Measurements of jet substructure using the CMS detector

Cristian Baldenegro

LLR-École Polytechnique

Deep Inelastic Scattering 2022
Santiago de Compostela

Laboratoire Leprince-Ringuet

ÉCOLE POLYTECHNIQUE

IP PARIS

Jet substructure

- New insights on the strong force via detailed studies of the radiation pattern inside jets.

- In last years, we have achieved the experimental precision to challenge state-of-the-art pQCD analytical calculations and constrain parton shower & hadronization models of MC generators.

- Additional inputs for quark-gluon jet discriminators, and to discriminate V/H/t jets from quark/gluon jets.

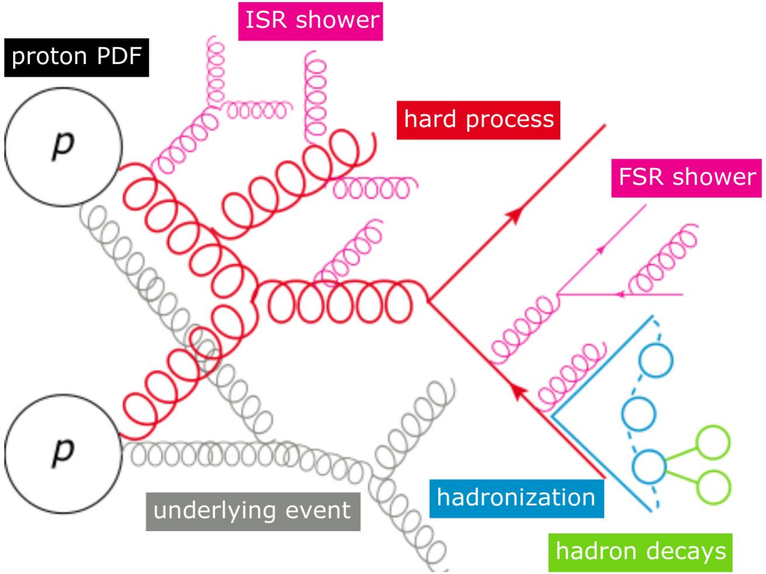
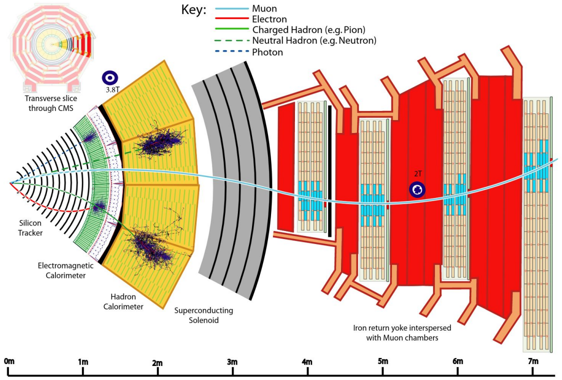


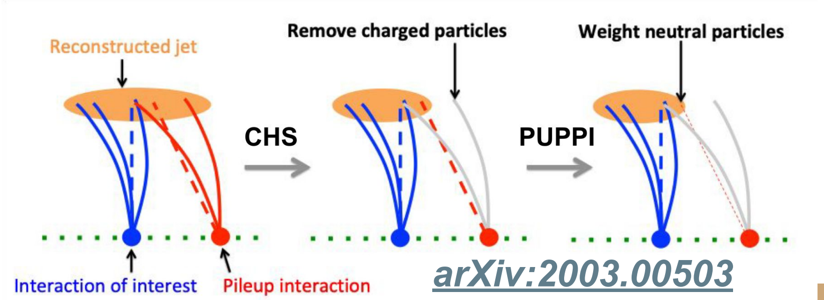
diagram by M. Seidel, LHCP2021

Jet substructure in CMS

Particle-flow (PF) candidates are clustered into jets with anti-kt algorithm, (R = 0.4 & 0.8 as standard sizes)



PUPPI algorithm: weigh down neutral clusters not close to PV tracks, remove tracks not associated to PV.



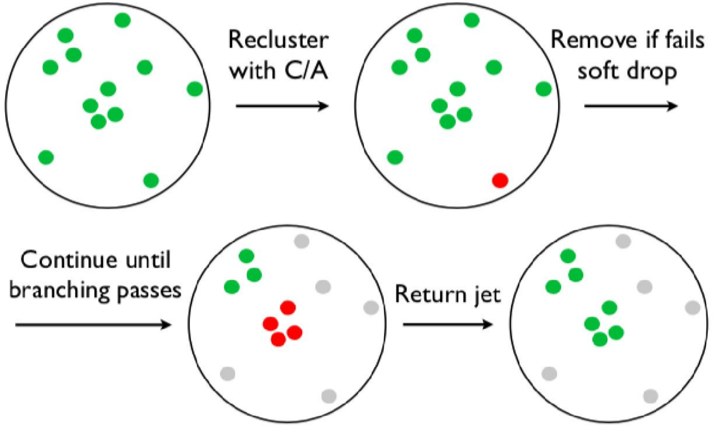
arXiv:2003.00503

Optional: iterative soft drop declustering to remove soft and wide-angle radiation.

Standard choice: β=0 (mMDT) for zCut = 0.1

Soft-drop criterion

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0} \right)^{\beta}$$

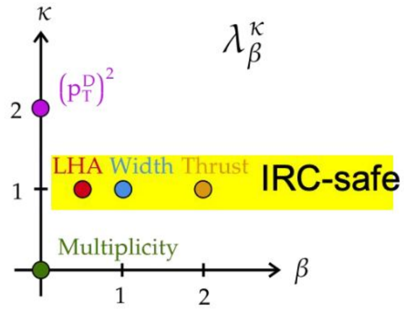


Jet angularities in Z+jets and dijets

arXiv:2109.03340

Generalized angularities, suggested by theorists to separate quark and gluon jets:

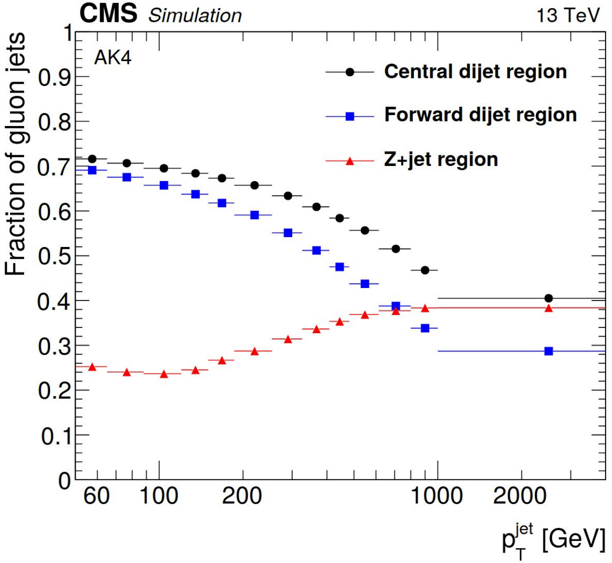
$$\lambda_\beta^\kappa = \sum_{i \in \text{jet}} z_i^\kappa \left(\frac{\Delta R_i}{R}\right)^\beta \quad z_i \equiv \frac{p_{Ti}}{\sum_{j \in \text{jet}} p_{Tj}}$$



Events classified in multiple categories:

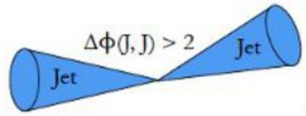
Dimension	Variants
Region	Z+jet vs. central dijet vs. forward dijet
Observable λ_β^κ	LHA, width, thrust, multiplicity, $(p_T^D)^2$
Jet p_T	$50 < p_T < 65\,\text{GeV}$, ..., $p_T > 1000\,\text{GeV}$
Jet size parameter R	0.4 vs. 0.8
Constituents	Charged+neutral vs. charged-only
Grooming	Ungroomed vs. groomed

Z+jet, quark enriched ***dijet, gluon enriched***

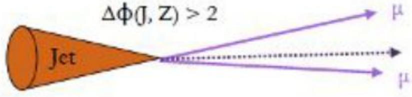


arXiv:2109.03340

Event samples

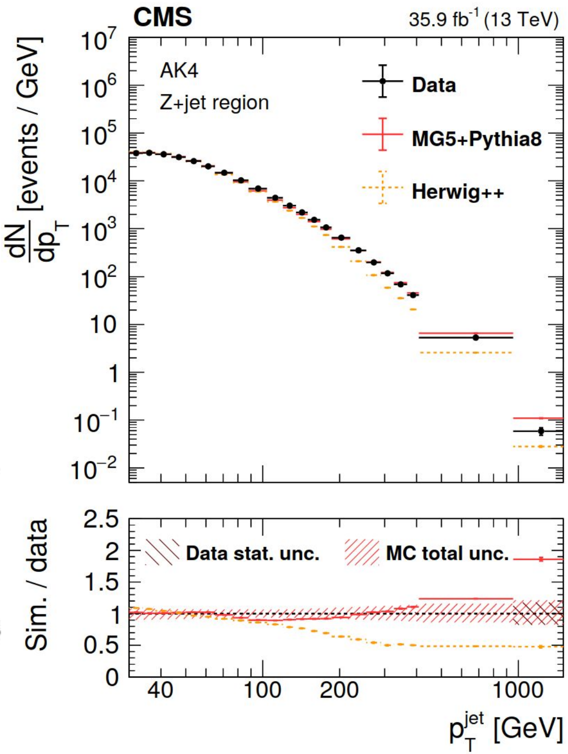


$\geq$2 jets with $|y| < 1.7$ and $p_T^j > 30\,\mathrm{GeV}$
$\Delta\phi(j_1, j_2) > 2$
$|p_T^{j_1} - p_T^{j_2}| / (p_T^{j_1} + p_T^{j_2}) < 0.3$

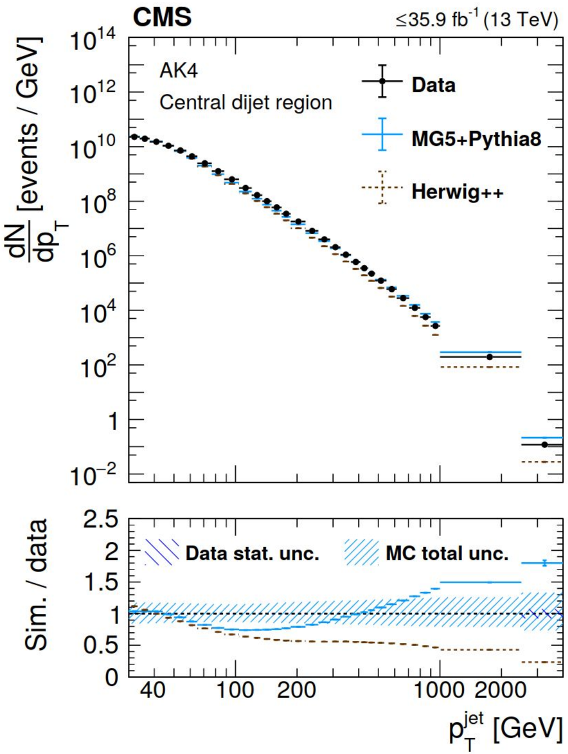


$\geq$2 muons with $|\eta| < 2.4$ and $p_T^{\mu} > 26\,\mathrm{GeV}$
Opposite charge muons
$|m_{\mu\mu} - m_Z| < 20\,\mathrm{GeV}$
$\geq$1 jet with $|y| < 1.7$ and $p_T^j > 30\,\mathrm{GeV}$,
not overlapping with muons of the Z boson candid
$\Delta\phi(j_1, Z) > 2$
$|p_T^{j_1} - p_T^{Z}| / (p_T^{j_1} + p_T^{Z}) < 0.3$

Z+jet, quark enriched



dijet, gluon enriched



Les Houches Angularity (LHA) distribution

arXiv:2109.03340

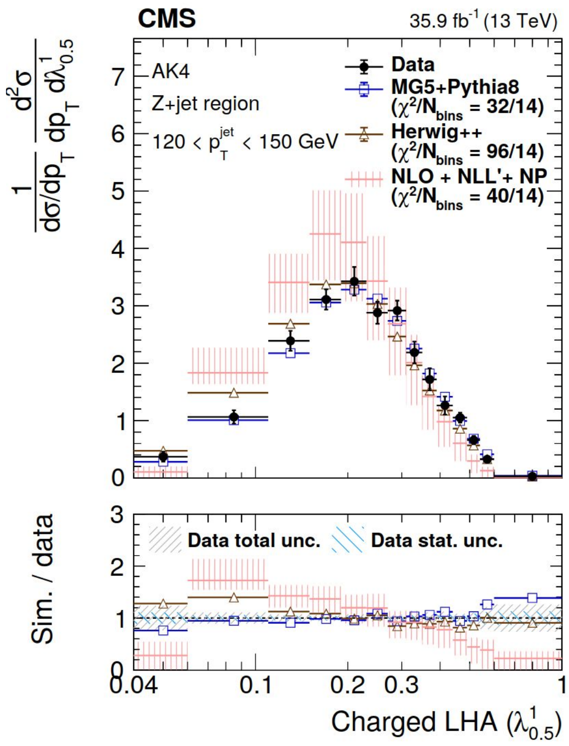
Measurement unfolded to stable-particle level.

MG5+Pythia8 and **Herwig++** describe well the quark-enriched sample. They “envolve” the gluon-enriched data.

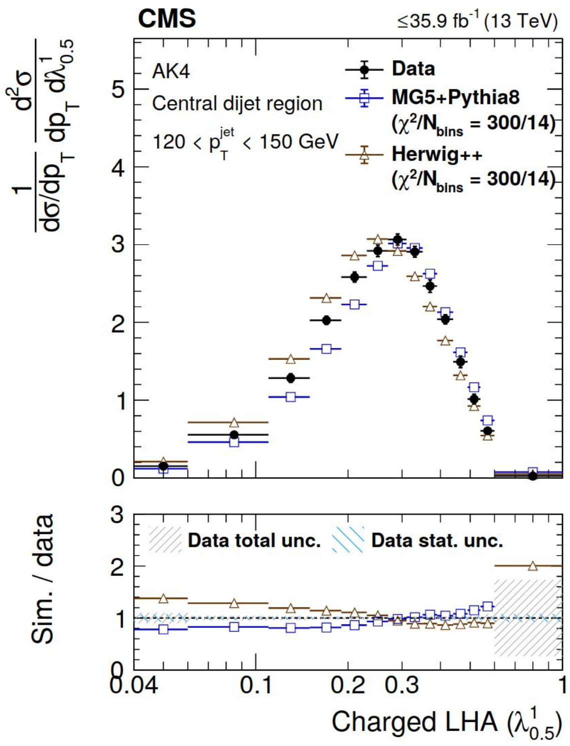
For Z+jet: analytical resummation at NLL accuracy matched to fixed order NLO matrix element, supplemented w/ non-perturbative corrections from Sherpa (**NLO+NLL’+NP**)

D. Reichelt, S. Caletti, O. Fedkevych, S. Marzani, S. Schumann, G. Soyez
arXiv:2112.09545

Z+jet, quark enriched



Dijet, gluon enriched



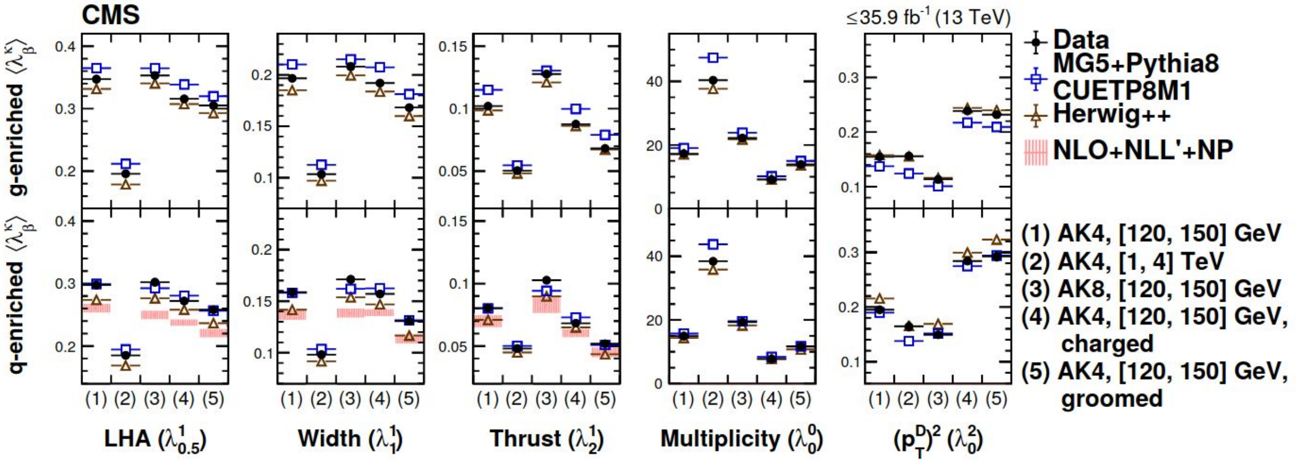
summary plot

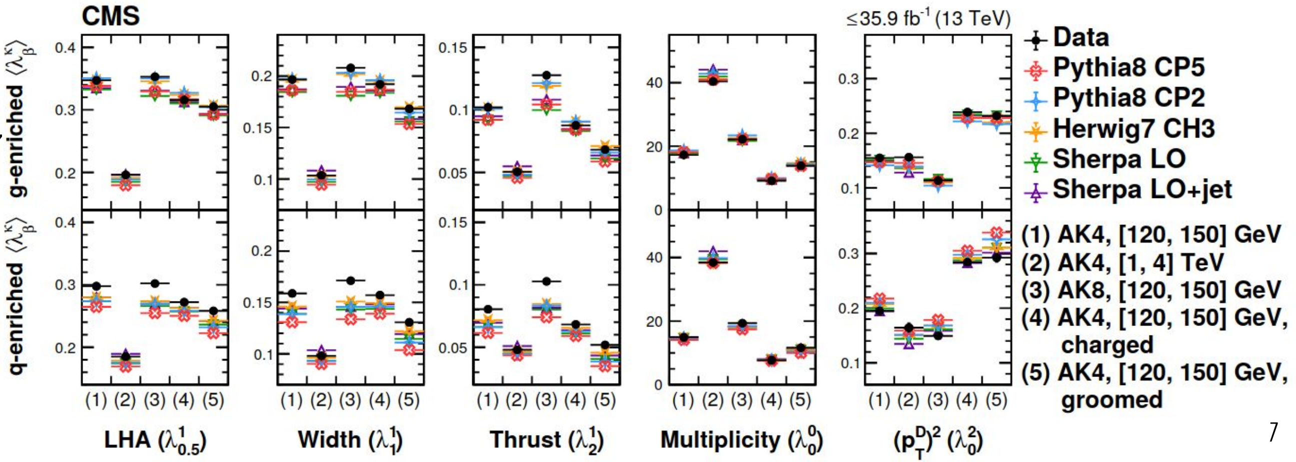
arXiv:2109.03340

Angularities are generally larger in gluon-enriched jet samples, consistent with LO picture.

Quark- and gluon-initiated parton showers are not well described by generators, ***important for flavor tagging developments.***

Newest Pythia8 (CP2, CP5) and Herwig7 (CH3) tunes: description of gluon-like jets improves, not much improvement for quark-like jets.





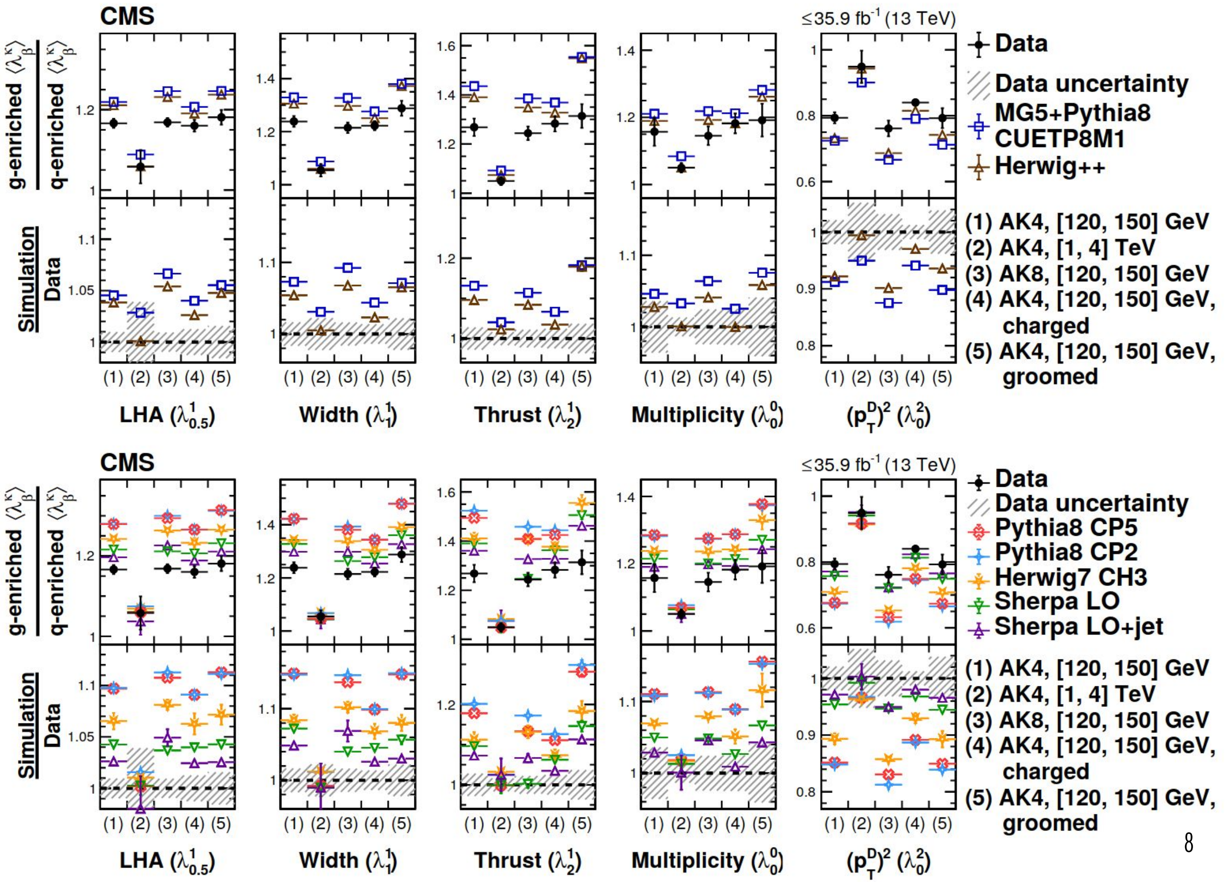
Dijet/Z+jet ratio (gluon-like/quark-like jet ratio)

arXiv:2109.03340

- experimental uncertainties partially cancel in the Z+jet / dijet ratio

- generators in LO+PS mode overestimate the gluon-enriched/quark-enriched ratio

- Description of gluon-enriched / quark-enriched ratio worsens with newest Pythia8 and Herwig7 tunes.



Jet substructure in top quark pair + jet events

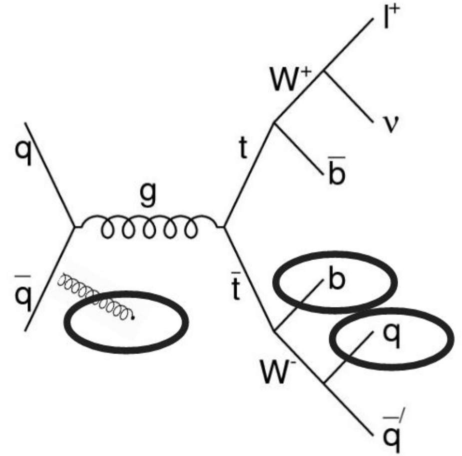
arXiv:1808.07340

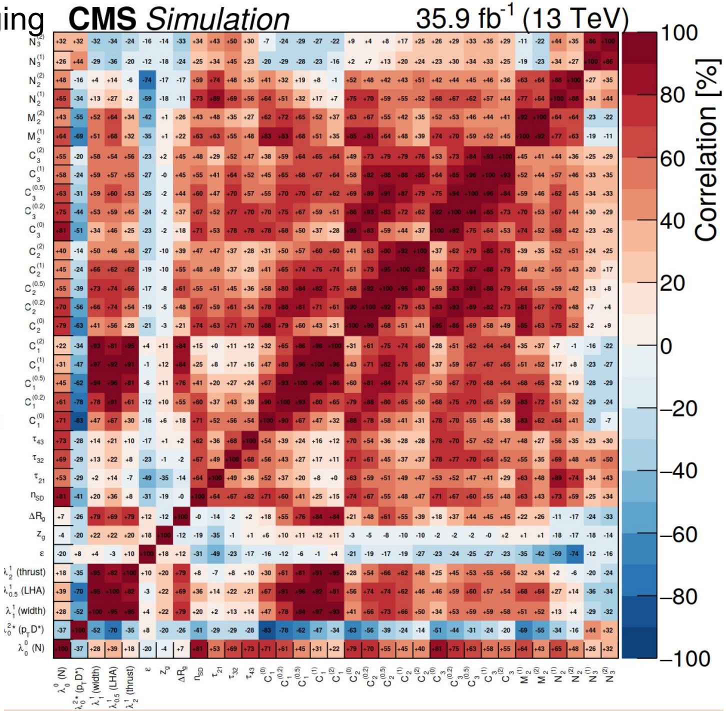
Top quark pair production provides: bottom, light-quark enriched, and gluon-enriched jet samples:

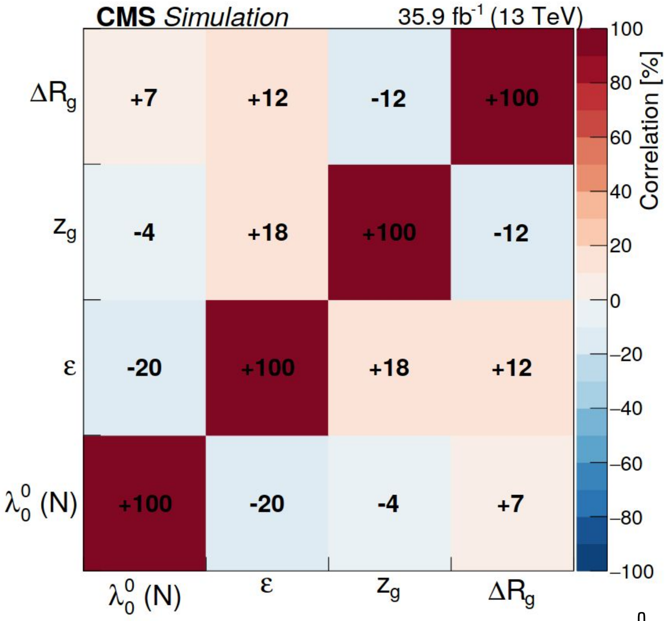
- ***Light-enriched:*** non b-tagged jets with |mjj − mW | < 15 GeV
- ***Gluon-enriched:*** non b-tagged jets with |mjj − mW | > 15 GeV
- ***B-quark* enriched** from b-tagging

33 observables were tested (Nsubjettiness, energy correlators, gen. angularities, …). A set of minimally-correlated variables is analyzed in detail:

ΔR_g, z_g, multiplicity (λ_0^0), eccentricity (ϵ)

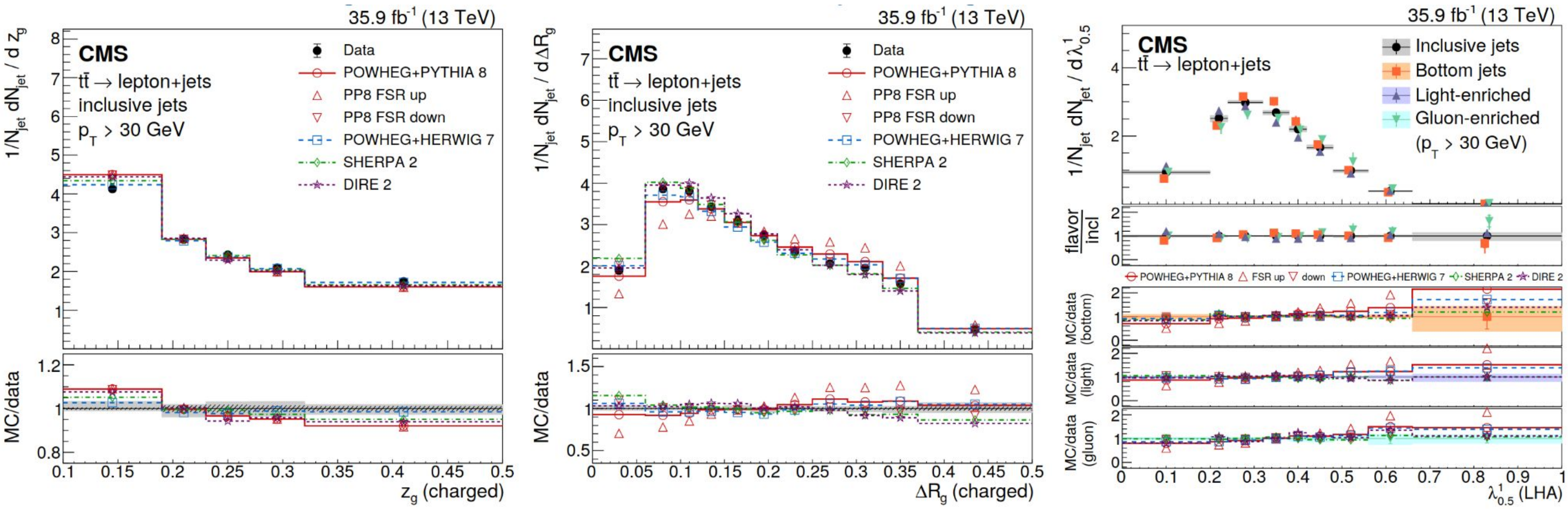






Measurement of jet substructure in top quark pair+jets

arXiv:1808.07340



- Groomed momentum fraction zg, related to the splitting function of QCD, at LO insensitive to αS
- Angle between groomed subjets ΔRg: sensitivity to αS, robust against non-perturbative corrections.
- LHA is expected to be larger for gluon-jets than quark jets. B-quark jets are expected to have smaller LHA.

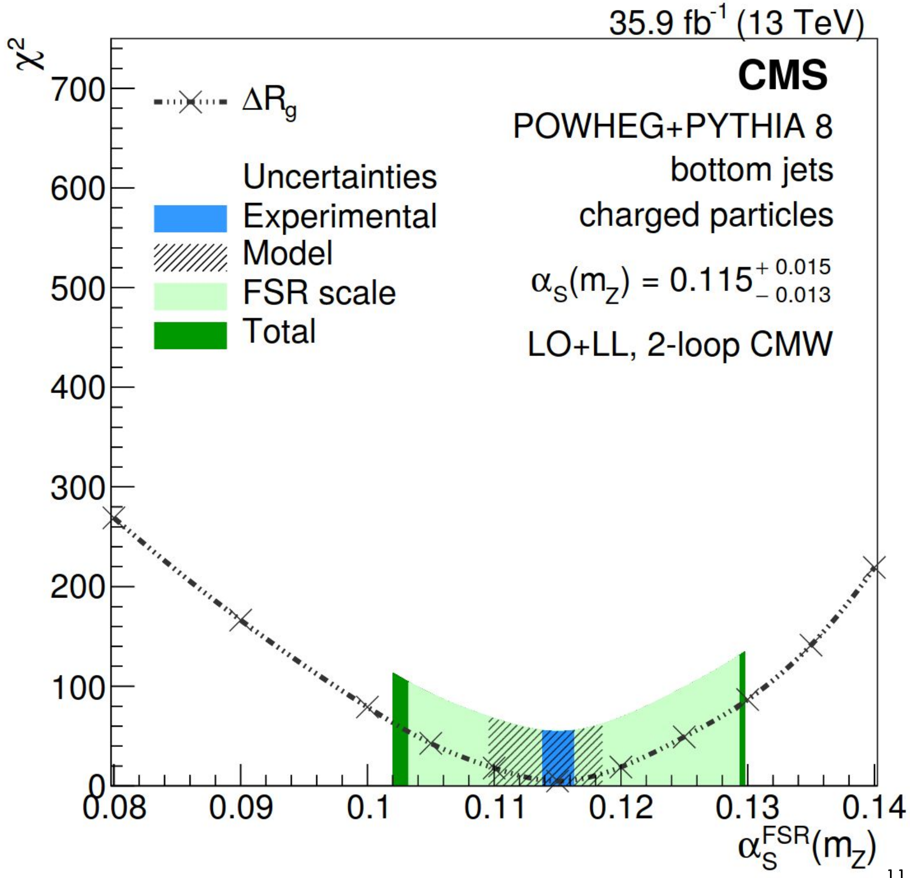
arXiv:1808.07340

α_S extraction from ttbar lepton + jets

- ∆Rg is expected to be the least sensitive to NP effects.
- Substructure of b-jets is used for the extraction of αS (NLO matrix element of top quark decay in PYTHIA8, effectively LO+LL b-jet substructure)
- Higher-order corrections supplemented in an effective way using 2-loop running and CMW rescaling of ΛQCD.
- Result:

$$\alpha_S(m_Z) = 0.115^{+0.015}_{-0.013}$$

experimental uncertainties are less than 1% of αS, uncertainty in αS is dominated by FSR scale uncertainties.

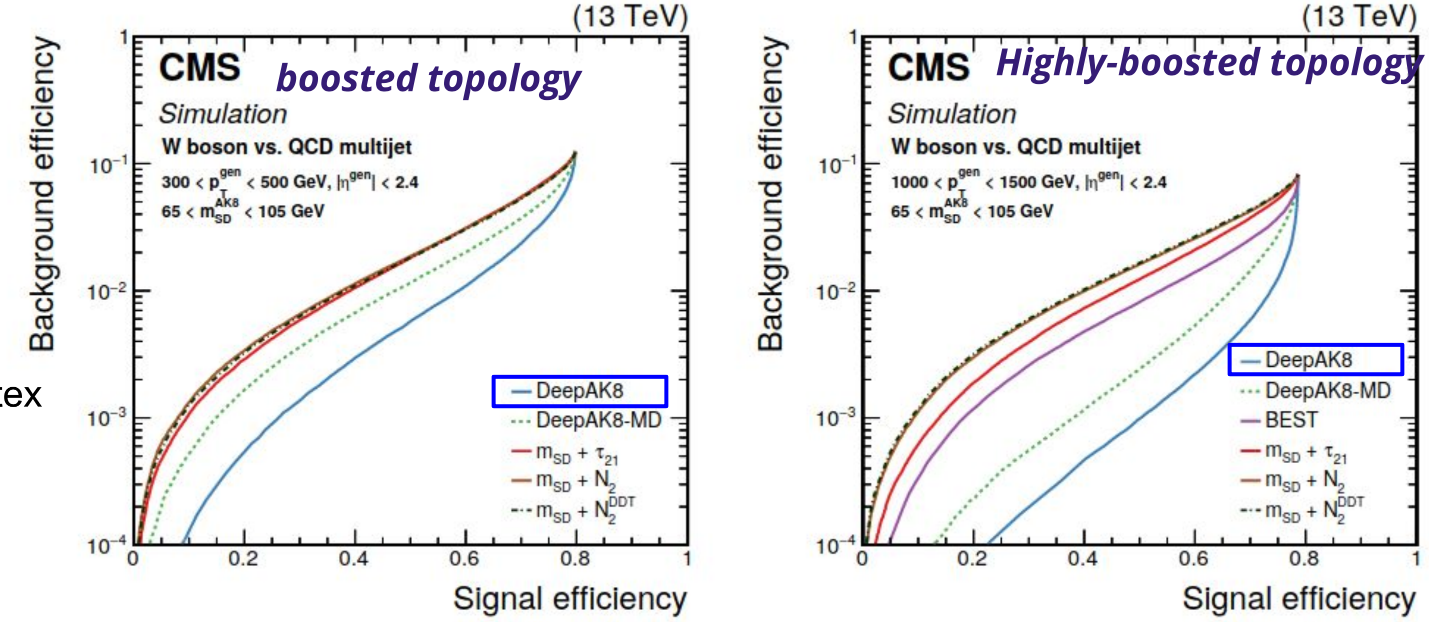


Boosted object tagging using machine learning algorithms

- Improved tagging performance with neural networks (NN) over cutoff methods
- DeepAK8 (convolutional NN) gives best performance for highly boosted topologies.

arXiv:2004.08262

DeepAK8:
100 particles
(4-momenta, charge,
track info, secondary vertex
info, angular separation
relative to jet axis, …),
42 variables total



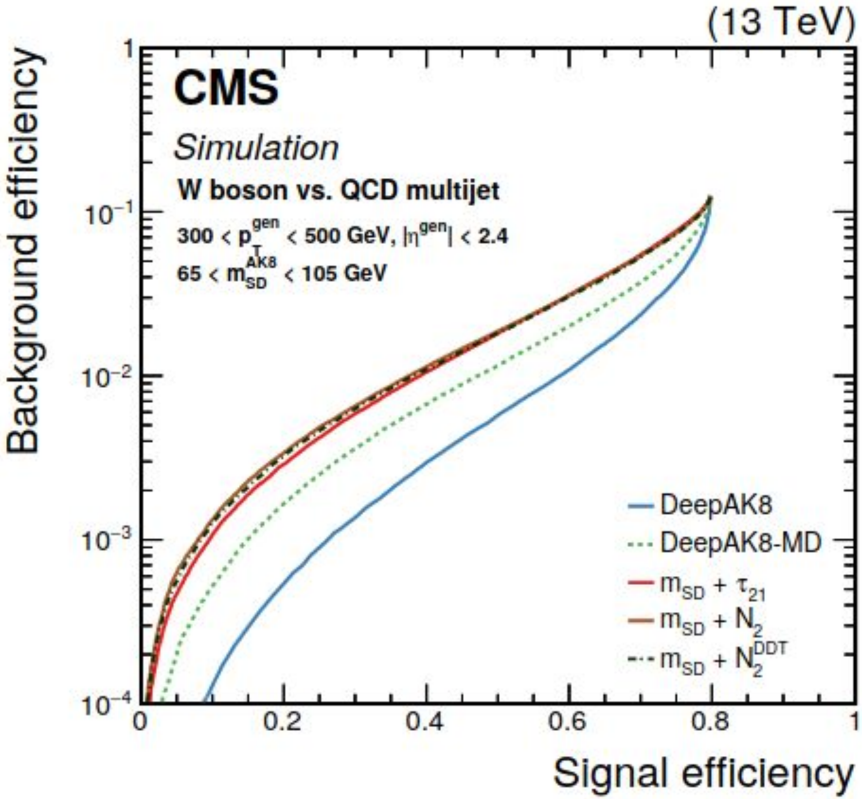
Performance of taggers & scale factors is highly sensitive to modeling of quark- and gluon-initiated jet parton showers. Important to have as many experimental inputs to constrain these uncertainties!!

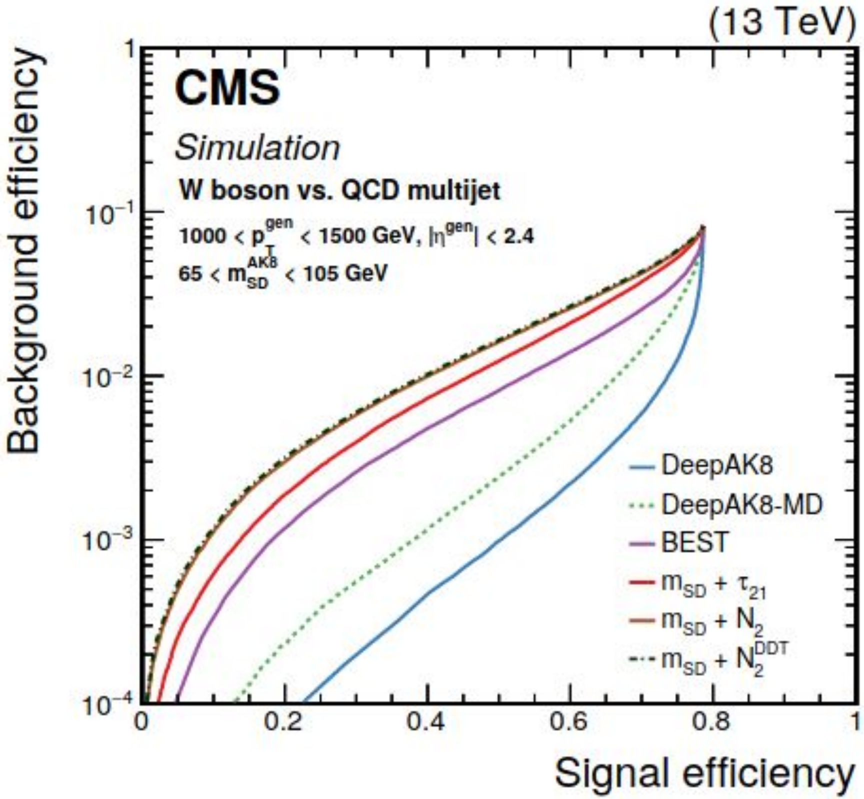
Summary

- Measurements of jet substructure in Z+jet, dijet, and top quark pair production expose building blocks of QCD (splitting function, α_S)

- Valuable input for a better understanding of quark-jet and gluon-jet substructure

- Data also provides additional input for quark-gluon jet discriminators, or to discriminate V/H/t jets from light-quark jets.


(13 TeV)
CMS
Simulation
W boson vs. QCD multijet
300 < p_T^{gen} < 500 GeV, $|\eta^{gen}|$ < 2.4
65 < m_{SD}^{AK8} < 105 GeV
DeepAK8
DeepAK8-MD
$m_{SD} + \tau_{21}$
$m_{SD} + N_2$
$m_{SD} + N_2^{DDT}$
Background efficiency
Signal efficiency


(13 TeV)
CMS
Simulation
W boson vs. QCD multijet
1000 < p_T^{gen} < 1500 GeV, $|\eta^{gen}|$ < 2.4
65 < m_{SD}^{AK8} < 105 GeV
DeepAK8
DeepAK8-MD
BEST
$m_{SD} + \tau_{21}$
$m_{SD} + N_2$
$m_{SD} + N_2^{DDT}$
Background efficiency
Signal efficiency

Algorithm	Subsection	jet p_T [GeV]	t quark	W boson	Z boson	H boson
$m_{SD} + \tau_{32}$	6.1	400	✓			
$m_{SD} + \tau_{32} + b$	6.1	400	✓			
$m_{SD} + \tau_{21}$	6.1	200	✓	✓		
HOTVR	6.2	200	✓			
N_3-BDT (CA15)	6.3	200	✓			
$m_{SD} + N_2$	6.3	200		✓	✓	✓
BEST	6.5	500	✓	✓	✓	✓
ImageTop	6.6	600	✓			
DeepAK8$^{(*)}$	6.7	200	✓	✓	✓	✓
	Jet mass decorrelated algorithms					
$m_{SD} + N_2^{DDT}$	6.3	200		✓	✓	✓
double-b	6.4	300			✓	✓
ImageTop-MD	6.6	600	✓			
DeepAK8-MD$^{(*)}$	6.7	200	✓	✓	✓	✓