen et al. PRL 124, 182001 (2020)

Omnifold



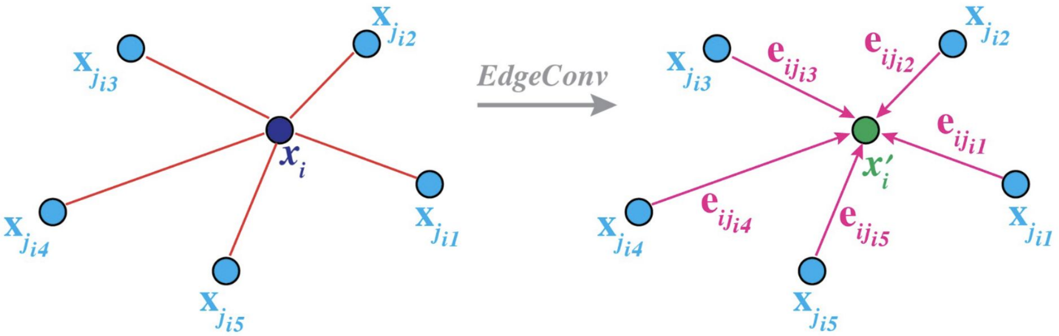
Different input levels for each step

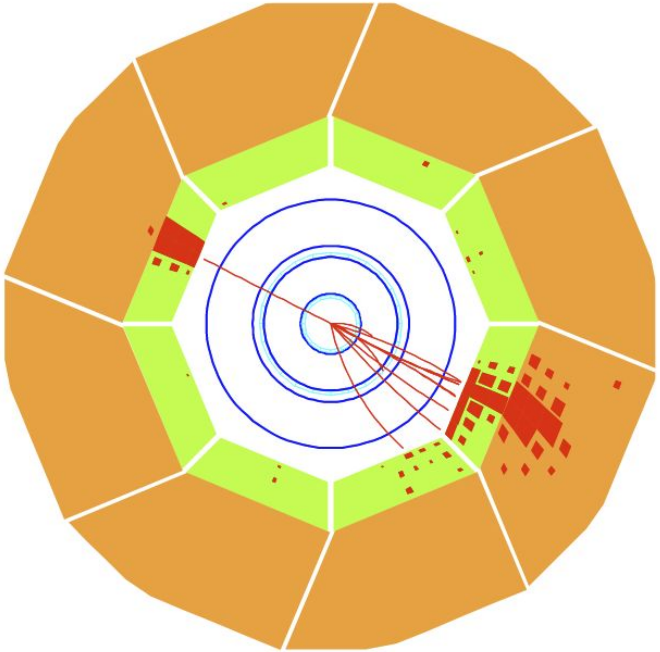
- **Step 1** particles are used as inputs
- **Step 2** uses the set of observables planned to unfold



Extracting particle information

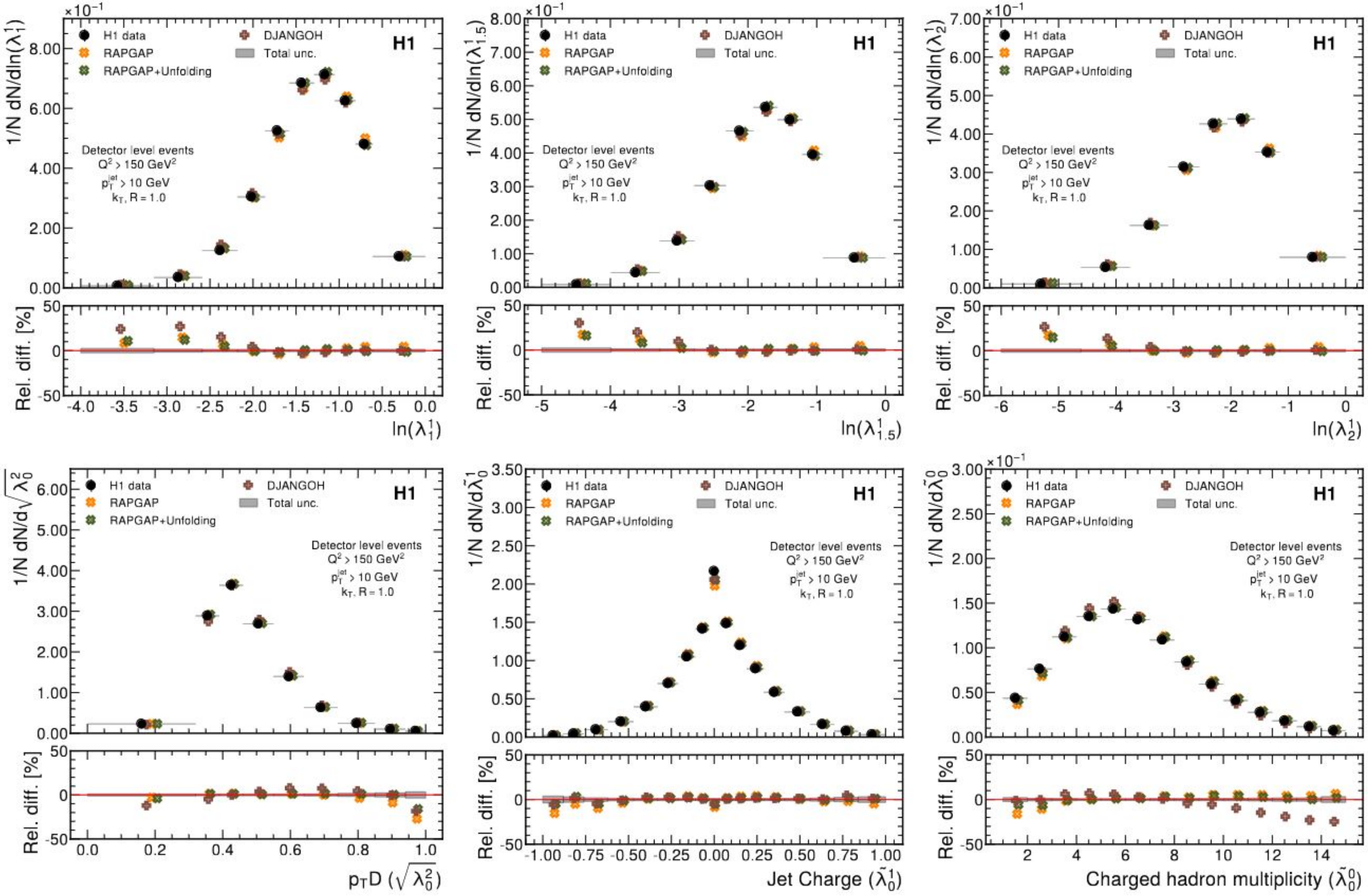
- Particle information is extracted using a **Point cloud transformer*** model
- Model takes **kinematic properties** of particles and use the distance between particles in η-φ to learn the relationship between particles
- Built in symmetries: **permutation invariance**
- Consider up to **30** particles per jet





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

All distributions are **simultaneously** unfolded.



Outputs of the unfolding methodology are **weights** that are applied to the simulation

- **Green markers** represent the unfolded results **at reco level**
- Agreement with data **improves** compared to **initial Rapgap simulation**

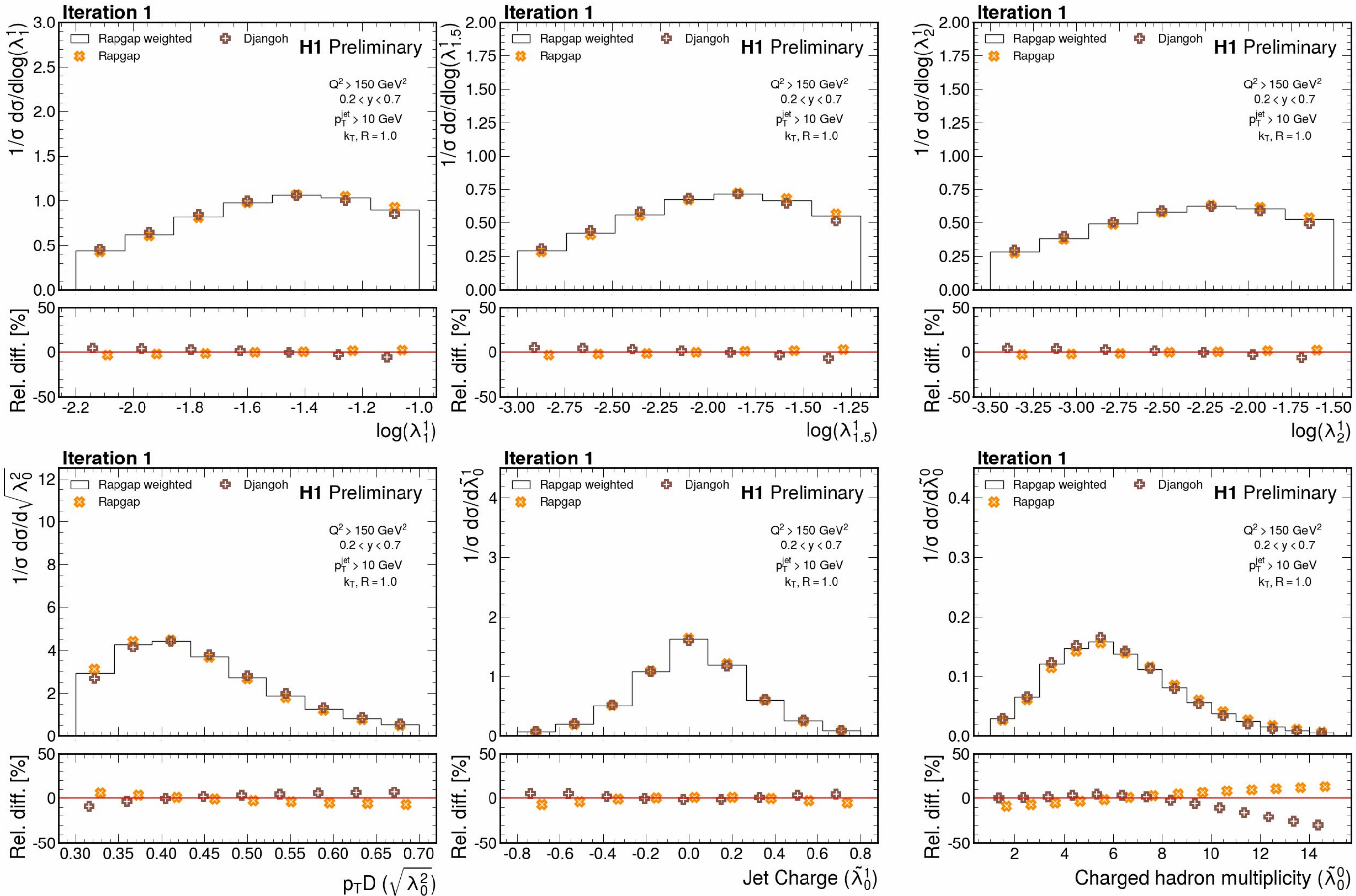
Systematic uncertainties

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
- **QED uncertainty**: Use the variation of measured quantities when radiation is turned off in the simulation
- **Statistical uncertainty:** Standard deviation of 100 bootstrap samples with replacement

Closure test

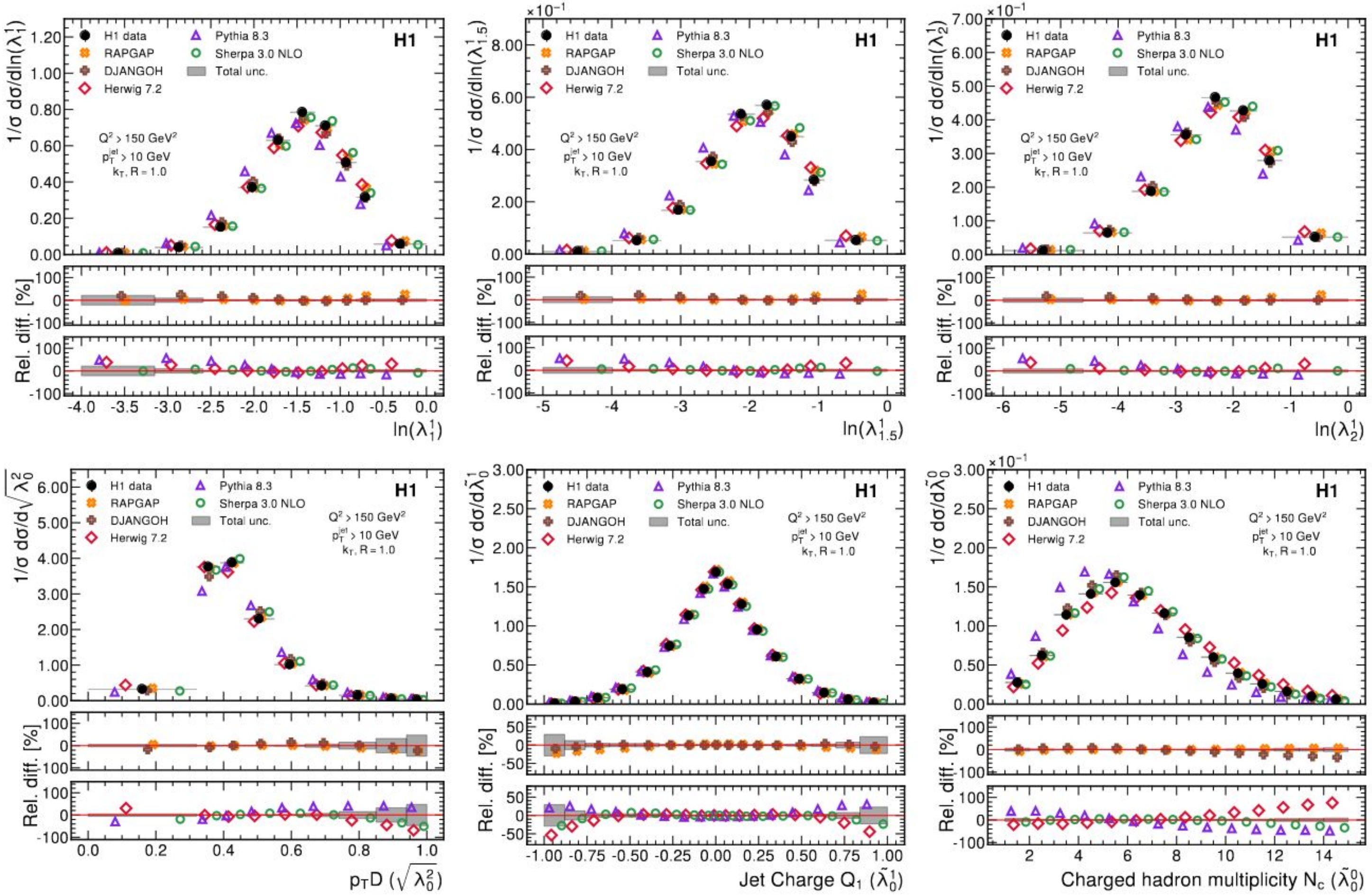
All distributions are unfolded **simultaneously without binning**



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

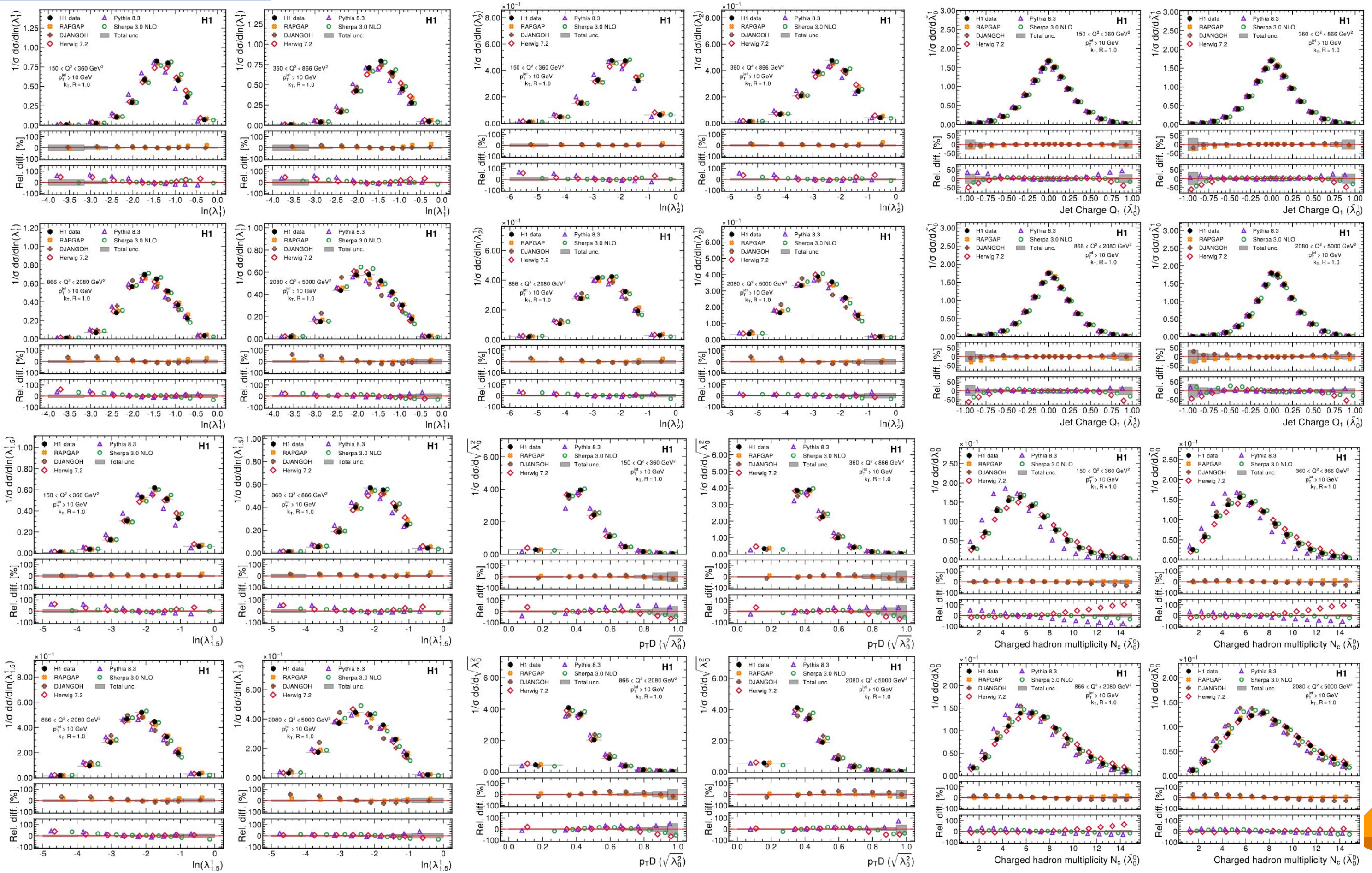
Inclusive



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

Herwig, **Pythia**, and (yet **unreleased update** to) **Sherpa** do a decent job for most distributions

Multi-differential

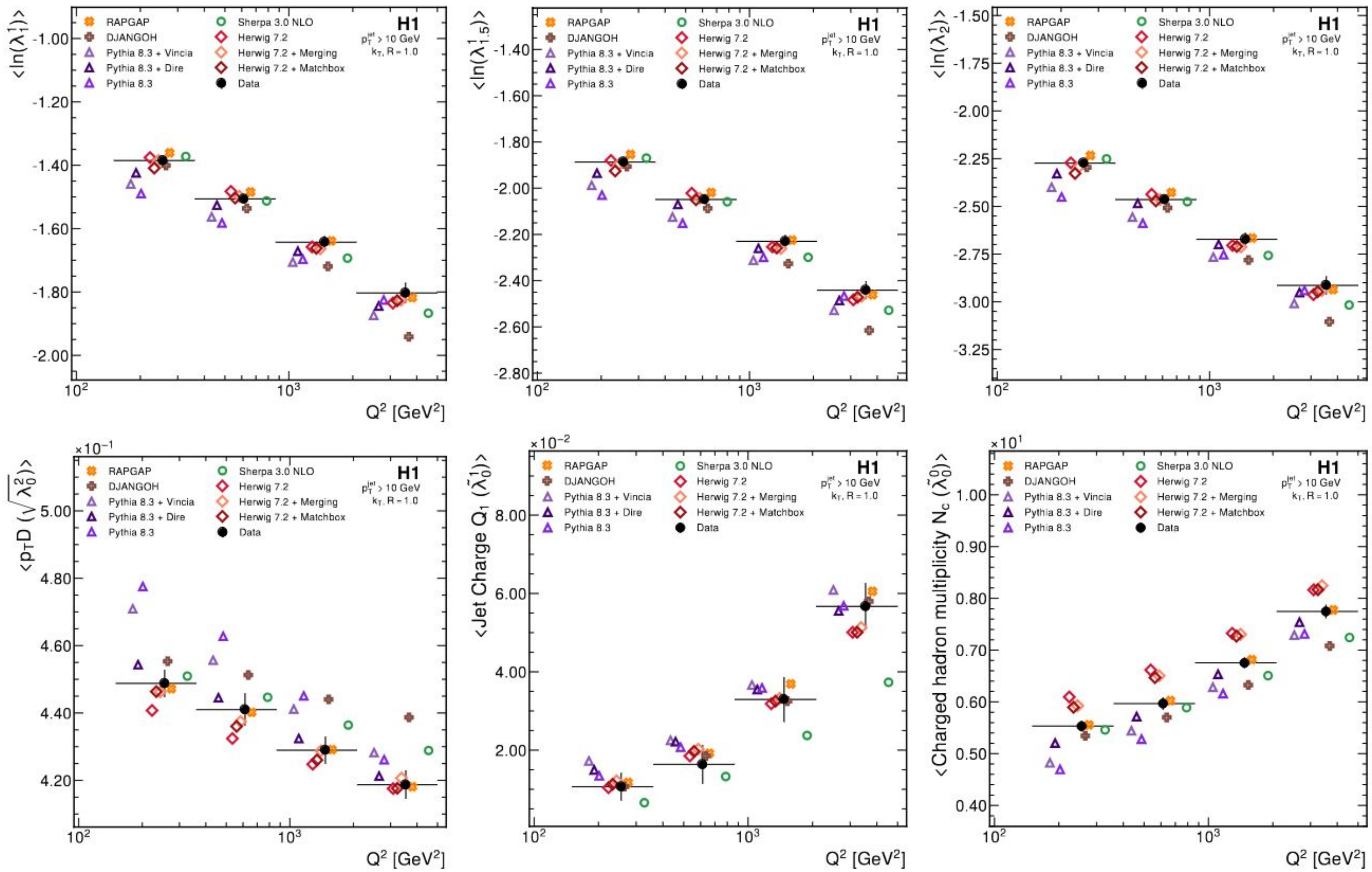


Q^2 distribution is simultaneously unfolded, displaying the energy scale dependence of the observables, resulting in more than **30 unfolded distributions provided**

- **Jet p_T** and **eta** also used during the **unfolding**

Multi-differential

Mean value of all distributions also unfolded for free

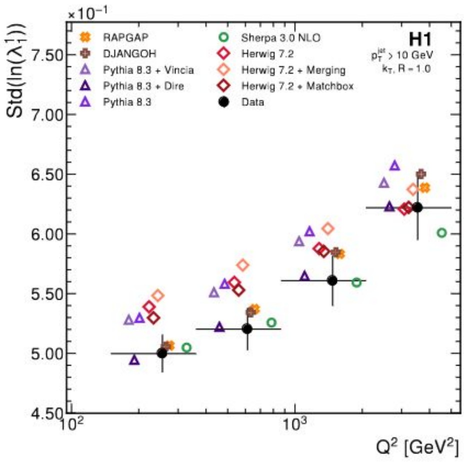


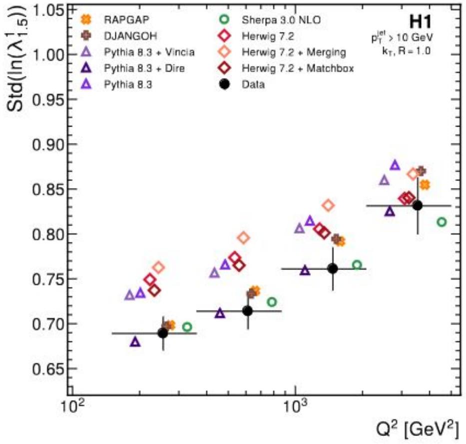
More quark-like behaviour at higher energies: mean jet charge becomes more positive

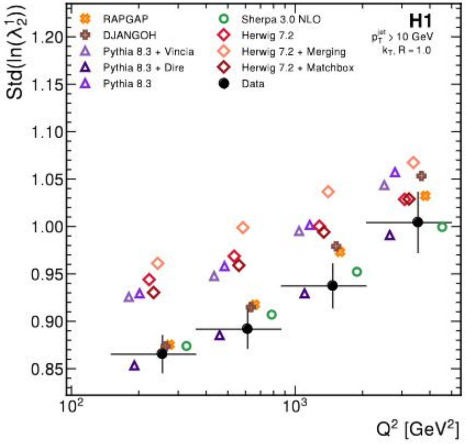
Agreement between general purpose generators **improve** at higher Q^2

Multi-differential

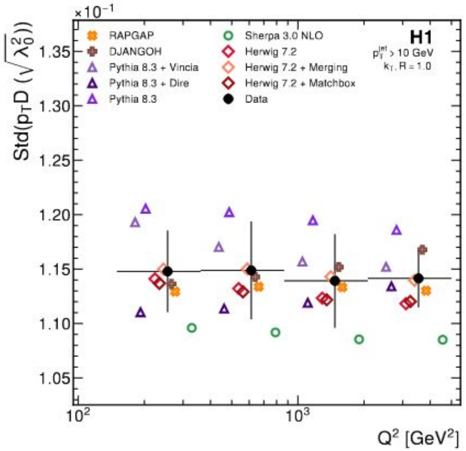
Standard deviation of all distributions also unfolded for free

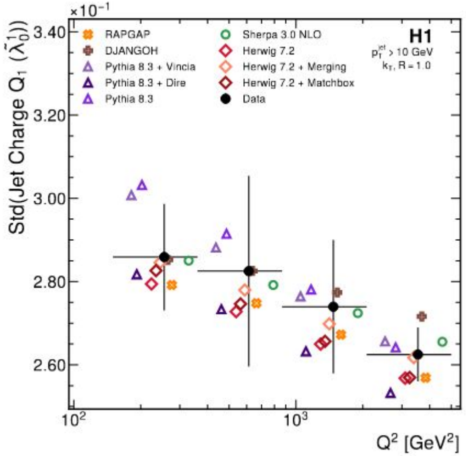


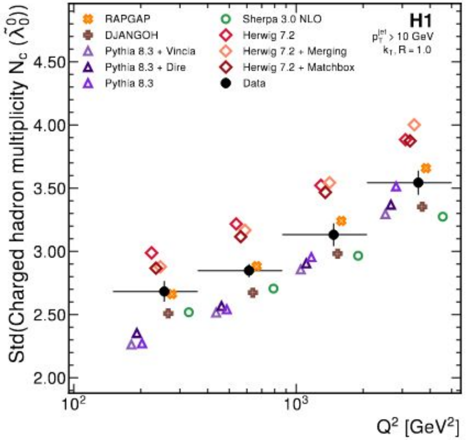




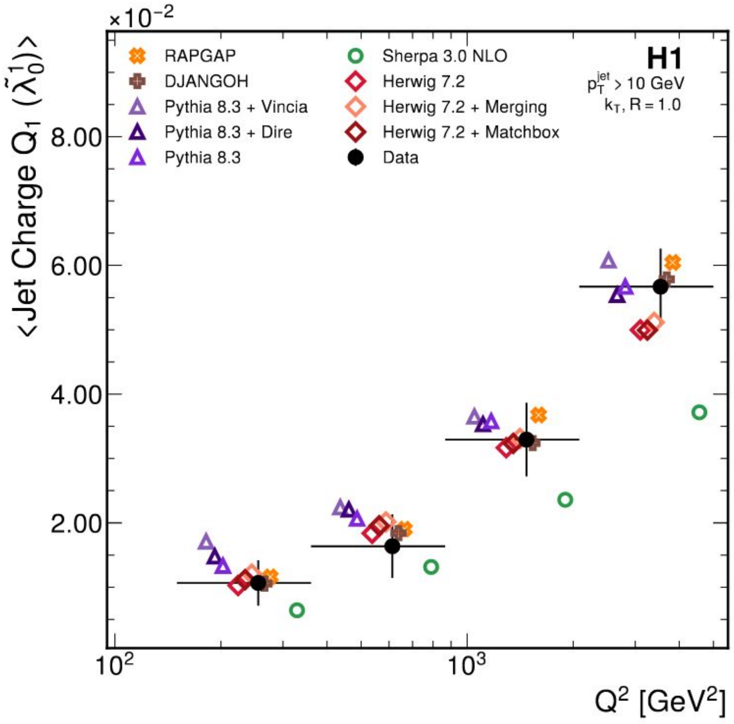
Worse general agreement between data and simulations



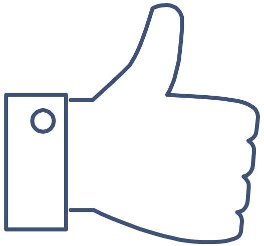




Conclusions and prospects



- 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
- Unfolded the means and standard deviations without bin artifacts
- Good agreement for dedicated DIS generators, **worse** agreement for general purpose simulators
- Public results available at: **DESY-23-034** and **PLB 844 (2023) 138101**



THANKS!

Any questions?

Backup

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

Sherpa 3.0: Cluster hadronization pQCD at NLO accuracy for the 1 & 2 jet final states and LO for the 3 jet contribution. Dipole shower.