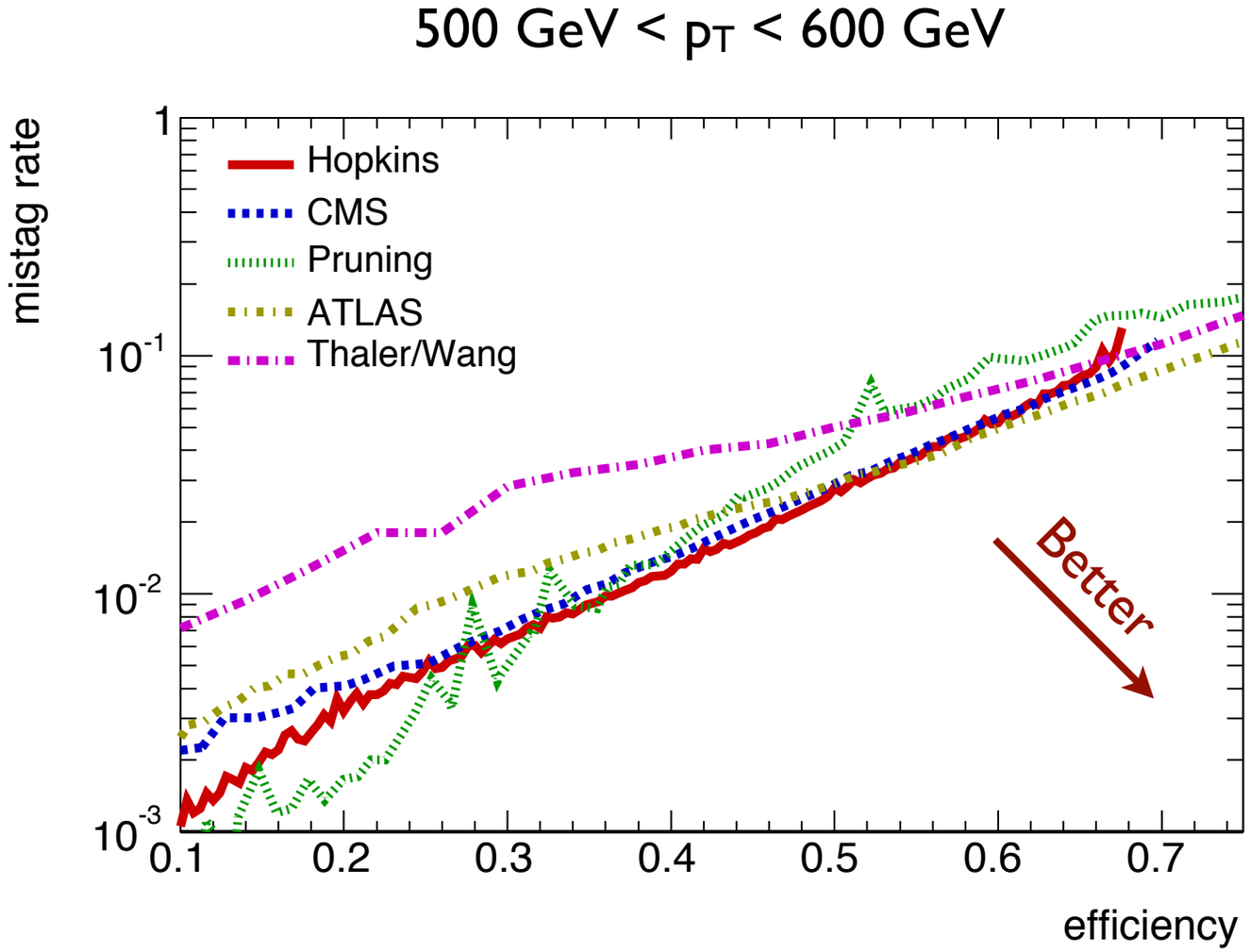
N-subjettiness

Jesse Thaler

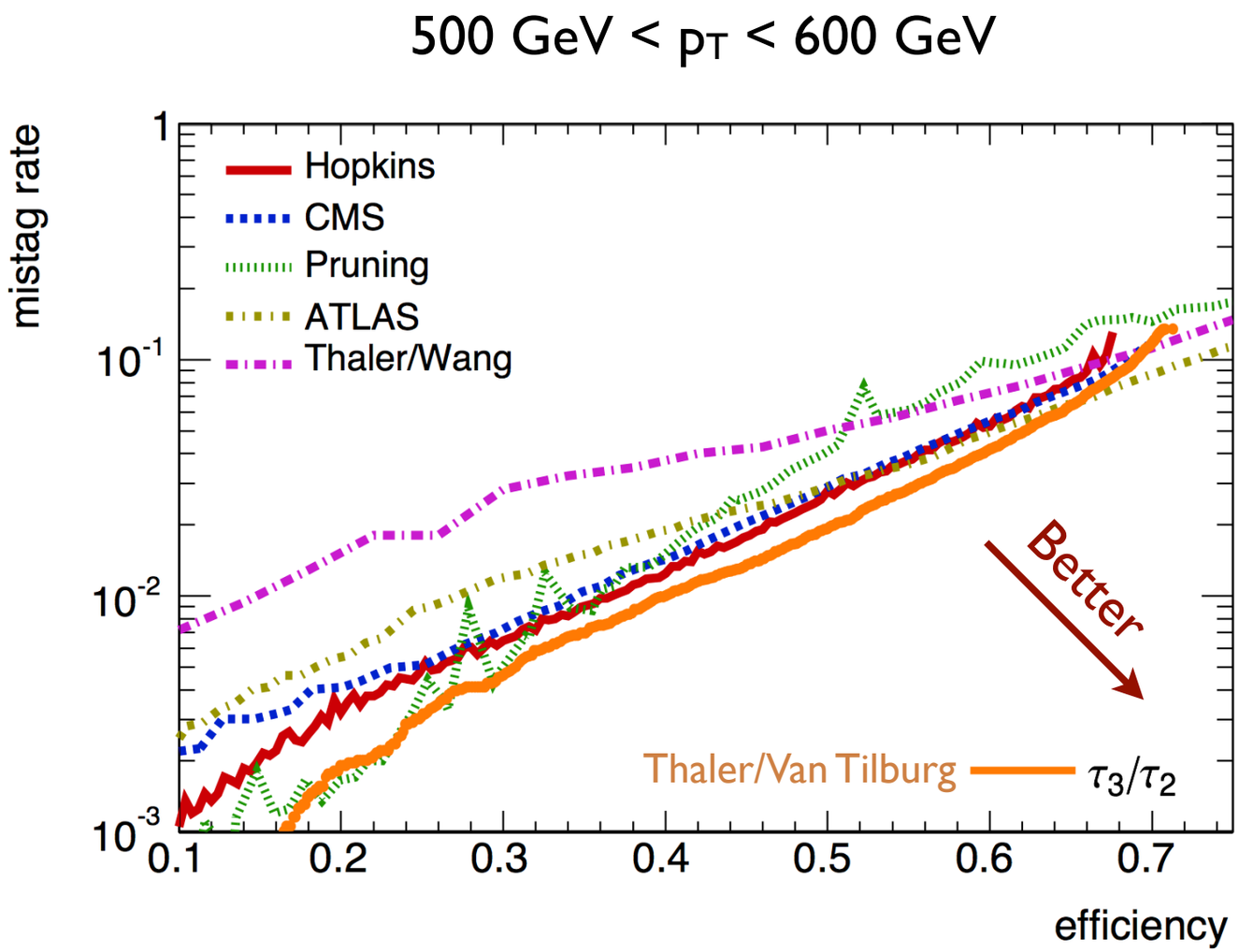
Boost 2011 Workshop

Based on work with Ken Van Tilburg: 1011.2268 & Ken's MIT Senior Thesis

Top Tagging c. 2010



Top Tagging c. 2011



fixed 160 GeV < m_{jet} < 240 GeV cut
one-dimensional cut on τ_3/τ_2

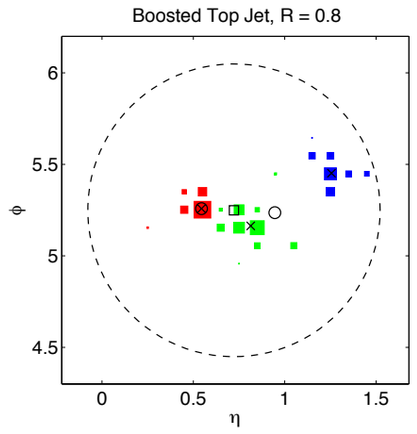
N-subjettiness

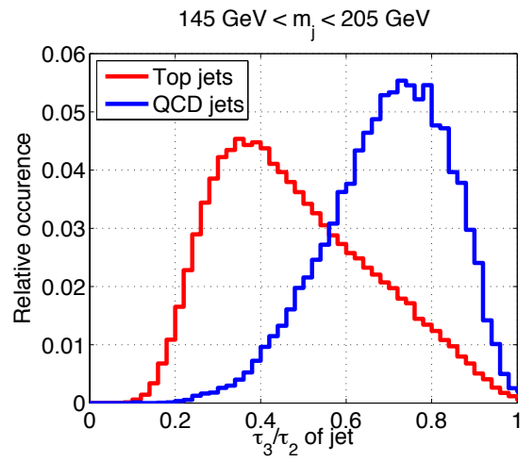
A New Substructure Measure

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min_A \{\Delta R_{A,k}\}$$

Top Tagging with τ_3/τ_2

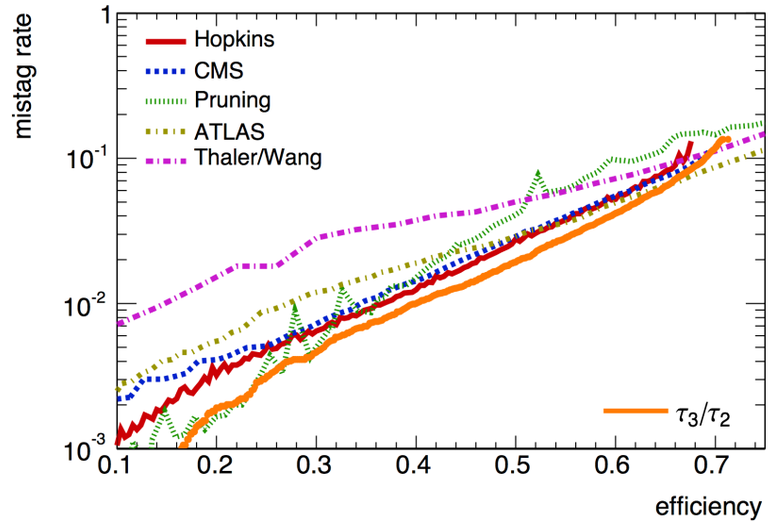
(W/Z/H Tagging with τ_2/τ_1)





Minimization & Boost2010

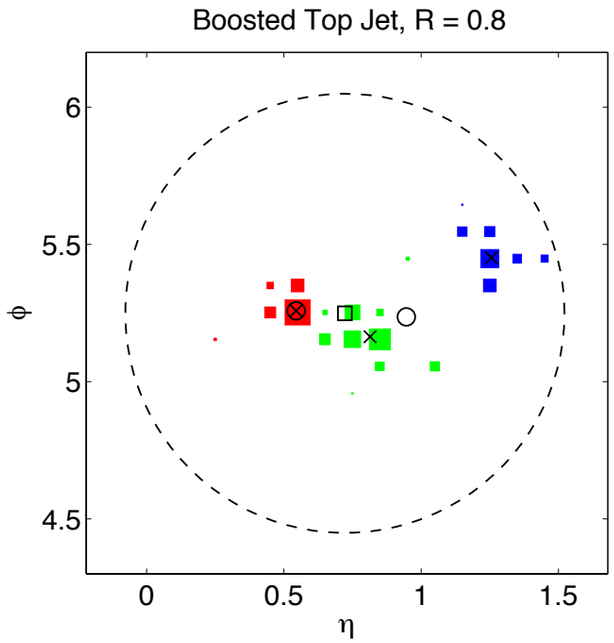
(Thoughts on Jet Algorithms)

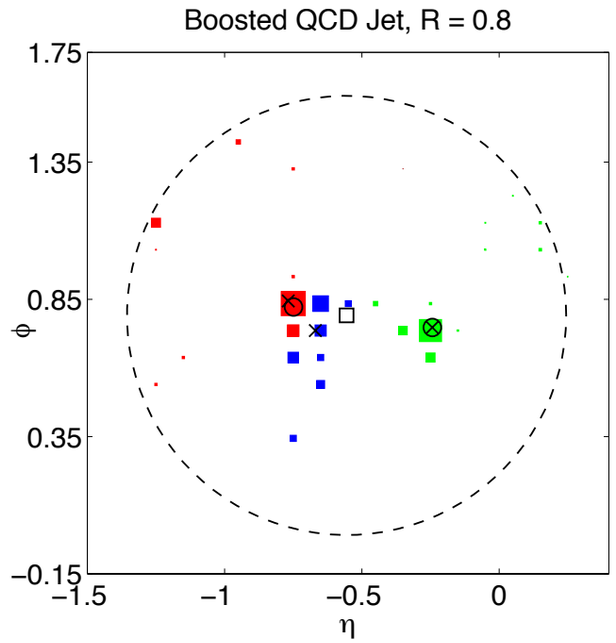


[Thaler, Van Tilburg: 1011.2268; See also J.-H. Kim: 1011.1493]

Jet Substructure by Eye

Two jets with $m_{jet} = m_{top}$. Coloring by exclusive k_T.





Tagging Methods:

"Algorithmic" $\{p_i\} \to p_{\text{jet}}, \text{yes/no}$

Jet Shape Cut $f(\{p_i\}) < f_{\text{cut}}$ (Eventually)

Hybrid $f(\{p_i\}, \text{alg. info}) < f_{\text{cut}}$

Introducing N-subjettiness

"There ... seems to be a rule in physics that the longer you let theorists play with an idea, the more likely it is that they'll give it a silly name."

— Flip Tanedo, USLHC Blog, April 22, 2011

N-subjettiness: Degree to which a jet has N subjets!

$$\tau_N \simeq 0 \quad \Rightarrow \quad \leq N \text{ subjets}$$

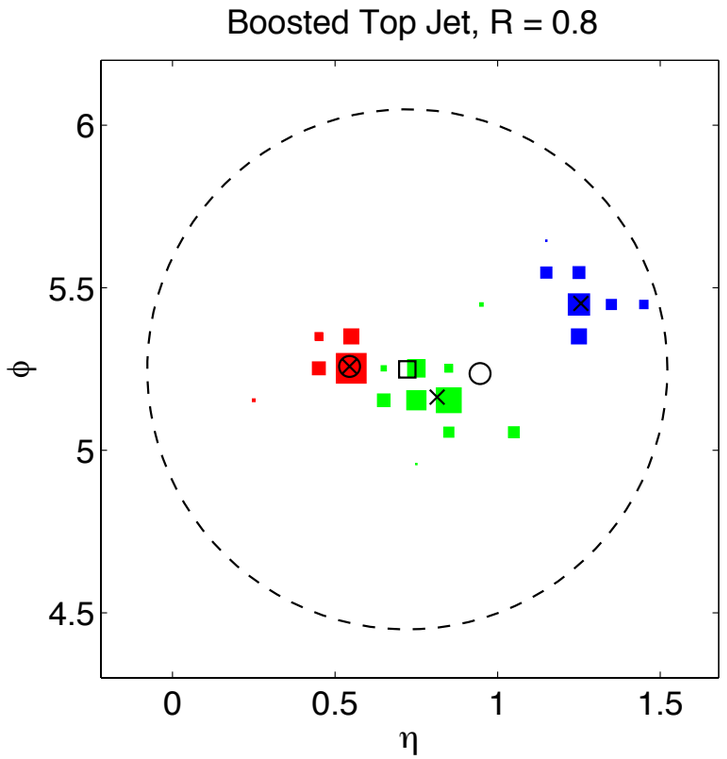
$$\tau_N \simeq 1 \quad \Rightarrow \quad > N \text{ subjets}$$

(You prefer "Voronoi-Tessellated Angularities"?)

Adapted from "N-jettiness" (See Iain's talk)
[Stewart, Tackmann, Waalewijn: 1004.2489]

Introducing N-subjettiness

A Hybrid Jet Shape: $\tau_N = f(\{p_k\}, \{p_{\text{axes}}\})$



$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min_A \{\Delta R_{A,k}\}$$

Sum over constituents

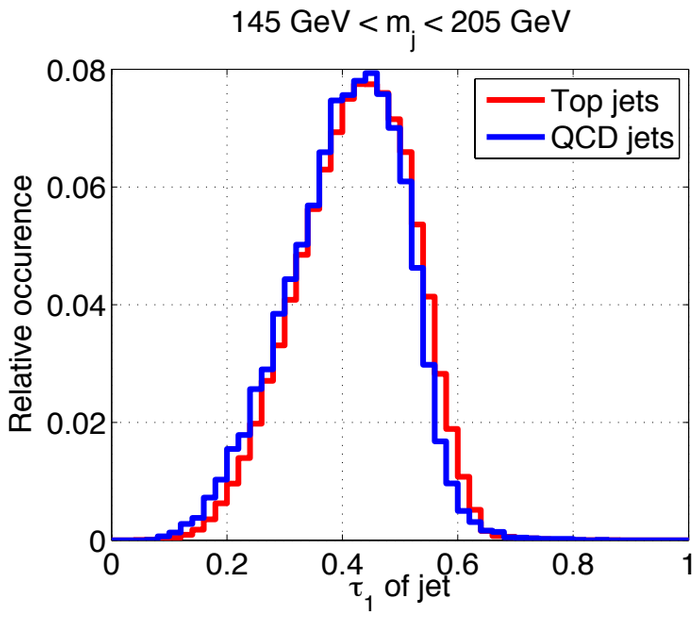
Minimize distance to candidate subjet axes

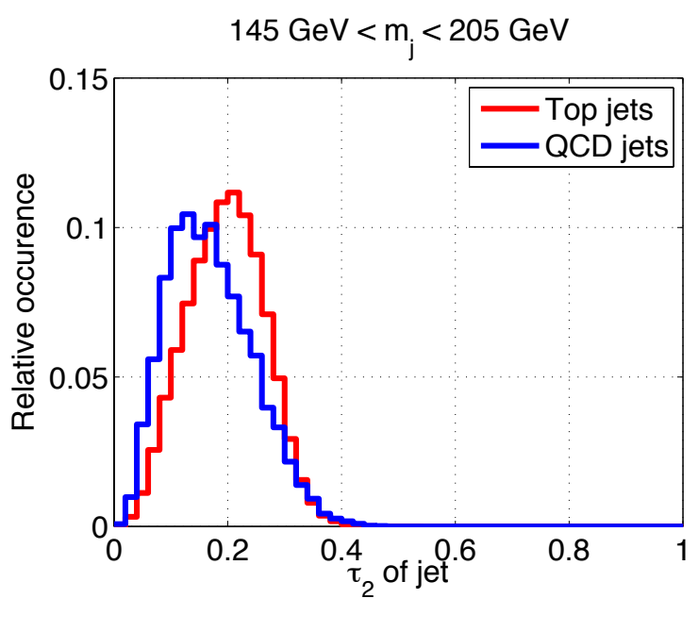
Later: Minimize τ_N over all possible subjet axes, becomes a true jet shape

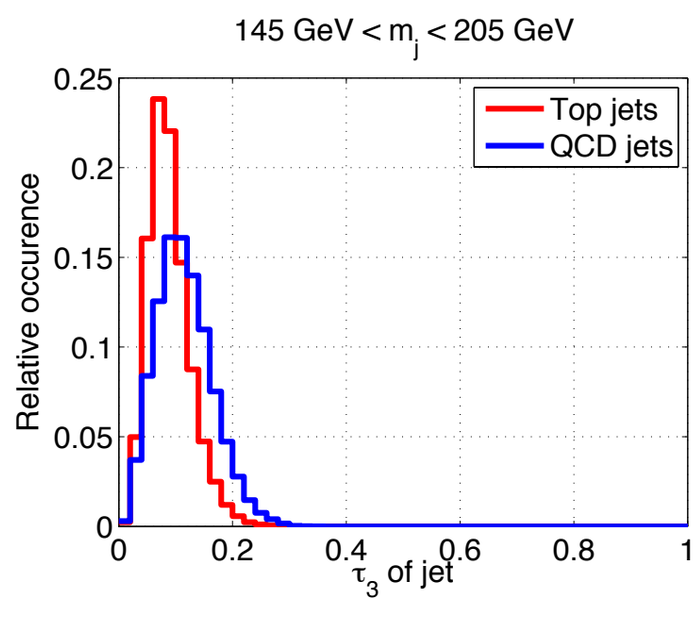
Jet shape that "counts" number of subjets!

# subjets:	≤ N	> N
τ_N:	0	1

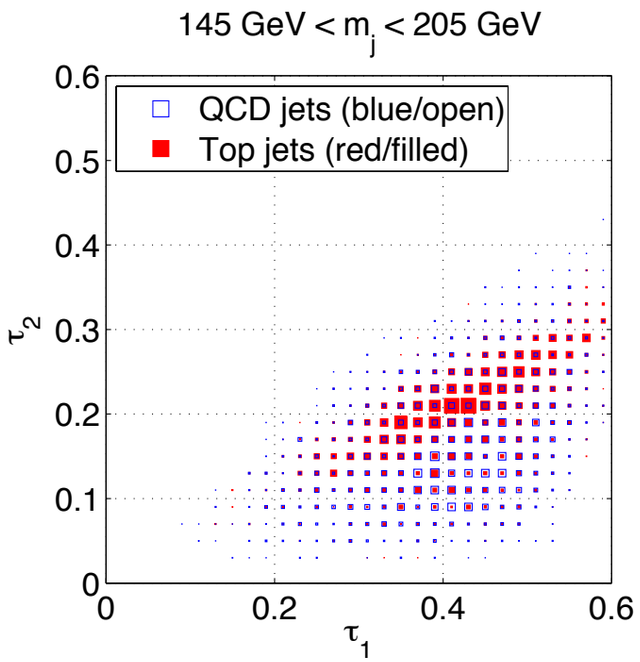
Raw N-subjettiness for Tops

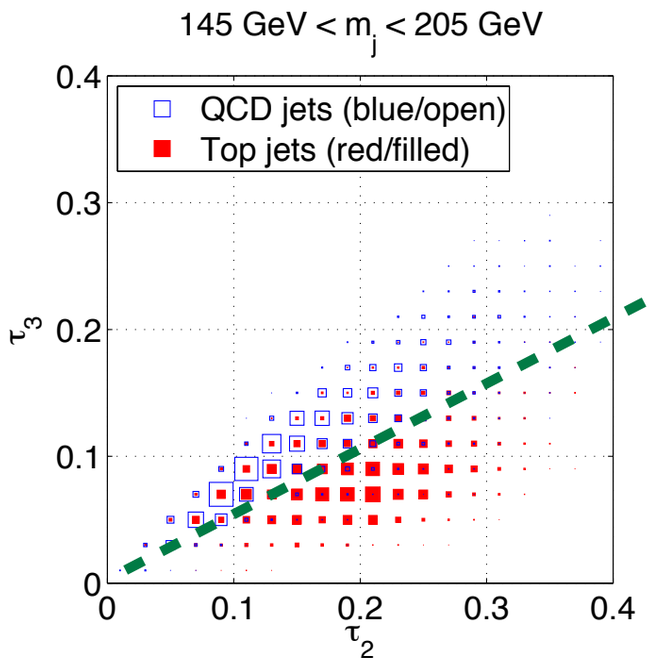






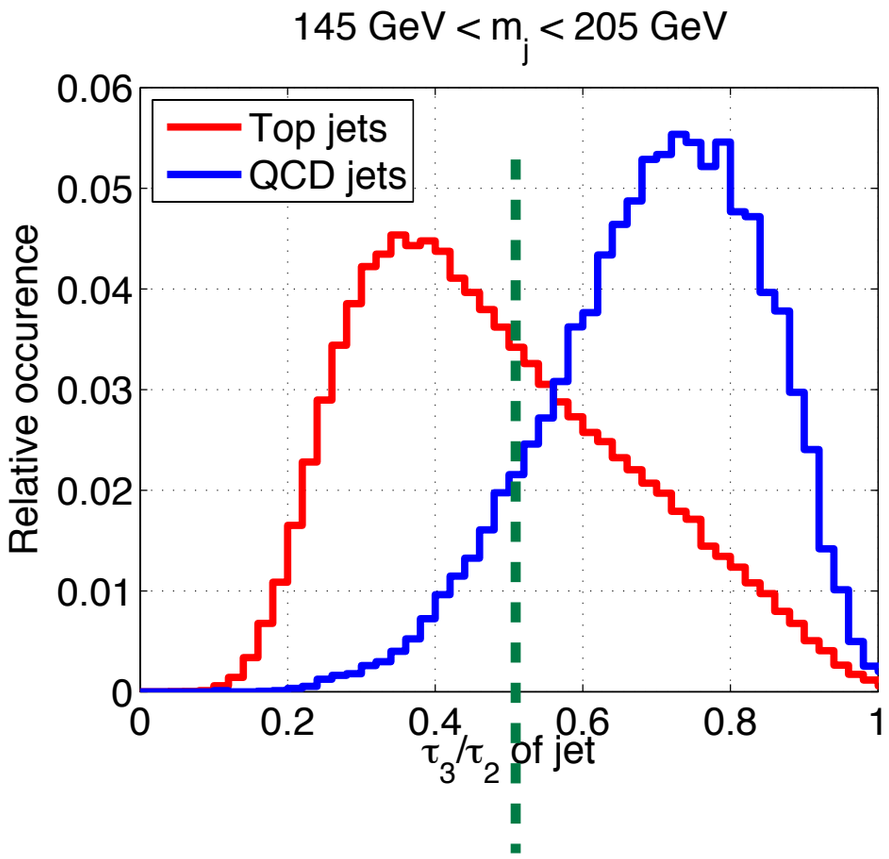
Some raw distinguishing power....





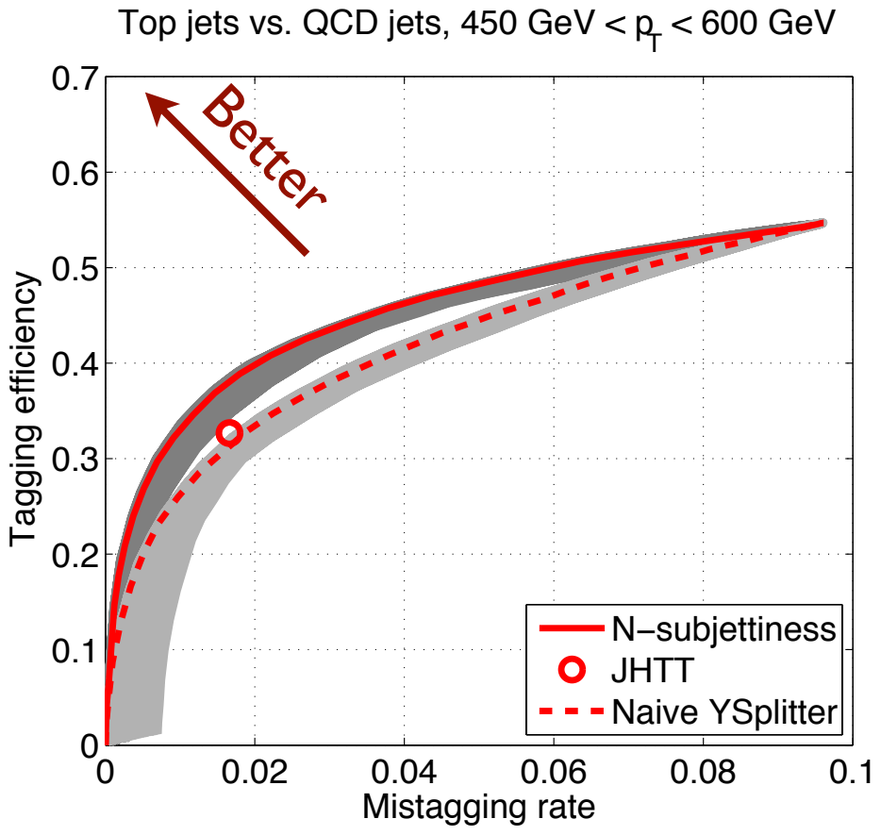
...but suggests multivariate cut!

N-subjettiness for Boosted Tops



Flexible cut to adjust signal acceptance vs. background rejection

τ_3/τ_2: Boosted Tops
τ_2/τ_1: Boosted W/Z/H



Compares favorably to existing methods!

(But "Theorist Detector". Our study before the Boost2010 Proceedings)

Variations/Improvements

Different Angular Weighting

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min_A \{\Delta R_{A,k}^{\beta}\}$$

β = 1 : ≈ Jet Broadening
β = 2 : ≈ Thrust
β = 2-a : ≈ Angularities

(Can also add dependence on axis energy. Doesn't seem to help tagging.)

Different Subjet Axes

Initial Study: Exclusive k_T Subjets
Better Choice: Axes that Minimize τ_N!

Minimization

Classic problem in computer science!

$$\tau_N = \min_{\{p_{\text{axes}}\}} \tilde{\tau}_N \qquad \tilde{\tau}_N = \frac{1}{d_0} \sum_k p_{T,k} \min_A \{\Delta R^{\beta}_{A,k}\}$$

True Jet Shape

Hybrid Jet Shape

β = 2 (≈ thrust): "k-means clustering"

≈ iterative jet finding, fixed number of jets, voronoi boundaries

$$\frac{\partial \tilde{\tau}_1}{\partial \vec{p}_{\text{axis}}} = 0 \qquad \Rightarrow \qquad \vec{p}_{\text{axis}} = \sum_k \vec{p}_k$$

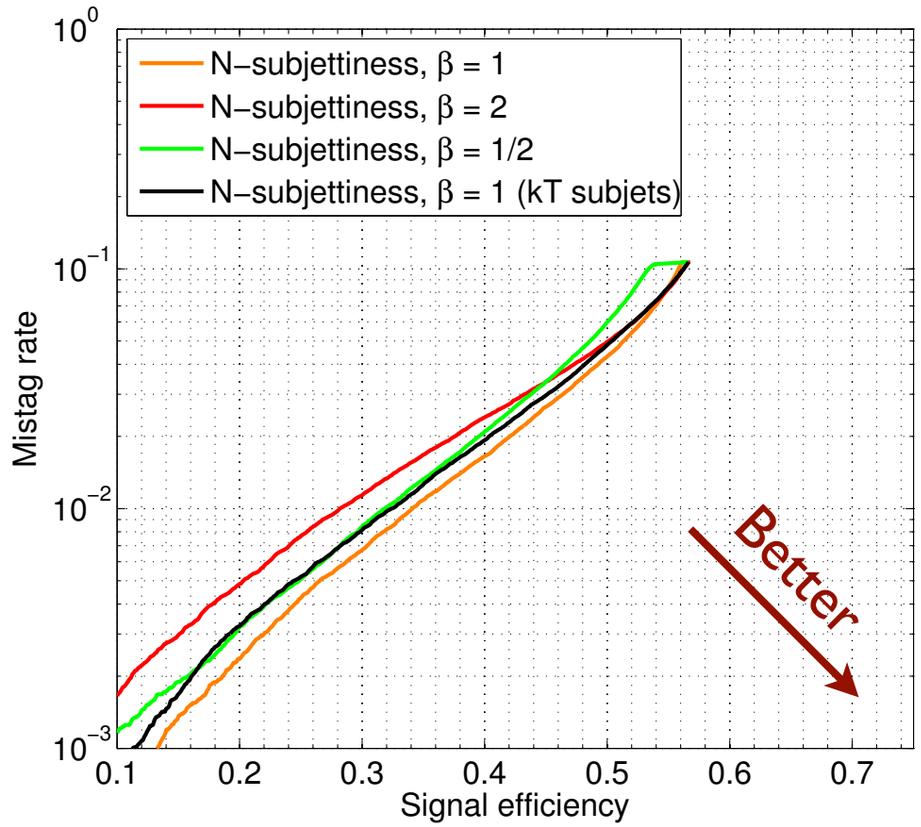
i.e. jet axis = jet momentum

β = 1 (≈ jet broadening): We developed new minimization algorithm, jet axis ≠ jet momentum, kind of like a "median" jet

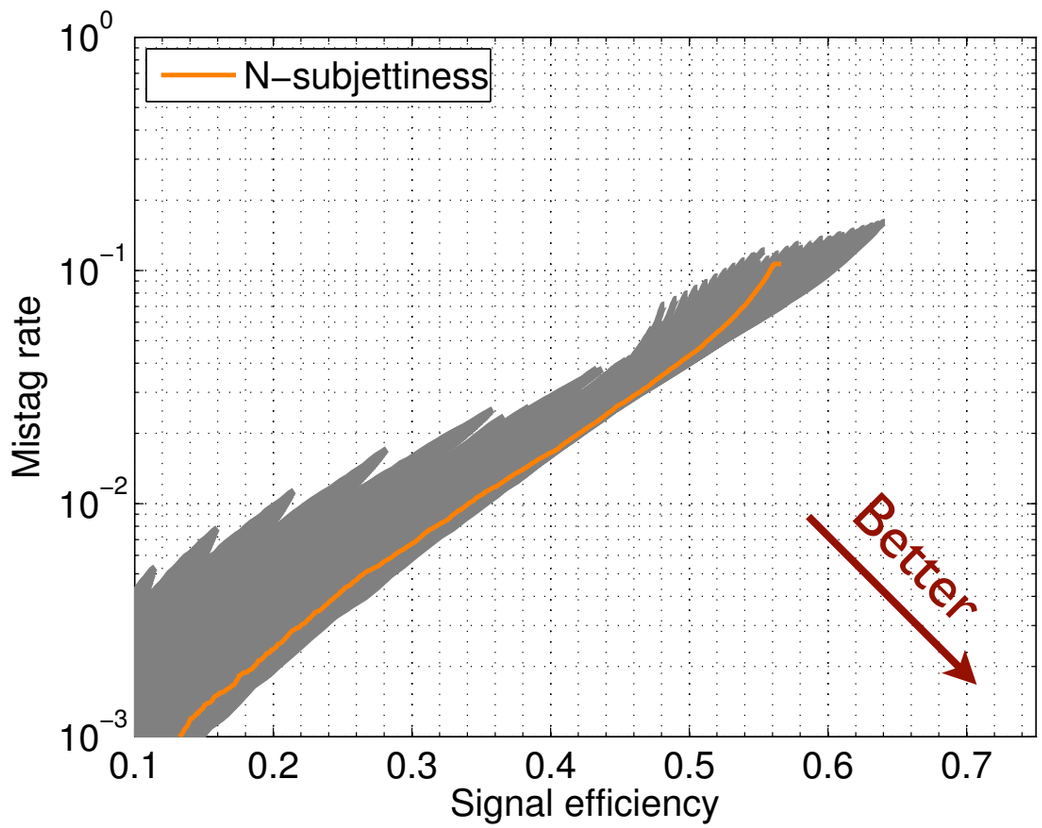
N-subjettiness Variations

Using Boost 2010 Samples

200 GeV < p_T < 800 GeV



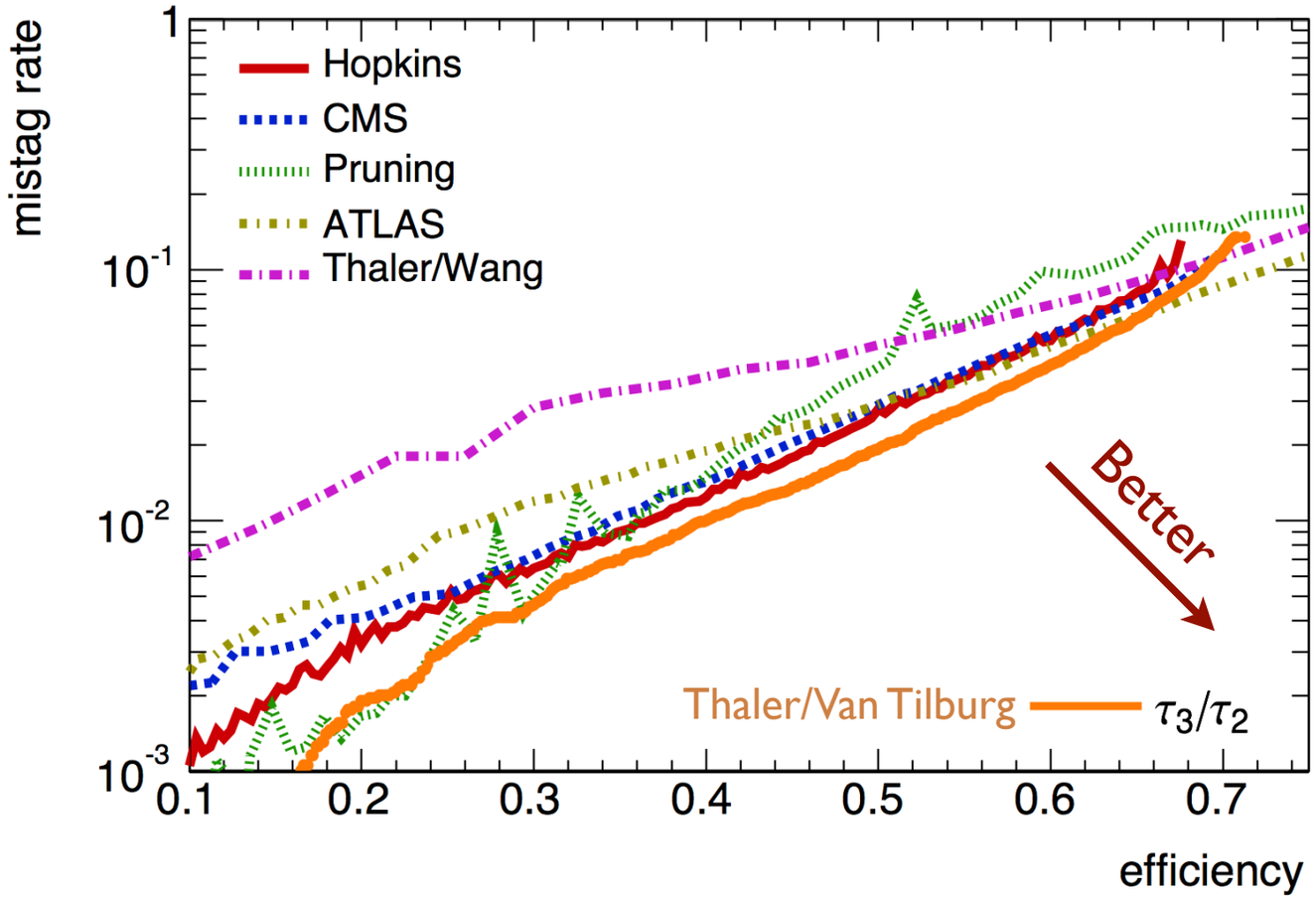
200 GeV < p_T < 800 GeV



Best for Top Tagging:
τ_3/τ_2 Cut
Jet Broadening Measure ($\beta = 1$)
Axes from Minimization

Can improve by adjusting slope of τ_3, τ_2 cut.
Many multivariate options.

Top Tagging c. 2011



500 GeV < p_T < 600 GeV

fixed jet mass cut:
160 GeV to 240 GeV

one-dimensional cut
on τ_3/τ_2 (from minimum)

N-subjettiness:

Intuitive measure of "number of subjets"
Based on true jet shape (hybrid still an option)
Plays well with multivariate methods
Calculable (resummable?) in perturbation theory
Excellent tagging performance out of the box
Ratio τ_N/τ_{N-1} has reduced JES sensitivity
Plays well with grooming

Thoughts on Jet Algorithms

Hybrid Jet Shape: $\tau_N = f(\{p_k\}, \{p_{\text{axes}}\})$

Minimization of any hybrid jet shape defines a jet algorithm!

Not a new idea... Stable cone finding:

Minimize over axis A: $\tau_1(R_0) = \sum_k p_{Tk} \min(\Delta R_{A,k}, R_0)^2$ ← Key!

Solution: $p_{\text{jet}} = \sum_k p_k$ in cone

[Ellis, Huston, Tonnesmann]

In CS optimization: cluster finding = minimization

N-jettiness Jets?

Minimize τ_N ($\beta = 2$) = k-means clustering

[See Chekanov]

Quasi-minimize $\tau_N(R_0)$ ($\beta = 2$) $\approx$ anti-k_T (!)

(if N known, because jet axis = jet momentum)

Potential Benefits of Jets from τ_N

τ_N defines jet directions *and* quality measure
Controllable split-merge (allows "recoil" like in anti-k_T)
With $\beta \neq 2$, can study jet axis $\neq$ jet momentum
And anti-k_T is very close to doing this already

Potential Issues with Jets from τ_N

Have to know N ahead of time (is this really a problem?)
Hard to guarantee global minimum (recall seeded cone IRC non-safety)
R_0 leads to local minima and slow convergence

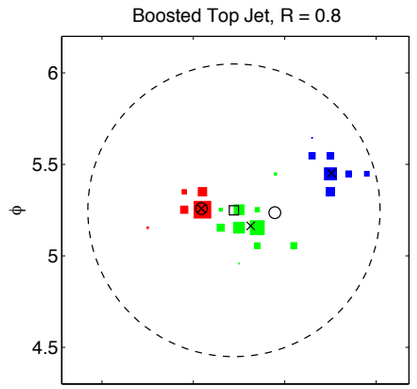
N-subjettiness

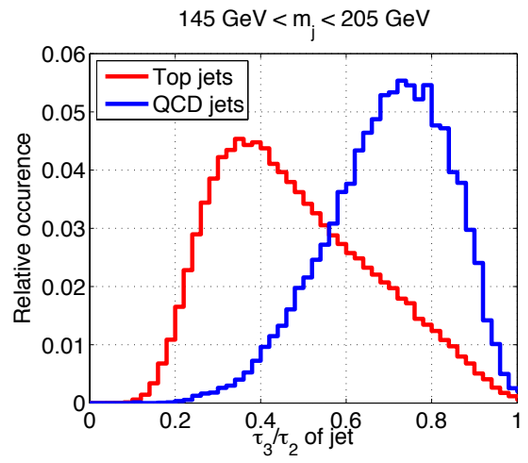
A New Substructure Measure

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min_A \{\Delta R_{A,k}\}$$

Top Tagging with τ_3/τ_2

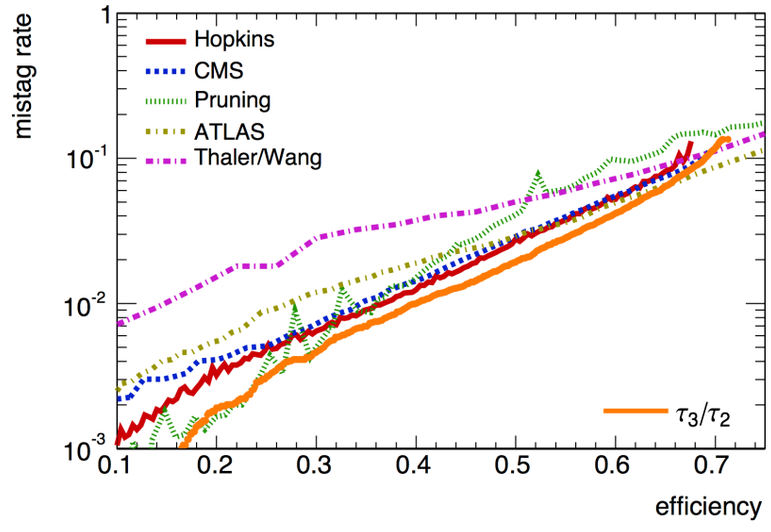
(W/Z/H Tagging with τ_2/τ_1)





Minimization & Boost2010

(Thoughts on Jet Algorithms)



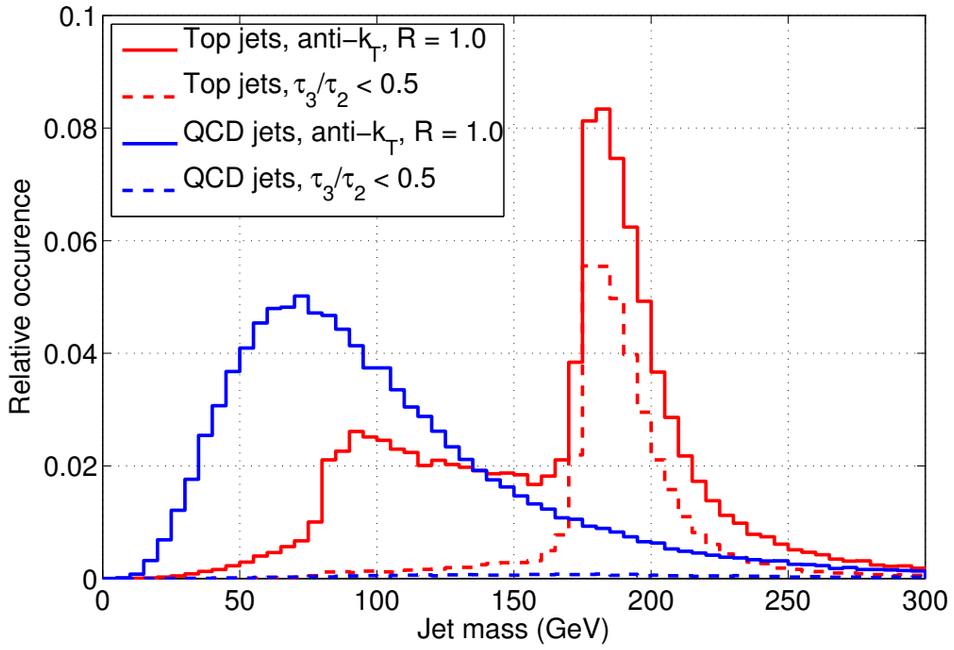
[Thaler, Van Tilburg: 1011.2268; See also J.-H. Kim: 1011.1493]

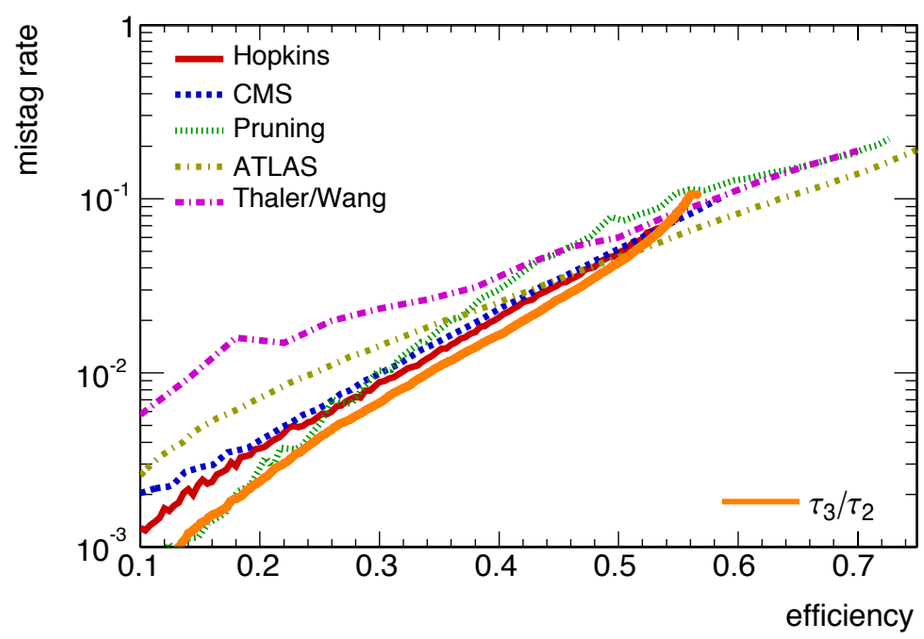
Backup Slides

More Plots

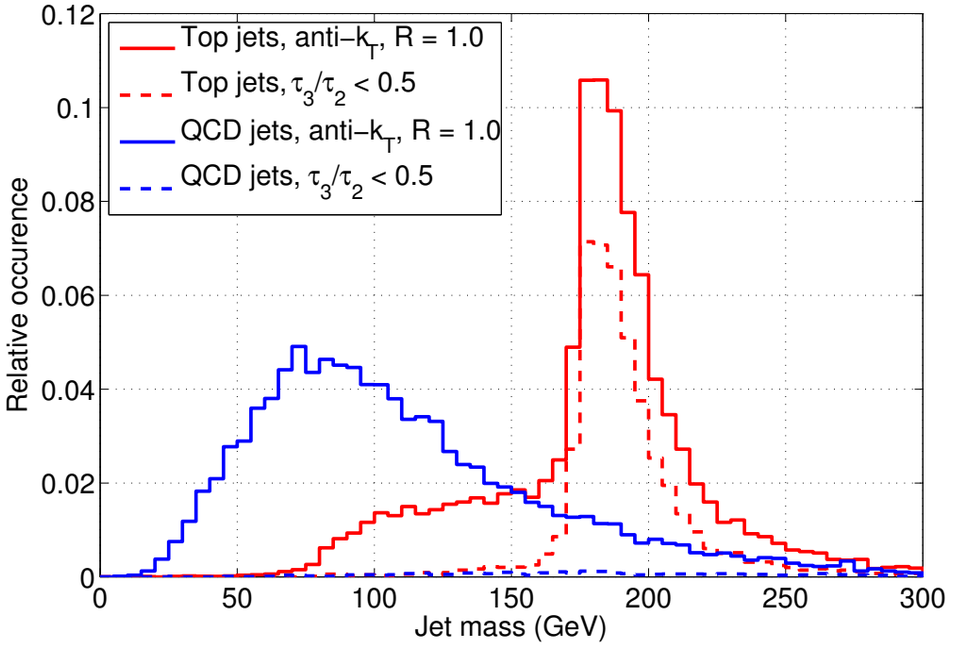
Using Boost 2010 Samples

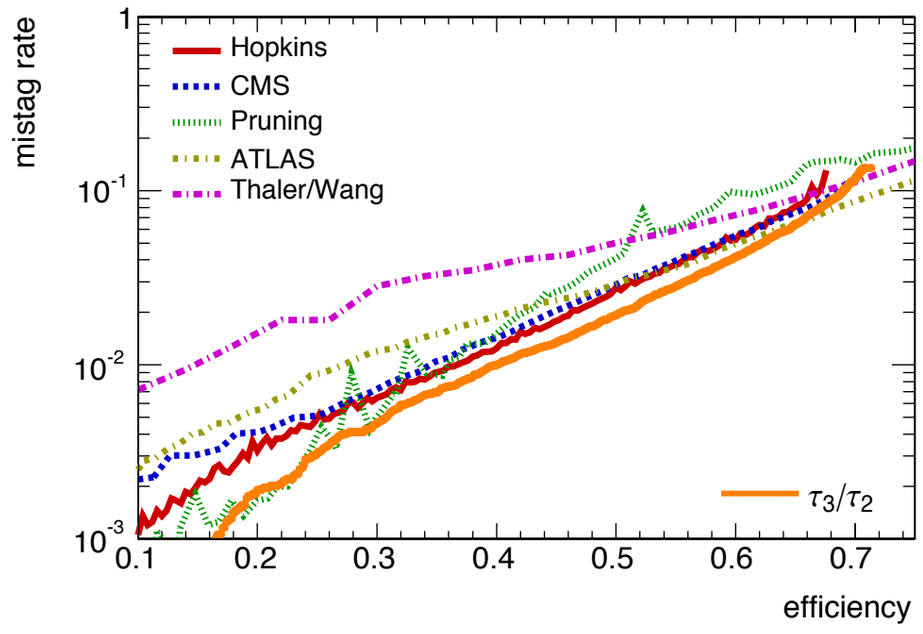
200 GeV < p_T < 800 GeV



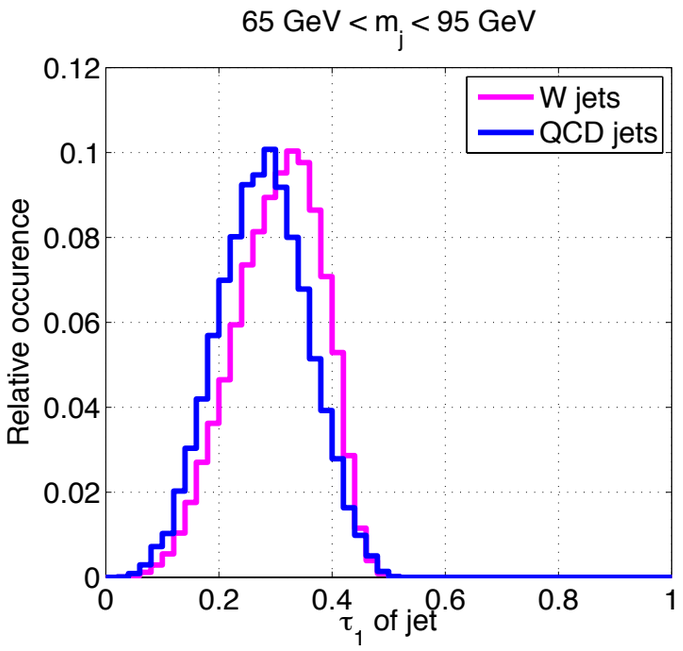


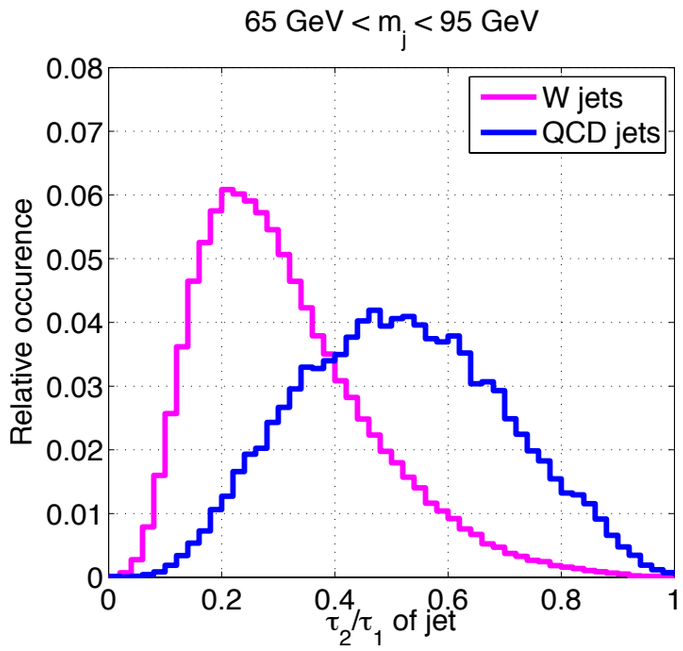
500 GeV < p_T < 600 GeV

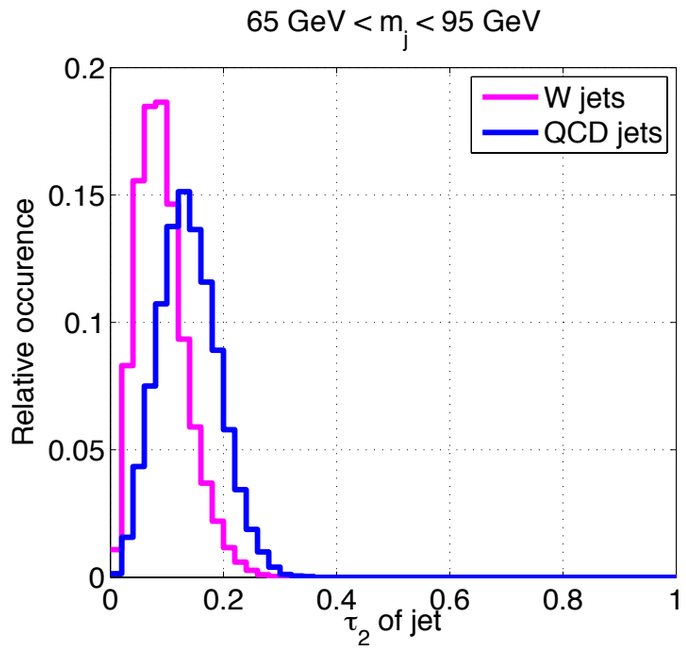


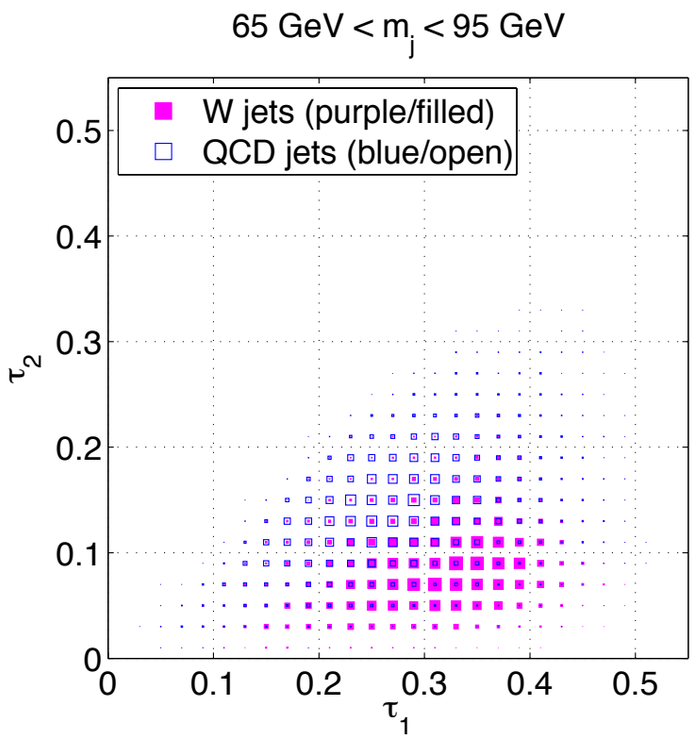


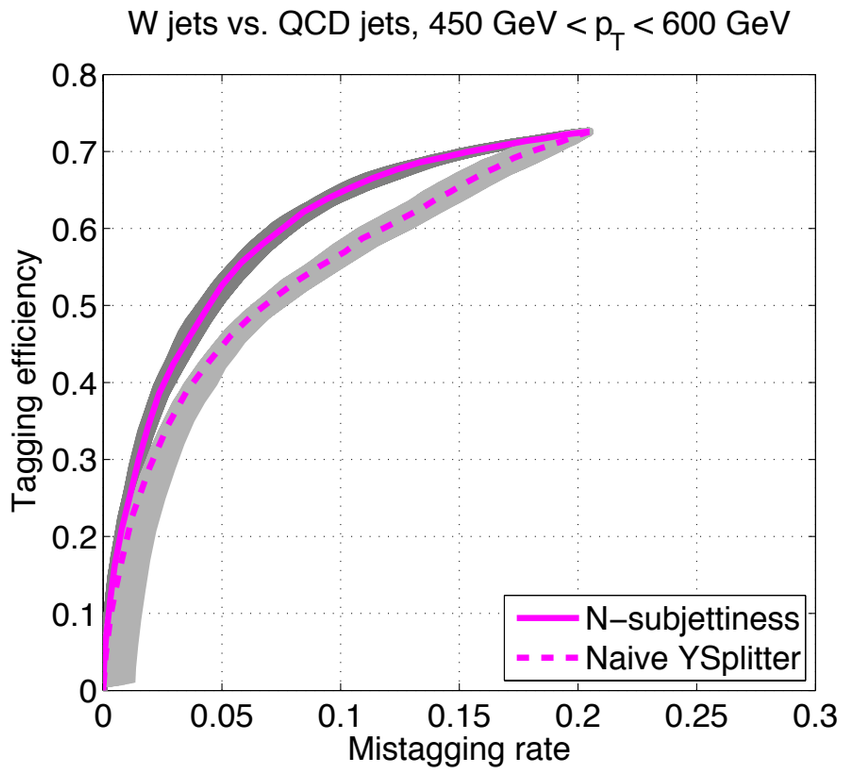
W Tagging


65 GeV < m_j < 95 GeV
W jets
QCD jets
Relative occurence
τ_1 of jet


65 GeV < m_j < 95 GeV
W jets
QCD jets
Relative occurence
τ_2/τ_1 of jet


65 GeV < m_j < 95 GeV
W jets
QCD jets
Relative occurence
τ_2 of jet


65 GeV < m_j < 95 GeV
W jets (purple/filled)
QCD jets (blue/open)
τ_2
τ_1


W jets vs. QCD jets, 450 GeV < p_T < 600 GeV
N−subjettiness
Naive YSplitter
Tagging efficiency
Mistagging rate