(Recent) jet substructure from archival data

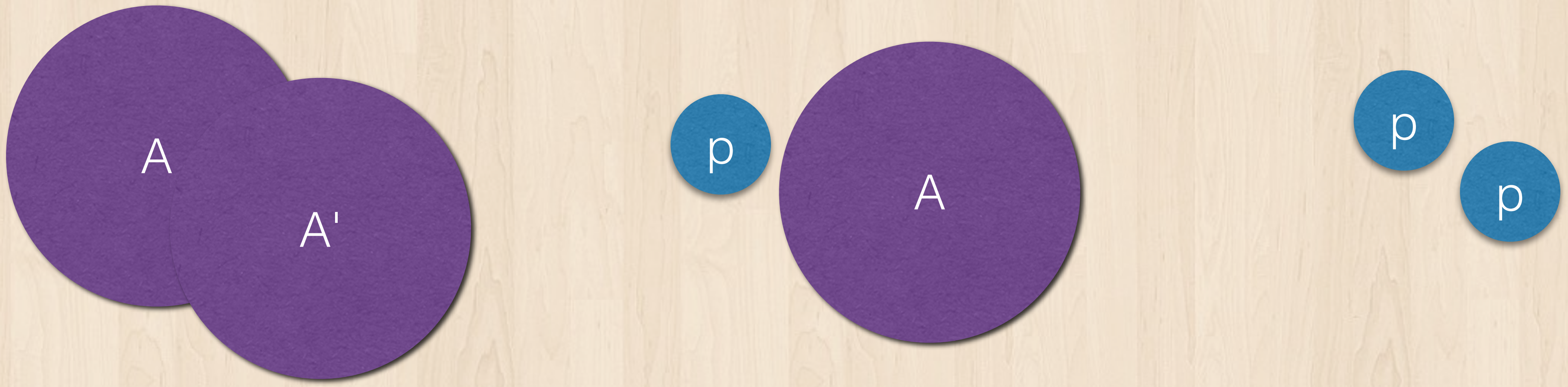
Yi Chen (MIT)

Sep 19, 2022. Non-Perturbative QCD Energy Flow

The MITHIG's work is supported by US DOE-NP

Preserved data

Collision systems

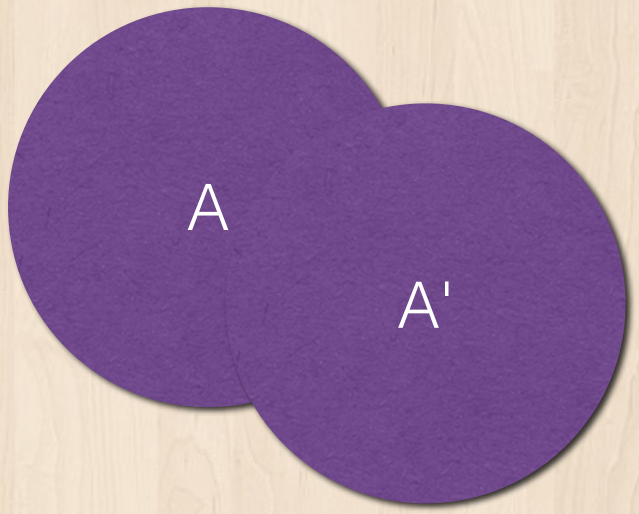


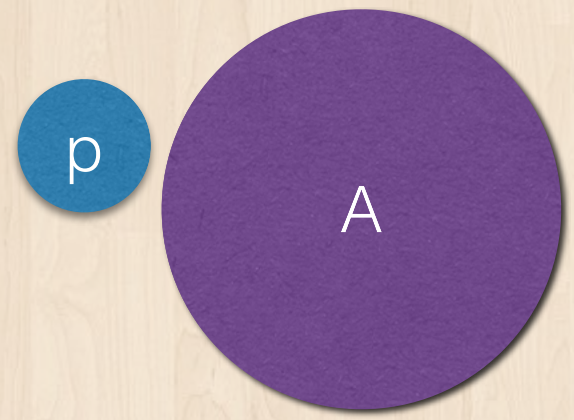
Proliferation of recent jet results from LHC and RHIC

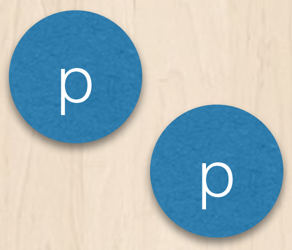
Always with hadronic initial states on both sides

Harder sometimes to disentangle various effects

Collision systems

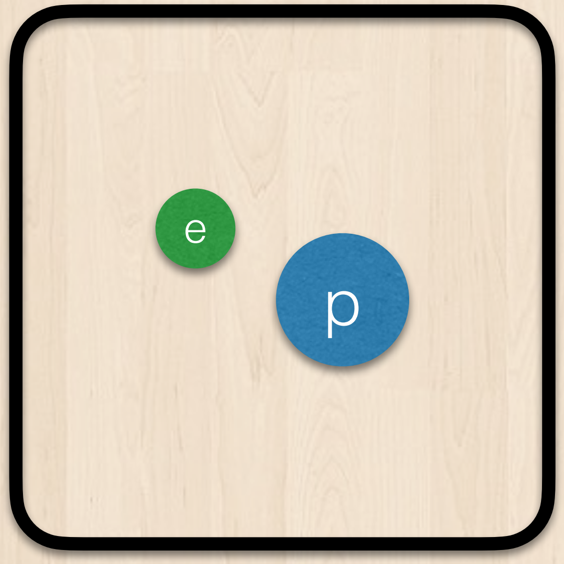


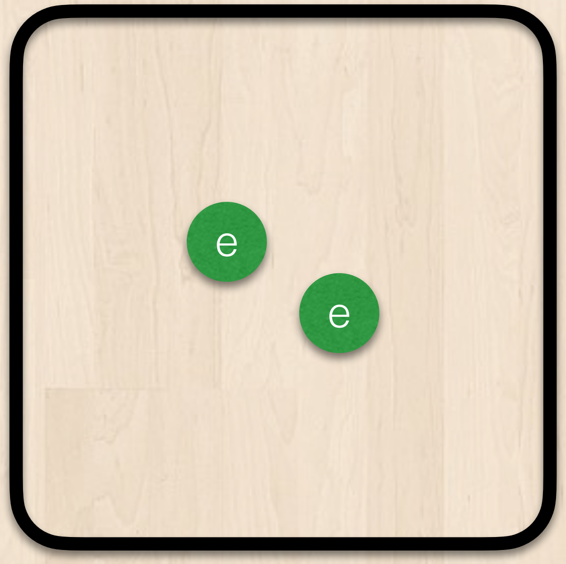




Much cleaner
& complementary

Today: some recent jet results with H1 and ALEPH



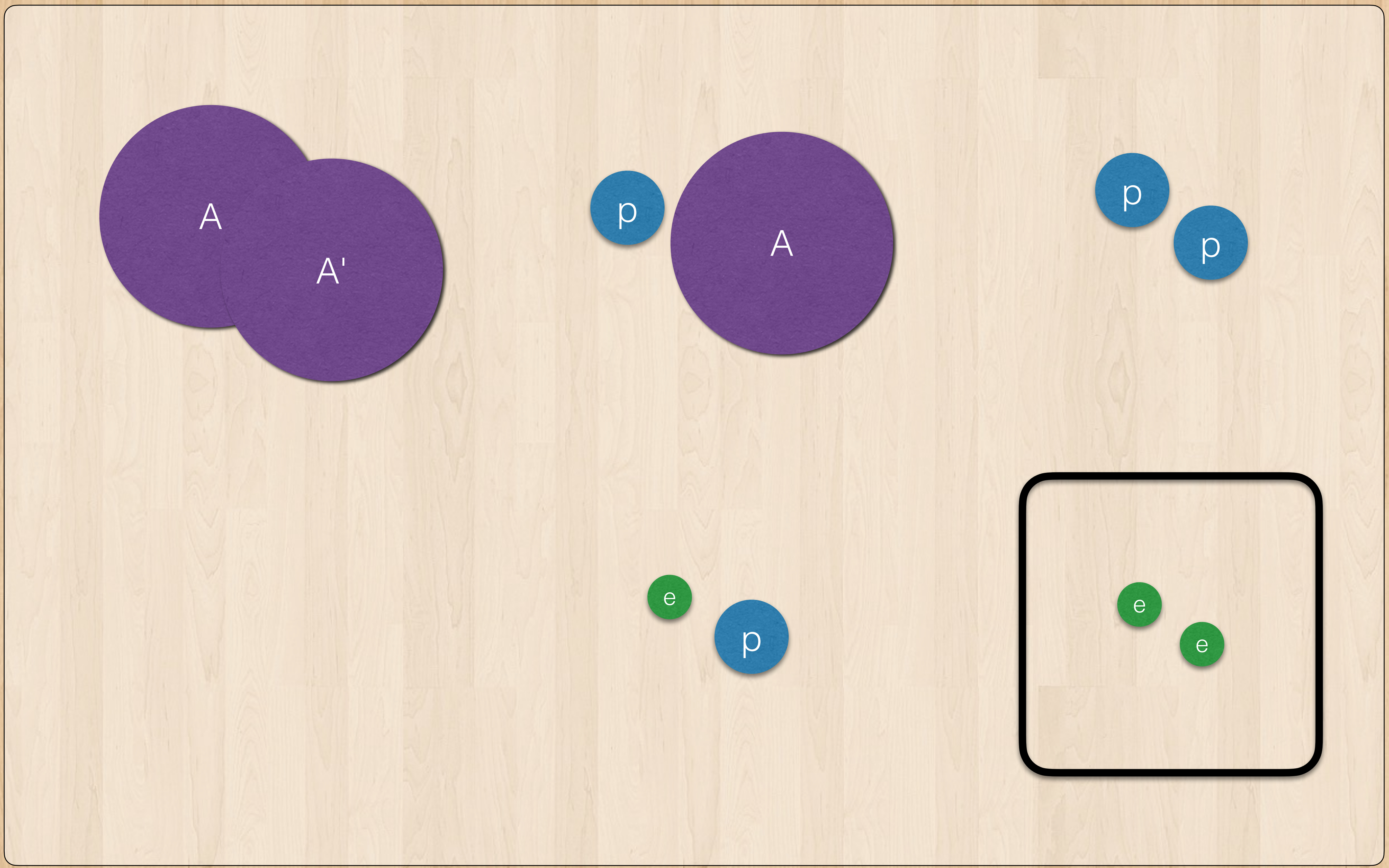


Early jet measurements

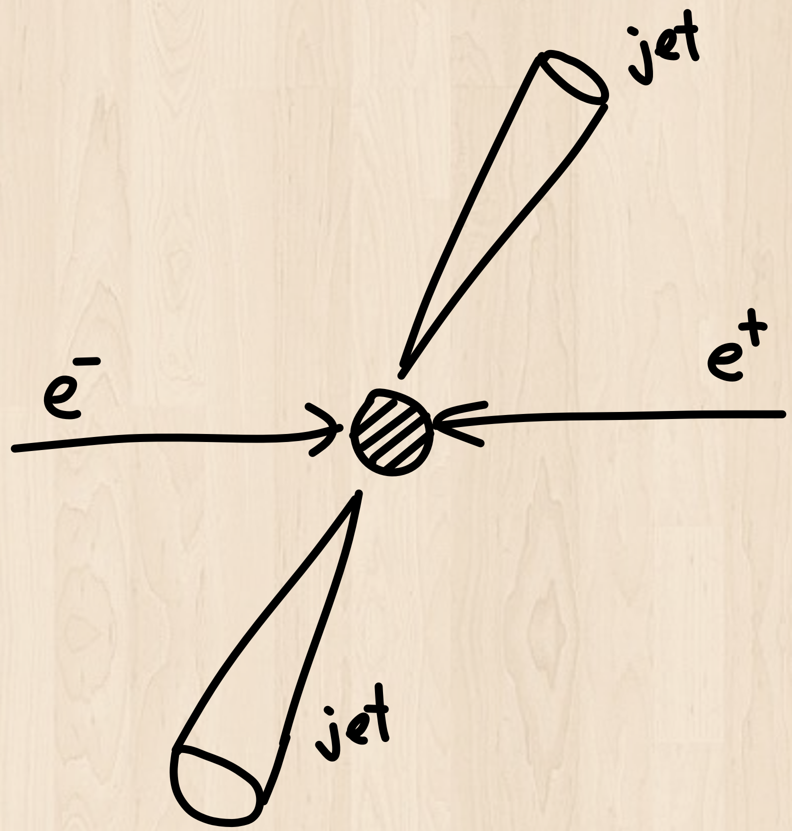
Many jet measurements are done with previous generation of jet algorithms
=> Not ideal for LHC/RHIC/EIC comparisons

Excellent opportunity for re-analysis with new algorithms

Collision systems

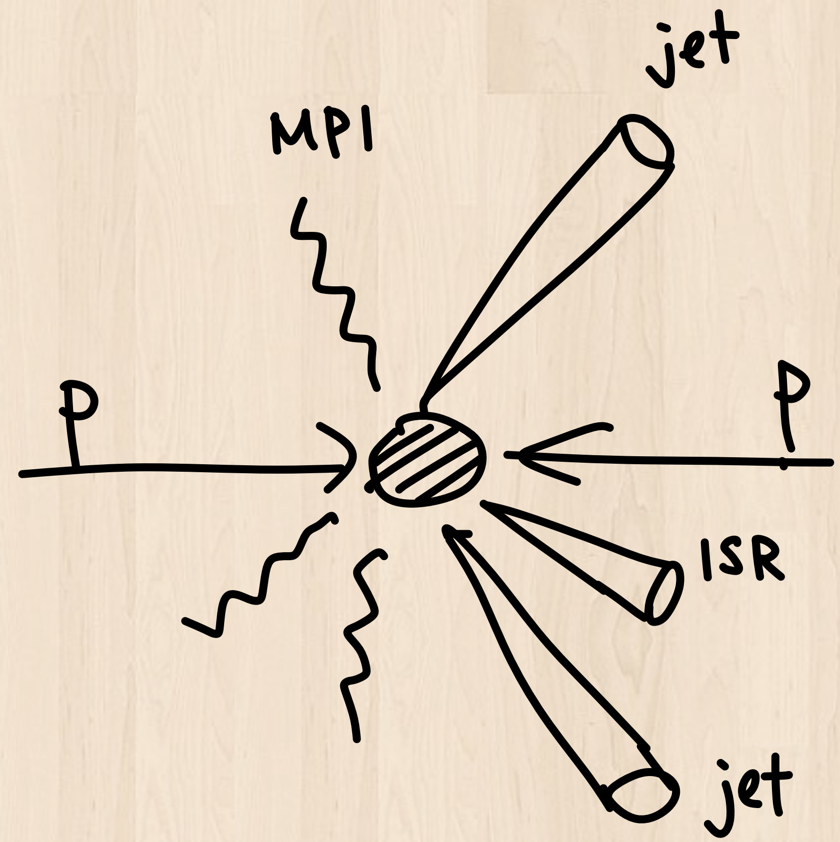


e^+e^-: clean



Better control of event kinematics

Cleanest test of pQCD and models

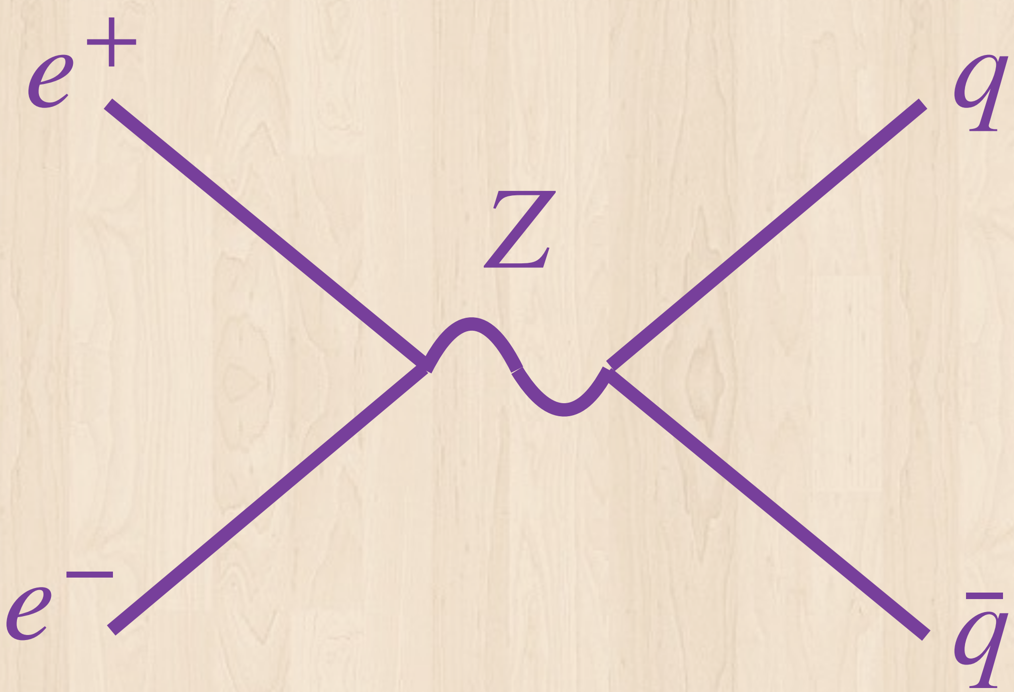


PDF convolution
No longitudinal control
More ISR
MPI

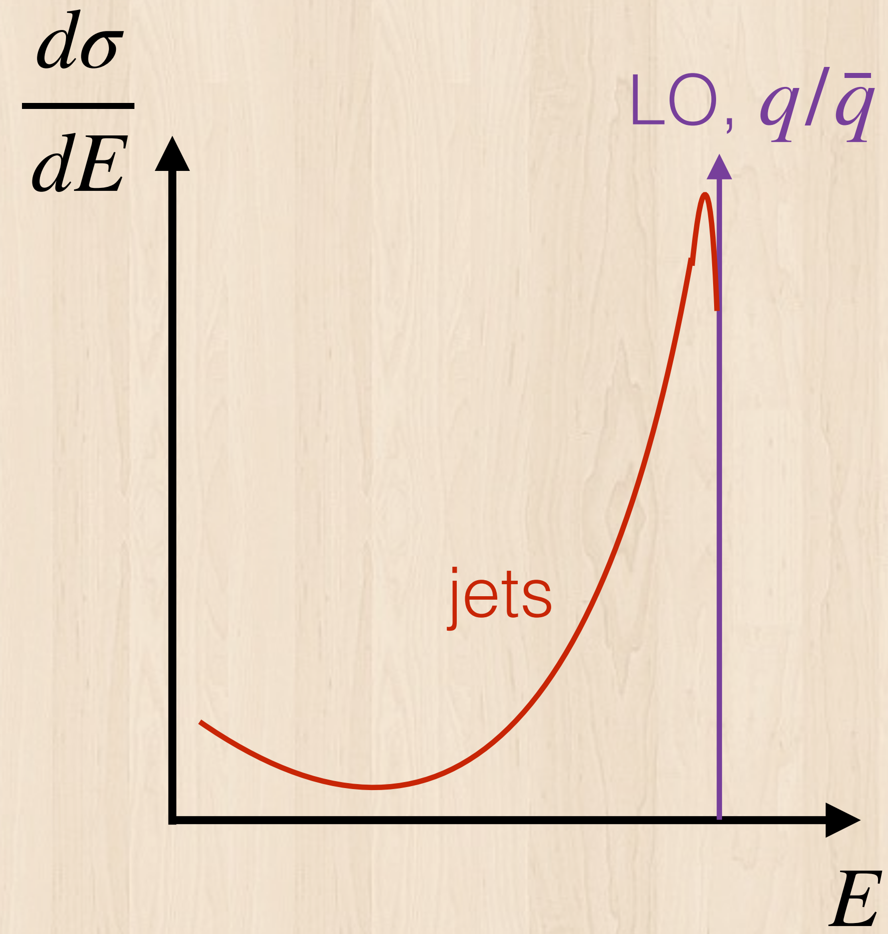
Complements well measurements from other systems

Peaked structure

91.2 GeV collisions



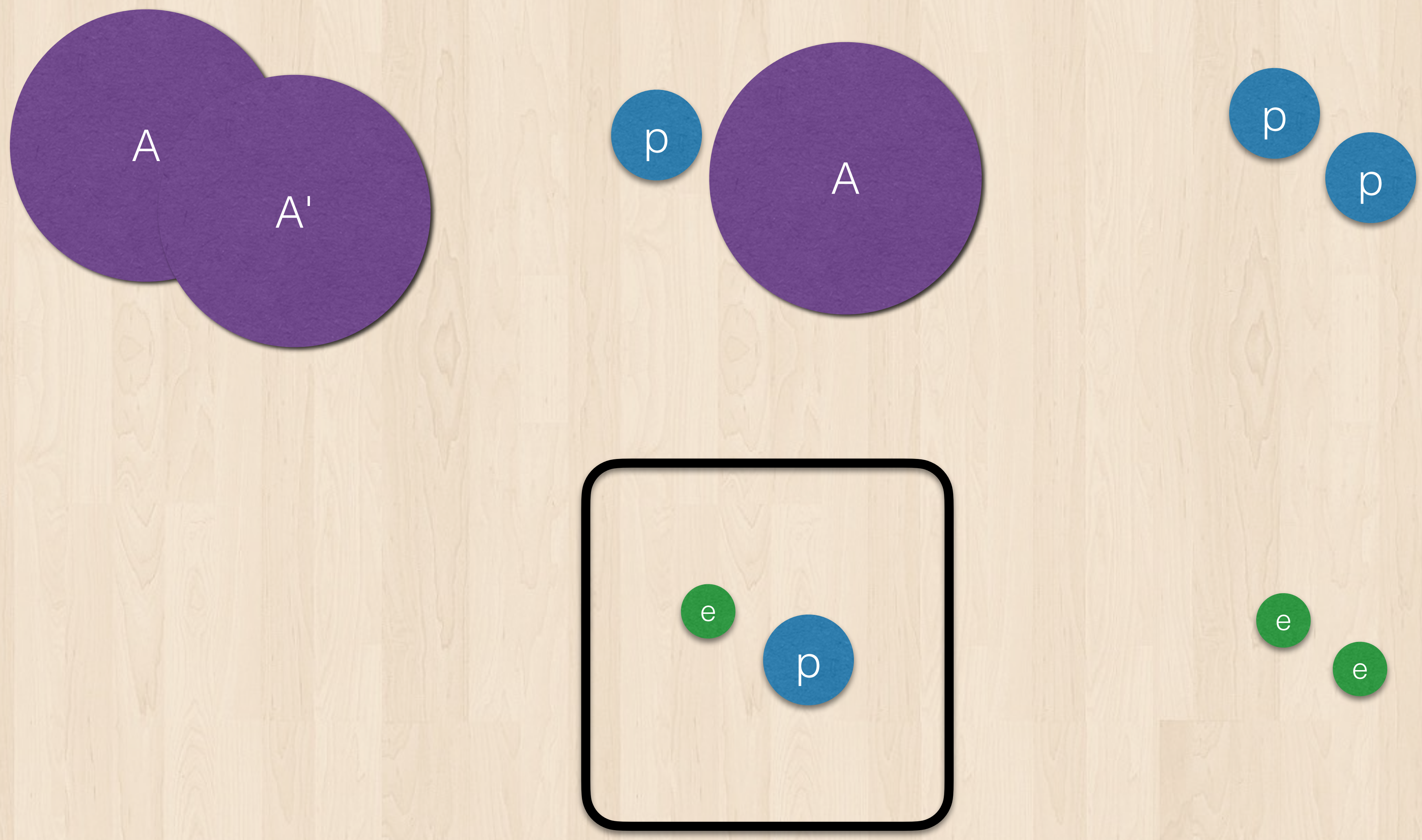
Dominant jet diagram



Peaked structure is useful for studying jets

Out-of-cone energy => “energy loss”

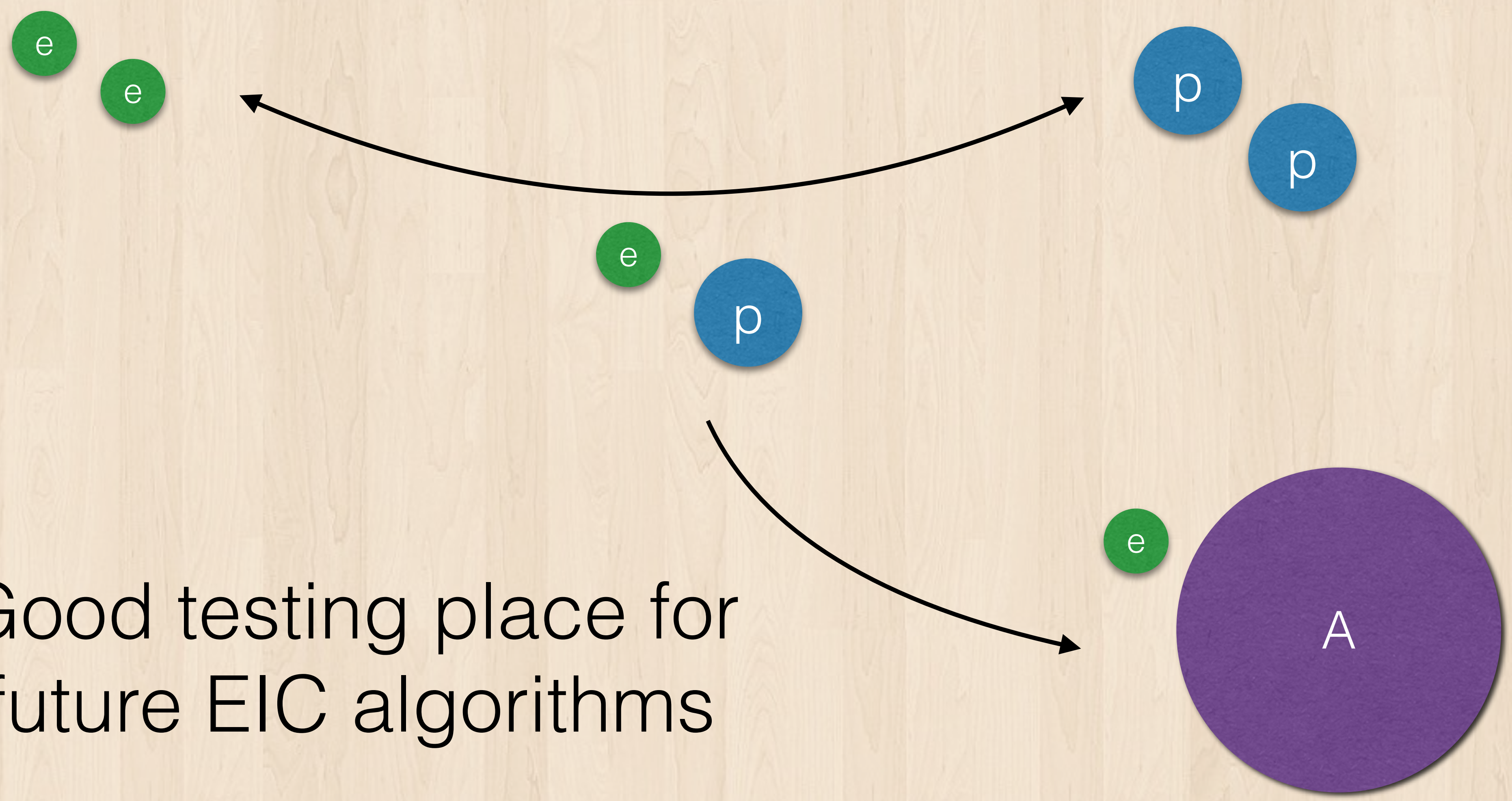
Collision systems



ep: bridging e^+e^- and $pp(\bar{p})$

Only one side with hadronic initial state

→ excellent middle-ground between ee and pp



Good testing place for future EIC algorithms

Using the data:
ALEPH as example

Accessibility of data

ALEPH internal format

Need ALEPH member for this step

Exported text format

MIT open data effort

Plain root

Analyses, ...

Big thanks to both the ALEPH collaboration and MIT OpenData

Jet clustering

1994 archived data & simulation analyzed

Energy-flow objects (combining tracker, calorimeter and muon chambers) are used as input

In order to compare with LHC/RHIC

anti-"k_T" jet, R = 0.4

Hadron-hadron collider

$$d_{ij} = \min\left(p_{T,i}^{-2}, p_{T,j}^{-2}\right)\frac{\Delta R_{ij}^2}{R^2}$$

$$d_{iB} = p_{T,i}^{-2}$$

e^+e^- distance measure

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

$$d_{iB} = E_i^{-2}$$

θ_{ij} = real opening angle

Jet calibration

"MC calibration" "Residual"

Jet energy scale in data = Jet energy scale in simulation × (Jet energy scale in data / Jet energy scale in simulation)

Inclusive Selection

Strategy: first go 99% of the way there with simulation

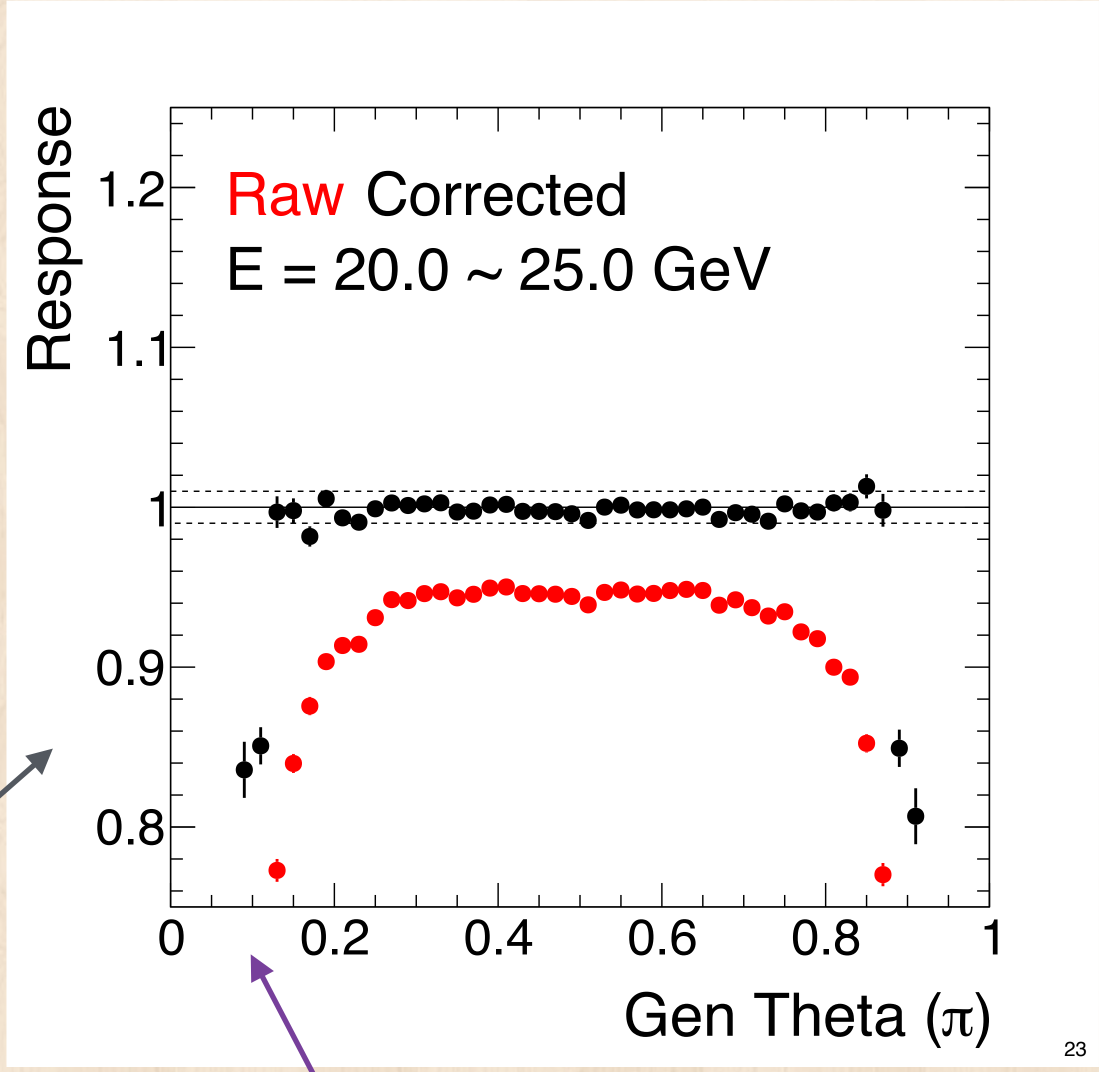
Then data and MC difference in restricted phase spaces

Simulated energy scale

Correct detector jet energy in bins of jet direction (θ_{jet})

Good closure with
E > 10 GeV
$0.2\pi < \theta_{\text{jet}} < 0.8\pi$

Example raw and corrected response (= reconstructed/generated)

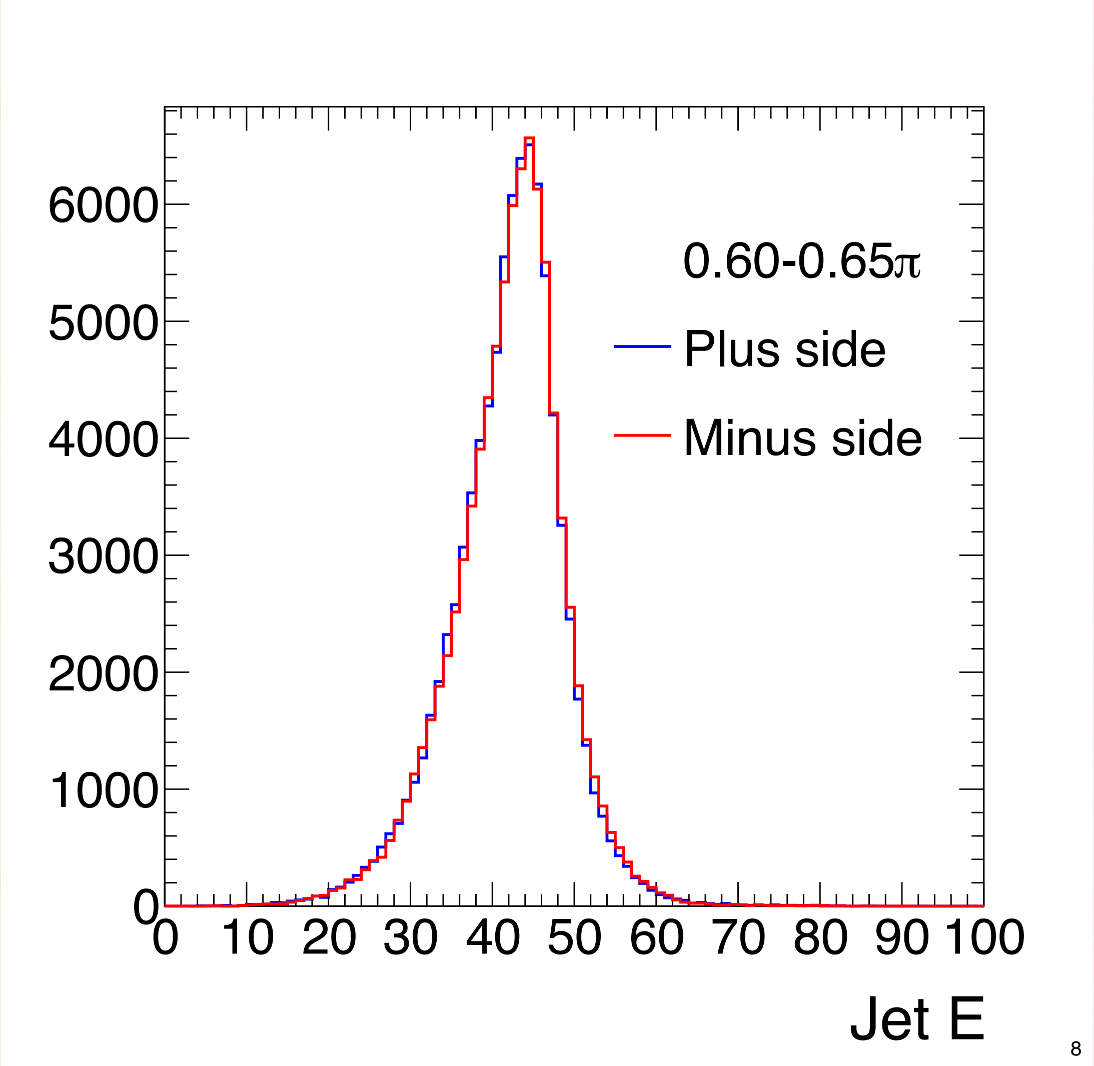


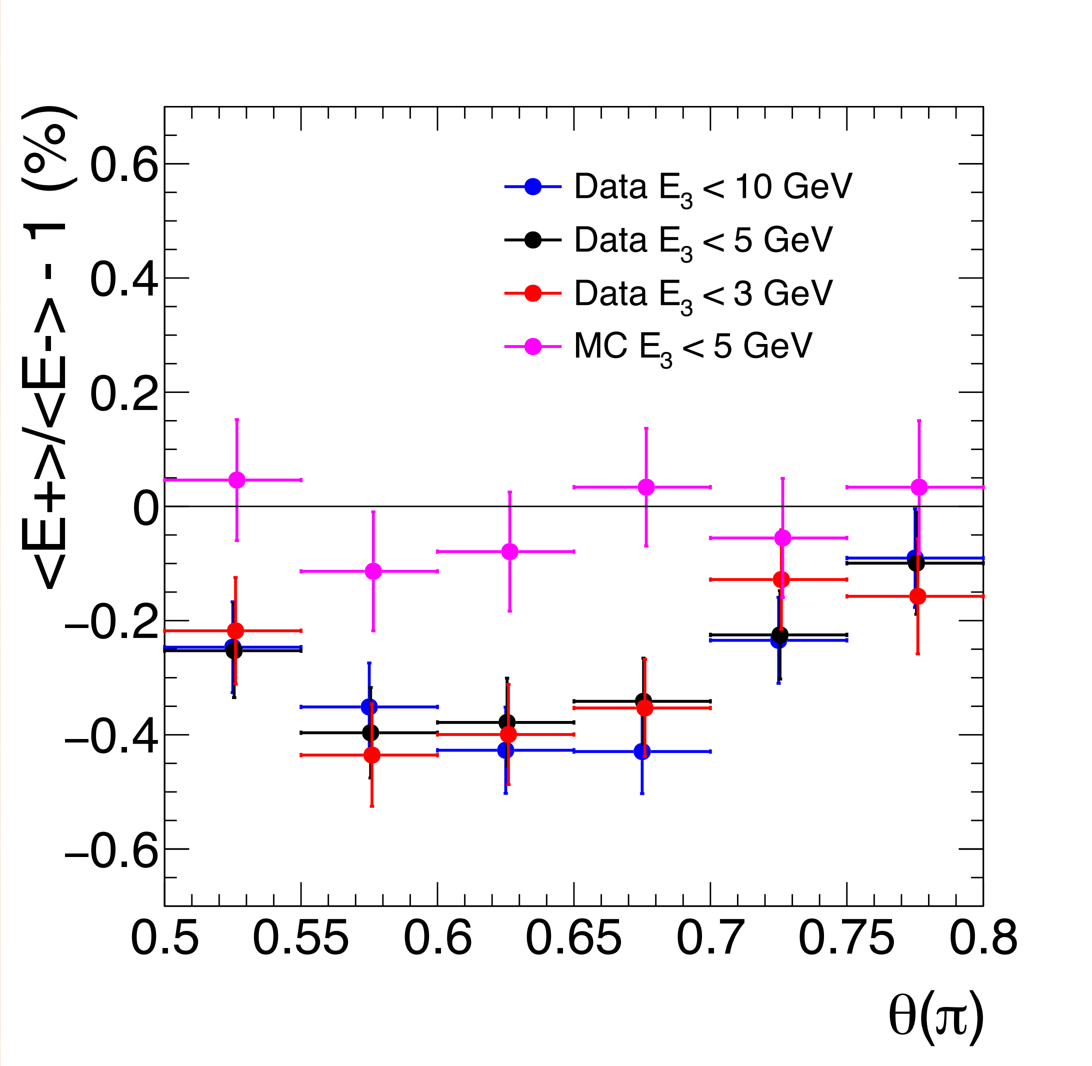
Energy leaking out around beam direction

arXiv 2108.04877

Residual calibration: step 1

Fiducial dijet, two sides of the detector



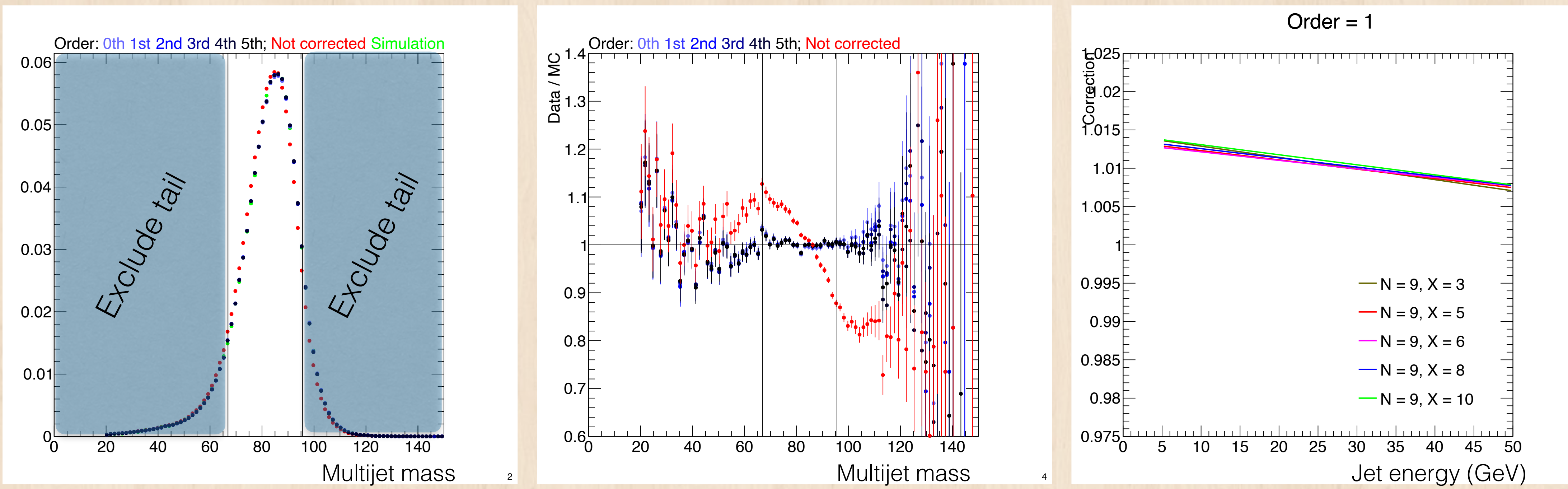


Look at data only, and calibrate out the difference between e^-- and e^+-going sides

arXiv 2108.04877

Residual calibration: step 2

Fiducial multijet invariant mass



Take up to leading N jet above X GeV

Vary N and X for systematics

Fit jet energy correction function parameters

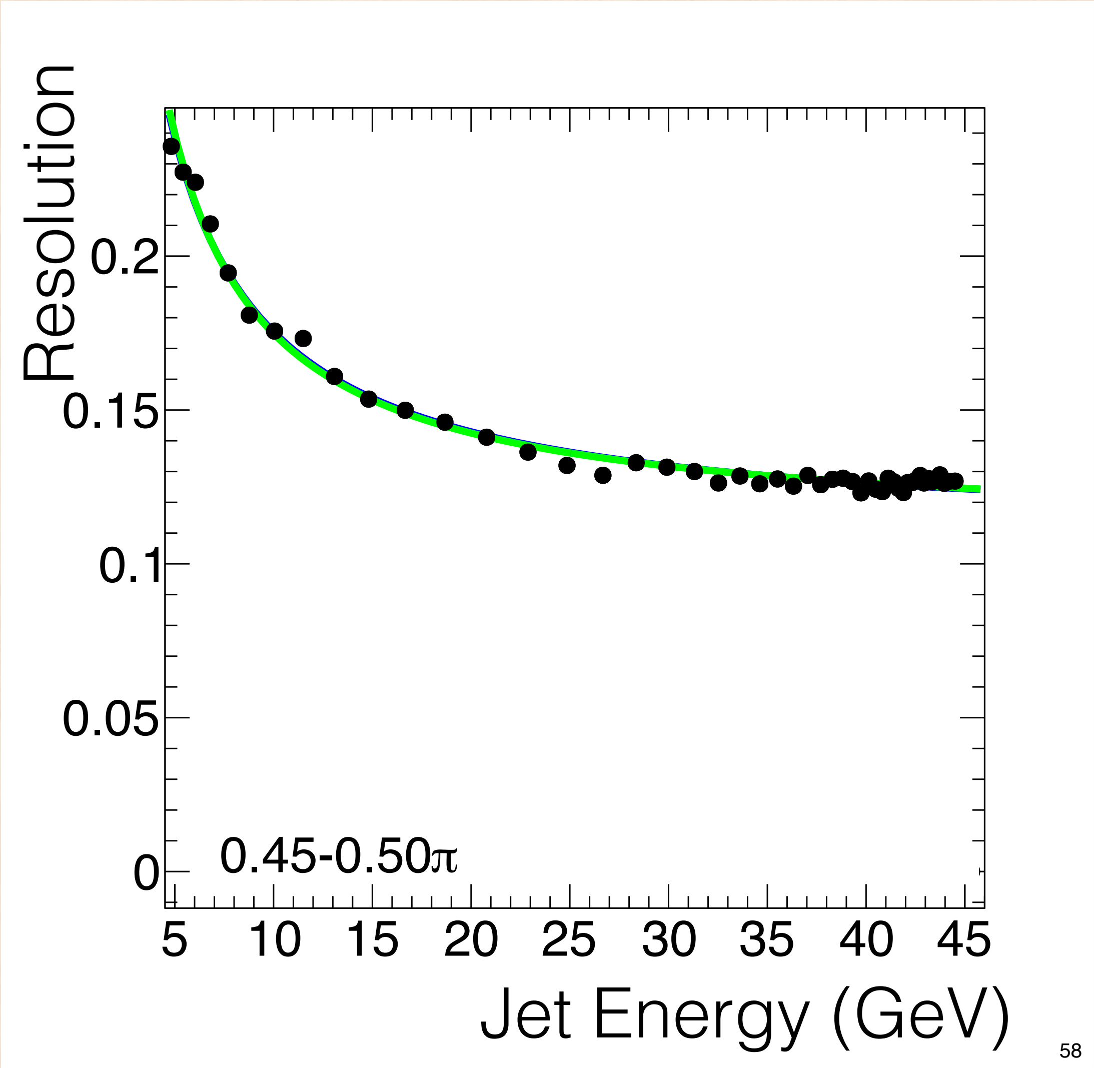
Minimize “quantile difference” (~KS) between data and MC curves

Nominal: linear correction as a function of energy

arXiv 2108.04877

Jet resolution

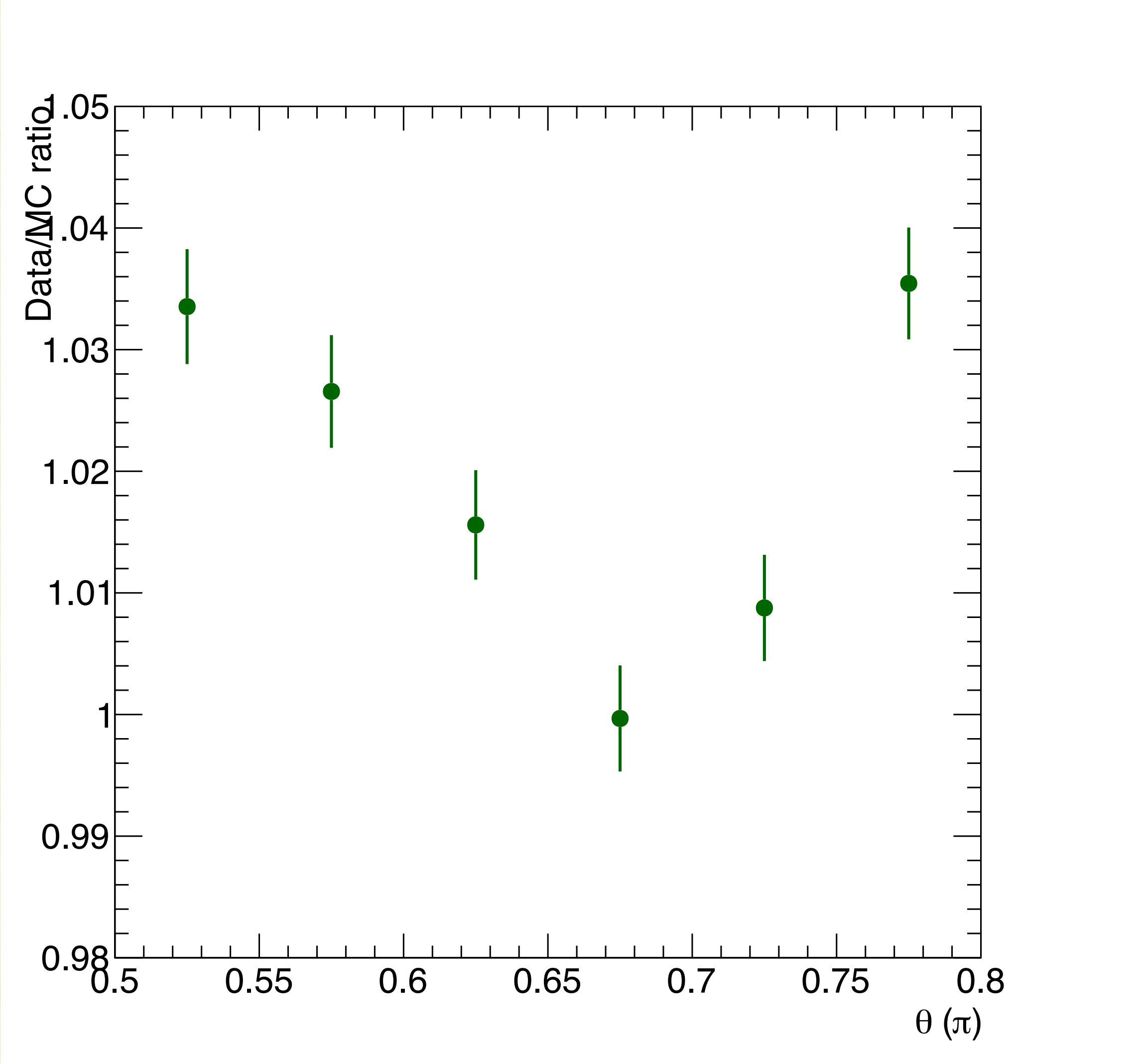
Jet resolution in simulation



Energy resolution: 10-25%

(Angular resolution: 0.01-0.05)

Fiducial dijet —
vary 3rd-leading jet as systematics



Up to 5% difference in energy resolution between data and MC

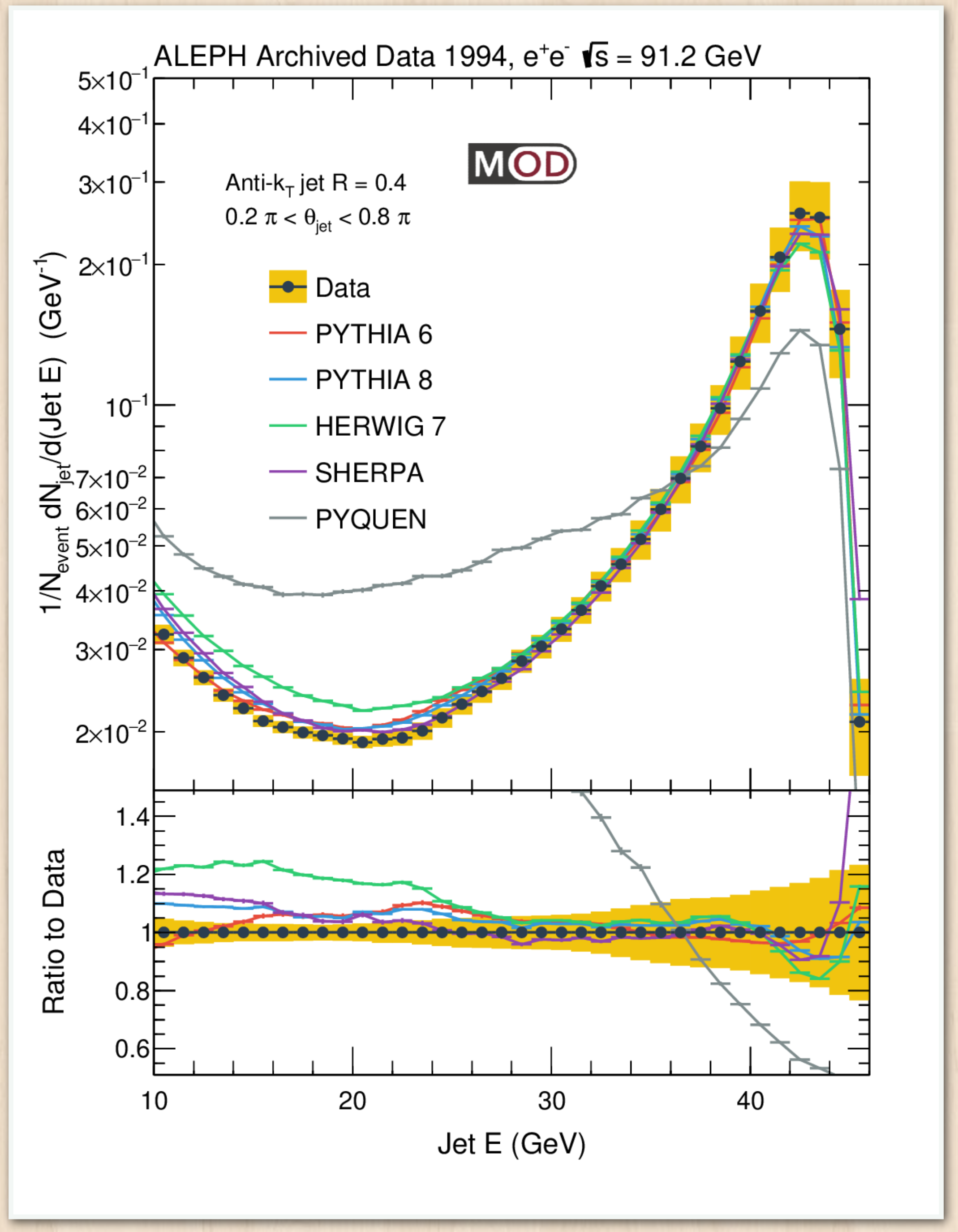
arXiv 2108.04877

Accessing the data: lessons

- Mileage vary a lot depending on the experiment
  - Some are in obscure **format** and hard to use
  - Sometimes it's not the most easy to get **control of stored information** (e.g. PID percentages) -- some lower level information will be useful
  - Sometimes only **one set of simulations** available
  - Enough information for end-to-end measurements?
- Many lessons for current & future experiments
  - Best to do some "user tests" for open data as we go

Some recent results

e^+e^-: jet energy



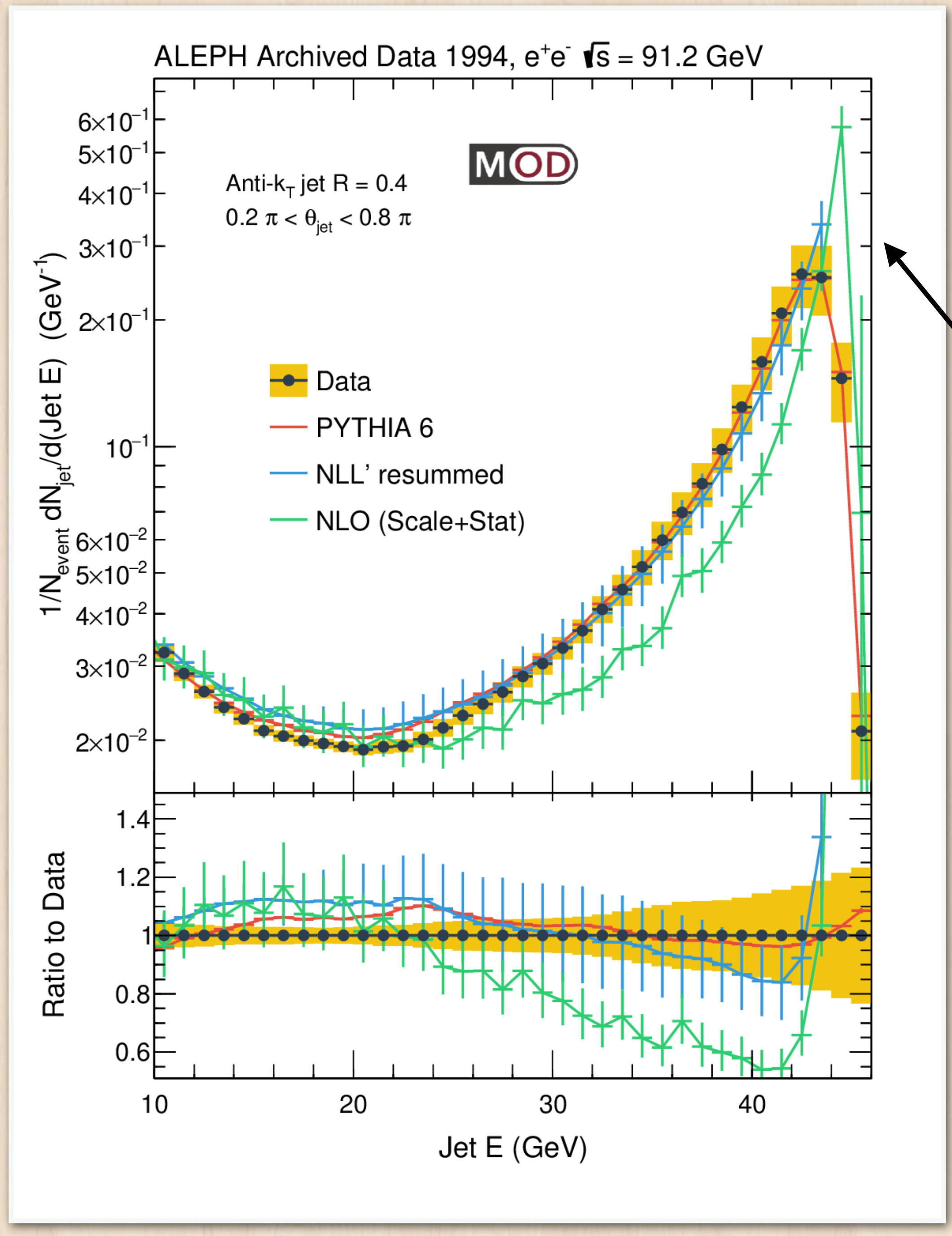
Comparison to MC and theory calculations (not shown in this plot)

Most generators can describe the peak region

Up to 10-20% disagreement at low E

→ out-of-cone energy, wide angle emission, …

e^+e^-: jet energy



LO parton level
= delta function at 45 GeV
not too interesting to plot

NLO parton level sharper than measured data

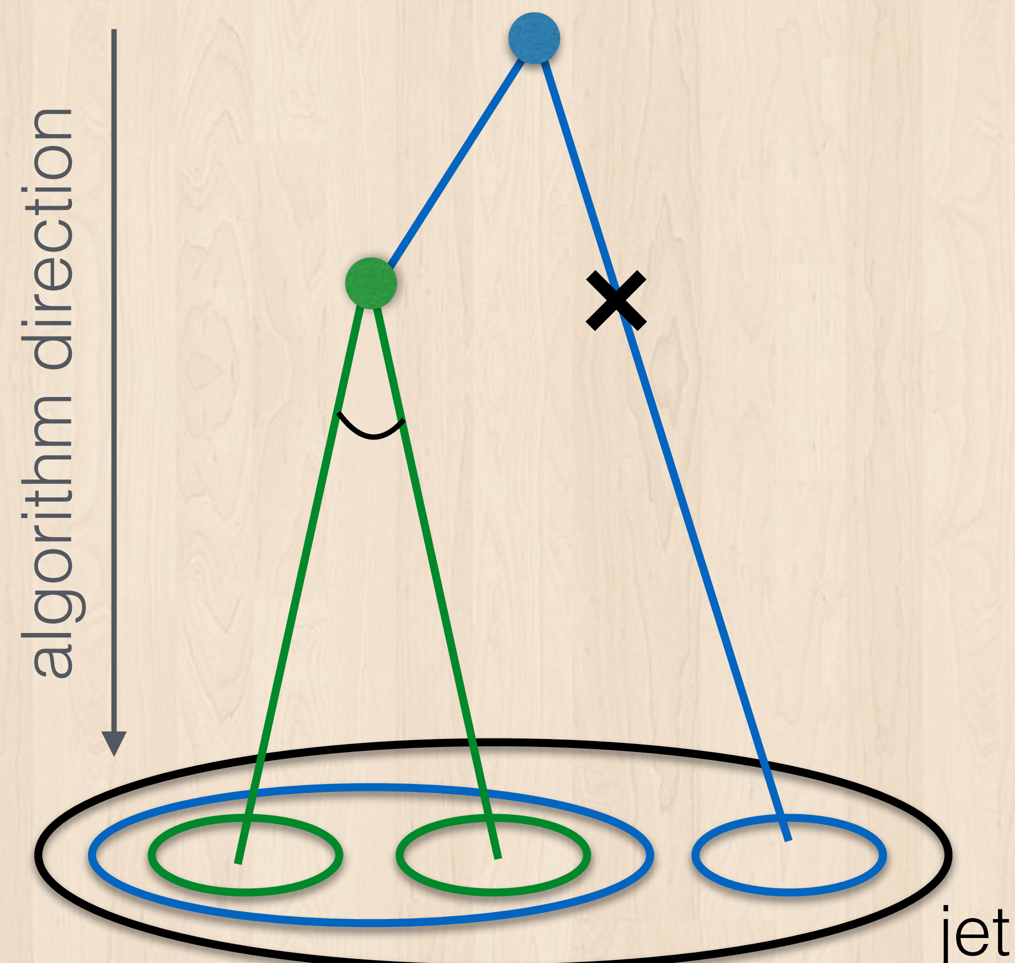
NLL' resummed generally describe data

Private communication with João Pires

Ringer, Sato, Neill JHEP 07 (2021) 041

Jet grooming

Soft drop/mMDT grooming



$(z_{\text{cut}}, \beta) = (0.1, 0.0)$

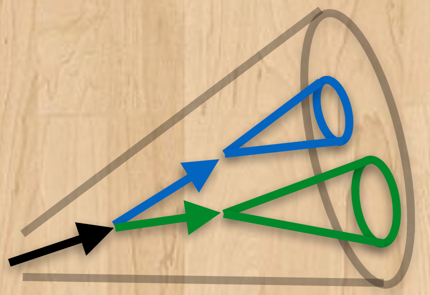
Recluster jet constituents with C/A algorithm

Sequentially open up jet until condition is met

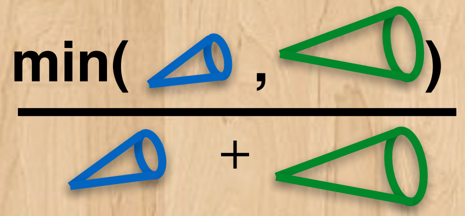
$$z \equiv \frac{\min(E_1, E_2)}{E_1 + E_2} > z_{\text{cut}} \left(\frac{\theta_{12}}{R} \right)^{\beta}$$

E instead of p_T $\quad$ θ_{12} = real angle

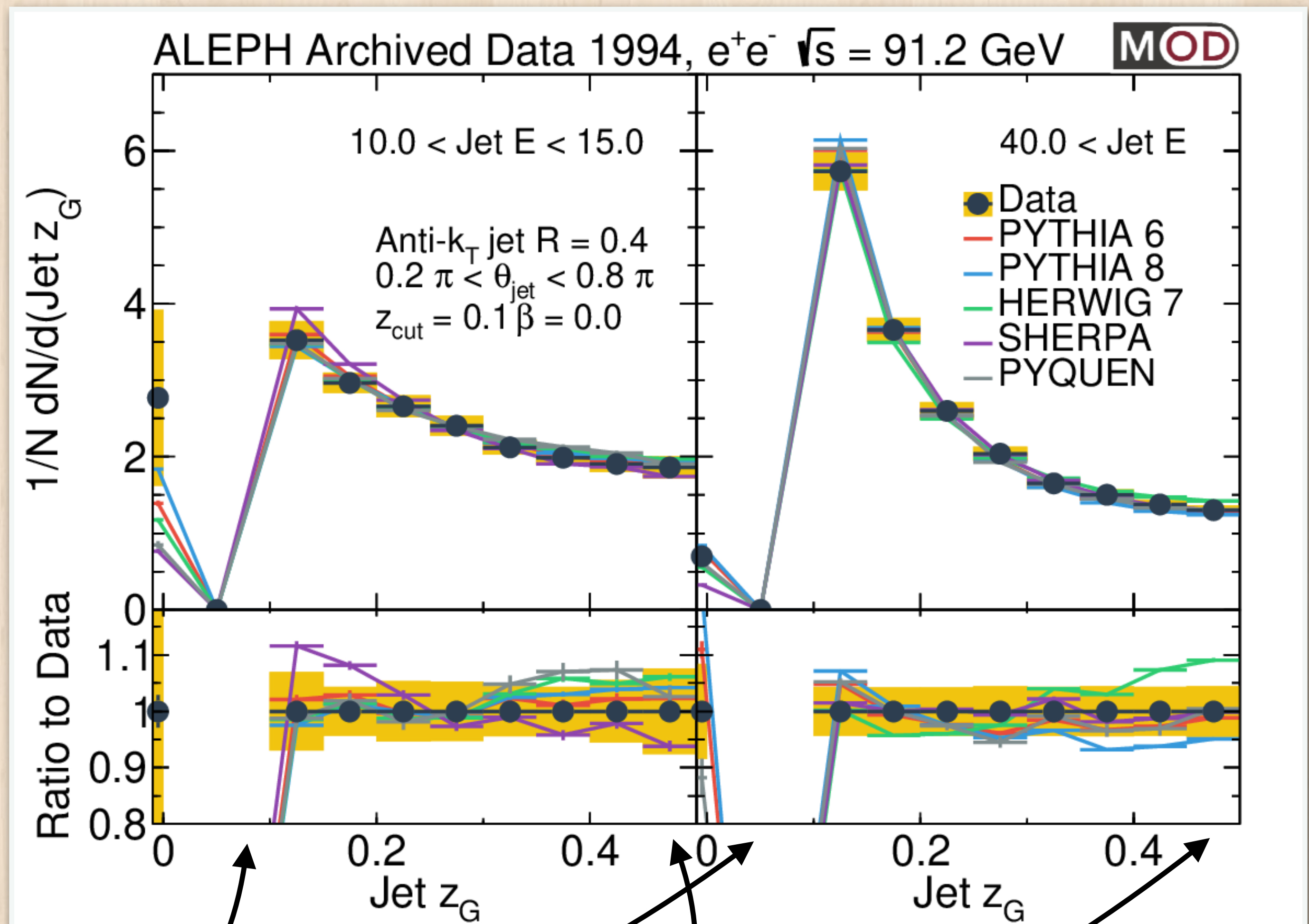
r_g = opening angle
z_g = energy sharing
M_g = invariant mass

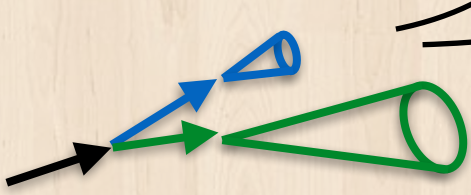


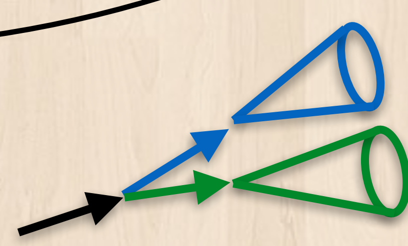
Energy sharing z_G



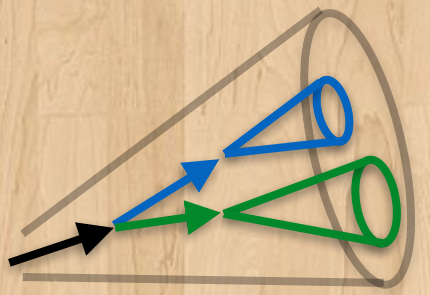
Measurement binned in energy (most not shown)



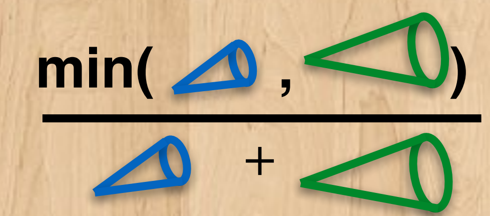


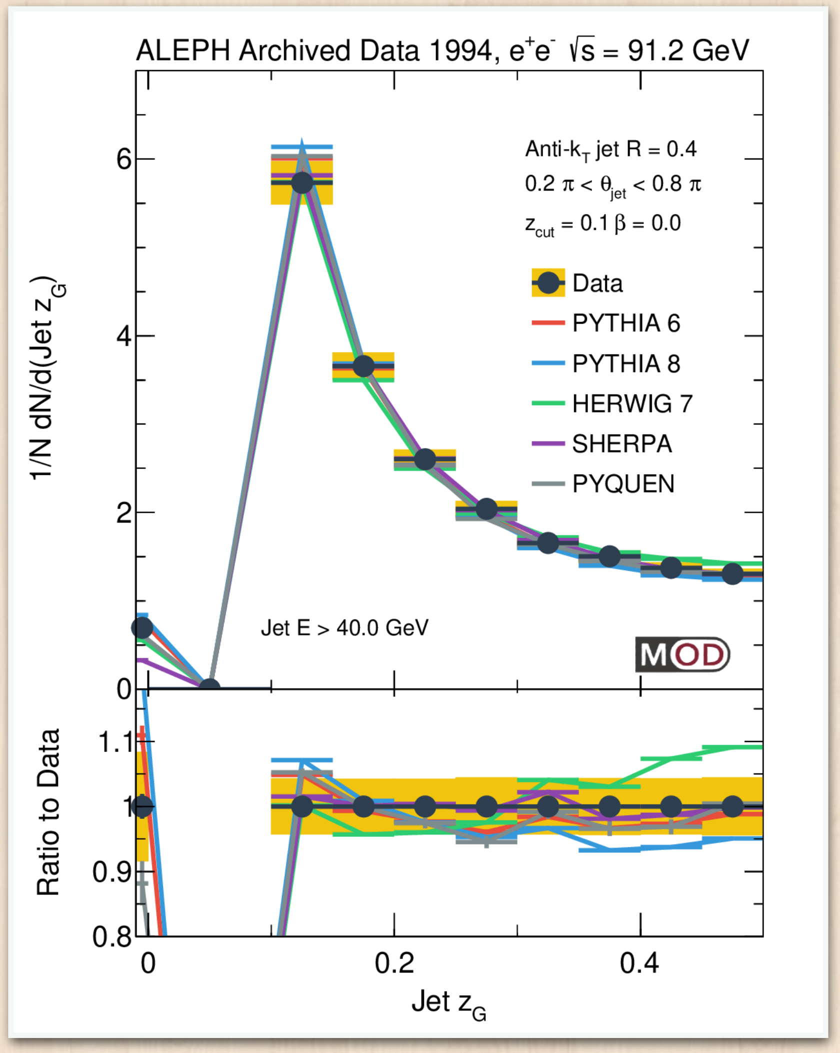


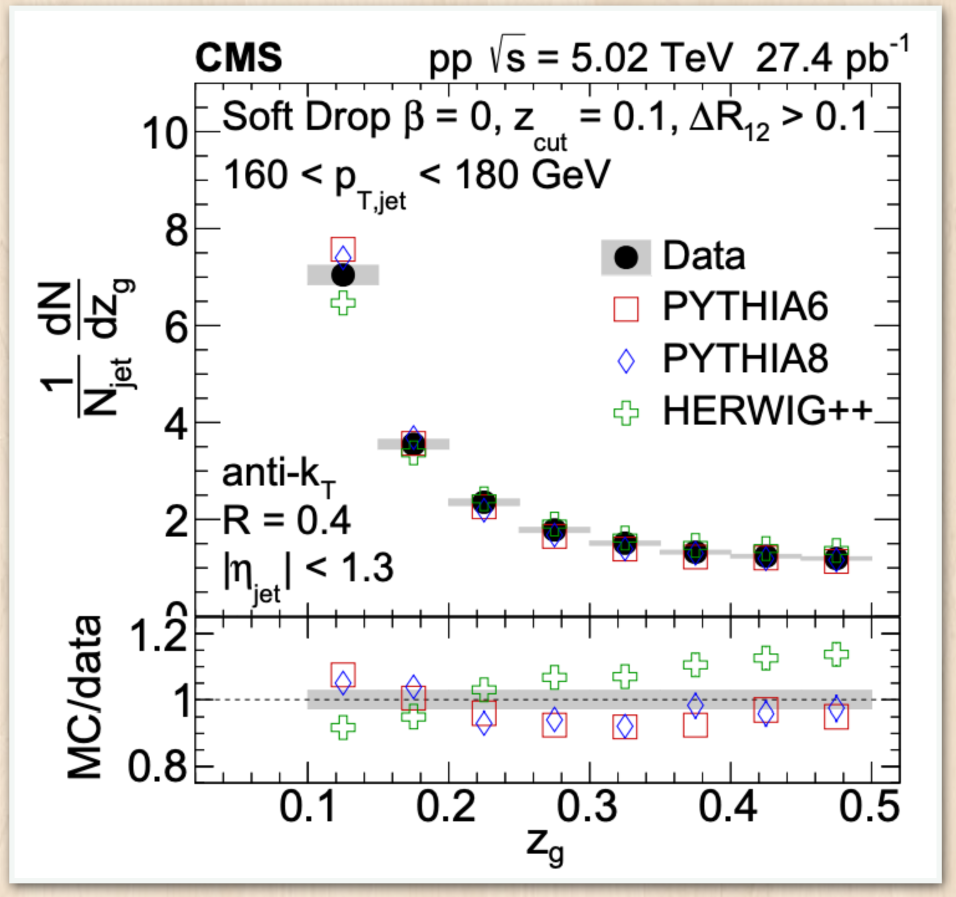
High energy
More similar to the $1/z$
from splitting function



Energy sharing z_G



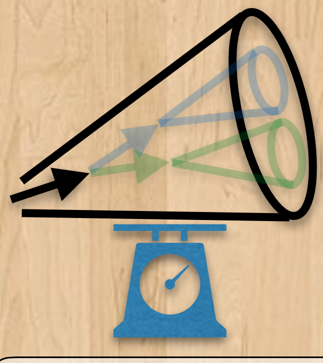




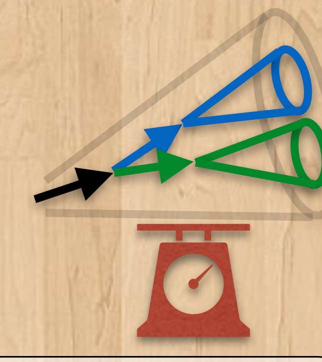
At high energy similar to LHC results

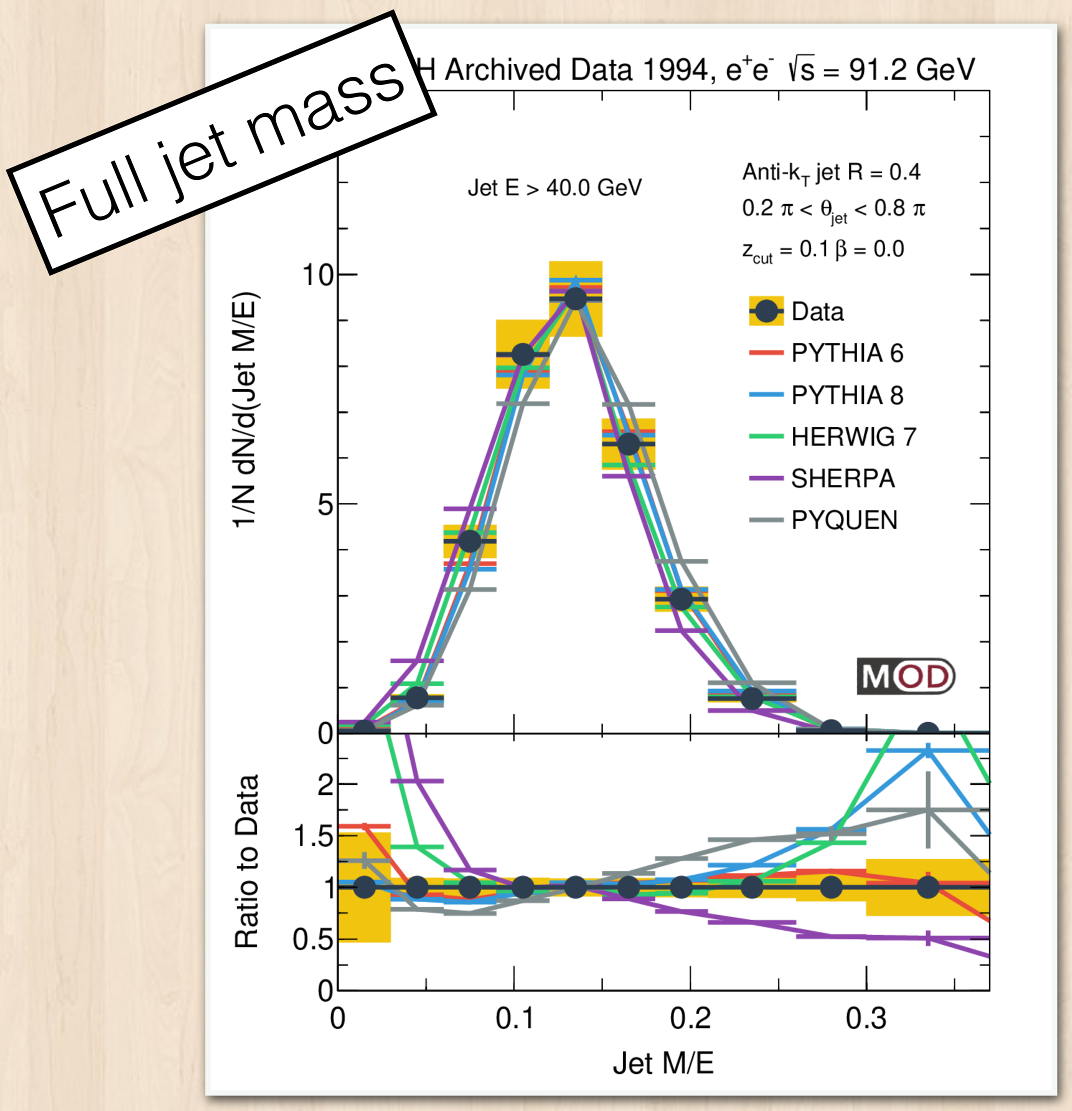
Comparison to **PYTHIA** and **HERWIG** also similar

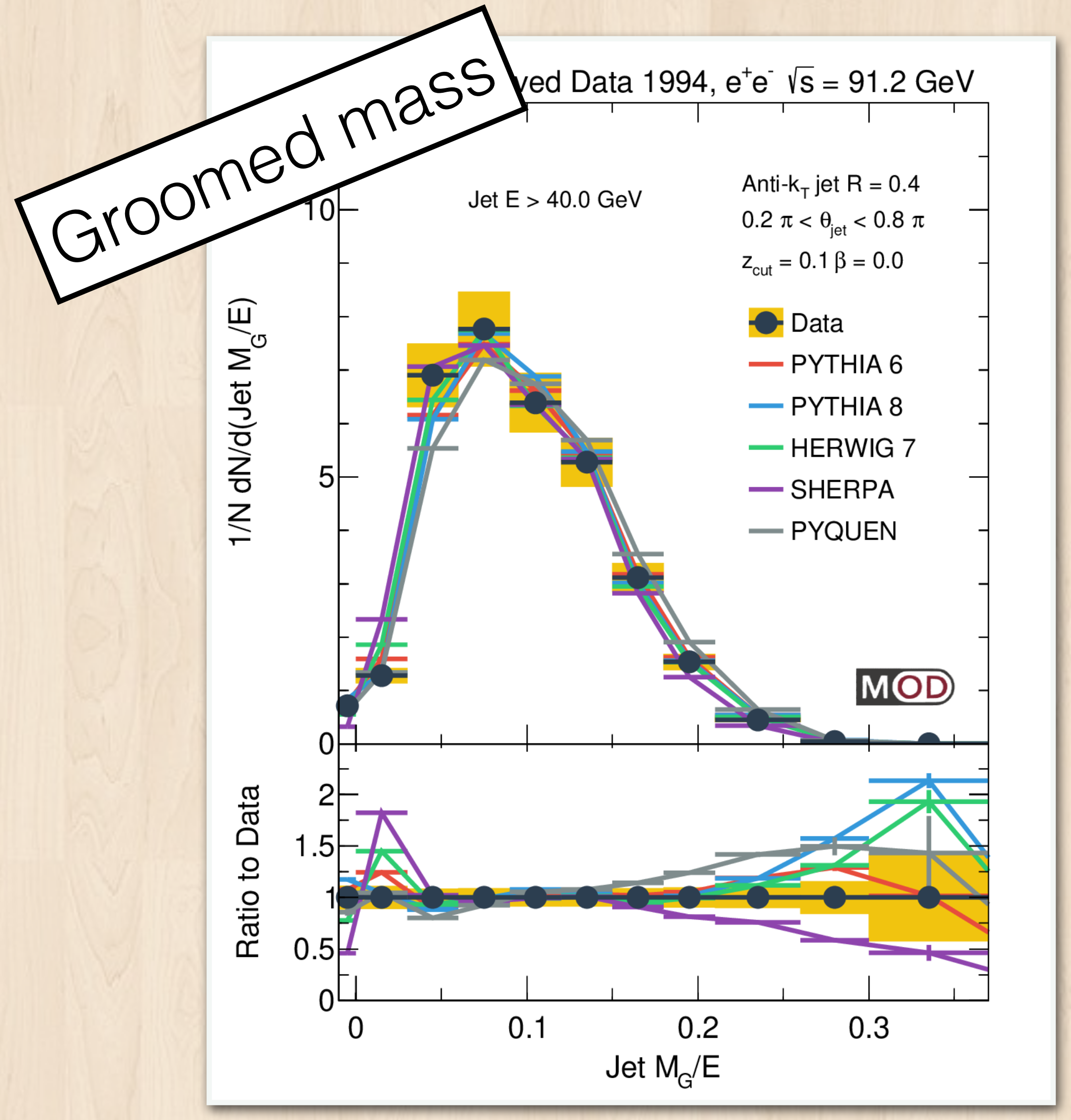
Disagreement in LHC can be improved by e^+e^- input



Jet mass





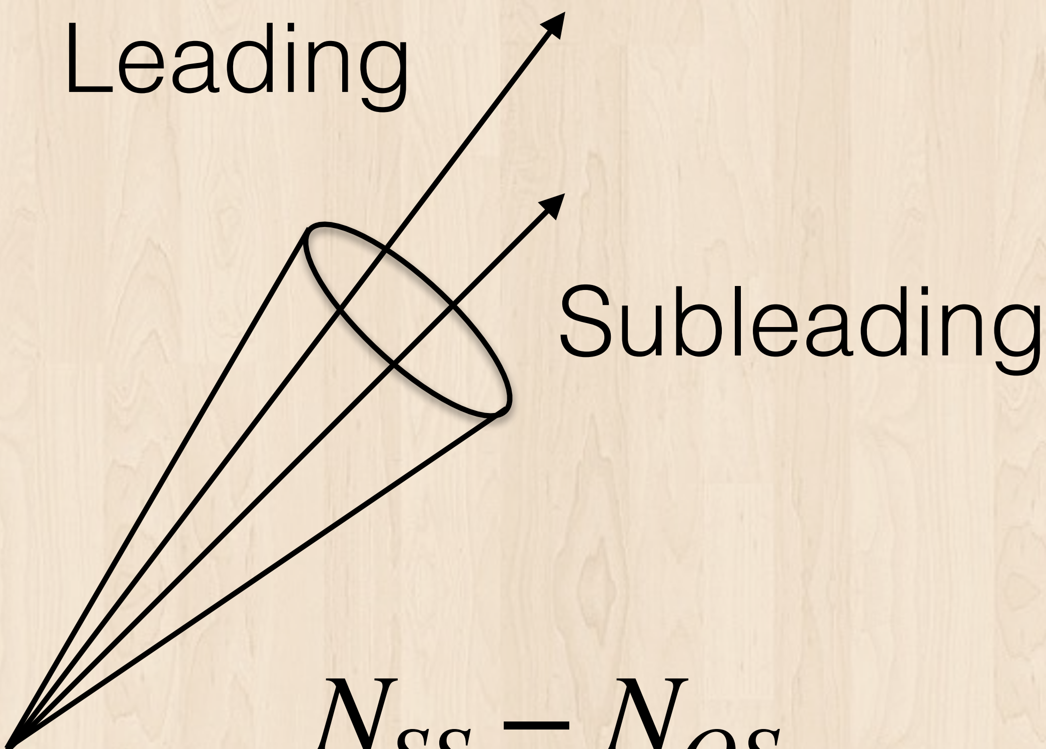


General shape vs tail

Explicit $(M - M_G)/E$

Interesting to compare to higher order generators

Leading particles (in jet) as probe

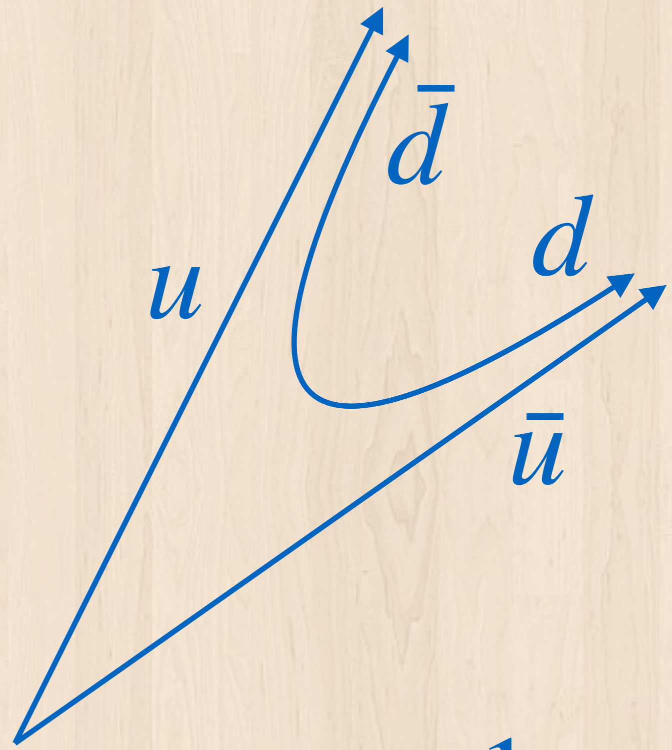


$$r_c \equiv \frac{N_{SS} - N_{OS}}{N_{SS} + N_{OS}}$$

anti-k_T jet, $R = 1.0$

Sensitive to hadronization

“Alternating” picture



$r_c = -1$

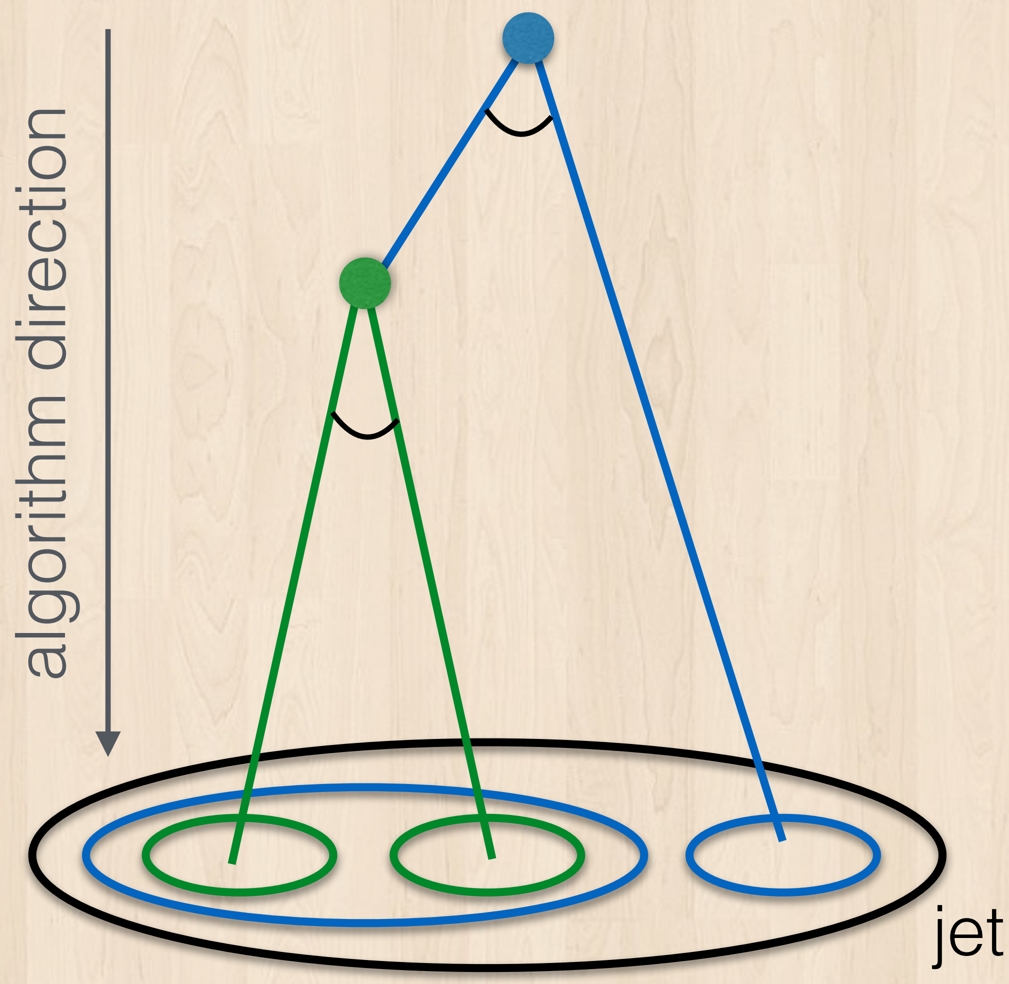
“Random” picture

$r_c = 0$

PRD 105, 051502

Also: subjet as proxy

Recursive soft drop



$(z_{\text{cut}}, \beta) = (0.2, 1.0)$

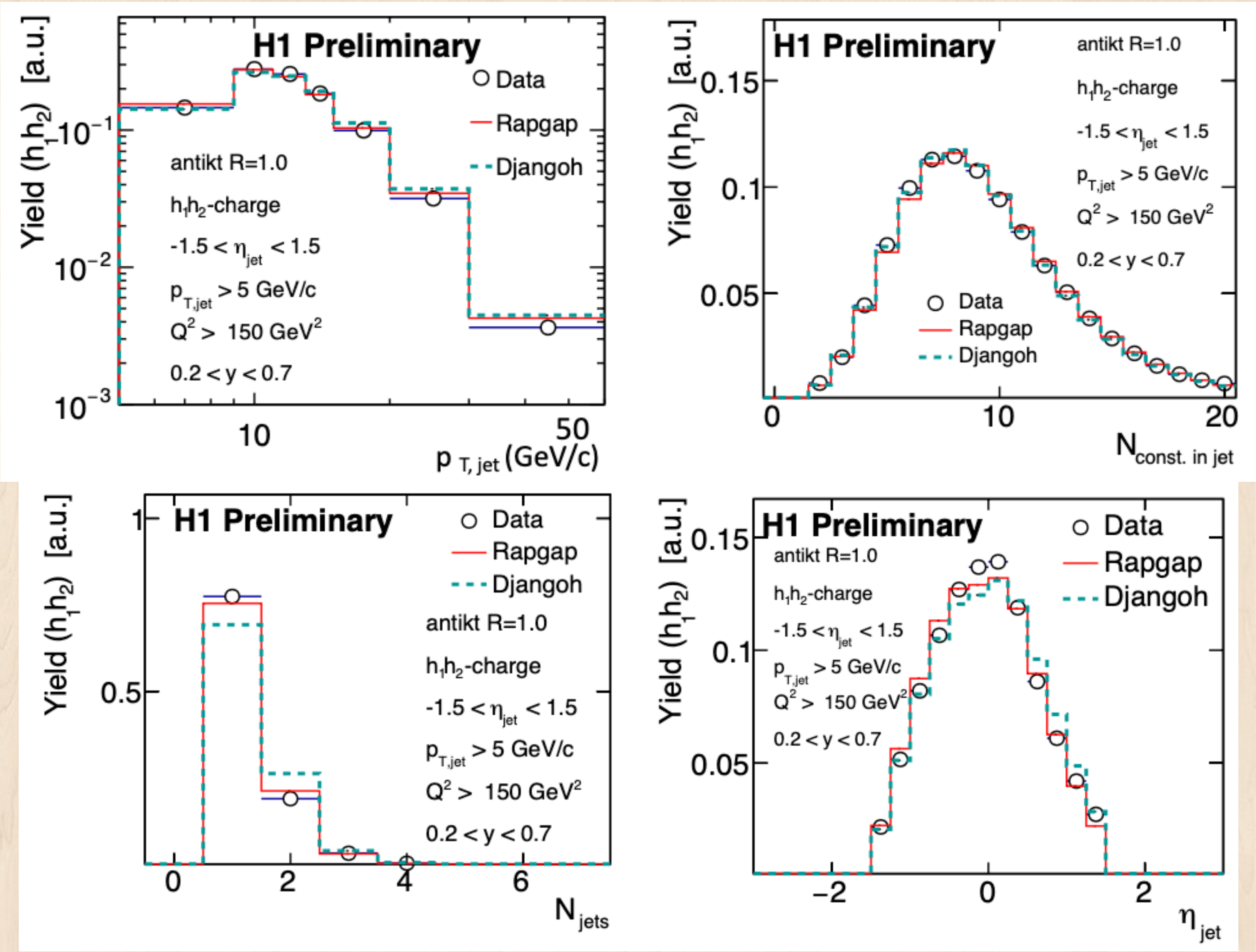
In addition to particles, also look at subjets as proxy to partons

Study first split and later splits

JHEP 06 (2018) 093

ep: jet kinematics (detector level)

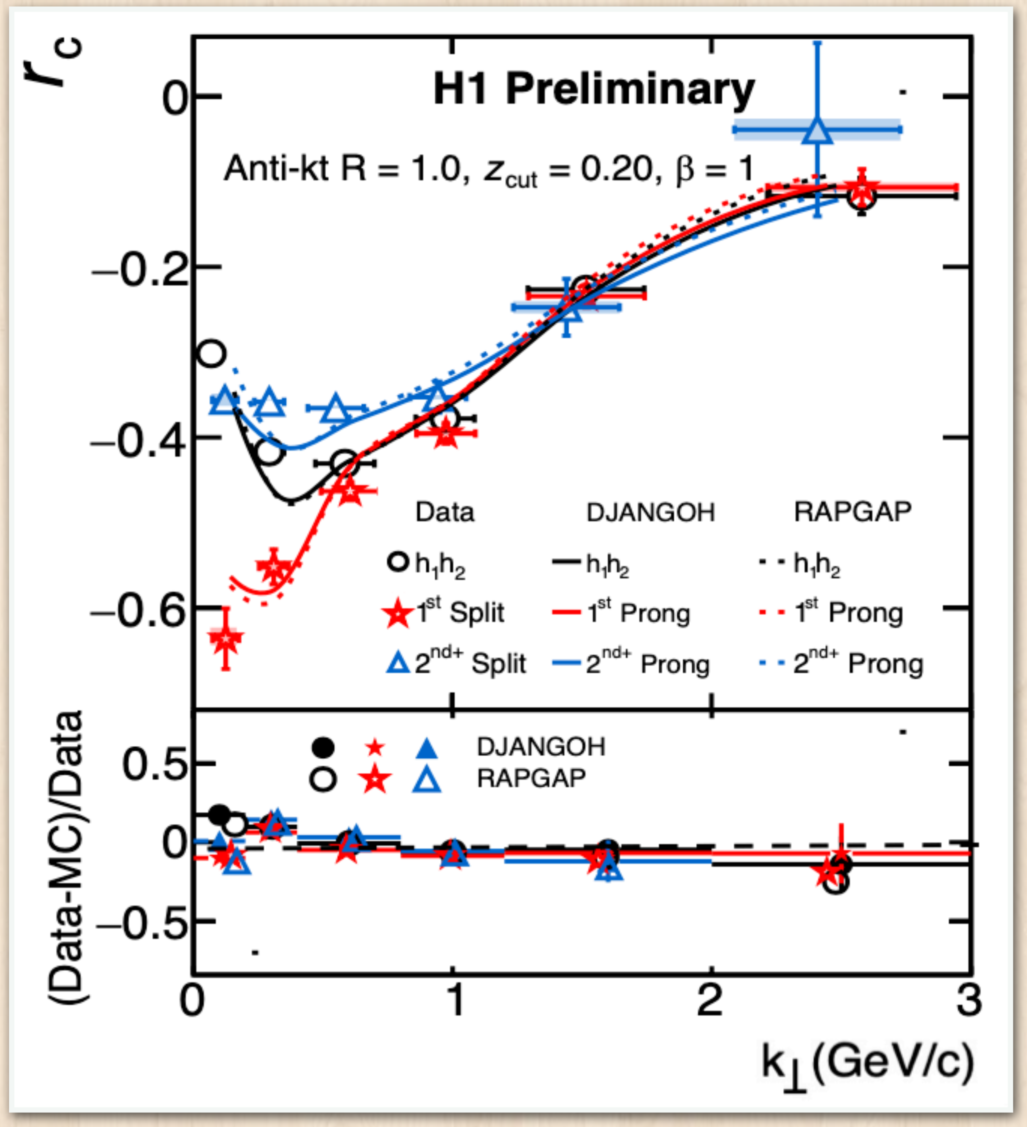
$E_{CM} = 319$ GeV



Well-reproduced by simulation

ep: r_c vs. $k_\perp$

$E_{CM} = 319$ GeV



Data generally decently reproduced by simulation

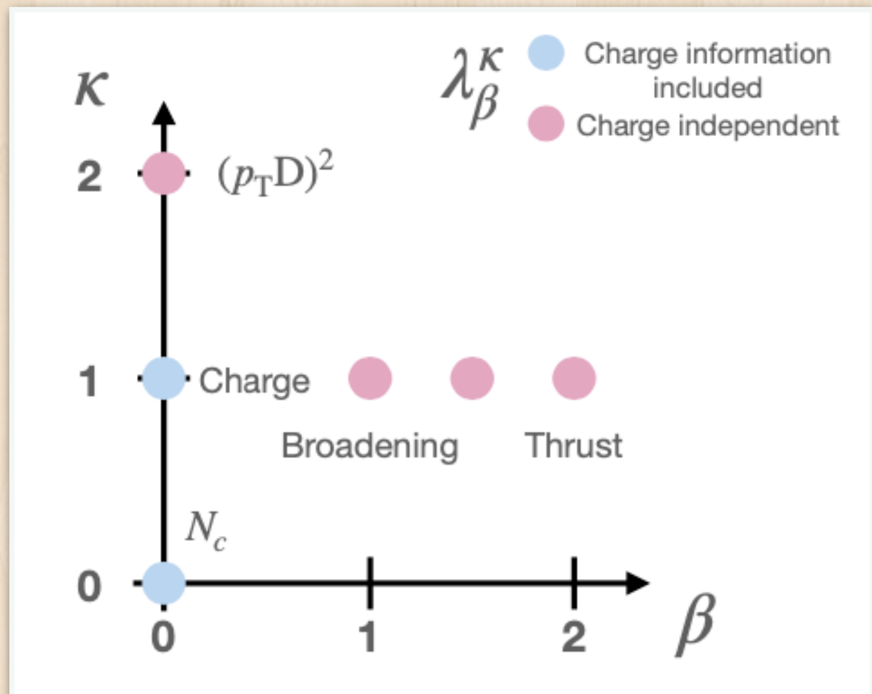
At large $k_\perp$, $r_c \sim 0$

r_c tends negative at small $k_\perp$

$k_\perp$ = relative transverse momentum

H1Prelim-22-032

Generalized jet angularities

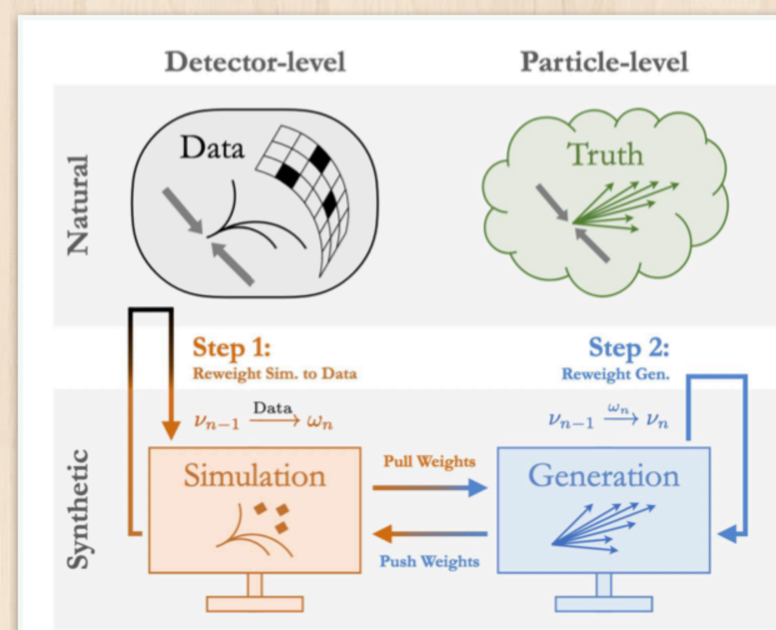


$$\lambda_\beta^\kappa = \sum_i z_i^\kappa \left(\frac{R_i}{R_0}\right)^\beta$$

$$\tilde{\lambda}_0^\kappa = \sum_i q_i z_i^\kappa$$

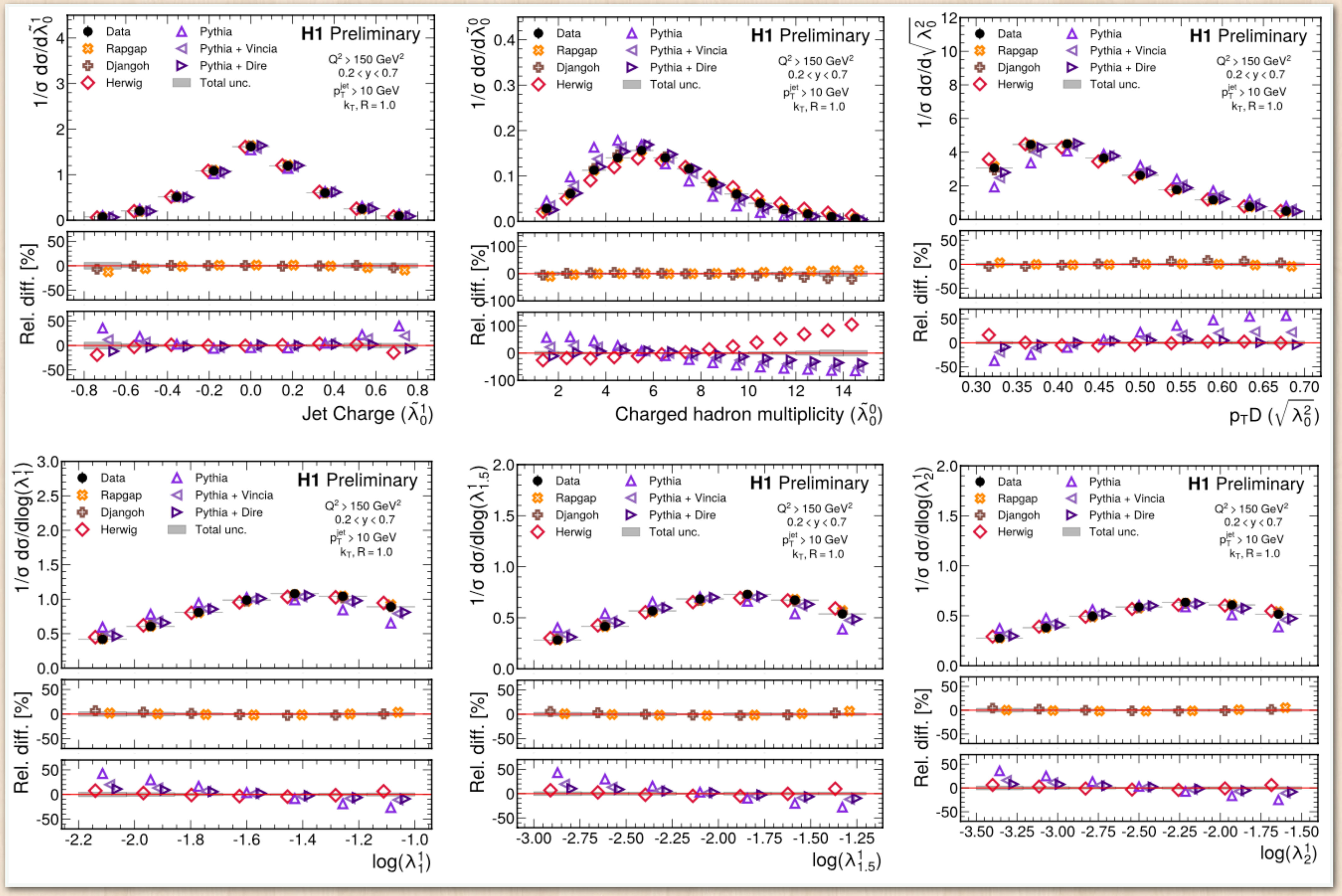
k_T jet, $R = 1.0$

Measurements unfolded using the Omnifold algorithm



ep: angularity results

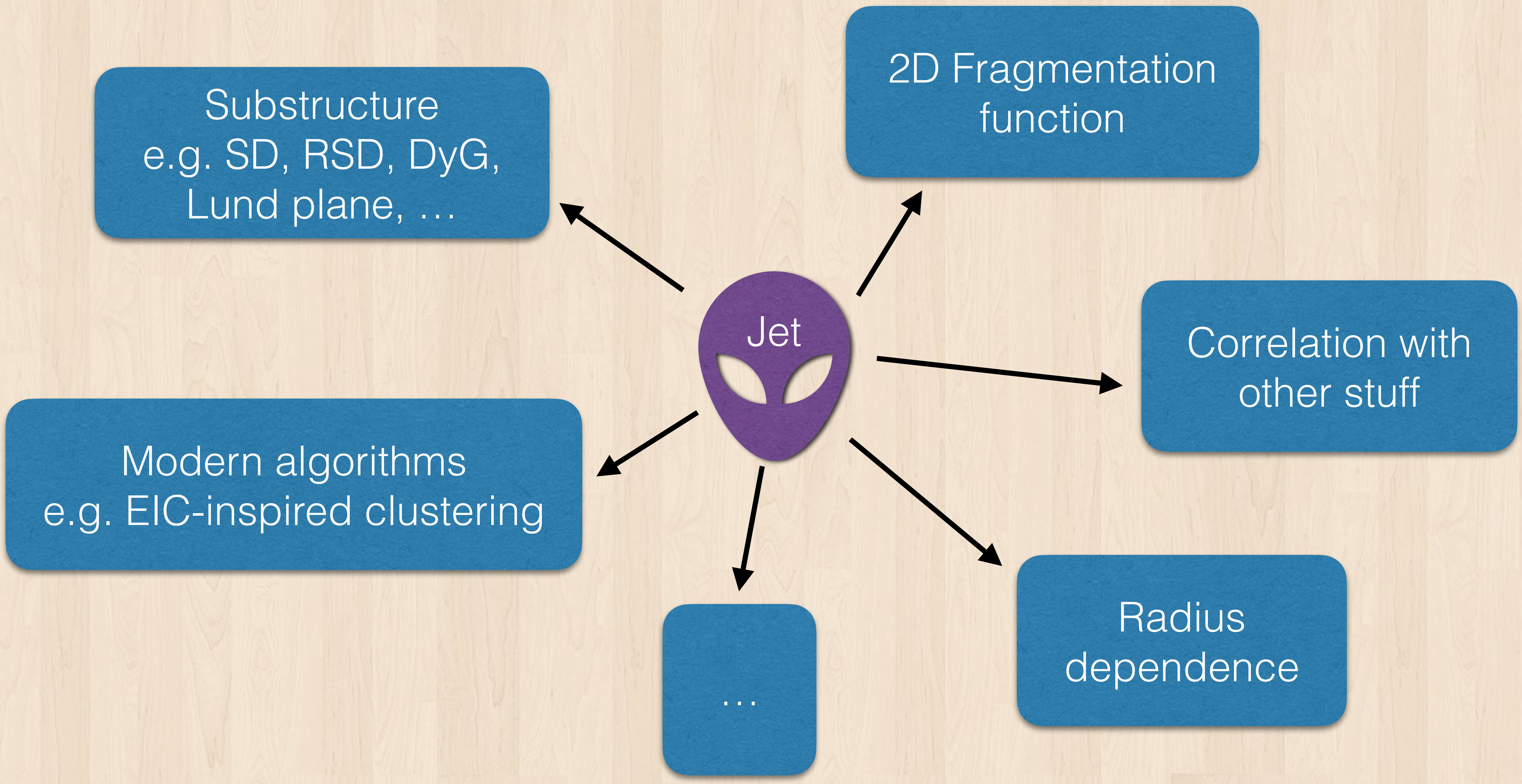
$E_{CM} = 319$ GeV



H1Prelim-22-034

Concluding remarks

A lot of potential



Testing ground for new algorithm developments

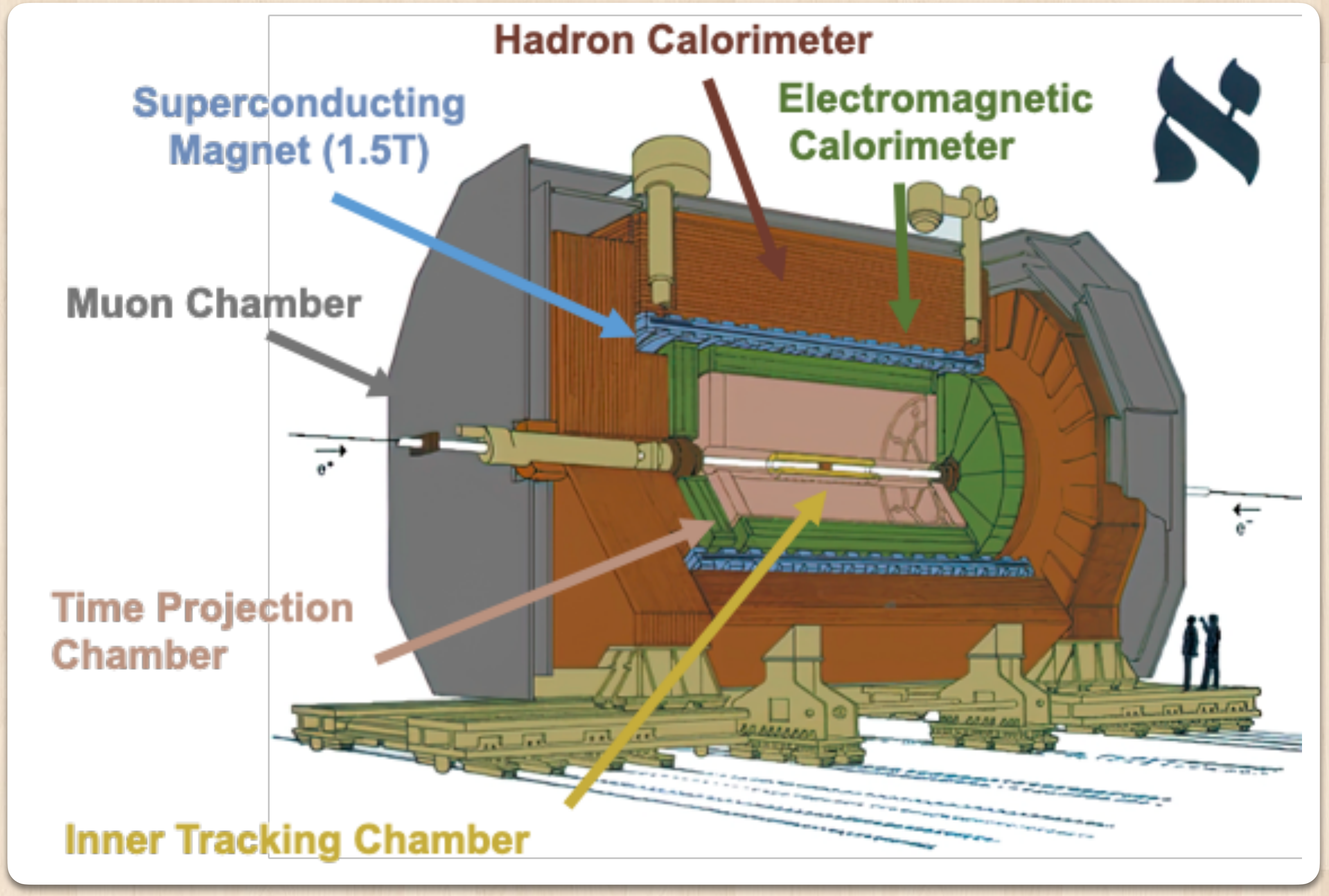
Provide reference measurements

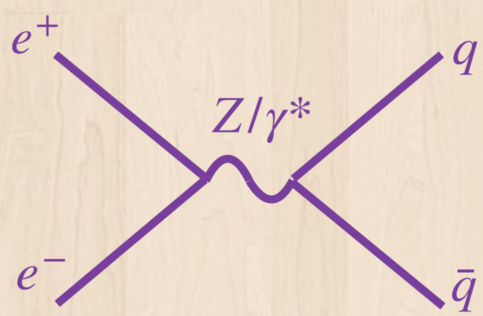
Summary

- Interesting recent jet substructure results from e^+e^- and ep
    - Other efforts ongoing
- A lot of untapped potential
- Lessons for the ongoing & future experiments in terms of data archiving

Backup Slides Ahead

ALEPH detector





LEP1 e^+e^- data taken at 91.2 GeV from 1992-1995

About 2.5M recorded hadronic events with ALEPH

Recent ZEUS jet alphas