Seeing Half the Story: Tagging Partially Reconstructed Jet Substructure

Marat Freytsis

Harvard University
UC-Berkeley Lawrence Berkeley National Lab

November 1, 2012

FNAL Theory Seminar

With Tomer Volansky and Jon Walsh, arXiv:121X.XXXX

Jet substructure & its limits

- Varying approaches have appears over the last several $\approx$ 5 years
- But, all available methods assume full decay chain reconstruction
  - Need mass resonance cuts to cut down on background
- If part of the decay of a boosted object is not reconstructed, what then?
  - Is there any hope for useful substructure in this case?
    - All the time, *e.g.* booste,d colored NP $\rightarrow$ SM + invisible
    - Part of the time if decays kinematically unresolvable, *e.g.* tops

Jet substructure & its limits

- Varying approaches have appears over the last several ≈ 5 years
- But, all available methods assume full decay chain reconstruction
  - Need mass resonance cuts to cut down on background
- If part of the decay of a boosted object is not reconstructed, what then?
  - Is there any hope for useful substructure in this case?
    - All the time, *e.g.* booste,d colored NP $\rightarrow$ SM + invisible
    - Part of the time if decays kinematically unresolvable, *e.g.* tops

Outline

Top quarks as standard candles

Top quarks are a good testing ground for boosted techniques

- Boosted tops have already been observed at ATLAS and CMS
- Mature field of taggers allows for comparison with standard techniques
- No new physics to add – understand how tops decay
- Two step decay through W provides rich spectrum of subjet behavior
  - Complex enough for hope with partially reconstructed tops *vs.* QCD

A couple of examples . . .

Top quarks as standard candles

Top quarks are a good testing ground for boosted techniques

- Boosted tops have already been observed at ATLAS and CMS
- Mature field of taggers allows for comparison with standard techniques
- No new physics to add – understand how tops decay
- Two step decay through W provides rich spectrum of subjet behavior
  - Complex enough for hope with partially reconstructed tops *vs.* QCD

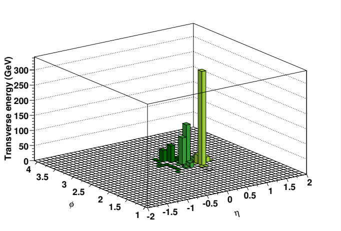
A couple of examples . . .

Johns Hopkins Top Tagger

Jet declustering

Basic idea: Traverse backwards through jet merging procedure, identify W, b candidates

- Take one step back in clustering algorithm
- Only keep subjet if $p_{T,i}/p_T > \delta_p$
  "Hard subjets only"
- Only keep pairs if $\Delta R_{ij} > \delta_r$
  "Well separated subjets"
- Repeat declustering until 3 or 4 subjets present
  4^{th} jet interpreted as hard gluon emission



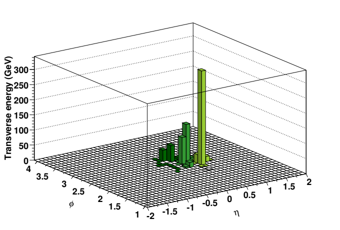
Kaplan *et al.*, 0806.0848

Johns Hopkins Top Tagger

Jet declustering(2)

- Check that all subjets together reconstruct close to the top mass
- Check that 2 subjets reconstruct close to the W mass
- Check that angle of W decay products to top direction in W rest frame has $\cos\theta < \cos\theta_h$

  "W helicity angle: cut on collinear divergence in QCD"



Kaplan *et al.*, 0806.0848

Choices of cuts can be fixed, or varied on an event-by-event basis with kinematics of event

τ_3 / τ_2 – Tagging through event shapes

Defining N-subjettiness

For N directions in the η-ϕ plane define by

$$\tilde{\tau}_N^{(\beta)} = \frac{1}{d_0} \sum_i p_{T,i} \min\left\{ (\Delta R_{1,i})^\beta, (\Delta R_{2,i})^\beta, \ldots, (\Delta R_{N,i})^\beta \right\}$$

where

$$d_0 = \sum_i p_{T,i} R_0^\beta .$$

N-subjettiness is then an observable given by

$$\tau_N^{(\beta)} = \min_{\hat{n}_1, \hat{n}_2, \ldots, \hat{n}_N} \tilde{\tau}_N^{(\beta)}$$

τ_N is a measure of distinguishability from a pure N-prong jet

τ_3 / τ_2 – Tagging through event shapes

Defining N-subjettiness

For N directions in the η-ϕ plane define by

$$\widetilde{\tau}_N^{(\beta)} = \frac{1}{d_0} \sum_i p_{T,i} \min \left\{ (\Delta R_{1,i})^\beta, (\Delta R_{2,i})^\beta, \ldots, (\Delta R_{N,i})^\beta \right\}$$

where

$$d_0 = \sum_i p_{T,i} R_0^\beta .$$

N-subjettiness is then an observable given by

$$\tau_N^{(\beta)} = \min_{\hat{n}_1, \hat{n}_2, \ldots, \hat{n}_N} \widetilde{\tau}_N^{(\beta)}$$

τ_N is a measure of distinguishability from a pure N-prong jet

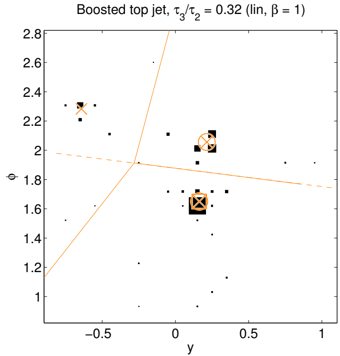
τ_3/τ_2 – Tagging through event shapes

A top tagger

Value of τ_N can vary substantially event-by-event – poor discriminant

Instead, cut on $\tau_3/\tau_2 < \delta$

Want events that are far more 3-subjet like than 2-subjet like



Thaler, Van Tilburg, 1108.2701

Which tops do we tag?

Many top taggers now on the market:

- ATLAS
- CMS
- Hopkins tagger
- HEPTopTagger
- N-subjettiness
- Thaler/Wang
- trimming
- pruning
- ASF

Methodologies vary significantly
Ultimately look for same structures: 3 visible/separable subjets w/ mass cuts
Actually, jet grooming methods make no such assumption in general,
but as implemented include specific cuts on number + form of subjets

How well can we ultimately hope to do with this approach?

Which tops do we tag?

Many top taggers now on the market:

- ATLAS
- CMS
- Hopkins tagger
- HEPTopTagger
- N-subjettiness
- Thaler/Wang
- trimming
- pruning
- ASF

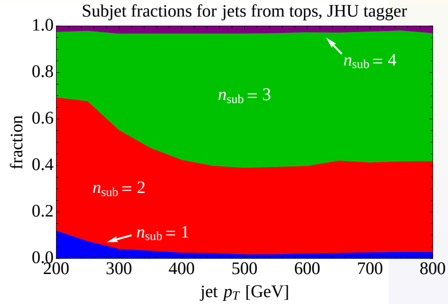
Methodologies vary significantly

Ultimately look for same structures: 3 visible/separable subjets w/ mass cuts

Actually, jet grooming methods make no such assumption in general,
but as implemented include specific cuts on number + form of subjets

How well can we ultimately hope to do with this approach?

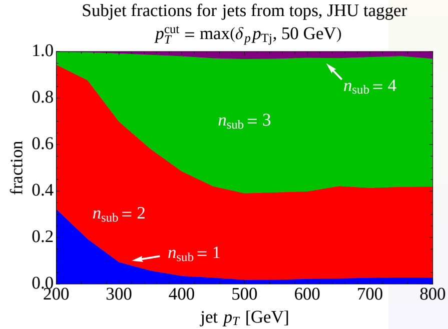
How many subjets in a hadronic top?



What do unreconstructed tops look like?

- Look at # of jets reconstructed by JHU tagger
- Most have at least 2 jets
- At intermediate p_T, such jets become the majority of the sample

How many subjets in a hadronic top?



What do unreconstructed tops look like?

- Look at # of jets reconstructed by JHU tagger
- Most have at least 2 jets
- At intermediate p_T, such jets become the majority of the sample

Identifying missing tops

Check our claims *vis-à-vis* other taggers

- Are the same tops being reconstructed?
- Are the mistags coming from the same QCD jets?

Look at Hopkins tagger, N-subjettiness, trimming –
Gives range across algorithmic, jet shape, and grooming approaches

taggers	tops	QCD
N-sub/Hopkins	94.6%	82.4%
N-sub/trim	87.1%	69.0%
Hopkins/trim	90.6%	73.1%

overlap at 50% efficiency points and p_T of 400-500 GeV

Identifying missing tops

Check our claims *vis-à-vis* other taggers

- Are the same tops being reconstructed?
- Are the mistags coming from the same QCD jets?

Look at Hopkins tagger, N-subjettiness, trimming –
Gives range across algorithmic, jet shape, and grooming approaches

taggers	tops	QCD
N-sub/Hopkins	94.6%	82.4%
N-sub/trim	87.1%	69.0%
Hopkins/trim	90.6%	73.1%

overlap at 50% efficiency points and p_T of 400-500 GeV

Tagger efficiencies and kinematics

Do untagged tops have any common features?

The common element of discussed top taggers is the presence of 3+ well-separated hard jets. Expect this assumption to break down if decay axis is close to boost axis or at lower p_T:

- W/b produced in opposite direction to top boost can be soft/out of jet
- Even if hard enough to pass cut, might be too close to another jet to be resolved
- At lower p_T, range of angles yielding soft decay products increased

Tagger efficiencies and kinematics

Do untagged tops have any common features?

The common element of discussed top taggers is the presence of 3+ well-separated hard jets. Expect this assumption to break down if decay axis is close to boost axis or at lower p_T:

- W/b produced in opposite direction to top boost can be soft/out of jet
- Even if hard enough to pass cut, might be too close to another jet to be resolved
- At lower p_T, range of angles yielding soft decay products increased

Tagger efficiencies and kinematics

Can calculate tagged distribution at parton level

Demand that all 3 partons be within one jet, and that W be resolvable from its decay products

Other distributions show similar behavior, but less helpful as initial distributions not flat

Tops with boost of $\gamma = 3$

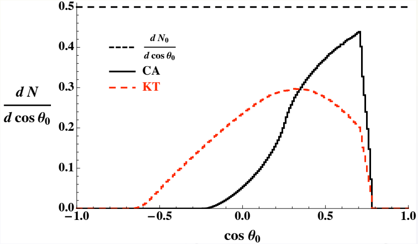
Tagger efficiencies and kinematics

Can calculate tagged distribution at parton level

Demand that all 3 partons be within one jet, and that W be resolvable from its decay products

Other distributions show similar behavior, but less helpful as initial distributions not flat

Tops with boost of $\gamma = 3$



Ellis, Vermilion, Walsh, 0912.0033

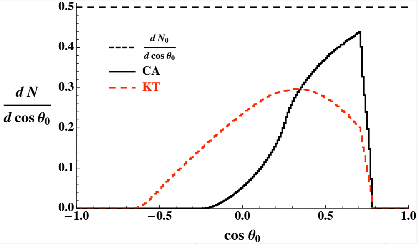
Tagger efficiencies and kinematics

Can calculate tagged distribution at parton level

Demand that all 3 partons be within one jet, and that W be resolvable from its decay products

Other distributions show similar behavior, but less helpful as initial distributions not flat

Tops with boost of $\gamma = 3$



Ellis, Vermilion, Walsh, 0912.0033

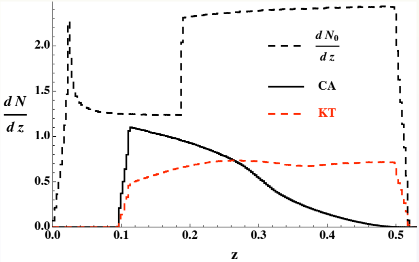
Tagger efficiencies and kinematics

Can calculate tagged distribution at parton level

Demand that all 3 partons be within one jet, and that W be resolvable from its decay products

Other distributions show similar behavior, but less helpful as initial distributions not flat

Tops with boost of $\gamma = 3$



Ellis, Vermilion, Walsh, 0912.0033

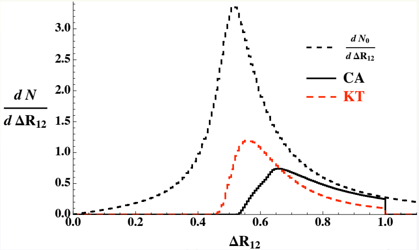
Tagger efficiencies and kinematics

Can calculate tagged distribution at parton level

Demand that all 3 partons be within one jet, and that W be resolvable from its decay products

Other distributions show similar behavior, but less helpful as initial distributions not flat

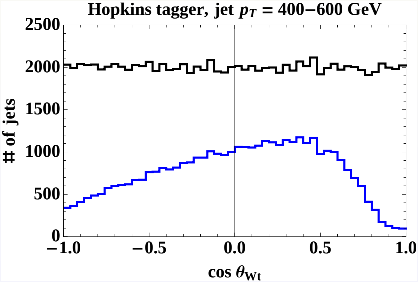
Tops with boost of $\gamma = 3$

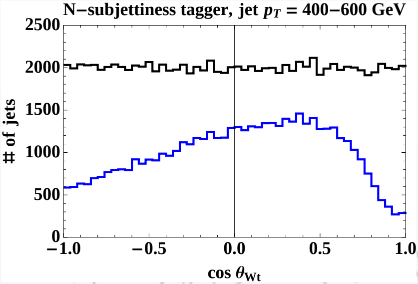


Ellis, Vermilion, Walsh, 0912.0033

Tagger efficiencies and kinematics

Similar behavior for taggers running on full jets
However, now cutoffs as functions of angles less pronounced & distributions smeared over range of boost factors





Looking for missing substructure

Finding the top

Understand the kinematic configurations likely to give untagged top
Do they provide any handles of their own?

- For boosted top, soft W/b may still shower partially in jet
- There will be typically 2 hard subjets, not necessarily aligned with the direction of the top

Want to look for "lopsided" fat jets

Looking for missing substructure

Finding the top

Understand the kinematic configurations likely to give untagged top
Do they provide any handles of their own?

- For boosted top, soft W/b may still shower partially in jet
- There will be typically 2 hard subjets, not necessarily aligned with the direction of the top

Want to look for "lopsided" fat jets

Looking for missing substructure

Cutting out QCD

- Massive QCD jets typically due to perturbitive splitting of g, q
- Hard splittings from QCD more likely to be uneven in energy (high z) than from massive objects due to kinematic effects

Want to look for uneven splittings in subjets of remaining sample

Lopsided jets?

Need a measure for lopsidedness of jet even though $p_J^\mu = \sum p_i^\mu$

Accomplished by defining subjets by axes minimizing measure with different ΔR weighting from original jet

This is exactly what the β weighting of N-subjettiness allows us to do!

$$\tau_N^{(\beta)} = \min_{\hat{n}_1,\hat{n}_2,\ldots,\hat{n}_N} \frac{1}{d_0} \sum_i p_{T,i} \min\left\{(\Delta R_{1,i})^\beta, (\Delta R_{2,i})^\beta, \ldots, (\Delta R_{N,i})^\beta\right\}$$

Lopsided jets?

Need a measure for lopsidedness of jet even though $p_J^\mu = \sum p_i^\mu$

Accomplished by defining subjets by axes minimizing measure with different ΔR weighting from original jet

This is exactly what the β weighting of N-subjettiness allows us to do!

$$\tau_N^{(\beta)} = \min_{\hat{n}_1,\hat{n}_2,\ldots,\hat{n}_N} \frac{1}{d_0} \sum_i p_{T,i} \min\left\{(\Delta R_{1,i})^\beta, (\Delta R_{2,i})^\beta, \ldots, (\Delta R_{N,i})^\beta\right\}$$

Defining the cut

$$\widetilde{\tau}_N^{(\beta)} = \frac{1}{d_0} \sum_i p_{T,i} \min\left\{ (\Delta R_{1,i})^\beta, (\Delta R_{2,i})^\beta, \ldots, (\Delta R_{N,i})^\beta \right\}$$

As $\beta \to 2$, measure approaches that of thrust

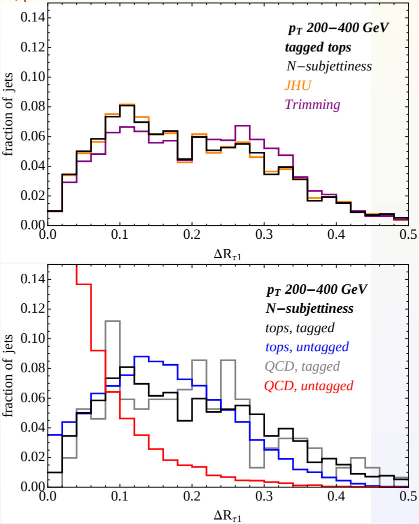
As $\beta \to 1$, we give more heavily weight to higher $p_{T,i}$ partons, even if they are away from center of jet; like jet broadening.

For $\beta = 1$:

- $\Delta R_{\tau_1} > \delta_R$:
  distance from axis defined by τ_1 to center of jet

Cuts and conventional top taggers

ΔR_{τ_1}

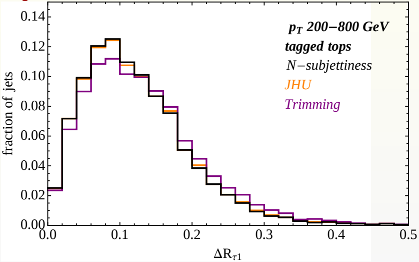


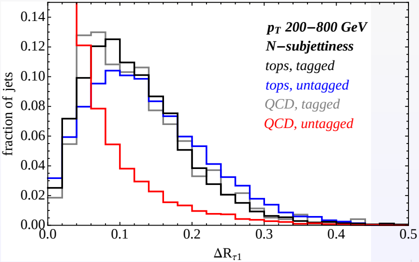
ΔR_{τ_1} cut independent of cuts by standard top taggers

N.B. Separation better at lower p_T

Cuts and conventional top taggers

ΔR_{τ_1}



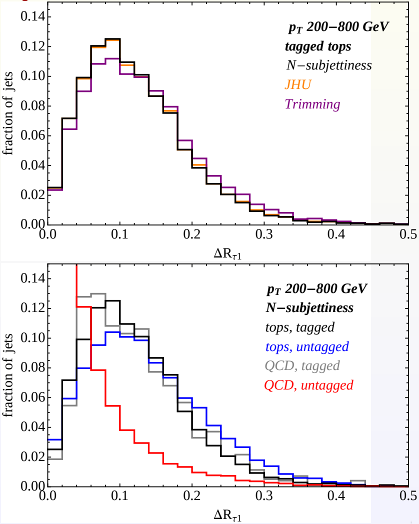


ΔR_{τ_1} cut independent of cuts by standard top taggers

N.B. Separation better at lower p_T

Cuts and conventional top taggers

ΔR_{τ_1}



ΔR_{τ_1} cut independent of cuts by standard top taggers

N.B. Separation better at lower p_T

A parametric separation

β and τ_N axes

Necessity of $\beta = 1$ can be understood from qualitatively different behavior of N-subjettiness axes from $\beta = 2$

For $\beta = 2$, the N-subjettiness measure becomes that of k_T jets. τ_1 axis will then be close to the jet axis by construction.

For $\beta = 1$, axis is instead pulled toward hardest subject. Can be widely separated from jet axis.

A parametric separation

Power counting and radiation

QCD has long tail on distribution from 1-subjet-like jets.
Behavior at $\beta = 1$ can be understood in the framework of soft collinear effective theory (SCET).

N-subjettiness is a generalization of the angularities jet shapes, which have been studied in SCET. Around $\beta = 2$, contributions to shift in axes $O(\lambda^2)$ (SCET$_I$). Anomalous dimensions for angularities have form

$$\Gamma_s = \frac{1}{\beta - 1}\Gamma_{\text{cusp}}, \quad \Gamma_c = -\frac{\beta/2}{\beta - 1}\Gamma_{\text{cusp}},$$

Hornig, Lee, Ovanesyan, 0901.3780

As $\beta \to 1$, need to work in different effective theory (SCET$_{II}$) where soft modes shift axes by $O(\lambda)$.

$O(\lambda)$ deviations in N-subjettiness axes from (sub)jet axes at $\beta = 1$, only $O(\lambda^2)$ for $\beta > 1$. Distributions for axes locations tend to widen as $\beta \to 1$.

Applications for other substructure studies

In general, location of N-subjettiness axes may contain physical information for jets with $N+1$ subjets or more.

This information appears to be orthogonal to that used by other top taggers, including the τ_3/τ_2 tagger itself.

For N-subjettiness can be implemented as an additional set of cuts at zero computation cost (provided you're finding the minimum axes).

Reducing QCD

ΔR_{τ_1} efficiently removes QCD with no hard splittings
Remainder will appear to resolve into two or more subjets
Typically untagged tops will also have two subjets

Kinematics of the two samples very similar
Will differentiate based on likelihood of uneven splitting
In practice what seems to work best is very aggressive trimming

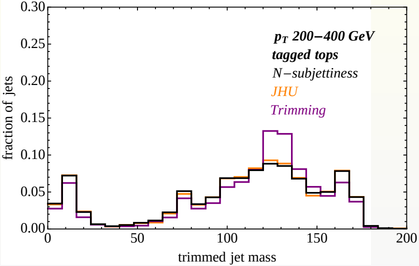
Reducing QCD

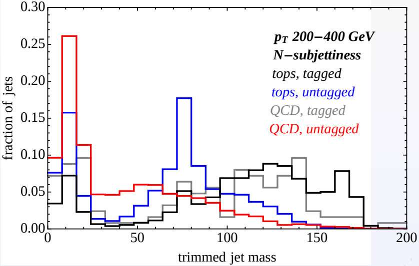
ΔR_{τ_1} efficiently removes QCD with no hard splittings
Remainder will appear to resolve into two or more subjets
Typically untagged tops will also have two subjets

Kinematics of the two samples very similar
Will differentiate based on likelihood of uneven splitting
In practice what seems to work best is very aggressive trimming

Cuts and conventional top taggers

Trimmed jet mass





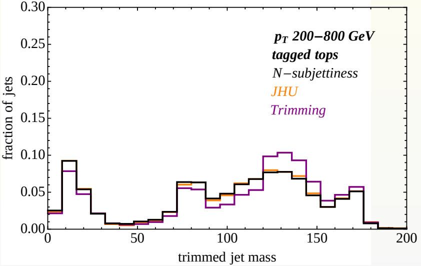
Cut on m_{trim}, using f_{cut} as tunable parameter

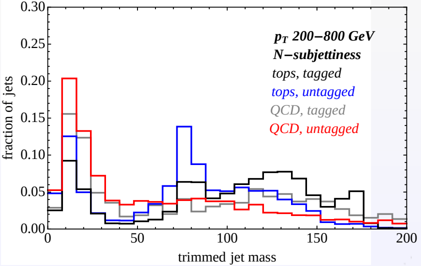
Keep untrimmed jet as output

Performance degrades at higher p_T

Cuts and conventional top taggers

Trimmed jet mass





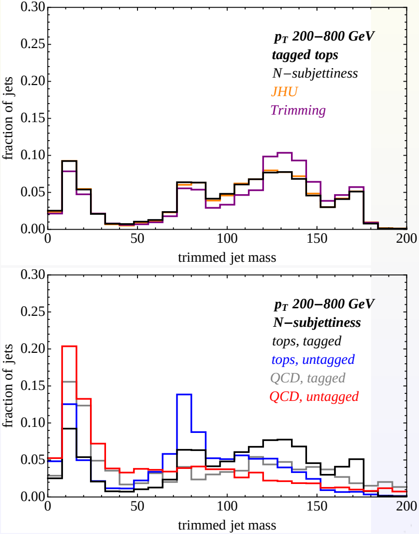
Cut on m_{trim}, using f_{cut} as tunable parameter

Keep untrimmed jet as output

Performance degrades at higher p_T

Cuts and conventional top taggers

Trimmed jet mass

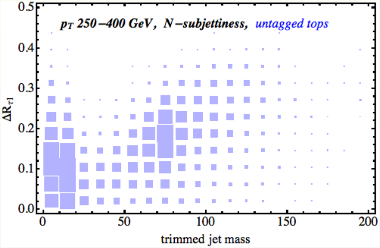


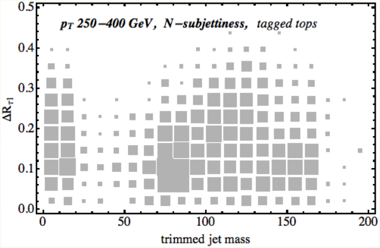
Cut on m_{trim}, using f_{cut} as tunable parameter

Keep untrimmed jet as output

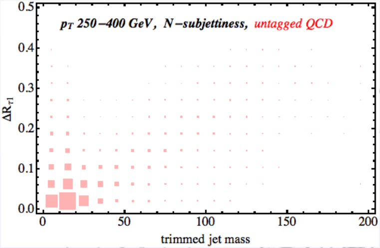
Performance degrades at higher p_T

ΔR_{τ_1} vs. trimmed mass



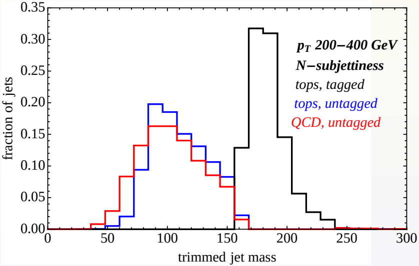


Tops tagged with τ_3/τ_2



Cuts can be tuned to select for typically untagged tops

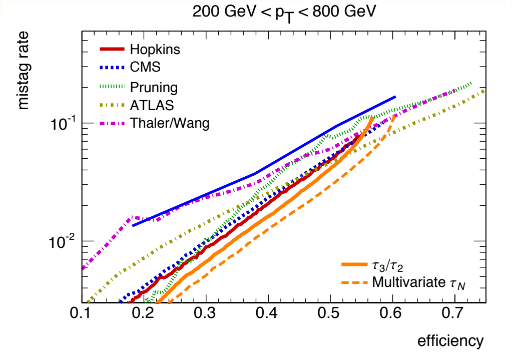
Mass reconstruction



Before trimmed mass cut QCD and untagged tops look very similar

Ultimately keep untrimmed jets after cutting on trimmed versions

Overall performance

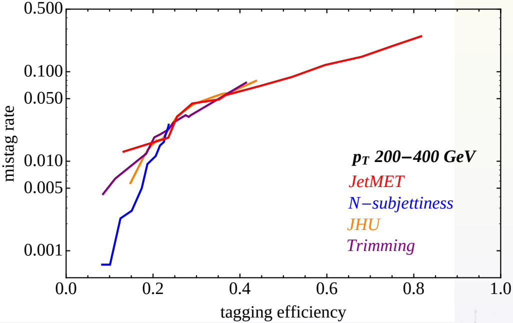


Efficiencies compared to standard laid out in BOOST 2010 report

Performance on full top samples, fairly mediocre, but . . .

Overall performance

Low p_T bins



Efficiencies compared to standard laid out in BOOST 2010 report

. . . when tops become harder to reconstruct, performance of other taggers quickly degrades

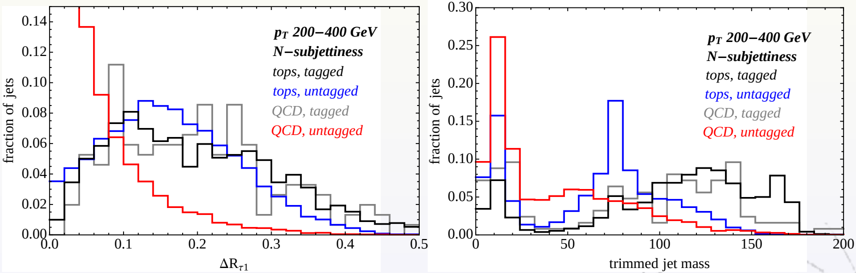
Coincidences in QCD mistags

For 50% efficiencies, $\approx$ 50% of tagged tops are new
This increases at lower efficiencies.
What about mistagged QCD jets?

Mistags occur mostly on same jets as conventional taggers, $\approx$ 70% overlap at 50% efficiency

Coincidences in QCD mistags

For 50% efficiencies, ≈ 50% of tagged tops are new
This increases at lower efficiencies.
What about mistagged QCD jets?



Mistags occur mostly on same jets as conventional taggers, ≈ 70% overlap at 50% efficiency

A two-pronged algorithm

Can imaging running our tagger in tandem with another

- Tops tagged by two methods significantly different
- Greater correlation in mistags leads to greater significance
- So far, this only makes sense at low $p_T < 400$ GeV

Natural to consider in parallel with τ_3/τ_2 tagger
Both require the same jet shapes.

- Tag with standard N-subjettiness tagger
- If untagged, run our tagger

A two-pronged algorithm

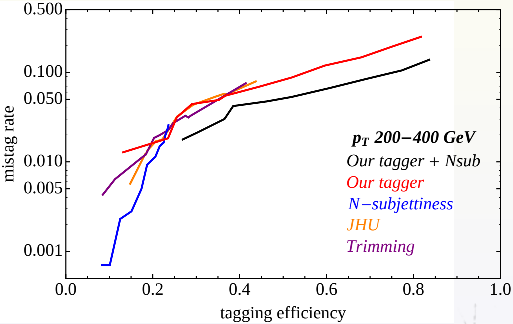
Can imaging running our tagger in tandem with another

- Tops tagged by two methods significantly different
- Greater correlation in mistags leads to greater significance
- So far, this only makes sense at low $p_T < 400$ GeV

Natural to consider in parallel with τ_3/τ_2 tagger
Both require the same jet shapes.

- Tag with standard N-subjettiness tagger
- If untagged, run our tagger

Overall Performance



Improvements in significance can be significantly increased if our tagger could be made to work at lower mistag rates

Future directions

General tagging improvements

Improvements to m_J reconstruction

- The jets passing our tagger miss some of the top by design
- We cannot hope to fully reconstruct the missing 4-momentum, but we measure the momentum transverse to jet axis
- Requires understanding of scaling from true transverse momentum to observed offset in ΔR_{τ_1}, other features
- Reminiscent of $\not{E}_T$ measured on whole events

Improve performance at higher p_T

- Currently looking at adding information from b tagging and tracking

Future directions

BSM applications

For top tagging, imbalance was caused by imperfect reconstruction. Can ask if same techniques would work if parts of jet were actually invisible

- New colored particles can be copiously produced at the LHC, and decay to $q, g+$ invisible (SUSY, RS, etc., etc.)
- If colored particles are long-lived, they will radiate before decay
  NB: Long-lived here $\neq$ displaced vertices
- May then get a fat jet containing radiation of BSM particle, and subsequent QCD decay product
- Better understanding of mass offsets might allow us to say something about mass of original particle

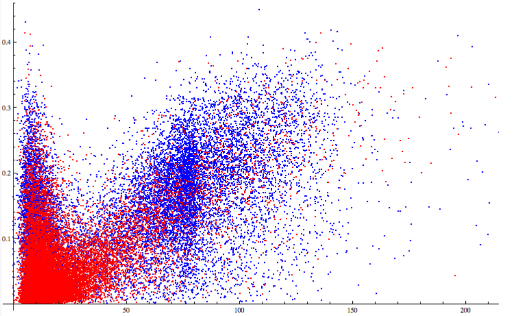
Conclusions

- It is possible to identify boosted hadronic tops by an orthogonal set of characteristics to those used thus far
- Top tagging along these lines provides performance comparable to that of standard taggers at low p_T with greater statistics
- Mistags occur on highly correlated sets of of QCD events, allowing for the construction of chained top taggers with improved performance
- Directions of N-subjettiness axes carry information that can also provide cuts for standard tagging analyses
- We hope to extend these technique into BSM applications soon, where their use may prove unavoidable

Backup slides

ΔR_{τ_1} and trimmed mass

Some (J. Wacker) have thought that ΔR_{τ_1} and trimmed mass are probing the same physics.



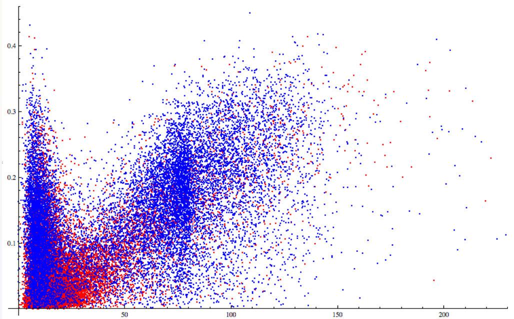
tops are blue, QCD is red

We see here they are mostly uncorrelated.

Backup slides

ΔR_{τ_1} and trimmed mass

Some (J. Wacker) have thought that ΔR_{τ_1} and trimmed mass are probing the same physics.



tops are blue, QCD is red

We see here they are mostly uncorrelated.