Measurement of collinear drop jet mass and its correlation with groomed jet substructure observables in pp collisions

Youqi Song (youqi.song@yale.edu)
for the STAR Collaboration

DIS 2023, East Lansing, Michigan

3/28/2023

WG4: QCD with Heavy Flavours and Hadronic Final States

Supported in part by U.S. DEPARTMENT OF ENERGY Office of Science

Motivation

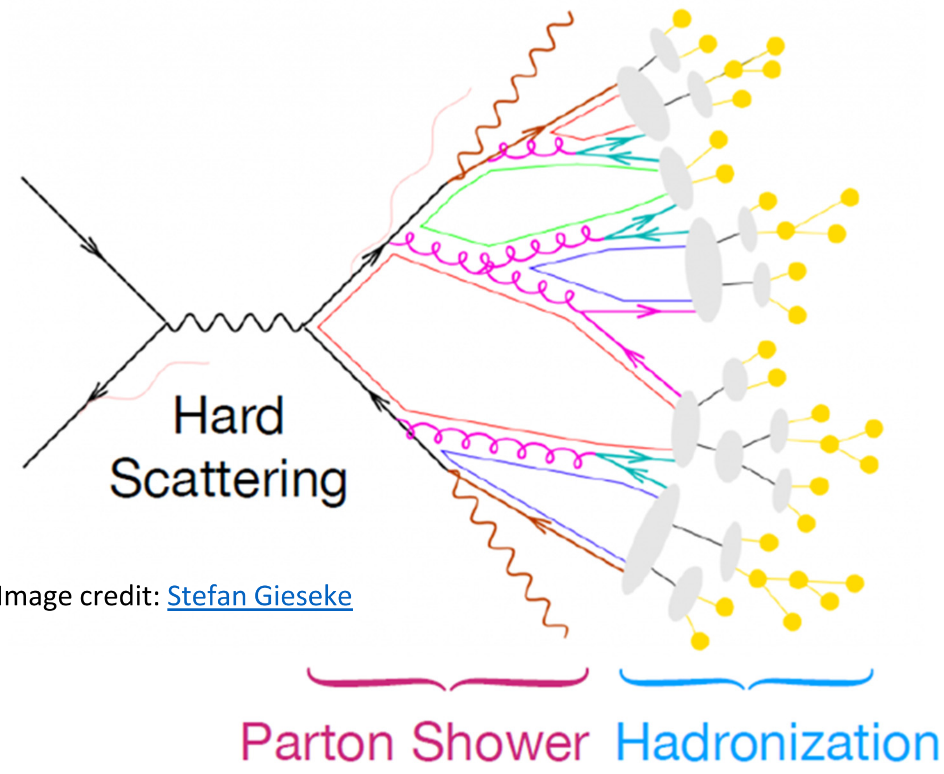
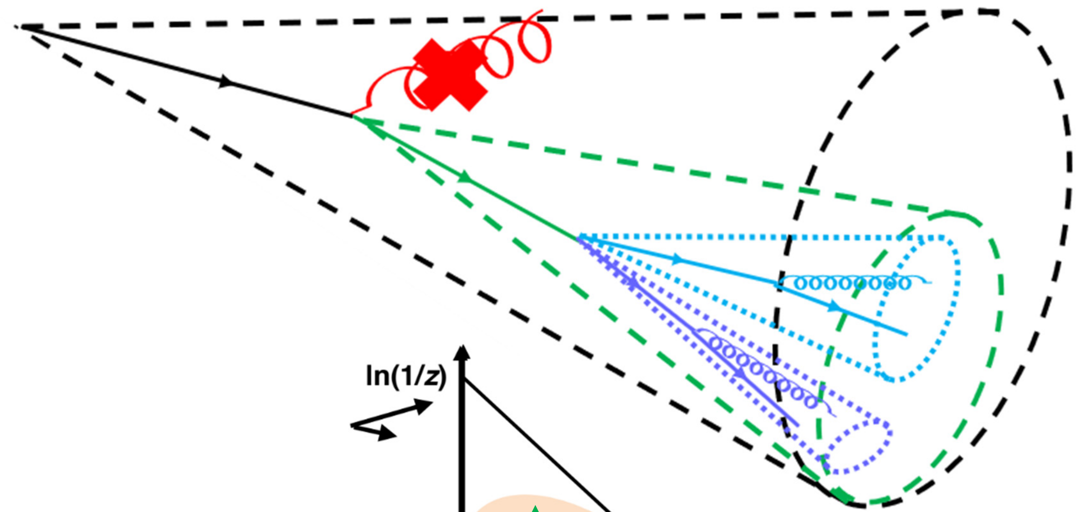


Image credit: Stefan Gieseke

To enhance perturbative contributions, SoftDrop grooming is often used to remove soft wide-angle radiation



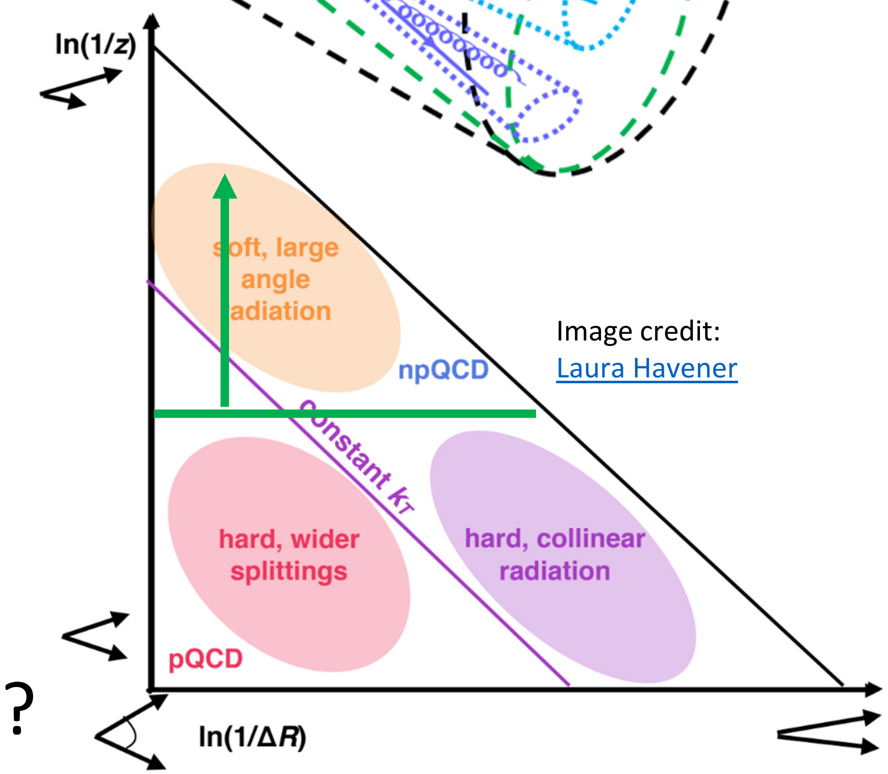


Image credit: Laura Havener

Jet is a multi-scale object

What about the wide-angle soft radiation?

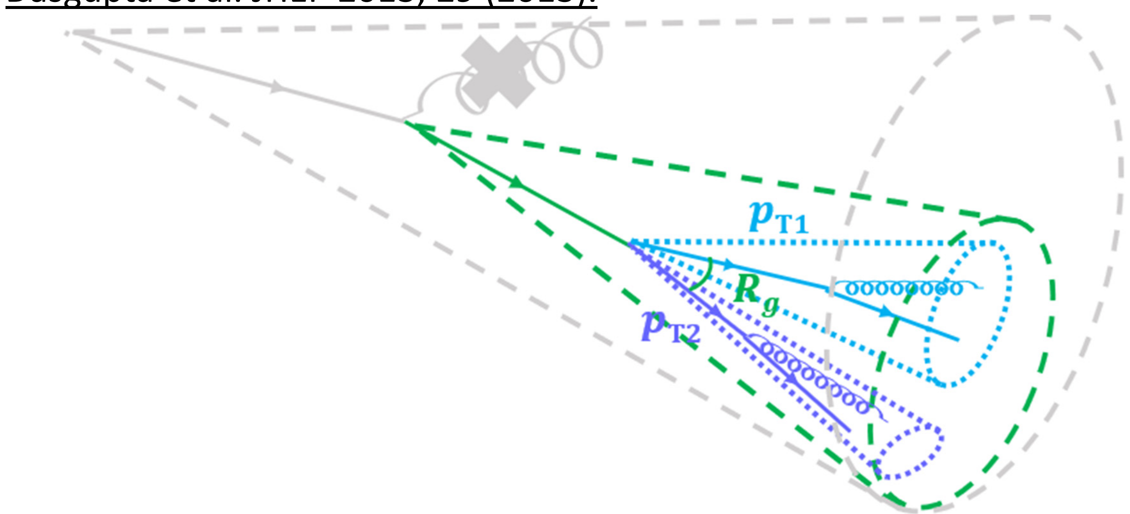
What is the soft-hard correlation within a jet?

Motivation

1. soft radiation within the jet → collinear drop jet observable

SoftDrop: removes wide-angle soft radiation

Larkoski, et al. JHEP 2014, 146 (2014).
Dasgupta et al. JHEP 2013, 29 (2013).



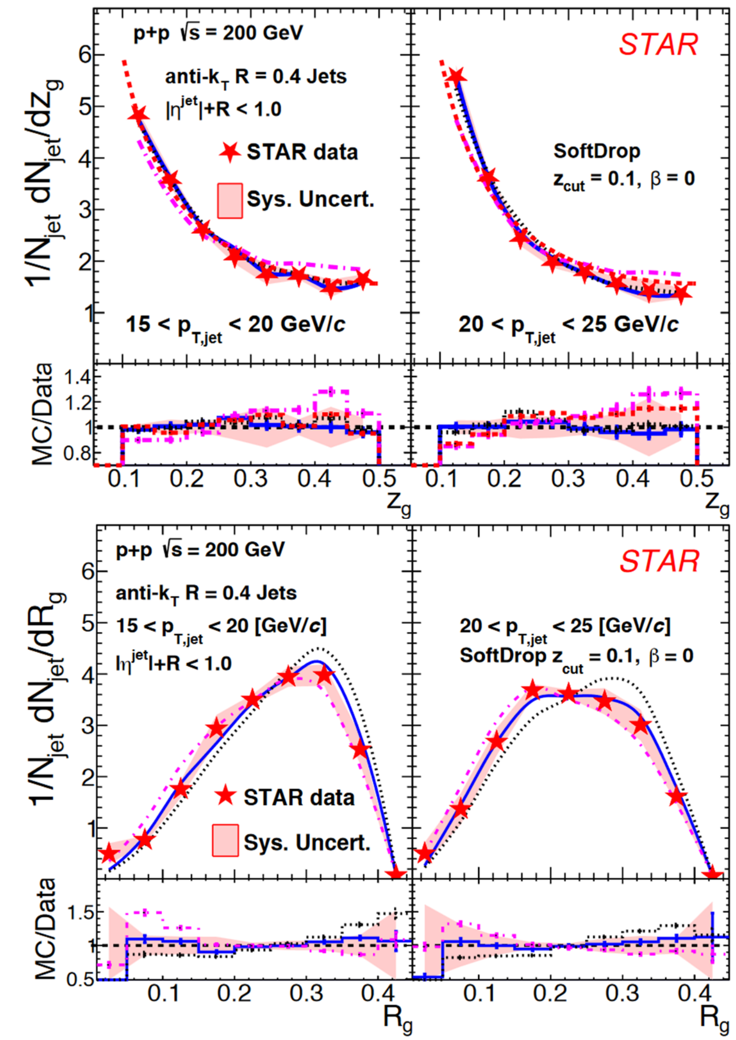
Require subjet momentum fraction to pass

$$z_g = \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut}(R_g/R_{jet})^\beta \qquad z_{cut} = 0.1 \quad \beta = 0$$

$$\rightarrow \quad \frac{\min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > 0.1$$

z_g

R_g



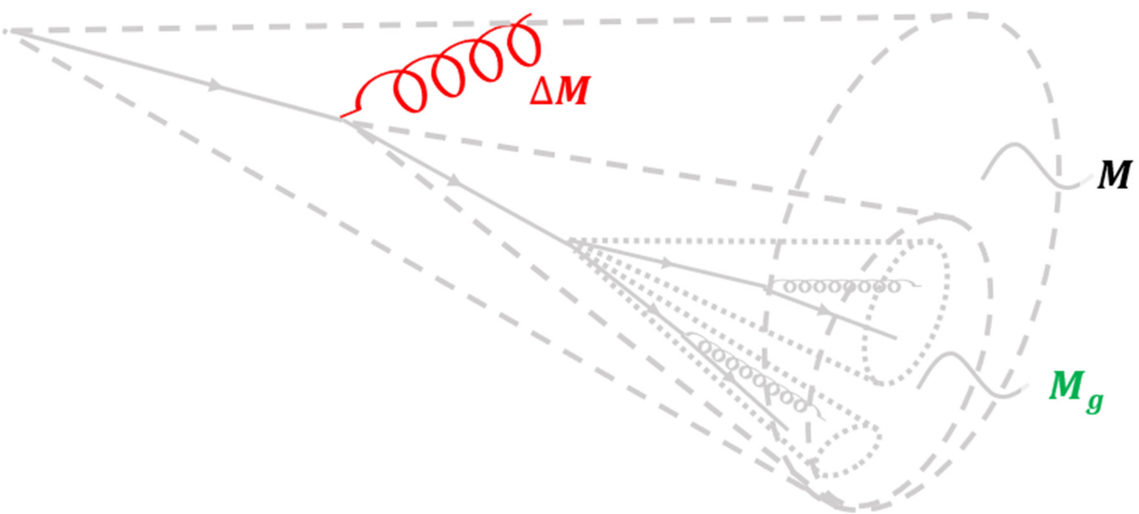
STAR. PLB 811 (2020) 135846

Motivation

1. soft radiation within the jet → collinear drop jet observable

Collinear Drop: probes the soft component

Chien and Stewart JHEP 2020, 64 (2020).



- General case: difference of an observable with two different SoftDrop selections $(z_{cut\ 1}, \beta_1)$ and $(z_{cut\ 2}, \beta_2)$
- For this analysis, $(z_{cut\ 1}, \beta_1) = (0,0)$ and $(z_{cut\ 2}, \beta_2) = (0.1,0)$: difference in the original and SoftDrop groomed observable
- Observables: e.g., $$\Delta M/M = \frac{M - M_g}{M}$$

where $$M = \left|\Sigma_{i\in \mathrm{jet}}\ p_i\right| = \sqrt{E^2 - |\vec{p}|^2}$$

Motivation

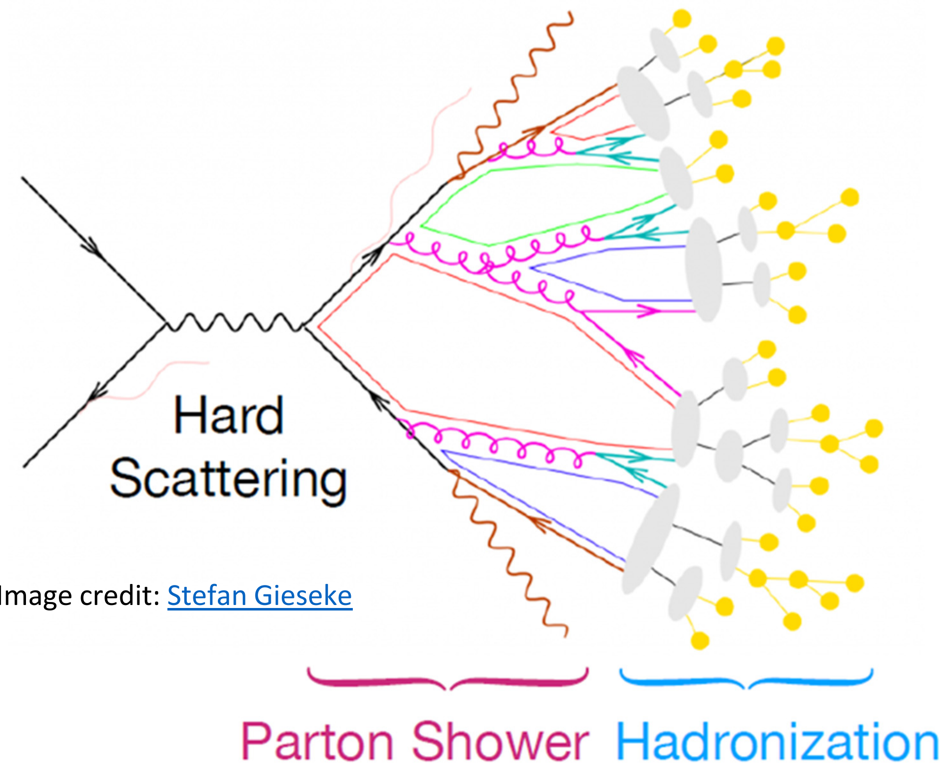
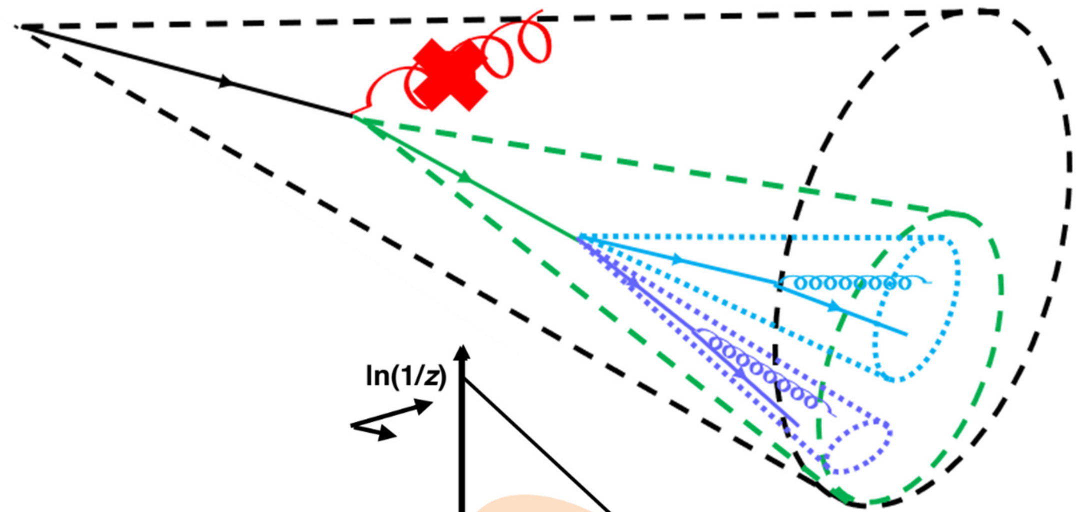


Image credit: Stefan Gieseke

Jet is a multi-scale object

To enhance perturbative contributions, SoftDrop grooming is often used to remove soft wide-angle radiation



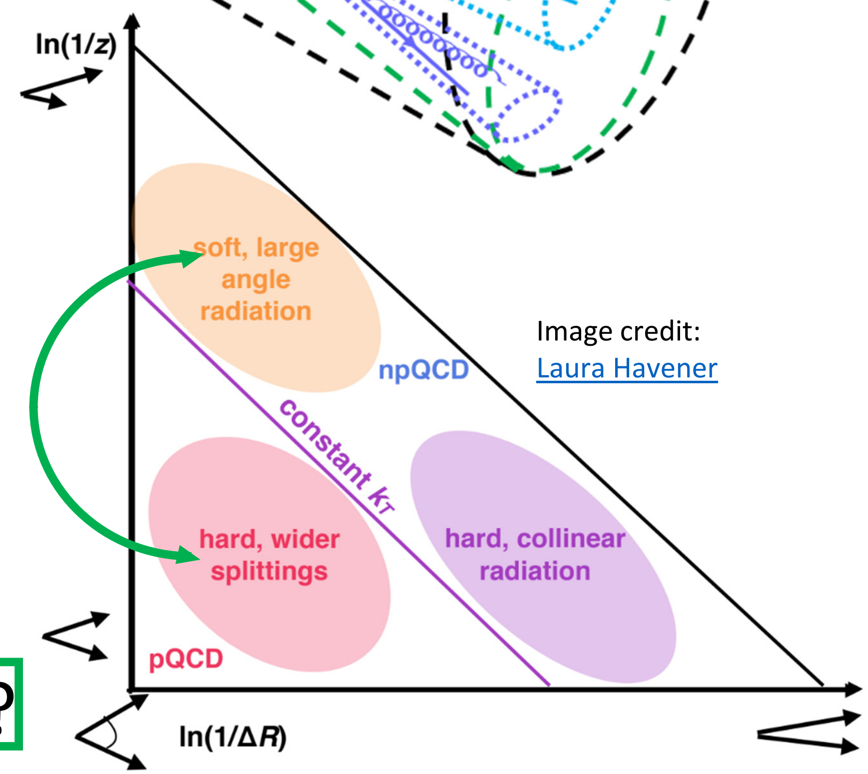


Image credit: Laura Havener

What about the wide-angle soft radiation?

What is the soft-hard correlation within a jet?

Motivation

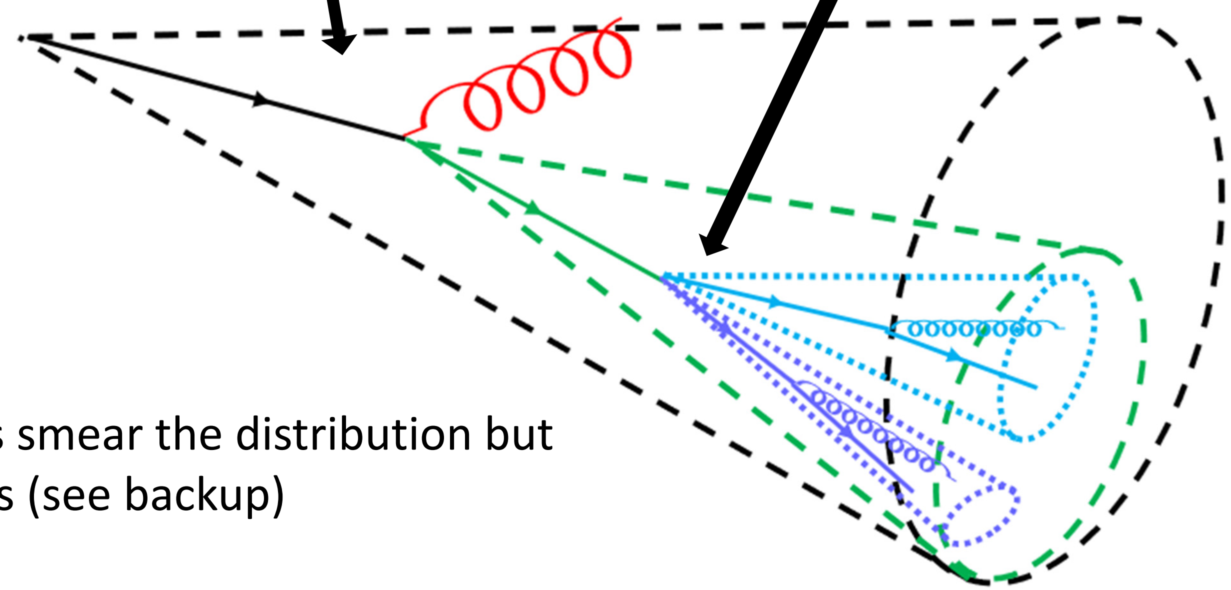
2. soft-hard correlation → collinear drop jet vs groomed jet observables

- How does the amount of soft radiation correlate with the angular and momentum scale of a hard splitting? ~ how an **early** emission affects a **later** splitting

Captured by dM/M:
~ amount of **initial** wide and soft emission

Assuming angular ordering of parton shower

Captured by Zg and Rg:
~ imbalance and angular scale of a **later** splitting



Note: Hadronization effects smear the distribution but don't affect the correlations (see backup)

Jet reconstruction at STAR

Important subdetectors for **200 GeV pp** collisions data-taking during 2012 RHIC run

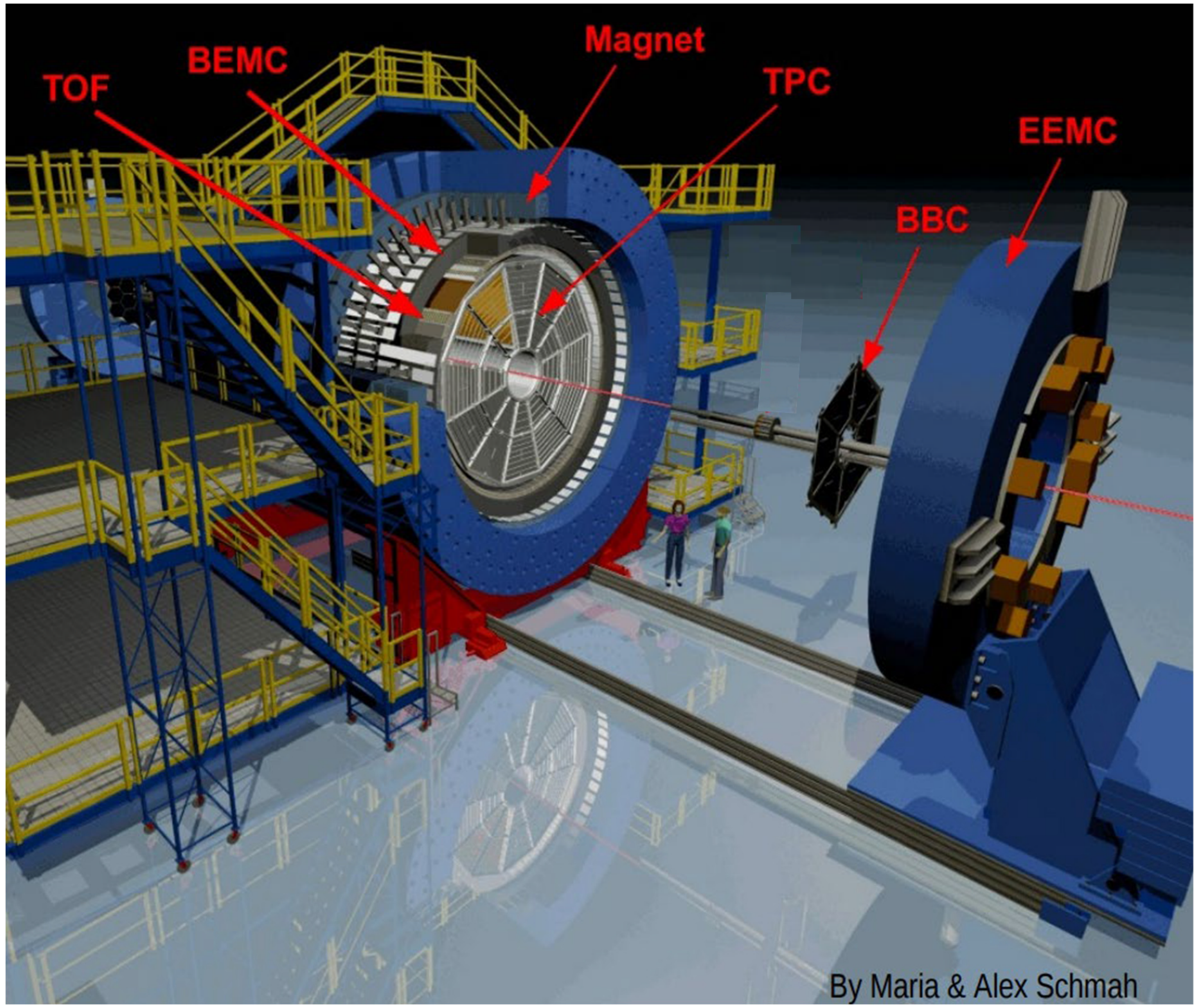
- **TPC** (Time Projection Chamber)
  - For **charged** particle track reconstruction
  - $|\eta| < 1$, full azimuthal coverage

- **BEMC** (Barrel ElectroMagnetic Calorimeter)
  - For **neutral** energy measurement and triggering
  - $|\eta| < 1$, full azimuthal coverage

➢ Reconstruct anti-k_T **full jets**
  - Jet resolution parameter **R=0.4**
  - $|\eta_{jet}| < 0.6$

Additional selections

- Tracks (Towers): $0.2 < p_{\mathrm{T}}(E_{\mathrm{T}}) < 30$ GeV/c
- Jets
  - $p_{\mathrm{T}} > 15$ GeV/c, $M > 1$ GeV/c^2
  - Passes SoftDrop with $z_{cut} = 0.1$ and $\beta = 0$



Unfolding method

- Jet measurements need to be corrected for detector effects for comparison with theory/model
- Unfolding methods:
  - Iterative Bayesian unfolding (D'Agostini. arXiv:1010.0632(2010))
  - **MultiFold** (Andreassen et al. PRL 124, 182001 (2020))
    - Machine learning driven
    - Unbinned
    - **Simultaneously** unfolds multiple observables → **Correlation** information is retained!
- First application of MultiFold on RHIC data!

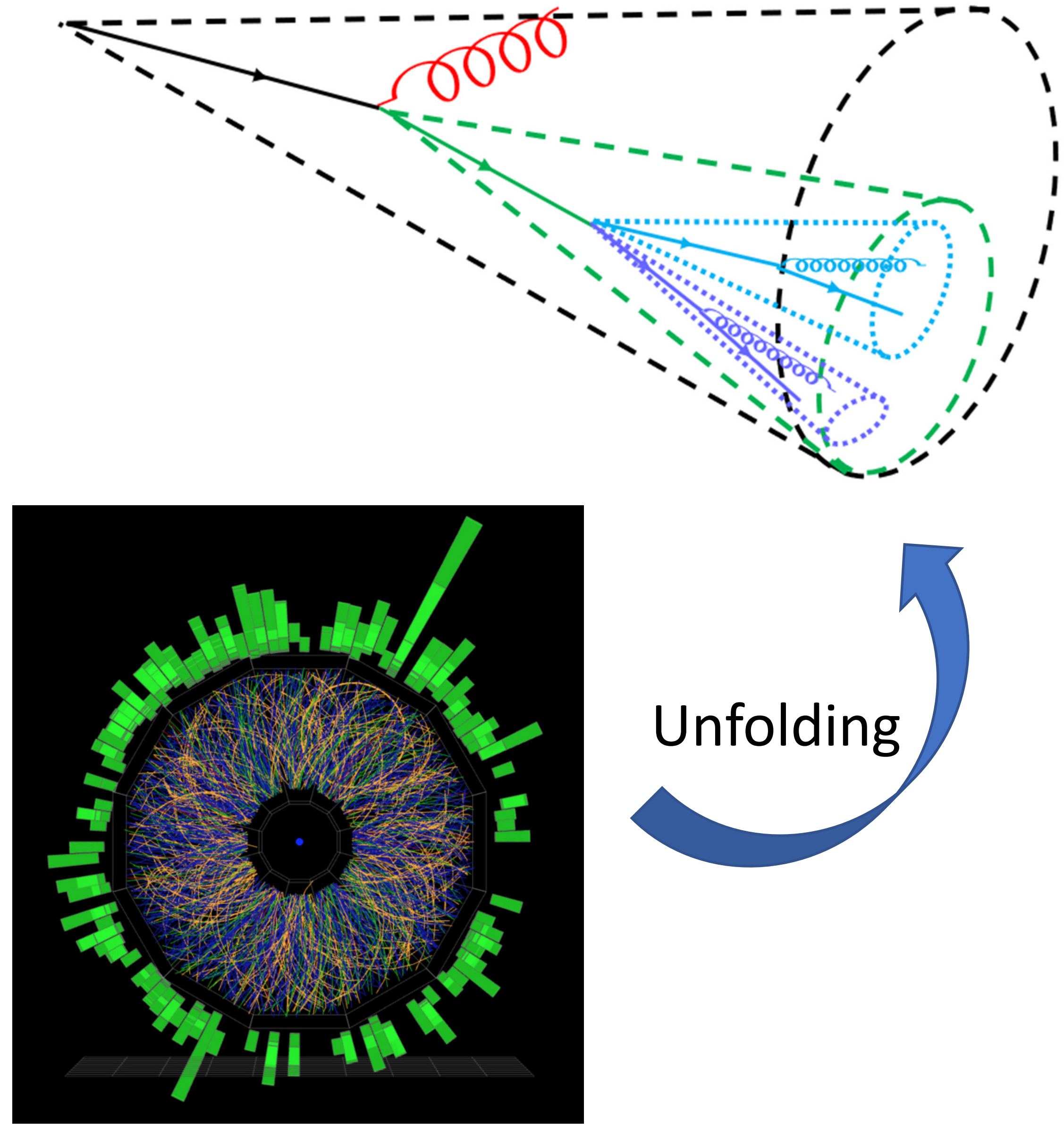
- Jet observables
  - p_{T}: transverse momentum
  - $Q^{\kappa} = \frac{1}{(p_{\mathrm{T}\,\mathrm{jet}})^{\kappa}} \sum_{i \in \mathrm{jet}} q_i \cdot (p_{\mathrm{T}\,i})^{\kappa}$ → Choose K=2
  - $M = |\Sigma_{i \in \mathrm{jet}}\ p_i| = \sqrt{E^2 - |\vec{p}|^2}$

    4-momentum of the constituent i
  - R_g: groomed jet radius
  - z_g: shared momentum fraction
  - M_g: groomed jet mass

All 6 observables are simultaneously unfolded in an unbinned way!

- Uncertainties due to prior choice accounted for through 6D reweighting based on PYTHIA8 or HERWIG (see backup)

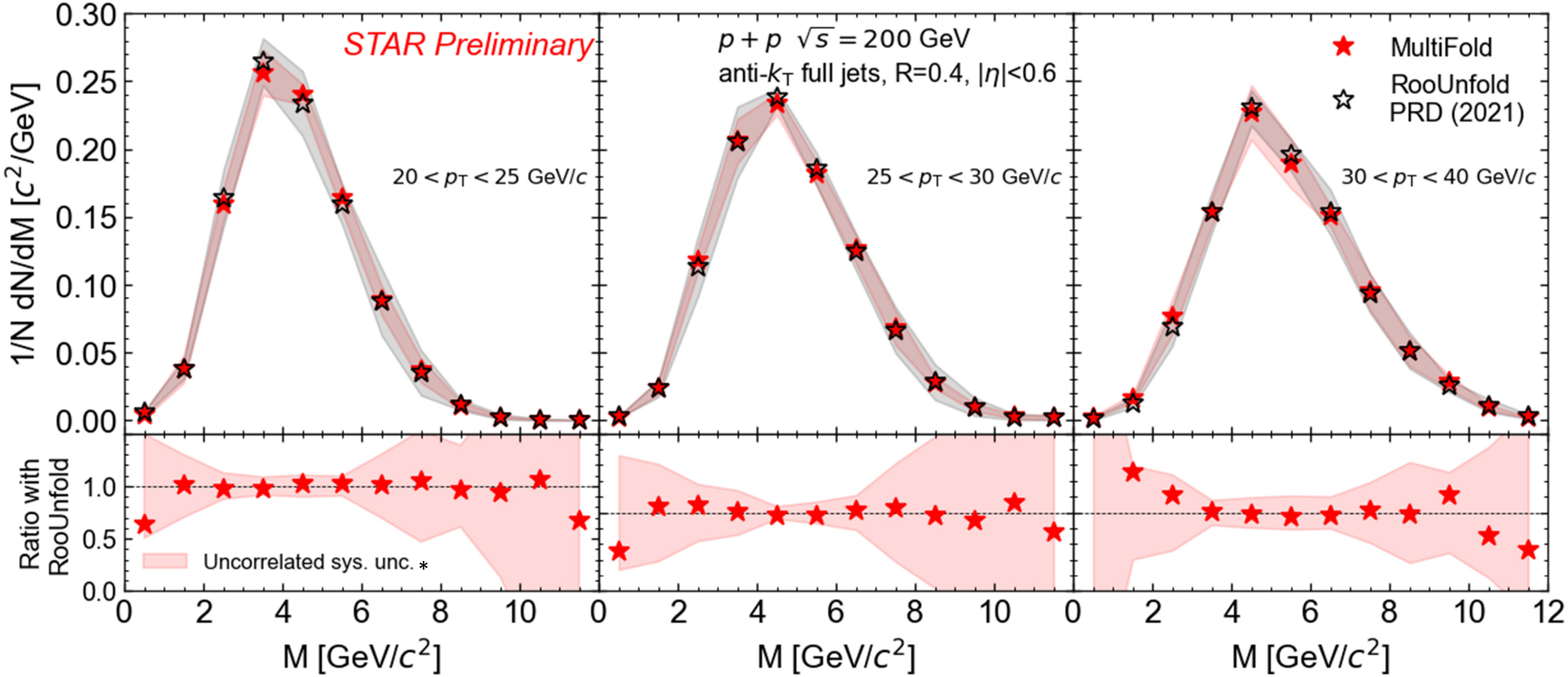
Does our method work?



Fully corrected jet *M*

$$M = |\Sigma_{i\in \text{jet}}\ p_i| = \sqrt{E^2 - |\vec{p}|^2}$$

MultiFold result agrees with **RooUnfold** result (STAR Collaboration. PRD 104, 052007(2021)) HEPData



… but **MultiFold** also gives us correlation between observables!

* 2D reweighting used for prior variation, to be consistent with RooUnfolded measurement

Looking at wide-angle soft radiation

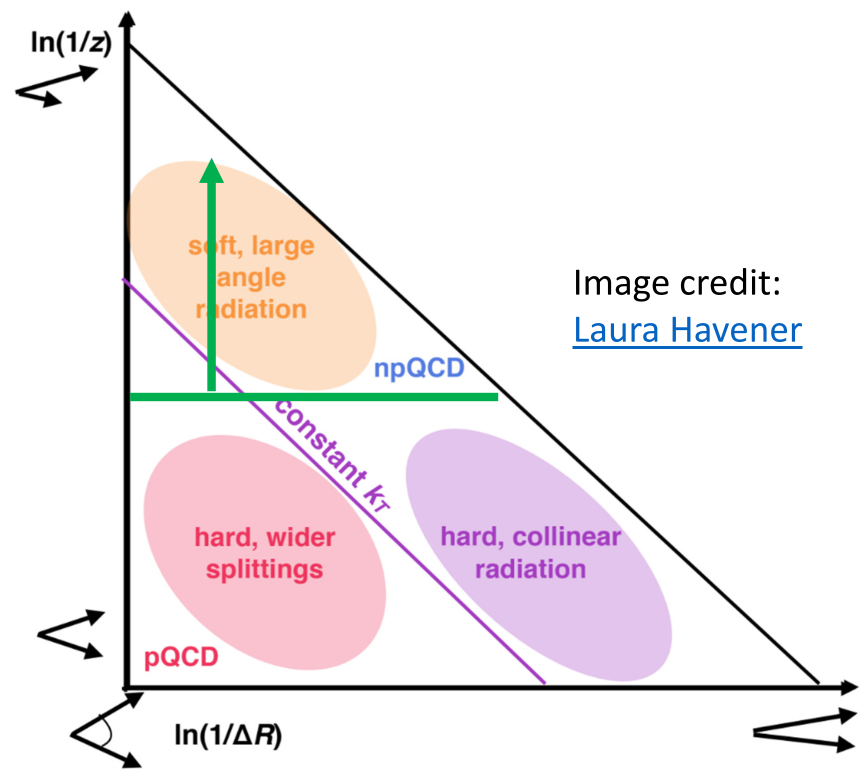
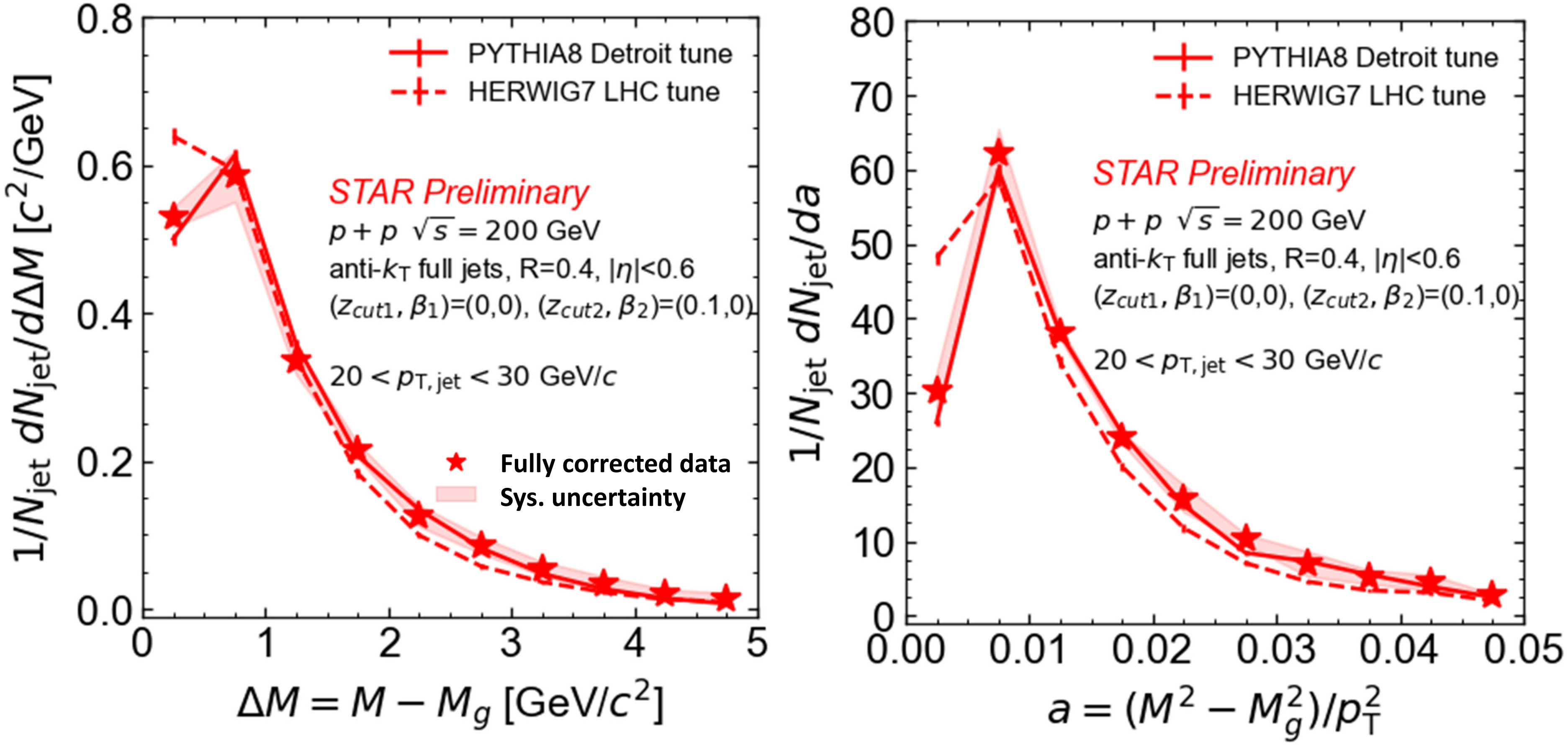


Image credit: Laura Havener

Collinear drop groomed jet mass

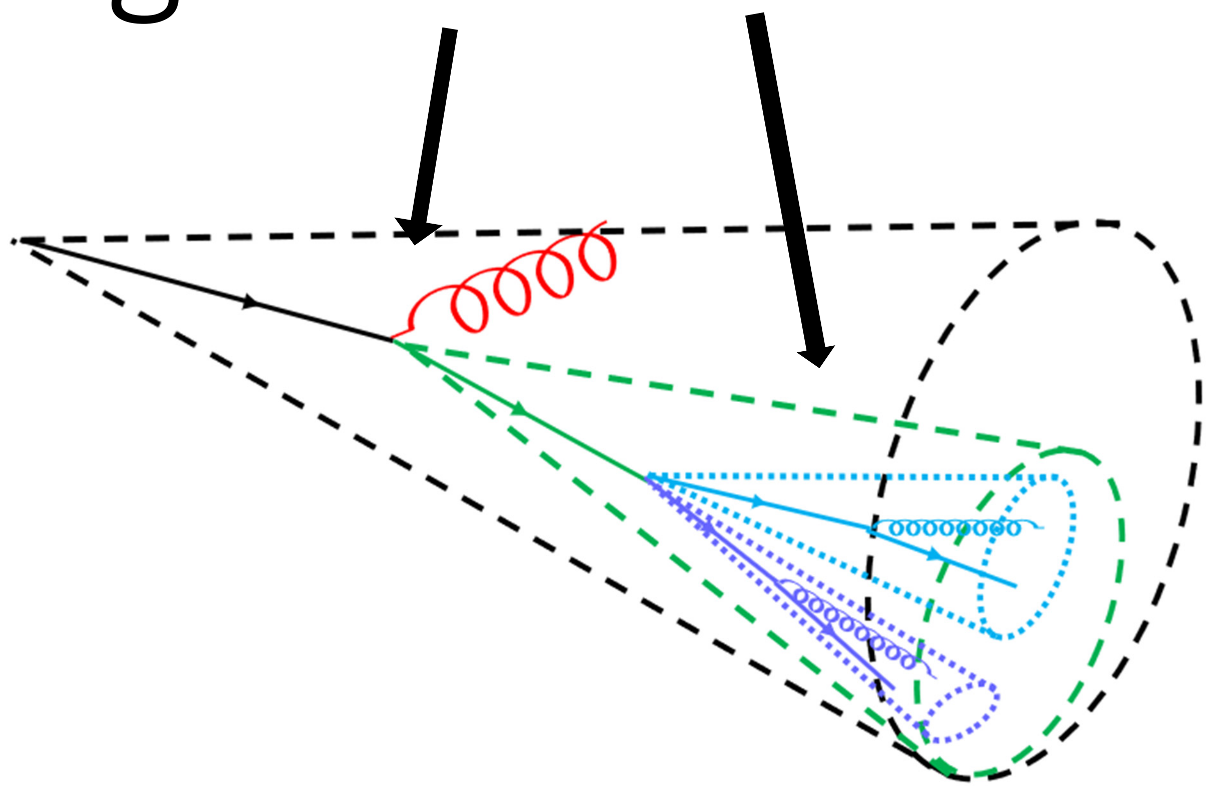


PYTHIA6 Perugia 2012 STAR tune: Skands, PRD 82, 074018 (2010)
PYTHIA8 Detroit tune: Aguilar et al. PRD 105, 016011(2022)
HERWIG7: Bellm, et al. PRC 76, 1-8 (2016)

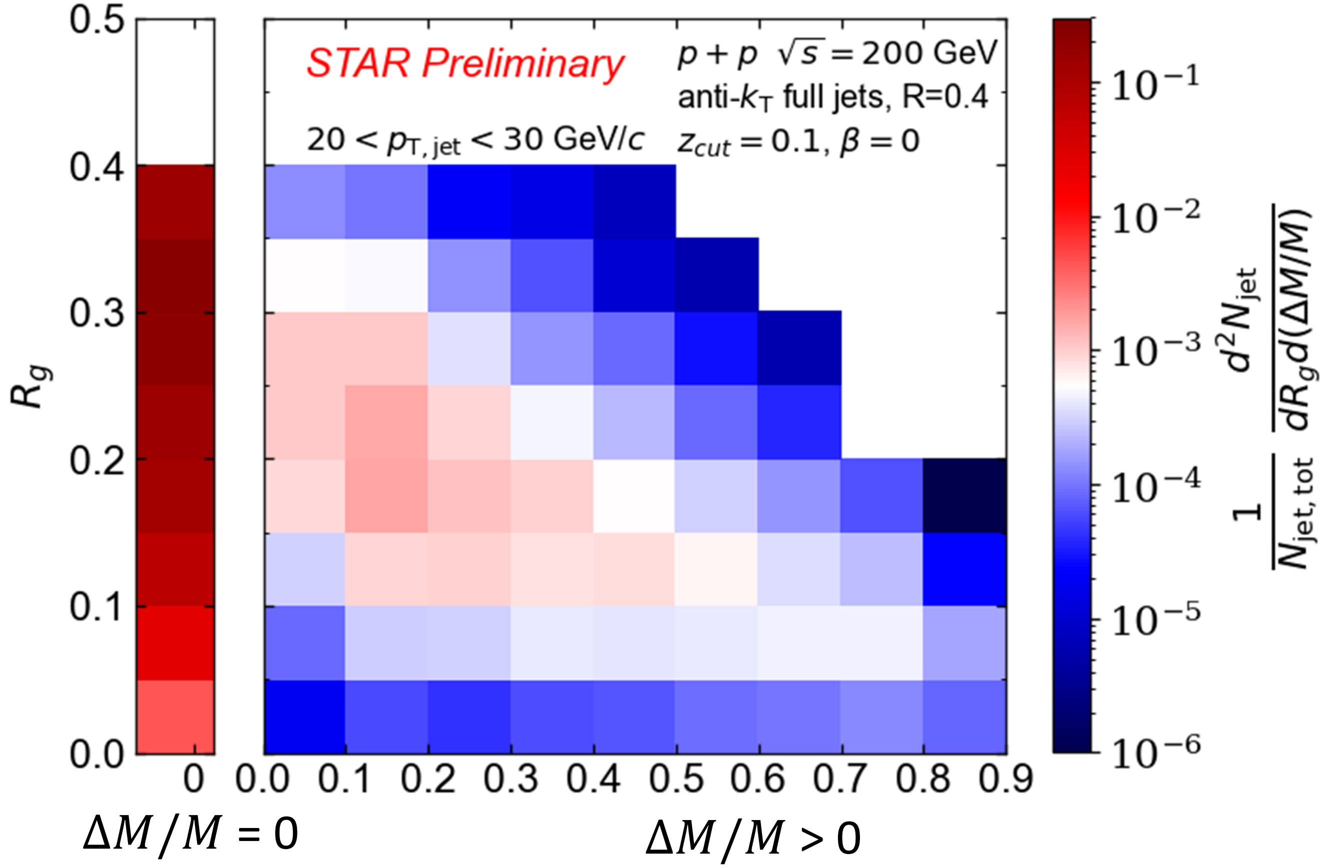
Measurement excludes jets with ΔM=0 (45.5% of jets in this jet p_T range)

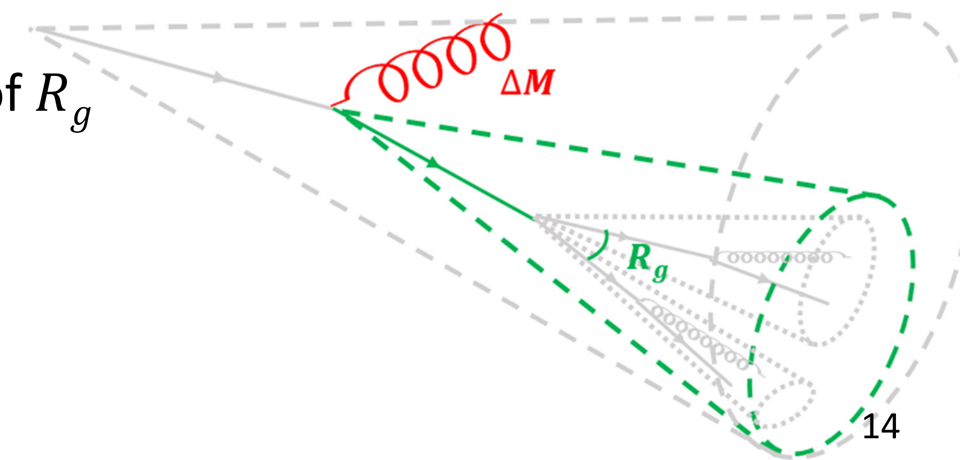
- First collinear drop groomed jet measurement, sensitive to soft radiation within jets
- MC predictions qualitatively consistent with data; some tension from HERWIG in small ΔM region
- MultiFold allows us to correlate (combinations of) unfolded quantities

Probing soft-hard correlation



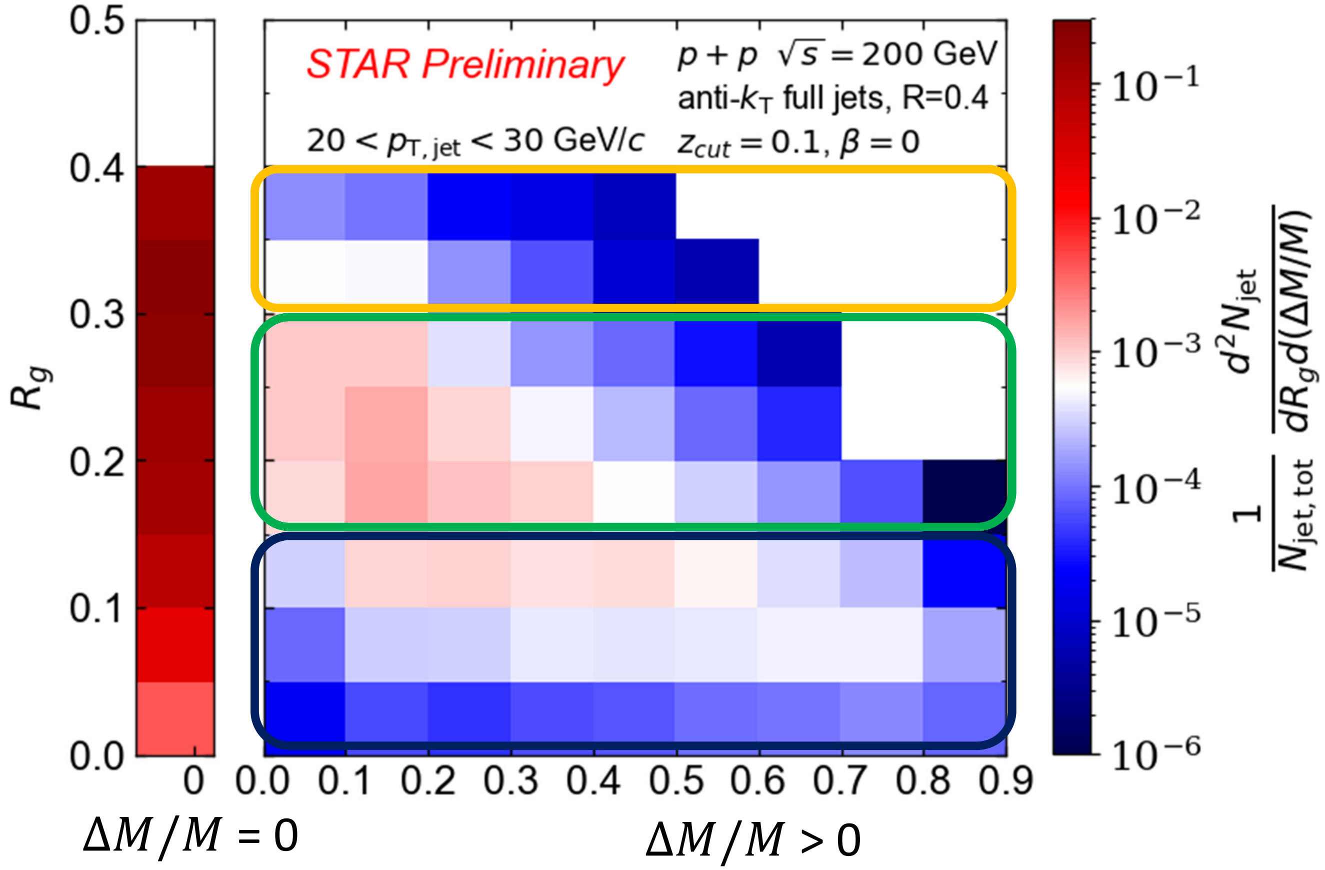
Soft radiation vs hard splitting angle

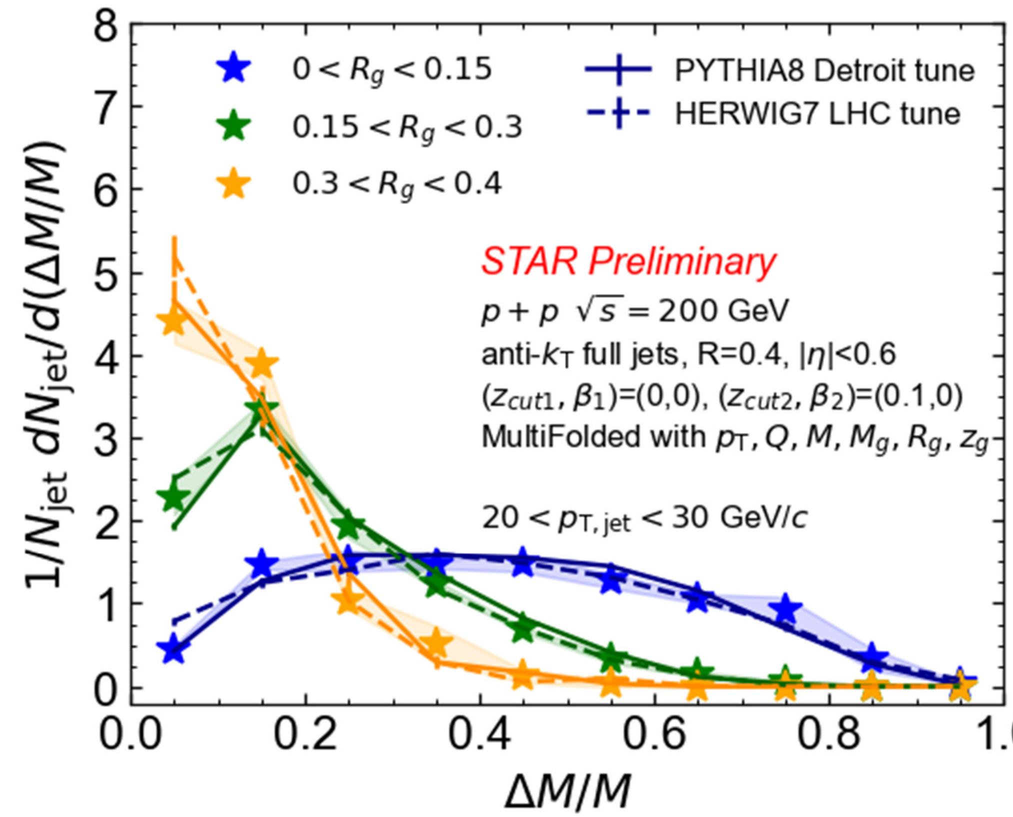




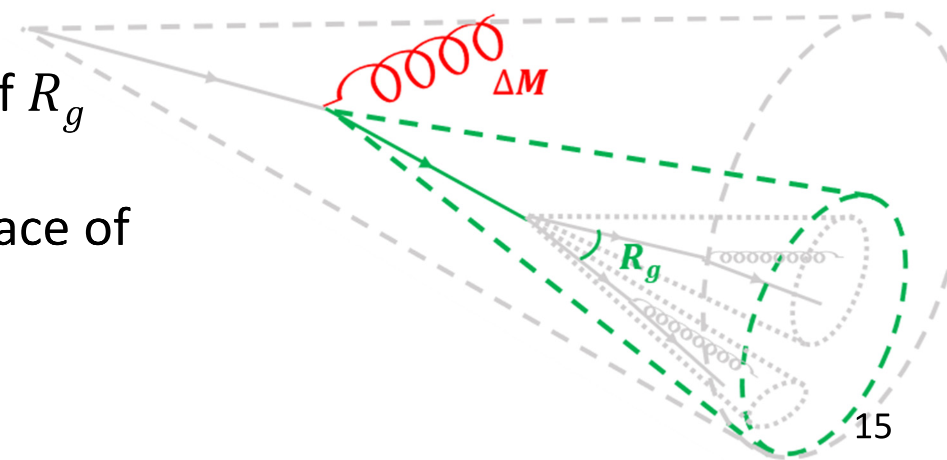
- The mean of $\Delta M/M$ distribution is <u>anti-correlated</u> with mean of R_g
  → consistent with angular ordered parton showers

Soft radiation vs hard splitting angle





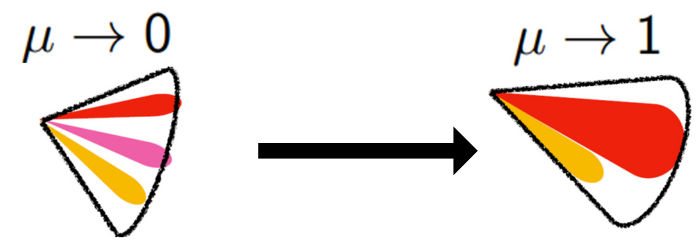
- The mean of $\Delta M/M$ distribution is <u>anti-correlated</u> with mean of R_g
  → consistent with angular ordered parton showers
- **Early** soft wide-angle radiation constrains the angular phase space of **later** splittings
- MC models describe the trend of data

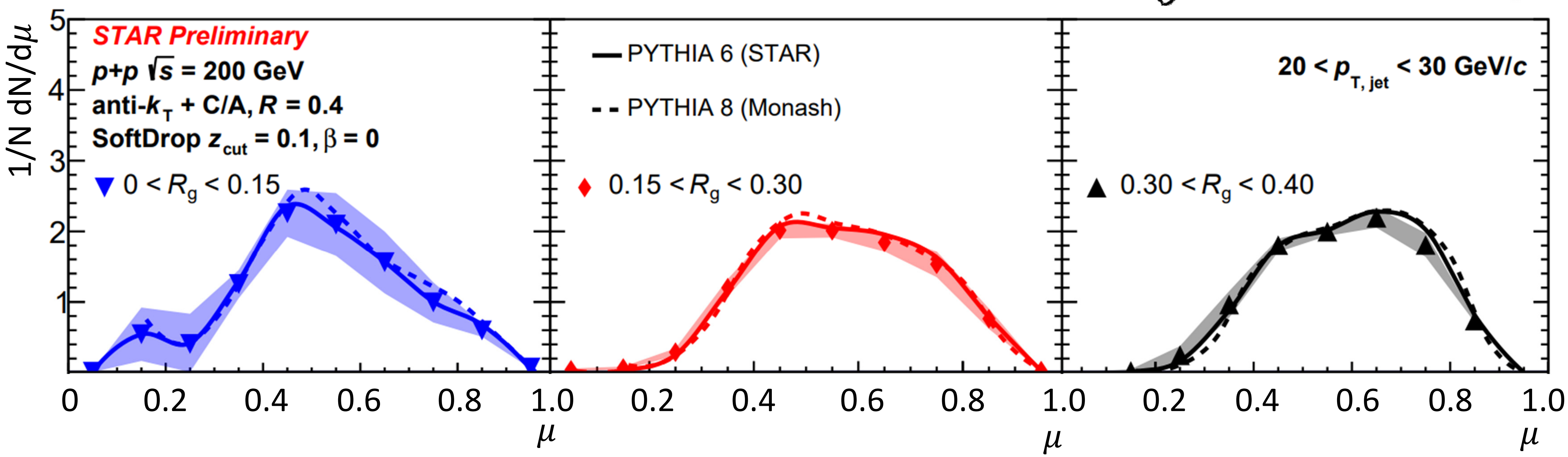


Hard radiation vs hard splitting angle

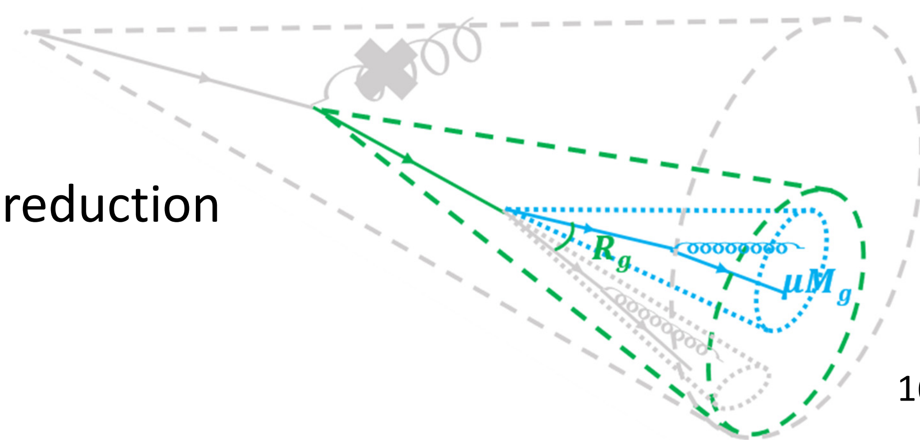
μ	groomed mass fraction	$\mu \equiv \frac{\max(m_{j,1}, m_{j,2})}{m_g}$

Dasgupta M. et al, JHEP (2013) 2013:29



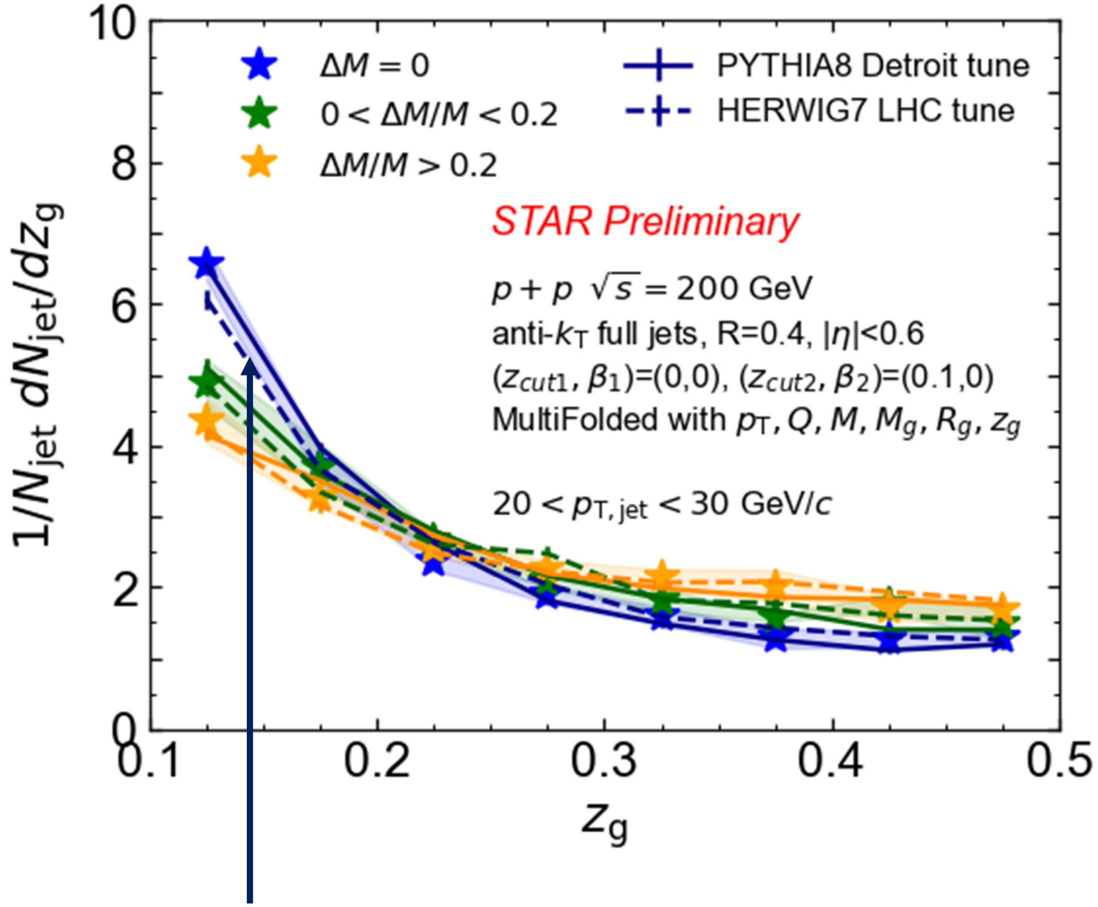


- μ has a weaker dependence on R_g compared to $\Delta M/M$. MC models describe the trend of data.
- Shift of μ to smaller values at smaller R_g indicates a faster reduction of virtuality in the jet shower.



Soft radiation vs hard splitting momentum imbalance

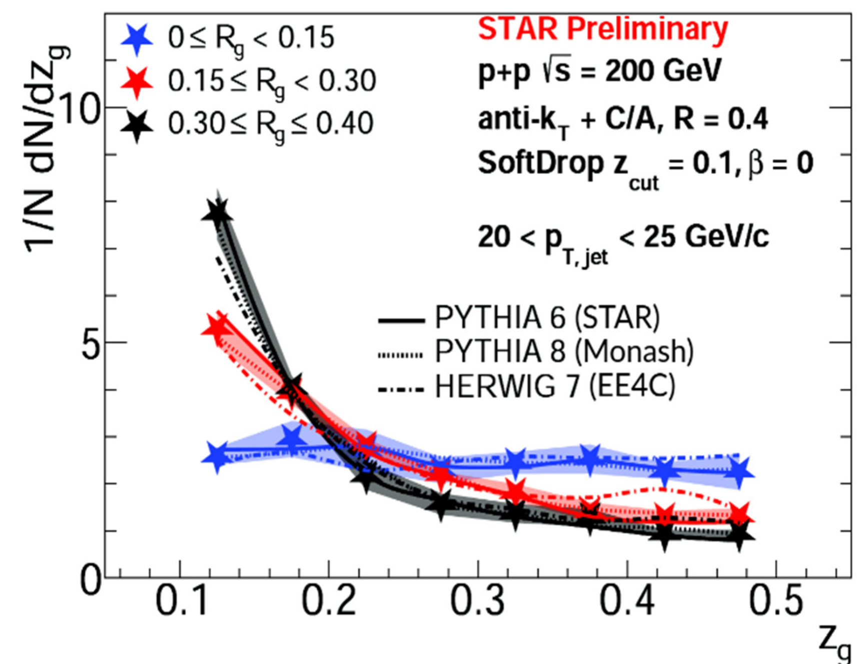
STAR



Steeply falling ~ DGLAP 1/z: pQCD

R_g and z_g are correlated, $\Delta M/M$ affects R_g
→ correlation between $\Delta M/M$ and z_g

- The more mass that is groomed away relative to the original mass, the flatter the z_g distribution is
  - Demonstrates that **early** soft wide angle radiation constrains the momentum imbalance of & the amount of npQCD contributions to **later** splittings
- MC models describe the trend of data

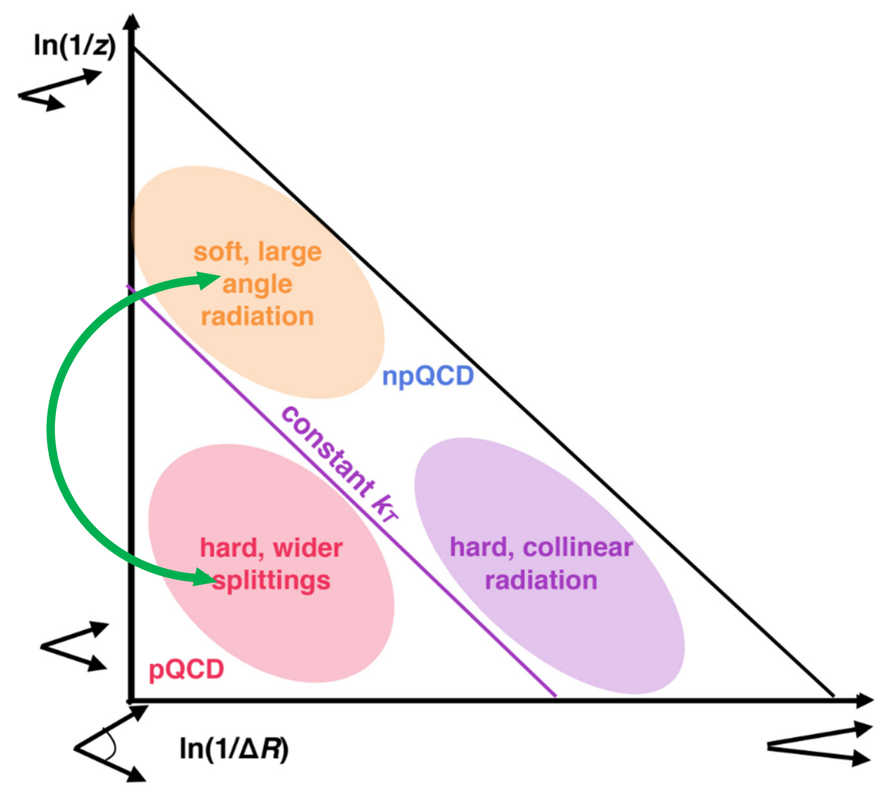


Conclusions

- Probing **soft** wide-angle radiation within jets
  - First fully corrected collinear drop jet measurement is presented

Improve understanding of jet substructure and the correlations between different substructure observables!

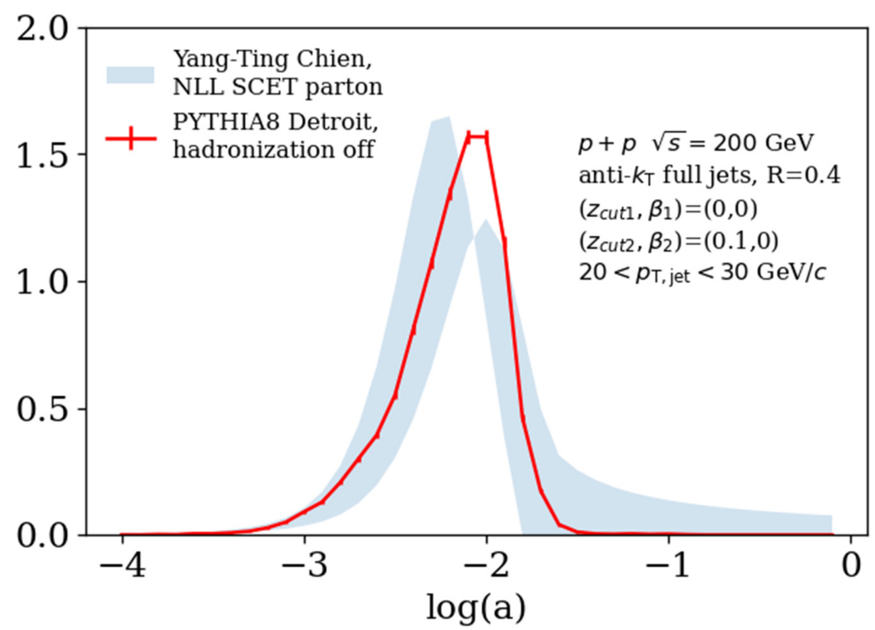
- Probing **soft-hard** correlation within jets
  - MultiFold allows for access of multi-dimensional correlations on a jet-by-jet basis. First application to RHIC data!
  - Jets with a large perturbative contribution (DGLAP splitting) are more likely to have small early-stage radiation
  - Anti-correlation between the amount of early-stage radiation and the angular scale of a later-stage splitting is observed



Backup

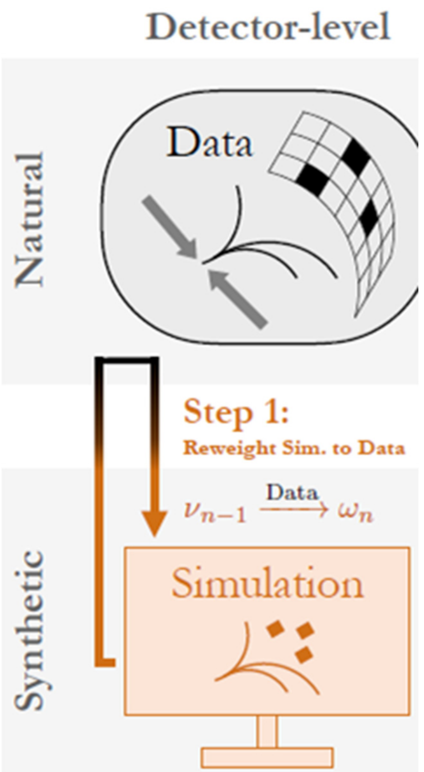
Collinear drop groomed jet mass

- Theoretical calculation (next-to-leading log precision, using SCET calculational framework, and not including hadronization) agrees with PYTHIA8

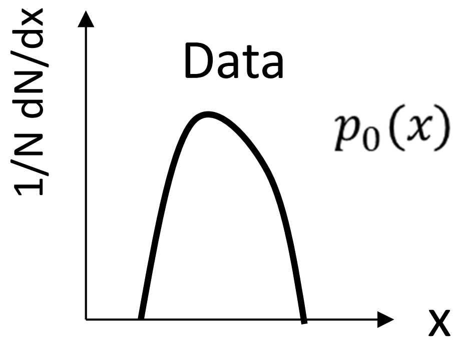


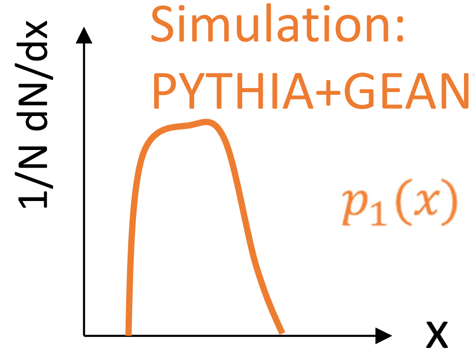
$$a = (M^2 - M_g^2)/p_T^2$$

Method: machine learning



Detector-level





Where does the machine learning part come in?

E.g., Iteration 1, step 1:

Weights: $w(x) = p_0(x)/p_1(x)$ Ok for 1D

$$\approx f(x)/(1 - f(x))$$

(Andreassen and Nachman PRD 101, 091901 (2020))

where f(x) is a neural network and trained with the binary cross-entropy loss function

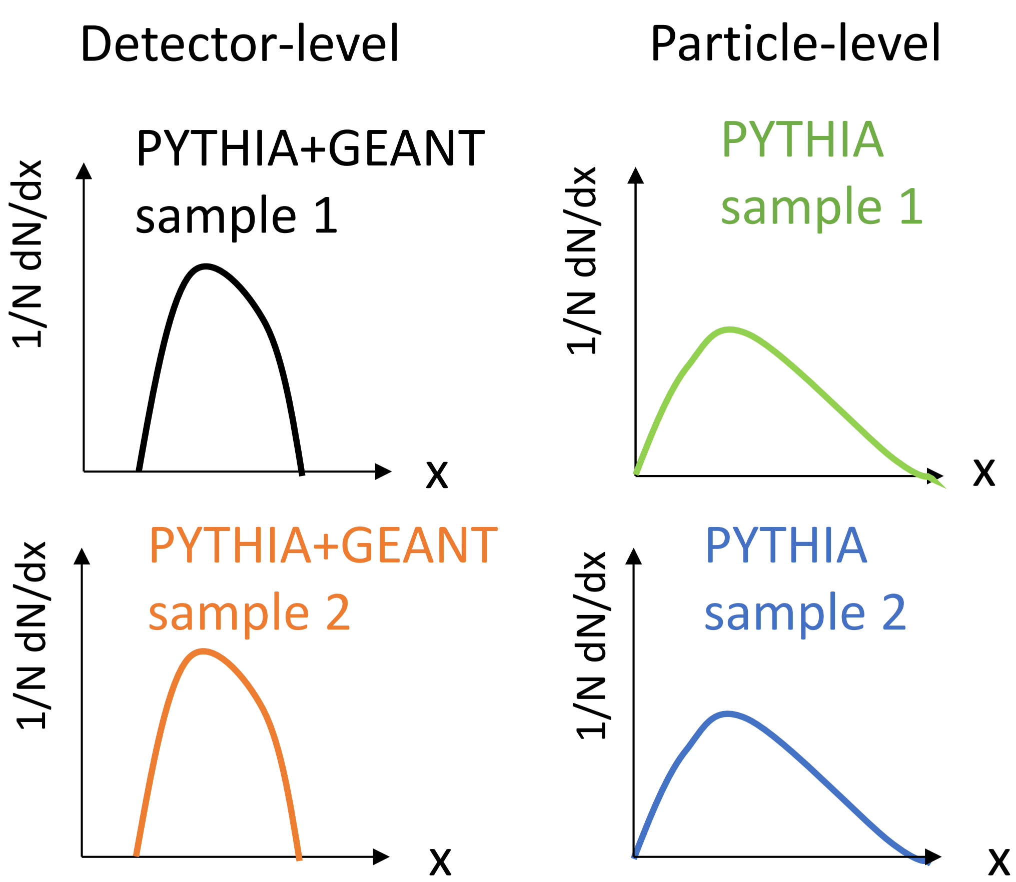
to distinguish jets coming from data vs from simulation

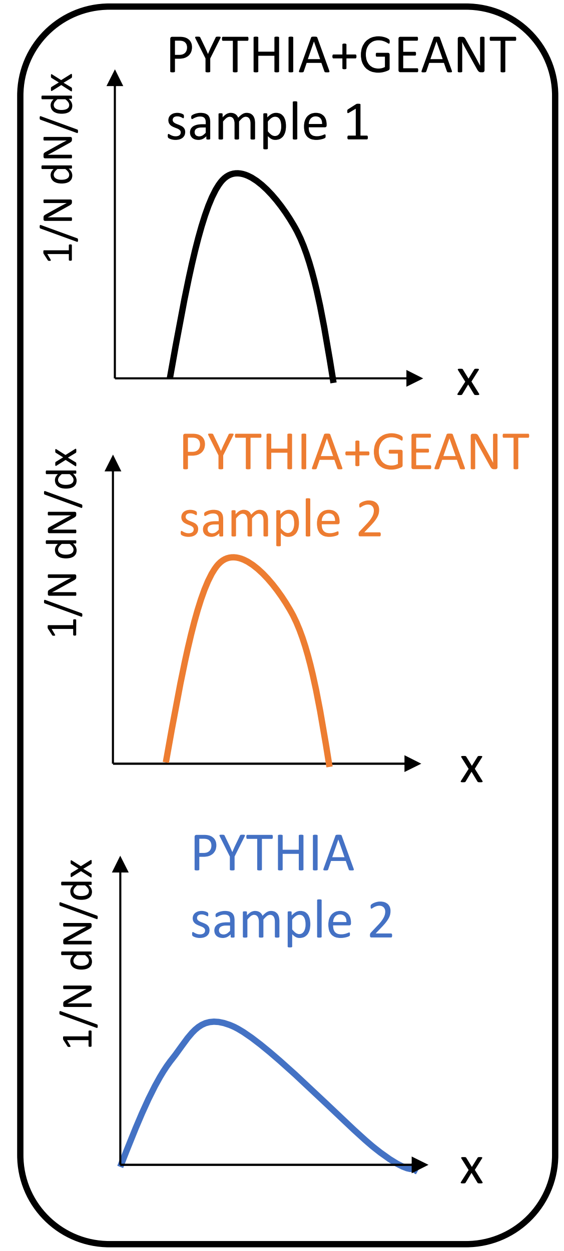
Unfolding → Reweighting histograms → Classification → Neural network

See backup slides for details of the neural networks.

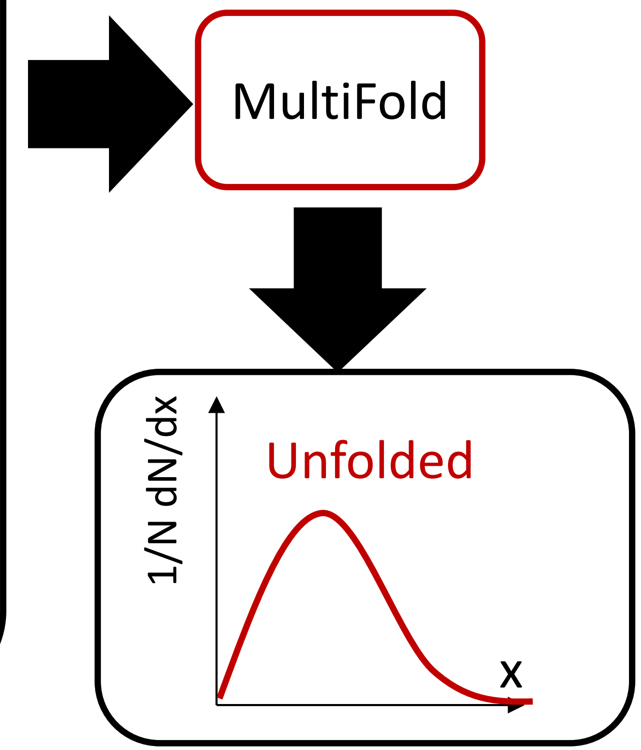
Closure test for unfolding

- Step 1: Separate matched jets from PYTHIA and PYTHIA+GEANT into 2 samples



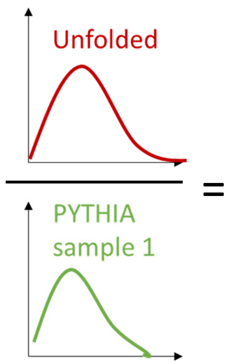


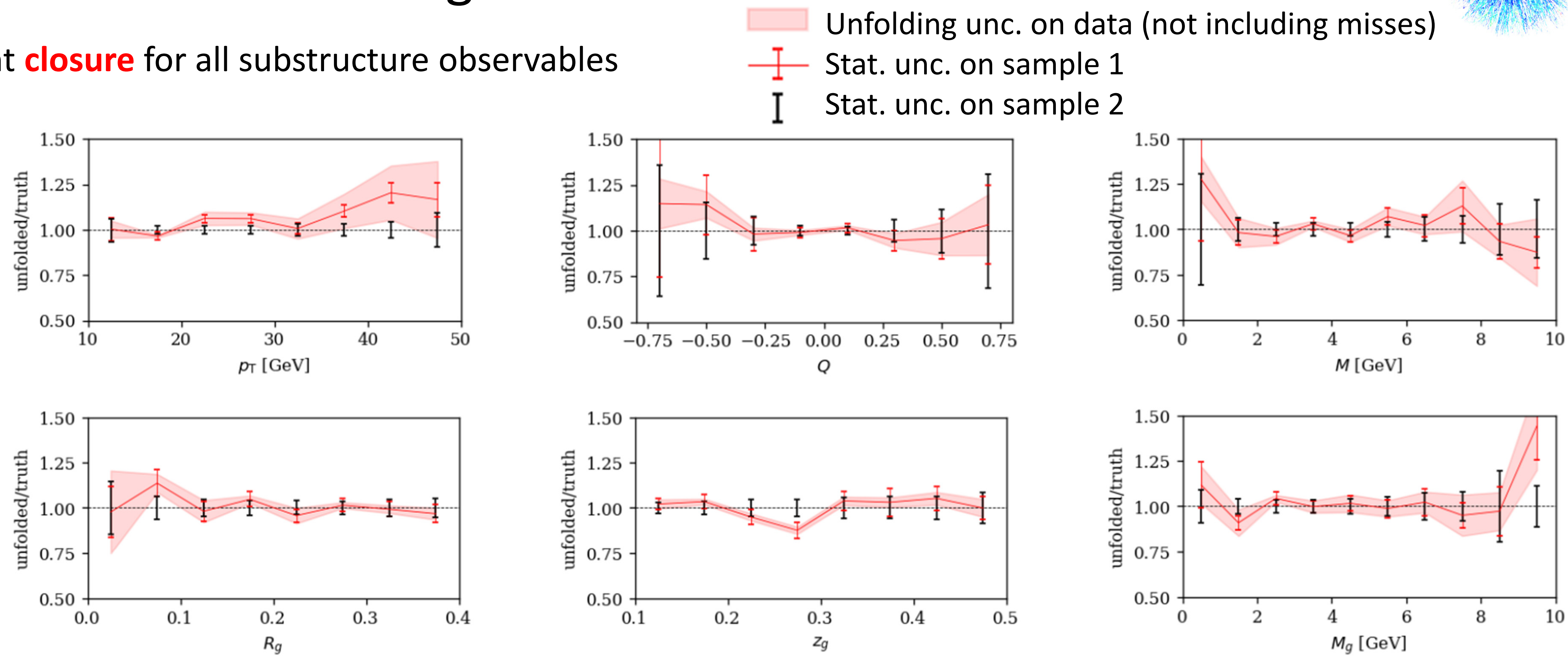
- Step 2: Unfold



Closure test for unfolding

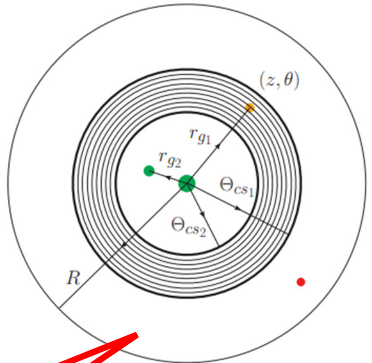
- Decent **closure** for all substructure observables

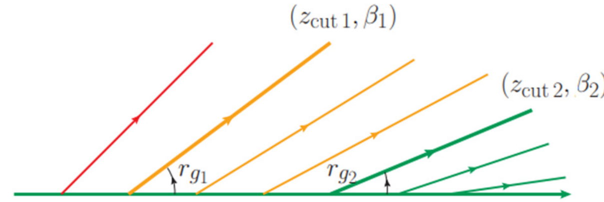




* 2D reweighting used for prior variation

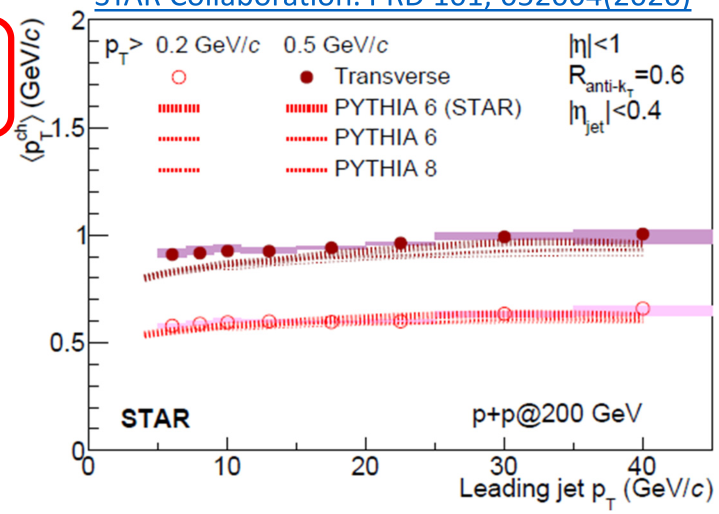
Underlying event

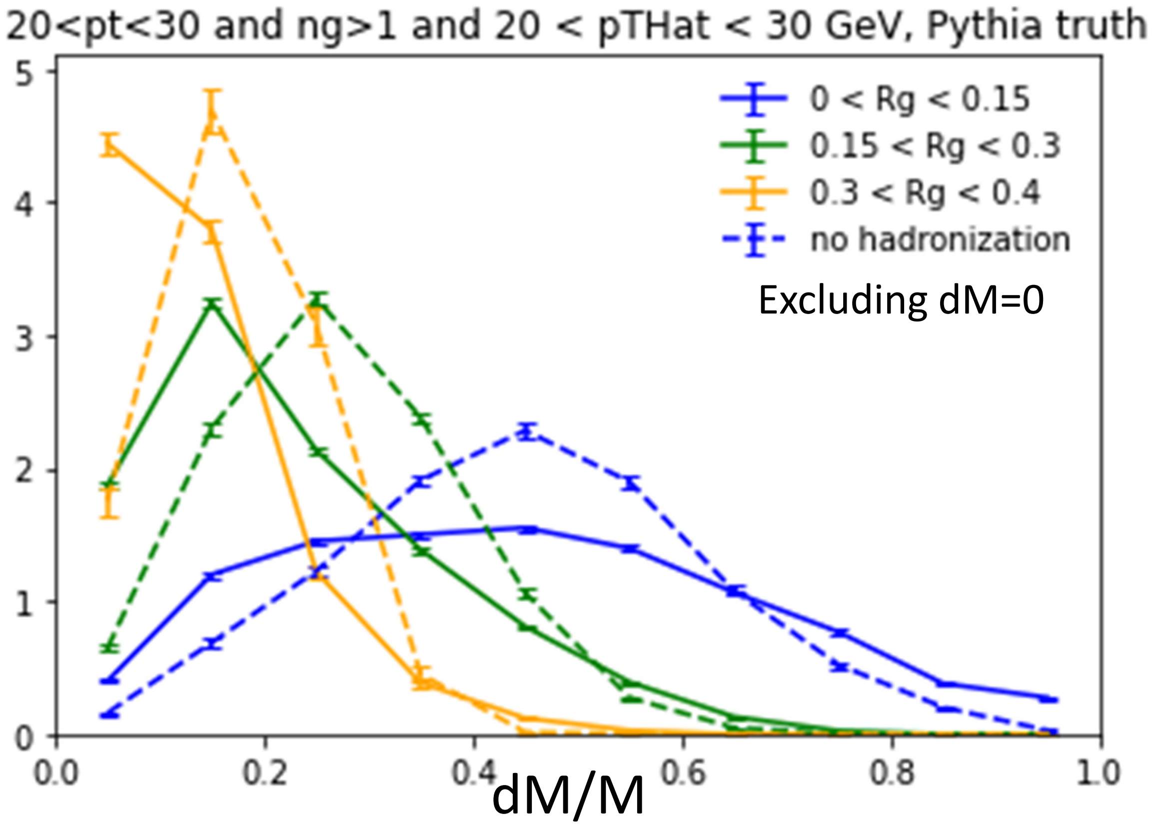


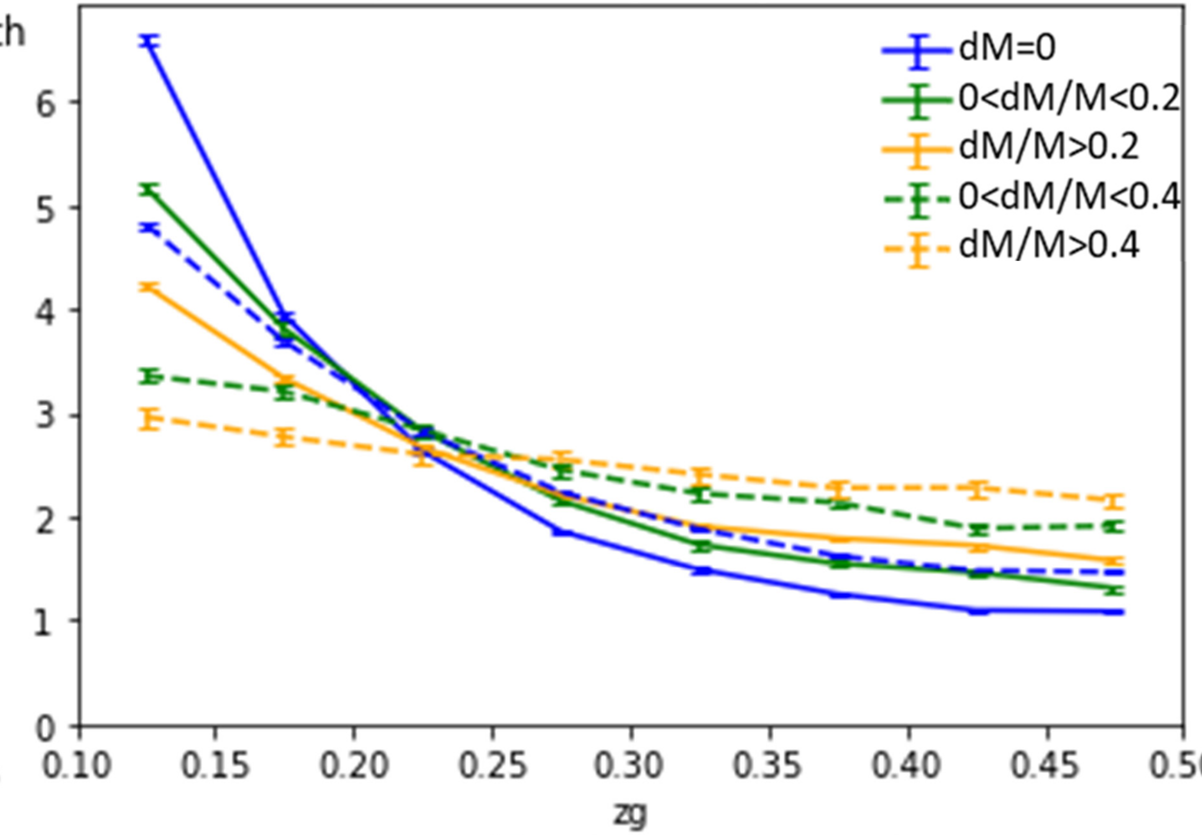


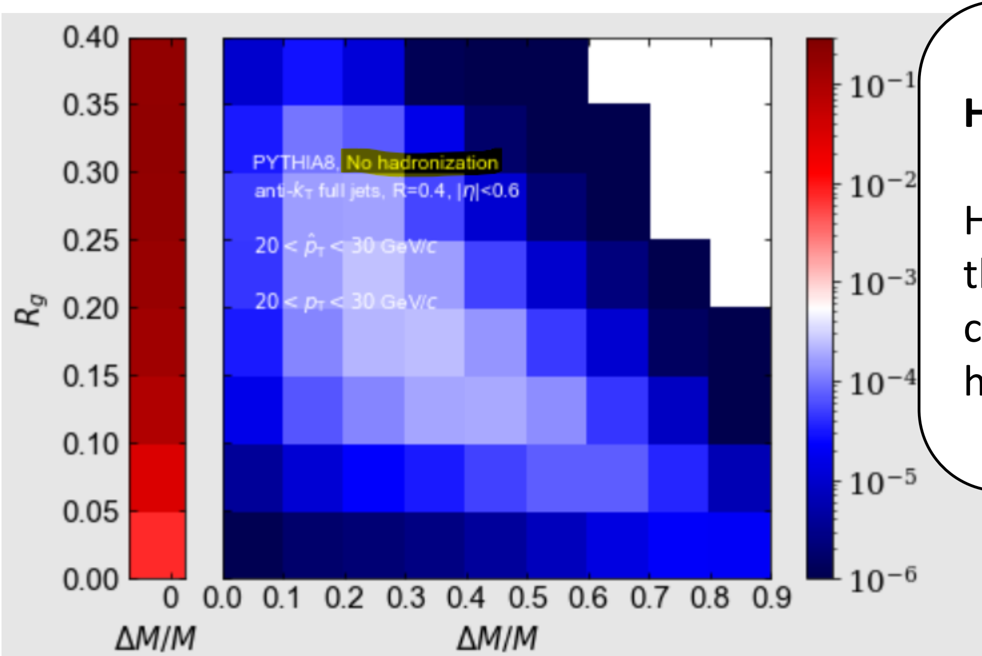
Remove underlying effects and pileup, not necessary for 200 GeV collisions

[STAR Collaboration. PRD 101, 052004(2020)]



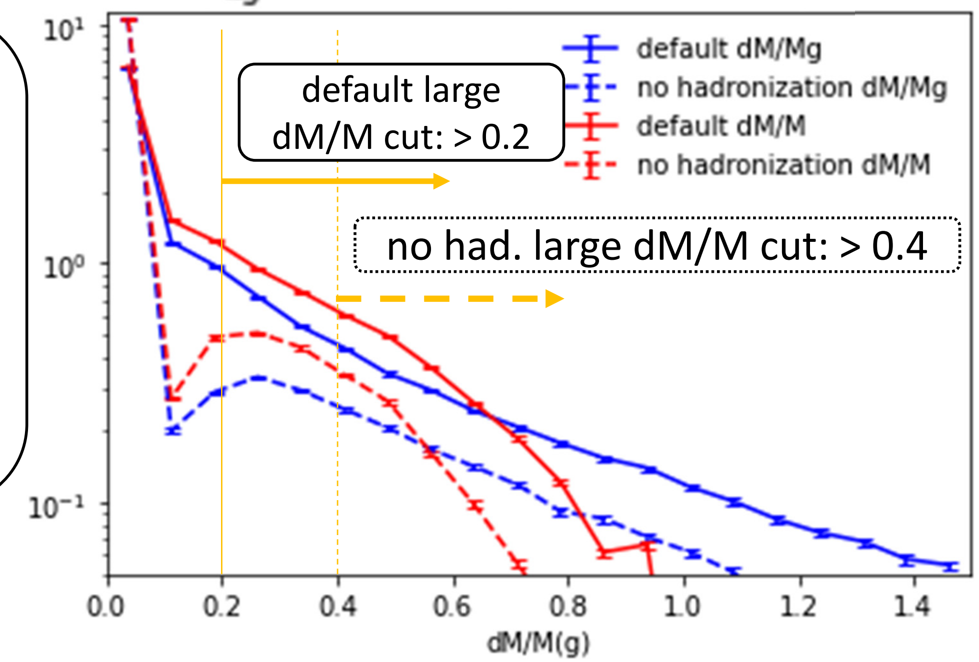






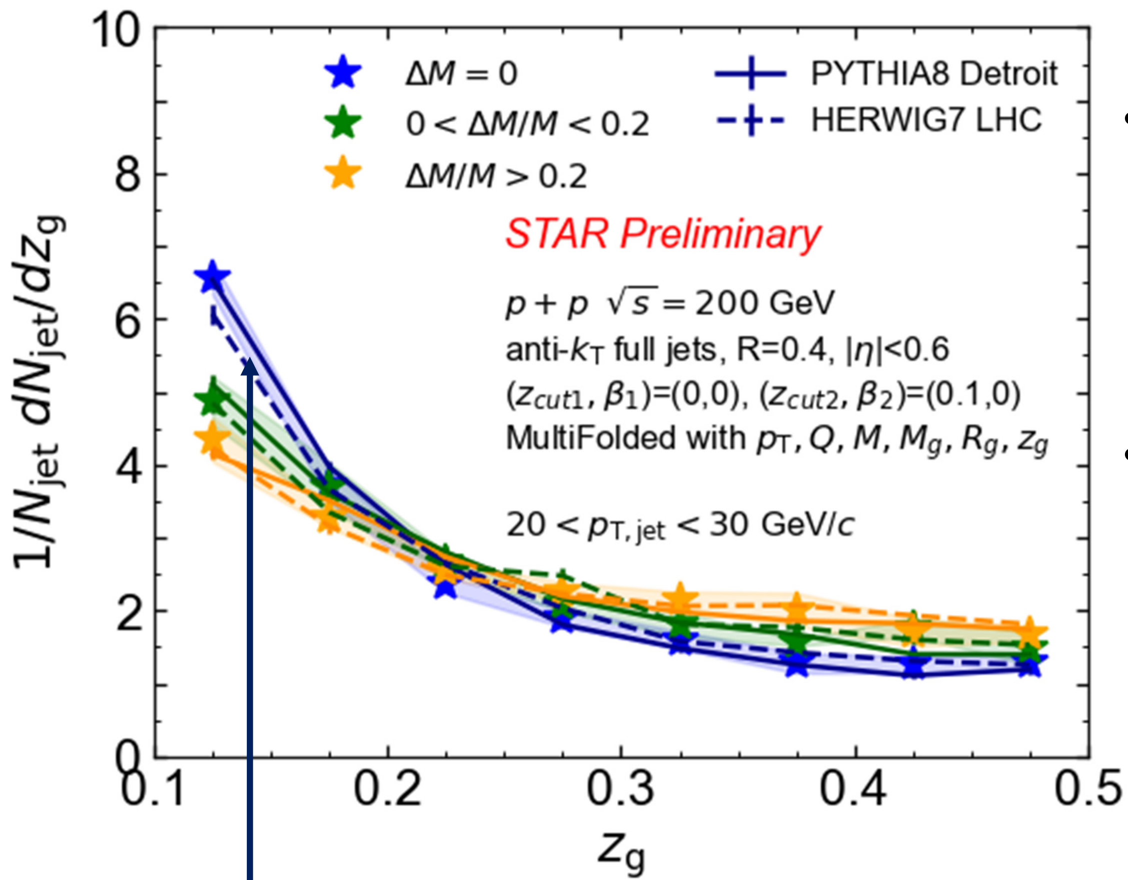
Hadronization effects study:

Hadronization smears/shifts the distributions, but the correlation with and without hadronization is the same.



- Among truth jets with $20 < pT < 30$ GeV, 44% of PYTHIA6 jets, 48% of PYTHIA8 jets, and 43% of HERWIG jets have dM=0.
- Among reco jets with $20 < pT < 30$ GeV, 36% of PYTHIA6 embedding jets and 37% of data jets have dM=0. (Higher pT jets have a smaller fraction of dM=0).

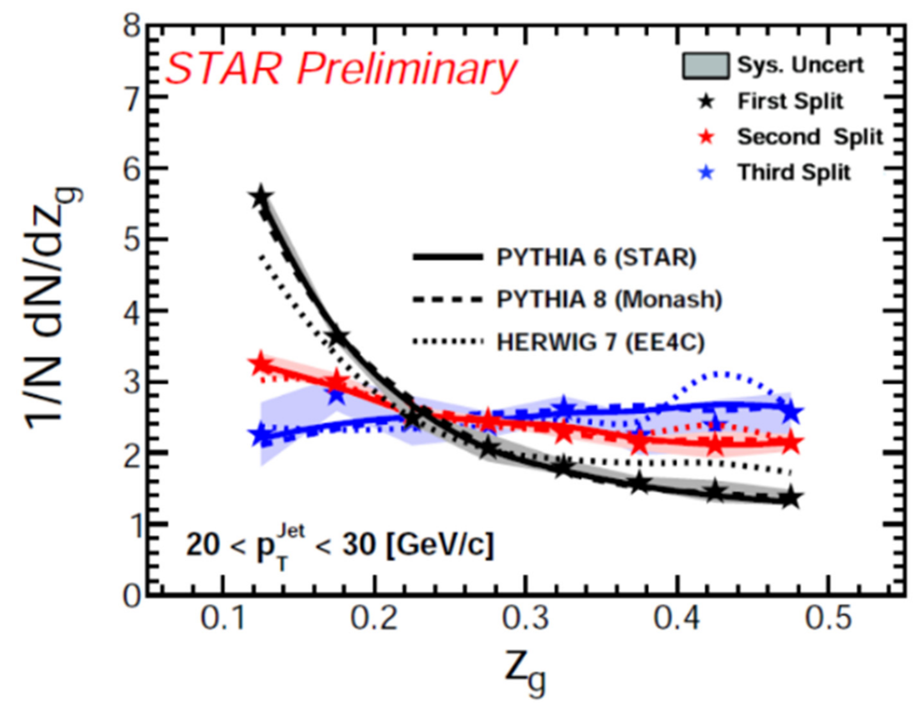
Soft radiation vs **hard** splitting momentum imbalance



Steeply falling ~ DGLAP 1/z: pQCD

- The more mass that is groomed away relative to the original mass, the flatter the Zg distribution is
  - Demonstrates that **early** soft wide angle radiation constrains the momentum imbalance of & the amount of npQCD contributions to **later** splittings
- MC models describe the trend of data

→ The first splitting that passes SoftDrop can still be non-perturbative, but if we apply the $\Delta M = 0$ selection, we can filter out some npQCD contribution due to the parton splitting



Systematic uncertainties

- Detector systematics
  - Hadronic correction 100% → 50%
  - Tower scale +3.8%
  - Tracking uncertainty -4%
- Unfolding systematics
  - Unfolding seed
  - Iteration number variation
  - Prior shape variation to HERWIG7 and PYTHIA8
    - Nominal: prior = (generation, simulation)
      = (PYTHIA6, PYTHIA6 + GEANT3 + embedding)
    - Varied to: $\text{prior} \rightarrow \text{reweight} \bigotimes \text{nominal prior}$,
      with $\text{reweight}(p_T, Q, M, M_g, R_g, z_g) = \frac{\text{Herwig truth}(p_T, Q, M, M_g, R_g, z_g)}{\text{Pythia6 truth}(p_T, Q, M, M_g, R_g, z_g)}$