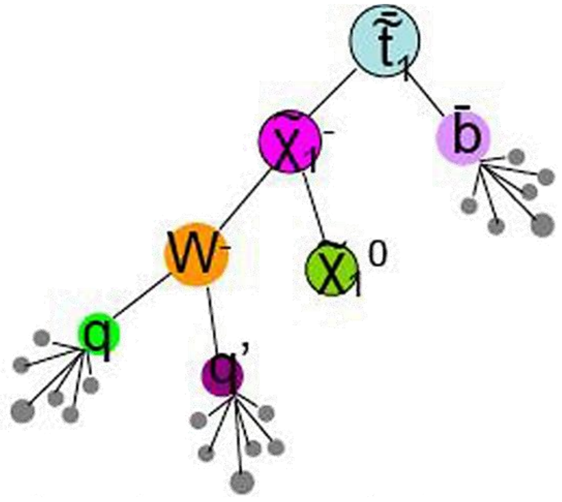
Fat Jets

Nick Amin

09/12/14

Overview/Motivation

- **Study jet substructure for “fat” jets to learn about composition in cases where jets are not fully resolved**
- In stop squark decay to hadronic products the 2 quarks from the W can result in merged jets if the W is highly boosted
  - How often does this occur (as a function of stop mass)?
- Study jet substructure variables
  - Explore jet “toolbox”
  - Can original constituents of overlapping jets be identified using substructure variables?

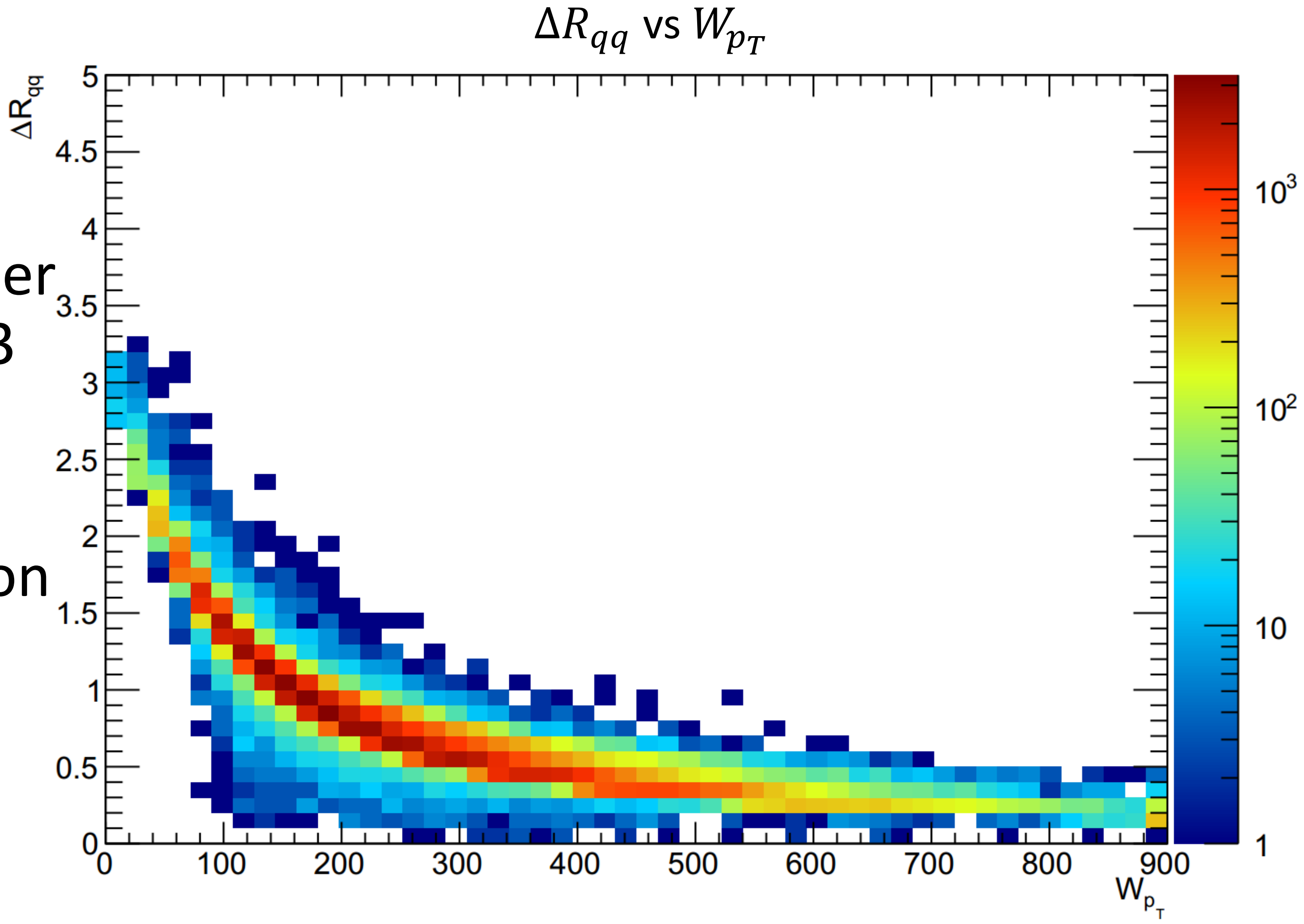


Samples Used

- For **rate studies** (first set of slides), used centrally-produced T2tt (https://twiki.cern.ch/twiki/bin/viewauth/CMS/SMST2ttMadgraph8TeV)
  - Stop mass in [200,900] (increments of 100 GeV)
  - LSP mass of 1 GeV
  - AK5 jets
- For **substructure studies** (second set of slides), used privately-produced T2tt with $m_{stop} = 850$, $m_{LSP} = 100$ to gain access to AK8 jets

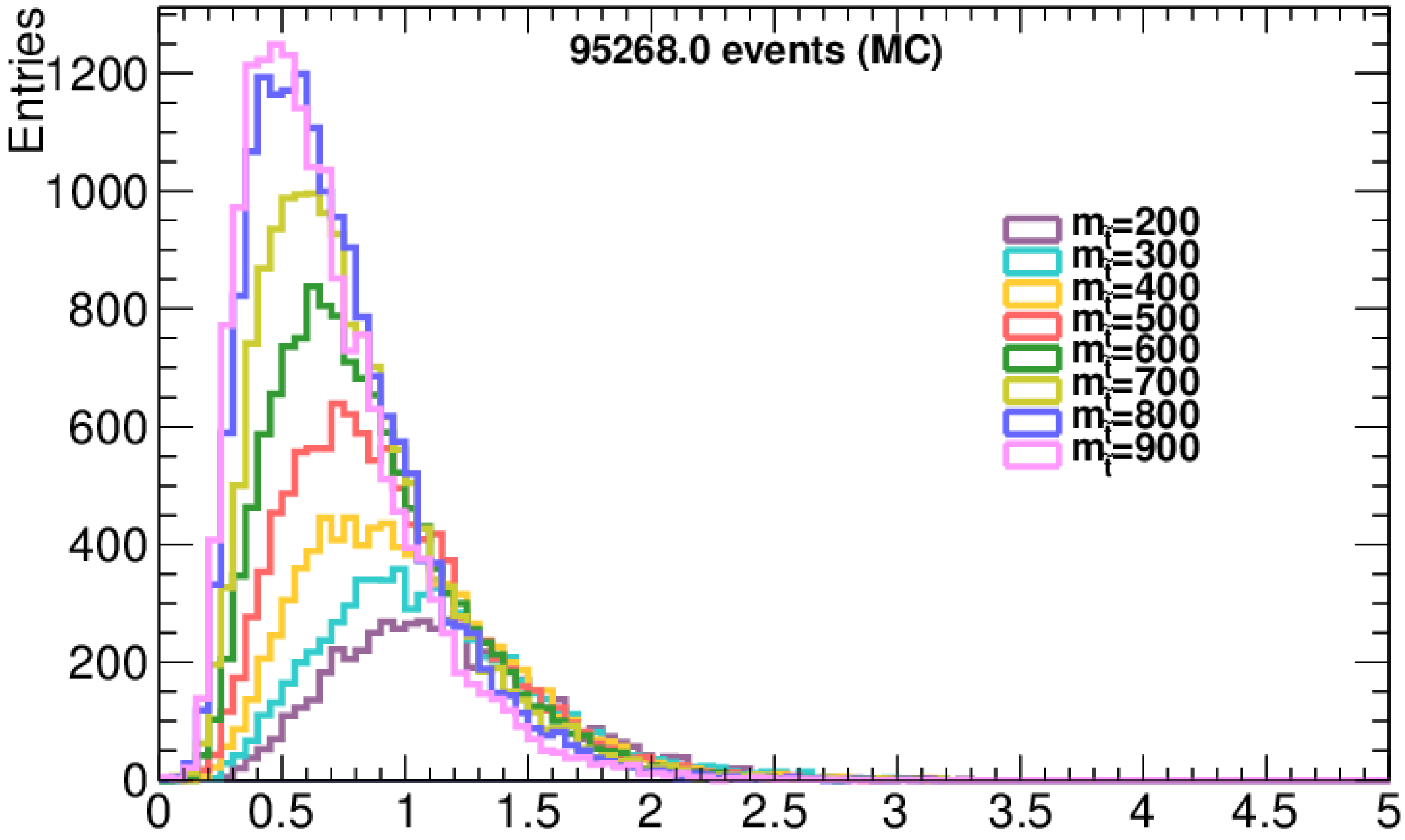
Quark separation

- Consider T2tt events where 1 W decays leptonically and the other decays hadronically (2/3 of the time)
- We see that boosted W→qq means separation decreases (a lot)

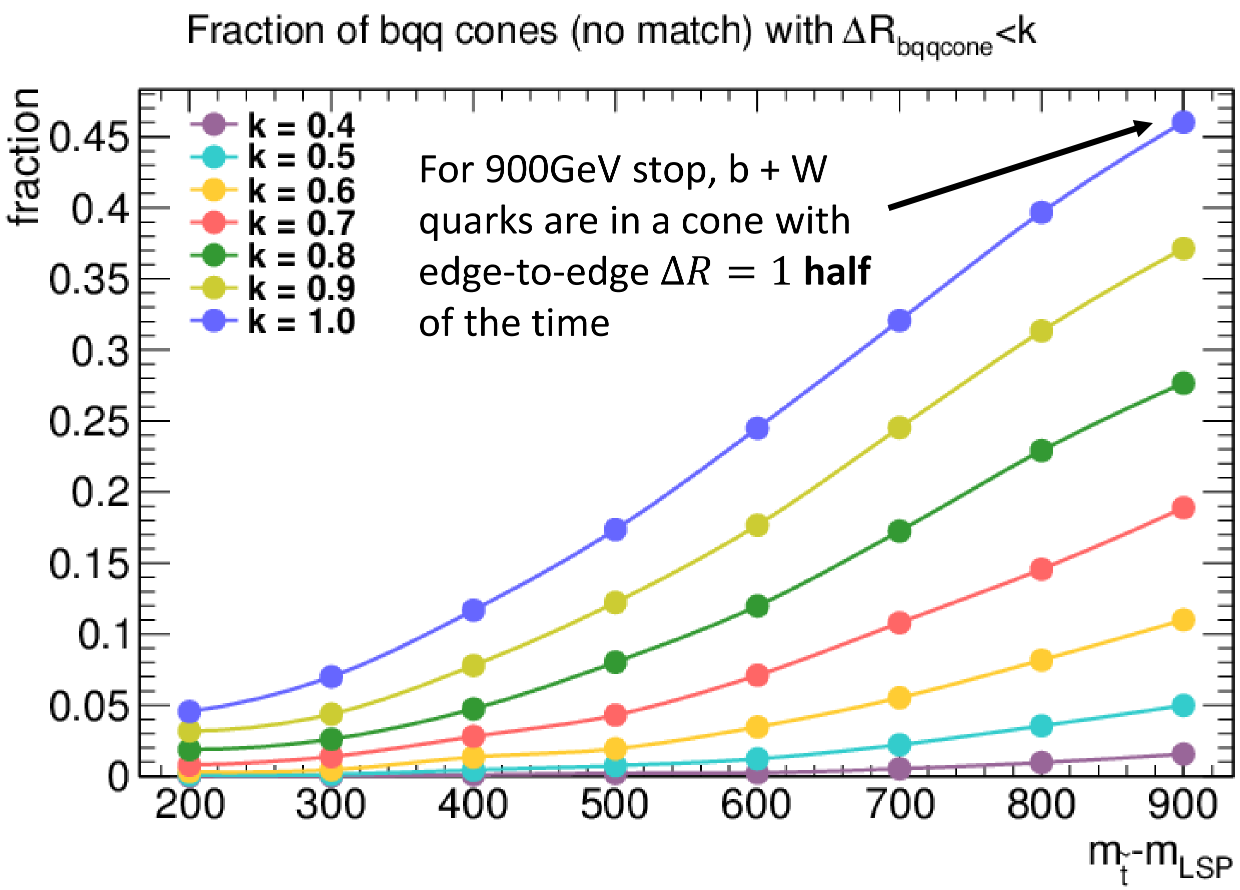


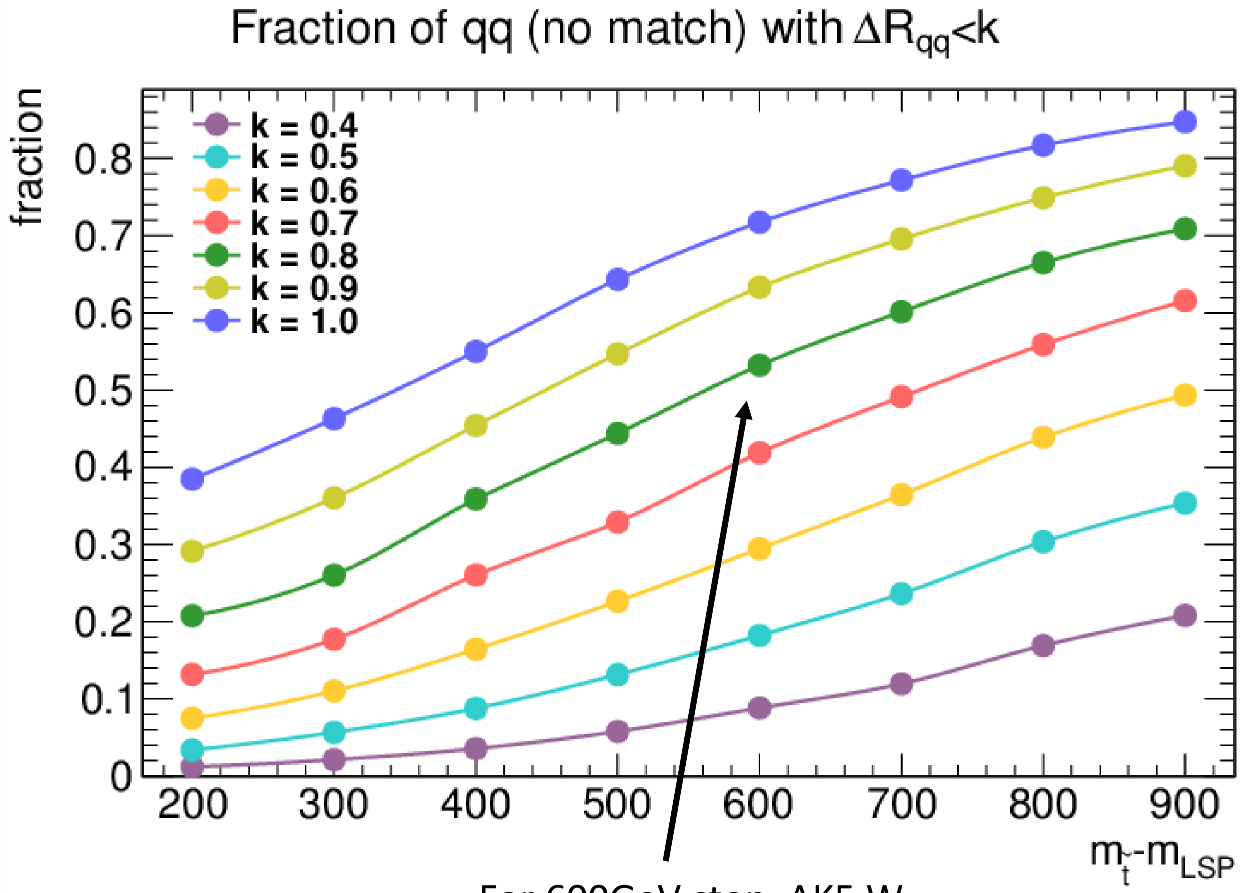
Quark separation

ΔR_{qq}



How common is the overlap?



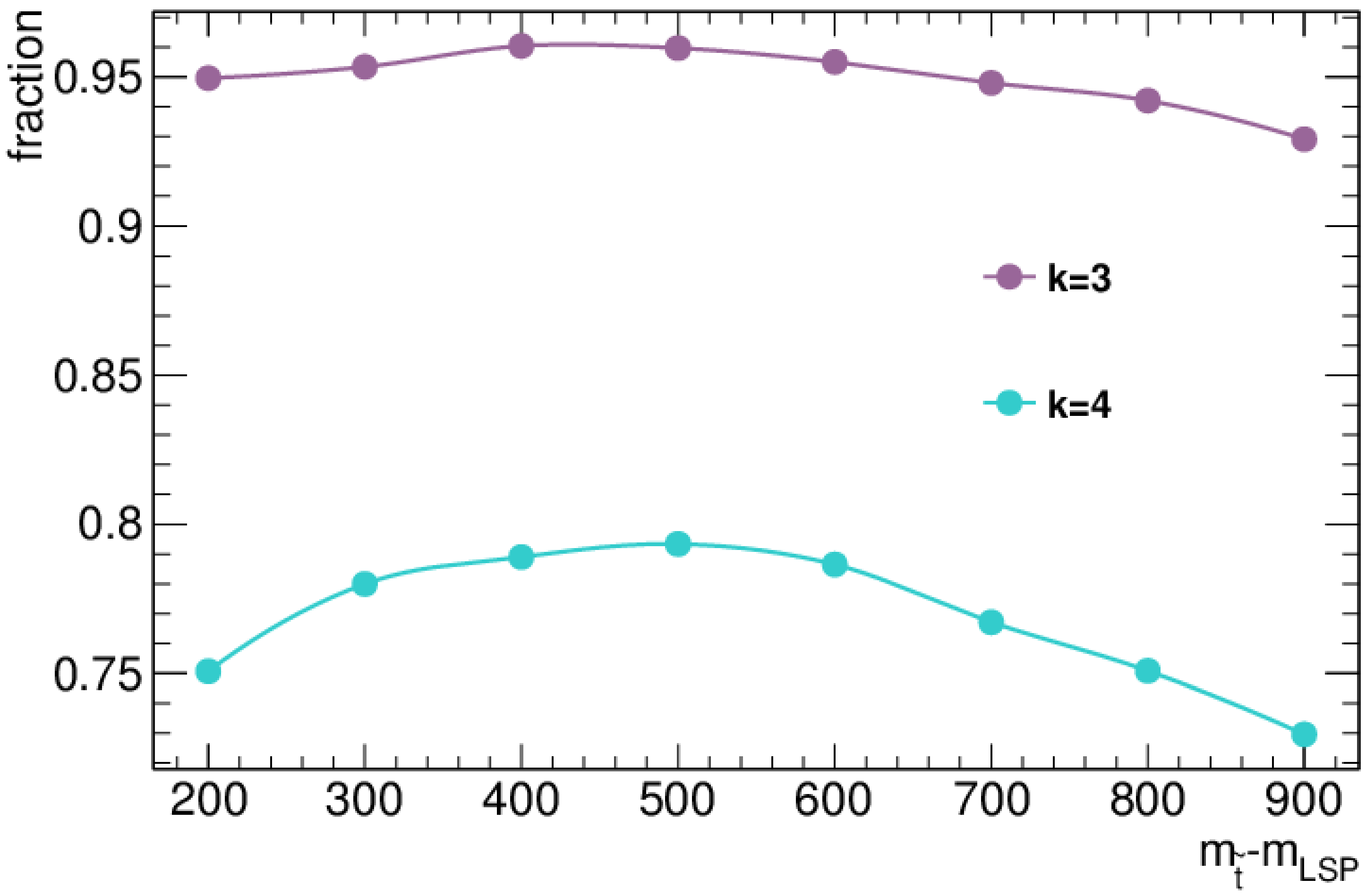


For 600GeV stop, AK5 W quark jets will touch **half** of the time

Njet impact

- As mass difference increases, Njets should decrease as more jets get merged
- Looks like ISR/FSR mostly compensates for this effect

Fraction of events with Nrecojets ≥ k

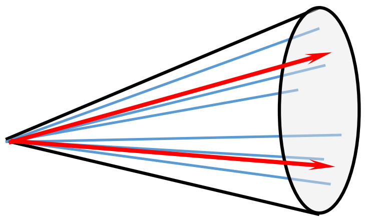


Jet Toolbox

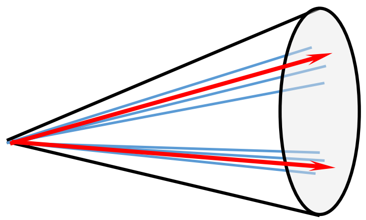
- Jet toolbox information at https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetToolbox
  - Takes miniAOD (or AODSIM) and adds various substructure userFloats
  - ~0.3 Hz for miniAOD→miniAOD+substructure
- For AK8, it includes:
  - Q-jets Volatility
  - N-subjettiness (τ_1, τ_2, τ_3)
  - Masses of pruned, trimmed, filtered jets
  - Top-tagged jet mass
    - If an ungroomed jet is matched to top-tagged jet, this branch is filled with the jet mass
- AK4 has additional user floats detailed on Twiki (pileup jet ID, quark/gluon tagging)
- CA8/CA4 counterparts for these values are calculated as well

N-subjettiness

- Given *N* subjet axes in a fat jet, N-subjettiness, τ_N, is given by
  - $\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}\}$
    - $\Delta R_{N,k}$ is the angular separation between constituent *k* and candidate subjet *N*
    - d_0 is a normalization factor given by $d_0 = \sum_k p_{T,k} R_0$, so $0 \leq \tau_N \leq 1$
    - $R_0 = 0.8$ for AK8 clustering



High τ_2 (consitutents spread out)



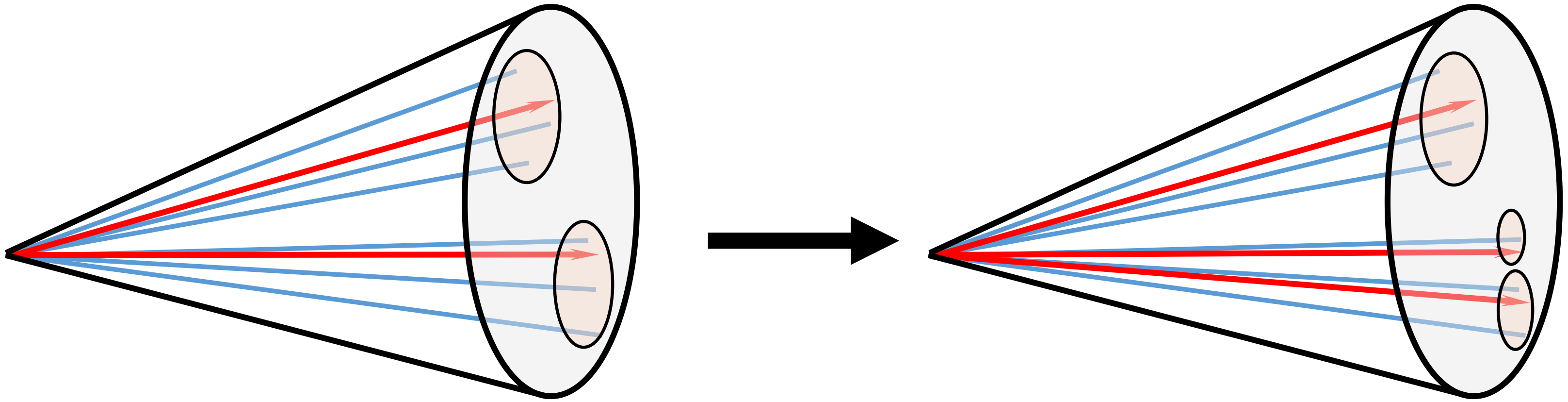
Low τ_2 (constituents close to subjet axes)

Clusters with exactly *N* subjets will have small τ_N

If $\tau_N \approx 1$, cluster most likely has more than *N* subjets

N-subjettiness

- When comparing N-subjettiness with (N+1)-subjettiness, adding extra candidate subjet allows minimum angular separation to decrease
- Thus, in most cases, $\tau_1 > \tau_2 > \tau_3$.
- Ratio $\tau_{NM} = \tau_N / \tau_M$, where *N*=*M*+*1*, is useful



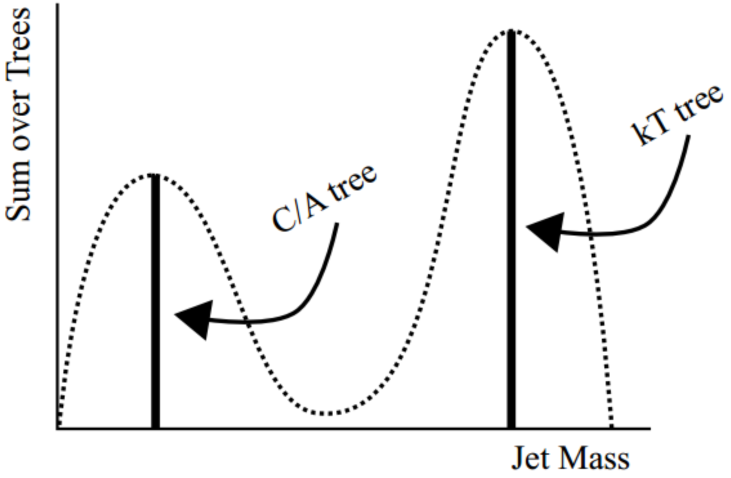
$\boldsymbol{\tau_2} \quad > \quad \boldsymbol{\tau_3}$

Clusters with small τ_{NM} most likely have *N* subjets.

We can use this quantity for cuts.

Q-jets Volatility

- Associate clustering history ("tree") with jets
  - Depending on the route (AK8-like, CA8-like, etc.) taken to get a jet, the mass can change, so you get a mass **distribution** for a jet when considering the space of possible trees
- For a pruned jet mass distribution, define **volatility** as $V = \frac{\sqrt{\langle m^2 \rangle - \langle m \rangle^2}}{\langle m \rangle}$
  - Physically, dependence of jet mass on clustering method governs volatility (large mass fluctuations over different algorithms/routes means the jet is **volatile**/fuzzy)



A. Hornig

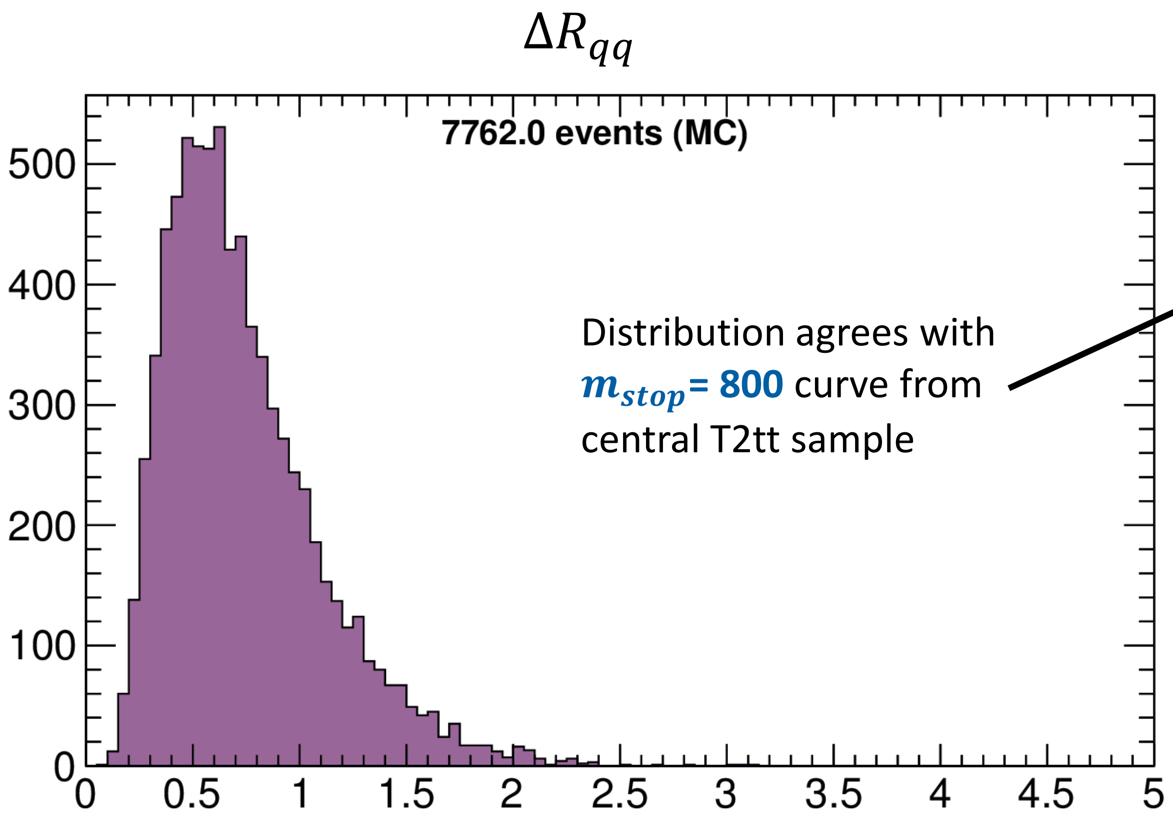
Other variables

- Groomed masses
  - **Pruned**
  - Trimmed
  - Filtered
  - These have essentially the same performance
- https://indico.cern.ch/event/334796/contribution/0/material/slides/0.pdf
  - Slide 37 shows efficacy of various variables (by adding them one-by-one to a BDT)
  - “Alternative” to n-subjettiness is the Energy Correlation Function (slide 26)
    - Gets number of subjets using constituent p_T and pairwise angular displacements without reclustering the jets
    - Decent variable, but not included in toolbox
- **Pruned mass** and **n-subjettiness** have good discriminative power, so I will focus on them

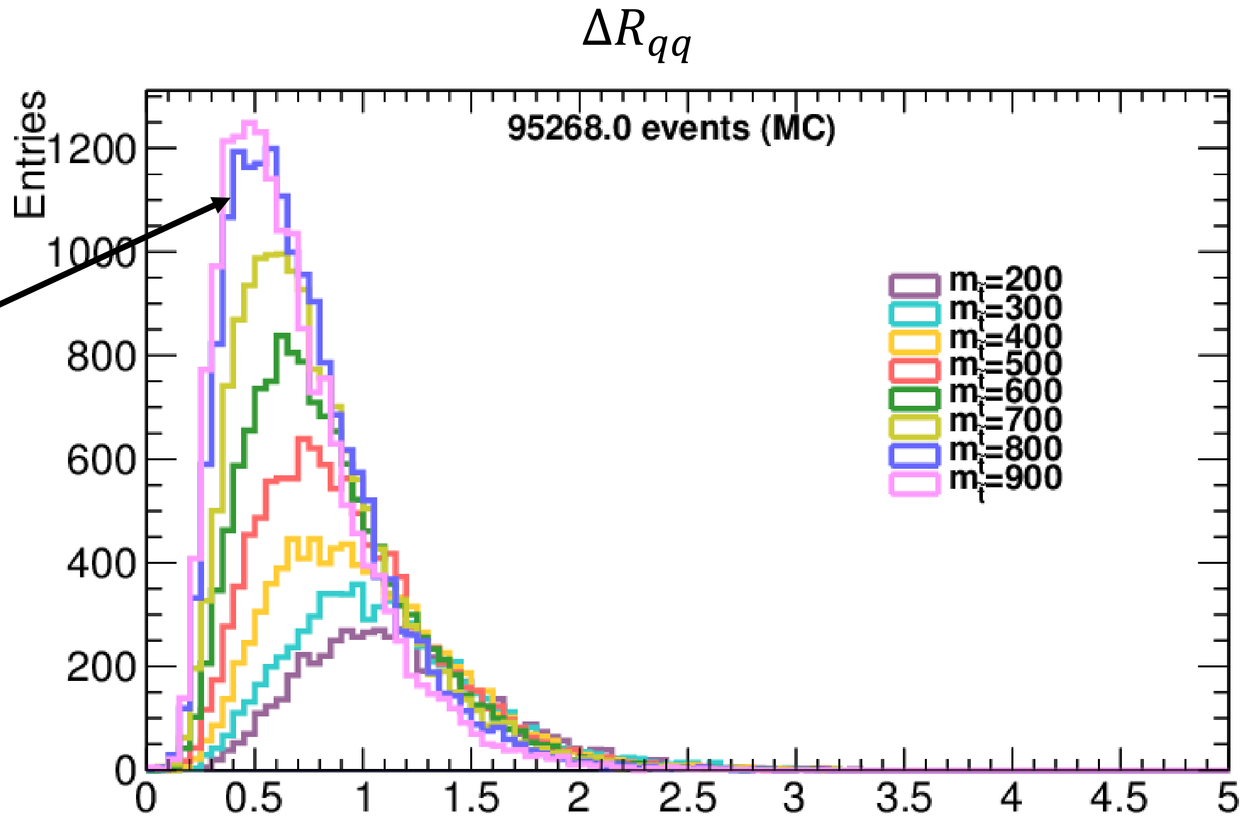
Setup

- 10k-event T2tt sample (m_{stop} = 850, m_{LSP} = 100)
  - Feed “slimmedJetsAK8” jet collection into the jet toolbox
  - Toolbox spits out userFloats mentioned on Twiki
- Added “SubJetMaker” to CMS3 makers, which stores these userFloat substructure variables from toolboxed miniAOD
- Use AK8 jets with $p_T > 150$

Quark separation

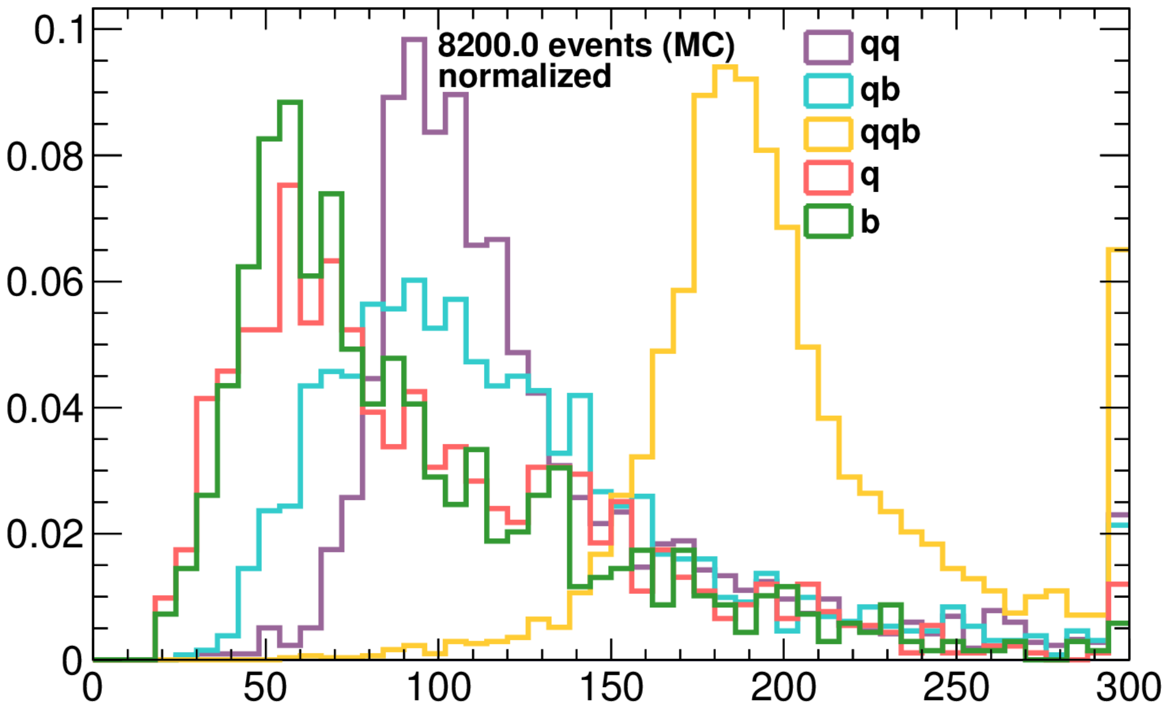


Distribution agrees with m_{stop}= **800** curve from central T2tt sample



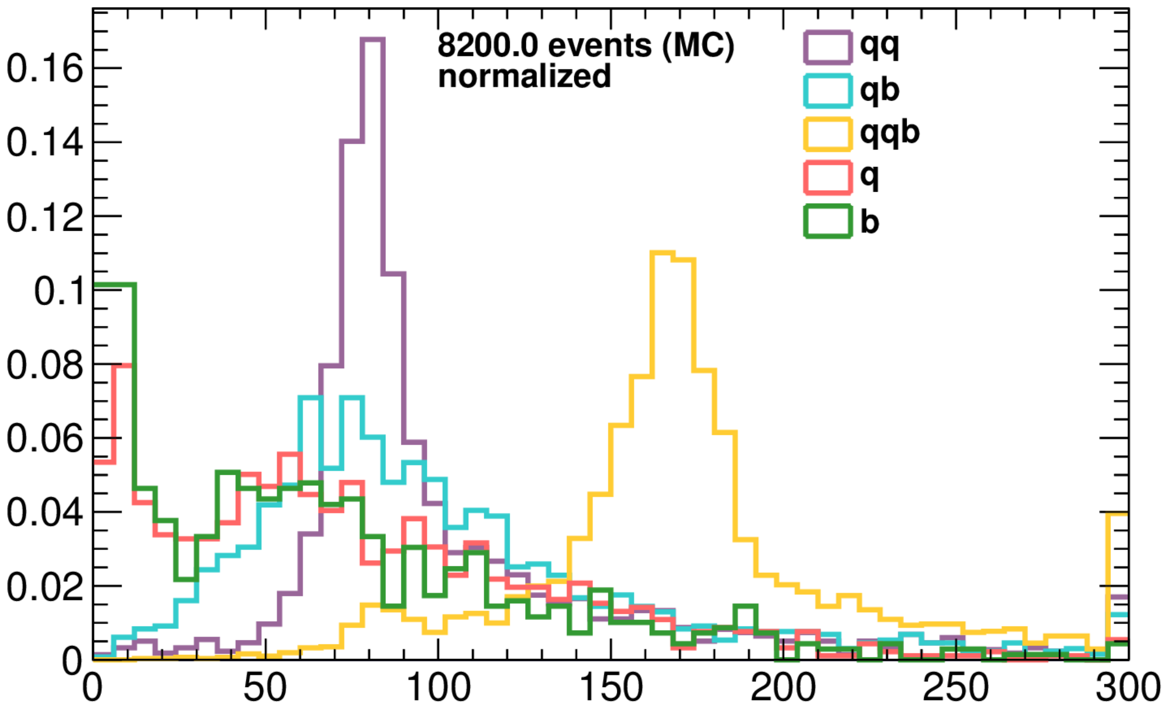
Jet matching and pruning

jet mass



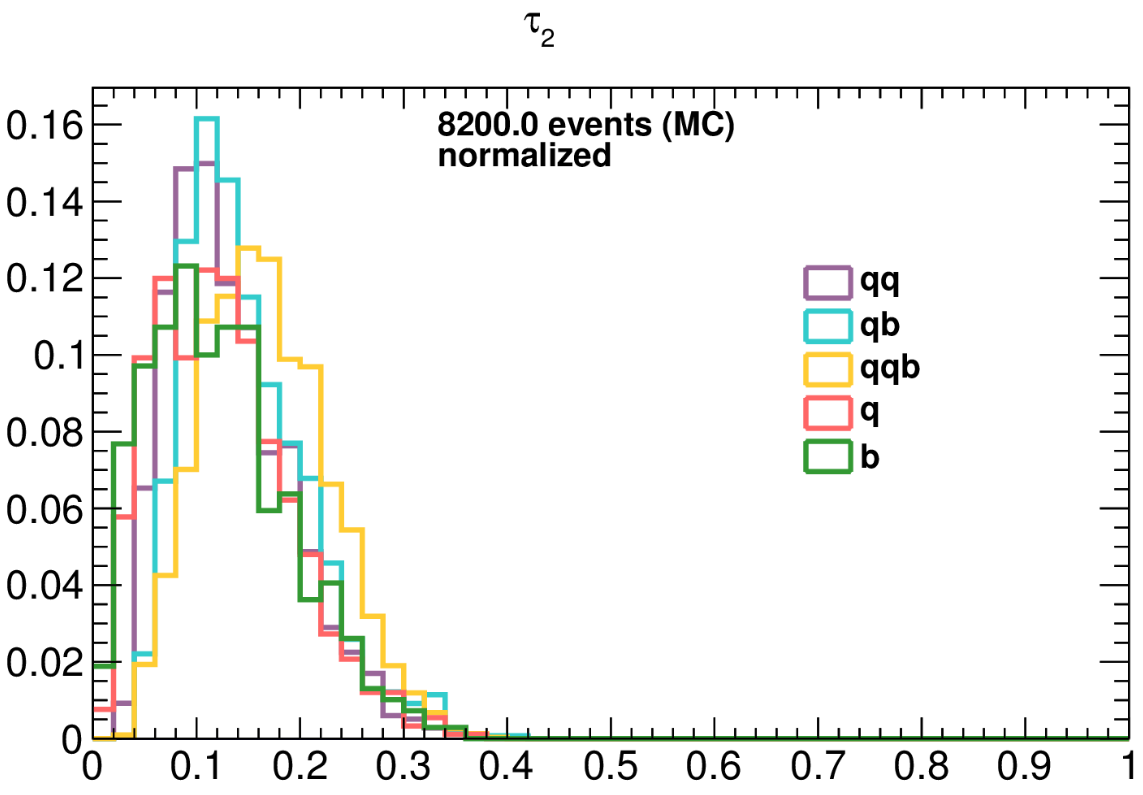
Link jets to gen partons (q's from W and/or b), so a jet can be ΔR matched to 1, 2, or 3 quarks.

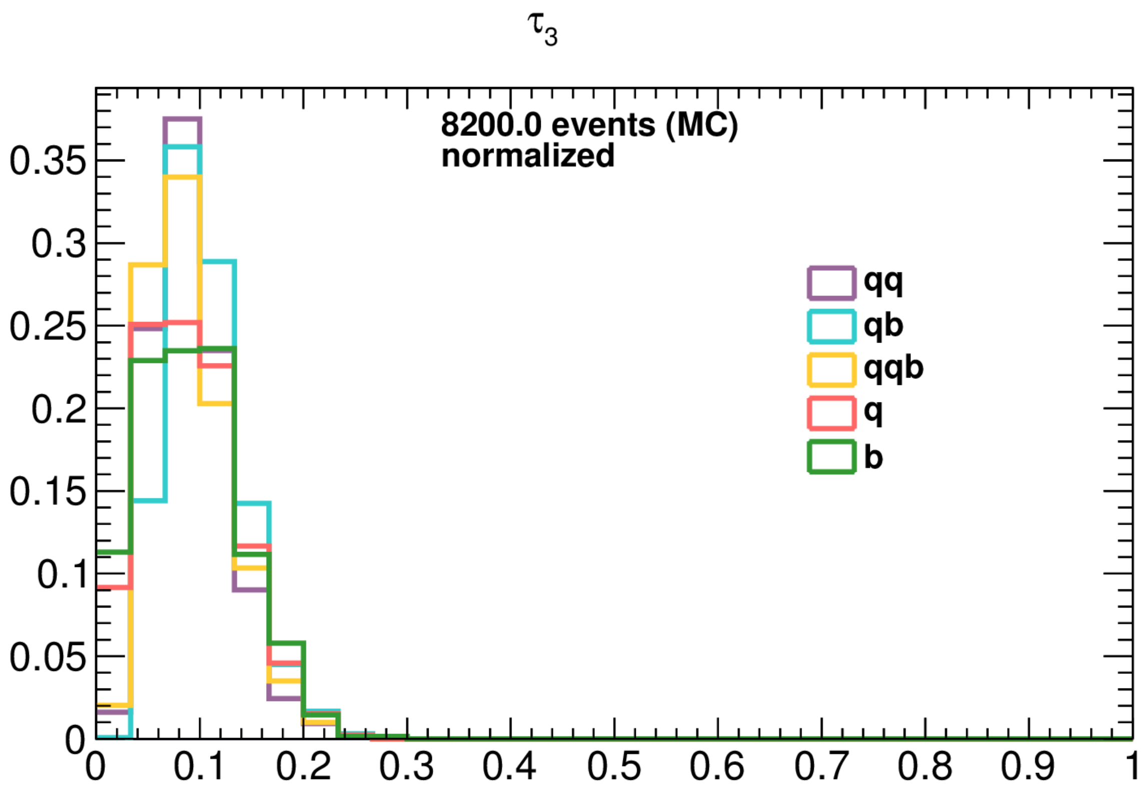
pruned jet mass



Using pruned masses gives more defined W, top peaks, as well as shifting peaks down to correct m_W, m_t values

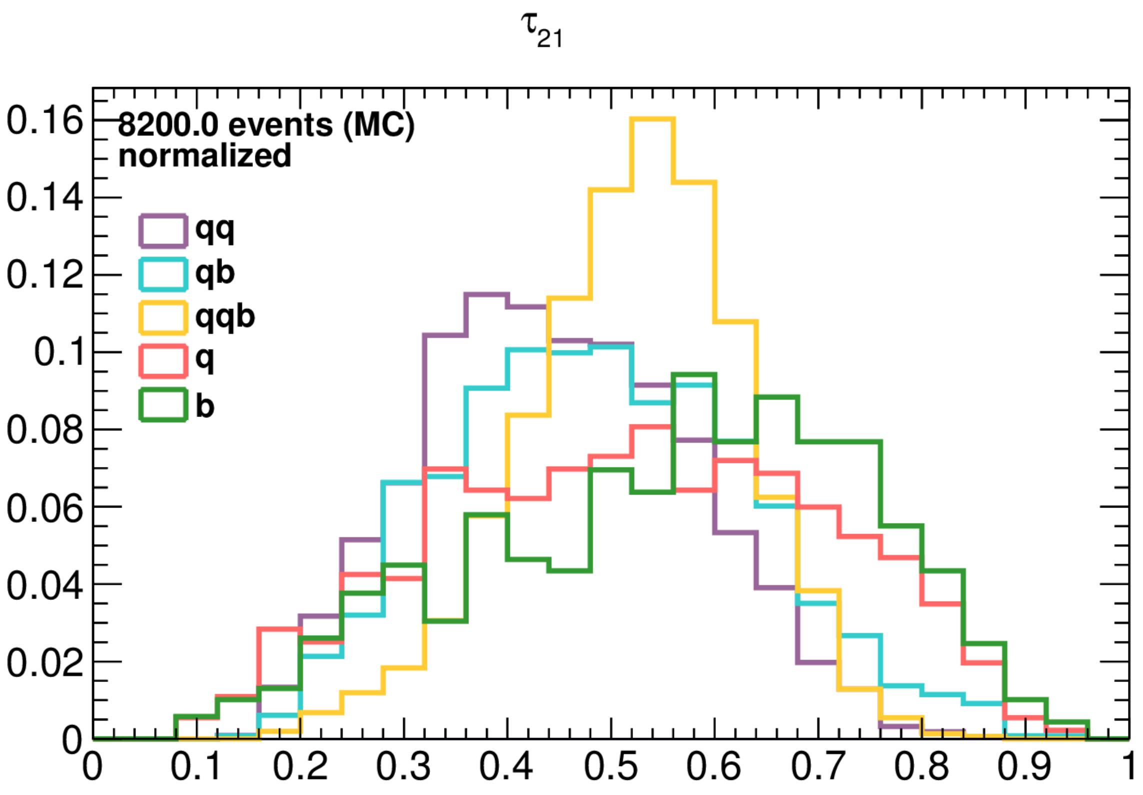
N-subjettiness

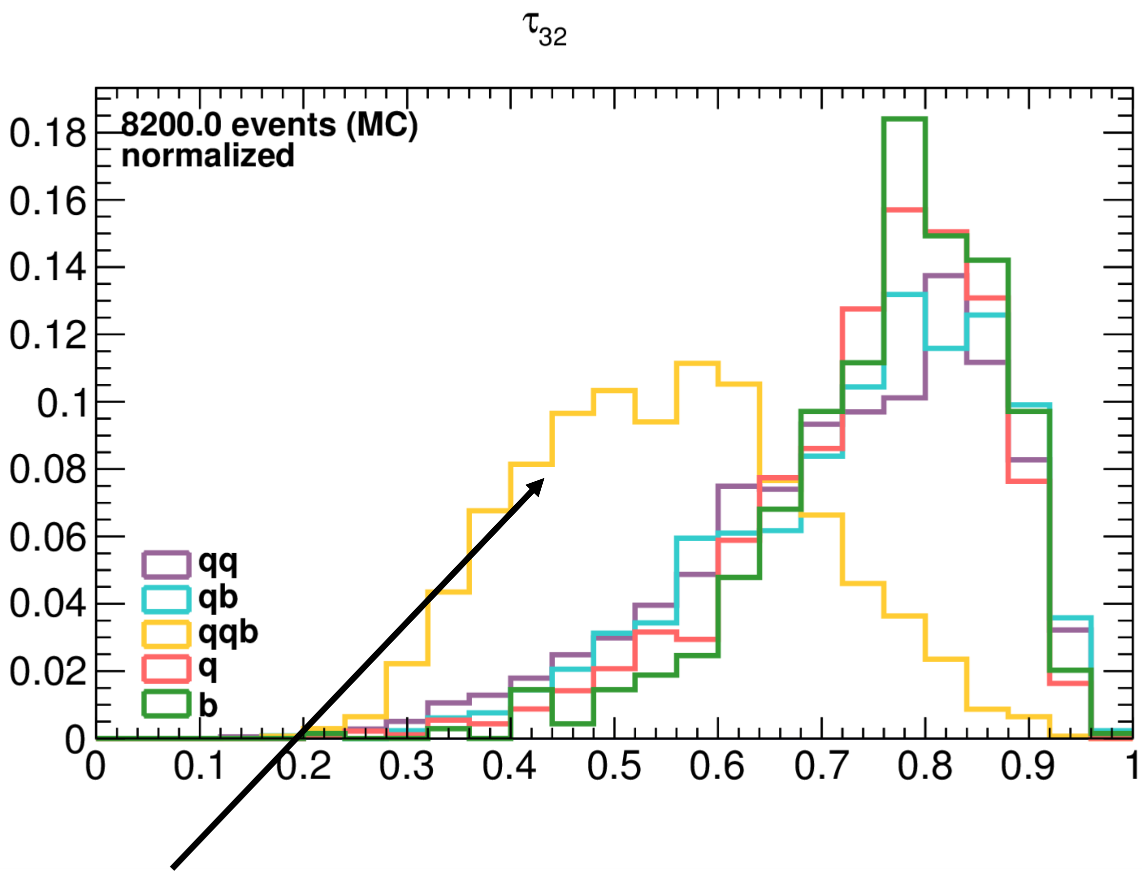




By themselves, these variables have pretty useless discriminative qualities

τ_{NM}





However, the ratio is useful. τ_{32} values for qqb curve are smaller than the rest. This indicates that qqb-matched jets most likely have 3 subcomponents.

Conclusions

- Using AK8 to calculate substructure parameters, we can then use n-subjettiness ratios to discriminate between 3-jets and 2- or 1-jets
- Pruned jet masses can be considered as well to determine if a jet is consistent with a top jet, W jet, etc.
- Calculating these parameters (in the toolbox) seems to be quite slow