Jet Substructure at CMS

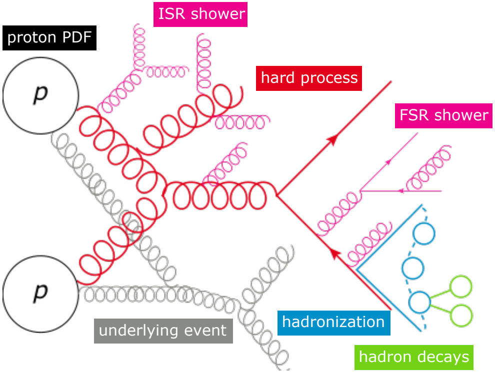
Markus Seidel

June 7, 2021 | LHCP

Measuring jet substructure ~~and correlations in hadronic final states~~[1]

Jet substructure

- Evolution of partonic states to measurable stable particles
- Precise knowledge necessary for:
  - Precision in measurements involving jets (Higgs, top quarks)
  - Flavor tagging, pileup jet ID
- Distinctive features if heavy resonance is contained in a jet
  - Tools for BSM searches with boosted objects

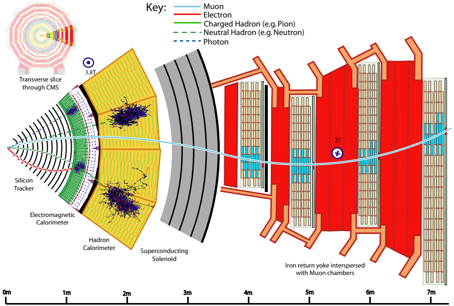


[1] overlap removal

Particle flow reconstruction and grooming

CMS Run 2 default

- PF algorithm keeps all tracks, and removes their energy from calorimeter towers



- Charged hadron subtraction (CHS) removes tracks from PU vertices
- Jets clustered with anti-k_t, $R = 0.4, 0.8$

Optional

- PUPPI algorithm: weigh down neutral clusters not close to PV tracks arXiv:1407.6013
- Iterative soft drop declustering $j_0 \to j_1 + j_2, j_1 \to j_0$, stop if $z_g = p_T(j_2)/p_T(j_0) > 0.1$ arXiv:1402.2657

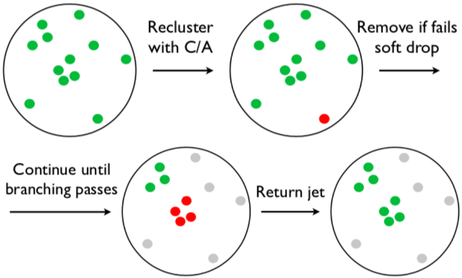
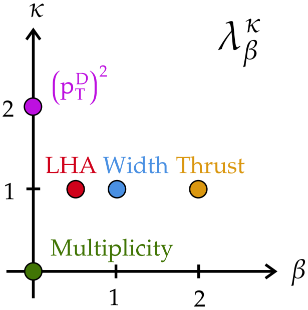


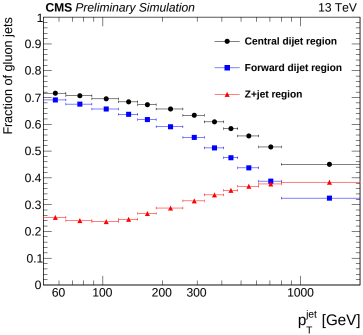
Figure from A. Larkoski (LPC 2014)

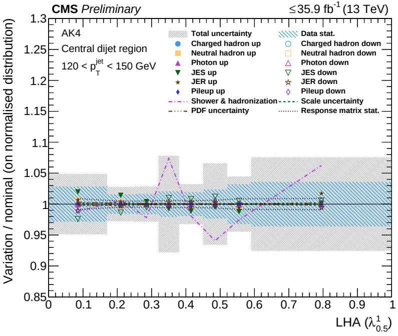
Quark and gluon jet substructure in dijet and Z+jet events

PAS SMP-20-010

- Measure generalized angularities $\lambda_\beta^\kappa = \sum_i z_i^\kappa \left(\frac{\Delta R(i,\hat{n}_r)}{R}\right)^\beta$, $z_i = p_T^i / \sum_i p_T^i$ arXiv:1408.3122
- Dimensions: jet p_T, R=0.4↔0.8, charged+neutral↔charged-only, groomed↔ungroomed, dijet events: gluon-enriched ↔ Z+jet events: quark-enriched





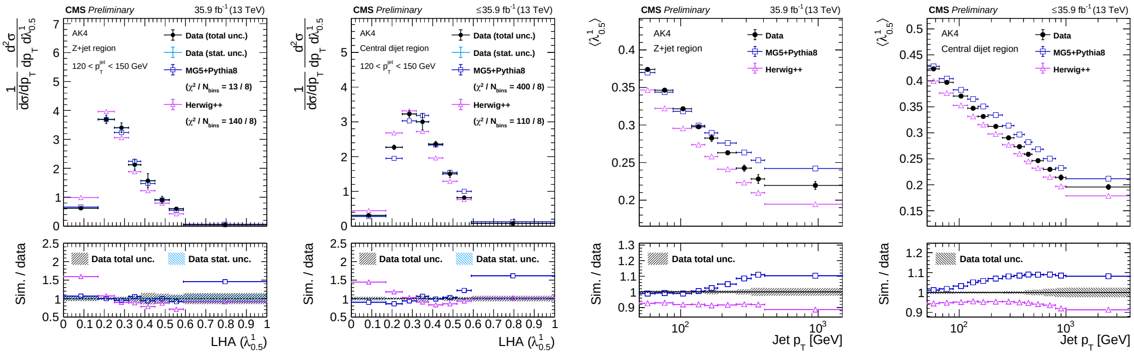


- Largest uncertainties: statistics in Z+jets, shower & hadronization

Quark and gluon jet substructure in dijet and Z+jet events

PAS SMP-20-010

- Unfolded data to the particle level and compared to MC

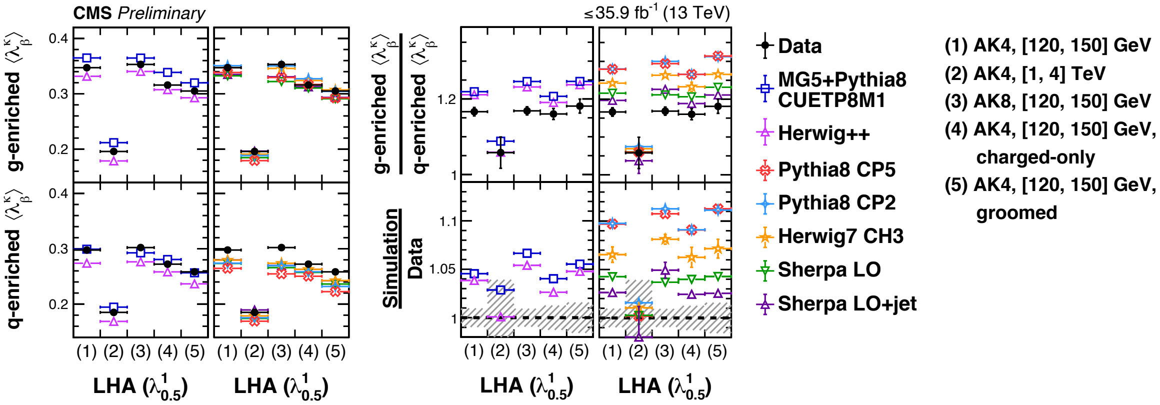


- Left: LHA distribution for $120 < p_T < 150$ GeV, for Z+jet and dijet
- Right: $\langle$LHA$\rangle$ vs. jet p_T, for Z+jet and dijet
- Jets in Z+jet (quark-enriched) narrower than in dijet events (gluon-enriched)
- MG+Pythia8 and Herwig++ bracket the data

Quark and gluon jet substructure in dijet and Z+jet events

PAS SMP-20-010

- Mean values and ratios in different p_T regions, jet radius, jet constituents, grooming

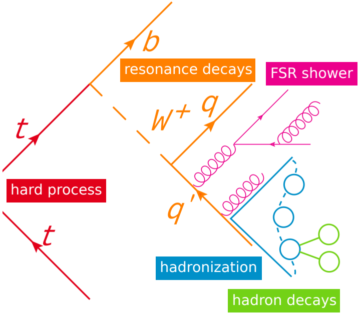


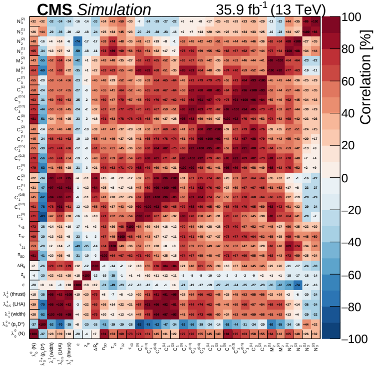
- Newer MC tunes: improved gluons but quark data described less well
- Little impact in ratios from jet radius, jet constituents, grooming
  $\rightarrow$ insensitive to soft radiation, computable with better precision
- All MC tunes/generators overestimate quark/gluon separation

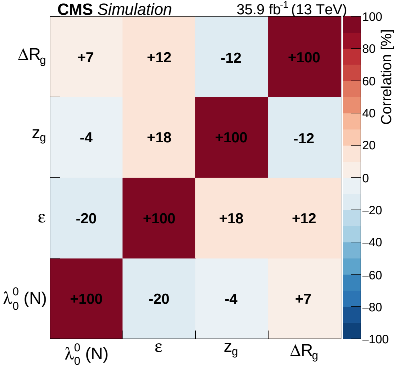
Measurement of jet substructure in $t\bar{t}$ lepton+jets

TOP-17-013 RIVET

- $t\bar{t}$ as standard candle: provides bottom, light-enriched and gluon-enriched jet samples
  - Bottom: b-tagged (ghost tagging at particle level)
  - Light-enriched: non b-tagged jets with $|m_{jj} - m_W| < 15$ GeV
  - Gluon-enriched: non b-tagged jets with $|m_{jj} - m_W| > 15$ GeV



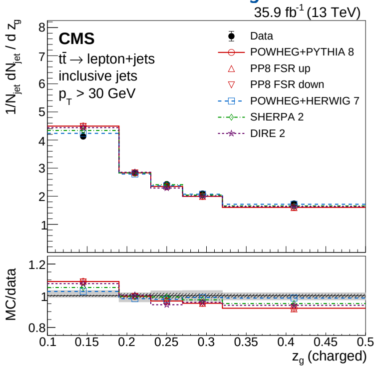


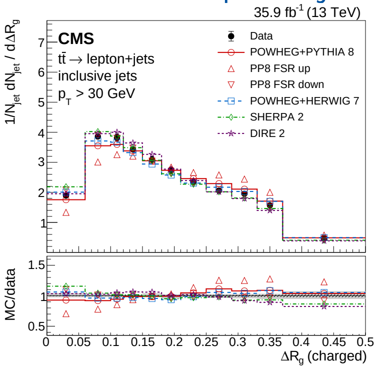


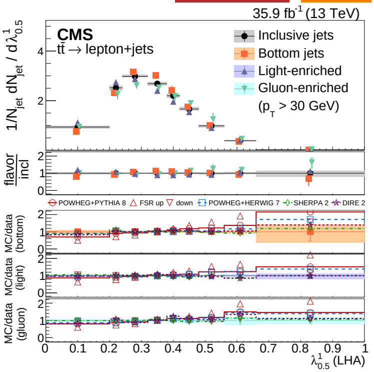
- Observables: angularities, soft drop, N-subjettiness, energy correlations (33 in total)
- Many observables correlated → find set of 4 low-correlation observables

Measurement of jet substructure in $t\bar{t}$ lepton+jets

TOP-17-013 RIVET



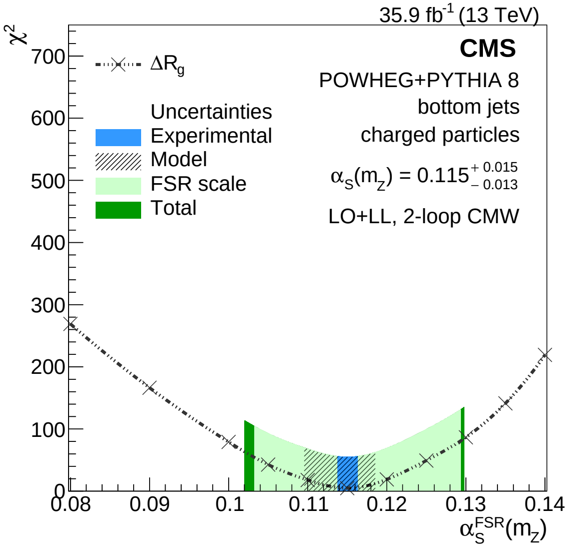




- Groomed momentum fraction z_g
  - Related to QCD splitting function, independent of α_S, best described by Herwig
- Angle between groomed subjets ΔR_g
  - Clearly disfavors high α_S in Pythia 8 FSR up variation
  - Very narrow b jets in early version of Dire nLL shower
    $\rightarrow$ inclusion of missing $b \rightarrow bg$ splitting functions gives good agreement
- LHA: Gluon-enriched jets > bottom jets > light-quark jets

Measurement of jet substructure in $t\bar{t}$ lepton+jets

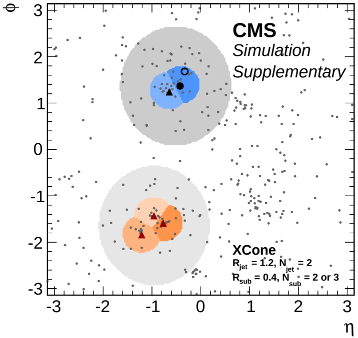
- Extraction of strong coupling α_S
- ΔR_g expected least sensitive to NP effects
- B jets most suitable: Pythia ME corrections for top decays at NLO
  $\rightarrow$ LO+LL for b jet substructure
- Higher-order corrections to soft gluon emissions incorporated in an effective way using 2-loop running and CMW rescaling of Λ_{QCD} Nucl. Phys. B
- Result: $\alpha_S = 0.115 \pm 0.001\ (\text{exp})^{+0.006}_{-0.003}\ (\text{model})^{+0.014}_{-0.012}\ (\text{scale})$
- Experimental uncertainties competitive, would need more precise prediction (top decay at NNLO)

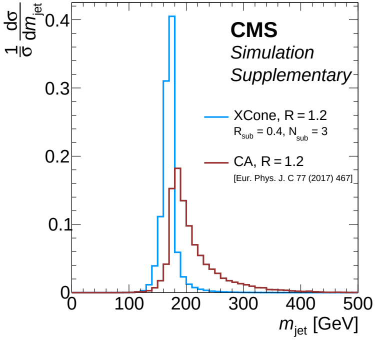


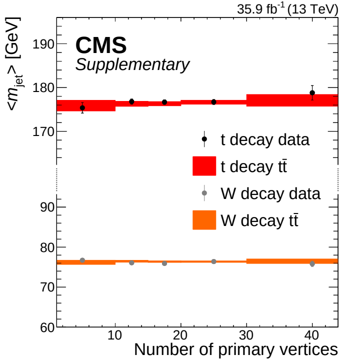
XCone jet mass distribution in boosted top events

TOP-19-005 RIVET

- Select boosted $t\bar{t} \rightarrow$lepton+jets events, reconstruct with XCone algorithm
- Finds exactly 2 large jets and 2-3 subjets, matching event signature



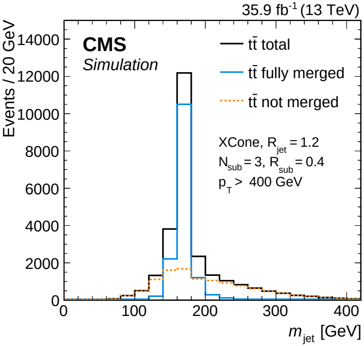


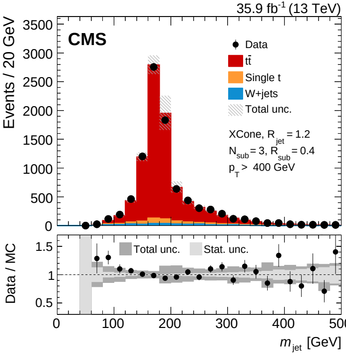


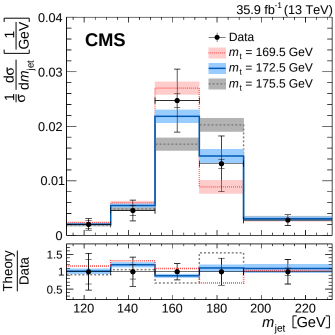
- XCone mass resolution far superior wrt CA jets (8 TeV publication)
- Reconstructed mass stable vs. pileup

XCone jet mass distribution in boosted top events

TOP-19-005 RIVET





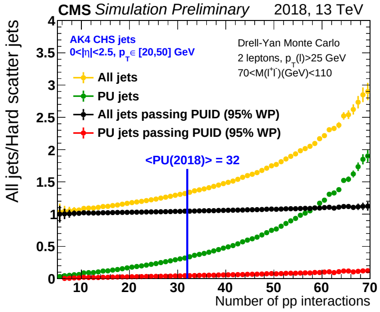


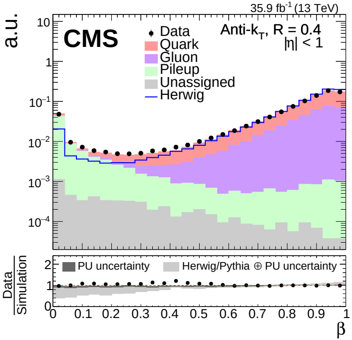
- Mass peak widened by unmerged $t\bar{t}$, good agreement with data
- Unfolded to the particle level to be compared to future SCET calculations in boosted regime with m_t in well-defined mass scheme 1708.02586 arXiv:1803.02321 2012.12304
- With Pythia prediction: $m_t = 172.6 \pm 1.6\ (\text{exp}) \pm 1.6\ (\text{model}) \pm 1.0\ (\text{theory})$ GeV

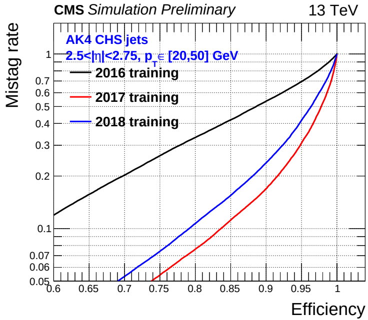
Substructure as a tool: pileup jet ID

DP-20-020 JME-18-001

- At high luminosity: growing number of PU jets from overlapping low-p_T PU interactions
- BDT with 15 input variables: substructure-based and track/vertex-based (LV fraction β)





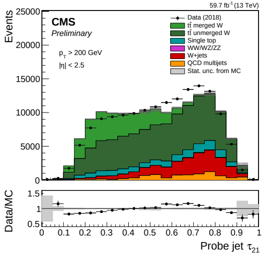


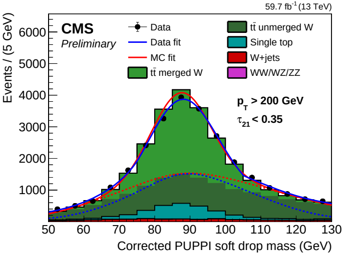
- 2017: pixel detector upgrade extends tracking coverage from $|\eta| < 2.5$ to $|\eta| < 2.7$
- Improved stability of jet multiplicity vs. pileup

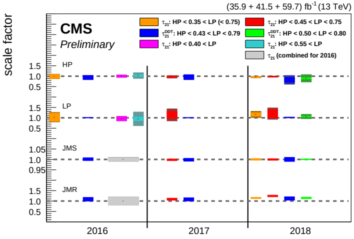
Boosted W & top tagging using N-subjettiness ratio

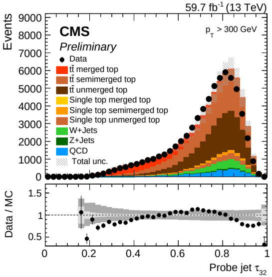
DP-20-025

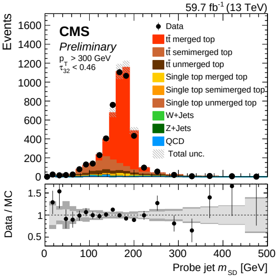
- τ_{NM} distinguishes jets with $N \leftrightarrow M$ subjets: τ_{21} for W tagging, τ_{32} for top tagging

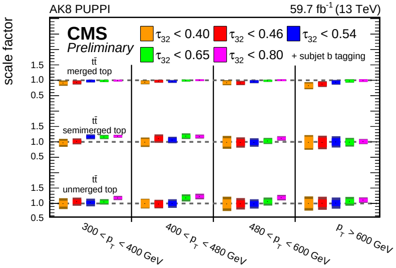










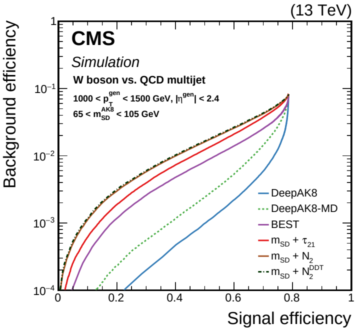


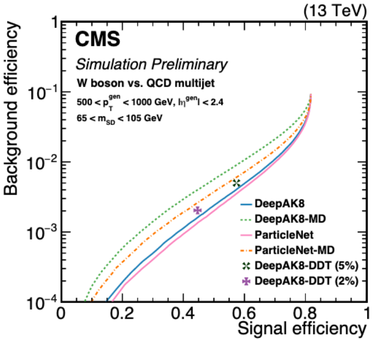
- Determined data/MC scale factors via tag & probe

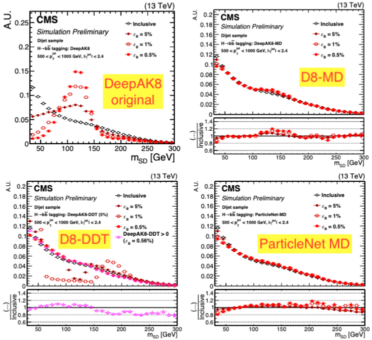
Boosted object tagging using ML algorithms

DP-20-002

- Improve boosted resonance tagging by feeding jet constituents into neural networks
- DeepAK8 (Convolutional NN) JME-18-002 ParticleNet (Graph NN) arXiv:1902.08570







- NNs shape background mass distribution → employ different decorrelation methods
  - DeepAK8-DDT: transform NN output to get flat background efficiency arXiv:1603.00027
  - DeepAK8-MD: training jets reweighted to be flat in p_T and m_{SD} + use adversarial network
  - ParticleNet-MD: train on $X \to jj$ sample with flat m_X + jets reweighted to be flat in p_T, m_{SD}

Summary

Jet substructure measurements

- Several measurements of jet substructure in dijet, Z+jet, $t\bar{t}$
- Sensitive to parton shower models, strong coupling, and top mass
- RIVET implementations available or in progress

Jet substructure as a tool

- Jet substructure used for PU jet ID and identifying heavy resonances
- Improved performance by machine learning
- For applications, please see BSM sessions :-)

More precision QCD by CMS

- today, 17:18: "Nucleon Structure and Soft QCD from CMS" by Rajat Gupta
- Thursday, 14:33: "Precision QCD Measurements from CMS" by Salim Cerci