Going off topics to demix quarks and gluons in α_S extractions

arXiv:2206.10642 [hep-ph]

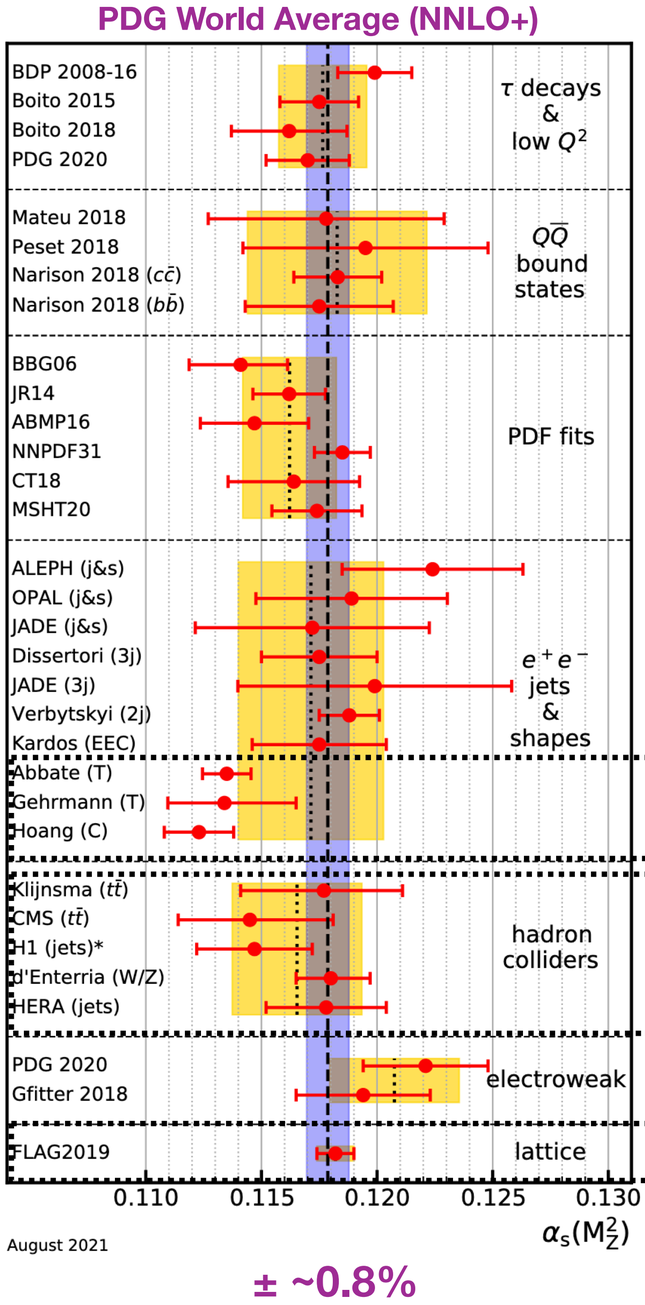
M. LeBlanc (CERN), B. Nachman (LBNL), C. Sauer (Heidelberg)

Strong coupling

The QCD coupling, α_S, is known with far less precision than the couplings of other gauge field theories.

$\delta\alpha \sim 10^{-10} \ll \delta G_F \sim 10^{-7} \ll \delta G \sim 10^{-5} \ll \delta\alpha_S \sim 10^{-3}$

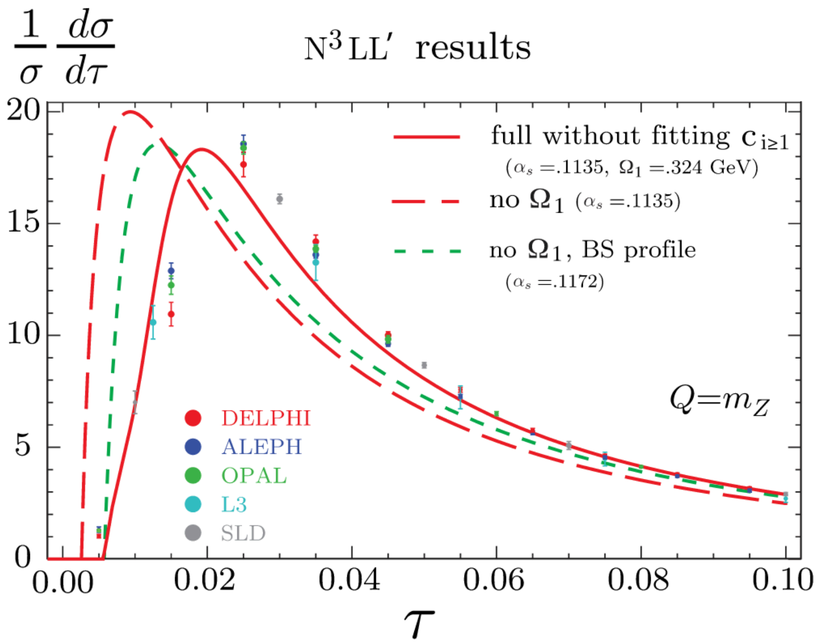
This uncertainty is driven by tensions within the α_S world average, and it is becoming increasingly relevant in predictions related to Higgs & top production, EWPOs, etc.



Most precise collider extractions (LEP) in tension with other results…

No LHC extraction from jet cross sections at NNLO+!

Lattice average very precise



$\alpha_S(M_Z) = 0.1135 \pm 0.0010$ (e^+e^- thrust)

vs.

$\alpha_S(M_Z) = 0.1182 \pm 0.0008$ (lattice)

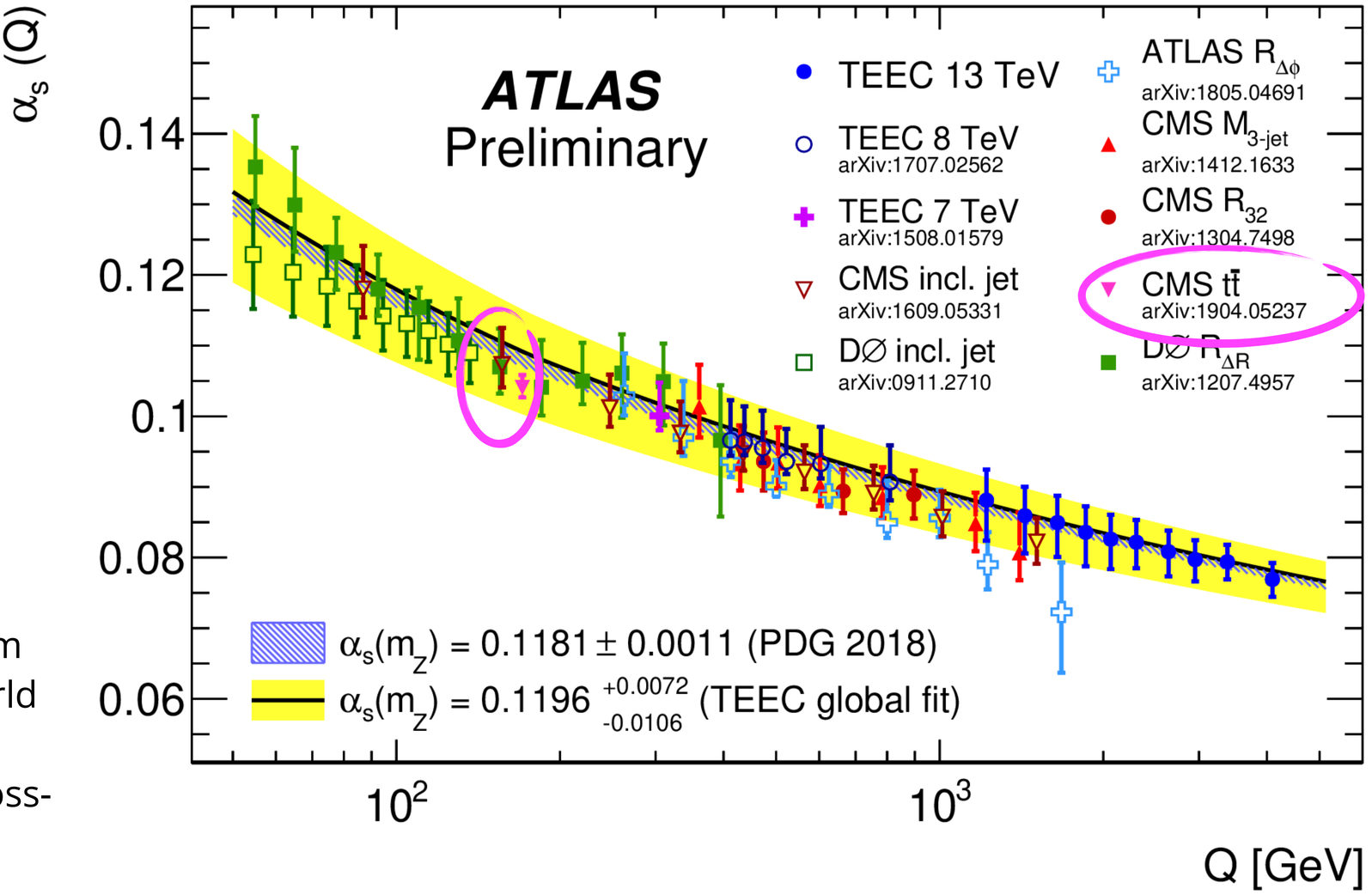
Tension > 3σ

Abbate et al. Phys.Rev.D 83 (2011) 074021

P.A. Zyla et al. (Particle Data Group), Prog. Theor. Exp. Phys. 2020, 083C01 (2020) and 2021 update.

α_S extractions from jets at the LHC

- Sensitivity to strong coupling at colliders from the **jet multiplicity** and **the jet rate** when performing extractions from jet cross-sections.
  - NNLO theory for 3-jet cross-section has only recently been computed at leading colour (Czakon *et al.* 2106.05331)
- The only test of asymptotic freedom at the LHC that enters the PDG world average so-far is the **single point** from CMS's top pair production cross-section measurement, at $Q=m_t$



THIS POINT IS AT NNLO!

How can we perform an extraction like the LEP thrust one at the LHC?

LEP Extractions	LHC Extractions
e^+e^- 	Underlying event / colour reconnection ✗
Thrust extraction sensitive in resummation-dominated region	Jet rate or cross-section ratio usually compared to fixed-order calculations ✗

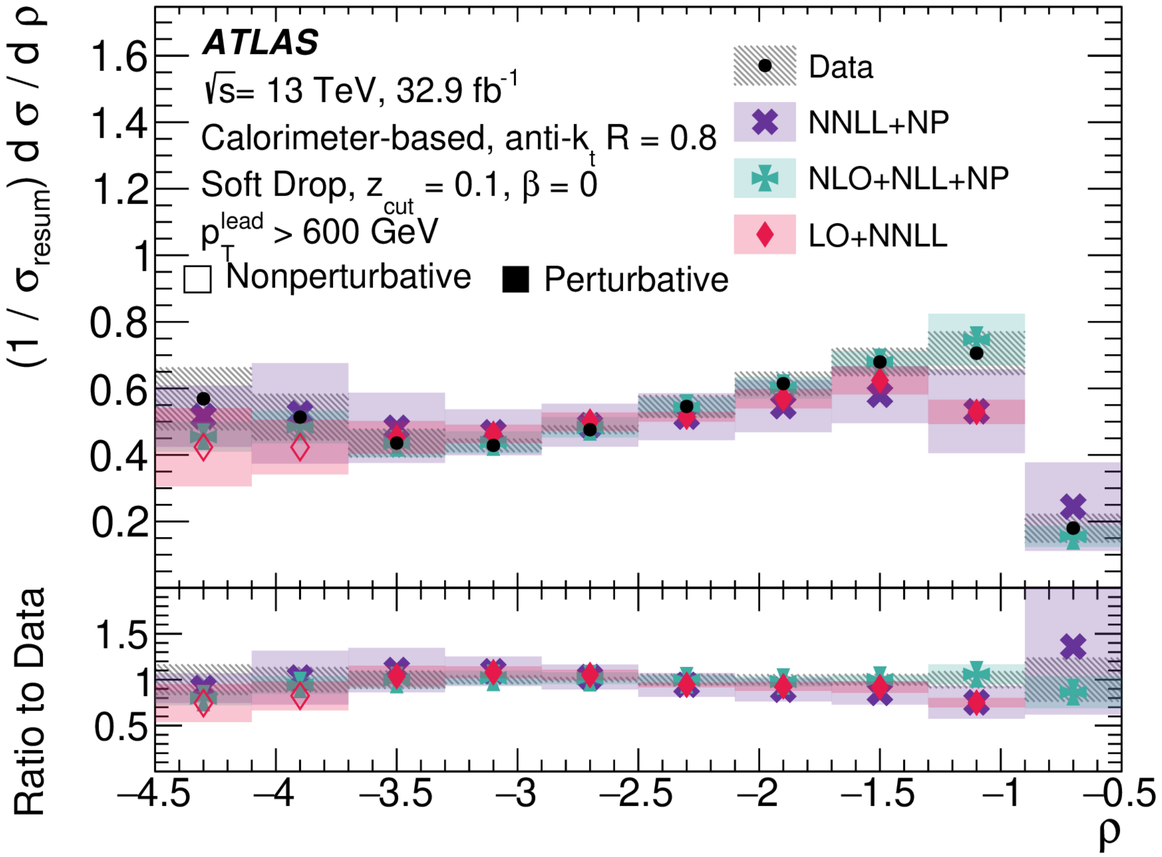
How can we perform an extraction like the LEP thrust one at the LHC?

LEP Extractions	LHC Extractions	LHC Extractions from JSS
e^+e^- 	Underlying event / colour reconnection ✗	soft-drop grooming → mitigate soft/wide-angled contributions ✓
Thrust extraction sensitive in resummation-dominated region	Jet rate or cross-section ratio usually compared to fixed-order calculations ✗	Soft-drop mass prediction → precise resummation ✓

Soft-drop jet mass

ATLAS, CMS references

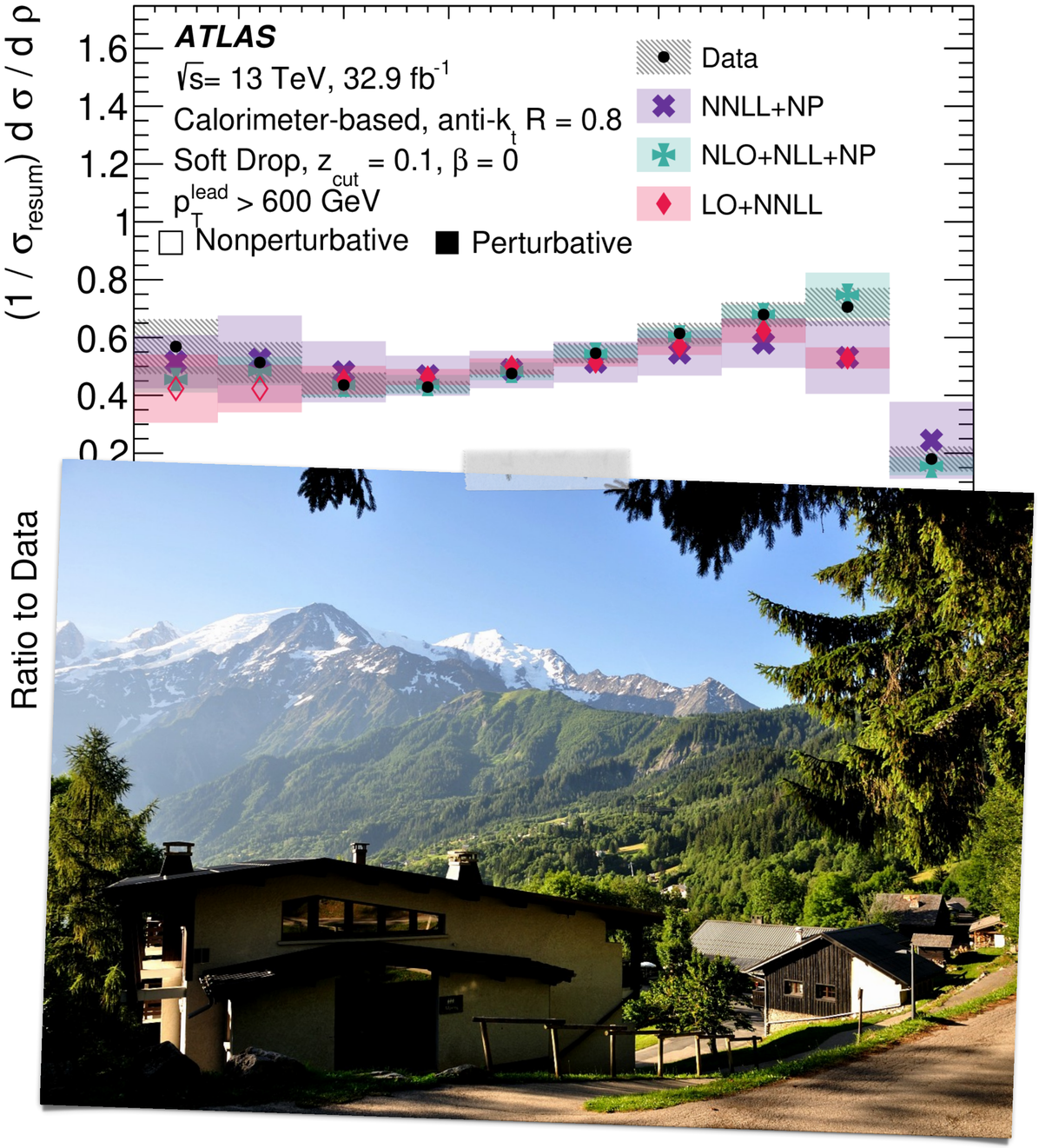
- The soft-drop jet mass has been measured during Run 2 and compared to >LL theory.
  - Great experimental and theoretical precision in the resummation-dominated region of the observable.
- Leading behaviour: $m_{SD} \sim C_i \alpha_S$.
  - Sensitive to the strong coupling.
  - Also sensitive to the q/g colour factor C_i.



Soft-drop jet mass

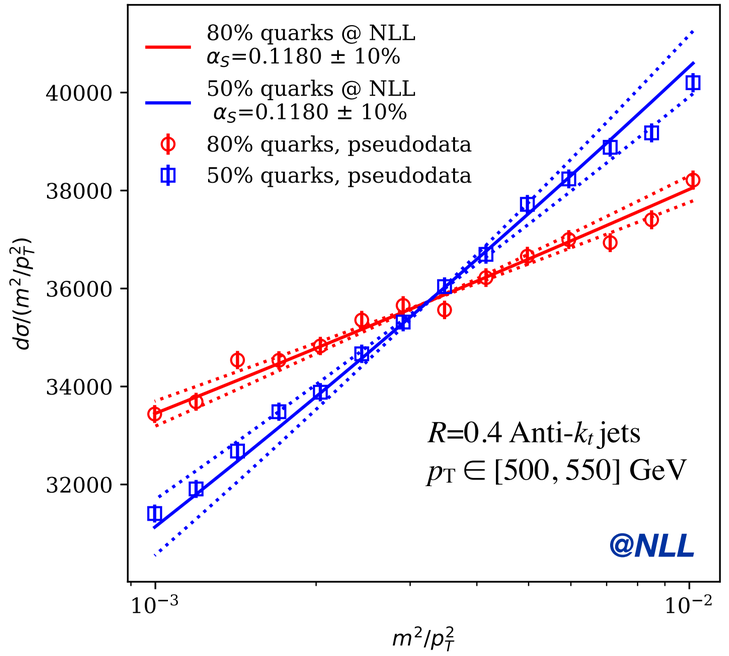
ATLAS, CMS references

- The soft-drop jet mass has been measured during Run 2 and compared to >LL theory.
  - Great experimental and theoretical precision in the resummation-dominated region of the observable.
- Leading behaviour: $m_{SD} \sim C_i \alpha_S$.
  - Sensitive to the strong coupling.
  - Also sensitive to the q/g colour factor C_i.
- Detailed prospective study for α_S extractions performed at Les Houches in 2017 (1803.07977).



Repeating the LH17 Fit

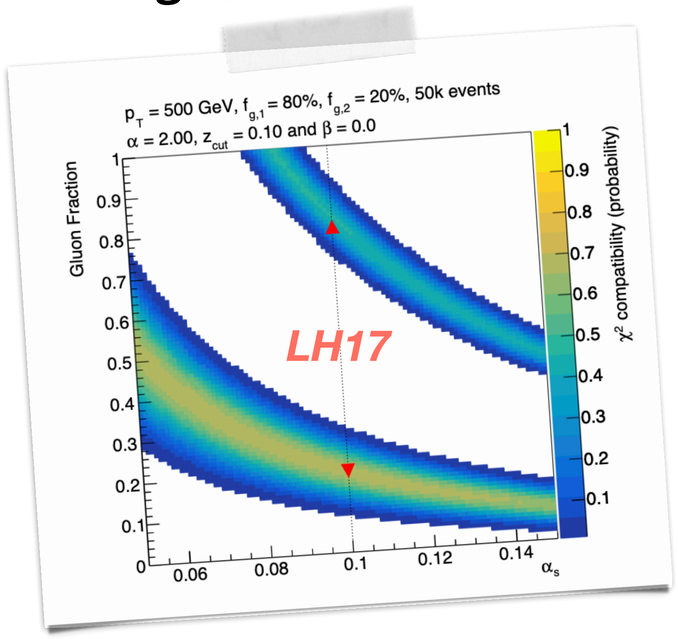
- Generate **pseudodata** from NLL prediction for quark- & gluon-initiated jets.

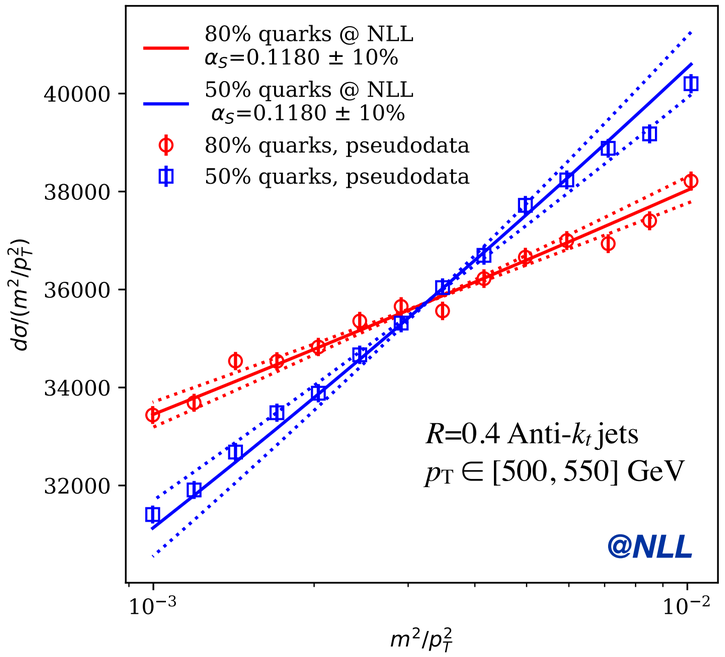


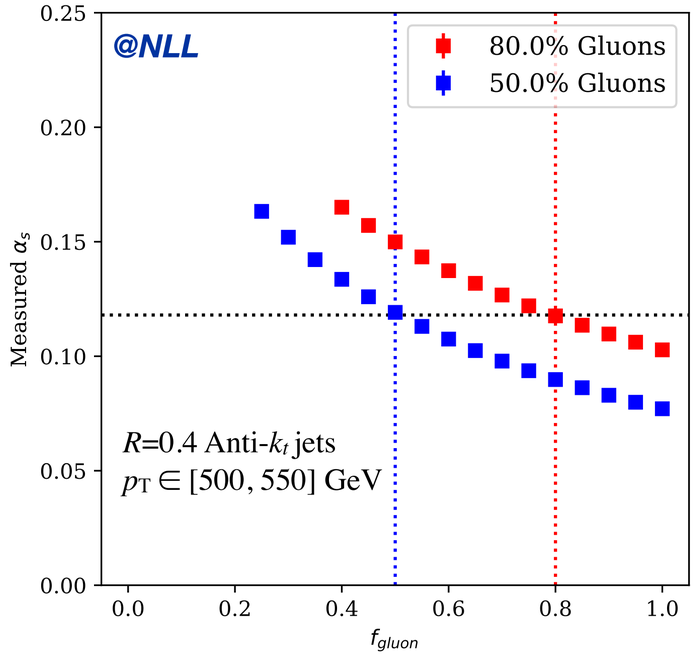
NLL prediction from Soyez et al. 1712.05105

Repeating the LH17 Fit

- Generate **pseudodata** from NLL prediction for quark- & gluon-initiated jets.
- Characteristic curve demonstrates degeneracy between f_{gluon} and α_S.
  - Obtained by **altering the value of f_{gluon} in the fit** for a given mixed sample **and extracting α_S.**



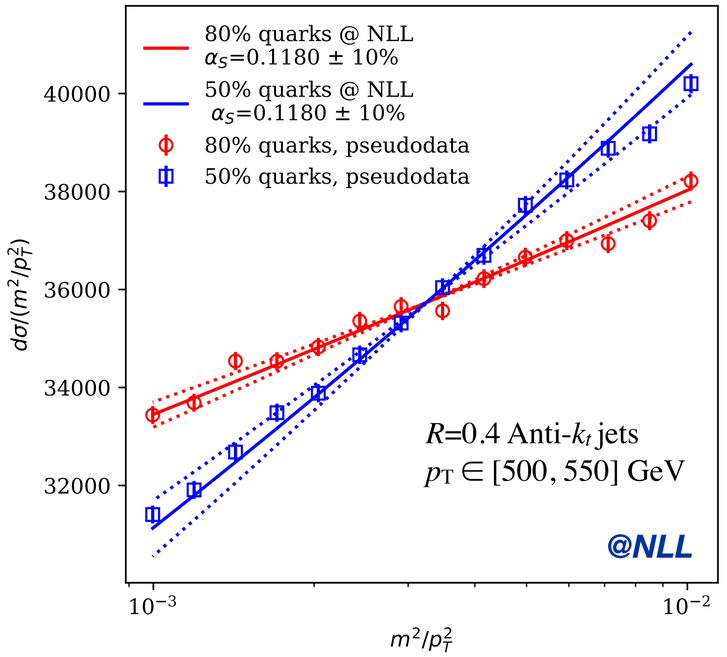


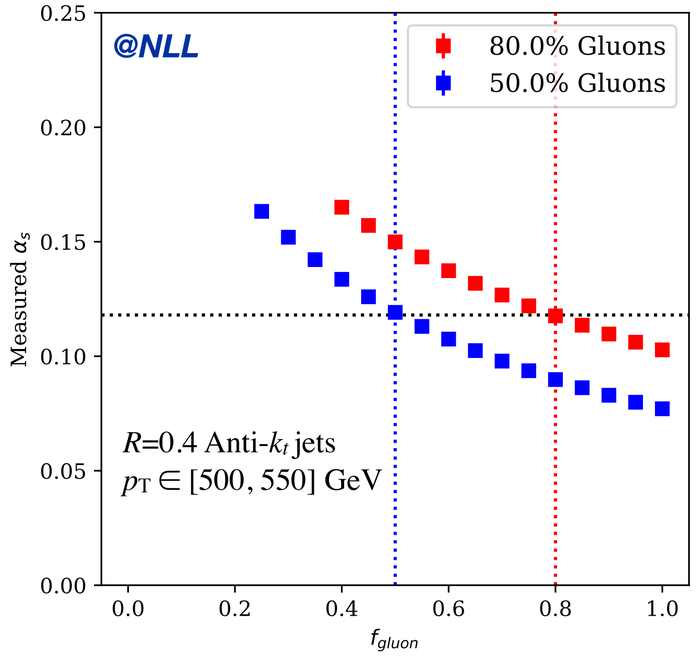


NLL prediction from Soyez et al. 1712.05105

Repeating the LH17 Fit

- Generate **pseudodata** from NLL prediction for quark- & gluon-initiated jets.
- Characteristic curve demonstrates degeneracy between f_{gluon} and α_S.
  - Obtained by **altering the value of f_{gluon} in the fit** for a given mixed sample **and extracting α_S.**
- Could break this degeneracy in a couple of ways:
  - Theoretical calculation of f_{gluon} → sensitive to PDFs, **could fit PDFs at the same time** (more complex).
  - Could **fit multiple higher-order correlators** simultaneously (degeneracy breaks beyond NLL).
    - *Requires simultaneous theoretical prediction of multiple JSS observables at very high precision …*

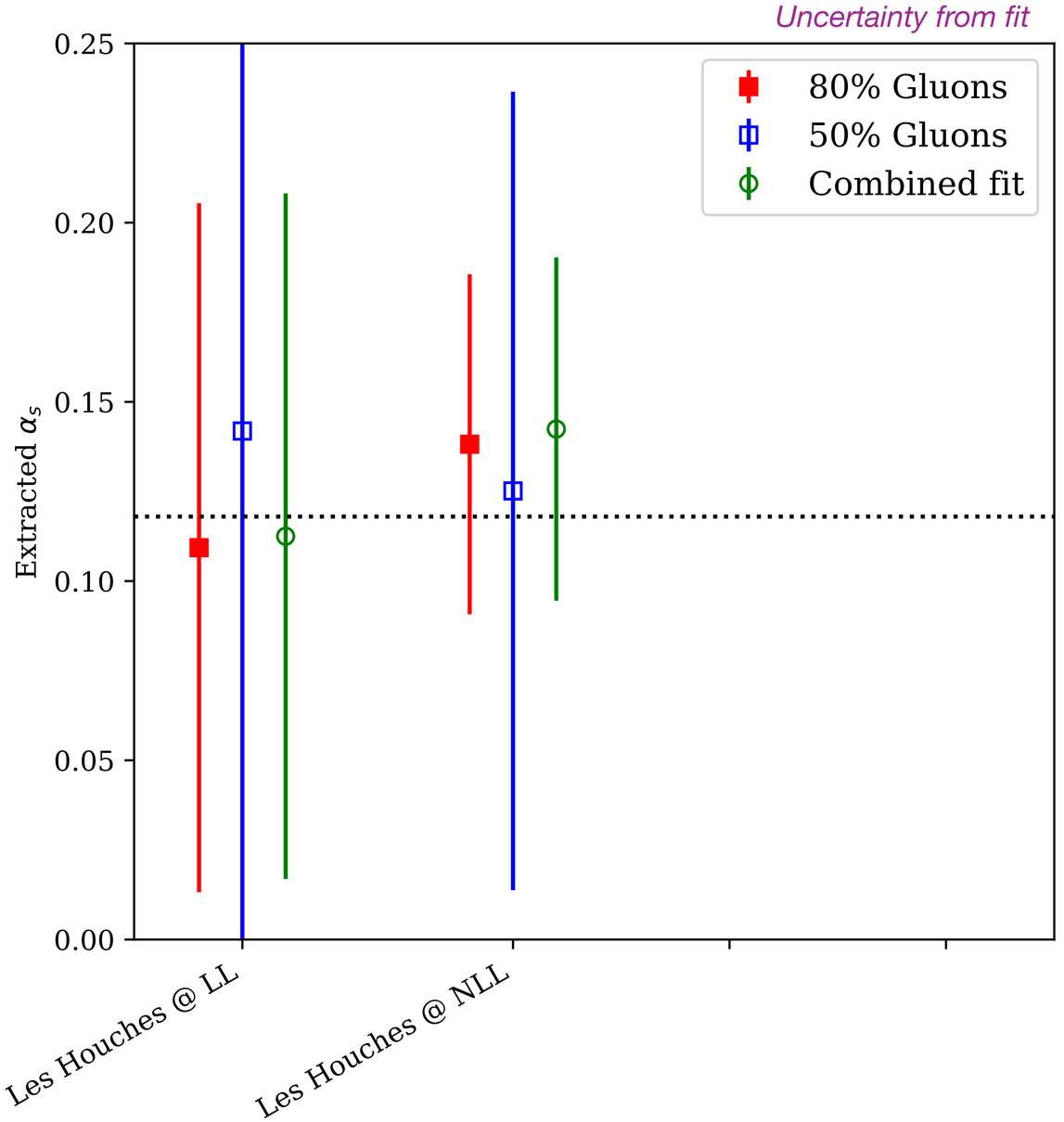




NLL prediction from Soyez et al. 1712.05105

Results: LH17 Approach

- When left unconstrained, this degeneracy results in large instabilities / uncertainties in the extracted α_S value.
  - Can fit both mixed samples at once to improve slightly, but errors still on the order of ~35%.
- The LH Jets Working Group became a bit pessimistic …



"Without some kind of conceptual breakthrough, though, we expect that the quark/gluon fraction will be a limiting aspect of α_S extractions from jet substructure at the LHC."

— Les Houches Jets Working Group, arXiv:1803.07977 [hep-ph]

How do we break this degeneracy?

Simultaneously fit α_S and the *q*/*g* fractions —
explored in LH17

Calculate the *q*/*g* fractions —
depends on PDFs

How do we break this degeneracy?

Simultaneously fit α_S and the *q*/*g* fractions —
explored in LH17

Calculate the *q*/*g* fractions —
depends on PDFs

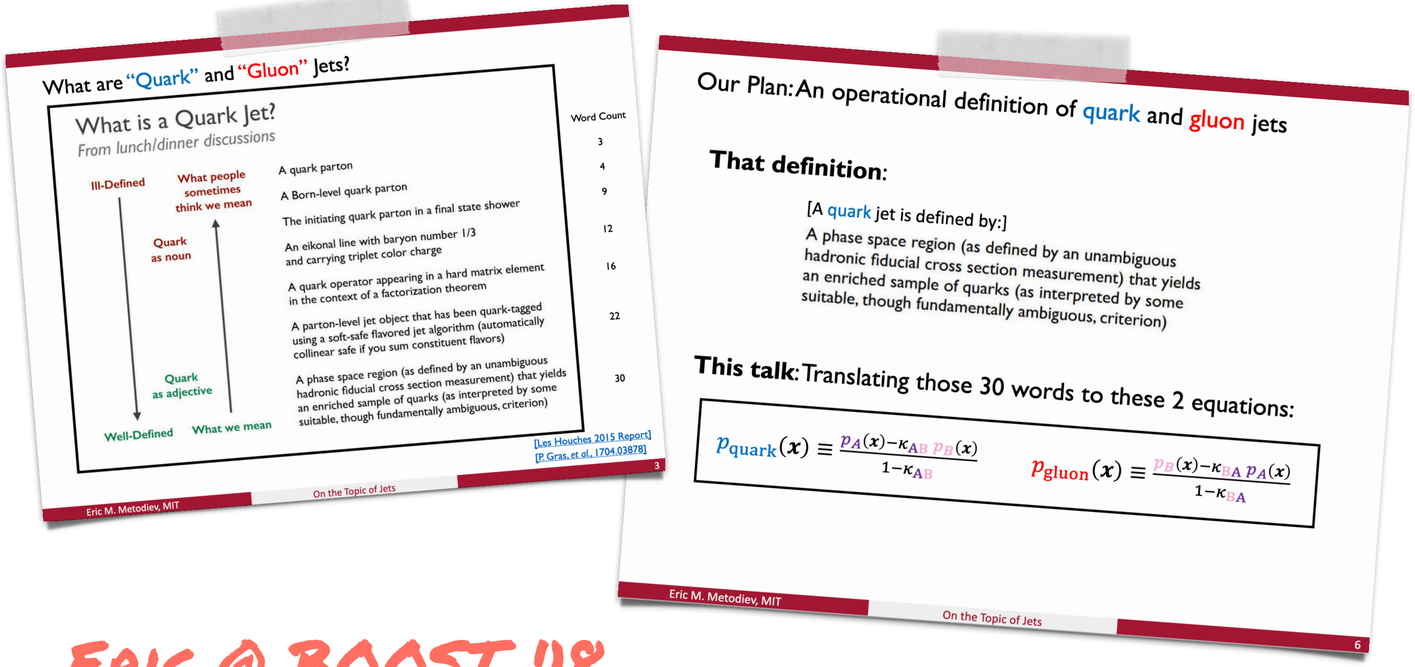
Demix underlying *q*/*g* distributions from data!

Jet topics

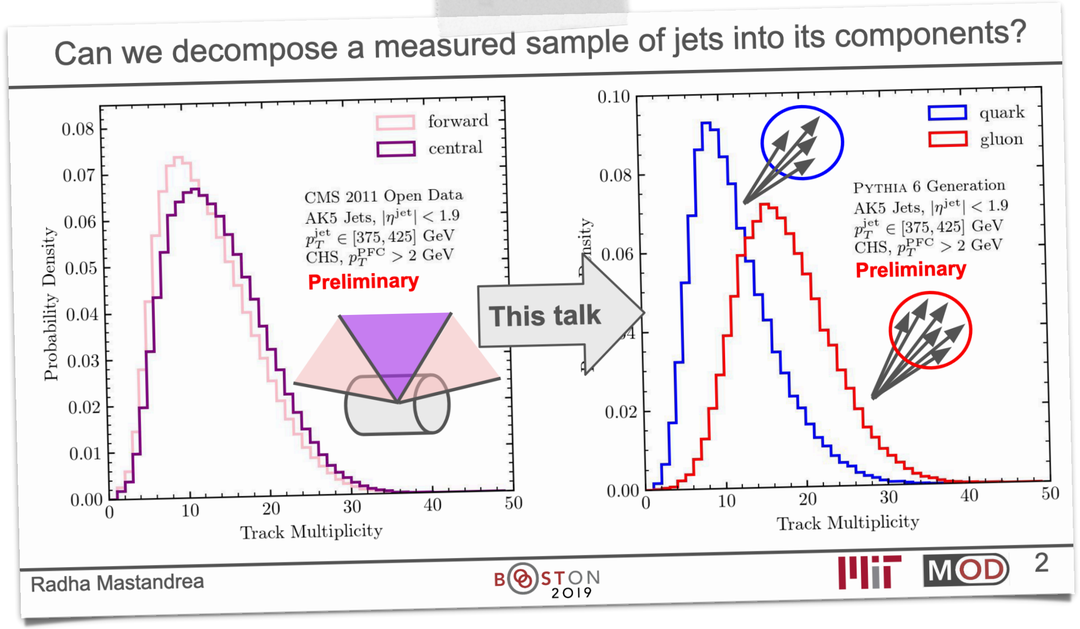
E. Metodiev & J. Thaler 1802.00008,
P. Komiske, E. Metodiev & J. Thaler 1809.01140
P. Komiske, S. Kryhin & J. Thaler 2205.04459

Key idea: *statistically* **`demix`** *two samples of q/g jets by maximally subtracting the two mixtures from one another, such that the zeros of the subtracted distributions correspond to the anchor regions.*

Data-driven: *does not depend on predictions/labels!*



Eric @ BOOST '18



Radha @ BOOST '19

Jet topics

- Obtain **jet topics** from likelihood ratios.

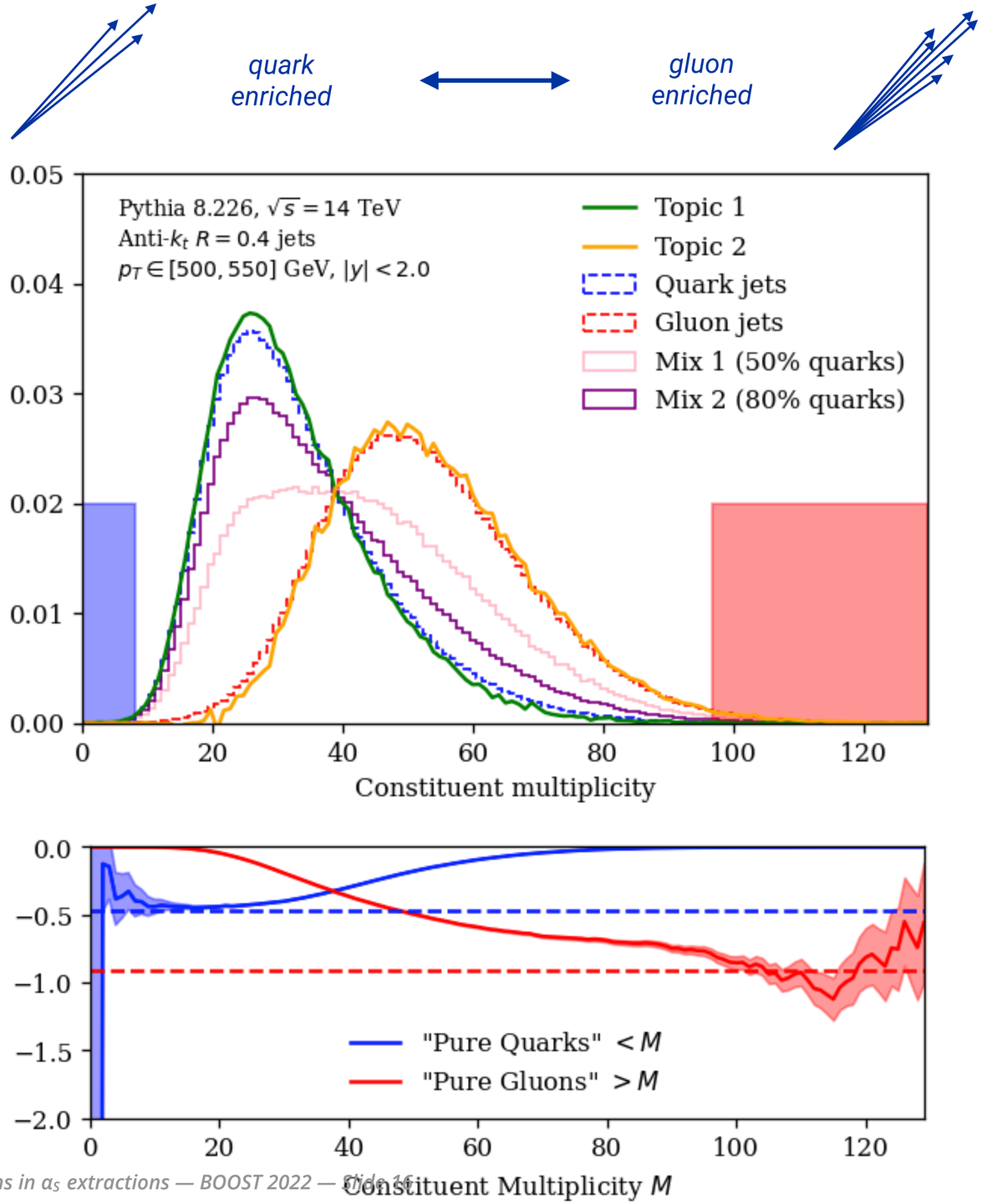
$$p_{\text{quark}}(x) = \frac{p_A(x) - \kappa_{AB}\, p_B(x)}{1 - \kappa_{AB}}$$

$$p_{\text{gluon}}(x) = \frac{p_B(x) - \kappa_{BA}\, p_A(x)}{1 - \kappa_{BA}}$$

  - Here, using two mixed samples of 50% (~dijets) and 80% quarks (~gluons).

- Extracted **topic 1** and **topic 2** recover original **quark** and **gluon** distributions!

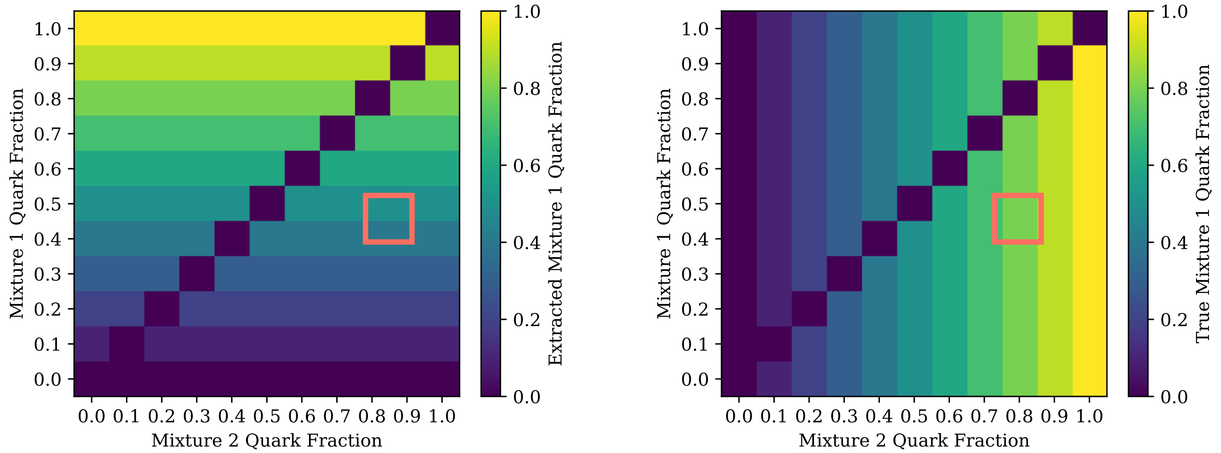
  - Can use extracted topics to obtain the q/g fractions of the two mixed samples (next slide).



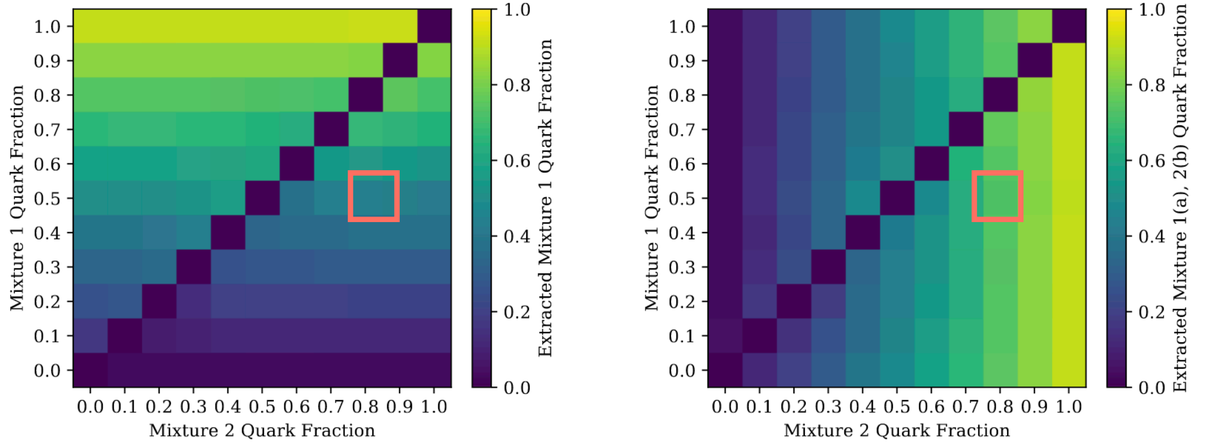
q/g fractions

- Results of q/g fraction extraction using $N_{\text{constitutents}}$.
  - Able to recover true fractions to within ~10-15%, performance worse for certain mixtures
    - *e.g.* pure in one flavour or degenerate.
- Using same anchor region for all bins — performance could probably be improved by someone patient.
  - Better methodologies to select anchor regions proposed while we were finalising the study : Komiske *et al.* 2205.04459.

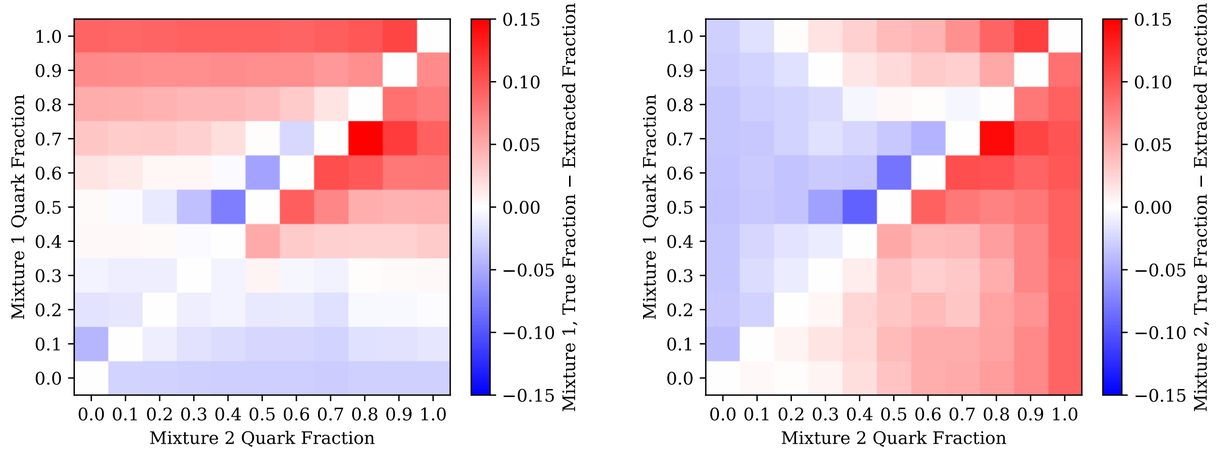
True fractions (trivial)



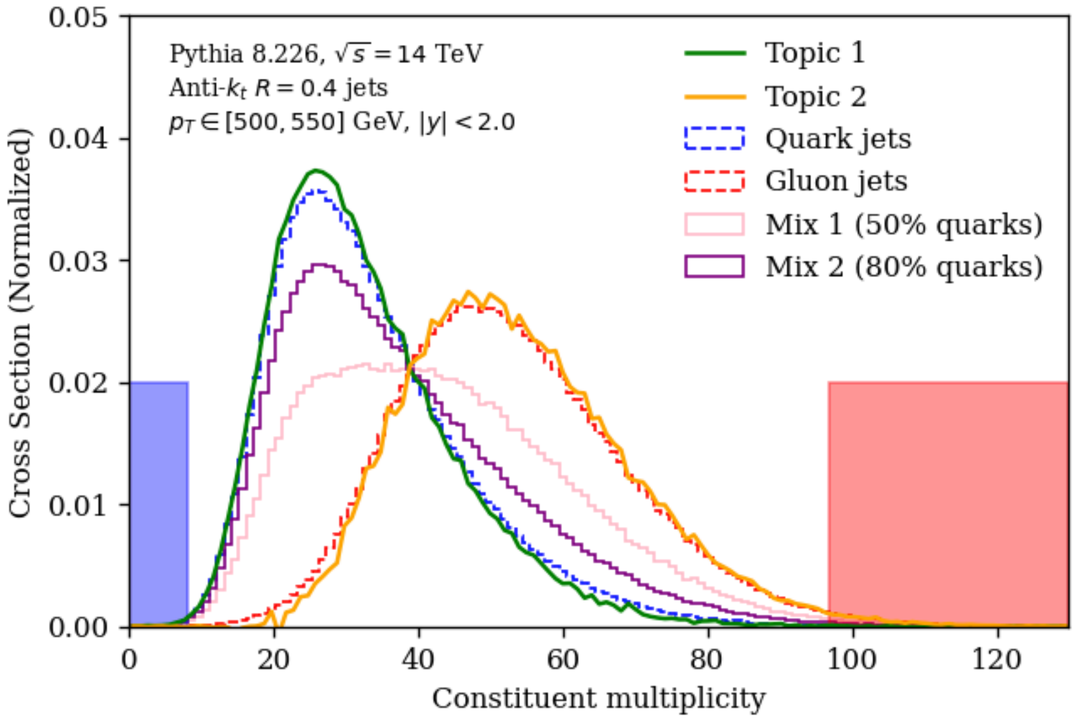
Extracted fractions (NConst.)

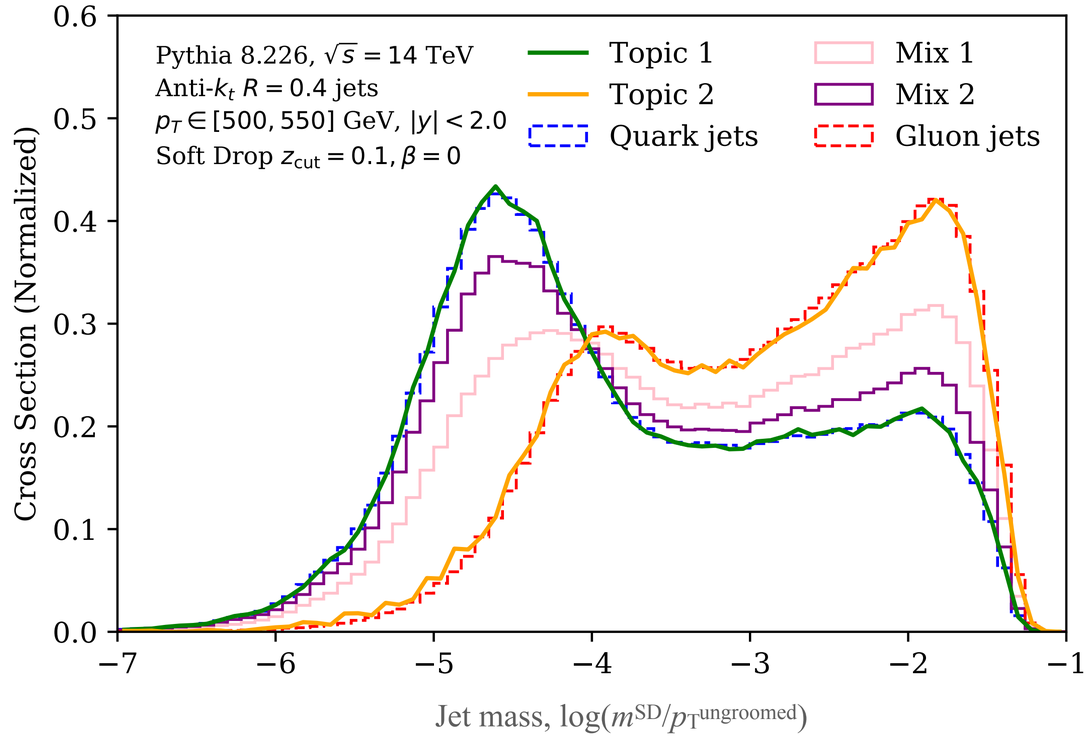


Non-closure (NConst.)



Going ‘off topics’ (get it now?)

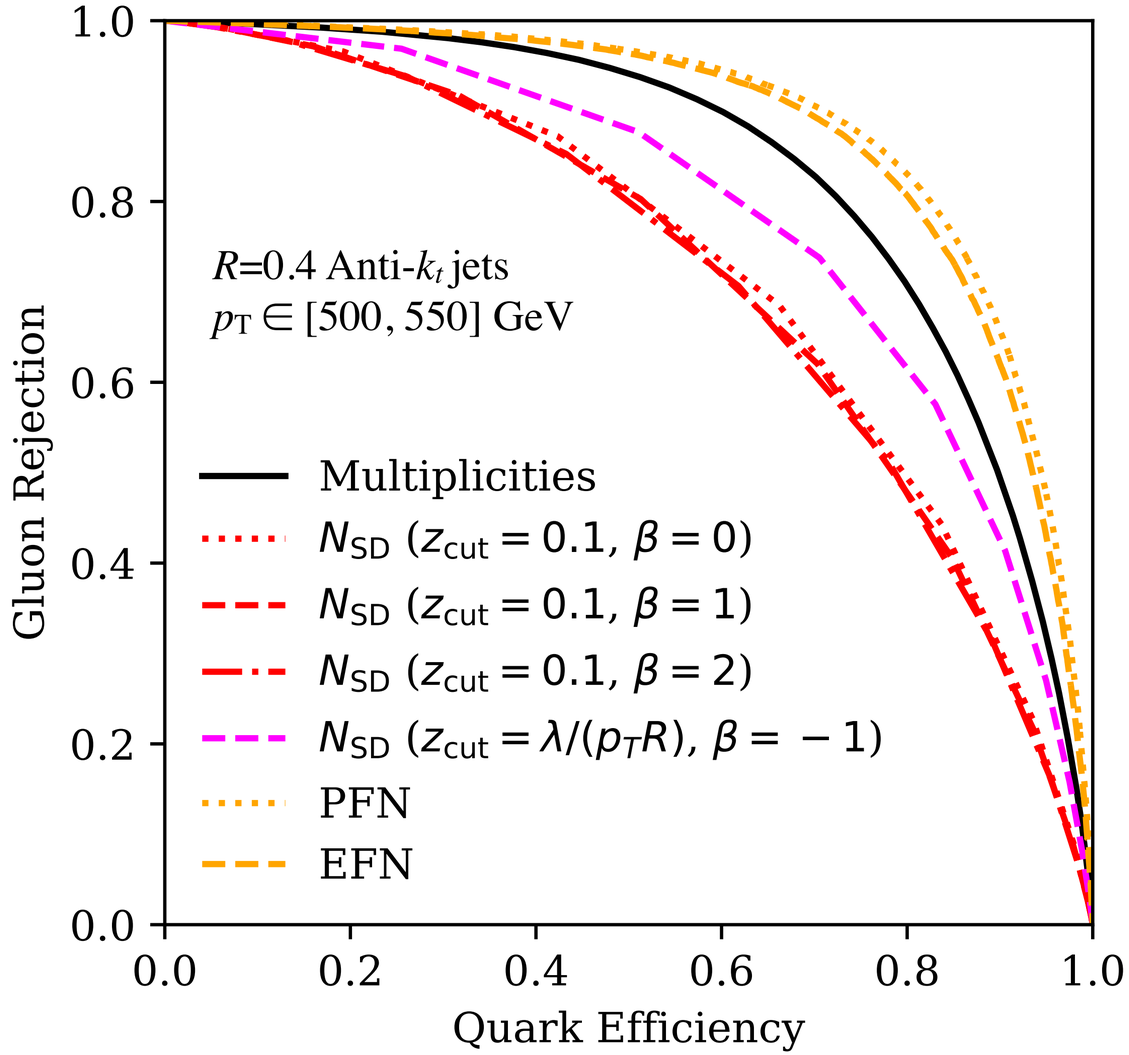




- Reducibility factors are a property of the ***event ensemble,*** rather than the jet/event.
  - Once we extract them from a mutually-irreducible observable, we can apply them to obtain **topics for *non-mutually-irreducible* observables.**

Other *q*/*g* discriminants

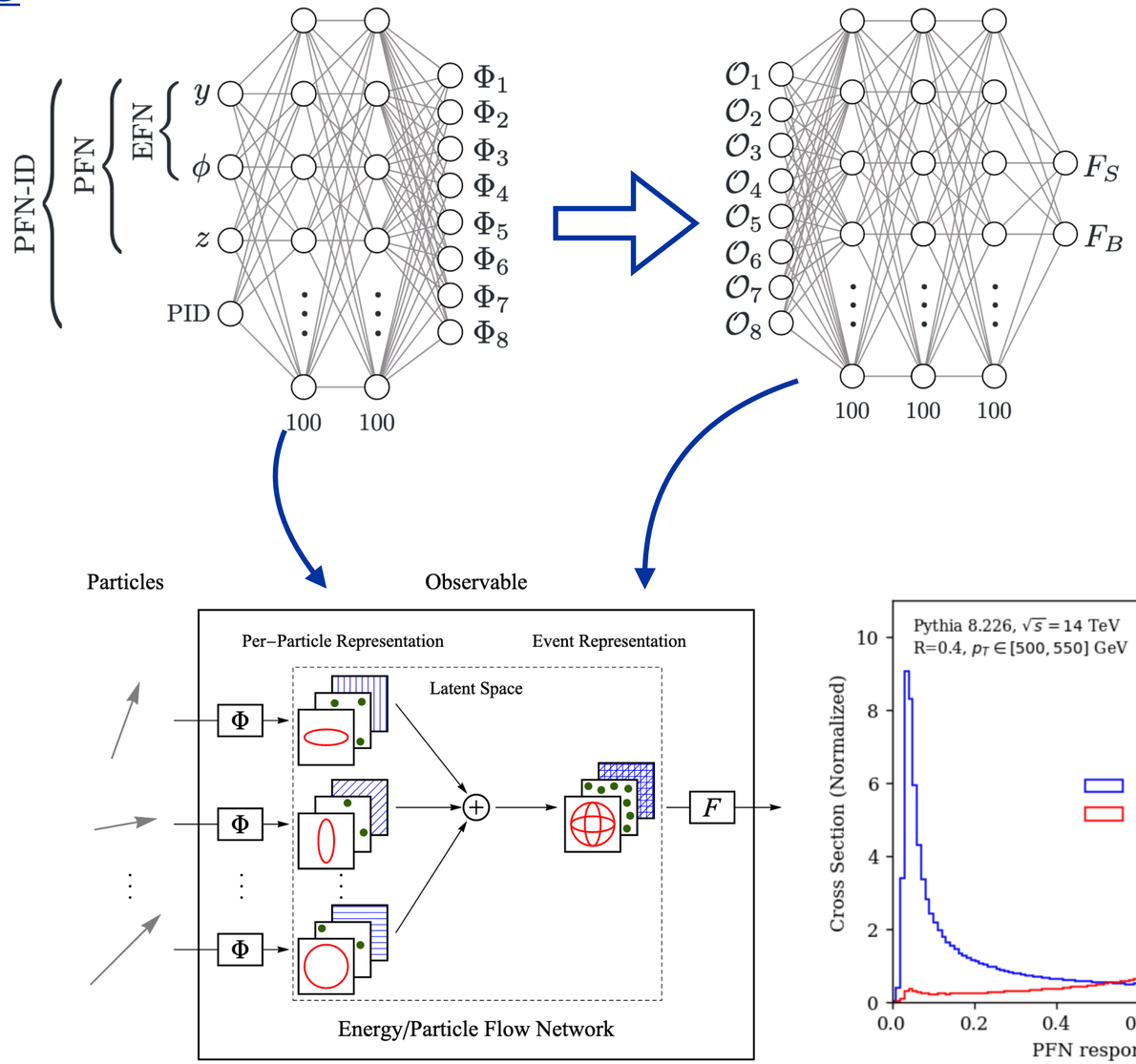
- $N_{\text{constituents}}$ may not be a desirable observable experimentally.
  - Systematics may be difficult to define in a robust way (cluster efficiency, *etc.*)
- — but any mutually-irreducible observable can be used.
  - In practice, whatever observable allows for the most precise extraction is probably the best one to try.
- Studied a few other options, and repeated the study in-detail with a PFN *q*/*g* classifier.
  - Could design an ML/AI discriminant with correlated response at particle- and detector-level, better experimental precision than $N_{\text{constituents}}$.



Particle Flow Networks (Deep Sets)

[Komiske, Metodiev & Thaler 1810.05165](#)
[Zaher et al. 1703.06114](#)

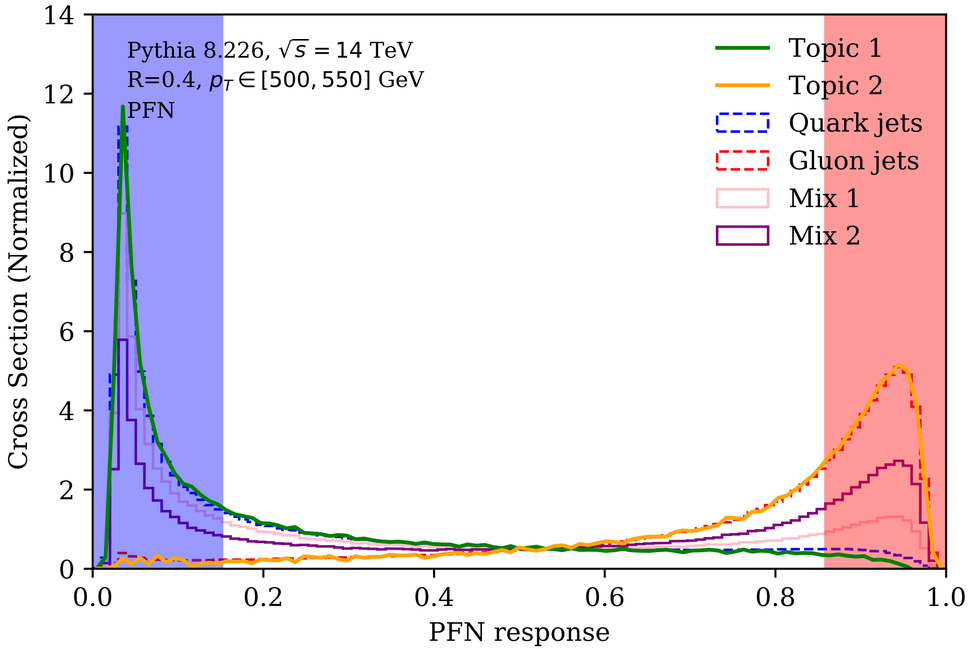
- Jets are ***unordered*** particle sets of ***variable length.***
  - Point-clouds are a natural representation for hadronic showers:
  - Deep sets framework is naturally permutation invariant due to internal sum layer.
- ***This approach has already been explained at BOOST, so we'll spare you the details this time.***

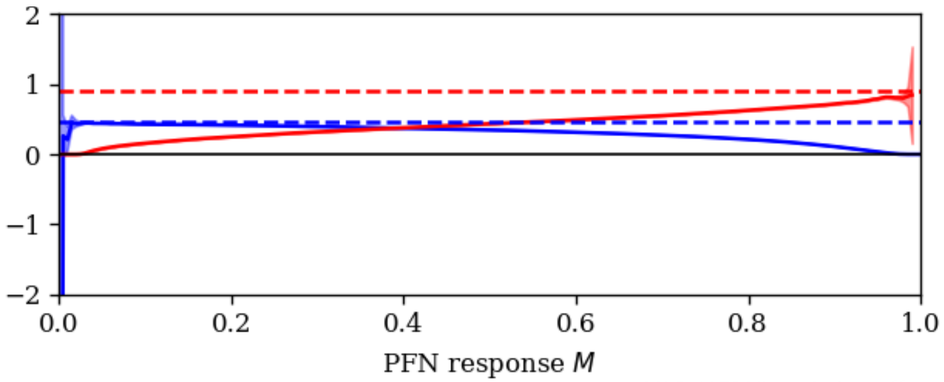


Particle Flow Networks (Deep Sets)

Komiske, Metodiev & Thaler 1810.05165

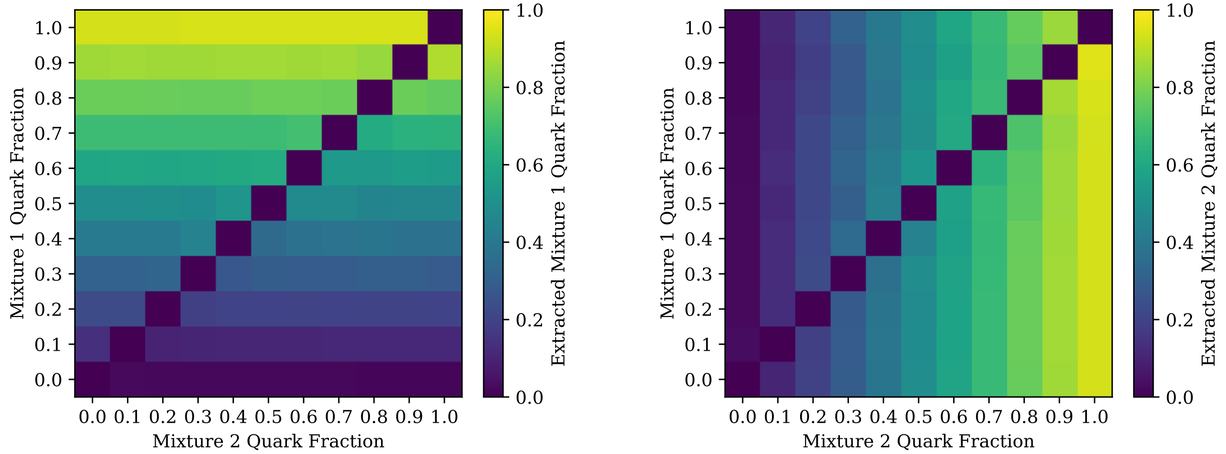
Zaher et al. 1703.06114





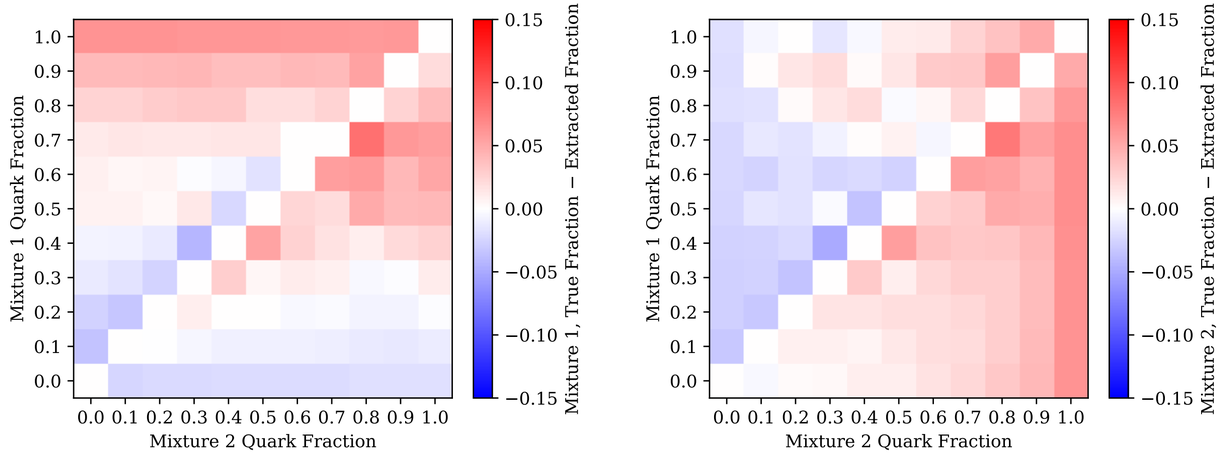
Follow same procedure

Extracted fractions



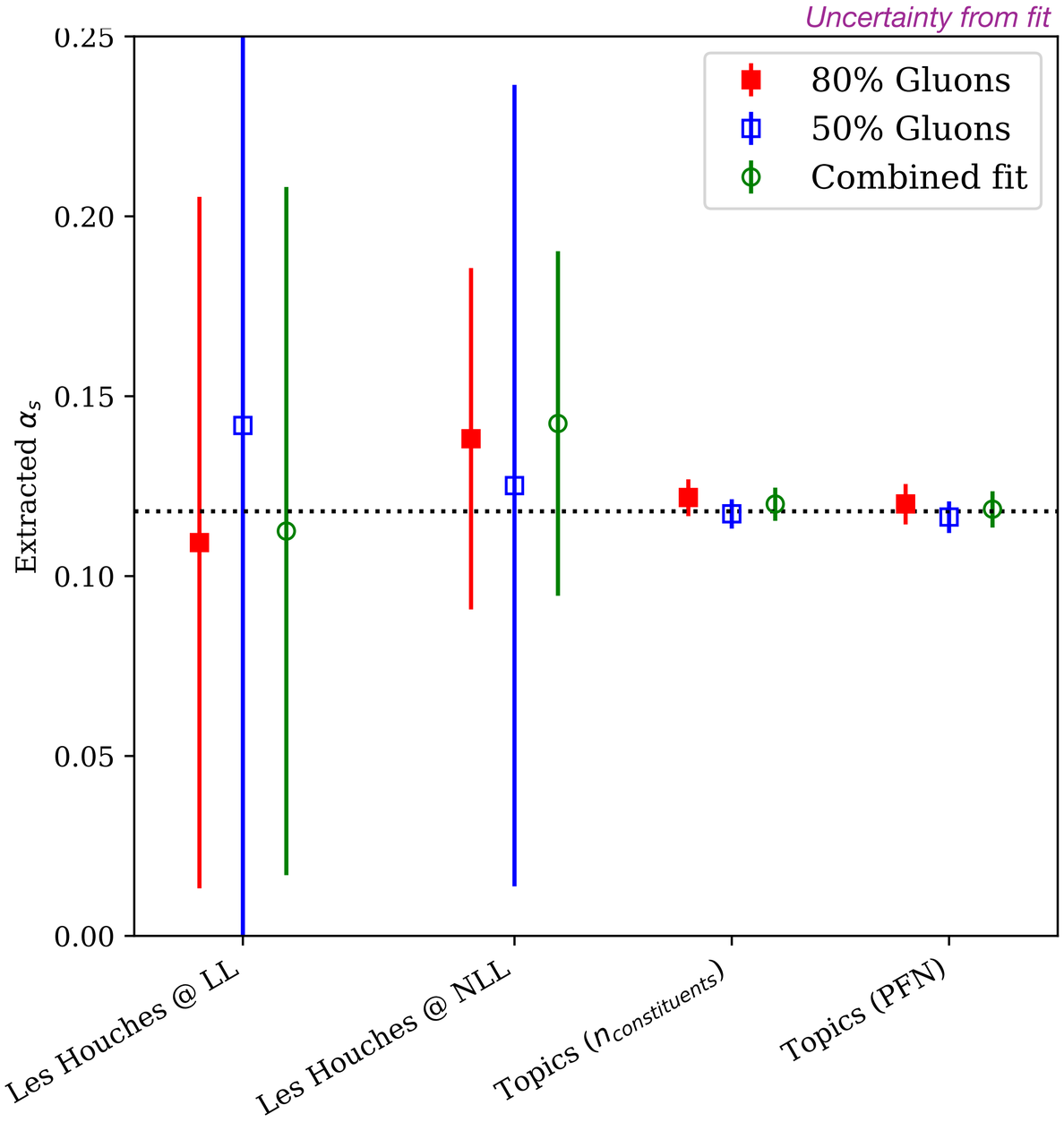
Closure

(Better than Nconstituents.)

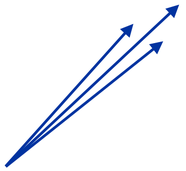


Improved Results

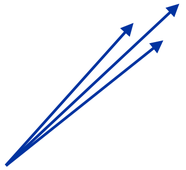
- Repeat fitting procedure, fixing *q*/*g* fractions to the values extracted via topics.
  - Large covariances in fit essentially removed: remaining uncertainty is either statistical in nature or related to anchor bin selection.
- Relative uncertainty in this fit reduced to ~7%.
  - Topics extractions can be further refined.
    - Anchor region selection still rudimentary here: better strategies in Komiske et al. 2205.04459.
- Other experimental & theoretical challenges remain, but this approach opens a new way forward toward the first extractions of α_S from JSS.



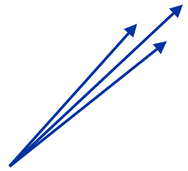
Summary



- We have outlined a proposal to apply statisical demixing procedures in the context of α_S extractions from groomed JSS observables.
  - This work builds upon the strategy outlined by the Jets working group at LH '17 (1803.07977).
  - Degeneracy between α_S and sample *q*/*g* composition is broken by this method, removing the major source of methodological uncertainty.



- It is possible to use IRC-unsafe information to extract the sample *q*/*g* composition without compromising theoretical tractability.
  - *i.e.* we can incorporate **ML/AI-based observables** and **precision QCD**!



- Many experimental & theoretical steps to go before complete α_S extractions from JSS are competitive.
  - *Can topics be identified as the "quarks" and "gluons" used in calculations?*

Thanks for thinking about these topics!